

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM SENSORIAMENTO REMOTO

LUIZ FERNANDO GUIMARÃES SCHWARTZMAN

**Análise do desempenho de classificadores para o
mapeamento de cobertura do solo em regiões de
floresta tropical; estudo de caso do município de Tefé-
AM**

Porto Alegre

2017

LUIZ FERNANDO GUIMARÃES SCHWARTZMAN

**Análise do desempenho de classificadores para o
mapeamento de cobertura do solo em regiões de
floresta tropical; estudo de caso do município de Tefé-
AM**

Dissertação de mestrado
apresentada ao Programa de Pós-
Graduação em Sensoriamento Remoto, da
Universidade Federal do Rio Grande do
Sul, como parte dos requisitos para a
obtenção do título de Mestre em
Sensoriamento Remoto.

Área de concentração:
Sensoriamento Remoto Aplicado

ORIENTADOR: Dejanira Luderitz Saldanha

Porto Alegre

2017

LUIZ FERNANDO GUIMARÃES SCHWARTZMAN

Análise do desempenho de classificadores para o mapeamento de cobertura do solo em regiões de floresta tropical; estudo de caso do município de Tefé-AM

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Sensoriamento Remoto e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: _____

Prof. Dra. Dejanira Luderitz Saldanha, UFRGS -
Porto Alegre, Brasil)

Banca Examinadora:

Prof. Dr. Dirce Maria Antunes Suertegaray, UFRGS
Doutora pela USP - São Paul, Brasil

Prof. Dr. Daniel Capella Zanotta, IFRS
Doutor pelo INPE - São José dos Campos, Brasil

Prof. Dr. Tatiana Silva da Silva, UFRGS
Doutor pela FURG - Rio Grande, Brasil

Coordenador do PPGRS:

Prof. Dr. Laurindo Antonio Guasselli

Porto Alegre, março, 2017.

AGRADECIMENTOS

Ao Cnpq pela bolsa que permitiu o desenvolvimento desse trabalho.

A professora Dejanira pelo apoio, paciência e carinho durante toda essa jornada.

Ao professor Rafael Ribeiro pelo apoio e conselhos que foram fundamentais para o desenvolvimento deste trabalho.

A equipe do ICMBio da sede em Tefé-AM pelo apoio técnico e logístico para realização das atividades de campo.

Ao meu pai pela motivação e ajuda financeira.

A minha namorada Raquel pela ajuda e companheirismo durante toda essa jornada.

Aos colegas Daniel Gomes e Mábila Makrakis pelas conversas e contribuições que permitiram o desenvolvimento desse trabalho.

A grande amiga Eliane Neves que me recebeu em sua residência durante as atividades de campo e tanto colaborou para o desenvolvimento desse trabalho.

Ao amigo e assistente de campo Sarney Martins de Oliveira pelo excelente trabalho e por compartilhar seu profundo conhecimento da região estudada.

Resumo

A Amazônia é a maior floresta tropical do mundo e desempenha papel fundamental na regulação climática do continente sul-americano. A expansão de atividades antrópicas ocorridas na Amazônia nos últimos 70 anos provocou mudanças nos padrões de cobertura do solo na região que levaram ao surgimento de diversas políticas de conservação em nível nacional e internacional. Uma das principais técnicas para o monitoramento e acompanhamento do desenvolvimento dessas políticas são as técnicas fornecidas pelo sensoriamento remoto. Esse trabalho tem por objetivo avaliar o desempenho de três classificadores de imagens digitais (Mínima Distância Euclidiana, Máxima Verossimilhança e *Extraction and Classification of Homogeneous Objects*) na região central da Amazônia brasileira. Foram utilizadas duas abordagens para realizar as classificações, na primeira os classificadores foram aplicados sobre uma imagem do satélite *Landsat - 8* que teve seus números digitais convertidos para valores de refletância e na segunda abordagem os classificadores foram aplicados sobre imagens fração resultantes da aplicação do modelo linear de mistura espectral na imagem do *Landsat - 8*. Para avaliar o desempenho dos classificadores e das abordagens propostas aplicou-se uma matriz de confusão e foi calculado o índice Kappa para cada um dos resultados obtidos. Os resultados mostraram que o Máxima verossimilhança apresentou o melhor desempenho nas duas abordagens seguido pelo ECHO com resultados próximos. O Mínima Distância teve o pior desempenho nas duas abordagens. Os resultados também indicaram que a redução da dimensionalidade dos dados provocadas pelo modelo linear de mistura espectral foi capaz de melhorar o desempenho apenas do ECHO.

Palavras-chave: Amazônia, Sensoriamento Remoto, Classificadores

Abstract

The Amazon is the largest tropical forest in the world and plays a key role in the climate regulation in South America. In the last 70 years, the expansion of anthropogenic activities in this biome has led to changes in land's cover patterns in that region which have led to the emergence of several conservation policies in national and international levels. One of the main techniques of monitoring the development of these policies is the technique provided by remote sensing. This thesis aims to evaluate the performance of three digital image classifiers (Minimum Distance, Maximum Verification and Extraction and Classification of Homogeneous Objects) in the Brazilian's Amazon central region. Two approaches were used to perform the classifications. To the first one the classifiers were applied on the satellite Landsat – 8's image that had its digital numbers converted to values of reflectance. To the second approach the classifiers were applied on fraction images resulting from the application of the linear model of spectral mixing in the Landsat-8 image. To evaluate the performance of the proposed classifiers and approaches a confusion matrix the Kappa index was adopted to calculate each of the obtained results. The results showed that the maximum probability has presented the best performance in the two approaches followed by ECHO with close results. The Minimum Distance had the worst performance in both approaches. The results also indicated that the reduction of data dimensionality caused by the linear spectral mixing model was able to improve the performance of only ECHO.

Key-words: Amazon, Remote Sensing, Classifiers

SUMÁRIO

1. INTRODUÇÃO.....	15
1.1 Amazônia	15
1.2 Antecedentes e Justificativas	16
1.3 O Sensoriamento Remoto na Amazônia	20
2. OBJETIVOS.....	22
2.1 Objetivo Geral	22
2.2 Objetivo Específico.....	22
3. REVISÃO BIBLIOGRÁFICA	22
3.1 Sensoriamento Remoto Óptico Aplicado em Estudos de Uso e Cobertura do Solo em Florestas Tropicais.....	22
3.2 Classificação de Imagens	25
3.2.1 Distância Mínima (Euclidiana)	28
3.2.1 Máxima Verossimilhança.....	29
3.2.1 <i>Extraction and Classification of Homogeneous Objects</i> (ECHO)	33
3.3 Modelo Linear de Mistura Espectral	37
3.4 Avaliação de algoritmos de classificação de imagens para o mapeamento da cobertura da terra	40
3.4.1 Avaliação de Acurácia	42
4. METODOLOGIA	53
4.1 Área de Estudo.....	54
4.2 Seleção de Imagens.....	59
4.2.1 Pré-processamento	60

4.3	Aplicação do Modelo Linear de Mistura Espectral.....	60
4.4	Classificação	61
4.4.1	Esquema de Classificação	62
4.4.2	Dados de Referência	64
4.4.3	Classificação Abordagem 01	71
4.4.4	Classificação Abordagem 02	71
4.6	Análise de Acurácia.....	71
5.	RESULTADOS E DISCUÇÃO	72
5.1	Classificação da Área de Estudo.....	72
5.2	Avaliação dos Classificadores.....	75
5.2.1	Abordagem 1	78
5.2.2	Abordagem 2	83
5.3	Cobertura do Solo	85
5.3.1	Classes Naturais.....	85
5.3.2	Classes Antrópicas	87
6.	CONCLUSÕES.....	90
7.	REFERÊNCIAS BIBLIOGRÁFICAS	93
ANEXO 1	100
ANEXO 2	101
ANEXO 3	102
ANEXO 4	103
ANEXO 5	104
ANEXO 6	105
ANEXO 7	106
ANEXO 8	107
ANEXO 9	108

LISTA DE ILUSTRAÇÕES

Figura 01: Gráfico de dispersão das médias das classes e do valor do pixel a ser classificado.....	29
Figura 02: Diagrama de dispersão por meio de isolinhas de probabilidade.....	32
Figura 03: Fluxograma das etapas de processamento do ECHO (Adaptado de Kast, <i>et al.</i> 1977).....	35
Figura 04: Matriz de Confusão.....	49
Figura 05: Fluxograma das etapas metodológicas do trabalho.....	53
Figura 06: Mapa Hidrográfico do estado do Amazonas com destaque para o município de Tefé.....	55
Figura 07: Mapa de localização da sede municipal de Tefé.....	57
Figura 08: Mapa da Área de Estudo.....	58
Figura 09: Foto das áreas referentes aos <i>pixels</i> selecionados como membros de referencia.....	61
Figura 10: Curvas espectrais dos <i>pixels</i> selecionados como membros de referencia.....	62
Figura 11: Fluxograma do Sistema para a identificação das classes de cobertura do solo em campo.....	63
Figura 12: Mapa de localização dos pontos de amostragem.....	65
Figura 12: Coleta de dados em campo.....	67
Figura 14: Comparação entre as composições de bandas RGB432 e RGB543.....	69

Figura 15: Gráfico de comportamento espectral da média das área de treinamento das classes da imagem Landsat-8, sendo o eixo X as bandas do sensor e o eixo Y os valores de refletância.....	69
Figura 16: Gráfico de comportamento espectral da média das classes e em relação a uma amostra na imagem Resourcesat-2, sendo o eixo X as bandas do sensor e o eixo Y os valores de refletância.....	70
Figura 17: Classificação produzida pelo Mínima Distância abordagem1.....	72
Figura 18: Classificação produzida pelo Mínima Distância abordagem2.....	73
Figura 19: Classificação produzida pelo Máxima Verossimilhança abordagem1.....	73
Figura 20: Classificação produzida pelo Máxima Verossimilhança abordagem2.....	74
Figura 21: Classificação produzida pelo ECHO abordagem1.....	74
Figura 22: Classificação produzida pelo ECHO abordagem2.....	75
Figura 23: Efeito "sal pimenta" na imagem classificada pelo MaxVer na abordagem1.....	79
Figura 24: Área de cobertura homogênea classificada pelo ECHO abordagem1.....	80
Figura 25: Efeito "sal pimenta" na imagem classificada pelo MD na abordagem1.....	82

Figura 26: Comparativo do efeito de nuvens e sombras nas classificações (A - Landsat-8 R4G3B2, B - Classificação MD, C - Classificação MaxVer, D - Classificação ECHO).....	82
Figura 27: Intensificação do efeito "sal pimenta" na imagem classificada pelo MD na abordagem 2.....	85
Figura 28: Núcleos de Floresta Primária em destaque na imagem classificada MaxVer A2.....	86
Figura 29: Diferentes estágios sucessionais de Floresta Secundária.....	87
Figura 30: Diferentes tipos de Campinarana encontrados na área de estudo.....	87
Figura 31: Cultivo familiar de mandioca.....	88
Figura 32: Produção da farinha de mandioca.....	88
Figura 33: Área destinada à construção de tanques para a criação de peixes.....	89
Figura 34: Centro da Cidade de Tefé-AM.....	89

LISTA DE TABELAS

Tabela 01: Número de amostras desejáveis em relação a extensão territorial da classe (Adaptado de LEA; CURTIS, 2010).....	45
Tabela 02: Tabela descritiva do sistema de classificação.....	63
Tabela 03: Datas de realização das expedições para coleta dos pontos de amostragem.....	67
Tabela 4: Matriz de Confusão Mínima Distancia Abordagem 1.....	75
Tabela 5: Matriz de Confusão Mínima Distancia Abordagem 2.....	76
Tabela 6: Matriz de Confusão Máxima Verossimilhança Abordagem 1.....	76
Tabela 7: Matriz de Confusão Máxima Verossimilhança Abordagem 2.....	76
Tabela 8: Matriz de Confusão ECHO Abordagem 1.....	77
Tabela 9: Matriz de Confusão ECHO Abordagem 2.....	77
Tabela 10: Indicadores de acurácia e índice Kappa dos classificadores.....	78
Tabela 11: Área das classes.....	78

LISTA DE ABREVIATURAS

AP: Acurácia do Produtor

ARPA: Áreas Protegidas da Amazônia

AU: Acurácia do Usuário

CDB: Convenção da Diversidade Biológica

EC: Esquema de Classificação

ECHO: *Extraction and Classification of Homogeneous Objects*

ETM+: *Enhanced Thematic Mapper Plus*

FAO: Organização das Nações Unidas para a Alimentação e a
Agricultura

FIRMS: *Fire Information for Resource Management System*

FRA: Global Forest Resources Assessments

GFC: *Global Forest Change*

IBGE: Instituto Brasileiro de Geografia e Estatística

INPE: Instituto Nacional de Pesquisas Espaciais

IPCC: Painel Intergovernamental sobre Mudanças Climáticas

JRC: *European Commission Joint Research Center*

MaxVer: Máxima Verossimilhança

MD: Mínima Distância

MLME: Modelo Linear de Mistura Espectral

OLI: *Operational Land Imager*

ONU: Organização das Nações Unidas

PAS: Programa Amazônia Sustentável

PIN: Plano de Integração Nacional

PDI: Processamento Digital de Imagens

Prodes:

RMSE: *Root Mean Square Error*

SIG: Sistema de Informações Geográficas

SNUC: Sistema Nacional de Unidades de Conservação

SUDAM: Superintendência de Desenvolvimento da Amazônia

TIRS: *Thermal Infrared Sensor*

UC's: Unidades de Conservação

USGS: Serviço Geológico do Estados Unidos

1. INTRODUÇÃO

1.1 Amazônia

A Amazônia, maior floresta tropical do mundo, expande-se sobre nove países da América do Sul, sendo eles: Brasil, Venezuela, Colômbia, Equador, Peru, Bolívia, Guiana, Guiana Francesa e Suriname. O Brasil é o que detém em seu território a maior parte da Amazônia (aproximadamente 60%). A Amazônia brasileira se estende por 5 milhões de km² e recobre nove estados (Amazonas, Acre, Amapá, Roraima, Rondônia, Tocantins, Pará, Mato Grosso e Maranhão), ocupando quase 60% do território nacional (Instituto Brasileiro de Geografia e Estatística, 2004).

Devido à sua dimensão continental, a floresta amazônica desempenha um papel fundamental na regulação climática do continente sul-americano. A floresta se apresenta como um elemento crítico para a regulação do clima na região. No restante do continente sul-americano, ela detém o principal mecanismo de regulação térmica para o intenso calor radiativo que incide sobre sua superfície. Por meio da evapotranspiração, a floresta contribui com grandes volumes de vapor d'água na atmosfera. A convecção tropical, por sua vez, redistribui esse calor e umidade para a zona temperada do continente. O resultado desse processo é um resfriamento da região amazônica e a liberação de calor na atmosfera, que, pelo seu grande volume, desempenha um papel importante também na circulação atmosférica global (Nobre, 2014).

Além do seu importante papel para o clima, a Amazônia ainda abriga grande parte da biodiversidade existente na região neotropical (Ayres *et al.* 2005) e aproximadamente um terço do estoque genético do planeta (Albagli, 2001).

Nesse panorama de inter-relações complexas, onde ainda não se conhece a totalidade das interações das variáveis climáticas, as mudanças no uso e cobertura do solo na Amazônia surgem como fatores que podem afeta-lá e, por consequência, o clima do continente.

1.2 Antecedentes e Justificativas

Para compreender a atual configuração do uso do solo e os principais vetores de transformação de sua cobertura é necessário fazer um breve resgate do seu histórico de ocupação.

Do começo do século XVI, início da ocupação portuguesa na Amazônia, até meados do século XX, estopim da ditadura militar no Brasil, a região passou por diferentes ciclos de exploração. Segundo Loureiro (2002) os principais ciclos exploratórios foram baseados em diversos recursos, como castanha, cacau, tabaco, frutos exóticos e peles de animais que tinham grande valor comercial no mercado europeu até metade do século XIX, além da borracha que abasteceu esse mesmo mercado.

Apesar de ter passado por esses ciclos de exploração que abasteceram os principais mercados do mundo durante séculos, nenhum deles desmatou tanto quando as políticas de desenvolvimento que tiveram início na década de 1960. Amparados no discurso de uma eminente internacionalização da Amazônia, os militares deram início ao período de implantação de grandes projetos na região (Loureiro, 2002).

Com o objetivo de criar pólos de desenvolvimento na região foi criada a Superintendência de Desenvolvimento da Amazônia (SUDAM) em 1968. Incentivos fiscais e créditos para investimento impulsionaram a instalação dos setores de mineração, extração de madeira e agropecuário na Amazônia brasileira. Outro grande projeto desse período foi a Rodovia Transamazônica, que tinha como objetivo integrar o sul da Amazônia cortando-a de leste a oeste. Ainda a fim de estimular a migração para a região, o governo militar estabeleceu o Plano de Integração Nacional (PIN), que garantia reservados 100 km de cada lado da Transamazônica para o assentamento prioritário da população proveniente dos estados do Nordeste (Loureiro 2002).

O resultado dessas políticas permitiu a mineração na serra dos Carajás, a construção de hidrelétricas, a implantação do pólo tecnológico e industrial da Zona Franca de Manaus e a construção de rodovias. Entretanto, o resultado mais evidente foi a rapidez com a qual a Amazônia foi desmatada. Destas frentes de expansão simultâneas formou-se a região hoje conhecida como

“Arco do Desmatamento”, envolvendo mais de cem municípios onde se encontram mais de 90% de área desmatada da Amazônia (Loureiro, 2002; Nobre, 2014).

O desmatamento por corte raso, até 2014, beirava os 20% da cobertura original na Amazônia brasileira, e a degradação florestal, estima-se, já teria perturbado a floresta remanescente em variados graus, afetando adicionalmente mais 20% da cobertura original (Nobre, 2014).

A extensão acumulada do desmatamento amazônico no Brasil, até 2016, foi de aproximados 421.871 km² (Instituto Nacional de Pesquisas Espaciais, 2016), área equivalente a quase dez estados do Rio de Janeiro. Sendo assim, a expansão dos centros urbanos e a intensificação de atividades do setor agropecuário e de mineração na região criaram uma situação de vulnerabilidade para a maior floresta tropical do planeta (Barreto et al. 2006; Magalhães et al. 2014).

Frente a essa conjuntura de rápida e intensa conversão da cobertura original da Amazônia para diversos usos, as Áreas Protegidas (AP's) se apresentam como um instrumento importante para preservar a integridade dos ecossistemas, a biodiversidade e os serviços ambientais associados, e ainda o direito de permanência e reprodução cultural das populações tradicionais e indígenas que habitam a região (ARPA, 2011).

Segundo o Decreto n.º 5.758/2006, que instituiu o Plano Estratégico Nacional de Áreas Protegidas (PNAP) e a Lei 9.985 de 2000 que criou o Sistema Nacional de Unidades de Conservação (SNUC), são consideradas Áreas Protegidas as Unidades de Conservação (UC's), as Terras Indígenas (TI's) e as terras ocupadas por remanescentes das comunidades dos quilombos. O surgimento das UC's na Amazônia tem início em 1959 com a criação do Parque Nacional do Bananal, no Tocantins, ocupando uma área de 20.000km². Nas décadas seguintes mais UC's foram criadas e no ano de 1984 essas áreas somavam 124.000km², sendo maior parte delas sob jurisdição Federal (ARPA, 2011). A partir de 1985 o movimento de criação de UC's tomou força no âmbito estadual devido às exigências de dois programas de

desenvolvimento sustentável (Polonoroeste e Planaflores) financiados pelo Banco Mundial (Millikan, 1998).

A Convenção da Diversidade Biológica (CDB), tratado da Organização das Nações Unidas (ONU) do qual o Brasil é signatário, estabelecido durante a realização da ECO-92, é um dos mais importantes instrumentos internacionais relacionados ao meio ambiente e serviu como um novo combustível para o surgimento de UC's na Amazônia. Afim de cumprir os objetivos de proteção da diversidade biológica estabelecidos na CDB, o governo federal, via Ministério do Meio Ambiente (MMA), criou o Programa de Áreas Protegidas da Amazônia, que tem como principais objetivos expandir e fortalecer o Sistema Nacional de Unidades de Conservação (SNUC) na Amazônia, proteger 60 milhões de hectares, assegurar recursos financeiros para a gestão destas áreas a curto e longo prazo e promover o desenvolvimento sustentável (ARPA, 2011).

Segundo dados do Instituto IMAZON (2011), até dezembro de 2010, o resultado desses investimentos na criação, estruturação e manutenção dessas áreas permitiu que 2.197.485 km² fossem definidos como Áreas Protegidas, cobertura equivalente a 43,9% do território da Amazônia Legal.

Não obstante aos esforços para a conservação da Amazônia, a gestão de Áreas Protegidas é ainda um grande desafio na região. No contexto amazônico grande parte das Áreas Protegidas enfrenta problemas, como: desmatamento para variados fins, pesca, caça, agricultura, mineração, invasão de terras e problemas de titulação (Ayres *et al*, 2005). Existem ainda aquelas que estão localizadas em áreas remotas e inacessíveis para a agência governamental responsável. O número reduzido das equipes de gestão dessas áreas, aproximadamente uma pessoa para cada 15.000km², também se mostra um obstáculo para o cumprimento dos objetivos de preservação e conservação (Ayres *et al*, 2005).

É comum se deparar com a expressão "vazio demográfico" quando se trata da Amazônia, mas é preciso ter cuidado e observar mais atentamente as nuances da distribuição de populações humanas nessa região. Com as políticas de desenvolvimento e integração para a região norte do país desenvolvidas pelo governo federal nas últimas décadas, houve um aumento

significativo no número de residentes na região. A região norte deu um salto de 3.603.860 milhões de habitantes em 1970 para 15.864.454 milhões de habitantes em 2010, sendo que nesses 40 anos a região mais que quadruplicou o seu número de habitantes (IBGE, 2010).

O estado do Amazonas merece destaque quando estamos falando de Amazônia e uso da terra, pois além de ser o maior estado brasileiro e estar totalmente inserido na floresta amazônica ele ainda detém a maior área de floresta contínua e também a maior extensão de Áreas Protegidas da Amazônia, recobrando em 2010 uma área de 798.808 km² (IMAZON, 2011). Em 1970 a população do estado era de 960.934 mil habitantes, sendo que desses a maioria (551.656 mil) moravam em áreas rurais, já no ano de 2016 o estado triplicou sua população em relação a 1970. Acompanhando o movimento nacional de migração para os centros urbanos ocorrido no Brasil no início do XIX e final do século XX, menos de um terço dos quatro milhões de habitantes do estado viviam em áreas rurais, no ano 2016 (IBGE, 2010).

Apesar dos dados do censo demográfico indicarem um adensamento populacional nos núcleos urbanos e áreas rurais pouco habitadas é interessante verificar a distribuição espacial desses dados. O Amazonas tem 62 municípios que se distribuem por toda a calha do Rio Amazonas e de seus afluentes. Além das sedes municipais, são abundantes nas margens da rede hidrográfica de toda a bacia amazônica a presença de pequenas comunidades rurais. Apenas na região do médio Solimões, que abrange as Reservas de Desenvolvimento Sustentável Mamirauá e Amanã, que representa aproximadamente 20% da área do Amazonas, são mais de 300 comunidades (Moura *et al*, 2016).

Um dos mais antigos sítios de ocupação humana do período colonial no interior da Amazônia é onde hoje se encontra o município de Tefé. Localizado na margem direita do Rio Tefé, no sentido foz-nascente com saída próxima ao rio Solimões, o município encontra-se numa posição geográfica central no estado (Rodrigues, 2011).

Devido à sua localização, Tefé apresenta forte vocação comercial, por viabilizar um porto próximo da principal artéria fluvial amazônica e o controle da

circulação de mercadorias na calha do rio Solimões, caracterizando-se como um entreposto de abastecimento na região (Souza, 1989).

Esses fatores levaram a região onde hoje se encontra o município de Tefé a se consolidar como o principal núcleo urbano na região do médio Solimões e um importante centro regional na rede de cidades do Amazonas (Programa Amazônia Sustentável, 2007), sendo ele o responsável por fornecer serviços (bancos, hospitais, universidade) à população da região e servir de elo comercial entre os produtores locais e os mercados nacional e internacional.

Compreender a distribuição e a configuração das atividades antrópicas e suas relações com o ambiente se mostra fundamental para a preservação da floresta tropical na região central da Amazônia. Identificar se as atividades desenvolvidas nessa região estão de acordo com as políticas de conservação é o primeiro passo para verificar a efetividade dessas políticas.

Nesse contexto, identificar o desempenho das técnicas utilizadas para a classificação da cobertura do solo por meio de imagens orbitais para essa região se mostra fundamental para que se possa optar pelo método mais adequado para a identificação das classes de cobertura do solo de um dos principais núcleos populacionais na região mais preservada da Amazônia brasileira.

1.3 O Sensoriamento Remoto na Amazônia

O sensoriamento remoto se apresenta como uma importante ferramenta para atingir os objetivos de conservação da Amazônia. Fornecendo dados que podem recobrir milhares de quilômetros em apenas uma cena, a um baixo custo, quando comparado com levantamentos de campo, as imagens de satélite tornaram-se recorrente no estudo das dinâmicas de uso do solo (Jensen, 2009).

Entretanto, alguns problemas são enfrentados ao utilizarmos imagens de satélites em regiões tropicais. As características climáticas dessas áreas contribuem para a formação de nuvens durante todo o ano, sendo um dos principais fatores a dificultar o uso de imagens produzidas por sensores óticos (Achard et al. 2002).

Devido aos intensos efeitos de absorção e espalhamento de radiação, provocados pelas nuvens e pela atmosfera, a correção desses efeitos torna-se indispensável em grande número de estudos que visa trabalhar com a temática de uso e cobertura do solo (Song et al. 2001). A dificuldade em encontrar cenas com baixo percentual de cobertura de nuvens, principalmente em duas cenas da mesma área de estudo, é um fator limitante para estudos que fazem análises temporais do padrão do uso do solo nesta área.

A presença de diferentes fitofisionomias em variados estágios sucessionais, diversas culturas e formas diferentes de manejo agrícola, áreas urbanas e uma hidrografia complexa que tem três "tipos" distintos de água (branca, clara e preta) tornam o uso de classificadores de imagens digitais uma árdua tarefa (Silva et al. 2011; Adams et al. 1995; Dutra et al. 2009).

A dificuldade em obter-se bons resultados nos estudos de classificação de cobertura da terra na Amazônia pode ser atribuída a uma série de elementos como: a dificuldade em obter classes consistentes em imagens de datas diferentes e a variabilidade na iluminação e na resposta dos sensores embarcados nos satélites (Adams et al., 1995).

Contudo, metodologias de classificação vêm sendo desenvolvidas a fim de solucionar ou ao menos minimizar as dificuldades encontradas na classificação da cobertura do solo em florestas tropicais. Métodos como o *extraction and classification of homogeneous objects* (ECHO) e a aplicação do modelo linear de mistura espectral vêm apresentando bons resultados (Lu, 2005). Nesse trabalho pretende-se analisar o desempenho de classificadores na região central da Amazônia utilizando imagens de resolução espacial média (30 metros). Assim, objetiva-se contribuir com uma melhor compreensão do desempenho e das limitações no uso de algoritmos de classificação de cobertura do solo em imagens ópticas para a região central da Amazônia.

2. OBJETIVOS

2.1 Objetivo Geral

Analisar o desempenho dos classificadores Mínima Distância, Máxima Verossimilhança e ECHO na região central da Amazônia utilizando imagens de resolução espacial média.

2.2 Objetivo Específico

- Identificar e analisar a cobertura do solo do município de Tefé por meio de dados de sensoriamento remoto orbital.
- Comparar os resultados obtidos pelos diferentes métodos de classificação aplicados.

3. REVISÃO BIBLIOGRÁFICA

Neste capítulo serão apresentadas a fundamentação teórica, que consiste na revisão da literatura, pertinente a temática do trabalho e a sua execução, bem como o embasamento científico utilizado para o desenvolvimento de suas etapas metodológicas.

3.1 Sensoriamento Remoto Óptico Aplicado em Estudos de Uso e Cobertura do Solo em Florestas Tropicais

O sensoriamento remoto tornou-se nas últimas décadas uma das mais bem-sucedidas tecnologias de coleta automática de dados para levantamento e monitoramento de recursos naturais em escala global. Sua definição mais usual é a proposta por Avery; Berlin (1992) e Jensen (2009): técnica para obter informações sobre objetos através de dados coletados por instrumentos que não estejam em contato físico com os objetos investigados. Desse conciso conceito é possível extrair a premissa básica do sensoriamento remoto que é a ausência de contato entre o alvo e o sensor, sendo a radiação eletromagnética (REM) a responsável por transmitir a informação sensor-alvo e alvo-sensor (Menezes, 2012).

Devido às suas características as imagens dos satélites da série Landsat são utilizadas de forma recorrente em estudos e trabalhos na região amazônica. O projeto PRODES do INPE é um dos mais bem sucedidos na utilização de imagens orbitais fornecidas pela série Landsat na identificação de áreas de corte raso da vegetação na Amazônia brasileira. Desde 1988 o projeto monitora as atividades de corte raso em áreas maiores de 6,25 hectares e é o responsável por calcular a taxa anual de desmatamento da Amazônia no Brasil.

Ao longo dos anos o projeto sofreu algumas alterações na metodologia utilizada para realizar o cálculo das áreas que sofreram o corte raso. Inicialmente o método utilizado para a detecção dos polígonos de áreas desmatadas era a interpretação visual. Para recobrir toda a Amazônia eram utilizadas 220 cenas Landsat impressas coloridas que serviam de base para a identificação das áreas desmatadas que posteriormente teriam seus polígonos digitalizados manualmente no sistema de informações geográficas (SIG) desenvolvido pela Divisão de Processamento de Imagens do INPE (CÂMARA; VALERIANO; SOARES, 2013).

Somente entre os anos de 2003 e 2005 que o projeto adotou o processo de interpretação via classificação digital assistida por computador para realizar a identificação das áreas desmatadas. Atualmente o método aplicado na identificação das áreas desmatadas é bem diferente do utilizado nos anos iniciais do projeto. Utiliza-se o modelo linear de mistura espectral (MLME) para estimar as proporções dos componentes vegetação, solo e sombra em cada imagem e concentrar a informação sobre desmatamento nas imagens-fração solo e sombra. Essas imagens-fração são então seguidas e posteriormente classificadas pelo algoritmo classificador não supervisionado de agrupamento de dados ISOSEG (CÂMARA; VALERIANO; SOARES, 2013).

Apesar do sucesso obtido pelo PRODES no monitoramento das alterações na cobertura florestal da Amazônia brasileira o recente estudo publicado por Richards et al (2016) traz dados que mostram que o sistema pode ter deixado de detectar 9.000 km² de desmatamento entre os anos de 2008 e 2012. Para chegar a esse número os autores realizaram um teste *t*

pareado de diferenças anuais no desmatamento da Amazônia com os dados fornecidos pelo Prodes, pelo *Global Forest Change* (GFC) e pelo *Fire Information for Resource Management System* (FIRMS) e analisaram, onde, quando e como os dados divergiram e ainda examinaram os padrões espaciais subjacentes as divergências.

Outro projeto de destaque no monitoramento de florestas é o *Global Forest Resources Assessments* (FRA) desenvolvido a partir de uma parceria entre a *European Commission Joint Research Center* (JRC) e a Organização das Nações Unidas para a Alimentação e a Agricultura (FAO). O projeto visa entender as condições e as mudanças dos recursos florestais globais através do tempo (FAO & JCR, 2012). Devido a sua abrangência global o FRA se tornou uma importante fonte de informação para subsidiar os esforços de manejo florestal sustentável, de redução na concentração dos gases do efeito estufa e em promover outras iniciativas internacionais. De acordo com as diretrizes para inventários de gases de efeito estufa nacionais publicados pelo Painel Intergovernamental sobre Mudanças Climáticas (IPCC) (PAUSTIAN; RAVINDRANATH; VAN AMSTEL, 2006), a FAO é a principal fonte de dados de atividade e fatores de emissão para a floresta e outras categorias de uso da terra. As diretrizes do IPCC sugerem que, sempre que dados mais detalhados do país não estejam disponíveis, informações agregadas pode ser obtidas a partir de fontes de dados internacionais, como a FRA.

Um dos mais relevantes trabalhos a utilizar imagens de satélite no monitoramento das florestas tropicais em escala global foi o desenvolvido por Achard *et al.*, (2002). Nesse trabalho foram utilizadas imagens com diferentes resoluções espaciais com o objetivo de calcular a taxa de desmatamento nas florestas tropicais do mundo entre os anos de 1990 e 1997. O resultado deste trabalho indicou uma taxa anual de desmatamento das florestas tropicais estimada em $5,8 \pm 1,4 \times 10^6$ ha somado ainda a $2,3 \pm 0,7 \times 10^6$ ha de áreas cuja a degradação florestal pode ser inferida visualmente por meio das imagens de satélite. Dos três continentes analisados pelos autores (Sudeste da Ásia, América Latina e África) o sudeste asiático foi o que apresentou a maior taxa percentual de desmatamento $2,5 \pm 0,8 \times 10^6$ (0,91%) seguido pela África com

aproximadamente metade da taxa da Ásia. A América Latina foi o continente com a menor taxa percentual (0,38%) mas em números absolutos a média da extensão de área desmatada no período estudado ($2,5 \pm 1,4 \times 10^6$) foi similar a apresentada pelo sudeste asiático.

Além desses dados o trabalho traz informações importantes sobre os vetores de conversão da floresta tropical para outros usos e as limitações na utilização de sensores orbitais para a identificação de processos de degradação florestal como o corte seletivo de madeira em escala global.

Partindo para uma abordagem local do uso de sensores remotos em estudos de florestas tropicais, as imagens de satélite também são uma importante fonte de dados que possibilita a realização de diversos estudos, como a classificação de vegetação realizada por Lu (2005) no noroeste do estado de Rondônia no município de Machadinho d'Oeste, onde a partir da integração de dados de um inventário de vegetação e imagens do sensor *Thematic Mapper* do Landsat foi possível identificar e diferenciar três diferentes estágio sucessionais de floresta. Outro exemplo da utilização de imagens de satélite é encontrado no trabalho de Renó *et al.* (2011) que utilizou cenas da série Landsat para avaliar as mudanças provocadas pelo desmatamento em um período de 30 anos nas várzeas amazônicas localizadas entre os municípios de Parintins - AM e Almeirim - PA.

3.2 Classificação de Imagens

Os dados de sensoriamento remoto orbital ótico permitem numerosas análises. O resultado único da interação de cada objeto da superfície terrestre com a energia eletromagnética em seus diferentes comprimentos de onda permite ao analista identificar, diferenciar e obter informações sobre o alvo observado pelo sensor (Novo, 2008; Jensen, 2009). Identificação de padrões hidrológicos e de geomorfologia, taxas de sedimentação e de conversão de uso do solo, quantificação de estoques de carbono entre outros são alguns dos dados atualmente extraídos das imagens de sensores orbitais (Mertes *et al.* 1995; Soares *et al.* 2001; Achard *et al.* 2002; Rosendo; Rosa, 2011; Cassol, 2013).

O processamento digital de imagens (PDI) pode ser entendido como a manipulação de imagens por computador de modo que a entrada e a saída do processo sejam imagens. O intuito de se usar PDI é melhorar o aspecto visual de algumas características estruturais da imagem para o analista e fornecer outros subsídios para a sua interpretação, inclusive gerando produtos que possam ser posteriormente submetidos a outros processamentos (CÂMARA; VALERIANO; SOARES, 2013).

Tendo como sua base uma série de algoritmos a idéia por trás do processamento digital de imagens é trabalhar em um ambiente computacional com a matriz de *pixels* que formam as imagens afim de extrair rapidamente informações de interesse. Como resultado da aplicação de equações que comparam e agrupam os *pixels*, uma nova imagem digital é gerada destacando características e permitindo manipulações adicionais (Shimabukuro, 1996). É grande o número de processamentos permitidos pelas técnicas de PDI incluindo: realce, composição colorida de bandas, filtragem, classificações entre outras. Essas diferentes formas de manipulação de imagens digitais podem ser segundo Lillesand; Kiefer (1994) divididas em cinco grupos: modelagem biofísica, retificação e registro, integração, classificação e melhoramento. Devido as características desse trabalho, nesse item será abordado apenas o grupo de classificação.

Na recente história dos satélites de exploração de recursos terrestres, um dos principais objetivos é a extração de informações contidas nas imagens afim de traduzi-las em documentos, tabelas, gráficos e mapas (LILLESAND; KIEFER 1994). Convergindo com esse objetivo as técnicas de classificação de imagens digitais são responsáveis por categorizar os *pixels* presentes em uma imagem em classes de cobertura da terra ou temas, tornando o processo de mapeamento ou reconhecimento de características da superfície terrestre menos subjetivo e com maior potencial de repetição em situações subseqüentes (Novo 1992).

Para Shimabukuro (1996) o processo de classificação pode ser descrito como o reconhecimento de padrões associado a cada *pixel* da imagem no que

tange as características dos objetos em determinado ponto na superfície terrestre.

Existem diversos métodos de classificação, e esses podem ser divididos segundo diferentes critérios: classificação paramétrica e não paramétrica, classificação espectral e espacial, e classificação supervisionada e não supervisionada (Mather, 1987).

Classificadores paramétricos partem da premissa que os objetos no espaço de classificação seguem uma distribuição conhecida, geralmente normal, já os classificadores não paramétricos não demandam que os dados sigam alguma distribuição em particular, ou seja, os parâmetros estatísticos não definem o comportamento de uma classe como um todo (Lu *et al.* 2012).

Os classificadores espectrais geralmente utilizam somente a informação espectral de cada pixel para encontrar regiões homogêneas, a partir de medidas de distâncias ou de probabilidades de um pixel pertencer a uma classe específica (Mather, 1987).

Nos classificadores espaciais o processo de decisão leva em consideração agrupamentos de pixels que é usado como unidade de classificação. Classificadores espaciais podem considerar aspectos como textura, proximidade do pixel, tamanho, forma, direção e contexto e com isso procura-se simular o comportamento de um fotointérprete, ao reconhecer áreas homogêneas nas imagens baseadas nas propriedades espectrais e espaciais do objeto (Mather, 1987).

Em classificadores supervisionados o analista indica na imagem áreas que servirão como amostras estatísticas de classes previamente conhecidas, geralmente esses classificadores são utilizados quando o usuário tem um conhecimento prévio da área a ser classificada (Mather, 1987).

Em classificadores não supervisionados se utiliza um algoritmo de agrupamento ("clustering") afim de determinar o agrupamento dos dados, em uma feição espacial de dimensão igual ao número de bandas presentes assumindo que cada grupo ("cluster") representa a distribuição de probabilidade de uma classe (KURY; VIEIRA, 2001).

Independente do tipo de classificador, as características ou variáveis, tais como as bandas dos satélites de observação da terra, escolhidas para definir as bases dos padrões, são denominadas atributos. Neste contexto, um padrão é simplesmente um conjunto de medidas definidas, de natureza espectral ou espacial, no universo composto pelos atributos escolhidos (Swain, 1978; Mather, 1987). Para o reconhecimento desses padrões, segue-se o seguinte roteiro, primeiro, selecionar o atributo com a melhor capacidade de caracterizar padrões, segundo escolher um método adequado para realizar a comparação dos padrões descritivos dos objetos a serem classificados com os padrões não identificados presentes na imagem. O processo se encerra com a verificação do grau de exatidão do processo de alocação das categorias selecionadas (Mather, 1987).

3.2.1 Distancia Mínima (Euclidiana)

Considerado um classificador paramétrico, espectral e supervisionado, o Distancia Mínima funciona da seguinte forma: Primeiro os vetores médios das classes propostas são calculados através dos números digitais médios nas amostras de treinamento em todas as bandas selecionadas. Em seguida, a distancia de cada pixel para cada vetor médio é determinada. Posteriormente, os pixels de toda a imagem são definidos como pertencentes a classe cujo o vetor médio é mais próximo, separando o espaço n-dimensional em regiões de decisão, com divisões retilíneas eqüidistantes de dois ou mais centróides (Mather, 1987).

Dessa maneira é possível afirmar que cada pixel da imagem é associado a determinada classe através da análise da medida de similaridade de distância euclidiana, que é dada pela equação 01:

$$D(i) = \sqrt{(X_i - M_i)^2} \quad (01)$$

onde **X_i** é o valor do pixel a ser classificado, **M_i** é a média das classes e **N** é o número de bandas.

A figura 01 a baixo ilustra a disposição dos valores médios de 3 classes hipotéticas e valor do pixel a ser classificado. O classificador compara a distância euclidiana de cada pixel a média de cada agrupamento. O pixel então é designado a classe com a média mais próxima, isto é, à classe com a menor distancia Euclidiana.

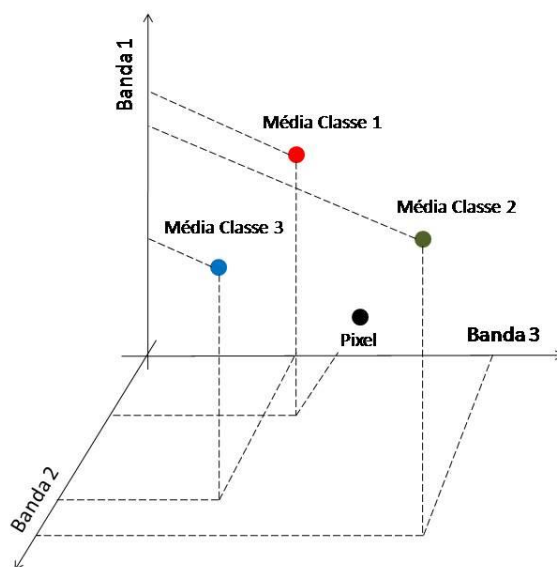


Figura 01: Gráfico de dispersão das médias das classes e do valor do pixel a ser classificado

Esse método apresenta como vantagem ser computacionalmente um dos mais simples, o que torna pouco dispendioso para o seu processamento, fator que colaborou para torná-lo um dos mais comumente utilizados (Jensen, 1996). Outra vantagem desse método é a não existência de pixels não classificados visto que sempre haverá uma media de classe mais próxima. Entretanto o classificador de Mínima Distancia também apresenta desvantagens como: a atribuição de classes a pixels que deveriam permanecer não classificados, isto é, pixels que não estão espectralmente próximos a nenhum das amostras dentro de limites razoáveis e ser insensível a variância espectral de cada classe, atribuindo classes erroneamente a pixels que tem valores distantes da média de sua classe original (Curran, 1985).

3.2.1 Máxima Verossimilhança

Assim como o Distância Mínima, a Máxima Verossimilhança (MaxVer) é um classificador paramétrico, espectral e supervisionado. O MaxVer considera a ponderação das distâncias às médias de cada classe no processo de designação dos pixels (Crósta, 1992). Assume que os dados tem uma distribuição normal ou gaussiana e calcula parâmetros estatísticos para os padrões de resposta espectral de cada classe (Castleman, 1979) a partir de vetores x que representam cada pixel nas áreas de treinamento:

$$x = \begin{pmatrix} \mathbf{B}_{xy1} \\ \vdots \\ \mathbf{B}_{xyk} \end{pmatrix} \quad (2)$$

onde \mathbf{B}_{xyk} é o número digital para (x,y)-ésimo pixel na banda k.

Com esses vetores é obtida a distribuição da assinatura espectral de cada classe por meio de um vetor médio μ_i e da matriz de covariância Σ_i :

$$\mu_i = \begin{pmatrix} \mu_{i1} \\ \vdots \\ \mu_{ik} \end{pmatrix} \quad (3)$$

$$\Sigma_i = \begin{pmatrix} \sigma_{i11} & \dots & \sigma_{i1l} \\ \vdots & \ddots & \vdots \\ \sigma_{ik1} & \dots & \sigma_{ikl} \end{pmatrix} \quad (4)$$

onde μ_{ik} representa o valor médio dos números digitais coletados para a classe i nas k bandas das amostras de treinamento e σ_{ikl} a covariância da classe i entre as bandas de k a l, definida por:

$$\sigma_{ikl} = \sum_{p=1}^N \frac{[x_k(p) - \mu_{ki}][x_l(p) - \mu_{li}]}{(N-1)} \quad (5)$$

onde N é o número de pixels na classe i da amostra de treinamento p e $x_k(p)$ e $x_l(p)$ são os números digitais dos pixels nas bandas k e l.

Depois de calcular a probabilidade estatística do vetor médio ligado a um pixel desconhecido pertencer a cada uma das classes definidas no processo de treinamento, o algoritmo, com base nesses parâmetros, associa o pixel a classe com maior probabilidade ou o mantém sem classe caso os valores sejam inferiores a um limiar estabelecido pelo analista.

A regra de decisão da máxima verossimilhança é baseada matematicamente na teoria de Bayes e pode ser apresentada de maneira sintética da seguinte forma: $x \in \omega_i$ se $g_i(x) \geq g_j(x)$ para $j \neq i$ e $i = 1, \dots, Z$ onde ω_i representa uma possível classe da imagem, Z o número de classes e $g_i(x)$ a função discriminante:

$$g_i(x) = -\ln/\Sigma_i / -(x - m_i)^t \Sigma_i^{-1} (x - m_i) \quad (6)$$

Recebe o nome de função de densidade de probabilidade, a representação gráfica da distribuição dos pixels de determinada classe nas suas respectivas áreas de treinamento em um diagrama de dispersão por meio de isolinhas de probabilidade. Os contornos gerados ao redor da média de cada classe estão relacionados a probabilidade da existência de um pixel na população definida pelo processo de treinamento com uma certa disposição de números digitais (Figura 02). Admitindo-se uma amostra adequada a variabilidade de uma determinada classe, esses elipsóides n-dimensionais também descreveriam a probabilidade de um pixel desconhecido e localizado em um ponto no espaço de atributos estar associado a uma determinada categoria. A forma dessas figuras equiprobabilísticas ainda indica o quão sensível é o classificador a covariância das áreas de treinamento (Crósta, 1992; LILLESAND; KIEFER, 1994).

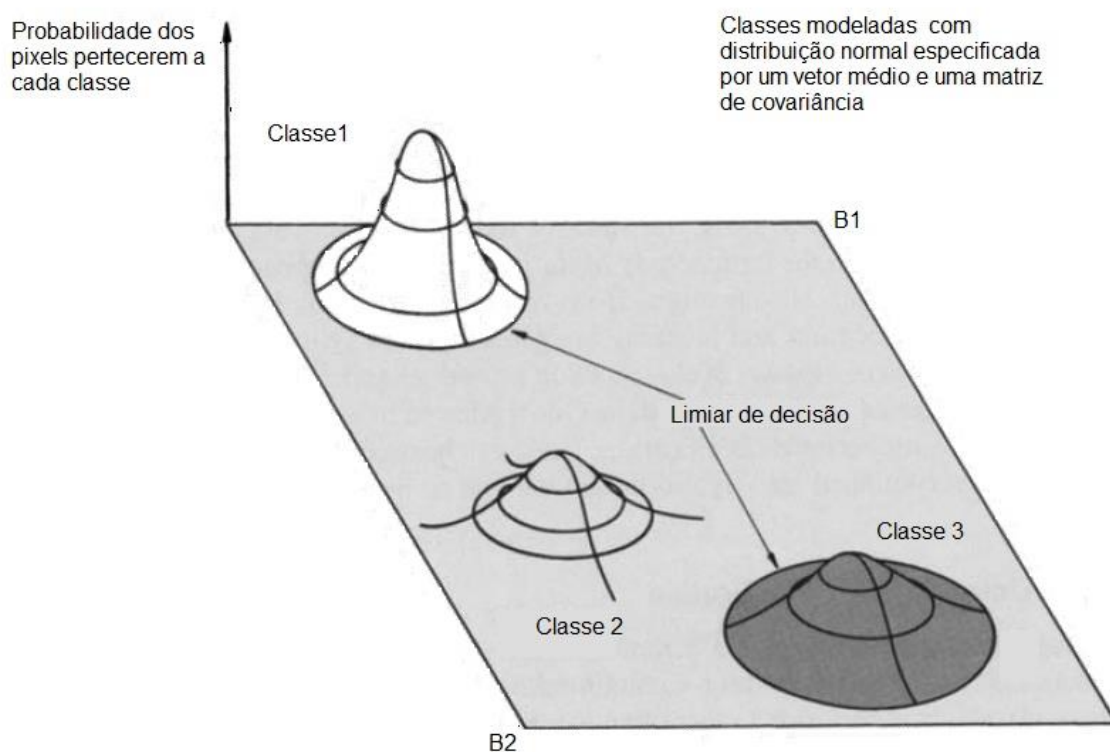


Figura 02: Diagrama de dispersão por meio de isolinhas de probabilidade

Além dessa abordagem de máxima verossimilhança, em situações onde se tem um conhecimento prévio de diferenças na probabilidade de ocorrência das classes é recomendável o emprego de uma regra de decisão Bayesiana (Richards, 1986). Alguns autores (Jensen, 1996; LILLESAND; KIEFER, 1994) propõe que esse classificador é uma extensão da técnica MaxVer. Apesar dessa perspectiva, Richards (1986) já havia demonstrado matematicamente que a regra de decisão de máxima verossimilhança é uma simplificação da regra Bayesiana onde, devido a inexistência de estimativas de probabilidade de ocorrência a priori de cada classe, se assume uma igualdade de seus valores.

O MaxVer ao longo dos anos, se consolidou como um método seguro e estável em situações onde a pressuposição de normalidade na distribuição dos dados no espaço de atributos n-dimensional se confirma (Shimabukuro, 1996).

3.2.1 *Extraction and Classification of Homogeneous Objects (ECHO)*

Proposto por Kettig (1975) o *Extraction and Classification of Homogeneous Objects (ECHO)* é um classificador espectral-espacial multi-estágio. O ECHO foi concebido para identificar objetos em dados multiespectrais, reunir as estatísticas desses objetos e, sempre que possível, realizar a classificação dos dados com base nos objetos identificados.

Classificadores *pixel-a-pixel*, como o Máxima Verossimilhança e o Distância Mínima, comparam as medidas espectrais de cada característica de cada *pixel* às estatísticas de classe, computando uma probabilidade ou função discriminante associada a cada classe e categorizando o *pixel* como pertencente à classe com a maior discriminante (Menezes et al, 2012). Cada *pixel* é classificado individualmente, com base apenas na medição espectral. Uma premissa desta técnica é que os objetos de interesse são maiores em comparação com o tamanho do *pixel*. Se não fosse assim, uma grande parte dos *pixels* seriam compostos de várias classes, tornando a classificação de padrões estatísticos pouco confiável, uma vez que categorias pré-especificadas seriam inadequadas para descrever a realidade da superfície classificada (Menezes et al, 2012). A partir desta premissa, entende-se que os objetos são representados por matrizes de *pixels*, e que existe uma dependência estatística entre *pixels* consecutivos. Os classificadores *pixel-a-pixel* não conseguem explorar a dependência estatística entre pontos adjacentes ao atribuir classes.

O ECHO se beneficia da informação espacial agregando em grupos *pixels* cujas respostas espectrais não são significativamente diferentes num sentido estatístico e aplicando uma regra de classificação de máxima verossimilhança a estes grupos homogêneos (Kettig, 1975). Objetos homogêneos são identificados em um processo de três etapas. Primeiro, as células são formadas sistematicamente dividindo os dados em N por N blocos de tamanho de *pixels*. As estatísticas de cada célula são então comparadas com um limiar de homogeneidade. Os *pixels* que não compreendem células homogêneas (isto é, *pixels* constituintes de células que não atendem ao critério de homogeneidade) são classificados em base *pixel a pixel*. As estatísticas de células homogêneas adjacentes são então comparadas. As células adjacentes

que parecem pertencer à mesma população estatística com base nos limiares de anexação fornecidos pelo analista são combinadas em um único objeto e classificadas (Kast, *et al.* 1977). A figura 03 desmonstra o processo acima descrito.

A definição dos limiares de homogeniedade e anexação pelo usuário exerce grande influencia no resultado obtido pelo ECHO pois eles determinam de que forma os dados serão agrupados em células e posteriormente anexados ou não para formar um novo objeto (Kast, *et al.* 1977). O limiar de homogeniedade funciona da seguinte forma:

$$Q_j(\mathbf{y}) = \sum_{i=1}^s (\mathbf{y}_i - \mathbf{m}_j)^t \mathbf{k}_j^{-1} (\mathbf{y}_i - \mathbf{m}_j) = \quad (07)$$

$$\text{tr}(\mathbf{k}_j^{-1} \sum_{i=1}^s \mathbf{y}_i \mathbf{y}_i^t) - 2\mathbf{m}_j^t \mathbf{k}_j^{-1} \sum_{i=1}^s \mathbf{y}_i + \mathbf{s} \mathbf{m}_j^t \mathbf{k}_j^{-1} \mathbf{m}_j \quad (08)$$

onde:

\mathbf{y}_i é o enésimo vetor do pixel na célula sendo testada

\mathbf{s} é o número de pixels na célula ($s = N^2$)

\mathbf{k}_j é a matriz de covariância da amostra para a j-ésima classe

\mathbf{m}_j é o vetor médio da amostra para a j-ésima classe de treinamento

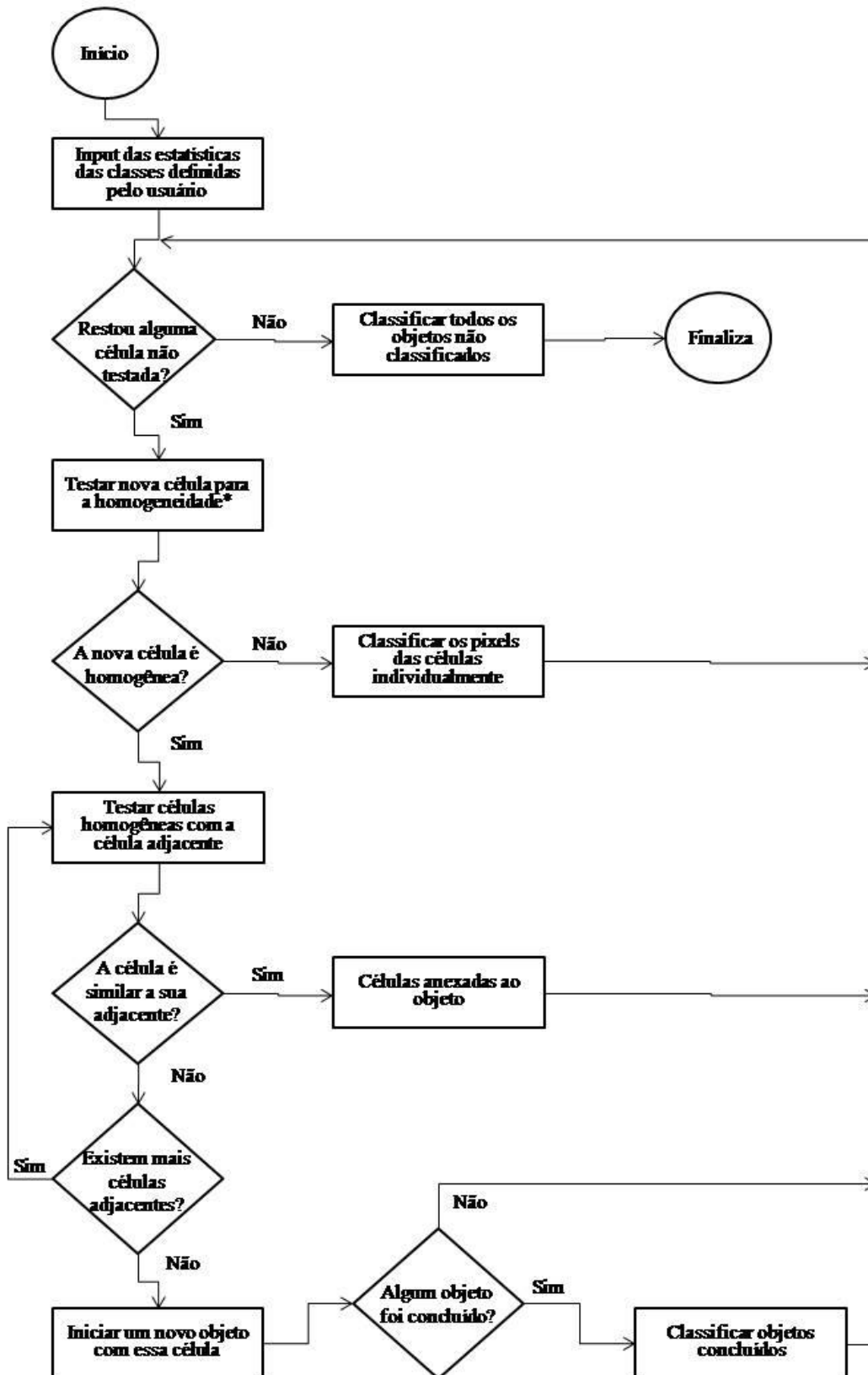


Figura 03: Fluxograma das etapas de processamento do ECHO (Adaptado de Kast, *et al.* 1977)

Esta forma quadrática é uma medida da distância estatística do conjunto de dados contidos na célula a partir da distribuição dos dados de treinamento para a j -ésima classe. Agora se w^* for a classe para a qual o "log-verossimilhança" da célula é máximo; Isto é, $\ln p(\underline{y}|w^*) = \max \ln p(\underline{y}|w_j) = \max_j [-s/2 \ln |2\pi k_j| - 1/2 Q_j(\underline{y})]$ e seja $Q^*(\underline{y})$ o valor da forma quadrática correspondente. Uma célula é definida como singular (e seus pixels classificados individualmente) se $Q^*(\underline{y}) > c$, onde c é um valor de limite especificado pelo usuário. Caso contrário, aceita-se a hipótese de que a célula \underline{y} é homogênea.

O uso deste critério tem a particular vantagem de que tende a "rejeitar" não apenas células não-homogêneas, mas também células "irreconhecíveis" (células muito improváveis de pertencer a qualquer uma das classes de treinamento). A distribuição dos valores Q_j pode ser mostrada como sendo Chi-quadrado com $s \cdot n$ graus de liberdade. Esse fato é usado na determinação de valores apropriados do parâmetro de limiar c (Kettig, 1975).

O limiar de anexação tem o papel de comparar um célula a um "campo" adjacente, que é simplesmente um grupo de uma ou mais células ligadas espacialmente que já foram fundidas. Se as duas amostras são estatisticamente semelhantes, de acordo com o teste abaixo (Equação 09), então a célula é mesclada ou "anexada" no campo. Caso contrário, a célula é comparada a outro campo adjacente, se existir, ou se torna um novo campo por si só (Kettig, 1975).

$$\Lambda = \frac{\max_i p(\mathbf{x}|\mathbf{W}_i) p(\mathbf{y}|\mathbf{w}_i)}{\max_i p(\mathbf{x}|\mathbf{W}_i) \max_j p(\mathbf{y}|\mathbf{w}_j)} \quad (09)$$

onde \mathbf{x} é o conjunto de pixels que formam o campo e \mathbf{y} é o conjunto de pixels da célula. Os valores de Λ ficam entre 0 e 1, e é mais próximo de 1 quando $p(\mathbf{x}|\mathbf{w}_i)$ e $p(\mathbf{y}|\mathbf{w}_i)$ têm seu valor máximo para a mesma classe.

Assim, o critério de anexação pode ser estabelecido da seguinte forma, considera-se que a célula pertence à mesma classe do campo é anexada ao

campo se $\Lambda \geq T$, em que T é um valor limiar ($0 \leq T \leq 1$). Caso contrário, a célula é considerada significativamente diferente do campo e não ocorre a anexação.

O ECHO tem mostrado bons resultados na classificação de cobertura do solo, em estudo ainda a ser publicado Gomes *et al.*(2017) identificaram e diferenciaram graus de infestação da gramínea exótica *Eragrostis plana* Nees na região dos pampas no interior do estado do Rio Grande do Sul. Utilizando uma metodologia que teve como base para a atribuição das estatísticas das classes de interesse um levantamento florístico e de densidade de infestação da espécie, os autores obtiveram resultados altamente satisfatórios com acurácia global de 90,3% e um índice Kappa de 88,5%. Na região amazônica o classificador apresentou desempenho similar, Mausel *et al.* (1993), Moran *et al.* (1994b), Brondízio *et al.* (1996) e Lu *et al.* (2004) o utilizaram para identificar estágios sucessionais da vegetação nas regiões de Altamira - PA, Ponta das Pedras na Ilha de Marajó - PA e Machadinho d'Oeste - RO e obtiveram valores superiores a 80% de acurácia global das classificações.

3.3 Modelo Linear de Mistura Espectral

Imagens de sensoriamento remoto quase sempre apresentam uma combinação de pixels puros e misturados. Dependendo da resolução espacial do sensor, os materiais na superfície podem ser menores do que um único elemento de cena da imagem (*pixel*). Desta forma, a radiância observada em um dado ponto é resultado de uma mistura espectral originada pela resposta de cada um destes componentes contidos no pixel (SHIMABUKURO; SMITH, 1995). Outra situação comum é quando esse pixel encontra-se no limite entre dois ou mais materiais diferentes. Nas duas situações, a radiância dos alvos é combinada pelo sensor para gerar o número digital referente aquele pixel nas diferentes bandas, dificultando a sua classificação.

Apesar da constante evolução tecnológica e do desenvolvimento de sensores orbitais com resoluções cada vez mais refinadas, a maioria dos dados ainda apresenta características espectralmente misturadas, isso ocorre devido a uma série de fatores como, a variação espacial dos fenômenos naturais sobre a superfície terrestre, a variação da geometria de aquisição dos

dados pelo sensor e as diferenças de iluminação na superfície devido a topografia da região imageada (SHIMABUKURO; SMITH, 1991)

Visando contornar esse problema o Modelo Linear de Mistura Espectral (MLME) tem como objetivo identificar a contribuição de cada alvo presente nos pixels de uma cena, permitindo que a imagem de satélite seja analisada de forma "suave", em uma escala *subpixel* (SCHWEIK; GREEN, 1999). A expressão "suave" se refere ao conceito de *soft classification* (classificação suave), onde os *pixels* não são necessariamente identificados como pertencentes a uma determinada classe temática, mas sim pela maior ou menor proporção dos componentes analisados pelo modelo de mistura.

O MLME está fundamentado na hipótese que o espectro misturado, numa imagem qualquer pode ser modelado por meio de uma combinação linear de espectros puros, conhecidos por membros de referência (SHIMABUKURO; SMITH, 1995). Dessa forma, para qualquer imagem, sendo as respostas espectrais dos componentes conhecidas, as proporções destes podem ser estimadas, de forma individual, por meio da geração de Imagens Fração (SHIMABUKURO; SMITH, 1991; Schowengerdt, 1997). Portanto, os membros de referência são o conjunto de espectros "puros" que serão utilizados em um modelo de mistura e que, quando combinados por regra apropriada, são capazes de reproduzir os demais espectros encontrados na imagem (Adams, 1995).

Uma variedade de métodos pode ser utilizada para determinar um membro de referência como, uma biblioteca espectral disponível para a área estudada, medidas de refletância obtidas em campo ou em laboratório, análise de componentes principais, índice de pureza do pixel e seleção manual diretamente da imagem (Quarmby *et al.* 1992; Boardman, 1993; Boardman, 1995; BATESON; CURTISS, 1996; Lu *et al.* 2003)

Pode se dizer que o MLME possibilita que o analista compreenda as variações encontradas nas assinaturas espectrais de diversos materiais, além de reduzir a dimensionalidade dos dados de uma imagem com n bandas, tornando-a apropriada para um novo processamento por meio das Imagens fração geradas (Aguiar *et al.* 1999).

A qualidade das imagens fração depende da seleção adequada dos pixels que serão utilizados como membros de referência e do número de membros de referência utilizados. Usualmente, o uso de 3 a 7 membros de referência se mostra apropriado para grande parte das aplicações, dependendo do número de bandas utilizadas e da variabilidade espectral dos componentes da imagem (MUSTARD; SUNSHINE, 1999).

A obtenção das imagens fração se dá por meio da aplicação do modelo linear de mistura apresentado abaixo:

$$R_i = \sum_{k=1}^n f_k R_{ik} + E_i \quad (10)$$

Onde:

i (i = 1,...,m): Número de bandas utilizadas

k = 1,...,m: Número de membros de referência (endmembers)

R_i: Refletância de um *pixel* na banda i, que pode conter um ou mais membros de referência

F_k: Proporção do membro de referência k dentro do pixel

R_{ik}: Refletância do membro de referência K dentro do pixel na banda i

E_i: Erro para banda i

O modelo deve seguir as restrições $0 \leq f_k \leq 1$ e $\sum f_k = 1$, que norteiam a análise das proporções de cada componente amostrado na imagem. Imagens fração podem ser produzidas a partir de diferentes métodos, sendo os mais populares o WLS (Weighted Least Square) e o CLS (Constrained Least Square), ambos voltados para minimizar a soma dos quadrados dos erros (valores residuais), sujeitos a aparecer devido as restrições do modelo (Aguiar et al., 1999; SHIMABUKURO; SMITH, 1995)

Assim, os valores dos pixels nas imagens fração variam de 0 a 1, indicando, respectivamente, a menor e a maior proporção do componente identificado dentro da imagem.

Além das imagens fração, o processamento do MLME também gera uma imagem resíduo, que tem como principal função indicar a qualidade dos

membros de referência utilizados no modelo. Analisada de forma inversa as imagens fração, na imagem RSME quanto maior o valor do pixel, maior o erro atribuído a ele (AGUIAR et al., 1999). Baseando-se na imagem RSME é possível verificar se os membros de referência usados na aplicação do modelo foram corretamente selecionados.

O uso de imagens fração tem sido recorrente, substituindo e complementando outros métodos de processamento. O principal destaque para o seu uso é a metodologia do PRODES que utiliza imagens fração das componentes solo e sombra para estimar a taxa de desmatamento da Amazônia brasileira (CÂMARA; VALERIANO; SOARES, 2013). Freitas e Shimabukuro (2008) utilizaram as frações vegetação e solo geradas pela aplicação do MLME para o desenvolvimento de um método para análise de mudanças de uso da terra. Lu et al (2004) desenvolveram uma árvore de decisão baseada nos valores das imagens fração para identificar três diferentes estágios sucessionais de floresta tropical na Amazônia brasileira.

3.4 Avaliação de algoritmos de classificação de imagens para o mapeamento da cobertura da terra

O principal objetivo no uso de classificadores é produzir dados e mapas de cobertura da terra que serão utilizados como referência na tomada de decisão em projetos futuros. Erros são introduzidos no processo de classificação e por consequência no seu resultado final. Esses erros são derivados das várias etapas que envolvem o processo de classificação como: a aquisição da imagem, etapas de pré-processamento, amostras utilizadas, conversão de dados entre outros (Lunetta, 1991).

A utilização de dados oriundo de classificações digitais, dada a sua complexidade, exige que sejam conhecidos a confiabilidade de tais produtos (Congalton, 1991). Avaliar a acurácia dos resultados gerados por classificadores, sobre tudo aqueles destinados a mapear a cobertura terrestre, é de extrema importância para o usuário, pois permite que ela possa julgar se esses produtos são adequados para as aplicações a que se destinam (LILLESAND;KIEF, 1994).

Embora exista uma vasta literatura sobre avaliação da acurácia em áreas como engenharia e estatística, para Smits *et al.* (2010) ainda existe uma necessidade de entendimento da comunidade do sensoriamento remoto, em especial os usuários dos produtos gerados por classificadores, sobre o tema. A falta de consciência da comunidade tem duas implicações o uso ineficiente de dados e o uso inadequado dos classificadores que adiciona erros desnecessários.

Congalton (1991) e Janssen & van der Wel (1994) reafirmam a importância de usuários e pesquisadores que utilizam dados de sensoriamento remoto e principalmente aqueles fornecidos por classificadores tenham um bom conhecimento dos fatores envolvidos bem como das técnicas usadas para se realizar a avaliação de acurácia.

Até o começo dos anos de 1980, grande parte dos trabalhos que realizavam a classificação de cobertura do solo por imagens de satélite apresentavam apenas um único número para expressar a acurácia de suas classificações, em grande parte desses estudos esse valor era chamado de acurácia de local não específico, referindo-se apenas ao montante total das categorias identificadas, desconsiderando a referencia espacial (Congalton, 2008). Esse tipo de abordagem geralmente produz resultados muito altos de acurácia, entretanto geralmente enganosos devido a superestimação provocada pelo método (Congalton, 1991).

No decorrer desta década novas técnicas foram propostas por Rosenfeld *et al.* (1982) e Congalton *et al.* (1983), e resultando desses trabalhos Story; Congalton (1986) propuseram o uso da matriz de erro como método para avaliação da acurácia das classificações. A matriz de erro tornou-se o padrão para registrar a acurácia de classificações de dados de sensoriamento remoto e é recomendada em adição o uso da estatística Kappa (Cohen, 1960) como uma medida aceitável da acurácia (Shimabukuro, 1996).

3.4.1 Avaliação de Acurácia

A avaliação da acurácia dos dados produzidos por métodos de classificação tem por objetivo examinar a qualidade dos dados produzidos para que os usuários dos produtos oriundos desses processos possam avaliar a sua utilidade para as aplicações pretendidas (Stephen et al., 1998).

No clássico trabalho de Stephen et al (1998) o procedimento para realizar a avaliação de acurácia para mapeamentos temáticos é dividido em três componentes básicos: design de amostragem, design de resposta e procedimentos de análise e estimativa.

3.4.1.1 Design de Amostragem

O design de amostragem é responsável por definir as amostras que serão usadas como referência para realizar a avaliação de acurácia e é um dos elementos principais para realizar uma avaliação estatisticamente rigorosa dos resultados obtidos com um classificador (LEA ; CURTIS, 2010).

O desenvolvimento de um design de amostragem requer o conhecimento da distribuição das classes na paisagem, a determinação do tipo e número de amostras a serem coletadas e a definição de um esquema de amostragem para a seleção de amostras (Congalton, 2008).

3.4.1.1.1 Esquema de Classificação (EC)

A forma como será realizada amostragem é parcialmente direcionada por como as classes a serem identificadas estão distribuídas na paisagem. A distribuição dessas classes segundo Congalton (2008) é na realidade uma função de como o analista definiu o esquema de classificação (EC), que por sua vez, é uma forma de organizar as informações espaciais de maneira ordenada e lógica.

Um EC é composto por dois componentes fundamentais: um grupo de rótulos (ex. Floresta, água, agricultura, pasto) e uma série de regras que irá definir como esses rótulos serão aplicados. A falta de um conjunto de regras pode tornar a atribuição dos rótulos arbitrária e sem consistência, podendo

provocar erros e confusões na interpretação do produto produzido (Congalton, 2008). No desenvolvimento do EC também são definidos, o nível de detalhe da classificação (ex. número e complexidade das classes) e a unidade mínima de mapeamento, ou seja, o menor valor de área que poderá ser definido como uma classe.

3.4.1.1.2 Unidade da Amostra

A unidade da amostra é a unidade fundamental na qual a avaliação da acurácia é baseada, ela é o elo entre a localização espacial da classificação e sua correspondente espacial na superfície (Stehman et al. 1998). Existem quatro unidades de amostra possíveis de serem definidas: um pixel, um aglomerado de pixels, um polígono e um aglomerado de polígonos Stehman et al (1998) e Congalton (2008).

3.4.1.1.3 Tamanho da Amostra

O número adequado de amostras que é necessário para se realizar a avaliação da acurácia depende de fatores como: nível de erro aceitável nos dados, nível de confiança na estimativa, tempo para realizar a aquisição e o custo do procedimento (Baile et al. 1994)

O modelo matemático da distribuição binomial foi reconhecido por Aronoff (1982); Ginevan (1979) e Congalton (1988) como apropriado para avaliar a acurácia. O número de amostras necessário pode ser aferido a partir da distribuição binomial por meio da seguinte fórmula:

$$n = \frac{pq}{\left(\frac{E}{z\alpha}\right)^2} \quad (11)$$

onde **p** é a exatidão requerida; **q** é o resultado de 1-p; **E** é o erro permitido ou grau de precisão desejado e **zα** é a abcissa da curva normal onde define-se o início da área relacionada ao nível de significância α (Goodchild et al. 1994)

Taylor (1989) propõe que a determinação do número ideal de amostras deve ser baseado em um percentual da área total estudada. Entretanto o tamanho desse percentual não é consenso e varia de 0.48% (Congalton 1988b) a 2.7% (Taylor, 1989).

Outra abordagem é a apresentada por Justice & Townshend (1981) que define a área de uma amostra com base na resolução espacial e na precisão geométrica da imagem gerada:

$$A = [PD[1 + 2PG]]^2 \quad (12)$$

A formula indica a área mínima de um local de amostragem **A** em relação ao diâmetro do pixel no terreno **PD** em metros e a precisão geométrica de um pixel **PG** em unidades de pixels.

Congalton (2009) relata que o pensamento tradicional sobre técnicas de amostragem não se adéqua ao grande volume de dados das imagens de sensoriamento remoto, e propões que um mínimo de 50 amostras por classes, para áreas classificadas menores que 1 milhão de acres e com menos de 12 classes, seria o ideal.

Na segunda versão do guia de procedimentos para a avaliação da acurácia temática do inventário de vegetação produzido pelo Serviço de Parques Nacionais dos Estados Unidos (LEA; CURTIS, 2010), recomenda-se que sejam definidos o número de amostras por classes baseando-se na abundância e na fragmentação de cada classes na área mapeada. O guia estabelece três cenário possíveis para a definição do número de amostras por classes. No primeiro cenário a classe é abundante e recobre uma área maior que 50 hectares, no segundo cenário a classe é relativamente abundante em relação as outras e recobre uma área de pelo menos 8,33 hectares, mas não maior que 50 hectares no total e no terceiro cenário a classe é relativamente escassa e recobre uma área menor que 8,33 hectares no total. A definição do número de amostras em cada um desses cenários se dá conforme a tabela 01 abaixo:

Tabela 01: Número de amostras desejáveis em relação a extensão territorial da classe (Adaptado de LEA;

Área Total da Classe	Número de Amostras por Classe
> 50 hectares	30
8,33 a 50 hectares	0,6 por hectare
< 8,33 hectares	5

CURTIS, 2010)

Para Rocha (1992) apesar dos estudos realizados, ainda não foi encontrado um método consensual, sendo a seleção do número de amostras usualmente feita através de considerações práticas, balanceando a exeqüibilidade e a adequação estatística.

3.4.1.1.4 Esquemas de Amostragem

Dentre os esquemas de amostragem existentes, cinco são mais comumente utilizados: amostragem aleatória simples, amostragem regular, amostragem aleatória estratificada, amostragem por aglomerados (cluster) e a amostragem estratificada regular não alinhada (Congalton, 2009).

- Amostragem Aleatória Simples

Em uma amostra aleatória simples, cada unidade de amostra na área de estudo tem uma chance igual de ser selecionada. Na maioria dos casos, um gerador de números aleatórios é usado para pegar coordenadas x, y aleatórias para identificar as amostras a serem coletadas. A principal vantagem da amostragem aleatória simples são as boas propriedades estatísticas que resultam da seleção aleatória de amostras (Congalton, 2009).

- Amostragem Regular ou Sistemática

A amostragem regular é um método no qual as unidades de amostra são selecionadas em algum intervalo especificado e regular na área de estudo. Na maioria dos casos, a primeira amostra é selecionada aleatoriamente e cada amostra sucessiva é tomada em algum intervalo especificado a partir da

amostra inicial. A principal vantagem da amostragem regular é a facilidade de amostragem de forma um tanto uniforme em toda a área de estudo (Congalton, 2009). Para Bary e Baker (1968) a maior desvantagem desse método é a não atribuição de chances iguais de inclusão da amostra para cada unidade da população, resultando em uma estimativa de erro de amostragem não totalmente confiável.

- Amostragem Aleatória Estratificada

A amostragem aleatória estratificada é similar à amostragem aleatória simples, no entanto, utiliza-se algum conhecimento prévio sobre a área de estudo para dividi-la em grupos ou estratos e, em seguida, cada estrato é aleatoriamente amostrado. A principal vantagem da amostragem aleatória estratificada é que todos os estratos (isto é, classes de mapas), não importa quão pequenos, serão incluídos na amostra. Este fator é especialmente importante para garantir que amostras suficientes de classes com menor extensão espacial sejam coletadas. (Stephen *et al.* 1998).

- Amostragem por Aglomerados (Clusters)

Na amostragem por aglomerados as amostras são grupos de pixels onde cada elemento (pixel) deve ser único para somente uma unidade amostral (aglomerado) (Kish, 1968). Stephen *et al.* (1998) descrevem algumas estruturas comumente utilizadas para gerar os agrupamentos como blocos de pixels de variados tamanhos, arranjos lineares de pixels e parcelas com dimensões previamente definidas.

- Amostragem Estratificada Regular Não Alinhada

Esse método combina as vantagens da aleatoriedade e da estratificação da população com a facilidade de uma amostragem regular, enquanto busca evitar a inclusão de tendências devido a periodicidade (Congalton, 2009). Este

método é uma abordagem combinada que introduz mais aleatoriedade do que apenas um começo aleatório dentro de cada estrato.

3.4.1.2 Design de Resposta

O design de resposta é um protocolo para determinar a que classe de cobertura de terra uma unidade de amostra se refere. Stephen *et al.* (1998) dividem esse processo em duas componentes: protocolo de avaliação e protocolo de rotulagem.

3.4.1.2.1 Protocolo de Avaliação

O protocolo de avaliação consiste em uma série de procedimentos utilizados para coletar as informações que irão contribuir para a determinação da classificação de referência (Stephen *et al.* 1998). Sua primeira etapa é definir o tamanho, a geometria e a orientação do espaço que será escolhido como amostra, Atkson e Curran (1995) apud Stephen *et al.* (1998) caracterizam esse processo como a escolha da região de suporte espacial.

A segunda etapa do protocolo de avaliação é determinar de que forma a região de suporte espacial será analisada para determinar a classificação de referência. Stephen *et al.* (1998) listam uma série de opções que são utilizadas de forma recorrente na determinação da classificação de referência como: a avaliação visual da área de suporte espacial seguida de dados qualitativos coletados pelo analista em campo, registro da composição florística, estrutura do dossel vegetal e a distribuição do tamanho de árvores. Os autores ainda reforçam que, sejam os métodos para determinar a classificação de referência qualitativos ou quantitativos, o protocolo de avaliação deve estar de acordo com o conceito do analista de classificação livre de erros.

3.4.1.2.2 Protocolo de Rotulagem

O protocolo de rotulagem tem o papel de atribuir uma classe de cobertura da terra à unidade de amostragem com base nas informações fornecidas pelo protocolo de avaliação (Stephen *et al.* 1998). Sua premissa

básica é que cada amostra deve ser rotulada como pertencente a apenas uma classe de cobertura do solo, como nem sempre as amostras estarão localizadas em áreas homogêneas, o protocolo de rotulagem deve especificar de que forma essas amostras serão rotuladas a partir de um conjunto de regras baseadas nos dados fornecidos pelo protocolo de avaliação (Stephen *et al.* 1998).

3.4.1.3 Análises e Estimativa

Ao observarmos a literatura que aborda avaliação de acurácia no sensoriamento remoto é possível notar alguma convergência, grande parte das pesquisas aplica uma matriz de confusão e em seguida utiliza coeficientes para analisá-la (Congalton 1991; Lu *et al.* 2004; SUAREZ; CANDEIAS 2012; SILVA *et al.* 2011).

A matriz de confusão (Figura 04) é uma matriz quadrada de números dispostos em linhas e colunas que expressa o número de unidades de amostra atribuídas a uma categoria particular numa classificação em relação ao número de unidades de amostra atribuídas a uma outra categoria. Usualmente um dos conjuntos de dado é considerado como sendo correto (classificação de referencia), esse geralmente é representado nas colunas, enquanto a classificação gerada a partir dos dados de sensoriamento remoto são representados nas linhas (Congalton, 2009).

Uma matriz de confusão permite análises tanto descritivas quanto analíticas. As técnicas descritivas são relativamente simples e incluem cálculo da acurácia global que pode ser descrita como a probabilidade de um ponto selecionado de forma aleatória no mapa esteja corretamente classificado (Congalton, 1991). Seu calculo se dá pela soma do total dos pixels corretamente classificados (diagonal principal da matriz) dividida pelo número total de pixels da matriz de confusão.

		Dados de Referência				Total
		1	2	...	c	
Classificação	1	x_{11}	x_{12}	...	x_{1c}	x_{1+}
	2	x_{21}	x_{22}	...	x_{2c}	x_{2+}

	c	x_{c1}	x_{c2}	...	x_{cc}	x_{c+}
	Total	x_{+1}	x_{+2}	...	x_{+c}	N

Figura 04: Matriz de Confusão

onde:

x_{c1} : número de pontos da classe c (referência) classificados na classe 1(classificação)

x_{cc} : número de pontos corretamente classificados na classe c

x_{+c} : número total de pontos avaliados da classe c na referência

x_{c+} : número total de pontos avaliados da classe c na classificação

Ainda falando sobre as técnicas descritivas, é possível obter os valores de acurácia de cada classe individualmente. Diferentemente da acurácia global, que apresenta um valor único de acurácia para toda a classificação, o calculo da acurácia das classes individualmente pode ser realizado de duas formas que como consequência retornam dois valores de acurácia diferentes para a mesma classe. Esse dois valores são a Acurácia do Produtor (AP) e a Acurácia do Usuário (AU) (STEMAN; CZAPLEWSKI, 1998).

A acurácia do produtor pode ser entendida como a probabilidade de um ponto selecionado aleatoriamente classificado como c nos dados de referencia seja classificado como c pelo classificador e é obtida por meio da divisão do número de pixels corretamente classificadas na categoria c pelo número total de pixels dessa categoria nos dados de referencia.

$$AP = \frac{x_{cc}}{x_{+c}} \quad (13)$$

A acurácia do usuário é definida como a probabilidade de um ponto selecionado aleatoriamente classificado como c pelo classificador seja classificado como c nos dados de referencia e é obtida por meio da divisão do número de pixels corretamente classificadas na categoria c pelo número total de pixels que foram classificados como c .

$$AU = \frac{x_{cc}}{x_{c+}} \quad (14)$$

As técnicas analíticas são úteis para comparar métodos de classificação diferentes. Congalton (1991) enfatiza que o uso técnicas analíticas multivariadas discretas é apropriado quando se trabalha com dados oriundo do sensoriamento remoto. Tal afirmação fundamenta-se nas características desse tipo de dado que geralmente tem valores discretos ao invés de contínuos e distribuição binomial ou multinomial, portanto muitas técnicas estatísticas baseadas na distribuição normal não se aplicam a esses dados.

Uma das técnicas analíticas é a normalização ou padronização da matriz de confusão. Ao normalizar a matriz de confusão por meio de um processo iterativo e proporcional que força que o resultado da soma das linhas e das colunas seja 1, é possível comparar diretamente tanto a diagonal principal da matriz quanto os valores fora da diagonal obtidos por diferentes algoritmos de classificação (Congalton, 1991). Desta forma, as diferenças nos tamanhos de amostra usados para gerar as matrizes são eliminadas, além disso, uma vez que a soma de cada linha e coluna resulta em 1, o valor de cada célula individual da matriz pode rapidamente ser convertido em uma porcentagem multiplicando por 100. Portanto, o processo de normalização fornece uma maneira conveniente de comparar valores de células individuais entre matrizes de erro independentemente do número de amostras utilizadas para gerar a matriz (Congalton, 1998)

Outra técnica analítica é a análise Kappa proposta por Cohen (1960), uma das mais populares técnicas multivariadas para avaliar acurácia, a análise

Kappa é utilizada para determinar estatisticamente se uma matriz de confusão é significativamente diferente de outra (Bishop et al. 1975).

O resultado da aplicação de Kappa é a estatística KHAT, que retorna uma medida de concordância global, baseada na diferença entre a concordância real da classificação e a probabilidade de concordância (concordância ao acaso). A concordância real corresponde à concordância entre os dados da classificação e os dados de referência, na diagonal principal da matriz. A concordância ao acaso é dada pelo produto dos valores das linhas e colunas marginais (Congalton et al. 1983).

A medida de concordância KHAT é calculada por:

$$\mathbf{K} = \frac{N \sum_{i=1}^r \mathbf{n}_{ii} - \sum_{i=1}^r (\mathbf{n}_{i+} \mathbf{n}_{+i})}{N^2 - \sum_{i=1}^r (\mathbf{n}_{i+} \mathbf{n}_{+i})} \quad (15)$$

onde r é o número de linhas de uma matriz; \mathbf{n}_{ii} , o número de observações na linha i e coluna i ; \mathbf{n}_{i+} e \mathbf{n}_{+i} respectivamente os totais marginais das linhas i e coluna i e N , o número total de observações (Bishop et al. 1975).

Um valor KHAT é calculado para cada matriz de erro e é uma medida da concordância entre a classificação e os dados de referência. Intervalos de confiança em torno do valor de KHAT podem ser calculados usando a variância aproximada da amostra (Equação 16), esse fato fornece um meio para testar a significância da estatística KHAT para uma matriz de confusão ,permitindo assim determinar se a concordância entre a classificação e os dados de referencia é significativamente maior que 0 (classificação aleatória) (Congalton, 2009).

$$\mathbf{var}(\mathbf{K}) = \frac{1}{n} \left\{ \frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2-\theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4-4\theta_2^2)}{(1-\theta_2)^4} \right\} \quad (16)$$

Sendo:

$$\theta_1 = \frac{1}{n} \sum_{i=1}^k \mathbf{n}_{ii} \quad (17)$$

$$\theta_2 = \frac{1}{n^2} \sum_{i=1}^k \mathbf{n}_{i+} \mathbf{n}_{+i} \quad (18)$$

$$\theta_3 = \frac{1}{n^2} \sum_{i=1}^k \mathbf{n}_{ii} (\mathbf{n}_{i+} + \mathbf{n}_{+i}) \quad (19)$$

$$\theta_4 = \frac{1}{n^3} \sum_{i=1}^k \sum_{j=1}^k \mathbf{n}_{ij} (\mathbf{n}_{j+} + \mathbf{n}_{+j})^2 \quad (20)$$

Comparar valores de KHAT independentes, ou seja, oriundos de matrizes de confusão diferentes, permite segundo Congalton (2009) confrontar estatisticamente classificações produzidas por: dois analistas, o mesmo analista ao longo do tempo, dois algoritmos, dois tipos de imagem e até mesmo duas imagens de datas diferentes e ver quais produzem a maior acurácia.

Tanto a matriz de confusão simples como os testes de significância de matriz de confusão pareados dependem do desvio normal padrão da seguinte forma: Sendo K1 e K2 estimativas da estatística Kappa para as matrizes de confusão 1 e 2 respectivamente e var(K1) e var(K2) as estimativas correspondentes de variancia. O teste estatístico para verificar a significancia de uma unica matriz de confusão é expresso por:

$$\mathbf{Z} = \frac{\mathbf{K}_1}{\sqrt{\text{var}(\mathbf{K}_1)}} \quad (21)$$

Z é padronizado e tem distribuição normal. Dada a hipótese nula H0: K1 = 0, e a alternativa H1: K1 ≠ 0, H0 é rejeitada se $Z \geq Z_{\alpha/2}$, onde $\alpha/2$ é o nível de confiança do teste Z de duas caldas e os graus de liberdade são assumidos como infinitos.

O teste estatístico para verificar a significancia de duas matrizes de erro independentes é expresso por:

$$\mathbf{Z} = \frac{\mathbf{K}_1 - \mathbf{K}_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad (22)$$

Z é padronizado e tem distribuição normal. Dada a hipótese nula H0:(K1-K2)=0, e a alternativa H1:(K1-K2) ≠ 0, H0 é rejeitada se $Z \geq Z_{\alpha/2}$.

4. METODOLOGIA

Nesse capítulo serão descritas as etapas metodológicas utilizadas no desenvolvimento deste trabalho. O fluxograma abaixo (Figura 05) ilustra todas as etapas presentes nesse capítulo.

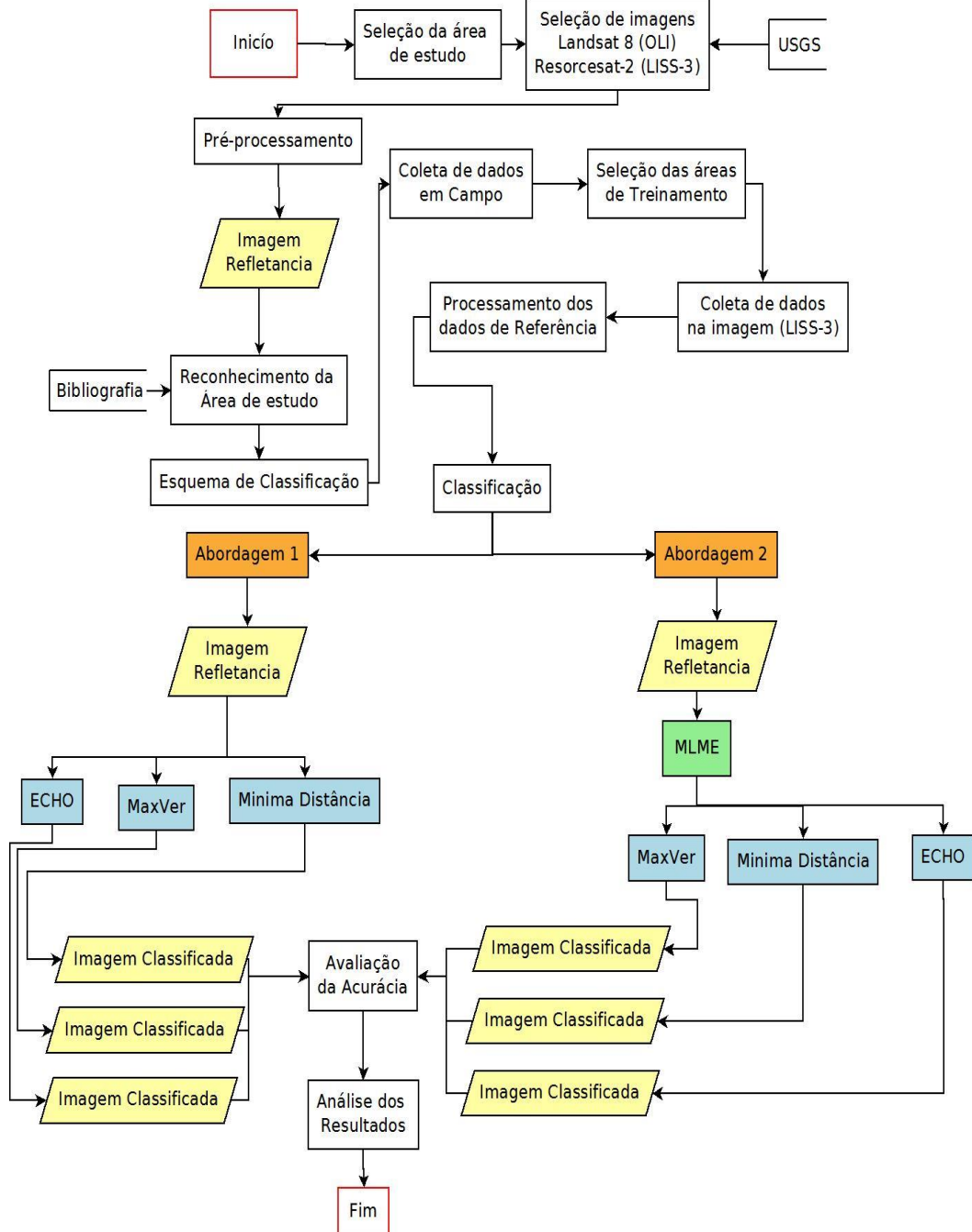


Figura 05: Fluxograma das etapas metodológicas do trabalho

4.1 Área de Estudo

A escolha da área de estudo gerou um processo de três etapas.

1. Identificação de uma região ou área que apresente grande variabilidade de tipos de cobertura do solo e que se encontra-se na região central da Amazônia.

Nesse contexto, o município de Tefé se destacou. Com uma área de 23.732 km², o município tem uma população estimada de 62.444 habitantes com uma densidade demográfica de 2,59 habitantes por km² (IBGE, 2016). Dentro de seus limites estão presentes a Terra Indígena Barreira da Missão, que ocupa uma área de 17,72 km² (Fundação Nacional do Índio, 2016), e parte da Floresta Nacional de Tefé.

Apesar de 81% de sua população residir na área urbana, a base da economia municipal fica dividida entre o setor agropecuário e o setor de serviços (IBGE, 2010). Devido à sua localização, às margens do rio Tefé e próximo à confluência dos rios Solimões e Japurá (Figura 06), principais vias fluviais para escoar a produção da região oeste da Amazônia (PAS, 2008), o município torna-se o principal entreposto comercial da região do médio Solimões, fator que favoreceu o desenvolvimento de uma infraestrutura que é inexistente em outras cidades da região. A presença de provedores de internet, agências bancárias, antenas de telefonia móvel e da Universidade Estadual do Amazonas viabilizaram o crescimento do setor de serviços, que elevou Tefé ao posto de pólo regional no Médio Solimões (PAS, 2008).

No setor agropecuário nota-se que a utilização do solo é desenvolvida por pequenas e médias propriedades, que no geral apresentam baixa produtividade. A produção é bastante diversificada em função das características da região, sendo os principais produtos a mandioca, utilizada principalmente na produção da farinha que tem grande demanda regional, e a banana. Outros produtos que têm uma produção em menor escala, como o milho, tomate, cana de açúcar e maracujá, são geralmente destinados às escolas municipais para a merenda escolar (IBGE, 2010).

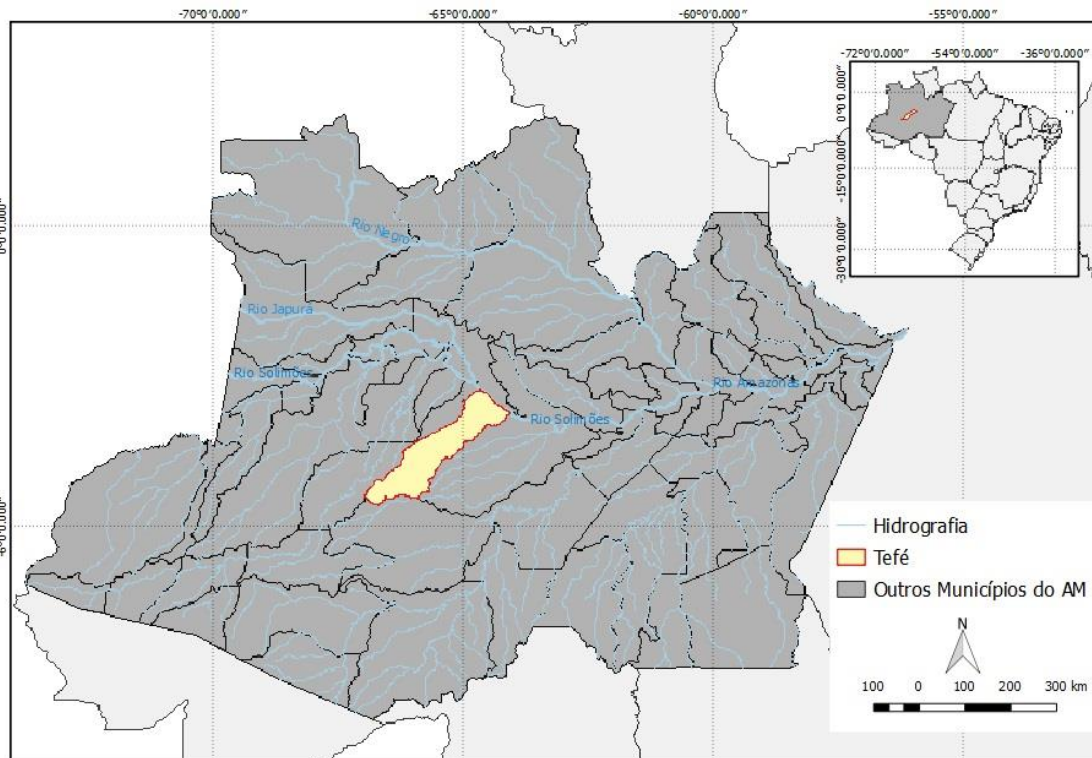


Figura 06: Mapa Hidrográfico do estado do Amazonas com destaque para o município de Tefé

Entretanto, a principal força do setor agropecuário do município são as atividades extrativistas que têm como principais produtos o açaí, a Castanha-do-Brasil e a madeira (IBGE, 2010). Esses produtos, devido ao seu elevado valor comercial no mercado nacional, impulsionam a atividade na região. As atividades do setor agropecuário se estabelecem ao longo das margens das duas estradas rurais do município (EMADE e Agrovila), fronteiras móveis de expansão das atividades antrópicas dentro do município, e em pequenas comunidades rurais localizadas nas margens do Rio Tefé e do Lago Caiambé.

Além das áreas onde são desenvolvidas atividades antrópicas o município conta ainda com uma vasta extensão de formações florestais, compreendendo um complexo mosaico de fitofisionomias devido à diversidade de condições ecológicas, políticas e socioeconômicas (Becker, 2001).

Segundo o mapeamento de vegetação realizado pelo projeto RADAM BRASIL (1978), Tefé apresenta três categorias de unidade de vegetação distintas em seu território, sendo elas: Floresta Ombrófila Densa, Floresta Ombrófila Aberta e Campinarana Arborizada.

O Manual técnico da vegetação brasileira (2012) caracteriza essas unidades da seguinte forma:

- Floresta Ombrófila Densa: Vegetação caracterizada por fanerófitos - subformas de vida macro e mesofanerófitos, além de lianas lenhosas e epífitas em abundância, que a diferenciam das outras classes de formações. Sua característica ecológica principal reside nos ambientes ombrófilos "amigo da chuva" que marcam muito a "região florística florestal". Assim, a característica ombrotérmica da Floresta Ombrófila Densa está presa a fatores climáticos tropicais de elevadas temperaturas (médias de 25 graus centígrados) e de alta precipitação, bem distribuída durante o ano (de 0 a 60 dias secos), o que determina uma situação bioecológica praticamente sem período biologicamente seco.
- Floresta Ombrófila Aberta: Considerada uma vegetação de transição entre a Floresta Amazônica e as áreas extra-amazônicas. Apresenta quatro faciações florísticas que alteram a fisionomia ecológica da Floresta Ombrófila Densa, imprimindo-lhe claros, daí advindo o nome adotado, além dos gradientes climáticos com mais de 60 dias secos por ano, assinalados na curva ombrotérmica.
- Campinarana Arborizada: Esta fitofisionomia mostra-se como uma transição das formas savânicas para as florestas de Campinarana. Nesta formação encontram-se principalmente espécies fanerófitas (*Pradosia schomburgkiana* subsp. *schomburgkiana*, *Clusia nemorosa*, *Ouratea spruceana*, *Talisia ghilleana*, *Protium heptaphyllum* subsp. *ulei* e *Pagamea duckei*) de troncos finos e elevada densidade. As lianas (*Dolioscarpus spraguei*, *Odontadenia verrucosa* e *Gnetum leyboldii*) são ocorrentes no local. Nessa fitofisionomia observa-se a formação de um dossel irregular que varia de 6 a 10m de altura. Espécies caméfitas e hemicriptófitas surgem por ainda haver incidência de

luz causada pela irregularidade do dossel, permitindo principalmente a presença de *Lagenocarpus verticillatus*, *Psychotria hoffmannseggiana* e *Encyclia mapuerae*. Observa-se também a abundante presença de epífitas da família Orchidaceae (NOGUEIRA, 2014).

2. Identificar no município regiões que apresentam grande variabilidade nos tipos de uso e cobertura do solo

Comparar o desempenho de classificadores para toda a extensão territorial do município seria pouco eficiente, visto que, assim como grande parte dos municípios do Amazonas, Tefé tem a maior parte do seu território coberto por grandes áreas de floresta contínua, ou seja, um único tipo de cobertura (Figura 07).

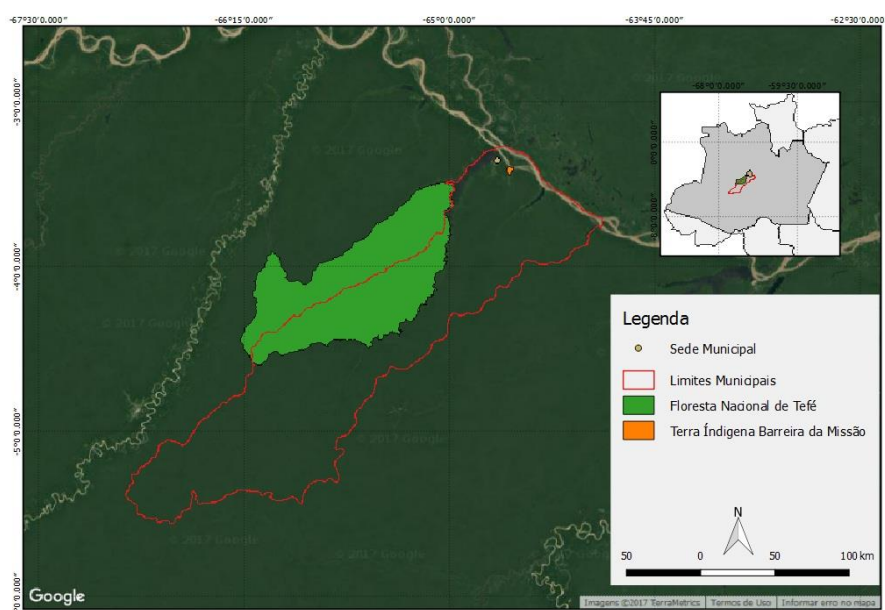


Figura 07: Mapa de localização da sede municipal de Tefé.

Avaliar o desempenho dos classificadores para área total do município exigiria um grande volume de recursos e tempo e seria pouco representativo quanto ao desempenho dos classificadores considerando que a variabilidade dos tipos de cobertura na área de estudo seria baixa.

A fim de otimizar o processo de avaliação de desempenho procurou-se identificar dentro do município uma região que apresentasse grande quantidade e variação de tipos de uso e cobertura do solo em um recorte espacial reduzido.

Para encontrar essa região foi utilizado o macro zoneamento proposto pelo plano diretor municipal. O plano diretor divide o município em 9 macro zonas distintas, sendo 8 definidas como zonas rurais e 1 definida como urbana (Anexo 01).

Dentre as zonas rurais, a Macro zona Novo Horizonte apresentou o maior número de atividades humanas sendo desenvolvidas. Dentro desta zona estão presentes as duas principais estradas rurais do município, a maior concentração das áreas de produção agrícola, a Terra Indígena Barreira das Missões e áreas de floresta.

3. Definição da Área de Estudo

A partir das etapas anteriores definiu-se o seguinte recorte espacial para a área de estudo (Figura 08) abrangendo a totalidade da zona urbana, parte do lago Tefé e a região de entorno das principais estradas rurais do município.

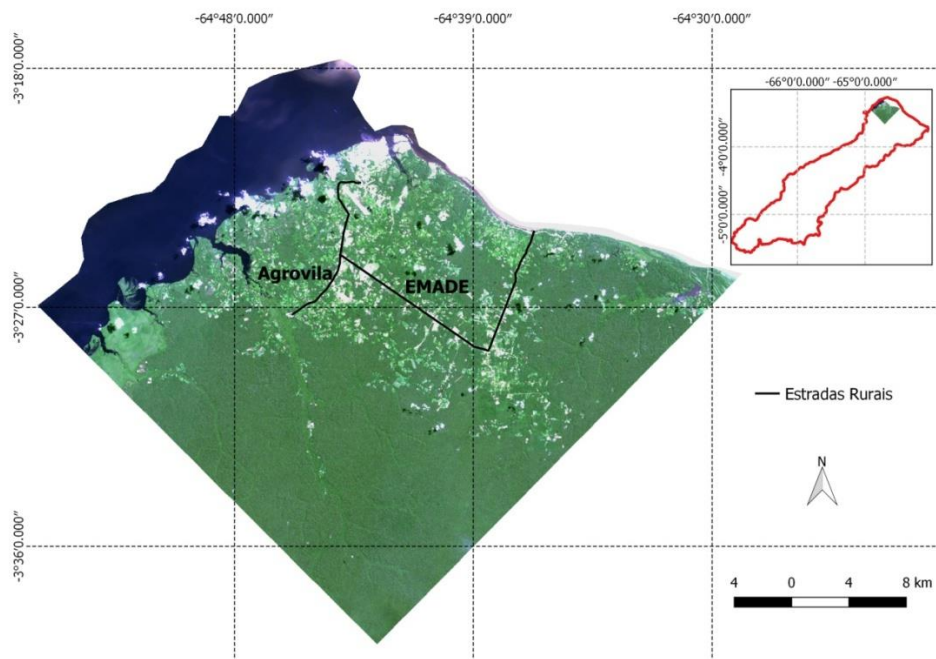


Figura 08: Área de Estudo

4.2 Seleção de Imagens

A seleção da imagem para esse trabalho seguiu os seguintes critérios: custo para aquisição das imagens, disponibilidade de cenas com baixa cobertura de nuvens para a área de estudo, frequência com a qual as imagens são utilizadas em estudos de classificação da cobertura da terra na região amazônica e resolução espectral.

Com base nesses critérios as imagens da série Landsat apresentaram aptidão para a realização deste trabalho. A série Landsat teve seu primeiro satélite lançado em 1972 e é a mais antiga série de satélites destinados exclusivamente a observação de recursos naturais terrestres.

Embarcados com sensores multiespectrais de resolução espacial média e com um período de revisita adequado para a realização de estudos de alteração da paisagem, os satélites da série se consolidaram como a principal fonte de imagens de satélite para estudos dedicados ao monitoramento de recursos naturais na superfície terrestre. Outro fator que corroborou para a consolidação do uso de imagens da série para estudos ambientais foi o acesso gratuito as imagens que podem ser adquiridas diretamente do site do Serviço Geológico do Estados Unidos (USGS) com diferentes níveis de processamento.

Lançado em 11 de fevereiro de 2013, o Landsat 8, último satélite da série, passou por uma atualização em suas especificações quando comparado com as versões anteriores da série. Nele estão embarcados dois sensores multiespectrais, o *Operational Land Imager* (OLI) e o *Thermal Infrared Sensor* (TIRS). O OLI tem além de sensores posicionados em faixas espectrais similares as presentes no sensor *Enhanced Thematic Mapper Plus* (ETM+) embarcado no Landsat 7, duas novas bandas destinadas a estudos de áreas costeiras, aerossóis e detecção de nuvens. Outra melhoria foi o aumento na resolução radiométrica das imagens geradas pelos sensores que agora são de 12 bits frente aos 8bits entregues pelo ETM+.

Devido a essas características optou-se pelo uso da imagens geradas pelo sensor OLI embarcado no Landsat 8 para a realização deste estudo. A orbita-ponto selecionada foi a 01-61 por recobrir a área de estudo. A seleção

da data da imagem foi baseada na proximidade com as datas das atividades de coleta de dados de referencia em campo e no percentual de cobertura de nuvens da cena (< 10%) sendo escolhida a imagem gerada no dia 10/06/2015 código LC80010622016198LGN00 com nível de processamento L1T.

4.2.1 Pré-processamento

Com o objetivo de aplicar os classificadores com base em dados físicos da superfície, os valores dos números digitais da imagem selecionada foram convertidos para valores de refletância segundo a equação 23.

$$p = d^2 \cdot L \cdot R_p \quad (23)$$

onde, d é a distancia entre a terra e o sol na data de aquisição da imagem, R_p é o coeficiente de conversão e L é a imagem de radiância.

A área de estudo foi então recortada da imagem de refletância, reduzindo assim o volume de dados a ser processado nas etapas seguintes.

4.3 Aplicação do Modelo Linear de Mistura Espectral

Os membros de referência utilizados para gerar as imagens fração de solo, sombra e vegetação foram extraídos diretamente da imagem (Shimabukuro, 1991). A seleção dos pixels que forneceram o comportamento dos referidos membros de referencia baseou-se nas atividades realizadas em campo que permitiram identificar regiões onde havia pouca ou nenhuma mistura de componentes na área referente ao pixel.

O pixel utilizado como membro de referência da fração solo tem a totalidade de sua área coberta por solo exposto (Figura 09A). O membro de referencia referente a fração sombra foi extraído de um pixel localizado no lago Tefé sobre influência de sombra provocada por nuvens no momento da aquisição da imagem (Figura 09B). O membro de referência para a fração vegetação foi extraído de uma área com vegetação secundária arbustiva densa, onde há pouca influência de sombreamento provocado por extratos de

vegetação mais desenvolvidos reduzindo assim a influência de possíveis sombreamentos (Figura 09C). A figura 10 mostra o comportamento espectral dos membros de referência selecionados. Com base nos pixels selecionados como membros de referência foram geradas as imagens fração.



Figura 09: Foto das áreas referentes aos *pixels* selecionados como membros de referência.

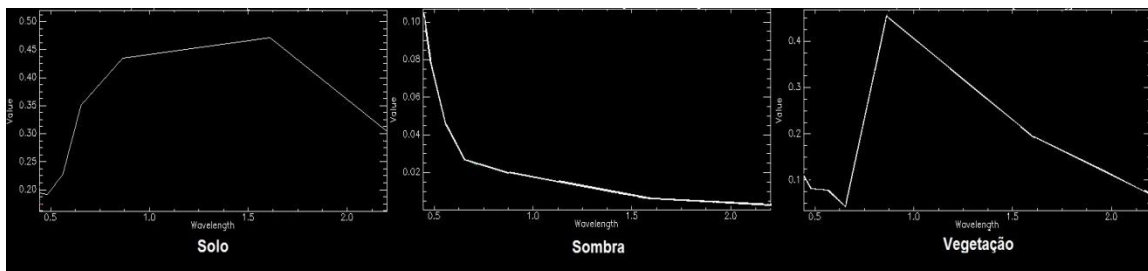


Figura 10: Curvas espectrais dos *pixels* selecionados como membros de referência.

4.4 Classificação

Com o objetivo de classificar a cobertura da terra segundo as classes de cobertura propostas no esquema de classificação apresentado no item 4.4.1, foram aplicados na imagem os seguintes classificadores: Mínima Distância, Máxima Verossimilhança e ECHO.

Para cada um dos classificadores foram utilizadas duas abordagens. A primeira abordagem foi aplicar os classificadores diretamente na imagem que teve seus números digitais convertidos para valores de refletância.

Na segunda abordagem os classificadores foram aplicados em uma composição das imagens fração (solo, sombra, vegetação), que foram geradas por meio da aplicação do Modelo Linear de Mistura espectral na imagem com valores de refletância.

Com o intuito de avaliar e comparar o desempenho dos classificadores e das abordagens propostas foram utilizados em ambas as abordagens as mesmas áreas de treinamento para os classificadores aplicados.

4.4.1 Esquema de Classificação

O esquema de classificação utilizado nesse estudo fundamentou-se nas diretrizes propostas pelo Manual Técnico de Uso da Terra desenvolvido pelo IBGE em 2013. Nesse contexto o esquema de classificação proposto destinou-se a identificar e diferenciar os principais tipos de cobertura terrestre presentes na área de estudo.

O sistema utilizado para identificar as classes de cobertura do solo é apresentado no fluxograma 03 (Figura 11), onde cada quadro amarelo representa uma classe de cobertura. Cada classe de cobertura foi rotulada segundo a tabela 02.

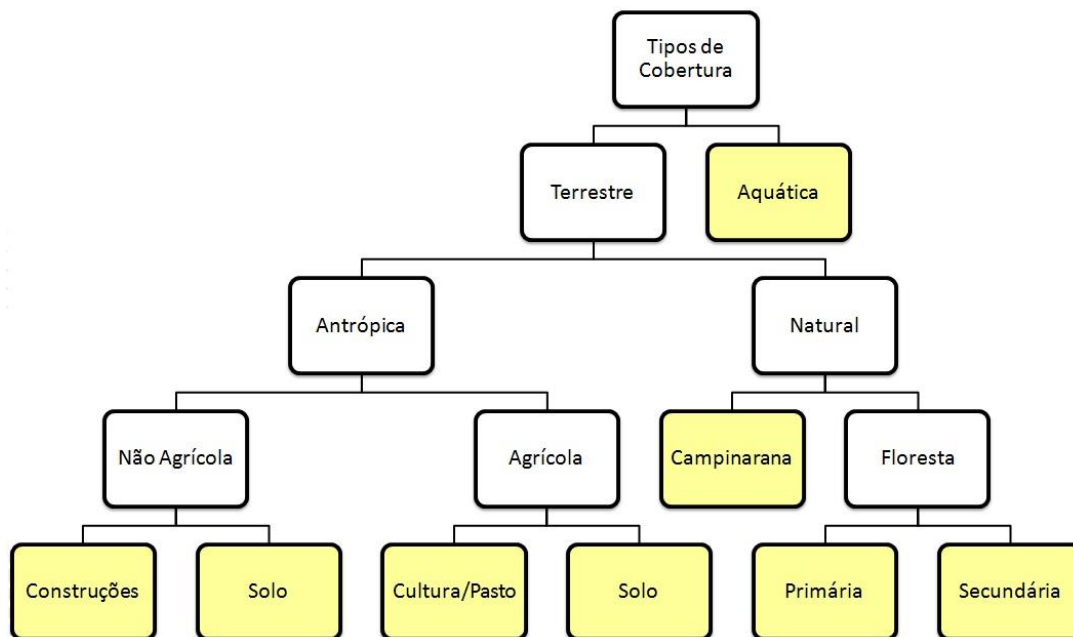



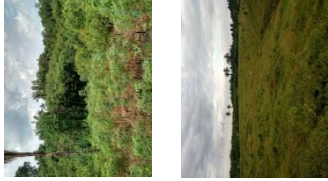





Figura 11: Fluxograma do Sistema para a identificação das classes de cobertura do solo em campo.

Tabela 02: Tabela descritiva do sistema de classificação.

Classes de Cobertura	Aquática	Construções	Solo	Cultura/ Pasto	Campinarana	Primária	Secundária
Rótulos	Água	Área Antrópica Urbana	Solo Expostos	Área Antrópica Agrícola	Campinarana	Floresta Primária	Floresta Secundária
Descrição	Toda e qualquer superfície recoberta por água como cursos de água e canais (rios, riachos, canais e outros corpos de água lineares), corpos d'água naturalmente fechados, sem movimento (lagos naturais regulados) e reservatórios artificiais	Superfície de uso intensivo, estruturadas por edificações e sistema viário, onde predominam superfícies artificiais não agrícolas	Superfície que não apresenta qualquer tipo de cobertura aquática, vegetal ou artificial	Superfície utilizada para a produção de alimentos, fibras e commodities do agronegócio. Está o inseridas nessa categoria lavouras (temporárias e permanentes) e pastagens.	Ecossistema amazônico sujeito a alagamentos periódicos com fisionomias florestais caracterizadas pela presença de um sub-bosque relativamente aberto e escassez de cipós e lianas, com indivíduos arbóreos podendo atingir 30 metros.	Superfície florestal caracterizada como de máxima expressão local, com grande diversidade biológica, onde os efeitos de ações antrópicas não afetaram significativamente a características originais de estrutura e de espécies.	Superfície florestal onde houve intervenção humana para o uso da terra, seja com finalidade agrícola, pecuária ou mineradora e que posteriormente foi abandonada descaracterizando a floresta primária.
Imagem							

A unidade mínima de mapeamento foi definida como a área correspondente ao pixel, visto que esta é a unidade básica das informações espectrais fornecidas pelas imagens.

4.4.2 Dados de Referência

A ida a campo é o método mais indicado para a coleta de dados de referência (CONGALTON, 2009) embora outros métodos como a utilização de mapeamentos previamente realizados e as imagens de sensores com melhor resolução espacial em comparação com a imagem utilizada para realizar a classificação também sejam indicados (LEA ; CURTIS, 2010).

Visando balancear da melhor forma possível o volume de recursos disponíveis para a realização desse estudo e a adequabilidade estatística do número de amostras necessárias para a uma avaliação de acurácia estatisticamente confiável, nesse estudo foram utilizados como dados de referência amostras coletadas em campo e extraídas de uma imagem do sensor LISS-3 do satélite Resourcesat-2, que tem 23 metros de resolução espacial e bandas posicionadas em comprimentos de onda similares aos presentes no Landsat 8.

4.4.2.1 Esquema de Amostragem

Para este estudo foi utilizado o esquema de amostragem aleatório simples. A opção por esse método deu-se devido a possibilidade de utilizar as amostras produzidas por esse esquema para avaliar os resultados produzidos por diferentes classificadores para a mesma região, sem a necessidade de gerar um novo conjunto de amostras para cada classificador testado, melhorando a relação custo/amostra para esse estudo.

Dada a grande extensão territorial da área de estudo e ao bioma na qual esta inserida, foi necessário definir um conjunto de regras para a realização do sorteio dos pontos de amostragem visto que a baixa densidade de rodovias, áreas com vegetação densa e a necessidade de documentos permitindo a

entrada em áreas protegidas elevam o custo e o tempo necessários para a realização das atividades em campo.

Buscou-se então identificar um raio partindo do ponto central da área urbana do município que englobasse as principais vias de acesso a área rural e seu entorno bem como um balanço entre o tempo de deslocamento até as amostras e o número mínimo de 10 amostras a serem coletadas por dia.

Dessa forma foi então definido que as amostras seriam sorteadas à uma distancia máxima de 25km do ponto central da área urbana do município excluindo-se a Terra Indígena Barreira das Missões.

A unidade amostral utilizada foi a área equivalente a resolução espacial das imagens geradas pelo Landsat 8, ou seja, uma área de 900 m². Utilizou-se o software QGis 2.14 para gerar uma camada vetorial com um total de 420 pares de coordenadas escolhidos de forma aleatória que posteriormente foram vinculados aos pixels na imagem com as coordenadas correspondentes (Figura 12).

Das 420 amostras que foram utilizadas como dados de referência, 182 tiveram seus dados coletados em campo e 238 foram extraídas de uma imagem do sensor LISS-3 do satélite Resourcesat-2.

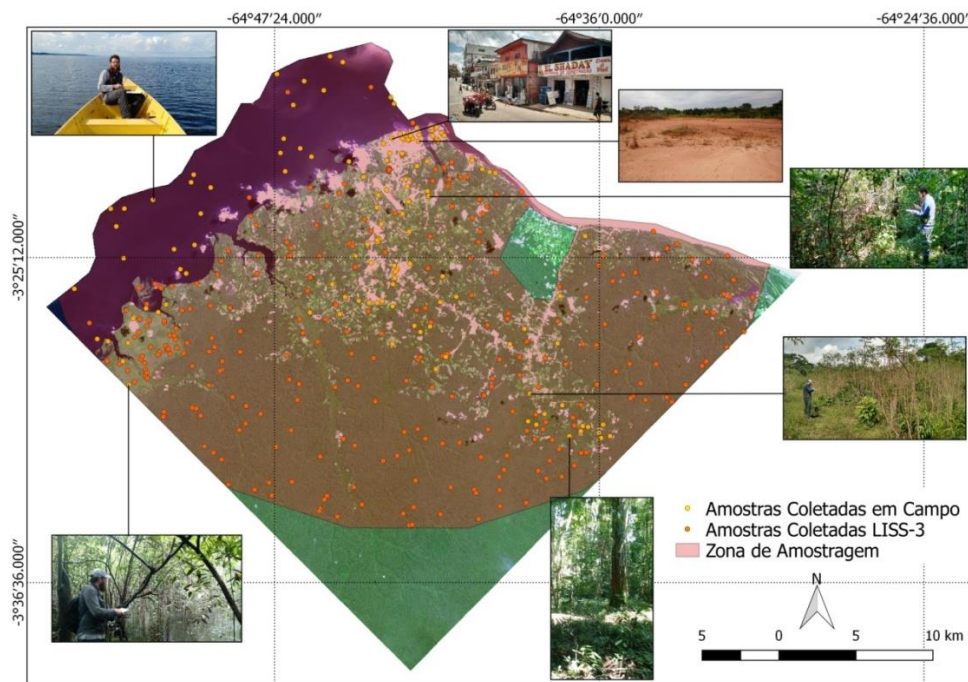


Figura 12: Mapa de localização dos pontos de amostragem

O uso da imagem gerada pelo sensor LISS-3 como fonte para os dados de referência se baseou na similaridade nas faixas do espectro eletromagnético a qual os seus sensores são sensíveis quando comparadas com os sensores do OLI (Venkateswarlu, 2014), na sua melhor resolução espacial e na disponibilidade de acesso gratuito as suas imagens.

4.4.2.2 Coleta de Dados em Campo

Localizada no entorno da Floresta Nacional de Tefé, a área de estudo está submetida a uma série de normativas legais quanto ao desenvolvimento de atividades de pesquisa. Afim de garantir a legalidade das atividades de coleta de dados em campo e ainda obter um suporte logístico e técnico para o seu desenvolvimento, foi solicitado junto ao ICMBio via Sistema de Autorização e Informação em Biodiversidade (SISBIO) uma autorização para a realização de atividades com finalidade científica (Anexo 2).

As atividades de reconhecimento in loco da área de estudo e coleta de dados em campo foram realizadas entre os dias 06 de julho de 2016 e 10 de agosto de 2016. Durante esse período foram realizadas um total de 13 expedições de campo conforme mostra a tabela 03. Todas as expedições foram realizadas com a presença de um auxiliar de campo com experiência em realizar esse tipo de atividade na região.

O acesso aos pontos de amostragem foi realizado utilizando um aparelho GPS Garmim modelo etrex20 onde foram colocados todos os pontos de amostragem .

O deslocamento até os pontos de amostragem foi realizado por via terrestre e por via aquática. Para acessar os pontos por via terrestre foi utilizada uma motocicleta. Quando não haviam vias de acesso que permitissem o uso da motocicleta, o acesso foi realizado pé (Figura 13 D,E,F). Os deslocamentos realizados por via aquática foram realizados em barcos de alumínio com motores de popa, popularmente chamados de "voadeiras" na região, e por canoas (Figura 13 A,B). Da mesma forma que os deslocamentos realizados por vias terrestres quando os meios de transporte anteriormente

citados se tornavam inviáveis devido as condições do ambiente, o acesso ao ponto de amostragem era feito a pé (Figura 12 C).

Tabela 03: Datas de realização das expedições para coleta dos pontos de amostragem

Número da expedição	Data	Número da expedição	Data
1	20/06/2016	8	12/07/2016
2	22/06/2016	9	13/07/2016
3	24/06/2016	10	15/07/2016
4	30/06/2016	11	18/07/2016
5	04/07/2016	12	21/07/2016
6	06/07/2016	13	27/07/2016
7	07/07/2016		



Figura 13: Coleta de dados em campo

O protocolo de avaliação das amostras deu-se da seguinte forma: verificação do local de amostragem via gps, registro do local por meio de 4

fotografias, sendo cada uma delas voltada para um ponto cardinal (Norte, Leste, Sul, Oeste). A região de suporte espacial foi definida como um raio de 15 metros partindo do par de coordenadas da amostra. Quando a visibilidade da área era inferior ao raio previamente estabelecido, se realiza um deslocamento de 10 metros a partir da coordenada nas direções norte, leste, sul e oeste.

Posteriormente era realizada uma análise visual da região de suporte espacial e o preenchimento da ficha de campo (Anexo 3). O protocolo de rotulagem das amostras seguiu o fluxograma 03 e a tabela 02 apresentados no item 4.4.1 baseando-se na análise visual, nas fichas de campo e nas fotografias retiradas durante o protocolo de avaliação.

4.4.2.3 Seleção das Áreas de Treinamento

Para que as classificações sejam executadas com sucesso é desejável que as classes de interesse sejam selecionadas cuidadosamente. Nesse contexto é necessário que sejam consideradas as diferenças entre as classes de informação, definidas pelo analista, e as classes espectrais, inerentes aos dados da imagem (JENSEN, 1996). Nesse contexto procurou-se traduzir as classes espectrais em classes de informação que atendessem os requisitos estipulados no esquema de classificação, analisando as resoluções, espacial e espectral do sensor bem como as características e proporções de cada tipo de alvo presente na imagem.

A definição das áreas de treinamento foi realizada com base na análise conjunta dos dados coletados em campo e na visualização de duas composições (RGB 432 e RGB 543) (Figura 14) da imagem da área de estudo. Optou-se pelo uso de diferentes composições nos canais RGB para facilitar a discriminação de classes que apresentaram respostas similares na composição RGB 432.

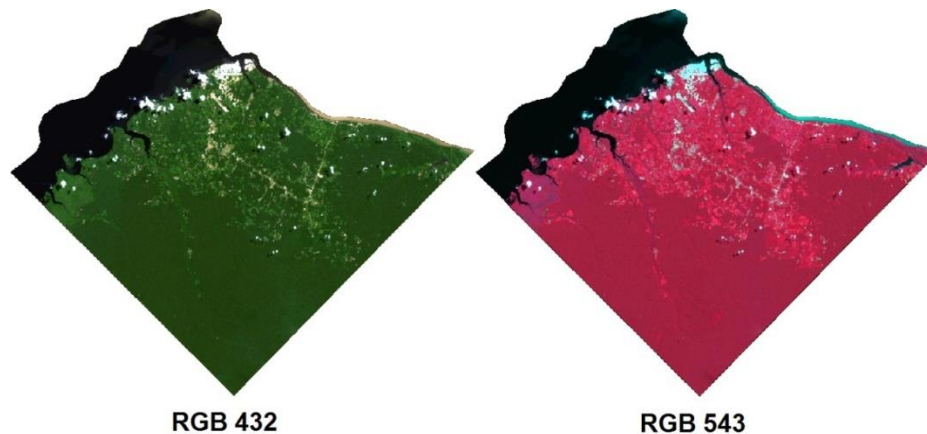


Figura 14: Comparação entre as composições de bandas RGB 432 e RGB 543.

Os polígonos referentes as áreas de treinamento foram delimitados cuidadosamente por meio do cursor, evitando a inclusão de pixels próximos as bordas das classes e buscando sempre um padrão homogêneo para cada classe. O tamanho médio das áreas de treinamento foi de aproximadamente 800 pixels.

Posteriormente a seleção das áreas de treinamento, foram calculadas a média e o desvio padrão dos valores de refletância dos seus pixels e suas curvas de comportamento espectral (Figura 15) afim de verificar se as amostras escolhidas para as classes eram espectralmente distintas.

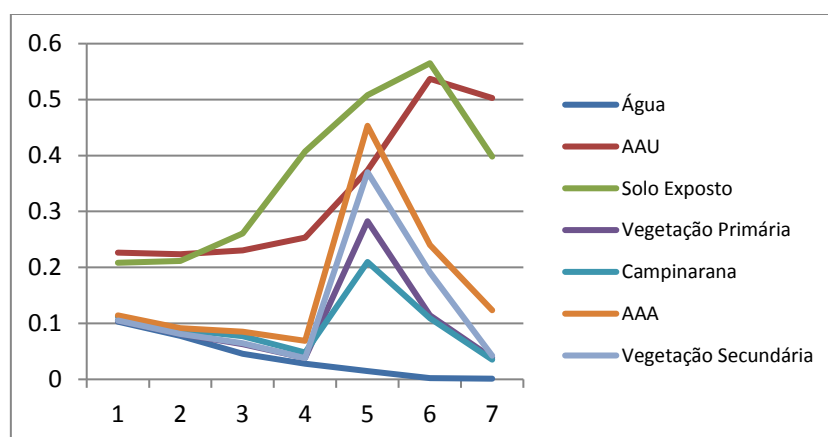


Figura 15: Gráfico de comportamento espectral da média da área de treinamento das classes da imagem Landsat-8, sendo o eixo X as bandas do sensor e o eixo Y os valores de refletância.

4.4.2.4 Coleta de Dados na Imagem Resurcesat-2

A imagem utilizada foi gerada no dia 37 de julho de 2016, data concomitante a realização das atividades de coleta de dados em campo. O protocolo de avaliação para as amostras extraídas da imagem baseou-se na comparação da resposta espectral apresentada pelo pixel identificado como amostra com a resposta espectral média apresentada pelas áreas de treinamento utilizadas para a aplicação dos classificadores.

Para realizar a comparação foram gerados gráficos de comportamento espectral a partir do valor apresentado por cada um dos 238 pixels e a partir da média dos valores apresentados pelos pixels pertencentes as áreas de treinamento utilizada para a identificação de cada uma das classes.

O protocolo de rotulagem das amostras balizou-se na similaridade entre o comportamento apresentado pela média das amostras e o comportamento de cada pixel individualmente. A avaliação de similaridade foi realizada com a sobreposição das curvas das classes com a curva do pixel, dessa forma, o pixel recebeu o rotulo a partir do comportamento a qual mais se aproximou no gráfico (Figura 16).

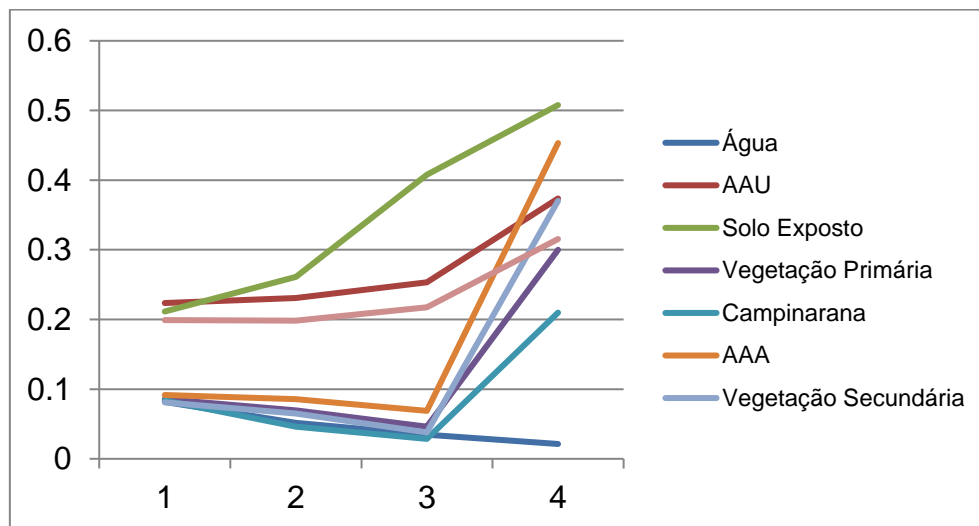


Figura 26: Gráfico de comportamento espectral da média das classes e em relação a uma amostra na imagem Resourcesat-2, sendo o eixo X as bandas do sensor e o eixo Y os valores de refletância.

4.4.2.5 Processamento dos Dados de Referência

Os dados de referência foram tabulados em uma planilha contendo seu código de identificação, suas respectivas coordenadas e sua classe correspondente. No software Qgis a planilha foi convertida para um arquivo vetorial de pontos

4.4.3 Classificação Abordagem 01

Nessa abordagem os classificadores foram aplicados diretamente sobre a imagem com os valores de refletância. A seqüência das etapas metodológicas dessa abordagem estão dispostas no Fluxograma 02.

4.4.4 Classificação Abordagem 02

Nessa abordagem os classificadores foram aplicados sobre uma composição das três imagens fração (solo, sombra, vegetação) obtidas como resultado da aplicação do MLME. A seqüência das etapas metodológicas dessa abordagem estão dispostas no fluxograma 02.

4.6 Análise de Acurácia

Nessa etapa o resultado produzido por cada um dos métodos de classificação foi submetido a uma avaliação de acurácia com base nos dados de referência. Para essa avaliação foi gerada uma matriz de confusão para cada um dos resultados obtidos, que originou valores de acurácia global, acurácia do usuário e acurácia do produtor. Posteriormente as foram submetidas a uma análise Kappa.

5. RESULTADOS E DISCUÇÃO

5.1 Classificação da Área de Estudo

Foram aplicados três classificadores (Mínima Distância, Máxima Verossimilhança, ECHO) baseados nas duas abordagens propostas, resultando em seis imagens classificadas da área de estudo. Todos os classificadores em ambas as abordagens foram capazes de identificar as sete classes de coberturas propostas (Água, Área Antrópica Urbana, Solo Exposto, Floresta Primária, Campinarana, Área Antrópica Agrícola e Floresta Secundária). Entretanto houve uma significativa variação nos resultados obtidos entre os classificadores e as abordagens. As figuras 17,18,19,20, 21, 22 e o anexos 4, 5, 6, 7, 8 e 9 mostram os resultados obtidos na seguinte ordem: Mínima Distância abordagem 1, Mínima Distância abordagem 2, Máxima Verossimilhança abordagem 1, Máxima Verossimilhança abordagem 2, ECHO abordagem 1 e ECHO abordagem 2.

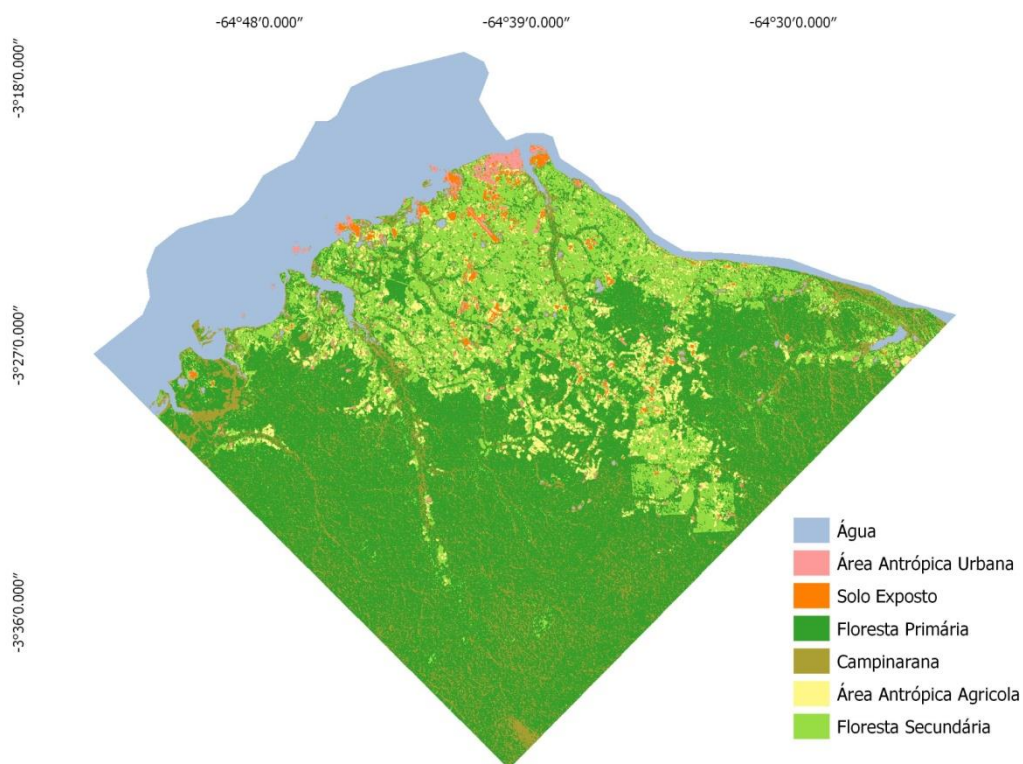


Figura 17: Classificação produzida pelo Mínima Distância abordagem 1

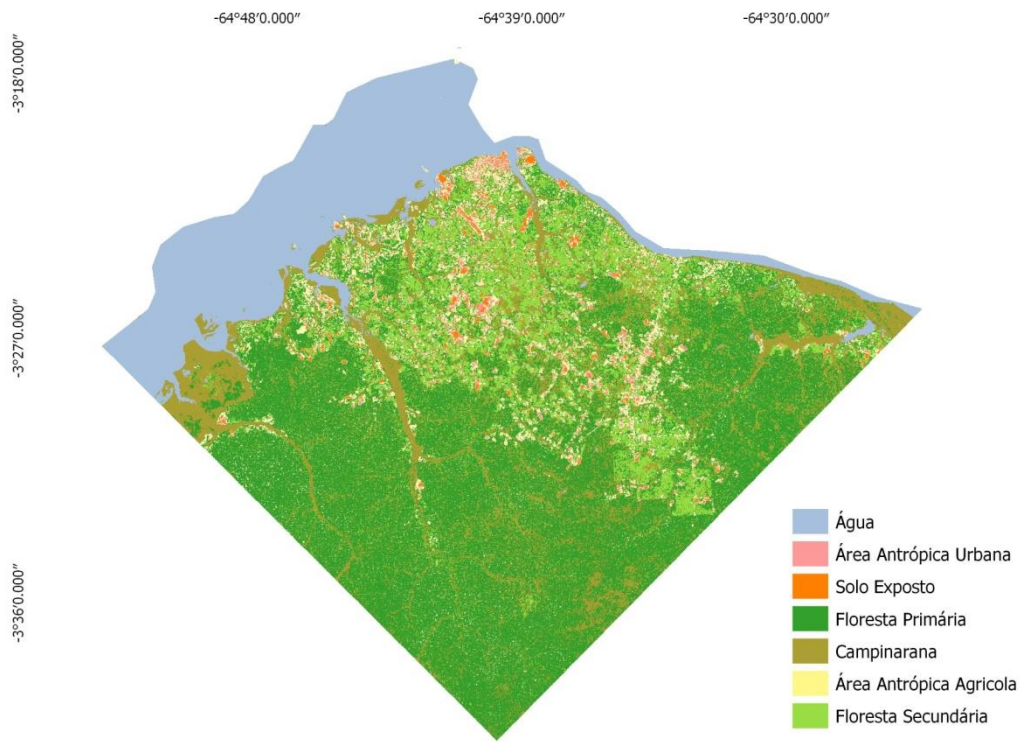


Figura 18: Classificação produzida pelo Mínima Distância abordagem 2

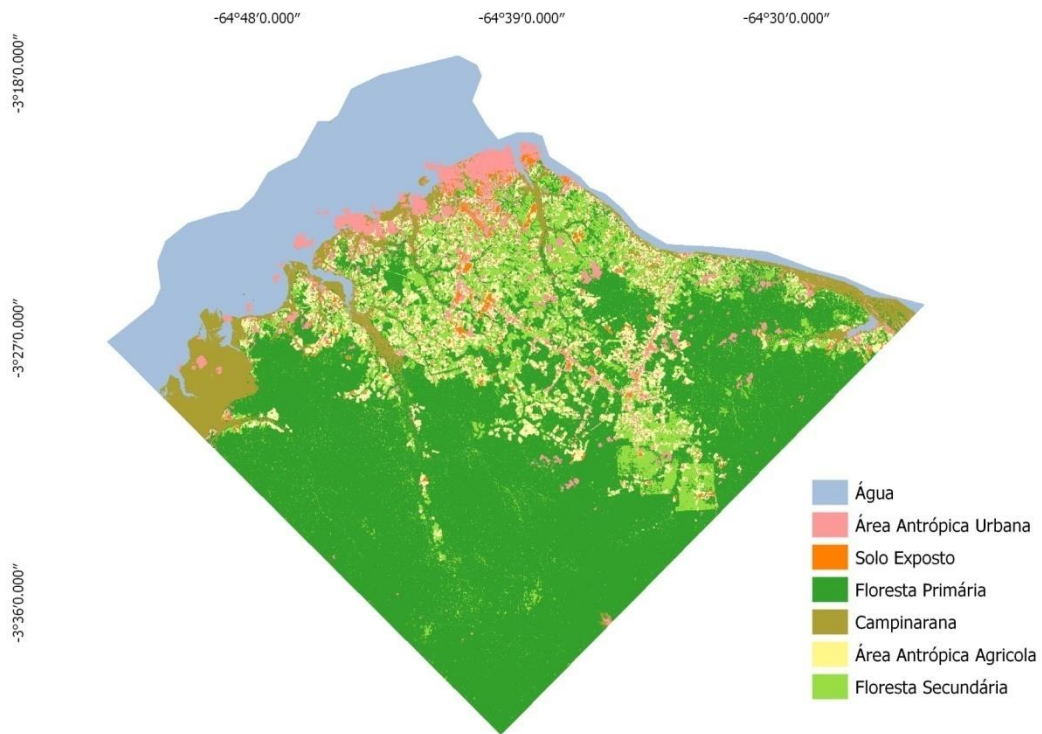


Figura 19: Classificação produzida pelo Máxima Verossimilhança abordagem 1

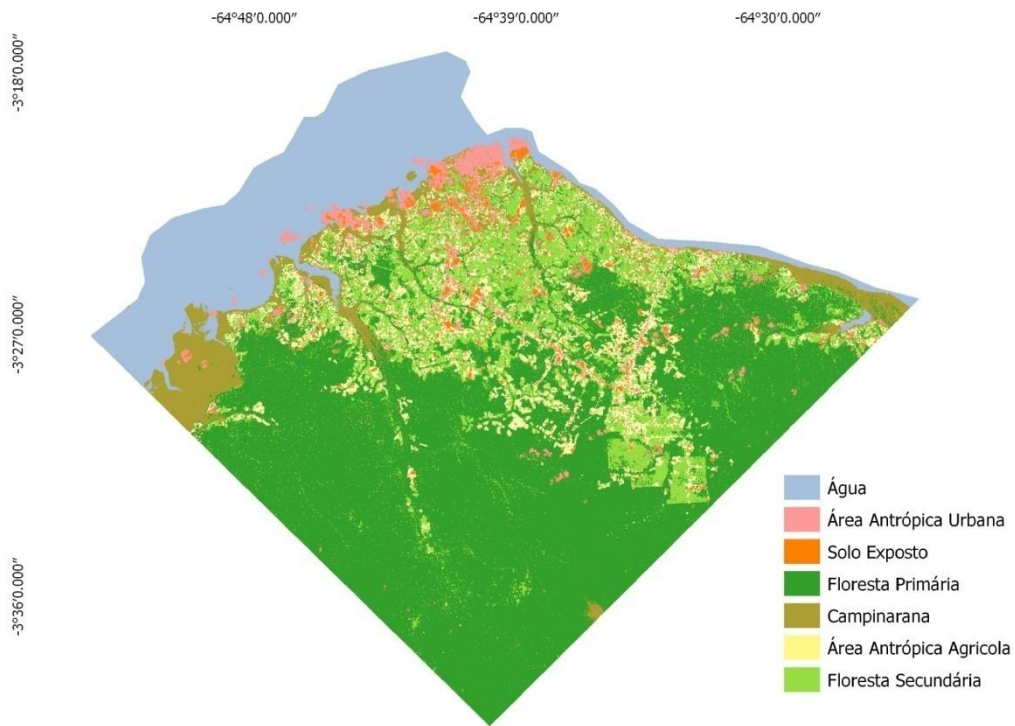


Figura 20: Classificação produzida pelo Máxima Verossimilhança abordagem 2

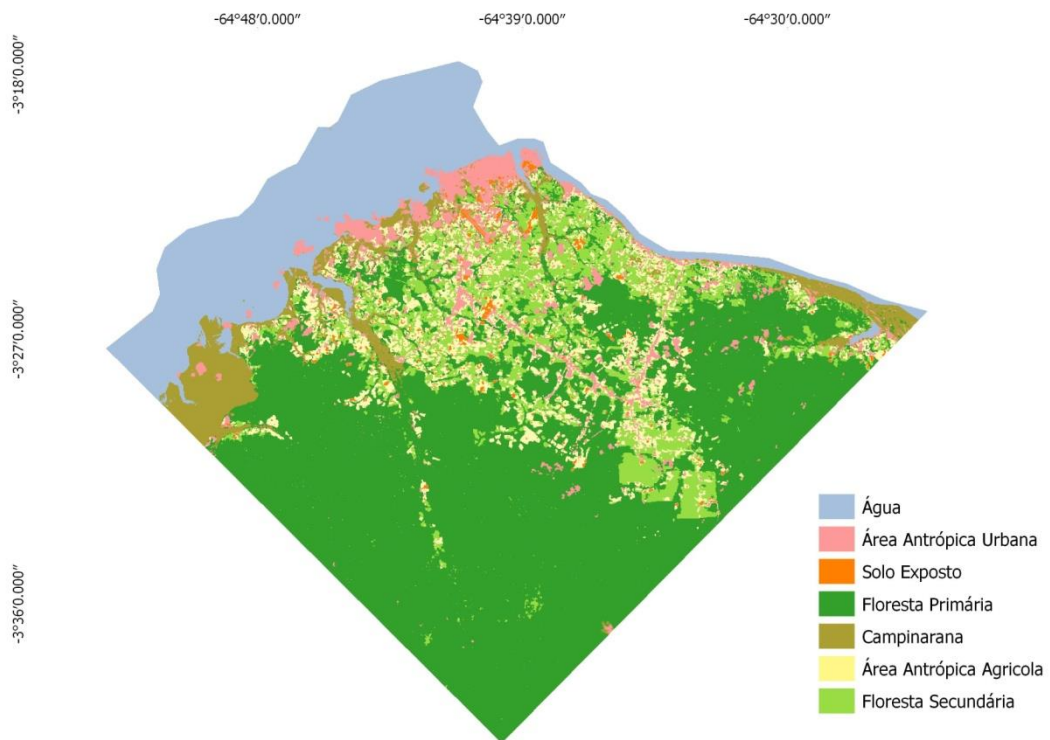


Figura 21: Classificação produzida pelo ECHO abordagem 1

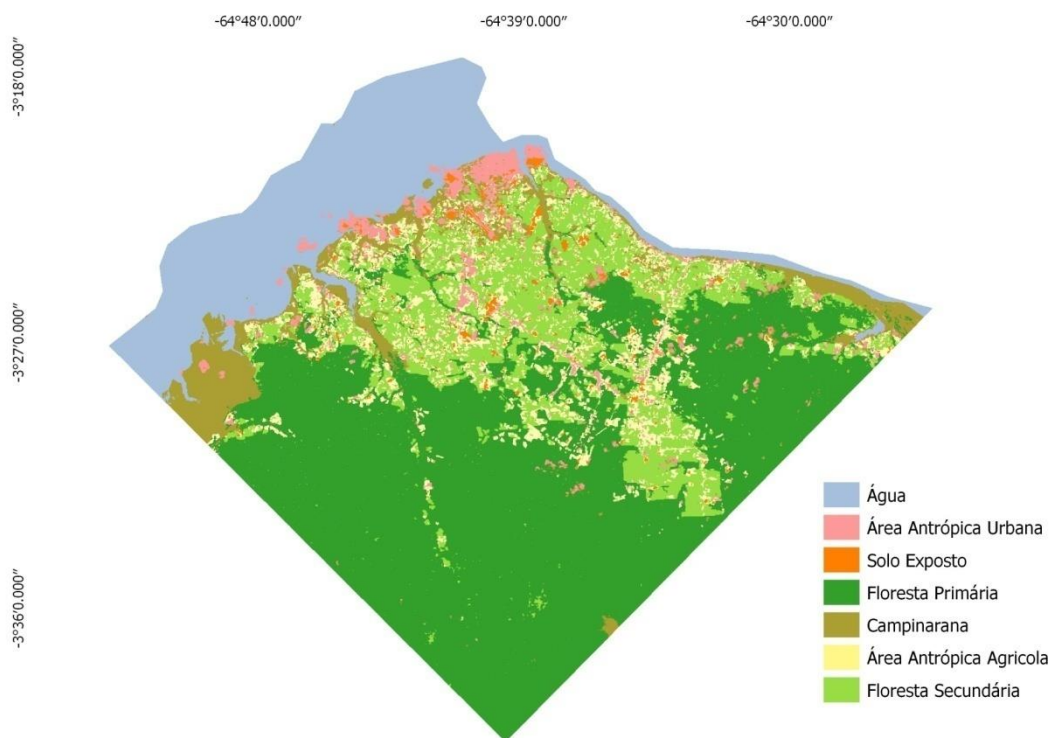


Figura 22: Classificação produzida pelo ECHO abordagem 2

5.2 Avaliação dos Classificadores

Para avaliar os resultados foi gerada com base nos dados de referência uma matriz de confusão para cada uma das imagens classificadas (Tabelas 4, 5, 6, 7, 8 e 9).

Tabela 4: Matriz de Confusão Mínima Distancia Abordagem 1

		Referência							Total
		Água	AAU	Solo Exposto	Floresta P	Campinarana	AAA	Floresta S	
Classes	Água	34	0	0	0	0	0	0	34
	AAU	0	24	10	0	1	0	0	35
	Solo Exposto	0	17	26	0	0	0	0	43
	Floresta P	0	0	0	103	20	0	4	127
	Campinarana	1	0	0	29	11	0	4	45
	AAA	0	0	0	0	0	49	4	53
	Floresta S	0	0	0	6	0	1	76	83
	Total	35	41	36	138	32	50	88	420

Tabela 5: Matriz de Confusão Mínima Distancia Abordagem 2

		Referência							
		Água	AAU	Solo Exposto	Floresta P	Campinarana	AAA	Floresta S	Total
Classes	Água	33	0	0	0	0	1	0	34
	AAU	0	19	7	0	0	9	0	35
	Solo Exposto	0	5	30	1	3	4	0	43
	Floresta P	0	0	0	88	22	4	13	127
	Campinarana	1	0	0	5	38	0	1	45
	AAA	0	1	0	9	2	19	22	53
	Floresta S	0	0	0	20	5	2	56	83
	Total	34	25	37	123	70	39	92	420

Tabela 6: Matriz de Confusão Máxima Verossimilhança Abordagem 1

		Referência							
		Água	AAU	Solo Exposto	Floresta P	Campinarana	AAA	Floresta S	Total
Classes	Água	34	0	0	0	0	0	0	34
	AAU	0	30	5	0	0	0	0	35
	Solo Exposto	0	4	39	0	0	0	0	43
	Floresta P	0	0	0	122	0	1	4	127
	Campinarana	0	1	0	0	44	0	0	45
	AAA	0	0	0	0	3	50	0	53
	Floresta S	0	0	0	8	1	1	73	83
	Total	34	35	44	130	48	52	77	420

Tabela 7: Matriz de Confusão Máxima Verossimilhança Abordagem 2

		Referência							
		Água	AAU	Solo Exposto	Floresta P	Campinarana	AAA	Floresta S	Total
Classes	Água	34	0	0	0	0	0	0	34
	AAU	0	28	7	0	0	0	0	35
	Solo Exposto	0	2	41	0	0	0	0	43
	Floresta P	0	0	0	123	0	0	4	127
	Campinarana	1	0	0	4	38	0	2	45
	AAA	0	0	0	0	0	50	3	53
	Floresta S	0	0	0	8	0	2	73	83
	Total	35	30	48	135	38	52	82	420

Tabela 8: Matriz de Confusão ECHO Abordagem 1

		Referência							
		Água	AAU	Solo Exposto	Floresta P	Campinarana	AAA	Floresta S	Total
Classes	Água	34	0	0	0	0	0	0	34
	AAU	0	30	5	0	0	0	0	35
	Solo Exposto	0	4	39	0	0	0	0	43
	Floresta P	0	4	0	115	0	0	8	127
	Campinarana	0	5	0	0	40	0	0	45
	AAA	0	2	0	0	2	49	0	53
	Floresta S	0	3	0	3	1	7	69	83
	Total	35	48	44	118	43	56	77	420

Tabela 9: Matriz de Confusão ECHO Abordagem 2

		Referência							
		Água	AAU	Solo Exposto	Floresta P	Campinarana	AAA	Floresta S	Total
Classes	Água	34	0	0	0	0	0	0	34
	AAU	0	26	9	0	0	0	0	35
	Solo Exposto	0	0	43	0	0	0	0	43
	Floresta P	0	1	0	113	3	0	10	127
	Campinarana	0	1	0	1	43	0	0	45
	AAA	0	0	0	0	4	48	1	53
	Floresta S	0	0	0	3	1	7	72	83
	Total	34	28	52	117	51	55	83	420

A partir das matrizes de confusão geradas foram calculados para cada uma das classes individualmente os valores de acurácia do produtor e acurácia do usuário, e os valores de acurácia global e o índice Kappa para aos classificadores e abordagens utilizadas (Tabela 10).

Posteriormente as imagens resultantes das classificações aplicadas foram convertidas para vetor e as classes tiveram suas áreas calculadas como apresentado na tabela 11.

Tabela 10: Indicadores de acurácia e índice Kappa dos classificadores

Classes	Classificadores											
	MD A1		MD A2		MaxVer A1		Maxver A2		ECHO A1		ECHO A2	
	AP	AU	AP	AU	AP	AU	AP	AU	AP	AU	AP	AU
Água	97,1%	100%	97,1%	97,1%	100%	100%	97,1%	100%	100%	100%	100%	100%
AAU	58,5%	68,5%	76%	54,2%	85,7%	85,7%	93,3%	80%	62,5%	85,7%	92,9%	74,2%
Solo Exposto	72,2%	60,4%	81,1%	69,7%	88,6%	90,6%	85,4%	95,3%	88,6%	90,6%	82,7%	100%
Floresta P	74,6%	81,1%	71,1%	69,2%	93,8%	96%	91,1%	96,8%	97,5%	90,5%	96,6%	88,9%
Campinarana	34,4%	24,4%	54,3%	84,4%	91,7%	97,7%	100%	84,4%	93%	88,8%	84,3%	95,5%
AAA	98%	92,4%	48,7%	35,8%	96,2%	94,3%	96,1%	94,3%	87,5%	92,4%	87,3%	90,5%
Floresta S	86,4%	91,5%	60,9%	67,4%	94,8%	87,9%	89%	87,9%	89,6%	93,1%	86,7%	96,7%
Acurácia Global	76,9%		67,4%		93,3%		92,1%		89,5%		90,2%	
Kappa(x100)	71,6%		60,1%		91,9%		90,4%		87,3%		88,1%	

Tabela 11: Área das classes

Classes	Classificadores					
	MD A1	MD A2	MaxVer A1	Maxver A2	ECHO A1	ECHO A2
	Área Ha					
Água	16.910	16.565	15.694	15.784	15.613	15.743
AAU	1.064	1.602	3.371	3.065	4.852	4.809
Solo Exposto	897	839	1.087	1.600	674	1.369
Floresta P	44.821	40.109	49.799	48.301	47.422	45.836
Campinarana	12.275	17.060	6.613	7.204	5.520	6.586
AAA	4.284	7.277	7.205	7.009	9.006	8.857
Floresta S	16.204	13.003	12.686	13.492	13.368	13.255
Total	96.455	96.455	96.455	96.455	96.455	96.455

5.2.1 Abordagem 1

O classificador que obteve o melhor desempenho na primeira abordagem foi o MaxVer apresentando uma acurácia global de 93,3% e um Kappa de 91,9%. Esse resultado indica o grande potencial que esse classificador tem para a classificação de cobertura do solo na região central da Amazônia brasileira. Com percentuais superiores a 85% de acurácia do produtor (AP) e acurácia do usuário (AU) em todas as classes testadas, a

aplicação do MaxVer seguindo a abordagem proposta se mostrou eficiente em diferenciar classes de cobertura com comportamento espectral muito similares como Floresta Primária, Campinarana e Floresta Secundária. Apesar do bom desempenho apresentado pelo classificador é possível notar o efeito "sal pimenta" em áreas que deveriam apresentar cobertura homogênea (Figura 12).

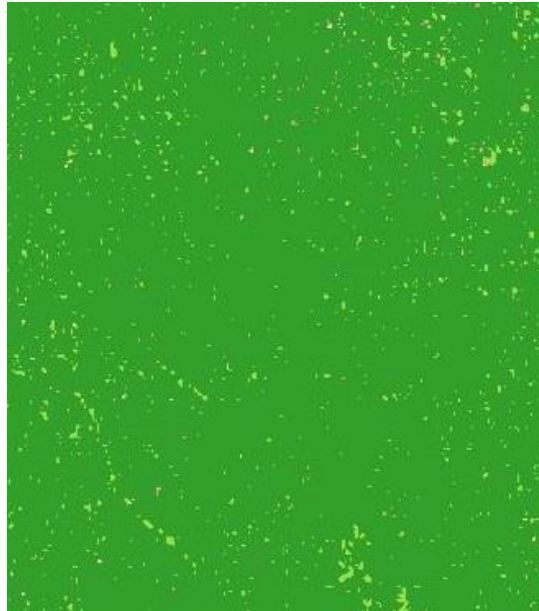


Figura 23: Efeito "sal pimenta" na imagem classificada pelo MaxVer na abordagem1

O ECHO teve o segundo melhor desempenho nesta abordagem e assim como o MaxVer apresentou bons valores de acurácia global (AG) (89,5%) e Kappa (87,3%). A pequena diferença de desempenho entre o ECHO e o MaxVer pode ser atribuída em grande parte à dificuldade que o ECHO apresentou em identificar a classe AAU. Dentre as sete classes essa foi a que apresentou os menores valores de AP (62,5%) e de AU (85,7%).

Ao se observar a coluna referente a classe AAU na matriz de confusão (Tabela 8) é possível notar que o baixo valor de AP apresentado é oriundo da dificuldade do classificador em diferenciar essa classe das demais sendo Água a única classe que não apresentou amostras que foram confundidas com AAU. É difícil explicar teoricamente esse fato, visto que o ECHO é um classificador que usa de informações espectrais e de textura para realizar as suas classificações, a Classe AAU tem um comportamento espectral próximo apenas à classe Solo Exposto e na área de estudo apresenta um distribuição

principalmente concentrada no núcleo urbano do município. Uma possível explicação pode ser a presença de nuvens, classificadas como AAU, e sua distribuição na imagem. Esse fenômeno pode ter influenciado o processo de agregação de células realizado pelo classificador, contribuindo assim para que pixels vizinhos às nuvens fossem identificados como tal.

Contudo o ECHO teve resultados bem próximos aos apresentados pelo MaxVer e obteve êxito em reduzir o efeito "sal pimenta" em áreas de cobertura homogênea (Figura 23).



Figura 23: Área de cobertura homogênea classificada pelo ECHO abordagem 1

Dentre os classificadores testados o Mínima Distância foi o que apresentou o pior desempenho, com uma AG de 76,9% e um índice Kappa de 71,6%. Apesar desses valores consideravelmente mais baixos em relação aos outros classificadores testados o desempenho do MD pode ser considerado satisfatório. Ao observarmos os valores de AP e AU das classes nota-se que o classificador apresentou um bom desempenho na identificação das classes Água (AP 97,1% e AU 100%), AAA (AP 98%, AU 92,4%) e Floresta Secundária (AP 86,4% AU 91,5%). As demais classes apresentaram alto grau de confusão com classes de comportamento espectral similar, nesse sentido a classe Campinarana foi a que teve o pior desempenho com os valores mais baixos de AP (34,4%) e AU (24,4%). Esse fenômeno de confusão entre classes

espectralmente próximas de certa forma era esperado visto que o MD tem como critério para a determinação da classe a qual o pixel vai pertencer apenas a distância euclidiana do valor do pixel em relação às médias dos valores apresentados pelas áreas de treinamento, impossibilitando assim a atribuição correta de classe a um pixel que esteja mais próximo do valor médio de uma classe espectralmente similar à sua classe do que de sua classe original.

Na matriz de confusão (Tabela 4) nota-se esse fato nas classes AAU, onde quase metade dos dados de referência foi atribuída a Classe Solo Exposto. Essa atribuição indevida também ocorre de maneira inversa com a classe Solo Exposto que teve 17 das suas 43 amostras atribuídas à classe AAU. Essa atribuição indevida dos dados amostrais aconteceu de maneira mais intensa com a classe Campinarana que teve 29 das suas 45 amostras, identificadas como Floresta Primária.

Assim como o MaxVer o MD apresentou um intenso efeito "sal pimenta" no seu resultado (Figura 20).

Entretanto apesar de ter tido o pior desempenho entre os classificadores o MD foi o classificador que melhor minimizou os efeitos provocados pela cobertura de nuvens (Figura 21). Diferente do que ocorreu no MaxVer e no ECHO, que classificaram nuvens e suas sombras como AAU, o MD classificou as nuvens como Solo Exposto e suas sombras como Água. Além dessa diferença quanto à classe destinada a nuvens e sombras, o MD foi capaz de classificar corretamente a classe de cobertura do solo mesmo quando esta estava sob a influência de nuvens e sombras de tamanho reduzido (Figura 21).

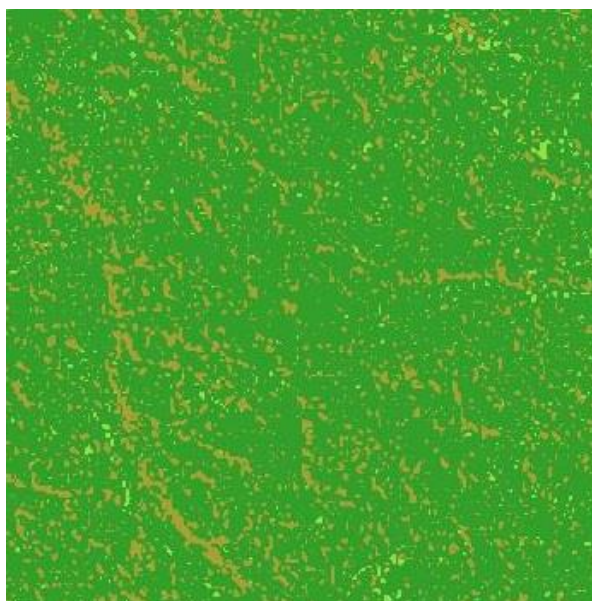


Figura 24: Efeito "sal pimenta" na imagem classificada pelo MD na abordagem1

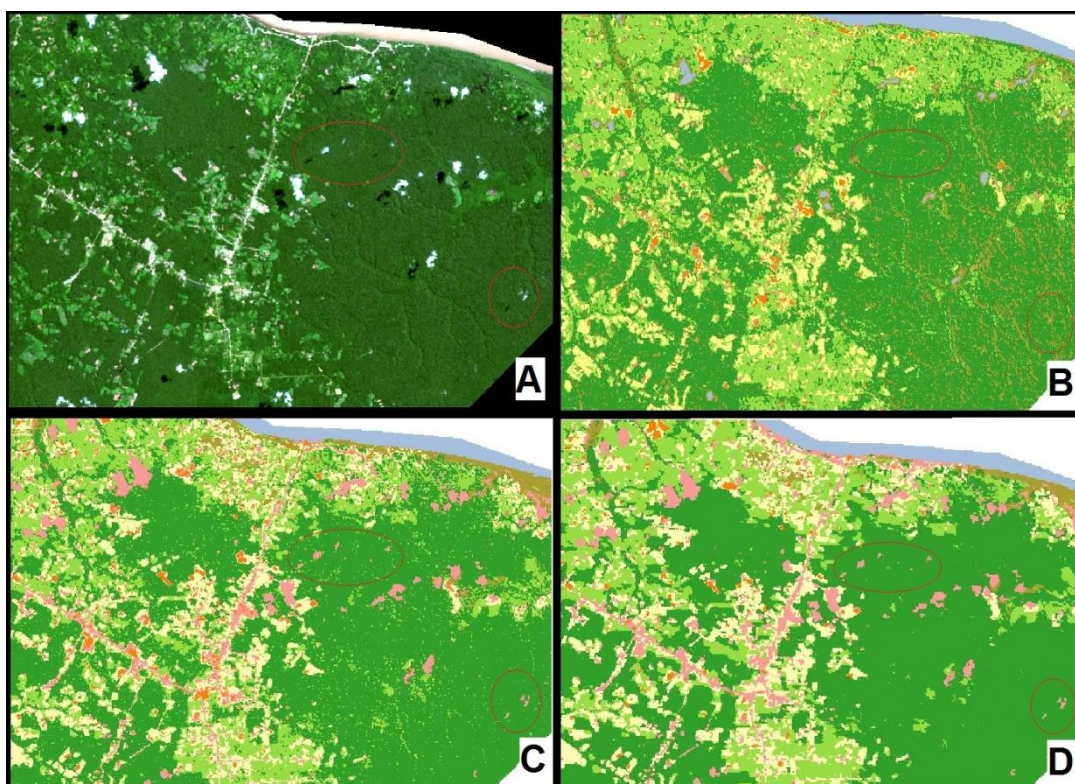


Figura 25: Comparativo do efeito de nuvens e sombras nas classificações (A - Landsat-8 R4G3B2, B - Classificação MD, C - Classificação MaxVer, D - Classificação ECHO)

5.2.2 Abordagem 2

Assim como na primeira abordagem o MaxVer foi o classificador que apresentou o melhor desempenho, reforçando seu bom potencial para a identificação de classes de cobertura do solo na região central da Amazônia brasileira. Com uma AG de 92,1% e índice Kappa de 90,4% novamente o MaxVer se mostrou eficiente em discriminar classes espectralmente similares. Todas as classes apresentaram valores de AP e AU igual ou superiores a 80%, reforçando assim a afirmação do bom desempenho do MaxVer na identificação e diferenciação das classes de cobertura desse trabalho.

O uso das imagens fração como base para a classificação apresentou pouca influência no resultado apresentado pelo MaxVer, tanto os valores de AG e Kappa como de AP e AU sofreram poucas alterações com a mudança de abordagem. De maneira geral a abordagem 1 obteve resultados um pouco melhores que a abordagem 2 para esse classificador.

Para o MaxVer a abordagem 2 melhorou os valor de AP para a classe AAU, que apresentou 85,7% na abordagem 1 e subiu para 93,3% na abordagem 2 e para a classe Campinarana que tinha 91,7% na abordagem 1 e subiu para 100% na abordagem 2. As classes que apresentaram melhoras nos valores de AU foram Solo Exposto (90,6% na abordagem 1 e 95,3% na abordagem 2) e Floresta Primária (96% na abordagem 1 e 96,8% na abordagem 2). Apesar da mudança de abordagem o efeito "sal pimenta" manteve-se presente na imagem classificada pelo MaxVer.

O ECHO teve o segundo melhor desempenho desta abordagem e diferentemente do MaxVer apresentou uma melhora em relação à abordagem 1. Sua AG teve acréscimo de 0,7% indo para 90,2%, o índice Kappa acompanhou a melhor da AG e teve um incremento de 0,8% atingindo 88,1% nessa abordagem. A classe AAU que foi a qual o ECHO encontrou mais dificuldade para identificar na primeira abordagem apresentou uma melhoria no seu valor de AP, que nessa abordagem foi de 92,9%. Em contrapartida, a AU para a classe teve uma redução, saindo de 85,7% na abordagem 1 para 74,2% nessa abordagem.

Essas diferenças encontradas nos valores de AP e AU para a classe AAU no ECHO apontam para uma melhora de desempenho da segunda abordagem na identificação e na diferenciação da classe AAU mesmo com a redução do valor de AU.

Assim como visto na abordagem 1, o ECHO teve resultados bem próximos aos apresentados pelo MaxVer na abordagem 2 e também obteve êxito em reduzir o efeito "sal pimenta" em áreas de cobertura homogênea.

O MD foi o classificador que apresentou o pior desempenho nessa abordagem, apresentando uma AG de 67,4% e um índice Kappa de 60,1%. Diferente dos outros classificadores testados o MD sofreu uma forte redução de desempenho nessa abordagem quando comparado com a primeira. As classes AAA, Floresta Primária e Floresta Secundária apresentaram uma intensa redução de seus valores de AP e AU. As únicas classes que apresentaram melhoras em ambos os valores de AP e AU foram as classes Solo Exposto e Campinarana (Tabela 10).

O efeito observado na abordagem 1 de minimização dos efeitos provocados pela cobertura de pequenas nuvens na área de estudo também se mostrou presente nessa abordagem. Em relação ao efeito "sal pimenta" notou-se um aumento nessa abordagem (Figura 22).

Ao observarmos o desempenho dos classificadores nas diferentes abordagens nota-se um indicativo de que a redução da dimensionalidade dos dados produzida pela aplicação do MLME provocou pouca alteração nos resultados obtidos pelos classificadores MaxVer e ECHO.

Para o MaxVer foi observada uma pequena redução do seu desempenho de maneira geral. Ao analisar o desempenho de cada classe individualmente houve pouca variação nos valores de AU e AP de forma que optar por uma abordagem frente à outra pouco irá contribuir para uma melhora significativa no desempenho na identificação de determinada classe.

Diferentemente do MaxVer, o ECHO apresentou uma pequena melhora no seu desempenho Global e é possível atribuir tal melhora ao grande incremento na AP da classe AAU na abordagem que utiliza as imagens fração como base para a classificação. Outro destaque para o ECHO em ambas as

abordagens foi a diminuição do efeito "sal pimenta" encontrado nos outros classificadores testados em ambas as abordagens.

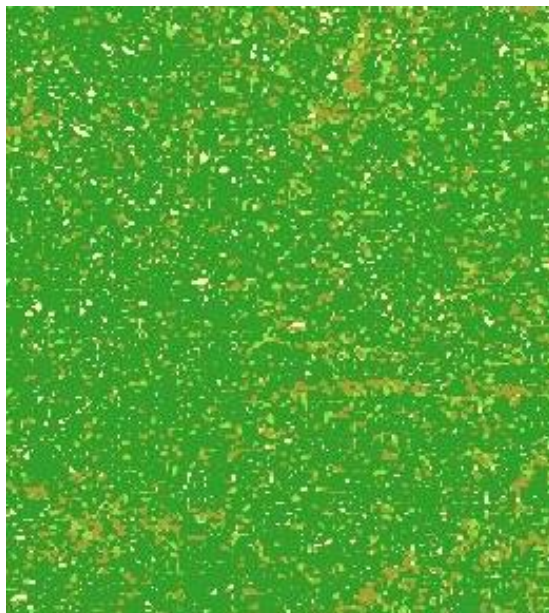


Figura 26: Intensificação do efeito "sal pimenta" na imagem classificada pelo MD na abordagem 2

O MD foi o classificador mais sensível à mudança de abordagem proposta. Utilizar as imagens geradas pelo MLME teve grande impacto nos resultados obtidos. O classificador não obteve êxito em utilizar a dimensionalidade reduzida dos dados e como consequência apresentou uma forte redução nos seus parâmetros de avaliação globais e individuais.

5.3 Cobertura do Solo

Nesse item as classes de cobertura do solo serão tratadas em dois grupos: naturais e antrópicas

5.3.1 Classes Naturais

A classe de cobertura do solo natural com maior extensão na área de estudo foi a Floresta Primária. Localizada em sua maior parte nas porções sul, nordeste e sudoeste da área de estudo, a Floresta Primária tem grandes áreas contínuas e que ainda sofrem pouca influência das atividades antrópicas

desenvolvidas na região. Além dessas extensas áreas contínuas dois núcleos de Floresta Primária cercados por áreas onde são ou foram desenvolvidas atividades antrópicas, principalmente ligadas ao setor agrícola, foram identificados (Figura 23).

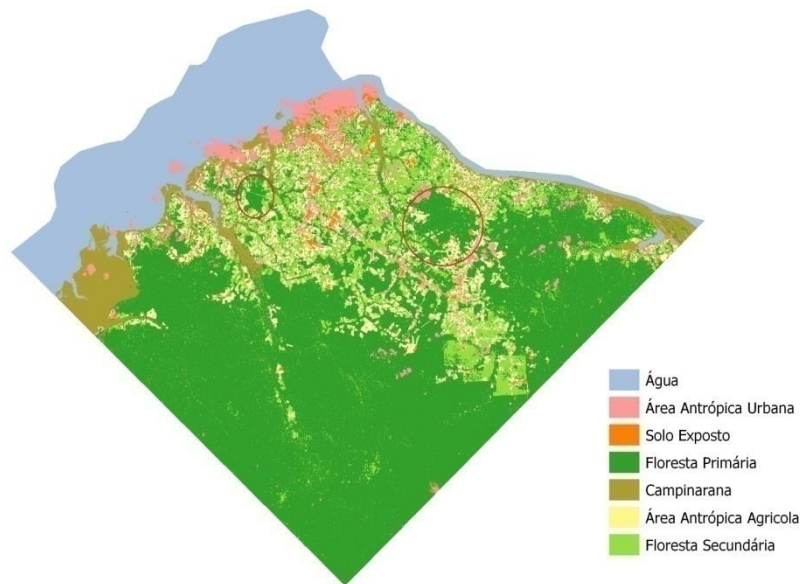


Figura 27: Núcleos de Floresta Primária em destaque na imagem classificada MaxVer A2

Na área de estudo a classe Floresta Secundária ocorre principalmente em áreas vizinhas às classes Solo Exposto e AAA. Esse padrão é justificável visto que a classe tem como origem o abandono dessas duas classes de uso antrópico. Nas atividades de campo foi possível notar uma grande diversidade de estágios sucessionais de Floresta Secundária (Figura 24), indicando que possivelmente existam diferentes formas de manejo agrícola e ciclos de cultivo.

A classe Campinarana apresentou uma distribuição esperada visto que devido às suas características biológicas essa classe está recorrentemente disposta próxima a corpos d'água e em solos arenosos. Apesar do comportamento esperado, notou-se em campo que possivelmente existam diferentes formações de Campinarana na área de estudo (Figura 25), indicando que a região apresenta áreas potencialmente importantes para a conservação dessa formação vegetal.



Figura 28: Diferentes estágios sucessionais de Floresta Secundária



Figura 29: Diferentes tipos de Campinarana encontrados na área de estudo

5.3.2 Classes Antrópicas

Dentre as classes oriundas do desenvolvimento de atividades antrópicas a classe AAA foi a que apresentou a maior extensão territorial. As áreas identificadas como pertencentes à classe se encontram distribuídas principalmente no entorno das estradas rurais e de corpos d'água com um tamanho médio de 3 hectares.

A localização e o tamanho das áreas classificadas como AAA sofre grande influência de fatores, como a facilidade de acesso à área, facilidade de escoamento da produção, fertilidade do solo e custo de implementação do cultivo. Conforme observado em campo, a maior parte dessas áreas é oriunda

da agricultura familiar e se destina à produção de alimentos para a subsistência e de mandioca (Figura 26) para a produção de farinha (Figura 27).



Figura 30: Cultivo familiar de mandioca



Figura 31: Produção da farinha de mandioca

A classe Solo Exposto teve uma distribuição irregular na área de estudo. Invariavelmente essa classe está ligada ao processo de transição de uma determinada área que será destinada a uma nova cultura agrícola ou a um novo empreendimento imobiliário. Essa classe pode ser considerada um estágio transitório entre uma cobertura natural e o uso antrópico ou entre diferentes tipos de uso antrópico (Figura 28).



Figura 32: Área destinada à construção de tanques para a criação de peixes

A classe AAU apresentou a maior concentração na região norte da área de estudo, onde se encontra a sede municipal (Figura 29). Além dessa concentração, pequenas áreas ao longo das estradas rurais do município também foram classificadas como AAU. Essas regiões, conforme observado em campo, são pequenas comunidades rurais que têm como principal fonte de renda a agricultura.



Figura 33: Centro da Cidade de Tefé-AM

6. CONCLUSÕES

O mapeamento e a classificação da cobertura do solo na Amazônia brasileira, principalmente em sua região central, são fundamentais para a avaliação e estruturação das políticas de gestão territorial e conservação da região e do continente.

Nesse contexto as imagens de satélite e as técnicas de classificação têm se destacado como ferramentas importantes no fornecimento de subsídios técnicos para a implementação de sistemas de monitoramento e avaliação da cobertura do solo na região.

A constante presença de nuvens na região se mostrou um forte impeditivo no uso de imagens ópticas para a classificação da cobertura da terra. Visto que todos os classificadores em ambas as abordagens testadas, atribuíram aos pixels referentes às nuvens e suas sombras a alguma classe dificultando assim o processo de análise dos resultados obtidos.

A pesar da ambas as abordagens testadas terem apresentados problemas na classificação quando há a presença de nuvens na cena, sugere-se que o método proposto na primeira abordagem seja utilizado em trabalhos futuros. Tal sugestão baseia-se no melhor desempenho apresentado pelos classificadores testados em relação a outra abordagem utilizada nesse trabalho.

Dentre os classificadores testados o MaxVer e o ECHO apresentaram um bom desempenho e obtiveram sucesso em identificar, diferenciar e classificar sete diferentes classes de cobertura da terra na área de estudo. O MD obteve resultado satisfatório apenas na abordagem 1, mas ainda assim bastante inferior ao apresentado pelos outros classificadores.

A aplicação do MLME reduziu a acurácia das classificações geradas pelo MD e pelo MaxVer, indicando que a redução da dimensionalidade dos dados não teve efeito positivo no resultado final apresentado por esses classificadores. Em contrapartida, o uso das imagens fração melhorou a acurácia apresentada pelo ECHO, indicando que esse método pode efetivamente melhorar o resultado produzido pelo classificador.

É válido ressaltar que a confiabilidade dos resultados obtidos por meio das matrizes de confusão está diretamente relacionado ao método utilizado para realizar a avaliação da acurácia da classificação. Recomenda-se que os designs de amostragem e resposta sejam planejados com algum conhecimento prévio da área há ser classificada e que sejam selecionados um número adequado de amostras. É recomendável que a construção dos protocolos de avaliação e rotulagem, sejam as amostras coletadas em campo ou no escritório, apresentem regras que excluam toda possibilidade de confusão entre as classes e que permita que qualquer analista as aplique.

Quanto a coleta de amostras, é de fundamental importância a realização de expedições a campo para um melhor reconhecimento da área de estudo a fim de melhor caracterizar as classes de cobertura auxiliando assim o processo de construção das regras a serem utilizadas nos protocolos de avaliação e rotulagem. A realização de expedições a campo para a coleta de amostras também aumenta a confiabilidade dos resultados obtidos na avaliação de acurácia quando comparado com o uso de amostras coletadas apenas do escritório.

Com os resultados obtidos nesse trabalho, notou-se que o uso do solo no município de Tefé corresponde com o que está previsto em seu plano diretor. Nesse contexto a macro zona rural Novo Horizonte se destaca. Nessa zona são encontrados as estradas rurais do município, que tem no seu entorno a principal concentração da classe AAA cujo apresentou a maior extensão territorial entre as classes de cobertura antrópicas.

A zona rural Novo Horizonte ainda abriga em sua extensão uma vasta área coberta por Florestas Primárias e Secundárias. Nota-se que é necessário que haja um acompanhamento por parte do estado quanto ao modelo de expansão das atividades antrópicas dentro dessa zona. Durante as atividades de campo percebeu-se que o processo abertura de novas áreas para desenvolvimento atividades agrícolas e agropecuárias é intenso e nem sempre está de acordo com as legislações previstas.

Dessa forma é possível afirmar que o zoneamento proposto no plano diretor está de acordo com a realidade encontrada em campo e com os

resultados obtidos pelos classificadores. Entretanto se mostrou necessário que haja um acompanhamento das atividades de expansão do município para que o planejamento proposto mantenha-se adequado.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- AGUIAR, A.P.D.; SHIMABUKURO, Y.E.; MASCARENHAS, N.D.A. Use of synthetic bands from mixing models in the multispectral classification of remote sensing images. *International Journal of Remote Sensing*, v. 20, n. 4, p. 647-657. (1999).
- ALBAGLI, Sarita. Amazônia: fronteira geopolítica da biodiversidade. *Revista Parcerias Estratégicas*, No 12. Ministério da Ciência e Tecnologia, ISSN 1413-9375.(2001).
- ACHARD, F, et al. Determination of Deforestation of the World`s Humid Tropical Forests. *Science*. 297, 999-1002. (2002).
- ADAMS, J,B, et al. Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*. 52, 137-154. (1995).
- Áreas Protegidas na Amazônia brasileira avanços e desafios / [organizadores Adalberto Veríssimo... [et al.]].Belém : Imazon ; São Paulo : Instituto Socioambiental.(2011).
- ATLAS DO DESENVOLVIMENTO HUMANO NO BRASIL. PNUD, IPEA, Fundação João Pinheiro. Rio de Janeiro. (2013).
- ARONOFF, S. Classification accuracy: a user approach. *Photogrammetric Engineering & Remote Sensing*. Bethesda, v 48, n8, p1299-1307. (1982a).
- AVERY, T. E.; BERLIN, G. L. *Fundamentals of Remote Sensing and Air photo Interpretation*. 5 ed. New Jersey: Prentice Hall. (1992).
- AYRES, J.M, et al. Os corredores ecológicos das florestas tropicais do Brasil. Belém, Pará. Sociedade Civil Mamirauá, 256p. (2005).
- BAILEY, M.; CONVIS, C.; DAVIS, F.; LI, X.; GOODIN, K.; GROSS<AM, D. Accuracy assessment procedures. Redlands: USDI-NBS/NPS/ESRI/University of California-NCGIA/The Nature Conservancy.108p. (1994).
- BALL, G.H; HALL, D.J. *Isodata: a method of data analysis and pattern classification*. Office of Naval Research. Stanford Research Institute, Menlo Park, U.S. (1965).
- BARRETO, P, et al. Human Pressure on the Brazilian Amazon Forest. World Resources Institute, Belém, Pará, Brasil. (2006).

BARTHOLOMÉ, E.; BELWARD, A.S. GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*. 26, 1959-1977. (2005).

BECKER, B. K. "Revisão das políticas de ocupação da Amazônia: é possível identificar modelos para projetar cenários?" *Parcerias Estratégicas* 12: 135-159. (2001).

BRASIL. Decreto n.º 5.758 de 2006

BRASIL. Lei 9.985 de 2000

BRASWELL, et al. A multivariable approach for mapping sub-pixel land cover distributions using MISR and MODIS: Application in the Brazilian Amazon region. *Remote Sensing of Environment*. 87, 243-256. (2003).

CÂMARA, G, et al. *Anatomia de Sistemas de Informação Geográfica*. Escola de Computação, SBC, 1996. (1996).

CÂMARA, G; SOUZA, R.C.M; FREITAS, U.M; GARRIDO, J. "Spring: Integrating remote sensing and GIS by object-oriented data modelling". *Computers & Graphics*, 20: (3) 395-403. (2013).

CASSOL, H.L.G. Estimativa de biomassa e estoque de carbono em um fragmento de floresta ombrófila mista com uso de dados ópticos de sensores remotos. 143f. Dissertação de Mestrado – Universidade Federal do Rio Grande do Sul. Programa de Pós-graduação em Sensoriamento Remoto. Instituto de Geociências. Porto Alegre, RS – BR. (2013).

CHAVEZ, P. Image-Based Atmospheric Corrections – Revised and Improved. *Photogrammetric Engineering & Remote Sensing*. 62, 1025-1036. (1996).

CRÓSTA, A.P. *Processamento digital de imagens de sensoriamento remoto*. Campinas: IG-UNICAMP. 170p. (1992).

CURRAN, P.J. *Principles of remote sensing*. Londres, Longman, 256p. (1985).

DUTRA, L.V. et al. Land Cover Classification in Amazon Using Alos Palsar Full Polarimetric Data. *Anais XIV Simposio Brasileiro de Sensoriamento Remoto*. Natal - RN. p 7259-7264. (2009).

GOODCHILD, M.F.; BIGING, G.S.; CONGALTON, R.G.; LANGLEY, P.G.; CHRISTMAN, N.R.; DAVIS, F.W. Final report of the accuracy assessment task force. Santa Barbara: University of California-NCGIA. (1994).

HARALICK, R.M.; FU, K. Pattern recognition and classification. In: COLWELL, R.N. Manual of Remote Sensing. Falls Church: American Society of Photogrammetry. p. 793-805. (1983).

HORD, R.M.; BROONER, W. Land-use map accuracy criteria. Photogrammetric Engineering & Remote Sensing. Bethesda, v42, n5, p.671-677. (1976).

Instituto Brasileiro de Geografia e Estatística (IBGE). Indicadores de desenvolvimento sustentável: Brasil 2004. Rio de Janeiro: IBGE.(2004).

Instituto do Meio Ambiente e dos Recursos Naturais Renováveis. Corredores Ecológicos – experiências em planejamento e implementação / Ministério do Meio Ambiente, Secretária de Biodiversidade e Florestas. – Brasília: MMA. (2007).

JENSEN, J.R. Introductory digital image processing - a remote sensing perspective. Englewood Cliffs: Prentice-Hall, 316p. (1996).

JENSEN, J.R. Sensoriamento Remoto do Ambiente: Uma Perspectiva em Recursos Terrestres. Tradução português 2 Ed. Parentese Editora. São José dos Campos SP, 598p. (2009).

JUSTICE, C.O.; TOWNSHEND, J.R.G. A comparison of unsupervised classification procedures using LANDSAT MSS data for an area of complex surface conditions in Basilicata, southern Italy. Remote Sensing of Environment. New York, v 12, n4, p.407-420. (1982).

KARDOULAS, N.G, et al. Geometric Correction of SPOT and Landsat Imagery: A Comparison of Map- and GPS-Derived Control Points. Photogrammetric Engineering & Remote Sensing. 62, 1173-1177. (1996).

KAST, J. L.; SWAIN, P. H.; DAVIS, B. J.; and SPENCER, P. W., "ECHO User's Guide" LARS Technical Reports. Paper 102.(1977).

KETTIG, R.L; LANDGRABE, D.A. Computer classification of remotely sensed multispectral image data by extraction and classification of homogeneous objects. IEEE Transactions on Geoscience Electronic, GE-14:19-26. (1976).

KETTING, R. L. Computer Classification of Remotely Sensed Multispectral Image Data by Extraction and Classification of Homogeneous Objects. Ph.D. Thesis, School of Electrical Engineering, Purdue University, West Lafayette, Indiana. (1975).

KURY, E.B. & Vieira, C. A. O. Classificação automática de imagens de alta resolução. In UFVGEO(2001). Acessado em: 19 nov. 2016.

LIANG, S, et al. Atmospheric Correction of Landsat ETM+ Land Surface Imagery-Part I: Methods. IEEE Transactions on Geoscience and Remote Sensing, vol 39, n11. (2001).

LOUREIRO, V.R. Amazônia: uma história de perdas e danos, um futuro a (re)construir. ESTUDOS AVANÇADOS, vol 16, 107 - 121. (2002).

LU, D. Integration of vegetation inventory data and Landsat TM image for vegetation classification in western Brazilian Amazon. Forest Ecology and Management. 213, 369-383. (2005).

MAGALHÃES, J.L.L, et al. Development of a Flooded Forest Anthropization Index (FFAI) applied to Amazonian areas under pressure from different human activities. Ecological Indicators. 48, 440-447. (2014).

MATHER, P.M. Computer processing of remote sensing imagens. Chicester: John Wiley & Sons. 352p. (1987).

MENEZES, P.R, et al. Introdução ao Processamento de Imagens de Sensoriamento Remoto. Universidade de Brasília, Brasília – DF, 266p. (2012).

MENESES, P. R. Fundamentos de Radiometria Óptica Espectral. In: MENESES, P. R.; NETTO, J. S. M. Sensoriamento Remoto: Reflectância dos alvos naturais. Brasília, DF: UnB; Planaltina: Embrapa Cerrados. (2001).

MERTES, L, et al. Spatial patterns of hydrology, geomorphology, and vegetation on the floodplain of the Amazon River in Brazil from a remote sensing perspective. Geomorphology. 13, 215-232. (1995).

MERTES, L. Rates of flood-plain sedimentation on the central Amazon River. Geology. 22, 171-174. (1994).

MILLIKAN, B. Políticas Públicas e Desenvolvimento Sustentável em Rondônia: situação atual e abordagem para um planejamento participativo de estratégias para o Estado. Governo do Estado de Rondônia. Secretaria de Planejamento e Coordenação Geral, Projeto de Cooperação Técnica ao Planaflo, Projeto BRA/94/007, Porto Velho. p. 112. (1998).

MORÁN, Emílio F. Ecologia Humana das populações da Amazônia. Rio de Janeiro: Editora Vozes. (1990).

MOURA, Edila A. F.; NASCIMENTO, Ana Claudeise do; CORRÊA, Dávila; ALENCAR, Edna; SOUSA, Isabel S. de. Sociodemografia da Reserva de Desenvolvimento Sustentável Mamirauá: 2001-2011. Tefé, AM: IDSM; Belém: NAEA. (2016).

MUSTARD, J. F.; SUNSHINE, J. M. Spectral Analysis for Earth Science: Investigation using remote sensing data. Remote Sensing for Earth Science: Manual of Remote Sensing, vol 3. New York: Wiley. (1999).

NOVO, E.M.L.M. Sensoriamento Remoto: Princípios e Aplicações. Ed. Edgard. (2008).

NOBRE CA, Obregón G, Marengo J, Fu R, Poveda G. Characteristics of Amazonian climate: main features. AGU Geophysical Monograph Series 186:149–162 Blucher Ltda, 3a Edição, São Paulo, 363 p. (2009).

NOBRE, A.D. O. Futuro Climático da Amazônia, Relatório de Avaliação Científica. Patrocinado por ARA, CCST-INPE, e INPA. São José dos Campos, Brasil, 42p. (2014).

MALHI Y, Sassan Saatchi, Cecile Girardin, Luiz E. O. C. Aragão. The Production, Storage, and Flow of Carbon in Amazonian Forests. AGU Geophysical Monograph Series 186:355-371. (2009)

PONS, X; SOLÉ-SUGRAÑES, L. A simple radiometric correction model to improve automatic mapping of vegetation from multispectral satellite data. Remote Sensing of Environment. 48, 191-204. (1994).

RENÓ, V.F; NOVO, E.M.L.M; SUEMITSU, C; RENNÓ, C. D; SILVA, T.S.F. Assessment of deforestation in the Lower Amazon floodplain using historical Landsat MSS/TM imagery. Remote Sensing of environment. 115, 3446-3456. (2011).

RICHARDS, J.A. Remote sensing digital image analysis. Berlim: Springer-Verlag, 281p. (1986).

RICHARDS, P; ARIMA, E; VanWEY, L; COHN, A; BHATTARAR, N. Are Brazil's Deforesters Avoiding Detection?. Conservation Letters. (2016)

ROCHA, J.A. The influence of ground survey size on accuracy of area estimates from satellite imagens. Silsoe: Cranfield Institute of Technology-Silsoe College. 244p. Tese de Doutorado em Filosofia.

RODRIGUES, E. A. Rede urbana do Amazonas: Tefé como cidade média de responsabilidade territorial na calha do Médio Solimões. Dissertação de mestrado. Universidade Federal do Amazonas. (2011).

ROSENDO, J.S; ROSA, R. Estoque de carbono em solos sob vegetação nativa e uso antrópico na bacia do Rio Araguari-MG. Anais XV SBSR, Curitiba, PR. p 9173-9180. (2011).

SAATCHI, S.S, et al. Mapping land cover types in the Amazon Basin using 1km JERS-1 mosaic. International Journal of Remote Sensing. 21, 1201-1234. (2000).

SCHOWENGERDT, R.A. Techniques for image processing and classification in remote sensing. Orlando: Academic Press. 249p. (1983).

SCHOWENGERDT, R. A. Remote sensing: models and methods for image processing. 2. ed. San Diego: Academic Press, 522 p. (1997).

SCHWEIK, C. M.; GREEN, G. M. The use of spectral mixture analysis to study human incentives, actions, and environmental outcomes. Social Science Computer Review, Thousand Oaks, v. 17, n. 1, p. 40-63. (1999).

SHIMABUKURO, Y.E.; SMITH, J.A. Fraction images derived from Landsat TM and MSS data for monitoring reforested areas. Canadian Journal of Remote Sensing, v. 21, p 67-74. (1995).

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).

SILVA .B.B, et al. Landcover Discrimination at Brazilian Amazon Using Region Based Classifier and Stochastic Distance. International Geoscience and Remote Sensing Symposium. p 2900-2903. (2011).

SOARES, S. C, et al. Mapeamento e análise multitemporal do uso e cobertura do solo em região do município de Santarém-PA através de imagens Landsat 5-TM. Anais XV SBSR, Curitiba, PR. p 6910-6916. (2011).

SOMBROEK, W. G. Amazon soils. A reconnaissance of the soils of the Brazilian Amazon region. Wageningen, NL, Centre for Agricultural Publishing and Documentation(PUDOC). (1966).

SONG, C, et al. Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects?. *Remote of Environment*, v.75 p.230-244. (2001).

SWAIN, P.H. Fundamentals of pattern recognition in remote sensing. IN SWAIN, P.H., DAVIS, S.M. (Ed.) *Remote sensing: the quantitative approach*. New York: McGraw-Hill, 1978. 136-187p. (1978).

TAYLOR, J.C. Evaluation of SPOT satellite images for measuring landscape features. In: *Monitoring Landsacpe Changes In The national Parks*. Silsoe; Cranfiled Institute of Tecnology-Silsoe College. (1989).

VERMOTE, E.F, et al. Atmospheric correction of MODIS data in the visible to middle infrared: first results. *Remote Sensing of Environment*. 83, 97-111. (2002).

VOLGELMANN, J,E. et al. Effects of Landsat 5 Thematic Mapper and Landsat 7 Enhanced Thematic Mapper Plus radiometric and geometric calibrations and corrections on landscape characterization. *Remote Sensing of Environment*. 78, 55-70. (2001).

ANEXO 2



Ministério do Meio Ambiente - MMA
 Instituto Chico Mendes de Conservação da Biodiversidade - ICMBio
 Sistema de Autorização e Informação em Biodiversidade - SISBIO

Autorização para atividades com finalidade científica

Número: 54027-1	Data da Emissão: 04/05/2016 09:43	Data para Revalidação*: 03/06/2017
<small>* De acordo com o art. 28 da IN 03/2014, esta autorização tem prazo de validade equivalente ao previsto no cronograma de atividades do projeto, mas deverá ser revalidada anualmente mediante a apresentação do relatório de atividades a ser enviado por meio do Sisbio no prazo de até 30 dias a contar da data do aniversário de sua emissão.</small>		

Dados do titular

Nome: Luiz Fernando Guimarães Schwartzman	CPF: 070.235.679-41
Título do Projeto: Análise da Aptidão de classificadores para o mapeamento de cobertura e uso do solo em regiões de floresta tropical; estudo de caso do município de Tefé-AM	
Nome da Instituição: UFRGS - UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL	CNPJ: 92.969.856/0001-98

Cronograma de atividades

#	Descrição da atividade	Início (mês/ano)	Fim (mês/ano)
1	Verificação em campo das classes identificadas pelos algoritmos de classificação de imagens.	06/2016	08/2016

Observações e ressalvas

1	As atividades de campo exercidas por pessoa natural ou jurídica estrangeira, em todo o território nacional, que impliquem o deslocamento de recursos humanos e materiais, tendo por objeto coletar dados, materiais, espécimes biológicos e minerais, peças integrantes da cultura nativa e cultura popular, presente e passada, obtidos por meio de recursos e técnicas que se destinem ao estudo, à difusão ou à pesquisa, estão sujeitas a autorização do Ministério de Ciência e Tecnologia.
2	Esta autorização NÃO exime o pesquisador titular e os membros de sua equipe da necessidade de obter as anuências previstas em outros instrumentos legais, bem como do consentimento do responsável pela área, pública ou privada, onde será realizada a atividade, inclusive do órgão gestor de terra indígena (FUNAI), da unidade de conservação estadual, distrital ou municipal, ou do proprietário, arrendatário, posseiro ou morador de área dentro dos limites de unidade de conservação federal cujo processo de regularização fundiária encontra-se em curso.
3	Este documento somente poderá ser utilizado para os fins previstos na Instrução Normativa ICMBio nº 03/2014 ou na Instrução Normativa ICMBio nº 10/2010, no que especifica esta Autorização, não podendo ser utilizado para fins comerciais, industriais ou esportivos. O material biológico coletado deverá ser utilizado para atividades científicas ou didáticas no âmbito do ensino superior.
4	O titular de licença ou autorização e os membros da sua equipe deverão optar por métodos de coleta e instrumentos de captura direcionados, sempre que possível, ao grupo taxonômico de interesse, evitando a morte ou dano significativo a outros grupos; e empregar esforço de coleta ou captura que não comprometa a viabilidade de populações do grupo taxonômico de interesse em condição in situ.
5	O titular de autorização ou de licença permanente, assim como os membros de sua equipe, quando da violação da legislação vigente, ou quando da inadequação, omissão ou falsa descrição de informações relevantes que subsidiaram a expedição do ato, poderá, mediante decisão motivada, ter a autorização ou licença suspensa ou revogada pelo ICMBio, nos termos da legislação brasileira em vigor.
6	Este documento não dispensa o cumprimento da legislação que dispõe sobre acesso a componente do patrimônio genético existente no território nacional, na plataforma continental e na zona econômica exclusiva, ou ao conhecimento tradicional associado ao patrimônio genético, para fins de pesquisa científica, bioprospecção e desenvolvimento tecnológico. Veja maiores informações em www.mma.gov.br/cgen .
7	Em caso de pesquisa em UNIDADE DE CONSERVAÇÃO, o pesquisador titular desta autorização deverá contactar a administração da unidade a fim de CONFIRMAR AS DATAS das expedições, as condições para realização das coletas e de uso da infra-estrutura da unidade.

Outras ressalvas

1	Ressalvo que o pesquisador deve encaminhar com antecedência cronograma de campo para a equipe gestora da FLONA de Tefé, bem como cópia de resumos, reportagens, artigos e outras publicações que sejam originadas desta pesquisa para compor o acervo bibliográfico da UC.
---	--

Locais onde as atividades de campo serão executadas

#	Município	UF	Descrição do local	Tipo
1	TEFE	AM	Tefe	Fora de UC Federal
2		AM	FLORESTA NACIONAL DE TEFE	UC Federal

Este documento (Autorização para atividades com finalidade científica) foi expedido com base na Instrução Normativa nº 03/2014. Através do código de autenticação abaixo, qualquer cidadão poderá verificar a autenticidade ou regularidade deste documento, por meio da página do Sisbio/ICMBio na Internet (www.icmbio.gov.br/sisbio).

Código de autenticação: 75979753



Página 1/3

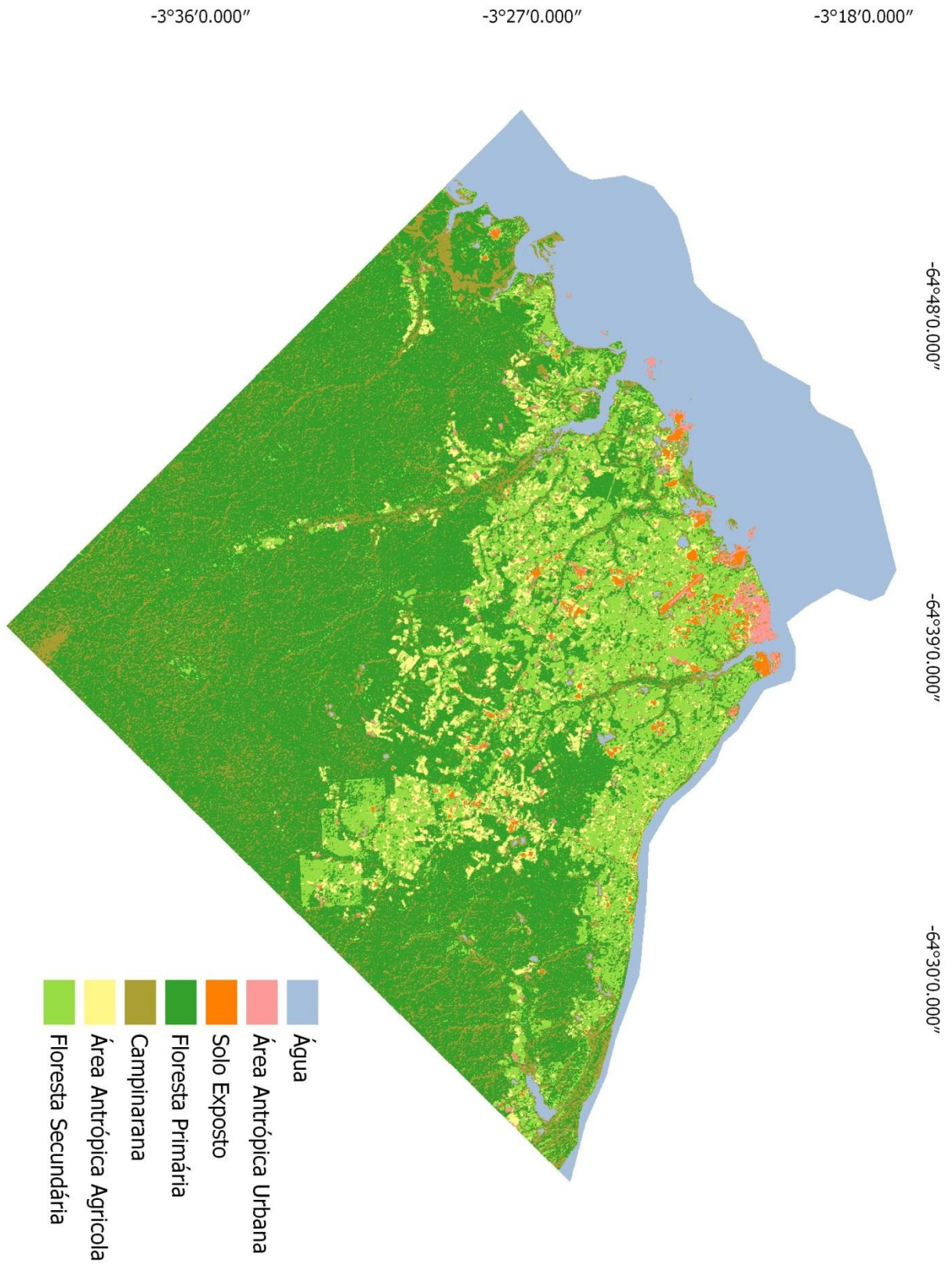
ANEXO 3

Ficha de Campo

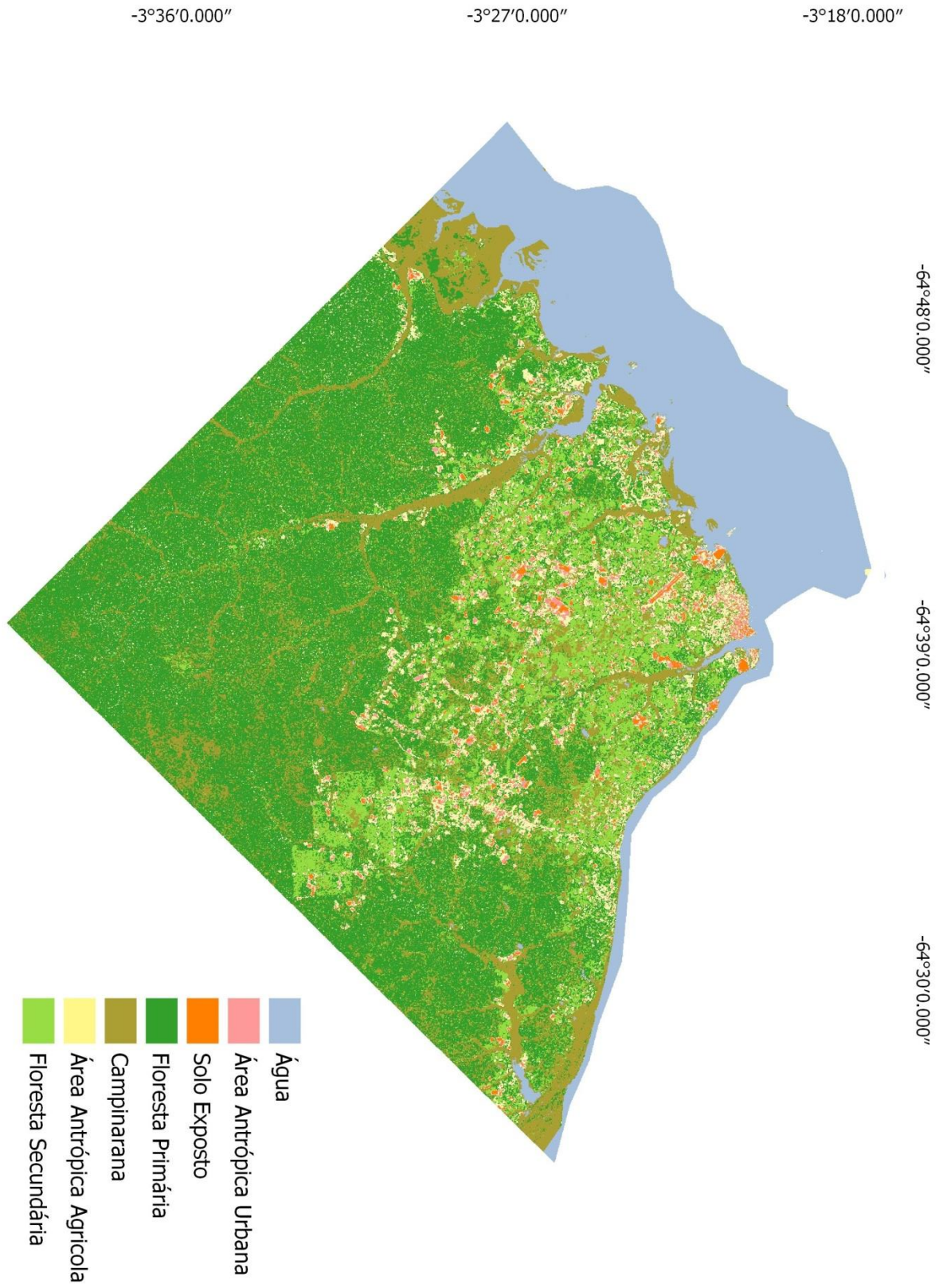
Nome do Observador:	
Data:	
Código da Amostra:	
Latitude:	
Longitude:	
Precisão do GPS:	
Forma da área observada:	
Tamanho da área observada:	
Classe de cobertura:	
Código da Foto:	

Observações:

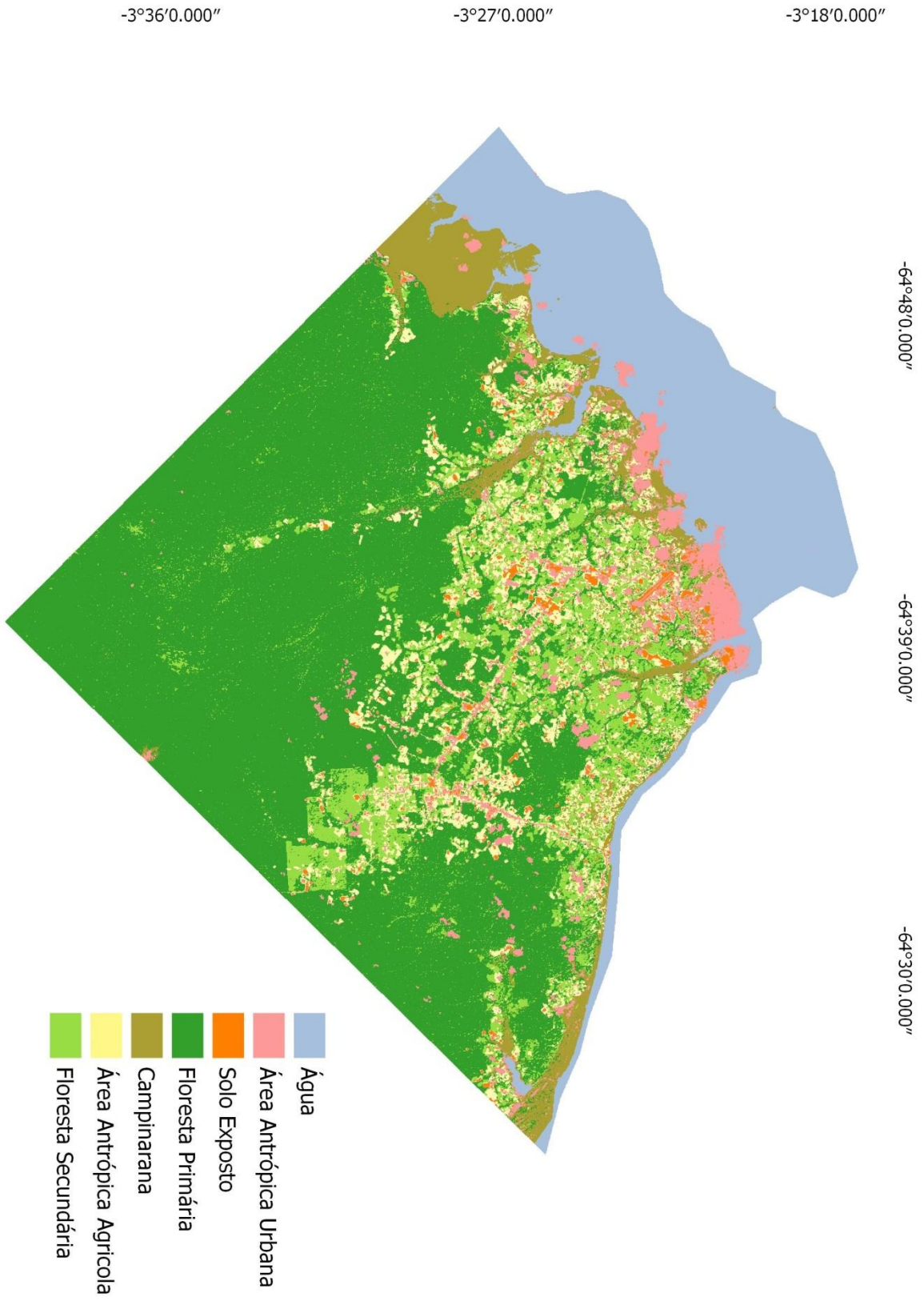
ANEXO 4



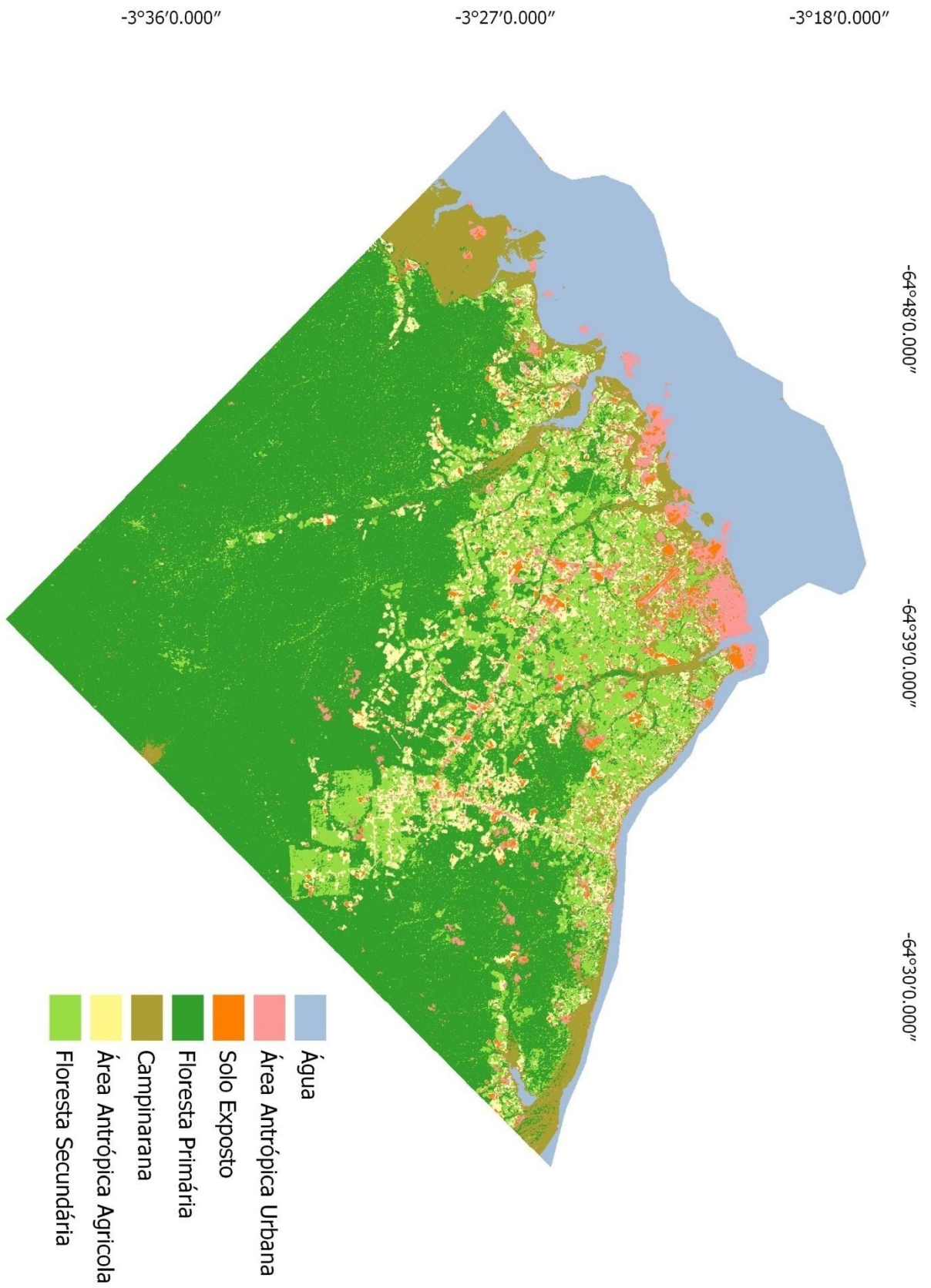
ANEXO 5



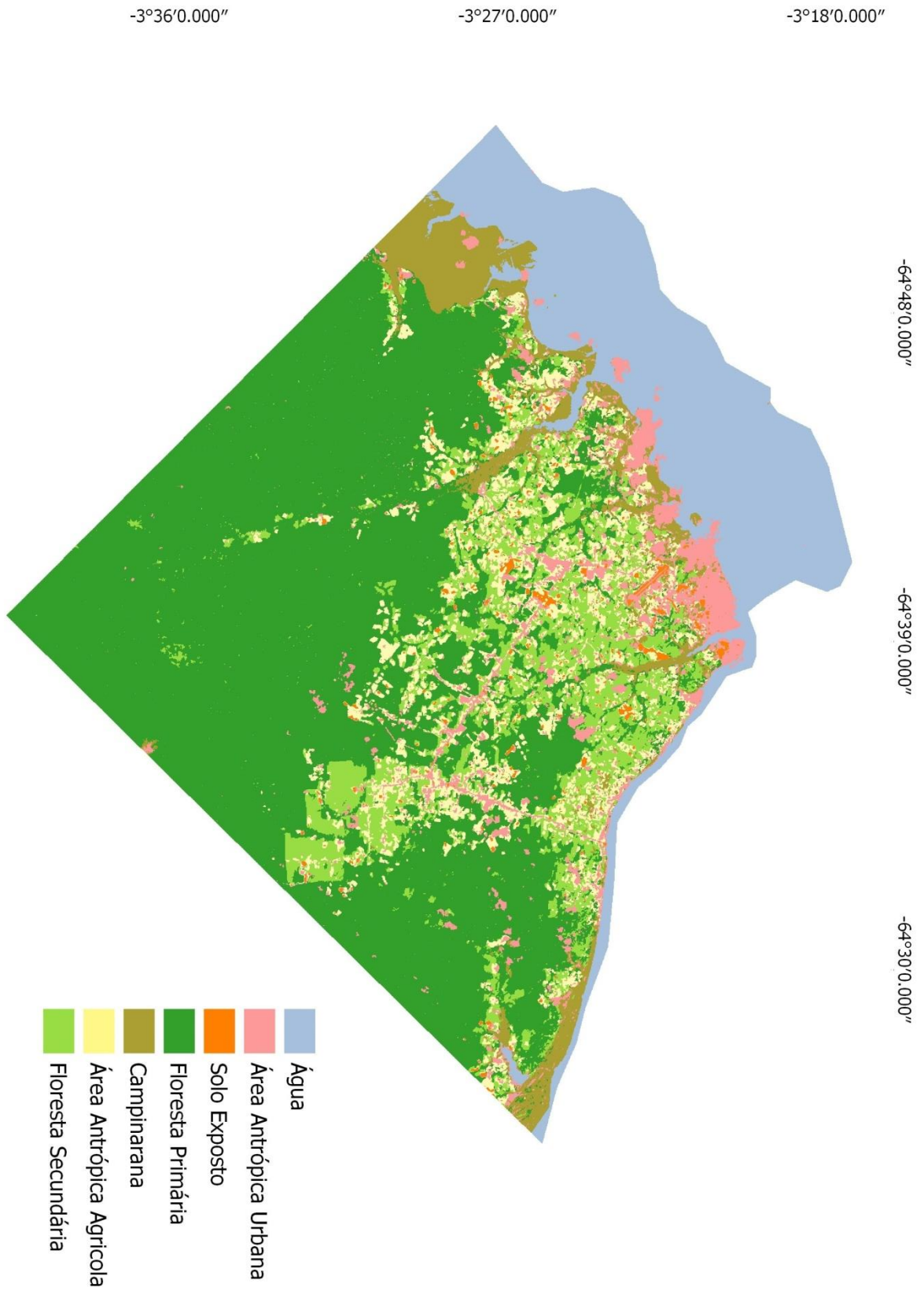
ANEXO 6



ANEXO 7



ANEXO 8



ANEXO 9

