

**COMPARAÇÃO ENTRE TÉCNICAS DE
CLASSIFICAÇÃO DIGITAL EM IMAGENS
LANDSAT NO MAPEAMENTO DO USO E
OCUPAÇÃO DO SOLO EM CAPIXABA, ACRE**

SYMONE MARIA DE MELO FIGUEIREDO

2005

59424

050616

SYMONE MARIA DE MELO FIGUEIREDO

**COMPARAÇÃO ENTRE TÉCNICAS DE CLASSIFICAÇÃO DIGITAL
EM IMAGENS LANDSAT NO MAPEAMENTO DO USO E OCUPAÇÃO
DO SOLO EM CAPIXABA, ACRE**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do curso de Mestrado em Engenharia Florestal, área de concentração em Manejo Ambiental, para a obtenção do título de “Mestre”.

Orientador

Prof. Luis Marcelo Tavares de Carvalho

LAVRAS
MINAS GERAIS - BRASIL
2005

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Figueiredo, Symone Maria de Melo.

Comparação entre técnicas de classificação digital em imagens landsat no mapeamento do uso e ocupação do solo em Capixaba, Acre / Symone Maria de Melo Figueiredo. -- Lavras : UFLA, 2005.

127p. : il.

Orientador: Luis Marcelo Tavares de Carvalho.

Dissertação (Mestrado) - UFLA.

Bibliografia.

1. Análise de mistura espectral. 2. NDVI. 3. Mineração de dados. 4. Exatidão do mapeamento. 5. Amazônia. I. Universidade Federal de Lavras. II. Título.

CDD-526.982
-621.3678

SYMONE MARIA DE MELO FIGUEIREDO

**COMPARAÇÃO ENTRE TÉCNICAS DE CLASSIFICAÇÃO DIGITAL
EM IMAGENS LANDSAT NO MAPEAMENTO DO USO E OCUPAÇÃO
DO SOLO EM CAPIXABA, ACRE**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do curso de Mestrado em Engenharia Florestal, área de concentração em Manejo Ambiental, para obtenção do título de “Mestre”.

Aprovada em 31 de janeiro de 2005.

Prof. Dr. José Aldo Alves Pereira - UFLA

Prof. Dr. Dalton de Morrison Valeriano - INPE



Prof. Dr. Luis Marcelo Tavares de Carvalho - UFLA
(Orientador)

**LAVRAS
MINAS GERAIS - BRASIL**

À minha mãe, Marlene.

Ao meu pai, Jediel.

DEDICO

A Deus, a meu esposo Evandro e ao nosso primeiro e querido bebê, cuja chegada aguardamos para os próximos dias, mas que nos últimos oito meses me acompanha em meu ventre

OFEREÇO

AGRADECIMENTOS

A Deus, por ter me guiado e apoiado em todas as etapas de minha vida.

Ao professor Dr. Luis Marcelo Tavares de Carvalho, pela orientação e confiança.

Aos professores Dr. José Aldo Alves Pereira, Dr. Dalton de Morrison Valeriano e Dr. Natalino Calegario, pela atenção e valiosas sugestões.

Ao Secretário de Estado de Meio Ambiente do Acre, Carlos Edgard de Deus, pela amizade, confiança e apoio, sem os quais não poderia realizar o curso de mestrado.

Aos professores do curso de mestrado em Engenharia Florestal da UFLA e a todos os funcionários do Departamento de Engenharia Florestal, que contribuíram para a realização deste trabalho.

Ao Instituto Nacional de Pesquisas Espaciais – INPE, pela colaboração, cedendo gentilmente as imagens de satélite.

Aos amigos da SEMA e IMAC, que colaboram e me apoiaram durante a realização dos trabalhos, em especial Cynara e Marcos, que prestaram importantes informações.

Ao meu grande amigo e incentivador Arnaldo Júnior, que me ajudou de forma incondicional em todas as horas e em tudo o que foi preciso.

Aos amigos da pós-graduação e de Lavras, Sandro e Silvinha, Flávio e Sônia, Breno, Zélia, Marcela, Cíntia, Simone, Edmilson, Daniel, Elvis, Lilian, Totonho, Samurai, Débora, Ivana, Ana Cecília, Rodrigo, Zigoto, Lucas, Nádia, Charles, Gentil, Claudião e Pedro, pelos bons momentos e agradável convivência.

Ao Luciano (Bodinho), Gleyce e Adauta, pela amizade e preciosa ajuda em diversas fases desse trabalho.

A todos os meus irmãos: Paulo, João, Marizete, Socorro, Júnior e Daniel, por tudo que significam.

À minha segunda família: Tereza, José Figueiredo, Lilian e Gustavo.

Ao Evandro, meu amado esposo, pelas sugestões, críticas, apoio e paciência, que foram imprescindíveis durante todo o tempo, e, especialmente, pelo amor a mim dedicado, sem o qual não teria conseguido alcançar meus objetivos.

A todos que ajudaram, acreditaram e torceram...

SUMÁRIO

RESUMO	i
ABSTRACT	ii
1 INTRODUÇÃO	1
2 REFERENCIAL TEÓRICO	4
2.1 Processo de ocupação no estado do Acre.....	4
2.2 Classificação digital de imagens de sensoriamento remoto.....	7
2.2.1 Princípios do reconhecimento de padrões.....	7
2.2.2 Transformações nas imagens digitais para extração de informações.....	11
2.2.2.1 Análise de mistura espectral	12
2.2.2.2 Índice de vegetação	16
2.2.3 Algoritmos de classificação	18
2.2.3.1 Máxima verossimilhança.....	18
2.2.3.2 Isodata	24
2.2.3.3 Árvore de decisão.....	27
2.3 Aplicação de técnicas de classificação no mapeamento do uso e ocupação do solo na Amazônia.....	30
2.5 Avaliação da acurácia da classificação	34
3 MATERIAL E MÉTODO.....	38
3.1 Área de estudo	38
3.1.1 Localização geográfica.....	38
3.1.2 Meio físico e biológico	39
3.1.3 Aspectos sócio-econômicos e situação fundiária	40
3.2 Material	41
3.3 Metodologia.....	43
3.3.1 Classes temáticas para o mapeamento do uso e ocupação do solo .	43
3.3.1.1 Caracterização das classes temáticas	43
3.3.1.2 Amostras das classes temáticas	49
3.3.2 Técnicas de pré-processamento das imagens	52
3.3.2.1 Correção geométrica	52
3.3.2.2 Calibração radiométrica.....	54
3.3.3 Técnicas de processamento das imagens.....	56
3.3.3.1 Extração de informações das imagens	56
3.3.3.1.1 Análise de mistura espectral.....	56
3.3.3.1.2 Índice de vegetação.....	58
3.3.3.2 Classificação digital	59
3.3.3.2.1 Máxima verossimilhança.....	59

3.3.3.2.2 Isodata.....	60
3.3.3.2.3 Classificação supervisionada e não supervisionada usando imagens fração oriundas do modelo linear de mistura espectral	61
3.3.3.2.4 Árvore de decisão.....	66
3.3.3.4 Homogeneização do resultado da classificação	68
3.3.4 Comparação entre os algoritmos de classificação.....	68
3.3.4.1 Análise de exatidão	68
3.3.4.2 Estimativa das áreas das classes de uso e ocupação do solo	70
3.3.4.3 Teste de significância.....	71
4 RESULTADOS E DISCUSSÕES.....	73
4.1 Correção geométrica.....	73
4.2 Calibração radiométrica	73
4.3 Extração de informações das imagens	76
4.3.1 Modelo linear de mistura espectral	76
4.3.2 Índice de vegetação	77
4.4 Classificação digital das imagens.....	78
4.4.1 Máxima verossimilhança	79
4.4.2 Isodata	81
4.5.3 Classificação usando as imagens fração oriundas do modelo linear de mistura espectral	82
4.5.4 Árvore de decisão.....	86
4.6 Comparação entre os algoritmos de classificação.....	87
4.6.1 Estimativa das áreas das classes de uso e ocupação do solo	87
4.6.3 Teste de significância.....	99
5 CONCLUSÕES	104
6 REFERÊNCIAS BIBLIOGRÁFICAS	106
ANEXOS.....	114

LISTA DE TABELAS

TABELA 1	Conceitos do coeficiente de concordância Kappa.....	37
TABELA 2	Situação fundiária do município de Capixaba, 1999.....	41
TABELA 3	Satélite, cena, data de aquisição (dia/mês/ano) e resolução espacial das imagens usadas neste estudo	42
TABELA 4	Tamanho das amostras de treinamento utilizadas na classificação por máxima verossimilhança.....	59
TABELA 5	Número de casos das amostras de treinamento e de teste utilizados na classificação por árvore de decisão.....	66
TABELA 6	Tamanho da amostra de validação por classe temática.....	69
TABELA 7	Valores médios do conjunto de pixels dos componentes água, vegetação e solo, nas bandas 1, 2, 3, 4, 5 e 7 do TM/Landsat em 2003	74
TABELA 8	Exatidão por classe de mapeamento obtido pelo classificador de máxima verossimilhança nos anos de 2003, 2002 e 2000	80
TABELA 9	Exatidão por classe de mapeamento obtido pelo classificador isodata nos anos de 2003, 2002 e 2000.....	81
TABELA 10	Exatidão por classe de mapeamento obtido pelo classificador isodata usando as imagens fração solo e sombra, nos anos de 2003, 2002 e 2000	84
TABELA 11	Exatidão por classe de mapeamento obtido pelo classificador isodata usando as imagens fração solo, sombra e vegetação, nos anos de 2003, 2002 e 2000.....	85
TABELA 12	Exatidão por classe de mapeamento obtido pelo classificador de máxima verossimilhança usando as imagens fração solo, sombra, vegetação e de erro, nos anos de 2003, 2002 e 2000	85
TABELA 13	Análise de exatidão por classe de mapeamento obtido pelo classificador árvore de decisão, nos anos de 2003, 2002 e 2000	87
TABELA 14	Análise de Kappa e variância de Kappa das matrizes de erros que resultaram dos métodos de classificação testados.....	99
TABELA 15	Comparação da análise de Kappa entre as matrizes de erro que resultaram dos métodos de classificação testados em 2003	100

TABELA 16 Comparação da análise de Kappa entre as matrizes de erro que resultaram dos métodos de classificação testados em 2002	102
TABELA 17 Comparação da análise de Kappa entre as matrizes de erro que resultaram dos métodos de classificação testados em 2000	103

LISTA DE FIGURAS

FIGURA 1 Espaço de feições representado pelas bandas 4 e 5 do satélite TM/Landsat, cujos valores dos pixels são expressos em número digital (ND)	8
FIGURA 2 Etapas básicas da classificação supervisionada. (a) imagem multiespectral com cinco números digitais (ND) por pixel; (b) amostras de treinamento: padrão espectral das classes de uso e ocupação do solo; (c) pixel a ser classificado com base em “b” e (d) mapa temático resultado da classificação..	9
FIGURA 3 Três <i>endmembers</i> ou componentes puros formando um triângulo num espaço bidimensional.	14
FIGURA 4 Funções densidade de probabilidade para três classes de padrões unidimensionais.....	20
FIGURA 5 Funções densidade de probabilidade definidas pelo classificador de máxima verossimilhança	22
FIGURA 6 Cluster circular usando a medida de distância euclidiana. ...	25
FIGURA 7 Representação de uma árvore de decisão binária.....	29
FIGURA 8 Localização do município de Capixaba, estado do Acre	38
FIGURA 9 Ilustração das classes de mapeamento de uso e ocupação do solo: (a) floresta, (b) e (c) capoeira em diferentes estágios de sucessão secundária, (d) pasto alto, (e) pasto baixo, (f) solo – área urbana, (g) água – rio Acre e (h) queimada.....	48
FIGURA 10 Imagem Landsat de 2003 na composição colorida das bandas 5, 4 e 3, com a rota percorrida para coleta de amostras das classes temática no município de Capixaba, AC	50
FIGURA 11 Dispersograma com a localização das amostras dos pixels puros no espaço bidimensional das bandas 4 e 5 da imagem TM/Landsat do município de Capixaba, AC, 2003, com valores em reflectância aparente.....	57
FIGURA 12 Esquema da classificação não supervisionada isodata usando as imagens fração solo e sombra.....	62
FIGURA 13 Esquema da classificação não supervisionada isodata usando as imagens fração solo, sombra e vegetação	64

FIGURA 14 Esquema da classificação supervisionada por máxima verossimilhança usando as imagens fração solo, sombra, vegetação e de erro.....	65
FIGURA 15 Curvas espectrais dos componentes água, vegetação e solo, expressas em (a) número digital e (b) reflectância aparente nas bandas 1, 2, 3, 4, 5 e 7 do TM/Landsat em 2003.....	75
FIGURA 16 Imagem TM/Landsat, ano 2003, na composição colorida RGB das bandas 5, 4 e 3 (a), imagem fração sombra (b), imagem fração solo (c) e imagem fração vegetação (d).....	76
FIGURA 17 Imagem TM/Landsat, ano 2003, na composição colorida RGB das bandas 5, 4 e 3 (à esquerda) e imagem NDVI (à direita).....	78
FIGURA 18 Índice de concordância Kappa das seis técnicas de classificação digital para o mapeamento do uso e ocupação do solo em Capixaba, Acre, nos anos de 2000, 2002 e 2003: (1) máxima verossimilhança usando imagens Landsat, (2) isodata usando imagens Landsat, (3) árvore de decisão, (4) isodata usando imagens fração solo e sombra, (5) isodata usando imagens fração solo, sombra e vegetação e (6) máxima verossimilhança usando imagens fração solo, sombra, vegetação e de erro.	79
FIGURA 19 Imagem TM/Landsat na composição colorida RGB das bandas 5, 4 e 3 (a) e mapa resultado da classificação isodata (b).....	82
FIGURA 20 Percentual de área ocupada pelas classes de uso e ocupação do solo em Capixaba, Acre, 2003, segundo a técnica de classificação utilizada	88
FIGURA 21 Percentual de área ocupada pelas classes de uso e ocupação do solo em Capixaba, Acre, 2002, segundo a técnica de classificação utilizada	89
FIGURA 22 Percentual de área ocupada pelas classes de uso e ocupação do solo em Capixaba, Acre, 2000, segundo a técnica de classificação utilizada	89
FIGURA 23 Comparação entre os classificadores de máxima verossimilhança, isodata e árvore de decisão, usando imagem Landsat e imagens fração oriundas do modelo	

linear de mistura espectral no mapeamento da classe temática capoeira, em Capixaba, Acre.....	91
FIGURA 24 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método de máxima verossimilhança usando imagem Landsat, 2003	93
FIGURA 25 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método isodata usando imagem Landsat, 2003	94
FIGURA 26 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método árvore de decisão, 2003	95
FIGURA 27 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método isodata usando as imagens fração solo e sombra, 2003	96
FIGURA 28 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método isodata usando as imagens fração solo, sombra e vegetação, 2003	97
FIGURA 29 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método de máxima verossimilhança usando as imagens fração solo, sombra, vegetação e de erro, 2003	98

RESUMO

FIGUEIREDO, Symone Maria de Melo. Comparação entre técnicas de classificação digital em imagens Landsat no mapeamento do uso e ocupação do solo em Capixaba, Acre. 2005. 127p. Dissertação (Mestrado em Engenharia Florestal) - Universidade Federal de Lavras, Lavras, MG.

A aplicação de diferentes algoritmos de classificação digital gera estimativas variáveis das áreas ocupadas pelas categorias de uso e ocupação da terra ou temas. O objetivo dessa dissertação é avaliar o efeito do uso de vários métodos de classificação de imagens Landsat para o mapeamento do uso e cobertura do solo no município de Capixaba, estado do Acre, por meio da análise comparativa entre os algoritmos de máxima verossimilhança, isodata e árvore de decisão. Na classificação foram utilizadas como atributos as bandas multiespectrais do satélite Landsat, imagens sintéticas oriundas da análise de mistura espectral e imagem gerada pelo índice de vegetação por diferença normalizada – NDVI para mapear floresta, capoeira, pasto alto, pasto baixo, água, queimada e solo. Foram usadas seis técnicas de classificação: a) máxima verossimilhança, utilizando imagens Landsat; b) máxima verossimilhança, utilizando imagens fração; c) isodata, utilizando imagens Landsat; d) isodata, utilizando imagens fração solo e sombra; e) isodata, utilizando imagens fração solo, sombra e vegetação e e) árvore de decisão, utilizando imagens Landsat, imagens sintéticas e NDVI. Os resultados demonstraram que os diferentes métodos de classificação são eficientes, pois apresentaram uma classificação muito boa a excelente, segundo o coeficiente de concordância Kappa. No entanto, entre os métodos testados, a classificação por máxima verossimilhança foi superior às demais técnicas de classificação, com exatidão global superior a 91%.

Palavras-chave: análise de mistura espectral, NDVI, mineração de dados, exatidão do mapeamento, Amazônia.

ABSTRACT

FIGUEIREDO, Symone Maria de Melo. Comparison among digital classification techniques in Landsat images of soil use and occupation in Capixaba, Acre. 2005, 127 p. Dissertation (Master in Forestry Engineering) - Federal University of Lavras, Lavras, MG.

The application of different schemes for digital image classification results in variable estimates of the area occupied by categories of land cover and use. The purpose of this work was to evaluate the effect of using several image classification schemes for mapping land cover and use in the municipality of Capixaba-AC, Brazil. Multi-spectral Landsat data, fraction images resulting from linear spectral unmixing and images representing the Normalized Difference Vegetation Index (NDVI) were used as attributes in the classification procedures in order to map forest, brushwood, high and low pasture, water, burned areas and bare soil. Six classification schemes were used: a) maximum likelihood using Landsat images; b) maximum likelihood using fraction images; c) isodata using Landsat images; d) isodata using fraction images of bare soil and shadow; e) isodata using soil, shadow and vegetation fraction images, and e) decision tree using Landsat images, fraction images and NDVI. Results showed that all classification methods used were efficient ranging from very good to excellent classification according to the Kappa coefficient. Nevertheless, the scheme using the maximum likelihood classification algorithm performed better than the others with an overall precision over 91%.

Key words: linear spectral unmixing, NDVI, data mining, mapping accuracy, Amazon.

1 INTRODUÇÃO

O modelo de desenvolvimento econômico implantado a partir da década de 1970 na Amazônia, estimulado por investimentos públicos, incentivos fiscais, apoio creditício e transferências intergovernamentais de receitas, proporcionou a ocupação do “vazio demográfico”, causando danos irreparáveis aos ecossistemas naturais.

O crescimento econômico foi acompanhado pela abertura de grandes eixos rodoviários, forte incremento populacional, impulsionado pelas migrações e pelas políticas de assentamentos rurais e expansão da fronteira agrícola. A concepção estratégica do poder público não buscou conciliar as dimensões econômica, social e ambiental de desenvolvimento.

As conseqüências desse modelo são desastrosas, marcadas pelos desmatamentos e queimadas para a implantação de grandes áreas de pastagem e agricultura, principalmente em projetos de assentamentos rurais, destruindo a exuberante floresta tropical. A preocupação internacional com a preservação da Amazônia tem pressionado o governo brasileiro a buscar políticas públicas mais eficazes para reduzir esses impactos.

O processo de ocupação no estado do Acre não foi diferente, sobretudo na mesorregião do Vale do Acre, impulsionado pela abertura e pavimentação das rodovias federais BR-317 e BR-364, e das estradas vicinais que dão acesso aos projetos de assentamento e as grandes propriedades rurais (acima de 1.000 hectares). O estado possui superfície territorial de 153.149,9 km², dos quais aproximadamente 10% encontram-se desmatados, incluindo as sedes municipais (Acre, 2000b).

Nesse contexto, a utilização de técnicas de sensoriamento remoto e de sistema de informação geográfica tem-se mostrado uma ótima ferramenta para o

monitoramento ambiental, principalmente em regiões onde há certa carência de informações e uma necessidade de agilizar o processo de gerenciamento dos recursos naturais, como no caso do estado do Acre.

A alteração da cobertura florestal (desmatamento) tem sido adotada como o principal indicador da intensidade e magnitude das atividades humanas. Desde 1988, o governo do Brasil monitora a situação da cobertura florestal na Amazônia. Anualmente, o Instituto Nacional de Pesquisas Espaciais (INPE) divulga o percentual de alteração da cobertura florestal da região. Essa informação é amplamente utilizada como parâmetro para avaliar o desempenho das ações do governo no controle do ritmo do desmatamento.

A taxa de desmatamento é uma informação útil e importante, porém, é insuficiente para orientar as ações de monitoramento e controle ambiental. É necessário saber não apenas o quanto foi alterado (taxa de desmatamento), mas também onde a alteração ocorreu.

A experiência mais recente do Acre em monitoramento da cobertura florestal é o levantamento feito pela Fundação de Tecnologia do Acre (FUNTAC) para o ano de 1996. Esse trabalho foi utilizado como base para a primeira etapa do Zoneamento Econômico-Ecológico (Acre, 2000a). O Instituto de Meio Ambiente do Acre (IMAC), por meio do Projeto de Gestão Ambiental Integrada, financiado pelo Programa Piloto para Proteção das Florestas Tropicais do Brasil, vem realizando, juntamente com o Instituto do Homem e Meio Ambiente da Amazônia (Imazon), o mapeamento da cobertura florestal do estado desde 1999, para auxiliar nas ações de controle dos recursos florestais.

As metodologias de classificação de imagens digitais adotadas pelas diferentes organizações governamentais e não governamentais resultaram em diferentes índices de desmatamento no estado, dificultando a implementação de políticas públicas de controle ambiental e desenvolvimento sustentável. Um exemplo, citado por Rocha (2002), refere-se ao ano de 1996, quando o INPE,

FUNTAC e Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) definiram a porcentagem de áreas desflorestadas no estado do Acre em 9,0%, 7,4% e 5,0%, respectivamente. Segundo o autor, considerando a taxa média de desflorestamento no Acre de 600 km²/ano, dada pelo INPE no período de 1978 a 1999, a diferença de 1,6% entre os índices do INPE e FUNTAC em 1996 equivalem a quatro anos de desflorestamento para o estado. Essas diferenças são resultado dos conceitos utilizados por cada instituição, bem como das técnicas de classificação aplicadas.

Nesse contexto, este estudo teve como objetivo avaliar o efeito do uso de vários métodos de classificação de imagens Landsat para o mapeamento do uso e cobertura do solo no município de Capixaba, estado do Acre, por meio da análise comparativa entre os algoritmos de máxima verossimilhança, isodata, e árvore de decisão, usando como entrada de dados imagens Landsat e imagens originadas de técnicas de extração de informações. Esta relação poderá subsidiar os órgãos ambientais na definição de uma metodologia mais apropriada às condições regionais e aplicável no licenciamento, fiscalização e monitoramento ambiental, como um importante instrumento de gestão e controle dos recursos naturais, bem como subsidiar políticas públicas de desenvolvimento sustentável.

2 REFERENCIAL TEÓRICO

2.1 Processo de ocupação no estado do Acre

A área correspondente ao atual estado do Acre não pertencia ao Brasil antes do século XIX e foi fruto de litígio entre Brasil e Bolívia. A descoberta da vulcanização em 1844 e a invenção dos pneumáticos em 1888 motivaram a ocupação do território pelos brasileiros, prioritariamente por nordestinos, para extração do látex com o objetivo de suprir a demanda mundial da borracha, que era a matéria-prima cobiçada por países da Europa e dos Estados Unidos. Em 1852, a área foi anexada à Província do Amazonas, porém, em 1898 foi estabelecido o Departamento Boliviano do Acre. Somente em 1903, após a assinatura do tratado de Petrópolis, o território foi anexado ao Brasil (Acre, 2000b).

No início da década de 1940, a assinatura dos Acordos de Washington promoveu o segundo despertar da economia extrativista da seringueira (Homma, 2003). Os Estados Unidos investiram na produção do látex para suprir a indústria bélica dos aliados na Segunda Guerra Mundial e houve a migração de 20 a 30 mil nordestinos castigados pela seca para impulsionar os seringais da Amazônia. Terminada a Segunda Guerra Mundial, a Amazônia passou a fornecer também minérios, petróleo, algodão e produtos da pecuária (Acre, 2000b).

Nos anos 1970, o Acre sofreu uma experiência colonizadora, por parte do Instituto Nacional de Colonização e Reforma Agrária (INCRA), com a implantação de projetos de colonização como, por exemplo, Pedro Peixoto (250,0 mil hectare), Boa Esperança (200,0 mil hectares), Quixadá, Humaitá e Santa Luzia, entre outros, que alteraram o quadro fundiário do estado. Antes, só

havia a grande propriedade seringueira, fruto da convalidação de títulos dos mais variados matizes e das mais variadas formas de falsificação. Já em 1992, os dados do INCRA evidenciaram a reconcentração fundiária, fruto do abandono de muitos Projetos Integrados de Colonização (PICs) e da agregação de suas parcelas (Cardim et al., 2004).

Os planos e programas governamentais no início da década de 1970 (I Plano de Integração Nacional – I PIN, Programa de Redistribuição de Terras e Estímulo à Agroindústria do Norte e Nordeste – PROTERRA, I Plano de Desenvolvimento da Amazônia – I PDA, Programa de Incentivo à Produção de Borracha Vegetal I – Probor I e Programa de Pólos Agropecuários e Agrominerais da Amazônia – Polamazônia) visavam o desenvolvimento da região por meio dos assentamentos em projetos de colonização ao longo das rodovias Transamazônica (inaugurada em 1972) e Cuiabá-Santarém. Eles previam o desenvolvimento da agricultura e da pecuária, a integração e a ocupação da Amazônia, bem como abriu caminho para investimento do grande capital para ocupação de áreas, privilegiando projetos agrominerais e agropecuários (Acre, 2000b; Homma, 2003).

Com essa forma de ocupação aliada à decadência dos seringais que eram vendidos a baixo custo no sul do país, fomentou-se a compra de terras e a implantação da pecuária extensiva em substituição a áreas extrativistas, modificando sobremaneira a forma de uso da terra. Os investimentos na região e os assentamentos de colonos demandaram a abertura de novas vias de acesso, como as rodovias federais BR-364 e BR-317. Segundo Acre (2000b), a criação de estradas sem um plano ambiental coerente trouxe complicações sociais, principalmente devido aos impactos causados ao meio ambiente. As políticas desenvolvimentistas do governo federal causaram sérios problemas para a Amazônia, pois produziram uma ocupação não sistemática da terra e colaboraram com atividades que envolviam o desmatamento de extensas áreas

de floresta. Em 1978, o desmatamento na Amazônia Legal atingiu 15.220 ha, ou 3,6% da floresta original (Homma, 2003).

Na década de 1980, o acontecimento mais marcante foi o assassinato do líder seringueiro Chico Mendes, em 22 de dezembro de 1988 em Xapuri, Acre, com repercussão internacional. A partir de então, o governo brasileiro se viu pressionado a mudar a postura quanto às questões ambientais.

No período compreendido entre 1970 a 2000, segundo Acre (2000b), os dados demográficos apontam um crescimento acelerado da população total do estado, saltando de 215 mil habitantes, em 1970, para 547 mil, em 2000, e uma inversão de predominância da população rural para população urbana. Por outro lado, a partir da década de 1970, verificou-se mudanças estruturais no setor primário do estado, com redução drástica do extrativismo na composição do PIB do Acre e aumento da contribuição da atividade agropecuária. Os dados referentes ao imposto de circulação de mercadorias e serviços (ICMS) mostram que as atividades agropecuárias vem representando mais de 50% do total arrecadado do setor primário desde 1997 (Acre, 2000b).

Segundo Sawyer (2001), a maior parte do desmatamento na Amazônia no período de 1970 a 2000 deveu-se à implantação de fazendas e projetos agropecuários empresariais. No entanto, os agricultores familiares respondem por parcela significativa e crescente do desmatamento e os assentamentos da reforma agrária, em sua forma atual de colonização, promovem o desmatamento atual nos lotes e futuramente em áreas adjacentes.

Diante das políticas e posições conflitivas de ocupação e desenvolvimento da Amazônia, Becker (2001) afirma que o desafio do presente é definir e implementar um novo padrão de desenvolvimento capaz de melhorar as condições de vida da população, de estancar o desflorestamento e utilizar o seu patrimônio natural com formas conservacionistas.

2.2 Classificação digital de imagens de sensoriamento remoto

2.2.1 Princípios do reconhecimento de padrões

A análise de imagens digitais é um processo de descobrimento, de identificação e de entendimento de padrões que sejam relevantes ao desempenho de uma determinada tarefa baseada em imagem. Dessa forma, o objetivo da classificação digital de imagens de sensoriamento remoto é categorizar automaticamente todos os pixels da imagem em classes relacionadas ao uso e ocupação do solo ou temas.

Normalmente, dados multiespectrais são usados para realizar a classificação, ou seja, o padrão espectral presente nos dados de cada pixel é empregado como base numérica para a categorização em classes de uso e ocupação do solo (Lillesand & Kiefer, 2000).

Vários autores definem o conceito de padrão ou reconhecimento de padrão na análise de imagens digitais. Gonzalez & Woods (2000) definem como uma descrição quantitativa ou estrutural de um objeto ou alguma entidade de interesse de uma imagem. Buiten (1996) descreve reconhecimento de padrão como a atribuição de um rótulo (nome de objeto, categoria ou classe) para o pixel a ser classificado, normalmente baseado no padrão espectral de áreas de treinamento selecionadas na imagem.

Para Lillesand & Kiefer (2000), o termo padrão refere-se ao conjunto de radiância medidas nos vários comprimentos de ondas ou bandas para cada pixel e, dessa forma, o reconhecimento de padrão espectral refere-se à família de procedimentos de classificação que utiliza a informação espectral pixel-a-pixel como base para a classificação automática do uso e ocupação do solo.

De acordo com Tso & Mather (2001), trata-se de vetor de feições que descrevem um objeto. Este padrão é feito por meio da mensuração de um

conjunto de feições representadas pelos eixos de um espaço multidimensional chamado espaço de feições ou espaço de atributos.

Na utilização de imagens de satélite, os eixos do espaço de feições são representados pelas respectivas bandas espectrais, cujos valores podem ser expressos em número digital ou reflectância (Figura 1). Dessa forma, o processo de reconhecimento de padrões envolve a subdivisão do espaço de feições em regiões homogêneas separadas por fronteiras de decisão.

Segundo Tso & Mather (2001), o objetivo do reconhecimento de padrões é o de estabelecer uma relação entre o padrão e a classe de padrão. O objeto poder ser um único pixel ou um conjunto de pixels adjacentes formando uma entidade geográfica.

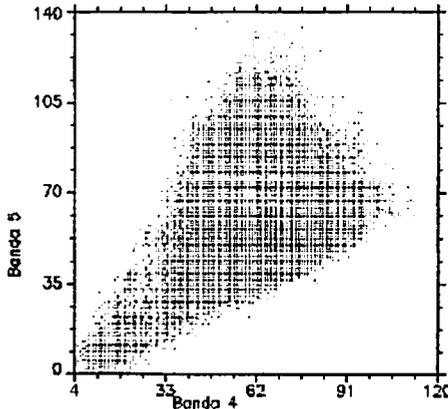


FIGURA 1 Espaço de feições representado pelas bandas 4 e 5 do satélite TM/Landsat, cujos valores dos pixels são expressos em número digital (ND)

O analista pode conhecer ou não a classe de padrão, devendo, no primeiro caso, estar apto a conhecer todas as classes presentes na área de estudo. Em outros casos, o analista deve determinar o número de categorias separáveis, sua localização e extensão. Usando essa informação, a separabilidade de classes é atribuída ao conhecimento do analista.

Esses dois métodos são conhecidos como classificação supervisionada e não supervisionada. Métodos supervisionados requerem o uso de coleção de amostra para o “treinamento” do classificador e, desta maneira, para determinar a fronteira de decisão no espaço de feições (Tso & Mather, 2001).

Na classificação supervisionada, o analista “supervisiona” o processo de categorização. Para isso, ele fornece, baseado em seu conhecimento, locais de amostras representativas dos tipos de uso e ocupação do solo presentes na cena, chamadas áreas de treinamento, usadas como “interpretação do padrão” e que descrevem os atributos espectrais de cada tipo de uso e ocupação do solo de interesse (Lillesand & Kiefer, 2000). Cada pixel do conjunto de dados é comparado numericamente com cada categoria de “interpretação do padrão” e rotulado com o nome da categoria mais semelhante.

A Figura 2 representa as principais etapas realizadas no procedimento de classificação supervisionada.

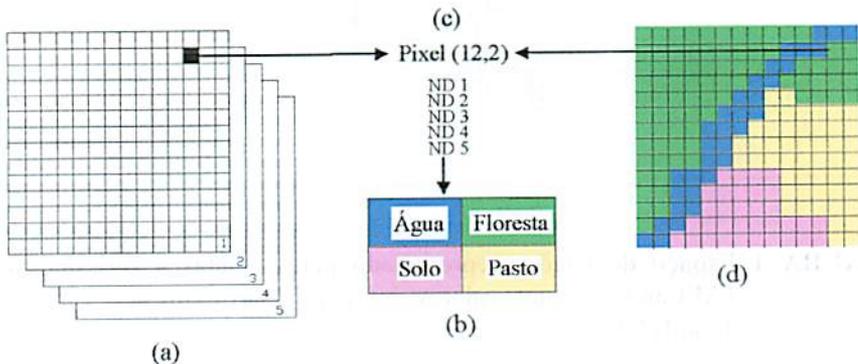


FIGURA 2 Etapas básicas da classificação supervisionada. (a) imagem multiespectral com cinco números digitais (ND) por pixel; (b) amostras de treinamento: padrão espectral das classes de uso e ocupação do solo; (c) pixel a ser classificado com base em “b” e (d) mapa temático resultado da classificação. Fonte: adaptado de Lillesand & Kiefer, 2000.

Na classificação supervisionada, as fronteiras de decisão entre as classes de uso e ocupação do solo são afetadas significativamente pelo tamanho das amostras usadas para treinar o classificador, bem como do conhecimento do analista para construí-las (Tso & Mather, 2001). Portanto, o analista deve ter conhecimento suficiente do tipo e número de classes da área de estudo e, dessa forma, decidir pelo conjunto de amostras de treinamento que sejam representativas dos pixels da imagem, pois a qualidade das amostras no processo de treinamento determina o sucesso da classificação (Novo, 1998; Lillesand & Kiefer, 2000; Tso & Mather, 2001; Moreira, 2003).

Em contraste, os métodos de reconhecimento de padrões não supervisionados são menos dependentes e não requerem conhecimento prévio da área de estudo. Normalmente, os classificadores não supervisionados “aprendem” as características de cada classe diretamente da entrada de dados, ou seja, da imagem em análise (Novo, 1998; Lillesand & Kiefer, 2000; Tso & Mather, 2001).

Neste caso, segundo Tso & Mather (2001), usando o critério da distância mínima entre o objeto e a classe média, o procedimento não supervisionado estima a média de cada classe e refina essa média iterativamente. A cada iteração, o conjunto de classes médias estimado anteriormente é refinado até o processo convergir, usualmente quando a média permanecer aproximadamente no mesmo lugar no espaço de feições após sucessivas iterações.

Os resultados de saída gerados por métodos não supervisionados são chamados agrupamentos (“*clusters*”) ou classes de dados. O processo de reconhecimento de padrões é concluído quando cada agrupamento é identificado, isto é, ligado a uma classe característica. A acurácia da classificação não supervisionada, geralmente, é menor em relação aos métodos supervisionados (Tso & Mather, 2001).

Outra questão relevante em classificações é que as classes são, muitas vezes, sobrepostas. No domínio espectral, isso implica que a reflectância emitida ou refletida por diferentes objetos podem ser similares. No domínio espacial, um pixel geralmente contém áreas representativas de mais de uma classe, ocasionado o problema de mistura espectral. A sobreposição espacial e espectral de classes consiste na principal barreira para elevar a acurácia da classificação, fornecendo, muitas vezes, estimativas imprecisas das áreas ocupadas por cada classe representativa da imagem (Aguilar, 1991; Tso & Mather, 2001).

Por mais de uma década, os métodos de reconhecimento de padrões aplicados ao sensoriamento remoto baseavam-se principalmente nas técnicas estatísticas convencionais, como os procedimentos de máxima verossimilhança e distância mínima. Esses métodos geralmente possuem habilidade limitada de resolver confusões entre classes. Recentemente, avanços nas técnicas computacionais propõem estratégias alternativas na classificação de imagens digitais, conhecidas como mineração de dados. Para Tso & Mather (2001), particularmente redes neurais artificiais, árvores de decisão e métodos derivados da teoria de conjuntos fuzzy são usados no procedimento de classificação com o propósito de melhorar a acurácia.

2.2.2 Transformações nas imagens digitais para extração de informações

A complexidade de informações adquiridas pelos sensores remotos aumentou com a capacidade de coletar simultaneamente imagens em diferentes regiões do espectro eletromagnético (Novo, 1998). Dessa forma, são aplicadas técnicas para a extração de informações das imagens, que conseguem realçar a diferença do comportamento espectral dos alvos na superfície terrestre, auxiliando na interpretação e análise digital.

2.2.2.1 Análise de mistura espectral

A mistura espectral consiste na integração da energia espectral refletida ou emitida por todos os objetos, denominados componentes da mistura, contidos no pixel, devido à variação e proximidade dos alvos na superfície da terra (Lillesand & Kiefer, 2000; Aguiar, 1991; Ferreira et al., 2003). Os pixels misturados resultam quando o campo de visada instantâneo, do inglês *Instantaneous Field of View* (IFOV), de um sensor observa a combinação da radiância desses diversos materiais.

Lillesand & Kiefer (2000) consideram que a extensão de pixels misturados contidos numa imagem é função da resolução espacial do sistema de sensoriamento remoto usado para adquirir a imagem e da escala espacial das feições na superfície. Segundo Aguiar (1991), o fenômeno de mistura ocorre tanto nos casos em que o tamanho dos objetos ou feições que estão sendo imageados pelo sensor seja menor que o pixel, ou maior, quando seus limites encontram-se dentro do pixel. Além disso, outros fatores, como a contaminação atmosférica e variações na geometria de aquisição, também contribuem com a mistura (Ferreira et al., 2003).

A mistura espectral é considerada, por vários autores, como um difícil problema para classificação de imagens, levando a estimativas imprecisas da área ocupada por cada classe de uso e ocupação do solo. Assim, é proposto na literatura o procedimento de análise de mistura espectral como um meio para efetuar a classificação sub pixel.

Para Tso & Mather (2001), a análise de mistura espectral consiste numa alternativa de determinação da proporção relativa dos componentes ou tipos de cobertura que cobrem o terreno dentro do pixel. O método usa um conjunto de componentes puros (*endmembers*), definidos como classes de informação puras, representando tipos ideais para todas as misturas presentes na imagem.

Dessa forma, a análise de mistura espectral envolve técnicas em que assinaturas espectrais misturadas são comparadas com um conjunto de referência espectral “pura” (Lillesand & Kiefer, 2000). A hipótese básica é que a variação espectral numa imagem é causada por misturas de um número limitado de materiais na superfície. O resultado é uma estimativa de proporções aproximadas das áreas em cada pixel que são ocupadas por cada classe de referência.

A qualidade das imagens fração geradas pela análise de mistura espectral depende da seleção apropriada dos componentes puros e do número suficiente de componentes (Lu et al., 2003). A obtenção de componentes puros pode ser realizada por dois caminhos (Meer & Jong, 2000):

- a) selecionando-os de uma biblioteca espectral, obtidos por meio de mensurações de reflectância em laboratório ou no campo, ou
- b) derivando-os de pixels puros na imagem.

O segundo método tem a vantagem dos componentes puros selecionados ocorrerem em condições atmosféricas similares às da imagem de dados. Lu et al. (2003) citaram, ainda, outros procedimentos para obter os componentes puros, como a identificação de pixels puros da imagem derivada na primeira ordem da análise de componentes principais e usando o índice de pureza do pixel (PPI), disponível no software ENVI.

Para muitas aplicações e pela facilidade de obtenção, a identificação dos componentes puros na imagem ocorre utilizando-se o espaço de feições. Cada componente puro tem uma localização no espaço de feições n -dimensional, no qual n é o número de bandas espectrais usadas para descrever o pixel. Um poliedro é formado pela ligação das posições dos componentes puros no espaço de feições. Todas as possíveis misturas com proporções não-negativa de componentes puros devem estar inclusas no poliedro (Settle & Drake, 1993 citado por Tso & Mather, 2001).

A Figura 3 ilustra a localização de três componentes puros ou *endmembers* no espaço de feições bidimensional, formando um triângulo. No interior do triângulo encontram-se os pixels misturados, enquanto nos seus vértices estão localizados os pixels puros.

A análise de mistura espectral difere de outros métodos de classificação de imagem, pois conceitualmente consiste num método mais determinístico do que estatístico, já que é baseado num modelo físico de resposta de padrões espectrais discretos de mistura (Lillesand & Kiefer, 2000).

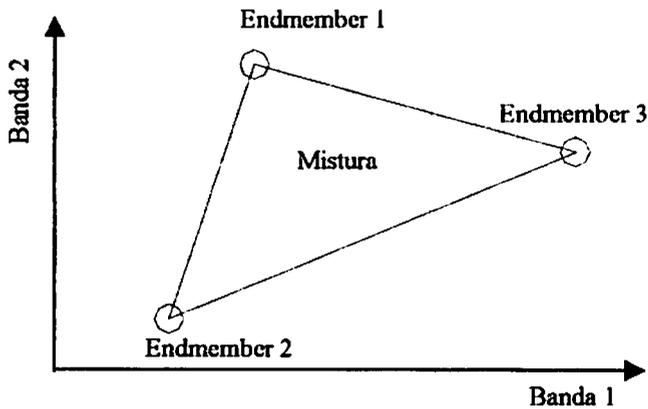


FIGURA 3 Três *endmembers* ou componentes puros formando um triângulo num espaço bidimensional. Fonte: Tso & Mather (2001).

Muitas aplicações de análise de mistura espectral utilizam modelos lineares de mistura, nos quais são observadas as respostas espectrais de uma área do terreno pela mistura linear de assinaturas espectrais individuais dos vários tipos de uso e ocupação do solo presentes na área (Lillesand & Kiefer, 2000). Essas assinaturas espectrais puras, os *endmembers*, representam os casos em

que, teoricamente, 100% do campo de visada do sensor são ocupados por um único tipo de cobertura.

Nesse modelo, o peso de cada *endmember* é proporcional à área ocupada pela classe correspondente. A entrada de dados para o modelo linear de mistura consiste na assinatura espectral observada para cada pixel numa imagem, a partir das quais são geradas, pelo modelo, as imagens de abundância ou fração de cada pixel que é ocupada por cada *endmember*.

A formulação matemática do modelo linear de mistura pode ser expressa por (Lu et al., 2003):

$$R_i = \sum_{k=1}^n f_k R_{ik} + \varepsilon_i$$

em que:

$i = 1, \dots, m$ (número de bandas espectrais);

$k = 1, \dots, n$ (número de componentes puros ou *endmembers*);

R_i = reflectância espectral na banda i de um pixel, contendo um ou mais *endmembers*;

f_k = proporção do *endmember* k dentro de um pixel;

R_{ik} = reflectância espectral do *endmember* k dentro do pixel na banda i ;

e,

ε_i = erro da banda i .

Por ser tratar de proporções de área de um pixel, o modelo linear de mistura é submetido às seguintes restrições:

$$\sum_{k=1}^n f_k = 1 \text{ e } 0 \leq f_k \leq 1$$

Em dados multiespectrais, segundo Lillesand & Kiefer (2000), as equações são calculadas para cada banda espectral. Dessa forma, são $m + 1$ equações para resolver as várias proporções de *endmembers* (f_1, f_2, \dots, f_k). Se o número de *endmembers* é igual ao número de bandas espectrais mais 1, o conjunto de equações pode ser resolvido simultaneamente para produzir uma solução exata. Se o número de bandas $m + 1$ é maior que o número de *endmembers* k , a magnitude do termo de erro junto com a cobertura proporcional para cada *endmember* pode ser estimado utilizando-se os princípios da regressão dos mínimos quadrados. Por outro lado, se o número de *endmembers* presentes na cena excede $i + 1$ bandas, o conjunto de equações não pode ser submetido a uma única solução.

2.2.2.2 Índice de vegetação

A maioria dos sensores remotos opera na faixa de $0,38 - 15 \mu\text{m}$ do espectro eletromagnético, a qual está subdividida em quatro sub-regiões: visível (azul, verde e vermelho), infravermelho próximo, médio e termal (Gurgel, 2000). O comportamento espectral da vegetação fotossinteticamente ativa, segundo Novo (1998), é caracterizado por baixa reflectância ($< 0,2$) até $0,7 \mu\text{m}$, causada pela absorção da radiação incidente pelos pigmentos da planta, responsável pela cor verde da vegetação. Na região entre $0,7 \mu\text{m}$ a $1,3 \mu\text{m}$, a reflectância é alta ($0,3 < \rho < 0,4$), devido a ausência da absorção pelos pigmentos da planta. Entre $1,3 \mu\text{m}$ e $2,5 \mu\text{m}$, a reflectância da vegetação é dominada pelo conteúdo de água das folhas e esta região corresponde também às bandas de absorção atmosférica e, por isso, os sensores desenvolvidos têm suas faixas espectrais deslocadas para regiões sujeitas à atenuação atmosférica.

Os índices de vegetação foram criados com o objetivo de ressaltar o comportamento espectral da vegetação em relação aos outros alvos da superfície

terrestre, além de reduzir a dimensionalidade dos dados orbitais por meio do realce de informações espectrais da vegetação no menor número de bandas (Moreira, 2003).

As bandas espectrais normalmente usadas para a determinação dos índices de vegetação são correspondentes às bandas do vermelho e infravermelho próximo, devido ao comportamento espectral da vegetação nessas faixas do espectro eletromagnético. São citadas na literatura várias metodologias para o cálculo do índice de vegetação (Lillesand & Kiefer, 2000; Moreira, 2003), porém, o índice NDVI ou *Normalized Difference Vegetation Index* ou Índice de Vegetação por Diferença Normalizada, é bastante utilizado no monitoramento da vegetação, uma vez que ajuda a compensar, entre outros, mudanças nas condições de iluminação, declividade da superfície e geometria de aquisição, que influenciam a resposta espectral das bandas dos sensores remotos (Shimabukuro et al., 1998; Lillesand & Kiefer, 2000). O cálculo do NDVI é obtido pela razão entre a diferença da resposta espectral do pixel nas bandas do infravermelho próximo e do vermelho e a soma dessas bandas.

Os valores de NDVI variam entre 0,1 e 0,6 para as áreas com vegetação em função da sua arquitetura, densidade e umidade. As coberturas vegetais mais densas possuem os maiores valores de NDVI; nuvens e água, geralmente, apresentam valores negativos de NDVI; rochas e solos expostos possuem valores próximos a zero, ocasionados pela reflectância similar nas duas regiões espectrais utilizadas para o cálculo do índice (Holben, 1986 citado por Gurgel, 2000).

O NDVI é aplicado para estudos de dinâmica da vegetação em áreas de clareiras de florestas tropicais, mensuração de índice de área foliar, estimacão de biomassa, determinacão do percentual do uso e ocupacão do solo e estimacão da radiação fotossinteticamente ativa (Lillesand & Kiefer, 2000; Moreira ,2003).

2.2.3 Algoritmos de classificação

Existe na literatura uma variedade de algoritmos para reconhecimento e interpretação de imagens, disponível aos usuários nos diversos softwares de processamento de imagens, desde os mais tradicionais até os baseados em recentes avanços no campo de aprendizado de máquina.

Os algoritmos de máxima verossimilhança e isodata são os mais empregados no procedimento de classificação supervisionada e não supervisionada, respectivamente.

A mineração de dados por árvore de decisão, apesar de ser uma técnica relativamente nova, é relatada como promissora na análise de dados de sensoriamento remoto (Carvalho et al., 2004).

2.2.3.1 Máxima verossimilhança

O método de máxima verossimilhança é um procedimento estatístico supervisionado de reconhecimento de padrões, no qual a probabilidade de cada pixel pertencer ao conjunto de classes pré-definidas é calculada, sendo o pixel atribuído para a classe de maior probabilidade (Tso & Mather, 2001). Para isso, assume-se que a distribuição dos agrupamentos de pixels que formam as amostras de treinamento é gaussiana, ou seja, normalmente distribuída (Lillesand & Kiefer, 2000). A distribuição das classes de padrões espectrais é descrita por meio de dois parâmetros, o vetor média e a matriz de covariância, pelos quais calcula-se a probabilidade estatística de um dado valor de pixel pertencer a uma determinada classe de uso e ocupação do solo.

Tso & Mather (2001) descrevem a formulação estatística da máxima verossimilhança que é baseada na probabilidade Bayesiana:

$$P(x, w) = P(w/x) P(x) = P(x/w) P(w) \quad (1)$$

Em que x e w são geralmente chamados eventos; $P(x,w)$ é a probabilidade de co-existência (ou interseção) dos eventos x e w ; $P(x)$ e $P(w)$ são as probabilidades a priori dos eventos x e w ; $P(w/x)$ é a probabilidade condicional de um dado evento x em relação ao evento w e $P(x/w)$ é interpretado da mesma maneira. Se o evento x_i é o i -ésimo vetor padrão e w_j representa as informações da j -ésima classe, de acordo com a equação (1), a probabilidade de que x_i pertença à classe w_j dada por:

$$P(w_j/x_i) = \frac{P(x_i/w_j) P(w_j)}{P(x_i)} \quad (2)$$

Como, geralmente, $P(x)$ é uniformemente distribuída, isto é, a probabilidade de ocorrência é a mesma para todos os pixels (Tso & Mather (2001), o cálculo pode ser simplificado, pois $P(x)$ é o denominador de todas as classes (Moreira, 2003). Assim, a Equação (2) pode ser reescrita por:

$$P(w_j/x_i) = P(x_i/w_j) P(w_j) \quad (3)$$

A quantidade $P(x_i/w_j)$, que é uma probabilidade, é chamada verossimilhança de w_j em relação à x_i . Uma regra de decisão simplificada é classificar o vetor x_i como pertencente à classe w_j de maior valor de $P(x_i/w_j)$, ou seja, de máxima verossimilhança.

Como exemplo, Moreira (2003) considerou três classes w_1 , w_2 e w_3 , com as respectivas funções de probabilidade mostradas na Figura 4. Aplicando-se o método da máxima verossimilhança, os pontos no gráfico até o limite de decisão x_1 seriam classificados como da classe w_1 , pois possuem maior probabilidade de

pertencer a essa classe. Os pontos localizados entre os limites x_1 e x_2 pertenceriam à classe w_2 , enquanto aqueles localizados entre x_2 e x_3 à classe w_3 , e, por último, os pontos acima do limite de decisão x_3 , classificados como da classe w_2 .

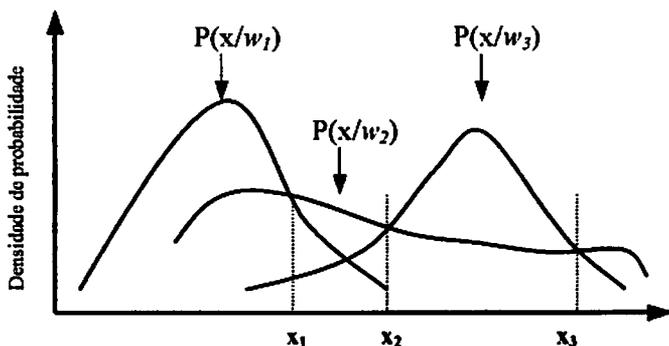


FIGURA 4 Funções densidade de probabilidade para três classes de padrões unidimensionais. Fonte: Moreira, 2003.

No caso n -dimensional, a densidade gaussiana dos vetores de j -ésima classe de padrões tem a seguinte forma (Gonzalez & Woods, 2000):

$$P(x/w_j) = \frac{1}{(2\pi)^{n/2} |C_j|^{1/2}} \exp \left[-\frac{1}{2} (x - m_j)^T C_j^{-1} (x - m_j) \right] \quad (4)$$

Em que x é um vetor n -dimensional; m_j o vetor média n -dimensional; C_j a matriz de covariância; $|C_j|$ o determinante de C_j ; e, C_j^{-1} sua inversa.

Uma distribuição normal multivariável estará completamente especificada se forem determinados o vetor média m_j e a matriz de covariância C_j (Moreira, 2003). Estes dois parâmetros são estimados, a partir das amostras de treinamento, pelas fórmulas:

$$\mathbf{m}_j = \frac{1}{N_j} \sum_{\mathbf{x} \in w_j} \mathbf{x} \quad e, \quad (5)$$

$$C_j = \frac{1}{N_j} \sum_{\mathbf{x} \in w_j} (\mathbf{x} - \mathbf{m}_j)(\mathbf{x} - \mathbf{m}_j)^T$$

em que N_j é o número de vetores de padrões da classe w_j , e a soma é feita sobre esses vetores.

Na prática, é mais conveniente trabalhar com o logaritmo natural da equação (4), devido à forma exponencial da densidade gaussiana, reduzindo a expressão para:

$$\ln[P(\mathbf{x}_i / w_j)] = -\frac{n}{2} \ln 2\pi - \frac{1}{2} \ln |C_j| - \frac{1}{2} [(\mathbf{x}_i - \mathbf{m}_j)^T C_j^{-1} (\mathbf{x}_i - \mathbf{m}_j)] \quad (6)$$

O termo $(n/2) \ln 2\pi$ é o mesmo para todas as classes, de maneira que pode ser eliminado da equação (6). Finalmente, multiplicando a constante -2 na equação (6), tem-se:

$$-\ln[P(\mathbf{x}_i / w_j)] = \ln |C_j| + (\mathbf{x}_i - \mathbf{m}_j)^T C_j^{-1} (\mathbf{x}_i - \mathbf{m}_j) \quad (7)$$

A equação (7) representa as funções de decisão de Bayes para classes de padrões gaussianas.

A Figura 5 mostra os valores de probabilidade no gráfico tridimensional. O eixo vertical é associado com a probabilidade de um pixel pertencer a uma das classes. A superfície em formato de sino são as funções de densidade probabilidade, uma para cada classe ou categoria espectral. A forma geométrica formada pelo conjunto de pixels pertencentes a uma determinada classe é elipsóide, delineando contornos de igual probabilidade, com forma expressa pela covariância que podem ser vistas como limites de decisão

(Lillesand & Kiefer, 2000; Tso & Mather, 2001). Os contornos ao redor de cada classe na Figura 5 podem ser entendidos como a probabilidade de qualquer pixel desconhecido pertencer a uma determinada classe.

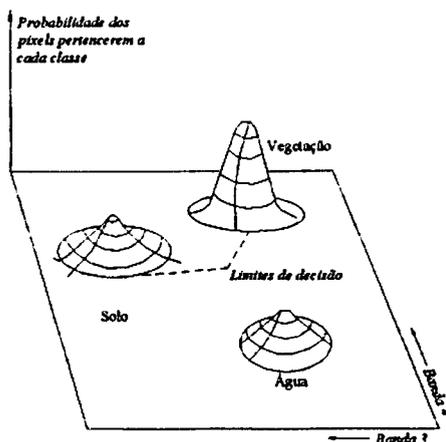


FIGURA 5 Funções densidade de probabilidade definidas pelo classificador de máxima verossimilhança. Fonte: Aguiar, 1991.

Segundo Crosta (1993), essas probabilidades representam uma ferramenta de classificação poderosa e flexível, pois é possível escolher, por exemplo, classificar apenas os pixels desconhecidos, que são bastantes semelhantes à média de uma classe ou os pixels que se situam a qualquer distância da média dentro da área geral definida pela amostra de treinamento. Dessa forma, pode-se classificar todos os pixels de uma imagem como pertencente a uma classe pré-definida, mesmo que eles tenham uma baixa probabilidade.

Devido ao fato do classificador utilizar o vetor média e a matriz de covariância como parâmetros estatísticos para definir cada classe, é importante

selecionar amostras de treinamento de tamanho adequado para fornecer estimativas eficientes desses parâmetros.

Tso & Mather (2001) consideram que o tamanho mínimo do conjunto de amostras de treinamento depende consideravelmente da extensão e número de feições usadas para caracterizar os objetos a serem classificados. Por exemplo, o ETM+/Landsat fornece sete bandas com 256 níveis de cinza por banda. Teoricamente, é possível um pixel ocupar uma das $(2^8)^7 = 7 \times 10^{16}$ posições discretas no espaço de feições. Esse número é maior que o total de pixels da imagem. Para um dado tamanho de amostra de treinamento, um incremento na dimensionalidade no espaço de feições (por exemplo, adicionando-se bandas espectrais) implica no aumento do número de parâmetros requeridos na classificação estatística. No procedimento de máxima verossimilhança, o número de parâmetros é dado por $k(k-1)/2 + 2k$, em que k = elementos a serem estimados. Para um aumento de k , o número de parâmetros a serem estimados aumenta desproporcionalmente. Para $k = 6$, o número de parâmetros é 27. Para $k = 12$ o valor aumenta para 90.

O fato de que uma estimação eficiente dos parâmetros estatísticos requer uma amostra representativa de tamanho suficiente é conhecido. Conseqüentemente, com o aumento do número de parâmetros, para um dado tamanho de amostra, a eficiência da estimação decresce. Isso é conhecido como fenômeno Hughes (Tso & Mather, 2001). Dessa forma, para obterem-se resultados eficientes na classificação de imagens de sensoriamento remoto, a relação entre dimensionalidade e tamanho de amostra de treinamento precisa ser considerada.

Na prática, diversos autores sugerem o tamanho mínimo das amostras de treinamento. Para Tso & Mather (2001), em dados com baixa dimensionalidade, resultados satisfatórios podem ser obtidos, na maioria dos casos, com tamanho de 30 x número de bandas espectrais. Lillesand & Kiefer (2000) recomendam o

mínimo de $10n$ a $100n$ pixels, em que n é o número de bandas espectrais, para as estimativas dos vetores média e matrizes de covariância.

2.2.3.2 Isodata

O método de classificação isodata é, provavelmente, o mais conhecido e é descrito como um meio de interpretação de imagens de sensoriamento remoto assistida por computador (Sulsoft, 2004). Segundo Tso & Mather (2001), é o método não supervisionado mais usado para a classificação de imagens digitais e consiste numa técnica de classificação não supervisionada que utiliza algoritmos de análise de agrupamento e migração para o reconhecimento de padrões.

O método não utiliza amostras de treinamento como base para classificação. O algoritmo isodata examina pixels desconhecidos na imagem e agrega-os num número de classes baseados em arranjos naturais ou agrupamentos (*clusters*) presente nos valores da imagem (Novo, 1998; Lillesand & Kiefer, 2000), considerando sua distribuição num espaço de n dimensões, no caso, bandas espectrais (Sulsoft, 2004).

O algoritmo isodata estima um número inicial de *clusters* n , presente nos dados e determina a posição dos n *clusters* médios no espaço de feições, gerando aleatoriamente vetores média. Cada pixel é então associado com o centro mais próximo do *cluster*, definido normalmente pela medida de distância euclidiana (Figura 6). No próximo estágio, a localização de cada *cluster* médio é recalculada baseada no novo conjunto de pixels alocados para os *clusters*. A cada iteração, recalculam-se as médias e reclassificam-se os pixels, considerando-se os novos valores médios. O processo é repetido até que a mudança entre iterações, ou seja, migração dos pixels de uma classe ou *cluster* para outra, tornar-se menor que o limiar pré-definido ou não ocorrer mais mudanças.

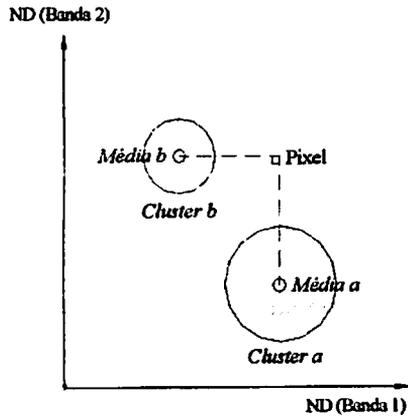


FIGURA 6 Cluster circular usando a medida de distância euclidiana. Fonte: Tso & Mather (2001).

O cálculo da distância entre um pixel e o centro do *cluster* normalmente usa a medida de distância euclidiana D_E representada por (Tso & Mather, 2001):

$$D_E^2 = (X_i - m_j)^2$$

em que X_i é o vetor observado no i -ésimo pixel e m_j é o vetor média do j -ésimo *cluster*. A dimensão do vetor X_i é igual ao número de bandas usadas na imagem.

O procedimento isodata pode ser refinado pela definição de desvio padrão máximo de um *cluster* (σ_{max}), e também pela distância mínima (d_{min}) entre *clusters*. Se um *cluster* tem desvio padrão maior que σ_{max} em alguma dimensão, o *cluster* é então dividido. Uma fusão ocorre onde a distância d_{ij} entre centro dos *clusters* i e j é menor que d_{min} . As operações de divisão e fusão permitem o refinamento do número de *clusters* (Tso & Mather, 2001). A seleção dos parâmetros σ_{max} e d_{min} não é uma tarefa fácil, pode ser feita através de

tentativas de valores para os dois parâmetros até que um resultado satisfatório seja alcançado. No entanto, a definição de “satisfatório” depende da percepção do analista e desta forma, o procedimento não é inteiramente automático e independente do usuário.

Uma alternativa para fusão de *clusters* pode ser realizado iterativamente usando a seguinte equação (Tso & Mather, 2001):

$$\min\{d_{ij}, \forall i, j, e i \neq j\}$$

d_{ij} é definido como:

$$d_{ij} = \frac{n_i \cdot n_j}{n_i + n_j} (m_i - m_j)^2$$

em que n_i e n_j são os números de pixels no *cluster* i e j , e m_i e m_j são vetores média dos *clusters* i e j , respectivamente. Assumindo que o número de *clusters* no começo da primeira iteração é k , em cada subsequente iteração, os dois *clusters* são unidos e o processo repete com $k - 1, k - 2, \dots$ *clusters*, até que o número mínimo de *cluster* préestabelecido seja alcançado. A solução que melhor satisfaz à necessidade do propósito específico é, então, selecionada.

Uma das principais vantagens da classificação não supervisionada, segundo Novo (1998), é não ser necessário que o analista tenha conhecimento prévio da área de estudo. Por outro lado, o analista tem pouco controle sobre a separação entre classes e a determinação do número final das classes (Novo, 1998; Sulsoft, 2004). O resultado é um conjunto de classes espectrais que devem ser rotulados pelo analista no final do processo em relação ao uso e cobertura do solo.

2.2.3.3 Árvore de decisão

Os métodos baseados em conhecimento representam um esforço para o desenvolvimento de métodos automáticos de reconhecimento de padrões que buscam simular os mecanismos de inferência do cérebro (Tso & Mather, 2001). Esses métodos são chamados de *data mining*, ou mineração de dados, cuja principal característica é a extração de informações previamente desconhecidas e de máxima abrangência a partir de grandes bases de dados, usando-as para efetuar decisões.

Para Coutinho (2003), a mineração de dados descende, em ordem cronológica, da estatística clássica, da inteligência artificial e do aprendizado de máquina. Por um lado, a estatística clássica e seus conceitos são o alicerce das mais avançadas análises estatísticas e, dessa forma, desempenha um papel fundamental nas técnicas de mineração de dados. A inteligência artificial é construída a partir dos fundamentos da heurística que, em oposição à estatística, tenta imitar a maneira como o homem pensa na resolução dos problemas estatísticos. A terceira e última linhagem da mineração de dados é o chamado aprendizado de máquina (*machine learning*), que pode ser melhor descrito como o casamento entre a estatística e a inteligência artificial.

O aprendizado de máquina tenta fazer com que os programas de computador “aprendam” com os dados que eles estudam, tal que esses programas tomem decisões diferentes baseadas nas características dos dados estudados, usando a estatística para os conceitos fundamentais e adicionando mais heurística avançada da inteligência artificial e algoritmos para alcançar os seus objetivos (Coutinho, 2003).

Dessa forma, pode-se descrever a mineração de dados como a união dos históricos e dos recentes desenvolvimentos em estatística, em inteligência

artificial e aprendizado de máquina. Essas técnicas são usadas juntas para estudar os dados e achar tendências e padrões nos mesmos.

Existem vários métodos de mineração de dados, como, por exemplo, redes neurais, indução de regras e árvores de decisão. Segundo Tso & Mather (2001) o interesse pelo uso de métodos de construção e uso de árvores de decisão tem crescido nos últimos anos. O algoritmo ID3 (e seus progressos C4.5, C5.0, See5) criado por Quilan, foi um dos primeiros para a extração de conhecimento por meio de árvore de decisão. A estratégia para a construção das árvores de decisão é “dividir para conquistar”, usando um procedimento recursivo com o objetivo de gerar uma pequena árvore.

Para Gonzalez & Woods (2000) uma árvore T é um conjunto finito de um ou mais nós para os quais existe um único nó chamado raiz, e os nós restantes são particionados em m conjuntos disjuntos T_1, \dots, T_m , cada qual sendo uma árvore chamada de subárvore de T .

A árvore de decisão utiliza mecanismos de categorização usando divisão hierárquica dos dados. O objetivo do uso de uma estrutura hierárquica para a classificação de objetos é aumentar o conhecimento da relação entre os objetos em diferentes escalas de observação ou diferentes níveis de detalhe. É simplesmente a representação da forma de uma árvore invertida com diferentes níveis de classificação representados por níveis ou hierarquia. Na aplicação em dados multiespectrais, o desenho da árvore de decisão é baseado no conhecimento das propriedades espectrais de cada classe e na relação entre as classes (Tso & Mather, 2001).

O classificador hierárquico árvore de decisão é caracterizado pelo fato de que um padrão desconhecido é rotulado usando-se uma seqüência de decisões. A árvore de decisão é composta de um nó raiz, um conjunto de nós interiores e nós terminais chamados folhas. O nó raiz e os nós interiores são conhecidos coletivamente como nós não-terminais e são representados por estágios de

decisão, enquanto o nó terminal representa a classificação final. O processo de classificação é implementado por um conjunto de regras que determina o caminho a ser seguido, iniciando no nó raiz e finalizando em um nó terminal que representa a classe temática da classificação. Em cada nó não-terminal, a decisão deve ser feita sobre o caminho para o próximo nó.

De acordo com Aurélio et al. (1999), inicialmente todos os dados que estão sendo minerados são associadas ao nó raiz da árvore. Então, o algoritmo seleciona uma partição de atributos e divide o conjunto de dados no nó raiz de acordo com o valor do atributo selecionado, gerando dois subconjuntos mais homogêneos que o anterior. O objetivo desse processo é separar os dados para que classes distintas tendam a ser associadas a diferentes partições. Esse processo é recursivamente aplicado a subconjuntos criados pelas partições, produzindo subconjuntos de dados cada vez menores, até que um critério de parada seja satisfeito. Para Carvalho (2001), teoricamente, o procedimento se repete até que subconjuntos “puros” sejam obtidos.

A Figura 7 ilustra a forma esquemática de uma árvore de decisão binária.

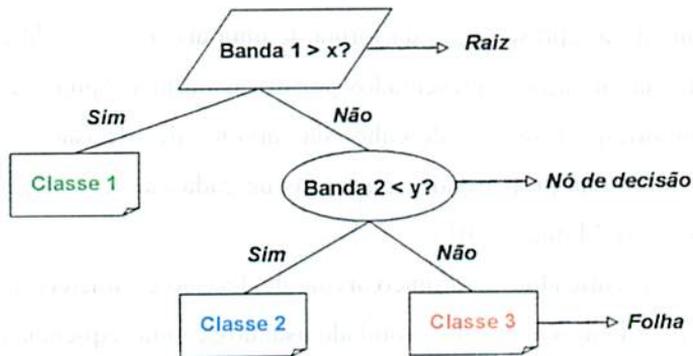


FIGURA 7 Representação de uma árvore de decisão binária

O nó raiz contém o conjunto de dados (pixels a serem classificados) que é dividido em subconjuntos mais homogêneos (nós de decisão) até os nós terminais ou folhas, de acordo com as regras de decisão. O caminho percorrido do nó raiz até cada folha consiste numa regra de decisão. As folhas da árvore resultam nas classes temáticas ou classes de uso e ocupação do solo.

A utilização de árvores de decisão binárias para classificação pode ser considerada uma abordagem não-paramétrica para reconhecimento de padrões (Quilan, 1986). Uma árvore de decisão faz uma representação hierárquica do espaço de feições, em que padrões x_i são alocados às classes w_j ($j=1, 2, \dots, k$), conforme o resultado encontrado depois de percorridos os ramos da árvore.

Uma árvore de decisão com alto poder preditivo e um pequeno número de nós constitui uma situação altamente desejável (Clarke & Bittencourt, 2003). A seleção da estrutura da árvore, de apropriados subconjuntos de feições em cada nó terminal e das regras de decisão, reflete na performance (acurácia da classificação) e eficiência (tempo computacional usado na classificação) do classificador (Swain & Hauska, 1977; Tso & Mather, 2001). As principais vantagens de algoritmos baseados em árvores de decisão são sua eficiência computacional e simplicidade (Aurélio et al., 1999).

2.3 Aplicação de técnicas de classificação no mapeamento do uso e ocupação do solo na Amazônia

Durante as últimas décadas, com a intensificação do desmatamento na Amazônia para a implantação de projetos agropecuários e minerais, aumentou a preocupação das instituições brasileiras e internacionais, pesquisadores, organismos não-governamentais e da sociedade em geral, com o ritmo acelerado desse processo. Várias experiências e políticas públicas têm sido aplicadas para reverter este cenário, entre elas, as que objetivam o monitoramento ambiental e

utilizam a aplicação de técnicas de classificação no mapeamento do uso e ocupação do solo.

Nesse contexto, o INPE desenvolve, desde 1988, o Projeto de Estimativa do Desflorestamento Bruto da Amazônia (PRODES), produzindo uma série histórica de estimativas anuais da extensão e taxas de desflorestamento para a Amazônia Legal, utilizando imagens do satélite Landsat na escala 1:250.000, que permite a identificação de desflorestamentos maiores que 6,25 hectares (Krug, 2001).

Até o ano de 2003, estas estimativas foram produzidas por interpretação visual de imagens. Após esse período, o INPE implantou o PRODES digital, cujo procedimento metodológico para identificar e mapear áreas desflorestadas, por meio de processamento digital de imagens Landsat, é feita pela segmentação por regiões e classificação não supervisionada, utilizando o algoritmo de agrupamentos de dados isoseg, das imagens fração sombra e solo oriundas da aplicação do modelo linear de mistura espectral (Duarte et al., 2003; INPE, 2004a). Os produtos gerados são um mapa contendo os incrementos anuais de desflorestamento e um segundo mapa contendo toda a extensão dos desflorestamentos acumulados dos anos anteriores.

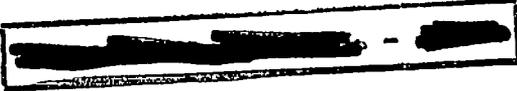
No Acre, a FUNTAC e o IMAC realizaram, em 1990 e 2001, respectivamente, estimativas para o monitoramento da cobertura florestal. No trabalho desenvolvido pela FUNTAC, utilizou-se a interpretação visual de imagens TM/Landsat, gerando mapas na escala de 1:100.000 (Funtac, 1990). Para mapear e quantificar o desmatamento no estado do Acre, em 1999, o IMAC por meio do Imazon, utilizou imagens de satélite ETM+/Landsat e escala da análise de 1:50.000. Essa escala, segundo Imazon (2001), possibilita maior precisão na identificação das áreas desmatadas, bem como fornece as informações necessárias para as ações de monitoração e controle ambiental. Na classificação usou-se o algoritmo de classificação não supervisionada isodata e,

em seguida, foi realizada a interpretação visual das imagens, corrigindo-se os erros de classificação.

Um convênio assinado entre o governo do Acre e o INPE, cujo objeto é o mapeamento das mudanças da cobertura vegetal do estado, relacionadas ao antropismo, para o período de 1988 a 2003, visa comparar a metodologia do Prodes Digital com a metodologia do Imazon. Os resultados deste acordo de cooperação técnico-científica não foram divulgados até o momento.

Além dessas experiências que objetivam o monitoramento em longo prazo, existem várias outras descritas na literatura para o mapeamento do uso e ocupação do solo na Amazônia. Roberts et al. (2003) listam uma série de pesquisas desenvolvidas no âmbito do Experimento de Grande Escala da Biosfera-Atmosfera na Amazônia (*Large Scale Biosphere-Atmosphere Experiment in Amazonia*, LBA) para o mapeamento e caracterização da cobertura da terra e mudanças de uso e ocupação na Amazônia. A escala espacial dos estudos variam de regional até áreas menores que 1.600 km², assim como os sensores utilizados, incluindo Landsat, Ikonos, Spot-4, SAR e Modis. Um tema comum em mais da metade dos doze artigos estudados pelo autor diz respeito à definição de cobertura da terra e mensuração da acurácia do mapeamento. As diversas técnicas utilizadas, segundo os autores, mostram-se promissoras e de fácil aplicação. Novas técnicas, como árvore de decisão e análise de mistura espectral, têm um potencial para mapear a cobertura da terra e quantificar as mudanças ocorridas em grandes áreas.

Vários autores aplicaram o modelo linear de mistura espectral e utilizaram as imagens fração na classificação digital para mapear o uso e ocupação do solo, realizar levantamento de áreas desflorestadas e caracterizar áreas florestais degradadas pela exploração madeireira na Amazônia (Shimabukuro et al., 1999; Graça, 2002; Espírito-Santo, 2003; Escada, 2003; Freitas et al., 2003; Lu et al., 2003; Mello et al., 2003). Nesses estudos, as



imagens fração foram segmentadas e, posteriormente, fez-se a classificação não supervisionada ou supervisionada ou, ainda, a combinação de ambas.

Lucas et al. (1998) realizaram o mapeamento da alteração de cobertura vegetal usando séries multitemporais do sensor Landsat para auxiliar no estudo de florestas em regeneração no norte de Manaus, AM, usando para a classificação dos tipos de cobertura vegetal um algoritmo de distância mínima. Comparando-se as classificações, foi possível gerar padrões de regeneração de florestas e a idade de todas as florestas em regeneração. Nobre et al. (1998) usaram o algoritmo de máxima verossimilhança para o mapeamento da vegetação na região de Manaus e obtiveram uma melhora significativa na definição de fronteiras entre áreas perturbadas e não perturbadas.

Silva (2003) aplicou a técnica análise do vetor de mudança com o objetivo de detectar, caracterizar e quantificar as mudanças ocorridas na cobertura vegetal e uso da terra, entre os anos de 2001 e 2002, numa área de floresta tropical, explorada pela atividade madeireira, localizada no norte do estado do Mato Grosso. Por meio da diferença entre as imagens frações multitemporais, originadas do modelo linear de mistura espectral para as duas datas, foram geradas a imagem magnitude do vetor de mudança e duas matrizes angulares, detectando a intensidade e a natureza da mudança ocorrida na cobertura vegetal e uso da terra.

Lorena et al. (2003) comparou a utilização da análise do vetor de mudança como técnica de detecção de transformações e a técnica tradicional de classificação por regiões Bhattacharyya, no estudo da dinâmica do uso e da cobertura da terra na região do Peixoto, estado do Acre, utilizando, para tanto, dados temporais e multiespectrais do TM/Landsat.

Espirito-Santo et al. (2003) utilizaram técnicas de processamento de imagens para a geração das imagens de índice de vegetação por diferença normalizada (NDVI) e imagens fração vegetação, solo e sombra do modelo

linear de mistura espectral, para detectar e caracterizar a mudança na cobertura vegetal da floresta de bambu numa área da porção sudeste do Acre, em um período de morte e crescimento vegetativo da floresta de bambu.

Clarke & Bittencourt (2003) aplicaram a técnica de CART na construção de árvore de decisão para classificar um segmento de imagem Landsat em três classes temáticas: água, cultura e vegetação. A árvore construída considerou apenas duas bandas espectrais, apesar dos dados de entrada contemplarem seis bandas espectrais, alcançando uma acurácia de 99%.

2.5 Avaliação da acurácia da classificação

Com a intensificação do uso de técnicas de sensoriamento remoto para a elaboração de mapas temáticos, aumentou a necessidade de estabelecer critérios mínimos de qualidade para os produtos gerados (Fidalgo, 1995). Segundo Lillesand & Kiefer (2000), historicamente, a habilidade de gerar produtos digitais oriundos da classificação de imagens de sensoriamento remoto excede a habilidade de quantificar significativamente sua acurácia. Dessa forma, Lillesand & Kiefer (2000) e Tso & Mather (2001) consideram que a classificação não está completa até a acurácia ser avaliada.

Os princípios e práticas geralmente usadas e aqui descritas para avaliar a acurácia da classificação baseiam-se nas propostas de Congalton & Green (1999).

Na classificação de imagens digitais, o termo acurácia informa o nível de concordância entre as classes especificadas pelo classificador e as classes de referência ou amostras de validação, baseadas nos dados de campo coletados pelo usuário. A ferramenta mais comum usada para determinar a acurácia da classificação é a matriz de erro, também chamada de matriz de confusão ou

tabela de contingência (Fidalgo, 1995; Lillesand & Kiefer, 2000; Tso & Mather, 2001).

A matriz de erro é uma matriz quadrada de dimensão $n \times n$, em que n é o número de classes. A matriz compara, categoria por categoria, a relação entre as classes de referência (amostras de validação) e os resultados correspondentes da classificação automática. As colunas da matriz representam os dados de referência, enquanto as linhas indicam os dados classificados pelo algoritmo de classificação. A diagonal principal indica a concordância entre estes dois grupos de dados. A exatidão global é calculada pela divisão da soma dos elementos da diagonal principal pelo número total de amostras tomadas, conforme a seguinte expressão:

$$P_o = \frac{\sum_{i=1}^m n_{ij}}{N}$$

em que:

P_o = exatidão global do mapeamento;

N = número total de unidades amostrais;

n_{ij} = total de elementos da diagonal principal, ou seja, número de pixels corretamente classificados;

m = número de classes ou categorias presentes na matriz de erro; e

i e j = linhas e colunas da matriz de erro, respectivamente.

Por meio da matriz de erro pode-se determinar a exatidão para cada categoria de uso e ocupação do solo mapeada. De acordo com Mascaro & Ferreira (2003), quando o número de amostras classificadas corretamente em uma dada categoria é dividido pelo número total das amostras de referência desta categoria, o resultado percentual indica a probabilidade da amostra de

referência ter sido classificada corretamente, estimando-se os erros de omissão e a exatidão do produtor. Quando o número de amostras classificadas corretamente em uma dada categoria é dividido pelo número total das amostras que foram classificadas nesta categoria, a porcentagem resultante indica a probabilidade da amostra realmente representar aquela categoria, estimando os erros de inclusão e a exatidão do usuário.

Segundo Moreira (2003), por meio da matriz de erro é possível utilizar técnicas de análise multivariada para determinar a concordância do mapeamento com a amostra de validação, sendo a estatística Kappa um dos métodos mais utilizados para essa avaliação. Gaboardi (2002) faz referência ao coeficiente de concordância Kappa como um dos mais utilizados quando se quer verificar se as matrizes de erro obtidas de classificações de uma mesma cena, por meio de diferentes métodos, diferem significativamente. Smits et al. (1999) consideram que este estimador é um dos mais adequados para se realizar comparações entre classificações, eliminando a subjetividade que é incluída durante o processo de avaliação.

Moreira (2003) cita que a grande vantagem da estatística Kappa é que no cálculo do coeficiente Kappa incluem-se todos os elementos da matriz de erro e não somente os elementos da diagonal principal, como no caso da exatidão geral.

Para a estimativa do coeficiente Kappa e da sua variância, empregam-se as equações:

$$K = \frac{P_o - P_c}{1 - P_c}, \quad P_o = \frac{\sum_{i=1}^M n_{ij}}{N},$$

$$P_c = \frac{\sum_{i=1}^M n_{i,i}}{N^2}, \quad e \quad \sigma_K^2 = \frac{P_o(1 - P_o)}{N(1 - P_c)^2}$$

sendo:

P_o = exatidão global;

P_c = proporção de unidades que concordam por casualidade;

M = número de classes presentes na matriz de erro;

n_{ij} = número de observações na linha i e coluna j ;

n_{i+} e n_{+i} = totais marginais da linha i e da coluna j , respectivamente; e

N = número total de unidades amostrais contempladas pela matriz.

O valor do coeficiente de Kappa (K) varia entre 0 e 1, em que o valor 1 representa uma completa concordância entre a imagem classificada com o dado de referência. O valor K pode ser comparado como os valores e conceitos contidos na Tabela 1, desenvolvidos por Landis & Koch (1977) e, dessa forma, avaliar o mapa temático obtido da classificação da imagem.

TABELA 1 Conceitos do coeficiente de concordância Kappa

<i>Valor de K</i>	<i>Qualidade da classificação</i>
<0,0	Péssima
0,0 - 0,2	Ruim
0,2 - 0,4	Razoável
0,4 - 0,6	Boa
0,6 - 0,8	Muito boa
0,8 - 1,0	Excelente

Fonte: Adaptada de Landis & Koch (1977).

3 MATERIAL E MÉTODO

3.1 Área de estudo

3.1.1 Localização geográfica

A área de estudo está localizada no sudeste do estado do Acre, situada em torno de $10^{\circ}11'52''$ de latitude sul e $68^{\circ}04'30''$ de longitude oeste (Figura 8) e corresponde ao município de Capixaba, com superfície territorial de 1.713 km², equivalendo a 1,1% da área total do estado (IBGE, 2004).

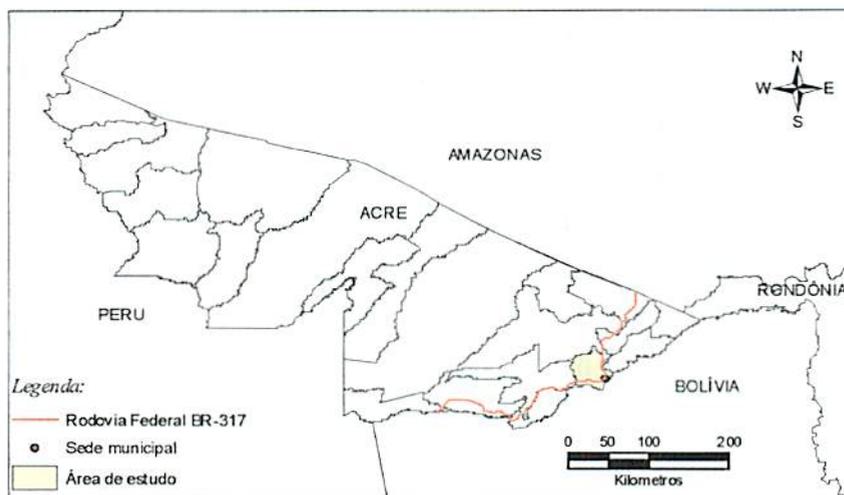


FIGURA 8 Localização do município de Capixaba, estado do Acre

O município de Capixaba está situado às margens da rodovia federal BR-317, com distância de 77 km de Rio Branco, capital do estado, oferecendo facilidade de acesso durante todo o ano e fazendo fronteira com a Bolívia e com os municípios acreanos de Rio Branco, Xapuri e Plácido de Castro.

3.1.2 Meio físico e biológico

O clima é do tipo equatorial quente e úmido, caracterizado por altas temperaturas, elevados índices de precipitação pluviométrica e alta umidade relativa do ar. Segundo classificação bioclimática (método de Bagnouls/Gausson), caracteriza-se por xeroquimênica subtermaxérica severa, com 1 a 3 meses de período seco e temperatura média do mês mais frio superior a 15°C. A precipitação anual varia em torno de 1.773 a 1.877 mm, sendo que no trimestre mais chuvoso (janeiro, fevereiro e março), gira em torno de 800 a 850 mm. A temperatura média anual é de 24,9°C a 25,1°C, com médias máximas de 34°C a 38°C nos meses de setembro, outubro e novembro (IMAC, 1991).

As características geomorfológicas do município são de feição colinosa da unidade morfoestrutural Depressão Amazônica, com relevo de topo pouco convexo separado por vales em V e, eventualmente, por vales de fundo plano, mapeado com índice de dissecação muito fraca (Acre, 2000a).

Os solos predominantes são argissolos com horizonte B textural e baixa atividade da argila, muito deles com alta saturação de alumínio, bem drenados e com relevo plano a suave ondulado (Acre, 2000a).

A rede hidrográfica é composta pelos rios Acre, Rapirã, Iquiri e Xipamanu (na divisa com a Bolívia), e vários igarapés permanentes, como Nova Amélia, Carupaty, Jarina e Mocó (Acre, 2000a).

A vegetação predominante é floresta ombrófila aberta de palmceiras com fasciações de floresta ombrófila densa (Acre, 2000a). A floresta ombrófila densa ou floresta pluvial tropical, segundo IBGE (1992), é caracterizada pela presença de fanerófitos, além de lianas lenhosas e epífitas em abundância que a diferencia das outras classes de formação. Sua característica ecológica principal reside nos ambientes ombrófitos atrelados às condições climáticas tropicais de elevada temperatura e alta precipitação bem distribuídas durante o ano. A

floresta ombrófila aberta apresenta fasciações florísticas que alteram a fisionomia ecológica da floresta ombrófila densa, imprimindo-lhes claros, advindo daí o nome adotado, além dos gradientes climáticos com mais de 60 dias secos por ano.

3.1.3 Aspectos sócio-econômicos e situação fundiária

Capixaba passou a município em 1992, por meio da Lei nº 1.027 de 28 de abril de 1992 (Acre, 2000b), desmembrado dos municípios vizinhos de Rio Branco e Xapuri. Segundo dados do censo demográfico de 2000 (IBGE, 2004), a população residente é de 5.206 habitantes, dos quais 70,8% vivem na zona rural e 29,2% na zona urbana.

A economia do município é baseada, principalmente, na pecuária extensiva e na produção extrativista. Segundo dados do IBGE (2004) para o ano de 2002, o efetivo do rebanho bovino totalizou 78.826 cabeças e a exploração de castanha-do-brasil, látex, açaí e carvão vegetal alcançou 588 toneladas. A exploração de madeira é também de grande representatividade na economia do município, possuindo em sua sede duas indústrias de desdobro e beneficiamento de madeira .

A produção agrícola, predominantemente familiar, tem como seus principais produtos arroz, feijão, milho, mandioca, banana, café e mamão. As lavouras localizam-se principalmente nos projetos de assentamento e pequenas propriedades rurais com a utilização de mão-de-obra familiar.

A estrutura fundiária do município envolve os projetos de assentamento (PA) São Gabriel e Alcoobrás, o projeto de assentamento agroextrativista (PAE) Remanso, parte da reserva extrativista (RESEX) Chico Mendes, 175 imóveis rurais com cadastro no INCRA, entre pequenas, médias e grandes propriedades, conforme Tabela 2. Além dessas, possui áreas discriminadas, porém em estudo

para criação de PAE e outras sem estudo de discriminação pelo INCRA. Os dados cadastrais do INCRA não correspondem ao universo das propriedades rurais do município (Acre, 2000b).

TABELA 2 Situação fundiária do município de Capixaba, 1999

Denominação	Quantidade	Área (ha)	% do município
Projetos de Assentamento do INCRA			
Projeto de Assentamento	2	20.200	11,8
Projeto de Assentamento Agroextrativista	1	39.570	23,2
Unidade de Conservação			
Reserva Extrativista (parte)*	1	5.802	3,4
Áreas discriminadas pelo INCRA e outras			
Propriedades com área até 100 ha	165	10.310	6,0
Propriedades com área de 100,1 até 500 ha	7	1.249	0,7
Propriedades com área de 500,1 até 1.000 ha	1	705	0,4
Propriedades com área de maior de 1.000 ha	2	20.100	11,7
Outras **	-	73.364	42,8
Total do município de Capixaba		171.300	100

* A Resex Chico Mendes possui área total de 976.570 ha distribuída entre os municípios de Assis Brasil, Brasiléia, Xapuri, Rio Branco, Sena Madureira e Capixaba.

**áreas sem estudo de discriminação e discriminadas pelo INCRA, área urbana e outras.

Fonte: Acre, 2000b.

3.2 Material

Para a realização deste estudo, as fonte de dados foram as imagens multiespectrais do satélite Landsat TM (*Thematic Mapper*) e ETM+ (*Enhanced*

Thematic Mapper Plus), bandas 1 (0,45 – 0,52 μm), 2 (0,53 – 0,61 μm), 3 (0,63 – 0,69 μm), 4 (0,76 – 0,90 μm), 5 (1,55 – 1,75 μm) e 7 (2,08 – 2,35 μm), dos anos de 2000, 2002 e 2003, utilizadas na classificação digital do uso e ocupação do solo e imagens do satélite sino-brasileiro de recursos terrestres Cbers-2 do ano de 2004, sensor CCD (*High resolution CCD camera*), bandas 2 (0,52 – 0,59 μm), 3 (0,63 – 0,69 μm), e 4 (0,77 – 0,89 μm), para auxiliar na determinação das amostras de amostra de validação. As informações das imagens Landsat e Cbers-2 estão descritas na Tabela 3.

TABELA 3 Satélite, cena, data de aquisição (dia/mês/ano) e resolução espacial das imagens usadas neste estudo

Satélite	Sensor	Órbita/ponto	Data de passagem	Resolução espacial
Landsat-5	TM	002/67	29/09/2000	30 x 30 m
Landsat-7	ETM+	002/67	10/08/2002	30 x 30 m
Landsat-5	TM	002/67	04/07/2003	30 x 30 m
Cbers-2	CCD	179/111	25/02/2004	20 x 20 m
Cbers-2	CCD	179/112	25/02/2004	20 x 20 m

As imagens Landsat são do período seco, em que se verifica pouca cobertura de nuvens. As imagens Cbers-2 utilizadas são do mesmo período de realização dos trabalhos de campo, que foi realizado em janeiro de 2004.

As imagens Landsat de 2000, 2001 e 2002 foram subdivididas em imagens de menor tamanho (1775 colunas x 1902 linhas), tomando-se por base os limites do município de Capixaba.

Para os cálculos de estatísticas das classes de mapeamento de uso e ocupação do solo e para a geração dos mapas resultantes das metodologias de

classificação digital aplicou-se uma máscara, construída com o vetor do município da área de estudo.

O processamento das imagens digitais e a mineração de dados para o classificador árvore de decisão foram realizados nos aplicativos disponíveis no Laboratório de Geoprocessamento e Análise de Imagens do Departamento de Ciências Florestais da Universidade Federal de Lavras.

No trabalho de campo foram usados os seguintes equipamentos: GPS (Sistema de Posicionamento Global), máquina fotográfica digital, notebook, trena, fichas de campo e veículo para deslocamento.

3.3 Metodologia

3.3.1 Classes temáticas para o mapeamento do uso e ocupação do solo

3.3.1.1 Caracterização das classes temáticas

Para a definição das classes de mapeamento do uso e ocupação do solo foram levados em consideração vários fatores: levantamento bibliográfico sobre o mapeamento do uso e ocupação do solo na Amazônia, atividades econômicas predominantes na área de estudo, dados e sazonalidade das imagens de satélite Landsat.

De acordo com Escada (2003), o comportamento espectral da floresta primária é caracterizado pela baixa reflectância na banda 3 (vermelho) e valores moderados na banda 4 (infravermelho próximo) devido à grande quantidade de sombras associadas a uma estrutura mais complexa do dossel. As florestas secundárias nos estágios iniciais têm um alto valor de reflectância na banda 4, decrescendo nos estágios mais avançados, devido ao aumento da complexidade do dossel. Na banda 5, a reflectância decresce gradualmente e, na banda 3, a

reflectância permanece baixa durante todo o período de sucessão. Em estágios avançados de sucessão, o comportamento espectral é parecido com o de florestas primárias. As pastagens na estação seca apresentam valores mais altos nas bandas 5 e 3, quando comparadas com as outras duas coberturas vegetais, podendo apresentar valores parecidos com o de vegetação secundária, no período chuvoso devido à presença de pastos altos, uniformes e bem manejados.

Dessa forma, foram definidas seis classes temáticas, cujas principais características são:

- **Floresta**

A região do município de Capixaba apresenta o domínio da floresta ombrófila aberta, com palmeiras, cipós ou bambus, dependendo das feições fitoecológicas de floresta densa ou de floresta aberta, em função das características das formas de relevo (Acre, 2000a). A floresta apresenta manchas com o dossel aberto e presença de espécies de palmeiras e espécies arbóreas citadas em anexo. Durante o trabalho de campo observou-se também grande presença de *Bertholletia excelsa* H.B.K. (castanheira). Distinguem-se três estratos: emergente, com dominância de árvores mais altas (35 a 40 metros de altura), com copas elevadas e que não formam um dossel aberto; intermediário, com predominância de breu vermelho e dossel mais fechado e inferior, com aspecto aberto ou limpo (Acre, 2000a).

Devido à proximidade de Rio Branco, principal centro consumidor do Acre, e à facilidade de acesso pela Rodovia BR-317 e ramais vicinais, a área florestal de Capixaba sofre as conseqüências da exploração predatória.

- **Capoeira**

A classe denominada capoeira é caracterizada pela ocorrência de vegetação em vários estágios de sucessão secundária, plantios florestais com seringueira (*Hevea brasiliensis* Muel. Arg.) ou mistos com várias espécies

florestais e sistemas agroflorestais com consórcio de culturas perenes e espécies florestais e frutíferas. O agrupamento das subclasses (estágios de sucessão secundária, plantios florestais e agroflorestais) ocorreu devido à heterogeneidade da estrutura da vegetação e à limitação da resolução espacial dos sensores remotos, como, por exemplo, TM e ETM+/Landsat que, segundo Lu et al. (2003), são as principais dificuldades para a difícil tarefa de classificar corretamente estágios de sucessão secundária.

De acordo com Lu et al. (2003), são três os estágios de sucessão que ocorrem na floresta tropical após o desflorestamento, cujas características principais são descritas a seguir.

No estágio sucessional inicial, ocorrem espécies pioneiras com alta demanda de luz, ciclo de vida curto, alta taxa de crescimento e reprodução. São espécies de plantas herbáceas, gramíneas, trepadeiras, mudas e varas. Após o segundo ou terceiro ano, algumas espécies arbóreas são importantes, assim como palmeiras comumente associadas a este período, incluindo *Vismia* sp e *Cecropia* sp. Porém, a densidade das varas é maior do que a das árvores, indicando sua importância nas características estruturais desta fase.

No estágio sucessional intermediário, a estrutura da vegetação resulta da mistura da densa cobertura vegetal das varas e árvores jovens e pouca diferença interna entre o dossel e o sub-bosque. Um dossel mais fechado altera o microclima, tornando-se um ambiente propício para o desenvolvimento das espécies arbóreas tolerantes à sombra e desfavorável para as espécies pioneiras.

No estágio sucessional avançado, as árvores ocupam o dossel e a estrutura da vegetação é estratificada, com aparência similar à da floresta primária. É nesta fase que ocorre a maior alteração na estrutura, diferenciando o dossel e sub-bosque em termos de altura e densidade. A presença das árvores é mais significativa do que a das varas, com a clara distinção de árvores dominantes. Nesse estágio é possível encontrar árvores com diâmetro à altura

do peito, altura e biomassa acima do solo similares às de árvores de florestas maduras.

- **Pasto alto**

Nesta classe foram agrupadas as áreas de pasto limpo com gramíneas exóticas de porte alto e textura mais homogênea; os pastos degradados ou pastos sujos, onde foram desenvolvidas atividades agropecuárias e após o seu abandono, o estabelecimento de vegetação natural, causando confusão com o estágio de sucessão secundária em sua fase inicial; e os plantios agrícolas de culturas anuais, predominantemente agricultura familiar de subsistência no projeto de assentamento e colonização Alcoobrás e São Gabriel, além de esparsas e pequenas áreas de cultivo agrícola às margens da BR-317 e no projeto de assentamento agroextrativista Remanso. Na imagem Landsat na composição colorida RGB 5,4,3, a classe de pasto alto assume tonalidades de verde-claro com a tendência de textura mais homogênea.

- **Pasto baixo**

Encontrado nas grandes fazendas de pecuária bovina, com atividade de gado de corte e de melhoramento genético do rebanho, localizadas principalmente às margens da rodovia federal BR-317 e ramais de acesso.

As gramíneas mais cultivadas são *Brachiaria brizantha*, *Brachiaria decumbens* e *Panicum maximum*. A pressão de pastejo maior que a capacidade de produção forrageira, principalmente na época da seca, que coincide com a data das imagens de satélite da área de estudo, confere o porte mais baixo ao pasto neste período. Dependendo do índice mensal de precipitação pluviométrica no período de julho a setembro, é identificado um percentual maior ou menor da área ocupada por esta classe de mapeamento, determinado pela influência da disponibilidade de água na capacidade de rebrota da pastagem.

Na imagem Landsat na composição colorida RGB 5,4,3, a classe de pasto baixo apresenta uma variação de tonalidades de rosa, às vezes com uma transparência para a cor verde em função da cobertura do solo por gramíneas.

- **Solo**

Representa a pequena área do perímetro urbano de Capixaba com população de 1.521 habitantes, segundo dados do censo 2000 (IBGE, 2004), composta de áreas construídas de uso residencial, comercial, industrial e público. Contempla ainda as edificações nas propriedades rurais, a antiga usina de produção de álcool no projeto de assentamento Alcoobrás, as estradas e ramais com ou sem pavimentação asfáltica.

- **Queimada**

Caracteriza-se pelas áreas que foram recentemente queimadas, seja para limpeza de pastos sujos, áreas de cultivo agrícola ou de áreas desmatadas de floresta ou capoeira. A prática da queima é comum na região, acontecendo nos meses mais secos, geralmente no período de julho a setembro. Essa classe temática é facilmente identificada, nas imagens Landsat, na composição colorida das bandas 5, 4 e 3, apresentando tonalidade magenta escura a púrpura.

- **Água**

Classe que representa os corpos d'água da área de estudo, composta pelo leito dos rios Acre, Rapirrã, Iquiri e Xipamanu, vários igarapés permanentes e temporários como Nova Amélia, Carupaty, Jarina e Mocó, lagos e açudes das propriedades rurais. Os rios e igarapés com leito muito estreitos ou sob o dossel da floresta não foram representados no mapeamento final, em consequência da resolução espacial das imagens, implicando em pixels com mistura espectral.

As classes de mapeamento descritas anteriormente estão representadas na Figura 9.



FIGURA 9 Ilustração das classes de mapeamento de uso e ocupação do solo: (a) floresta, (b) e (c) capoeira em diferentes estágios de sucessão secundária, (d) pasto alto, (e) pasto baixo, (f) solo – área urbana, (g) água – rio Acre e (h) queimada

3.3.1.2 Amostras das classes temáticas

As amostras utilizadas na classificação digital e verificação da exatidão do mapeamento constituem-se de conjunto de pixels, divididos em três grupos: treinamento, teste e validação.

As amostras de treinamento foram utilizadas na classificação supervisionada por máxima verossimilhança e para descrever os casos de treinamento que compõem o arquivo de dados para a mineração de dados por árvore de decisão.

As amostras de teste foram aplicadas para descrever os casos que foram utilizados na avaliação do erro da árvore de decisão gerada pela mineração de dados.

Para avaliar a exatidão do mapeamento que resultou da classificação digital por meio de diferentes métodos, utilizou-se o conjunto de amostras de validação ou dados de referência.

A preparação das amostras de treinamento e teste foi realizada em fases: realização de trabalho de campo e análise visual das imagens Landsat e Cbers-2. Primeiramente, foram coletados 290 pontos com GPS, distribuídos entre as classes temáticas definidas anteriormente.

Para isso, foram percorridos as estradas e ramais com condições de acesso e trafegabilidade no município de Capixaba, conforme a rota apresentada na Figura 10. Os pontos coletados com GPS foram alocados nas imagens de satélite, em subgrupos correspondentes às classes temáticas e, a partir da análise visual, foram selecionadas as coleções de amostras. Nos casos em que não foi possível determinar a localização do polígono da classe temática referente aos pontos de amostra de validação, devido à resolução espacial da imagem Landsat, utilizaram-se as imagens Cbers-2 com melhor resolução espacial para auxiliar na seleção das amostras.



FIGURA 10 Imagem Landsat de 2003 na composição colorida das bandas 5, 4 e 3, com a rota percorrida para coleta de amostras das classes temática no município de Capixaba, AC

As amostras de validação foram selecionadas numa imagem de referência estratificada por classe temática. Nessa imagem foram selecionados aleatoriamente os pixels por classe de mapeamento para compor o conjunto de amostras a serem usadas como dados de referência ou amostras de validação com base nos levantamentos realizados em campo e análise visual das imagens. O mesmo conjunto de amostras de validação foi utilizado na geração da matriz de erro para todas as técnicas de classificação digital.

Posteriormente, foi mensurada a separação estatística entre as categorias de padrões espectrais, computadas para todos os pares de classes temáticas. Na avaliação da separabilidade espectral entre as classes temáticas foi utilizada a medida de distância Jeffries-Matusita, pois existe uma relação inversa entre a separabilidade estatística e a probabilidade de erro. Em geral, quanto maior a

distância estatística entre as amostras de treinamento, maior a probabilidade de classificação correta das classes (Lillesand & Kiefer, 2000).

Os valores da medida de distância Jeffries-Matusita variam entre 0 a 2, indicando como os pares de amostras das classes temáticas selecionados são separados estatisticamente. Os valores maiores que 1,9 indicam que os pares de classes têm boa separabilidade e valores menores indicam similaridade (Sulsoft, 2004).

Essa medida de separabilidade estatística entre as classes é apropriada para situações que envolvam múltiplas classes e se mostra um bom modo de prever a precisão de classificação (Aguiar,1991; Lillesand & Kiefer, 2000; Tso & Mather, 2001).

A medida de distância Jeffries-Matusita (distância J-M) sob hipótese gaussiana é dada por (Tso & Mather, 2001):

$$JM_{ij} = \sqrt{2(1 - e^{-\alpha})}$$

em que:

$$\alpha = \frac{1}{8}(\mathbf{u}_i - \mathbf{u}_j)^T \left(\frac{\mathbf{C}_i + \mathbf{C}_j}{2} \right)^{-1} (\mathbf{u}_i - \mathbf{u}_j) + \frac{1}{2} \ln \left(\frac{\mathbf{C}_i + \mathbf{C}_j}{2 \sqrt{|\mathbf{C}_i| |\mathbf{C}_j|}} \right)$$

sendo:

i e j: classes que estão sendo comparadas;

\mathbf{u}_i : vetor média da classe *i*;

\mathbf{u}_j : vetor média da classe *j*;

T : matriz transposta;

\mathbf{C}_i : matriz de covariância da classe *i*;

\mathbf{C}_j : matriz de covariância da classe *j*; e

ln: logaritmo neperiano.

3.3.2 Técnicas de pré-processamento das imagens

3.3.2.1 Correção geométrica

As imagens produzidas por sensores remotos apresentam distorções espaciais, não possuindo, portanto, precisão cartográfica quanto ao posicionamento dos objetos, superfícies ou fenômenos nelas representados (Sulsoft, 2004). A correção geométrica consiste na transformação das imagens originais segundo propriedades de escala e de projeção de uma mapa, visando eliminar os erros que ocorrem devido ao movimento do satélite e determinados pela curvatura da terra (Moreira, 2003) e reorganizar os pixels da imagem em relação a determinado sistema de projeção cartográfica (Novo, 1998).

As etapas de processamento da correção geométrica são a transformação de coordenadas em pixels e a reamostragem para determinar o valor de pixel na imagem a ser registrada.

Inicialmente, realizou-se a correção geométrica das imagens de 2000, 2002 e 2003, utilizando como referência para o registro a imagem de 2003, que foi georreferenciada por meio de 17 pontos de controle no terreno coletados com GPS e alocados na imagem de 2003, por suas coordenadas na projeção cartográfica *Universal Transverse Mercator* ou UTM, datum SAD 69/Brasil. O método de transformação usado foi polinomial de 1º grau e o de reamostragem dos pixels a técnica de alocação do vizinho mais próximo (*nearest neighbour*). Por meio dessa técnica, o valor do nível de cinza a ser atribuído ao pixel z terá o mesmo valor do nível de cinza do pixel que se encontrar mais próximo da posição ocupada por z . As vantagens são processamento rápido, fácil implementação e não haver alteração dos valores originais de cinza (Novo, 1998).

O registro das imagens dos anos de 2000 e 2002 foi realizado com base na imagem de 2003, corrigida geometricamente. Neste caso, os pontos de controle na imagem de referência de 2003 foram alocados em pontos parecidos, distribuídos nas extremidades esquerda e direita, inferior e superior, e no centro das imagens a serem registradas. Também se utilizou a transformação polinomial de 1º grau e a técnica do vizinho mais próximo para a reamostragem dos pixels.

Para finalizar o registro das imagens, calculou-se o erro médio quadrático (RMS), que representa uma medida de ajuste do modelo de transformação. O erro RMS é estimado pela diferença entre amostras dos valores calculados e seus valores originais, elevadas ao quadrado e somadas. A soma é dividida pelo número de medidas, obtendo-se a média, cuja raiz quadrada fornece uma medida de erro na mesma unidade das medidas originais (Sulsoft, 2004). A expressão do erro RMS é dada por:

$$\text{RMS}_{\text{error}} = \left[\left(x_1 - x_{\text{orig}} \right)^2 + \left(y_1 - y_{\text{orig}} \right)^2 / n \right]^{1/2}$$

em que:

x_1 : coordenada de linha calculada na imagem não-registrada;

y_1 : coordenada de coluna calculada na imagem não-registrada;

x_{orig} : coordenada original de linha do ponto de controle na imagem original; e

y_{orig} : coordenada original de coluna do ponto de controle na imagem original.

3.3.2.2 Calibração radiométrica

A calibração consiste na transformação dos valores digitais brutos (número digital – DN) das imagens Landsat para reflectância, com o objetivo de atenuar os efeitos da interferência atmosférica e características do sensor (Meer, 1999).

Os números digitais, embora relacionados, não representam quantitativamente valores físicos reais (Aguiar, 1991; Pereira et al., 1996) e são usados apenas por conveniência no processamento dos dados.

Os números digitais são relacionados, por um modelo linear (calibração), à intensidade da energia radiante refletida. Contudo, não são diretamente comparáveis entre satélites diferentes devido às diferenças na calibração dos sensores, os quais determinam as diferentes relações entre as radiâncias e os números digitais, dependendo da banda analisada, da localização da área e da data de passagem (Aguiar, 1991; Pereira et al., 1996).

A utilização de valores físicos, sobretudo da reflectância, de acordo com Pereira et al. (1996), possibilita um melhor monitoramento do comportamento de uma determinada cobertura, uma vez que mudanças na sua resposta espectral estariam relacionadas somente às mudanças nas propriedades estruturais/espectrais da cobertura e não às mudanças no ângulo de iluminação solar ou variações dos sensores.

A conversão dos números digitais para reflectância gera uma base mais normalizada em cenas simples e entre imagens tomadas de diferentes sensores ou datas. A reflectância foi calculada utilizando-se as relações propostas por Markham & Barker (1987), em que inicialmente, convertem-se os números digitais em valores de radiância e depois para valores de reflectância aparente, usando as seguintes equações:

$$L(\lambda) = L_{\min \lambda} + \frac{(L_{\max \lambda} - L_{\min \lambda})}{QCAL_{\max \lambda}} QCAL$$

em que:

$L(\lambda)$ = radiância espectral, em $\text{mW cm}^{-2} \text{sr}^{-1} \mu\text{m}^{-1}$;

$L_{\min \lambda}$ = radiância espectral equivalente ao número digital mínimo;

$L_{\max \lambda}$ = radiância espectral equivalente ao número digital máximo;

$QCAL_{\max \lambda}$ = amplitude de variação dos números digitais, equivalente a 255 para todos os dados TM e ETM+; e

$QCAL$ = número digital.

Os valores de radiância obtidos nesta equação são convertidos em reflectância aparente pela seguinte equação:

$$\rho_p = \frac{\pi L_\lambda d^2}{E_{sol\lambda} \cos \theta_s}$$

em que:

ρ_p = reflectância aparente, adimensional;

L_λ = radiância espectral ($\text{mW cm}^{-2} \text{sr}^{-1} \mu\text{m}^{-1}$);

d = distância Terra-Sol em unidades astronômicas;

$E_{sol\lambda}$ = irradiância solar exoatmosférica média ($\text{mW cm}^{-2} \mu\text{m}^{-1}$); e

θ_s = ângulo zenital solar, em graus.

As novas imagens de reflectância aparente não constituem os dados de reflectância real dos alvos, pois, no cálculo das novas imagens, foram utilizados valores da irradiância ($E_{sol\lambda}$) no topo da atmosfera (Moreira, 2003).

Os valores de $L_{\min\lambda}$, $L_{\max\lambda}$ e $E_{sol\lambda}$ utilizados na calibração radiométrica do Landsat 5 e Landsat 7 estão relacionados em anexo. Os valores de θ_s e d foram obtidos a partir dos dados dos arquivos, em formato texto, que acompanham as imagens (INPE, 2000; INPE, 2002; INPE, 2003).

3.3.3 Técnicas de processamento das imagens

3.3.3.1 Extração de informações das imagens

Consiste em etapas preliminares à aplicação dos algoritmos de classificação digital com o propósito de extração de informações que serão utilizadas na fase de classificação digital. Estão incluídos a geração de imagens fração ou sintéticas oriundas da análise de mistura espectral e do índice de vegetação por diferença normalizada (NDVI).

3.3.3.1.1 Análise de mistura espectral

A análise de mistura espectral foi realizada nas imagens Landsat TM e ETM+, bandas 1, 2, 3, 4, 5 e 7, convertidas de número digital para reflectância. Os componentes puros selecionados foram solo, sombra e vegetação, amplamente utilizados no mapeamento do uso e ocupação do solo na Amazônia (Shimabukuro et al., 1999; Graça, 2002; Espírito-Santo, 2003; Duarte et al., 2003; Escada, 2003; Freitas et al., 2003; Lu et al., 2003; Mello et al., 2003; Silva, 2003).

Para a seleção das amostras dos componentes puros foi utilizada a teoria do simplex (Aguiar, 1991; Bajjouk et al., 1998), em que os componentes puros localizam-se nos vértices do triângulo formado no espaço de feições bidimensional, representando a distribuição dos pixels nas bandas 4 e 5,

correspondente ao comprimento de onda do infravermelho próximo e infravermelho distante. O simplex de sinais pode ser entendido como a figura geométrica formada quando são ligados os pontos determinados pelos vetores média dos componentes do pixel (Aguiar, 1991).

Na Figura 11 observa-se a distribuição das amostras dos pixels puros dos componentes sombra, vegetação e solo no espaço de feições coletados na imagem Landsat do ano de 2003.

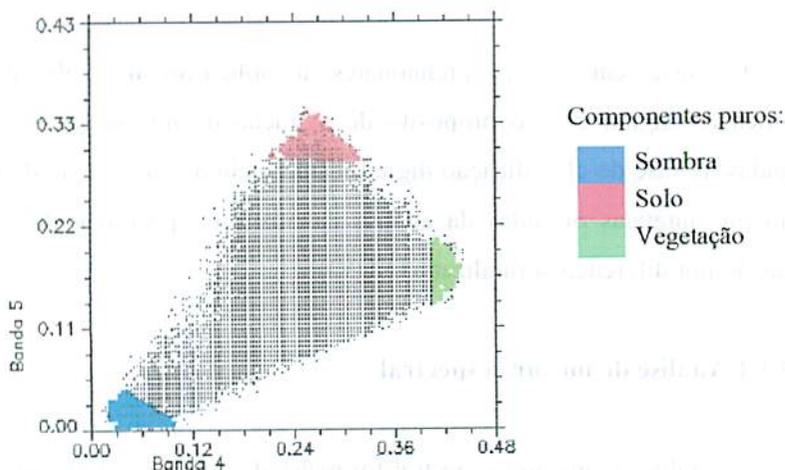


FIGURA 11 Dispersograma com a localização das amostras dos pixels puros no espaço bidimensional das bandas 4 e 5 da imagem TM/Landsat do município de Capixaba, AC, 2003, com valores em reflectância aparente

O componente sombra foi amostrado em áreas do rio Acre, de igarapés, de lagos e de açudes existentes no município. As amostras do componente vegetação foram obtidas de áreas de pastagem alta e capoeira em variados estágios de sucessão secundária. Para o componente solo, a seleção das amostras aconteceu em áreas de solo exposto com pouca ou nenhuma cobertura vegetal, assim como em ramais de acesso sem pavimentação asfáltica.

Depois de selecionados os componentes puros, a análise espectral foi realizada pelo método de mapeamento, utilizando-se o algoritmo de decomposição espectral linear, em que a reflectância de cada pixel da imagem é considerada uma combinação linear da reflectância de cada material ou componente puro presente dentro do pixel. A decomposição linear é resolvida por valores de abundância de cada componente puro para todo pixel.

3.3.3.1.2 Índice de vegetação

O índice de vegetação por diferença normalizada (NDVI) foi usado para transformar os dados multiespectrais do Landsat em uma imagem com única banda representando a distribuição da vegetação. Os valores do NDVI indicam a quantidade de vegetação verde presente no pixel, ou seja, quanto maior o valor do NDVI, mais vegetação verde possui o pixel. O NDVI foi gerado por meio do seguinte algoritmo padrão:

$$\text{NDVI} = \frac{(\text{NIR} - \text{R})}{(\text{NIR} + \text{R})}$$

em que:

NIR: valor do pixel na faixa do espectro eletromagnético correspondente à região do infra-vermelho próximo (0,7 a 0,9 μm) e

R: valor do pixel na faixa do espectro eletromagnético correspondente a região do vermelho (0,6 a 0,7 μm).

3.3.3.2 Classificação digital

3.3.3.2.1 Máxima verossimilhança

As amostras de treinamento das classes temáticas floresta, capoeira, pasto alto, pasto baixo, solo, água e queimada foram selecionadas conforme descrito no item 3.3.1.2. As amostras são proporcionais ao tamanho das feições presentes nas imagens Landsat, ou seja, as feições correspondentes às classes temáticas de maior tamanho receberam maior número de amostras. Esse procedimento foi adotado com o objetivo de representar bem as feições, manter a homogeneidade de cada classe temática e alcançar o índice de Jeffries-Matusita entre a combinação dos pares de classe que indica boa separabilidade entre elas.

A quantidade de amostras por classe temática utilizadas nos três anos objeto de análise está descrita na Tabela 4.

TABELA 4 Tamanho das amostras de treinamento utilizadas na classificação por máxima verossimilhança

Classe temática	Número de pixels da amostra de treinamento/ano		
	2003	2002	2000
Floresta	6.520	12.600	9.000
Capoeira	1.645	1.792	2.130
Pasto alto	3.226	2.220	1.856
Pasto baixo	3.625	2.253	2.125
Solo	589	1.600	1.400
Água	1.200	1.321	1.300
Queimada	-	1.200	1.127
Total	16.805	22.986	18.938

No ano de 2003, não foi identificada na imagem a ocorrência de áreas queimadas e, portanto, não houve a seleção de amostras de treinamento.

Na aplicação do algoritmo de máxima verossimilhança não foi definido nenhum limite de probabilidade, para que todos os pixels da imagem fossem associados a uma das classes de mapeamento previamente definidas.

3.3.3.2.2 Isodata

A classificação não supervisionada isodata calcula classes médias eventualmente distribuídas no espaço multidimensional definido pelas n bandas espectrais e, iterativamente, forma agrupamentos de pixels remanescentes usando a técnica de distância mínima.

Para isso, utilizou-se o mínimo de 5 e o máximo de 15 classes para o cálculo do algoritmo de agrupamento, com até 10 iterações, procedimento pelo qual são recalculadas as médias e os pixels inicialmente agrupados são reclassificados considerando-se os novos valores médios. A variação do número de classes é usada porque o algoritmo isodata separa e une classes baseado nos parâmetros estabelecidos e não em um número fixo de classes.

Os parâmetros estabelecidos foram também os seguintes: o mínimo de 10 pixels para formar uma classe; mínimo de 1 classe de desvio padrão; máximo de 5 classes de distância e máximo de 2 pares de classes a serem unidas. Com base nestes limites as classes são iterativamente divididas, unidas ou excluídas. Quando o número de pixels na classe é menor que 10, a classe é excluída e os pixels são agrupados à classe mais próxima. Se o desvio padrão da classe é maior que o limite estabelecido, então, esta é dividida em duas classes. Se a distância entre as classes médias é menor que 5, as classes são unidas, considerando-se o número máximo de pares de classes determinado.

Todos os pixels são classificados para a classe próxima, considerando-se os parâmetros especificados e a classificação é considerada pronta quando menos de 5% dos pixels mudou de classe ou quando o número máximo de iterações foi concluído.

Por último, as classes espectrais geradas pelo algoritmo isodata foram agrupadas com a mesma legenda das classes temáticas definidas para o mapeamento do uso e ocupação do solo.

3.3.3.2.3 Classificação supervisionada e não supervisionada usando imagens fração oriundas do modelo linear de mistura espectral

Foram testadas três modalidades de classificação das imagens fração geradas pelo algoritmo de decomposição espectral linear, entre as técnicas supervisionada e não supervisionada, aplicando o algoritmo isodata e de máxima verossimilhança, respectivamente. As modalidades de classificação referem-se as imagens fração usadas para o mapeamento das classes temáticas e o algoritmo de classificação. Foram utilizadas: 1) apenas as imagens fração solo e sombra e o algoritmo isodata. A aplicação dessas imagens fração são as mesmas utilizadas pelo PRODES digital para estimar o índice de desmatamento para a Amazônia Legal (Duarte et al., 2003); 2) as imagens fração solo, sombra e vegetação e, o algoritmo isodata. Nesse caso, foi incluída a imagem fração vegetação com o propósito de verificar seu desempenho no mapeamento das classes de vegetação secundária e pasto; e, 3) as imagens fração solo, sombra, vegetação e de erro, com aplicação do algoritmo de máxima verossimilhança. Essas três modalidades foram designadas por classificação não supervisionada usando as imagens fração solo e sombra; classificação não supervisionada usando as imagens fração solo, sombra e vegetação; classificação supervisionada usando imagens fração, respectivamente, cujos passos metodológicos são descritos a seguir.

- **Classificação não supervisionada usando as imagens fração solo e sombra**

As etapas realizadas para a classificação das imagens fração solo e sombra oriundas da aplicação do modelo linear de mistura espectral foram feitas segundo o esquema representado na Figura 12.

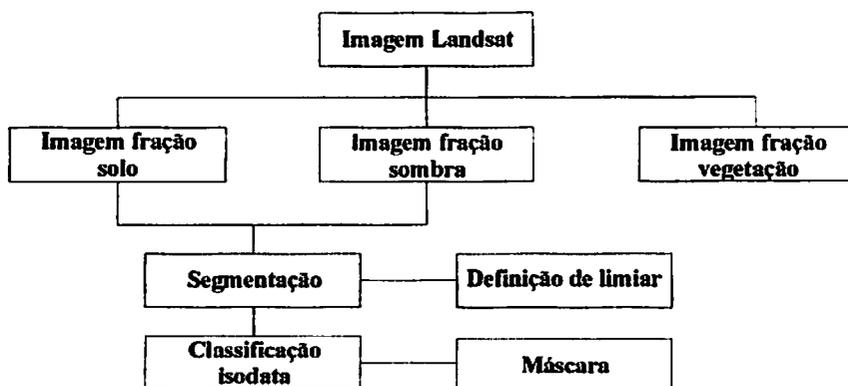


FIGURA 12 Esquema da classificação não supervisionada isodata usando as imagens fração solo e sombra

Inicialmente, por meio da segmentação pixel a pixel foram extraídas das imagens fração solo e sombra as feições mais realçadas em cada uma delas, geradas pela aplicação do modelo linear de mistura espectral.

A segmentação subdivide uma imagem em partes ou objetos de interesse e pode ser orientada a pixel ou a regiões. Nesse contexto, a limiarização é uma das abordagens mais importantes e consiste em extrair objetos através da seleção de um limiar T que os separe (Gonzalez & Woods, 2000). O limiar T pode ser entendido como um intervalo na escala de níveis de cinza do histograma da imagem a ser particionado. A segmentação é então efetuada, varrendo-se a imagem, pixel por pixel, e rotulando-se cada pixel como sendo ou não do objeto, dependendo se o nível de cinza for maior ou menor que o limiar T .

A definição do limiar nas imagens fração solo e sombra foi realizada por meio da verificação das estatísticas dos pixels encontrados nas feições mais realçadas em cada imagem. No caso da imagem fração solo, essas feições correspondem às áreas de pastagem, capoeira e solo. A imagem fração sombra foi segmentada utilizando-se um limiar que separou os rios, igarapés, lagos e açudes. Para isso, foi considerado, como população mínima de pixels a ser segmentada, o conjunto de 4 pixels e, para a análise do pixel, seus vizinhos de 8.

Após a limiarização, foi realizada a classificação não supervisionada isodata nas imagens fração solo e sombra com os seguintes parâmetros: mínimo de 5 e máximo de 10 classes, máximo de 5 iterações, mínimo de 10 pixels para formar uma classe, mínimo de 1 classe de desvio padrão, máximo de 5 classes de distância e máximo de 2 pares de classes a serem unidas.

Para agilizar o processamento digital e preservar as feições realçadas nas imagens geradas pelo modelo linear de mistura espectral, foram criadas máscaras com as classes temáticas obtidas do resultado da classificação não supervisionada isodata em cada imagem fração (solo e sombra). As classes definidas pela imagem fração sombra foram mascaradas na imagem fração solo e vice-versa.

Após a realização da classificação não supervisionada isodata, resultaram as seguintes classes de uso e ocupação do solo:

- floresta, capoeira e água, extraídas da imagem fração sombra e
- pasto alto, pasto baixo, solo exposto e capoeira, adquiridas da imagem fração solo.

Para gerar uma única imagem classificada e calcular as estatísticas e acurácia do mapeamento, os resultados das classificações das imagens fração solo e sombra foram combinados, obtendo-se uma imagem final da classificação com todas as classes identificadas anteriormente.

- **Classificação não supervisionada usando as imagens fração solo, sombra e vegetação**

A segunda modalidade de classificação usou as imagens fração solo, vegetação e sombra, conforme Figura 13, seguindo os mesmos passos descritos anteriormente para construção das máscaras e classificação da imagem fração sombra para determinar as classes de floresta, capoeira e água.

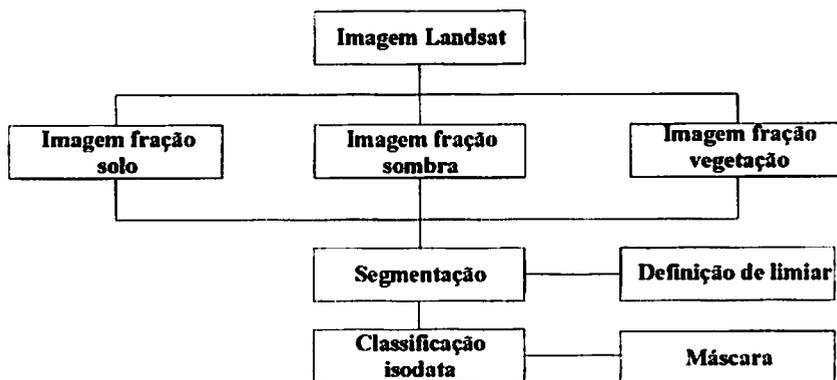


FIGURA 13 Esquema da classificação não supervisionada isodata usando as imagens fração solo, sombra e vegetação

A classe solo foi obtida pela segmentação da imagem fração solo, por meio da definição do limiar que separou as áreas de solo exposto, rodovias, ramais e edificações e classificação não supervisionada isodata.

Posteriormente, as classes de floresta, água, capoeira e solo foram utilizadas para a criação da máscara que foi aplicada na imagem fração vegetação. A classificação isodata da imagem fração vegetação, com os mesmos parâmetros descritos anteriormente, resultou na combinação das classes: pasto alto, pasto baixo, queimada e capoeira.

A classificação final foi composta pelas seguintes classes:

- floresta, capoeira e água, oriundas da imagem fração sombra;
- solo, adquirida da imagem fração solo e

- pasto alto, pasto baixo, queimada e capoeira, oriundas da imagem fração vegetação.
- **Classificação supervisionada usando as imagens fração solo, sombra, vegetação e de erro**

A terceira modalidade de classificação com as imagens geradas pela análise de mistura espectral foi realizada utilizando-se as imagens fração solo, sombra, vegetação e de erro, e a técnica de classificação supervisionada por máxima verossimilhança (Figura 14).

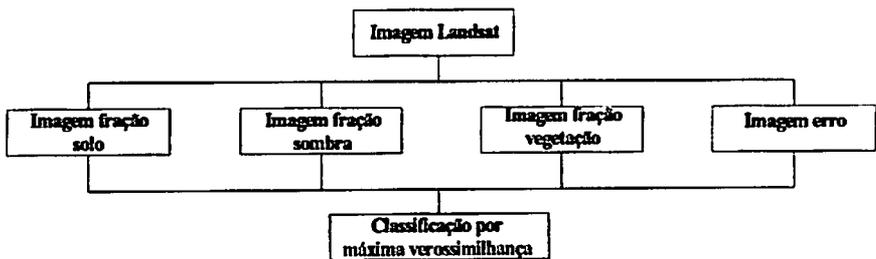


FIGURA 14 Esquema da classificação supervisionada por máxima verossimilhança usando as imagens fração solo, sombra, vegetação e de erro

Na fase de treinamento do classificador foram utilizadas as mesmas amostras de treinamento citadas no item 3.3.3.2.1, com o intuito de comparar a performance das metodologias de classificação digital.

Na aplicação do algoritmo de máxima verossimilhança não foi definido nenhum limite de probabilidade, para que todos os pixels da imagem fossem associados a uma das classes de mapeamento.

3.3.3.2.4 Árvore de decisão

Para a classificação pelo algoritmo de árvore de decisão foi utilizado um aplicativo de mineração de dados, na geração do conjunto de regras da árvore de decisão e, posteriormente, um software de processamento de imagens para a classificação digital.

Os atributos usados nesta técnica de classificação foram: bandas 1, 2, 3, 4, 5 e 7 do Landsat; imagens fração solo, sombra, vegetação e de erro geradas pela decomposição linear espectral e índice de vegetação NDVI. Essas onze variáveis foram organizadas em um único arquivo de imagem.

Foram utilizadas amostras de treinamento e amostras de teste na mineração de dados. O número de casos das amostras usadas para a classificação por árvore de decisão estão citados na Tabela 5. Cada pixel das amostras e seus respectivos valores nas onze imagens do arquivo de dados constituem um caso a ser analisado pelo algoritmo de aprendizado de máquina.

TABELA 5 Número de casos das amostras de treinamento e de teste utilizados na classificação por árvore de decisão

Classe temática	Número de pixels da amostra/ano					
	2003		2002		2000	
	Treinamento	Teste	Treinamento	Teste	Treinamento	Teste
Floresta	3.260	3.260	6.000	6.600	4.500	4.500
Capoeira	823	822	1.100	692	1.100	1.030
Pasto alto	1.613	1.613	1.110	1.110	928	928
Pasto baixo	1.813	1.812	900	873	1.100	1.025
Solo	295	294	800	800	700	700
Água	600	600	808	513	650	650
Queimada	-	-	600	600	600	527
Total	8.404	8.401	11.318	11.188	9.578	9.360

Segundo Curroto (2000), não existe nenhum limite prático no número de classes, atributos, valores de atributos ou casos nos arquivos de dados e de teste.

Na preparação dos dados necessários para a implementação da mineração de dados, foram organizados três arquivos, denominados arquivo de nomes, arquivo de dados e arquivo de teste.

O arquivo de nomes contém o nome das classes temáticas, dos atributos e valores dos atributos. Os atributos foram cada uma das onze imagens usadas e os valores dos atributos denominados *continuous*, que indica que os atributos têm valores numéricos, inteiros ou reais.

O arquivo de dados é utilizado para descrever os casos de treinamento. Cada pixel da amostra foi especificado numa linha descrevendo um caso, contendo os valores para todos os atributos (imagens usadas) e a classe do caso. Os valores dos atributos de cada pixel da amostra foram ordenados conforme definido no arquivo de nomes.

O arquivo de teste é utilizado para descrever os casos de teste que foram utilizados para avaliar a árvore de decisão produzida pela mineração de dados. O formato é o mesmo do arquivo de dados, porém usando os pixels da amostra de teste.

O aplicativo de mineração de dados gerou a árvore de decisão final podada. Foi utilizada a opção avançada de poda, pela qual o processo de geração do classificador pode ser refinado (Rulequest Research, 2004). A poda faz a remoção de partes que podem apresentar relativamente alta taxa de erros. O processo da poda é primeiramente aplicado em cada subárvore ou galho para decidir se deve ser substituída pela folha ou galho menor. A poda é efetuada se a substituição do galho pela folha ou galho menor conduzir a uma taxa de erro menor, desde que a taxa de erro global da árvore decresça. Esse processo leva a uma árvore de taxa de erro mínima (Curroto, 2000).

3.3.3.4 Homogeneização do resultado da classificação

Geralmente, no resultado da classificação, ocorrem pixels isolados ou pequenos grupos de pixels pertencentes a várias classes, próximo a áreas homogeneamente classificadas, gerando uma imagem com ruído (Crosta, 1993).

Com o propósito de homogeneizar o resultado da classificação pelos diferentes algoritmos, foi aplicado um filtro de mediana de tamanho 3 x 3, para a redução do ruído nas imagens classificadas. O valor de cada pixel (que corresponde à classe a qual pertence) foi substituído pela mediana dos valores do pixels da sua vizinhança e não pela sua média. No filtro de mediana de tamanho 3 x 3 que foi aplicado, os valores do pixel e dos seus vizinhos foram ordenados, determinado-se a mediana e atribuído este valor ao pixel central. Dessa forma, obtém-se a redução do ruído da imagem, sem o efeito de borramento e preservando as bordas (Gonzalez & Woods, 2000).

3.3.4 Comparação entre os algoritmos de classificação

3.3.4.1 Análise de exatidão

Um método comum para avaliação da acurácia da classificação é o uso da matriz de erro e das estatísticas derivadas, o que permite o cômputo da exatidão global, exatidão do produtor, exatidão do usuário e do coeficiente Kappa (Lu et al., 2003). O uso da matriz de erro em dados de sensoriamento remoto é relatado desde 1970 (Skidmore, 1999).

Dessa forma, para quantificar os erros gerados pelas diferentes técnicas de classificação utilizadas neste estudo, foram construídas as matrizes de erro ou de confusão que permitem a comparação entre os dados provenientes das amostras de amostra de validação e o resultado da classificação. O conjunto de

pixels por classe temática, coletados com GPS, foi denominado de amostras de validação.

Em todas as modalidades de classificação realizadas neste estudo, as matrizes de erros foram calculadas com o mesmo conjunto de amostras de validação. Além da exatidão global ou acurácia do mapeamento, por meio da matriz de erro, foram calculados também os erros de inclusão e de omissão por classe temática e o coeficiente Kappa, que auxilia na avaliação do classificador.

A exatidão global do mapeamento foi, então, obtida pela razão entre a somatória do número de pixels corretamente classificados em todas as classes temática e o número total de pixels da amostra de validação. Os pixels corretamente classificados são os componentes da diagonal da matriz de confusão. A quantidade de pixels amostrados por classe temática, usados na construção das matrizes de erros, está listada na Tabela 6.

TABELA 6 Tamanho da amostra de validação por classe temática

Classe temática	Número de pixels da amostra de validação/ano		
	2003	2002	2000
Floresta	3.000	5.000	1.000
Capoeira	687	557	823
Pasto alto	947	600	600
Pasto baixo	2.193	842	1.921
Solo	311	1.000	600
Água	562	600	473
Queimada	-	601	600
Total	7.700	9.200	6.017

Os erros de inclusão foram obtidos pela razão entre o número de pixels classificados erroneamente em uma classe, pelo número total de pixels

classificados nesta classe. Esse erro indica o percentual de pixels incluídos em uma classe quando, na verdade, pertence a outra(s). A acurácia do usuário foi obtida pela razão entre o número de pixels corretamente classificados em uma classe pelo número total de pixels classificados nesta classe.

Os erros de omissão foram calculados pela razão entre o número de pixels classificados erroneamente em uma classe pelo número total de pixels da amostra de validação da respectiva classe. Esse erro indica o percentual de pixels omitidos da classe correta que foram atribuídos a outra(s). A acurácia do produtor foi calculada pela razão entre o número de pixels corretamente classificados em uma classe pelo número total de pixels da amostra de validação da respectiva classe.

O valor do coeficiente Kappa e sua variância foram calculados aplicando-se as equações citadas no item 2.5, incluindo no cálculo os pixels que foram erroneamente classificados, além dos que foram classificados corretamente, como na exatidão global. Os valores de Kappa variam de 0 a 1, indicando a concordância entre a imagem classificada e as amostras de amostra de validação. A variância de Kappa foi usada para calcular se existe diferença significativa entre as modalidades de classificação.

3.3.4.2 Estimativa das áreas das classes de uso e ocupação do solo

No levantamento do uso e ocupação do solo e para a comparação entre os algoritmos de classificação computou-se a estatística das classes temáticas para todos os anos e métodos de classificação, calculando-se o percentual de área coberta pelas classes de floresta, capoeira, pasto alto e baixo, água e solo em relação à área total do município de Capixaba, Acre.

Foram utilizadas as informações geradas a partir da matriz de erro para analisar a acurácia por classe de mapeamento, como os erros de inclusão e omissão, conforme descrito no item 3.3.4.1.

3.3.4.3 Teste de significância

Para comparar o desempenho entre os diferentes métodos de classificação foi usada a variância de Kappa e estatística Z, comumente utilizada na literatura (Fidalgo, 1995; Carvalho, 2001; Gaboardi, 2002; Lu et al., 2003; Carvalho et al., 2004).

A variância de Kappa foi calculada a partir das matrizes de erro das classificações. Na comparação dos resultados de Kappa entre os métodos de classificação, a estatística Z foi analisada para os níveis de confiança de 99%, 95% e 90%. O teste Z é indicado para duas amostras independentes quando as variâncias paramétricas são conhecidas e as variáveis nas populações apresentam distribuições normais (Ayres et al., 2000). Segundo Skidmore (1999), este teste estatístico pode ser aplicado em combinações pares de matrizes de erro, determinando se as matrizes são significativamente diferentes.

Para se fazer a inferência estatística, utilizou-se o teste de hipótese, que é um procedimento matemático denominado regra de decisão, para aceitar ou rejeitar uma hipótese estatística com base nos elementos da amostra (Fonseca & Martins, 1996).

Consideraram-se K_1 e K_2 como as estimativas de Kappa e σ_{k1}^2 e σ_{k2}^2 suas respectivas variâncias, representando dois classificadores, respectivamente, para testar a igualdade de duas matrizes de confusão e a estatística Z, expressa por:

$$Z = \frac{|K_1 - K_2|}{\sqrt{\sigma_{k1}^2 + \sigma_{k2}^2}}$$

As hipóteses para o teste de significância foram:

$H_0: K_1 = K_2$ (hipótese nula) e

$H_1: K_1 > K_2$ (hipótese alternativa).

A hipótese H_0 foi testada para os níveis de significância (α) de 0,01; 0,05 e 0,10. O valor tabelado de Z foi o da hipótese alternativa unilateral à direita. Assim, à significância α , H_0 será rejeitada se $Z > Z_\alpha$.

4 RESULTADOS E DISCUSSÕES

4.1 Correção geométrica

A correção geométrica das imagens pelo método de transformação polinomial de 1º grau e reamostragem dos pixels pela alocação do vizinho mais próximo (*nearest neighbour*) obteve resultados satisfatórios, gerando um erro inferior a meio pixel para todas as imagens. Para as imagens de 2000, 2002 e 2003, os erros médios quadráticos foram, respectivamente, 0,2774, 0,4767 e 0,3924 pixel.

4.2 Calibração radiométrica

A conversão dos números digitais para reflectância aparente, foi calculada utilizando-se as relações propostas por Markham & Barker (1987), empregando-se os parâmetros citados na seção 3.3.2.2, cujos resultados são de atenuação dos efeitos da interferência atmosférica e característica dos sensores no comportamento espectral dos alvos na superfície terrestre.

Nas imagens cujos pixels são expressos em números digitais, com resolução radiométrica de 8 bits, cada pixel pode ser representado por $2^8 = 256$ níveis de cinza discretos, ou seja, 1 byte (8 bits) por pixel, numa escala de 0 a 255. Nas imagens calibradas para reflectância aparente, cada pixel pode ser representado por 4 bytes ou 2^{32} , que é igual a aproximadamente, 4 bilhões de combinações, numa escala de valores decimais entre 0 e 1.

Se, por um lado, a calibração radiométrica possui a vantagem de representar valores físicos de reflectância dos alvos, por outro, é necessário maior recurso computacional para as etapas de processamento das imagens,

visto que o número de bits para armazenar uma imagem digital é diretamente proporcional ao número de pixels e ao número de níveis de cinza permitidos para cada pixel. Enquanto a imagem em número digital ocupou 20.256,80 Kb (mil bytes), em reflectância aparente foram necessário 81.025,20 Kb.

Na Tabela 7 são apresentados valores médios de pixels puros dos componentes solo, água e vegetação expressos em número digital e reflectância, coletados na imagem Landsat do ano de 2003.

TABELA 7 Valores médios do conjunto de pixels dos componentes água, vegetação e solo, nas bandas 1, 2, 3, 4, 5 e 7 do TM/Landsat em 2003

Banda	Água		Vegetação		Solo	
	ND	Ra	ND	Ra	ND	Ra
1	58	0,0949	54	0,0878	62	0,1024
2	28	0,0963	24	0,0811	29	0,0997
3	28	0,0733	19	0,0465	34	0,0905
4	12	0,0471	100	0,4110	63	0,2605
5	06	0,0142	65	0,1684	107	0,2782
7	04	0,0108	18	0,0573	42	0,1392

em que: ND - número digital e Ra - reflectância aparente.

Observando-se os pixels de vegetação, verifica-se que a diferença entre os valores nas bandas 3 e 4 é de aproximadamente 5 vezes e 9 vezes, respectivamente, em termos de número digital e reflectância aparente. Comportamento semelhante ocorre também nos pixels de solo e água, porém, com menor amplitude.

Os pixels expressos em valores físicos, ou seja, em reflectância oferecem um comportamento mais próximo do real dos alvos ou componentes

que estão sendo analisados. Esse comportamento é evidenciado na Figura 15, em que estão representadas as curvas espectrais dos mesmos pixels descritos na Tabela 6.

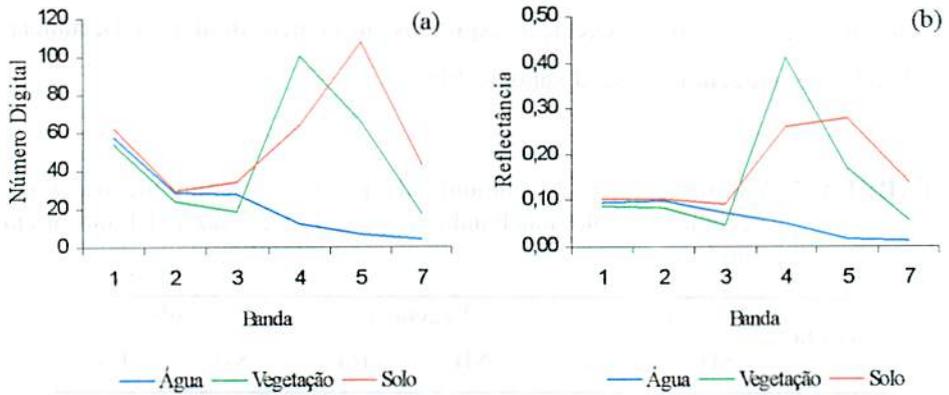


FIGURA 15 Curvas espectrais dos componentes água, vegetação e solo, expressas em (a) número digital e (b) reflectância aparente nas bandas 1, 2, 3, 4, 5 e 7 do TM/Landsat em 2003

Analisando-se a Figura 15, observa-se o aumento na inclinação da curva da vegetação em relação aos outros componentes puros, quando comparado o comportamento espectral desses alvos em número digital e reflectância aparente. A assinatura espectral do solo apresentou comportamento diferenciado expresso em número digital e reflectância. Em número digital, ocorreu um aumento acentuado entre as bandas 4 e 5, enquanto que, em valores de reflectância aparente, a curva foi mais suavizada. As assinaturas espectrais expressas em reflectância aparente representam o comportamento dos alvos mais próximo do real.

4.3 Extração de informações das imagens

4.3.1 Modelo linear de mistura espectral

A aplicação do algoritmo de decomposição espectral linear gerou as imagens sintéticas sombra, solo e vegetação, que realçaram as feições dos componentes puros, além da imagem de erro que contém a distribuição espacial dos erros residuais dos alvos que não se enquadram no padrão dos componentes puros selecionados.

Na Figura 16 são apresentadas a imagem Landsat na composição colorida das bandas 5, 4 e 3 e as imagens geradas pelo modelo linear de mistura espectral de uma mesma área contendo as feições dos componentes puros.

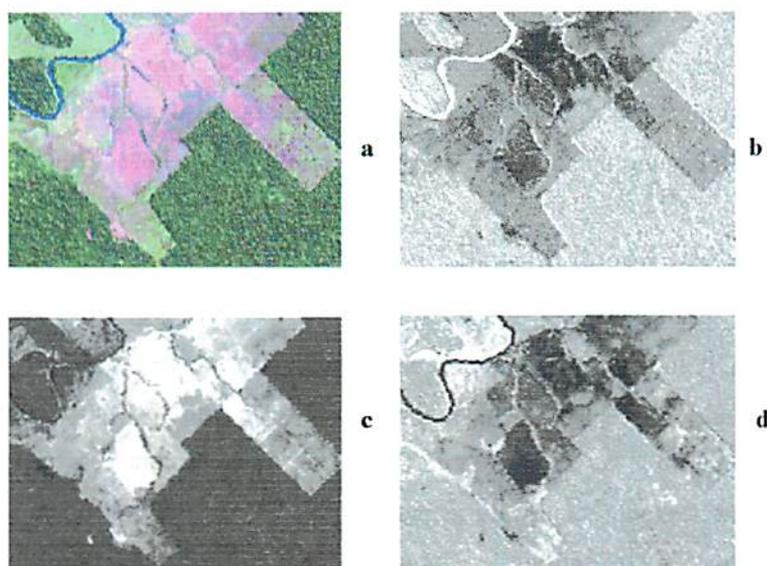


FIGURA 16 Imagem TM/Landsat, ano 2003, na composição colorida RGB das bandas 5, 4 e 3 (a), imagem fração sombra (b), imagem fração solo (c) e imagem fração vegetação (d)

Na imagem Landsat (Figura 16a), a área de floresta possui cor verde-escuro, água na cor azul, solo exposto na tonalidade magenta mais claro e áreas com vegetação arbustiva e gramíneas em tons de verde-claro.

Na imagem fração sombra (Figura 16b), foram realçadas água e floresta, que apresentam tons de cinza mais claros. Neste caso, floresta foi mais realçada que as demais formas de vegetação, devido à maior quantidade de sombra sob o dossel, diferenciando-a da vegetação onde penetra maior quantidade de luz.

As áreas de solo exposto foram bem destacadas na imagem fração solo (Figura 16c) em tons de cinza claro. Quanto mais claro o pixel na imagem, maior a proporção do componente puro que está presente no pixel. Onde ocorre vegetação mais rasteira com relativa exposição de solo, foram verificados tons de cinza-médio e nas áreas com maior cobertura do solo, como no caso das florestas, tons de cinza mais escuro.

A imagem fração vegetação (Figura 16d) realçou mais a vegetação secundária em estágio intermediário de regeneração, com cobertura de solo mais homogênea, como as observadas na área marginal das curvas do rio na parte superior esquerda da imagem e em algumas regiões de borda com floresta.

4.3.2 Índice de vegetação

A imagem resultante do NDVI gerou um arquivo com uma única banda, reduzindo a dimensionalidade dos dados e realçando as feições com vegetação, conforme pode ser verificado na Figura 17. As áreas com vegetação arbustiva e arbórea possuem tons de cinza mais claro, enquanto as áreas com exposição do solo tons de cinza-escuro. A imagem NDVI foi utilizada como um importante atributo no algoritmo árvore de decisão, auxiliando na divisão do conjunto de dados em subconjuntos mais homogêneos até a diferenciação entre as classes temáticas.

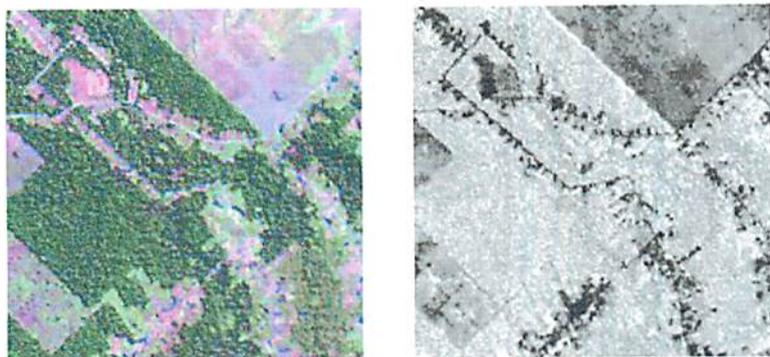


FIGURA 17 Imagem TM/Landsat, ano 2003, na composição colorida RGB das bandas 5, 4 e 3 (à esquerda) e imagem NDVI (à direita)

4.4 Classificação digital das imagens

Nos seis esquemas de classificação digital foram avaliados a exatidão global e coeficiente de concordância Kappa obtidos por meio das matrizes de erro dos seis métodos de classificação analisados que se encontram em anexo. Também foi considerado a exatidão por classe de mapeamento, bem como os respectivos erros de inclusão e omissão.

Em geral, a classificação supervisionada apresentou maior acurácia no mapeamento do uso e ocupação do solo em Capixaba em relação às técnicas de classificação não supervisionada. Esse resultado é evidenciado na Figura 18, na qual verifica-se que a classificação por máxima verossimilhança usando as imagens Landsat e as imagens fração geradas pelo análise de mistura espectral, apresentam, em média, os maiores valores do coeficiente de concordância Kappa, em torno de 0,9. Nas modalidades de classificação usando o algoritmo isodata, com as imagens Landsat ou imagens fração, o valor de Kappa variou entre 0,74 a 0,85.

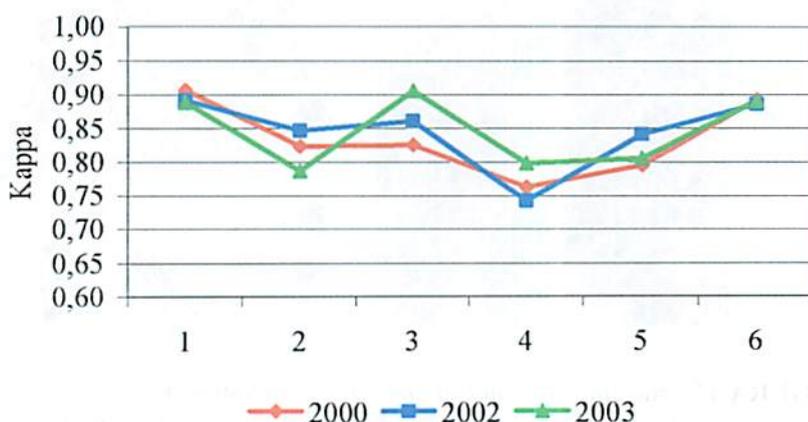


FIGURA 18 Índice de concordância Kappa das seis técnicas de classificação digital para o mapeamento do uso e ocupação do solo em Capixaba, Acre, nos anos de 2000, 2002 e 2003: (1) máxima verossimilhança usando imagens Landsat, (2) isodata usando imagens Landsat, (3) árvore de decisão, (4) isodata usando imagens fração solo e sombra, (5) isodata usando imagens fração solo, sombra e vegetação e (6) máxima verossimilhança usando imagens fração solo, sombra, vegetação e de erro.

A eficiência obtida por cada uma das técnicas de classificação digital aplicadas são detalhadas a seguir.

4.4.1 Máxima verossimilhança

A exatidão global da classificação utilizando o algoritmo de máxima verossimilhança variou entre 91,1% e 92,5%, com valor do coeficiente Kappa em torno de 0,9, representando uma excelente classificação (Landis & Koch, 1977).

Em referência à exatidão por classe de mapeamento, os resultados obtidos são mostrados na Tabela 8. A acurácia do usuário representa a proporção

de pixels corretamente classificados numa classe em relação ao número total de pixels classificados na classe. A acurácia do produtor representa o percentual de pixels corretamente classificados em relação ao número total de pixels da amostra de validação. A classe queimada não foi incluída no mapeamento do ano de 2003.

TABELA 8 Exatidão por classe de mapeamento obtido pelo classificador de máxima verossimilhança nos anos de 2003, 2002 e 2000

Classe temática	Ano/Acurácia					
	2003		2002		2000	
	Usuário	Produtor	Usuário	Produtor	Usuário	Produtor
Capocira	91,03%	85,54%	76,00%	95,00%	79,17%	91,57%
Floresta	95,57%	99,49%	98,79%	99,39%	99,13%	97,45%
Solo	100,00%	53,85%	75,44%	86,00%	100,00%	83,95%
Água	100,00%	100,00%	98,36%	100,00%	98,41%	100,00%
Pasto alto	86,41%	93,68%	96,97%	64,00%	80,95%	78,46%
Pasto baixo	79,13%	89,22%	81,40%	70,00%	86,25%	89,61%
Queimada	-	-	100,00%	100,00%	95,00%	100,00%

As classes água, queimada e floresta obtiveram excelente acurácia do usuário e do produtor, acima de 95%. O pior resultado foi na classe solo e pasto alto, com acurácia do produtor de 54% no ano de 2003 e 64% no ano de 2002. No entanto, o desempenho foi melhorado nos demais anos.

Os maiores erros de omissão foram pasto alto, pasto baixo e solo, entre 22% e 46%, indicando, percentualmente, os pixels dos dados de referência de uma classe que foram classificados erroneamente em outras classes.

4.4.2 Isodata

A exatidão global da classificação utilizando o algoritmo isodata variou entre 82,9% e 87,4%, com valor de Kappa entre 0,79 a 0,85, representando uma classificação muito boa a excelente (Landis & Koch, 1977).

Os melhores resultados de exatidão por classe de mapeamento foram das classes floresta e solo, conforme mostrado na Tabela 9.

TABELA 9 Exatidão por classe de mapeamento obtido pelo classificador isodata nos anos de 2003, 2002 e 2000

Classe temática	Ano/Acurácia					
	2003		2002		2000	
	Usuário	Produtor	Usuário	Produtor	Usuário	Produtor
Capoeira	70,73%	34,94%	85,94%	91,67%	93,55%	69,88%
Floresta	96,00%	98,46%	97,01%	98,78%	97,61%	86,81%
Solo	82,26%	98,08%	100,00%	100,00%	97,50%	96,30%
Água	100,00%	93,51%	57,14%	100,00%	69,66%	100,00%
Pasto alto	55,17%	84,21%	93,48%	86,00%	59,09%	100,00%
Pasto baixo	91,67%	75,49%	100,00%	94,00%	93,62%	57,14%
Queimada	-	-	100,00%	25,00%	86,36%	100,00%

Nas demais classes, os resultados foram variados nos três anos de análise. A classe capoeira demonstrou desempenho inferior, em termos de acurácia do produtor, de 35% a 70% no ano de 2003 e 2000, respectivamente.

Em 2002, o pior resultado foi de 25% de acurácia do produtor da classe queimada, isso porque 75% dos pixels da amostra de referência foram erroneamente omitidos da classe queimada e incluídos na classe água.

Verificou-se também confusão com as classes de capoeira, água, pasto alto e pasto baixo, apresentando erros de inclusão ou omissão entre 41% e 65%, conforme Tabela 4C em anexo.

Particularmente no ano de 2000, houve também confusão entre as classes de floresta e água, quando pixels de floresta em tons de verde mais escuro na composição colorida das bandas 5, 4 e 3, que representam maior quantidade de sombra, foram classificados erroneamente como da classe água.

A Figura 19 ilustra essa situação: na imagem classificada pelo algoritmo isodata, as cores azul, verde-escuro, verde-claro, lilás e rosa representam, respectivamente, as classes temáticas água, floresta, capoeira, queimada e solo.

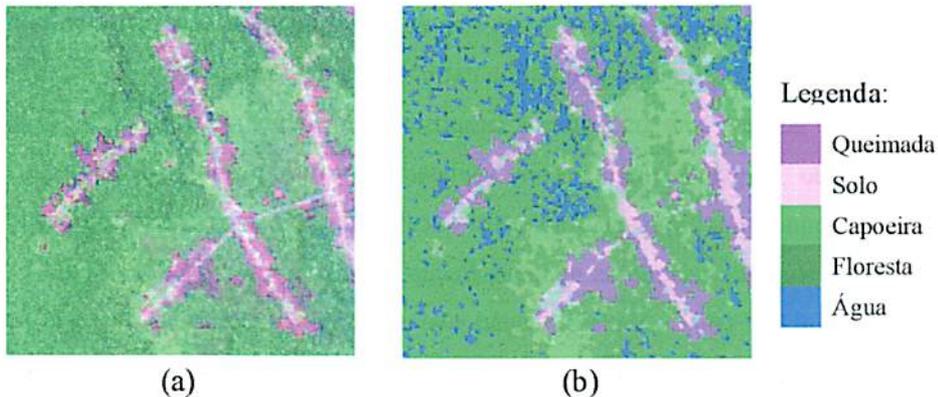


FIGURA 19 Imagem TM/Landsat na composição colorida RGB das bandas 5, 4 e 3 (a) e mapa resultado da classificação isodata (b)

4.5.3 Classificação usando as imagens fração oriundas do modelo linear de mistura espectral

As classificações usando as imagens fração oriundas do modelo linear de mistura espectral são: i) classificação não supervisionada, com aplicação do algoritmo isodata nas imagens fração solo e sombra; ii) classificação não

supervisionada, com aplicação do algoritmo isodata nas imagens fração solo, sombra e vegetação e iii) classificação supervisionada, com aplicação do algoritmo de máxima verossimilhança nas imagens fração solo, sombra, vegetação e de erro.

A classificação supervisionada por máxima verossimilhança usando as imagens fração obteve desempenho superior com exatidão global de 91% em todos os anos. O pior resultado foi da classificação não supervisionada, com aplicação do algoritmo isodata nas imagens fração solo e sombra, entre 78,9% e 83,8%, devido ao fato do algoritmo não separar a classe temática queimada em 2000 e 2002, com erro de omissão de 100% na respectiva classe. No entanto, o uso da imagem fração vegetação juntamente com as imagens fração solo e sombra melhorou esse resultado, visto que nesse esquema de classificação não supervisionada a exatidão global variou entre 83,8% e 87,0%.

De acordo com os critérios propostos por Landis & Koch (1977), a classificação supervisionada, com aplicação do algoritmo de máxima verossimilhança nas imagens fração foi excelente, assim como a classificação não supervisionada usando as imagens fração solo, sombra e vegetação para os anos de 2002 e 2003. A classificação não supervisionada usando apenas as imagens fração solo e sombra foi considerada muito boa, pois os valores de Kappa variaram de 0,7407 a 0,7969.

Em termos de acurácia por classe de mapeamento, a classificação não supervisionada usando apenas as imagens fração solo e sombra apresentou resultado superior a 93% para floresta e 85% para água (Tabela 10). No entanto, ocorreu maior confusão com a classe de pasto alto, com acurácia entre 44% e 83%. A classe de pasto baixo apresentou bom desempenho em termos de acurácia do produtor, com percentual superior a 92%. A classe queimada não foi mapeada.

TABELA 10 Exatidão por classe de mapeamento obtido pelo classificador isodata usando as imagens fração solo e sombra, nos anos de 2003, 2002 e 2000

Classe temática	Ano/Acurácia					
	2003		2002		2000	
	Usuário	Produtor	Usuário	Produtor	Usuário	Produtor
Capoeira	89,66%	62,65%	80,00%	66,67%	72,73%	67,47%
Floresta	97,94%	97,44%	92,66%	100,00%	93,20%	99,15%
Solo	89,29%	96,15%	100,00%	86,00%	98,18%	66,67%
Água	100,00%	98,70%	100,00%	88,33%	100,00%	85,48%
Pasto alto	59,15%	44,21%	61,19%	82,00%	73,97%	83,08%
Pasto baixo	64,43%	94,12%	47,12%	98,00%	53,38%	92,21%
Queimada	-	-	0,00%	0,00%	0,00%	0,00%

Na classificação não supervisionada usando as imagens fração solo, sombra e vegetação, conforme Tabela 11, a exatidão do usuário e do produtor variou de 79% a 100% nas classes de floresta, água e queimada. Na classe capoeira, o desempenho foi, em média, de 53% de acurácia do produtor e 72% de acurácia do usuário. Para pasto baixo, obteve-se acima de 95% para exatidão do produtor e em torno de 71% para exatidão do usuário. Solo obteve 97% e 84%, para acurácia do usuário e produtor, respectivamente. Pasto alto apresentou os piores resultados; seu desempenho médio foi de 64%.

Em geral, a exatidão por classe de mapeamento do classificador de máxima verossimilhança usando todas as imagens fração (Tabela 12) foi superior, alcançando média de 81% em todas as classes de mapeamento, exceto para solo e pasto baixo, com acurácia do produtor de 78% a 79%, respectivamente. Na classe água, o desempenho foi de 100% e, para floresta, acima de 96%.

TABELA 11 Exatidão por classe de mapeamento obtido pelo classificador isodata usando as imagens fração solo, sombra e vegetação, nos anos de 2003, 2002 e 2000

Classe temática	Ano/Acurácia					
	2003		2002		2000	
	Usuário	Produtor	Usuário	Produtor	Usuário	Produtor
Capoeira	81,13%	51,81%	62,75%	53,33%	70,97%	53,01%
Floresta	97,44%	97,44%	92,66%	100,00%	92,86%	99,57%
Solo	94,34%	96,15%	100,00%	86,00%	98,25%	69,14%
Água	100,00%	93,51%	100,00%	90,00%	100,00%	85,48%
Pasto alto	61,54%	58,95%	60,78%	62,00%	71,88%	70,77%
Pasto baixo	70,71%	97,06%	79,31%	92,00%	62,71%	96,10%
Queimada	-	-	100,00%	100,00%	85,71%	78,95%

TABELA 12 Exatidão por classe de mapeamento obtido pelo classificador de máxima verossimilhança usando as imagens fração solo, sombra, vegetação e de erro, nos anos de 2003, 2002 e 2000

Classe temática	Ano/Acurácia					
	2003		2002		2000	
	Usuário	Produtor	Usuário	Produtor	Usuário	Produtor
Capoeira	92,21%	85,54%	75,71%	88,33%	93,44%	68,67%
Floresta	95,57%	99,49%	95,88%	99,39%	96,28%	99,15%
Solo	100,00%	51,92%	79,25%	84,00%	94,12%	98,77%
Água	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%
Pasto alto	87,38%	94,74%	96,97%	64,00%	70,11%	93,85%
Pasto baixo	78,63%	90,20%	79,17%	76,00%	85,94%	71,43%
Queimada	-	-	100,00%	100,00%	95,00%	100,00%

Lu et al. (2003) obtiveram exatidão global de 78,2% aplicando modelo de mistura para classificar vegetação na Amazônia nos seguintes temas: floresta,

pasto e agricultura, sucessão secundária em fase inicial, intermediária e final, urbano e solo exposto e água. A classe de floresta obteve acurácia acima de 93%, enquanto que as classes de sucessão secundária apresentaram resultados variados entre 42% e 70%.

4.5.4 Árvore de decisão

A classificação pelo algoritmo árvore de decisão foi excelente, com valor de Kappa acima de 0,83. A exatidão global foi de 92,38%, 88,66% e 86,74%, em 2003, 2002 e 2000, respectivamente.

As classes de floresta, água e queimada apresentaram os melhores desempenhos, com exatidão do usuário e produtor acima de 94%. A classe de capoeira manteve-se com 82% a 88% de exatidão do produtor e a classe de solo acima de 72%.

As maiores variações na exatidão foram verificadas nas classes de pasto alto e pasto baixo. Em 2003, essas classes variaram de 82% a 95% de acurácia. No entanto, em 2000 e 2002, obtiveram acurácia do produtor entre 41% e 88% (Tabela 13).

A classe que apresentou a menor exatidão foi pasto alto, que inclui as subclasses pasto sujo, pasto limpo e agricultura. Dessa forma, foram cometidos erros de inclusão e omissão com a classe capoeira, nos estágios iniciais de sucessão, que se confunde com a subclasse pasto sujo. O mesmo ocorreu entre as classes de pasto limpo e pasto baixo. Essa situação ocorreu com maior intensidade no ano de 2000, possivelmente em função de maior quantidade de chuva ocorrida no período em relação aos demais anos analisados, bem como do espaço temporal entre a coleta das amostras e a data da imagem.

TABELA 13 Análise de exatidão por classe de mapeamento obtido pelo classificador árvore de decisão, nos anos de 2003, 2002 e 2000

Classe temática	Ano/Acurácia					
	2003		2002		2000	
	Usuário	Produtor	Usuário	Produtor	Usuário	Produtor
Capoeira	92,00%	83,13%	73,13%	81,67%	63,30%	83,13%
Floresta	94,63%	99,49%	94,22%	99,39%	96,27%	98,72%
Solo	97,78%	84,62%	72,73%	80,00%	100,00%	71,60%
Água	100,00%	100,00%	100,00%	100,00%	100,00%	98,39%
Pasto alto	81,82%	94,74%	94,12%	64,00%	69,23%	41,54%
Pasto baixo	91,30%	82,35%	77,27%	68,00%	71,58%	88,31%
Queimada	-	-	98,36%	100,00%	97,37%	97,37%

4.6 Comparação entre os algoritmos de classificação

4.6.1 Estimativa das áreas das classes de uso e ocupação do solo

O percentual da área ocupada pelas classes de floresta, capoeira, pasto alto, pasto baixo, água, queimada e solo exposto variou em função da técnica de classificação empregada no mapeamento do uso e ocupação do solo em Capixaba, conforme demonstra a Figura 20.

A diferença entre o menor e o maior percentual de área coberta com floresta foi de 5,1%. O menor índice foi obtido com a técnica de classificação não supervisionada isodata usando as imagens fração solo e sombra (52,8%) e o maior com o algoritmo árvore de decisão (57,9%). Esses classificadores alcançaram exatidão superior a 97% para a classe de floresta. Porém, o erro de inclusão foi de 5,4% a 2,1% para a classificação por árvore de decisão e

classificação não supervisionada isodata usando as imagens fração solo e sombra, respectivamente, enquanto a omissão foi de 0,5% a 2,7%.

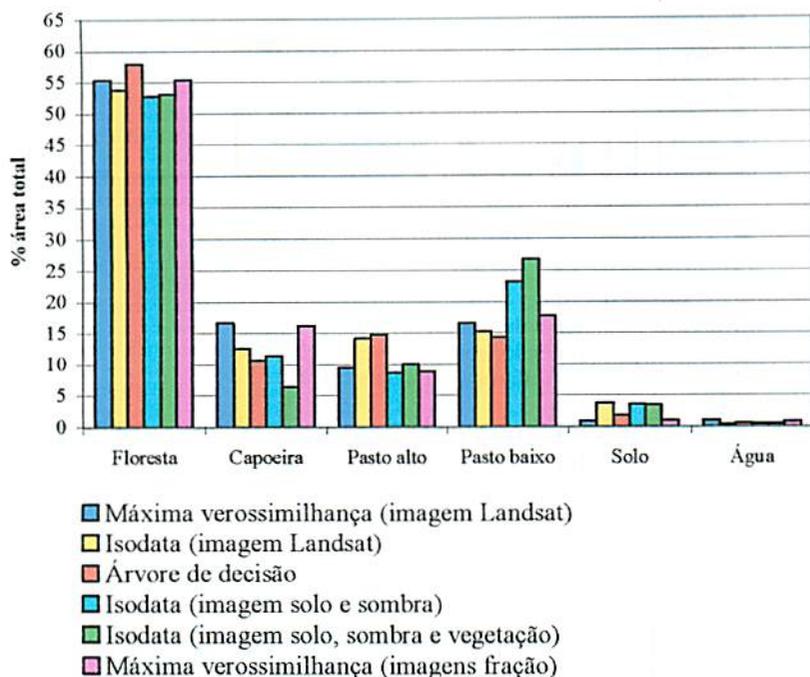


FIGURA 20 Percentual de área ocupada pelas classes de uso e ocupação do solo em Capixaba, Acre, 2003, segundo a técnica de classificação utilizada

O mesmo comportamento foi verificado nas demais classes mapeadas e nos anos de 2002 e 2000. Nas Figuras 21 e 22, observa-se o resultado da porcentagem de área ocupada por cada classe de mapeamento, segundo a técnica de classificação adotada. A variação entre as classes de capoeira, pasto alto e pasto baixo ocorreu devido a semelhança espectral na fronteira de decisão entre as classes, podendo ocorrer a sobreposição em alguma delas, e a habilidade de cada classificador em estabelecer o limite de decisão no espaço de feições.

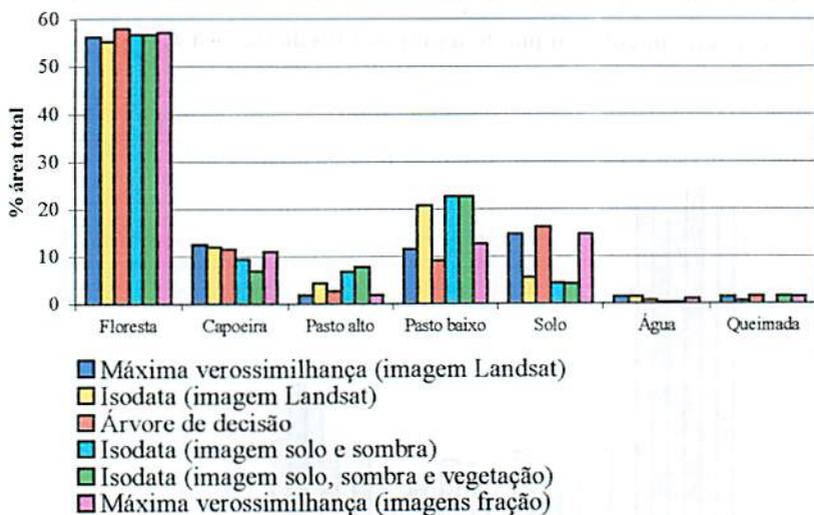


FIGURA 21 Percentual de área ocupada pelas classes de uso e ocupação do solo em Capixaba, Acre, 2002, segundo a técnica de classificação utilizada

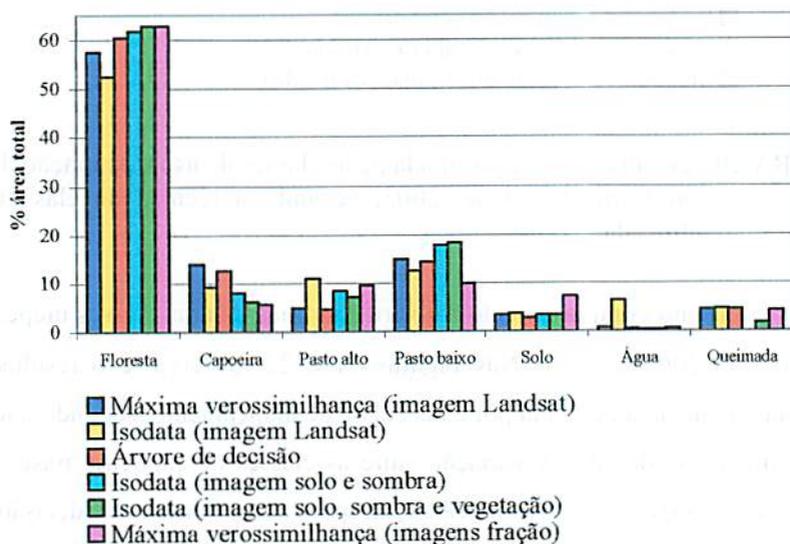


FIGURA 22 Percentual de área ocupada pelas classes de uso e ocupação do solo em Capixaba, Acre, 2000, segundo a técnica de classificação utilizada

A confusão entre as classes de mapeamento, causada pelos erros de inclusão ou omissão nas diferentes técnicas de classificação, constituem num dos fatores que levam a diferentes estimativas da proporção de área ocupada por essas classes. No ano de 2000, por exemplo, o classificador isodata teve 30,3% de erros de inclusão na classe água, quando, na verdade, os pixels eram da classe floresta. Em consequência, no ano de 2002, a área ocupada por floresta foi maior que em 2000, segundo a classificação isodata. Aplicando-se as imagens fração do modelo linear de mistura, esse problema não ocorreu.

Em áreas de capoeira em estágios avançados de sucessão secundária, por exemplo, o comportamento espectral é muito semelhante ao das florestas maduras, tornando-se uma tarefa difícil a separação entre elas.

Na Figura 23 é mostrada a interpretação das seis técnicas de classificação para uma região onde ocorre a classe capoeira em vários estágios sucessionais, cuja classificação do PRODES digital corresponde a desmatamento total (INPE, 2004b). As setas de cor amarela na imagem Landsat indicam áreas de capoeira em vários estágios de sucessão secundária; as setas de cor branca nos mapas resultado da classificação digital indicam áreas que foram classificadas como floresta; as áreas em amarelo indicam regiões classificadas corretamente como capoeira e o polígono com bordas em branco indica áreas classificadas como desmatamento total pelo INPE.

Analisando-se ainda a Figura 23, verifica-se que as áreas em estágios de sucessão secundária de inicial a intermediário foram mais facilmente classificadas corretamente como capoeira pelos diferentes algoritmos utilizados. Nos estágios de sucessão mais avançados, os classificadores incluíram as áreas de capoeira como floresta.

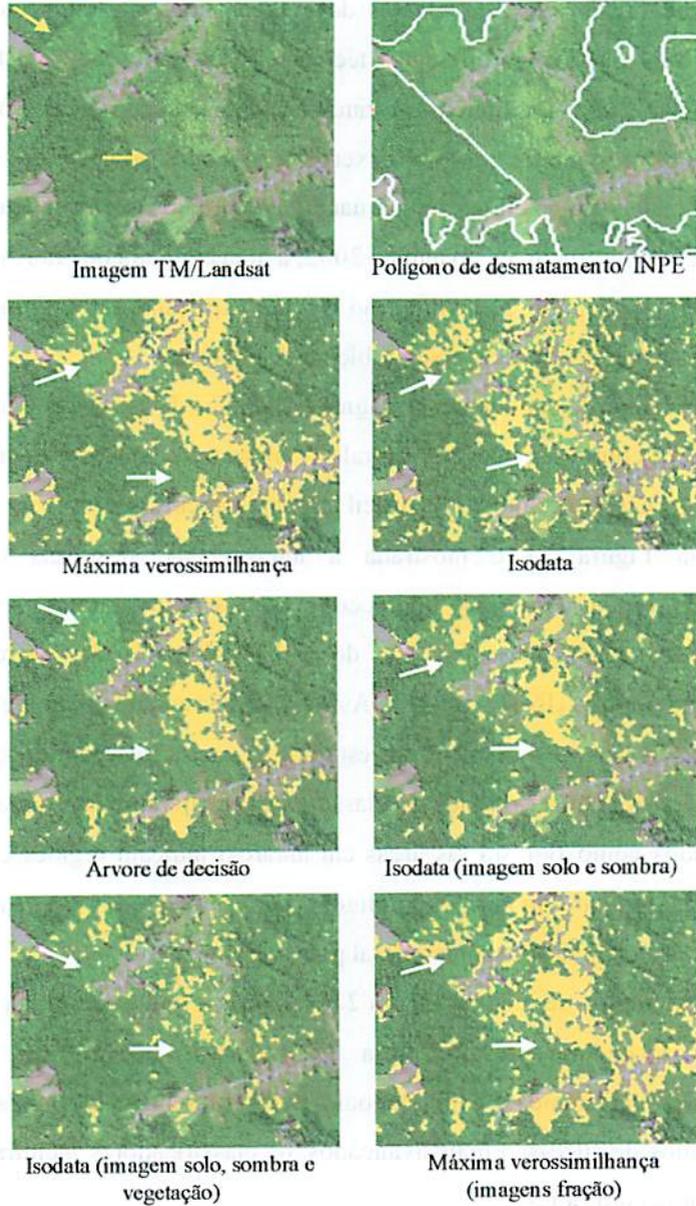


FIGURA 23 Comparação entre os classificadores de máxima verossimilhança, isodata e árvore de decisão, usando imagem Landsat e imagens fração oriundas do modelo linear de mistura espectral no mapeamento da classe temática capoeira, em Capixaba, Acre

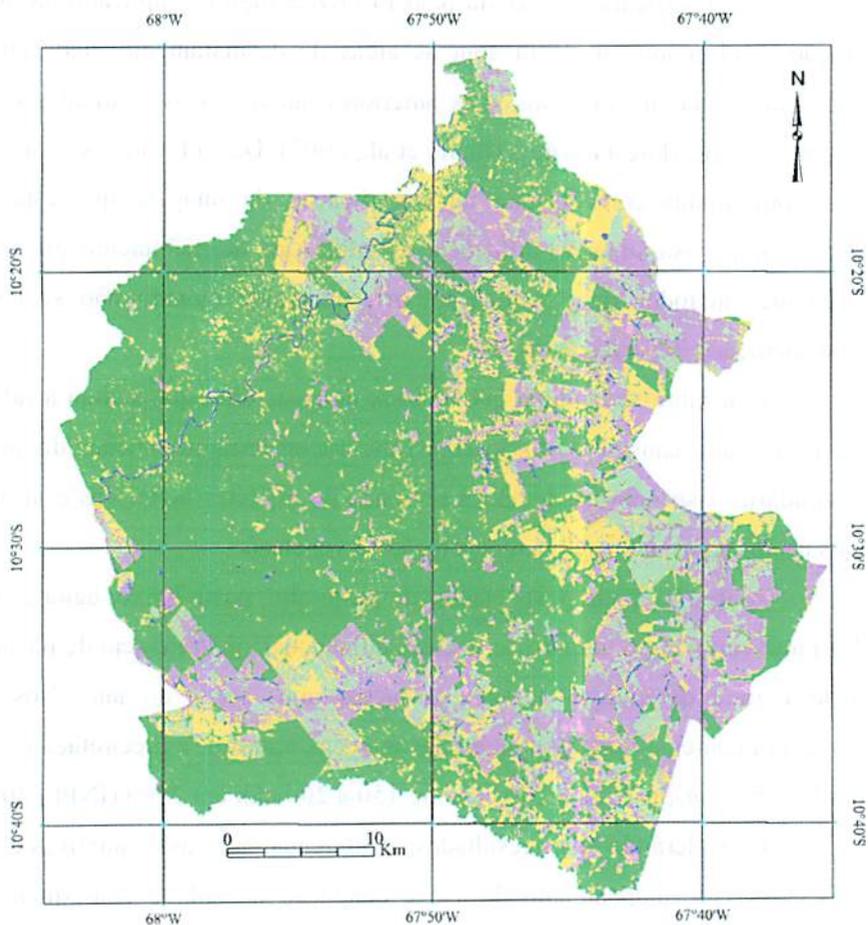
Na classificação realizada pelo PRODES digital é aplicado na imagem fração sombra uma máscara com as áreas de desmatamento, não floresta e hidrografia, classificadas nos anos anteriores, incluindo os resultados da série histórica de desflorestamento (Duarte et al., 2003). Dessa forma, essas áreas não são confundidas com floresta na classificação da imagem que está sendo processada. Essa situação influencia os índices de desmatamento gerados por diferentes métodos de classificação, principalmente quando não se leva em consideração a série histórica.

Constatou-se também que a classe de pasto alto, que incluiu a subclasse de pasto sujo, causou confusão com capoeira em estágios iniciais de sucessão secundária. Isso ocorreu devido à presença de espécies herbáceas e arbustivas nos pastos sujos, iniciando a regeneração da vegetação.

A área ocupada pelas classes de pasto alto, pasto baixo, água e solo foi bastante variável no período de análise, influenciada pelo manejo de pastagem e pelo regime de chuvas durante os meses mais secos do ano. Nos meses correspondentes a data de tomada das imagens Landsat, a precipitação total foi de 25 a 50 mm em 2002 e 2003 e entre 150 a 200 mm em 2000 (INPE, 2004c).

Considerando-se os resultados, a diferença entre as estimativas de áreas das classes de mapeamento de uso e ocupação do solo é consequência das diferentes técnicas de classificação e da habilidade e experiência do analista. Dessa forma, deve-se considerar os desempenhos da exatidão global e por classe temática entre as diferentes técnicas de classificação, bem como a importância das categorias de uso e ocupação do solo que se deseja mapear para selecionar o método mais adequado ao objetivo do trabalho a ser realizado.

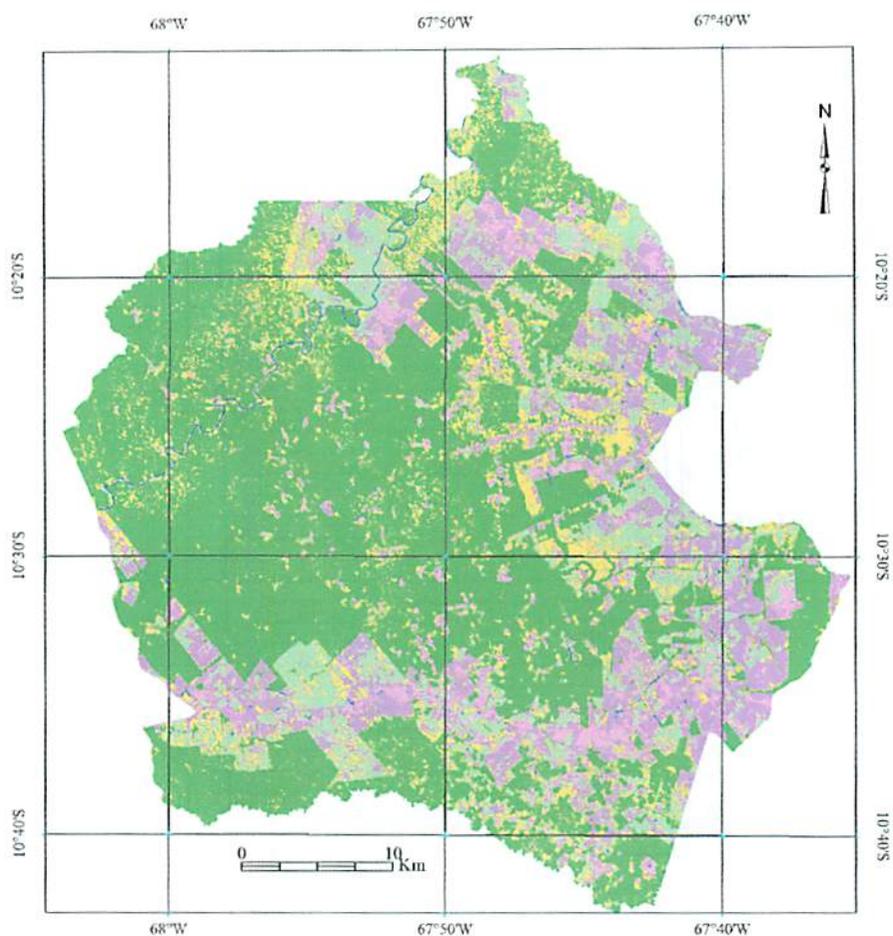
Os mapas de uso e ocupação do solo do município de Capixaba em 2003, segundo as diferentes técnicas de classificação digital aplicadas neste estudo, são apresentados nas Figuras 24, 25, 26, 27, 28 e 29.



Legenda:

- Floresta
- Capoeira
- Pasto alto
- Pasto baixo
- Solo
- Água

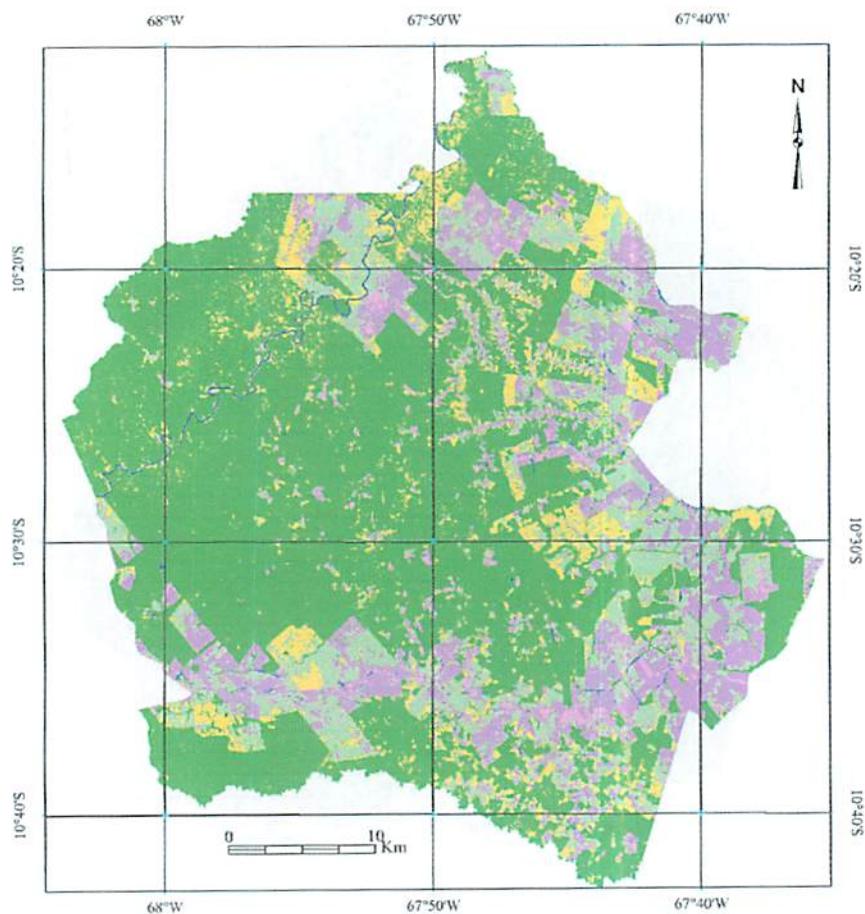
FIGURA 24 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método de máxima verossimilhança usando imagem Landsat, 2003



Legenda:

- Floresta
- Capoeira
- Pasto alto
- Pasto baixo
- Solo
- Água

FIGURA 25 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método isodata usando imagem Landsat, 2003



Legenda:

- Floresta
- Capoeira
- Pasto alto
- Pasto baixo
- Solo
- Água

FIGURA 26 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método árvore de decisão, 2003

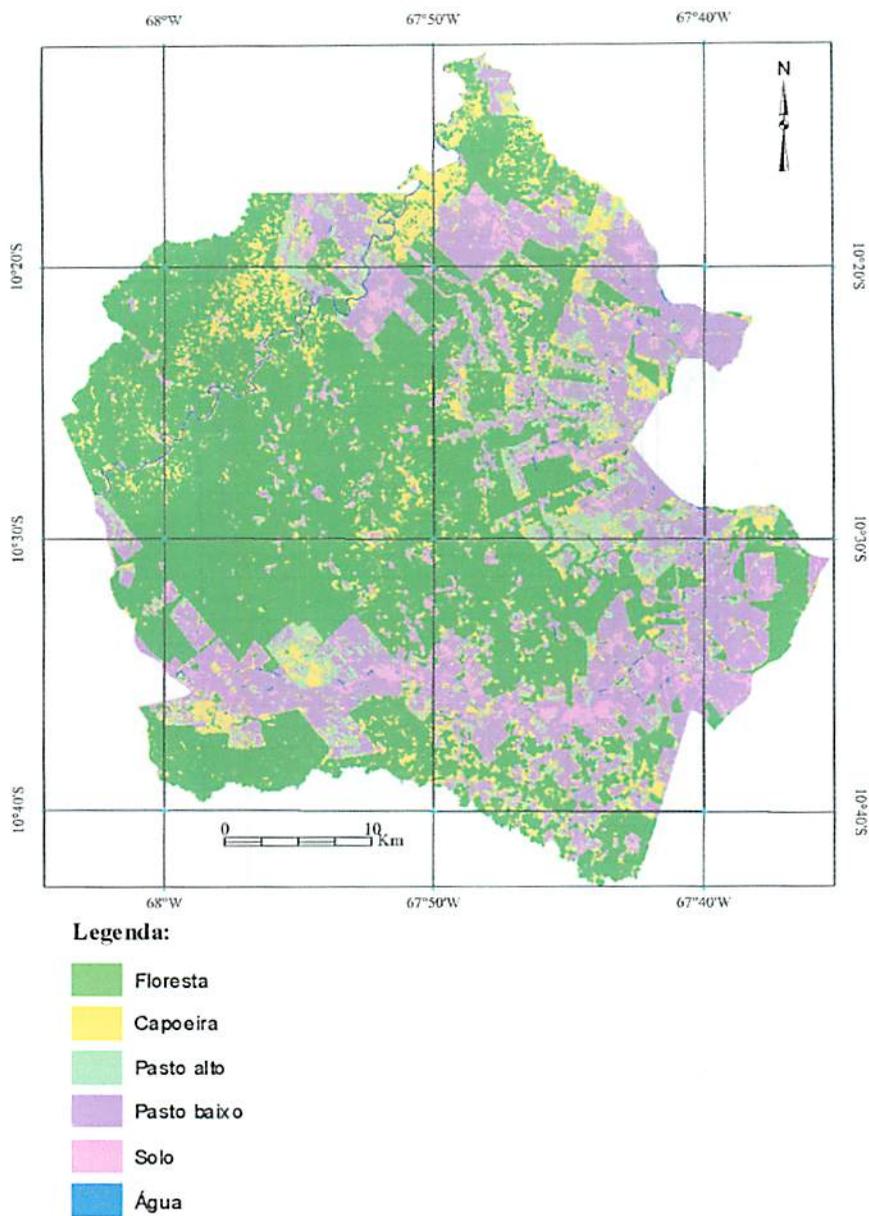


FIGURA 27 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método isodata usando as imagens fração solo e sombra, 2003

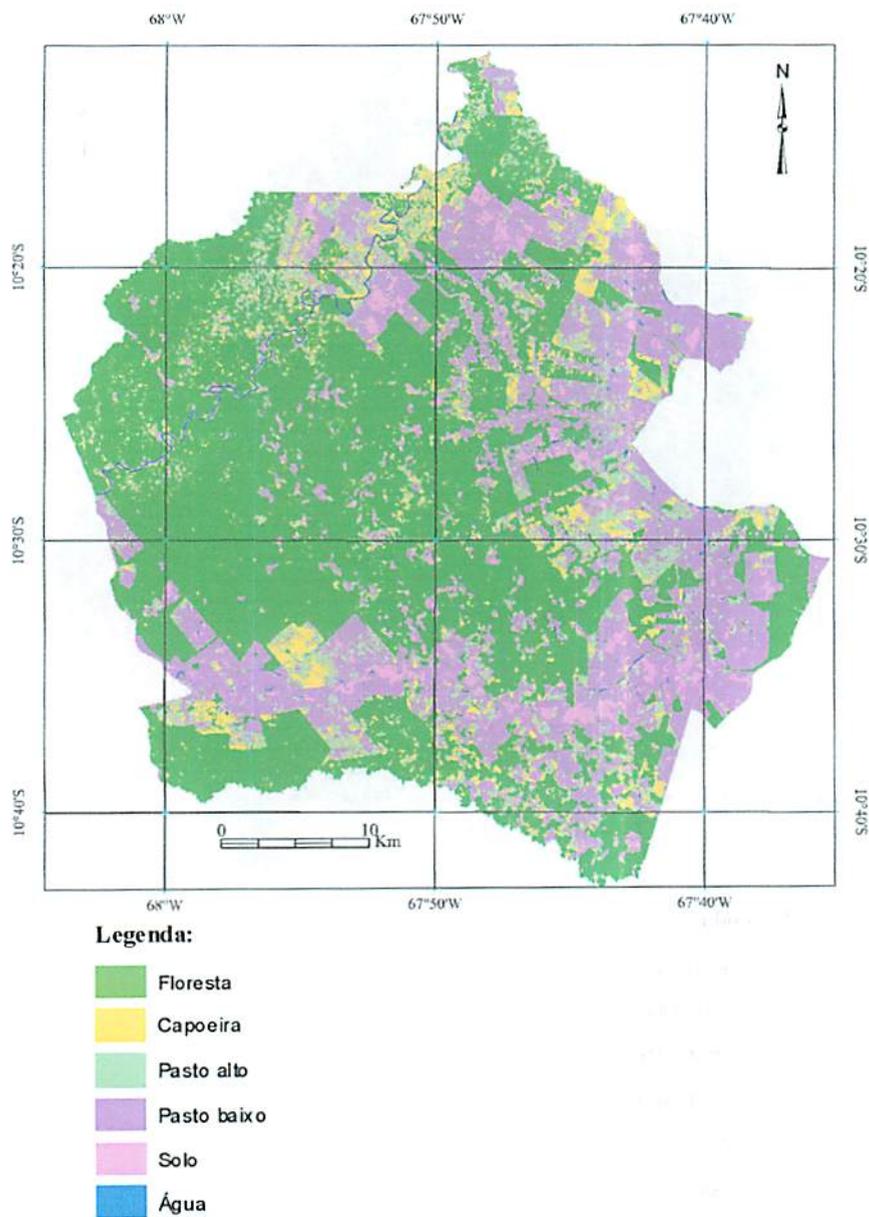


FIGURA 28 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método isodata usando as imagens fração solo, sombra e vegetação, 2003

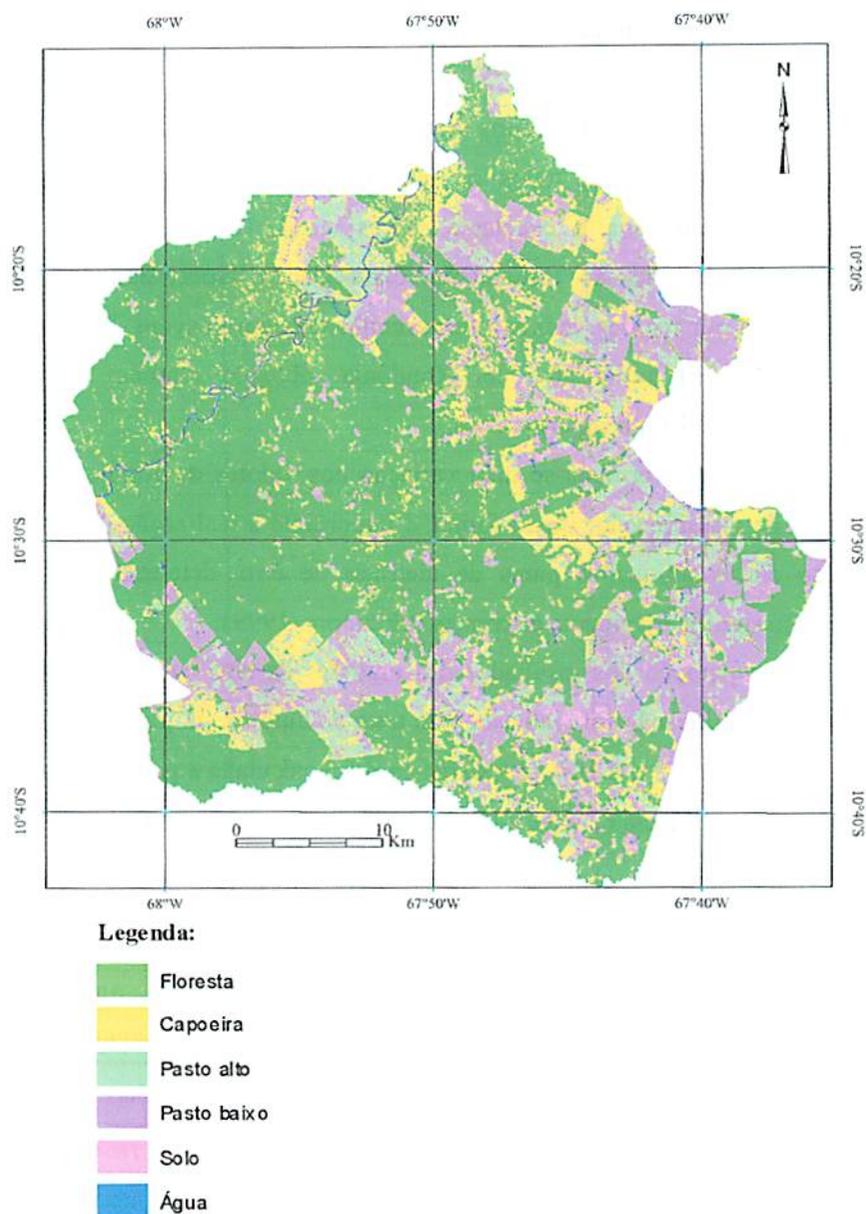


FIGURA 29 Mapa de uso e ocupação do solo em Capixaba, Acre, classificado pelo método de máxima verossimilhança usando as imagens fração solo, sombra, vegetação e de erro, 2003

4.6.3 Teste de significância

De acordo com os conceitos estabelecidos por Landis & Koch (1977) para o coeficiente Kappa, todos os métodos de classificação apresentam qualidade de muito boa a excelente. Alves et al. (1998) testaram cinco algoritmos de classificação (fatiamento, classificação supervisionada e não supervisionada por regiões, máxima verossimilhança com imagens TM/Landsat e imagens fração) numa área em Rondônia, obtendo também excelente classificação, com Kappa superior a 0,8.

No entanto, verificou-se uma variação nos valores de Kappa entre os algoritmos de classificação. Dessa forma, é importante a realização de um teste estatístico em combinações pares de matrizes de erro, determinando se as matrizes são significativamente diferentes (Skidmore, 1999).

A comparação dos resultados das matrizes de erro entre os diferentes métodos de classificação foi realizada pelo teste de hipótese Z, para os níveis de significância de 1%, 5% e 10%. A estatística Z foi calculada a partir dos valores de Kappa e sua variância (Tabela 14), obtidos das matrizes de erro.

TABELA 14 Análise de Kappa e variância de Kappa das matrizes de erros que resultaram dos métodos de classificação testados

Métodos de Classificação	Ano 2000		Ano 2002		Ano 2003	
	Kappa	σ^2 Kappa	Kappa	σ^2 Kappa	Kappa	σ^2 Kappa
Máxima verossimilhança (imagem Landsat)	0,9060	0,000170	0,8905	0,0002481	0,8875	0,000214
Isodata (imagem Landsat)	0,8225	0,000293	0,8457	0,0003359	0,7864	0,000367
Árvore de decisão	0,8247	0,000297	0,8601	0,0003099	0,9044	0,000184
Isodata (imagens solo e sombra)	0,7619	0,000384	0,7407	0,0005104	0,7969	0,000353
Isodata (imagens solo, sombra e vegetação)	0,7943	0,000341	0,8399	0,0003485	0,8050	0,000342
Máxima verossimilhança (imagens fração)	0,8919	0,000194	0,8853	0,0002595	0,8895	0,000210

Os valores tabelados de Z_{α} para a hipótese alternativa unilateral à direita, são os seguintes: $Z_{0,01} = 2,32634193$, $Z_{0,05} = 1,644853$ e $Z_{0,10} = 1,28155079$.

A Tabela 15 mostra o resultado da análise de Kappa entre as matrizes de erro dos classificadores testados em 2003.

TABELA 15 Comparação da análise de Kappa entre as matrizes de erro que resultaram dos métodos de classificação testados em 2003

Comparação	Estatística Z	Nível de confiança		
		99%	95%	90%
(1) vs (2)	4,1925	S	S	S
(1) vs (3)	0,8473	NS	NS	NS
(1) vs (4)	3,8063	S	S	S
(1) vs (5)	3,5005	S	S	S
(1) vs (6)	0,1000	NS	NS	NS
(2) vs (3)	5,0239	S	S	S
(2) vs (4)	0,3906	NS	NS	NS
(2) vs (5)	0,6975	NS	NS	NS
(2) vs (6)	4,2908	S	S	S
(3) vs (4)	4,6399	S	S	S
(3) vs (5)	4,3353	S	S	S
(3) vs (6)	0,7473	NS	NS	NS
(4) vs (5)	0,3071	NS	NS	NS
(4) vs (6)	3,9047	S	S	S
(5) vs (6)	3,5991	S	S	S

em que: (1) Máxima verossimilhança: imagem Landsat, (2) Isodata: imagem Landsat, (3) Árvore de decisão, (4) Isodata: imagem solo e sombra, (5) Isodata: imagem solo, sombra e vegetação, (6) Máxima verossimilhança: imagens fração solo, sombra, vegetação e de erro, (S) diferença significativa e (NS) diferença não significativa

Para o ano de 2003, foi indicada diferença significativa entre os métodos de classificação supervisionados (por máxima verossimilhança e árvore de decisão) e os métodos de classificação não supervisionados com o algoritmo isodata (usando as imagens Landsat ou imagens fração geradas pelo modelo linear de mistura espectral), para todos os níveis de confiança. Portanto, a

classificação supervisionada possui um desempenho superior em relação a classificação não supervisionada.

Ainda em 2003, os métodos de classificação supervisionados: máxima verossimilhança usando imagens Landsat, máxima verossimilhança usando as imagens fração da análise de mistura espectral e árvore de decisão foram idênticos, pois não houve diferença significativa entre eles.

O mesmo resultado também foi verificado comparando-se os métodos de classificação não supervisionados aplicados com o algoritmo isodata (usando imagem Landsat ou imagens fração oriundas do modelo linear de mistura espectral), ou seja, foram considerados idênticos, pois não apresentaram diferença significativa entre suas matrizes de erro.

No ano de 2002, os resultados da comparação entre as matrizes de erro dos métodos de classificação são mostrados na Tabela 16. Analisando seus resultados, pode-se concluir que, o método de classificação não supervisionado isodata usando as imagens fração solo e sombra foi significativamente inferior a todos os outros, para todos os níveis de confiança testados. O algoritmo de máxima verossimilhança foi significativamente melhor que os métodos usando o algoritmo isodata, com 95% de confiança.

O método de classificação isodata em 2002 (usando imagem Landsat e as imagens fração solo, sombra e vegetação) apresentou o mesmo desempenho obtido pelo algoritmo árvore de decisão, em todos os níveis de confiança, ao contrário do que ocorreu em 2003. No entanto, não houve diferença significativa entre os classificadores de máxima verossimilhança e de árvore de decisão, com 99% e 95% de confiança, nos anos de 2003 e 2002.

TABELA 16 Comparação da análise de Kappa entre as matrizes de erro que resultaram dos métodos de classificação testados em 2002

Comparação	Estatística Z	Nível de confiança		
		99%	95%	90%
(1) vs (2)	1,8565	NS	S	S
(1) vs (3)	1,2887	NS	NS	S
(1) vs (4)	5,4407	S	S	S
(1) vs (5)	2,0720	NS	S	S
(1) vs (6)	0,2339	NS	NS	NS
(2) vs (3)	0,5676	NS	NS	NS
(2) vs (4)	3,6084	S	S	S
(2) vs (5)	0,2196	NS	NS	NS
(2) vs (6)	1,6226	NS	NS	S
(3) vs (4)	4,1689	S	S	S
(3) vs (5)	0,7860	NS	NS	NS
(3) vs (6)	1,0548	NS	NS	NS
(4) vs (5)	3,3859	S	S	S
(4) vs (6)	5,2102	S	S	S
(5) vs (6)	1,8387	NS	S	S

em que: (1) Máxima verossimilhança: imagem Landsat, (2) Isodata: imagem Landsat, (3) Árvore de decisão, (4) Isodata: imagem solo e sombra, (5) Isodata: imagem solo, sombra e vegetação, (6) Máxima verossimilhança: imagens fração solo, sombra, vegetação e de erro, (S) diferença significativa e (NS) diferença não significativa

No entanto, em 2000, esse resultado não se repetiu, havendo uma superioridade do método de máxima verossimilhança (com imagem Landsat e com imagens fração oriundas do modelo linear de mistura espectral) em relação aos demais, para todos os níveis de significância. Esse resultado pode ser verificado na Tabela 17, na qual compara-se o desempenho das matrizes de erro entre os diferentes métodos de classificação em 2000.

TABELA 17 Comparação da análise de Kappa entre as matrizes de erro que resultaram dos métodos de classificação testados em 2000

Comparação	Estatística Z	Nível de confiança		
		99%	95%	90%
(1) vs (2)	3,8811	S	S	S
(1) vs (3)	3,7618	S	S	S
(1) vs (4)	6,1236	S	S	S
(1) vs (5)	4,9436	S	S	S
(1) vs (6)	0,7407	NS	NS	NS
(2) vs (3)	0,0915	NS	NS	NS
(2) vs (4)	2,3283	S	S	S
(2) vs (5)	1,1197	NS	NS	NS
(2) vs (6)	3,1438	S	S	S
(3) vs (4)	2,4067	S	S	S
(3) vs (5)	1,2042	NS	NS	NS
(3) vs (6)	3,0308	S	S	S
(4) vs (5)	1,2034	NS	NS	NS
(4) vs (6)	5,4068	S	S	S
(5) vs (6)	4,2199	S	S	S

em que: (1) Máxima verossimilhança: imagem Landsat, (2) Isodata: imagem Landsat, (3) Árvore de decisão, (4) Isodata: imagem solo e sombra, (5) Isodata: imagem solo, sombra e vegetação, (6) Máxima verossimilhança: imagens fração solo, sombra, vegetação e de erro, (S) diferença significativa e (NS) diferença não significativa

Oliveira (2004), obteve resultado diferente no mapeamento de fisionomias de cerrado, havendo mais eficiência do algoritmo de árvore de decisão em relação ao de máxima verossimilhança. Porém, no mapeamento de remanescentes de floresta atlântica realizado por Carvalho (2001), a classificação por máxima verossimilhança foi melhor que por árvore de decisão.

Em 2000, o método de classificação com o algoritmo isodata usando as imagens fração solo e sombra foi significativamente inferior a todos os demais, exceto quando comparado com o método isodata que incluiu a imagem fração vegetação gerada pelo modelo linear de mistura espectral.

5 CONCLUSÕES

Os resultados obtidos neste trabalho de comparação entre técnicas de classificação de imagens Landsat no mapeamento do uso e ocupação do solo em Capixaba, Acre, permitem as seguintes conclusões:

- os diferentes métodos de classificação com os algoritmos de máxima verossimilhança, isodata e árvore de decisão utilizando imagens Landsat e imagens derivadas de técnicas de extração de informações, são eficientes para o mapeamento do uso e ocupação do solo em Capixaba, pois apresentaram uma classificação muito boa a excelente;
- entre os métodos testados, a classificação por máxima verossimilhança foi superior às demais técnicas de classificação;
- a classificação por árvore de decisão demonstrou ser promissora no mapeamento de classes de uso e ocupação do solo, visto que em dois dos três anos analisados obteve eficiência idêntica à classificação por máxima verossimilhança;
- em relação à utilização das imagens fração geradas pela análise de mistura espectral para a classificação digital, o melhor resultado foi com o algoritmo de máxima verossimilhança;
- a utilização da imagem fração vegetação juntamente com as imagens fração solo e sombra na classificação não supervisionada isodata melhorou a exatidão global do mapeamento utilizando as imagens derivadas do modelo linear de mistura espectral;
- as classes floresta, água e solo obtiverem bom desempenho em todas as técnicas de classificação aplicadas;
- as classes de capoeira, pasto alto e pasto baixo apresentaram confusão entre os classificadores, devido à semelhança espectral entre elas;

- o mapeamento das classe capoeira e pasto alto pelo algoritmo de máxima verossimilhança foi mais eficiente, pois apresentou os menores erros de omissão e inclusão;
- o mapeamento da classe pasto baixo na classificação por máxima verossimilhança e por árvore de decisão obteve maior exatidão em relação ao algoritmo isodata;
- a classe queimada apresentou excelente acuracidade no mapeamento realizado com os algoritmos de máxima verossimilhança e árvore de decisão;
- a decisão sobre qual técnica de classificação representa melhor o mapeamento depende da importância relativa da categoria de uso e ocupação do solo que se deseja mapear, de acordo com os objetivos do trabalho a ser realizado.

6 REFERÊNCIAS BIBLIOGRÁFICAS

ACRE. Governo do Estado do Acre. Programa Estadual de Zoneamento Ecológico-Econômico do Estado do Acre. **Zoneamento ecológico-econômico: recursos naturais e meio ambiente – documento final**. Rio Branco: SECTMA, 2000a. v.1. 116p.

ACRE. Governo do Estado do Acre. Programa Estadual de Zoneamento Ecológico-Econômico do Estado do Acre. **Zoneamento ecológico-econômico: aspectos socioeconômicos e ocupação territorial – documento final**. Rio Branco: SECTMA, 2000b. v.2. 313p.

AGUIAR, A.P.D. **Utilização de atributos derivados de proporções de classes dentro de um elemento de resolução de imagem (“pixel”) na classificação multispectral de imagens de sensoriamento remoto**. 1991. 227p. Dissertação (Mestrado em Sensoriamento Remoto e Aplicações)-Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP.

ALVES, D.S. et al. **Análise comparativa de técnicas de classificação de imagens do sensor Landsat/TM para caracterização de áreas desflorestadas**. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 9., 1998, Santos. Anais... Santos: SBSR/INPE, 1998. Disponível em: <sid.inpe.br/deise/1999/02.09.09.16>. Acesso em: 20 nov. 2004.

ARAÚJO, H.J.B. de; SILVA, I.G. **Lista de espécies florestais do Acre: ocorrência com base em inventários florestais**. Rio Branco: Embrapa Acre, 2000. 77p. (Documentos, 48).

AURÉLIO, M.; VELLASCO, M.; LOPES, C.H. **Descoberta de conhecimento e mineração de dados**. Rio de Janeiro: PUC-Rio, 1999. 102p. Apostila

AYRES, M. et al. **BioEstat 2.0: aplicações estatísticas nas áreas das ciências biológicas e médicas**. Belém: Sociedade Civil Mamirauá; Brasília: CNPq, 2000. 272p.

BAJJOUK, T.; POPULUS J.; GUILLAUMONT, B. Quantification of subpixel cover fractions using principal component analysis and a linear programming method: application to the Coastal Zone of Roscoff (France). **Remote Sensing of Environmental**, v.64, p.153-165, 1998.

BECKER, B.K. Síntese do processo de ocupação da Amazônia: lições do passado e desafios do presente. In: BRASIL-MMA. **Causas e dinâmica do desmatamento na Amazônia**. Brasília: MMA, 2001. p.5-28.

BUITEN, H.J. General view of remote sensing as a source of information. In: BUITEN, H.J. & Clevers, J.G.P.W. **Land observation by remote sensing: theory and applications**. Amsterdam: Gordon and Breach Science, 1996. p.9-26.

CARVALHO, L.M.T. de. **Mapping and monitoring forest remnants: a multi-scale analysis of spatio-temporal data**. 2001. 138p. Thesis (Doctor in Remote Sensing)-Wageningen University, Wageningen.

CARVALHO, L.M.T. de et al. Selection of imagery data and classifiers for mapping brazilian semidecuiduos atlantic forests. **International Journal of Applied Earth Observation and Geoinformation**, n.5, p.173-186. 2004.

CARDIM, S.E. de C.S.; VIEIRA, P. de T.L.; VIEGAS, J.L.R. **Análise da estrutura fundiária brasileira**. Brasília: INCRA, 2004. Disponível em: <<http://incra.gov.br>>. Acesso em: 27 abr. 2004.

CHANDER, G.; MARKHAM, B. **Revised Landsat 5 TM radiometric calibration procedures and post-calibration dynamic ranges**. [S.l.]: USGS, [2003?]. Disponível em: <<http://landsat.usgs.gov/documents/L5TMCAL2003.pdf>> Acesso em: 11 out. 2004.

CLARKE, R.T.; BITTENCOURT, H.R. Uso de árvores de decisão na classificação de imagens digitais. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 2003, Belo Horizonte. **Anais... Belo Horizonte: SBSR/INPE, 2003. p.2043-2045.**

COLGATON, R.G.; GRENN, K. **Assessing the accuracy of remotely sensed data: principles and practices**. New York: Lewis, 1999. 137p.

COUTINHO, V.C. **Data mining**. Brasília: DwBrasil, 2003. Disponível em: <<http://www.dwbrasil.com.br/html/dmining.html>>. Acesso em: 07 jun. 2004.

CROSTA, A.P. **Processamento digital de imagens de sensoriamento remoto**. 3 ed. Campinas, SP: Unicamp/IG, 1993. 170p.

CURROTO, C.L. **Árvore de decisão**. Rio de Janeiro: COOPE/UFRJ, 2000. 59p. Relatório técnico. Disponível em: <www.curotto.com/doc>. Acesso em: 20 out. 2004.

DUARTE, V.; SHIMABUKURO, Y.E.; AUCILINO, L.C.M. Metodologia para atualizar e padronizar o banco de dados do projeto “PRODES digital”. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 2003, Belo Horizonte. Anais... Belo Horizonte: SBSR/INPE, 2003. p. 2705-2712.

ESCADA, M.I.S. Evolução de padrões da terra na região centro-norte de Rondônia. 2003. 264p. Tese (Doutorado em Sensoriamento Remoto)-Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP.

ESPÍRITO-SANTO, F.D.B. Caracterização e mapeamento da vegetação da região da Floresta Nacional do Tapajós através de dados óticos, radar e de inventários florestais. 2003. 277p. Dissertação (Mestrado em Sensoriamento Remoto)-Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP.

ESPÍRITO-SANTO, F.D.B.; SILVA, B.S.G.; SHIMABUKURO, Y.E. Detecção da dinâmica da floresta de bambu no sudeste do Acre com o uso de técnicas de processamento de imagens de satélite. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 2003, Belo Horizonte. Anais... Belo Horizonte: SBSR/INPE, 2003. p.649-656.

FIDALGO, L.C.C. Exatidão no processo de mapeamento temático da vegetação de uma área de mata atlântica no estado de São Paulo, a partir de imagens TM-Landsat. 1995. 184p. Dissertação (Mestrado em Sensoriamento Remoto e Aplicações)-Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP.

FERREIRA, E.M. et al. Uso do modelo linear de mistura espectral para o mapeamento sistemático e operacional do bioma cerrado: possibilidades, implicações e procedimentos metodológicos. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 2003, Belo Horizonte. Anais... Belo Horizonte: SBSR/INPE, 2003. p.657-664.

FONSECA, J.S.da; MARTINS, G.de A. Curso de estatística. 6.ed. São Paulo: Atlas, 1996. 320p.

FREITAS, R.M. de; NOVO, E.M de L.M.; SHIMABUKURO, Y.E. Mapeamento de ecossistemas alagáveis do rio Amazonas a partir do mosaico digital TM-Landsat – escala regional. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 2003, Belo Horizonte. Anais... Belo Horizonte: SBSR/INPE, 2003. p.2745-2752.

FUNDAÇÃO DE TECNOLOGIA DO ACRE. Monitoramento da cobertura florestal do Estado do Acre: desmatamento e uso atual da terra. Rio Branco, 1990. 214p.

GABOARDI, C. Utilização de imagem de coerência SAR para classificação do uso da terra: Floresta Nacional do Tapajós. 2002. 137p. Dissertação (Mestrado em Sensoriamento Remoto e Aplicações)-Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP.

GONZALEZ, R.C.; WOODS, R.E. Processamento de imagens digitais. Tradução de Roberto Marcondes Cesar Junior e Luciano da Fontoura Costa. São Paulo: Edgard Blücher, 2000. 509p. Título original: Digital image processing.

GRAÇA, P.M.L. de A. Procedimentos automatizados para monitoramento de áreas sob impacto da exploração madeireira. In: SENSORIAMENTO REMOTO E SISTEMAS DE INFORMAÇÕES GEOGRÁFICAS APLICADOS À ENGENHARIA FLORESTAL: seminário de atualização, 5., 2002, Curitiba. Anais... Curitiba: UFPR, 2002. p. 39-43.

GURGEL, H.da C. Variabilidade espacial e temporal do NDVI sobre o Brasil e suas conexões com o clima. 2000. 118p. Dissertação (Mestrado em Sensoriamento Remoto)-Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP.

HOMMA, A.K.O. História da agricultura na Amazônia: da era pré-colombiana ao terceiro milênio. Brasília: Embrapa Informação Tecnológica, 2003. 274p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Diretoria de Geociências. Manual técnico da vegetação brasileira. Rio de Janeiro, 1992. 92p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Cidades@. Rio de Janeiro, 2004. Disponível em: <<http://www.ibge.gov.br/cidadesats>>. Acesso em: 05 set. 2004.

INSTITUTO DE MEIO AMBIENTE DO ACRE. Núcleo de Cartografia. Atlas geográfico ambiental do Acre. Rio Branco, 1991. 48p.

INSTITUTO DO HOMEM E MEIO AMBIENTE DA AMAZÔNIA. Mapeamento da cobertura florestal do estado do Acre – 199p: relatório final. [Belém], 2001. 21p.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. TM/Landsat-7 ORB 2 P. 67. Data Pass. 29-09-2000. Imagem de satélite. Cachoeira Paulista, 2000. CD ROM.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. ETM⁺/Landsat-7 ORB 2 P. 67. Data Pass. 10-08-2002. Imagem de satélite. Cachoeira Paulista, 2002. CD ROM

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. TM/Landsat-7 ORB 2 P. 67. Data Pass. 04-07-2003. Imagem de satélite. Cachoeira Paulista, 2003. CD ROM.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. Coordenação-Geral de Observação da Terra. Banco de dados do projeto Prodes digital. Disponível em: <<http://www.obt.inpe.br/prodes>>. Acesso em: 25 set. 2004a.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. Coordenação-Geral de Observação da Terra. Banco de dados do projeto Prodes digital. São José dos Campos, 2004. Disponível em: <<http://www.obt.inpe.br/prodesdigital>>. Acesso em: 15 set. 2004b.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. Centro de Previsão de Tempo e Estudos Climáticos. Apresenta informações sobre o clima. São José dos Campos, 2004. Disponível em: <<http://www.cptec.inpe.br/clima/monit>>. Acesso em: 18 nov. 2004c.

KRUG, T. O quadro do desflorestamento da Amazônia. In: BRASIL-MMA. Causas e dinâmica do desmatamento na Amazônia. Brasília: MMA, 2001. p. 91-98.

LANDIS, J.; KOCH, G.G. The measurements of agreement for categorical data. Biometrics, v.33, n.3, p.159-179, mar. 1977.

LILLESAND, T.M.; KIEFER, R.W. Remote sensing and image interpretation. 4.ed. New York: J. Wiley & Sons, 2000. 724p.

LORENA, R.B.; SANTOS, J.R dos; SHIMABUKURO, Y.E. Evolução do uso da terra em porção da Amazônia ocidental (Acre), através da técnica de detecção de mudanças. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 2003, Belo Horizonte. Anais... Belo Horizonte: SBSR/INPE, 2003. p.673-680.

LU, D.; MORAN, E. BATISTELLA, M. Linear mixture model applied to Amazonian vegetation classification. *Remote Sensing of Environment*. v.87. p.456-469. 2003.

LUCAS, M.R. et al. Avaliação da composição florística, biomassa e estrutura de florestas tropicais em regeneração: a contribuição do sensoriamento remoto. In: INSTITUTO NACIONAL DE PESQUISA DA AMAZÔNIA. *Floresta Amazônica: dinâmica, regeneração e manejo*. Manaus: INPA, 1998. p.61-82.

MARKHAM, B.L.; BARKER, J.L. Radiometric properties of U.S. processes Landsat MSS data. *Remote Sensing of Environment*, v.17, p.39-71, 1987.

MASCARO, S. de A.; FERREIRA, M.C. Análise comparativa entre algoritmos de classificação digital com base na exatidão do mapeamento do uso e cobertura do solo: um exemplo na área de influência do reservatório de Jurumirim-SP. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 2003, Belo Horizonte. *Anais...* Belo Horizonte: SBSR/INPE, 2003. p.1365-1372.

MEER, F. van der. Physical principles of optical remote sensing. In: STEIN, A.; MEER, F. van der; GORTE, B. (Ed.). *Spatial statistics for remote sensing*. Dordrecht: Kluwer, 1999. p.27-40.

MEER, F. van der; JONG, S.M.de. Improving the results of spectral unmixing of Landsat Thematic Mapper imagery by enhancing the orthogonality of end-members. *International Journal of Remote Sensing*, v.21, n.15, p.2781-2797, 2000.

MELLO, E.M.K. et al. Técnicas de modelo de mistura espectral, segmentação e classificação de imagens TM-Landsat para o mapeamento do desflorestamento da Amazônia. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 2003, Belo Horizonte. *Anais...* Belo Horizonte: SBSR/INPE, 2003. p.2807-2814.

MOREIRA, M.A. *Fundamentos do sensoriamento remoto e metodologias de aplicação*. 2.ed. Viçosa: UFV, 2003. 307p.

NASA. United States Geological Survey. Apresenta informações sobre o Landsat 7. Disponível em: <http://ftpwww.gsfc.nasa.gov/IAS/handbook/handbook_htmls>. Acesso em: 11 out. 2004.

NOBRE, A.D. et al. Classificação e mapeamento preliminar da vegetação na região de Manaus, Amazonas, utilizando sensoriamento remoto passivo. In: INSTITUTO NACIONAL DE PESQUISA DA AMAZÔNIA. **Pesquisas florestais para a conservação da floresta e reabilitação de áreas degradadas da Amazônia**. Manaus: INPA, 1998. p. 27-49.

NOVO, E.M.L de M. **Sensoriamento remoto: princípios e aplicações**. 2.ed. São Paulo: Edgard Blucher, 1998. 308p.

OLIVEIRA, L.T. **Fusão de imagens de sensoriamento remoto e mineração de dados geográficos para mapear as fitofisionomias do bioma cerrado**. 2004. 131p. Dissertação (Mestrado em Manejo Ambiental)-Universidade Federal de Lavras, Lavras, MG.

PEREIRA, J.L.G.; BATISTA, G.T.; ROBERTS, D. Reflectância de coberturas vegetais na Amazônia. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 8., 1996, Salvador. **Anais...** Salvador: SBSR/INPE, 1996. p. 411-416.

—QUILAN, J.R. Induction of decision trees. **Machine learning**, Dordrecht, v.1, p.81-106, 1986.

ROBERTS, D.A.; KELLER, M.; SOARES, J.V. Studies of land-cover, land-use, and biophysical properties of vegetation in the Large Scale Biosphere Atmosphere experiment in Amazônia. **Remote Sensing of Environment**, v.87, p.377-388, 2003.

ROCHA, J.A. Comparação de métodos para estimativa de áreas desflorestadas: em estudo de caso no Acre. In: BRASIL-MMA. **Aplicações do sensoriamento remoto e sistemas de informações geográficas no monitoramento e controle do desmatamento da Amazônia brasileira**. Brasília: MMA/SCA/PPG7, 2002. p. 27-28.

RULEQUEST RESEARCH. **See5: an informal tutorial**. Sydney: Rulequest Research, 2004. 26p. Disponível em: <<http://www.rulequest.com/see5-win.html>>. Acesso em: 01 jun. 2004.

SAWYER, D. Evolução demográfica, qualidade de vida e desmatamento na Amazônia. In: **BRASIL-MMA. Causas e dinâmica do desmatamento na Amazônia.** Brasília: MMA, 2001. p.73-89.

SKIDMORE, A.K. Accuracy assessment of spatial information. In: **STEIN, A.; MEER, F. van der; GORTE, B. (Ed.). Spatial statistics for remote sensing.** Dordrecht: Kluwer, 1999. p.197-209.

SHIMABUKURO, Y.E. et al. Índice de vegetação e modelo linear de mistura espectral no monitoramento da região do pantanal. **Pesquisa Agropecuária Brasileira**, v.33, n. especial, p.1729-1737, out. 1998.

SHIMABUKURO, Y.E. et al. Levantamento de áreas desflorestadas na Amazônia através de processamento digital de imagens orbitais. **Floresta e Ambiente**, v.6, n.1, p.38-44, jan./dez. 1999.

SILVA, P.G. da. Detecção de mudanças em áreas de florestas tropicais sob exploração madeireira através da técnica análise do vetor de mudança. 2003. 111p. Dissertação (Mestrado em Sensoriamento Remoto)-Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP.

SMITS, P. C.; DELLEPIANE, S. G.; SHOWENGERT, R. A. Quality assessment of image classification algorithms for land-cover mapping: review and a proposal for a cost-based approach. **International Journal of Remote Sensing**, v.20, n.8, p.1461-1486. 1999.

SULSOFT. Guia do envi em português: ENVI 4.0. [Porto Alegre]: [s.n.], [2004]. Disponível em: <<http://www.envi.com.br>>. Acesso em: 04 maio 2004.

SWAIN, P.H.; HAUSKA, H. The decision tree classifier: desing and potential. **IEEE Transactions on geoscience electronics**, v.ge-15, n.3, p.142-147, jul. 1977.

TSO, B.; MATHER, P.M. Classification methods for remotely sensed data. London: Taylor & Francis, 2001. 332p.

ANEXOS

ANEXO A		Página
TABELA 1A	Valores de $L_{min\lambda}$ e $L_{max\lambda}$ em $W/(m^2.sr.\mu m)$ utilizados na calibração radiométrica das imagens TM/Landsat 5	115
TABELA 2A	Valores de $L_{min\lambda}$ e $L_{max\lambda}$ em $W/(m^2.sr.\mu m)$ utilizados na calibração radiométrica das imagens ETM+/Landsat 7	115
TABELA 3A	Valores de irradiância solar exoatmosférica média em $mW\ cm^{-2}\ \mu m^{-1}$	115

TABELA 1A Valores de $L_{min\lambda}$ e $L_{max\lambda}$ em $W/(m^2.sr.\mu m)$ utilizados na calibração radiométrica das imagens TM/Landsat 5

Data do processamento	De 01/03/1984 a 04/03/2003		Depois de 05/03/2003	
	L_{min}	L_{max}	L_{min}	L_{max}
1	-1,52	152,10	-1,52	193,00
2	-2,84	296,81	-2,84	365,00
3	-1,17	204,30	-1,17	264,00
4	-1,51	206,20	-1,51	221,00
5	-0,37	27,19	-0,37	30,20
7	-0,15	14,38	-0,15	16,50

Fonte: Chander & Markham (2003)

TABELA 2A Valores de $L_{min\lambda}$ e $L_{max\lambda}$ em $W/(m^2.sr.\mu m)$ utilizados na calibração radiométrica das imagens ETM+/Landsat 7

Data do processamento	Depois de 1/7/2000	
	L_{min}	L_{max}
1	-6,2	191,6
2	-6,4	196,5
3	-5,0	152,9
4	-5,1	157,4
5	-1,0	31,06
7	-0,35	10,8

Fonte: NASA (2004)

TABELA 3A Valores de irradiância solar exoatmosférica média em $mW\ cm^{-2}\ \mu m^{-1}$

Banda	Satélite	
	TM Landsat 5	ETM+ Landsat 7
1	195,7	196,9
2	182,9	184,0
3	155,7	155,1
4	104,7	104,4
5	21,93	22,57
7	7,452	8,207

Fonte: Chander & Markham (2003); NASA (2004)

TABELA 1B	Lista das espécies florestais da fisionomia floresta aberta com palmeiras e floresta densa	117
-----------	--	-----

TABELA 1B Lista das espécies florestais da fisionomia floresta aberta com palmeiras e floresta densa

<i>Nome vulgar</i>	<i>Nome científico</i>	<i>Família</i>
Ubim	<i>Geonoma diversa</i>	Arecaceae
Ubim-galope	<i>Geonoma caulis</i> Mart.	Arecaceae
Açai	<i>Euterpe precatória</i> Mart.	Arecaceae
Patauá	<i>Oenocarpus bataua</i> Mart.	Arecaceae
Marajá-de-terra-firme	<i>Bactris</i> sp.	Arecaceae
Bacaba	<i>Oenocarpus bacaba</i> M.	Arecaceae
Paxiubão	<i>Iriartea deltoidea</i> Ruiz & Pav	Arecaceae
Paxiubinha	<i>Socratea exorrhiza</i> Mart	Arecaceae
Murmuru	<i>Astrocaryum murumuru</i> Mart.	Arecaceae
Uricuri	<i>Attalea phalerata</i> Mart.	Arecaceae
Jarina	<i>Phytelephas macrocarpa</i> Ruiz & Pav.	Arecaceae
Amarelão	<i>Aspidosperma vargasii</i> A. DC.	Apocynaceae
Caucho	<i>Castilla ulei</i> Warburg.	Moraceae
Cedro-vermelho	<i>Cedrela odorata</i> L.	Meliaceae
Cerejeira	<i>Torresea acreana</i> Ducke	Fabaceae
Maçaranduba	<i>Manilkara surinamensis</i> (Miq.) Dub.	Sapotaceae
Marupá	<i>Jacaranda copaia</i> (Aubl.) D. Don	Bignoniaceae
Matamatá	<i>Eschweilera odorata</i> (Poepp.) Miers.	Lecythidaceae
Pau-d'arco-amarelo	<i>Tabebuia serratifolia</i> (Vahl.) Nichols.	Bignoniaceae
Cumaru-ferro	<i>Dipteryx ferrea</i> Ducke	Fabaceae
Jutai	<i>Hymenaea oblongifolia</i> Huber	Caesalpiniaceae

Fonte: Acre (2000a); Araújo & Silva (2000).

TABELA 1C	Matriz de erro do classificador máxima verossimilhança em 2000.....	119
TABELA 2C	Matriz de erro do classificador máxima verossimilhança em 2002.....	119
TABELA 3C	Matriz de erro do classificador máxima verossimilhança em 2003.....	120
TABELA 4C	Matriz de erro do classificador isodata em 2000.....	120
TABELA 5C	Matriz de erro do classificador isodata em 2002.....	121
TABELA 6C	Matriz de erro do classificador isodata em 2003.....	121
TABELA 7C	Matriz de erro do classificador árvore de decisão em 2000.....	122
TABELA 8C	Matriz de erro do classificador árvore de decisão em 2002.....	122
TABELA 9C	Matriz de erro do classificador árvore de decisão em 2003.....	123
TABELA 10C	Matriz de erro do classificador isodata usando imagem fração solo e sombra em 2000.....	123
TABELA 11C	Matriz de erro do classificador isodata usando imagem fração solo e sombra em 2002.....	124
TABELA 12C	Matriz de erro do classificador isodata usando imagem fração solo e sombra em 2003.....	124
TABELA 13C	Matriz de erro do classificador isodata usando imagem fração solo, sombra e vegetação em 2000...	125
TABELA 14C	Matriz de erro do classificador isodata usando imagem fração solo, sombra e vegetação em 2002...	125
TABELA 15C	Matriz de erro do classificador isodata usando imagem fração solo, sombra e vegetação em 2003...	126
TABELA 16C	Matriz de erro do classificador máxima verossimilhança usando imagens fração em 2000.....	126
TABELA 17C	Matriz de erro do classificador máxima verossimilhança usando imagens fração em 2002.....	127
TABELA 18C	Matriz de erro do classificador máxima verossimilhança usando imagens fração em 2003.....	127

TABELA 1C Matriz de erro do classificador máxima verossimilhança em 2000

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	76	6	0	0	14	0	0	96	20,8%
Floresta (FL)	2	229	0	0	0	0	0	231	0,9%
Solo (SO)	0	0	68	0	0	0	0	68	0,0%
Água (AG)	1	0	0	62	0	0	0	63	1,6%
Pasto alto (PA)	4	0	0	0	51	8	0	63	19,0%
Pasto baixo (PB)	0	0	11	0	0	69	0	80	13,8%
Queimada (QU)	0	0	2	0	0	0	38	40	5,0%
Total	83	235	81	62	65	77	38	641	
Omissão (%)	8,4%	2,6%	16,0%	0,0%	21,5%	10,4%	0,0%		

Exatidão global = 92,51% Kappa = 0,9060

TABELA 2C Matriz de erro do classificador máxima verossimilhança em 2002

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	57	1	1	0	16	0	0	75	24,0%
Floresta (FL)	2	163	0	0	0	0	0	165	1,2%
Solo (SO)	0	0	43	0	0	14	0	57	24,6%
Água (AG)	1	0	0	60	0	0	0	61	1,6%
Pasto alto (PA)	0	0	0	0	32	1	0	33	3,0%
Pasto baixo (PB)	0	0	6	0	2	35	0	43	18,6%
Queimada (QU)	0	0	0	0	0	0	60	60	0,0%
Total	60	164	50	60	50	50	60	494	
Omissão (%)	5,0%	0,6%	14,0%	0,0%	36,0%	30,0%	0,0%		

Exatidão global = 91,09% Kappa = 0,8905

TABELA 3C Matriz de erro do classificador máxima verossimilhança em 2003

Classificação	Amostras de amostra de validação (pixels)						Total	Inclusão (%)	
	CA	FL	SO	AG	PA	PB			
Capoeira (CA)	71	1	0	0	6	0	78	9,0%	
Floresta (FL)	9	194	0	0	0	0	203	4,4%	
Solo (SO)	0	0	28	0	0	0	28	0,0%	
Água (AG)	0	0	0	77	0	0	77	0,0%	
Pasto alto (PA)	3	0	0	0	89	11	103	13,6%	
Pasto baixo (PB)	0	0	24	0	0	91	115	20,9%	
Total	83	195	52	77	95	102	604		
Omissão (%)	14,5%	0,5%	46,2%	0,0%	6,3%	10,8%			
Exatidão global = 91,06%		Kappa = 0,8875							

TABELA 4C Matriz de erro do classificador isodata em 2000

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	58	4	0	0	0	0	0	62	6,5%
Floresta (FL)	5	204	0	0	0	0	0	209	2,4%
Solo (SO)	0	0	78	0	0	2	0	80	2,5%
Água (AG)	0	27	0	62	0	0	0	89	30,3%
Pasto alto (PA)	14	0	0	0	65	31	0	110	40,9%
Pasto baixo (PB)	0	0	3	0	0	44	0	47	6,4%
Queimada (QU)	6	0	0	0	0	0	38	44	13,6%
Total	83	235	81	62	65	77	38	641	
Omissão (%)	30,1%	13,2%	3,7%	0,0%	0,0%	42,9%	0,0%		
Exatidão global = 85,65%		Kappa = 0,8225							

TABELA 5C Matriz de erro do classificador isodata em 2002

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	55	2	0	0	7	0	0	64	14,1%
Floresta (FL)	5	162	0	0	0	0	0	167	3,0%
Solo (SO)	0	0	50	0	0	0	0	50	0,0%
Água (AG)	0	0	0	60	0	0	45	105	42,9%
Pasto alto (PA)	0	0	0	0	43	3	0	46	6,5%
Pasto baixo (PB)	0	0	0	0	0	47	0	47	0,0%
Queimada (QU)	0	0	0	0	0	0	15	15	0,0%
Total	60	164	50	60	50	50	60	494	
Omissão (%)	8,3%	1,2%	0,0%	0,0%	14,0%	6,0%	75,0%		
Exatidão global = 87,55%		Kappa = 0,8457							

TABELA 6C Matriz de erro do classificador isodata em 2003

Classificação	Amostras de amostra de validação (pixels)						Total	Inclusão (%)	
	CA	FL	SO	AG	PA	PB			
Capoeira (CA)	29	3	0	0	9	0	41	29,3%	
Floresta (FL)	6	192	0	2	0	0	200	4,0%	
Solo (SO)	0	0	51	0	0	11	62	17,7%	
Água (AG)	0	0	0	72	0	0	72	0,0%	
Pasto alto (PA)	48	0	0	3	80	14	145	44,8%	
Pasto baixo (PB)	0	0	1	0	6	77	84	8,3%	
Total	83	195	52	77	95	102	604		
Omissão (%)	65,1%	1,5%	1,9%	6,5%	15,8%	24,5%			
Exatidão global = 82,95%		Kappa = 0,7864							

TABELA 7C Matriz de erro do classificador de árvore de decisão em 2000

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	73	3	0	0	39	0	0	115	36,5%
Floresta (FL)	9	232	0	1	0	0	0	242	4,1%
Solo (SO)	0	0	58	0	0	0	0	58	0,0%
Água (AG)	0	0	0	61	0	0	0	61	0,0%
Pasto alto (PA)	0	0	0	0	26	9	0	35	25,7%
Pasto baixo (PB)	1	0	19	0	0	68	0	88	21,6%
Queimada (QU)	0	0	4	0	0	0	38	42	9,5%
Total	83	235	81	62	65	77	38	641	
Omissão (%)	12,0%	1,3%	28,4%	1,6%	60,0%	11,7%	0,0%		

Exatidão global = 86,74% Kappa = 0,8325

TABELA 8C Matriz de erro do classificador árvore de decisão em 2002

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	49	0	1	0	17	0	0	67	26,9%
Floresta (FL)	10	163	0	0	0	0	0	173	5,8%
Solo (SO)	1	0	40	0	0	14	0	55	27,3%
Água (AG)	0	1	0	60	0	0	0	61	1,6%
Pasto alto (PA)	0	0	0	0	32	2	0	34	5,9%
Pasto baixo (PB)	0	0	9	0	1	34	0	44	22,7%
Queimada (QU)	0	0	0	0	0	0	60	60	0,0%
Total	60	164	50	60	50	50	60	494	
Omissão (%)	18,3%	0,6%	20,0%	0,0%	36,0%	32,0%	0,0%		

Exatidão global = 88,66% Kappa = 0,8601

TABELA 9C Matriz de erro do classificador de árvore de decisão em 2003

Classificação	Amostras de amostra de validação (pixels)						Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB		
Capoeira (CA)	69	1	0	0	5	0	75	8,0%
Floresta (FL)	11	194	0	0	0	0	205	5,4%
Solo (SO)	0	0	44	0	0	1	45	2,2%
Água (AG)	0	0	0	77	0	0	77	0,0%
Pasto alto (PA)	3	0	0	0	90	17	110	18,2%
Pasto baixo (PB)	0	0	8	0	0	84	92	8,7%
Total	83	195	52	77	95	102	604	
Omissão (%)	16,9%	0,5%	15,4%	0,0%	5,3%	17,6%		
Exatidão global = 92,38%		Kappa = 0,9044						

TABELA 10C Matriz de erro do classificador isodata usando imagem fração solo e sombra em 2000

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	56	2	0	9	9	0	1	77	27,3%
Floresta (FL)	17	233	0	0	0	0	0	250	6,8%
Solo (SO)	0	0	54	0	0	0	1	55	1,8%
Água (AG)	0	0	0	53	0	0	0	53	0,0%
Pasto alto (PA)	9	0	0	0	54	6	4	73	26,0%
Pasto baixo (PB)	1	0	27	0	2	71	32	133	46,6%
Queimada (QU)	0	0	0	0	0	0	0	0	-
Total	83	235	81	62	65	77	38	641	
Omissão (%)	32,5%	0,9%	33,3%	14,5%	16,9%	7,8%	100%		
Exatidão global = 81,28%		Kappa = 0,7619							

TABELA 11C Matriz de erro do classificador isodata usando imagem fração solo e sombra em 2002

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capocira (CA)	40	0	0	4	6	0	0	50	20,0%
Floresta (FL)	12	164	0	1	0	0	0	177	7,3%
Solo (SO)	0	0	43	0	0	0	0	43	0,0%
Água (AG)	0	0	0	53	0	0	0	53	0,0%
Pasto alto (PA)	8	0	0	1	41	1	24	75	45,3%
Pasto baixo (PB)	0	0	7	1	3	49	36	96	49,0%
Queimada (QU)	0	0	0	0	0	0	0	0	-
Total	60	164	50	60	50	50	60	494	
Omissão (%)	33,3%	0,0%	14,0%	11,7%	18,0%	2,0%	100%		
Exatidão global = 78,95%		Kappa = 0,7407							

TABELA 12C Matriz de erro do classificador isodata usando imagem fração solo e sombra em 2003

Classificação	Amostras de amostra de validação (pixels)						Total	Inclusão (%)	
	CA	FL	SO	AG	PA	PB			
Capocira (CA)	52	4	0	0	2	0	58	10,3%	
Floresta (FL)	3	190	0	1	0	0	194	2,1%	
Solo (SO)	0	0	50	0	0	6	56	10,7%	
Água (AG)	0	0	0	76	0	0	76	0,0%	
Pasto alto (PA)	28	1	0	0	42	0	71	40,8%	
Pasto baixo (PB)	0	0	2	0	51	96	149	35,6%	
Total	83	195	52	77	95	102	604		
Omissão (%)	37,3%	2,6%	3,8%	1,3%	55,8%	5,9%			
Exatidão global = 83,77%		Kappa = 0,7969							

TABELA 13C Matriz de erro do classificador isodata usando imagem fração solo, sombra e vegetação em 2000

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	44	1	0	1	16	0	0	62	29,0%
Floresta (FL)	18	234	0	0	0	0	0	252	7,1%
Solo (SO)	0	0	56	0	0	0	1	57	1,8%
Água (AG)	0	0	0	53	0	0	0	53	0,0%
Pasto alto (PA)	15	0	0	0	46	3	0	64	28,1%
Pasto baixo (PB)	6	0	23	5	3	74	7	118	37,3%
Queimada (QU)	0	0	2	3	0	0	30	35	14,3%
Total	83	235	81	62	65	77	38	641	
Omissão (%)	47,0%	0,4%	30,9%	14,5%	29,2%	3,9%	21,1%		

Exatidão global = 83,78% Kappa = 0,7943

TABELA 14C Matriz de erro do classificador isodata usando imagem fração solo, sombra e vegetação em 2002

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	32	0	0	0	19	0	0	51	37,3%
Floresta (FL)	12	164	0	1	0	0	0	177	7,3%
Solo (SO)	0	0	43	0	0	0	0	43	0,0%
Água (AG)	0	0	0	54	0	0	0	54	0,0%
Pasto alto (PA)	16	0	0	0	31	4	0	51	39,2%
Pasto baixo (PB)	0	0	7	5	0	46	0	58	20,7%
Queimada (QU)	0	0	0	0	0	0	60	60	0,0%
Total	60	164	50	60	50	50	60	494	
Omissão (%)	46,7%	0,0%	14,0%	10,0%	38,0%	8,0%	0,0%		

Exatidão global = 87,04% Kappa = 0,8399

TABELA 15C Matriz de erro do classificador isodata usando imagem fração solo, sombra e vegetação em 2003

Classificação	Amostras de amostra de validação (pixels)						Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB		
Capoeira (CA)	43	2	0	0	8	0	53	18,9%
Floresta (FL)	4	190	0	1	0	0	195	2,6%
Solo (SO)	0	0	50	0	0	3	53	5,7%
Água (AG)	0	0	0	72	0	0	72	0,0%
Pasto alto (PA)	32	3	0	0	56	0	91	38,5%
Pasto baixo (PB)	4	0	2	4	31	99	140	29,3%
Total	83	195	52	77	95	102	604	
Omissão (%)	48,2%	2,6%	3,8%	6,5%	41,1%	2,9%		
Exatidão global = 84,44%		Kappa = 0,8050						

TABELA 16C Matriz de erro do classificador máxima verossimilhança usando imagens fração em 2000

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	57	2	0	0	2	0	0	61	6,6%
Floresta (FL)	9	233	0	0	0	0	0	242	3,7%
Solo (SO)	0	0	80	0	0	5	0	85	5,9%
Água (AG)	0	0	0	62	0	0	0	62	0,0%
Pasto alto (PA)	9	0	0	0	61	17	0	87	29,9%
Pasto baixo (PB)	7	0	0	0	2	55	0	64	14,1%
Queimada (QU)	1	0	1	0	0	0	38	40	5,0%
Total	83	235	81	62	65	77	38	641	
Omissão (%)	31,3%	0,9%	1,2%	0,0%	6,2%	28,6%	0,0%		
Exatidão global = 91,42%		Kappa = 0,8919							

TABELA 17C Matriz de erro do classificador máxima verossimilhança usando imagens fração em 2002

Classificação	Amostras de amostra de validação (pixels)							Total	Inclusão (%)
	CA	FL	SO	AG	PA	PB	QU		
Capoeira (CA)	53	1	1	0	15	0	0	70	24,3%
Floresta (FL)	7	163	0	0	0	0	0	170	4,1%
Solo (SO)	0	0	42	0	0	11	0	53	20,8%
Água (AG)	0	0	0	60	0	0	0	60	0,0%
Pasto alto (PA)	0	0	0	0	32	1	0	33	3,0%
Pasto baixo (PB)	0	0	7	0	3	38	0	48	20,8%
Queimada (QU)	0	0	0	0	0	0	60	60	0,0%
Total	60	164	50	60	50	50	60	494	
Omissão (%)	11,7%	0,6%	16,0%	0,0%	36,0%	24,0%	0,0%		
Exatidão global = 90,69%		Kappa = 0,8853							

TABELA 18C Matriz de erro do classificador máxima verossimilhança usando imagens fração em 2003

Classificação	Amostras de amostra de validação (pixels)						Total	Inclusão (%)	
	CA	FL	SO	AG	PA	PB			
Capoeira (CA)	71	1	0	0	5	0	77	7,8%	
Floresta (FL)	9	194	0	0	0	0	203	4,4%	
Solo (SO)	0	0	27	0	0	0	27	0,0%	
Água (AG)	0	0	0	77	0	0	77	0,0%	
Pasto alto (PA)	3	0	0	0	90	10	103	12,6%	
Pasto baixo (PB)	0	0	25	0	0	92	117	21,4%	
Total	83	195	52	77	95	102	604		
Omissão (%)	14,5%	0,5%	48,1%	0,0%	5,3%	9,8%			
Exatidão global = 91,06%		Kappa = 0,8875							

