

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Geociências
Programa de Pós-Graduação em Análise e Modelagem de Sistemas Ambientais

Fernanda Paula Bicalho Pio

**AVALIAÇÃO DA TÉCNICA DE MODELO LINEAR DE MISTURA ESPECTRAL
COMO SUBSÍDIO À CLASSIFICAÇÃO DO USO E OCUPAÇÃO DO SOLO**

Belo Horizonte
2023

Fernanda Paula Bicalho Pio

**AVALIAÇÃO DA TÉCNICA DE MODELO LINEAR DE MISTURA ESPECTRAL
COMO SUBSÍDIO À CLASSIFICAÇÃO DO USO E OCUPAÇÃO DO SOLO**

Versão final

Dissertação apresentada ao Programa de Pós-Graduação em Análise e Modelagem de Sistemas Ambientais da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Mestre em Análise e Modelagem de Sistemas Ambientais.

Orientador: Prof. Dr. Plinio da Costa Temba

Belo horizonte

2023

P662a
2023

Pio, Fernanda Paula Bicalho.

Avaliação da técnica de modelo linear de mistura espectral como subsídio à classificação do uso e ocupação do solo [manuscrito] / Fernanda Paula Bicalho Pio. – 2023.

74 f., enc.: il. (principalmente color.)

Orientador: Plínio da Costa Temba.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Geociências, 2023.

Bibliografia: f. 69-74.

1. Modelagem de dados – Aspectos ambientais – Teses. 2. Solo – Uso – Minas Gerais – Teses. 3. Processamento de imagens – Teses. 4. Classificação – Teses. I. Temba, Plínio da Costa. II. Universidade Federal de Minas Gerais. Instituto de Geociências. III. Título.

CDU: 911.2:519.6(815.1)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ANÁLISE E MODELAGEM DE SISTEMAS AMBIENTAIS

FOLHA DE APROVAÇÃO

AVALIAÇÃO DA TÉCNICA DE MODELO LINEAR DE MISTURA ESPECTRAL COMO SUBSÍDIO À CLASSIFICAÇÃO DO USO E OCUPAÇÃO DO SOLO

FERNANDA PAULA BICALHO PIO

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ANÁLISE E MODELAGEM DE SISTEMAS AMBIENTAIS, como requisito para obtenção do grau de Mestre em ANÁLISE E MODELAGEM DE SISTEMAS AMBIENTAIS.

Eliane Maria Vieira - UNIFEI

Helder Lages Jardim - UFMG

Marcelo Antônio Nero - UFMG

Marcos Antônio Timbó Elmiro - UFMG

Plínio da Costa Temba - UFMG (Orientador)

Belo Horizonte, 24 de março de 2023.



Documento assinado eletronicamente por Marcos Antonio Timbo Elmiro, Professor do Magistério Superior, em 31/03/2023, às 13:17, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543 de 13 de novembro de 2020](#).



Documento assinado eletronicamente por Marcelo Antonio Nero, Professor do Magistério Superior, em 31/03/2023, às 14:21, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543 de 13 de novembro de 2020](#).



Documento assinado eletronicamente por Plínio da Costa Temba, Professor do Magistério Superior, em 31/03/2023, às 20:07, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por Helder Lages Jardim, Membro, em 03/04/2023, às 13:18, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por Eliane Maria Vieira, Usuária Externa, em 11/04/2023, às 09:43, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador 2194859 e o código CRC 49F9F689.

AGRADECIMENTOS

Inicialmente, agradeço a Deus pela oportunidade, força e capacidade fornecida para desenvolver mais um projeto em minha vida e concretizar mais um sonho que me permitirá convertê-lo em boas ações.

Aos meus familiares e às pessoas mais próximas que fazem parte da minha convivência diária, que me incentivaram nesta jornada e que compreenderam a dedicação que era necessária durante mais este processo de formação.

A todos os professores do Programa de Pós-Graduação em Análise e Modelagem de Sistemas Ambientais pelo conhecimento repassado, em especial ao professor Plinio Temba que demonstrou considerável presteza em conduzir a orientação deste trabalho e aos professores Marcelo Antônio Nero e Marcos Antônio Timbó pelo auxílio e simpatia durante todo o processo de formação.

Aos professores do curso de engenharia ambiental da Universidade Federal de Itajubá, *campus* de Itabira, que sempre me incentivaram no desenvolvimento da pesquisa científica.

Às colegas de turma Gabriela Rausse, Natália Lopes, Thaisa Nogueira e Raisa Costa pela amizade construída, por terem tornado o período do mestrado mais leve e divertido e por terem sido parceiras em todos os trabalhos desenvolvidos.

RESUMO

O constante desenvolvimento das tecnologias para aquisição e processamento de dados de sensoriamento remoto tem incentivado a aplicação de imagens digitais em variadas circunstâncias, inclusive em estudos ambientais e do monitoramento da superfície terrestre. Considerando o desenvolvimento das ações antrópicas e a intensificação do uso alternativo do solo para diversas atividades, os estudos do uso da Terra a partir de dados de sensoriamento remoto têm se tornado mais comum e as geotecnologias são consideradas como importantes ferramentas para fornecimento de dados que auxiliam os trabalhos de gestão territorial. A aplicação de imagens fornecidas com acesso livre também demonstra maior uso, inclusive com os dados do satélite sentinel. As imagens multiespectrais obtidas através do sensoriamento remoto permitem a aplicação de técnicas de análises para a obtenção de informações qualitativas e quantitativas de uma determinada região. Uma significativa aplicação destes dados é observada em análises do uso e cobertura da superfície terrestre sendo este um conhecimento de relevante importância para a gestão territorial. Para os estudos de classificação do uso e ocupação da Terra diversas técnicas de processamento e classificação de imagens podem ser aplicadas. Estas, por sua vez, variam em função das características dos sensores de aquisição das cenas e da finalidade do experimento. Em estudos de classificação de imagens é comum se deparar com problemas de mistura espectral que são observados devido à resolução espacial dos sensores. Dependendo da finalidade e parâmetros do estudo a ser realizado, a mistura espectral poderá se caracterizar como uma limitação dos dados e processos realizados. Assim, métodos de trabalho são utilizados para permitir extrair informações das imagens com maior riqueza de detalhes considerando as propriedades dos materiais presentes dentro de um pixel. O estudo da mistura espectral é utilizado para auxiliar as técnicas de classificação das imagens digitais. Na análise da mistura espectral ocorre então a decomposição do espectro apresentado dentro de um pixel e diferentes métodos podem ser aplicados para esta finalidade. O método de maior frequência de uso é o do modelo linear de mistura espectral (MLME). A facilidade de operação do método faz com que ele seja empregado em estudos de diversas partes do mundo tendo bons resultados. O presente trabalho objetiva abordar o modelo linear de mistura espectral baseado no método dos mínimos quadrados com restrição. Este modelo assume que a resposta espectral em cada pixel em qualquer banda espectral é decorrente de uma combinação linear das respostas espectrais de cada componente presente na mistura. Assim, neste estudo buscou-se comparar dois resultados de classificação do uso e ocupação do solo para a região da microbacia do Ribeirão Jirau, localizada nos municípios de Itabira e de Santa Maria de Itabira, no interior de Minas Gerais, sendo um mapeamento gerado a partir de classificação convencional com cenas de imagens digitais do satélite Sentinel-2 e o outro resultado gerado a partir de imagens fração obtidas no modelo linear de mistura espectral. Os resultados foram comparados a partir de coeficientes de concordância extraídos da matriz de erros onde foram consideradas análises gerais de precisão (exatidão global e coeficiente *Kappa*) e a análise individualizada por classe com o estudo das precisões do usuário e do produtor, buscando-se conhecer detalhadamente o desempenho dos classificadores. Os valores de precisão geral para as classificações originárias dos dados do Sentinel-2 e do MLME foram de 65% e 62%, respectivamente. O coeficiente *Kappa* apresentou resultado de 0,53 para a classificação obtida a partir dos dados espectrais do satélite e 0,62 para a classificação obtida a partir do modelo de mistura indicando que a qualidade dos dois mapeamentos pode ser considerada como moderada, suficientemente boa ou boa. A análise das classificações por categoria revelou que os dois mapeamentos apresentaram especificidades, principalmente para as classes de floresta natural e floresta plantada. O estudo permitiu atestar a funcionalidade do uso de imagens fração do MLME como subsídio à classificação de imagens e proporcionou possibilidades de novos tipos de análises para a aplicação do modelo.

Palavras-chave: Mistura espectral. Modelo linear de mistura espectral. Classificação de imagens. Sentinel-2.

ABSTRACT

The constant development of technologies for the acquisition and processing of remote sensing data has encouraged the application of digital images in various circumstances, including environmental studies and monitoring of the Earth's surface. Considering the development of anthropic actions and the intensification of alternative land use for various activities, studies of land use based on remote sensing data have become more common and geotechnologies are considered important tools for providing data that help territorial management work. The application of images provided with free access also demonstrates greater use, including data from the sentinel satellite. The multispectral images obtained through remote sensing allow the application of analysis techniques to obtain qualitative and quantitative information on a given region. A significant application of these data is observed in analyzes of the use and coverage of the earth's surface, which is knowledge of great importance for territorial management. For land use and occupation classification studies, several image processing and classification techniques can be applied. These, in turn, vary depending on the characteristics of the scene acquisition sensors and the purpose of the experiment. In image classification studies, it is common to encounter spectral mixing problems that are observed due to the spatial resolution of the sensors. Depending on the purpose and parameters of the study to be carried out, the spectral mixture may be characterized as a limitation of the data and processes carried out. Thus, working methods are used to extract information from images with greater detail, considering the properties of the materials present within a pixel. The study of the spectral mixture is used to help the classification techniques of digital images. In the analysis of the spectral mixture, the spectrum presented inside a pixel is decomposed and different methods can be applied for this purpose. The most frequently used method is the linear model of spectral mixing (MLME). The ease of operation of the method makes it used in studies in different parts of the world with good results. The present work aims to address the linear model of spectral mixing based on the method of least squares with restriction. This model assumes that the spectral response at each pixel in any spectral band is due to a linear combination of the spectral responses of each component present in the mixture. Thus, in this study, we sought to compare two classification results of land use and occupation for the Ribeirão Jirau microbasin region, located in the municipalities of Itabira and Santa Maria de Itabira, in the interior of Minas Gerais, with a mapping generated from conventional classification with scenes of digital images from the Sentinel-2 satellite and the other result generated from fraction images obtained in the linear model of spectral mixture. The results were compared based on concordance coefficients extracted from the error matrix, where general precision analyzes were considered (global accuracy and Kappa coefficient) and the individualized analysis by class with the study of user and producer precisions, seeking to know detail the performance of the classifiers. The overall precision values for classifications derived from Sentinel-2 and MLME data were 65% and 62%, respectively. The Kappa coefficient presented a result of 0,53 for the classification obtained from the satellite spectral data and 0,62 for the classification obtained from the mixture model, indicating that the quality of the two mappings can be considered as moderate, good enough or good. The analysis of the classifications by category revealed that the two mappings presented specificities, mainly for the classes of natural forest and planted forest. The study allowed attesting the functionality of using MLME fraction images as a subsidy for image classification and provided possibilities for new types of analysis for the application of the model.

Keywords: Spectral mixing. Linear spectral mixing model. Image classification. Sentinel-2.

LISTA DE ILUSTRAÇÕES

Figura 1. Representação da imagem digital em formato matricial.....	18
Figura 2. Representação do IFOV dos sensores.	19
Figura 3. Ilustração explicativa do processo de mistura espectral para três sensores com resoluções espaciais diferentes e quatro classes de cobertura no terreno.....	26
Figura 4. 10.(A) mostra a composição colorida das imagens fração ($R_{\text{solo}}G_{\text{vegetação}}B_{\text{sombra/água}}$) derivadas do sensor OLI/Landsat-8 para a região de Manaus (AM). As imagens fração correspondentes são:(B) solo, (C) vegetação e (D) fração sombra/água.....	30
Figura 5. Fluxograma do processo metodológico utilizado.	32
Figura 6. Localização da área de estudo.....	33
Figura 7. Caracterização da área de estudo com destaque para regiões de vegetação nativa características da Mata Atlântica e pastagens.....	34
Figura 8. Caracterização da área de estudo com destaque para regiões com presença de água (barragem de Santana).....	35
Figura 9. Caracterização da área de estudo com destaque para regiões com presença de água (barragem de Santana).....	35
Figura 10. Caracterização da área de estudo com destaque para a área urbana do município.....	36
Figura 11. Caracterização da área de estudo com destaque para área de pastagem degradada.....	36
Figura 12. Caracterização da área de estudo com destaque para áreas de mineração e urbana.....	37
Figura 13. Composição colorida RGB-432 para a área de estudo.....	39
Figura 14. Demonstração das amostras puras de componentes utilizados para o MLME.	40
Figura 15. Curva de resposta espectral (em μm) para as regiões dos componentes de referência adquiridos para o MLME. A linha de cor laranja representa a resposta espectral do solo, a verde da vegetação e a azul da água.....	47
Figura 16. Visualização das imagens fração. (A) Composição colorida RGB-432 do sensor MSI/Sentinel-2; (B) Composição colorida das imagens fração do modelo de mistura; (C) imagem fração de vegetação; (D) imagem fração do solo; e (E) imagem fração de sombra/água.	48
Figura 17. Seleção de um pixel em região de solo para exemplificação das proporções de <i>endmembers</i> representada. O MLME resultou em 100% de solo para o pixel analisado.	49
Figura 18. Seleção de um pixel em região de vegetação para exemplificação das proporções de <i>endmembers</i> representadas. O MLME resultou em 100% de vegetação para o pixel analisado.	49

Figura 19. Seleção de um pixel em região de água para exemplificação das proporções de <i>endmembers</i> representadas. O MLME resultou em 100% de água para o pixel analisado.	50
Figura 20. Seleção de um pixel em área de afloramento rochoso com presença de vegetação para observação do efeito de mistura. Proporções obtidas: 0,21 (solo), 0,31 (vegetação) e 0,46 (água/sombra).	50
Figura 21. Seleção de um pixel em área de vegetação nativa para observação do efeito de mistura sendo obtido 100% de vegetação.	51
Figura 22. Seleção de um pixel em área de solo exposto para observação do efeito de mistura. Proporções obtidas: 0,88 (solo), 0,12 (vegetação) e 0,0 (água/sombra).	51
Figura 23. Seleção de um pixel em área de floresta plantada para observação do efeito de mistura sendo obtidas frações de vegetação (0,83) e de sombra (0,17).	51
Figura 24. Mapa de uso e ocupação do solo gerado a partir das cenas originais do satélite Sentinel-2.	52
Figura 25. Mapa de uso e ocupação do solo gerado a partir do resultado do MLME.	53
Figura 26. Comparação entre os resultados de precisão geral.	55
Figura 27. Comparação entre os resultados do coeficiente <i>kappa</i>	56
Figura 28. Representação dos resultados de precisão do usuário obtidos a partir do mapeamento realizado com as cenas espectrais originais do Sentinel-2.	58
Figura 29. Representação dos resultados de precisão do produtor obtidos a partir do mapeamento realizado com as cenas espectrais originais do Sentinel-2.	59
Figura 30. Representação dos resultados de precisão do usuário obtidos a partir do mapeamento realizado com as imagens fração do MLME.	61
Figura 31. Representação dos resultados de precisão do produtor obtidos a partir do mapeamento realizado com as imagens fração do MLME.	61
Figura 32. Comparação entre os resultados de precisão do usuário.	61
Figura 33. Comparação entre os resultados de precisão do produtor.	62
Figura 34. Comparação dos mapeamentos de área de floresta natural cujo MLME resultou em 100% de proporção de vegetação. Destaca-se a maior proporção de área considerada como sendo de floresta plantada no mapeamento gerado a partir do MLME sendo a área de floresta natural.	62
Figura 35. Comparação dos mapeamentos de área de solo exposto cujo MLME resultou divisão entre as proporções de solo e vegetação e gerou como classificação final a categoria de solo (de maior proporção no pixel selecionado). Destaca-se o mapeamento do MLME com maior abrangência de pixels de solo exposto.	63

Figura 36. Comparação dos mapeamentos de área de solo exposto cujo MLME 100% de solo para o pixel selecionado. Destaca-se para a região o mapeamento do Sentinel-2 que apresentou maior uniformidade da área de solo sendo o mapeamento do MLME com mistura entre as classes de solo, área urbana e mineração..... 64

LISTA DE QUADROS

Quadro 1. Especificação das bandas Sentinel-2.	21
Quadro 2. Relação de materiais utilizados para o desenvolvimento do estudo.....	38
Quadro 3. Quantitativo de amostras de treinamento.	41
Quadro 4. Tamanho das amostras para a classificação obtida a partir dos dados espectrais do Sentinel-2.....	42
Quadro 5. Tamanho das amostras para a classificação obtida a partir dos dados espectrais do MLME.	42
Quadro 6. Modelo ilustrativo da matriz de erros.....	44
Quadro 7. Classificação do coeficiente Kappa conforme os principais estudos da literatura. .	46
Quadro 8. Relação de áreas para cada categoria classificada.....	53
Quadro 9. Códigos representativos das classes.	54
Quadro 10. Matriz de confusão gerada a partir da análise de precisão do mapa gerado a partir das cenas originais do Sentinel-2.....	54
Quadro 11. Matriz de confusão gerada a partir da análise de precisão do mapa gerado a partir das imagens fração do MLME.....	55
Quadro 12. Interpretação do coeficiente <i>kappa</i> conforme autores.....	56

LISTA DE ABREVIATURAS SIGLAS

ANA	Agência Nacional de Águas e Saneamento Básico
AcATaMa	<i>Accuracy Assessment of Thematic Maps</i>
CBH-SANTO ANTÔNIO	Comitê de Bacia Hidrográfica do Rio Santo Antônio
Cód.	Código
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
ESA	<i>European Space Agency</i>
FEAM	Fundação Estadual do Meio Ambiente de Minas Gerais
IBGE	Fundação Instituto Brasileiro de Geografia e Estatística
IDE-SISEMA	Infraestrutura de Dados Espaciais do Sistema Estadual de Meio Ambiente e Recursos Hídricos
IFOV	<i>Instantaneous Field of View</i>
INPE	Instituto Nacional de Pesquisas Espaciais
L1C	Nível 1 C
L2A	Nível 2 A
LMSM	<i>Linear Model of Spectral Mixing</i>
M	Metro
MLME	Modelo Linear de Mistura Espectral
MODIS	<i>Moderate Resolution Imaging Spectroradiometer</i>
MSI	<i>Instrumento MultiSpectral</i>
MSS	<i>Multispectral Scanner</i>
Nm	Nanômetro
ND	Número Digital
OLI	<i>Operational Land Imager</i>
RGB	<i>Red, Green, Blue</i>
Sitim	Sistema de Tratamento de Imagens
Spring	Sistema de Processamento de Informações Geográficas
TM	<i>Thematic Mapper</i>
UFV	Universidade Federal de Viçosa
USGS	<i>United States Geological Survey</i>
UTM	<i>Universal Transverse de Mercator</i>
Km	Quilômetro

SUMÁRIO

1. INTRODUÇÃO	14
1.1 Objetivo Geral.....	16
1.2 Objetivos Específicos	16
2. FUNDAMENTAÇÃO TEÓRICA	17
2.1 Sensoriamento Remoto	17
2.2 Imagens Digitais e Sensores Orbitais	18
2.3 Classificação de Imagens	22
2.4 Mistura Espectral	23
2.5 Modelo Linear de Mistura Espectral.....	27
2.6 Escolha dos <i>Endmembers</i>	29
2.7 Imagens Fração	29
3. METODOLOGIA	31
3.1 Área de Estudo.....	32
3.2 Materiais	38
3.3 Classificação do Uso e Ocupação do Solo.....	38
3.4 Aquisição e Preparação de Dados.....	39
3.5 Modelo Linear de Mistura Espectral.....	40
3.6 Validação das Classificações	41
4. RESULTADOS	46
5. DISCUSSÃO.....	64
6. CONCLUSÃO	67
REFERÊNCIAS	69

1. INTRODUÇÃO

O avanço das tecnologias para aquisição e processamento de dados de sensoriamento remoto tem favorecido o uso de imagens digitais e incentivado a aplicação de diferentes técnicas de processamento (PEREIRA *et al.*, 2016; SILVA, 2018; MOKHTARI, 2019). A aplicação dos produtos de sensoriamento remoto em estudos ambientais e de monitoramento da Terra ganhou ainda mais destaque quando utilizados dados de acesso livre, apresentando crescente utilização das imagens de satélites como o *Moderate Resolution Imaging Spectroradiometer* (MODIS), Landsat 8 e o Sentinel-2 (RADOČAJ *et al.*, 2020), dentre outros.

A missão Sentinel-2 é operada a partir do Instrumento MultiSpectral (MSI) e foi projetada para realizar o monitoramento de superfícies terrestres globais e águas costeiras com sensores multiespectrais de melhor resolução (GASCON *et al.*, 2017). Atualmente, os seus dados são disponibilizados para os usuários nos níveis L1C (dados de reflectância no topo da atmosfera) e L2A (dados de reflectância corrigidas para reflectância de superfície).

Os dados obtidos das imagens digitais, registradas via sensoriamento remoto, fornecem a possibilidade de extrair informações qualitativas e quantitativas de uma determinada região de interesse. Esta possibilidade dá-se por meio da aplicação de técnicas de processamento que analisam a resposta espectral obtida nas diferentes bandas disponibilizadas pelos sensores, sendo estas comparadas com a resposta espectral padrão dos alvos de interesse. Um dos procedimentos frequentemente utilizados para a caracterização de áreas é a classificação das cenas para interpretação do uso e ocupação do solo, podendo ela ser realizada a partir de diversos métodos e em diferentes condições, como no âmbito do pixel, de objeto e da cena (CHENG *et al.*, 2020).

Dependendo do tipo de análise a ser realizada a classificação do pixel, que é uma das formas frequentemente aplicadas em análises da superfície da terra, pode apresentar limitações devido à resolução espacial dos sensores. Com a evolução do sensoriamento remoto e das técnicas de trabalho com dados multiespectrais, alguns métodos foram estudados para permitir extrair informações com maior riqueza de detalhes sobre as propriedades dos materiais presentes dentro de um pixel, surgindo assim uma outra vertente com estudos que buscam analisar as substâncias constituintes de um pixel, ou seja, uma análise em escala de subpixel (KESHAVA e MUSTARD, 2002).

Para estas situações, estudos de análise em nível de subpixel foram desenvolvidos a fim de obter técnicas que podem gerar resultados que auxiliem a classificação correta de uma região fornecendo uma estimativa das proporções de área de cada classe de uso e ocupação do solo dentro do pixel utilizando os dados de sensoriamento remoto (LI *et al.*, 2014). Este tipo de

análise é conhecido como o estudo da mistura espectral. Considera-se que um elemento de área (pixel) pode estar apresentando uma combinação de diferentes objetos e para analisá-lo detalhadamente utiliza-se de técnicas de decomposição espectral dos pixels (KESHAVA e MUSTARD, 2002).

Na análise da mistura espectral ocorre então a decomposição do espectro apresentado dentro de um pixel. Os objetos familiares na cena (vegetação, solo, água, etc) são conhecidos como *endmembers* (KESHAVA e MUSTARD, 2002; SMALL, 2004; PIROMAL, 2007) e a fração correspondente de cada *endmember* dentro do pixel pode ser nomeada como abundância e é indicativa da proporção de cada membro final dentro do pixel (KESHAVA e MUSTARD, 2002; SHI, 2014). Assim, o estudo da mistura espectral surge a partir do interesse de se classificar corretamente um pixel buscando auxiliar as técnicas de classificação de imagens digitais (SHIMABUKURO e PONZONI, 2017; SHIMABUKURO; DUTRA; ARAI, 2020).

Diferentes métodos podem ser aplicados para a análise de mistura espectral, no entanto, o presente trabalho objetiva abordar o modelo linear de mistura espectral baseado no método dos mínimos quadrados com restrição. Para o uso deste modelo considera-se a premissa básica de que um pixel pode conter materiais diferentes dentro de si. Estes materiais, como o solo, água, vegetação, etc., são os membros finais que possuem suas respectivas abundâncias fracionárias correspondentes, logo, considera-se a existência de uma combinação das propriedades espectrais do pixel para modelar a sua variabilidade espectral (KESHAVA e MUSTARD, 2002).

Assim, o modelo assume que a resposta espectral em cada pixel em qualquer banda espectral é decorrente de uma combinação linear das respostas espectrais de cada componente presente na mistura (SHIMABUKURO e PONZONI, 2017; SHIMABUKURO; DUTRA; ARAI, 2020). O modelo exige ainda a definição prévia de *endmembers* que são espectros de referência utilizados durante a modelagem (pixels puros) (SHIMABUKURO e PONZONI, 2017; DUTRA; SHIMABUKURO; ARAI, 2019). Como resultado, o modelo linear de mistura espectral gera imagens fração que representam as proporções dos componentes considerados na mistura espectral. Na imagem digital correspondente às imagens fração ocorre o destaque das informações dos alvos da superfície terrestre e através delas tem-se a representação da informação espectral em valores de proporção dos componentes no pixel (SHIMABUKURO e PONZONI, 2017). Os alvos comumente presentes na superfície terrestre e na maioria das vezes representados a partir das imagens fração são: vegetação; solo; e sombra/água. Dessa forma, na imagem fração cada um destes alvos é destacado e pode ser visualizado em coloração mais clara.

O uso do modelo linear de mistura espectral para gerar imagens fração de componentes da superfície terrestre pode apresentar diversas aplicações como no monitoramento de desmatamento, mapeamento de áreas queimadas, identificação do uso e cobertura da terra, detecção de corte seletivo e agricultura (SHIMABUKURO e PONZONI, 2017). Dentre as diversas aplicações que podem ser consideradas, destacam-se os estudos do uso e ocupação da superfície terrestre como sendo de extrema importância para a gestão territorial. Neste tipo de análise busca-se ao máximo aproximar o mapa digital à realidade de uma região e os seus resultados devem ser avaliados de acordo com o desempenho do classificador (VALE *et al.*, 2018).

Pelo exposto, o presente trabalho objetiva contribuir com as pesquisas de classificação de imagens digitais fazendo uma abordagem da técnica de desagregação espectral utilizando o modelo linear de mistura espectral em dados do satélite Sentinel-2. Com esse propósito, considerou-se os componentes vegetação, solo e água/sombra, para auxiliar no processo de classificação do uso e ocupação do solo de uma área localizada nos municípios de Itabira e de Santa Maria de Itabira, correspondente à microbacia do Ribeirão Jirau. O resultado da classificação utilizando os dados gerados no modelo foi comparado com uma outra classificação realizada para a mesma área utilizando o método convencional de classificação a partir das cenas espectrais originais do satélite Sentinel-2.

Para comparação dos resultados foram analisados os coeficientes de concordância obtidos para cada classificação gerados a partir da matriz de erros. Cabe ressaltar que o modelo linear de mistura espectral não deve ser considerado um classificador, mas, uma técnica de transformação de imagem para facilitar a obtenção de informações (SHIMABUKURO e PONZONI, 2017).

1.1 Objetivo Geral

Avaliar o uso do resultado do modelo linear de mistura espectral como subsídio à classificação do uso e ocupação do solo sendo este comparado com uma abordagem convencional de classificação realizada a partir de cenas espectrais originais do satélite Sentinel-2.

1.2 Objetivos Específicos

São considerados objetivos específicos do presente estudo:

a) Classificar o uso e ocupação do solo utilizando as bandas B2, B3, B4 e B8 do sensor MSI a bordo do satélite Sentinel-2;

- b) Classificar o uso e ocupação do solo utilizando as imagens fração correspondentes aos componentes vegetação, solo e água/sombra resultantes do modelo linear de mistura espectral;
- c) Validar os mapeamentos gerados a partir de cenas originais do sensor MSI e das imagens fracionadas obtendo coeficientes de concordância originários das matrizes de erros; e
- d) Comparar a precisão das classificações de uso e ocupação do solo utilizando os dados espectrais originais do satélite Sentinel-2 e as imagens fração obtidas no modelo linear de mistura espectral.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Sensoriamento Remoto

Os sensores a bordo de satélites, conhecidos com sensores orbitais, são responsáveis por coletar informações a respeito da superfície terrestre e integram um conjunto de processos comumente resumidos pelo termo sensoriamento remoto onde ocorre o registro da interação entre a matéria e a energia eletromagnética por meio de técnicas de aquisição, processamento e interpretação de imagens e dados relacionados que são coletados por aeronaves e satélites sem que haja o contato físico com os objetos (SABINS, 1999; NOVO e PONZONI, 2001).

Um fluxo de radiação eletromagnética refletida pela superfície terrestre e detectada pelos sensores possui diversos comprimentos de onda. Cada comprimento de onda do fluxo de radiação é caracterizado por uma intensidade que recebe o nome de radiância. Após as interações com a atmosfera e os alvos na superfície terrestre, o fluxo de radiação é detectado pelos sensores orbitais e a intensidade é recebida na forma de sinais elétricos, sendo convertida em Números Digitais (ND) que são valores numéricos proporcionais à intensidade/radiância seguindo critérios específicos de cada faixa espectral do sensor. Dessa forma, os ND's caracterizam espectralmente os objetos da superfície terrestre dando a eles uma assinatura espectral (SHIMABUKURO e PONZONI, 2017). A possibilidade de medir a reflectância dos objetos para cada tipo de radiação do espectro eletromagnético permite observar as diferentes respostas para diferentes comprimentos de onda e prever um comportamento típico de resposta do alvo (assinatura espectral) sendo este um conhecimento fundamental para a interpretação e extração de informação em produtos de sensoriamento remoto (PONZONI, 2001; STEFFEN, 2022;).

Um dos produtos do sensoriamento remoto é disponibilizado como imagens digitais (arquivos raster que são apropriados para o uso computacional) que podem subsidiar aplicações em diversas áreas do conhecimento, tais como: avaliação de recursos terrestres, monitoramento ambiental; avaliações de mudanças do uso e cobertura do solo, urbanização, agricultura,

exploração mineral, vegetação, saúde pública, gestão territorial (SABINS, 1999, BECK; LOBITZ; WOOD; 2000; CECCATO *et al.*, 2005; NAVALGUND *et al.*, 2007; USTIN e GAMON, 2010; MAS *et al.*, 2014; WHITE *et al.*, 2016, OSGOUEI *et al.*, 2019; XI; THINH; LI, 2019; ZHAO, 2019; PREIDL; LANGE; DOKTOR, 2020; MAIDJELELE *et al.*, 2020; RADOČAJ *et al.*, 2020; KOULI, 2023, LAVENDER e LAVENDER, 2023), dentre outras.

2.2 Imagens Digitais e Sensores Orbitais

As imagens digitais geradas a partir de dados de sensoriamento remoto são representações de parte da superfície terrestre adquirida do espaço. Elas são formadas por elementos (dígito) organizados em uma malha ou grade regular. As células desta grade que são unidades conhecidas como pixels representam uma área da superfície terrestre com dimensões definidas e com valores indicativos de intensidade e localização. A localização de cada pixel é organizada por linhas e colunas representadas por “x” e “y”, respectivamente, tendo como origem da grade o canto superior esquerdo (Figura 1). Cada pixel de uma imagem digital possui um número digital (ND) como atributo numérico que representa a intensidade de energia eletromagnética emitida ou refletida medida pelo sensor. Desta forma, o ND de um pixel representa a “média da intensidade de energia refletida ou emitida pelos diferentes materiais presentes neste pixel” (CRISP, 2001; CRÓSTA, 2002).

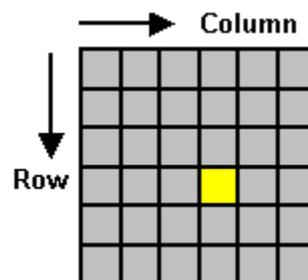


Figura 1. Representação da imagem digital em formato matricial.
Fonte: CRISP, 2001.

Os sistemas sensores são projetados para fornecer dados com detalhes específicos. O produto das aquisições que são as imagens digitais pode apresentar diferentes características de resoluções: resolução espacial; espectral; radiométrica e temporal. O tamanho dos pixels de uma imagem digital, definido previamente a partir das especificações dos sensores de aquisição, são responsáveis por definir a sua resolução espacial, ou seja, o tamanho do menor objeto capaz de ser identificado na imagem, assim, quanto menor o objeto possível de ser identificado, maior é a resolução espacial da imagem (CRÓSTA, 1992). Esta característica das imagens está diretamente relacionada com o campo de visão instantâneo do sensor (*Instantaneous Field of View* - IFOV). Na Figura 2 tem-se a representação esquemática do IFOV dos sensores onde se

destaca um cone de visibilidade do sensor (A) e a área da superfície terrestre (B) que é detectada pelo sensor em determinada altitude (C) (CANADÁ, 2019).

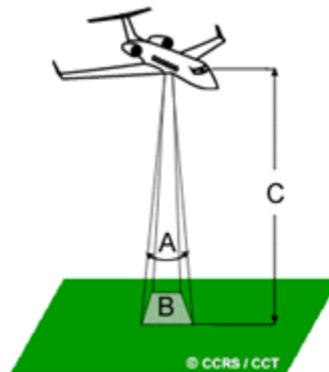


Figura 2. Representação do IFOV dos sensores.
Fonte: CANADÁ, 2019.

A resolução espectral dos dados de um sensor está relacionada com as imagens multiespectrais, ou seja, ao conjunto de imagens onde cada uma é obtida em uma determinada banda de comprimento de onda do espectro eletromagnético (CRISP, 2001). Desta forma, a resolução espectral é definida pelo “número de bandas espectrais de um sistema sensor e pela largura do intervalo de comprimento de onda coberto por cada banda” (CRÓSTA, 1992).

O uso de várias bandas do espectro facilita a distinção entre os alvos já que os objetos encontrados na superfície terrestre podem apresentar respostas espectrais semelhantes em um determinado comprimento de onda e respostas diferentes em outra faixa do espectro. Esta característica faz com que sistemas sensores com maior número de canais espectrais finos (estreitos) ofereçam maiores possibilidades de distinção entre os alvos (IBGE, 2001). Os sistemas de sensoriamento remoto projetados para registrar energia em várias faixas de comprimento de onda são chamados de sensores multiespectrais ou ainda de sensores hiperespectrais quando utilizam centenas de bandas espectrais muito estreitas na faixa do espectro eletromagnético visível, do infravermelho próximo e infravermelho médio (CANADÁ, 2019).

A resolução radiométrica das imagens digitais é uma propriedade definida pelo número de níveis digitais expresso em dígitos binários (*bits*) que representa o detalhamento de brilho da cena (energia radiante) coletado pelo sensor e que é posteriormente transformado em níveis digitais representados na cena em diferentes tonalidades de cinza. Assim, quanto maior o número de níveis, maior é a resolução radiométrica e melhor é a quantidade de detalhes perceptíveis na cena (CRÓSTA, 1992). Cada *bit* da imagem é representado por uma potência do número 2 onde 6 *bits*, por exemplo, são representados por 2^6 e significam 64 níveis de cinza

(CRÓSTA, 1992). A amplitude dos ND's indica a resolução radiométrica do sensor (SHIMABUKURO e PONZONI, 2017).

Por fim, as imagens digitais são caracterizadas também pela sua resolução temporal que condiz com a frequência de revisita do sensor em um único ponto da superfície terrestre, indicando o tempo necessário para que um satélite percorra o seu ciclo orbital completo. Esta configuração permite que os sensores adquiram dados multitemporais do local que podem ser aplicados no monitoramento da evolução das mudanças na superfície terrestre (MENEZES e ALMEIDA, 2012).

As diversas missões globais de satélites já promovidas dispõem de diferenças tecnológicas, principalmente nas características de resolução espectral (RADOČAJ *et al.*, 2020). Além disso, a fusão de imagens de diferentes satélites, principalmente os que disponibilizam dados abertos como o Sentinel-2, Landsat 8 e *Moderate Resolution Imaging Spectroradiometer* (MODIS) amplia as possibilidades de análises a serem realizadas (MOKHTARI, 2019).

Em um estudo realizado por Radočaj *et al.* (2020) os autores destacam que o uso de imagens de satélites de dados abertos para análises de monitoramento da superfície terrestre e estudos de conservação tem se destacado quando comparados com o uso de dados de satélites comerciais. A pesquisa realizada pelos autores baseou-se no banco de dados de publicações da *Web of Science Core Collection*, uma das principais bases referenciais de registro de periódicos. Neste estudo os autores destacaram o acelerado crescimento do uso dos dados dos satélites MODIS, Sentinel-2 e Landsat 8 e ressaltaram que para o período compreendido entre 2017 e 2019 as publicações que apresentavam o uso de dados da missão Sentinel-2 para pesquisas da Terra e ambiental mostraram crescimento de 371%.

Desde que os produtos do satélite Sentinel-2 começaram a ser disponibilizados, os dados têm subsidiado diversas pesquisas aplicadas às análises ambientais como cobertura do solo, agricultura, incêndios florestais, recursos hídricos, dentre outros (SOLA *et al.*, 2018).

As missões do satélite Sentinel são operadas pela Agência Espacial Europeia e integram o programa de estudos de monitoramento terrestre *Copernicus*. Esta missão possui como finalidade a observação da Terra a partir da disponibilização de séries históricas de registros comparáveis com as séries de dados de outras missões, como do Landsat 8. Cada uma das missões do Sentinel apresenta um objetivo específico e a missão Sentinel-2, que opera a partir do *MultiSpectral Instrument* (MSI), visa realizar o monitoramento de superfícies terrestres globais e águas costeiras com sensores multiespectrais de melhor resolução (GASCON *et al.*, 2017). O sensor óptico a bordo do satélite apresenta largura da faixa orbital de 290 km, sendo

capaz de amostrar 13 bandas espectrais na faixa do visível, infravermelho próximo e infravermelho de ondas curtas e com diferentes resoluções espaciais, conforme detalhado no Quadro 1 (PASQUALOTTO *et al.*, 2019; ESA, 2022).

Quadro 1. Especificação das bandas Sentinel-2.

Banda Espectral	Comprimento de Onda Central	Resolução Espacial
B02	Azul - 490 nm	10 m
B03	Verde - 560 nm	
B04	Vermelho - 665 nm	
B08	Infravermelho próximo - 842 nm	
B05	Vegetação - borda vermelha 1 - 705 nm	20
B06	Vegetação - borda vermelha 2 - 740 nm	
B07	Vegetação - borda vermelha 3 - 783 nm	
B8A	Vegetação - borda vermelha 4 - 865 nm	
B11	SWIR 1 - 1.610 nm	
B12	SWIR 2 - 2.190 nm	60 m
B01	Aerossol costeiro- 443 nm	
B09	Vapor de água - 945 nm	
B10	Infravermelho de ondas curtas (SWIR) - cirrus - 1.375 nm	

Fonte: Adaptado de Pasqualotto (2019).

Cada um dos sensores Sentinel-2 (A e B) possui tempo de revisita de 10 dias sendo a revisita combinada dos dois sensores resultante em 5 dias (resolução temporal). Já a resolução radiométrica do sensor MSI é de 12 bits (ESA, 2022).

A missão Sentinel-2 fornece dois tipos de produtos que são disponibilizados para os usuários, o Nível-1C (L1C) e o Nível-2A. (L2A). Os produtos nível 2A, corrigidos, começaram a ser disponibilizados em 2017 através do *Copernicus Open Access Hub*. A correção dos dados é realizada a partir do processador conhecido como SEN2COR (SOLA *et al.*, 2018) e os seus resultados são produtos validados conforme explicado por Gascon *et al.* (2017) sendo estes considerados consistentes e de alta qualidade (GASCON *et al.*, 2017; SOLA *et al.*, 2018). Os dados são disponibilizados em orto-imagens, em projeção *Universal Transverse de Mercator* (UTM), WGS 84 (ESA, 2022).

Frente ao desenvolvimento das ações antrópicas e à intensificação do uso alternativo do solo para diversas atividades, os estudos do uso da Terra a partir de dados de sensoriamento remoto têm se tornado mais comum e as geotecnologias são considerados como importantes ferramentas para fornecimento de dados que auxiliam os trabalhos de gestão territorial. Assim, o emprego de diferentes métodos de preparação e classificação de dados de sensores são

facilmente encontrados na literatura (PREIDL; LANGE; DOKTOR, 2020), inclusive em trabalhos desenvolvidos com os produtos Sentinel-2.

Em um trabalho realizado por Preidl, Lange e Doktor (2020) os autores utilizaram dados do Sentinel-2 para classificação da cobertura do solo baseada em pixels e obtiveram precisão geral de cerca de 88%. Os autores analisaram também a precisão de cada classe considerada no mapeamento e ressaltaram a possibilidade de alteração da precisão para uma mesma classe à medida em que se muda de região. Osgouei *et al.* (2019) utilizaram técnicas multi-índices de classificação aplicadas em dados do Sentinel-2 para distinguir diferentes classes de uso e ocupação da Terra em áreas urbanas complexas da Turquia obtendo 75% de precisão geral e 0,67 de coeficiente *Kappa*. Já Zhao *et al.* (2019) utilizaram uma combinação de dados do Sentinel-2 e de veículos aéreos não tripulados para classificação de uma área agrícola onde, para o resultado do Sentinel-2, obteve-se 71,93% de precisão geral e um coeficiente *kappa* de 0,65.

2.3 Classificação de Imagens

O mapeamento da superfície terrestre pode ocorrer de diversas formas, incluindo o mapeamento convencional com trabalhos de campo e o mapeamento realizado a partir de imagens de satélites. O mapeamento realizado a partir de imagens de satélite consiste na delimitação de áreas na superfície terrestre que apresentam a mesma resposta espectral. Assim, cada pixel da imagem recebe um rótulo que indica o seu enquadramento a uma determinada classe temática (REIS *et al.*, 2009; TALUKDAR *et al.*, 2020). Dessa forma, o processo de classificação de imagens digitais para mapeamento consiste na extração de informações com o objetivo de reconhecer padrões e objetos homogêneos na superfície terrestre permitindo assim, mapear áreas que correspondam aos temas de interesse e gerar como resultado um mapa temático indicando a distribuição geográfica das classes consideradas (CRÓSTA, 2002; INPE, 2006).

O desenvolvimento tecnológico dos sensores, das resoluções espacial, espectral, radiométrica e temporal dos dados e das ferramentas de trabalho de sensoriamento remoto tem favorecido o avanço de pesquisas que buscam gerar informações qualitativas e quantitativas da superfície terrestre (SHIMABUKURO *et al.*, 1997; GAO *et al.*, 2020). No entanto, para o uso e acompanhamento das evoluções na geração de dados é importante obter capacidade interpretativa e de processamento das imagens, o que torna as pesquisas de classificação de imagens cada vez mais ativas no campo dos estudos de sensoriamento remoto (CHENG *et al.*, 2020). À vista disso, Sheykhmousa *et al.* (2020) caracterizam a classificação de imagens como

sendo desafiadora devido ao grande volume de dados, à complexidade das paisagens e aos dados de treinamento normalmente limitados.

A classificação das imagens de sensoriamento remoto pode ser realizada no âmbito de pixel (cada pixel da cena é atribuído a um rótulo de categoria), de objeto (reconhecer os objetos na imagem) e de cena (classificar cada fragmento de imagem) (CHENG *et al.*, 2020).

Os algoritmos de classificação no contexto de pixel são geralmente segregados entre dois grupos: os algoritmos de classificação supervisionada e os algoritmos de classificação não supervisionada. Consideram-se como treinamento supervisionado as situações em que o analista reconhece regiões na imagem que permitem a identificação de classes de interesses, onde o usuário identifica na imagem áreas representativas de cada classe considerada. Já o processo não supervisionado é desenvolvido automaticamente sendo os pixels agrupados naturalmente em classes de acordo com os valores da imagem e não há o uso de dados de referência ou de conhecimento prévio da região de estudo (LI *et al.*, 2014; VALE *et al.*, 2018).

Diversos métodos classificadores supervisionados já foram até o momento desenvolvidos e apresentados em diferentes estudos. Podem ser citados: Máxima Verossimilhança, Distância Mínima; Paralelepípedo, Vizinho mais Próximo, dentre outros (LI *et al.*, 2014). Além destes, algoritmos baseados em aprendizado de máquina também são aplicados e aprimorados para a classificação de imagens, tais como redes neurais artificiais, classificação baseada em árvore de decisão, florestas aleatórias, máquina de vetores de suporte, algoritmos genéticos, dentre outros (LI *et al.*, 2014; CHENG *et al.*, 2020). O método de distância mínima, abordado neste trabalho, considera a média estatística para cada classe em cada banda espectral atribuindo cada pixel desconhecido a uma classe em que a média é a mais próxima ao pixel (VALE *et al.*, 2018).

Li *et al.* (2014) destacam ainda o tipo de classificação de imagem em nível de subpixel, onde considera-se a existência de mistura espectral nos pixels da cena devido às características de resolução espacial. Esta condição revela problemas de mistura de pixels. Os métodos de classificação subpixel buscam estimar as proporções de área de cada classe dentro do pixel, alguns deles são: classificação difusa, redes neurais, modelagem de regressão, análise de árvore de regressão e análise de mistura espectral.

2.4 Mistura Espectral

O desenvolvimento das tecnologias de sensoriamento remoto permite extrair informações com maior riqueza de detalhes sobre as propriedades dos materiais presentes dentro de um pixel, porém, além da classificação de pixels inteiros uma outra vertente surge

com estudos que buscam analisar as substâncias constituintes de um pixel, ou seja, uma análise em escala de subpixel (KESHAVA e MUSTARD, 2002).

Cada objeto presente na superfície terrestre apresenta propriedades espectrais diferentes que fazem com que eles possam refletir, transmitir e absorver a radiação eletromagnética de maneira específica. Esta condição faz com que cada elemento presente dentro do pixel possa refletir mais ou menos do que o outro considerando uma determinada faixa espectral, isto é, para uma faixa espectral um determinado elemento de resolução poderá ter seu valor de radiância influenciado pelo objeto que refletir mais intensamente nessa mesma faixa espectral mesmo que todos os objetos presentes ocupem as mesmas proporções de área no pixel (SHIMABUKURO e PONZONI, 2017).

Considerando que a radiância detectada pelos sensores é proporcionalmente representada em um número digital e que os sensores possuem capacidade de resolução espacial previamente definida entende-se que a resposta de radiância que chega ao sensor é uma média da resposta dos objetos inseridos dentro das dimensões de área do pixel fazendo com que um ND possa representar a integração (mistura) espectral de radiância referente aos diversos materiais que podem estar presentes dentro de cada pixel mais a contribuição atmosférica. O entendimento de que os pixels podem estar apresentando uma combinação de diferentes objetos leva a necessidade de, em alguns estudos, realizar a decomposição espectral dos pixels (KESHAVA e MUSTARD, 2002).

A presença de diferentes materiais dentro de um pixel leva a considerar a existência de um pixel misto (mistura espectral). A mistura espectral pode ocorrer devido a resolução espacial finita dos sensores representando áreas heterogêneas com diferentes objetos detectados dentro do campo de visão do sensor (quando os objetos representados dentro de um elemento de pixel são menores que o pixel ou quando o pixel se sobrepõe à fronteira entre dois ou mais objetos maiores do que ele ou, quando os diferentes materiais são encontrados em uma mistura homogênea (TSENG, 2000; KESHAVA e MUSTARD, 2002)). Quando existente, a mistura espectral pode gerar problemas para a classificação correta de um pixel se tornando assim um desafio para as análises de dados de sensoriamento remoto (NARVAES e SANTOS, 2007; SHI, 2014; SHIMABUKURO e PONZONI, 2017).

A análise dos componentes em âmbito de subpixel é conhecida como desagregação espectral. Neste processo ocorre a decomposição do espectro apresentado dentro de um pixel. Os objetos familiares na cena (vegetação, solo, água, etc) são conhecidos como *endmembers* e a fração correspondente de cada *endmember* dentro do pixel pode ser nomeada como

abundância e é indicativa da proporção de cada membro final dentro do pixel (KESHAVA e MUSTARD, 2002; SMALL, 2004; PIROMAL, 2007; SHI, 2014).

O problema da mistura espectral está centrado em técnicas de desagregação espectral onde um pixel misto é decomposto em membros finais e um conjunto de frações correspondentes representativas da proporção de cada membro final existente no pixel (KESHAVA e MUSTARD, 2002).

O estudo do processo de mistura espectral decorre do interesse de se classificar corretamente um pixel de acordo com os componentes ao qual ele representa já que um único pixel pode estar representando uma mistura de diversos materiais da superfície terrestre, como solo, água, vegetação, rocha, dentre outros (SHIMABUKURO e PONZONI, 2017; SHIMABUKURO; DUTRA; ARAI, 2020). Shimabukuro, Dutra, Arai (2020) explicam ainda que a mistura espectral é um complicador de técnicas de processamento digital de imagens visto que a maioria das técnicas se baseia nas características espectrais do pixel.

Quando da ocorrência de mistura espectral os métodos de classificação convencionais podem fazer com que este pixel seja mal classificado, pois, eles atribuem para cada pixel apenas uma classificação temática. Assim, a cobertura do solo observada em escala espacial inferior à resolução espacial do sensor não pode ser discriminada (SMALL, 2004).

A Figura 3 apresenta um exemplo ilustrativo de mistura espectral considerando três sensores com resoluções espaciais diferentes e cinco classes de cobertura terrestre (a, b, c, d e e). Nesta figura, os sensores 1, 2 e 3 apresentam a resolução espacial equivalente a 4a, 2a e a, respectivamente. Dessa forma, o sensor com pixels de maior dimensão, sensor 1, não apresenta nenhum pixel puro (com conteúdo único) enquanto nos sensores 2 e 3, que possuem pixels com menor área, os pixels puros podem ser identificados. Dessa maneira, quanto menor a resolução espacial de um sensor maior será a probabilidade de se observar pixels com mistura.

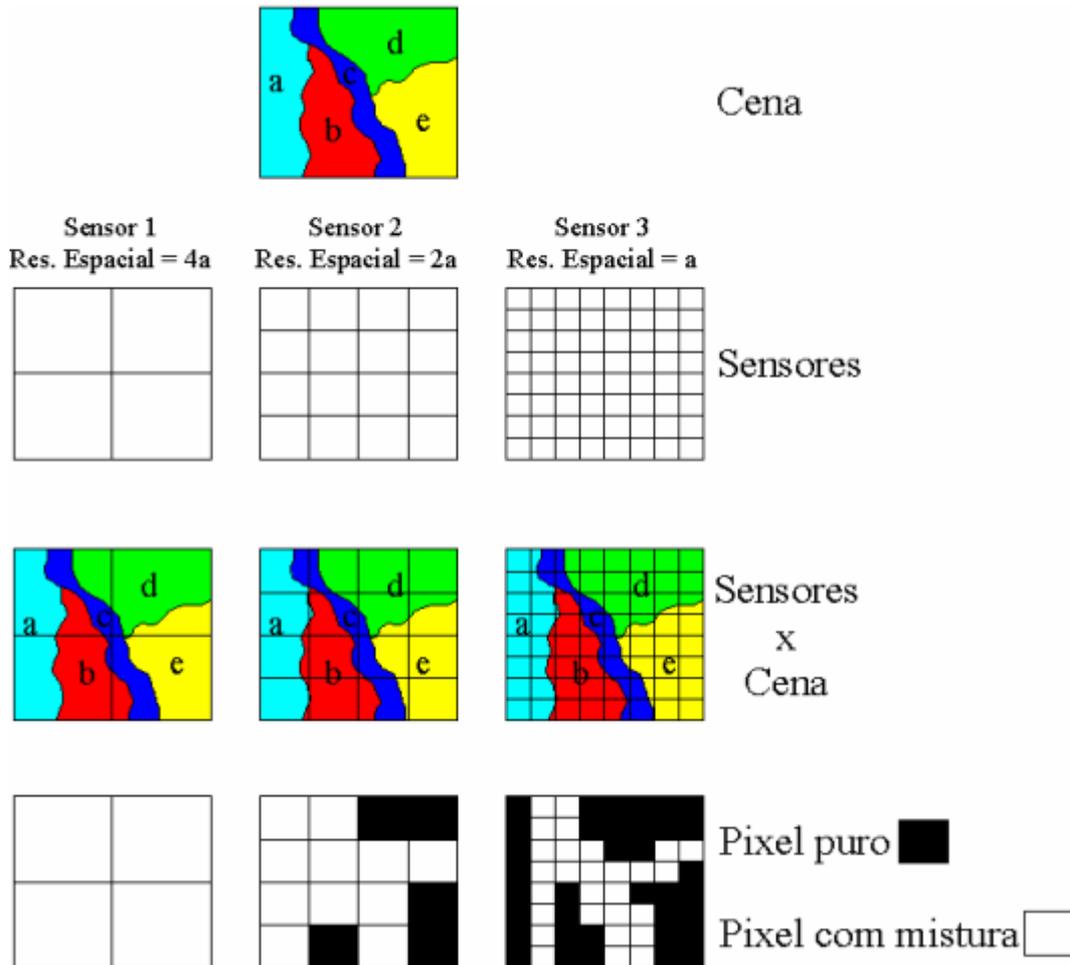


Figura 3. Ilustração explicativa do processo de mistura espectral para três sensores com resoluções espaciais diferentes e quatro classes de cobertura no terreno.

Fonte: Piromal, 2006.

O estudo da mistura espectral busca aplicar métodos de aprimoramento da técnica de classificação ou da modelagem entre os tipos e proporções de componentes dentro de um pixel. Diversos pesquisadores seguiram no pensamento de buscar estimar a proporção dos componentes dentro do pixel, como foi o caso de Shimabukuro e Smith (1991) que utilizaram um modelo de método dos mínimos quadrados para gerar imagens fração a partir de dados de sensoriamento remoto.

Os modelos analíticos fornecem a base para o estudo da mistura espectral. No entanto, são complexos para aplicação. Assim, a modelagem de mistura assume a premissa básica de que um elemento de área (pixel) pode conter materiais diferentes. Estes materiais, como o solo, água, vegetação, etc., são os membros finais que possuem suas respectivas abundâncias fracionárias correspondentes, logo, considera-se a existência de uma combinação das propriedades espectrais do pixel para modelar a sua variabilidade espectral (KESHAVA e MUSTARD, 2002).

Os métodos utilizados para análise de mistura espectral podem ser divididos entre modelos lineares e não lineares. Nos modelos lineares se assume que um pixel misto apresenta a soma ponderada dos espectros como resposta espectral final e que os pesos representativos dos membros finais são associados a cobertura de área proporcional correspondente no pixel. Esta é a proposta do modelo de mínimos quadrados indicado por Shimabukuro e Smith (1991) preparado para gerar imagens fração a partir de dados de sensoriamento remoto. Já os modelos não lineares são considerados mais complexos e são menos apresentados na literatura (KESHAVA e MUSTARD, 2002; SHI, 2014).

2.5 Modelo Linear de Mistura Espectral

Devido à maior facilidade de operação, os modelos lineares de mistura espectral (MLME) apresentam-se como sendo os métodos utilizados com maior frequência para análises de mistura espectral (KESHAVA e MUSTARD, 2002). O modelo assume que a resposta espectral em cada pixel em qualquer banda espectral é decorrente de uma combinação linear das respostas espectrais de cada componente presente na mistura. Assim, entende-se que conhecendo a proporção dos componentes no elemento de área é possível conhecer a resposta espectral de cada um deles (SHIMABUKURO e PONZONI, 2017; SHIMABUKURO; DUTRA; ARAI, 2020).

Shimabukuro e Ponzoni (2017) explicam que na década de 1980 Shimabukuro (1987) apresentou o modelo linear de mistura espectral aplicado a dados orbitais dos sensores *Thematic Mapper* (TM) e *Multispectral Scanner* (MSS) e, na década de 1990 sistemas de processamento de imagens como o Sistema de Tratamento de Imagens (Sitim) e o Sistema de Processamento de Informações Geográficas (Spring) começaram a concretizar o uso dos modelos lineares de mistura espectral, o que incentivou a sua implementação em outros sistemas.

Uma das premissas do modelo é a definição prévia de *endmembers*, que são espectros de referência utilizados durante a modelagem (pixels puros). Assim, a seleção dos *endmembers* exige o conhecimento da resposta espectral destes componentes (SHIMABUKURO e PONZONI, 2017; DUTRA; SHIMABUKURO; ARAI, 2019).

Partindo do pressuposto de que para aplicação de um modelo linear de mistura espectral considera-se que o resultado da resposta espectral de um pixel em qualquer comprimento de onda é uma combinação linear das assinaturas espectrais puras de seus componentes constituintes ponderados pela cobertura fracionada de subpixel entende-se que cada pixel apresenta informações sobre a proporção e resposta espectral de cada componente dentro da área do pixel (SHIMABUKURO e SMITH, 1991; INPE, 2022). Dessa forma, os autores explicam que em uma imagem multiespectral é possível modelar cada espectro dentro da

imagem como uma combinação linear de um conjunto finito de componentes (SHIMABUKURO e SMITH, 1991). O modelo pode ser escrito como:

$$\begin{aligned} r_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + e_1 \\ r_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + e_2 \\ &\dots \\ r_m &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n + e_m \end{aligned} \quad (1)$$

As equações podem ser expressas como:

$$r_i = \sum_{j=1}^n (a_{ij}x_j) + e_i \quad (2)$$

Sendo:

i: 1, ..., m (número de bandas);

j: 1, ..., n (número de componentes).

r_i : reflectância espectral média para a i-ésima banda espectral;

a_{ij} : reflectância espectral do j-ésimo componente no pixel para a i-ésima banda espectral.

x_j : valor de proporção do j-ésimo componente no pixel; e

e_i : erro para a i-ésima banda espectral.

Para o modelo assume-se a condição de que o número de componentes seja menor do que o número de bandas utilizadas ($n \leq m$).

Como os valores de x_j representam proporções de área dentro de um pixel de cena, as estimativas dos x_j estão sujeitas às seguintes restrições apresentadas nas equações 3 e 4.

$$r_i = \sum_{j=1}^n x_j = 1 \quad (3)$$

$$0 \leq x_j \leq 1 \quad (4)$$

Conforme exposto, o modelo linear de mistura espectral é formado por um sistema de equações sendo uma equação para cada banda do sensor considerado. Durante a aplicação do método não é necessário que o analista utilize todas as bandas disponibilizadas pelo sensor, mas, é necessário que o número de *endmembers* seja menor que o número de bandas espectrais utilizadas para a modelagem (SHIMABUKURO e PONZONI, 2017).

Para a resolução do modelo e estimativa das proporções é necessário conhecer a reflectância espectral do pixel em cada banda (r_i) e a reflectância espectral de cada componente (a_{ij}) em cada banda. O critério utilizado para a estimativa das proporções é o dos mínimos quadrados onde a estimativa das proporções busca minimizar a soma do quadrado dos erros. Três métodos podem ser aplicados e sua escolha será em função do número de componentes da mistura definido e na opção sobre aplicação da restrição. Os métodos que são comumente disponibilizados em *softwares* de processamento de imagens como o Spring, Envi e

TerraAmazon são: mínimos quadrados com restrição; mínimos quadrados ponderados e; componentes principais (SHIMABUKURO e PONZONI, 2017). A seguir, apresenta-se o conceito de cada um deles conforme os autores INPE (2022) e Shimabukuro e Ponzoni (2017):

a) Método dos Mínimos Quadrados com Restrição: este método utiliza a minimização da soma dos erros ao quadrado para estimar a proporção de cada componente dentro do pixel. Neste método, os resultados dos valores de proporção devem ser não negativos e o seu somatório igual a 1 (um).

b) Método dos Mínimos Quadrados Ponderados: método que aplica soluções iterativas onde o somatório das proporções também deve ser igual a 1 (um) e não negativo.

c) Método de Combinação entre Transformação de Componentes Principais e Mínimos Quadrados: este método utiliza uma transformação de componentes principais para diminuir o número de equações do sistema. O procedimento é seguido do método de estimação por mínimos quadrados e se destaca pela vantagem da rapidez computacional quando o número de componentes é diferente de três.

A estimativa das proporções dos pixels requer o uso de informações da assinatura espectral dos componentes considerados. Estes dados podem ser obtidos a partir de bibliotecas de curvas espectrais ou através de pixels puros obtidos na própria imagem de trabalho (SHIMABUKURO e PONZONI, 2017). Piromal (2007) ressalta ainda que a maior parte dos pesquisadores estudam as misturas que ocorrem entre os elementos vegetação, solo e água/sombra.

2.6 Escolha dos *Endmembers*

Os *endmembers* são os componentes puros da imagem (pixels puros, onde não há mistura), necessários para a aplicação dos algoritmos matemáticos. Eles podem ser obtidos diretamente na imagem (*image endmembers*) ou por meio de dados obtidos em laboratório ou campo (*reference endmembers*). Os *endmembers* escolhidos devem fazer sentido para a interpretação da imagem de forma que atendam aos critérios da fração de acordo com as equações de mistura espectral (SHIMABUKURO; DUTRA; ARAI, 2020). A escolha de *endmembers* adequados pode minimizar os erros residuais do modelo e fornecer frações razoáveis (positivas e significativas) na maior parte da imagem (ZANOTTA *et al.*, 2013).

2.7 Imagens Fração

As imagens fração representam as proporções dos componentes considerados na mistura espectral e são o produto das equações para o modelo apresentadas anteriormente. Elas realçam as informações dos alvos da superfície terrestre e através delas tem-se a representação da

informação espectral em valores de proporção dos componentes no pixel (SHIMABUKURO e PONZONI, 2017). Os alvos comumente presentes na superfície terrestre e na maioria das vezes representados a partir das imagens fração são: vegetação; solo; e sombra/água. Deste modo, a imagem fração de vegetação destaca as áreas de cobertura vegetal; a imagem fração de solo destaca as áreas de solo e a imagem fração sombra/água destaca áreas que contêm água, áreas queimadas e áreas sombreadas (SHIMABUKURO; DUTRA; ARAI, 2020). As classes de sombra, água e queimadas são normalmente representadas juntas por apresentarem resposta espectral semelhante para os sensores comuns de observação da Terra. O modelo linear de mistura espectral resulta então em cenas de imagens fração, portanto, não deve ser considerado um classificador, mas, uma técnica de transformação de imagem para facilitar a obtenção de informações (SHIMABUKURO e PONZONI, 2017).

Um exemplo para caracterização das imagem fração é apresentado na Figura 4 onde os pesquisadores Shimabukuro e Ponzoni (2017) utilizaram dados do sensor *Operational Land Imager* (OLI) do satélite Landsat 8 para estudos na região de Manaus. Na Figura 4.A tem-se uma composição colorida das imagens fração ($R_{\text{solo}}G_{\text{vegetação}}B_{\text{sombra/água}}$) seguida da imagem fração para cada um dos componentes considerados, sendo: 10.B solo; 10.C vegetação; e 10.D sombra/água. Estas imagens fração são apresentadas em escalas de cinza onde as tonalidades mais claras indicam as maiores proporções do componente.

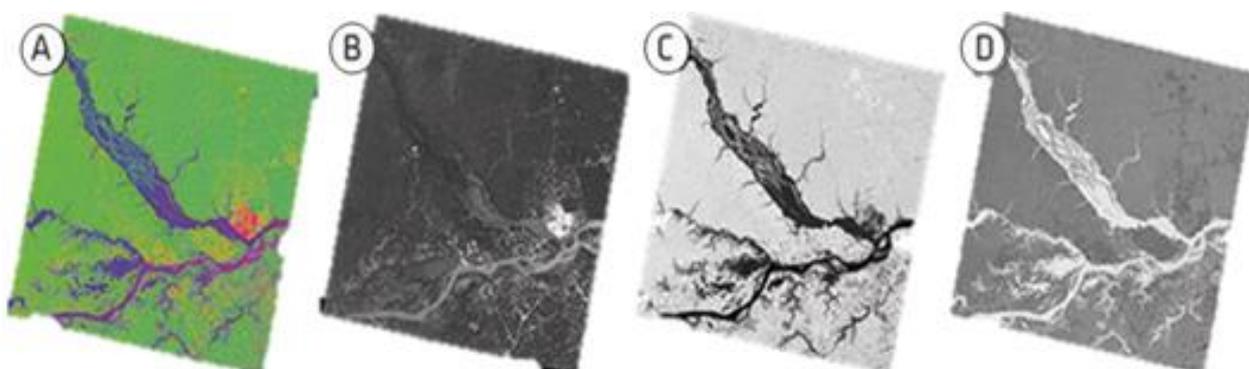


Figura 4. 10.(A) mostra a composição colorida das imagens fração ($R_{\text{solo}}G_{\text{vegetação}}B_{\text{sombra/água}}$) derivadas do sensor OLI/Landsat-8 para a região de Manaus (AM). As imagens fração correspondentes são:(B) solo, (C) vegetação e (D) fração sombra/água.

Fonte: Shimabukuro e Ponzoni, 2017.

Shimabukuro e Ponzoni (2017) ressaltam que o uso de imagens fração tem sido aplicado em diferentes estudos de diversas partes do mundo e têm gerado bons resultados, como em florestas, agricultura, uso do solo e áreas urbanas. Além disso, no uso do modelo linear de mistura espectral o número de bandas espectrais utilizadas pode ser limitado para cada tipo de

aplicação e a conversão dos dados espectrais em imagens fração pode reduzir a quantidade de dados a serem analisados.

A verificação do desempenho dos modelos na geração de imagens fração pode ser realizada a partir da formação de imagens de erro correspondentes para cada uma das bandas espectrais. Considerando a equação geral do modelo linear de mistura espectral (Equação 2), onde:

$$r_i = \sum (a_{ij}x_j) + e_i$$

O erro pode ser obtido por:

$$e_i = r_i - \sum (a_{ij}x_j) \quad (5)$$

Ou seja,

$$\begin{aligned} e_1 &= r_1 - a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ e_2 &= r_2 - a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ &\dots \\ e_m &= r_m - a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \quad (6) \end{aligned}$$

Nas imagens de erro geradas serão realçados os componentes não considerados na mistura. Dessa forma, as imagens de erro irão indicar ao usuário se o número de componentes previamente definido foi adequado para a imagem analisada, implicando na qualidade do modelo (SHIMABUKURO e PONZONI, 2017).

Considerando a aplicabilidade do modelo linear de mistura espectral na seção 3 apresenta-se a metodologia desenvolvida no presente trabalho tendo como foco a abordagem do modelo como subsídio à classificação do uso e ocupação do solo utilizando os dados do satélite Sentinel-2.

3. METODOLOGIA

A metodologia proposta e desenvolvida para o presente projeto de pesquisa consistiu em utilizar cenas espectrais do sensor MSI, a bordo do satélite Sentinel-2, para realizar a classificação do uso e ocupação do solo e comparar o resultado com um segundo mapeamento gerado a partir de imagens fração resultantes da aplicação do Modelo Linear de Mistura Espectral. Os procedimentos metodológicos necessários para a execução do estudo se resumem na seleção e caracterização da área de estudo, aquisição e preparação da base de dados vetorial e matricial, seleção das bandas do sensor a serem utilizadas, composição colorida das bandas do sensor, seleção de regiões de referência para a aplicação do modelo de mistura, obtenção das imagens fracionadas a partir do modelo, classificação supervisionada do uso e ocupação do solo

utilizando as cenas originais do satélite e as cenas resultantes da modelagem, validação dos dois mapeamentos e obtenção de coeficientes de concordância para comparação dos resultados. Um resumo da metodologia pode ser consultado Figura 5.

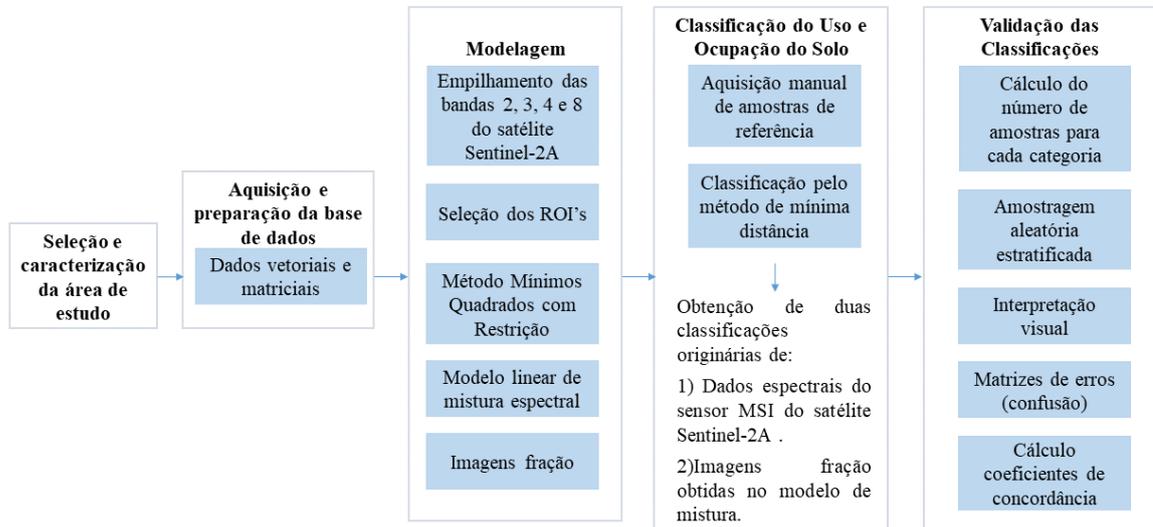


Figura 5. Fluxograma do processo metodológico utilizado.

3.1 Área de Estudo

A área de estudo utilizada para este trabalho compreende a região da bacia hidrográfica do Ribeirão Jirau, localizada nos municípios de Itabira e de Santa Maria de Itabira, no interior do estado de Minas Gerais, a cerca de 137 quilômetros da capital do estado (Figura 66).

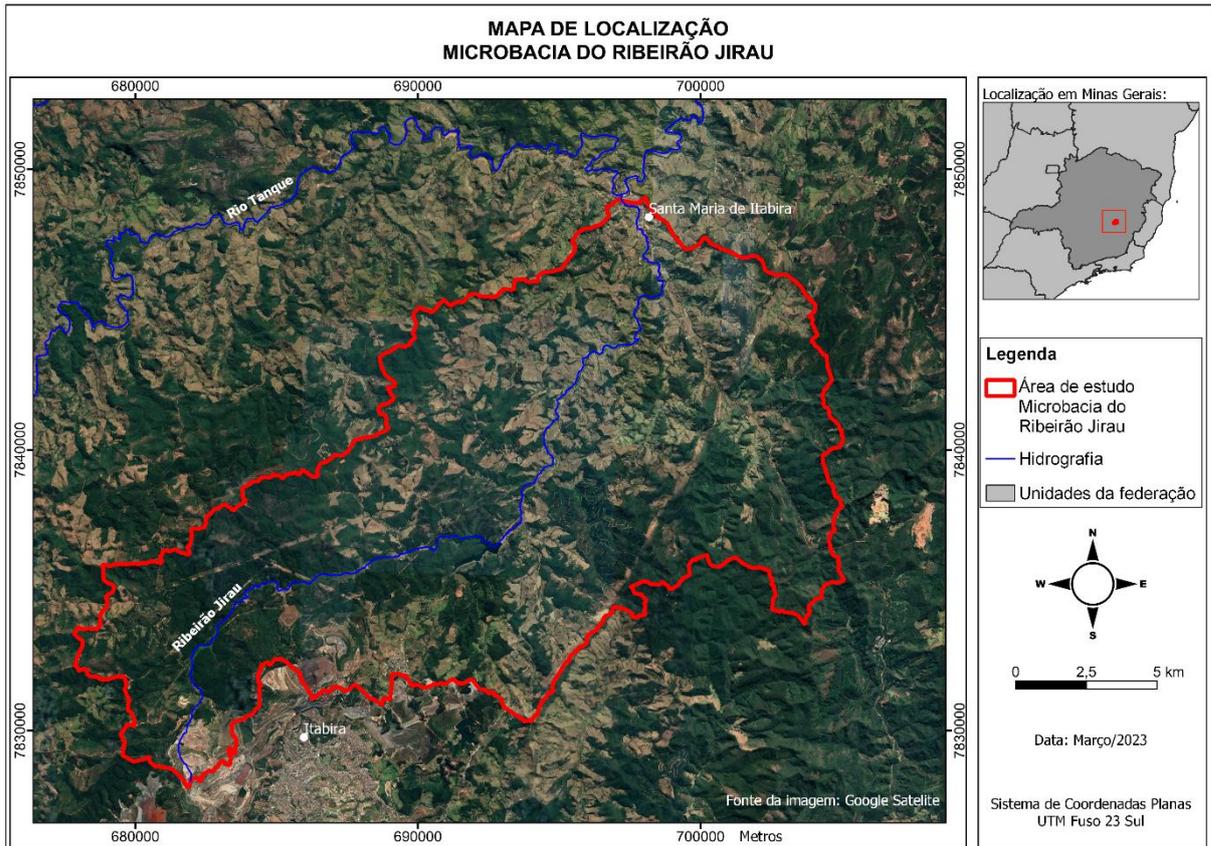


Figura 6. Localização da área de estudo.

A região está inserida na Unidade de Planejamento e Gestão de Recursos Hídricos do Rio Santo Antônio (DO3) pertencente à bacia do Rio Doce. O Ribeirão Jirau é um dos afluentes do Rio Tanque que deságua no Rio Santo Antônio que desemboca no Rio Doce. A nascente do Ribeirão Jirau está localizada no município de Itabira e a sua foz no município de Santa Maria de Itabira. Ao todo, o curso d'água apresenta cerca de 38 quilômetros e a sua bacia possui aproximadamente 28.262 hectares.

A região da bacia do Rio Santo Antônio é caracterizada por relevo acidentado, solos sensíveis à erosão e atividade mineradora, fatores estes que contribuem para uma forte suscetibilidade dos solos à erosão (CBH-SANTO ANTÔNIO, 2022).

Considerando o estudo geomorfológico realizado para o Estado de Minas Gerais e disponibilizado na Infraestrutura de Dados Espaciais do Sistema Estadual de Meio Ambiente e Recursos Hídricos (IDE-SISEMA, 2022), a área em foco está inserida no domínio dos Cinturões Móveis Neoproterozóicos, na região dos Planaltos do Leste de Minas e na unidade do Planalto da Zona Metalúrgica Mineira com categoria homogênea convexa e natureza de dissecação. Além disso, a área está inserida nas tipologias latossolo e neossolo (FEAM/UFV, 2022). Os latossolos caracterizam-se como solos bem intemperizados, profundos e

normalmente bem drenados. Já os neossolos são solos com pouco desenvolvimento pedogenético e que apresentam poucas alterações do material originário (EMBRAPA, 2018).

A região está inserida no bioma Mata Atlântica e apresenta áreas de remanescente de vegetação nativa, de floresta plantada, principalmente das culturas de eucalipto, pastagem, solo exposto e mineração, sendo esta uma forte característica da parcela inserida no município de Itabira. Além disso, uma fração da Área de Proteção Ambiental Municipal Santo Antônio está inserida dentro desta região (ITABIRA, 2018). A seguir, apresentam-se imagens para representação das características da região.

Na Figura 7 apresenta-se a vista de região localizada dentro da área de onde observa-se a paisagem composta por remanescentes de vegetação nativa, pastagens e água.

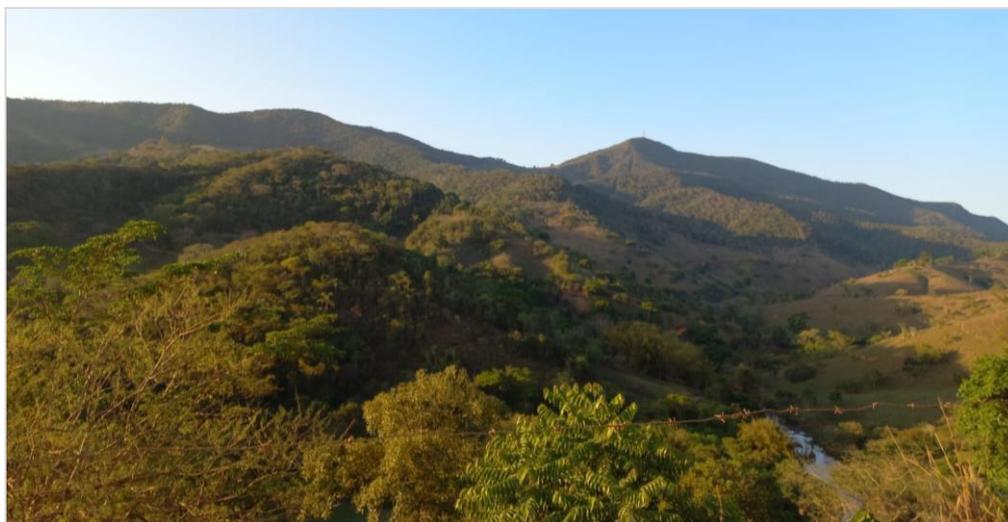


Figura 7. Caracterização da área de estudo com destaque para regiões de vegetação nativa características da Mata Atlântica e pastagens.

Fonte: Santa Maria de Itabira (2023).

Na Figura 8 observa-se a vegetação característica de pastagem e uma parcela de um barramento de água existente no local. Este barramento pode ser observado também na Figura 9 onde são observados fragmentos de vegetação no entorno do barramento.



Figura 8. Caracterização da área de estudo com destaque para regiões com presença de água (barragem de Santana).

Fonte: Figueiredo (2020).



Figura 9. Caracterização da área de estudo com destaque para regiões com presença de água (barragem de Santana).

Fonte: Figueiredo (2021).

Já a Figura 10 demonstra uma parcela da área urbana do município de Santa Maria de Itabira sendo a população distribuída no entorno da calha do Ribeirão Jirau.



Figura 10. Caracterização da área de estudo com destaque para a área urbana do município.
Fonte: Santa Maria de Itabira (2023).

Na Figura 11 mostra-se a presença de pastagens degradadas onde é possível visualizar trechos de solo exposto. Esta é uma característica facilmente observada na região.



Figura 11. Caracterização da área de estudo com destaque para área de pastagem degradada.
Fonte: Santa Maria de Itabira (2023).

Já na Figura 12, obtida a partir de um levantamento aerofotogramétrico realizado pela Prefeitura Municipal de Itabira, observa-se a presença de áreas de mineração, área urbana, pastagens e fragmentos de vegetação que se distribuem na região e são compostas por florestas de mata nativa (Mata Atlântica) e floresta plantada.



Figura 12. Caracterização da área de estudo com destaque para áreas de mineração e urbana.
Fonte: Itabira (2020).

A seleção da área de estudo deu-se em função das características locais que permitiam a identificação por meio de imagens de satélites de áreas de vegetação, água e solo exposto necessárias para o desenvolvimento da metodologia proposta no presente trabalho. Além disso, a região é constantemente impactada por eventos extremos de inundações devido ao transbordamento da calha do ribeirão Jirau além de eventos de movimentação de massa que aterrorizaram a população de Santa Maria de Itabira no ano de 2021. Estas condições somadas à carência de estudos para a área favoreceram a escolha do local.

3.2 Materiais

Para o desenvolvimento deste estudo foram utilizados os materiais detalhados no Quadro 2.

Quadro 2. Relação de materiais utilizados para o desenvolvimento do estudo.

Material	Fonte
Limite vetorial da região da bacia do Ribeirão Jirau	ANA (2022)
Imagens orbitais do satélite Sentinel, bandas B2, B3, B4 e B8, sensor MSI, já corrigidas para o nível 2A, coletadas no dia 15/09/2021, zona T23KPU, 100 x 100 km ² , UTM/WGS84	ESA (2022)
Base de dados da Infraestrutura de Dados Espaciais do Sistema Estadual de Meio Ambiente e Recursos Hídricos de Minas Gerais, utilizados para a caracterização da área de estudo	IDE-SISEMA (2022)
<i>Softwares</i> livres: QGIS versão 3.16.8 e <i>TerraAmazon</i> versão 7.3.2	QGIS <i>Development Team</i> (2021), INPE (2021)
Materiais bibliográficos citados durante o presente estudo.	-

Todos os dados geográficos utilizados neste trabalho foram adequados ao sistema de coordenadas planas, *Universal Transversa de Mercator* (UTM), fuso 23 sul, datum SIRGAS 2000.

3.3 Classificação do Uso e Ocupação do Solo

O desenvolvimento desta pesquisa deu-se a partir da aplicação de técnicas de sensoriamento remoto utilizando os dados e ferramentas citados na seção 3.2 tendo como foco o uso do modelo linear de mistura espectral. Ressalta-se que inicialmente foram realizados testes utilizando o *software Spring* e as imagens orbitais adquiridas com o sensor *Operational Land Imager* (OLI) do satélite Landsat 8 (USGS, 2022), para outras regiões do município de Itabira/MG, tendo como objetivo testar diferentes dados e ferramentas para o desenvolvimento da pesquisa. Assim, devido às diferentes características de vegetação encontradas nas áreas, à resolução espacial do sensor e às limitações do *software* para trabalhos com diferentes sensores não cadastrados identificou-se como conveniente realizar os trabalhos utilizando os dados do satélite Sentinel-2 e as ferramentas dos *softwares* QGIS e *TerraAmazon*, desenvolvendo todo o trabalho em ferramentas livres e de simples manuseio. Para o desenvolvimento metodológico do presente trabalho foi utilizado o período aproximado de 11 meses onde foram realizados os testes iniciais e se desenvolveu o presente processo.

Esta metodologia consiste em utilizar cenas do satélite Sentinel-2 para aplicar o modelo linear de mistura espectral e realizar a classificação do uso do solo na área de estudo a partir dos dados originais do Sentinel-2 e das imagens fração geradas a partir do modelo de mistura e, por fim, avaliar a precisão dos dois resultados gerados.

3.4 Aquisição e Preparação de Dados

Os dados orbitais do satélite Sentinel-2 foram adquiridos a partir da plataforma *online* de dados da *Copernicus Hub* gerenciada pela Agência Espacial Europeia (*European Space Agency - ESA*) (ESA, 2022). Foram utilizadas as cenas do nível L2A, da órbita ponto T23KPU, em projeção UTM/WGS84. As cenas foram escolhidas para um período mais recente de disponibilidade de dados e com menor cobertura de nuvens. Utilizaram-se as cenas capturadas no dia 15 de setembro de 2021 e foram selecionadas as bandas de 10 metros de resolução espacial, melhor resolução disponibilizada pelo sensor, sendo elas: B2 (azul), B3 (verde), B4 (vermelho) e B8 (infravermelho próximo). Realizou-se o empilhamento das referidas bandas e manteve-se para os trabalhos a composição colorida RGB-432 (Figura 13).

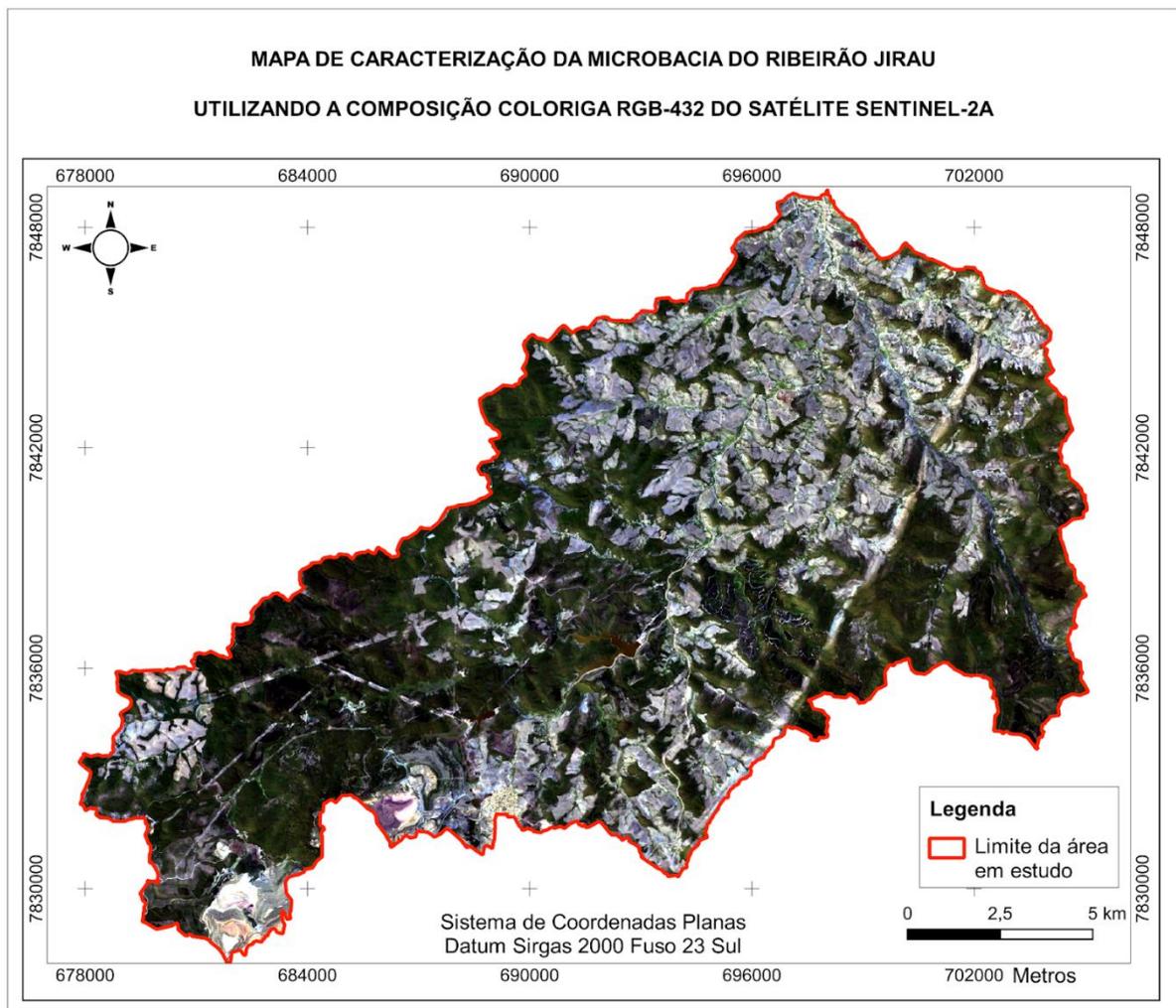


Figura 13. Composição colorida RGB-432 para a área de estudo.

Também foi utilizado o dado vetorial referente ao limite da bacia hidrográfica do Ribeirão Jirau, obtido junto à Agência Nacional de Águas e Saneamento Básico (ANA, 2022) e utilizado para o recorte da área de interesse. Todos os dados foram padronizados no sistema de coordenadas planas, UTM, fuso 23 sul, *datum* SIRGAS 2000.

3.5 Modelo Linear de Mistura Espectral

A aplicação do modelo linear de mistura espectral (MLME) foi realizada utilizando a ferramenta de processamento matricial do *software TerraAmazon* e o método dos mínimos quadrados com restrição. O procedimento prevê a seleção de *endmembers*, portanto, foi utilizada a visualização das cenas em uma composição de cor natural (RGB-432) permitindo assim facilitar a interpretação visual da região necessária para a seleção dos componentes.

Para a aplicação do modelo de mistura foram utilizados três componentes (água/sombra, vegetação e solo) e as quatro bandas do satélite Sentinel-2. A execução do modelo de mistura gera como resultado imagens fração de cada um dos componentes considerados. A seguir, tem-se a sua representação na Equação 13.

$$r_i = a * veg_i + b * solo_i + c * sombra/água_i + e_i \quad (13)$$

Onde r_i é a resposta do pixel na banda i ; a , b e c são as proporções de vegetação, solo e sombra (ou água), respectivamente; veg_i , $solo_i$ e $sombra/água_i$ são as respostas espectrais das componentes vegetação, solo e sombra (ou água), respectivamente; e_i é o erro na banda i e i indica a banda do sensor *MultiSpectral Instrument* (MSI) (SHIMABUKURO *et al.*, 1997). Utilizando a composição colorida, foram selecionados na imagem, a partir de interpretação visual, os pixels mais puros correspondentes aos componentes considerados, assim, os *endmembers* foram obtidos diretamente na imagem (técnica de *image endmembers*). A seguir, apresenta-se na Figura 14 as amostras utilizadas para o modelo de mistura.



Figura 14. Demonstração das amostras puras de componentes utilizados para o MLME.

A classificação do uso e ocupação do solo foi o procedimento realizado para comparação da aplicação do resultado do MLME. Utilizou-se as cenas do sensor MSI do satélite Sentinel-2 e as imagens fração geradas do modelo de mistura para se obter dois produtos de

classificação do uso e ocupação do solo para as classes de água, floresta natural, solo exposto, floresta plantada, afloramento rochoso, pastagem, área urbana e mineração. Estes procedimentos foram realizados no *software* QGis utilizando o complemento *Semi-Automatic Classification Plugin* (SCP) que é um *plugin* que permite realizar a classificação semiautomática ou supervisionada a partir de imagens de sensoriamento remoto (CONGEDO, 2021) e o algoritmo de classificação de mínima distância.

Para o treinamento do algoritmo realizou-se a aquisição de amostras de treinamento conforme categorias e quantitativo apresentado no Quadro 3.

Quadro 3. Quantitativo de amostras de treinamento.

Categoria	Água	Floresta natural	Solo exposto	Floresta plantada	Afloramento rochoso	Pastagem	Área urbana	Mineração
Quantidade de amostras	15	15	15	15	4	15	15	15

A classificação foi realizada utilizando as cenas do Sentinel-2 e as imagens fração obtidas no MLME gerando assim dois diferentes resultados de classificação. Ressalta-se que para as duas classificações foram utilizadas as mesmas áreas de amostras de treinamento.

3.6 Validação das Classificações

A validação do procedimento consistiu avaliar a precisão utilizando os resultados obtidos na matriz de confusão. Foram analisados dois dados de entrada: o mapeamento obtido com as imagens originais do sensor MSI/Sentinel-2 e o mapeamento referente ao mapeamento obtido a partir das imagens fração do modelo de mistura.

A acurácia temática do mapeamento foi realizada utilizando o *plugin* AcATaMa na interface do *software* QGis sendo necessário definir um projeto de amostragem conforme explicado por Llano (2022).

O cálculo do número de amostras necessárias para validação de cada categoria considerada de uso e ocupação do solo foi realizado conforme a metodologia proposta por Congalton e Green (2009), por meio de uma distribuição multinomial que considera o número de classes e a proporção da classe que representa a maior área no mapeamento. Dessa forma, o cálculo do número ideal de amostras foi realizado para o resultado da classificação utilizando os dados espectrais do Sentinel-2 e para o resultado da classificação utilizando as imagens fração geradas no MLME obtendo assim dois valores diferentes de unidades amostrais de referência. A seguir, apresenta-se a equação utilizada no cálculo do tamanho amostral.

$$N = B \Pi_i (1 - \Pi_i) / b^2 \quad (14)$$

Considera-se para o cálculo um nível de confiança desejado de 95% e precisão de 5% ($\alpha = 0,05$). Na Equação 14, N representa o número de amostras, B é um valor obtido a partir

de uma tabela de distribuição qui-quadrado com grau de liberdade $(1 - \alpha/k)$, k indica o número de classes do mapeamento, Π_i representa a proporção da classe com maior área no mapeamento e b é o erro máximo admissível quanto ao número apropriado de amostras ($b = 0,05$).

Considerando que para os dois mapeamentos foram utilizadas 8 categorias de uso e ocupação do solo tem-se $k = 8$, $B(x^2_{(1.0,99375)} = 7,568)$. Para o mapeamento realizado com os dados espectrais do Sentinel $\Pi = 0,33995$ e $N = 679$ amostras e, para o mapeamento com os dados do MLME tem-se $\Pi = 0,3866$ e $N = 718$ amostras.

No Quadro 4 e Quadro 5 é apresentada a relação de área e porcentagem de ocupação de cada categoria no espaço analisado bem como o resultado do número de amostras por extrato para cada uma das categorias.

Quadro 4. Tamanho das amostras para a classificação obtida a partir dos dados espectrais do Sentinel-2.

Classe	Área (h)	Percentual dos Extratos na População	Número de Amostra por Extrato
Corpos d'água	117,52	0,0042	3
Floresta natural	9.597,52	0,3395	231
Solo exposto	266,69	0,0094	6
Floresta plantada	7.226,94	0,2557	174
Afloramento rochoso	1.098,15	0,0388	26
Pastagem	8.081,83	0,2859	194
Área urbana	632,00	0,0224	15
Mineração	1.245,85	0,0441	30
Total	28.266,5	1	679

Quadro 5. Tamanho das amostras para a classificação obtida a partir dos dados espectrais do MLME.

Classe	Área (h)	Percentual dos Extratos na População	Número de Amostra por Extrato
Corpos d'água	115,03	0,0041	3
Floresta natural	6.171,65	0,2183	157
Solo exposto	704,79	0,0249	18
Floresta plantada	10.928,18	0,3866	278
Afloramento rochoso	1.183,01	0,0419	30
Pastagem	7.645,86	0,2705	194
Área urbana	642,33	0,0227	16
Mineração	875,65	0,0310	22
Total	28266,5	1	718

Tendo o número ideal de amostras definido para cada categoria analisada, utilizou-se as ferramentas disponibilizadas pelo *plugin* AcATaMa (*Accuracy Assessment of Thematic Maps*)

(LLANO, 2022) para se obter pontos para a amostragem aleatória estratificada para cada um dos mapeamentos. Os pontos de validação foram então distribuídos na região e rotulados conforme classificação observada na imagem digital considerando a checagem dos pontos por interpretação visual da composição colorida obtida com os dados do Sentinel-2. Esta fase do trabalho foi ainda auxiliada pelo conhecimento prévio da região.

Cabe ressaltar que durante a análise da qualidade do mapeamento houve o cuidado em garantir que as áreas utilizadas como amostra de referência não fossem consideradas no momento da avaliação, assim, manteve-se ativa no QGis a camada dos dados de referência evitando a ocorrência sobreposição das amostras.

O procedimento de indicação das classes para cada ponto foi seguido da análise de acurácia dos mapeamentos obtidos gerando como produto uma matriz de erros para cada uma das duas análises realizadas. Baseado nestes resultados, obteve-se a precisão geral e o índice *kappa*, que são análises globais da precisão, e as precisões do usuário e do produtor, que são análises que enfatizam a precisão particular de cada categoria da classificação.

A precisão de dados temáticos é uma das etapas fundamentais em procedimentos de mapeamento. A validação de dados temáticos é normalmente realizada a partir de índices de exatidão global da classificação e do coeficiente de concordância *kappa*. Estes índices são originados de uma matriz de confusão que é uma ferramenta de constante uso nos procedimentos de validação (GASCON *et al.*, 2017; OSGOUEI *et al.*, 2019; XI; THINH; LI, 2019; ZHAO *et al.*, 2019; PREIDL; LANGE; DOKTOR, 2020).

A matriz de confusão, ou, matriz de erros, como também é conhecida, é estruturada em um formato de matriz quadrada (matriz com o mesmo número de linhas e colunas) onde as colunas geralmente representam os dados de referência e as linhas representam a classificação obtida no mapeamento. Por permitir a obtenção de estatísticas descritivas e analíticas que representam a proporção de erros e acertos em um mapeamento, a matriz de confusão é uma das técnicas mais utilizadas para a avaliação da precisão (CONGALTON e GREEN, 2009). Além disso, uma matriz de confusão permite obter os coeficientes de concordância que indicam o desempenho da classificação do produtor e do usuário. Estes, por sua vez, contribuem para fornecer maior confiabilidade à análise (CONGALTON e GREEN, 2009; GASCON *et al.*, 2017).

As linhas e colunas da matriz de erros indicam o número de unidades de amostra associados a cada categoria definida em uma classificação em relação ao número de unidades de amostras definidas em outra classificação (ou em dados de sensoriamento remoto) permitindo assim a interpretação estatística das diferenças e correspondências existentes entre

elas. Dessa forma, a análise de acurácia do dado temático a partir da matriz de erros permite uma comparação cruzada entre os dados produzidos e os dados de referência utilizados (CONGALTON e GREEN, 2009; GASCON *et al.*, 2017;).

Geralmente, na matriz de erros os dados de referência são dispostos nas colunas enquanto os dados da classificação estão representados nas linhas. Já na diagonal principal da matriz tem-se os valores concordantes entre os dois mapeamentos analisados (CONGALTON e GREEN, 2009; GASCON *et al.*, 2017).

Conforme detalhado por Congalton e Green (2009), a matriz pode ser representada matematicamente supondo que n amostras estejam distribuídas em K^2 células, onde cada amostra é atribuída a uma das k categorias consideradas no mapa, geralmente representadas nas linhas e independentemente relacionada a uma das k categorias dos dados de referência comumente representados nas colunas. Seja n_{ij} o número de amostras classificadas na categoria i ($i = 1, 2, 3, \dots, k$) no mapa e na categoria j ($j = 1, 2, 3, \dots, k$) no conjunto de dados de referência conforme exemplo no Quadro 6.

Quadro 6. Modelo ilustrativo da matriz de erros.

		<i>i = Colunas</i> (Referência)			Linha total
		1	2	k	n_{i+}
<i>i = Linhas</i> (Classificação)	1	n_{11}	n_{12}	n_{1k}	n_{1+}
	2	n_{21}	n_{22}	n_{2k}	n_{2+}
	k	n_{k1}	n_{k2}	n_{kk}	n_{k+}
Coluna total		n_{+1}	n_{+2}	n_{+k}	n
		n_{+j}			

Fonte: Adaptado de Congalton e Green (2009).

Dessa forma, tem-se o número de amostras classificadas na categoria i na classificação de sensoriamento remoto sendo:

$$n_{i+} = \sum_{j=1}^k n_{ij} \quad (7)$$

e;

$$n_{+j} = \sum_{i=1}^k n_{ij} \quad (8)$$

sendo o número de amostras classificadas na categoria j no conjunto de dados de referência.

A precisão geral da classificação, ou, exatidão global, pode então ser obtida a partir da razão entre a soma dos valores da diagonal principal da matriz e o número total de amostras coletadas (n), conforme demonstrado na Equação 9 (CONGALTON e GREEN, 2009).

$$\text{Precisão geral} = \frac{\sum_{i=1}^k n_{ii}}{n} \quad (9)$$

As análises de precisão do produtor e do usuário estão associadas com os erros de omissão (erros de exclusão) e de comissão (erros de inclusão). O erro de comissão ocorre quando uma área é incluída em uma categoria, mas, ela não pertence a esta categoria e o erro de omissão ocorre quando uma área é excluída da classe ao qual ela pertence. Assim, conclui-se que todo erro é considerado uma omissão para a categoria correta e uma comissão para a categoria errada (CONGALTON e GREEN, 2009; MUÑOZ, 2016)

As precisões do produtor e do usuário são utilizadas para representar as categorias individualmente. A precisão do produtor é obtida a partir da divisão do número total de unidades amostrais classificadas corretamente para uma determinada categoria pelo número total de unidades amostrais desta mesma categoria conforme indicado nos dados de referência. Já a precisão do usuário é obtida a partir da divisão do número total de unidades amostrais classificadas corretamente em uma categoria pelo número total de unidades amostrais que foram classificadas para esta mesma categoria. Assim, é possível analisar as confusões entre categorias que podem ser encontradas no mapeamento tendo como base os dados de referência (CONGALTON e GREEN, 2009).

Seguindo o exemplo utilizado para explicação da estrutura de uma matriz de confusão apresentada no Quadro 6, tem-se que a precisão do produtor pode ser obtida a partir da Equação 10 e a do usuário pela Equação 11.

$$\text{Precisão do produtor } j = \frac{n_{jj}}{n_{+j}} \quad (10)$$

$$\text{Precisão do usuário } i = \frac{n_{ii}}{n_{i+}} \quad (11)$$

O coeficiente *kappa* é um índice de concordância entre os dados resultantes da classificação e os dados de referência considerando toda a matriz e não apenas os dados da sua diagonal principal. Ele pode ser obtido a partir da Equação 12 (FIGUEIREDO e CARVALHO, 2006; QUEIROZ *et. al.*, 2017).

$$k = \frac{n \sum_{i=1}^k n_{ii} - \sum_{i=1}^k n_{i+} n_{+i}}{n^2 - \sum_{i=1}^k n_{i+} n_{+i}} \quad (12)$$

onde:

k : número de linhas da matriz de erros;

n : número total de observações (amostras);

n_{ii} : número de observações na linha i e coluna i ;

n_{i+} : total da linha i ;

n_{+i} : total da coluna i .

O valor do coeficiente *kappa* pode variar entre 0, nenhuma concordância, e 1, concordância completa entre o resultado da classificação e os dados de referência (FIGUEIREDO e CARVALHO, 2006). Para interpretação do seu resultado é comum que se utilizem valores de referência como os apresentados pelos autores Fleiss, Cohen e Everitt (1969), Congalton e Green (2019) (baseado em Landis e Koch, 1977) e Monserud e Leemans (1992) apresentada por Foody (2020) (Quadro 7).

Quadro 7. Classificação do coeficiente Kappa conforme os principais estudos da literatura.

	Congalton e Green (2019)	Fleiss, Cohen e Everitt (1969)	Monserud e Leemans (1992)
Kappa	1,0	Quase perfeito	Excelente
	0,8		Muito bom
	0,6	Substancial	Bom
	0,4		Suficientemente bom
	0,2	Suficiente	Pobre
	0,0	Leve	Muito pobre
		Pobre	Nenhuma

Fonte: Adaptado (FOODY, 2020).

Na seção 4 apresenta-se os resultados obtidos a partir dos procedimentos descritos.

4. RESULTADOS

Após a seleção dos componentes de referência para solo exposto, vegetação e água/sombra foram obtidas as respostas espectrais de cada região de referência para cada banda do raster utilizado. Na Figura 15 é apresentado o gráfico de curva espectral gerado a partir da seleção dos componentes de referência para o modelo.

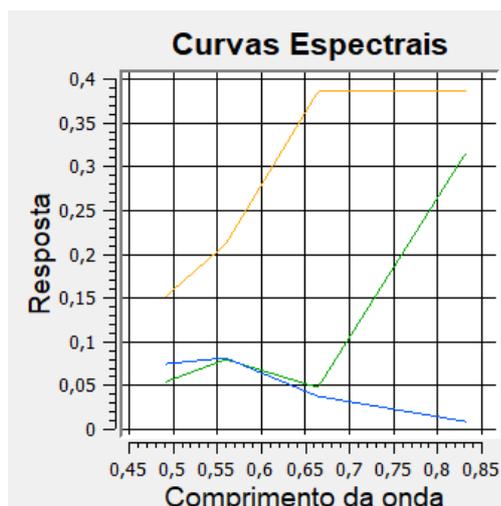


Figura 15. Curva de resposta espectral (em μm) para as regiões dos componentes de referência adquiridos para o MLME. A linha de cor laranja representa a resposta espectral do solo, a verde da vegetação e a azul da água.

A região de solo apresentou resposta espectral máxima nas bandas 4 e 8. Santori (2022) explica que a resposta espectral do solo pode ser influenciada por diversos fatores, como a quantidade de água presente no solo, a sua constituição mineralógica, concentrações de água e ar e o estado de conservação do solo. No entanto, a tendência comum dos solos é de apresentar alta reflectância a partir do comprimento de onda de 400 nm, alcançar um pico próximo de 50% entre 865 nm e reduzir a resposta espectral novamente. Para a região de solo selecionada neste estudo, a amostra apresentou aproximadamente 39% de resposta nas bandas 4 e 8, regiões de 664 nm (banda 4 - vermelho) e 832 nm (banda 8 - infravermelho próximo), respectivamente.

A resposta espectral da vegetação é normalmente influenciada pelas suas características biológicas, morfológicas e ambientais. Assim, a vegetação geralmente apresenta uma grande absorção (baixa reflectância) nas regiões de 400 nm (faixa do azul) e 600 nm (faixa do vermelho) e alta reflectância (cerca de 40%) no intervalo entre 700 nm e 1.200 nm que é a região do infravermelho próximo (SANTORI, 2022). Para a região de vegetação selecionada neste estudo, obteve-se aproximadamente 32% de resposta espectral na banda do infravermelho próximo do satélite Sentinel-2 (banda 8) e baixas respostas (que indicam alta absorção) nas bandas do azul (banda 2; 5,3%), verde (banda 3; 7,9%) e vermelho (banda 4; 4,7%).

Por fim, as áreas de água são geralmente caracterizadas pela alta absorção da radiação eletromagnética, principalmente acima da região do visível, e reflectância inferior à 10%, quando profundas e limpas, no intervalo entre 380 nm e 700 nm (SANTORI, 2022). Neste estudo, a amostra de água resultou em uma menor resposta espectral na banda do infravermelho próximo (banda 8), sendo esta de cerca de 0,9%. Já nas bandas do visível obteve 7,3% de resposta na banda azul (banda 8), 8% na banda do verde (banda 3) e 3,6% na banda do vermelho (banda 4).

O resultado do modelo de mistura pode ser observado nas Figura 16-A, B, C, D e E onde apresenta-se cada imagem fração obtida para os componentes solo, vegetação e água/sombra em comparação com a composição colorida das bandas multiespectrais do Sentinel-2.

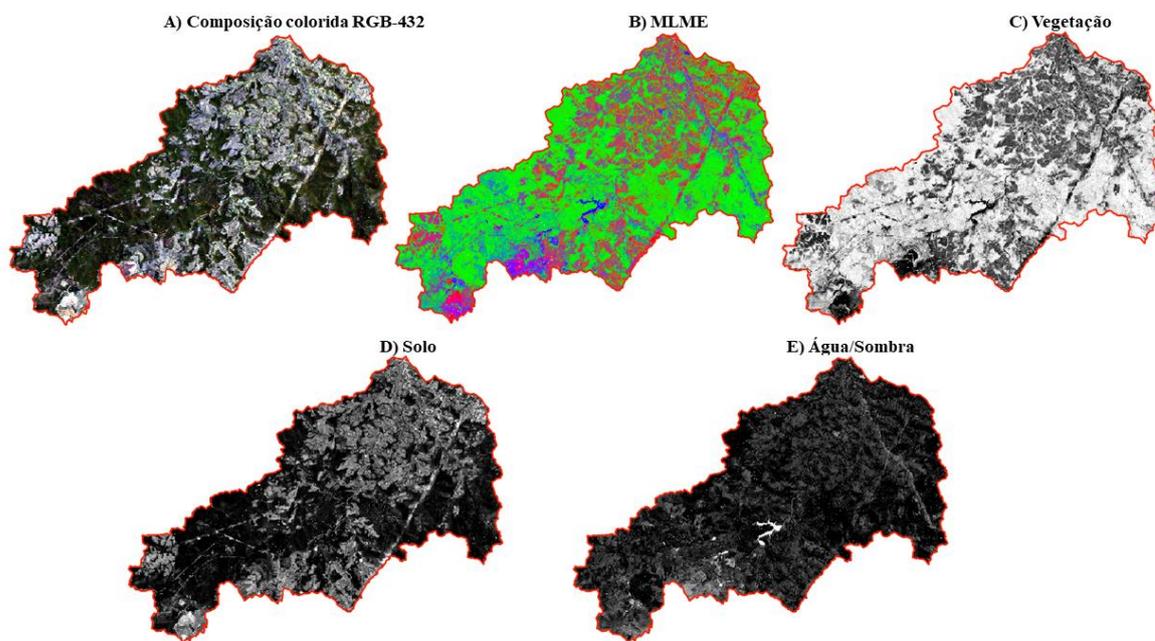


Figura 16. Visualização das imagens fração. (A) Composição colorida RGB-432 do sensor MSI/Sentinel-2; (B) Composição colorida das imagens fração do modelo de mistura; (C) imagem fração de vegetação; (D) imagem fração do solo; e (E) imagem fração de sombra/água.

As imagens fração indicadas na Figura 16 são representativas dos componentes vegetação (C), solo (D) e água/sombra (E), respectivamente. Observa-se que elas são geradas em escala de tons de cinza onde o valor de cada pixel está associado com as proporções de cada um dos componentes utilizados no modelo de mistura. Dessa forma, quanto maior for o valor de um pixel em uma imagem fração de um determinado componente, maior será a proporção deste componente dentro do pixel e a sua tonalidade será mais clara (SHIMABUKURO; DUTRA; ARAI, 2020). Diante do exposto, observa-se os pixels com maior abundância de vegetação destacados em tons mais claros na Figura 16-C, os pixels com maior abundância de solo destacados na Figura 16-D e os pixels com maior abundância de água ou sombra destacados em tons mais claros na Figura 16-E.

Na Figura 16-B apresenta-se mais um resultado do modelo de mistura que é uma composição colorida das imagens fração obtidas. Nesta composição é possível realizar a seleção de um pixel e obter a proporção de cada um dos *endmember*, conforme exemplificado nas figuras a seguir onde realiza-se a seleção de um pixel puro selecionado dentro das áreas utilizadas como amostras de referência para os componentes do modelo.

Na Figura 17 houve a seleção de um pixel (ponto indicado na cor verde) localizado em uma região de solo exposto onde o modelo resultou em 100% de solo para o ponto analisado.

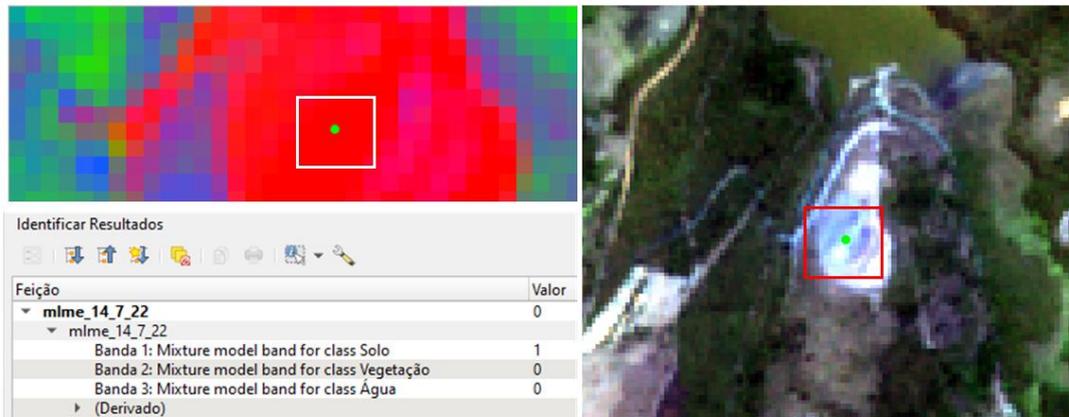


Figura 17. Seleção de um pixel em região de solo para exemplificação das proporções de *endmembers* representada. O MLME resultou em 100% de solo para o pixel analisado.

Na Figura 18 a seleção do pixel (ponto vermelho) foi em uma região de vegetação natural, resultando em 100% de vegetação conforme análise dos resultados do modelo.



Figura 18. Seleção de um pixel em região de vegetação para exemplificação das proporções de *endmembers* representadas. O MLME resultou em 100% de vegetação para o pixel analisado.

Na Figura 19 selecionou-se um pixel na região de água obtendo 100% de água para o ponto analisado.

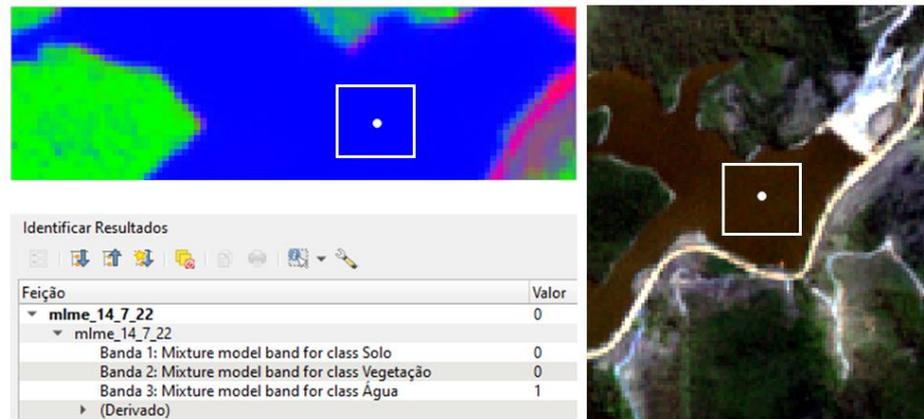


Figura 19. Seleção de um pixel em região de água para exemplificação das proporções de *endmembers* representadas. O MLME resultou em 100% de água para o pixel analisado.

Cabe ressaltar que os pixels selecionados na Figura 17, Figura 18 e Figura 19 são localizados dentro das áreas de referência escolhidas para seleção dos *endmembers* que foram regiões consideradas como pixels puros. Para observar o efeito da mistura entre os componentes, selecionou-se regiões onde são observadas maior heterogeneidade dos objetos na superfície. Nestes locais é possível observar a mistura dos componentes solo, vegetação e água/sombra dentro dos pixels, tendo cada um destes uma determinada proporção, conforme exemplificado na Figura 20.

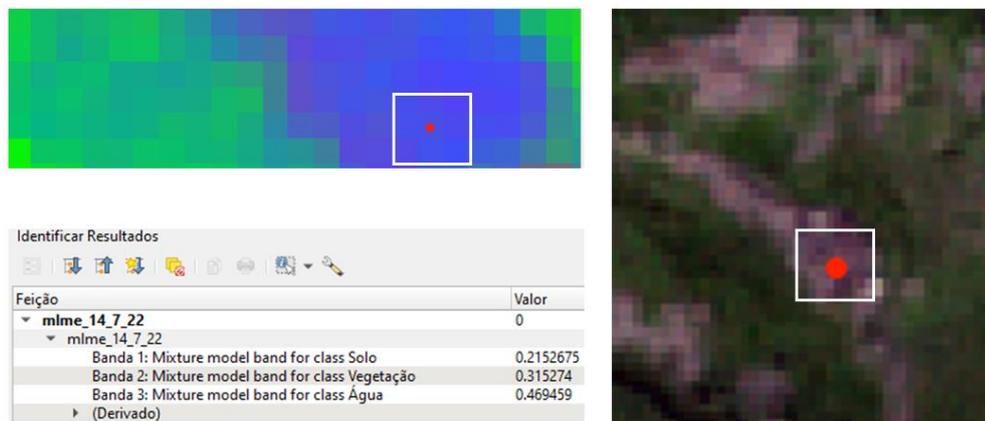


Figura 20. Seleção de um pixel em área de afloramento rochoso com presença de vegetação para observação do efeito de mistura. Proporções obtidas: 0,21 (solo), 0,31 (vegetação) e 0,46 (água/sombra).

Na Figura 21 apresenta-se um exemplo de seleção em área correspondente à presença de vegetação nativa que se localiza fora da região considerada para pixels puros. Neste local foi possível obter pixels com 100% de resposta da vegetação, conforme demonstrado.



Figura 21. Seleção de um pixel em área de vegetação nativa para observação do efeito de mistura sendo obtido 100% de vegetação.

Ao observar uma região de solo exposto, provocada pela ação de movimentação de massa na região (Figura 22), identifica-se a mistura entre frações de solo e de vegetação.

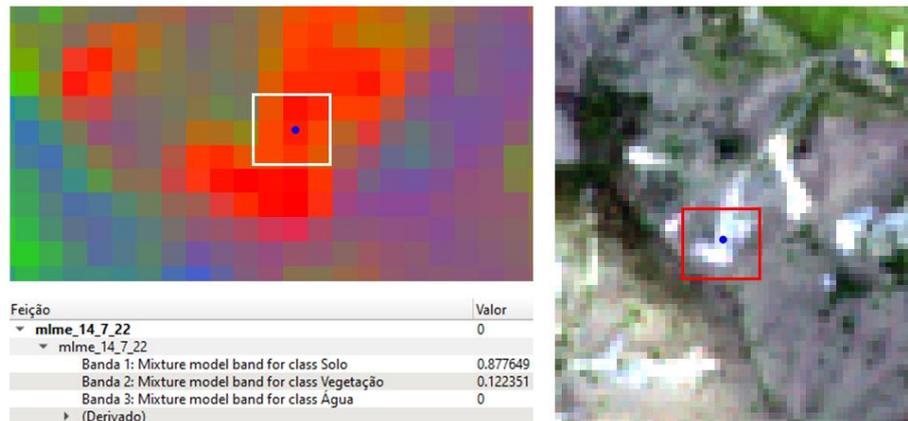


Figura 22. Seleção de um pixel em área de solo exposto para observação do efeito de mistura. Proporções obtidas: 0,88 (solo), 0,12 (vegetação) e 0,0 (água/sombra).

A mesma consideração do efeito de mistura é observada para áreas de floresta plantada (Figura 23). Nesta, verifica-se os valores correspondentes à resposta da vegetação e do possível efeito de sombreamento.



Figura 23. Seleção de um pixel em área de floresta plantada para observação do efeito de mistura sendo obtidas frações de vegetação (0,83) e de sombra (0,17).

Na Figura 24 e Figura 25 são apresentados os resultados do mapeamento da classificação do uso e ocupação do solo obtido a partir das cenas espectrais originais do Sentinel-2 e das imagens fração do modelo de mistura.

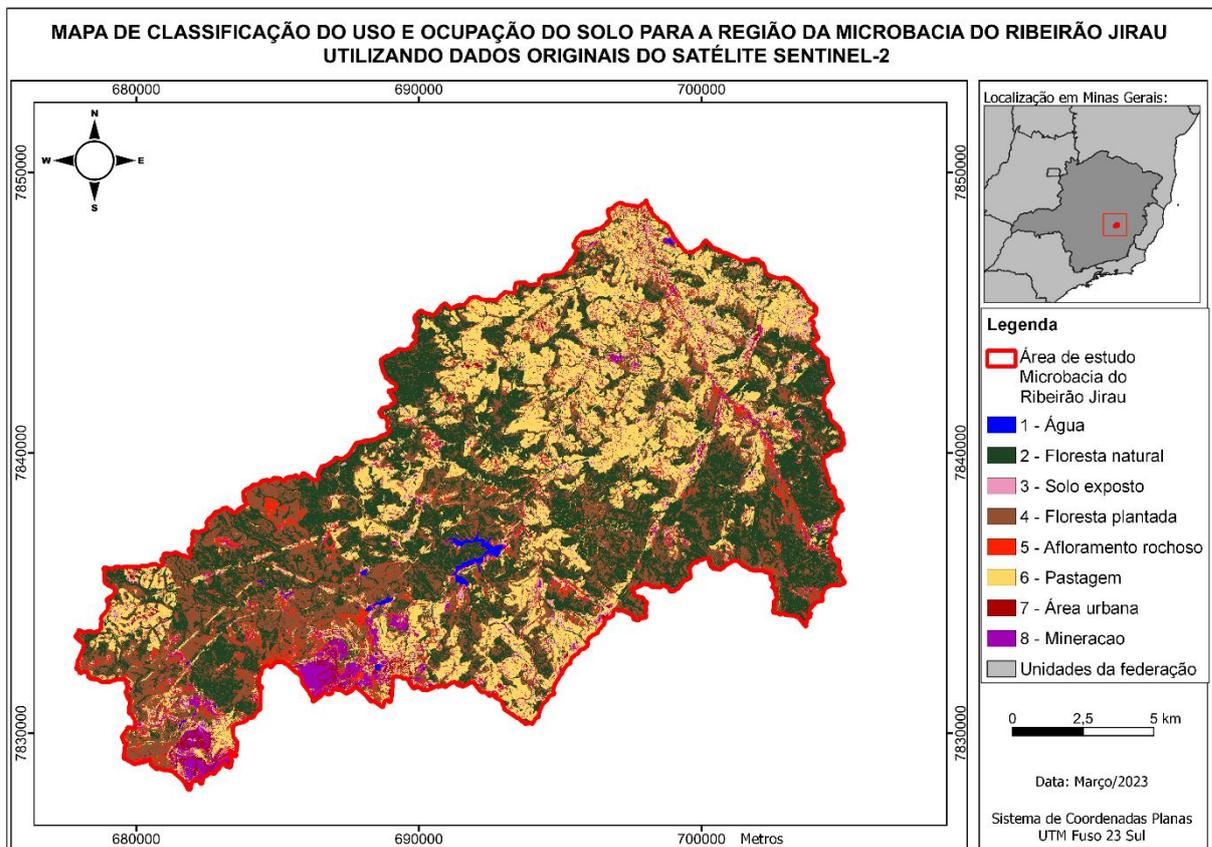


Figura 24. Mapa de uso e ocupação do solo gerado a partir das cenas originais do satélite Sentinel-2.

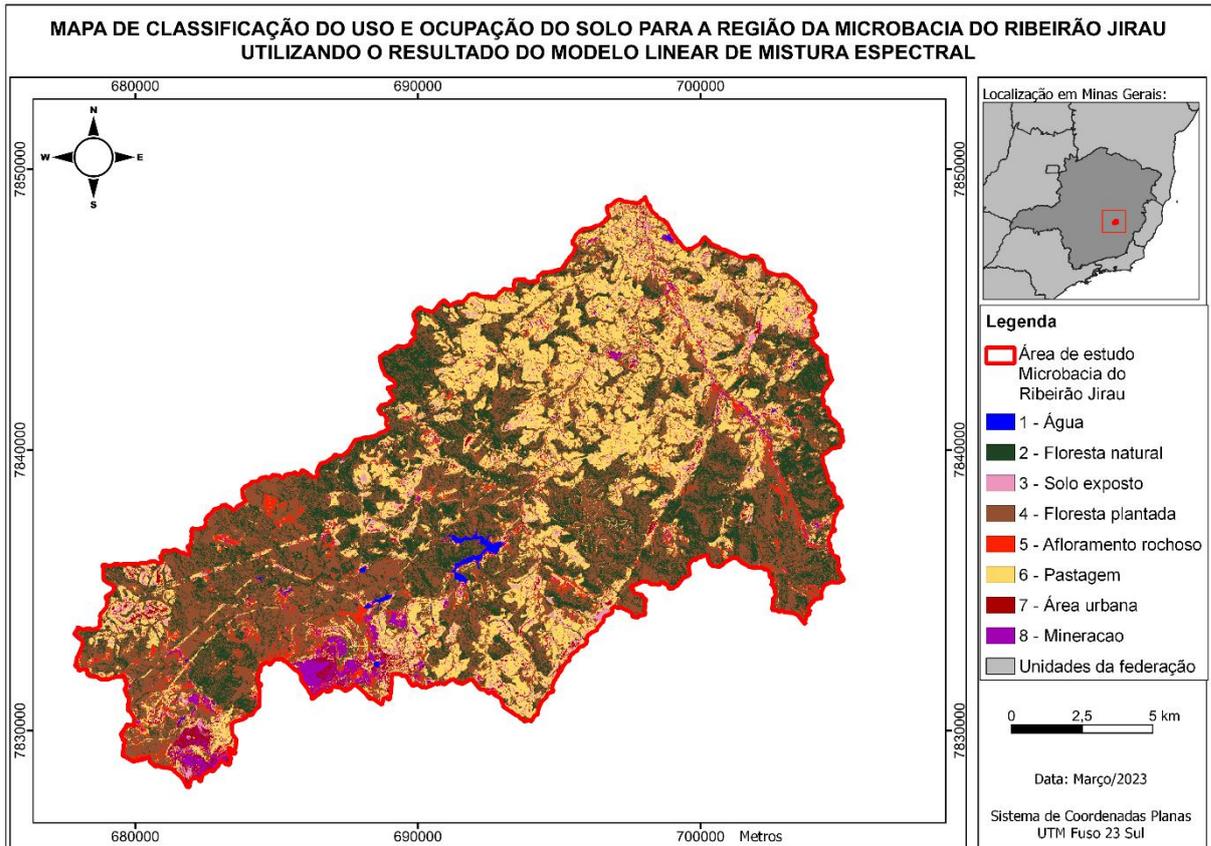


Figura 25. Mapa de uso e ocupação do solo gerado a partir do resultado do MLME.

Como pode ser observado na Figura 24 e na Figura 25, todas as classes consideradas foram mapeadas nos dois processamentos. No entanto, é possível notar diferenças no quantitativo de regiões classificadas para cada categoria. Esta consideração é detalhada no Quadro 8, onde apresenta-se uma relação das áreas (em hectare) obtidas para cada categoria de classificação.

Quadro 8. Relação de áreas para cada categoria classificada.

Classe	Código	Classificação imagens fração do MLME		Classificação cenas originais do Sentinel-2		Diferença (%)
		Área (ha)	Porcentagem	Área (ha)	Porcentagem	
Água	1	115,03	0,41%	117,52	0,42%	0,01
Floresta natural	2	6.171,65	21,83%	9.597,52	33,95%	12,12
Solo exposto	3	704,79	2,49%	266,69	0,94%	1,55
Floresta plantada	4	10.928,18	38,66%	7.226,94	25,57%	13,09
Afloramento rochoso	5	1.183,01	4,19%	1.098,15	3,88%	0,31
Pastagem	6	7.645,86	27,05%	8.081,83	28,59%	1,54
Área urbana	7	642,33	2,27%	632,00	2,24%	0,03
Mineração	8	875,65	3,10%	1.245,85	4,41%	1,31

A partir da relação de áreas obtidas para cada uma das classificações originárias dos dois dados utilizados (Sentinel-2 e MLME) é possível analisar cada categoria individualmente e quantitativamente. Observa-se que a categoria com maior representatividade de área no dado originário do MLME é de floresta plantada, representando cerca de 38% do total classificado. Já nos dados originários do Sentinel-2 esta mesma consideração é observada para a classe de floresta natural que representou cerca de 33% do total classificado.

As demais categorias analisadas (água, solo exposto, afloramento rochoso, pastagem, área urbana e mineração) apresentaram menor divergência de área entre as duas classificações obtidas (Quadro 8).

A partir do arquivo raster referente aos dois mapas gerados foi realizado o procedimento de análise de precisão para as duas classificações. Como resultado, obteve-se as matrizes de confusão apresentadas no Quadro 10 e Quadro 11. Ressalta-se que para esta análise, deve-se considerar os códigos de classes indicados no Quadro 9.

Quadro 9. Códigos representativos das classes.

Código	Classe
1	Água
2	Floresta natural
3	Solo exposto
4	Floresta plantada
5	Afloramento rochoso
6	Pastagem
7	Área urbana
8	Mineração

Quadro 10. Matriz de confusão gerada a partir da análise de precisão do mapa gerado a partir das cenas originais do Sentinel-2.

		Valores Classificados								Total	Precisão do Usuário
		1	2	3	4	5	6	7	8		
Classes do raster temático	1	3	0	0	0	0	0	0	0	3	1,00
	2	0	186	0	38	0	7	0	0	231	0,81
	3	0	0	5	0	0	1	0	0	6	0,83
	4	0	91	0	66	2	14	0	1	174	0,38
	5	2	4	2	4	6	4	0	4	26	0,23
	6	0	1	10	3	1	166	4	9	194	0,86
	7	0	0	1	1	0	3	2	8	15	0,13
	8	0	1	5	0	1	8	1	14	30	0,47
Total		5	283	23	112	10	203	7	36	679	Precisão geral: 0,65
Precisão do Produtor		0,60	0,66	0,22	0,59	0,60	0,82	0,29	0,39		Kappa: 0,53

Quadro 11. Matriz de confusão gerada a partir da análise de precisão do mapa gerado a partir das imagens fração do MLME.

	Valores Classificados								Total	Precisão do Usuário
	1	2	3	4	5	6	7	8		
Classes do raster temático	1	3	0	0	0	0	0	0	3	1,00
	2	0	141	0	15	0	1	0	157	0,90
	3	0	1	13	1	0	0	2	18	0,72
	4	0	172	1	78	0	27	0	278	0,28
	5	1	7	0	5	8	7	0	30	0,27
	6	0	1	3	1	0	186	1	194	0,96
	7	0	0	5	0	0	5	5	16	0,31
	8	0	0	6	0	0	8	0	22	0,36
Total	4	322	28	100	8	234	8	14	718	Precisão geral: 0,62
Precisão do Produtor	0,75	0,44	0,46	0,78	1,00	0,79	0,63	0,57		Kappa: 0,62

A precisão geral para as classificações originárias dos dados do Sentinel-2 e do MLME foi de 65% e 62%, respectivamente. Esta informação revela a proporção de dados classificados corretamente e indica um melhor desempenho para os dados classificados a partir do Sentinel-2. Já o coeficiente *kappa* apresentou resultado de 0,53 para a classificação obtida a partir dos dados espectrais do satélite e 0,62 para a classificação obtida a partir do modelo de mistura. Na Figura 26 e Figura 27 apresenta-se gráficos ilustrativos da análise de precisão geral de classificação realizada.

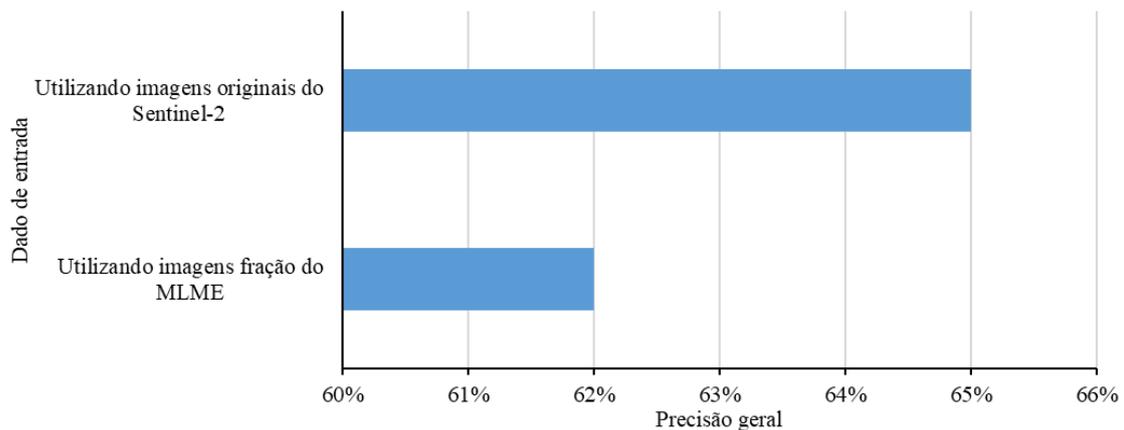


Figura 26. Comparação entre os resultados de precisão geral.

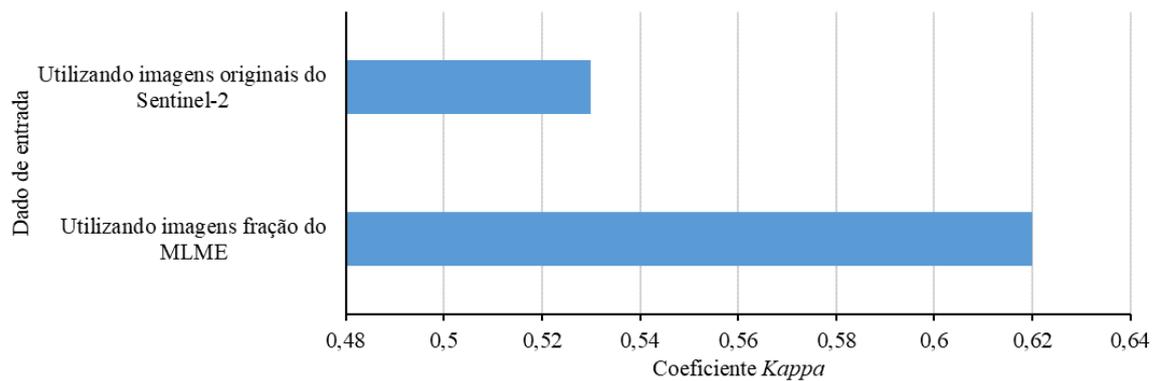


Figura 27. Comparação entre os resultados do coeficiente *kappa*.

Considerando que o valor do coeficiente *kappa* representa a concordância do dado classificado com o dado de referência, percebe-se que o dado classificado a partir do MLME apresentou maior concordância, pois, encontra-se mais próximo do valor 1 da escala de referência utilizada (de 0 a 1 sendo 0 nenhuma concordância e 1 concordância completa). Para interpretação deste valor, comparou-se os valores obtidos com as referências citadas no Quadro 7. Tendo como base as considerações apresentadas por Congalton e Green (2019) ambas as classificações apresentam concordância moderada. Para Fleis, Cohen e Everitt (1969), suficientemente boa e para Monserud e Leemans (1992) como boa (Quadro 12).

Quadro 12. Interpretação do coeficiente *kappa* conforme autores.

Dado utilizado para o mapeamento	Coeficiente <i>kappa</i>	Classificação conforme Congalton e Green (2019)	Classificação conforme Fleis, Cohen e Everitt (1969)	Classificação conforme Monserud e Leemans (1992)
Sentinel-2	0,53	Moderado	Suficientemente bom	Bom
MLME	0,62			

Analisando individualmente as classes consideradas para o resultado obtido a partir dos dados multiespectrais do satélite Sentinel-2 verifica-se que para a classe de Água (1) o valor de unidade de amostra de referência foi de 3 amostras e todas elas foram confirmadas com os dados de referência resultando em 100% de acurácia do usuário. No entanto, esta classificação gerou 5 unidades de amostras classificadas com este código, revelando que duas das áreas classificadas a partir dos dados de sensoriamento remoto como Água (1) eram, na verdade, áreas de afloramento rochoso conforme observações de referência. Assim, para a categoria de água a análise resultou em 60% de acurácia do produtor, tendo 40% das unidades amostrais classificadas como Água (1) pertencentes a outra categoria.

Para a classe de Floresta Natural (2) foram avaliadas 231 unidades amostrais. Destas, 186 foram corretamente classificadas e um total de 283 foram consideradas nesta categoria

conforme o resultado da classificação do dado de sensoriamento remoto resultando em 66% de acurácia do produtor (unidades amostrais classificadas corretamente) e 81,5% de acurácia do usuário (proporção de unidades amostrais de floresta natural na classificação que são desta categoria conforme dados de referência). Para a categoria de Floresta Natural (2) observou-se que 97 unidades amostrais do mapa de classificação foram incluídas nesta categoria (erro de comissão) e 45 unidades de amostra consideradas pelo dado de referência como sendo pertencentes a esta categoria foram classificadas como sendo de Floresta Plantada (4) (38 unidades de amostra) e Pastagem (6) (7 unidades de amostra), revelando assim os erros de omissão.

Para a classe de Solo Exposto (Cód. 3) foram utilizadas 6 unidades de amostras. Destas, 5 foram classificadas corretamente e uma incluída na classe de Pastagem (Cód. 6). Nesta categoria o mapa de classificação considerou 23 unidades amostrais, sendo: 2 unidades amostrais de Afloramento Rochoso (Cód. 5), 10 de Pastagem (Cód. 6), 1 de Área Urbana (Cód. 7) e 5 de Mineração (Cód. 8). Esta categoria teve 22% de precisão do produtor, mostrando baixa relação de unidades amostrais classificadas corretamente e 83% de precisão do usuário.

Para a classe de Floresta Plantada (Cód. 4) 66 das 174 unidades amostrais consideradas foram classificadas corretamente, resultando em 38% de precisão do usuário. Já nos valores classificados 112 unidades amostrais foram incluídas nesta categoria, obtendo assim 59% de precisão do produtor. De acordo com os valores classificados e os dados de referência, 38 unidades de amostra de Floresta Natural (Cód. 2), 4 de Afloramento Rochoso (Cód. 5), 3 de Pastagem (Cód. 6) e 1 de Área Urbana (Cód. 7) foram incluídas na categoria de Floresta Plantada (Cód. 4). Considerando os erros de omissão, observou-se 2 amostras da categoria de Afloramento Rochoso (Cód. 5), 14 de Pastagem (Cód. 6) e 1 de Mineração (Cód. 8).

Para a classe de Afloramento Rochoso (Cód. 5) foram utilizadas 26 unidades amostrais das quais 6 foram corretamente classificadas. A categoria obteve 23% de precisão do usuário e 60% de precisão do produtor. De acordo com a classificação, 2 amostras de Floresta Plantada (Cód. 4), 1 de Pastagem (Cód. 6) e 1 de Mineração (Cód. 8) foram incluídas nesta categoria (erros de comissão) e, considerando as amostras de referência, foram omitidas da classe de afloramento rochoso 2 amostras classificadas como Água (Cód. 1), 4 amostras classificadas como Floresta Natural (Cód. 2), 2 amostras classificadas como Solo Exposto (Cód. 3), 4 classificadas como Floresta Plantada (Cód. 4), 4 classificadas como Pastagem (Cód. 6) e 4 classificadas como Mineração (Cód. 8).

Para a categoria de Pastagem (Cód. 6) foram utilizadas 194 unidades amostrais. Destas, 166 foram classificadas corretamente enquanto que 1 foi incluída na classe de Floresta Natural

(Cód. 2), 10 na classe de Solo Exposto (Cód. 3), 3 na classe de Floresta Plantada (Cód. 4), 1 na classe de Afloramento Rochoso (Cód. 5), 4 na classe de Área Urbana (Cód. 7) e 9 na classe de Mineração (Cód. 8). No resultado da classificação foram incluídas nesta categoria 7 unidades amostrais de Floresta Natural (Cód. 2), 1 de Solo Exposto (Cód. 3), 14 de Floresta Plantada (Cód. 4), 4 de Afloramento Rochoso (Cód. 5), 3 de Área Urbana (Cód. 7) e 8 de Mineração (Cód. 8), resultando em 203 amostras do mapa classificadas nesta categoria. A classe de Pastagem (Cód. 6) obteve 82% de precisão do produtor e 86% de precisão do usuário.

Para a classe de Área Urbana (Cód. 7) foram consideradas 15 unidades amostrais das quais 2 foram classificadas corretamente enquanto que as demais eram, conforme dados de referência: 1 Solo Exposto (Cód. 3), 1 Floresta Plantada (Cód. 4), 3 Pastagem (Cód. 6) e 8 Mineração (Cód. 8). A classificação incluiu na categoria de Área Urbana (Cód. 7) 4 amostras de Pastagem (Cód. 6) e 1 de Mineração (Cód. 8). Esta categoria obteve 13% de precisão do usuário e 29% de precisão do produtor.

Por fim, para a classe de Mineração (Cód. 8) foram consideradas 30 unidades amostrais sendo 14 classificadas corretamente. As demais amostras eram referentes às classes de Floresta Natural (Cód. 2) (1 amostra), Solo Exposto (Cód. 3) (5 amostras), Afloramento Rochoso (Cód. 5) (1 amostra), Pastagem (Cód. 6) (8 amostras) e Área Urbana (Cód. 7) (uma amostra). O resultado da classificação incluiu na classe de Mineração (Cód. 8) 1 amostra de Floresta Plantada (Cód. 4), 4 de Afloramento Rochoso (Cód. 5), 9 de Pastagem (Cód. 6) e 8 de Área Urbana (Cód. 7). Para esta categoria obteve-se 47% de precisão do usuário e 39% de precisão do produtor.

Buscando facilitar a visualização dos resultados de precisão do usuário e do produtor obtidos a partir do mapeamento realizado com as bandas espectrais do Sentinel-2 os valores aqui apresentados foram graficamente demonstrados na Figura 28 e Figura 29.

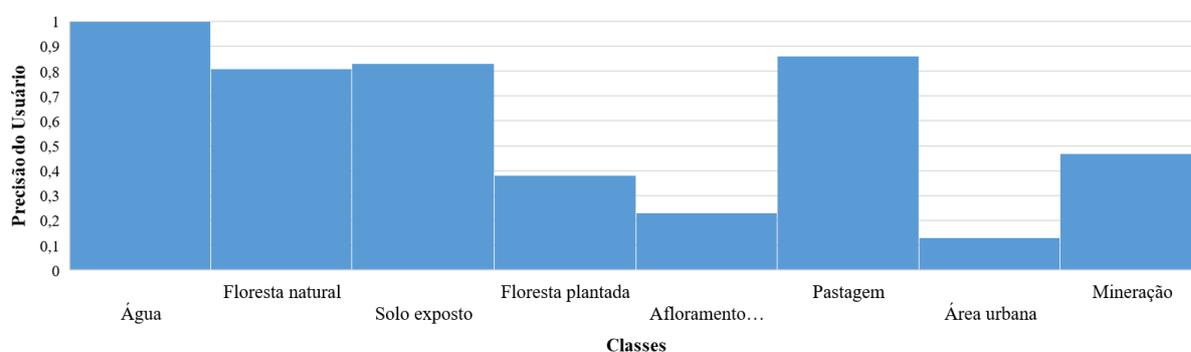


Figura 28. Representação dos resultados de precisão do usuário obtidos a partir do mapeamento realizado com as cenas espectrais originais do Sentinel-2.

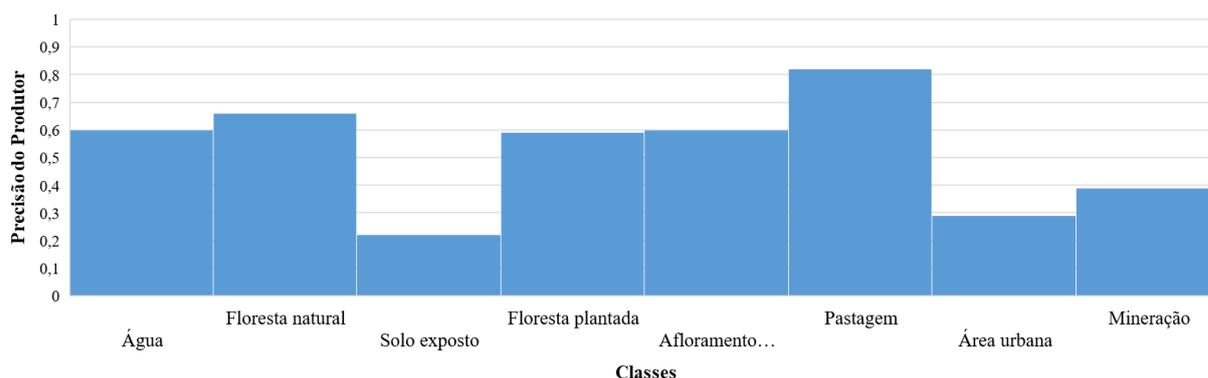


Figura 29. Representação dos resultados de precisão do produtor obtidos a partir do mapeamento realizado com as cenas espectrais originais do Sentinel-2.

Analisando individualmente as classes consideradas para o resultado obtido a partir dos dados do modelo de mistura espectral (imagens fração) verifica-se que para a classe de Água (Cód. 1) também foram utilizadas 3 unidades amostrais sendo todas elas corretamente classificadas e, portanto, com precisão do usuário de 100%. No entanto, no arquivo classificado foi incluída nesta categoria 1 amostra de Afloramento Rochoso (Cód. 5), gerando assim 75% de precisão do produtor.

Para a categoria de Floresta Natural (Cód. 2) foram utilizadas 157 unidades amostrais sendo 141 classificadas corretamente, 15 omitidas e incluídas na classe de Floresta Plantada (Cód. 4) e uma na classe de Pastagem (Cód. 6) resultando em uma precisão do usuário de 90% para a categoria. Os resultados da classificação incluíram nesta classe 1 amostra de Solo Exposto (Cód. 3), 172 amostras de Floresta Plantada (Cód. 4), 7 amostras de Afloramento Rochoso (Cód. 5) e 1 amostra de Pastagem (Cód. 6), resultando em 322 unidades amostrais incluídas nesta categoria conforme o resultado da classificação e 44% de precisão do produtor.

Para a categoria de Solo Exposto (Cód. 3) foram utilizadas 18 unidades amostrais das quais 13 foram corretamente classificadas, 1 omitida para a classe de Floresta Natural (Cód. 2), 1 para Floresta Plantada (Cód. 4), 2 para Área Urbana (Cód. 7) e 1 para Mineração (Cód. 8). Foram incluídas nesta categoria, conforme dados classificados, 1 amostra de Floresta Plantada (Cód. 4), 3 amostras de Pastagem (Cód. 6), 5 de Área Urbana (Cód. 7) e 6 de Mineração (Cód. 8). Para esta categoria obteve-se 72% de precisão do usuário e 46% de precisão do produtor.

Para a categoria de Floresta Plantada (Cód. 4) foram consideradas 278 unidades amostrais sendo 78 classificadas corretamente. 172 amostras desta classe foram confundidas com Floresta Natural (Cód. 2), 1 com Solo Exposto (Cód. 3) e 27 com Pastagem (Cód. 6). Foram incluídas nesta categoria 15 amostras de Floresta Natural (Cód. 2), 1 de Solo Exposto (Cód. 3), 5 de Afloramento Rochoso (Cód. 5) e 1 de Pastagem (Cód. 6). Esta categoria resultou em 28% de precisão do usuário e 78% de precisão do produtor.

Para a categoria de Afloramento Rochoso (Cód. 5) foram consideradas 30 unidades amostrais sendo 8 classificadas corretamente e não foram incluídas amostras de outras classes nesta categoria resultando em 100% de precisão do produtor. No entanto, outras unidades amostrais de afloramento rochoso foram classificadas no mapeamento como Floresta Plantada (Cód. 4) (5 unidades), Floresta Natural (Cód. 2) (7 unidades), Água (Cód. 1) (1 unidade), Pastagem (Cód. 6) (7 unidades) e Mineração (Cód. 8) (2 unidades), resultando em 27% de precisão do usuário.

Para a categoria de Pastagem (Cód. 6) foram consideradas 194 unidades amostrais sendo 186 classificadas corretamente. As demais unidades foram confundidas com as classes de Floresta Natural (Cód. 2) (1 unidade), Solo Exposto (Cód. 3) (3 unidades), Floresta Plantada (Cód. 4) (1 unidade), Área Urbana (Cód. 7) (1 unidade) e Mineração (Cód. 8) (2 unidades). Foram incluídas nesta categoria 1 unidade de Floresta Natural (Cód. 2), 27 de Floresta Plantada (Cód. 4), 7 de Afloramento Rochoso (Cód. 5), 5 de Área Urbana (Cód. 7) e 8 de Mineração (Cód. 8). Esta categoria resultou em 96% de precisão do usuário e 79% de precisão do produtor.

Para a categoria de Área Urbana (Cód. 7) foram consideradas 16 unidades amostrais sendo 5 classificadas corretamente. As demais amostras foram classificadas como Solo Exposto (Cód. 3) e Pastagem (Cód. 6), sendo 5 unidades cada e Mineração (Cód. 8) com 1 unidade. Foram incluídas nesta categoria 2 unidades amostrais de Solo Exposto (Cód. 3) e 1 de Pastagem (Cód. 6). Obteve-se 31% de precisão do usuário e 63% de precisão do produtor.

Por fim, para a categoria de Mineração (Cód. 8) foram consideradas 22 unidades amostrais sendo 8 delas classificadas corretamente e as demais confundidas com as categorias de Solo Exposto (Cód. 3) (6 unidades) e Pastagem (Cód. 6) (8 unidades). Além disso, foram incluídas nesta categoria 1 unidade amostral de Solo Exposto (Cód. 3) e de Área Urbana (Cód. 7) e 2 unidades de Afloramento Rochoso (Cód. 5) e de Pastagem (Cód. 6). Obteve-se 36% de precisão do usuário e 57% de precisão do produtor.

Os dados das precisões do usuário e do produtor do mapeamento realizado com as imagens fração do MLME também foram reunidos em gráficos para melhor visualização dos resultados (Figura 30 e Figura 31).

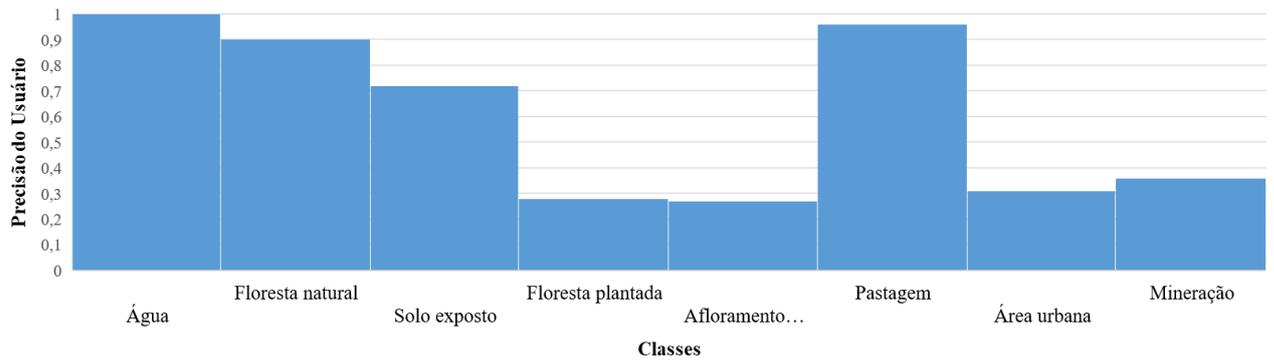


Figura 30. Representação dos resultados de precisão do usuário obtidos a partir do mapeamento realizado com as imagens fração do MLME.

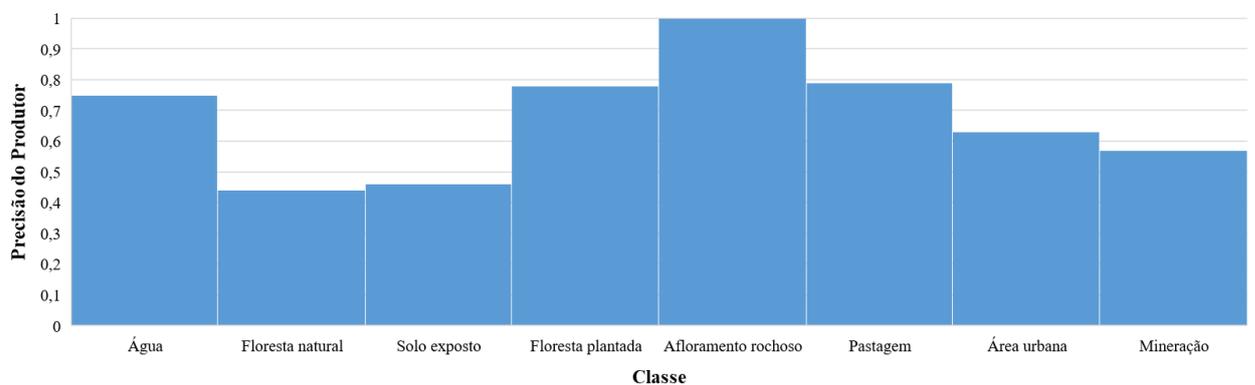


Figura 31. Representação dos resultados de precisão do produtor obtidos a partir do mapeamento realizado com as imagens fração do MLME.

Buscando ainda melhorar a comparação de ambos os mapeamentos na Figura 32 e Figura 33 apresenta-se as precisões dos dois resultados juntos tornando possível a visualização das diferenças identificadas entre o resultado gerado a partir das cenas espectrais e das imagens fração.

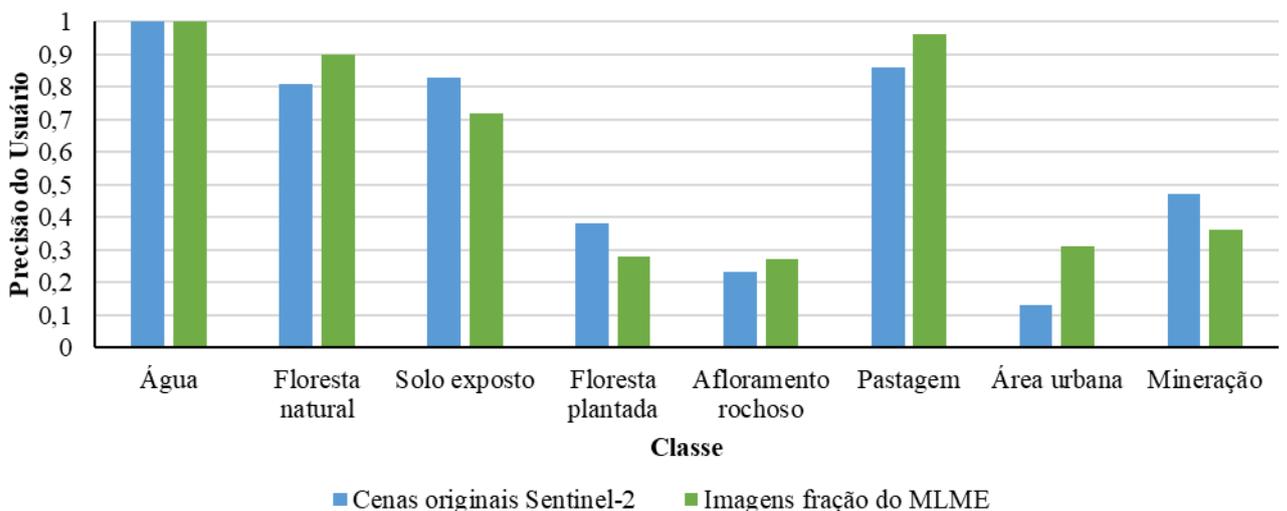


Figura 32. Comparação entre os resultados de precisão do usuário.

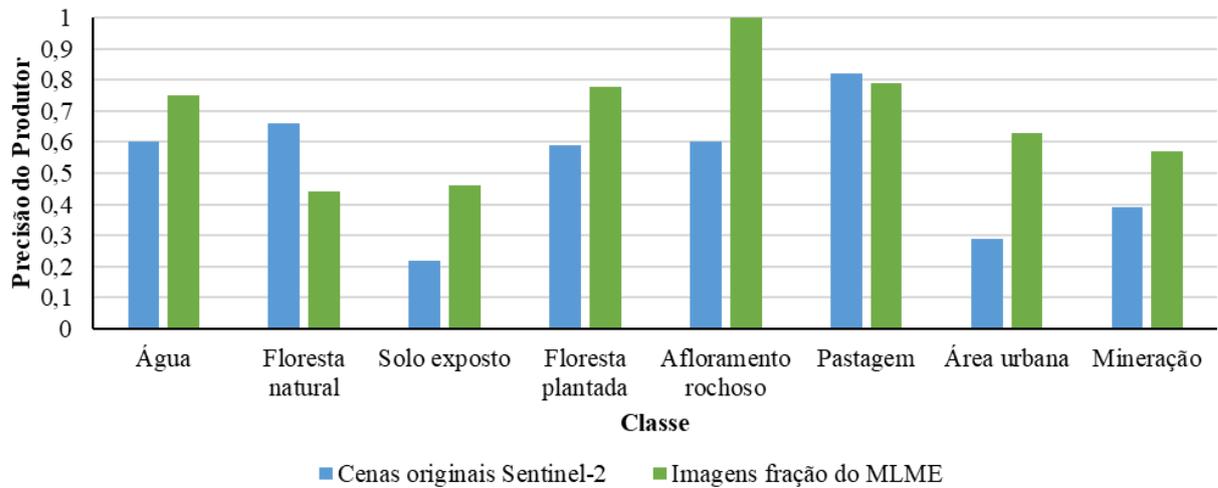


Figura 33. Comparação entre os resultados de precisão do produtor.

Para complementar a análise dos resultados obtidos, destaca-se nas figuras a seguir alguns pontos observados durante o estudo dos mapeamentos elaborados com destaque para as classificações e imagem de referência.

Na Figura 34 houve a seleção de um pixel em área de floresta natural onde a classificação gerada a partir das imagens fração do MLME resultou em 100% de proporção de vegetação. Nesta figura destaca-se ainda a maior proporção de área considerada como sendo de floresta plantada no mapeamento gerado a partir do MLME sendo a área de floresta natural.

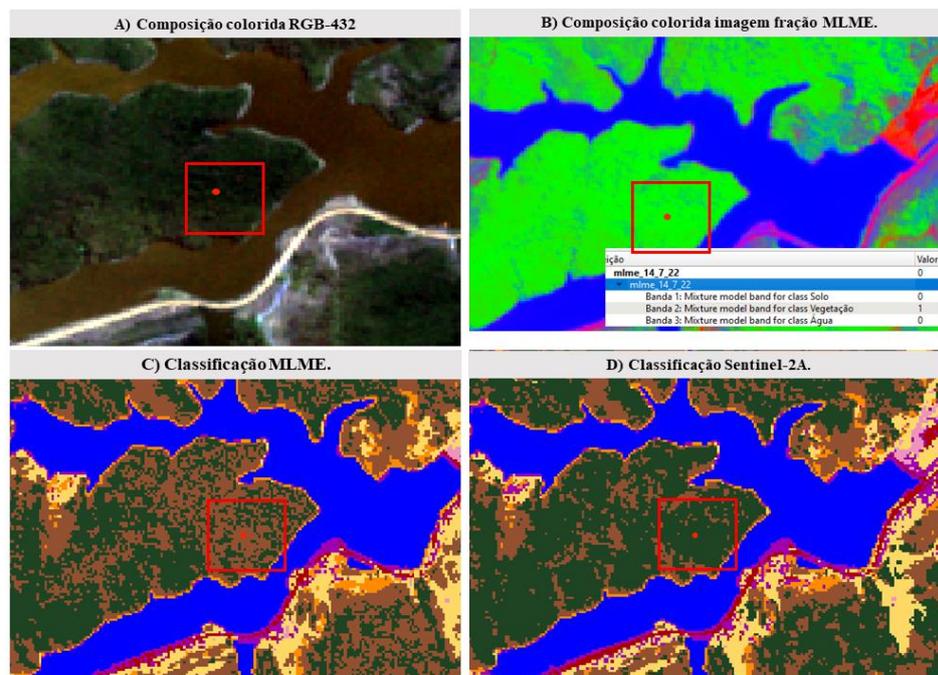


Figura 34. Comparação dos mapeamentos de área de floresta natural cujo MLME resultou em 100% de proporção de vegetação. Destaca-se a maior proporção de área considerada como sendo de floresta plantada no mapeamento gerado a partir do MLME sendo a área de floresta natural.

Na Figura 35 houve a seleção de um pixel em área de solo exposto onde obteve-se as 0,88 e 0,12 para as classes de solo e de vegetação, respectivamente. Para esta região, destaca-se que o mapeamento gerado a partir do resultado do MLME resultou em uma maior abrangência de pixels de solo exposto na área.

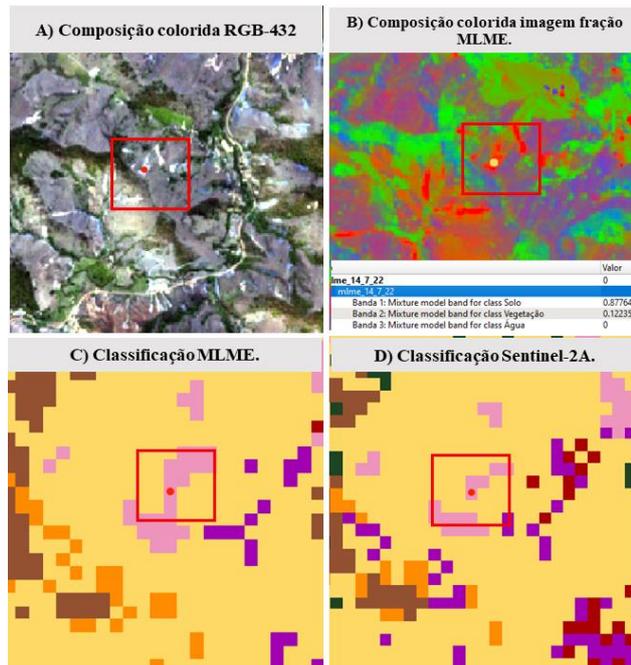


Figura 35. Comparação dos mapeamentos de área de solo exposto cujo MLME resultou divisão entre as proporções de solo e vegetação e gerou como classificação final a categoria de solo (de maior proporção no pixel selecionado). Destaca-se o mapeamento do MLME com maior abrangência de pixels de solo exposto.

Na Figura 36 houve a seleção de um pixel localizado também em uma área de solo exposto, com 100% de abundância de solo conforme análise do MLME. Nesta figura, destaca-se o mapeamento gerado a partir das cenas originais do Sentinel-2 que apresentou maior uniformidade da área de solo e o mapeamento gerado a partir do resultado do MLME com presença de pixels das classes de solo, área urbana e mineração.

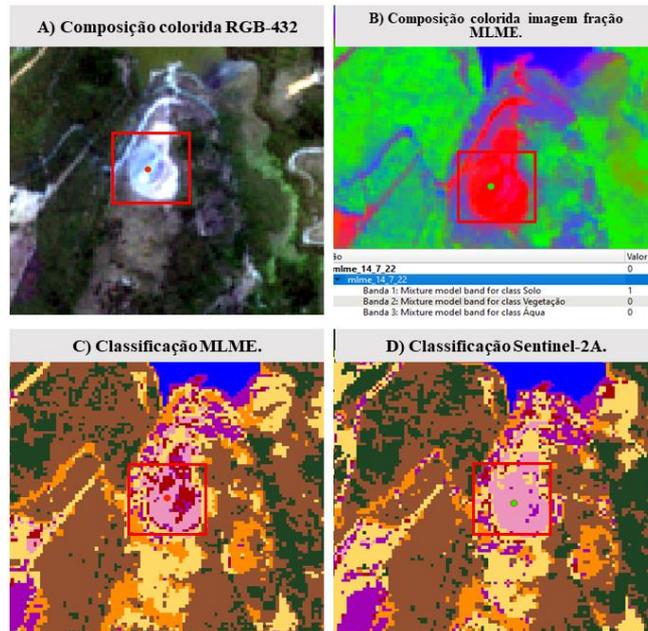


Figura 36. Comparação dos mapeamentos de área de solo exposto cujo MLME 100% de solo para o pixel selecionado. Destaca-se para a região o mapeamento do Sentinel-2 que apresentou maior uniformidade da área de solo sendo o mapeamento do MLME com mistura entre as classes de solo, área urbana e mineração.

5. DISCUSSÃO

As imagens fração obtidas a partir do modelo linear de mistura espectral resultaram em arquivos digitais que destacaram adequadamente as áreas dos componentes analisados no modelo (solo, vegetação e água/sombra), deixando estas regiões apresentadas com pixels de maior abundância e tonalidades mais claras em cada imagem fração correspondente.

Ao observar o arquivo de composição das três imagens fração obtidas no modelo verifica-se ainda que o processamento utilizado de fato resultou em frações de composição de cada componente considerado, podendo ser possível identificar aquele de maior predominância dentro do pixel conforme exemplificado a partir da seleção de pixels na seção 4. Além disso, nas áreas em que houve a seleção de pixels puros para o MLME não se identificou mistura espectral nos pixels selecionados.

Os dois tipos de dados considerados para a classificação do uso e ocupação do solo mostraram-se funcionais para o processamento realizado, no entanto, os seus respectivos resultados evidenciam, até mesmo com uma breve análise visual, diferenças de classificação ao comparar os dois mapeamentos. Esta divergência é notória, principalmente, para as classes de floresta plantada e de floresta natural dado que as respectivas representatividades de ocupação da área revelam maiores diferenças, tendo o resultado originário do Sentinel-2 maior representatividade da classe de floresta natural (33,95%) e o resultado originário do MLME maior representatividade da classe de floresta plantada (38,66%).

Ressalta-se que apenas a análise quantitativa de representação do total de área classificada não é suficiente para se obter considerações dos dois dados de entrada utilizados para o processo de classificação. Assim, a análise dos coeficientes de concordância obtidos a partir da matriz de erros permitiu investigar profundamente a precisão dos dois mapeamentos, em condições gerais e específicas para cada classe.

A precisão geral para as classificações originárias dos dados do Sentinel-2 e do MLME foi de 65% e 62%, respectivamente. Esta informação revela a proporção de dados classificados corretamente e indica um melhor desempenho para os dados classificados a partir do Sentinel-2. Segundo Osgouei *et al.* (2019) as classificações originárias de imagens de satélite apresentam desempenhos melhores, com precisão acima de 85%, quando se trabalha apenas com as principais classes de cobertura do solo como vegetação, solo e água. Ao se considerar outras classes que contemplam regiões mais heterogêneas, como é o caso deste estudo, obter melhores precisões torna-se um desafio visto que as semelhanças espectrais e espaciais e a heterogeneidade das áreas dificulta a classificação. Os desafios são justificados ainda pela deficiência na obtenção de amostras de treinamento, elevando as dificuldades de classificação de dados de sensoriamento remoto independente do algoritmo que se esteja fazendo o uso.

O coeficiente *kappa* apresentou resultado de 0,53 para a classificação obtida a partir dos dados espectrais do satélite e 0,62 para a classificação obtida a partir do modelo de mistura. Considerando que o valor deste coeficiente representa a concordância do dado classificado com o dado de referência percebe-se que o dado classificado a partir do MLME apresentou maior concordância, pois, encontra-se mais próximo do valor 1 da escala de referência utilizada (de 0 a 1 sendo 0 nenhuma concordância e 1 concordância completa). Cabe ressaltar que no cálculo do coeficiente *kappa* considera-se uma maior representatividade dos dados disponibilizados na matriz de erros do que no cálculo da precisão geral. Conforme a referência de interpretação do *kappa* apresentada no Quadro 7 as duas classificações apresentam resultado de qualidade moderada, suficientemente boa ou boa.

De acordo com as duas matrizes de erros obtidas, a classificação gerada a partir das bandas espectrais originais do Sentinel-2 gerou melhor desempenho de precisão do usuário para as classes de água (100%), floresta natural (81%), solo exposto (83%) e pastagem (86%), assim, estas são as probabilidades de um pixel classificado no mapeamento de fato representar estas respectivas categorias conforme análise do dado de referência. A classe que apresentou menor desempenho foi a de área urbana (13%), que teve parte de suas amostras classificadas principalmente como mineração (aproximadamente 53%) sendo esta seguida das classes de

afloramento rochoso (23%) e floresta plantada (38%) que representou grande confusão com a classe de floresta natural.

Considerando a precisão do produtor, que faz relação com as amostras que não foram corretamente classificadas como pertencendo à categoria analisada e omitidas de suas devidas categorias, este mapeamento revelou melhor desempenho para a classe de pastagem (82%) e piores desempenhos para as classes de solo exposto (22%) e novamente de área urbana (29%).

Analisando a matriz de erros obtida para a classificação originada do resultado do modelo linear de mistura espectral, a classe de água novamente se destaca com 100% de precisão do usuário, seguida pelo bom desempenho das classes de pastagem (96%) e de floresta natural (90%). Observou-se que este mapeamento não apresentou desempenho satisfatório para a classe de floresta plantada já que a precisão do usuário para esta categoria foi de 28% devido ao fato de 62% das amostras de validação classificadas como floresta plantada eram, na verdade, floresta natural conforme análise do dado de referência. Notou-se baixa precisão do usuário também para a categoria de afloramento rochoso (27%), indicando confusão principalmente com as classes de floresta plantada, floresta natural e de pastagem. Para esta categoria, considera-se a possibilidade de erros do analista durante a validação pois, as áreas de afloramento rochoso são normalmente recobertas por vegetação e este fato pode ter influenciado durante a análise das amostras.

Neste mapeamento, a precisão do produtor resultou em maior valor para a classe de afloramento rochoso (100%), já que o mapeamento não incluiu nenhuma outra amostra de outra categoria como sendo de afloramento rochoso, pastagem com 79% e floresta plantada com 78%.

Nos mapas de classificação e quadros do quantitativo de área e de representatividade de cada classe considerada foi verificada a existência de relevante divergência entre as classes de floresta natural e de floresta plantada. A análise da matriz de erros revelou que o mapeamento gerado a partir dos dados do modelo de mistura resultou em melhor classificação para a categoria de floresta plantada pois houve uma menor representatividade de amostras de outra categoria que foram incluídas nesta classificação. No entanto, 62% das amostras consideradas para a validação foram confundidas com a classe de floresta natural.

Para a categoria de floresta natural a matriz de erros revelou que o mapeamento gerado a partir dos dados do modelo de mistura resultou em melhor desempenho de classificação já que 90% das amostras analisadas realmente pertenciam a esta categoria (precisão do usuário) enquanto que no mapeamento gerado a partir do dado de satélite esta análise resultou em 81%.

No entanto, o mapeamento obtido nos dados espectrais do Sentinel-2 incluiu uma menor representatividade de outras categorias nesta classificação.

Para a categoria de floresta plantada, observou-se que o melhor desempenho foi para o dado originário das cenas de satélite (38% das amostras classificadas corretamente), enquanto para o segundo mapeamento essa precisão foi de 28%. Cabe ressaltar que esta categoria apresentou maior confusão tendo a maior parte das amostras analisadas omitidas desta categoria e incluídas na categoria de floresta natural para os dois resultados.

Segundo Zanotta *et al.* (2013), os erros de classificação obtidos a partir das imagens fração podem estar associados aos *endmembers* selecionados para o modelo já que o resíduo do MLME pode ser representativo também de membros finais ausentes. A qualidade da classificação obtida para ambos os dados de entrada pode ser influenciada também pela deficiência das amostras de referência e de validação já que todo o processo foi desenvolvido tendo como base a composição colorida da imagem multiespectral.

Cabe destacar também a categoria de área urbana que normalmente apresenta desafios para a sua adequada classificação pois a heterogeneidade característica desta área dificulta o desempenho dos classificadores e, para o mapeamento obtido a partir dos resultados do MLME esta categoria apresentou maior desempenho de classificação quanto comparada com os dados das cenas dos satélites, sendo 31% e 13% das amostras classificadas corretamente, respectivamente. O mapeamento do modelo também se mostrou mais eficiente ao incluir menos amostras de outras categorias nesta classe.

Contudo, cabe ressaltar que a qualidade do mapeamento para cada categoria individualizada gerado a partir de dados espectrais pode ser influenciada ainda pelo tipo de componente analisado uma vez que alguns objetos apresentam respostas bem específicas (OSGOUEI *et al.*, 2019).

6. CONCLUSÃO

Os resultados obtidos neste trabalho demonstram a funcionalidade da aplicação de imagens fracionadas derivadas de um modelo linear de mistura espectral como subsídio à classificação do uso e ocupação do solo sendo uma complementação para fase inicial de pré-processamento para posterior aplicação de técnicas apropriadas de classificação, agrupamentos de dados, fatiamentos, limiarizações e outros. Além disso, o estudo enfatiza o uso de dados gratuitos e programas computacionais de acesso livre que permitiram realizar todos os procedimentos apresentados neste trabalho.

O resultado do modelo de mistura que é fornecido através de imagens fração de cada componente considerado para a modelagem indica que de fato a metodologia evidencia os

componentes da área de estudo a partir da abundância do componente obtido em cada elemento de área. Neste estudo foi possível observar que as áreas de vegetação, solo e água foram destacadas em suas respectivas imagens fração.

Ambos os tipos de dados de entrada utilizados para a classificação do uso e ocupação do solo resultaram em mapeamentos de qualidade moderada, suficientemente boa ou boa conforme avaliação do coeficiente *kappa*. Verificou-se que as análises de precisão geral não apresentaram grandes variações entre os dois resultados, diferentemente do que foi observado para as precisões individuais de cada categoria.

Neste estudo é possível avaliar que o problema da mistura espectral pode então ser reduzido a partir do uso de sensores de melhor resolução espacial e a compreensão da representação das cenas pode ser melhorada a partir do uso de técnicas de tratamento de imagens, no entanto, mesmo com a melhoria da resolução espacial dos sensores a mistura espectral continuará existindo.

Assim, o uso de técnicas de decomposição espectral, como o modelo linear de mistura espectral, são potenciais ferramentas para auxiliar os processos de classificação de imagens digitais. Salienta-se que a classificação detalhada com mais categorias, como a realizada neste trabalho, pode ter sua eficiência prejudicada levando a recomendar a aplicação do modelo para a classificação de um menor número de classes e de classes gerais como solo, vegetação e água/sombra, sendo este resultado utilizado para detalhamento da classificação em outras categorias. Além disso, o MLME não se mostrou eficiente para discriminação entre os diferentes tipos de vegetação observados na região (floresta natural e floresta plantada), o que leva a sugerir uma segmentação apenas da fração de vegetação para a discriminação das diferentes tipologias de vegetação.

Este estudo pode ser aprimorado a partir do uso de dados de referência coletados em campo ou com imagens digitais de melhor resolução espacial permitindo melhor qualidade durante a validação, da aquisição de maior número de amostras de treinamento e da aplicação de procedimentos de pós classificação e correção vetorial.

Em síntese, o estudo do uso e ocupação do solo é de grande relevância para auxílio em práticas de gestão territorial. A aplicação de imagens digitais de sensoriamento remoto para este tipo de análise contribui para a possibilidade de alcance de grandes áreas territoriais e para a obtenção de informações detalhadas que variam de acordo com a melhoria dos sensores. O modelo linear de mistura espectral mostra-se recomendável para o subsídio de técnicas de classificação do uso e ocupação do solo e se apresenta como uma importante ferramenta para estudo e conciliação com outras técnicas de processamento digital de imagens.

REFERÊNCIAS

- ANA. Agência Nacional de Águas e Saneamento Básico. Bacias Hidrográficas Otto Nível 5. Disponível em: <https://sigaaguas.org.br/portal/home/item.html?id=6d279a59a0a3494997cf5f6ba94c24b7#overview>. Acesso em 25 nov. 2022.
- BECK, L.; LOBITZ, B. M.; WOOD, B. L. *Remote Sensing and Human Health: new sensors and new opportunities. Emerging Infectious Diseases*, v. 6, n. 3, p. 217-227, jun. 2000. Centers for Disease Control and Prevention (CDC).
- CANADÁ. *Remote Sensing Tutorials*. 2019. Tutorial elaborado pelo Governo do Canadá. Disponível em: <https://www.nrcan.gc.ca/maps-tools-and-publications/satellite-imagery-and-air-photos/tutorial-fundamentals-remote-sensing/9309>. Acesso em: 21 mar. 2022.
- CBH-SANTO ANTÔNIO - Comitê de Bacia Hidrográfica do Rio Santo Antônio. A Bacia. Disponível em: <https://www.cbhsantoantonio.org.br/a-bacia>. Acesso em 25 nov. 2022.
- CECCATO, P.; CONNOR, S. J.; JEANNE, E.; THOMSON, M. C. *Application of geographical information systems and remote sensing technologies for assessing and monitoring malaria risk. Parassitologia*, v. 47, n. 1, p. 81-96, 2005.
- CHENG, G.; XIE, X. HAN, J.; GUO, L.; XIA, GS. *Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities. ArXiv*, 2020.
- CONGALTON, R.G.; GREEN, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. 2nd Edition, Lewis Publishers, Boca Raton, 2009.
- CONGALTON, R. G. GREEN, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. 3ª Ed.; Nova York, Boca Raton: CRC Press, 328 p.; 2019.
- CONGEDO, L (2016). Semi-Automatic Classification Plugin Documentation. DOI: <http://dx.doi.org/10.13140/RG.2.2.29474.02242/1>.
- CRISP, Centre for Remote Imaging Sensing and Processing. *Analog and Digital Images*. 2001. Disponível em: <https://crisp.nus.edu.sg/~research/tutorial/image.htm#:~:text=Remote%20sensing%20images%20are%20representations,two-dimensional%20array%20of%20pixels>. Acesso em: 24 mar. 2022.
- CRÓSTA, A. P. Processamento Digital de Imagens de Sensoriamento Remoto. UNICAMP - Instituto de Geociências Departamento de Metalogênese e Geoquímica. Edição Revisada, 1992.
- DUTRA, A. C.; SHIMABUKURO, Y. E.; ARAI, E. *Linear Spectral Mixing Model Applied in Images from PROBA-V SENSOR: a spatial multiresolution approach. Raega - O Espaço Geográfico em Análise, Curitiba*, v. 46, n. 3, p. 48-62, 2019. Universidade Federal do Paraná.
- EMBRAPA. Empresa Brasileira de Pesquisa Agropecuária. Sistema Brasileiro de Classificação de Solos / Humberto Gonçalves dos Santos... [et al.]. – 5. ed., rev. e ampl. – Brasília, DF: Embrapa, 2018. 356 p.: il. color.; 16 cm x 23 cm.
- ESA. *European Space Agency. Copernicus Open Access Hub*. Disponível em: <https://scihub.copernicus.eu/dhus/#/home>. Acesso em 14 jul. 2022.
- _____. *European Space Agency. Sentinel-2 MSI User Guide*. <https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi>. Acesso em 14 jul. 2022.

- _____. *European Space Agency. Sentinel-2*. Disponível em: <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2>. Acesso em 14 jul. 2022.
- FEAM/UFV. Fundação Estadual do Meio Ambiente. Universidade Federal de Viçosa. Mapa de Solos de Minas Gerais. Disponível em: <https://idesisema.meioambiente.mg.gov.br/webgis>. Acesso em 25 nov. 2022.
- FIGUEIREDO, A. V. T. [Sem título]. 2020. Coleção particular.
- FIGUEIREDO, A. V. T. [Sem título]. 2021. Coleção particular.
- FIGUEIREDO, S. M. M.; CARVALHO, L. M. T. de. Avaliação da Exatidão do Mapeamento da Cobertura da Terra em Capixaba, Acre, Utilizando Classificação por Árvore de Decisão. *Cerne, Lavras*, v. 12, n. 1, p. 38-47, mar. 2006.
- FLEISS, J. L.; COHEN, J.; EVERITT, B. S. *Large sample standard errors of kappa and weighted kappa. Psychological Bulletin*, v. 72, n. 5, p. 323-327, nov. 1969. *American Psychological Association (APA)*.
- FOODY, G. M. *Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. Remote Sensing of Environment*, v. 239, p. 111630, mar. 2020. Elsevier BV.
- IBGE – Instituto Brasileiro de Geografia e Estatística. Introdução ao Processamento Digital de Imagens. Manuais Técnicos em Geociências, nº 9. Primeira Divisão de Geociências do Nordeste, Rio de Janeiro, 2001.
- IDE-SISEMA. Infraestrutura de Dados Espaciais do Sistema Estadual de Meio Ambiente e Recursos Hídricos. Disponível em: <https://idesisema.meioambiente.mg.gov.br/webgis>. Acesso em 25 nov. 2022.
- INPE. Modelos de Mistura. Disponível em: <http://www.dpi.inpe.br/spring/portugues/tutorial/modelos.html>. Acesso em: 4 mar. 2022.
- _____. *TerraAmazon Copyright (c) 2005-2018 INPE*. 2021.
- ITABIRA. Plano de Manejo da Área de Proteção Ambiental Municipal Santo Antônio. Encarte 1. Itabira, MG, 2018. 412 p.
- ITABIRA. Sistema de Informação Territorial. 2020.
- GAO, L.; WANG, X.; JOHNSON, B. A.; TIAN, Q.; WANG, Y.; VERRELST, J.; MU, X.; GU, X. *Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review. ISPRS Journal of Photogrammetry and Remote Sensing*, v. 159, p. 364-377, jan. 2020. Elsevier BV.
- GASCON, F.; BOUZINAC, C.; THÉPAUT, O.; JUNG, M.; FRANCESCONI, B.; LOUIS, J.; LONJOU, V.; LAFRANCE, B.; MASSERA, S.; GAUDEL-VACARESSE, A. *Copernicus Sentinel-2 Calibration and Products Validation Status. Remote Sensing*, v. 9, n. 6, p. 584, 10 jun. 2017. MDPI AG.
- KESHAVA, N.; MUSTARD, J. F. *Spectral unmixing. IEEE signal processing magazine*, v. 19, n. 1, p. 44-57, 2002.
- KOULI, M. *Editorial for the Special Issue “Application of Satellite Remote Sensing in Solving Urban Geo-Environmental Issues”*. *Remote Sensing*, v. 15, n. 1, p. 63, 2023.

- LANDIS, J. R.; KOCH, G. G. *The Measurement of Observer Agreement of Categorical Data. Biometrics*, v. 33, n. 1, p. 159–174, 1977.
- LAVENDER, S.; LAVENDER, A. *Practical handbook of remote sensing*. CRC Press, 2023.
- LI, M.; ZANG, S.; ZHANG, B.; LI, S.; WU, C. *A Review of Remote Sensing Image Classification Techniques: the role of spatio-contextual information. European Journal Of Remote Sensing*, v. 47, n. 1, p. 389-411, jan. 2014.
- LLANO, X. C. SMBYc-IDEAM. AcATaMa - QGIS plugin for Accuracy Assessment of Thematic Maps, 2021. Disponível em: <https://github.com/SMBYc/AcATaMa>.
- MAIDJELELE, G. B. A. M.; PACHECO, J. A. A.; DRUMMOND, J. A. L.; SILVA, E. V.; CARLOS, D. A. I.; MARTORANO, L. G. Proposta de planejamento de uso e ocupação do solo a partir de análise espaço-temporal com imagens orbitais do município de Chimoio, Moçambique. *Brazilian Journal of Development*, v. 6, n. 11, p. 84952-84965, 2020.
- MAS, J.-F.; KOLB, M.; PAEGELOW, M.; OLMEDO, M. T. C.; HOUET, T. *Inductive pattern-based land use/cover change models: a comparison of four software packages. Environmental Modelling & Software*, v. 51, p. 94-111, jan. 2014. Elsevier BV.
- MENEZES, P. R.; ALMEIDA, T. *Introdução ao Processamento de Imagens de Sensoriamento Remoto*. Universidade de Brasília, Brasília, 2012.
- MOKHTARI, A.; NOORY, H.; POURSHAKOURI, F.; HAGHIGHATMEHR, P.; AFRASIABIAN, Y.; RAZAVI, M.; FERAYDOONI, F.; NAENI, A. S. *Calculating potential evapotranspiration and single crop coefficient based on energy balance equation using Landsat 8 and Sentinel-2. Isprs Journal Of Photogrammetry And Remote Sensing*, v. 154, p. 231-245, ago. 2019. Elsevier BV.
- MONSERUD, R. A.; LEEMANS, R. *Comparing global vegetation maps with the Kappa statistic. Ecological Modelling*, v. 62, n. 4, p. 275-293, ago. 1992. Elsevier BV.
- MUÑOZ, J. M. S. *Análisis de Calidad Cartográfica mediante el estudio de la Matriz de Confusión. Pensamiento Matemático*, v. 6, n. 2, p. 9-26, 2016.
- NARVAES, I. S.; SANTOS, J. R. A utilização da segmentação de imagens-fração como técnica de classificação da cobertura vegetal na região central e entorno da FLONA do Tapajós, utilizando imagem CCD/CBERS-2. *In: Simpósio Brasileiro de Sensoriamento Remoto*, v. 12, p. 993-1000, 2007.
- NAVALGUND, R. R.; JAYARAMAN, V.; ROY, P. S. *Remote sensing applications: An overview. current science*, p. 1747-1766, 2007.
- NOVO, E. M. L. M.; PONZONI, F. J. *Introdução ao Sensoriamento Remoto*. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2001.
- OSGOUEI, P. E.; KAYA, S.; SERTEL, E.; ALGANCI, U. *Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2 Imagery. Remote Sensing*, v. 11, n. 3, p. 345, 10 fev. 2019. MDPI AG.
- PASQUALOTTO, N.; DELEGIDO, J.; VAN WITTENBERGHE, S.; RINALDI, M.; MORENO, J. *Multi-Crop Green LAI Estimation with a New Simple Sentinel-2 LAI Index (SeLI). Sensors*, v. 19, n. 4, p. 904, 21 fev. 2019. MDPI AG.
- PEREIRA, L. E.; LASTORIA, G.; PARANHOS FILHO, A. C.; MARCATO JÚNIOR, J.. *Application Techniques of Digital Image Processing in Different Orbital Sensors. Anuário do Instituto de Geociências - UFRJ*, v. 39, n. 3, p. 15, 12 set. 2016. Instituto de Geociências - UFRJ.

- PONZONI, F. J. Sensoriamento Remoto no Estudo da Vegetação: Diagnosticando a Mata Atlântica. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2002.
- PIROMAL, R.A.S. Avaliação do Modelo 5-Scale para Simular Valores de Reflectância de Unidades de Paisagem da Floresta Nacional do Tapajós. 151f. 2006. Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2006.
- PREIDL, S.; LANGE, M.; DOKTOR, D. *Introducing APiC for regionalised land cover mapping on the national scale using Sentinel-2 imagery. Remote Sensing Of Environment*, v. 240, p. 111673, abr. 2020. Elsevier BV.
- QGIS Development Team, 2021. QGIS *Geographic Information System. Open Source Geospatial Foundation Project*. <http://qgis.osgeo.org>.
- QUEIROZ, T. B.; SOUSA, R dos S. S.; BALDIN, T.; BATISTA, F. de J.; MARCHESAN, J.; PEDRALI, L. D.; PEREIRA, R. S. Avaliação do Desempenho da Classificação do Uso e Cobertura da Terra a partir de Imagens Landsat 8 e Rapideye na Região Central do Rio Grande do Sul. *Geociências*, São Paulo, v. 36, n. 3, p. 569-587, jan. 2017.
- RADOČAJ, D.; OBHOĐAŁ, J.; JURILIĆ, M.; GAŁPAROVIĆ, M. *Global Open Data Remote Sensing Satellite Missions for Land Monitoring and Conservation: a review. Land*, v. 9, n. 11, p. 402, 23 out. 2020. MDPI AG.
- REIS, R. B.; CARDOSO, P. V.; CRUZ, C. B. M.; VICENS, R. S. Classificação do Uso e Cobertura do Solo da APA do São João em uma Abordagem Orientada a Objeto. In: Simpósio Brasileiro de Sensoriamento Remoto, XIV, 2009, Natal, Anais. Natal: INPE, 2009, p. 7084-7094.
- SABINS, F. F. *Remote sensing for mineral exploration. Ore Geology Reviews*, v. 14, n. 3-4, p. 157-183, set. 1999. Elsevier BV.
- SANTA MARIA DE ITABIRA. [Sem título]. 2023. Coleção particular.
- SANTORI, A. R. Amostragem de Componentes Puros (vegetação, solo, água/sombra) em Imagens de Satélite Sentinel-2 como subsídio à Interpretação do NDVI na Bacia Hidrográfica do Ribeirão das Cruzes, Selvíria/MS. *Boletim Paulista de Geografia*, nº 107, 2022.
- SILVA, R. C. B. A. Revisão da Realidade Brasileira em Termos da Aplicabilidade de Pequenos Satélites para Missões de Sensoriamento Remoto. In: 1º Congresso Aeroespacial Brasileiro – Edição Especial Plêiade, Foz do Iguaçu, 2018.
- SHEYKHMUSA, M.; MAHDIANPARI, M.; GHANBARI, H.; MOHAMMADIMANESH, F.; GHAMISI, P.; HOMAYOUNI, S. *Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: a meta-analysis and systematic review. Ieee Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, v. 13, p. 6308-6325, 2020. Institute of Electrical and Electronics Engineers (IEEE).
- SHI, C.; WANG, L. *Incorporating spatial information in spectral unmixing: a review. Remote Sensing Of Environment*, v. 149, p. 70-87, jun. 2014. Elsevier BV.
- SHIMABUKURO, Y. E.; DUTRA, A. C.; ARAI, E. Modelo Linear de Mistura Espectral: Conceitos Teóricos, Algoritmos e Aplicações em Estudos na Amazônia Legal. *Revista Brasileira de Cartografia*, v. 72, p. 1140-1169, 30 dez. 2020. EDUFU - Editora da Universidade Federal de Uberlândia.
- SHIMABUKURO, Y. E.; MELLO, E. M. K.; MOREIRA, J. C.; DUARTE, V. Segmentação e classificação da imagem sombra do modelo de mistura para mapear desflorestamento na

Amazônia. São José dos Campos: INPE, 1997. 24 p. Republicado em 12-05-2010 por causa de acerto na numeração de controle da capa. (INPE-6147-PUD/83).

SHIMABUKURO, Y. E.; PONZONI, F. J. *Mistura Espectral: modelo linear e aplicações*. São Paulo: Oficina de Textos, 2017. 127 p.

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.

SMALL, C. *The Landsat ETM+ Spectral Mixing Space*. *Remote Sensing of Environment*, v. 93, n. 1-2, p. 1-17, out. 2004. Elsevier BV.

SOLA, I.; GARCÍA-MARTÍN, A.; SANDONÍS-POZO, L.; ÁLVAREZ-MOZOS, J.; PÉREZ-CABELLO, F.; GONZÁLEZ-AUDÍCANA, M.; LLOVERÍA, R. M. *Assessment of atmospheric correction methods for Sentinel-2 images in Mediterranean landscapes*. *International Journal Of Applied Earth Observation And Geoinformation*, v. 73, p. 63-76, dez. 2018. Elsevier BV.

STEFFEN, C. A. *Introdução ao Sensoriamento Remoto*. São José dos Campos, Instituto Nacional de Pesquisas Espaciais, 2022.

TALUKDAR, S.; SINGHA, P.; MAHATO, S.; SHAHFAHAD; PAL, S.; LIOU, Y.; RAHMAN, A. L. *Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review*. *Remote Sensing*, v. 12, n. 7, p. 1135, 2 abr. 2020. MDPI AG.

TSENG, Y. *Spectral unmixing for the classification of hyperspectral images*. *International Archives of Photogrammetry and Remote Sensing*, v. 33, n. B7/4; PART 7, p. 1532-1538, 2000.

USGS. *United States Geological Survey. EarthExplorer*. Disponível em: <https://earthexplorer.usgs.gov/>. Acesso em 14 jul. 2022.

USTIN, S. L.; GAMON, J. A.. *Remote sensing of plant functional types*. *New Phytologist*, v. 186, n. 4, p. 795-816, 10 maio 2010. Wiley.

VALE, J. R. B.; COSTA, J. A. da; SANTOS, J. F. dos; SILVA, E. L. S. da; FAVACHO, A. T. *Análise Comparativa de Métodos de Classificação Supervisionada aplicada ao Mapeamento da Cobertura do Solo no Município de Medicilândia, Pará*. *Interespaço: Revista de Geografia e Interdisciplinaridade*, v. 4, n. 13, p. 26, 23 abr. 2018. Universidade Federal do Maranhão.

WHITE, J. C.; COOPS, N. C.; WULDER, M. A.; VASTARANTA, M.; HILKER, T.; TOMPALSKI, P. *Remote Sensing Technologies for Enhancing Forest Inventories: a review*. *Canadian Journal of Remote Sensing*, v. 42, n. 5, p. 619-641, 27 jul. 2016.

XI, Y.; THINH, N. X.; LI, C. *Preliminary comparative assessment of various spectral indices for built-up land derived from Landsat-8 OLI and Sentinel-2 MSI imageries*. *European Journal of Remote Sensing*, v. 52, n. 1, p. 240-252, 1 jan. 2019.

ZANOTTA, Daniel Capella et al. *Linear spectral mixing model for identifying potential missing endmembers in spectral mixture analysis*. *IEEE transactions on geoscience and remote sensing*, v. 52, n. 5, p. 3005-3012, 2013.

ZANOTTA, D. C.; HAERTEL, V.; SHIMABUKURO, Y. E.; RENNO, C. D. *Linear Spectral Mixing Model for Identifying Potential Missing Endmembers in Spectral Mixture Analysis*.

Ieee Transactions On Geoscience And Remote Sensing, v. 52, n. 5, p. 3005-3012, maio 2014.
Institute of Electrical and Electronics Engineers (IEEE).

ZHAO, L.; SHI, Y.; LIU, B.; HOVIS, C.; DUAN, Y.; SHI, Z. *Finer Classification of Crops by Fusing UAV Images and Sentinel-2 Data. Remote Sensing*, v. 11, n. 24, p. 3012, 14 dez. 2019. MDPI AG.