

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA AGRÍCOLA

APLICAÇÃO DE LÓGICA *FUZZY* PARA ESTIMATIVA DE
ÁREA PLANTADA DA CULTURA DE SOJA UTILIZANDO
IMAGENS AVHRR-NOAA

JOÃO FRANCISCO GONÇALVES ANTUNES

CAMPINAS - SP
SETEMBRO DE 2005

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA AGRÍCOLA

APLICAÇÃO DE LÓGICA *FUZZY* PARA ESTIMATIVA DE
ÁREA PLANTADA DA CULTURA DE SOJA UTILIZANDO
IMAGENS AVHRR-NOAA

Dissertação apresentada à banca examinadora
para obtenção do título de Mestre em Engenharia
Agrícola na área de concentração de Planejamento
e Desenvolvimento Rural Sustentável.

AUTOR: JOÃO FRANCISCO GONÇALVES ANTUNES

Orientador: Prof. Dr. Jurandir Zullo Júnior

CAMPINAS - SP
SETEMBRO DE 2005

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

An89a	<p>Antunes, João Francisco Gonçalves</p> <p>Aplicação de lógica <i>fuzzy</i> para estimativa de área plantada da cultura de soja utilizando imagens AVHRR-NOAA / João Francisco Gonçalves Antunes. --Campinas, SP: [s.n.], 2005.</p> <p>Orientador: Jurandir Zullo Júnior.</p> <p>Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Agrícola.</p> <p>1. Agricultura - Estatística. 2. Sensoriamento remoto. 3. Reconhecimento de padrões. 4. Vegetação - Classificação. 5. Lógica difusa. I. Zullo Júnior, Jurandir. II. Universidade Estadual de Campinas. Faculdade de Engenharia Agrícola. III. Título.</p>
-------	--

Titulo em inglês: Application of fuzzy logic for soybean crop area estimation using AVHRR-NOAA images

Palavras-chave em inglês: Harvest forecast; Remote sensing; NDVI; NDMI; Pattern recognition; Image classification; Fuzzy set; Fuzzy logic

Área de concentração: Planejamento e Desenvolvimento Rural Sustentável

Titulação: Mestre em Engenharia Agrícola

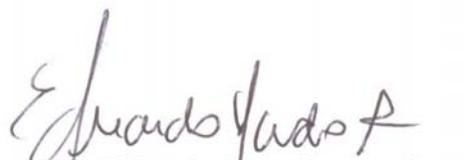
Banca examinadora: Eduardo Delgado Assad, Rubens Augusto de Camargo Lamparelli e Luiz Henrique Antunes Rodrigues

Data da defesa: 01 de setembro de 2005.

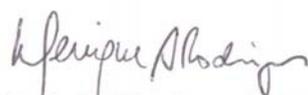
Este exemplar corresponde a redação final da Dissertação de Mestrado defendida por
João Francisco Gonçalves Antunes, aprovada pela Comissão Julgadora em
01 de setembro de 2005.

Campinas, 03 de outubro de 2005


Prof. Dr. Jurandir Zullo Júnior - Presidente
CEPAGRI/UNICAMP


Dr. Eduardo Delgado Assad
Embrapa Informática Agropecuária


Prof. Dr. Rubens Augusto de Camargo Lamparelli
CEPAGRI/UNICAMP


Prof. Dr. Luiz Henrique Antunes Rodrigues
FEAGRI/UNICAMP

Dedico
à minha mulher GLÁUCIA e
nosso filhinho GABRIEL,
com muito amor e carinho.

AGRADECIMENTOS

À Empresa Brasileira de Pesquisa Agropecuária por essa oportunidade de capacitação profissional.

Ao Prof. Dr. Jurandir Zullo Jr. pela orientação e ao Dr. Eduardo Delgado Assad pelas reuniões técnicas que auxiliaram no desenvolvimento do trabalho.

Ao Prof. Dr. Fernando José Von Zuben da Faculdade de Engenharia Elétrica e Computação pelo apoio oferecido no início do mestrado, o que me fez acreditar no potencial do trabalho.

À Faculdade de Engenharia Agrícola, em nome do Prof. Dr. Jansle Vieira Rocha, pela utilização da infra-estrutura do Laboratório de Geoprocessamento, com apoio técnico de Agmon Rocha. E, também, pelo suporte computacional de André A. Nogueira do Laboratório de Informática.

Ao Prof. Dr. Rubens Augusto de Camargo Lamparelli do Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura pela ajuda prestada na área de sensoriamento remoto e geoprocessamento para a elaboração do plano de trabalho inicial.

Ao Colorado Center for Astrodynamics Research, da Universidade do Colorado, EUA, nas pessoas do Dr. Daniel G. Baldwin e do Dr. William Emery por repassarem a tecnologia de software para o processamento das imagens AVHRR-NOAA utilizada neste trabalho.

Ao Jülich Research Center, Alemanha, na pessoa do Dr. Morton J. Canty pelo repasse da tecnologia de software para classificação não-supervisionada *fuzzy* utilizada neste trabalho.

Aos meus grandes amigos Edmundo Capelas de Oliveira e José Gilberto Jardine pela força que sempre me passaram para cursar a pós-graduação.

Aos meus amigos de turma, em especial a Erivelto Mercante e Júlio C. D. M. Esquerdo pela amizade e pelas valiosas discussões técnicas, sugestões e publicações. Juntamente com Luis A. Nery, Fernando A. Benvenuti, Liane Weber, Leonardo S. de Lourenço, Pedro P. B. Pontes, Juliana R. Ganan, Lídia S. Bertolo e Enéias C. Tonhasca Jr., formamos um verdadeiro grupo de pesquisa, fazendo do Laboratório de Geoprocessamento um ambiente sempre agradável e descontraído para se trabalhar.

À secretária da pós-graduação Ana Paula Montagner pela disponibilidade ininterrupta de atendimento aos detalhes administrativos do curso sempre realizado de forma bem-humorada e pela amizade.

À minha mulher Gláucia pela compreensão e apoio em tantas ocasiões.

Aos meus pais Jacinto e Benice por terem sempre me incentivado nos estudos.

E a Deus pela vida.

SUMÁRIO

LISTA DE TABELAS	ix
LISTA DE FIGURAS.....	x
RESUMO.....	xii
ABSTRACT	xiii
1. INTRODUÇÃO	1
1.1. Problema.....	3
1.2. Hipótese.....	3
1.3. Objetivos.....	3
2. REVISÃO BIBLIOGRÁFICA	5
2.1. A cultura da soja.....	5
2.2. A previsão de safras.....	8
2.3. O sensoriamento remoto e o uso na agricultura	10
2.3.1. Comportamento espectral da vegetação	11
2.3.2. Índices de vegetação.....	13
2.3.3. Satélites NOAA	15
2.3.4. Série multitemporais de imagens.....	17
2.4. O processamento das imagens AVHRR-NOAA	18
2.4.1. Conversão do formato bruto	18
2.4.2. Calibração radiométrica.....	19
2.4.3. Georreferenciamento preciso.....	19
2.4.4. Mascaramento de nuvens	22
2.5. Conceitos de conjuntos e lógica <i>fuzzy</i>	23
2.5.1. Integração da lógica <i>fuzzy</i> com geotecnologias.....	26

3. MATERIAL E MÉTODOS	30
3.1. Material.....	30
3.1.1. Área de estudo	30
3.1.2. Imagens AVHRR-NOAA.....	32
3.1.3. Imagens Landsat 5/TM.....	32
3.1.4. Programas computacionais e microcomputadores	33
3.1.5. Informações cartográficas.....	34
3.2. Métodos	34
3.2.1. Processamentos das imagens AVHRR-NOAA	34
3.2.2. Estimativas de área	42
4. RESULTADOS E DISCUSSÕES	54
4.1. Processamento das imagens AVHRR-NOAA.....	54
4.2. Perfis espectrais temporais dos índices de vegetação.....	61
4.3. Estimativa de área de soja de referência por meio de matriz de erros.....	66
4.4. Estimativa de área de soja de referência por expansão direta	72
4.5. Estimativa de área de soja baseada em classificação <i>fuzzy</i>	72
4.6. Comparação entre as estimativas de área de soja.....	78
5. CONCLUSÕES.....	82
6. REFERÊNCIAS BIBLIOGRÁFICAS	84

LISTA DE TABELAS

Tabela 2.1: Estádios de desenvolvimento da cultura da soja.....	7
Tabela 2.2: Características espectrais do sensor AVHRR, versão 3.	16
Tabela 3.1: Área plantada de soja de oito municípios do oeste do Paraná, safra 2003/2004. ...	30
Tabela 3.2: Ocupação do solo dos pontos amostrais de Cascavel.	45
Tabela 3.3: Matriz de erros para as classes Soja e Não Soja.	45
Tabela 3.4: Matriz de proporções para as classes Soja e Não Soja.	46
Tabela 4.1: Imagens AVHRR-NOAA processadas na safra 2003/2004.	54
Tabela 4.2: Métricas de acurácia da classificação e área de soja em cada município.	66
Tabela 4.3: Qualidade da classificação associada aos valores do Índice Kappa.	67
Tabela 4.4: Matriz de erros e proporções - Assis Chateaubriand.	67
Tabela 4.5: Matriz de erros e proporções - Cascavel.....	68
Tabela 4.6: Matriz de erros e proporções - Mamborê.	68
Tabela 4.7: Matriz de erros e proporções - Maripá.	69
Tabela 4.8: Matriz de erros e proporções - Palotina.....	69
Tabela 4.9: Matriz de erros e proporções - Rancho Alegre do Oeste.....	70
Tabela 4.10: Matriz de erros e proporções - Toledo.....	70
Tabela 4.11: Matriz de erros e proporções - Ubitatã.	71
Tabela 4.12: Área de soja calculada por contagem de pixels e corrigida por matriz de erros. 71	
Tabela 4.13: Área de soja estimada por expansão direta.....	72
Tabela 4.14: Área de soja estimada com base na classificação <i>fuzzy</i>	74
Tabela 4.15: Correlação entre as proporções e os graus de pertinência de soja.....	77
Tabela 4.16: Estimativas de área de soja.	78
Tabela 4.17: Correlação e erro relativo entre as estimativas de área de soja.	79

LISTA DE FIGURAS

Figura 2.1: Produção de soja em nível mundial na safra 2003/2004.....	5
Figura 2.2: Evolução da área, produção e produtividade da soja no Paraná.	6
Figura 2.3: Períodos em dias dos estádios vegetativos e reprodutivos da soja.	8
Figura 2.4: Interação da radiação eletromagnética solar com a folha.	10
Figura 2.5: Interação da radiação eletromagnética fonte-alvo-sensor.....	11
Figura 2.6: Curva padrão de reflectância de uma vegetação verde.....	12
Figura 2.7: Ângulos de atitude da plataforma orbital.....	21
Figura 2.8: Funções de pertinência - conjuntos clássicos (a) e conjuntos <i>fuzzy</i> (b).....	24
Figura 2.9: Diferentes formatos das funções de pertinência - triangular (a), trapezoidal (b), Gaussiana (c), conjunto unitário (d).....	26
Figura 2.10: Exemplo de graus de pertinência de um pixel.	28
Figura 3.1: Oito municípios da região oeste do Paraná cobertos pela cena Landsat 5/TM, órbita/ponto 223/77.	31
Figura 3.2: Máscara de soja da região oeste do Paraná referente à safra 2003/2004.....	33
Figura 3.3: Fluxograma do sistema de processamento e geração de produtos das imagens AVHRR-NOAA.	35
Figura 3.4: Janela de moldura e de busca da MCC.....	39
Figura 3.5: Fluxograma das etapas utilizadas para estimativas de área de soja.....	43
Figura 3.6: Pontos amostrais no município de Cascavel.....	44
Figura 4.1: Efeito da correção dos erros de atitude do satélite no georreferenciamento.....	55
Figura 4.2: Composição RGB-121.....	56
Figura 4.3: Máscara de nuvens.....	56
Figura 4.4: NDVI.	57
Figura 4.5: NDMI.....	58
Figura 4.6: NDVI com pixels inaptos.	59
Figura 4.7: NDMI com pixels inaptos.....	59
Figura 4.8: Composição Máxima de NDVI.	60
Figura 4.9: Composição Máxima de NDMI.....	61
Figura 4.10: Perfil espectral do NDVI e do NDMI - Assis Chateaubriand.....	62

Figura 4.11: Perfil espectral do NDVI e do NDMI - Cascavel	62
Figura 4.12: Perfil espectral do NDVI e do NDMI - Mamborê	63
Figura 4.13: Perfil espectral do NDVI e do NDMI - Maripá	63
Figura 4.14: Perfil espectral do NDVI e do NDMI - Palotina	64
Figura 4.15: Perfil espectral do NDVI e do NDMI - Rancho Alegre do Oeste.	64
Figura 4.16: Perfil espectral do NDVI e do NDMI - Toledo.	65
Figura 4.17: Perfil espectral do NDVI e do NDMI - Ubiratã.....	65
Figura 4.18: Imagem fração soja da região oeste do Paraná referente à safra 2003/2004.	73
Figura 4.19: Proporções e graus de pertinência de soja - Assis Chateaubriand.....	74
Figura 4.20: Proporções e graus de pertinência de soja - Cascavel.....	75
Figura 4.21: Proporções e graus de pertinência de soja - Mamborê.	75
Figura 4.22: Proporções e graus de pertinência de soja - Maripá.	75
Figura 4.23: Proporções e graus de pertinência de soja - Palotina.....	76
Figura 4.24: Proporções e graus de pertinência de soja - Rancho Alegre do Oeste.....	76
Figura 4.25: Proporções e graus de pertinência de soja - Toledo.....	76
Figura 4.26: Proporções e graus de pertinência de soja - Ubiratã.....	77

RESUMO

A estimativa precisa com antecedência à época da colheita de áreas plantadas com culturas agrícolas, como a soja, é de fundamental importância para a economia brasileira. A previsão do escoamento e comercialização da produção agrícola é estratégica para o Brasil, pois estão diretamente relacionados com o planejamento, custos e preço. Com o recente avanço tecnológico na obtenção de dados por sensoriamento remoto orbital é possível melhorar a previsão de safras, diminuindo cada vez mais o nível de subjetividade. Embora designadas para fins meteorológicos, as imagens AVHRR-NOAA de elevada repetitividade temporal, têm sido utilizadas para o monitoramento agrícola. Porém, a sua baixa resolução espacial faz com que possa ocorrer a mistura espectral das classes de cobertura do solo dentro de um mesmo pixel e isso pode acarretar problemas de imprecisão na estimativa de área plantada de uma cultura agrícola. O objetivo principal do trabalho foi desenvolver uma metodologia de classificação automática com a aplicação de lógica *fuzzy* para o reconhecimento de padrões em imagens AVHRR-NOAA, utilizando índices de vegetação para estimar a área plantada de soja no nível sub-pixel. Para oito municípios produtores de soja da região oeste do Estado do Paraná, foi possível obter a estimativa de área no final de janeiro de 2004, com antecedência em relação à época da colheita, ao contrário dos levantamentos oficiais que se estendem até o final da safra, além de utilizarem dados subjetivos vindos do campo. As estimativas de área de soja baseadas em classificação *fuzzy* mostraram-se altamente correlacionadas com as estimativas de área de referência obtidas a partir da máscara de soja e por expansão direta, sendo um indicativo de boa precisão. E também apresentaram alta correlação, balizadas com as estimativas oficiais da SEAB/DERAL e do IBGE. Em ambas comparações, o nível de erro relativo geral foi aceitável. O sistema desenvolvido para processamento e geração de produtos das imagens AVHRR-NOAA mostrou-se uma ferramenta fundamental de infra-estrutura, por aliar automação e precisão a metodologia do trabalho.

Palavras-chave: Previsão de safras; Sensoriamento remoto; NDVI; NDMI; Reconhecimento de padrões; Classificação de imagens; Conjunto nebuloso; Lógica nebulosa.

APPLICATION OF FUZZY LOGIC FOR SOYBEAN CROP AREA ESTIMATION USING AVHRR-NOAA IMAGES

ABSTRACT

An early accurate estimation of agricultural crop areas, such as soybean, is fundamental for the Brazilian economy. The draining forecast and the estimation of agricultural production commercialization are strategic to Brazil, since they are directly related to planning, costs and price. Recent technological progress of data acquisition from orbital remote sensing makes possible to improve harvest forecast, reducing more and more the level of subjectivity. Although designed for meteorological aims, the AVHRR-NOAA images of high temporal resolution, have been used for the crop monitoring. However, its low spatial resolution might cause the spectral mixture of the different land cover classes within the same pixel and it can lead to accuracy problems on crop area estimation. The main objective of the work was to develop an automatic classification methodology with the application of fuzzy logic for pattern recognition in AVHRR-NOAA images, using vegetation indices to estimate the soybean crop areas at sub-pixel level. For eight soybean producer counties in the West region of the Paraná State, it was possible to obtain the crop area estimation at the end of January 2004, prior to the harvest period, on the contrary of the official surveys that extend until the end of the harvest, besides using subjective data collected on the field. The soybean crop area estimation based on fuzzy classification showed to be highly correlated with the reference area estimation obtained from the soybean mask and by direct expansion, being an indicative of good accuracy. And also presented high correlation, marked out with the official estimations from SEAB/DERAL and IBGE. In both comparisons, the level of general relative error was acceptable. The system developed for processing and products generation of AVHRR-NOAA images had proved to be a fundamental infrastructure tool, due to its capacity to combine automation and accuracy to the work methodology.

Keywords: Harvest forecast; Remote sensing; NDVI; NDMI; Pattern recognition; Image classification; Fuzzy set; Fuzzy logic.

1. INTRODUÇÃO

O planejamento e o gerenciamento do setor agrícola têm se tornado cada vez mais sofisticado, requerendo informações precisas e com antecedência cada vez maior em relação à época da safra. As indústrias necessitam de informações que possibilitem o conhecimento prévio dos seus mercados compradores, o planejamento do transporte e o beneficiamento da safra, para estabelecer metas de produção. Além disso, os governos devem funcionar como facilitadores dos processos iniciados no setor agrícola e industrial, planejar exportações e importações e administrar possíveis crises nos diversos setores agrícolas.

Para o desenvolvimento de modelos de previsão de safras é necessário ter informações sobre o modelo de produtividade agrícola da cultura e a estimativa da área plantada. As duas informações têm importância singular, pois pouco adianta ter-se modelos de rendimento acurados, se a estimativa da área plantada for imprecisa.

Nesse sentido, estimativas confiáveis de produção agrícola são ferramentas poderosas para orientar o produtor nas questões ligadas ao plantio, auxiliar as indústrias no setor operacional e de comercialização, indicando ao governo números confiáveis que permitam sua intervenção para reduzir impactos negativos na economia ou aproveitar antecipadamente os benefícios de uma situação favorável.

As estimativas de área plantada são, em sua maioria, realizadas por meio de métodos indiretos, através de entrevistas realizadas junto aos produtores e entidades relacionadas à atividade agrícola. Tais métodos possuem um custo elevado e execução demorada, tendo em vista a dinâmica e a extensão agrícola e, por isso, estão sujeitos a imprecisões.

A soja destaca-se como sendo o principal produto de exportação do país pois responde por 10% do total das receitas cambiais brasileiras. A soja é cultivada em grandes extensões, em praticamente todo o Brasil, sendo que as regiões Sul e Centro-Oeste são as maiores produtoras e que também possuem as maiores áreas plantadas. O Estado do Paraná é o segundo maior produtor de soja, responsável por quase 20% da produção brasileira. A região oeste do Estado responde por mais de 25% desse total.

Nos últimos anos, os dados de sensoriamento remoto orbital têm sido utilizados na agricultura, especialmente para mapear o uso do solo, ou seja, na discriminação das espécies, culturas vegetais e, também, na qualificação e quantificação de área plantada.

Com o advento dos satélites ambientais operando de forma sistemática, atenção considerável tem sido dada ao desenvolvimento de modelos de previsão de safras com a participação de dados de sensoriamento remoto orbital. O Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura (CEPAGRI) da Universidade Estadual de Campinas (UNICAMP) é uma das instituições no Brasil que recebe diariamente as imagens do satélite meteorológico NOAA, fornecendo coberturas de grandes áreas a custos baixos. Tais imagens podem ser utilizadas para auxiliar na estimativa de área plantada de culturas agrícolas, como a soja, por possibilitarem uma visão sinóptica e freqüente da superfície terrestre.

Para o monitoramento agrícola em tempo real é necessário georreferenciar as imagens recebidas para garantir que cada pixel na imagem corresponda à mesma localização geográfica em qualquer uma das imagens disponíveis no período. Antes disso, deve-se proceder à calibração radiométrica para geração de produtos quantitativos. Também se faz necessário detectar os pixels contaminados com nuvens e separá-los dos não contaminados nas imagens para obtenção de produtos confiáveis gerados para o usuário final, tais como os índices de vegetação.

Entretanto, a baixa resolução espacial das imagens AVHRR-NOAA faz com que possa ocorrer a mistura espectral das classes de cobertura do solo dentro de um mesmo pixel, dificultando o detalhamento e a precisão necessária para a estimativa de área plantada de uma cultura agrícola.

Diversos trabalhos têm sido conduzidos procurando estabelecer relações entre dados espectrais de satélites para identificação da cultura com o objetivo de melhorar a estimativa de área plantada. Os resultados destes trabalhos enfatizam a necessidade de se buscar novos estimadores, de desempenho mais favorável, como o reconhecimento de padrões dos dados digitais em imagens de satélites. A aplicação de técnicas como a lógica *fuzzy* (nebulosa) é bastante promissora para a estimativa de área plantada no nível sub-pixel, a partir da classificação automática de imagens de satélites.

A seguir é apresentado o problema a ser investigado, o enunciado da hipótese, o objetivo geral e os objetivos específicos do trabalho. A partir desses elementos é desenvolvida a revisão bibliográfica sobre os tópicos relacionados ao tema de pesquisa. Após isso são descritos os materiais e métodos utilizados no estudo. A seguir são relatados os resultados, as

discussões e as conclusões do trabalho. Ao final são listadas as referências bibliográficas utilizadas.

1.1. Problema

Os satélites NOAA fornecem coberturas de grandes áreas a custos baixos com alta periodicidade. Porém, a baixa resolução espacial das imagens AVHRR-NOAA faz com que possa ocorrer a mistura espectral das classes de cobertura do solo dentro de um mesmo pixel e isso pode acarretar problemas de imprecisão na estimativa da área plantada de culturas agrícolas, como a soja.

1.2. Hipótese

A aplicação de lógica *fuzzy* para o reconhecimento de padrões na classificação automática de imagens AVHRR-NOAA pode auxiliar na obtenção de uma estimativa de área plantada de soja mais objetiva, precisa e com antecedência à época da colheita, porque pode descrever as proporções de cada classe de cobertura do solo dentro do pixel.

1.3. Objetivos

O objetivo geral do trabalho foi desenvolver uma metodologia de classificação automática baseada em lógica *fuzzy* para obter a estimativa de área plantada de soja no nível sub-pixel, de oito municípios da região oeste do Estado do Paraná, ao longo da safra 2003/2004.

Dentre os objetivos específicos, destacam-se:

- Desenvolver um sistema automático para conversão de formato bruto, calibração radiométrica, georreferenciamento preciso e geração de produtos das imagens AVHRR-NOAA;
- Estimar a área plantada de soja de referência a partir de imagens Landsat 5/TM obtidas ao longo do ciclo de desenvolvimento da cultura na região em estudo;

- Implementar o algoritmo do classificador *fuzzy* definindo a função de pertinência a ser utilizada;
- Estimar a área plantada de soja a partir do resultado da classificação *fuzzy* utilizando índices de vegetação de imagens AVHRR-NOAA, realizada no período de maior vigor vegetativo da cultura na região em estudo;
- Comparar os dados de estimativa de área de soja obtidos por meio da classificação *fuzzy*, confrontando-os com as estimativas de área de soja oficiais e de referência.

2. REVISÃO BIBLIOGRÁFICA

2.1. A cultura da soja

A soja (*Glycine max (L.) Merrill*) é originária da China, onde é cultivada há séculos. No Brasil, chegou vinda dos Estados Unidos em 1882. O primeiro registro de plantio foi em 1914, no município de Santa Rosa - RS, mas foi só em 1941 que adquiriu alguma importância econômica, merecendo o primeiro dado estatístico nacional no anuário agrícola do Rio Grande do Sul, com área cultivada de 640 hectares e produção de 450 toneladas (EMBRAPA SOJA, 2004).

A soja vem se consolidando como a principal cultura do agronegócio brasileiro. Na safra 2003/2004, o país foi o segundo maior produtor mundial do grão, alcançando a marca de 50 milhões de toneladas, o que representou 28% da produção mundial, como pode ser observado na Figura 2.1. O Brasil é responsável por cerca de 35% das exportações mundiais de soja, atrás somente dos Estados Unidos (CONAB, 2004).

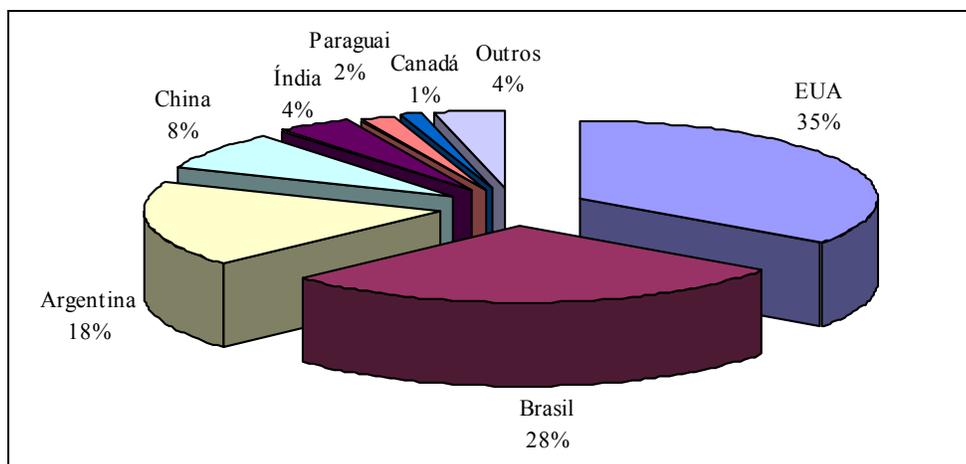


Figura 2.1: Produção de soja em nível mundial na safra 2003/2004.

Fonte: Adaptado de CONAB (2004).

O Estado do Mato Grosso é o líder nacional na produção de soja, com mais de 15 milhões de toneladas, 30% do total produzido no país, e área plantada estimada em 5,15 milhões de hectares. O Estado do Paraná é o segundo maior produtor, com 10,03 milhões de toneladas, 20% do total no Brasil, e área plantada estimada em 3,5 milhões de hectares. A

região oeste do Paraná, local de estudo deste trabalho, responde por mais de 25% do total da produção do Estado. O município de Cascavel é o maior produtor de soja do Paraná e o segundo maior do Brasil, sendo responsável por 2,3% do total produzido no Estado (CONAB, 2004). Na Figura 2.2, pode-se observar a evolução do cultivo da soja no Paraná.

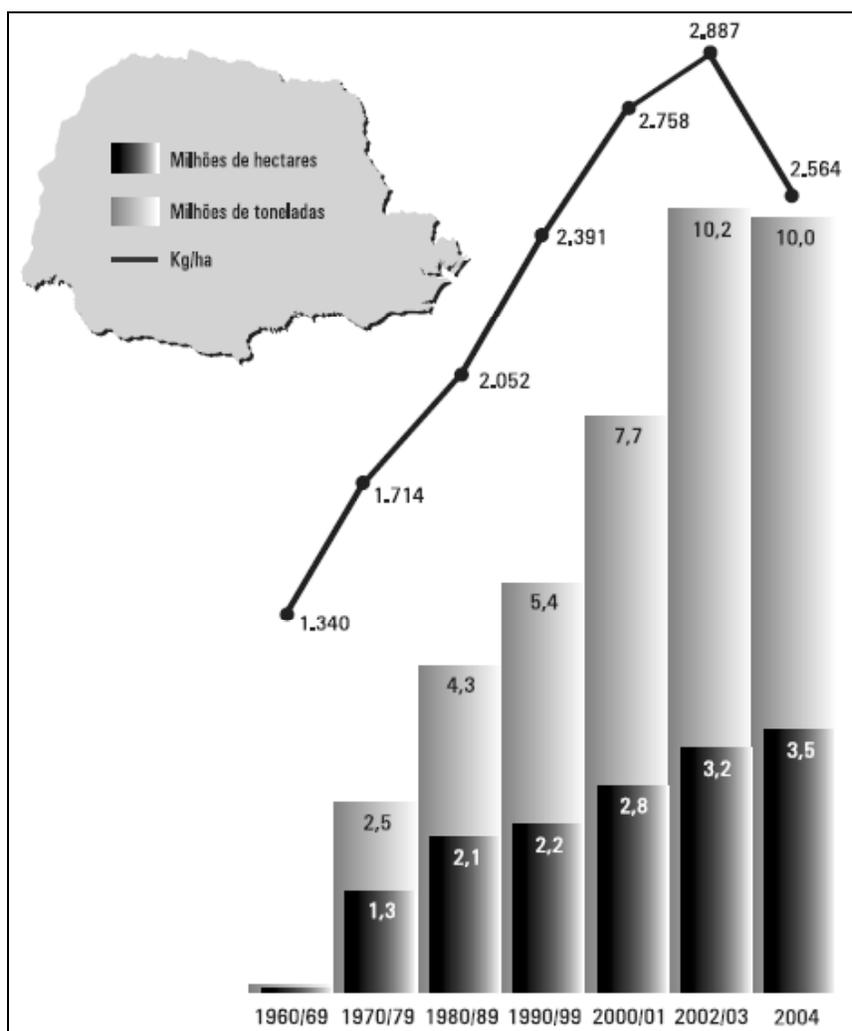


Figura 2.2: Evolução da área, produção e produtividade da soja no Paraná.
Fonte: Adaptado de EMBRAPA SOJA (2004).

A soja é uma planta leguminosa, dicotiledônea, predominantemente utilizada para o processamento do grão em óleo e proteína. De um modo geral, o ciclo fenológico da planta varia de 115 a 150 dias, desde a sua germinação até a maturação. Os cultivares podem ser classificados em precoce (até 115 dias), semiprecoce (116 a 125 dias), médio (126 a 137 dias), semitardio (138 a 150 dias) e tardio (mais que 150 dias). A altura da planta varia entre 0,45 m

a 1,20 m, dependendo da variedade, pluviosidade, temperatura e época de semeadura. Normalmente, a época de semeadura da soja indicada no Paraná estende-se de 15 de outubro a 15 de novembro e o período de colheita de 15 de fevereiro a 15 de março (EMBRAPA SOJA et al., 2004).

A produção da soja está diretamente relacionada com parâmetros fisiológicos da planta, tais como taxa de crescimento, índice de área foliar (IAF) e produção de matéria seca (biomassa). O crescimento e o desenvolvimento da soja são medidos pela quantidade de matéria seca acumulada na planta que é produzida por meio da fotossíntese. Com exceção da água, a matéria seca é composta por carboidratos, proteínas, lipídeos e nutrientes minerais.

O desenvolvimento da cultura da soja pode ser dividido em estádios (fases) vegetativos (V) e reprodutivos (R), conforme as subdivisões apresentadas na Tabela 2.1.

Tabela 2.1: Estádios de desenvolvimento da cultura da soja.

Estádios Vegetativos	Estádios Reprodutivos
VE - emergência	R ₁ - início do florescimento
VC - cotilédone	R ₂ - pleno florescimento
V ₁ - primeiro nó	R ₃ - início da formação das vagens
V ₂ - segundo nó	R ₄ - plena formação das vagens
V ₃ - terceiro nó	R ₅ - início do enchimento das sementes
*	R ₆ - pleno enchimento das vagens
*	R ₇ - início da maturação
V _(n) - enésimo nó	R ₈ - maturação plena

Fonte: Adaptado de FEHR e CAVINESS (1981).

O crescimento vegetativo e a produção de novos nós continuam durante alguns estádios reprodutivos, em que os estádios de R1 a R6 descrevem melhor o desenvolvimento da planta. A partir dos estádios R3 e R5 há uma gradativa transferência de matéria seca das partes vegetativas da planta para as vagens e grãos em formação. No estádio R6, a semente está com aproximadamente 80% de sua matéria seca total (FEHR e CAVINESS, 1981).

De forma geral, o desenvolvimento fenológico da soja, com a duração dos períodos em dias, nos respectivos estádios vegetativos e reprodutivos é esquematizado na Figura 2.3.

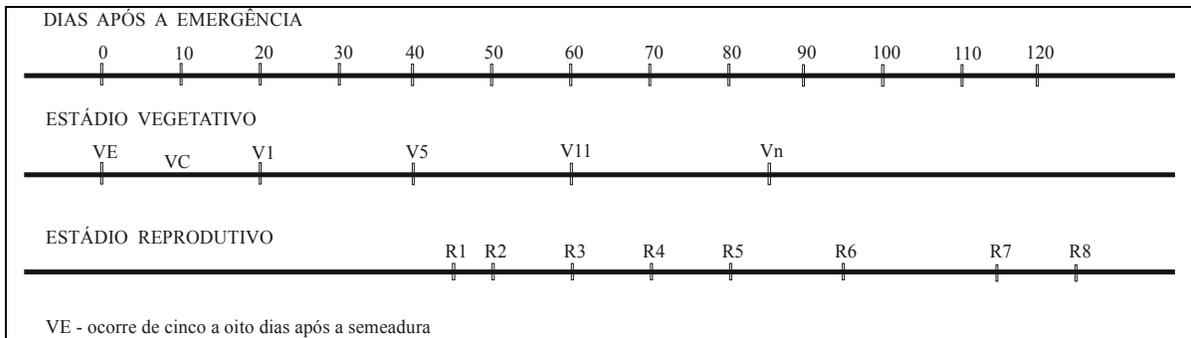


Figura 2.3: Períodos em dias dos estádios vegetativos e reprodutivos da soja.
Fonte: Adaptado de FEHR e CAVINESS (1981).

2.2. A previsão de safras

A previsão de safras é uma atividade estratégica no estabelecimento de instrumentos de políticas e negócios agrícolas, tanto para o setor público quanto para o privado, com a finalidade de fornecer informações aos formuladores de políticas públicas.

IPPOLITI-RAMILO et al. (1999) cita que as previsões obtidas com antecedência suficiente à época de colheita são importantes para o planejamento do transporte, processamento e armazenamento da produção, bem como para a tomada de decisões relacionadas às exportações e importações agrícolas e ao abastecimento do mercado interno.

PINO (1999) menciona que a informação estatística sobre a agropecuária tem se tornado extremamente importante para que empresas atuem numa economia globalizada. Os países que pretendem ter desenvolvimento pleno necessitam contar com previsões de safras eficazes, precisas e com rápida percepção das mudanças que eventualmente possam ocorrer, não somente sobre si mesmo, mas sobre os demais países.

Os sistemas de previsão de safras convencionais baseiam-se no produto da área ocupada pelo rendimento médio para obtenção da produção final das culturas. Os valores de cada uma dessas variáveis podem ser obtidos por meio de censos agropecuários ou pela amostragem da população agrícola. O levantamento censitário tem como objetivo fornecer uma classificação detalhada da estrutura agrícola de um país, onde o valor de cada variável para a área total do levantamento é obtido a partir dos valores das variáveis em todas as unidades. No levantamento amostral, o valor de cada variável para a área total do

levantamento é obtido a partir de inferências feitas sobre uma amostra da população (FAO, 1996).

No Brasil, o Instituto Brasileiro de Geografia e Estatística (IBGE) realiza o Levantamento Sistemático da Produção Agrícola (LSPA) para previsão de safras. Trata-se de um método censitário de acompanhamento das safras agrícolas em âmbito nacional, abrangendo todos os municípios brasileiros, fornecendo estimativas de área, rendimento médio e produção de 35 produtos. As estimativas são divulgadas mensalmente calculadas com base em dados coletados em questionários aplicados ao setor agrícola, como produtores rurais, bancos de crédito agrícola, cooperativas, secretarias de agricultura, agentes de assistência técnica e comerciantes de insumos (IBGE, 2002).

A CONAB faz o levantamento de safra no país por meio de pesquisas de intenção de plantio e colheita realizadas seis vezes durante o ano safra, também por meio de enquetes com as entidades ligadas à produção agrícola. O diferencial é que este método é realizado em forma de amostragem pré-estabelecida nos municípios que sejam os mais representativos das culturas em cada região do país (CONAB, 2004).

Segundo LUIZ (2003), esses levantamentos têm caráter subjetivo porque os valores de uma variável são obtidos de forma indireta, sujeitos a interferências não controladas como nos realizados por entrevistas. Ainda, como são obtidos de forma não-probabilística, não possibilitam realizar uma análise quantitativa dos erros envolvidos por meio de avaliações estatísticas.

De acordo com FONTANA et al. (2000), métodos tradicionais envolvendo coleta de informações por meio de questionários aplicados aos agentes ligados ao meio rural são formas difíceis de se obter dados confiáveis, além de serem demoradas, onerosas e subjetivas, o que pode causar imprecisões no processo de estimativa de produção.

Nesse contexto, IPPOLITI-RAMILO et al. (1999) destaca que o caráter multiespectral, sinóptico, repetitivo e global do sensoriamento remoto orbital, em conjunto com as tecnologias de geoprocessamento, têm grande potencial de uso com os sistemas de estimativas agrícolas para um levantamento mais rápido e uma quantificação da ocupação agrícola mais objetiva e precisa.

2.3. O sensoriamento remoto e o uso na agricultura

COLWELL (1983) define sensoriamento remoto como a ciência que permite a aquisição de informações sobre objetos da superfície terrestre sem contato físico com eles, através de imagens obtidas à distância.

A obtenção das informações da superfície terrestre depende da energia eletromagnética que interage com a matéria. O Sol é a principal fonte de energia radiante, a qual se propaga no espaço e incide sobre a superfície em diversos comprimentos de onda, que constituem o espectro eletromagnético.

Um fluxo de radiação eletromagnética (I), ao se propagar pelo espaço, pode interagir com os elementos da superfície terrestre, sendo por estes refletido (R), absorvido (A) e transmitido (T), como no caso de uma folha, exemplificado na Figura 2.4. As variações que essas interações produzem no fluxo eletromagnético dependem fortemente das propriedades físico-químicas desses objetos representadas em termos de reflectância, absorptância e transmitância. A soma dos fluxos de energia refletida, absorvida e transmitida é igual ao fluxo de energia incidente, ou seja, sempre guardam o princípio da conservação de energia. A absorção e o espalhamento causados por gases e aerossóis presentes na atmosfera reduzem a radiação solar incidente em certos comprimentos de onda.



Figura 2.4: Interação da radiação eletromagnética solar com a folha.
Fonte: Adaptado de MOREIRA (2003).

Os sensores remotos orbitais são desenvolvidos para medir a energia eletromagnética provenientes dos alvos na superfície terrestre, como ilustrado na Figura 2.5. A característica multiespectral desses sensores permite medir e registrar a energia eletromagnética em determinadas faixas de comprimentos de onda, chamadas de bandas espectrais, transformando-a numa imagem digital que pode ser interpretada.

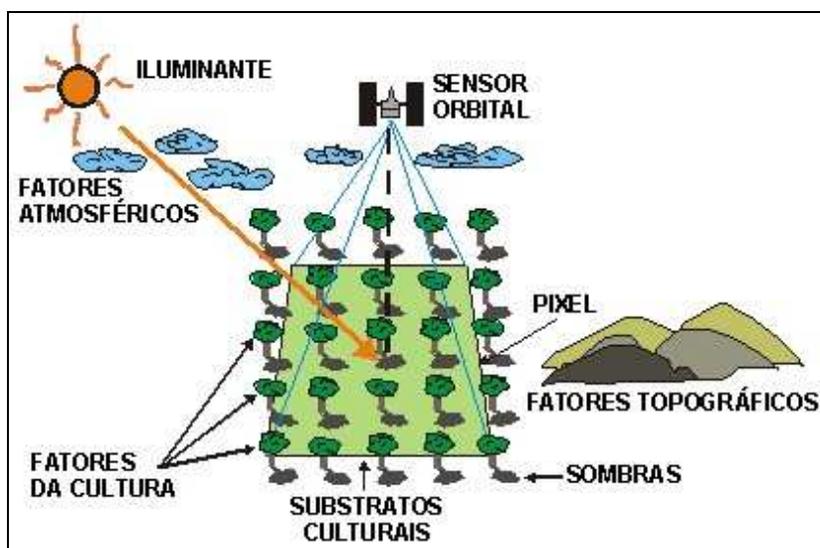


Figura 2.5: Interação da radiação eletromagnética fonte-alvo-sensor.

O sensoriamento remoto vem sendo utilizado no monitoramento agrícola devido à sua visão sinóptica da superfície terrestre obtida em diversos comprimentos de onda. Dentre as várias aplicações na agricultura, pode-se citar o monitoramento da extensão, vigor e tipo de cobertura vegetal (BATISTA et al., 1990, PRICE et al., 2001) e a estimativa de rendimento agrícola (LIU e KOGAN, 2002, XIN et al., 2002).

No entanto, para a realização de estudos envolvendo culturas agrícolas, é necessário conhecer o comportamento espectral dessas superfícies nas várias faixas de comprimento de onda do espectro eletromagnético.

2.3.1. Comportamento espectral da vegetação

O acompanhamento de uma cultura agrícola pode ser feito por meio do estudo da radiação eletromagnética por ela refletida ou emitida, sendo que as folhas são a estrutura de

maior influência no comportamento espectral das plantas. A reflectância de uma vegetação verde sadia é distinta e completamente variável com o comprimento de onda nas regiões do espectro eletromagnético do visível, do infravermelho próximo e do infravermelho médio, como ilustrado na Figura 2.6.

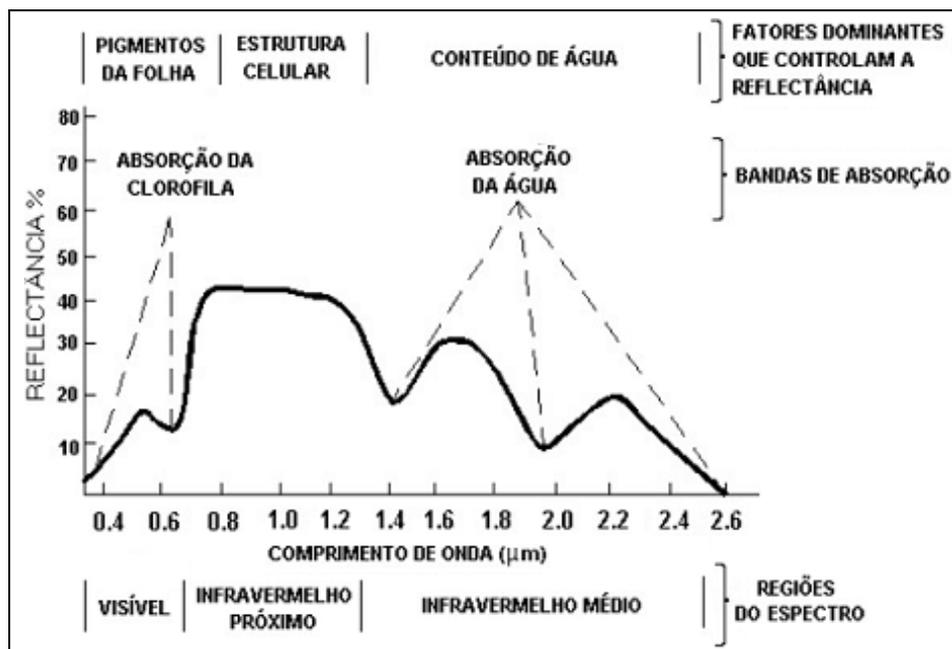


Figura 2.6: Curva padrão de reflectância de uma vegetação verde.
Fonte: Adaptado de HOFFER (1978).

Na região do visível (0,4 μm a 0,7 μm), a presença de pigmentos clorofilados da planta é responsável pela maior absorção da radiação nas porções azul e vermelho, sendo menos absorvida na porção do verde. O pico de reflectância máxima no visível ocorre no comprimento de onda do verde (0,54 μm) e não ultrapassa os 20%. No infravermelho próximo (0,7 μm a 1,3 μm), a estrutura celular é responsável por altos valores de reflectância e transmitância (quase 50%) e baixos valores de absorção (menor que 5%). No infravermelho médio (1,3 μm a 2,6 μm), as propriedades ópticas internas das plantas e a presença de água provocam alta absorção da radiação nos comprimentos de onda de 1,4 μm, 1,9 μm e 2,6 μm (HOFFER, 1978).

Segundo FORMAGGIO (1989), a energia refletida por uma área cultivada é uma integração dos valores de reflectância da planta e do solo. O índice de cobertura vegetal

(porcentagem de solo coberto por plantas) encontra-se diretamente relacionado com a reflectância da vegetação nas faixas do visível e infravermelho próximo. Quando a cobertura do solo é completa, as folhas são elementos dominantes nas plantas na definição das propriedades espectrais de dosséis vegetais.

KOLLENKARK et al (1982), estudando a cultura da soja, relatam que a reflectância espectral dos dosséis vegetais pode ser alterada por práticas culturais de espaçamento, população de plantas, data de plantio, bem como pelo porte, vigor e fase fenológica da cultura. No início do desenvolvimento, o comportamento espectral do dossel vegetal é altamente influenciado pelo solo, já que as plantas estão germinando e a cobertura vegetal praticamente não existe. À medida que a soja cresce e a cobertura vegetal aumenta, a influência do solo diminui e o comportamento do dossel tende a se ajustar ao comportamento das plantas verdes. Ao entrar em senescência, o porte das plantas e o vigor vegetal diminuem, reduzindo a cobertura vegetal e alterando o comportamento espectral do dossel que volta ao padrão do solo.

Em estudos de culturas agrícolas utilizando o sensoriamento remoto tem sido comum utilizar índices de vegetação para ressaltar a resposta espectral da vegetação em função da cobertura do dossel sobre o solo.

2.3.2. Índices de vegetação

WIEGAND et al. (1991) definem os índices de vegetação como o resultado de transformações lineares de reflectância entre duas ou mais bandas espectrais através de operações de soma, diferença ou razão, com o objetivo de desenvolver relações funcionais dos dados com as características da vegetação.

JACKSON e HUETE (1991) citam que um índice de vegetação ideal deve ser capaz de quantificar pequenas variações no estado fenológico da vegetação, amenizar a influência do solo e não ser muito influenciado por variações da geometria da cena e das condições atmosféricas.

Segundo TOWNSHEND (1994), os índices de vegetação guardam uma relação com a quantidade e o estado da vegetação na área onde foi feita a medida espectral. Por isso, têm sido empregados em estudos para caracterizar parâmetros biofísicos da vegetação por

indicarem sua presença e condição, visto que a biomassa tem estreita relação com a energia solar absorvida, o que define o crescimento e o desenvolvimento de uma cultura. Em geral, as medidas de índices de vegetação têm associação direta com o índice de área foliar (WANG et al., 2005), a quantidade de biomassa (MOREAU et al., 2003), a produtividade (MIKA et al., 2002) e a porcentagem de cobertura do solo (SENAY e ELLIOTT, 2002).

Na literatura são encontrados vários índices de vegetação. Os mais comuns são obtidos de medidas de reflectância nas faixas espectrais do vermelho e do infravermelho próximo. Também existem outros índices de vegetação utilizando dados de reflectância na faixa do infravermelho médio em trabalhos mais recentes.

De acordo com TUCKER (1979), os índices de vegetação exploram o contraste nos valores de reflectância das faixas espectrais, com o objetivo de melhorar a caracterização da vegetação e reduzir as interferências de fatores que prejudiquem a associação com os parâmetros biofísicos.

A origem dos índices de vegetação é atribuída a JORDAN (1969) que criou uma razão simples baseada na reflectância das faixas espectrais do infravermelho próximo e vermelho. Contudo, os pioneiros nesse estudo foram PEARSON e MILLER (1972) que realizaram pesquisas para medir o fator de reflectância de gramíneas e constataram que os valores obtidos nos comprimentos de onda de 0,68 μm a 0,78 μm foram os melhores para a separação entre a vegetação e o solo.

Após isso, ROUSE et al. (1973) propõem o *Normalized Difference Vegetation Index* (NDVI), com o objetivo de minimizar a interferência do solo na resposta da vegetação e, também, diminuir a influência da atmosfera e das variações sazonais do ângulo solar. O NDVI é calculado pela Equação (2.1) e, na teoria, os seus valores variam no intervalo de -1,0 a +1,0:

$$\text{NDVI} = \frac{\rho_{\text{IVP}} - \rho_{\text{VER}}}{\rho_{\text{IVP}} + \rho_{\text{VER}}} \quad (2.1)$$

onde,

ρ_{VER} = reflectância no vermelho (%);

ρ_{IVP} = reflectância no infravermelho próximo (%).

WILSON e SADER (2002) aplicaram o pouco conhecido *Normalized Difference Moisture Index* (NDMI) para estudar a influência hídrica em vegetações e concluíram que o NDMI explicou melhor a dinâmica da vegetação do que o NDVI, durante a análise de uma série multitemporal de imagens. O NDMI utiliza dados da faixa do infravermelho médio que é sensível à umidade por causa das regiões de absorção de água. Essa banda também é menos influenciada pelos efeitos de absorção e espalhamento dos aerossóis e vapor de água presentes na atmosfera devido ao tamanho do comprimento de onda. Em função disso, pode indicar a presença de umidade na vegetação e no solo permitindo obter um maior contraste entre tipos diferentes de vegetação, refletindo melhor as mudanças de biomassa que o NDVI.

Segundo JIN e SADER (2005), o aperfeiçoamento das técnicas de monitoramento da vegetação é particularmente relevante e um grande desafio. Nesse contexto, o NDMI é um bom índice para a detecção de mudanças porque proporciona uma melhor distinção espectral da vegetação na análise ao longo de uma série multitemporal de imagens. O NDMI é calculado pela Equação (2.2) e seus valores também podem variar de -1,0 a +1,0:

$$\text{NDMI} = \frac{\rho_{\text{IVP}} - \rho_{\text{IVM}}}{\rho_{\text{IVP}} + \rho_{\text{IVM}}} \quad (2.2)$$

onde,

ρ_{IVP} = reflectância no infravermelho próximo (%);

ρ_{IVM} = reflectância no infravermelho médio (%).

2.3.3. Satélites NOAA

A empresa norte-americana National Oceanic and Atmospheric Administration (NOAA) possui um programa de satélites meteorológicos, administrado pela National Environmental Satellite, Data and Information Service (NESDIS). A National Aeronautics and Space Administration (NASA) é responsável pela construção, integração e lançamento dos aparelhos.

Os chamados satélites NOAA são de cobertura diária, órbita quase polar e heliossíncrona, ou seja, sobrevoa a superfície terrestre com o Sol na posição aproximada da passagem anterior, numa altitude média de 850 km. O período médio para cruzar novamente a

linha do Equador é de 102 min com resolução temporal de 12 h. De 1978 a 2004 foram colocados quinze satélites em órbita. Em 2004 encontrava-se em funcionamento o NOAA-12, 14, 15, 16 e 17 (KIDWELL, 2004). O novo satélite NOAA-18 foi lançado no dia 11 de maio de 2005 em substituição ao NOAA 16.

Dentre os sensores existentes a bordo dos satélites NOAA, está o *Advanced Very High Resolution Radiometer* (AVHRR), versão 3, criado a partir do NOAA-15. O instrumento é um radiômetro de varredura transversal que mede a energia refletida em três bandas na região do vermelho e infravermelho próximo e, também, detecta a energia emitida em três bandas na região do infravermelho termal, conforme características espectrais apresentadas na Tabela 2.2.

Tabela 2.2: Características espectrais do sensor AVHRR, versão 3.

Banda	Faixa Espectral (μm)	Região do Espectro Eletromagnético
1	0,58 a 0,68	Vermelho
2	0,725 a 1,0	Infravermelho próximo
3A	1,58 a 1,64	Infravermelho médio
3B	3,55 a 3,93	Infravermelho termal
4	10,3 a 11,3	Infravermelho termal
5	11,5 a 12,5	Infravermelho termal

Fonte: Adaptado de GOODRUM et al. (2004).

Entretanto, somente cinco bandas podem ser transmitidas simultaneamente. Nos satélites em órbitas vespertinas, as bandas 3A e 3B são alternadas para operação de dia e de noite, respectivamente, enquanto que para órbitas matutinas, a banda 3B fica permanentemente ativa.

Os dados analógicos de cada banda do sensor são digitalizados a bordo do satélite com resolução radiométrica de 10 bits, isto é, de 0 a 1023 números digitais (ND). Com um campo de visada instantânea (IFOV) de 1,3 mrad, taxa de 39.936 amostras por segundo por banda e ângulo de abertura igual a $\pm 55,4^\circ$, a varredura da superfície terrestre é feita a uma taxa de seis linhas por segundo, gerando uma imagem com largura aproximada de 2.700 km com resolução espacial de 1,1 x 1,1 km no nadir, podendo chegar a 2,4 x 6,9 km nos extremos da imagem.

Um mesmo local na superfície terrestre é rastreado aproximadamente a cada seis horas por um mesmo satélite AVHRR-NOAA, com coberturas diurnas e noturnas. Com cinco satélites operacionais pode-se obter uma imagem a cada duas horas, sendo pelo menos dez por dia (GOODRUM et al., 2004).

Os métodos de monitoramento agrícola em nível global, continental e regional podem ser realizados com o auxílio de imagens diárias geradas a partir de plataformas orbitais com elevada repetitividade temporal como as AVHRR-NOAA, embora tenham sido projetados para fins meteorológicos, geográficos e ambientais, por proverem a necessária visão da superfície terrestre.

TOWNSHEND e TUCKER (1984) iniciaram o emprego de dados AVHRR-NOAA para mapeamento da vegetação da superfície terrestre no início de 1980 utilizando o NDVI. JAKUBAUSKAS et al. (2002) relatam que as imagens multitemporais AVHRR-NOAA podem ser utilizadas para discriminação de áreas cultivadas, uma vez que cada cultura apresenta variações distintas de reflectância ao longo do tempo durante o seu ciclo fenológico.

2.3.4. Série multitemporais de imagens

Uma das grandes vantagens que as imagens AVHRR-NOAA apresenta é a sua elevada resolução temporal, com garantia de cobertura diária e de aquisição gratuita. Por isso, possuem características fundamentais no estudo de alvos com grande dinâmica espectral e podem ser empregadas para o monitoramento de culturas agrícolas. Porém, para obter um produto de qualidade com valores confiáveis dos índices de vegetação é necessário dispor de uma série multitemporal de imagens com determinadas características geométricas e radiométricas.

Para isso utiliza-se a técnica *Maximum Value Compositie* (MVC) que é obtida de uma série multitemporal de imagens georreferenciadas de índices de vegetação, onde o valor máximo de um dado pixel numa seqüência de imagens é utilizado para a construção de uma composição de um dado período (NOAA, 1983).

Segundo HOLBEN (1986), as composições máximas tem o objetivo de minimizar a influência de nuvens, das variações do ângulo de iluminação solar, dos efeitos de sombras, da geometria de visada e, também, diminuir os efeitos dos aerossóis e vapor de água nas imagens.

Quanto mais longo o período da composição, maior a atenuação dos efeitos atmosféricos. Entretanto, séries muito longas reduzem a quantidade de dados para análise da dinâmica espectral das vegetações, de forma que o período das composições deve considerar as características atmosféricas da área de estudo, assim como o ciclo fenológico da cultura.

A partir das composições máximas é possível gerar perfis dos índices de vegetação com objetivo de realizar o monitoramento de culturas agrícolas. Os perfis espectrais descrevem a variação do vigor vegetativo de uma cultura ao longo do ciclo de desenvolvimento das plantas. Num trabalho utilizando imagens AVHRR-NOAA, FONTANA et al. (2001) analisaram as diferenças entre os perfis de NDVI em regiões produtoras de soja no Rio Grande do Sul e conseguiram realizar o acompanhamento da evolução da biomassa da soja, em nível regional.

2.4. O processamento das imagens AVHRR-NOAA

As imagens AVHRR-NOAA continuam sendo de uso amplo em função da disponibilidade de séries temporais longas armazenadas em acervos históricos em todo mundo. Para estudos envolvendo esses dados, primeiramente, é necessário que as imagens passem por algumas etapas de processamento até ficarem aptas à geração de produtos confiáveis e de qualidade.

2.4.1. Conversão do formato bruto

As imagens AVHRR-NOAA têm sido transmitidas diretamente para os usuários espalhados ao redor do planeta. O modo mais comum, chamado *High Resolution Picture Transmission* (HRPT), permite a aquisição direta das imagens em alta resolução por estações de recepção localizadas no campo de visada dos satélites. As transmissões incluem dados do processador de informações, do processador do instrumento e do sensor AVHRR (KIDWELL, 2004). Porém, o formato HRPT não é compatível com softwares de processamento de imagens e, por isso, necessita ser convertido.

2.4.2. Calibração radiométrica

De acordo com DI e RUNDQUIST (1994), a radiância medida por um sistema de sensoriamento remoto para um dado alvo é influenciada por fatores como mudanças de iluminação da cena, condições atmosféricas, geometria de visada e características de resposta dos sensores. Para se comparar dados de uma mesma cena que tenham sido gerados pelo mesmo sensor, porém em momentos distintos, é necessário remover os erros relacionados a esses fatores.

Os dados de sensoriamento remoto possuem características geométricas e radiométricas que visam identificar alvos na superfície terrestre. A geometria de seus produtos fornece material para análises qualitativas e geração de mapas temáticos. Já a radiometria atende a outro objetivo e com a devida calibração, os valores de ND podem ser transformados em reflectância no topo da atmosfera que, por sua vez, possibilita realizar uma caracterização quantitativa relacionada com os parâmetros biofísicos dos objetos da superfície.

Em estudos que exijam uma análise multitemporal, espera-se verificar variações no alvo ao longo do tempo. Porém, existe a necessidade de se eliminar os efeitos de degradação do sensor que acontece com o passar do tempo, para que os dados se tornem úteis para uma análise quantitativa.

A calibração radiométrica é o processo que expressa a relação entre o nível de radiância que chega ao sensor e ND gerado por este. Os coeficientes associados aos dados são obtidos por análise de regressão e precisam ser redefinidos freqüentemente (DINGUIRARD e SLATER, 1999).

2.4.3. Georreferenciamento preciso

Em imagens AVHRR-NOAA utilizadas para fins meteorológicos normalmente pode-se tolerar erros de georreferenciamento, não necessitando uma alta precisão geométrica. Porém, em trabalhos que envolvam o mapeamento da superfície terrestre, como os de monitoramento agrícola que utilizam imagens multitemporais, tais erros não são aceitáveis e requerem, obrigatoriamente, um procedimento de georreferenciamento preciso para que os resultados obtidos sejam confiáveis e consistentes durante um período analisado.

ROSBOROUGH et al. (1994) descrevem o método de navegação indireta para georreferenciamento de imagens AVHRR-NOAA. Um pixel com coordenadas geográficas na superfície terrestre é mapeado no campo de visada do satélite, em termos de linha e coluna relacionada a este pixel. O processo inicia-se com a definição de parâmetros da geometria da imagem. Esses parâmetros permitem que seja atribuído um par de coordenadas geográficas para cada pixel da imagem que alimentam o algoritmo da navegação indireta para calcular sua correspondente posição de linha e coluna no campo de visada do satélite. Para isso são consideradas a posição e a velocidade do satélite obtidas por meio dos dados de efemérides propagados por modelos orbitais, prevendo a localização do satélite num dado momento no tempo durante a geração da imagem. Partindo desta localização e do conhecimento prévio da geometria do sensor, a imagem é corrigida para considerar a curvatura e a rotação da Terra, para depois ser reamostrada e registrada numa projeção cartográfica conhecida.

De acordo com EMERY et al. (1989), problemas inevitáveis na predição da posição exata do satélite NOAA devem-se principalmente a variações de atitude provocadas por forças variáveis perturbadoras, atraso do relógio interno e modelos orbitais imprecisos. Estas são as principais fontes de erros de navegação das imagens. Como é intrinsecamente difícil de se prever a contribuição de cada erro, o ajuste da navegação normalmente é realizado numa etapa complementar através de vários pontos de controles coletados manualmente. Esse processamento torna-se inviável devido ao enorme trabalho quando aplicado a uma série multitemporal de imagens. E ainda, quando automatizado, pode não garantir a mesma precisão de georreferenciamento em todas as imagens.

ROSBOROUGH et al. (1994) partem do princípio que a maior fonte dos problemas de georreferenciamento é devida aos erros de atitude do satélite, ou seja, a variação nos três eixos de movimentação da plataforma orbital, denominados *roll* (rotação em torno do próprio eixo), *pitch* (rotação em relação ao eixo horizontal) e *yaw* (rotação em relação ao eixo vertical), no momento da sua passagem, conforme ilustrado na Figura 2.7. Esses erros surgem quando a plataforma orbital não está perfeitamente alinhada aos três eixos do sistema de referência geodésico da Terra, afetando a precisão geométrica das imagens geradas. Cabe ressaltar que os erros de atitude permanecem constantes ao longo de cada passagem do satélite.

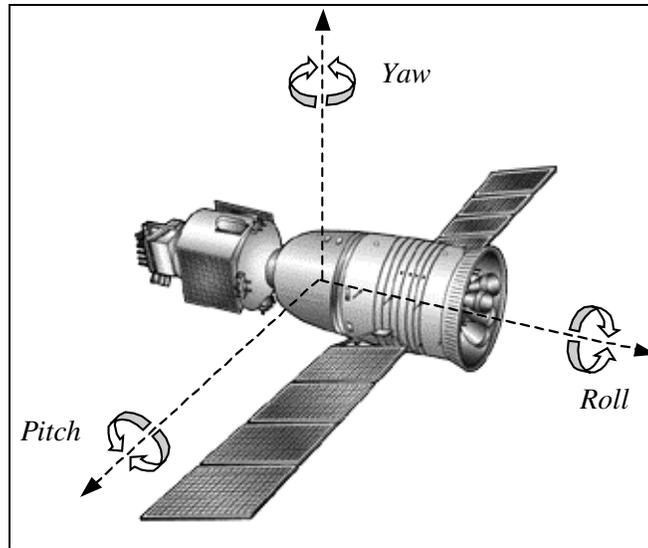


Figura 2.7: Ângulos de atitude da plataforma orbital.

Para originar imagens com precisão geométrica dentro de um pixel, essa metodologia requer o conhecimento dos erros de atitude no momento exato da geração da imagem. Como é possível determinar quanto os pixels diferem das feições geográficas tomadas como pontos de controle, em termos de deslocamento de latitude e longitude, então uma estimativa de erro de atitude do satélite pode ser calculada e, também, o erro equivalente de posição da órbita.

Baseado na primeira navegação da imagem sem qualquer correção dos erros de atitude, o método faz uso normalmente de dois pontos de controle coletados manualmente de imagens geometricamente precisas, para determinar os erros de atitude. A partir dos pares de coordenadas dos pontos de controle são gerados os parâmetros de correção dos erros de atitude. Com isso a imagem é renavegada considerando os erros de atitude, dando origem a uma imagem com georreferenciamento preciso, com erro máximo de um pixel.

A grande vantagem deste método é o reduzido número de pontos de controle necessários para a geração dos parâmetros de correção dos erros de atitude do satélite. Porém, não é totalmente automático, uma vez que necessita de intervenção humana no momento da coleta dos pontos.

PERGOLA e TRAMUTOLI (2000) desenvolveram um método automático de coleta de pontos de controle utilizando regiões costeiras que são automaticamente identificadas na imagem, em função de seu elevado gradiente de variação da reflectância e, então, são comparadas a uma imagem de referência composta por um modelo digital de elevação.

Coletados os pontos de controle, o método de ROSBOROUGH et al. (1994) é aplicado de forma a identificar os erros de atitude do satélite e refazer a navegação da imagem.

PERGOLA e TRAMUTOLI (2003) implementaram um sistema que utiliza tal abordagem, corrigindo progressivamente a posição do satélite e reduzindo os erros de navegação pixel a pixel. Os testes realizados em mais de 400 imagens AVHRR-NOAA de regiões da Europa obtidas durante dois anos, demonstraram a capacidade do sistema em obter de forma automática uma precisão de georreferenciamento sub-pixel, sem intervenção humana.

Uma outra técnica disponível para identificação automática de pontos de controle em imagens AVHRR/NOAA é a chamada Máxima Correlação Cruzada (MCC), aplicada com sucesso no cálculo da velocidade advectiva superficial de correntes oceânicas por EMERY et al. (1986) e, mais recentemente, por BOWEN et al. (2002). O método é aplicado num par de imagens termais, de modo a encontrar feições semelhantes e estimar o movimento de correntes oceânicas entre duas imagens consecutivas, encontrando assim a velocidade de deslocamento resultante dessas correntes. Este método também pode ser aplicado para fins de georreferenciamento, de modo a encontrar automaticamente feições geográficas semelhantes entre uma imagem alvo e a imagem base geometricamente precisa, como detalhado por EMERY et al. (2003) e utilizado por ANTUNES et al. (2005).

2.4.4. Mascaramento de nuvens

As imagens AVHRR-NOAA dificilmente são livres de nuvens em sua totalidade devido ao tamanho do pixel e, por conseqüência, da área de varredura. Em estudos de monitoramento de culturas agrícolas, os pixels contaminados por nuvens não representam a informação real da superfície e devem ser eliminados por meio de algum processo.

CHEN et al. (2003) observaram que pixels contaminados por nuvens causam diferenças em composições máximas de NDVI ao longo de uma série multitemporal de imagens. Baseado nisso, concluem que antes da elaboração de composições máximas é necessário o mascaramento de pixels contaminados por nuvens, como forma de obter produtos quantitativos confiáveis.

2.5. Conceitos de conjuntos e lógica *fuzzy*

A teoria dos conjuntos *fuzzy* (nebulosos), desenvolvida por Lofti A. Zadeh em 1965, é uma extensão da teoria de conjuntos clássicos e está associada aos conceitos básicos de funções de pertinência. A sua utilização é apropriada em sistemas que se caracterizam pela generalidade, ambigüidade e imprecisão para a extração de informações vagas.

A lógica *fuzzy* pode ser definida como a parte da lógica matemática dedicada aos princípios formais do raciocínio incerto ou aproximado, portanto mais próxima do pensamento humano e da linguagem natural. Devido a esta propriedade e à capacidade de realizar inferências, a lógica *fuzzy* tem encontrado grandes aplicações em várias áreas como modelagem de sistemas e reconhecimento de padrões (BEZDEK e PAL, 1992).

Enquanto os conjuntos tratados pela lógica booleana permitem apenas o uso de funções de associação binária, que identificam os elementos pertinentes ou não pertinentes a uma classe, a lógica *fuzzy* admite, também, a possibilidade de uma associação parcial, isto é, elementos que são parcialmente pertinentes a uma ou mais classes.

O conjunto A de valores *fuzzy* do universo X é definido pelos seguintes pares ordenados:

$$A = \{(x, \mu_A(x)) \mid x \in X\}.$$

A função $\mu_A(x)$ é denominada nível de pertinência do ponto x ao conjunto A e determina com que grau um objeto x pertence a um conjunto A . Em conjuntos clássicos, apenas dois valores para $\mu_A(x)$ são permitidos: o elemento pertence (verdadeiro = 1) ou não pertence a um determinado conjunto (falso = 0). Na teoria dos conjuntos *fuzzy*, a transição entre pertencer e não pertencer é gradual, ou seja, é um número real variando de forma contínua no intervalo de zero a um (ZADEH, 1965).

Como exemplo, considere $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ uma coleção de números inteiros. Seja A um conjunto nebuloso que define “números inteiros próximos a 5”, dado por:

$$A = \{(1, 0), (2, 0,4), (3, 0,6), (4, 0,8), (5, 1), (6, 0,8), (7, 0,6), (8, 0,4), (9, 0)\}.$$

O valor 0 representa a exclusão completa ao conjunto, 1 representa a pertinência completa e 0,5 é o ponto de passagem (*cross-over*) de uma classe a outra. Os conjuntos clássicos podem ser vistos como um caso particular dos conjuntos *fuzzy*, nos quais apenas os limites do intervalo são utilizados na definição da função de pertinência.

Segundo DELGADO (2002), considerando-se o universo contínuo $X \in \mathfrak{R}$, o exemplo do conceito de “números próximos a 5” pode ser expresso de forma diferente, dependendo da definição da função de pertinência associada.

Na concepção clássica:

$$\mu_A(x) = \begin{cases} 0 & \text{se } x \leq 4,5 \\ 1 & \text{se } 4,5 < x \leq 5,5 \\ 0 & \text{se } x > 5,5 \end{cases}$$

Na concepção de conjuntos *fuzzy*:

$$\mu_A(x) = \begin{cases} 0 & \text{se } x \leq 4,5 \\ \frac{x - 4,5}{0,5} & \text{se } 4,5 < x \leq 5 \\ \frac{5,5 - x}{0,5} & \text{se } 5 < x \leq 5,5 \\ 0 & \text{se } x > 5,5 \end{cases}$$

A Figura 2.8 ilustra as funções de pertinência no caso de conjuntos clássicos e *fuzzy*.

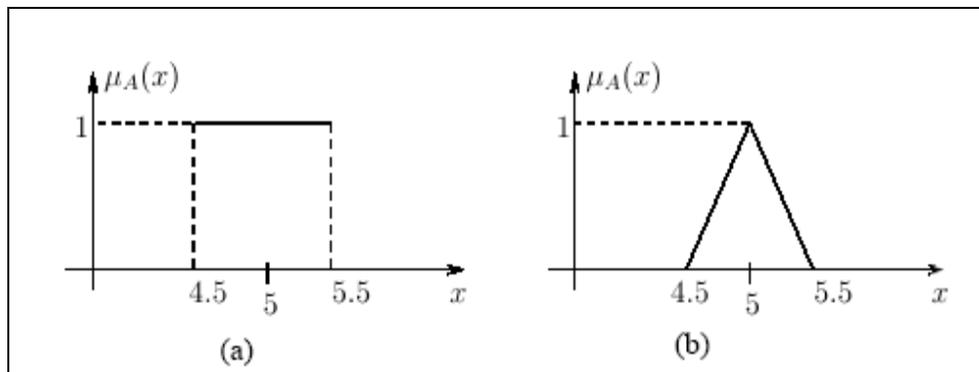


Figura 2.8: Funções de pertinência - conjuntos clássicos (a) e conjuntos *fuzzy* (b).
Fonte: DELGADO (2002).

Em geral, o formato das funções de pertinência é restrito a uma certa classe de funções, representadas por alguns parâmetros específicos. Os formatos mais comuns são:

Função triangular: parâmetros (a, m, b), com $a \leq m \leq b$.

$$\mu_A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{m-a} & \text{se } a < x \leq m \\ \frac{b-x}{b-m} & \text{se } m < x \leq b \\ 0 & \text{se } x > b \end{cases}$$

Função trapezoidal: parâmetros (a, m, n, b), com $a \leq m$, $n \leq b$ e $m < n$.

$$\mu_A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{m-a} & \text{se } a < x \leq m \\ 1 & \text{se } m < x \leq n \\ \frac{b-x}{b-n} & \text{se } n < x \leq b \\ 0 & \text{se } x > b \end{cases}$$

Função Gaussiana: parâmetros (m, σ_k), com $\sigma_k > 0$.

$$\mu_A(x) = \exp\left[-\sigma_k(x-m)^2\right]$$

Uma outra forma bastante utilizada é o conjunto unitário (*singleton*), com parâmetros (m, h).

$$\mu_A(x) = \begin{cases} h & \text{se } x = m \\ 0 & \text{caso contrário.} \end{cases}$$

A Figura 2.9 ilustra os formatos descritos anteriormente, especificando os parâmetros associados a cada caso.

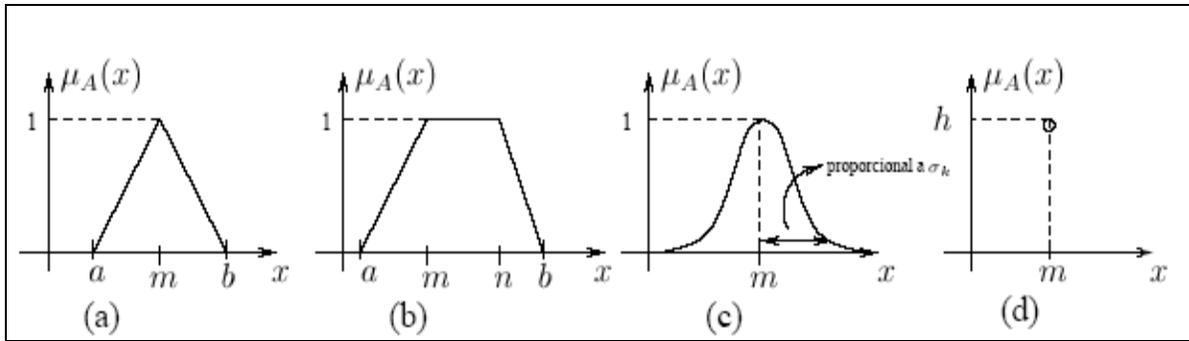


Figura 2.9: Diferentes formatos das funções de pertinência - triangular (a), trapezoidal (b), Gaussiana (c), conjunto unitário (d).
 Fonte: DELGADO (2002).

A escolha do formato mais adequado nem sempre é óbvia. A escolha de funções triangulares e trapezoidais é mais comum porque a idéia de se definirem regiões de pertinência total, média e nula é mais intuitiva do que a especificação do valor médio (m) e dispersão (σ_k), conceitos ligados às funções Gaussianas. Entretanto, as funções Gaussianas são bastante utilizadas devido às suas propriedades matemáticas de continuidade e diferenciação de classes.

De acordo com ZADEH (1990), os níveis de pertinência refletem uma ordem não baseada em probabilidade, mas sim em possibilidade, que são conceitos diferentes. As medidas de probabilidade estão associadas a conjuntos de limites precisos, enquanto que a possibilidade a conjuntos imprecisos. A questão que surge na determinação dos conjuntos *fuzzy* é a escolha de uma função de pertinência apropriada, uma vez que esta função deve ser flexível tanto para modelar a diversidade de formas de um atributo quanto no estabelecimento dos parâmetros do modelo.

2.5.1. Integração da lógica *fuzzy* com geotecnologias

O reconhecimento de padrões é a técnica que tem por objetivo procurar, detectar e explicitar estruturas associadas às regularidades ou propriedades presentes em um conjunto de dados, por meio da análise de classes (BEZDEK e PAL, 1992).

A classificação automática de imagens multiespectrais de sensoriamento remoto significa associar cada pixel da imagem a um rótulo, descrevendo um objeto real da superfície. Os valores de ND associados a cada pixel, definidos pela reflectância dos materiais que

compõem esse pixel, são identificados em termos de um tipo de cobertura da superfície terrestre, determinando as classes.

Existem duas abordagens na classificação de imagens. A primeira denomina-se classificação supervisionada onde o usuário identifica alguns dos pixels pertencentes às classes desejadas e o método localiza todos os demais pixels pertencentes àquelas classes, baseado em alguma regra estatística pré-estabelecida. A segunda chama-se classificação não-supervisionada, onde o próprio método decide, por análise de agrupamento, quais as classes a serem separadas e quais os pixels pertencentes a cada uma (RICHARDS e JIA, 1999).

Segundo WANG (1990a), várias ferramentas utilizadas na elaboração de mapas envolvem operações lógicas (E, OU, NÃO), geralmente realizadas com operadores lógicos dicotômicos, como na lógica clássica, booleana. Os classificadores dicotômicos permitem construir um mapa com pixels sim-não, isto é, cada pixel contém a informação relativa à existência de determinado atributo ou de sua não existência. A informação está representada no método um-pixel-uma-classe que pode não descrever adequadamente a mistura de objetos de distintas classes. Dessa forma, essa limitação pode reduzir o nível de precisão da classificação devido à pobre extração de informações da imagem.

ZHENG et al. (2000) relataram a integração de Sistemas de Informação Geográfica (SIG) com lógica *fuzzy* no mapeamento do uso do solo. Concluíram que as regras da lógica *fuzzy* têm, indubitavelmente, um grande potencial a ser considerado como um modelo inteligente para operações espaciais num SIG, mostrando que não somente o objetivo de classificação foi alcançado, mas também os dados puderam perfeitamente simular a variação de classes e inferir o resultado da classificação.

RANGSANSERI et al. (2000) trataram do problema de classificação de imagens com um algoritmo baseado em lógica *fuzzy*, mostrando a diminuição da imprecisão na modelagem dos dados que podem pertencer a mais de uma classe.

De acordo com WANG (1990a), um sistema classificador *fuzzy*, politômico, expressa o grau de pertinência de um pixel a uma, a outra ou a outras classes. No campo dos SIG, os conjuntos *fuzzy* podem ser utilizados para o processamento de problemas de mapeamento de áreas, ou seja, para determinar a que classe de uso do solo pertence uma área ou se uma área pode ser utilizada para um determinado fim.

Segundo WANG (1990b), a teoria de conjuntos *fuzzy* pode melhorar a representação de informações geográficas. Na representação *fuzzy* para análise de imagens de sensoriamento remoto, as classes de cobertura do solo são definidas como conjuntos *fuzzy* e os pixels os seus elementos. Cada pixel é vinculado com um grupo de graus de pertinência para indicar quanto um pixel pode pertencer a uma determinada classe. Os pixels com mistura de classes ou condições intermediárias podem ser descritos pelos graus de pertinência. Se num pixel existem três tipos de cobertura podem-se ter três graus de pertinência que indicam a associação do pixel com essas três classes, como exemplificado na Figura 2.10.

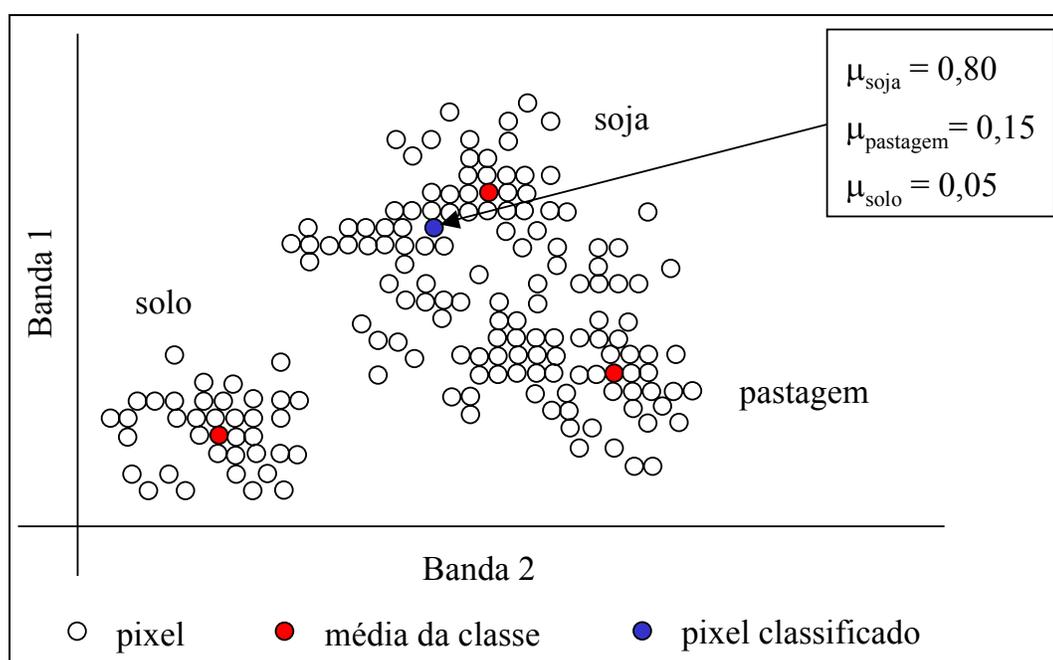


Figura 2.10: Exemplo de graus de pertinência de um pixel.

Os métodos tradicionais de classificação podem gerar resultados extremamente generalizados para o caso de imagens geradas por sensores de baixa resolução espacial como os AVHRR-NOAA. A elevada mistura espectral das classes pode ocorrer quando o tamanho do pixel é maior do que as feições terrestres as quais se deseja extrair informações. No monitoramento de culturas agrícolas, a vegetação pode estar em diferentes condições ao longo do seu ciclo de desenvolvimento, provocando mudanças na quantidade de energia refletida. Entretanto, tais condições podem não serem diferenciadas, a menos que mais classes sejam

definidas, o que torna o processamento mais custoso. Mesmo expandindo-se o número de classes para melhorar a discriminação, a variabilidade dentro de cada classe ainda pode existir.

Nesses casos, uma das técnicas que pode ser empregada é o classificador baseado em lógica *fuzzy* para a solução da mistura espectral pois permite a análise sub-pixel para estimar as proporções de cada superfície dentro dos pixels, gerando, como produto, imagens fracionadas de cada classe (FOODY, 1998).

FOODY e COX (1994) estimaram a composição sub-pixel de cobertura do solo usando os modelos de mistura linear e de regressão baseado em funções de pertinência *fuzzy*, aplicados à imagens Landsat 4/MSS. Imagens de videografia aérea foram classificadas por métodos convencionais e, a seguir, degradadas para gerar imagens de referência. Após a comparação dos resultados, verificou-se que os graus de pertinência *fuzzy* para cada classe de cobertura do solo eram fortemente e significativamente correlacionados com as proporções dessas classes na superfície terrestre dentro de um pixel.

BASTIN (1997) utilizou o modelo de classificação baseado em lógica *fuzzy*, o modelo de mistura linear e os valores de probabilidades derivados da classificação pelo método da Máxima Verossimilhança (MaxVer) para separar classes com mistura espectral em imagens Landsat 5/TM. A partir dessas imagens também foram obtidas imagens de referência por classificação não-supervisionada. Um conjunto de imagens fração de referência para cada classe foi produzido, mostrando a proporção de cada classe em cada pixel. As imagens fração geradas pelos três métodos foram comparadas com as imagens fração de referência e, também, com o resultado da classificação pelo método MaxVer, em que se concluiu que a classificação *fuzzy* obteve a melhor performance para localizar e quantificar os pixels misturados, produzindo uma boa estimativa sub-pixel das classes de cobertura do solo.

ATKINSON et al. (1997) aplicaram a técnica de redes neurais artificiais, o modelo de mistura linear e a classificação baseada em lógica *fuzzy* para mapeamento das proporções sub-pixel de classes de cobertura do solo, utilizando imagens AVHRR-NOAA. Imagens SPOT HRV foram classificadas pelo método MaxVer e utilizadas para gerar as proporções sub-pixel de referência de cada classe para comparação com cada pixel da imagem AVHRR-NOAA, por meio de funções de distribuição de probabilidade. Após a avaliação da acurácia, deduziram que a técnica de redes neurais artificiais produziu o resultado mais preciso e a classificação *fuzzy* teve um desempenho melhor que o modelo de mistura linear.

3. MATERIAL E MÉTODOS

3.1. Material

3.1.1. Área de estudo

A área de estudo é compreendida por oito municípios da região oeste do Paraná que corresponde a mais de 25% do total da área plantada de soja no Estado. A Tabela 3.1 apresenta os dados oficiais de área plantada de soja referente à safra 2003/2004.

Tabela 3.1: Área plantada de soja de oito municípios do oeste do Paraná, safra 2003/2004.

Município	Área Plantada (ha)	
	SEAB/DERAL ¹	IBGE ²
Assis Chateaubriand	72.000,00	73.568,00
Cascavel	82.000,00	78.200,00
Mamborê	54.000,00	54.060,00
Maripá	21.500,00	21.000,00
Palotina	43.300,00	43.600,00
Rancho Alegre do Oeste	15.400,00	15.505,00
Toledo	68.600,00	67.150,00
Ubiratã	51.200,00	51.000,00

¹Fonte: SEAB/DERAL (2005).

²Fonte: IBGE (2005).

A CONAB apenas disponibiliza informações de estimativa de área consolidadas por estado. Para isso, se serve dos dados em nível municipal da Secretaria da Agricultura e Abastecimento do Paraná (SEAB), Departamento de Economia Rural (DERAL), os quais foram utilizados neste trabalho juntamente com os dados municipais do IBGE.

Os municípios monitorados são cobertos pela cena Landsat 5/TM, órbita/ponto 223/77, a ser utilizada como referência na determinação das áreas de soja, conforme mostrado na Figura 3.1.

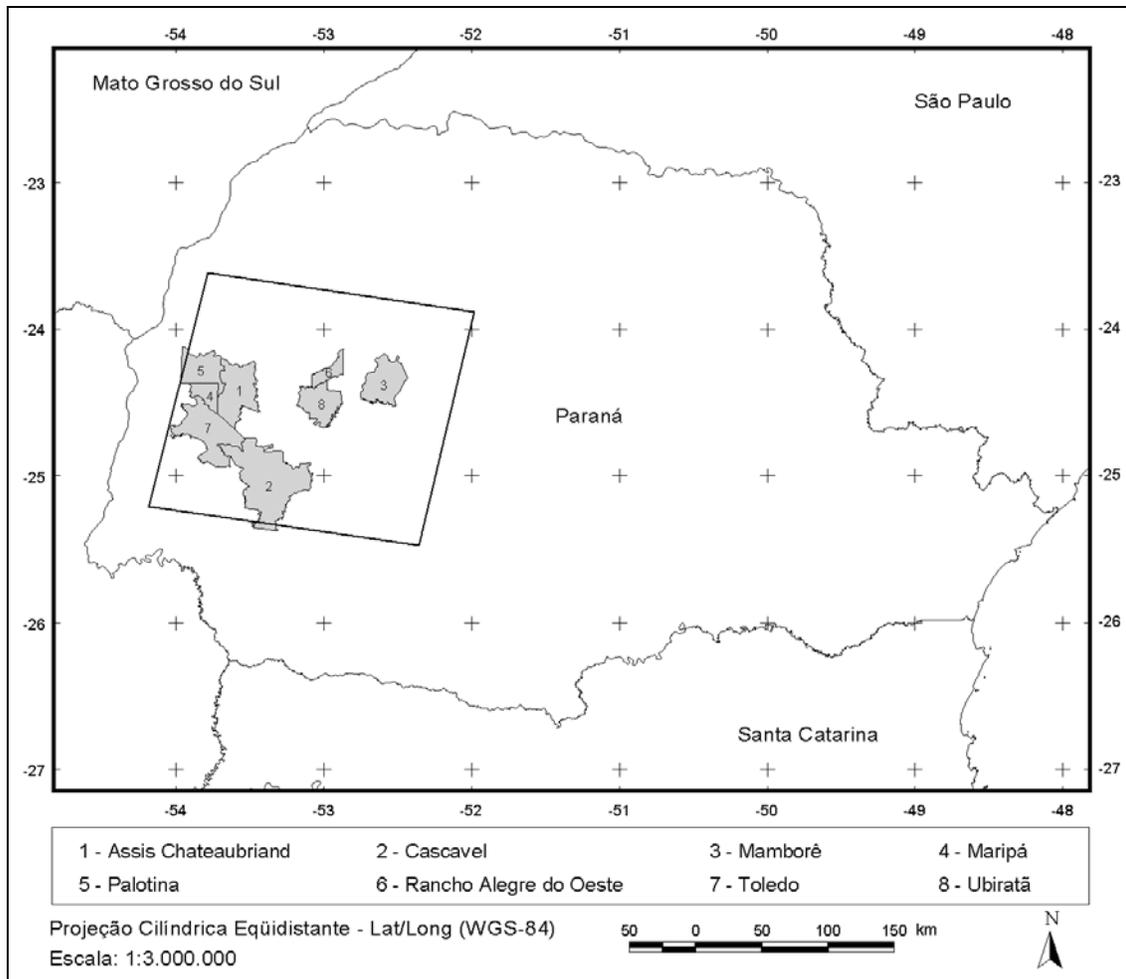


Figura 3.1: Oito municípios da região oeste do Paraná cobertos pela cena Landsat 5/TM, órbita/ponto 223/77.

Esses municípios foram escolhidos pelo fato de já estarem sendo monitorados no projeto GeoSafras - Aperfeiçoamento Metodológico do Sistema de Previsão de Safras no Brasil - coordenado pela CONAB e, também, por serem responsáveis por mais de 11% da área plantada de soja no Paraná. O GeoSafras tem como um dos seus objetivos a geração de metodologias que possam subsidiar estimativas de safras de caráter menos subjetivo e, dessa forma, reduzir incertezas.

3.1.2. Imagens AVHRR-NOAA

As imagens AVHRR-NOAA foram obtidas no acervo histórico do CEPAGRI/UNICAMP que conta com mais de 25.000 imagens armazenadas desde abril de 1995, superando um terabyte de dados. Até 2004, a recepção era feita por meio de uma antena marca Telonics, modelo THRPT-6 T-RIS, série HRPT Earth Stations, posicionada nas coordenadas 22°49'22" S e 47°03'59" W. Foram utilizadas imagens diurnas do satélite NOAA-17, em seu estado bruto, em função da sua passagem ocorrer no período da manhã, aproximadamente 14h GMT, quando a possibilidade de nebulosidade é menor. As datas de aquisição das imagens englobaram a época do ciclo de desenvolvimento da soja na região em estudo, desde os plantios mais precoces, em novembro de 2003, até as colheitas mais tardias, em março 2004.

Para substituição de imagens da série multitemporal devido a problemas de recepção, foi utilizado o *Comprehensive Large Array-data Stewardship System* (CLASS), de onde as imagens foram transferidas no formato padrão Nível 1B (NOAA, 2005). Como última opção, as imagens faltantes foram solicitadas junto ao Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), do Instituto Nacional de Pesquisas Espaciais (INPE), também obtidas no formato padrão Nível 1B.

3.1.3. Imagens Landsat 5/TM

Para determinar as áreas de referência com plantios de soja e auxiliar na análise sub-pixel das imagens AVHRR-NOAA, foram utilizadas seis imagens Landsat 5/TM, órbita/ponto 223/77, resolução radiométrica de 8 bits (0 a 255 ND), adquiridas em 21/11/2003, 07/12/2003, 23/12/2003, 24/01/2004, 09/02/2004 e 25/02/2004, de forma a abranger todo o período do ciclo de desenvolvimento da cultura na região em estudo.

GANAN et al. (2005) procederam ao trabalho de classificação dessas imagens pelos métodos supervisionados Paralelepípedo e MaxVer dando origem ao que foi denominado “máscara de soja”, apresentada na Figura 3.2, destacando os oito municípios produtores na região oeste do Paraná.

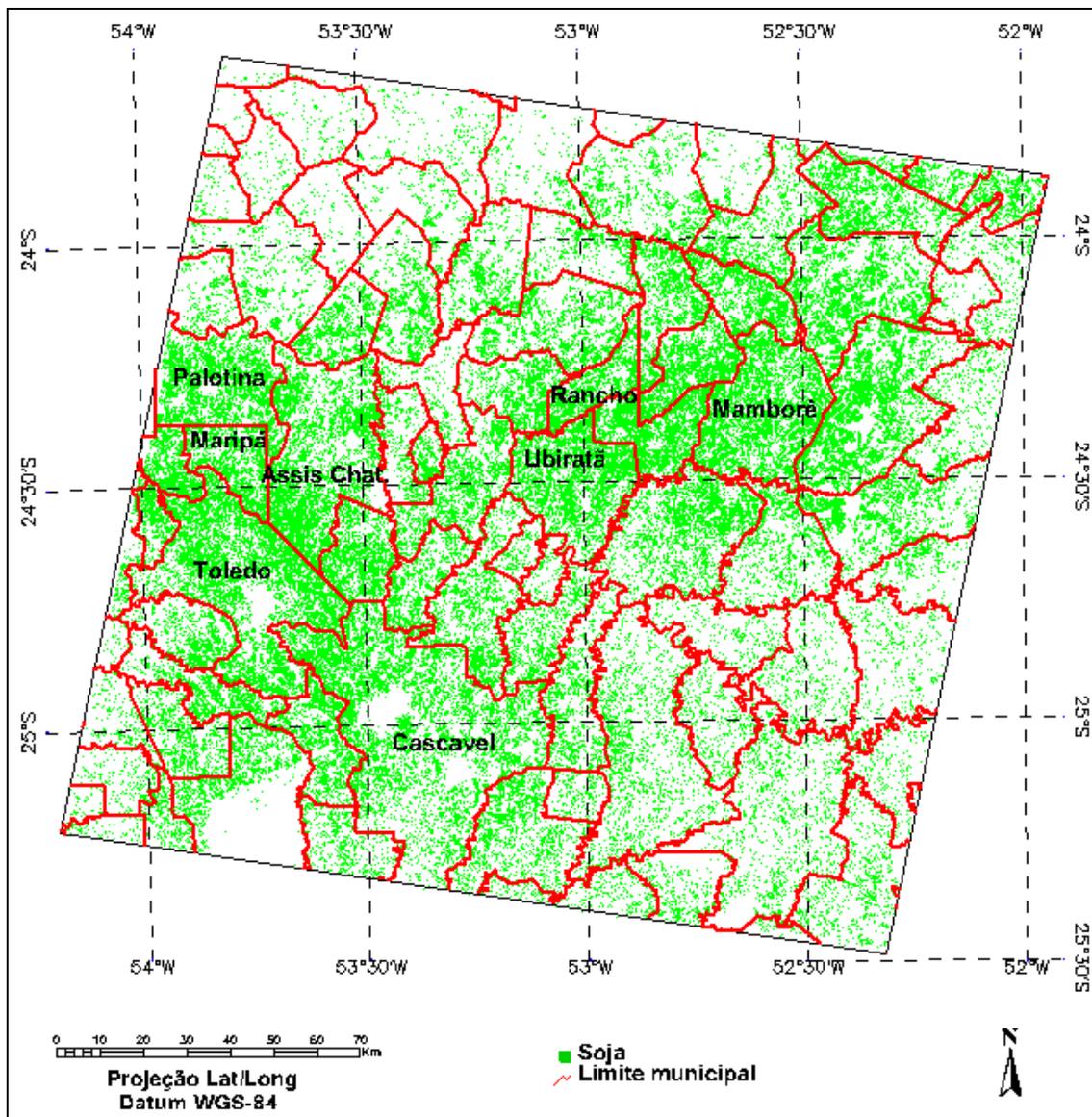


Figura 3.2: Máscara de soja da região oeste do Paraná referente à safra 2003/2004.

3.1.4. Programas computacionais e microcomputadores

Os programas para o processamento das imagens AVHRR-NOAA envolvendo as etapas de conversão do formato bruto, calibração radiométrica e georreferenciamento preciso foram os disponíveis no pacote de software NAV (*NAVigation*), desenvolvido pelo Dr. Daniel G. Baldwin, em linguagem C, sob a plataforma Linux. Essa tecnologia foi repassada gratuitamente pelo Colorado Center for Astrodynamics Research (CCAR), Aerospace Engineering Sciences, da Universidade do Colorado, Boulder, EUA.

A partir disso foi implementado um sistema em script *c-shell*, numa atividade conjunta com o aluno de doutorado Júlio César Dalla Mora Esquerdo, da FEAGRI/UNICAMP, cuja primeira parte é composta por um conjunto de comandos que acionam os módulos do NAV.

A segunda parte do sistema refere-se à geração de produtos e foi implementada utilizando-se a linguagem de programação IDL 4.0 (*Interactive Development Language*) do software ENVI 6.0 (*The Environment for Visualizing Images*), também acionada por meio de comandos do script *c-shell*.

O tratamento dos dados, a análise estatística e a geração de gráficos foi feita com a planilha Microsoft Excel 2000.

As atividades foram realizadas em dois computadores Pentium 4, com CPU de 2,4 GHz, 512 de RAM e 80 GB de disco rígido, um com o sistema operacional Linux Fedora Red Hat 9.0 e outro com o Microsoft Windows 2000 Professional, localizados no Laboratório de Geoprocessamento (LabGEO) da FEAGRI/UNICAMP.

3.1.5. Informações cartográficas

Foram utilizadas informações cartográficas dos limites estaduais e municipais do Paraná em formato DXF obtidos do IBGE, da malha referente ao ano 2001.

3.2. Métodos

3.2.1. Processamentos das imagens AVHRR-NOAA

A primeira parte da metodologia consistiu na execução do sistema desenvolvido para processamento das imagens AVHRR-NOAA utilizando os módulos do NAV e, também, na geração dos produtos quantitativos feita com programas implementados em IDL/ENVI. A Figura 3.3 apresenta o fluxograma das etapas que compõem o sistema de processamento das imagens e geração de produtos e, a seguir, a descrição de cada uma delas.

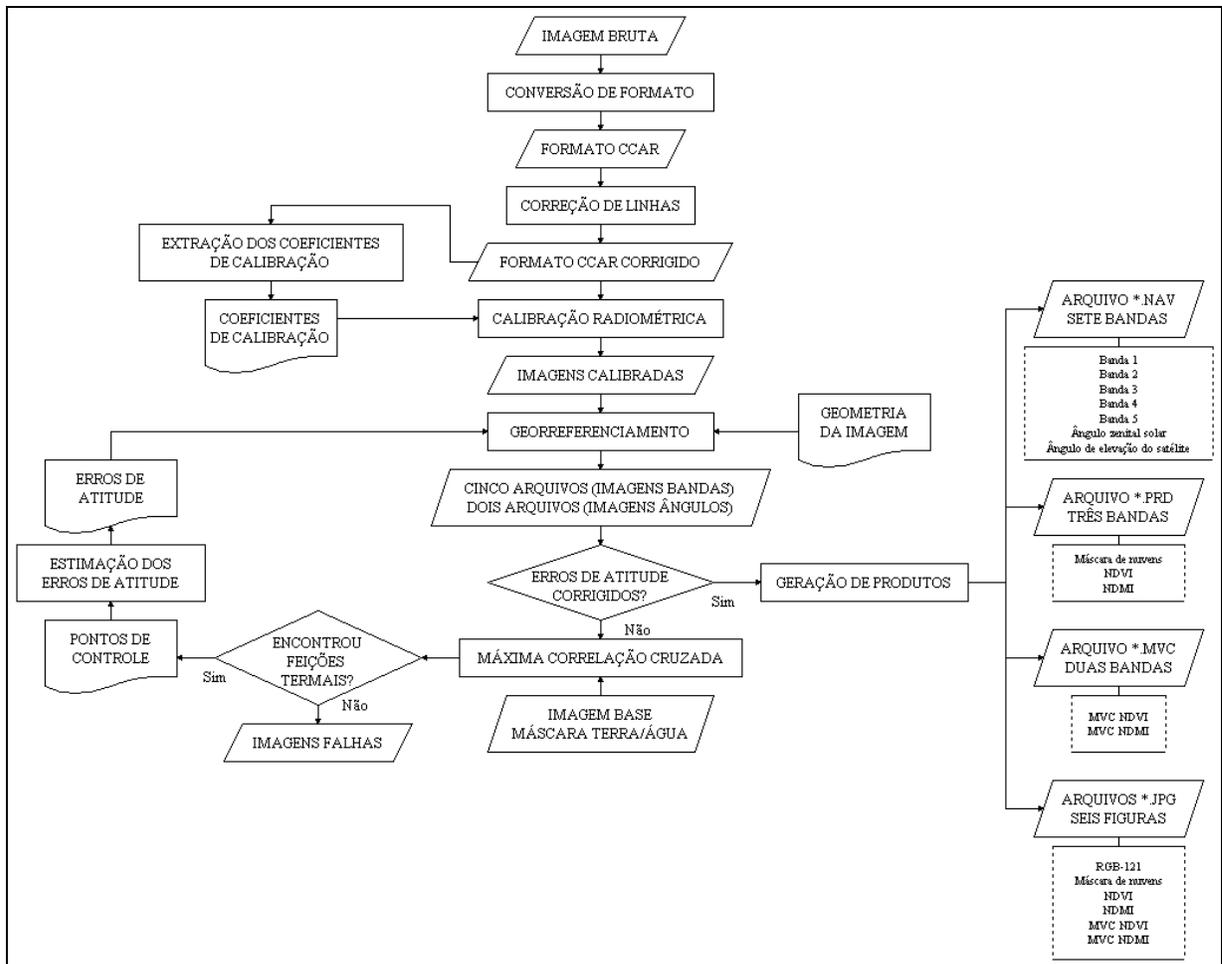


Figura 3.3: Fluxograma do sistema de processamento e geração de produtos das imagens AVHRR-NOAA.

3.2.1.1. Conversão do formato bruto

O processo iniciou-se com a conversão da imagem no formato bruto Telonics T-RIS ou Nível 1B para um formato genérico denominado CCAR, compatível com os módulos subsequentes do NAV. Falhas de recepção ocasionam a perda de linhas na imagem e a ausência desses valores afeta a precisão do georreferenciamento, uma vez que o procedimento requer necessariamente seis linhas por segundo para a realização do processo de navegação. Por causa disso, logo após a conversão, as linhas sem valores são compensadas com a atribuição de zeros. O novo número de linhas é recalculado e atualizado no cabeçalho da imagem.

3.2.1.2. Calibração radiométrica

Em seguida foi executado o processo de calibração radiométrica das imagens, onde os valores de ND foram transformados em reflectância aparente (no topo da atmosfera) para as bandas 1, 2 e 3A e temperatura de brilho, para as bandas 3B, 4 e 5, respectivamente, conforme os procedimentos descritos a seguir, válidos para os satélites da série a partir do NOAA-15.

- Calibração das bandas do vermelho, do infravermelho próximo e do infravermelho médio (1, 2 e 3A)

Devido à degradação do sensor AVHRR e como não existem equipamentos de calibração a bordo, foram desenvolvidas técnicas alternativas de calibração radiométrica pós-lançamento das bandas reflectivas dos satélites NOAA. Essas técnicas utilizam alvos terrestres radiometricamente estáveis e cálculos de predição da órbita do satélite para monitorar a calibração dos instrumentos em função do tempo.

O sensor AVHRR versão 3 usa um circuito de ganho dual para baixos níveis de luminosidade. Essa característica proporciona um aumento da resolução na detecção das radiâncias no intervalo mais baixo de albedo, de 1% a 25% para as bandas 1, 2 e 1% a 12,5% na banda 3A. Isso resulta numa menor perda da resolução nos valores mais elevados de albedo, de 26% a 100% para as bandas 1, 2 e 12,6% a 100% para a banda 3A. A metodologia de calibração radiométrica de uma determinada banda dá origem a dois valores de albedo, alto e baixo, como forma de acomodar os dois intervalos de ganhos.

A soma desses dois valores resulta no valor de albedo (fator de reflectância) total para a banda i , conforme as Equações (3.1), (3.2) e (3.3) de RAO e CHEN (1996):

$$A_{i,\text{baixo}} = S_{i,\text{baixo}} \times C_{i,\text{baixo}} + I_{i,\text{baixo}} \quad (3.1)$$

$$A_{i,\text{alto}} = S_{i,\text{alto}} \times C_{i,\text{alto}} + I_{i,\text{alto}} \quad (3.2)$$

$$A_i = A_{i,\text{baixo}} + A_{i,\text{alto}} \quad (3.3)$$

onde,

$A_{i,alto}$ = albedo no intervalo alto da banda i (%);
 $A_{i,baixo}$ = albedo no intervalo baixo da banda i (%);
 A_i = albedo total (%);
 $S_{i,baixo}$ = coeficiente *slope* no intervalo baixo da banda i ($mW/(m^2 \cdot sr \cdot cm^{-1})/ND$);
 $I_{i,baixo}$ = coeficiente *intercept* no intervalo baixo da banda i ($mW/(m^2 \cdot sr \cdot cm^{-1})$);
 $C_{i,baixo}$ = ND da banda i no intervalo baixo de 0 a 500;
 $S_{i,alto}$ = coeficiente *slope* no intervalo alto da banda i ($mW/(m^2 \cdot sr \cdot cm^{-1})/ND$);
 $I_{i,alto}$ = coeficiente *intercept* no intervalo alto da banda i ($mW/(m^2 \cdot sr \cdot cm^{-1})$);
 $C_{i,alto}$ = ND da banda i no intervalo alto de 501 a 1023.

A NOAA monitora as inclinações duais de calibração e os valores de *slope* e *intercept* altos e baixos das bandas 1, 2 e 3A são calculados em função do tempo do satélite em órbita e atualizados periodicamente em sua página na Internet (NOAASIS, 2004).

- Calibração das bandas termiais (3B, 4 e 5)

Segundo GOODRUM et al. (2004), os ND da imagem para as bandas termiais são convertidos primeiramente em radiância por um modelo linear, conforme a Equação (3.4):

$$R_{i,lin} = S_i \times C_i + I_i \quad (3.4)$$

onde,

$R_{i,lin}$ = radiância da banda i obtida de forma linear (%);
 S_i = coeficiente *slope* da banda i ($mW/(m^2 \cdot sr \cdot cm^{-1})/ND$);
 I_i = coeficiente *intercept* da banda i ($mW/(m^2 \cdot sr \cdot cm^{-1})$);
 C_i = ND da banda i .

O sensor AVHRR realiza medições de radiância de cada pixel na superfície e, ao mesmo tempo, do espaço e de um alvo interno de referência, com radiância conhecida. Os valores de S_i e I_i são obtidos com base nesses dois pontos, o que possibilita uma calibração continuada das bandas termiais. Esses valores estão armazenados no arquivo da imagem para cada um das bandas termiais e são extraídos por um módulo do NAV.

Os dados de radiância devem ser submetidos a um procedimento de correção que considera a não-linearidade entre os sinais produzidos pelo sensor. Para a obtenção da radiância corrigida, utiliza-se a Equação (3.5):

$$R_i = a_0 + a_1 \times R_{i,lin} + a_2 \times (R_{i,lin})^2 \quad (3.5)$$

onde,

R_i = radiância da banda i corrigida de forma não-linear ($W/(m^2 \cdot sr \cdot cm^{-1})$).

Os coeficientes a_0 , a_1 e a_2 estão presentes em cada linha do arquivo da imagem e são extraídos por meio de uma rotina do pacote NAV. O detector da banda 3B responde linearmente à radiância e, por isso, o coeficiente a_2 sempre será zero.

A conversão da radiância corrigida em temperatura de brilho pode ser calculada pela Equação Inversa da Radiação de Planck (3.5):

$$T_{Ei} = \frac{C_2 \nu_i}{\ln\left(1 + \frac{C_1 \nu_i^3}{R_i}\right)} \quad (3.6)$$

onde,

T_{Ei} = temperatura de brilho estimada da superfície da banda i (K);

ν_i = valor correspondente à porção central do comprimento de onda da banda i (cm^{-1});

$C_1 = 1,1910427 \times 10^{-5} \text{ mW}/(m^2 \cdot sr \cdot cm^{-4})$;

$C_2 = 1,4387752 \text{ cm} \cdot K$.

Após isso, deve ser realizado um ajuste por meio da Equação (3.7), para calcular a temperatura efetiva:

$$T_i = \frac{T_{Ei} - A}{B} \quad (3.7)$$

onde,

T_i = temperatura de brilho efetiva da superfície da banda i (K);

A e B são coeficientes de ajuste da temperatura de brilho das bandas 3B, 4 e 5 que podem ser obtidos em GOODRUM et al. (2004).

3.2.1.3. Georreferenciamento preciso

A partir da imagem calibrada radiometricamente foi executado o módulo de navegação indireta baseada na metodologia proposta por ROSBOROUGH et al. (1994) que faz a correção dos deslocamentos de georreferenciamento a partir da estimativa dos erros de atitude do satélite, combinando os passos de transformação relativa ao modelo de geometria orbital e reamostragem dos pixels por interpolação com o vizinho mais próximo.

O primeiro passo foi realizar uma navegação inicial, sem considerar os erros de atitude do satélite, com a definição dos parâmetros de geometria da imagem como o nome das imagens de entrada e saída, bandas a serem processadas, latitude e longitude do ponto central da imagem, número de linhas e colunas que englobam a área de interesse, projeção cartográfica, dentre outros opcionais. Deste processamento resultam cinco imagens individuais navegadas de cada banda e outras duas referentes ao ângulo zenital solar e ao ângulo de elevação do satélite para cada pixel da imagem, contendo apenas a região de interesse.

Para automatizar o processo de coleta de pontos de controle, o módulo utilizou a MCC na busca por feições termais semelhantes entre a imagem alvo que se pretende corrigir e uma imagem base geometricamente precisa, baseado na metodologia descrita por EMERY et al. (2003), conforme ilustrado na Figura 3.4.

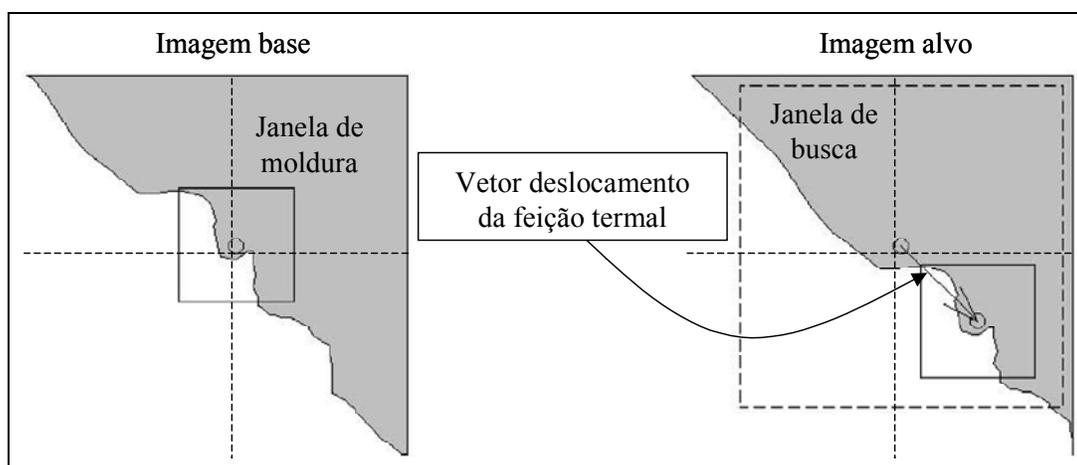


Figura 3.4: Janela de moldura e de busca da MCC.

O método computacional consiste na identificação da MCC entre matrizes que definem sub-áreas de um par de imagens divididas em janelas de moldura e de busca. Da primeira imagem extrai-se a sub-área que define a janela de moldura, cujo tamanho pode variar de acordo com a área de estudo. Na segunda imagem, uma sub-área mais ampla é selecionada e definida como janela de busca, cujo ponto central inicial coincide com o da janela de moldura. O algoritmo aplica um filtro de nuvens para assegurar que todos os pixels da imagem usados pela MCC não estão contaminados com nuvens nas imagens base e alvo. Nesse processo também é utilizada uma máscara terra (valor 1)/água (valor 0) para que a busca por feições seja realizada apenas sobre a superfície terrestre.

Uma vez reconhecidas as feições termais semelhantes entre as imagens base e alvo, o programa coleta automaticamente os pontos de controle. Para cada deslocamento realizado entre as duas imagens é calculada a correlação cruzada. O conjunto de vetores das feições que possuem alta correlação (maior que 0,7) é convertido em deslocamentos ortogonais de latitude e longitude para estimar os erros de atitude do satélite. Uma vez determinados tais erros, a segunda navegação é executada com os parâmetros de correção dos parâmetros de atitude conforme a disposição espacial dessas feições. Em seguida a imagem é ajustada a partir da média dos vetores de deslocamentos para corrigir eventuais erros residuais, gerando, deste modo, imagens registradas com georreferenciamento preciso dentro de um pixel.

A imagem base é de fundamental importância para o sucesso do processamento automático de georreferenciamento das imagens AVHRR-NOAA e por isso deve apresentar elevada precisão geométrica. O georreferenciamento da imagem base foi realizado por meio de no mínimo dois pontos de controle coletados manualmente, utilizando-se imagens do sensor MODIS (*Moderate Resolution Imaging Spectroradiometer*) que apresentam boa qualidade cartográfica. A imagem base deve ser da banda 4 do termal, porque sofre menor influência devido às condições de iluminação e por isso as feições termais são mais estáveis ao longo do tempo. Além disso, deve conter o mínimo de nuvens possível e ser próxima ao nadir.

A imagem alvo também deve ser da banda 4 do termal, ser temporalmente próxima à imagem base para que a distribuição de radiâncias entre as imagens seja similar e, também, não estar totalmente coberta por nuvens, o que implicaria no fracasso pela busca por feições termais semelhantes. Portanto, imagens da safra de verão deverão ser corrigidas com uma imagem base da mesma estação para evitar o efeito sazonal devido às mudanças térmicas. A

utilização de uma imagem base distante da época do estudo pode resultar em falhas na coleta de pontos de controle, por se tratarem de imagens com gradientes de temperatura muito diferentes. Os pontos de controle coletados a partir da banda 4 do termal são utilizados para a refazer a navegação das demais bandas da imagem.

3.2.1.4. Geração de produtos

Os arquivos contendo as bandas individuais das imagens não eram compatíveis com o software ENVI. A etapa final do processamento contava com rotinas desenvolvidas em linguagem IDL/ENVI para geração dos produtos quantitativos.

O primeiro produto era um arquivo contendo sete bandas, reunindo os sete arquivos gerados num único arquivo. Uma rotina foi implementada para extração das informações contidas no cabeçalho da imagem com a finalidade de definir suas características geométricas. Os dados calibrados da imagem de saída foram gravados em números inteiros, multiplicados por dez, para redução do tamanho final do arquivo. O mesmo ocorreu com os dados referentes ao ângulo zenital solar e ao ângulo de elevação do satélite.

O segundo produto gerado pela rotina IDL/ENVI foi um arquivo contendo a máscara de nuvens e os índices de vegetação. A máscara de nuvens foi criada utilizando-se uma combinação para satisfazer os critérios das Equações (3.8), (3.9) e (3.10), descritos por FRANÇA e CRACKNELL (1995) e por CHEN et al. (2002), após várias iterações e testes levando-se em consideração as características das imagens da área de estudo:

$$1^{\circ} \text{ critério: } \rho_{\text{VER}} > 15 \quad (3.8)$$

$$2^{\circ} \text{ critério: } 0,8 \leq \frac{\rho_{\text{IVP}}}{\rho_{\text{VER}}} \leq 1,6 \text{ e } T_4 < 270 \quad (3.9)$$

$$3^{\circ} \text{ critério: } T_5 < 280 \quad (3.10)$$

onde,

ρ_{VER} = reflectância no vermelho (%);

ρ_{IVP} = reflectância no infravermelho próximo (%);

T_4 = infravermelho termal 4 (K);

T_5 = infravermelho termal 5 (K).

Os índices de vegetação gerados para cada imagem foram o NDVI e o NDMI. Seguindo as recomendações de CHEN et al. (2003), a rotina para cálculo desses índices apresentou a opção para mascarar os pixels com ângulo zenital solar maior que 70° e ângulo de elevação do satélite maior que 42°, considerados inaptos para análise espectral, devido à baixa luminosidade solar e distorção do tamanho do pixel por estar muito fora do nadir.

Uma outra rotina IDL/ENVI gerou automaticamente o terceiro produto referente às composições máximas dos índices de vegetação a partir de cada imagem. Após a definição de um determinado período que poderia ser semanal, decenal, quinzenal ou mensal, a rotina fez a leitura automática de todos os arquivos produzidos no período e gerou um arquivo contendo o valor mais alto para cada pixel.

Além dos arquivos com informação espectral, ao final foram geradas figuras no formato JPEG (*Joint Pictures Expert Group*), sobrepostas com a malha do IBGE e com informações cartográficas: composição colorida RGB-121, máscara de nuvens, NDVI, NDMI e as composições máximas de NDVI e NDMI, estas últimas num período escolhido. Nas figuras relacionadas aos índices de vegetação pode ser feito um escalonamento dos valores de 0 a 255 e aplicada uma paleta de cores para melhorar a visualização da imagem.

A cultura da soja foi monitorada ao longo do seu ciclo de desenvolvimento por meio de composições máximas de NDVI e NDMI obtidos a partir das imagens AVHRR-NOAA.

3.2.2. Estimativas de área

A segunda parte da metodologia consistiu na realização da classificação *fuzzy* utilizando os índices de vegetação de imagens AVHRR-NOAA no período de maior vigor vegetativo da cultura na região em estudo, como forma de obter a proporção de soja em cada pixel e assim estimar a área plantada. Os resultados encontrados foram comparados com a estimativa de área de soja de referência obtida a partir das imagens Landsat 5/TM e as estimativas oficiais da SEAB/DERAL e do IBGE.

Como a análise foi realizada de forma pareada entre duas estimativas de área, então o erro relativo é equivalente ao coeficiente de variação. Baseado nisso, para avaliação dos resultados das comparações foi adotada a classificação de PIMENTEL-GOMES (2000), em que o erro relativo é considerado baixo quando for menor que 10%, médio de 10% a 20%, alto de 20% a 30% e muito alto quando for maior que 30%.

A Figura 3.5 apresenta o fluxograma das etapas utilizadas para obtenção das estimativas de área de soja e, a seguir, a descrição de cada um delas.

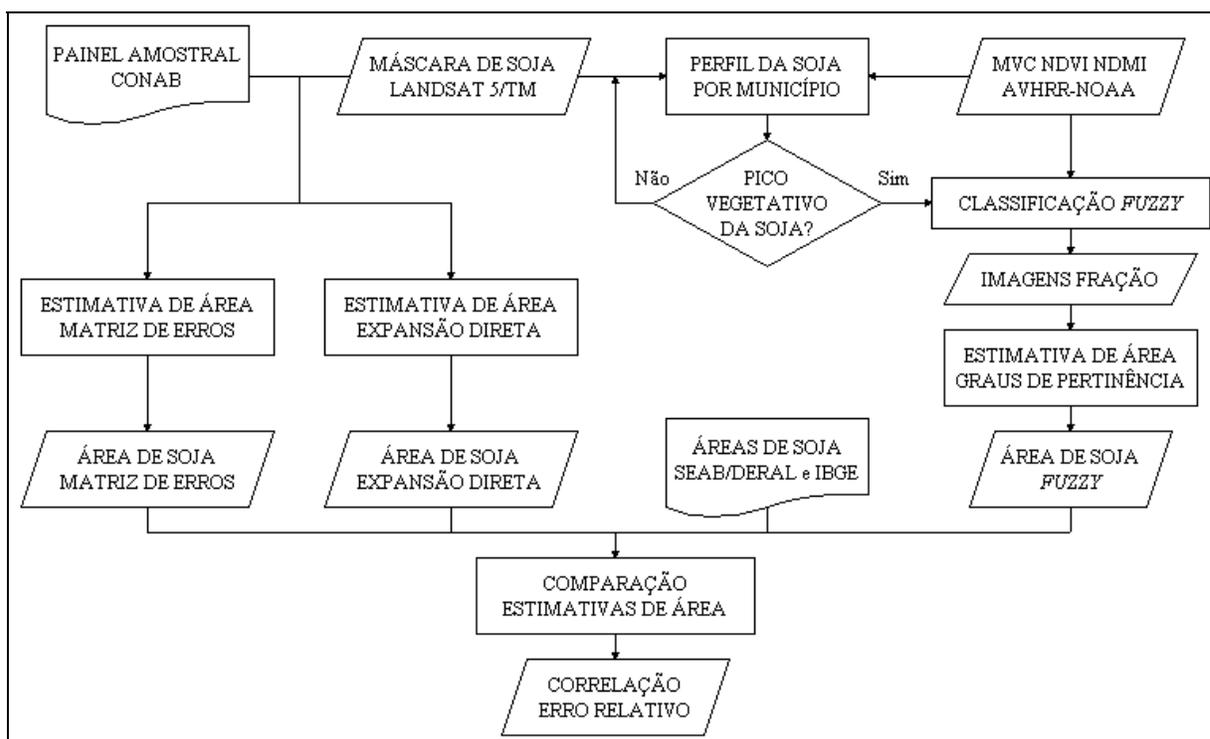


Figura 3.5: Fluxograma das etapas utilizadas para estimativas de área de soja.

3.2.2.1. Estimativa de área de soja de referência por meio de matriz de erros

Devido ao limitante de acurácia inerente dos classificadores booleanos, a área dos oito municípios da máscara de soja foi corrigida primeiramente por meio de matrizes de erros, elaboradas a partir dos pontos do painel amostral da CONAB no Paraná, tomados como referência terrestre.

O pixel de uma imagem de satélite é definido como o elemento amostral regularmente espaçado, de área conhecida, fixa e passível de identificação e de localização. A

classe a que ele pertence é a variável de interesse. A seleção dos pontos foi realizada por meio de amostragem aleatória simples que consistiu na seleção de uma amostra do total populacional de tal forma que qualquer possível amostra tivesse uma mesma probabilidade de ser escolhida. Com a distribuição aleatória dos pontos amostrais sobre as imagens dentro dos limites dos municípios e o posterior levantamento da classe de uso do solo de cada ponto, define-se um painel amostral (LUIZ, 2003).

Por fornecer uma boa relação custo-benefício, foram utilizados 100 pontos amostrais por município, onde foi possível fazer um levantamento de campo com menor gasto de recursos financeiros e em menor tempo, com um nível de confiança aceitável para a estimativa de área (EPIPHANIO et al., 2002). A Figura 3.6 ilustra os pontos amostrais sorteados no município de Cascavel e a Tabela 3.2 a identificação de ocupação do solo de alguns deles, referente à safra 2003/2004.

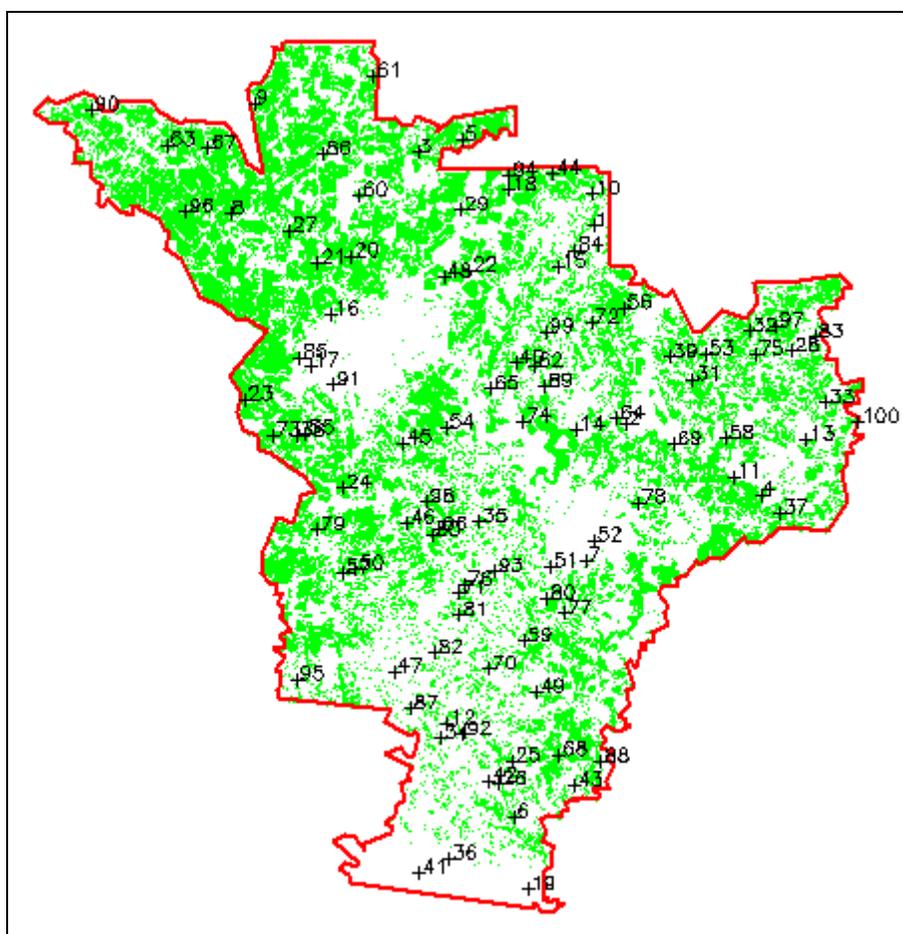


Figura 3.6: Pontos amostrais no município de Cascavel.

Tabela 3.2: Ocupação do solo dos pontos amostrais de Cascavel.

Ponto	Latitude	Longitude	Ocupação do Solo
1	-24,876	-53,284	mata
2	-25,019	-53,259	soja
3	-24,823	-53,422	soja
4	-25,070	-53,153	soja
5	-24,814	-53,387	mata
6	-25,302	-53,346	pasto
7	-25,117	-53,290	mata
8	-24,867	-53,569	mata
9	-24,787	-53,551	mata
10	-24,853	-53,285	soja
11	-25,057	-53,174	soja
12	-25,234	-53,399	mata
13	-25,030	-53,117	milho
14	-25,023	-53,298	mata
15	-24,905	-53,312	soja
16	-24,940	-53,490	área urbana
17	-24,976	-53,506	área urbana
18	-24,850	-53,351	mata
...

A Tabela 3.3 representa uma matriz de erros para as classes Soja e Não Soja, no padrão definido por RICHARDS e JIA (1999), com as colunas contendo dados de referência e linhas representadas pela classificação da imagem, construída para a correção das estimativas de áreas.

Tabela 3.3: Matriz de erros para as classes Soja e Não Soja.

Máscara (i)	Pontos de Referência (j)		Total Linha
	Soja (1)	Não Soja (2)	
Soja (1)	n_{11}	n_{12}	N_1
Não Soja (2)	n_{21}	n_{22}	N_2
Total			100

Sendo que:

n_{ij} = número de observações da classe de referência j ocorrendo na classe espectral i , na amostra;

N_i = número total de observações que ocorrem na classe espectral i , na amostra.

O método proposto por JUPP (1989), normaliza a matriz de erros pelas linhas e por isso é denominado direto. A probabilidade de um pixel que foi previamente atribuído a uma certa classe na realidade pertencer a todas as classes examinadas pode ser estimada pela padronização das linhas e o total de pixels das classes de uso do solo determinadas pela imagem de satélite. Dessa forma, o método direto é considerado mais natural para corrigir as probabilidades de designação das classes por meio das matrizes de proporções.

A matriz de proporções para as classes Soja e Não Soja representada na Tabela 3.4 é gerada pela divisão dos elementos das linhas por N_i , conforme a Equação (3.11):

$$p(r_j | c_i) = n_{ij}/N_i \quad (3.11)$$

onde,

$p(r_j | c_i)$ = proporção da classe de referência j representada na classe espectral i , na amostra.

Tabela 3.4: Matriz de proporções para as classes Soja e Não Soja.

Máscara (i)	Pontos de Referência (j)	
	Soja (1)	Não Soja (2)
Soja (1)	$p(r_1 c_1) = n_{11}/N_1$	$p(r_2 c_1) = n_{12}/N_1$
Não Soja (2)	$p(r_1 c_2) = n_{21}/N_2$	$p(r_2 c_2) = n_{22}/N_2$

Se o número total de pixels da classe espectral i na máscara de soja é T_i , por município, então a área estimada corrigida da classe de referência j em pixels é calculada pela Equação (3.12):

$$A_j = \sum_{i=1}^c p(r_j | c_i) * T_i \quad (3.12)$$

onde,

c = número total de classes.

Sejam:

T_1 = total de pixels da classe Soja na máscara, por município;

T_2 = total de pixels da classe Não Soja na máscara, por município.

Portanto, as áreas estimadas corrigidas das classes Soja e Não Soja são calculadas pelas Equações (3.13) e (3.14), respectivamente:

$$A_1 = p(r_1 | c_1) * T_1 + p(r_1 | c_2) * T_2 \quad (3.13)$$

$$A_2 = p(r_2 | c_1) * T_1 + p(r_2 | c_2) * T_2 \quad (3.14)$$

Para uma simples checagem neste caso específico, a soma da área das classes Soja e Não Soja tem que ser igual a área do município, conforme a Equação (3.15):

$$A_{\text{Município}} = A_1 + A_2 \quad (3.15)$$

Diante do exposto, as matrizes de erros são transformadas em matrizes de proporções com o intuito de corrigir todas as probabilidades de designação das classes simultaneamente e, assim, corrigir tendências nas estimativas de área.

3.2.2.2. Estimativa de área de soja de referência por expansão direta

As estimativas de área baseadas em levantamentos de campo podem conter erros de amostragem não tendenciosos. As áreas medidas por contagem de pixels podem ser tendenciosas por causa dos erros na classificação digital. A partir disso, uma outra relação entre a área da cultura e o resultado da classificação digital que deu origem à máscara de soja pode ser esperada. Isso representa a oportunidade de combinar essas duas fontes de informações para se obter estimativas de área agrícolas por amostragem, utilizando o método por expansão direta (COCHRAN, 1977).

Segundo LUIZ (2003), o estimador por expansão direta é aquele que, baseando-se apenas nos dados amostrais, reproduz para toda a população as proporções relativas encontradas nas amostras. Este método de estimativa de área plantada com uma certa cultura num determinado município combina os princípios estatísticos da teoria da amostragem com as características inerentes aos dados de sensoriamento remoto.

De acordo com EPIPHANIO et al. (2002), a estimativa por expansão direta utiliza os dados coletados nas imagens e em levantamento de campo a partir do painel amostral, a respeito das diferentes coberturas vegetais para os pontos amostrais de referência terrestre, permitindo rápida avaliação da área de uma certa classe de uso do solo. A proporção com que uma classe ocorre na amostra passa a equivaler à proporção de área amostral ocupada por essa

classe. Então, pressupõe-se que a mesma proporção dos pontos da classe encontrada no painel amostral vá se reproduzir na área total do município, numa expansão direta.

De forma resumida, a estimativa por expansão direta também utilizando a máscara de soja consiste na multiplicação da proporção de pontos identificados como soja a partir da referência terrestre pela área municipal, calculada pela Equação (3.16):

$$A_{\text{Soja}} = (p_{\text{Soja}} / p_m) * A_m \quad (3.16)$$

onde,

A_{Soja} = área estimada da cultura de soja no município (ha);

p_{Soja} = número de pontos amostrais identificados como soja no município;

p_m = número total de pontos amostrais no município, igual a 100 pontos neste estudo;

A_m = área total do município (ha).

3.2.2.3. Estimativa de área de soja baseada em classificação *fuzzy*

O método de classificação não supervisionado aplicado nas composições máximas dos índices de vegetação de imagens AVHRR-NOAA combina o método de agrupamento por partição *Fuzzy K-means* (FKM), proposto por BEZDEK et al. (1984), com a medida de similaridade adaptativa conhecida como *Expectation Maximization* (EM), definida por GATH e GEVA (1989).

Segundo TSO e MATHER (2001), o algoritmo trabalha a partir de um procedimento iterativo onde os pontos (pixels da imagem), inicialmente em posição aleatória, são classificados em grupos (classes de uso do solo). Dado o número de grupos desejados são calculados os centros de cada grupo com base na média dos atributos dos pontos. Num próximo passo os pontos são realocados entre os grupos de acordo com a similaridade entre eles. Como o algoritmo de agrupamento FKM é não paramétrico, não faz nenhuma suposição estatística sobre a distribuição dos dados.

Seja $X = \{x_1, x_2, \dots, x_n\}$ um subconjunto de \mathfrak{R}^d , onde d é dimensão do espaço de atributos, n o número de pontos e K é o número de grupos, com $2 \leq K \leq n$. Deste modo, uma partição *fuzzy* de X pode ser representada por uma matriz U de dimensão $K \times n$, onde cada elemento é denotado por μ_{ki} que define o grau de pertinência do ponto x_i com relação a cada

um dos K grupos, sendo que um ponto pode pertencer a mais de um grupo, satisfazendo as condições das Equações (3.17), (3.18) e (3.19):

$$1- \mu_{ki} \in [0, 1]; \quad (3.17)$$

$$2- \sum_{i=1}^K \mu_{ki} = 1, i = 1, 2, \dots, n; \quad (3.18)$$

$$3- 0 < \sum_{i=1}^k \mu_{ki} < n, k = 1, 2, \dots, K. \quad (3.19)$$

Ou seja, a soma dos graus de pertinência de um mesmo ponto em todos os grupos deve ser igual a um e, ainda, a soma dos graus de pertinência de todos os pontos de um grupo deve ser menor que o número de elementos existentes no conjunto X .

O algoritmo FKM procura agrupar os pontos criando uma partição *fuzzy* em X de modo a minimizar a soma dos quadrados dos erros dos graus de pertinência dentro dos grupos, onde os pontos que pertencem a um mesmo grupo sejam similares e os pontos de diferentes grupos sejam dissimilares em termos de seus atributos. O critério de agrupamento usado pela maioria dos algoritmos de agrupamentos por partição *fuzzy* é a otimização da função da Equação (3.20).

$$J_q(U, M) = \sum_{k=1}^K \sum_{i=1}^n (\mu_{ki})^q |d(i, k)| \quad (3.20)$$

O parâmetro q dos graus de pertinência μ_{ki} é chamado fator de “fuzzificação” e tem como objetivo controlar a associação contínua entre os grupos da partição *fuzzy* resultante. O valor de q varia no intervalo de um a ∞ . Quando q tende a um, a solução do algoritmo FKM converge para uma partição convencional, com os graus de pertinência aos grupos sendo um ou 0, na qual cada ponto pertence totalmente a um único grupo ou não. Quanto maior o valor de q , maior a associação entre os grupos resultantes, com diferentes graus de pertinência do ponto aos vários grupos. NIKHIL e BEZDEK (1995) sugerem que o parâmetro q esteja no intervalo de 1,5 a 2,5 para obtenção de um resultado de agrupamento válido. Não existe base teórica para uma escolha ótima do valor de q mas freqüentemente utiliza-se q igual a dois.

A seguir são descritos os passos do algoritmo, segundo DUDA e CANTY (2002).

1- O processamento é iniciado com a definição do número de grupos k , o fator de “fuzzificação” $q > 1$ e a condição de término $\varepsilon > 0$, atingida quando os centros dos grupos ou os elementos da matriz de partição *fuzzy* U não mais tiverem alterações significativas.

2- Os graus de pertinência μ_{ki} da matriz de partição *fuzzy* U são inicializados a partir de um número de amostras dos pixels selecionadas aleatoriamente na imagem. A partir deles são calculados os centros iniciais m_k dos grupos por meio da Equação (3.21):

$$m_k = \frac{\sum_{i=1}^n (\mu_{ki})^q x_i}{\sum_{i=1}^n (\mu_{ki})^q}, k = 1, 2, \dots, K. \quad (3.21)$$

3- Os graus de pertinência de cada x_i com relação a cada um dos k grupos são calculados segundo uma medida de similaridade $d(i,k)$ do ponto ao centro de cada grupo para atualização da matriz de pertinência *fuzzy* U , conforme a Equação (3.22):

$$\mu_{ki} = \frac{1}{\sum_{j=1}^K \left(\frac{|d(i,k)|}{|d(i,j)|} \right)^{\frac{2}{q-1}}}, k = 1, 2, \dots, K; i = 1, 2, \dots, n. \quad (3.22)$$

A distância de cada ponto x_i ao centro do grupo k da partição *fuzzy* atual é comparada à distância do mesmo ponto x_i ao centro do grupo j da partição *fuzzy* j , de modo a atribuir um grau de pertinência ao ponto que seja proporcional ao seu grau de pertinência aos demais grupos.

4- O algoritmo é executado iterativamente aplicando-se as Equações (3.21) e (3.22) e converge quando satisfizer a condição de término $\max_{i,j} |\mu_{ki}^{(it)} - \mu_{ki}^{(it-1)}| < \varepsilon$, onde it é o número da iteração atual. Usualmente ε é igual a 0,001.

A condição de término verifica se a máxima diferença entre o grau de pertinência de qualquer elemento x_i em um dado grupo k na iteração atual e na iteração anterior é menor que ε definido, sendo que isso é checado para todos os grupos.

O resultado final da matriz de partição *fuzzy* U representa os graus de pertinência de todos os pontos em todos os grupos e depende da definição de uma medida de similaridade que normalmente é uma métrica de distância.

O algoritmo FKM adota a distância Euclidiana como medida de similaridade entre o ponto e o centro do grupo, calculada por meio da Equação (3.23), que obtém bons resultados para grupos de formas circulares:

$$d(i,k) = (x_i - m_k)^T(x_i - m_k) \quad (3.23)$$

Porém, devido a falta de conhecimento sobre todas as classes onde pode haver a existência de grupos elipsoidais com grande variabilidade de formas e com densidade de pontos desproporcional, torna-se difícil a escolha do número de grupos. Para estes casos é utilizada como medida de similaridade a distância adaptativa exponencial EM que cria uma partição *fuzzy* otimizada para um certo número de grupos.

O algoritmo EM interpreta o grau de pertinência μ_{ki} como a probabilidade a posteriori $P_r(C_k | x_i)$ de selecionar um grupo C_k dada o ponto x_i , usando o teorema de Bayes:

$$\mu_{ki} \rightarrow P_r(C_k | x_i) \sim P_r(C_k) P_r(x_i | C_k)$$

A probabilidade a priori $P_r(C_k)$ é dada pela Equação (3.24):

$$P_r(C_k) = \frac{1}{n} \sum_{i=1}^n (\mu_{ki})^q, \quad k = 1, 2, \dots, K. \quad (3.24)$$

$P_r(x_i | C_k)$ é assumida como uma função de distribuição multivariada normal com média *fuzzy* m_k e matriz de covariância *fuzzy* F_k que podem ser estimadas pelas Equações (3.25) e (3.26):

$$m_k = \frac{1}{n P_r(C_k)} \sum_{i=1}^n (\mu_{ki})^q x_i, \quad k = 1, 2, \dots, K. \quad (3.25)$$

$$F_k = \frac{1}{n P_r(C_k)} \sum_{i=1}^n P_r(C_k) (x_i - m_k)(x_i - m_k)^T, k = 1, 2, \dots, K. \quad (3.26)$$

Este algoritmo é similar ao FKM exceto que a medida de similaridade adotada $d(i,k)$ do ponto ao centro do grupo é a distância EM, calculada por meio da Equação (3.27):

$$d(i,k) = P_r(C_k) \frac{1}{\sqrt{|F_k|}} \exp \left[-\frac{1}{2} (x_i - m_k)^T F_k^{-1} (x_i - m_k) \right] \quad (3.27)$$

A complementação dos passos do algoritmo EM consiste em determinar as probabilidades a priori $P_r(C_k)$ de cada grupo, os centros dos grupos m_k e, assim, a matriz de covariância *fuzzy* F_k . Com isso os graus de pertinência μ_{ki} são calculados segundo distância EM e a matriz de partição *fuzzy* U é atualizada até que não haja mais alterações significativas.

Devido à dependência exponencial da medida de distância EM na função de pertinência *fuzzy* que é medida de similaridade $d(i,k)$, o algoritmo é muito sensível às condições de inicialização e pode se tornar instável. Para evitar este problema, os valores iniciais de m_k e μ_{ki} devem ser obtidos com o algoritmo FKM.

Durante o agrupamento dos dados, os critérios de validação para definir uma partição *fuzzy* ideal baseiam-se nos três requisitos citados abaixo, conforme GATH e GEVA (1989):

- 1- Clara separação entre grupos resultantes;
- 2- Mínimo volume dos grupos;
- 3- Maior concentração de pontos coesos em torno do centro do grupo.

Para determinar os critérios de validação foram utilizadas as medidas de desempenho hypervolume e a densidade dos grupos para comparação entre boas partições com diferentes números de grupos. A cada iteração as medidas de desempenho são calculadas até a obtenção de um valor ótimo.

A função hypervolume é a quantidade proporcional ao volume no espaço ocupado pelos grupos e deve ser mínima para grupos de boa qualidade, definida pela Equação (3.28):

$$FHV = \sum_{i=1}^K \sqrt{|F_k|} \quad (3.28)$$

A densidade da partição é a relação entre a soma dos centros dos grupos (S) com a FHV e deve ser máxima para grupos de boa qualidade, calculada pela Equação (3.29):

$$P_D = \frac{S}{FHV}, S = \sum_{i=1}^n \sum_{k=1}^K \mu_{ik}, \forall i \in \{i | (x_i - m_k)^T F_k^{-1} (x_i - m_k) < 1\}. \quad (3.29)$$

Ao final, o algoritmo EM gera um conjunto de imagens fração, uma para cada classe, com os graus de pertinência dos pixels medidos entre zero e um, escalonados no intervalo de 0 a 255. Dessa forma, os pixels com graus de pertinência próximos a um aparecem em branco e os próximos a zero em preto, variando em tons de cinza (DUDA e CANTY, 2002).

Para proceder a classificação não-supervisionada *fuzzy* das composições máximas AVHRR-NOAA de NDVI e NDMI foi adaptado o programa em IDL/ENVI desenvolvido pelo Dr. Morton J. Canty, do Jülich Research Center, Jülich, Alemanha, tecnologia essa também repassada gratuitamente (CANTY, 2005).

As imagens fração geradas pelo algoritmo de classificação *fuzzy* permitem estimar as proporções de cobertura do solo misturadas dentro de cada pixel. As proporções de área de cada classe de uso do solo são contínuas variando num intervalo de 0% a 100%. Baseado nisso, identificando-se a imagem fração correspondente a classe Soja, pode-se estimar a proporção de área a partir dos graus de pertinência dos pixels, conforme a Equação (3.30):

$$A_{Soja} = A_{pixel} * \left(\frac{\sum_{i=1}^n \mu_{ci}}{255} \right) * 100 \quad (3.30)$$

onde,

A_{Soja} = área de soja em hectares (ha);

A_{pixel} = área do pixel AVHRR-NOAA igual a 1,21 km² (1,1 x 1,1 km);

μ_{ci} = graus de pertinência dos pixels da imagem fração referente a classe c, no caso, a classe Soja.

4. RESULTADOS E DISCUSSÕES

4.1. Processamento das imagens AVHRR-NOAA

A cultura da soja foi monitorada ao longo do seu ciclo de desenvolvimento na região oeste do Paraná que se estendeu de outubro de 2003 a março de 2004, com as imagens disponíveis do satélite NOAA-17 de passagem no período da manhã, próxima às 14h GMT.

Das 75 imagens AVHRR-NOAA processadas, 55 deram origem a produtos com precisão geométrica dentro de um pixel, conforme detalhado na Tabela 4.1. As imagens em que o processamento falhou apresentaram nível excessivo de ruídos que podem ser atribuídos a problemas de recepção, afetando o processo de calibração radiométrica e, conseqüentemente, o georreferenciamento. Outra fonte de erros foi a presença maciça de nuvens na região em estudo, prejudicando o processo de busca por feições semelhantes por meio da MCC.

Tabela 4.1: Imagens AVHRR-NOAA processadas na safra 2003/2004.

Mês	Imagens AVHRR-NOAA	
	Bruta	Georreferenciada
Nov/2003	16	9
Dez/2003	18	12
Jan/2004	19	14
Fev/2004	10	9
Mar/2004	12	11
Total	75	55

O procedimento implementado para o georreferenciamento utilizou quatro imagens base durante o período - 09/11/2003, 10/12/2003, 28/02/2004 e 02/03/2004, selecionadas devido à pouca cobertura de nuvens e à proximidade no nadir com a região que engloba o Paraná, centrada nas coordenadas geográficas 24°30' S e 51°20' W. Decidiu-se utilizar uma imagem base por mês para evitar o efeito sazonal devido às mudanças térmicas, como forma de garantir uma precisão maior no processo. Não foi possível criar uma imagem base para janeiro de 2004 por causa da grande nebulosidade presente no transcorrer do mês.

Os resultados mostraram que, quando livre de ruídos e da grande presença de nuvens, as imagens puderam ser processadas automaticamente, gerando produtos calibrados e geometricamente precisos, com erro de no máximo um pixel, sem qualquer intervenção

humana. A análise dos deslocamentos entre as imagens alvo e as imagens base da série multitemporal mostrou um erro médio de 0,76 pixel (menos de 1 km), considerado aceitável para estudos terrestres. A navegação indireta, aliada à busca automática de feições por meio da MCC, mostrou-se eficaz e contribuiu para a total automação do processamento das imagens.

A Figura 4.1 ilustra os efeitos da correção dos erros de atitude do satélite em três regiões com feições geográficas conhecidas, antes e depois do processamento.

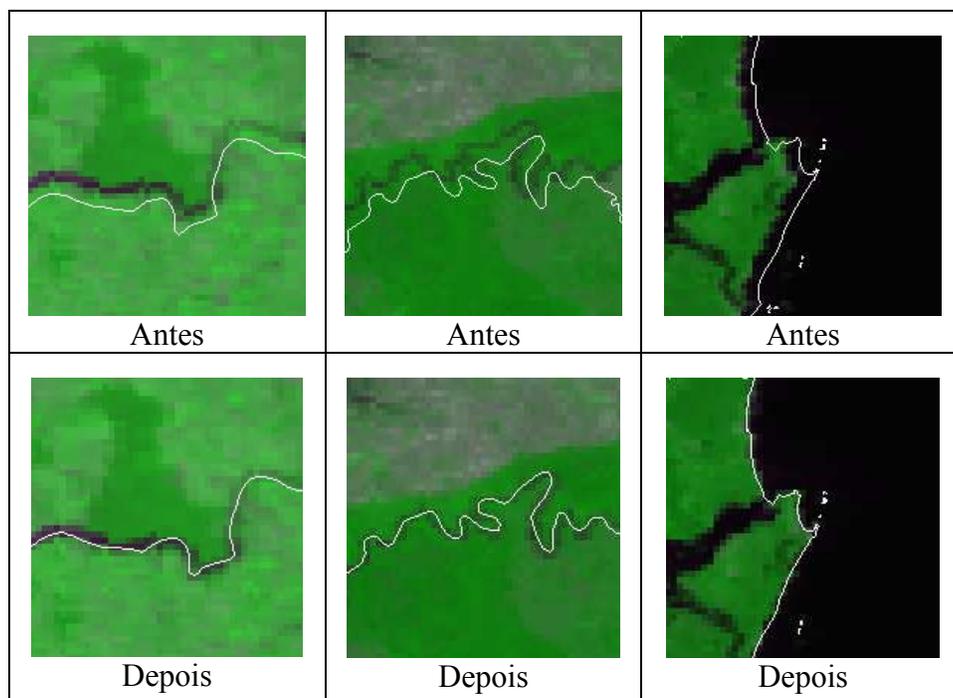


Figura 4.1: Efeito da correção dos erros de atitude do satélite no georreferenciamento.

A seguir é apresentado o conjunto de imagens JPEG do dia 09 de novembro de 2004, início do ciclo de desenvolvimento da soja na região, geradas pelas rotinas do sistema desenvolvido para processamento das imagens e geração de produtos AVHRR-NOAA. As informações referentes ao número do satélite, data e hora de aquisição da imagem, bem como à malha do IBGE e às informações cartográficas, são inseridas automaticamente pelo sistema. Também são representados os vetores dos oito municípios da região oeste do Paraná.

A Figura 4.2 mostra a composição colorida RGB-121, cuja coloração facilita a observação visual da presença de nuvens. A Figura 4.3 exibe o mascaramento de nuvens da mesma imagem onde as manchas escuras correspondem à localização das nuvens, têm valor zero e as áreas brancas são livres de nuvens, têm valor um.

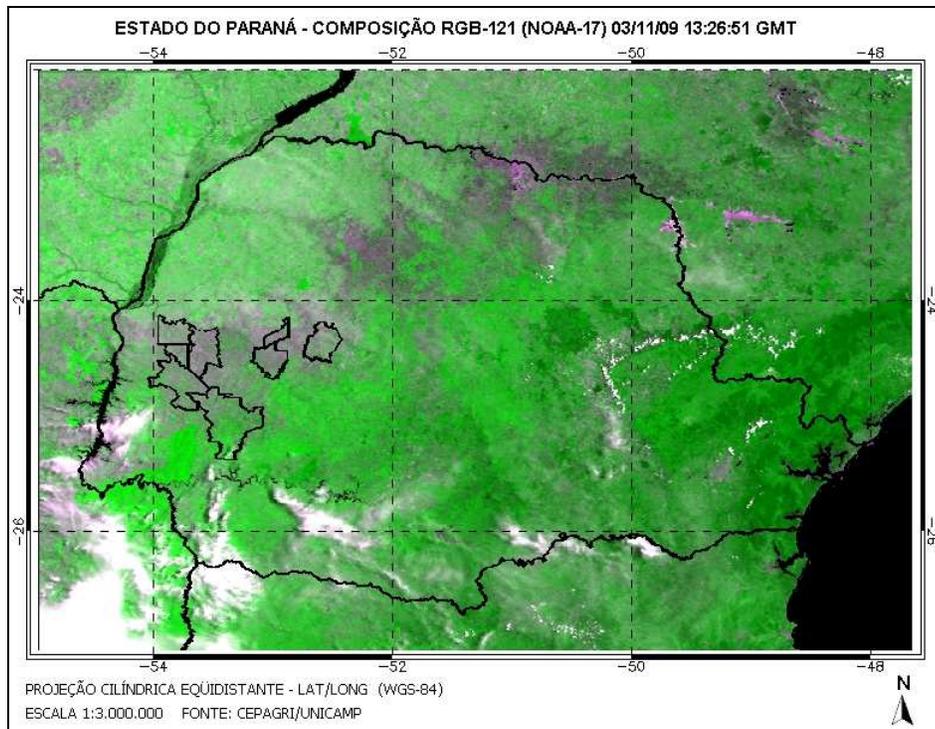


Figura 4.2: Composição RGB-121.

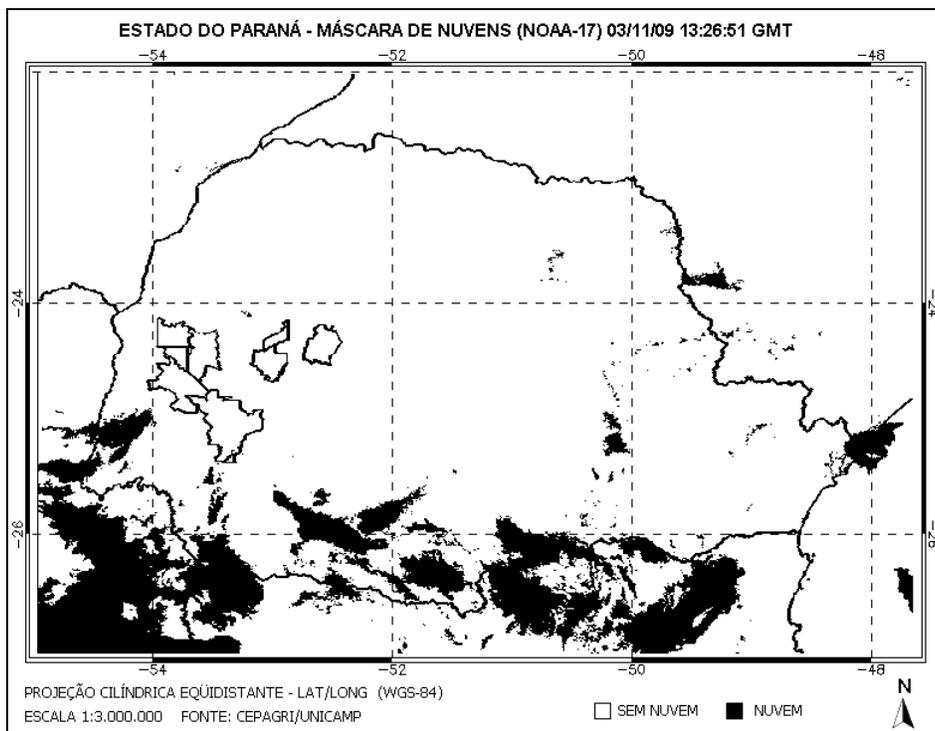


Figura 4.3: Máscara de nuvens.

Uma análise visual entre as composições RGB-121 e as máscaras geradas mostrou que os critérios adotados para detecção de nuvens se mostraram eficazes. Deve-se observar que esses critérios foram selecionados após um processo iterativo para teste dos valores limiares, de forma a otimizar a detecção das nuvens na região. Vários estudos têm mostrado que o mascaramento de nuvens é necessário para evitar a contaminação da análise espectral, principalmente na extração dos valores dos índices de vegetação nos casos em que o uso de composições máximas não foi suficiente para eliminar a presença das nuvens. O sistema pode ainda ser alterado de forma a acrescentar outros critérios para o mascaramento de nuvens em trabalhos que envolvam outras regiões.

A Figura 4.4 refere-se ao NDVI e a Figura 4.5 ao NDMI, ambos calculados para a mesma imagem. Para efeitos de escala foi utilizada a paleta de cores *RAINBOW* do ENVI, variando do preto, valores mais baixos, ao vermelho, valores mais altos.

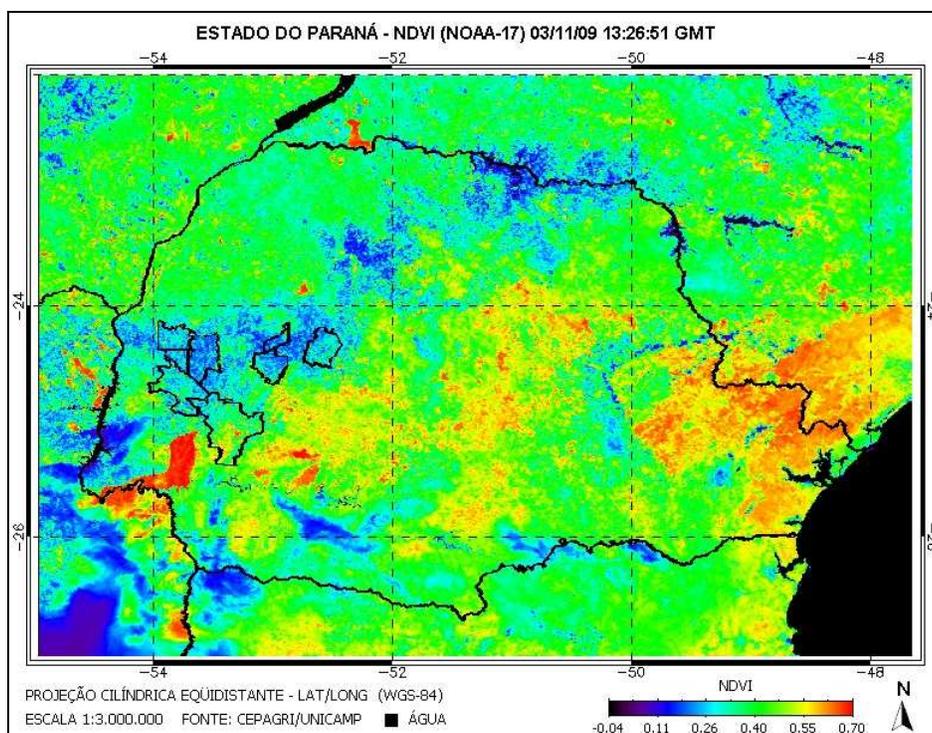


Figura 4.4: NDVI.

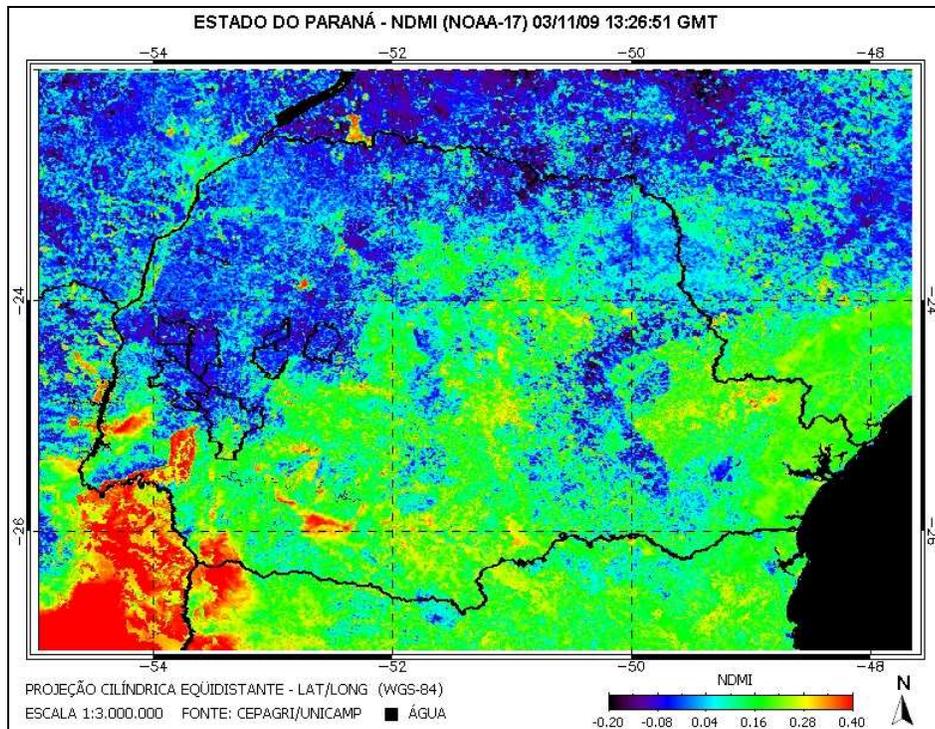


Figura 4.5: NDMI.

Verifica-se que a presença de nuvens reduz o valor do NDVI e pode causar confusão na interpretação pela semelhança com solo exposto e áreas urbanas. O contrário acontece com o NDMI, em que a existência de nuvens provoca aumento nos valores e pode ocasionar confusão pela semelhança com florestas e culturas agrícolas com grande vigor vegetativo.

Esse tipo de confusão é reduzido quando se mascaram os pixels cobertos por nuvens. Na Figura 4.4 pode-se observar que alguns pixels na coloração azulada na região de latitude $25^{\circ}05' S$ e longitude $54^{\circ}15' W$ não se referem à presença de nuvens, mas sim ao solo exposto das áreas de soja preparada para novos plantios. Na Figura 4.5 pode-se notar que vários pixels na coloração avermelhada na região de $25^{\circ}55' S$ e longitude $52^{\circ}40' W$ não representam vegetação porque estão totalmente cobertos por nuvens.

A Figura 4.6 mostra o NDVI e a Figura 4.7 o NDMI, ambos com o mascaramento de pixels inaptos, onde os pixels da imagem com ângulo zenital solar maior que o limite de 70° e com ângulo de elevação do satélite maior que o limite de 42° e, também, cobertos por nuvens, foram descartados para uma análise quantitativa. Para esses pixels inaptos foi atribuído o valor -1,1, propositalmente fora da escala de variação dos índices de vegetação para não contaminar a análise.

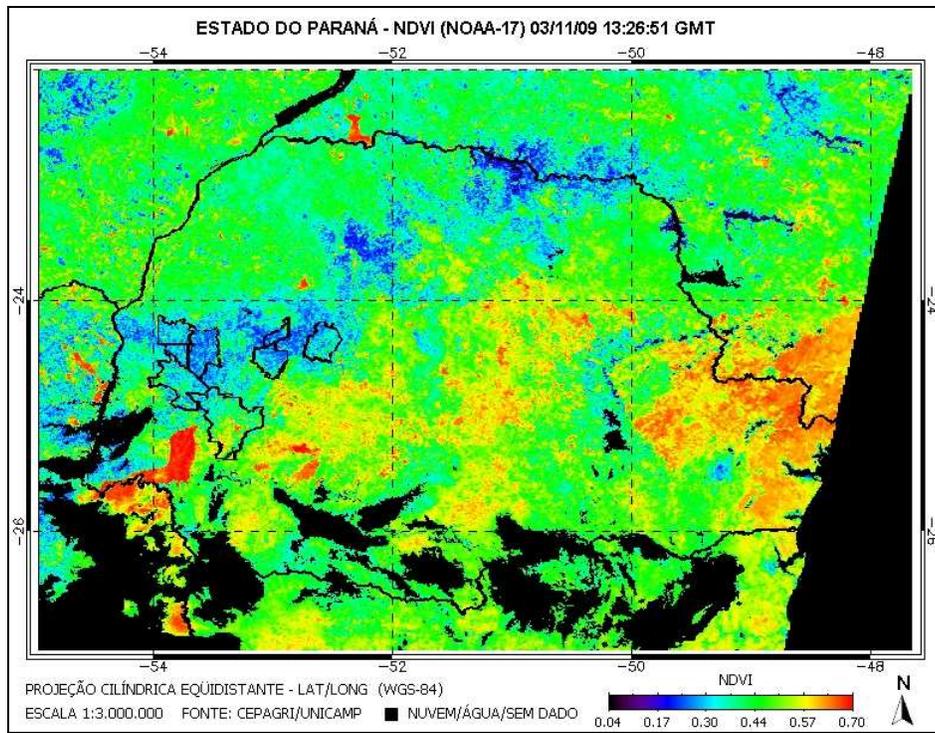


Figura 4.6: NDVI com pixels inaptos.

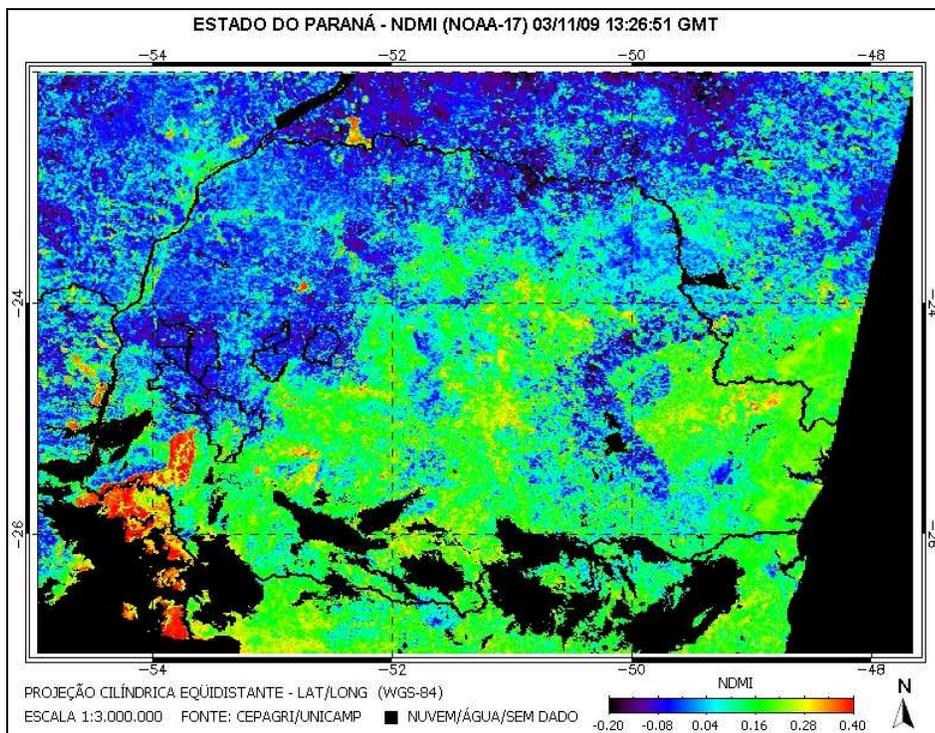


Figura 4.7: NDMI com pixels inaptos.

Para uma análise temporal dos índices de vegetação é aconselhado o uso de composições máximas durante um período determinado como forma de atenuar os efeitos atmosféricos. As Figuras 4.8 e 4.9 exibem o resultado da rotina desenvolvida para geração automática de composições máximas mensais do NDVI e NDMI, respectivamente.

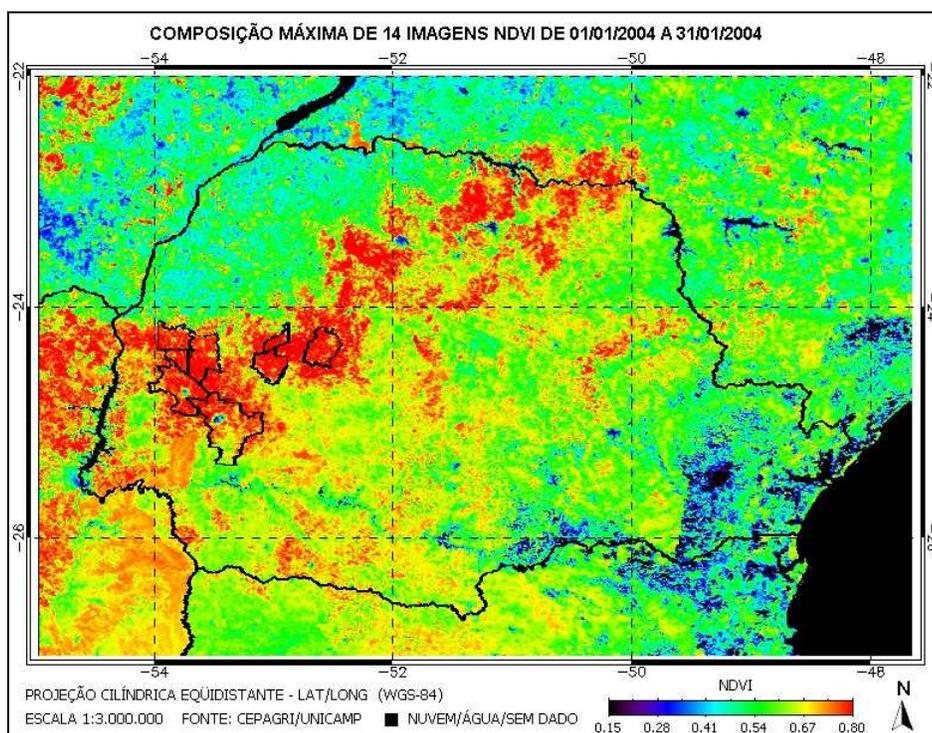


Figura 4.8: Composição Máxima de NDVI.

O mês a que se refere a imagem é janeiro de 2004 com um total de 14 imagens processadas, onde as áreas de tonalidade mais intensa representam valores mais altos do NDVI e NDMI e as áreas de tonalidade menos intensa, valores mais baixos. Tal comportamento é explicado pela presença da cultura da soja em sua época de maior vigor vegetativo na região. Todas as imagens dos índices de vegetação utilizadas nas composições foram processadas com o mascaramento de pixels inaptos.

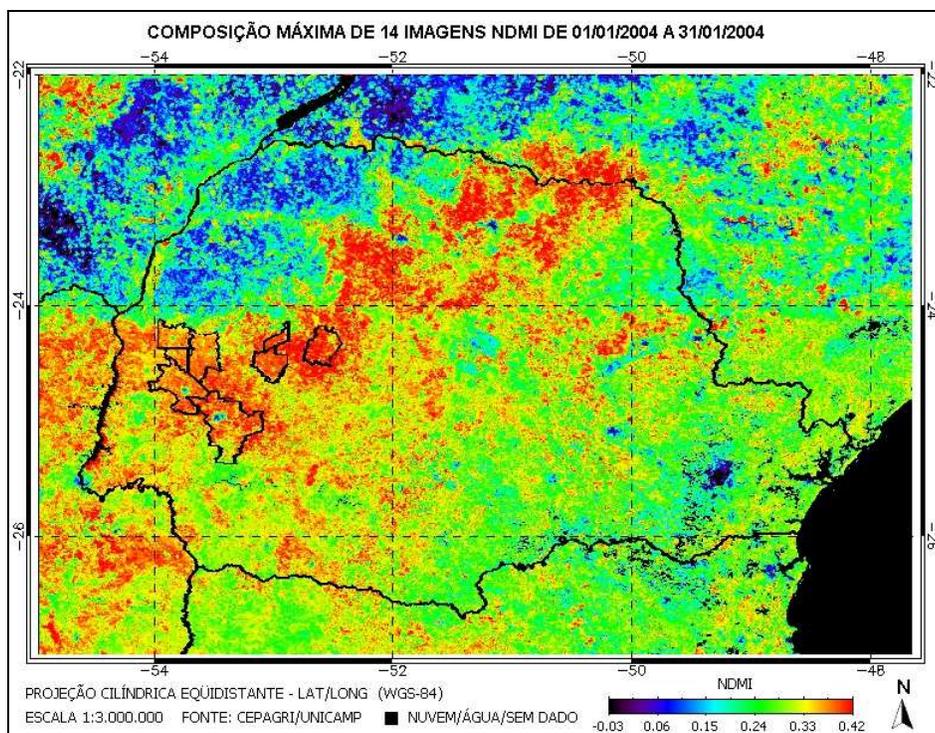


Figura 4.9: Composição Máxima de NDMI.

4.2. Perfis espectrais temporais dos índices de vegetação

Os valores dos índices de vegetação de imagens AVHRR-NOAA para a construção dos perfis espectrais foram extraídos de composições quinzenais geradas de outubro de 2003 a março de 2004. A extração foi feita em nível municipal sobre as áreas identificadas tomando-se como referência a máscara de soja. Para delimitação das áreas representativas foi utilizada uma janela amostral de 3 x 3 pixels devido ao erro máximo de georreferenciamento das imagens AVHRR-NOAA que pode ser de um pixel, sendo o pixel central da janela localizado sobre áreas com predominância de soja.

O perfil para cada um dos oitos municípios foi gerado a partir dos valores médios coletados em cinco áreas e posteriormente agrupados numa planilha. A partir disso foram gerados os gráficos que descrevem os perfis espectrais do NDVI e do NDMI ao longo do ciclo de desenvolvimento da soja, ou seja, a variação temporal dos índices de vegetação, conforme apresentado nas Figuras 4.10 a 4.17.

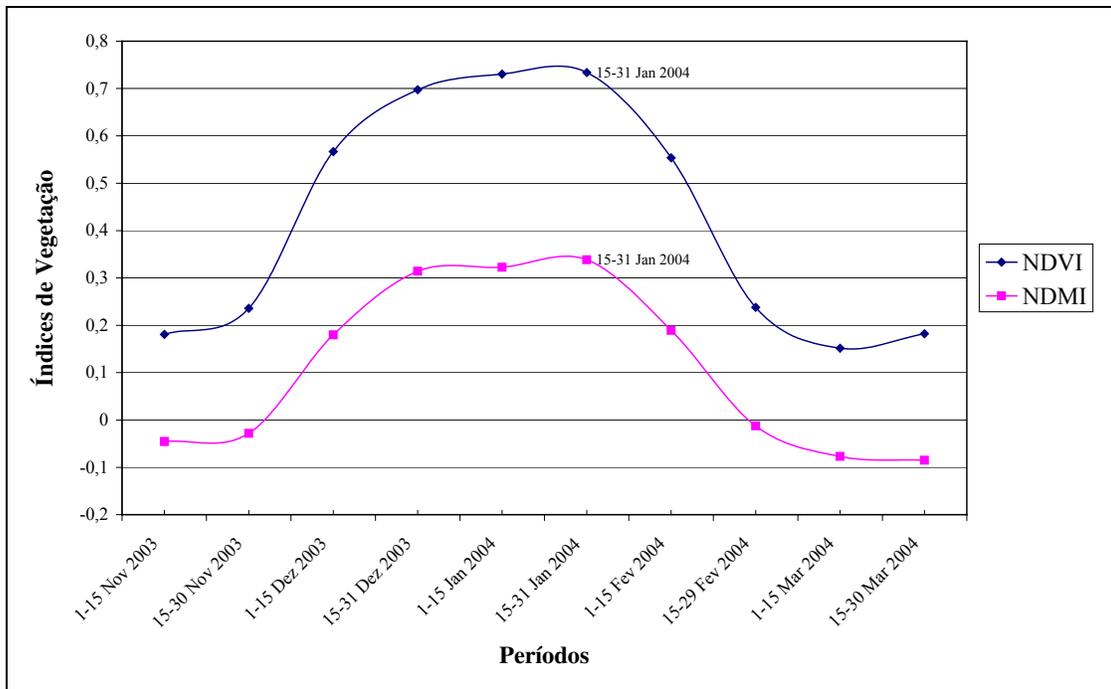


Figura 4.10: Perfil espectral do NDVI e do NDMI - Assis Chateaubriand.

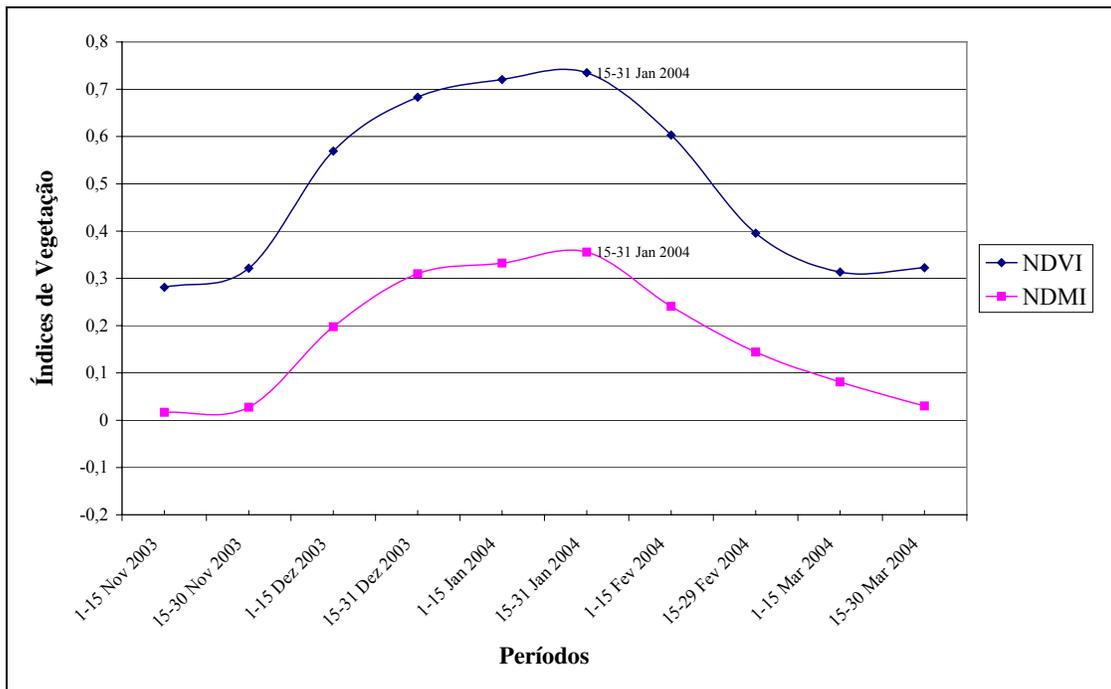


Figura 4.11: Perfil espectral do NDVI e do NDMI - Cascavel

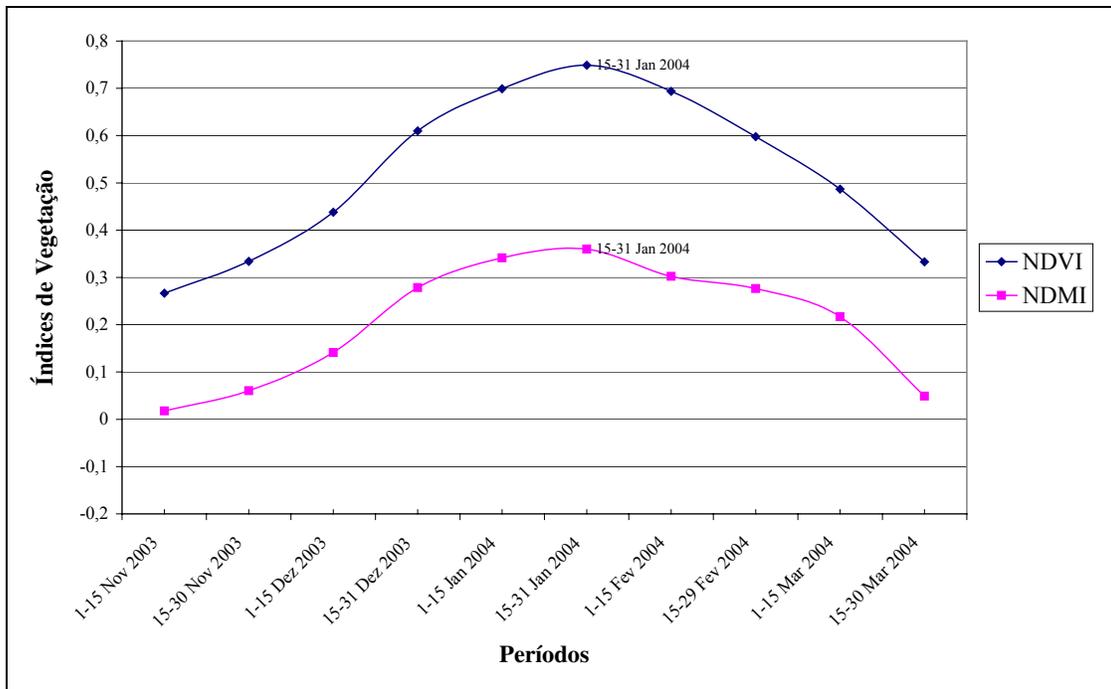


Figura 4.12: Perfil espectral do NDVI e do NDMI - Mambrorê

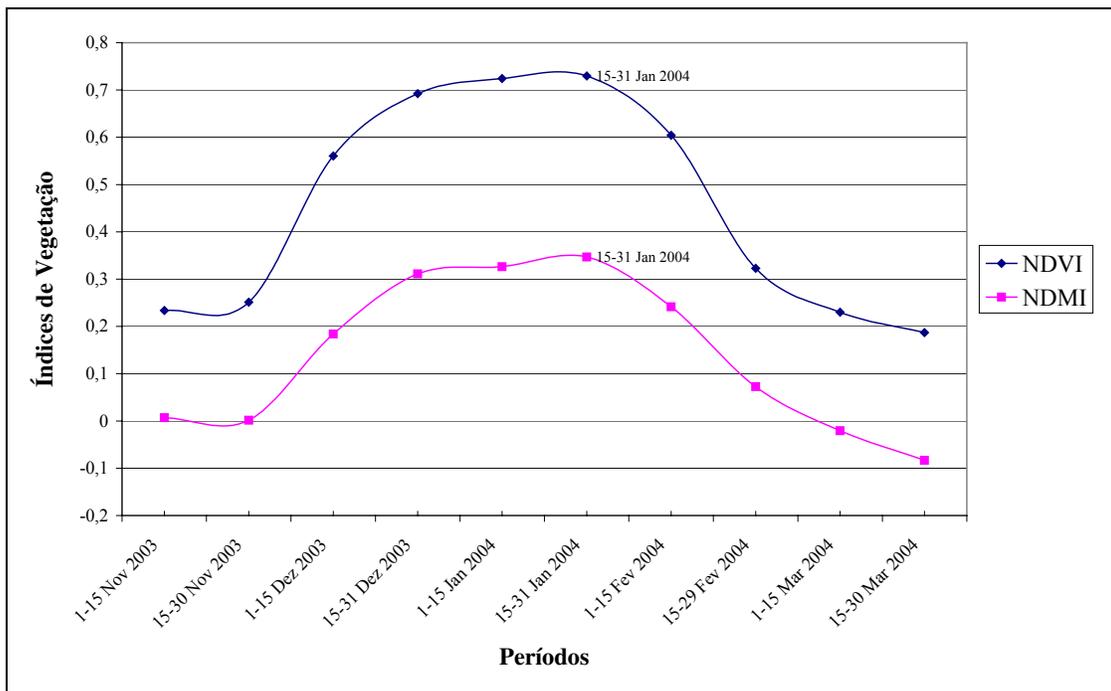


Figura 4.13: Perfil espectral do NDVI e do NDMI - Maripá

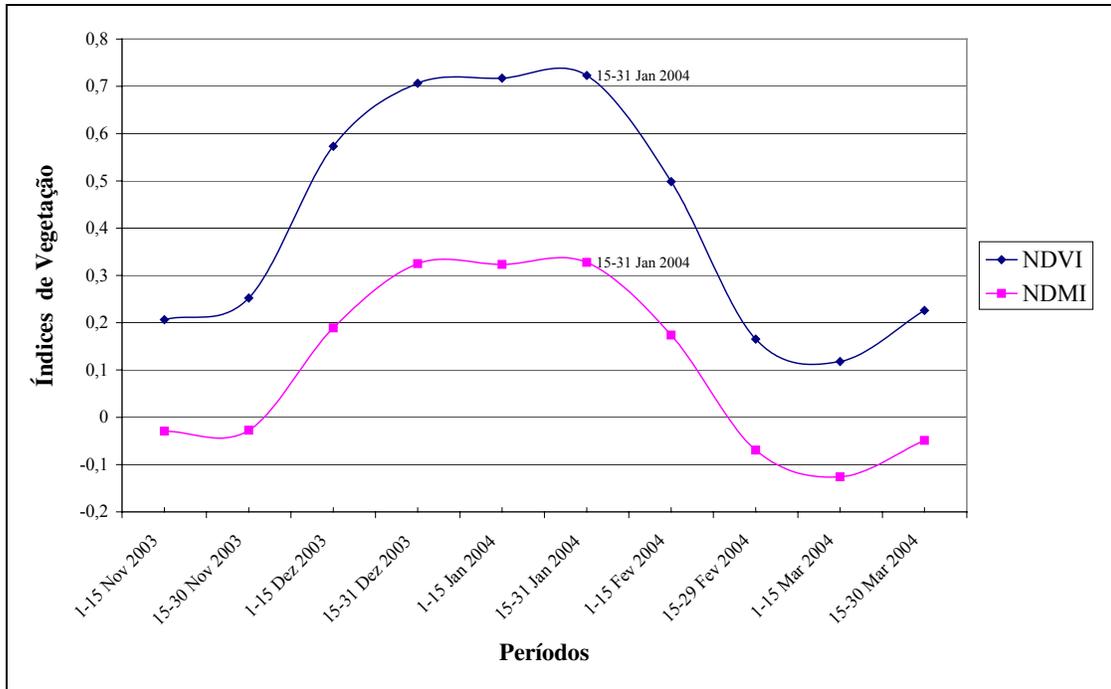


Figura 4.14: Perfil espectral do NDVI e do NDMI - Palotina

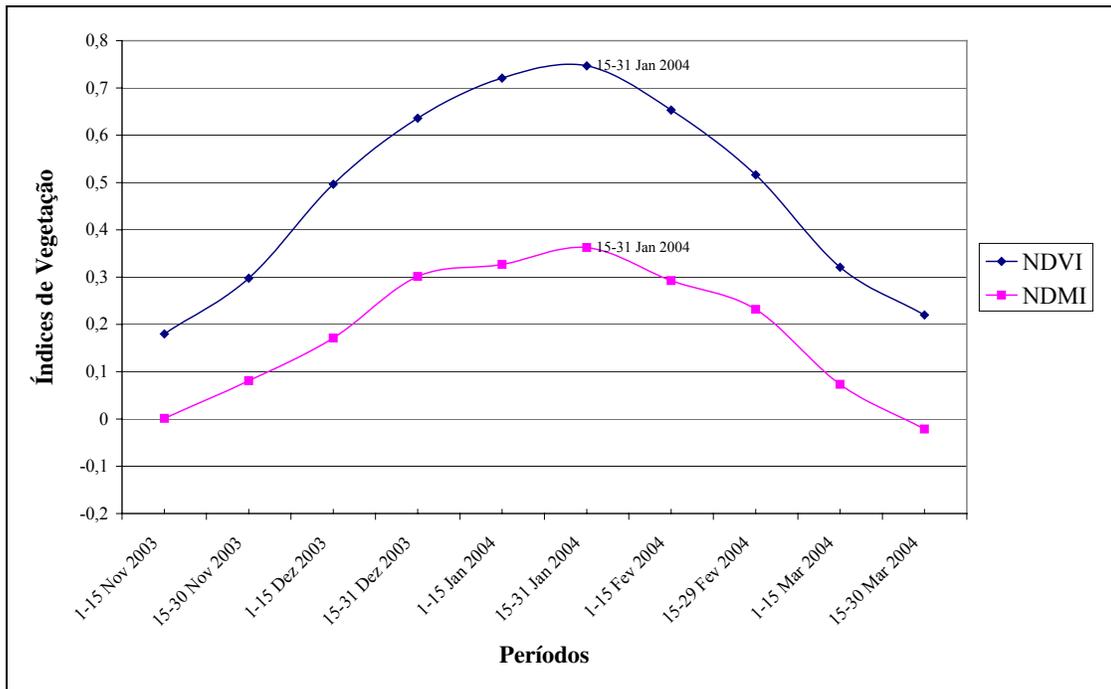


Figura 4.15: Perfil espectral do NDVI e do NDMI - Rancho Alegre do Oeste.

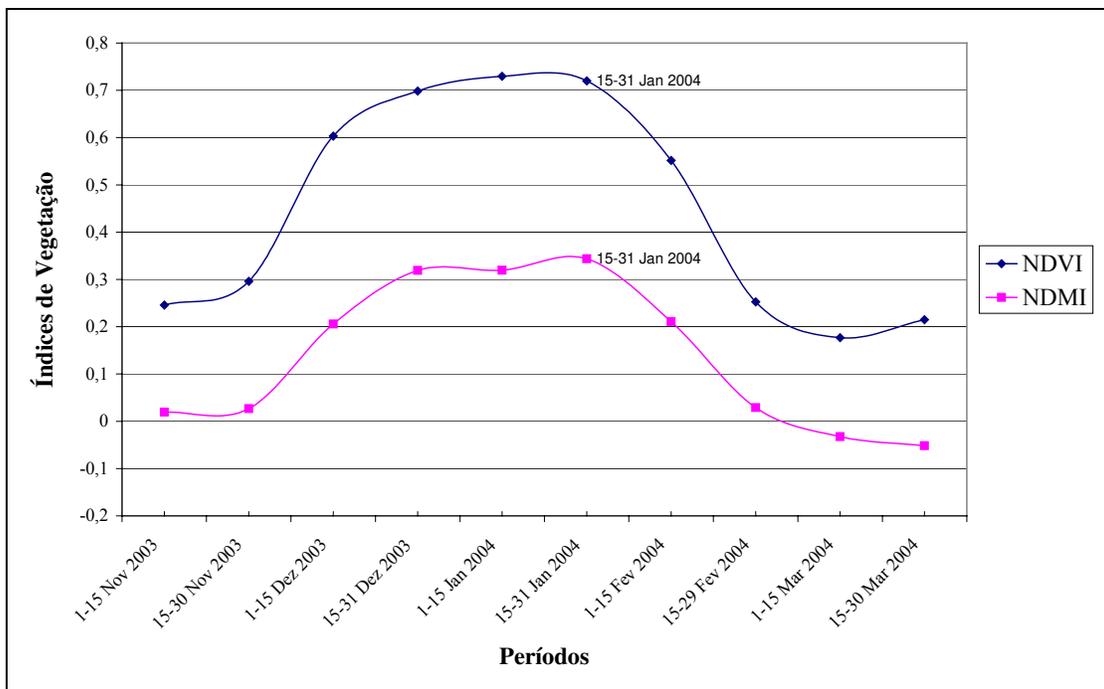


Figura 4.16: Perfil espectral do NDVI e do NDMI - Toledo.

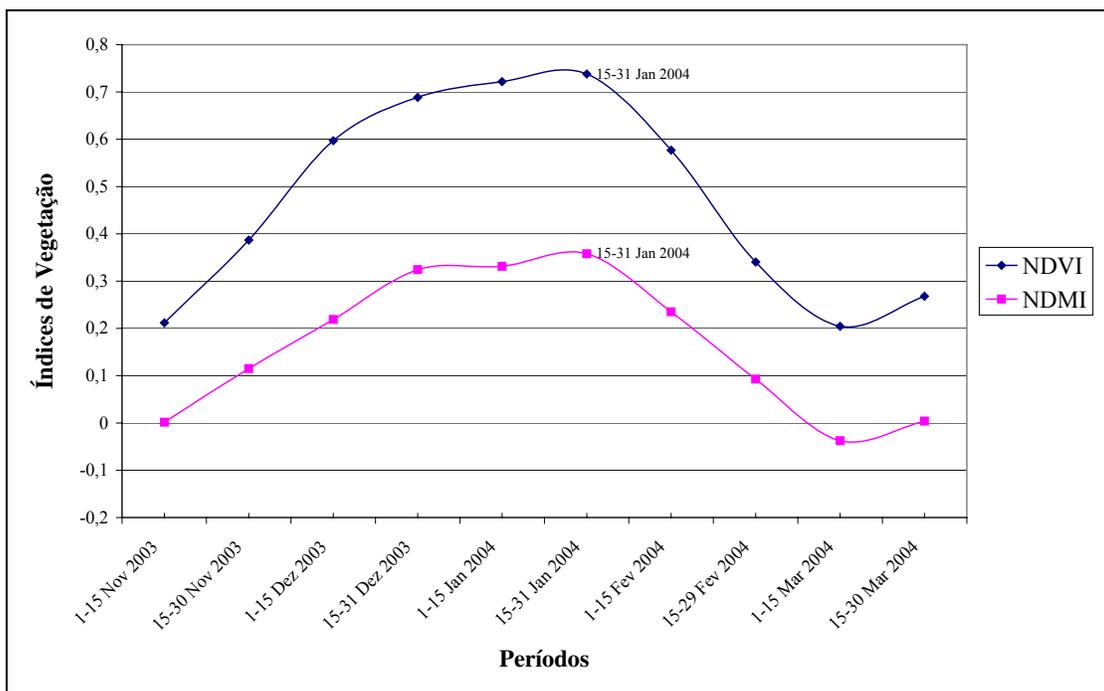


Figura 4.17: Perfil espectral do NDVI e do NDMI - Ubitatã.

A partir da análise dos gráficos é possível observar que o período de maior vigor vegetativo da soja nos oito municípios ocorreu de 15 a 31 de janeiro de 2004. Logo, essa é a

época em que a planta teve o maior índice de área foliar e assim estava cobrindo a maior área de superfície do solo. Portanto, foi nesse momento em que a classificação *fuzzy* deveria ser executada para estimar a área plantada da cultura.

4.3. Estimativa de área de soja de referência por meio de matriz de erros

Diante dos 100 pontos do painel amostral da CONAB tomados como referência terrestre em cada um dos oito municípios, foi possível elaborar as matrizes de erros e realizar a avaliação da qualidade (acurácia) da classificação digital que deu origem à máscara de soja.

A Tabela 4.2 apresenta os valores das métricas de Exatidão Global e do Índice Kappa e, também, os valores de área calculados por contagem de pixels na máscara de soja para cada um dos oito municípios. A área foi obtida multiplicando-se o número total de pixels da classe Soja pela área do pixel Landsat 5/TM igual a 900 m² (30 x 30 m), dividindo-se por 10.000 para conversão em hectares (ha).

Tabela 4.2: Métricas de acurácia da classificação e área de soja em cada município.

Município	Exatidão Global	Índice Kappa	Área de Soja (ha)
Assis Chateaubriand	0,710	0,708	48.724,11 (541.379 pixels)
Cascavel	0,810	0,808	82.434,42 (915.938 pixels)
Mamborê	0,800	0,799	46.961,46 (521.724 pixels)
Maripá	0,660	0,650	16.324,11 (181.379 pixels)
Palotina	0,646	0,635	34.002,09 (377.801 pixels)
Rancho Alegre do Oeste	0,768	0,767	14.565,33 (161.837 pixels)
Toledo	0,750	0,746	66.565,89 (739.621 pixels)
Ubiratã	0,790	0,789	36.069,84 (400.776 pixels)

A Exatidão Global corresponde à concordância entre os pontos classificados como Soja ou Não Soja na máscara e os da referência terrestre, presentes na diagonal principal da matriz de erros, por cada município. Os valores ficaram em torno de 0,742.

O Índice Kappa envolve a matriz de erros completa, incorporando os elementos das linhas e colunas marginais referentes aos erros de classificação na máscara de soja, além dos que se situam na diagonal principal. Os valores para todos os municípios estão acima de 0,635.

A Tabela 4.3 descreve a avaliação de qualidade da classificação proposta por LANDIS e KOCH (1977).

Tabela 4.3: Qualidade da classificação associada aos valores do Índice Kappa.

Índice Kappa	Qualidade
0	Péssima
0,01 a 0,20	Ruim
0,21 a 0,40	Razoável
0,41 a 0,60	Boa
0,61 a 0,80	Muito Boa
0,81 a 1,00	Excelente

Baseado nisso, pode-se dizer que a classificação que deu origem a máscara de soja foi caracterizada de qualidade muito boa. Ainda, FOODY (2002) revisando os métodos de análise de acurácia sugere que os níveis recomendados para as métricas de Exatidão Global e do Índice Kappa devem estar em torno de 0,85, para se obter uma boa classificação que forneça dados precisos.

De qualquer maneira, pode-se concluir que os classificadores booleanos apresentaram limitações ao efetuar o reconhecimento da cultura, apesar do padrão da soja ter sido bem delimitado nas amostras de treinamento.

Devido ao limitante de acurácia inerente desses classificadores, a área dos oito municípios da máscara de soja foi corrigida por meio de matrizes de erros que foram transformadas em matrizes de proporções, conforme o método proposto por JUPP (1989). Dessa forma, todas as probabilidades de designação das classes foram corrigidas simultaneamente e, conseqüentemente, as estimativas de área.

A Tabela 4.4 representa a matriz de erros e as proporções das classes Soja e Não Soja entre parênteses para o município de Assis Chateaubriand.

Tabela 4.4: Matriz de erros e proporções - Assis Chateaubriand.

Máscara	Pontos de Referência		Total Linha
	Soja	Não Soja	
Soja	47 (0,940)	3 (0,060)	50
Não Soja	26 (0,520)	24 (0,480)	50
Total			100

Tem-se que:

$$T_{\text{Soja}} = 48.724,11 \text{ ha.}$$

$$T_{\text{Não Soja}} = 48.304,44 \text{ ha.}$$

Então:

$$A_{\text{Soja}} = 0,940 * 48.724,11 + 0,520 * 48.304,44 = 70.918,97 \text{ ha.}$$

$$A_{\text{Não Soja}} = 0,060 * 48.724,11 + 0,480 * 48.304,44 = 26.109,58 \text{ ha.}$$

A Tabela 4.5 representa a matriz de erros e as proporções das classes Soja e Não Soja entre parênteses para o município de Cascavel.

Tabela 4.5: Matriz de erros e proporções - Cascavel.

Pontos de Referência			
Máscara	Soja	Não Soja	Total Linha
Soja	27 (0,730)	10 (0,270)	37
Não Soja	9 (0,143)	54 (0,857)	63
Total			100

Tem-se que:

$$T_{\text{Soja}} = 82.434,42 \text{ ha.}$$

$$T_{\text{Não Soja}} = 127.634,13 \text{ ha.}$$

Então:

$$A_{\text{Soja}} = 0,730 * 82.434,42 + 0,143 * 127.634,13 = 78.388,29 \text{ ha.}$$

$$A_{\text{Não Soja}} = 0,270 * 82.434,42 + 0,857 * 127.634,13 = 131.680,26 \text{ ha.}$$

A Tabela 4.6 representa a matriz de erros e as proporções das classes Soja e Não Soja entre parênteses para o município de Mamborê.

Tabela 4.6: Matriz de erros e proporções - Mamborê.

Pontos de Referência			
Máscara	Soja	Não Soja	Total Linha
Soja	57 (0,950)	3 (0,050)	60
Não Soja	17 (0,425)	23 (0,575)	40
Total			100

Tem-se que:

$$T_{\text{Soja}} = 46.961,46 \text{ ha.}$$

$$T_{\text{Não Soja}} = 30.896,37 \text{ ha.}$$

Então:

$$A_{\text{Soja}} = 0,950 * 46.961,46 + 0,425 * 30.896,37 = 57.744,34 \text{ ha.}$$

$$A_{\text{Não Soja}} = 0,050 * 46.961,46 + 0,575 * 30.896,37 = 20.113,49 \text{ ha.}$$

A Tabela 4.7 representa a matriz de erros e as proporções das classes Soja e Não Soja entre parênteses para o município de Maripá.

Tabela 4.7: Matriz de erros e proporções - Maripá.

Pontos de Referência			
Máscara	Soja	Não Soja	Total Linha
Soja	28 (0,583)	20 (0,417)	48
Não Soja	14 (0,269)	38 (0,731)	52
Total			100

Tem-se que:

$$T_{\text{Soja}} = 16.324,11 \text{ ha.}$$

$$T_{\text{Não Soja}} = 12.162,96 \text{ ha.}$$

Então:

$$A_{\text{Soja}} = 0,583 * 16.324,11 + 0,269 * 12.162,96 = 12.797,04 \text{ ha.}$$

$$A_{\text{Não Soja}} = 0,417 * 16.324,11 + 0,731 * 12.162,96 = 15.690,03 \text{ ha.}$$

A Tabela 4.8 representa a matriz de erros e as proporções das classes Soja e Não Soja entre parênteses para o município de Palotina.

Tabela 4.8: Matriz de erros e proporções - Palotina.

Pontos de Referência			
Máscara	Soja	Não Soja	Total Linha
Soja	37 (0,673)	18 (0,327)	55
Não Soja	17 (0,386)	27 (0,614)	44
Total			99*

*Observação: um ponto amostral em Palotina não pode ser identificado.

Tem-se que:

$$T_{\text{Soja}} = 34.002,09 \text{ ha.}$$

$$T_{\text{Não Soja}} = 31.352,13 \text{ ha.}$$

Então:

$$A_{\text{Soja}} = 0,673 * 34.002,09 + 0,386 * 31.352,13 = 34.987,46 \text{ ha.}$$

$$A_{\text{Não Soja}} = 0,327 * 34.002,09 + 0,614 * 31.352,13 = 30.366,76 \text{ ha.}$$

A Tabela 4.9 representa a matriz de erros e as proporções das classes Soja e Não Soja entre parênteses para o município de Rancho Alegre do Oeste.

Tabela 4.9: Matriz de erros e proporções - Rancho Alegre do Oeste.

Pontos de Referência			
Máscara	Soja	Não Soja	Total Linha
Soja	61 (0,968)	2 (0,032)	63
Não Soja	25 (0,676)	12 (0,324)	37
Total			100

Tem-se que:

$$T_{\text{Soja}} = 14.565,33 \text{ ha.}$$

$$T_{\text{Não Soja}} = 9.557,10 \text{ ha.}$$

Então:

$$A_{\text{Soja}} = 0,968 * 14.565,33 + 0,676 * 9.557,10 = 20.560,44 \text{ ha.}$$

$$A_{\text{Não Soja}} = 0,032 * 14.565,33 + 0,324 * 9.557,10 = 3.561,99 \text{ ha.}$$

A Tabela 4.10 representa a matriz de erros e as proporções das classes Soja e Não Soja entre parênteses para o município de Toledo.

Tabela 4.10: Matriz de erros e proporções - Toledo.

Pontos de Referência			
Máscara	Soja	Não Soja	Total Linha
Soja	42 (0,824)	9 (0,176)	51
Não Soja	16 (0,327)	33 (0,673)	49
Total			100

Tem-se que:

$$T_{\text{Soja}} = 66.565,89 \text{ ha.}$$

$$T_{\text{Não Soja}} = 53.212,86 \text{ ha.}$$

Então:

$$A_{\text{Soja}} = 0,824 * 66.565,89 + 0,327 * 53.212,86 = 72.194,60 \text{ ha.}$$

$$A_{\text{Não Soja}} = 0,176 * 66.565,89 + 0,673 * 53.212,86 = 47.584,15 \text{ ha.}$$

A Tabela 4.11 representa a matriz de erros e as proporções das classes Soja e Não Soja entre parênteses para o município de Ubiratã.

Tabela 4.11: Matriz de erros e proporções - Ubiratã.

Pontos de Referência			
Máscara	Soja	Não Soja	Total Linha
Soja	57 (0,966)	2 (0,034)	59
Não Soja	19 (0,463)	22 (0,537)	41
Total			100

Tem-se que:

$$T_{\text{Soja}} = 36.069,84 \text{ ha.}$$

$$T_{\text{Não Soja}} = 29.265,21 \text{ ha.}$$

Então:

$$A_{\text{Soja}} = 0,966 * 36.069,84 + 0,463 * 29.265,21 = 48.409,06 \text{ ha.}$$

$$A_{\text{Não Soja}} = 0,034 * 36.069,84 + 0,537 * 29.265,21 = 16.925,99 \text{ ha.}$$

A Tabela 4.12 sintetiza os valores de área de soja calculados por contagem de pixels e corrigidos por matriz de erros para cada um dos oito municípios.

Tabela 4.12: Área de soja calculada por contagem de pixels e corrigida por matriz de erros.

Município	Área de Soja (ha)	
	Contagem de pixels	Corrigida por matriz de erros
Assis Chateaubriand	48.724,11	70.918,97
Cascavel	82.434,42	78.388,29
Mamborê	46.961,46	57.744,34
Maripá	16.324,11	12.797,04
Palotina	34.002,09	34.987,46
Rancho Alegre do Oeste	14.565,33	20.560,44
Toledo	66.565,89	72.194,60
Ubiratã	36.069,84	48.409,06

4.4. Estimativa de área de soja de referência por expansão direta

Para fins de comparação, também foi estimada a área de soja de cada um dos oito municípios por expansão direta, a partir dos 100 pontos do painel amostral reconhecidos como soja nas imagens Landsat 5/TM e em levantamento de campo, com base na área total do município obtida do IBGE, conforme apresentado na Tabela 4.13.

Tabela 4.13: Área de soja estimada por expansão direta.

Município	Pontos de Soja	Área Municipal IBGE (ha)	Área de Soja por Expansão Direta (ha)
Assis Chateaubriand	73	97.000,00	70.810,00
Cascavel	37	210.000,00	77.700,00
Mamborê	74	77.900,00	57.646,00
Maripá	42	28.400,00	11.928,00
Palotina	54	65.100,00	35.154,00
Rancho Alegre do Oeste	86	24.100,00	20.726,00
Toledo	58	119.700,00	69.426,00
Ubiratã	76	65.300,00	49.628,00

4.5. Estimativa de área de soja baseada em classificação *fuzzy*

A classificação *fuzzy* deve ser executada no período de maior vigor vegetativo da soja na região em estudo que correspondeu à segunda quinzena de janeiro de 2004, segundo conclusão baseada nos gráficos de perfis espectrais temporais dos índices de vegetação de imagens AVHRR-NOAA. Portanto, foi utilizada a composição máxima de 15 a 31 de janeiro de 2004 para obter a estimativa de área de soja em cada um dos oito municípios.

O método de classificação não supervisionado começa com a definição do número de amostras dos pixels que são selecionados aleatoriamente na imagem para inicializar os graus de pertinência da matriz de partição *fuzzy* e, também, com o número de classes. Neste caso o número de amostras utilizadas foi de 10.000 e o número de classes definido em oito.

Para a definição do fator de “fuzzificação” q foram testados os valores 1,5, 2, 2,5 e 3, onde não foi observada diferença significativa nos agrupamentos obtidos, com relação aos graus de pertinência dos pixels às classes. Por isso foi utilizado o valor mais usual de q igual a 2.

O algoritmo é executado iterativamente onde é calculada a média *fuzzy*, a matriz de covariância *fuzzy* e, assim, a distância EM dos pixels aos centros das classes. Com isso os graus de pertinência são recalculados e os pixels realocados entre as classes até a convergência. A condição de término é atingida quando os centros das classes ou os elementos da matriz de pertinência não mais tiverem alterações significativas, ou seja, não forem maiores que ϵ com relação à iteração anterior.

Com a identificação da imagem fração da classe Soja com base no resultado da classificação *fuzzy*, conforme mostrado na Figura 4.18, foi obtida a proporção de soja em cada pixel a partir dos graus de pertinência, aplicando-se a Equação (3.30). A partir disso foram geradas as estimativas de área plantada em cada um dos oito municípios da região em estudo.

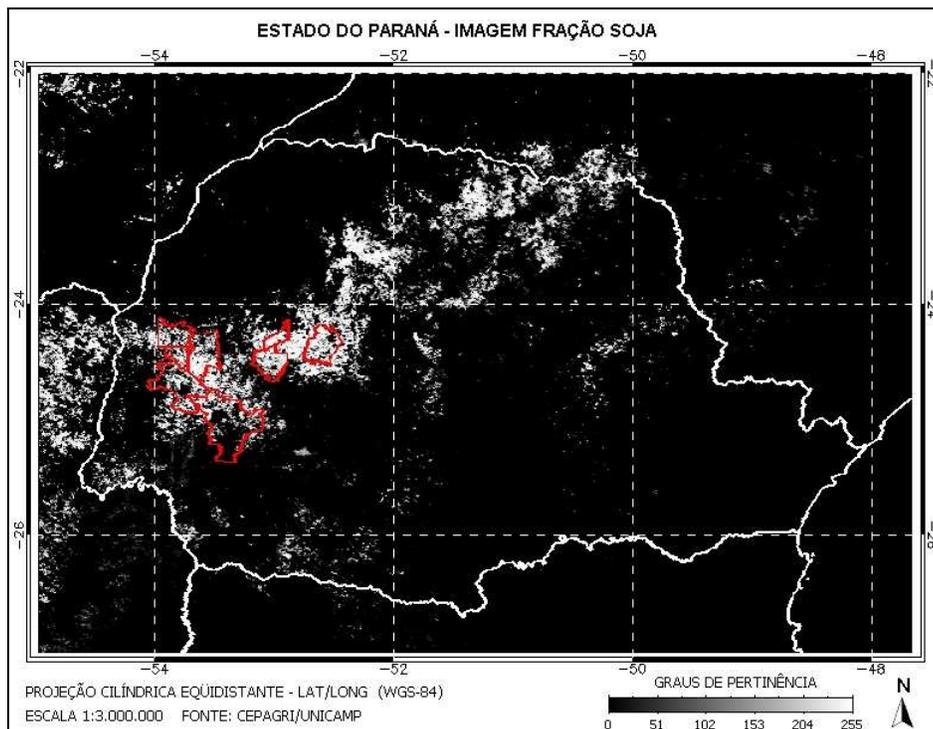


Figura 4.18: Imagem fração soja da região oeste do Paraná referente à safra 2003/2004.

O algoritmo é sensível às condições de inicialização porque a amostra aleatória dos pixels pode gerar uma matriz de partição *fuzzy* com diferentes valores dos graus de pertinência. Para reduzir esse problema, a classificação *fuzzy* foi repetida 10 vezes e as estimativas de área para os oito municípios foram obtidas considerando-se a média dos resultados, como mostrado na Tabela 4.14.

Tabela 4.14: Área de soja estimada com base na classificação *fuzzy*.

Município	Área de Soja por Classificação <i>Fuzzy</i> (ha)
Assis Chateaubriand	74.500,41
Cascavel	82.793,00
Mamborê	57.512,01
Maripá	18.553,24
Palotina	37.680,35
Rancho Alegre do Oeste	20.620,44
Toledo	69.253,47
Ubiratã	47.525,15

Para análise da acurácia da classificação *fuzzy*, os valores dos graus de pertinência da classe Soja foram relacionados com a proporção de soja de referência para cada pixel. Para isso, baseado em FOODY e COX (1994), a máscara de soja foi degradada espacialmente com um filtro de convolução passa baixa de dimensões 37 x 37 pixels, de forma que a resolução espacial ficasse similar à imagem AVHRR-NOAA, ou seja, 37 x 30 m que é igual a 1.110 m, aproximadamente 1,1 km. Em seguida foi reamostrada pelo método do vizinho mais próximo para um pixel de tamanho médio de 0,0098 graus.

As Figuras de 4.19 a 4.26 mostram lado a lado, para cada um dos oito municípios, as proporções (p_{soja}) de soja de referência da imagem degradada da máscara de soja e os graus de pertinência (μ_{soja}) da imagem fração soja da classificação *fuzzy*, respectivamente.

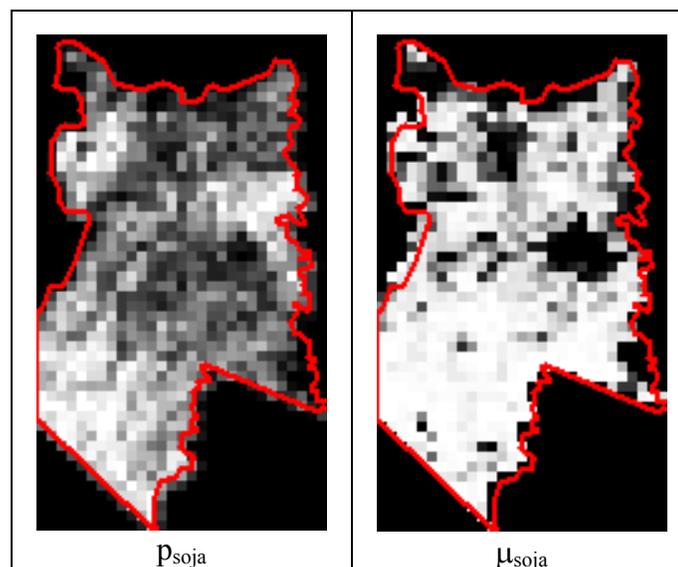


Figura 4.19: Proporções e graus de pertinência de soja - Assis Chateaubriand.

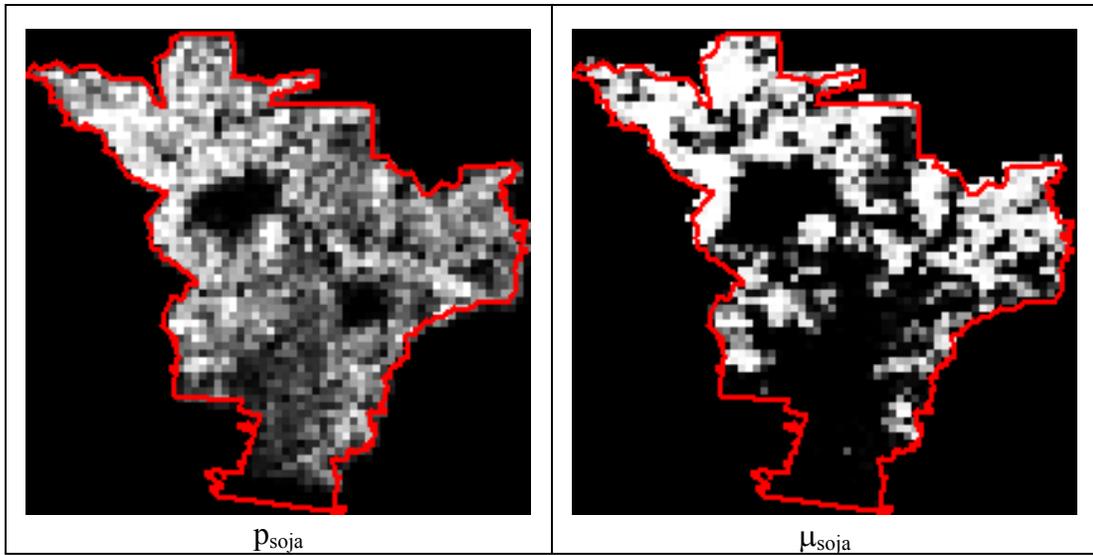


Figura 4.20: Proporções e graus de pertinência de soja - Cascavel.

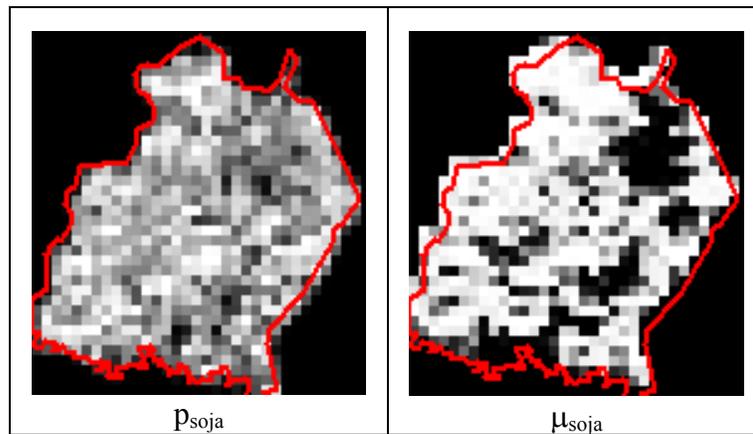


Figura 4.21: Proporções e graus de pertinência de soja - Mamborê.

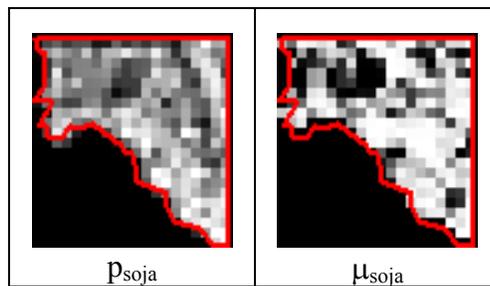


Figura 4.22: Proporções e graus de pertinência de soja - Maripá.

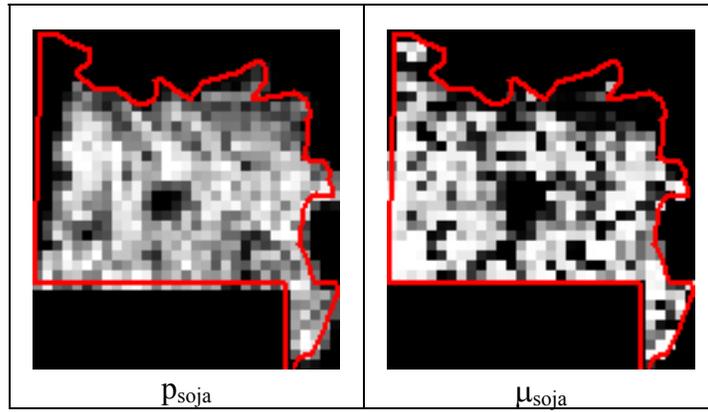


Figura 4.23: Proporções e graus de pertinência de soja - Palotina.

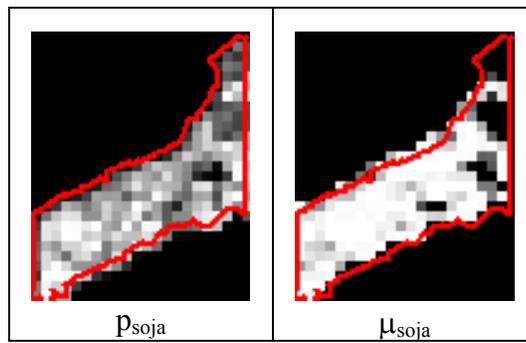


Figura 4.24: Proporções e graus de pertinência de soja - Rancho Alegre do Oeste.

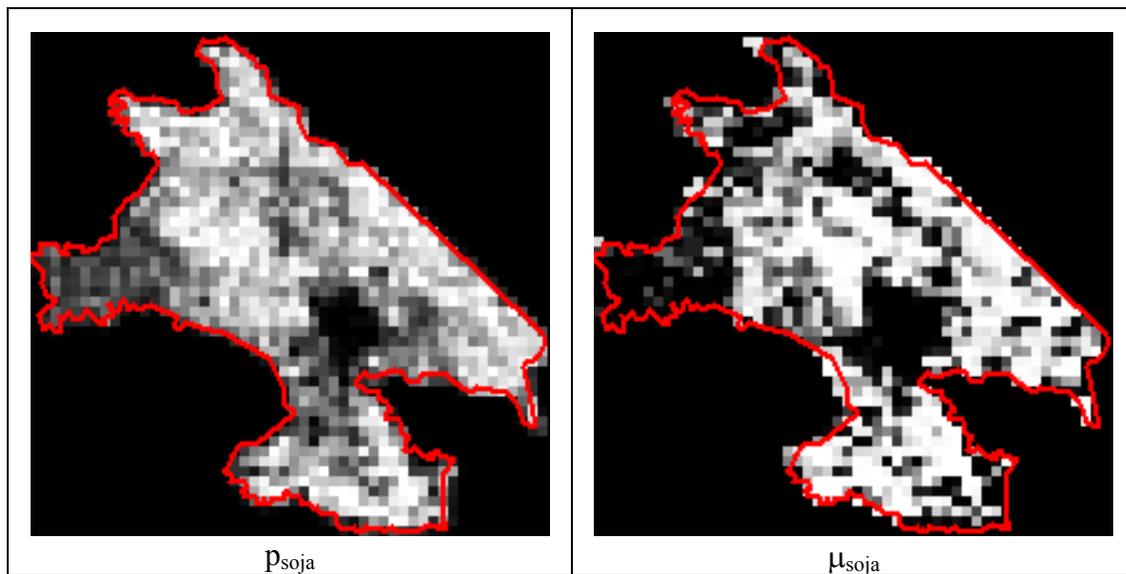


Figura 4.25: Proporções e graus de pertinência de soja - Toledo.

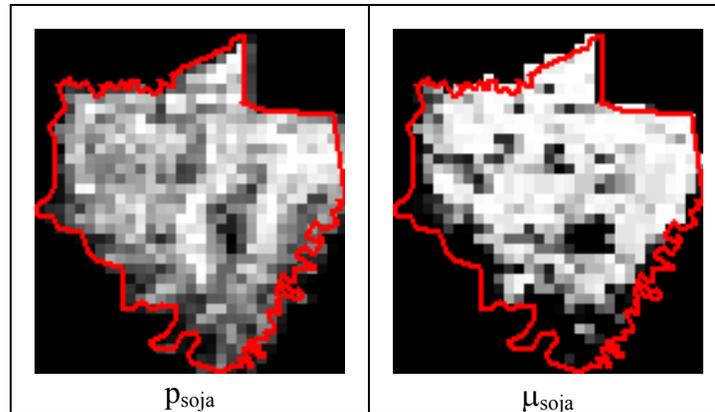


Figura 4.26: Proporções e graus de pertinência de soja - Ubiratã.

Para cada município foi calculado o coeficiente de correlação estatística (r) entre as proporções de soja de referência e os graus de pertinência da imagem fração soja, obtidas pixel a pixel, conforme apresentado na Tabela 4.15.

Tabela 4.15: Correlação entre as proporções e os graus de pertinência de soja.

Município	Correlação (r)
Assis Chateaubriand	0,546
Cascavel	0,617
Mamborê	0,565
Maripá	0,563
Palotina	0,521
Rancho Alegre do Oeste	0,585
Toledo	0,620
Ubiratã	0,686

Os valores de r obtidos foram considerados moderados, sendo o menor 0,521 para o município de Palotina e o maior de 0,686 para o município de Toledo. Como o p -value foi menor que 0,05 para todos os municípios, a associação entre as proporções e os graus de pertinência de soja foi considerada estatisticamente significativa, com um nível de confiança de 95%. Isso mostra uma tendência moderada e coerente dos dados, na mesma direção.

O grau das correlações obtidas pode ser explicado pelo fato de que o limitante de acurácia inerente dos classificadores booleanos que deram origem à máscara de soja, pode gerar proporções de soja que não correspondam exatamente à cobertura de soja em cada pixel. Por outro lado, mesmo sendo garantido que o erro de georreferenciamento das imagens

AVHRR-NOAA foi de no máximo um pixel, tal erro pode prejudicar a comparação das proporções de soja já que é realizada pixel a pixel.

4.6. Comparação entre as estimativas de área de soja

A Tabela 4.16 reúne as estimativas de área de soja oficiais da SEAB/DERAL e do IBGE, as de referência obtidas a partir da máscara de soja por meio de matriz de erros (ME) e por expansão direta (ED) e, por último, a baseada em classificação *fuzzy* que utilizou índices de vegetação de imagens AVHRR-NOAA, referentes aos oito municípios.

Tabela 4.16: Estimativas de área de soja.

Município	Área de Soja (ha)				
	SEAB	IBGE	ME	ED	Fuzzy
Assis Chateaubriand	72.000,00	73.568,00	70.918,97	70.810,00	74.500,41
Cascavel	82.000,00	78.200,00	78.388,29	77.700,00	82.793,00
Mamborê	54.000,00	54.060,00	57.744,34	57.646,00	57.512,01
Maripá	21.500,00	21.000,00	12.797,04	11.928,00	18.553,24
Palotina	43.300,00	43.600,00	34.987,46	35.154,00	37.680,35
Rancho A. do Oeste	15.400,00	15.505,00	20.560,44	20.726,00	20.620,44
Toledo	68.600,00	67.150,00	72.194,60	69.426,00	69.253,47
Ubiratã	51.200,00	51.000,00	48.409,06	49.628,00	47.525,15

O coeficiente de correlação estatística representa o resultado da comparação pareada das estimativas de área de soja e o *p-value* o nível de significância. O erro relativo (ER) pode ser utilizado para medir a precisão entre duas estimativas e é calculado pela subtração da segunda e a primeira, dividindo-se pelo valor da primeira estimativa. O valor de ER é multiplicado por 100 para ser expresso em porcentagem.

A Tabela 4.17 apresenta os valores de *r* com seu respectivo *p-value* e o ER na comparação duas a duas entre todas as estimativas de área de soja.

Tabela 4.17: Correlação e erro relativo entre as estimativas de área de soja.

Município	Erro Relativo (%)									
	SEAB	SEAB	SEAB	SEAB	IBGE	IBGE	IBGE	ME	ME	ED
	x IBGE	x ME	x ED	x Fuzzy	x ME	x ED	x Fuzzy	x ED	x Fuzzy	x Fuzzy
Assis Chateaubriand	2,18	-1,50	-1,65	3,47	-3,60	-3,75	1,27	-0,15	5,05	5,21
Cascavel	-4,63	-4,40	-5,24	0,97	0,24	-0,64	5,87	-0,88	5,62	6,55
Mamborê	0,11	6,93	6,75	6,50	6,82	6,63	6,39	-0,17	-0,40	-0,23
Maripá	-2,33	-40,48	-44,52	-13,71	-39,06	-43,20	-11,65	-6,79	44,98	55,54
Palotina	0,69	-19,20	-18,81	-12,98	-19,75	-19,37	-13,58	0,48	7,70	7,19
Rancho A. do Oeste	0,68	33,51	34,58	33,90	32,61	33,67	32,99	0,81	0,29	-0,51
Toledo	-2,11	5,24	1,20	0,95	7,51	3,39	3,13	-3,83	-4,07	-0,25
Ubiratã	-0,39	-5,45	-3,07	-7,18	-5,08	-2,69	-6,81	2,52	-1,83	-4,24
Correlação (r)	0,998	0,976	0,976	0,988	0,976	0,977	0,986	0,999	0,993	0,992
P-value	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

Os valores de r entre as estimativas de área de soja foram extremamente altos, superiores a 0,975. Como o *p-value* foi menor que 0,05 para todos os resultados das comparações, a associação entre as estimativas foi considerada estatisticamente significativa, com um nível de confiança de 95%. Isso indica que as estimativas são fortemente correlacionadas e possuem a mesma tendência, movendo-se em perfeita proporção na mesma direção.

A variação do ER é considerada baixa quando for menor que 10%. Então, entre as estimativas de área oficiais da SEAB/DERAL e do IBGE, não há disparidade significativa dos dados, onde o valor de r é igual a 0,998, apesar da subestimação de 4,63% com relação ao município de Cascavel.

A estimativa de área de soja de referência por meio de matriz de erros obtida para os municípios de Maripá e Palotina subestimaram as estimativas oficiais em torno de 40% e 19%, respectivamente. Já para o município de Rancho Alegre do Oeste, a estimativa de área superestimou por volta de 33% as estimativas oficiais. Para os demais municípios o ER ficou dentro da variação aceitável. O menor ER com relação a estimativa da SEAB/DERAL foi de -1,50% para o município de Assis Chateaubriand e em relação a estimativa do IBGE foi de 0,24% para o município de Cascavel.

Na comparação entre as estimativas de área de soja oficiais e de referência por expansão direta, os valores do ER seguem a mesma tendência da estimativa de área por meio de matriz de erros. As estimativas para os municípios de Maripá e Palotina subestimaram as estimativas oficiais por volta de 44% e 19%, respectivamente. A estimativa do município de Rancho Alegre do Oeste superestimou em torno de 34% as estimativas oficiais. O menor ER da estimativa em relação a estimativa da SEAB/DERAL foi 1,20% para o município de Toledo e com relação a estimativa do IBGE foi de -0,64% novamente para o município de Cascavel.

Ao se analisar o ER entre as estimativas de área de referência por meio de matriz de erros e por expansão direta, observam-se valores muito baixos para todos os municípios. O maior ER foi apresentado para o município de Maripá, onde a estimativa por expansão direta subestimou em 6,79% a por matriz de erros. Mesmo com as diferenças encontradas em relação as estimativas oficiais para os municípios de Maripá, Palotina e Rancho Alegre do Oeste, vale destacar a grande similaridade entre as estimativas para todos os municípios, apresentando o maior valor de r na análise pareada, igual a 0,999. Cabe lembrar que essas estimativas basearam-se em métodos totalmente objetivos utilizando imagens Landsat 5/TM com pixel de tamanho 30 x 30 m, ao contrário das estimativas de área de soja oficiais que são obtidos de forma subjetiva.

A estimativa de área de soja baseada em classificação *fuzzy* para os municípios de Maripá e Palotina subestimaram as estimativas oficiais em torno de 13%. A estimativa do município de Rancho Alegre do Oeste superestimou em torno de 34% as estimativas oficiais. Para os demais municípios o ER ficou dentro da variação aceitável. O menor ER com relação a estimativa da SEAB foi de 0,95% para o município de Toledo e em relação a estimativa do IBGE foi de 1,27% para o município de Assis Chateaubriand.

Na comparação com as estimativas de área de soja de referência por matriz de erros e por expansão direta, a estimativa de área baseada em classificação *fuzzy* obtida para o município de Maripá foi superestimada em torno de 50%. Para os demais municípios o ER também ficou dentro da variação aceitável. O menor ER com relação a estimativa por matriz de erros foi de 0,29% para o município de Rancho Alegre do Oeste e em relação a estimativa por expansão direta foi de -0,23% para o município de Mamborê.

A estimativa de área de soja baseada em classificação *fuzzy* também se aproximou mais das estimativas de referência por matriz de erros e por expansão direta, com exceção do município de Maripá, apresentando valores de r iguais a 0,993 e 0,992, respectivamente. Essa estimativa também se baseou num método objetivo utilizando índices de vegetação de imagens AVHRR-NOAA, porém o tamanho do pixel de 1,1 x 1,1 km diminui o detalhamento da superfície terrestre.

5. CONCLUSÕES

Para viabilizar esta pesquisa foi necessário buscar uma metodologia para o georreferenciamento automático e preciso das imagens AVHRR-NOAA, problema comum e sem solução em vários softwares comerciais. Aliar automação e precisão no georreferenciamento das imagens foi uma tarefa necessária de infraestrutura do trabalho, em função do grande número de imagens e da necessidade de precisão geométrica dentro de um pixel, para obtenção de estimativas de área de soja municipais confiáveis.

A partir dessa demanda foi desenvolvido um sistema automático de processamento das imagens para conversão de formato bruto, calibração radiométrica, georreferenciamento preciso e geração de produtos das imagens AVHRR-NOAA.

O sistema se mostrou uma ferramenta importante aos usuários de imagens AVHRR-NOAA, principalmente àqueles que fazem uso de grandes séries temporais, por automatizar o processamento e a geração de produtos das imagens.

O georreferenciamento é a etapa fundamental de todo o processo. A navegação indireta, aliada à busca automática de feições por meio da MCC, resultou em imagens geometricamente precisas e se mostrou eficaz para o sucesso da automação. Uma outra etapa de extrema importância foi o mascaramento de nuvens. Esses procedimentos visam garantir a precisão e a qualidade final dos produtos gerados pelo sistema.

A metodologia descrita neste trabalho sobre o georreferenciamento é de domínio público e está disponível a qualquer usuário. O sistema é aberto e pode ser aperfeiçoado para a geração de outros produtos úteis a diversas aplicações e, ainda, ser adaptado para georreferenciar imagens de outros sensores.

Este estudo mostrou que o menos conhecido índice de vegetação NDMI também pode ser utilizado para detecção de mudanças na cultura da soja ao longo de uma série multitemporal de imagens AVHRR-NOAA.

A máscara de soja gerada a partir de imagens Landsat 5/TM mostrou-se bastante útil na delimitação das áreas representativas de soja tomadas como referência para a geração dos perfis temporais do NDVI e do NDMI de imagens AVHRR-NOAA.

O método de estimativa de área de referência por meio de matriz de erros propiciou a correção das áreas estimadas pela máscara de soja, obtendo-se alta correlação com as estimativas oficiais da SEAB/DERAL e do IBGE.

As estimativas de área de soja de referência por meio de matriz de erros apresentaram valores muito similares às estimativas de área por expansão direta, ambos os métodos realizados de maneira objetiva, podendo ser um indicativo de melhor precisão devido a melhor resolução espacial das imagens Landsat 5/TM.

A estimativa de área baseada em classificação *fuzzy* utilizou o NDVI e o NDMI gerados a partir das imagens diárias e gratuitas AVHRR-NOAA. No caso da cultura de soja no oeste do Paraná, foi possível estimar as áreas de soja nos oito municípios até o final de janeiro de 2004, período de maior vigor vegetativo da soja na região em estudo, com antecedência em relação à época da colheita, ao contrário dos levantamentos oficiais que se estendem até o final da safra, além de utilizarem dados subjetivos vindos do campo. Essas estimativas mostraram-se altamente correlacionadas com as estimativas de área de referência obtidas a partir da máscara de soja e por expansão direta, sendo um indicativo de boa precisão. E, também, apresentaram alta correlação balizadas com estimativas oficiais da SEAB/DERAL e do IBGE. Em ambas as comparações, o nível de erro relativo geral foi aceitável. As proporções de soja de referência e os graus de pertinência da imagem fração soja mostraram-se moderadamente correlacionados.

A estimativa de área baseada em classificação *fuzzy* utilizando índices de vegetação de imagens de satélites AVHRR-NOAA mostrou ser viável, uma vez que é realizada de forma objetiva, possui menor custo e pode ser realizada com antecedência em relação às estimativas de área oficiais. Portanto, essa metodologia pode ser empregada com o objetivo de auxiliar os métodos convencionais.

Para resolver o problema da sensibilidade do algoritmo de classificação *fuzzy* com relação às condições de inicialização, sugere-se implementar o método supervisionado em que as amostras de treinamento são definidas pelo usuário.

O método de estimativa de área baseada em classificação *fuzzy* pode gerar melhores resultados quando aplicado a imagens de satélites com melhor resolução espacial, onde a mistura de classes dentro de um pixel também ocorre, porém com mais detalhamento da superfície terrestre.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ANTUNES, J. F. G.; ESQUERDO, J. C. D. M.; ZULLO JR, J. Desenvolvimento de um sistema automático para a geração de produtos derivados de imagens AVHRR-NOAA. [CD-ROM] In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, XII., **Anais...** Goiânia, 2005.

ATKINSON, P. M.; CUTLER, M. E. J.; LEWIS, H. Mapping sub-pixel proportional land cover with AVHRR imagery. **International Journal of Remote Sensing**, v. 18, n. 4, p. 917-935, 1997.

BASTIN, L. Comparison of fuzzy c-means classification, linear mixture modelling and MLC probabilities as tools for unmixing coarse pixels. **International Journal of Remote Sensing**, v. 18, n. 17, p. 3629-3648, 1997.

BATISTA, G. T.; TARDIN, A. T.; CHEN, S. C.; DALLEMAND, J. F. Avaliação de produtos HRV/SPOT e TM/LANDSAT na discriminação de culturas. **Pesquisa Agropecuária Brasileira**, v. 25, n. 3, p. 379-386, 1990.

BEZDEK, J. C.; PAL, S. K. **Fuzzy Models for Pattern Recognition**, IEEE Press, New York, 1992.

BEZDEK, J.; EHRLICH, R; FULL, W. FCM: the fuzzy c-means clustering algorithm. **Computers and Geosciences**, v. 10, n. 2, p. 191-203, 1984.

BOWEN, M.; EMERY, W. J.; WILKING, J. L.; TILDESLEY, P. C.; BARTON, I. J.; KNEWTSON, R. Extracting multiyear surface currents from sequential thermal imagery using the Maximum Cross-Correlation technique. **Journal of Atmospheric and Oceanic Technology**, v. 19, n. 10, p. 1665-1676, 2002.

CANTY, M. **Image analysis and pattern recognition for remote sensing with examples in ENVI/IDL**. Disponível em: <<http://www.fz-juelich.de/ste/datapool/page/210/zfl.html>>. Acesso em: 28. mar. 2005.

CHEN, P. Y., SRINIVASAN, R., FEDOSEJEVS, G., NARASIMHAN, B. An automated cloud detection method for daily NOAA-14 AVHRR data for Texas, USA. **International Journal of Remote Sensing**, v. 23, n. 15, p. 2939-2950, 2002.

CHEN, P. Y.; SRINIVASAN, R.; FEDOSEJEVS, G.; KINIRY, J. R. Evaluating different NDVI compositing techniques using NOAA-14 AVHRR data. **International Journal of Remote Sensing**, v. 24, n. 17, p. 3403-12, 2003.

COCHRAN, W. G. **Sampling Techniques**. 3 Edition. John Wiley & Sons, 413 p., 1977.

COLWELL, R. N. (Ed.) **Manual of remote sensing**. Falls Church, Virginia: American Society of Photogrammetry, 2. ed., v. 1, 1232 p., 1983.

COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). **Safras**. Disponível em: <<http://www.conab.gov.br>>. Acesso em 10. fev. 2004.

DELGADO, M. R D. B. da S. **Projeto automático de sistemas nebulosos: uma abordagem co-evolutiva**. UNICAMP, Campinas, SP, 2002. 186p. (Tese de Doutorado).

DI, L.; RUNDQUIST, D. C. A one-step algorithm for correction and calibration of AVHRR level 1B data. **Photogrammetric Engineering and Remote Sensing**, v. 60, n. 2, p.165-171, 1994.

DINGUIRARD, M.; SLATER, P. N. Calibration of space-multispectral imaging sensors: a review. **Remote Sensing of Environment**, v. 68, n.3, p.194-205, 1999.

DUDA, T.; CANTY, M. J. Unsupervised classification of satellite imagery: choosing a good algorithm, **International Journal of Remote Sensing**, v. 23, n. 11, p. 2193-2212, 2002.

EMBRAPA SOJA. **Tecnologia de produção de soja - Paraná 2005**. Londrina - PR, 224p., 2004.

EMBRAPA SOJA; EMBRAPA CERRADOS; EMBRAPA AGROPECUÁRIA OESTE; FUNDAÇÃO MERIDIONAL. **Tecnologia de produção de soja - região central do Brasil 2005**. Londrina - PR, 239 p., 2004.

EMERY, W. J.; BROWN, J.; NOVAK, Z. P. AVHRR image navigation: summary and review. **Photogrammetric Engineering and Remote Senging**, v. 55, n. 8, p.1175-1183, 1989.

EMERY, W. J.; BALDWIN, D.; MATTHEWS, D. Maximum Cross Correlation Automatic Satellite Image Navigation and Attitude Corrections for Open Ocean Image Navigation, **IEEE Transactions on Geoscience and Remote Sensing**, v. 41, n. 1, p. 33-42, 2003.

EMERY, W. J.; THOMAS, A. C.; COLLINS, M. J. An objective method for computing advective surface velocities from sequential infrared satellite images. **Journal of Geophysical Research**, v. 91, n. 11, p. 12865-12878, 1986.

EPIPHANIO, J. C. N.; LUIZ, A. J. B.; FORMAGGIO, A. R. Estimativa de Áreas Agrícolas Municipais utilizando Sistema de Amostragem Simples sobre Imagens de Satélite. **Bragantia**, Campinas, v. 61, n. 2, p. 187-197, 2002.

FEHR, W. R.; CAVINESS, C. E. **Stages of soybean development**. Ames: Iowa State University of Science and Technology, 12 p., 1981. (Special Report, 80).

FONTANA, D. C.; DUCATI, J. R.; WEBER, E.; BERLATO, M. A.; FIGUEIREDO, D. C.; BERGAMASCHI, H. **Monitoramento e previsão da safra de soja 1999/2000 no Brasil**. Porto Alegre: UFRGS/CONAB, 2001, 121P. (Série Relatório Técnico).

FONTANA, D. C.; WEBER, E.; DUCATI, J.; FIGUEIREDO, D. C.; BERLATO, M. A.; BERGAMASCHI, H. A case study for crop monitoring and harvest forecast in south Brazil. In: INTERNATIONAL SYMPOSIUM ON REMOTE SENSING OF ENVIRONMENT, 28. Cape Town. **Proceedings...** p. 91-94, 2000.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). **Multiple frame agricultural surveys: current surveys based on area and list sampling methods**. Rome: FAO, v. 1., 119 p., 1996. (FAO Statistical Development Series, 7).

FOODY, G. M. Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution. **International Journal of Remote Sensing**, v. 19, n. 13, p. 2593-2599, 1998.

FOODY, G. M. Status of Land Cover Classification Accuracy Assessment. **Remote Sensing of Environment**, v. 80, n. 1, p. 185-201, 2002.

FOODY, G. M., COX, D. P. Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. **International Journal of Remote Sensing**, v. 15, n. 3, p. 619-631, 1994.

FORMAGGIO, A. R. **Características agronômicas e espectrais para sensoriamento remoto de trigo e de feijão**. ESALQ, Piracicaba, SP, 161p., 1989. (Tese de Doutorado).

FRANÇA, G. B.; CRACKNELL, A. P. A simple cloud masking approach using NOAA AVHRR daytime data for tropical areas. **International Journal of Remote Sensing**, v. 16, n. 6, p. 1697-1705, 1995.

GANAN, J. R.; ROCHA, J. V.; MERCANTE, E.; ANTUNES, J. F. G. Mapeamento da cultura da soja com imagens LANDSAT 5/TM utilizando algoritmos de classificação supervisionada. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, XII., **Anais...** Goiânia, 2005.

GATH, I; GEVA, A. B. Unsupervised optimal fuzzy clustering. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 3, n. 3, p. 773-781, 1989.

GOODRUM, G.; KIDWELL, K.; WINSTON, W. (Eds.) **NOAA KLM user's guide**. September 2000 Revision. Disponível em: <<http://www2.ncdc.noaa.gov/docs/klm>>. Acesso em 02. fev. 2004.

HOFFER, R. M. Biological and physical considerations in applying computer-aided analysis techniques to remote sensor data. In: SWAIN, P.H.; DAVIS, S.M. (Ed.) **Remote sensing: the quantitative approach**. New York: McGraw-Hill, 1978.

HOLBEN, B. N. Characteristics of maximum value composite images from temporal AVHRR data. **International Journal of Remote Sensing**, v. 7, n. 11, p. 1417-1435, 1986.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Levantamento sistemático da produção agrícola**. Disponível em: <<http://www.ibge.gov.br/>>. Acesso em 27. abr. 2005.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Pesquisas agropecuárias**. Rio de Janeiro: IBGE, 96 p., 2002. (Relatórios metodológicos)

IPPOLITI-RAMILO, G.; EPIPHANIO, J. C. N.; SHIMABUKURO, Y. E.; FORMAGGIO, A. R. Sensoriamento remoto orbital como meio auxiliar na previsão de safras. **Agricultura em São Paulo**, v. 46, n. 1, p. 89-101, 1999.

JACKSON, R. D.; HUETE, A. R. Interpreting vegetation indices. **Preventive Veterinary Medicine**, v. 11, n. 2, p. 185-200, 1991.

JAKUBAUSKAS, M. E.; LEGATES, D. R.; KASTENS, J. H. Crop identification using harmonic analysis of time-series AVHRR NDVI data. **Computers and Electronics in Agriculture**, v. 37, n. 1, p. 127-139, 2002.

JIN, S.; SADER, S. A. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. **Remote Sensing of Environment**, v. 94, n. 3, p. 364-372, 2005.

JORDAN, C. F. Derivation of leaf-area index from quality of light on the forest floor. **Ecology**, v. 50, n.1, p. 663-666, 1969.

JUPP, D. L. B. The Stability of Global Estimates from Confusion Matrices. **International Journal of Remote Sensing**, v. 10, n. 9, p. 1563-1569, 1989.

KIDWELL, K. (Ed.) **NOAA polar orbiter data user's guide**. November 2000 Revision. Disponível em: <<http://www2.ncdc.noaa.gov/docs/podug/index.htm>> Acesso em 25. fev. 2004.

KOLLENKARK, J. C., C. S. T. DAUGHTRY, M. E. BAUER, T. L. HOUSLEY. Effects of cultural practices on agronomic and reflectance characteristics of soybean canopies. **Agronomy Journal**, v. 74, p. 751–758, 1982.

LANDIS, J. R.; KOCH, G. G. The Measurement of Observer Agreement for Categorical Data. **Biometrics**, v. 33, n. 1, p. 159-174, 1977.

LIU, W.T.; KOGAN, F. Monitoring Brazilian soybean production using NOAA/AVHRR based vegetation condition indices. **International Journal of Remote Sensing**, v. 23, n. 6, p. 1161-1179, 2002.

LUIZ, A. J. B. **Estatísticas agrícolas por amostragem auxiliadas pelo sensoriamento remoto**. INPE, São José dos Campos, SP, 2003. (Tese de Doutorado).

MIKA, J; KERÉNYI, J; RIMÓCZI-PAÁL, A.; MERZA, Á.; SZINELL, C.; CSISZÁR, I. On correlation of maize and wheat yield with NDVI: Example of Hungary (1985-1998), **Advances in Space Research**, v. 30, n. 11, p. 2399-2404, 2002.

MOREAU, S.; BOSSENO R; GU; X. F.; BARET, F. Assessing the biomass dynamics of Andean bofedal and totora high-protein wetland grasses from NOAA/AVHRR. **Remote Sensing of Environment**, v. 85, n. 4, p. 516-529, 2003.

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

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION (NOAA). **Comprehensive Large Array-data Stewardship System (CLASS)**. Disponível em <<http://www.class.noaa.gov/nsaa/products/welcome>>. Acesso em 11. mai. 2005.

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION (NOAA). **The global vegetation index user's guide**. SDS/NESDIS/NCDC. Washington D.C., 1983.

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION SATELLITE INFORMATION SYSTEM (NOAASIS). **Satellite Instrument Calibration**. Disponível em <<http://noaasis.noaa.gov/NOAASIS/ml/calibration.html>>. Acesso em 10. mar. 2004.

NIKHIL, R. P.; BEZDEK, J. C. On cluster validity for the fuzzy c-means model. **IEEE Transactions on Fuzzy Systems**, v. 3, n. 3, 1995 p. 370 - 379, 1995.

PEARSON, R L.; L. D. MILLER. 1972. Remote mapping of standing crop biomass for estimation of the productivity of the shortgrass prairie. In: INTERNATIONAL SYMPOSIUM ON REMOTE SENSING OF ENVIRONMENT, 8., Colorado. **Proceedings...** p. 1355-1379, 1972.

PERGOLA, N.; TRAMUTOLI, V. SANA: sub-pixel automatic navigation of AVHRR imagery. **International Journal of Remote Sensing**, v. 21, n. 12, p. 2519-2524, 2000.

PERGOLA, N.; TRAMUTOLI, V. Two years of operational use of Subpixel Automatic Navigation of AVHRR scheme: accuracy assessment and validation. **Remote Sensing of Environment**, v. 85, n. 2, p. 190-203, 2003.

PIMENTEL-GOMES, F. **Curso de estatística experimental**. 14 Edição. Piracicaba: ESALQ, 477 p., 2000.

PINO, F. A. Estatísticas agrícolas para o século XXI. **Agricultura em São Paulo**, v. 46, n. 2, p. 71-105, 1999.

PRICE, K. P.; CROOKS, T. J.; MARTINKO, E. A. Grasslands across time and scale: a remote sensing perspective. **Photometric Engineering and Remote Sensing**, v. 67, n. 4, p. 414-420, 2001.

RANGSANSERI, Y.; PROMCHAROEN, S.; THITIMAJSHIMA, P. A fuzzy neural network approach for multispectral imagery analysis. In: GEOSPATIAL INFORMATION IN AGRICULTURE AND FORESTRY, 2., **Proceedings...** Lake Buena Vista, Florida: ERIM, v.1, p.383-390, 2000.

RAO, C. R. N.; CHEN, J. Post-launch calibration of the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on the NOAA-14 spacecraft. **International Journal of Remote Sensing**, v. 17, p. 2743-2747, 1996.

RICHARDS, J. A.; JIA, X., **Remote sensing digital image analysis: an introduction**, 3 Edition. New York: Springer Verlag, 363p., 1999.

ROSBOROUGH, G. W.; BALDWIN, D. G.; EMERY, W. J. Precise AVHRR image navigation. **IEEE Transactions on Geoscience and Remote Sensing**, v. 32, n. 3, p.644-657, 1994.

ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the great plains with ERTS. In: EARTH RESOURCES TECHNOLOGY SATELLITE-1 SYMPOSIUM, 3., **Proceedings...** Washington, D. C.: NASA, v. 1, p. 309-317, 1973.

SEAB/DERAL. **Secretaria da Agricultura e do Abastecimento do Paraná/Departamento de Economia Rural**. Disponível em: <<http://www.pr.gov.br/seab/ativi2.shtml>>. Acesso em 27. abr. 2005.

SENAY, G.B.; ELLIOT, R.L. Capability of AVHRR data in discriminating rangeland cover mixtures. **International Journal of Remote Sensing**, v.23, n.2, p.299-312, 2002.

TOWNSHEND, J. R. G. Global data sets for land applications from the Advanced Very High Resolution Radiometer: an introduction. **International Journal of Remote Sensing**, v. 15, n. 17, p. 3319-3332, 1994.

TOWNSHEND, J. R. G.; TUCKER, C. J. Objective assessment of Advanced Very High Resolution Radiometer data for land cover mapping. **International Journal of Remote Sensing**, v. 5, n. 2, p. 497-504, 1984.

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

TUCKER, C. J. Red and photographic infrared linear combination for monitoring vegetation. **Remote Sensing of Environment**, v. 8, n. 2, p. 127-150, 1979.

WANG, F. Fuzzy supervised classification of remote sensing images. **IEEE Transactions on Geoscience and Remote Sensing**, v. 28, n. 2, p. 194-201, 1990a.

WANG, F. Improving remote sensing image analysis through fuzzy information representation. **Photogrammetric Engineering and Remote Sensing**, v. 56, n. 8, p.1163-1169, 1990b.

WANG, Q; ADIKU, S; TENHUNEN, J; GRANIER, A. On the relationship of NDVI with leaf area index in a deciduous forest site, **Remote Sensing of Environment**, v. 94, n. 2, p. 244-255, 2005.

WIEGAND, G. L.; RICHARDSON, A. J.; ESCOBAR, D. E. Vegetation indices in crop assessment. **Remote Sensing of Environment**, v. 35, n. 2, p. 105-119, 1991.

WILSON, E. H.; SADER, S. A. Detection of forest harvest type using multiple dates of Landsat TM imagery. **Remote Sensing of Environment**, v. 80, n. 3, p. 385-396, 2002.

XIN, J.; YU, Z.; VAN LEEUWEN, L.; DRIESSEN, P.M. Mapping crop key phenological stages in the North China Plain using NOAA time series images. **International Journal of Applied Earth Observation and Geoinformation**, v. 4, n. 9, p. 109-117, 2002.

ZADEH, L. A. Fuzzy sets. **Information and Control**, v. 8, p. 338-353, 1965.

ZADEH, L. A. Fuzzy sets and systems. **International Journal of General Systems**, v. 17, p. 129-138, 1990.

ZHENG, D.; KAINZ, W.; GROTEN, S. Integrated approach of GIS and neural fuzzy network for linguistic decision-making on land use management. In: GEOSPATIAL INFORMATION IN AGRICULTURE AND FORESTRY, 2., **Proceedings...** Lake Buena Vista, Florida: ERIM, v. 1, p. 95-102, 2000.