



JOÃO FRANCISCO GONÇALVES ANTUNES

CLASSIFICAÇÃO DE SÉRIES TEMPORAIS DE DADOS
MODIS BASEADA EM REDES NEURO-FUZZY PARA O
MONITORAMENTO SISTEMÁTICO DO CULTIVO DA
CANA-DE-AÇÚCAR

CAMPINAS

2014



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Agrícola

JOÃO FRANCISCO GONÇALVES ANTUNES

CLASSIFICAÇÃO DE SÉRIES TEMPORAIS DE DADOS
MODIS BASEADA EM REDES NEURO-FUZZY PARA O
MONITORAMENTO SISTEMÁTICO DO CULTIVO DA
CANA-DE-AÇÚCAR

Tese apresentada à Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas como parte dos requisitos exigidos para obtenção do título de Doutor em Engenharia Agrícola, na área de concentração de Planejamento e Desenvolvimento Rural Sustentável.

Orientador: Prof. Dr. Rubens Augusto Camargo Lamparelli
Coorientador: Prof. Dr. Luiz Henrique Antunes Rodrigues

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE
DEFENDIDA PELO ALUNO JOÃO FRANCISCO GONÇALVES ANTUNES
E ORIENTADO PELO PROF. DR. RUBENS AUGUSTO CAMARGO LAMPARELLI.

Assinatura do Orientador

A handwritten signature in blue ink, appearing to be "R.", is positioned above a horizontal line that serves as a signature line.

CAMPINAS

2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Luciana Pietrosanto Milla - CRB 8/8129

An89c Antunes, João Francisco Gonçalves, 1965-
Classificação de séries temporais de dados MODIS baseada em redes neuro-fuzzy para o monitoramento sistemático do cultivo da cana-de-açúcar / João Francisco Gonçalves Antunes. – Campinas, SP : [s.n.], 2014.

Orientador: Rubens Augusto Camargo Lamparelli.
Coorientador: Luiz Henrique Antunes Rodrigues.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Agrícola.

1. Sensoriamento remoto. 2. Processamento de imagens. 3. Análise harmônica. 4. Reconhecimento de padrões. 5. Estatística agrícola. I. Lamparelli, Rubens Augusto Camargo. II. Rodrigues, Luiz Henrique Antunes. III. Universidade Estadual de Campinas. Faculdade de Engenharia Agrícola. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Classification of time-series MODIS data based neuro-fuzzy networks for the systematic monitoring of sugarcane cultivation

Palavras-chave em inglês:

Remote sensing

Image processing

Harmonic analysis

Pattern recognition

Agricultural statistic

Área de concentração: Planejamento e Desenvolvimento Rural Sustentável

Titulação: Doutor em Engenharia Agrícola

Banca examinadora:

Rubens Augusto Camargo Lamparelli [Orientador]

Jansle Vieira Rocha

Stanley Robson de Medeiros Oliveira

Heloísa de Arruda Camargo

Jerry Adriani Johann

Data de defesa: 25-08-2014

Programa de Pós-Graduação: Engenharia Agrícola

Este exemplar corresponde à redação final da **Tese de Doutorado** defendida por **João Francisco Gonçalves Antunes**, aprovada pela Comissão Julgadora em 25 de agosto de 2014, na Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas.

R.

Prof. Dr. Rubens Augusto Camargo Lamparelli – Presidente e Orientador
Feagri/Unicamp

Prof. Dr. Jansle Vieira Rocha – Membro Titular
Feagri/Unicamp

Prof. Dr. Stanley Robson de Medeiros Oliveira – Membro Titular
CNPTIA/Embrapa

Prof. Dra. Heloísa de Arruda Camargo
UFSCar/São Carlos

Prof. Dr. Jerry Adriani Johann – Membro Titular
Unioeste

RESUMO

O setor agrícola brasileiro está sendo marcado por um novo ciclo de expansão do cultivo da cana-de-açúcar. O Brasil é hoje o maior produtor de cana-de-açúcar do mundo, sendo o Estado de São Paulo o maior produtor nacional, com grandes áreas de plantio. Nesse sentido, a estimativa confiável de área plantada da cana-de-açúcar é de fundamental importância para o agronegócio sucroalcooleiro. As geotecnologias têm sido empregadas nas estimativas de safras agrícolas para diminuir o nível de subjetividade dos métodos tradicionais. As imagens do sensor MODIS fornecem uma ampla cobertura da superfície da Terra com alta periodicidade, que possibilitam o monitoramento agrícola sistemático. Entretanto, a sua moderada resolução espacial faz com que possa ocorrer a mistura espectral de diferentes classes de cobertura do solo dentro de um mesmo pixel, acarretando problemas de acurácia na obtenção de estimativas de área. Nesse contexto, o objetivo principal do trabalho foi desenvolver uma metodologia de classificação automática baseada em redes neuro-fuzzy utilizando séries temporais de índices de vegetação MODIS, para estimar a área plantada de cana-de-açúcar no nível sub-pixel. Considerando as safras 2004/2005 a 2011/2012 analisadas no Estado de São Paulo, os mapeamentos gerados pelo modelo de classificação Fuzzy ARTMAP obtiveram uma alta exatidão e baixa discordância, aliado a uma forte correlação linear com as proporções de cana-de-açúcar de referência. Com isso, foi possível obter a estimativa de área plantada no final de março, com antecedência em relação à colheita, ao contrário dos levantamentos oficiais que se estendem até o final da safra, além de utilizarem dados subjetivos. As estimativas de área plantada de cana-de-açúcar baseadas nas medidas de Compromisso do classificador Fuzzy ARTMAP mostraram-se fortemente correlacionadas e em concordância com as estimativas de área de referência do Canasat, também bem balizadas com as estimativas oficiais do IBGE, em nível municipal, sendo um indicativo de boa precisão. O nível dos desvios entre as estimativas de área mostrou uma variação média municipal menor em relação aos dados do Canasat do que os dados do IBGE. O desempenho do classificador Fuzzy ARTMAP está intrinsecamente relacionado à caracterização da distribuição geográfica do cultivo da cana-de-açúcar nas mesorregiões do Estado de São Paulo, alcançando melhores ajustes onde o cultivo é mais expressivo. Na consolidação para o nível de mesorregiões, os ajustes alcançaram um desempenho ainda superior, indicado por valores extremamente altos de correlação e concordância. Na comparação da área plantada com os dados do Zoneamento Agroecológico da

cana-de-açúcar ao longo das oito safras, observou-se que a expansão do cultivo da cana-de-açúcar predomina essencialmente em regiões aptas, com a tendência recente de incorporar mais áreas para produção agrícola do que pastagens. A metodologia de estimativa de área plantada baseada nas medidas de Compromisso do classificador Fuzzy ARTMAP foi eficiente para o mapeamento da cana-de-açúcar, demonstrando grande potencial para a análise sub-pixel de séries temporais de índices de vegetação MODIS.

Palavras-chave: Sensoriamento remoto; Processamento de imagens; Análise harmônica; Reconhecimento de padrões; Estatística agrícola.

ABSTRACT

The Brazilian agricultural sector has been marked by a new cycle of expansion of sugarcane cultivation. Currently, Brazil is the largest sugarcane producer and the São Paulo State is the largest national producer, with an extensive cropping area. In this sense, a reliable estimation of sugarcane crop area is essential for the sugar-ethanol agribusiness. Geotechnologies have been employed on agricultural crop estimates to reduce the level of subjectivity of the traditional methodologies. MODIS sensor images provide a wide coverage of the Earth's surface with high periodicity, supporting the systematic agricultural monitoring. However, its moderate spatial resolution leads to the spectral mixture of different land cover classes within the same pixel, causing accuracy problems on crop area estimation. In this context, the main objective of the study was to develop an automatic classification methodology based on neuro-fuzzy networks using MODIS vegetation indices time-series to estimate the sugarcane crop areas at sub-pixel level. Considering the analyzed cropping years in the São Paulo State, from 2004/2005 to 2011/2012, the maps generated by the Fuzzy ARTMAP classification model showed high accuracy and low disagreement, in addition to a strong linear correlation with the proportions of the sugarcane reference. Then, it was possible to obtain the crop area estimation at the end of March, prior to the harvest period, unlike the official surveys, which extend until the end of the harvest and use subjective data. The sugarcane crop area estimation based on the Commitment measures from the Fuzzy ARTMAP classifier showed to be strongly correlated and in agreement with the reference area estimation of Canasat, also well marked out with the official estimation from IBGE, at municipal level, being an indicative of good accuracy. The deviation level between the estimations showed a lower municipal average variation in relation to Canasat data than IBGE data. The performance of the Fuzzy ARTMAP classifier is intrinsically related to the characterization of geographical distribution of the sugarcane cultivation in the mesoregions of the São Paulo State, reaching best adjustments where the cultivation is more expressive. In the data consolidation at the mesoregion level, the adjustments achieved even greater performance, demonstrated by extremely high values of correlation and agreement. Comparing the crop area estimation with the Sugarcane Agroecological Zoning in the eight crop seasons, it was observed that this cultivation predominates mostly in suitable regions, with the recent trend of incorporating more areas for agricultural production than pastures. The methodology for crop

area estimation, based on the Commitment measures of the Fuzzy ARTMAP classifier, was efficient for sugarcane crop area mapping, showing great potential for the sub-pixel analysis of MODIS time-series of vegetation indices.

Keywords: Remote sensing; Image processing; Harmonic analysis; Pattern recognition; Agricultural statistic.

SUMÁRIO

1. INTRODUÇÃO	1
1.1. Hipótese	3
1.2. Objetivos.....	3
2. REVISÃO BIBLIOGRÁFICA	5
2.1. A cultura da cana-de-açúcar	5
2.1.1. Importância socioeconômica	5
2.1.2. Fisiologia da produção	7
2.2. Estimativas de safras agrícolas	12
2.3. O sensoriamento remoto aplicado ao monitoramento agrícola	15
2.3.1. Comportamento espectral da vegetação	18
2.3.2. Índices de vegetação.....	20
2.3.3. Imagens do sensor MODIS	22
2.3.4. Análise de séries temporais de imagens	26
2.4. Inteligência artificial na classificação de imagens	29
2.4.1. Noções de redes neurais artificiais	31
2.4.2. Noções da teoria de conjuntos fuzzy	33
2.4.3. Classificação sub-pixel de imagens.....	37
3. MATERIAL E MÉTODOS	41
3.1. Material.....	41
3.1.1. Caracterização da área de estudo.....	41
3.1.2. Séries temporais MODIS.....	42
3.1.3. Mapas temáticos de referência	45
3.1.4. Dados meteorológicos	46
3.1.5. Estatísticas agrícolas.....	47
3.1.6. Informações cartográficas.....	48
3.1.7. Recursos computacionais	48

3.2.	Métodos	49
3.2.1.	Decomposição das séries temporais MODIS	49
3.2.2.	Reconhecimento do padrão espectro-temporal da cana-de-açúcar.....	54
3.2.3.	Extração de métricas fenológicas sazonais da cana-de-açúcar.....	58
3.2.4.	Classificação Fuzzy ARTMAP	61
3.2.5.	Estimativa sub-pixel de área plantada de cana-de-açúcar	66
3.2.6.	Análises Estatísticas	67
3.2.7.	Padronização do ano-safra.....	72
4.	RESULTADOS E DISCUSSÕES	75
4.1.	Reconhecimento do padrão espectro-temporal da cana-de-açúcar.....	75
4.2.	Decomposição das séries temporais MODIS	77
4.3.	Extração de métricas fenológicas sazonais da cana-de-açúcar.....	84
4.4.	Conversão de escala do Canasat.....	86
4.5.	Seleção do conjunto de amostras.....	89
4.6.	Classificação Fuzzy ARTMAP	93
4.7.	Estimativa sub-pixel de área plantada de cana-de-açúcar	108
4.8.	Comparação entre as estimativas de área plantada de cana-de-açúcar.....	113
4.9.	Cruzamento dos dados de área plantada de cana-de-açúcar com o ZAE	128
4.10.	Considerações finais e sugestões.....	130
5.	CONCLUSÕES.....	133
6.	REFERÊNCIAS BIBLIOGRÁFICAS	135

Dedico
à minha adorável esposa GLÁUCIA e
nossos lindos filhos GABRIEL e ARTHUR,
com muito amor e carinho.

AGRADECIMENTOS

A Deus, pela vida, saúde e proteção.

À Empresa Brasileira de Pesquisa Agropecuária por esta oportunidade excepcional de capacitação acadêmica. Em especial a Dra. Silvia M. F. S. Massruhá, como Chefe Adjunta de P&D, por ter permitido a utilização da infraestrutura física e computacional da Embrapa Informática Agropecuária, que auxiliou sobremaneira no transcorrer do curso de doutorado.

À Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas pela oportunidade de aprendizado com a pós-graduação na realização deste trabalho.

Ao Prof. Dr. Rubens A. C. Lamparelli e ao Prof. Dr. Luiz Henrique A. Rodrigues pela orientação, aprendizado, contribuições e grande parceria profissional estabelecida ao longo do trabalho, definitivamente consolidada para o futuro.

In memoriam ao Dr. José Demisio S. Silva, pesquisador do INPE falecido em 30 de abril de 2011, pela orientação e grande cooperação no início do projeto de pesquisa de doutorado, principalmente nas sugestões de utilização das técnicas de inteligência artificial.

Ao Dr. Zhe Li, pesquisador do Ministério do Meio Ambiente do Canadá, pela parceria técnica-científica estabelecida na utilização da tecnologia de software para classificação de imagens de sensoriamento remoto empregada neste trabalho.

Aos membros das bancas do exame de qualificação e de defesa da tese de doutorado pelas sugestões que contribuíram para a melhoria do trabalho. Em especial, ao Dr. Stanley R. M. Oliveira pelas reuniões técnicas com outros grupos de pesquisa após o falecimento do Dr. Demisio e ao Dr. Jerry A. Johan pelas ótimas dicas na análise estatística dos resultados.

Ao Dr. Bernardo F. T. Rudorff pelo repasse dos dados do Canasat que foram fundamentais para o desenvolvimento do trabalho.

Ao Dr. Marcos Adami, pesquisador do INPE, pelas providenciais informações sobre a qualidade temática dos dados Canasat.

A Raízen, nas pessoas de Fernando A. Benvenuti e Jefferson R. B. Mello, pelo fornecimento dos dados de campo de áreas de cultivo da cana-de-açúcar.

Ao Dr. Sandro E. M. Pereira da Embrapa Meio Ambiente pelo repasse dos dados do Zoneamento Agroecológico da cana-de-açúcar.

Ao Dr. Antonio R. Formaggio pela orientação, apoio e compreensão durante a minha passagem pelo INPE.

Ao Michael Steinmayer, Diretor Executivo da SulSoft, pelas ótimas dicas na utilização do ENVI/IDL.

A secretaria de Pós-graduação da Faculdade de Engenharia Agrícola pelo auxílio e atendimento prestados, em especial a sempre educada Rita C. C. Ferreira, a Ritinha.

Ao apoio técnico de Agmon Rocha do Laboratório de Geoprocessamento e de André A. Nogueira do Laboratório de Informática da Faculdade de Engenharia Agrícola.

A Ana P. Montagner da Coordenadoria Geral da Unicamp pela amizade e incentivo de sempre.

Ao meu amigo Aparecido T. Souza, o Cidão, pelo apoio de infraestrutura e ao Juliano Ribeiro pelo suporte computacional na Embrapa Informática Agropecuária.

Ao meu grande amigo Luiz Manoel S. Cunha pela parceria nas disciplinas do curso de doutorado e apoio de sempre.

Ao meu grande amigo Júlio C. D. M. Esquerdo pela atuação como tutor do trabalho, com valiosas discussões técnicas, sugestões e pela ajuda prestada na utilização do ArcGis.

Ao meu grande amigo Edmundo C. Oliveira pelo forte incentivo para cursar a pós-graduação.

Aos meus grandes amigos José G. Jardine e Alvaro S. Neto pelos ótimos momentos vividos durante esses meus 20 anos de Embrapa.

Ao Dr. Sebastião T. Silva pelo incansável zelo com os meus filhos que me deu muita tranquilidade durante o doutorado.

À minha esposa Gláucia pela compreensão, apoio, companheirismo, paciência e dedicação para cuidar dos nossos queridos filhos em tantos momentos da minha ausência.

Aos meus queridos pais Jacinto e Benice por terem sempre me incentivado nos estudos e a minha irmã Liliane e minha sobrinha afilhada Manuella pela convivência nesta fase de nossas vidas.

Muito Obrigado!

“No que diz respeito ao empenho, ao compromisso, ao esforço, à dedicação, não existe meio termo. Ou você faz uma coisa bem feita ou não faz.”

Ayrton Senna

LISTA DE FIGURAS

Figura 2.1: Mapa da produção de cana-de-açúcar no Brasil.....	6
Figura 2.2: Produção e área colhida de cana-de-açúcar nas últimas 10 safras.	6
Figura 2.3: Ciclo fenológico da cultura da cana-de-açúcar.	8
Figura 2.4: Ciclo da cultura de cana-de-açúcar com cinco cortes.	10
Figura 2.5: Zoneamento Agroecológico (ZAE) da cana-de-açúcar para o Estado de São Paulo.	13
Figura 2.6: Interação da radiação eletromagnética solar com uma folha.	16
Figura 2.7: Interação da radiação eletromagnética fonte-alvo-sensor.	17
Figura 2.8: Curva padrão de refletância da vegetação.....	18
Figura 2.9: Evolução da refletância de um dossel vegetal agrícola entre fases fenológicas.	19
Figura 2.10: Dinâmica do NDVI de um dossel vegetal agrícola entre as fases fenológicas de crescimento e senescência.	22
Figura 2.11: Representação dos 17 quadrantes MODIS que cobrem o Brasil.	25
Figura 2.12: Quadrantes MODIS que cobrem o Estado de São Paulo.	26
Figura 2.13: Perfis temporais de NDVI apresentados por diferentes alvos.....	27
Figura 2.14: Neurônio biológico.....	31
Figura 2.15: Neurônio artificial.	32
Figura 2.16: Formato das funções de pertinência - (a) Triangular, (b) Trapezoidal, (c) Gaussiana, (d) Conjunto unitário.	35
Figura 2.17: Operações básicas em conjuntos fuzzy - (a) Conjuntos fuzzy A e B, (b) União, (c) Interseção, (d) Complemento.	36
Figura 2.18: Esquematização de um classificador neuro-fuzzy.	37
Figura 3.1: Mapa de área plantada média de cana-de-açúcar (ha) das oito safras por mesorregiões do Estado de São Paulo.....	41
Figura 3.2: Processamento do produto MOD13Q1 para alimentação do banco de imagens.	42
Figura 3.3: Banco de Produtos MODIS.....	43
Figura 3.4: Produtos MODIS para o Estado de São Paulo.	44
Figura 3.5: Canasat.	45
Figura 3.6: Agritempo.....	46

Figura 3.7: SIDRA.....	47
Figura 3.8: Fluxograma das etapas da metodologia do trabalho.	49
Figura 3.9: Representação dos termos harmônicos.	50
Figura 3.10: Espaço de cores HLS.....	53
Figura 3.11: <i>Wavelet</i> da família Daubechies 8.	56
Figura 3.12: Ambiente para visualização de séries temporais MODIS.....	56
Figura 3.13: Áreas de cultivo da cana-de-açúcar de referência.	57
Figura 3.14: Métricas fenológicas sazonais extraídas pelo TIMESAT.	59
Figura 3.15: Parametrização do TIMESAT para extração das métricas fenológicas.....	61
Figura 3.16: Arquitetura da rede neural Fuzzy ARTMAP.	62
Figura 4.1: Perfis temporais do EVI2 nas áreas de cultivo da cana-de-açúcar de referência.	75
Figura 4.2: Ciclo fenológico característico da cana-de-açúcar no Estado de São Paulo.	77
Figura 4.3: Representação HLS dos termos harmônicos de EVI do período de outubro a março das oito safras.....	79
Figura 4.4: Representação HLS dos termos harmônicos de NDVI do período de outubro a março das oito safras.....	80
Figura 4.5: Representações HLS dos termos harmônicos de EVI e NDVI para a mesorregião de Ribeirão Preto de outubro a março da safra 2007/2008.....	81
Figura 4.6: Representações HLS dos termos harmônicos de EVI e NDVI para a mesorregião de São José do Rio Preto de outubro a março da safra 2007/2008.....	81
Figura 4.7: Mapas de precipitação acumulada.....	82
Figura 4.8: Métricas fenológicas sazonais nível base e taxa de senescência sínteses da série temporal de EVI.	84
Figura 4.9: Métricas fenológicas sazonais nível base e taxa de senescência sínteses da série temporal de NDVI.....	85
Figura 4.10: Conversão de escala do Canasat.....	86
Figura 4.11: Mapa de referência do Canasat da safra 2011/2012.....	87
Figura 4.12: Mapa de proporções de cana-de-açúcar da safra 2011/2012.....	88
Figura 4.13: Geração de polígonos homogêneos de treinamento da safra 2011/2012.	90
Figura 4.14: Amostras de treinamento da safra 2011/2012.	91

Figura 4.15: Amostras de teste da safra 2011/2012.....	92
Figura 4.16: Interface do Fuzzy ARTMAP no IDRISI.....	94
Figura 4.17: Modelo de classificação ART da safra 2011/2012.....	95
Figura 4.18: Medidas de acurácia do modelo de classificação ART da safra 2011/2012.....	96
Figura 4.19: Pós-processamento do modelo de classificação CMT Cana da safra 2011/2012. ...	98
Figura 4.20: Pós-processamento da imagem proporções Cana da safra 2011/2012.....	99
Figura 4.21: Relação entre as proporções Cana e as medidas CMT Cana após o procedimento de pós-processamento.....	100
Figura 4.22: Curva ROC entre o modelo de classificação CMT Cana e a imagem booleana de referência da safra 2011/2012.....	101
Figura 4.23: Modelos de classificação ART.....	102
Figura 4.24: Modelos de classificação CMT Cana.....	103
Figura 4.25: Imagens de proporções Cana.....	104
Figura 4.26: Imagens de referência Cana e Não-Cana.....	105
Figura 4.27: Relação entre a área plantada de cana-de-açúcar do IBGE e do Canasat em nível municipal nas oito safras.....	114
Figura 4.28: Relação entre a área plantada de cana-de-açúcar do Canasat e do Fuzzy ARTMAP CMT em nível municipal nas oito safras.....	116
Figura 4.29: Relação entre a área plantada de cana-de-açúcar do IBGE e do modelo Fuzzy ARTMAP CMT em nível municipal nas oito safras.....	118
Figura 4.30: Relação entre a área plantada de cana-de-açúcar do IBGE e do Canasat em nível de mesorregiões das oito safras.....	124
Figura 4.31: Relação entre a área plantada de cana-de-açúcar do Canasat e do Fuzzy ARTMAP CMT em nível de mesorregiões das oito safras.....	125
Figura 4.32: Relação entre a área plantada de cana-de-açúcar do IBGE e do modelo Fuzzy ARTMAP CMT em nível de mesorregiões das oito safras.....	126

LISTA DE TABELAS

Tabela 2.1: Especificações das bandas espectrais do sensor MODIS.....	23
Tabela 3.1: Matriz de erros.	68
Tabela 3.2: Matriz de proporções.....	69
Tabela 3.3: Anos-safra das fontes de dados.	73
Tabela 4.1: Conjuntos de amostras.	93
Tabela 4.2: Exatidão global, índice Kappa e discordância dos modelos de classificação ART.	106
Tabela 4.3: Erros de omissão e comissão das classes Cana e Não-Cana dos modelos de classificação ART.	106
Tabela 4.4: Correlação e curva ROC dos modelos de classificação ART Cana.....	107
Tabela 4.5: Estimativas de área plantada de cana-de-açúcar da safra 2004/2005.....	109
Tabela 4.6: Estimativas de área plantada de cana-de-açúcar da safra 2005/2006.....	109
Tabela 4.7: Estimativas de área plantada de cana-de-açúcar da safra 2006/2007.....	110
Tabela 4.8: Estimativas de área plantada de cana-de-açúcar da safra 2007/2008.....	110
Tabela 4.9: Estimativas de área plantada de cana-de-açúcar da safra 2008/2009.....	111
Tabela 4.10: Estimativas de área plantada de cana-de-açúcar da safra 2009/2010.....	111
Tabela 4.11: Estimativas de área plantada de cana-de-açúcar da safra 2010/2011.....	112
Tabela 4.12: Estimativas de área plantada de cana-de-açúcar da safra 2011/2012.....	112
Tabela 4.13: Medidas estatísticas médias das oito safras.....	121
Tabela 4.14: Expansão da área plantada de cana-de-açúcar em relação ao ZAE.	129

1. INTRODUÇÃO

O setor agrícola brasileiro está sendo marcado por um novo ciclo de expansão do cultivo da cana-de-açúcar para obtenção do álcool combustível, não apenas para substituição do petróleo devido à elevação dos preços internacionais, mas por questões relacionadas a mudanças climáticas, já que o etanol é uma fonte alternativa de energia renovável e menos poluidora. Dentre as fontes de energia usadas no Brasil, a cana-de-açúcar passou a ser a segunda matriz energética, superando a energia hidráulica, atrás apenas do petróleo.

A região Centro-Sul produz quase toda a cana-de-açúcar do país, em que o Estado de São Paulo é o maior produtor nacional, onde vários municípios têm grandes áreas de plantio e várias usinas instaladas, gerando milhares de empregos diretos e indiretos. Nos últimos anos, a crescente importância socioeconômica da cultura canavieira vem refletindo um aumento bastante significativo de área plantada.

A estimativa confiável de área plantada da cana-de-açúcar, obtida com antecedência à época da colheita, é de fundamental importância para o governo e setores envolvidos no agronegócio. A previsão de escoamento e a comercialização da produção agrícola são estratégicas para o país, pois estão diretamente relacionados com o planejamento, custos e preço. Além disso, a quantificação de áreas ocupadas pela cultura é um ponto crítico para a concessão de crédito agrícola e pagamento do seguro rural.

Em sua maioria, as estimativas de área plantada são realizadas de forma subjetiva, com dados levantados junto aos produtores e entidades relacionadas à atividade agrícola. Por isso, possuem custo elevado e execução demorada, tendo em vista a dinâmica e a extensão agrícola.

As geotecnologias têm sido empregadas nas estimativas de safras agrícolas principalmente para diminuir o nível de subjetividade. As imagens da superfície terrestre adquiridas por sensoriamento remoto, em conjunto com técnicas de geoprocessamento, apresentam grande potencial de discriminação de culturas agrícolas para obter estimativas objetivas de área plantada.

Os satélites de monitoramento ambiental TERRA e AQUA, que carregam a bordo o sensor MODIS (*Moderate Resolution Imaging Spectroradiometer*), vêm operando regularmente nos últimos anos, proporcionando uma visão sinóptica e repetitiva de grandes áreas para estudos da vegetação. Essas características possibilitam um monitoramento sistemático do

desenvolvimento da cana-de-açúcar, por ser uma cultura semiperene e plantada em áreas de grande extensão.

Entretanto, as imagens geradas pelo sensor MODIS possuem limitações quanto ao tamanho do pixel devido a sua moderada resolução espacial e isso faz com que possa ocorrer a mistura espectral de diferentes classes de cobertura do solo dentro de um mesmo pixel, dificultando o detalhamento necessário para a obtenção das estimativas de área plantada da cultura da cana-de-açúcar.

Os resultados mais recentes enfatizam a necessidade de se buscar novos estimadores de desempenho mais favorável para o reconhecimento de padrões em imagens de satélites. Nesse contexto, a inteligência artificial no âmbito do sensoriamento remoto pode ser compreendida como sendo o estado da arte. A aplicação de técnicas de aprendizado de máquina, combinando redes neurais e lógica fuzzy (nebulosa), é bastante promissora para a estimativa de área plantada no nível sub-pixel, a partir da classificação automática de séries temporais de imagens de satélite.

O potencial de manuseio de incertezas e de controle de sistemas complexos, tornados possíveis pela lógica fuzzy, estão sendo integrados às redes neurais artificiais, que por sua vez possuem características de adaptação e aprendizagem. Essa sinergia vem gerando novos tipos de classificadores híbridos que agregam os potenciais e as características individuais em sistemas adaptativos e inteligentes.

O presente texto está organizado como segue, apresentado o enunciado da hipótese científica que norteia esta tese de doutorado, o objetivo geral e específicos do trabalho.

No capítulo 2 é desenvolvida a revisão bibliográfica sobre os tópicos relacionados ao tema de pesquisa, apresentando a cultura da cana-de-açúcar, assim como a descrição dos métodos de estimativas de safras agrícolas no Brasil. Além disso, é feita uma explanação sobre o sensoriamento remoto aplicado ao monitoramento agrícola e, também, noções gerais das técnicas de inteligência artificial na classificação sub-pixel de imagens de sensoriamento remoto.

No capítulo 3 são descritos o material e métodos que foram utilizados no desenvolvimento do trabalho, incluindo a aquisição dos dados e os procedimentos de análise de séries temporais, da classificação sub-pixel de imagens e das análises estatísticas.

No capítulo 4 são relatados os resultados da classificação de séries temporais de índices de vegetação MODIS baseada em redes neuro-fuzzy para estimar a área plantada de cana-de-açúcar no nível sub-pixel no Estado de São Paulo, bem como a discussão desses resultados para o

monitoramento sistemático do cultivo da cana-de-açúcar. Também são apresentadas as considerações finais sobre os resultados gerais do trabalho e sugestões de estudos futuros.

No capítulo 5 são apresentadas as principais conclusões relacionadas aos objetivos do trabalho. Ao final são listadas as referências bibliográficas utilizadas na pesquisa.

1.1. Hipótese

O monitoramento sistemático da cultura da cana-de-açúcar poderá ser realizado a partir de séries temporais de imagens do sensor MODIS, por meio da combinação de técnicas de redes neurais e lógica fuzzy para o reconhecimento de padrões na classificação automática de imagens, para fornecer estimativas objetivas de área plantada de cana-de-açúcar, pois pode descrever as proporções de cada classe de cobertura do solo dentro do pixel.

1.2. Objetivos

O objetivo principal do trabalho foi desenvolver uma metodologia de classificação automática baseada em redes neuro-fuzzy utilizando séries temporais de índices de vegetação MODIS, para estimar a área plantada de cana-de-açúcar no nível sub-pixel no Estado de São Paulo, ao longo das safras 2004/2005 a 2011/2012.

Dentre os objetivos específicos, destacam-se:

- Mapear a cultura da cana-de-açúcar por meio do modelo de classificação neuro-fuzzy, com base nas séries temporais de índices de vegetação MODIS.
- Estimar a área plantada de cana-de-açúcar no nível sub-pixel a partir do modelo de classificação neuro-fuzzy, com antecedência em relação à colheita.
- Comparar as estimativas de área plantada de cana-de-açúcar obtidas pela classificação neuro-fuzzy com as estimativas oficial e de referência.
- Analisar a dinâmica e a tendência da expansão de área de cana-de-açúcar, cruzando os resultados com os dados do Zoneamento Agroecológico (ZAE).

2. REVISÃO BIBLIOGRÁFICA

2.1. A cultura da cana-de-açúcar

2.1.1. Importância socioeconômica

A cana-de-açúcar é originária da Ásia e foi trazida para o Brasil pelos portugueses, a partir do cultivo na Ilha da Madeira, em Portugal. Desde então exerce notável influência nos ciclos econômicos e no agronegócio brasileiro. Além de ser o maior produtor de cana-de-açúcar do mundo, o Brasil é, também, o maior exportador de açúcar e o segundo maior produtor de etanol. Metade da produção de açúcar é exportada tendo uma participação aproximada de 35% na balança comercial. Além disso, a agroindústria sucroalcooleira representa 6% dos empregos brasileiros. A região Centro-Sul do país é responsável por 90% do total da produção de cana-de-açúcar nacional, sendo o Estado de São Paulo o maior produtor, respondendo por mais de 60% de toda produção de açúcar e etanol (UNICA, 2014).

A safra 2013/2014 de cana-de-açúcar atingiu uma produção de 658,8 milhões de toneladas. Desse total, 54,8% foi destinada à extração de 27,9 bilhões de litros de etanol e 45,2% à fabricação de 37,9 milhões de toneladas de açúcar, além da bioeletricidade proveniente da queima da matéria seca do bagaço e da palha. A área colhida total foi de 8,8 milhões de hectares, concentrada nos estados da região Centro-Sul, com São Paulo tendo a participação de 51,7%, Goiás com 9,3%, Minas Gerais com 8,9%, Mato Grosso do Sul com 7,4% e Paraná com 6,7%, somando mais de 80% da produção brasileira (CONAB, 2014).

A crescente importância socioeconômica da cultura canavieira tem promovido um aumento bastante significativo do cultivo da cana-de-açúcar no Brasil. A Figura 2.1 mostra o mapa da produção de cana-de-açúcar, com as regiões onde estão concentradas as plantações e as unidades produtoras. Atualmente, o cultivo da cana-de-açúcar ocupa aproximadamente 2% das terras agrícolas do país, dos quais, 99,7% estão distantes da Floresta Amazônica. Nos últimos 25 anos, o maior crescimento da produção da cana-de-açúcar aconteceu nos Estados de Goiás, São Paulo e Mato Grosso do Sul, com a expansão ocorrendo principalmente em áreas já utilizadas para a produção agrícola e sobre pastagens, não promovendo de forma expressiva a criação de novas fronteiras agrícolas (MACEDO, 2007).

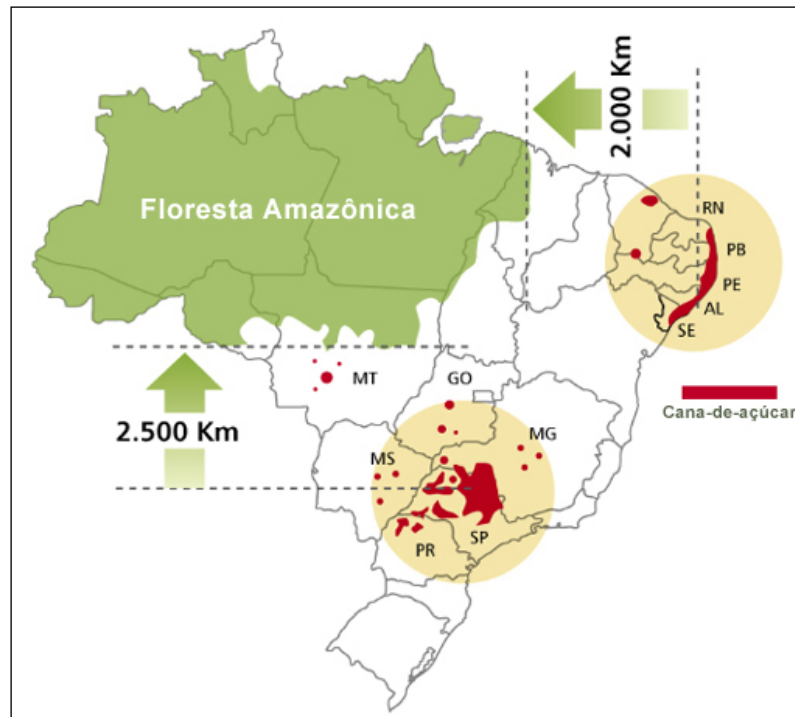


Figura 2.1: Mapa da produção de cana-de-açúcar no Brasil.
 Fonte: NIPE/Unicamp, IBGE, CTC, elaborado por UNICA (2014).

Segundo NEVES e TROMBIN (2014), na última década vem ocorrendo uma expansão bastante acentuada do cultivo da cana-de-açúcar no Brasil, onde a produção apresentou um crescimento de 84% e a área colhida de 64% nesses 10 anos, com uma taxa composta anual (CAGR - *Compound Annual Growth Rate*) de 7% e 6%, respectivamente, conforme representado pela Figura 2.2.

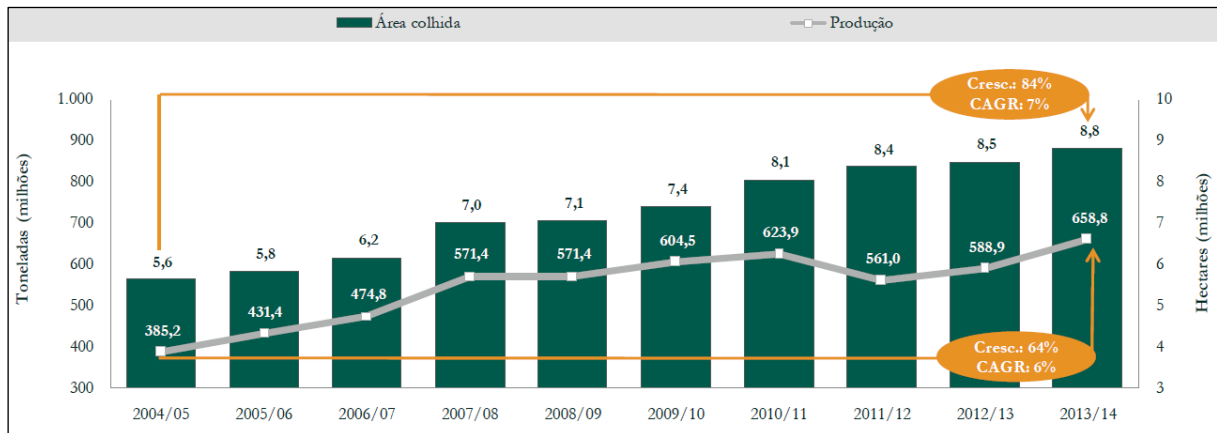


Figura 2.2: Produção e área colhida de cana-de-açúcar nas últimas 10 safras.
 Fonte: CONAB, elaborado por NEVES e TROMBIN (2014).

De acordo com CAMARGO et al. (2008), a expansão da cultura canavieira no Estado de São Paulo entre 2001 e 2006 teve um aumento de área plantada de 37,4%, determinando novas configurações regionais das explorações agropecuárias, principalmente na região Oeste do estado. De forma geral, nesse período a cana-de-açúcar incorporou áreas de pastagens e a diminuição de área de culturas alimentares ainda pouco afetou a produção de alimentos, apesar de significativa em alguns locais.

Nos próximos anos, a cana-de-açúcar deverá continuar se expandindo nas regiões Oeste e Noroeste do Estado de São Paulo, nos Estados de Goiás, Mato Grosso do Sul, Mato Grosso, Paraná e em Minas Gerais, principalmente no Triângulo Mineiro. Poderão ser consideradas ainda como novas fronteiras o Vale do São Francisco, a região Oeste da Bahia e, também, as regiões Sul do Maranhão e do Piauí. Contando com uma privilegiada situação territorial e climática, com oferta de áreas agrícolas em clima predominantemente tropical e com boa disponibilidade hídrica, o Brasil tem condições de continuar atuando de forma estratégica para liderar a produção mundial de cana-de-açúcar, sem prejuízo para a segurança alimentar e com impactos ambientais reduzidos, promovendo a intensificação de áreas já destinadas à produção agrícola, estimulando o manejo sustentável das terras e o aproveitamento de áreas degradadas (BRASIL, 2006).

2.1.2. Fisiologia da produção

A cana-de-açúcar pertence à família das poáceas (gramíneas), considerada um híbrido complexo derivado dos cruzamentos entre espécies do gênero *Saccharum* que possuem a capacidade de acumular sacarose nos colmos, própria de climas tropicais e subtropicais, tendo rápido crescimento, reprodução abundante e o aproveitamento econômico de toda a planta. Como pode ter vários ciclos de desenvolvimento é caracterizada como uma cultura semiperene, existindo dois ciclos básicos de produção. O ciclo da cana-planta começa com o plantio e termina após a primeira colheita e o ciclo da cana-soca que se inicia após a colheita da cana-planta e continua com sucessivas colheitas das soqueiras até a erradicação da cultura. O ciclo produtivo completo da cana-de-açúcar comumente dura de quatro a sete anos, perfazendo três a seis cortes antes de ser replantada, para posteriormente ser feita a reforma do canavial (SCARPARI e BEAUCLAIR, 2010).

O ciclo agrícola fenológico da cana-de-açúcar considerando a cana-planta e a cana-soca, ilustrado na Figura 2.3, é composto pelas fases de brotação e emergência dos brotos, perfilhamento e estabelecimento da cultura, crescimento dos colmos e maturação, que são descritas a seguir, baseado em CASAGRANDE e VASCONCELOS (2010).

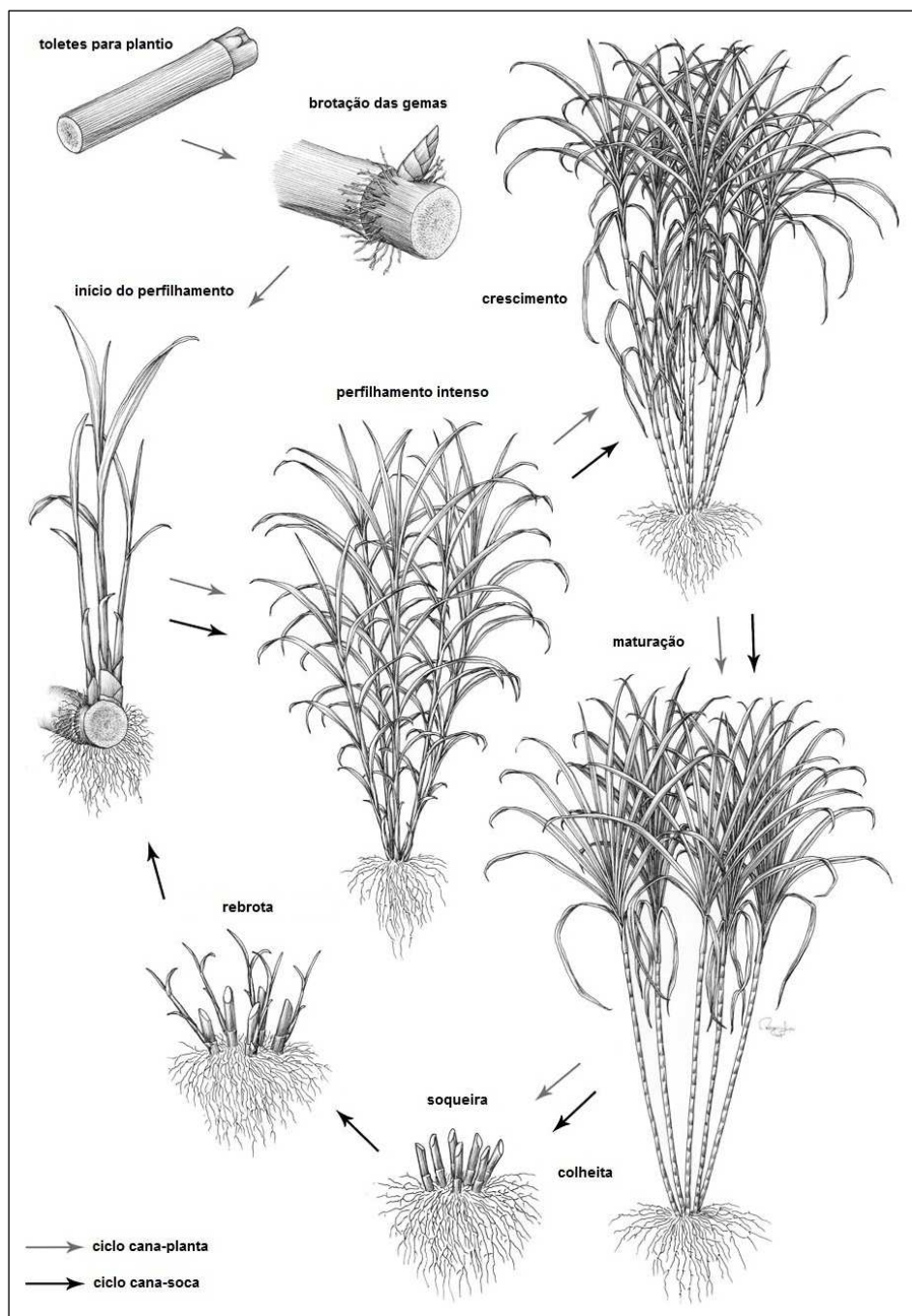


Figura 2.3: Ciclo fenológico da cultura da cana-de-açúcar.
Fonte: Adaptado de CHEAVEGATTI-GIANOTTO et al. (2011).

Na fase de brotação e emergência dos brotos, a propagação vegetativa da cana-de-açúcar é feita por toletes em que o broto rompe as folhas das gemas e se desenvolve em direção à superfície do solo, onde ocorre o enraizamento inicial e o aparecimento das primeiras folhas. A brotação começa de sete a 10 dias após o plantio e usualmente dura em torno de 30 a 35 dias. Esse é o estágio mais prematuro da cultura, no qual há predomínio do solo exposto.

Na fase de perfilhamento, a planta desenvolve-se de maneira abundante com a formação de outros rebentos nos toletes chamados de perfilhos. Esse processo é bastante influenciado principalmente pela luz solar, mas também é afetado por outros fatores como temperatura, umidade do solo, nutrientes, espaçamento, profundidade do plantio, doenças, pragas e época de plantio. Quando atinge o estabelecimento da cultura, o autosombreamento induz a inibição do perfilhamento e promove a aceleração do crescimento do colmo principal. A população máxima de perfilhos é alcançada ao redor de 90 a 120 dias após a brotação das gemas quando ocorre o fechamento do dossel vegetal, com quantidades entre 10 e 20 plantas por metro quadrado de solo, que formam a touceira da cana-de-açúcar e a população de colmos que será colhida. O auge do perfilhamento é quando ocorre a cobertura total do solo pela folhagem das plantas que é o momento em que cada touceira possui o máximo de perfilhos, sendo que após esse período cerca de 50% da população morre e a população final se estabelece com aproximadamente nove meses de idade. Segundo BENVENUTI (2005), um menor ou maior perfilhamento certamente influenciará a resposta espectral da cana-de-açúcar, pois está diretamente relacionado com a densidade do dossel e, conseqüentemente, à menor ou maior exposição do solo.

A fase de crescimento é a mais importante do cultivo, que começa a partir de 120 dias após o plantio e dura até 270 dias em cultivo de 12 meses, onde ocorre o alongamento dos colmos entre quatro a cinco entrenós por mês e o acúmulo de matéria seca, resultando na produção da cana-de-açúcar. A distribuição do crescimento da planta é variável e depende principalmente da época de plantio, que determinará em quais condições climáticas se dará o desenvolvimento da cultura. Em um período de estiagem, o alongamento dos colmos pode ser reduzido ou paralisado, resultando em menor volume final. O crescimento em altura continua até a ocorrência de alguma limitação no suprimento de água, ocorrência de baixas temperaturas ou ainda devido ao florescimento.

A fase final do desenvolvimento é a maturação que corresponde à senescência da planta, na qual, devido a restrições térmicas e hídricas, a taxa de crescimento é baixa e os produtos da fotossíntese são destinados principalmente ao armazenamento de sacarose nos colmos. Geralmente, a cana-de-açúcar requer de seis a oito meses com temperaturas elevadas, radiação solar intensa e precipitações regulares, durante o período entre a primavera e o verão, para que haja pleno crescimento vegetativo, seguidos de quatro a seis meses com estação seca e temperaturas mais baixas, durante o período entre o outono e o inverno, que é quando ocorre a maior atividade de maturação e a menor atividade de crescimento, com intenso acúmulo de sacarose. Dessa forma, a planta estará fisiologicamente madura, ou seja, em ponto de colheita.

Do plantio até a primeira colheita, a cana-planta pode ter um ciclo de desenvolvimento com duração de 18 meses (cana de ano-e-meio) ou 12 meses (cana de ano), dependendo da época e da região de plantio. Após a primeira colheita, a cana-soca passa a ter um ciclo de 12 meses, onde a cada corte ocorre a rebrota da soqueira e o início de um novo ciclo. Com os sucessivos cortes, ocorre uma perda natural e gradativa na produtividade agrícola. Os cortes da cana-soca se repetem em intervalos anuais até que a lavoura não seja mais rentável economicamente, quando é feito o manejo de reforma do canavial e o ciclo reinicia-se com o plantio de novas mudas (ANJOS e FIGUEIREDO, 2010). A Figura 2.4 esboça o ciclo da cultura de cana-de-açúcar com cinco cortes.

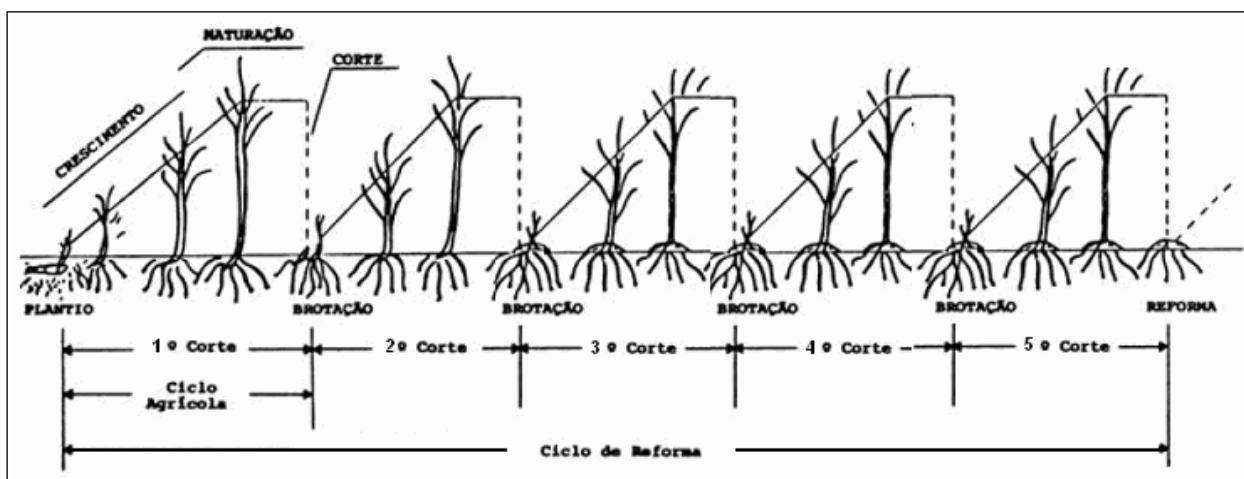


Figura 2.4: Ciclo da cultura de cana-de-açúcar com cinco cortes.
Fonte: Modificado de RODRIGUES (1995).

No Estado de São Paulo, predominam os ciclos de produção da cana de ano-e-meio e de ano, com a colheita estendendo-se de abril a dezembro. A cana de ano-e-meio, plantada de janeiro a maio, apresenta taxa de crescimento mínima de maio a setembro em função das condições pouco favoráveis do inverno, como pequena disponibilidade hídrica no solo, baixas temperaturas e menores intensidades de radiação. Já o período de maior crescimento acontece de outubro a abril, com o início das precipitações, aumento da intensidade luminosa e da temperatura, com o pico vegetativo ocorrendo em março. A cana de ano, plantada de setembro a dezembro, tem o maior período de crescimento de novembro a abril, também com pico vegetativo ocorrendo em março, diminuindo posteriormente devido às condições climáticas adversas do período de inverno, podendo ser colhida a partir de julho (RODRIGUES, 1995).

Como qualquer fenômeno natural, o ciclo da cana-de-açúcar está relacionado com as condições climáticas, variando de acordo com as fases de desenvolvimento. Para lavouras cultivadas em sequeiro, como é o caso no Estado de São Paulo, é feita uma descrição geral dos principais fatores hídricos, térmicos e de radiação solar que influenciam na resposta espectral do dossel da cana-de-açúcar, com base em CASAGRANDE e VASCONCELOS (2010).

Não só a quantidade de precipitação é importante, mas também a distribuição ao longo da safra. O excesso de chuvas desde o plantio até o alongamento dos colmos, que determinará a formação do dossel da cana-de-açúcar, pode reduzir as taxas de crescimento e diminuir a radiação disponível para a planta, afetando a produção final da cultura. Já o excesso de chuvas durante a fase de maturação pode acarretar no atraso da colheita. As deficiências hídricas nos primeiros oito meses do ciclo da cultura provocam as maiores reduções de produtividade, enquanto que na fase de maturação a influência é muito pequena para o acúmulo de sacarose nos colmos.

A cana-de-açúcar possui uma faixa de temperatura do ar ideal entre 25°C e 35°C, sendo que acima ou abaixo disso a brotação, o perfilhamento e o crescimento são quase nulos. Na fase de maturação, a temperatura média do ar necessita ser menor que 20°C. No período do frio, o desenvolvimento vegetativo é estacionado e a planta passa a preparar mais sacarose que será acumulada como substância de reserva, elevando seus teores no colmo.

Como a cana-de-açúcar é uma planta C4, cujo primeiro composto orgânico fixado da fotossíntese é uma molécula de quatro carbonos, as altas eficiências fotossintéticas devem-se às altas intensidades luminosas. Com elevadas taxas de radiação, os colmos são mais grossos e

curtos, as folhas mais longas e mais verdes e o perfilhamento mais intenso. Em condições de baixas irradiâncias os colmos são mais finos e longos, as folhas estreitas e amareladas.

A cana-de-açúcar mostra-se uma cultura bastante adaptável sendo cultivada em solos de fertilidade e textura bastante distintas, desde arenosas até muito argilosas. Os cultivares da cana-de-açúcar definem a capacidade de brotação de soqueiras, perfilhamento, altura, número de colmos por planta, diâmetro do colmo, comprimento e largura das folhas, arquitetura da parte aérea, resistência a pragas e doenças e produtividade. Além dessas características, existem cultivares de ciclo precoce, médio e tardio, de acordo com a duração da fase de maturação, fazendo com que o corte coincida com a máxima produção de matéria seca para que a colheita seja realizada durante um longo período, que se estende de abril, para o ciclo precoce, a dezembro, para o ciclo tardio, o que permite manter a qualidade da matéria prima entregue à indústria durante todo o período da safra (LANDELL e BRESSIANI, 2010).

2.2. Estimativas de safras agrícolas

Em função da grandeza dos números que envolvem a produção canavieira no Brasil, principalmente as extensas áreas cultivadas, percebe-se a importância do planejamento da produção da cultura, apoiado nas estimativas de safra.

A disponibilidade de informações confiáveis para a previsão de safras agrícolas obtidas com a maior antecedência possível, em relação à colheita, é de importância estratégica para os setores público e privado envolvidos no agronegócio, pois influenciam diretamente o planejamento do transporte, mão de obra, armazenamento e custo de comercialização da produção, tanto no mercado interno como nas exportações (KRÜGER et al., 2007).

Como forma de orientar a expansão sustentável da produção de cana-de-açúcar no território brasileiro, o Zoneamento Agroecológico (ZAE) da cana-de-açúcar fornece subsídios técnicos ao Governo Federal para formulação de políticas públicas na concessão de crédito agrícola e pagamento do seguro rural. A elaboração do ZAE da cana-de-açúcar foi baseada na integração de dados espaciais de solo e clima relacionados com os requerimentos da cultura, considerando também aspectos ambientais, econômicos e sociais para estabelecer regiões de alta, média e baixa aptidão e épocas favoráveis à produção agrícola rentável e com menor risco para o cultivo da cana-de-açúcar, com declividade inferior a 12% que é uma condição para a colheita

mecanizada. As áreas indicadas para a expansão pelo ZAE compreendem as de produção agrícola, lavouras perenes, anuais e pastagens. Nos estados da região Centro-Sul foram excluídas as áreas cultivadas com cana-de-açúcar na safra 2007/2008, utilizando-se o mapeamento do Canasat (MANZATTO et al., 2009). A Figura 2.5 representa o ZAE da cana-de-açúcar para o Estado de São Paulo.

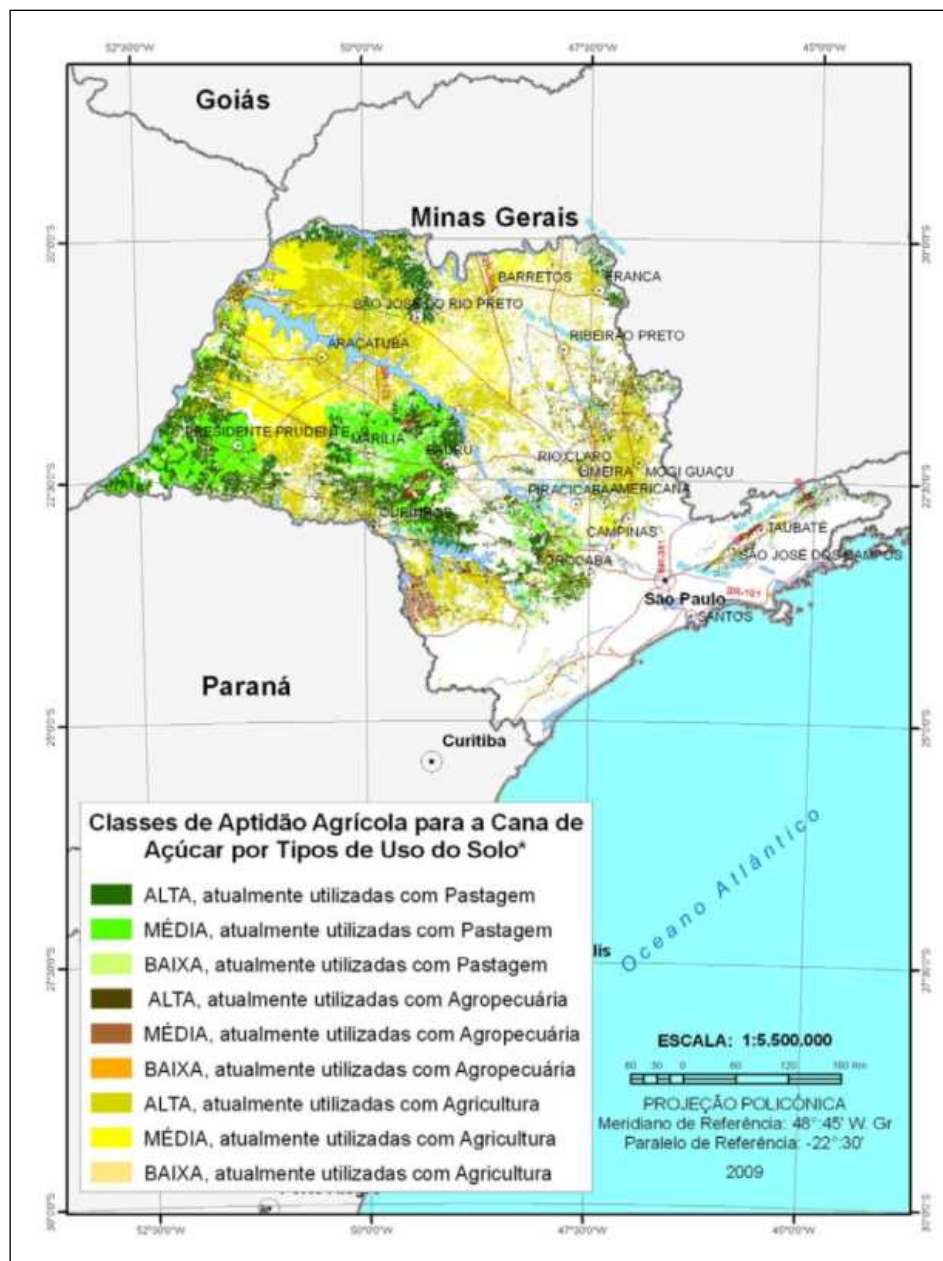


Figura 2.5: Zoneamento Agroecológico (ZAE) da cana-de-açúcar para o Estado de São Paulo.
 Fonte: MANZATTO et al. (2009).

As classes de uso da terra consideradas no ZAE são agricultura (Ac), agropecuária (Ag) e pastagem (Ap) conforme denominação do Sistema Brasileiro de Classificação de Vegetação do Brasil. Os dados de uso da terra são oriundos do Projeto de Conservação e Utilização Sustentável da Diversidade Biológica Brasileira (PROBIO), coordenado pelo Ministério do Meio Ambiente (MMA), que realizou o levantamento da cobertura vegetal nativa do Brasil, utilizando imagens de satélite Landsat do ano de 2002. O mapeamento da cobertura vegetal do bioma Mata Atlântica obteve exatidão global de 86,4% e do bioma Cerrado de 96,5%, ambos verificados por trabalhos de campo para validar as diferentes classes de cobertura da terra (ARAÚJO et al., 2007; SANO et al., 2007).

Nesse sentido, a obtenção de estimativas confiáveis de área plantada de cana-de-açúcar é uma necessidade dos atuais modelos de previsão de safras e, também, é um dos pontos críticos no zoneamento agrícola para a quantificação de áreas ocupadas em regiões tradicionais e de expansão da cultura.

Em geral, a estimativa de áreas de culturas agrícolas pode ser obtida por enquetes junto ao meio rural em conjunto com levantamentos de campo, que normalmente geram um grande volume de dados, porém, devido à extensão territorial do Brasil é um processo trabalhoso, demorado, oneroso e subjetivo, além de não permitir a detecção de variações sutis que ocorrem ao longo da safra (FAO, 1996; PINO, 2001).

O Instituto Brasileiro de Geografia e Estatística (IBGE) é o órgão oficial do governo que realiza o Levantamento Sistemático da Produção Agrícola (LSPA) da cana-de-açúcar em todos os municípios do Brasil, levando em consideração informações provenientes de agentes ligados ao setor agrícola, como produtores rurais, cooperativas, associações de classe e comerciantes de insumos. O LSPA é um censo da safra agrícola com periodicidade mensal que fornece, dentre outras estatísticas, as estimativas de área na fase de intenção de plantio e no final da colheita, fazendo a consolidação das informações em cada município, que são posteriormente agrupadas em todo território nacional (IBGE, 2002).

A Companhia Nacional de Abastecimento (CONAB) é outro órgão oficial do governo que também faz o acompanhamento da safra da cana-de-açúcar por meio de pesquisas de dados estimativos de área cultivada, entre outros, junto às unidades de produção sucroalcooleira nos vários estados produtores em cada região do país. Os levantamentos são realizados com

periodicidade quadrimestral durante a safra e divulgados em forma de boletins de monitoramento agrícola que descrevem as perspectivas agronômicas dos canaviais e as influências climáticas na produção (CONAB, 2014).

A União da Indústria de Cana-de-Açúcar (UNICA) é uma entidade de classe, mantida pela iniciativa privada, representativa do setor de açúcar e etanol do Brasil que divulga os dados da safra levantados junto às usinas de cana-de-açúcar, consolidados por estados e regiões (UNICA, 2014), em parceria com o Centro de Tecnologia Canavieira (CTC) e o Centro de Estudos Avançados em Economia Aplicada (CEPEA).

Esses métodos subjetivos provêm informações detalhadas sobre as áreas de cultivo da cana-de-açúcar, que além de apresentarem discrepâncias, não permitem a quantificação dos erros envolvidos nos levantamentos para análise do nível de precisão das estimativas de área, nem o conhecimento da distribuição espacial. IPPOLITI-RAMILO et al. (2003) destacam que o caráter multiespectral, sinóptico, repetitivo e global do sensoriamento remoto, em conjunto com as técnicas de geoprocessamento, têm grande potencial para auxiliar os sistemas de estimativas de safras agrícolas, permitindo a análise temporal e o mapeamento da distribuição espacial de áreas plantadas, de forma mais rápida e objetiva. Desde 2004, por meio do projeto GeoSafras, a CONAB tem tomado a iniciativa de utilizar geotecnologias para aprimorar as estatísticas de safras das principais culturas agrícolas do país e, assim, vem aperfeiçoando a metodologia do sistema de previsão de safras no Brasil (CONAB, 2010).

2.3. O sensoriamento remoto aplicado ao monitoramento agrícola

COLWELL (1983) define o sensoriamento remoto como a ciência que permite a aquisição de dados tomados à distância sobre alvos da superfície terrestre, por meio da utilização de sensores sensíveis à radiação eletromagnética, sem que exista contato físico entre eles. O Sol é a principal fonte de energia radiante e incide sobre a Terra em diversos comprimentos de onda do espectro eletromagnético.

Ao se propagar pelo espaço, um fluxo de radiação eletromagnética (I) interage com todos os alvos que compõem a superfície terrestre, sendo por estes refletido (R), absorvido (A) e transmitido (T), como no caso de uma folha, exemplificado na Figura 2.6.

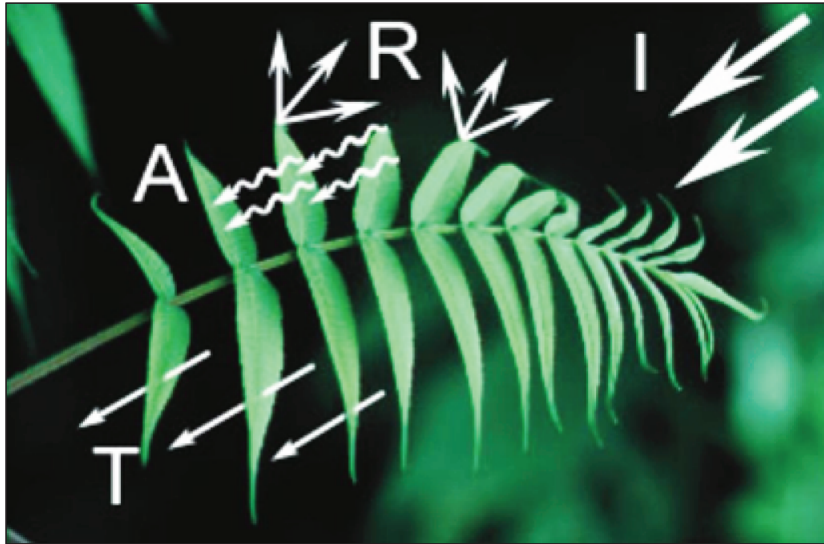


Figura 2.6: Interação da radiação eletromagnética solar com uma folha.
Fonte: MOREIRA (2003).

Essa interação ocorre de maneira diferente para cada comprimento de onda e depende das propriedades físico-químicas e biológicas de cada alvo da superfície por serem fenômenos naturais. A soma das componentes 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 partículas (gases e aerossóis) presentes na atmosfera afetam a radiação solar incidente em certos comprimentos de onda, bem como as condições das plantas, geométrica de iluminação e influência do solo (COLWELL, 1983).

Os sensores remotos orbitais são desenvolvidos para medir a energia eletromagnética que é refletida ou emitida pela superfície terrestre, conforme ilustrado pela Figura 2.7. A característica multiespectral desses sensores permite registrar essa energia em determinadas faixas de comprimentos de onda do espectro eletromagnético que são as bandas espectrais, transformando-a numa imagem digital, passível de interpretação visual (JENSEN, 2009).

O sensoriamento remoto tem se mostrado uma ferramenta valiosa no monitoramento agrícola devido à sua visão sinóptica e repetitiva da superfície terrestre, sabendo-se que os alvos apresentam comportamento espectral distinto ao longo do espectro eletromagnético. A cana-de-açúcar possui características favoráveis ao uso de sensoriamento remoto, pois se trata de uma cultura que ocupa extensas áreas e que apresenta um dossel vegetal regular.

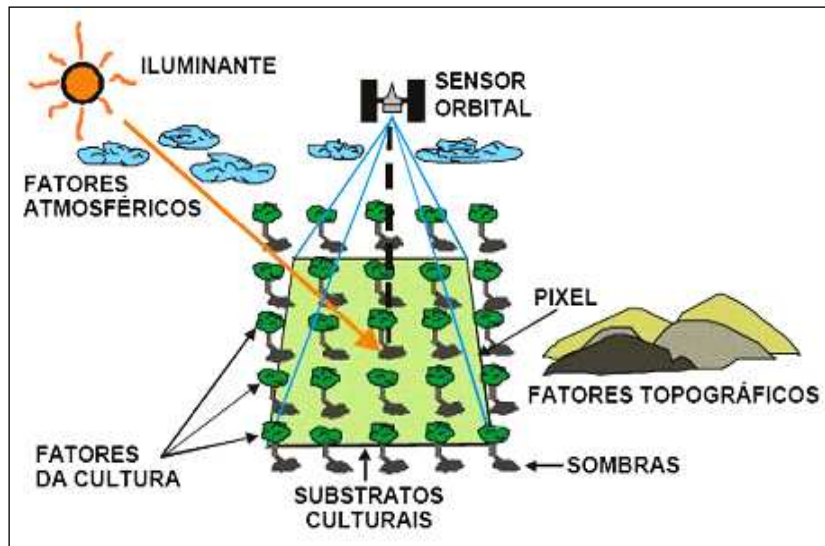


Figura 2.7: Interação da radiação eletromagnética fonte-alvo-sensor.

No Brasil, o monitoramento operacional da safra de cana-de-açúcar envolvendo o uso de sensoriamento remoto é realizado pelo projeto Canasat, executado pelo Instituto Nacional de Pesquisas Espaciais (INPE) com apoio da UNICA. A metodologia empregada no Canasat baseia-se em técnicas de segmentação e interpretação visual de imagens tipo Landsat, conduzidas por uma equipe de intérpretes, para mapear a distribuição espacial da área cultivada de cana-de-açúcar identificada como soca, expansão, em reforma e reformada (RUDORFF et al., 2010) e o tipo de colheita com ou sem a queima da palha da cana-de-açúcar (AGUIAR et al., 2011), em nível de talhão. Os resultados anuais são apresentados por municípios e consolidados pelos estados produtores em todo o território nacional, sendo publicados na forma de mapas, tabelas e gráficos no sítio do projeto na Internet (INPE, 2014).

Diversas pesquisas vêm mostrando o potencial de aplicação do sensoriamento remoto no estudo da cultura da cana-de-açúcar. VICENTE et al. (2012) fizeram o mapeamento e a quantificação de áreas de cana-de-açúcar no Estado de São Paulo utilizando imagens de satélite e obtiveram boa concordância com os dados oficiais do IBGE. XAVIER et al. (2006) procederam a classificação automática de imagens de satélite e conseguiram uma boa discriminação de áreas canavieiras em relação a outros alvos e à distinção dos diferentes ciclos de produção. MACHADO e LAMPARELLI (2007) e SIMÕES et al. (2009) demonstraram a existência de correlação significativa entre dados espectrais e variáveis biofísicas da cana-de-açúcar, mostrando o potencial de utilização na estimativa de produtividade da cultura. PICOLI et al.

(2012) e FERNANDES et al. (2011) mostraram ser promissor o uso de dados espectrais para identificar e discriminar diferentes fases fenológicas da cana-de-açúcar. IMMERZEEL et al. (2008) fizeram a integração de dados espectrais na calibração de um modelo hidrológico para irrigação da cana-de-açúcar em áreas propensas à seca. FORTES e DEMATTÊ (2006) mostraram ser possível fazer a identificação e discriminação de variedades de cana-de-açúcar a partir de dados espectrais, reduzindo o tempo de avaliação e o trabalho de verificação de campo. APAN et al. (2004) demonstraram que dados espectrais podem ser utilizados na detecção da doença da ferrugem amarela para proteção da cultura da cana-de-açúcar.

2.3.1. Comportamento espectral da vegetação

O monitoramento de uma cultura agrícola pode ser realizado por meio do estudo da radiação eletromagnética que interage com a vegetação e o solo, sendo que as folhas têm a maior influência no comportamento espectral dos dosséis vegetais. A curva de refletância típica da vegetação (verde e sadia) é distinta e variável com o comprimento de onda nas regiões do espectro eletromagnético do visível, do infravermelho próximo e do infravermelho médio, conforme mostrado na Figura 2.8.

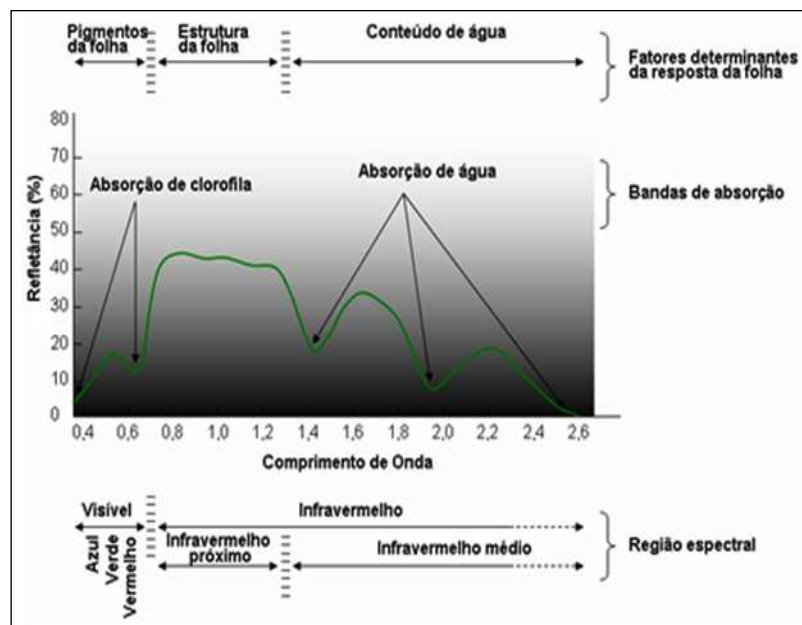


Figura 2.8: Curva padrão de refletância da vegetação.
Fonte: Adaptado de HOFFER (1978).

Nas faixas do azul ao vermelho do visível (0,4 μm a 0,7 μm), a maior parte da radiação é absorvida pelos pigmentos clorofilados presentes nas folhas, com máxima refletância na faixa do verde (0,54 μm), não ultrapassando 15%. Na faixa do infravermelho próximo (0,7 μm a 1,3 μm), a vegetação é caracterizada pela alta refletância (45% a 50%), alta transmitância (45% a 50%) e baixa absorptância (menos de 5%), influenciada pela estrutura celular das plantas. Na faixa do infravermelho médio (1,3 μm a 2,6 μm), as propriedades ópticas das plantas são afetadas por intervalos de alta absorção da radiação em função da presença de água nos comprimentos de onda de 1,4 μm , 1,9 μm e 2,6 μm (HOFFER, 1978).

A energia refletida por uma área cultivada é uma integração dos valores de refletância da planta e do solo. A cobertura vegetal do solo coberto por plantas encontra-se diretamente relacionado com a refletância da vegetação nas faixas do visível e infravermelho próximo. Quando a cobertura do solo é completa, as folhas são elementos dominantes das plantas na definição das propriedades espectrais de dosséis vegetais (PONZONI et al., 2012).

O comportamento espectral de culturas agrícolas caracteriza-se por apresentar fases fenológicas bem definidas durante o seu ciclo de desenvolvimento. A Figura 2.9 representa a evolução da refletância do dossel vegetal de uma cultura agrícola entre fases fenológicas do ciclo de desenvolvimento.

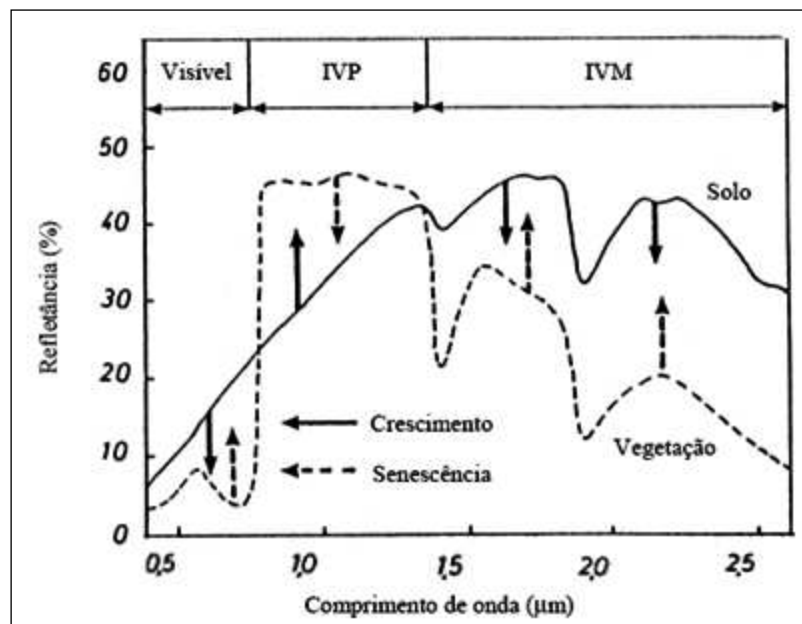


Figura 2.9: Evolução da refletância de um dossel vegetal agrícola entre fases fenológicas.
Fonte: ESQUERDO (2007).

De modo geral, observa-se que a planta, ao passar da fase de crescimento para a fase de senescência, aumenta a refletância na faixa do visível e diminui na faixa do infravermelho próximo devido a alterações nos mecanismos biológicos, se assemelhando assim a resposta espectral do solo. A energia captada pelo satélite proveniente de uma área agrícola é influenciada pelo solo devido às características da cultura, como espaçamento, porte, densidade da cobertura vegetal, vigor e fase fenológica. No início do desenvolvimento, o solo exposto prevalece entre as plantas. Com o crescimento da cultura, a cobertura vegetal passa a predominar, diminuindo assim a influência do solo. Em estudos de culturas agrícolas utilizando o sensoriamento remoto tem sido comum utilizar índices de vegetação para desenvolver relações funcionais dos dados com as características da vegetação (PONZONI et al., 2012).

2.3.2. Índices de vegetação

Os índices de vegetação sintetizam a informação contida em imagens multiespectrais, resultado de transformações lineares do fator de refletância envolvendo soma, diferença e razão entre as bandas espectrais, com a finalidade de ressaltar a resposta espectral da vegetação em função da cobertura do dossel sobre o solo (WIEGAND et al., 1991).

Um índice de vegetação ideal deve ser capaz de responder às pequenas variações das fases fenológicas da vegetação, amenizar a influência das condições e tipos de solo, da geometria de iluminação da cena e das condições atmosféricas (JACKSON e HUETE, 1991).

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 vegetal tem estreita relação com a energia solar absorvida, o que define o crescimento e o desenvolvimento de uma cultura.

Um dos mais conhecidos e utilizados índices de vegetação é o *Normalized Difference Vegetation Index* (NDVI), proposto por ROUSE et al. (1973) para atenuar a interferência do solo na resposta espectral da vegetação e diminuir a influência da atmosfera e do ângulo de elevação solar em imagens temporais. O NDVI explora o contraste entre as bandas do vermelho e do infravermelho próximo e é calculado pela Equação 2.1, variando 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} = fator de refletância na banda do vermelho;

ρ_{IVP} = fator de refletância na banda do infravermelho próximo.

Como o NDVI tem a possibilidade de saturação, tornando-se insensível ao aumento da biomassa vegetal, mesmo quando há aumento da densidade do dossel vegetal, HUETE et al. (1994) desenvolveram o índice de vegetação *Enhanced Vegetation Index* (EVI) para otimizar a resposta espectral da vegetação e melhorar a sensibilidade em regiões de maior biomassa vegetal, propiciando o monitoramento da vegetação através da redução dos efeitos de substrato do dossel vegetal e de influências atmosféricas. O EVI representa melhor a dinâmica da vegetação ao longo de uma série temporal de imagens e é calculado pela Equação 2.2, também variando de -1,0 a +1,0:

$$\text{EVI} = G \times \left[\frac{(\rho_{\text{IVP}} - \rho_{\text{VER}})}{(\rho_{\text{IVP}} + C_1 \times \rho_{\text{VER}} - C_2 \times \rho_{\text{AZU}} + L)} \right] \quad (2.2)$$

onde,

ρ_{AZU} = fator de refletância na banda do azul;

ρ_{VER} = fator de refletância na banda do vermelho;

ρ_{IVP} = fator de refletância na banda do infravermelho próximo;

$G = 2,59$: fator de ganho;

$L = 1$: fator de ajuste do substrato do dossel;

$C_1 = 6,0$ e $C_2 = 7,5$: coeficientes da resistência de aerossóis que utiliza a banda do azul para atenuar a influência dos aerossóis na banda do vermelho.

A Figura 2.10 exemplifica a dinâmica do NDVI de um dossel vegetal agrícola, revelando a passagem entre as fases fenológicas de crescimento e senescência, durante o ciclo de desenvolvimento.

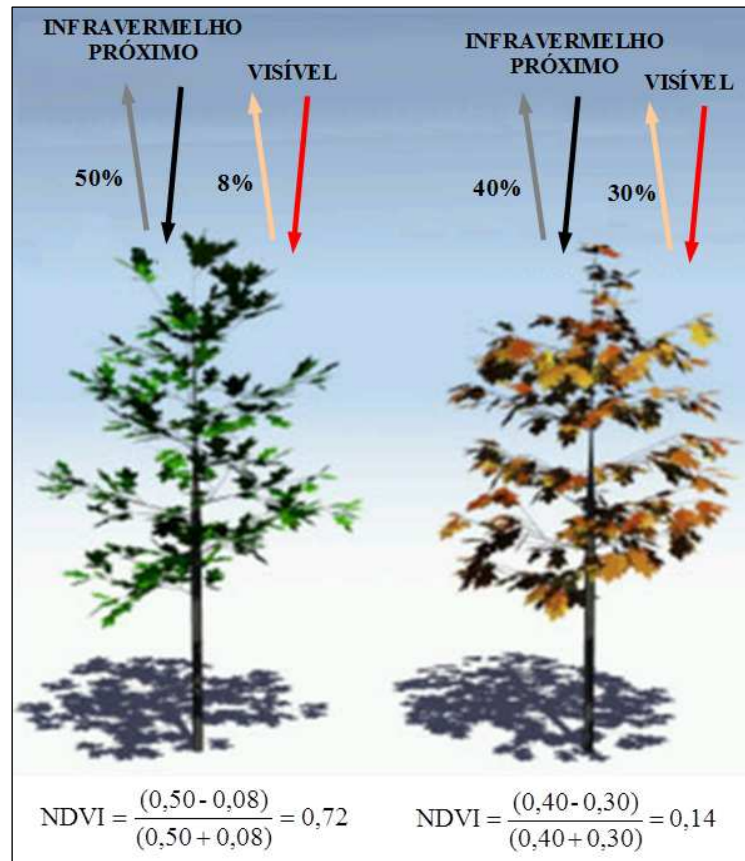


Figura 2.10: Dinâmica do NDVI de um dossel vegetal agrícola entre as fases fenológicas de crescimento e senescência.

Fonte: Adaptado de NASA/EOS (2014).

Na fase de crescimento (à esquerda), a refletância da vegetação é baixa na faixa do visível e alta na faixa do infravermelho próximo. Na fase de senescência (à direita), a refletância da vegetação aumenta na faixa do visível e diminui na faixa do infravermelho próximo. Dessa forma, os valores mais altos de NDVI estão associados com maior vigor e densidade da vegetação e os valores mais baixos a redução da atividade fotossintética.

2.3.3. Imagens do sensor MODIS

O sensor *MODerate Resolution Imaging Spectroradiometer* (MODIS) é o principal instrumento a bordo das plataformas orbitais do programa internacional *Earth Observing System* (EOS), liderado pela *National Aeronautics and Space Administration* (NASA) e formado por diversos países, para estudo dos processos de mudanças globais. O primeiro satélite do programa

EOS, chamado TERRA, foi lançado em 18 de dezembro de 1999 e tem passagem pelo Equador às 10h30 (horário local), em órbita descendente. O segundo, chamado AQUA, foi lançado em 4 de maio de 2002 e cruza o Equador às 13h30 (horário local), em órbita ascendente (SOARES et al., 2007).

O MODIS é um sensor multiespectral de sensibilidade radiométrica de 12 bits, ou seja, 4.096 níveis de cinza, com 36 bandas posicionadas nas faixas do visível, infravermelho próximo, infravermelho médio e infravermelho termal do espectro eletromagnético. Duas bandas são adquiridas na resolução espacial de 250 m, outras cinco bandas em 500 m e as demais 29 bandas em 1.000 m, conforme as especificações apresentadas na Tabela 2.1. A varredura de 55° para cada lado na órbita quase-polar e heliossíncrona de 705 km de altitude resulta numa faixa imageada aproximada de 2.330 km x 5.000 km, com período de 98,9 min., que recobre quase todo o globo terrestre diariamente e, a cada 16 dias, completa um ciclo de imageamento (NASA, 2014).

Tabela 2.1: Especificações das bandas espectrais do sensor MODIS.

Bandas Espectrais	Resolução Espectral (µm)	Resolução Espacial (m)	Bandas Espectrais	Resolução Espectral (µm)	Resolução Espacial (m)
1	0,620 - 0,670	250	19	0,915 - 0,965	1.000
2	0,841 - 0,876		20	3,660 - 3,840	
3	0,459 - 0,479	500	21	3,929 - 3,989	
4	0,545 - 0,565		22	3,929 - 3,989	
5	1,230 - 1,250		23	4,020 - 4,080	
6	1,628 - 1,652		24	4,433 - 4,498	
7	2,105 - 2,155		25	4,482 - 4,549	
8	0,405 - 0,420	1.000	26	1,360 - 1,390	
9	0,438 - 0,448		27	6,535 - 6,895	
10	0,483 - 0,493		28	7,175 - 7,475	
11	0,526 - 0,536		29	8,400 - 8,700	
12	0,546 - 0,556		30	9,580 - 9,880	
13	0,662 - 0,672		31	10,780 - 11,280	
14	0,673 - 0,683		32	11,770 - 12,270	
15	0,743 - 0,753		33	13,185 - 13,485	
16	0,862 - 0,877		34	13,485 - 13,785	
17	0,890 - 0,920		35	13,785 - 14,085	
18	0,931 - 0,941		36	14,085 - 14,385	

Fonte: Adaptado de NASA (2014).

Diversas inovações são apresentadas pelo sensor MODIS, desde a melhoria substancial na qualidade dos detectores, no sistema de imageamento, no aumento do número de bandas espectrais e na evolução dos algoritmos de processamento de dados para a geração de produtos de alta qualidade com correções atmosférica, radiométrica e geométrica. Contudo, o grande avanço na área de sensoriamento remoto se deve ao conceito de distribuição gratuita dos produtos, desde fevereiro de 2000 até o momento, o que vem viabilizando um grande número de aplicações multidisciplinares para monitorar as mudanças na superfície terrestre.

Os dados do sensor MODIS são divididos em três classes de produtos: terrestres, oceânicos e atmosféricos. A classe terrestre denominada MODLAND contém o produto MOD13 referente ao NDVI e EVI, produzidos globalmente na resolução espacial 250 m, 500 m e 1.000 m, em imagens compostas de 16 dias. Essas composições consistem dos pixels de alta qualidade radiométrica, boa geometria de observação, mínima presença de nuvens e aerossóis, ou seja, a melhor informação espectral dentro do período de tempo, o que favorece a obtenção de imagens com menores influências atmosféricas e ruídos (LATORRE et al., 2007).

Na prática, a seleção dos melhores pixels em regiões tropicais e subtropicais que têm a presença constante de nuvens, como é o caso do Brasil, é equivalente a aplicação da técnica da *Maximum Value Compositie* (MVC). O valor máximo de um determinado pixel numa sequência de imagens diárias é utilizado para a construção da imagem composição num dado período, que no caso do MODIS seria uma imagem composta de 16 dias. Como a presença de nuvens torna o valor dos índices de vegetação negativo, os pixels contaminados têm menor probabilidade de entrar na composição (HOLBEN, 1986).

Os produtos MODIS também vêm acompanhados da informação de confiabilidade do pixel (*pixel reliability*), o que permite ao usuário fazer uma análise da qualidade de cada pixel da composição (LATORRE et al., 2007).

Os 44 produtos MODIS são obtidos gratuitamente do *Land Processes Distributed Active Archive Center* (LP DAAC), pertencente ao EOS da NASA, em forma de quadrantes (*tiles*) que cobrem uma área geográfica de 10° x 10° de latitude/longitude (1.200 km por 1.200 km), no formato HDF (*Hierarchical Data Format*) e na projeção cartográfica sinusoidal. O sistema de coordenadas começa no quadrante (0, 0) (horizontal, vertical) no canto superior esquerdo e termina no (35, 17) no canto inferior direito. A Figura 2.11 representa os 17 quadrantes que cobrem o território brasileiro, com destaque para o h13v11.

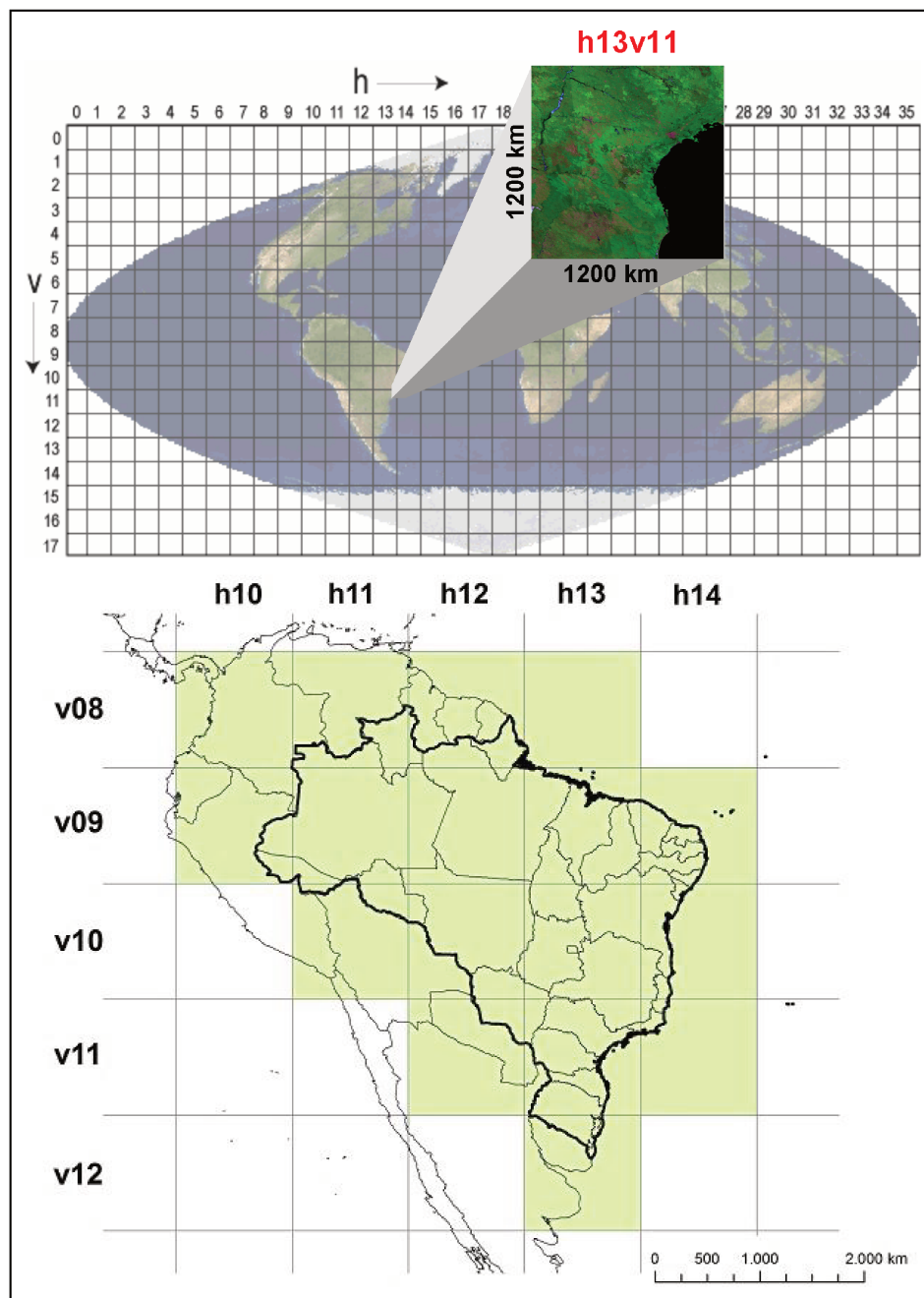


Figura 2.11: Representação dos 17 quadrantes MODIS que cobrem o Brasil.
 Fonte: Adaptado de NASA (2014).

A Figura 2.12 apresenta os quadrantes h13v10 e h13v11 que cobrem o Estado de São Paulo. A manipulação dos dados para seleção do produto, mosaico, recorte da área de interesse e definição da projeção cartográfica, é feita com o programa gratuito MODIS *Reprojection Tools* (MRT) (LP DAAC, 2014).

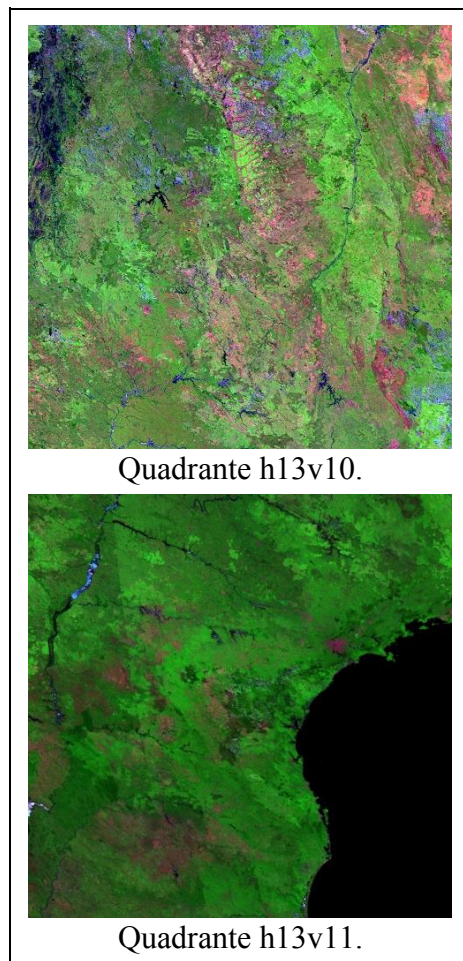


Figura 2.12: Quadrantes MODIS que cobrem o Estado de São Paulo.

2.3.4. Análise de séries temporais de imagens

Uma série temporal pode ser entendida como um conjunto de observações ordenadas sequencialmente no tempo e registradas em períodos regulares. A suposição básica é a existência de um mecanismo gerador da série temporal dependente do tempo, que tem relação com os dados passados e futuros, criando padrões que podem ser detectados.

A análise de séries temporais consiste em modelar os padrões de repetição das observações e fazer previsões futuras, a partir da decomposição nas componentes tendência (comportamento de longo prazo), ciclo (oscilações que se repetem de forma periódica) e sazonalidade (oscilações que ocorrem em intervalos regulares). A modelagem dos padrões pode ser feita no domínio do tempo por meio de modelos paramétricos ou no domínio da frequência utilizando modelos não paramétricos (MORETTIN e TOLOI, 2006).

Uma das principais vantagens que as imagens do sensor MODIS apresentam é a alta repetitividade de extensas áreas da superfície terrestre, aliado a boa qualidade radiométrica e excelente precisão geométrica. Essas são características fundamentais para o estudo de alvos com grande dinâmica espectral e temporal e, dessa forma, podem ser empregadas no monitoramento agrícola ao longo das fases do ciclo de desenvolvimento de uma cultura, desde o plantio ou rebrota até a colheita, em nível regional e local.

Nesse sentido, a abordagem espectro-temporal tem sido apontada como sendo a mais apropriada para monitorar e mapear culturas agrícolas por meio da análise do comportamento dos pixels das imagens ao longo de uma série temporal, com ênfase nas variações espectrais. Dessa forma, é possível identificar a dinâmica da vegetação em distintas escalas temporais, acompanhando as variações dos ciclos fenológicos e caracterizando a transição entre diferentes usos e coberturas do solo ao longo dos anos (JÖNSSON e EKLUNDH, 2002; WARDLOW et al., 2007; VERBESSELT et al., 2010; BROWN et al., 2013).

Das imagens sequenciais dos índices de vegetação podem ser gerados os perfis temporais, que descrevem o comportamento espectral característico de alvos localizados na superfície terrestre ao longo do tempo. A Figura 2.13 apresenta, como exemplo, perfis temporais que foram gerados a partir de composições de imagens NDVI decendiais suavizadas, obtidas do sensor SPOT Vegetation de resolução espacial de 1 km, em áreas com o conhecimento prévio de distintos alvos (FERNANDES, 2009).

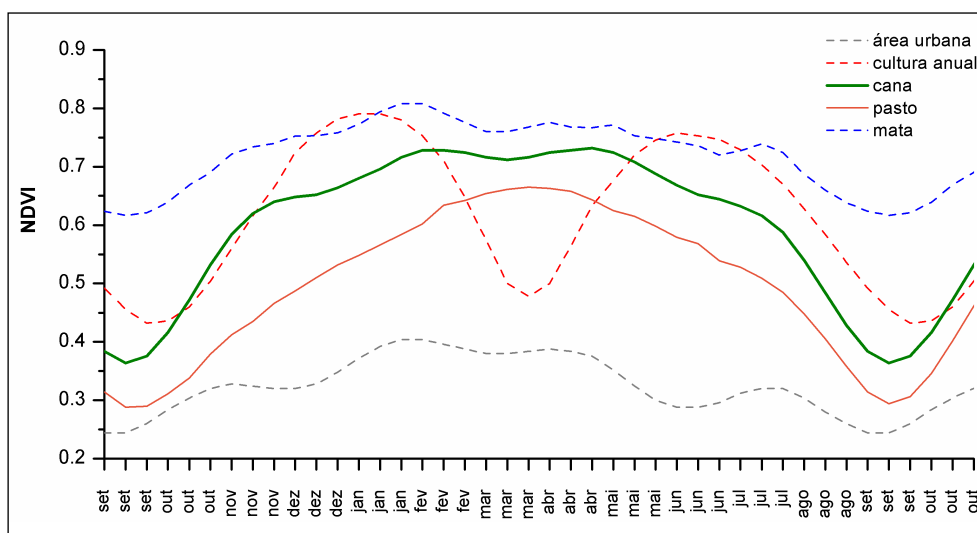


Figura 2.13: Perfis temporais de NDVI apresentados por diferentes alvos. Fonte: FERNANDES (2009).

A área urbana apresenta baixos valores do NDVI durante todo o período por causa da pouca presença de vegetação. A mata apresenta valores elevados do NDVI durante todo o período, o que está relacionado a um vigor da vegetação constantemente alto, com pequenas variações entre as estações seca e chuvosa. O pasto possui um comportamento espectro-temporal muito similar ao da cana-de-açúcar, com os valores do NDVI decrescendo na estação seca devido à diminuição do vigor da vegetação e aumentando durante a estação chuvosa.

No caso de culturas agrícolas, os perfis temporais descrevem a variação do vigor vegetativo ao longo do seu ciclo de desenvolvimento, onde são observados baixos valores do NDVI no início do ciclo e no final da safra, enquanto que os altos valores ocorrem na época de grande crescimento vegetativo com alta biomassa vegetal. A cultura anual apresenta dois picos vegetativos relacionados ao desenvolvimento de diferentes cultivos durante o período. Já a cana-de-açúcar apresenta apenas um período de máximo crescimento na série temporal. A partir dessa caracterização, pode-se acompanhar tanto o comportamento de uma cultura agrícola ao longo do seu ciclo fenológico para verificar mudanças sazonais como, também, comparar as condições atuais com as condições apresentadas em safras passadas.

Entretanto, uma série temporal oriunda das imagens do sensor MODIS pode estar sujeita à interferência de ruídos relacionados principalmente à presença de nuvens, condições de iluminação e geometria de visada, mesmo em composições dos pixels com a melhor informação espectral dentro de um período de tempo (HIRD e MCDERMID, 2009).

O comportamento temporal de um índice de vegetação tende a ser equilibrado e de baixa frequência. As oscilações de alta frequência são tratadas como ruídos que podem ser reduzidos com a aplicação de técnicas matemáticas de análise de séries temporais.

A análise de Fourier, também conhecida como análise harmônica, recebeu este nome em homenagem a Jean-Baptiste Joseph Fourier, matemático e físico francês que em 1822 iniciou a investigação sobre a decomposição de funções periódicas em séries trigonométricas convergentes e a sua aplicação aos problemas de propagação do calor, sendo muito úteis no processamento de sinais. A Transformada de Fourier permite representar uma série temporal por uma soma de componentes (termos harmônicos) senoidais ou cossenoidais, no domínio da frequência (MORETTIN e TOLOI, 2006).

As *wavelets* são definidas como ondas pequenas (ondaletas) com propriedades que as tornam adequadas a servirem de base para decomposição de outras funções. A Transformada de Wavelet permite decompor uma série temporal em diferentes escalas para obtenção de informações no domínio da frequência e do tempo, onde cada escala é representada por uma frequência específica. O que as diferencia da Transformada de Fourier é que a Transformada de Wavelet além da frequência, também considera a escala (MORETTIN, 1999).

A aplicação da Transformada de Fourier e da Transformada de Wavelet, nas séries temporais de imagens, tem o propósito de eliminar as variações bruscas nos valores espectrais dos índices de vegetação, permitindo uma melhor análise baseada nas mudanças que acontecem na superfície terrestre, o que facilita o entendimento da dinâmica temporal. As curvas suavizadas obtidas por essas transformações matemáticas normalmente são diferenciadas para cada tipo de cultura agrícola, o que possibilita reconhecer um padrão que possa ser usado na classificação das imagens. Também, pela análise dessas curvas, é possível identificar os picos fenológicos que caracterizam cada um dos diferentes tipos de cultura em uma mesma área ao longo da safra (BROWN et al., 2007; ZHANG et al., 2008; FREITAS e SHIMABUKURO, 2008; MARTÍNEZ e GILABERT, 2009).

2.4. Inteligência artificial na classificação de imagens

A inteligência artificial é parte da ciência da computação voltada para o desenvolvimento de aplicações em sistemas computadorizados que emulam habilidades cognitivas do homem, imitando aspectos do pensamento humano. O aprendizado de máquina é uma subárea da inteligência artificial dedicado ao desenvolvimento de modelos, algoritmos e técnicas que permitam ao computador aprender por experiência para aperfeiçoar seu desempenho em alguma tarefa. O mapeamento da cobertura do solo por meio de técnicas de aprendizado de máquina para reconhecimento de padrões em imagens de satélites é uma das principais aplicações na área de sensoriamento remoto. Essas técnicas que permitem ao programa de computador aprender um determinado padrão automaticamente a partir de exemplos, têm por objetivo procurar, detectar e extrair estruturas associadas às regularidades ou propriedades relevantes que estão implícitas num conjunto de dados, por meio da análise de classes (HAYKIN, 2001).

A discriminação de culturas agrícolas em imagens de satélite não é uma tarefa trivial, tornando-se um importante desafio devido ao comportamento espectral similar de várias espécies vegetais e em fases fenológicas distintas, o que resulta em grande variabilidade dentro de uma mesma classe. Entretanto, a variação temporal das características espectrais da vegetação associada à repetitividade de aquisição dos sensores orbitais é uma possível solução para a diferenciação do padrão de cultivo de uma determinada cultura agrícola, quando se faz a análise no decorrer do seu ciclo de desenvolvimento.

A tarefa de classificação de imagens de sensoriamento remoto refere-se à associação de cada pixel da imagem a um rótulo, descrevendo um alvo da superfície terrestre. Esse processo de rotulação é feito pela similaridade do padrão do pixel, descrito pelas bandas espectrais, com o padrão determinado para uma classe específica, normalmente utilizando uma medida de distância no espaço de atributos. Existem duas abordagens básicas de treinamento. Na classificação supervisionada, o usuário identifica alguns pixels pertencentes às classes desejadas e o método localiza todos os demais pixels pertencentes àquelas classes, baseado em alguma regra estatística pré-estabelecida. Na classificação não supervisionada, 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).

Em reconhecimento de padrões, em geral, assume-se que as classes são mutuamente excludentes. Com isso, os classificadores tradicionais baseados na análise pixel a pixel podem gerar resultados imprecisos no caso de imagens geradas por sensores de moderada resolução espacial, como o MODIS, que rotulam um determinado pixel a uma única classe, desconsiderando a possibilidade de haver mistura espectral na extensa área de 62.500 m^2 ($250 \text{ m} \times 250 \text{ m}$) representada pelo pixel.

Esse problema tem maior importância quando as regiões de cultivo não apresentam homogeneidade, onde a medida radiométrica representa uma combinação das medidas radiométricas de todos os alvos presentes neste pixel. Dessa forma, a mistura espectral das classes pode ocorrer quando o tamanho do pixel é maior do que as feições terrestres. No monitoramento de culturas agrícolas, a vegetação pode estar em diferentes condições ao longo do seu ciclo de desenvolvimento que provocam uma variabilidade espectral dentro de cada classe. Essas limitações podem reduzir o nível de acurácia da classificação devido à pobre extração de informações da imagem.

Nesse sentido, novas abordagens têm sido introduzidas na classificação de imagens de sensoriamento remoto com o propósito de se obter um desempenho mais favorável a partir do uso da inteligência artificial. Dentre as técnicas de aprendizado de máquina, tem havido um interesse considerável no uso de redes neurais e lógica fuzzy para o reconhecimento de padrões em imagens de satélites (TSO e MATHER, 2001).

2.4.1. Noções de redes neurais artificiais

Uma rede neural artificial é definida como um processador massivo e paralelamente distribuído, constituído de unidades de processamento simples, que tem a capacidade para armazenar conhecimento experimental e torná-lo disponível para uso. Devido a essas características, as redes neurais artificiais têm grande potencial de uso em aplicações que envolvam modelagem de sistemas e reconhecimentos de padrões. Um neurônio artificial tem o funcionamento análogo ao processo biológico realizado no cérebro humano. Os neurônios são organizados e ligados em rede, de modo a constituírem camadas, permitindo grande interconectividade (BISHOP, 1995).

A Figura 2.14 ilustra um neurônio biológico. No sistema neural do cérebro humano, um neurônio é excitado quando recebe sinais de outros neurônios através dos dendritos. Se ativado, caso o estímulo recebido ultrapasse seu limiar de excitação, irá disparar um impulso para outros neurônios através dos axônios. A sinapse é o efeito que ocorre quando um axônio de um neurônio libera substâncias para estimular os receptores dos dendritos de outros neurônios, fazendo a transmissão de um impulso nervoso de um neurônio ao outro.

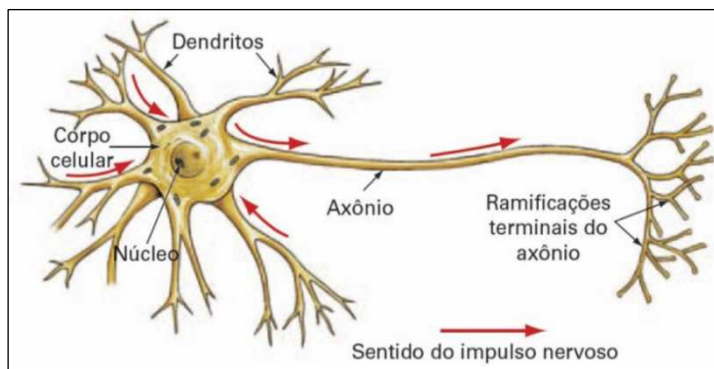


Figura 2.14: Neurônio biológico.
Fonte: Adaptado de WASSERMAN (1989).

A Figura 2.15 apresenta um neurônio artificial que é um modelo matemático análogo ao neurônio biológico, onde os elementos básicos são os nós de processamento. Os terminais de entrada correspondem aos dendritos que recebem os valores x e um terminal de saída y referente a um axônio. Para representar o comportamento das sinapses, as entradas dos neurônios têm pesos w . A representação do efeito de uma sinapse no neurônio artificial é dada pela ponderação $x * w$. Quando ocorre a entrada da informação na rede, o primeiro nó de processamento executa a função aditiva Σ da ponderação somada ao valor do bias b , comparando a função de ativação ϕ (sigmoide, tangente hiperbólica, logarítmica ou Gaussiana), para verificar a magnitude da ativação v e decidir se o neurônio k deve ou não disparar o impulso (BISHOP, 1995).

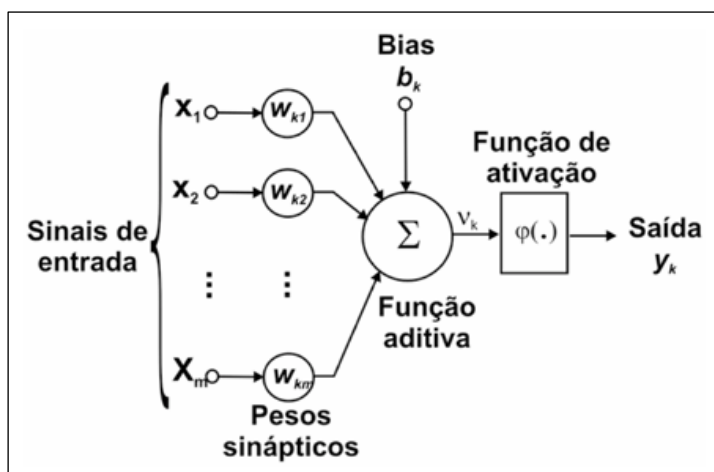


Figura 2.15: Neurônio artificial.
 Fonte: Adaptado de HAYKIN (2001).

Em 1962, Frank Rosenblatt melhorou esse modelo matemático e propôs o conceito de Perceptron, em que a rede neural artificial poderia ter sinapses ajustáveis e ser treinada para classificar certos tipos de padrões. As redes perceptrons de múltiplas camadas (MLP) são constituídas por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. A alimentação dessas redes é progressiva, onde o sinal é propagado camada por camada e processado por funções matemáticas.

O treinamento é o passo mais importante para as redes neurais e tem a finalidade de alcançar os pesos apropriados entre a camada de entrada, as camadas ocultas e a camada de saída para a classificação de padrões desconhecidos. Os algoritmos de retropropagação do erro (RPE) e de função de base radial (FBR) são os mais popularmente utilizados no treinamento de redes

MLP para o mapeamento do espaço multidimensional, fazendo aproximações arbitrárias e não lineares. Esses dois algoritmos podem ser vistos como a aplicação da técnica recursiva estatística de aproximação estocástica, onde aprender é encontrar uma superfície no espaço multidimensional que forneça o melhor ajuste para as amostras de treinamento, e generalizar é utilizar essa superfície para interpolar as amostras de teste.

As redes MLP treinadas com o algoritmo RPE podem ter mais de uma camada oculta, onde os parâmetros são determinados ao mesmo tempo e, ao final, divide o espaço de atributos em hiperplanos. Já com o algoritmo FBR, as redes MLP, com apenas uma camada oculta, são treinadas em dois estágios. Primeiro com a função de base sendo determinada por técnicas não supervisionadas, usando somente os dados de entrada. Segundo, ajustando-se os pesos da segunda camada com aprendizagem supervisionada. Por fim, divide o espaço de atributos em elipses (HAYKIN, 2001).

Os classificadores baseados em redes neurais artificiais possuem várias vantagens sobre os classificadores tradicionais, como serem não paramétricos não impondo suposições sobre a forma e distribuição dos dados, treinamento com poucas amostras, limites de decisão não lineares, capacidade de generalização sem a necessidade de coletar novas amostras de treinamento e aprendizagem de padrões complexos. Porém, a definição da arquitetura da rede e das regras de aprendizado é bastante trabalhosa e subjetiva. As redes MLP apresentam capacidade de aproximação universal, ou seja, podem aproximar virtualmente qualquer função de interesse para qualquer grau de precisão desejado, sempre que for possível definir uma taxa de convergência aceitável para o algoritmo de treinamento, avaliada pelos erros entre os resultados preditos e esperados (TSO e MATHER, 2001).

2.4.2. Noções da teoria de conjuntos 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, ambiguidade e imprecisão para a extração de informações vagas (ZADEH, 1965).

A lógica fuzzy pode ser definida como uma extensão 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 a teoria de conjuntos clássicos permite apenas o uso de funções de associação binária, que identificam os elementos pertencentes ou não pertencentes a uma classe, a teoria de conjuntos fuzzy admite a possibilidade de uma associação parcial, isto é, elementos que são parcialmente pertencentes a uma ou mais classes. Formalmente, um conjunto fuzzy A em um universo de discurso X é definido por uma função de pertinência μ_A que assume valores em um intervalo de zero a um, produzindo o seguinte conjunto de pares ordenados:

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

A função de pertinência μ_A determina o grau de pertinência do elemento x ao conjunto A . Em conjuntos clássicos, um elemento pertence (μ_A igual a um) ou não pertence (μ_A igual a zero) a uma determinada classe. Na teoria dos conjuntos fuzzy, a transição entre pertencer e não pertencer a uma classe é gradual, ou seja, μ_A é um número real variando de forma contínua no intervalo de zero a um. 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 (ZADEH, 1965).

Os graus de pertinência refletem uma ordem não baseada em probabilidade, mas sim em possibilidade. As medidas de probabilidade estão associadas a conjuntos de limites precisos, enquanto que a possibilidade a conjuntos imprecisos. A questão fundamental na determinação dos conjuntos fuzzy é a escolha de uma função de pertinência que seja apropriada para modelar a diversidade de formas de um atributo e estabelecer os parâmetros do modelo. Em geral, o formato das funções de pertinência é restrito a uma certa classe de funções, representadas por parâmetros específicos, conforme exposto a seguir e ilustrado pela Figura 2.16.

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

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

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

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

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

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

d) Conjunto unitário: parâmetros (m, h).

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

A Figura 2.16 ilustra os formatos mais comuns das funções de pertinência dos conjuntos fuzzy descritos anteriormente, especificando os parâmetros associados a cada caso.

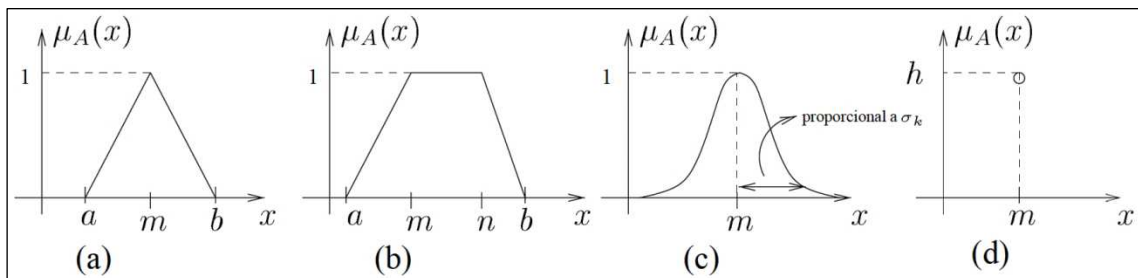


Figura 2.16: Formato das funções de pertinência - (a) Triangular, (b) Trapezoidal, (c) Gaussiana, (d) Conjunto unitário.

Fonte: DELGADO (2002).

As funções triangulares e trapezoidais definem de forma mais intuitiva as regiões de pertinência total, parcial e nula, enquanto as funções gaussianas têm propriedades matemáticas de continuidade para uma diferenciação suave entre as classes (ZADEH, 1990).

Assim como na teoria dos conjuntos clássicos, na teoria dos conjuntos fuzzy existem operações essenciais como união, intersecção e complemento. Considerando os conjuntos fuzzy A e B no universo de discurso X, estas operações básicas são definidas a seguir.

União de A e B: $\mu_{A \cup B}(x) = \max[\mu_A(x), \mu_B(x)] = \mu_A(x) \vee \mu_B(x)$.

Intersecção de A e B: $\mu_{A \cap B}(x) = \min[\mu_A(x), \mu_B(x)] = \mu_A(x) \wedge \mu_B(x)$.

Complemento de A: $\mu_{\bar{A}}(x) = 1 - \mu_A(x)$.

A Figura 2.17 representa as operações básicas de união, intersecção e complemento aplicadas aos conjuntos fuzzy A e B descritos anteriormente.

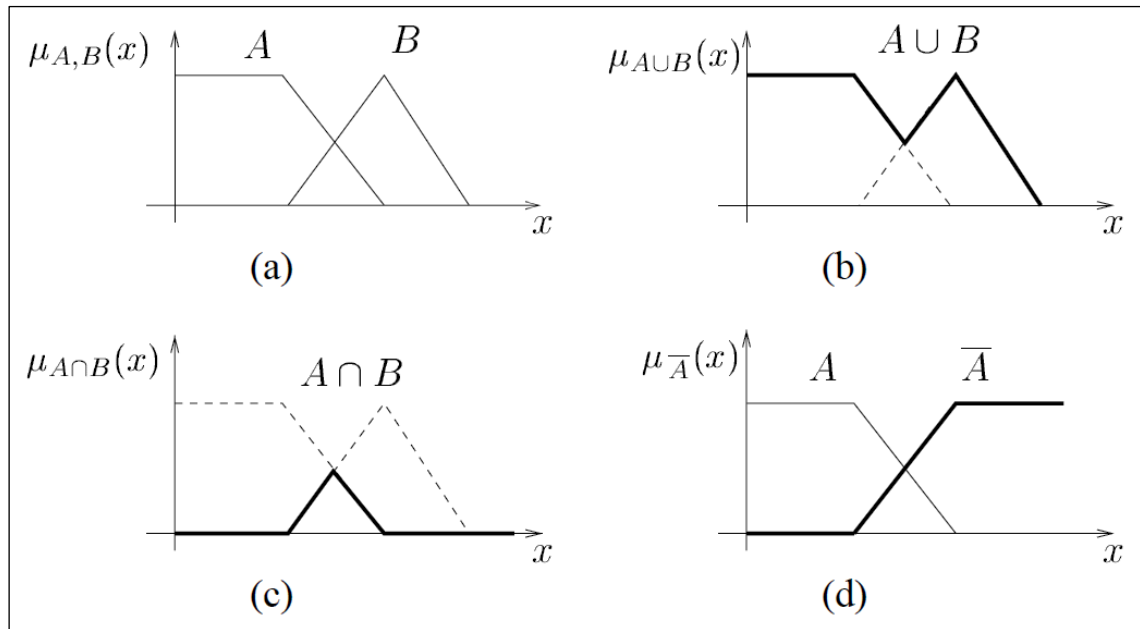


Figura 2.17: Operações básicas em conjuntos fuzzy - (a) Conjuntos fuzzy A e B, (b) União, (c) Intersecção, (d) Complemento.

Fonte: DELGADO (2002).

Devido à semelhança entre a teoria dos conjuntos clássicos e fuzzy, a união e a intersecção podem ser identificadas pela disjunção (OU) e pela conjunção (E), respectivamente, e assim, podem ser representadas pelos operadores \vee e \wedge (PEDRYCZ e GOMIDE, 1998).

2.4.3. Classificação sub-pixel de imagens

Enquanto os métodos de classificação tradicionais atribuem univocamente cada pixel para apenas uma classe, a abordagem de classificação sub-pixel tenta superar as limitações espaciais das imagens de satélites, de modo a revelar possíveis misturas e identificar para cada pixel as proporções das diferentes classes de cobertura da terra. A aplicação de técnicas de inteligência artificial tem sido uma alternativa atraente para o reconhecimento de padrões em imagens de satélite no nível sub-pixel.

O potencial de manuseio de incertezas e de controle de sistemas complexos tornados possíveis pela lógica fuzzy, podem ser integrados às redes neurais artificiais, que por sua vez possuem características de adaptação e aprendizagem. Os classificadores neuro-fuzzy gerados a partir dessa integração têm como propósito minimizar a deficiência da lógica fuzzy em aprender novas regras e, também, melhorar o entendimento do processo de treinamento feito pelas redes neurais artificiais através da observação do comportamento das funções de pertinência fuzzy (BARALDI et al., 2001; KAVZOGLU e MATHER, 2003).

A Figura 2.18 esboça o esquema genérico de um classificador neuro-fuzzy. Como exemplo, os pixels das composições de NDVI e EVI do MODIS são os elementos das classes solo, pastagem e cana-de-açúcar que são definidas como conjuntos fuzzy.

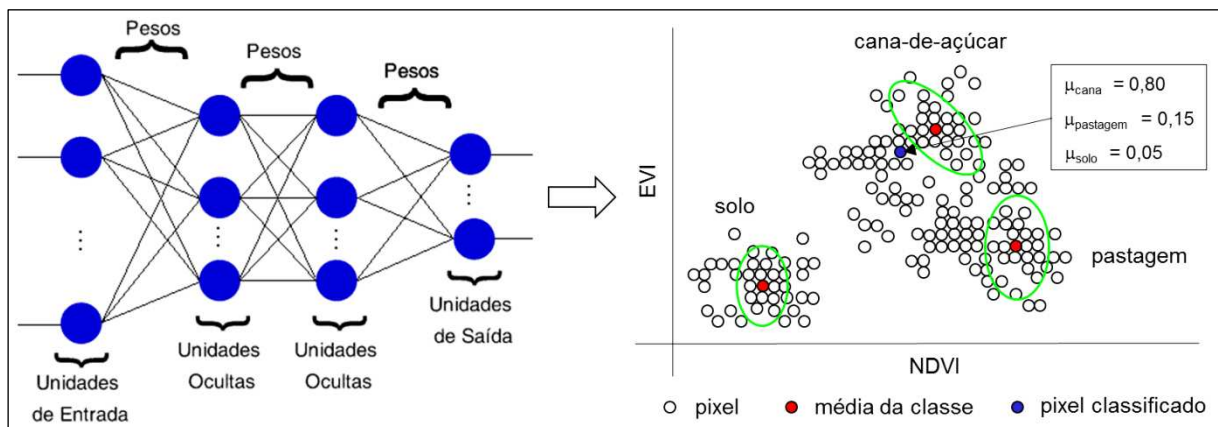


Figura 2.18: Esquemática de um classificador neuro-fuzzy.

A partir do treinamento de redes MLP com amostras coletadas em áreas representativas das classes de cobertura da terra, normalmente constituídas de pixels homogêneos, o algoritmo gera como resultado um conjunto de imagens, onde cada pixel pode estar vinculado às classes solo, pastagem e cana-de-açúcar, descrevendo a mistura espectral pelos seus graus de pertinência, medidos entre zero e um. Os classificadores neuro-fuzzy podem gerar tanto uma imagem de classificação tradicional como também as imagens de classificação sub-pixel, mostrando serem bastante eficazes, pois obtêm alto nível de acurácia com pouco conhecimento necessário para treinar a rede.

Nos últimos anos muito esforço tem sido empregado na compreensão e identificação da incerteza na classificação de imagens. A classificação tradicional pressupõe que cada pixel descreva uma cobertura do solo homogênea, mas na realidade, um pixel pode representar a mistura espectral de classes ou padrões desconhecidos. Nesse sentido, novos algoritmos têm sido desenvolvidos para a análise sub-pixel com o propósito de resolver o problema da mistura espectral. Assim, ao contrário da classificação tradicional que atribui um pixel a uma única classe, a classificação sub-pixel gera um conjunto de imagens que expressam pertinência, adesão ou probabilidade, visando estimar as proporções de mistura de cada classe dentro do pixel (WANG, 1990; BERNARD et al., 1997, ANTUNES, 2005). Essas abordagens incluem classificações sub-pixel com base em probabilidades posteriores Bayesianas (EASTMAN e LANEY, 2002), tipicidades Mahalanobis (FOODY et al., 1992), graus de pertinência fuzzy (FOODY, 1998) e níveis de ativação de redes neurais (ZHANG e FOODY, 2001).

No entanto, para os classificadores não paramétricos, a classificação sub-pixel foi estabelecida apenas para as redes MLP. Porém, a seleção adequada de uma topologia da rede e das propriedades de aprendizagem é considerada bastante difícil e subjetiva, por ser determinada mediante tentativa e erro. Além disso, as redes MLP incorporam novas informações sem esquecer o passado da aprendizagem e, quando utilizam o algoritmo de treinamento RPE, são potencialmente instáveis devido à tendência de encontrar soluções subótimas por ficarem presas em um mínimo local e ser susceptível ao sobreajuste reduzindo a generalização (FOODY, 2004; QIU e JENSEN, 2004).

Embora seja um membro importante da família das redes neurais, o modelo Fuzzy ARTMAP (*Adaptive Resonance Theory*) não foi explorado tão profundamente quanto às redes MLP. Essa arquitetura faz parte das redes neurais auto-organizáveis de agrupamento que atua de forma recorrente, em que os padrões não são apenas alimentados para frente, mas também realimentados da saída para os nós de entrada. O treinamento é realizado com a regra de aprendizagem competitiva “o-vencedor-leva-tudo”, que consiste de uma camada de neurônios com a mesma entrada e o neurônio com a melhor saída é o vencedor, onde um mecanismo de mapeamento interno maximiza a generalização das categorias e minimiza os erros de predição. Ambos os modelos podem realizar a classificação supervisionada e não supervisionada e vêm mostrando grande potencial no reconhecimento de padrões em imagens de satélite (LIU et al., 2004; ROGAN et al., 2008).

As redes neurais ART evoluíram da teoria inspirada no processamento humano da informação cognitiva e são consideradas modelos plausíveis dos neurônios biológicos. O modelo ART baseado na Teoria da Ressonância Adaptativa foi proposto inicialmente por Stephen Grossberg nos anos 70 para resolver o dilema “estabilidade x plasticidade”, que mostrou como um sistema treinado pode permanecer adaptativo a cada nova informação e ainda permanecer estável em resposta a informações irrelevantes, preservando o conhecimento já adquirido e ao mesmo tempo sendo flexível o suficiente para armazenar novas informações (GROSSBERG, 1976).

Esse dilema é resolvido com um mecanismo de vigilância que administra a inclusão de novas entradas em cada agrupamento. Uma regra de similaridade define onde agrupar um padrão por um grau de semelhança entre um padrão previamente armazenado. Em uma rede baseada no modelo ART, as informações são trocadas entre as camadas da rede. O treinamento é realizado quando ocorre a ressonância da atividade neural, ou seja, quando um padrão já treinado é apresentado e a rede o reconhece e, também, quando um novo padrão de entrada é apresentado e a rede percebe que o padrão constitui uma nova informação e, então, entra em estado ressonante para memorizá-lo (CARPENTER, 1989).

Essa teoria serviu de base para o desenvolvimento de arquiteturas da família ART que são redes neurais auto-organizáveis de agrupamento, que mantêm um equilíbrio em criar novas categorias de reconhecimento quando padrões desconhecidos estimulam a rede e a capacidade em agrupar padrões similares na mesma categoria de reconhecimento. A arquitetura ART1 foi

desenvolvida para realizar agrupamentos de dados binários realizando o treinamento não-supervisionado. A arquitetura ART2 é uma evolução da ART1 e pode manipular tanto dados binários como analógicos. A arquitetura Fuzzy ART é uma generalização da ART2 que usa as operações da teoria de conjuntos fuzzy ao invés das operações dos conjuntos clássicos (CARPENTER et al., 1991).

Nesse contexto, LI (2007) desenvolveu algoritmos de classificação sub-pixel para a rede neural Fuzzy ARTMAP, propondo novas medidas baseadas em neurocomputação de acionamento e frequência do neurônio para lidar com a incerteza espacial na classificação de imagens. A medida de Compromisso (*Commitment*) expressa o grau de compromisso de um pixel em relação a uma determinada classe e mostra-se de natureza muito semelhante às probabilidades posteriores Bayesianas. A medida de Tipicidade (*Typicality*) exprime quanto um pixel é característico de uma determinada classe considerando a variabilidade intraclasse e corresponde rigorosamente as tipicidades Mahalanobis. No entanto, além de gerar limites de decisão não lineares, esses algoritmos possuem a grande vantagem de serem não paramétricos (LI e EASTMAN, 2006; LI, 2008; LI e EASTMAN, 2010).

A inteligência artificial no âmbito do sensoriamento remoto pode ser compreendida como sendo o estado da arte. Pouca pesquisa está relacionada com a utilização de técnicas de aprendizado de máquina na classificação de imagens (GIACCO et al., 2010). Como os algoritmos implementados por LI (2007) foram propostos há poucos anos, ainda não existem trabalhos práticos visando à solução de problemas reais publicados na literatura. No entanto, essas novas abordagens de análise sub-pixel mostram ser uma alternativa oportuna, viável e inovadora com grande potencial de aplicação no monitoramento sistemático do cultivo da cana-de-açúcar, podendo trazer contribuições originais para o campo da classificação sub-pixel.

3. MATERIAL E MÉTODOS

3.1. Material

3.1.1. Caracterização da área de estudo

A área de estudo compreende o Estado de São Paulo que está localizado na região Sudeste do Brasil, entre as latitudes 19°40'S e 25°20'S e as longitudes 44°10'O e 53°10'O. Possui 645 municípios numa extensão territorial de 248.222,801 km², que são congregados segundo similaridades sociais, naturais e econômicas em 15 mesorregiões estabelecidas pelo IBGE. A Figura 3.1 ilustra a espacialização dos dados oficiais da área plantada média de cana-de-açúcar (ha) das safras 2004/2005 a 2011/2012, distribuída nas 15 mesorregiões do Estado de São Paulo (IBGE, 2012).

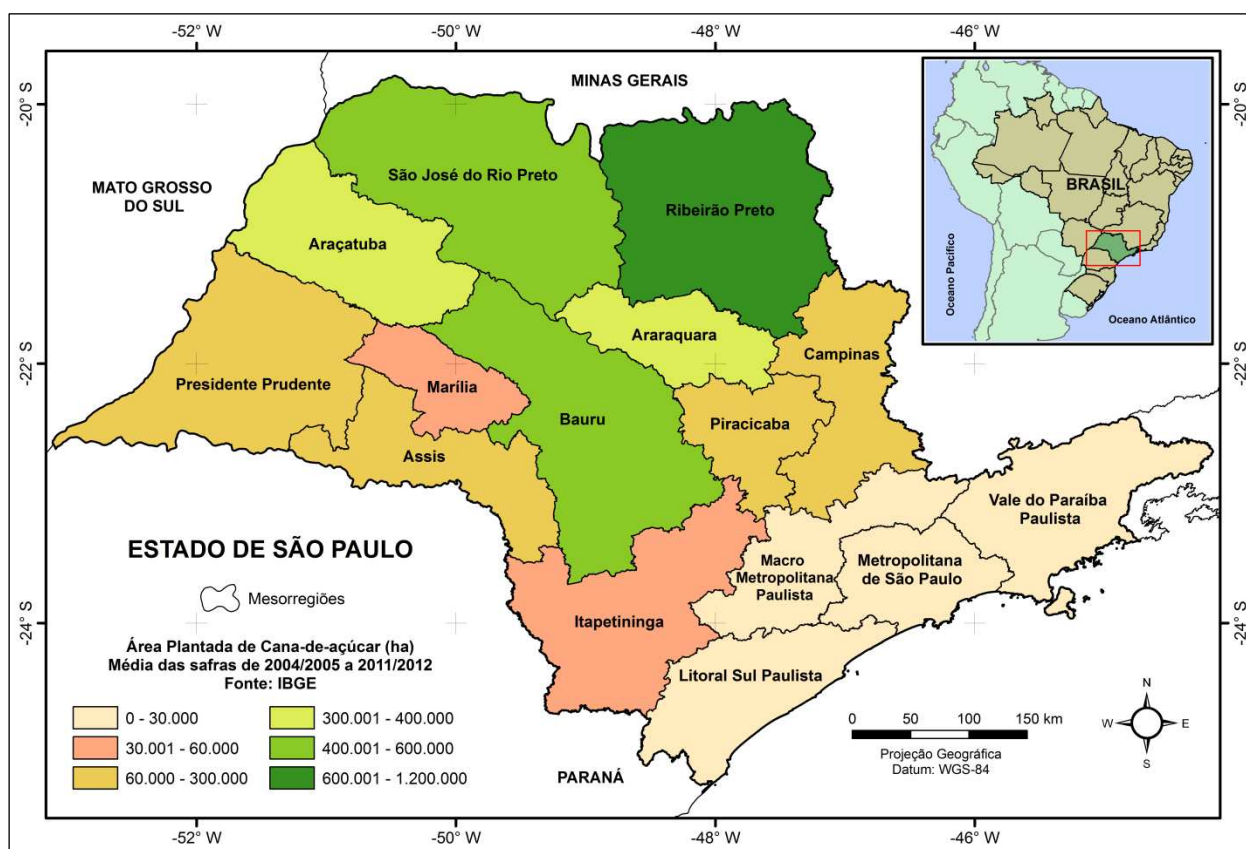


Figura 3.1: Mapa de área plantada média de cana-de-açúcar (ha) das oito safras por mesorregiões do Estado de São Paulo.

3.1.2. Séries temporais MODIS

As imagens do sensor MODIS foram utilizadas na classificação sub-pixel para estimar a área plantada de cana-de-açúcar em nível municipal no Estado de São Paulo, ao longo das safras 2004/2005 a 2011/2012.

As séries temporais de imagens do satélite TERRA foram obtidas do Banco de Produtos MODIS na base Estadual Brasileira (EMBRAPA INFORMÁTICA AGROPECUÁRIA, 2014) que armazena e disponibiliza na Internet imagens prontas para uso, sem a necessidade de processamentos complementares. Os produtos são fornecidos em recortes por estados, na projeção geográfica, datum WGS-84 e no formato GeoTIFF (*Geographic Tagged Image File Format*), compatível com a maioria dos softwares de processamento de imagens e sistemas de informação geográfica (ESQUERDO et al., 2010).

O processamento do produto MOD13Q1 inclui a geração do mosaico dos quadrantes que cobrem o território nacional, a reprojeção cartográfica e o recorte do mosaico de acordo com a base estadual brasileira, conforme o diagrama apresentado na Figura 3.2.

As etapas são executadas automaticamente por meio de rotinas implementadas na linguagem IDL (*Interactive Data Language*) do software ENVI (*The Environment for Visualizing Images*) (EXELIS, 2014), com a execução de programas do pacote computacional MRT (LP DAAC, 2014).

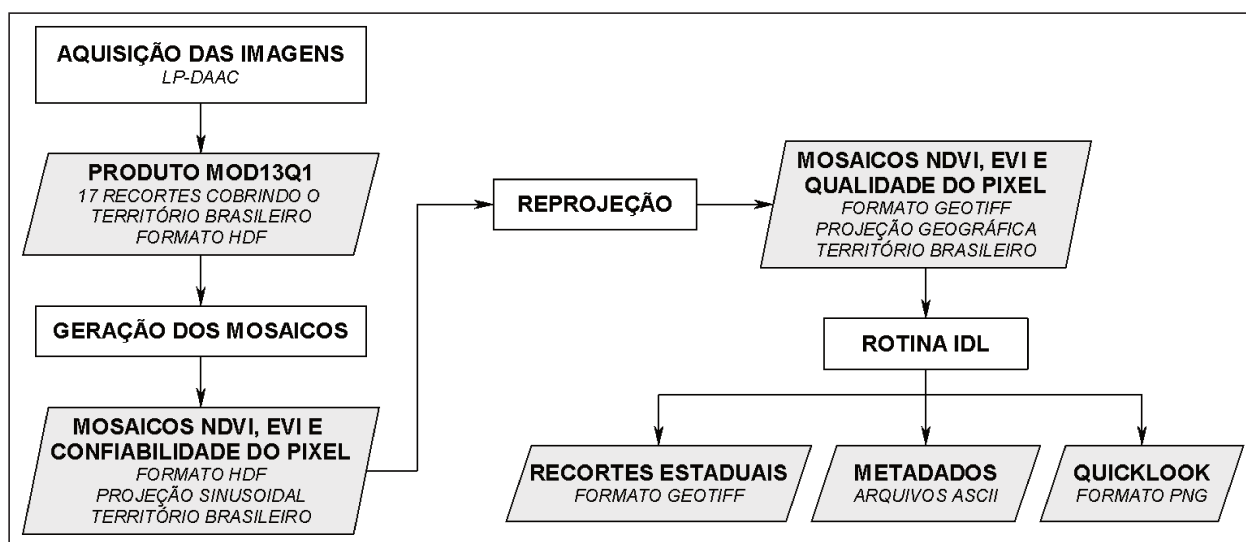


Figura 3.2: Processamento do produto MOD13Q1 para alimentação do banco de imagens.

O primeiro passo é a separação das imagens EVI, NDVI e confiabilidade do pixel presentes no produto MOD13Q1 e a geração de três mosaicos, reunindo os 17 quadrantes que cobrem o território brasileiro. Em seguida é realizado o processo de reprojeção do mosaico da projeção sinusoidal para a projeção geográfica, adotando-se o método de reamostragem pelo vizinho mais próximo, datum WGS-84 e formato de saída GeoTIFF.

No passo seguinte é chamada uma rotina IDL, cuja função é carregar os mosaicos reprojatados e fazer o recorte dos limites geográficos de cada estado. Nessa etapa também são gerados os metadados e as figuras para visualização rápida das imagens, que são utilizados no sistema Web implementado no ambiente GeoNetwork para disponibilização dos produtos na Internet (OSGEO, 2010), como pode ser visto na Figura 3.3.



Figura 3.3: Banco de Produtos MODIS.
 Fonte: EMBRAPA INFORMÁTICA AGROPECUÁRIA (2014).

Atualmente estão disponíveis as séries temporais completas de EVI, NDVI e confiabilidade do pixel, da coleção 5, do produto MOD13Q1 (satélite TERRA) e, também, MYD13Q1 (satélite AQUA), na resolução espacial de 250 m, atualizadas a cada 16 dias seguindo a distribuição do LP DAAC. Para cada satélite, anualmente são processadas 23 composições para

cada quadrante, totalizando 391 imagens para os 26 estados do Brasil. Neste trabalho foram utilizadas apenas as séries temporais de EVI e NDVI do sensor MODIS do satélite TERRA.

Como exemplo, a Figura 3.4 apresenta a composição colorida RGB-732 e a composição de NDVI do MODIS/TERRA, compostas no período de 16 dias, de 22 de março a 06 de abril de 2011, obtida do Banco de Produtos MODIS.

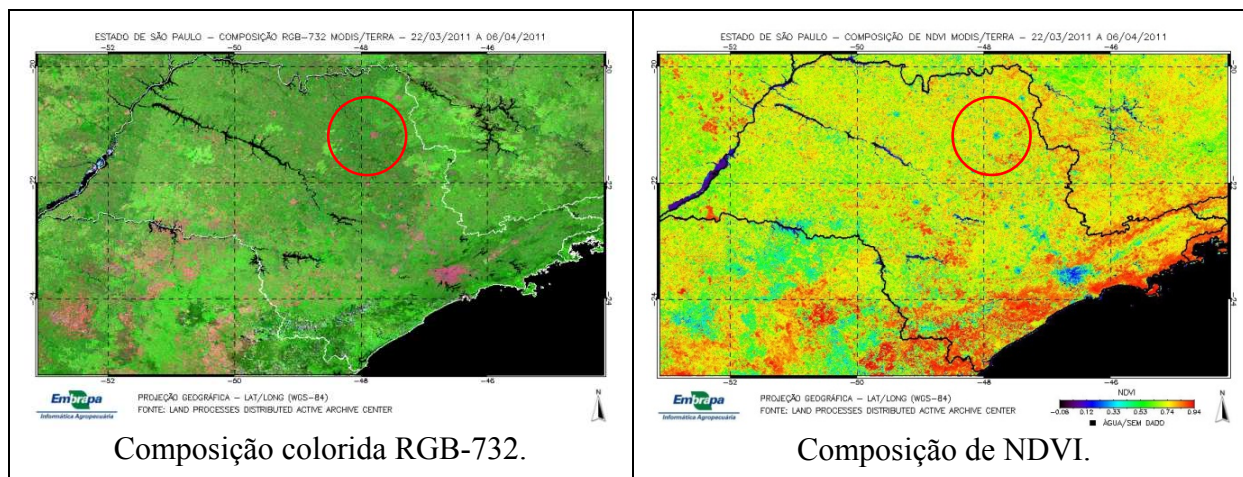


Figura 3.4: Produtos MODIS para o Estado de São Paulo.

A composição colorida RGB-732 oferece um bom contraste dos alvos da superfície terrestre. As áreas em tons de verde mais escuro correspondem à vegetação natural, as em tons de verde mais claro a áreas agrícolas e pastagens, as em tons de marrom a solo exposto, as em tons de magenta a áreas urbanas e as escuras a corpos d'água.

A composição de NDVI indica a presença e o vigor da vegetação, variando do preto referente a valores mais baixos, ao vermelho com valores mais altos. As áreas em tons de vermelho mais intenso indicam a presença de vegetação natural, as em tons de vermelho mais claro a áreas agrícolas, as em tons de amarelo a pastagens, as em tons de verde mais claro a solo exposto, as em tons de azul a áreas urbanas e as escuras a corpos d'água.

Na região de Ribeirão Preto em destaque, pode-se observar o aparecimento de áreas em tons de verde claro na composição colorida RGB-732 e em tons de vermelho claro na composição de NDVI, que correspondem à cultura da cana-de-açúcar na fase final de maturação quando é iniciada a colheita.

3.1.3. Mapas temáticos de referência

Para delimitação das amostras de treinamento a serem utilizadas pelo classificador e das amostras de teste para avaliação da acurácia dos modelos de classificação, foram utilizados os mapas de referência de cultivo da cana-de-açúcar do Canasat, no formato vetorial, em nível municipal no Estado de São Paulo, para as safras 2004/2005 a 2011/2012, solicitados a partir do sítio do projeto na Internet (INPE, 2012), mostrado na Figura 3.5.

Os dados de área plantada de cana-de-açúcar do Canasat, no formato de planilha eletrônica, também serviram de referência para análise comparativa dos resultados do trabalho entre as estimativas oficiais e as obtidas pela classificação sub-pixel.

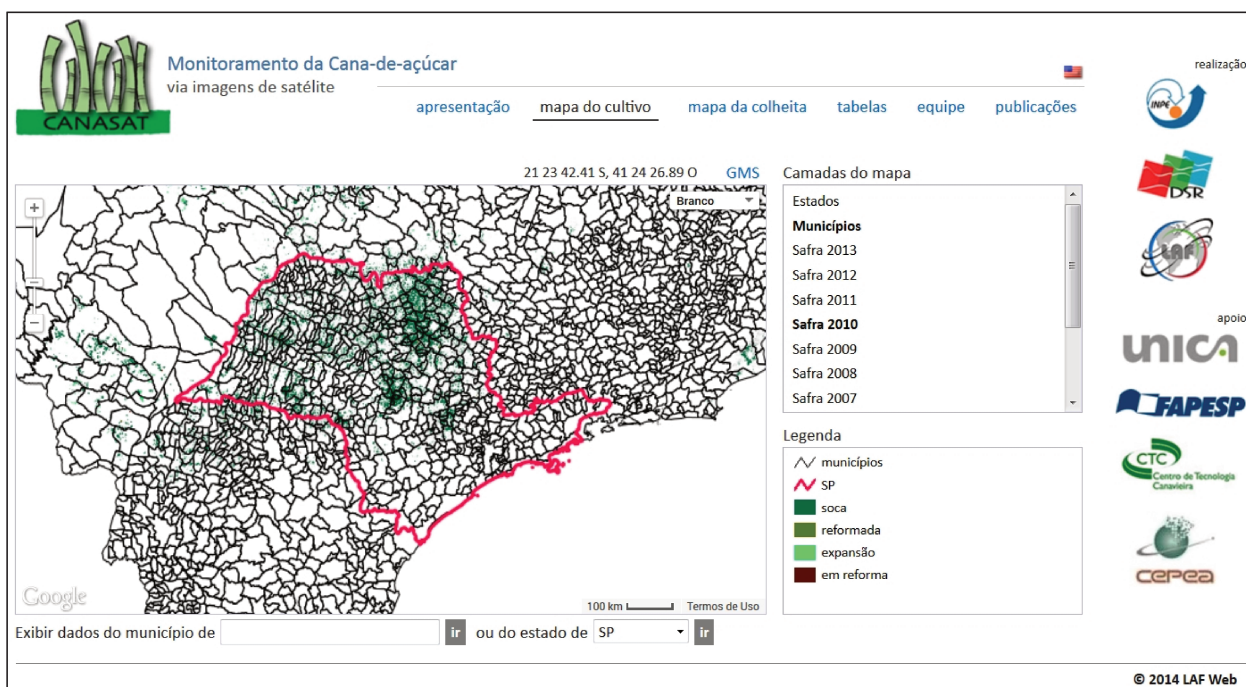


Figura 3.5: Canasat.
Fonte: INPE (2012).

O mapeamento temático produzido pelo Canasat da safra 2010/2011 na região Centro-Sul obteve exatidão global de 98%, com um erro de subestimativa de área plantada de cana-de-açúcar de 0,5%, verificada por meio de uma plataforma Web que integra diferentes tipos de imagens de sensoriamento remoto para interpretação visual de dados de referência. A validação foi realizada por uma campanha de campo nos estados de São Paulo, Minas Gerais e Paraná,

mostrando que o mapeamento é de excelente qualidade temática, pois oferece estimativas de área muito precisas e confiáveis para fins de estatísticas agrícolas. Dessa forma, também são esperados resultados com esse alto nível de acurácia para as demais safras, devido ao método consistente e a cuidadosa interpretação visual realizada por intérpretes experientes do Canasat (ADAMI et al., 2012).

3.1.4. Dados meteorológicos

A distribuição das chuvas durante a safra pode afetar os padrões fenológicos da cana-de-açúcar ao longo do seu ciclo de desenvolvimento, alterando o momento de maior vigor vegetativo. Para auxiliar na análise da dinâmica temporal da cana-de-açúcar, foram utilizados os dados diários de precipitação do Estado de São Paulo obtidos do Agritempo (Figura 3.6), acumulados para o período de 16 dias do MODIS, ao longo das safras 2004/2005 a 2011/2012.

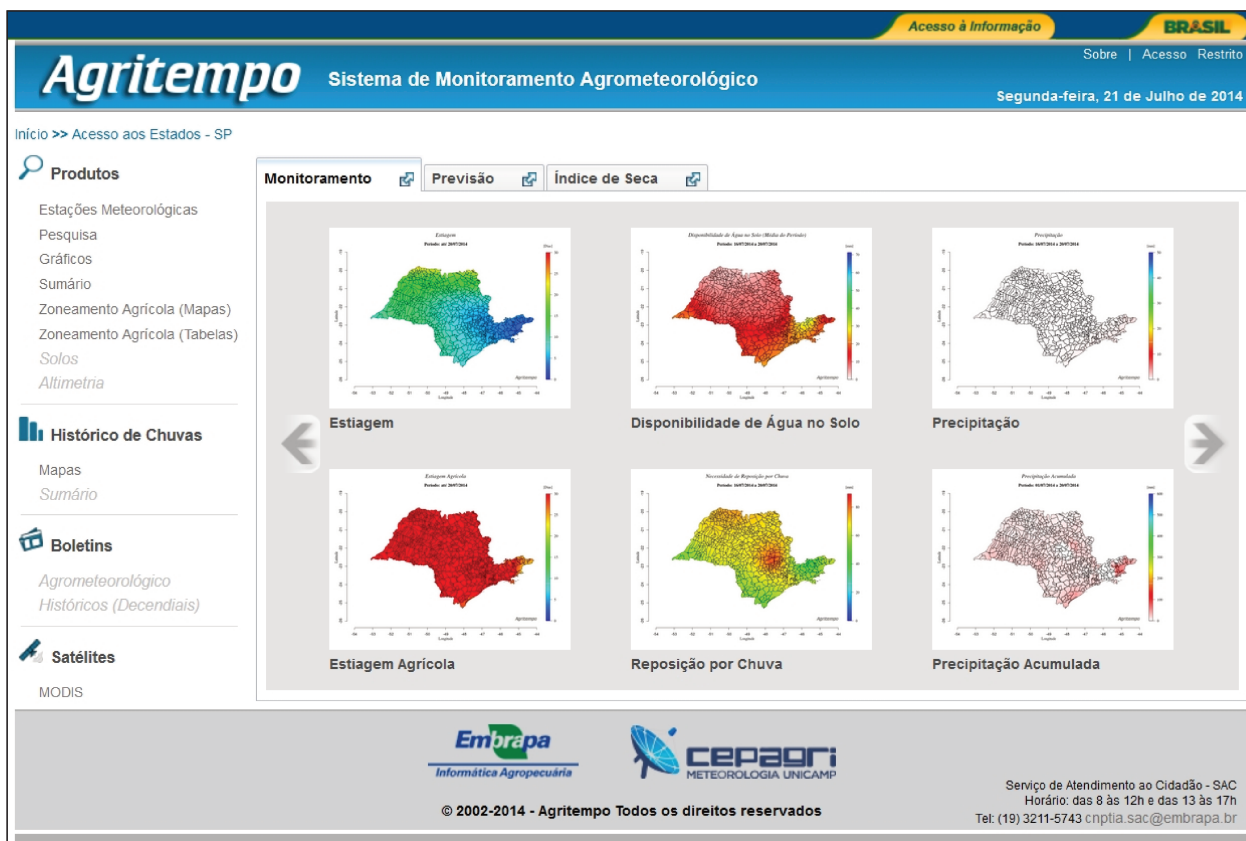


Figura 3.6: Agritempo.
Fonte: EMBRAPA INFORMÁTICA AGROPECUÁRIA e CEPAGRI/UNICAMP (2014).

O Agritempo é um sistema de monitoramento que disponibiliza na Internet dados meteorológicos e agrometeorológicos de diversos municípios e estados brasileiros, adquiridos de mais de 1.300 estações meteorológicas espalhadas por todo o país, possuindo um banco histórico de dados diários de chuvas coletados durante mais de 20 anos. No Estado de São Paulo, os dados de precipitação acumulada são calculados por interpolação de 195 estações pluviométricas. (EMBRAPA INFORMÁTICA AGROPECUÁRIA e CEPAGRI/UNICAMP, 2014).

3.1.5. Estatísticas agrícolas

Para análise comparativa dos resultados do trabalho, foram utilizadas as estatísticas agrícolas oficiais referentes aos dados de área plantada de cana-de-açúcar, em nível municipal do Estado de São Paulo, para as safras 2004/2005 a 2011/2012, obtidas do Sistema IBGE de Recuperação Automática (SIDRA) que disponibiliza na Internet dados agregados do LSPA sobre previsão e acompanhamento de safras agrícolas com estimativas de produção, rendimento e áreas cultivadas, tendo como unidade de coleta os municípios (IBGE, 2012), mostrado na Figura 3.7.

The screenshot shows the IBGE SIDRA system interface. The main content area is titled "Montar quadro" and displays the following configuration:

- Tabela:** 1612 - Área plantada, área colhida, quantidade produzida e valor da produção da lavoura temporária (vide Nota de Rodapé)
- Matriz:** multidimensional (7x32x23x6281) com 32.359.712 valores
- Variável (7):**
 - Área plantada (Selecção)
 - Área colhida (Selecção)
 - Lavoura temporária (32): (Selecção)
 - Ano (23): (Selecção)
 - Unidade Territorial (6281): (Selecção)
- Ordenação:** Ascendente (checkbox checked)
- Exibir:** código (checkbox checked), nome (checkbox checked)
- Níveis Territoriais:**
 - Brasil (1): Sim
 - Grande Região (5): Sudeste
 - Unidade da Federação (27): São Paulo
 - Mesorregião Geográfica (137): Em São Paulo (15)
 - Microrregião Geográfica (558): Não
 - Município (5553): Não

Figura 3.7: SIDRA.
Fonte: IBGE (2012).

Para o monitoramento do cultivo da cana-de-açúcar foram considerados todos os 645 municípios, onde o somatório da área plantada em nível municipal foi consolidado por mesorregiões e totalizado para o nível estadual.

3.1.6. Informações cartográficas

As informações cartográficas dos limites municipais, de mesorregiões e estadual para o Estado de São Paulo foram obtidas do IBGE no formato vetorial SHP (*shapefile*), referente à malha digital do ano 2001.

3.1.7. Recursos computacionais

Para processamento das imagens e vetores, visualização dos dados e análise espacial foram utilizados os softwares comerciais ENVI 5.1/IDL 8.3 (EXELIS, 2014), IDRISI 16.05 Taiga (CLARK LABS, 2014) e ArcGIS 10.2 (ESRI, 2014) e, também, o software livre QGIS 2.0.1 Dufour, distribuído sob a licença GNU (OSGEO, 2014).

Para intercâmbio entre os formatos de dados geográficos do ENVI e do IDRISI foi utilizada a biblioteca livre GDAL 1.10.0, distribuída sob a licença X/MIT (OSGEO, 2012).

Para o tratamento dos dados e geração de gráficos foi utilizado o software de planilha eletrônica Microsoft Excel 2010. Para as análises estatísticas foi utilizado o software livre Action, desenvolvido sob a plataforma R e que permite trabalhar de forma integrada ao Excel, distribuído sob a licença GPL (ESTATCAMP, 2014).

As atividades do trabalho de pesquisa foram realizadas em um computador Dell, processador Intel, CPU de 3,3 GHz, 8 GB de RAM, 1 TB de disco rígido e monitor de 21,5 polegadas, com sistema operacional de 64 bits Microsoft Windows 7 Professional, localizado na Embrapa Informática Agropecuária.

3.2. Métodos

A abordagem do projeto de pesquisa baseou-se na integração e manipulação das imagens de entrada para gerar os modelos de classificação, por meio de operações espaciais e aplicação de algoritmos de processamento de série temporais. A análise dos resultados consistiu na avaliação da acurácia da classificação e na comparação pareada entre as estimativas de área plantada de cana-de-açúcar. A Figura 3.8 apresenta o fluxograma da pesquisa realizada no trabalho e, a seguir, a descrição de cada etapa metodológica na ordem do círculo numerado.

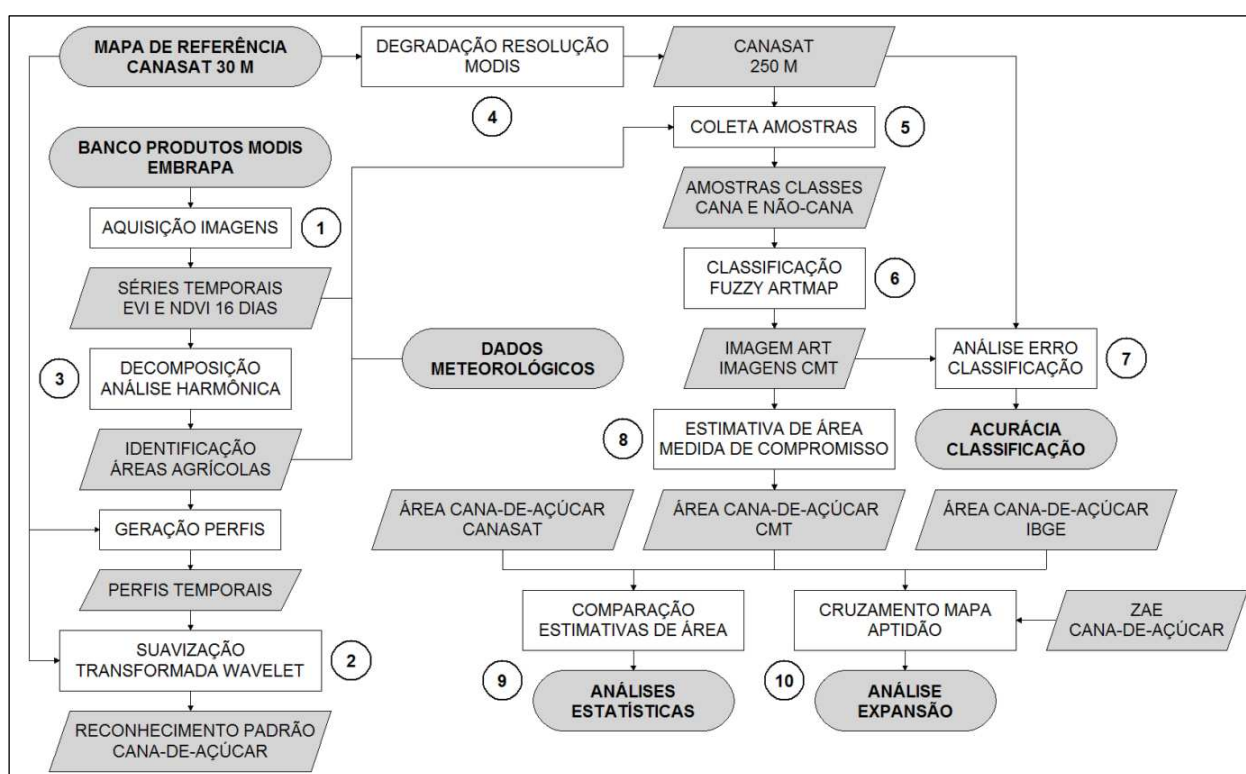


Figura 3.8: Fluxograma das etapas da metodologia do trabalho.

3.2.1. Decomposição das séries temporais MODIS

A análise harmônica foi utilizada com o propósito de decompor as séries temporais dos índices de vegetação EVI e NDVI em termos harmônicos, que são ondas cossenoidais ou senoidais, onde cada onda é definida por um valor único de amplitude e fase, conforme ilustrado na Figura 3.9.

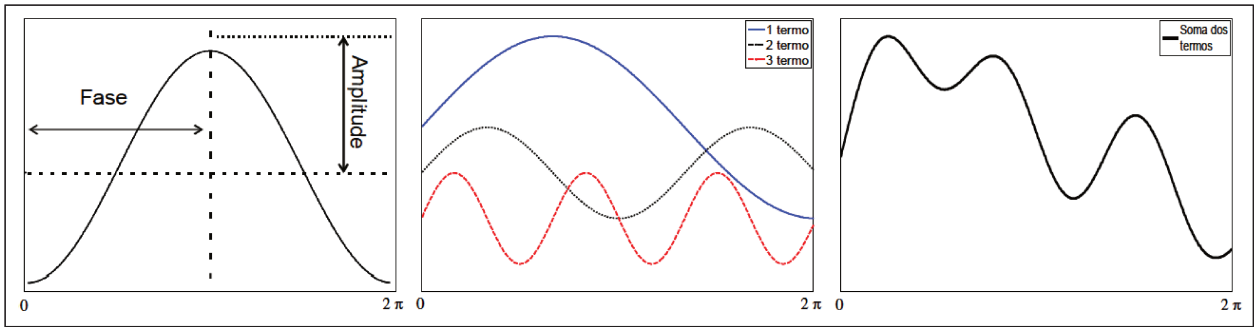


Figura 3.9: Representação dos termos harmônicos.
Fonte: ADAMI (2010).

A amplitude corresponde à metade do valor onde a função é maximizada e a fase é o deslocamento entre a origem e o pico da onda no intervalo de 0 até 2π . Cada termo harmônico representa o número de ciclos completados por uma onda num determinado intervalo de tempo e é responsável por um percentual da variância total da série temporal de dados original. Assim, o primeiro harmônico tem um período T igual ao período total, o segundo harmônico corresponde à metade do período do primeiro harmônico $T/2$, o terceiro harmônico a $T/3$ e assim sucessivamente (JAKUBAUSKAS et al., 2001).

Uma série temporal de dados de n amostras equidistantes pode ser representada pela Transformada de Fourier para a obtenção dos valores ao longo do tempo t , permitindo a reconstrução da curva, por meio da Equação 3.1:

$$f(x)_t = \bar{c} + \sum_{j=1}^{n/2} c_j \cos\left(\frac{2\pi t}{n} - \phi_j\right) \quad (3.1)$$

onde,

\bar{c} = média dos valores da série temporal;

c_j = amplitude do j -ésimo harmônico;

ϕ_j = fase do j -ésimo harmônico.

A amplitude é calculada com a Equação 3.2:

$$c_j = \sqrt{a_j^2 + b_j^2} \quad (3.2)$$

onde, a_j e b_j são as componentes de cosseno e seno do vetor amplitude c_j e são obtidas pelas Equações 3.3 e 3.4, respectivamente:

$$a_j = \frac{2}{n} + \sum_{t=1}^n y_t \cos\left(\frac{2\pi t}{N}\right), \text{ para } t \geq 0. \quad (3.3)$$

$$b_j = \frac{2}{n} + \sum_{t=1}^n y_t \sin\left(\frac{2\pi t}{N}\right), \text{ para } t \geq 1. \quad (3.4)$$

A fase é calculada pela Equação 3.5:

$$\varphi_j = \begin{cases} \arctan \frac{b_j}{a_j}, & \text{se } a_j > 0. \\ \arctan \frac{b_j}{a_j} \pm \pi, & \text{se } a_j < 0. \\ \frac{\pi}{2}, & \text{se } a_j = 0. \end{cases} \quad (3.5)$$

As imagens obtidas a partir da aplicação da Transformada de Fourier foram geradas numa base por pixel para cada imagem da série temporal. Na prática, considerando um ciclo anual, significa decompor a variação dos índices de vegetação em harmônicos que representem a média e as oscilações anuais, semestrais, quadrimestrais, etc., o que propicia a identificação de áreas agrícolas. A imagem harmônico zero (termo aditivo) corresponde à média dos índices de vegetação e representa o vigor total para cada tipo de cobertura vegetal durante o período. As imagens amplitude indicam a variação máxima dos índices de vegetação para todo o período. A fase, com intervalo de valores de 0° a 360° , indica o tempo, ao longo da série temporal em que acontece o máximo valor de amplitude, podendo ser associada com os meses do ano.

O comportamento desses parâmetros é indicativo do tipo de mudança que vem ocorrendo com o tempo. As mudanças sazonais somente na amplitude podem indicar variações no tipo de uso do solo ou nas condições da vegetação. As mudanças apenas de fase podem indicar variações no tempo de máximo vigor vegetativo que pode estar relacionado às variações na época de plantio ou de colheita. As mudanças na amplitude e na fase podem indicar

transformações significativas na superfície terrestre relacionadas com a alteração do manejo do solo ou substituição da vegetação.

Os altos valores de amplitude indicam um alto grau de variação e o harmônico onde essa variação acontece indica a periodicidade do evento. Os elevados valores de amplitude no primeiro harmônico indicam uma cultura agrícola com padrão temporal unimodal e no segundo harmônico, um padrão temporal bimodal, diferentemente de áreas de pastagens e vegetação natural cujo padrão temporal tem pequena variação (JAKUBAUSKAS et al., 2002).

O processamento das séries temporais de EVI e NDVI foi realizado por meio do software HANTS (*Harmonic Analysis of NDVI Time-Series*), desenvolvido originalmente por ROERINK et al. (2000) e implementado em linguagem IDL por DE WIT e SU (2005).

O conceito básico do algoritmo é que o desenvolvimento da vegetação apresenta um forte efeito sazonal que pode ser simulado por funções senoidais de baixa frequência, com diferentes fases e amplitudes. As nuvens provocam alterações nas séries temporais e podem ser consideradas ruídos de altas frequências. O HANTS trabalha de forma iterativa a partir de uma série temporal, definindo-se o número de harmônicos relevantes dos períodos para decompor e recompor os dados, o intervalo válido dos valores e a tolerância do ajuste entre a diferença dos valores da série temporal original e da filtrada. O algoritmo analisa a série temporal original e identifica os valores de baixas frequências referentes à vegetação. Os valores discrepantes da série temporal que estiverem acima do limiar de tolerância são considerados ruídos de altas frequências e, assim, são substituídos pelos valores filtrados da Transformada de Fourier. No entanto, ao substituir os valores pode ocorrer a mudança na média da série temporal e, assim, é necessária uma próxima iteração, repetindo o processamento até que não haja mais valores fora da tolerância estabelecida.

Para decomposição das séries temporais de EVI e NDVI foram utilizados os três primeiros harmônicos que são responsáveis por explicar a maioria da variância. Os dados estão no fator original de escala, no intervalo válido que varia de -2.000 a 10.000, sendo -3.000 para dados inválidos. A tolerância do ajuste entre a diferença dos valores da série temporal original e da filtrada foi definida em 1.000.

A representação visual eficaz das imagens amplitude e fase dos termos harmônicos pode auxiliar na interpretação das mudanças fenológicas da vegetação pela combinação de cores. Porém, a fase é quantificada numa escala circular, o que torna difícil a interpretação de imagens no espaço de cores RGB, onde ocorre uma mudança abrupta entre a cor escura para a cor brilhante. Já no espaço de cores HLS (*Hue, Lightness, Saturation*), a componente matiz é inerentemente circular, proporcionando um efeito visual de transição suave do azul para o magenta e para o vermelho. Os outros parâmetros harmônicos são lineares e podem ser bem representados nas componentes luminosidade ou saturação (HALL-BEYER, 2007).

Com base nisso, para analisar a condição de crescimento da vegetação de áreas agrícolas, em especial a informação de fase, a transformação HLS foi aplicada aos termos harmônicos obtidos com o HANTS, segundo o procedimento desenvolvido por FOLEY e VAN DAM (1984), que está disponível no software ENVI. Esse espaço de cores é formado por um duplo hexacone, conforme apresentado na Figura 3.10.

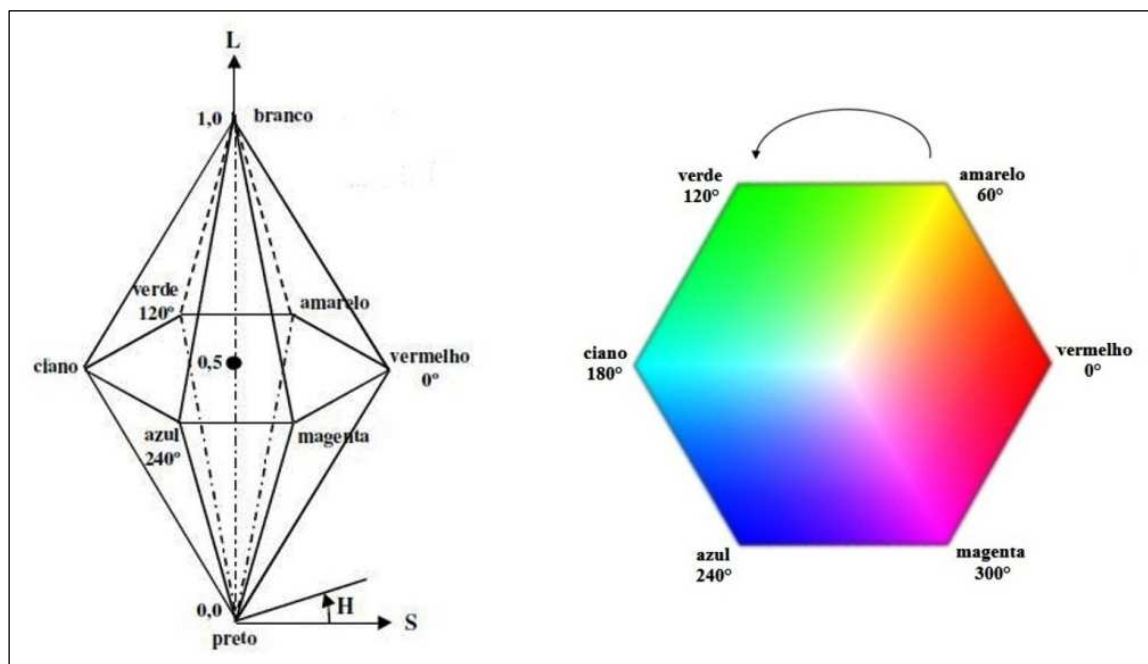


Figura 3.10: Espaço de cores HLS.

As cores primárias estão situadas no perímetro da base comum e as cores branca e preta no vértice de cada um dos cones. Os tons de cinza localizam-se sobre o eixo comum dos dois cones. Matiz (H) é o ângulo ao redor do eixo vertical do hexacone, percorrido no sentido anti-

horário, com o vermelho em 0°, amarelo em 60°, verde em 120°, ciano em 180°, azul em 240° e magenta em 300°. Saturação (S) é medida radialmente a partir do eixo vertical, sendo zero as cores brancas com os tons de cinza variando até um referente às cores puras. Luminosidade (L) pode variar entre zero para preto até um para branco, com as cores puras apresentando uma luminosidade de 0,5.

3.2.2. Reconhecimento do padrão espectro-temporal da cana-de-açúcar

Os perfis temporais foram gerados a partir das séries temporais de índices de vegetação em áreas de cultivo da cana-de-açúcar de referência terrestre. Porém, a possível presença de ruídos nas composições MODIS, pode afetar o reconhecimento do padrão espectral característico da cana-de-açúcar ao longo das safras.

A Transformada de Wavelet foi utilizada com a finalidade de suavizar os perfis temporais, fazendo o ajuste de uma curva para remover os ruídos de alta frequência e destacar as alterações de baixa frequência relacionadas à fenologia da vegetação. Segundo SAKAMOTO et al. (2005), essa transformação matemática é aplicada para decompor e recompor dados em diferentes escalas temporais, para obtenção de informações ocultas no domínio da frequência, onde cada escala é representada por uma frequência específica.

Com base em DAUBECHIES (2006), uma série temporal pode ser representada pela Transformada de Wavelet, onde $\phi(t)$ é uma função de oscilação com energia finita e média nula, conforme a Equação 3.6:

$$\int_{-\infty}^{+\infty} \phi(t) dt = 0 \quad (3.6)$$

A Transformada de Wavelet $W(a,b)$ é definida pela Equação 3.7:

$$W(a, b)_i = \frac{1}{\sqrt{|a|}} \int \phi^* \left(\frac{t-b}{a} \right) s(t) dt \quad (3.7)$$

onde, $s(t)$ é o sinal de entrada analisado e ϕ^* é a *wavelet* mãe ou base. Nesta equação, a largura da *wavelet* é determinada pelo parâmetro de escala “a” enquanto o seu centro é determinada pelo

parâmetro “b”, com $a, b \in \mathbb{R}$ e $a \neq 0$. Para trabalhar com sinais discretos, usualmente utilizam-se valores especiais para a e b , sendo $a = 2^j$ e $b = k2^j$, com $j, k \in \mathbb{Z}$. Mudando o valor do parâmetro “a” tem-se o efeito de dilatação ($a > 1$) ou de contração ($a < 1$), enquanto que mudanças no parâmetro “b” têm o efeito de translação para analisar a função $\phi(t)$ em torno desse ponto. A variável “t” representa o intervalo de tempo na série temporal na qual é realizada a integração. A Transformada de Wavelet tem a vantagem de manter as informações relacionadas à largura (escala) e a localização (tempo) das características presentes em $s(t)$. Conforme GENDRIN et al. (2006), a Equação 3.8 pode ser utilizada para reconstruir o sinal.

$$W = \sum_{i=1}^x W(a, b)_i \quad (3.8)$$

A série temporal de *wavelet* W é um somatório das *wavelets* de larguras decrescentes de i a x , em que x é número de transformações necessárias para alcançar o número definido de coeficientes retidos dos dados de entrada. A filtragem inicia-se com a aplicação da função de suavização na série temporal para remover os ruídos de alta frequência. A série temporal suavizada é então reconstruída com a Transformada de Wavelet discreta. As funções *wavelets* são obtidas por uma dilatação e uma translação, formando uma base que não precisa ser necessariamente ortogonal. Uma das vantagens de se trabalhar com funções de bases ortogonais é que permitem a reconstrução perfeita do sinal original a partir dos coeficientes das *wavelets* (MORETTIN, 1999).

A aplicação da Transformada de Wavelet numa série temporal requer a definição dos parâmetros ordem e potência da *wavelet* mãe que definem o comportamento da curva. Ordem é uma medida de suavidade, em que valores mais elevados produzem mais suavização. Potência corresponde ao número de coeficientes retidos da série temporal original. Uma maior potência gera uma *wavelet* mais estreita que inclui detalhes de escala mais fina, mas também pode reter mais ruído. Uma menor potência remove mais ruído pela forma mais larga da *wavelet*, captura tendências, mas perde detalhes fenológicos (GALFORD et al., 2008).

A Figura 3.11 ilustra a função *wavelet* mãe e escala da família Daubechies 8 (Db8). Por ser ortogonal, essa função assegura que o sinal decomposto é reconstruído sem a presença de resíduos devido à assimetria da *wavelet* mãe.

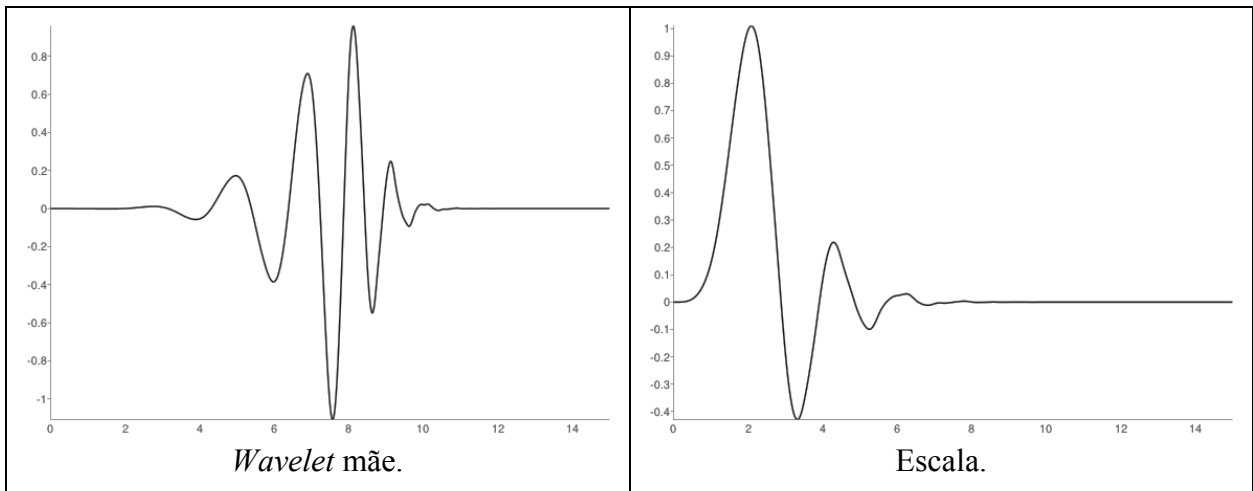


Figura 3.11: *Wavelet* da família Daubechies 8.
 Fonte: WASILEWSKI (2014).

Os perfis temporais suavizados foram gerados por meio do sistema *Series View* do INPE que é um ambiente para a visualização de séries temporais MODIS para dar suporte a estudos e análises de mudança de uso e cobertura do solo (INPE, 2014), mostrado na Figura 3.12.

series view
 ambiente para a visualização de séries temporais para análise de mudanças de uso e cobertura da terra

home informações publicações equipe tutorial

O LAF disponibiliza neste portal uma ferramenta para a visualização instantânea das séries temporais derivadas de imagens de sensoriamento remoto. Esta ferramenta foi desenvolvida dentro do conceito de um Laboratório Virtual de Sensoriamento Remoto (Freitas et al., 2011) para dar suporte a estudos e análises de mudança de uso e cobertura da terra.

Atualmente estão disponíveis as séries temporais EVI2-MODIS e precipitação acumulada mensal do TRMM para qualquer coordenada geográfica da América do Sul, desde o ano de 2000. Estas séries estão integradas às funcionalidades do globo virtual do Google Maps. Com um mínimo de conhecimento técnico sobre a dinâmica de diferentes coberturas de vegetação é possível recuperar o histórico de uso da terra desde o ano 2000, bem como da precipitação acumulada mensal.

O portal oferece ainda as funcionalidades:

- o usuário pode carregar os seus próprios arquivos no formato shapefiles e criar visualizações (camadas) deste arquivo que serão exibidas no Google Maps;
- o usuário criar e salvar os seus próprios polígonos sobre o Google Maps.

Atenção especial aos seguintes itens:

- Aceitar os Termos de Uso disponível em informações;
- The MODIS data are distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at the U.S. Geological Survey

Folder do projeto: [ícones]

© 2014 LAF Web

Figura 3.12: Ambiente para visualização de séries temporais MODIS.
 Fonte: INPE (2014).

Essa ferramenta gera instantaneamente perfis temporais a partir do índice de vegetação EVI2, calculado com dados das bandas espectrais das composições de 16 dias do MODIS, apresentando as curvas originais e as curvas suavizadas com a Transformada de Wavelet Daubechies (FREITAS et al., 2011). O EVI2, desenvolvido por JIANG et al. (2008), não utiliza a banda do azul por ser altamente influenciada pelos aerossóis da atmosfera e, deste modo, reduz os ruídos nas imagens, mostrando ter um comportamento muito similar ao EVI.

A curva gerada pela Transformada Wavelet Daubechies 8 propicia capturar as tendências dos ciclos vegetativos em toda a série temporal, no entanto, a suavização acentuada pode provocar uma diminuição dos valores mais altos do EVI2 e, assim, perder alguns detalhes da fenologia da vegetação.

A Figura 3.13 apresenta a localização geográfica no sistema *Series View* das 10 áreas de cultivo da cana-de-açúcar no Estado de São Paulo, de aproximadamente 400 hectares cada uma, referentes aos dados de campo fornecidos pela Raízen, que é uma das maiores empresas do setor de energia do Brasil, com o plantio realizado em terras arrendadas e 24 unidades produtoras de cana-de-açúcar distribuídas pelo país.

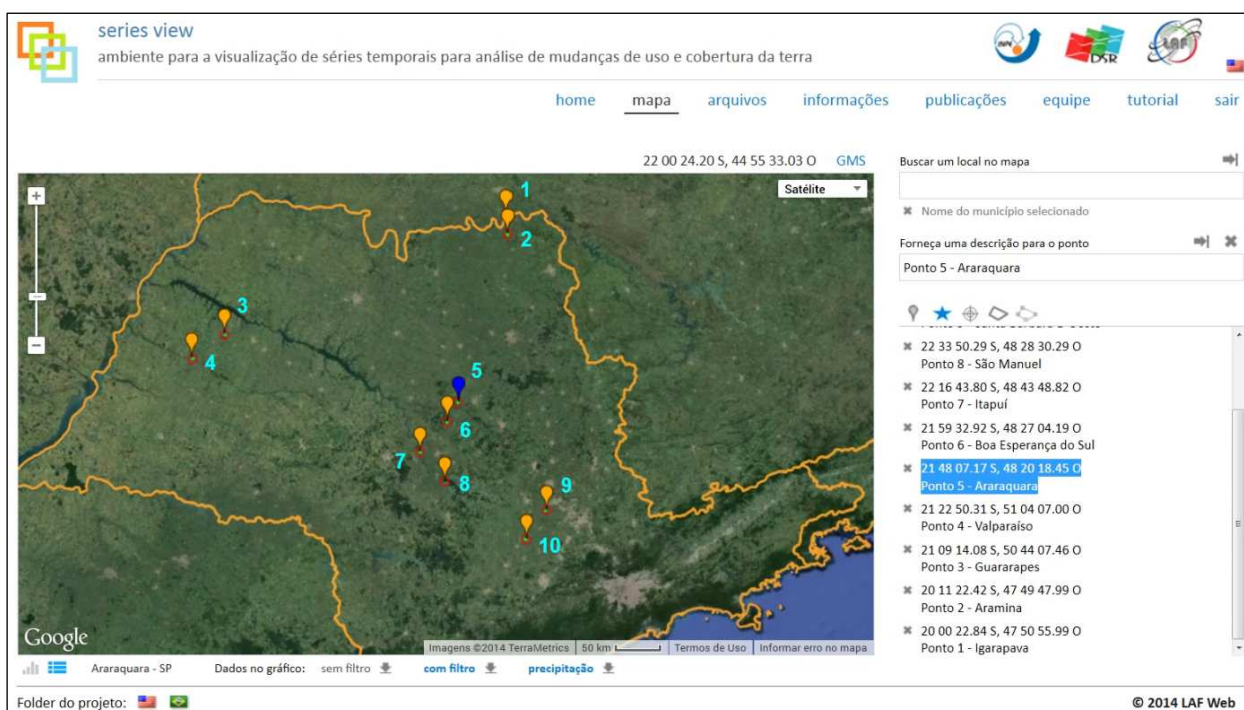


Figura 3.13: Áreas de cultivo da cana-de-açúcar de referência.
Fonte: INPE (2014).

As áreas de cultivo da cana-de-açúcar de referência (Figura 3.13) localizam-se nos seguintes municípios:

- (1) - Igarapava (20°00'S, 47°51'O);
- (2) - Aramina (20°11'S, 47°49'O);
- (3) - Guararapes (21°09'S, 50°44'O);
- (4) - Valparaíso (21°22'S, 51°04'O);
- (5) - Araraquara (21°48'S, 48°20'O);
- (6) - Boa Esperança do Sul (21°59'S, 48°27'O);
- (7) - Itapuí (22°16'S, 48°44'O);
- (8) - São Manuel (22°33'S, 48°28'O);
- (9) - Santa Bárbara D'Oeste (22°50'S, 47°26'O);
- (10) - Tietê (23°06'S, 47°38'O).

O pixel selecionado para compor a série temporal corresponde ao centroide geométrico de cada área de cultivo, referente às coordenadas geográficas que estão entre parênteses, representando assim o perfil espectro-temporal característico da respectiva área.

3.2.3. Extração de métricas fenológicas sazonais da cana-de-açúcar

As métricas fenológicas sazonais extraídas de séries temporais de imagens propiciam a identificação de ciclos da vegetação para discriminação de culturas agrícolas e, dessa forma, também podem melhorar o desempenho da classificação de imagens (LI et al., 2012; ALCANTARA et al., 2012; SENF et al., 2013).

Com o propósito de auxiliar na separabilidade entre as classes de cobertura do solo na classificação de imagens MODIS, foram utilizadas métricas descritivas da fenologia da cana-de-açúcar derivadas das séries temporais de EVI e NDVI.

As métricas fenológicas sazonais da vegetação foram extraídas com o pacote de software TIMESAT, desenvolvido por JÖNSSON e EKLUNDH (2004), que analisa séries temporais de índices de vegetação por meio das funções matemáticas de ajuste (dupla logística, Gaussiana assimétrica e Savitzky-Golay) para investigar a sazonalidade e a relação com o ciclo de desenvolvimento da vegetação ao longo do tempo.

A Figura 3.14 mostra a série temporal suavizada em vermelho que é resultado da filtragem da série temporal original em azul utilizando uma das três funções matemáticas de ajuste, em que os 11 parâmetros fenológicos são descritos a seguir.

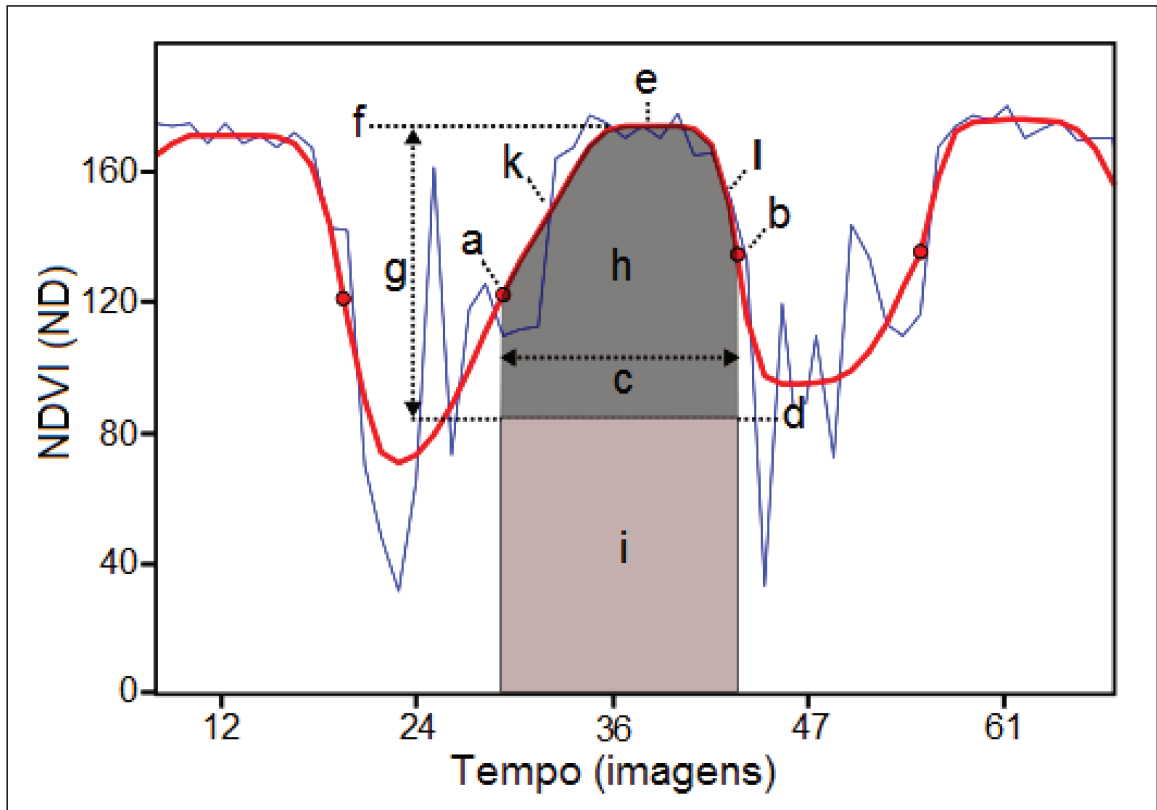


Figura 3.14: Métricas fenológicas sazonais extraídas pelo TIMESAT.
 Fonte: Adaptado de LUND UNIVERSITY e MALMÖ UNIVERSITY (2014).

- (a) Início do ciclo da vegetação: quando a função ajustada aumenta para 20% da amplitude.
- (b) Final do ciclo da vegetação: quando a função ajustada diminui para 20% da amplitude.
- (c) Comprimento do ciclo da vegetação: tempo entre o início e o final do período de crescimento.
- (d) Nível base: média dos valores mínimos antes e após o período de crescimento.
- (e) Período do pico vegetativo: média do tempo em que a função ajustada à esquerda aumentou 80% e à direita diminuiu em 80%.

(f) Valor máximo: valor máximo do índice de vegetação para a função ajustada durante o ciclo da vegetação.

(g) Amplitude sazonal: diferença entre o valor máximo e o nível base.

(h) Pequeno valor integrado: área sob a curva da função ajustada entre o início e o fim do ciclo da vegetação, a partir do nível base.

(j = h + i) Grande valor integrado: área total sob a curva da função ajustada entre o início e o final do ciclo da vegetação.

(k) Taxa de crescimento no início do ciclo da vegetação: razão da diferença entre os níveis de 20% e 80% do lado esquerdo e a correspondente diferença de tempo.

(l) Taxa de senescência no final do ciclo da vegetação: valor absoluto da razão da diferença entre os níveis de 20% e 80% do lado direito e a correspondente diferença de tempo.

A função matemática de ajuste utilizada para a filtragem de ruídos e suavização dos dados das séries temporais de EVI e NDVI foi o filtro adaptativo Savitzky-Golay (SAVITZKY e GOLAY, 1964), calculado pela Equação 3.9.

$$y_i = \sum_{i=-n}^n c_i x_{i+j} \quad (3.9)$$

onde, cada valor x_i , $i = 1, \dots, n$ é substituído pela média móvel dos valores subjacentes, com os pesos $c_i = 1/(2n + 1)$.

Segundo EKLUNDH e JÖNSSON (2012), o filtro Savitzky-Golay preserva a área e a posição média do pico de um ciclo pela média móvel e, também, a extensão e a altura pela aproximação do valor dos dados subjacentes com o método dos mínimos quadrados de uma função polinomial. Dessa forma, propiciou um melhor ajuste espectro-temporal do ciclo de desenvolvimento da cana-de-açúcar.

A Figura 3.15 apresenta a parametrização do TIMESAT utilizada na análise dos perfis temporais de EVI e NDVI, que foi determinada mediante testes realizados a partir de áreas de cultivo da cana-de-açúcar de referência de campo da Raízen, para posterior extração das métricas fenológicas.

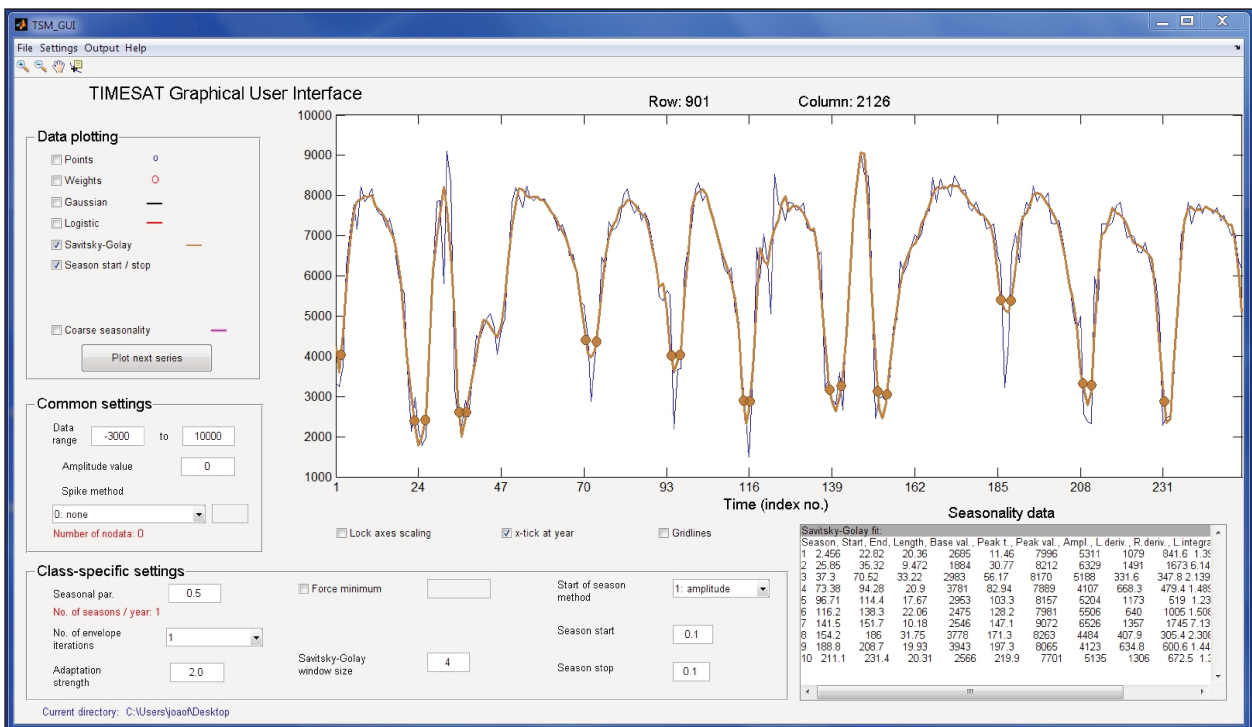


Figura 3.15: Parametrização do TIMESAT para extração das métricas fenológicas.

Os dados de EVI e NDVI estão no fator original de escala, no intervalo válido que varia de -2.000 a 10.000, sendo -3.000 para dados inválidos. O parâmetro de sazonalidade foi definido em um ciclo anual. O tamanho da janela do filtro Savitzky-Golay foi definido em quatro. Os valores de início e do final do período de crescimento foram definidos em 0,1, onde a curva ajustada em marrom atinge uma proporção da amplitude sazonal que representa o ciclo de desenvolvimento característico da cana-de-açúcar em toda a série temporal de imagens.

3.2.4. Classificação Fuzzy ARTMAP

A arquitetura ARTMAP é a versão mais avançada da família ART, projetada para realizar o mapeamento entre dois espaços de vetores, usando o treinamento supervisionado e auto-organizável com padrões binários ou analógicos. A arquitetura Fuzzy ARTMAP é uma generalização mais atraente da ARTMAP porque por um lado integra o agrupamento com treinamento supervisionado e, por outro, associa a lógica fuzzy com a Teoria da Ressonância Adaptativa, fazendo a aproximação de funções não lineares multidimensionais para classificar

vetores de entrada cujos elementos são valores dos conjuntos fuzzy, entre zero e um, que indicam o nível de adesão para cada categoria (CARPENTER et al. 1992).

A rede neural Fuzzy ARTMAP é composta por duas redes ART, uma camada de mapeamento de campo e uma camada de saída, conforme ilustrado na Figura 3.16.

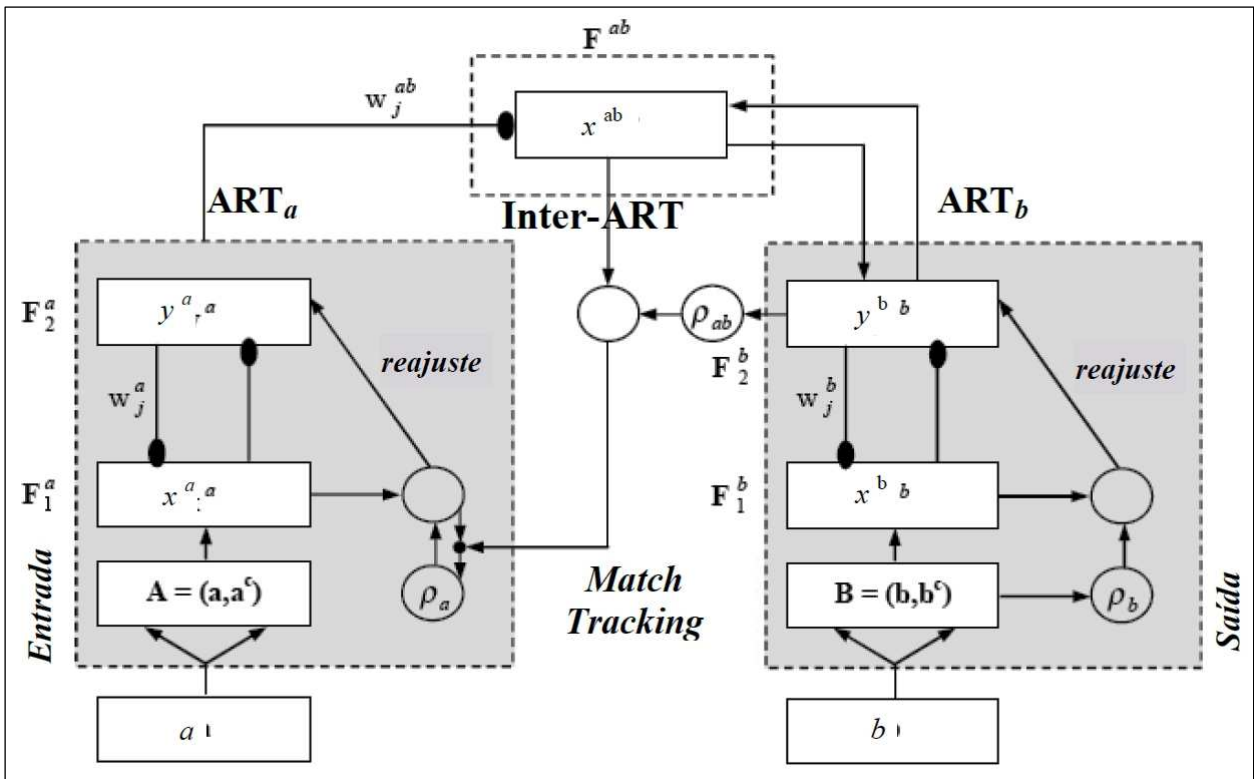


Figura 3.16: Arquitetura da rede neural Fuzzy ARTMAP.
Fonte: MALANGE (2010).

A rede ART_a contém duas camadas que recebe e processa os dados de entrada. A camada F_1 representa o vetor de entrada e tem neurônios para cada dimensão dos dados. A camada F_2 representa as categorias. A rede ART_b processa a resposta de saída desejada para o padrão apresentado à rede. A camada de mapeamento de campo $Inter-ART$ é um módulo de memória associativa que conecta os modelos ART_a e ART_b . As camadas $Inter-ART$ e ART_b consistem de m neurônios cada, onde m é a dimensão das categorias de saída, com uma conexão um-para-um neurônio entre essas camadas (CARPENTER et al. 1992).

A rede neural Fuzzy ARTMAP contém $2n$ entradas para realizar uma codificação complementar dos vetores de entrada e assim preservar a amplitude da informação. Por exemplo, se um vetor n -dimensional $\mathbf{x} \in \mathfrak{R}^n = [x_1, x_2, \dots, x_n]$ é apresentado para a rede, o número de neurônios da camada de entrada são duplicados para acomodar os $2n$ elementos, isto é, $\mathbf{x} \in \mathfrak{R}^n = [x_1, x_2, \dots, x_n, 1 - x_1, 1 - x_2, \dots, 1 - x_n]$. O número de neurônios na camada F_2 é determinado de forma dinâmica, ou seja, começa com um único neurônio e aumenta dinamicamente durante o processo de aprendizagem (MANNAN et al., 1998).

O treinamento supervisionado do Fuzzy ARTMAP é realizado com a regra de aprendizagem competitiva “o-vencedor-leva-tudo” e a estratégia de votação dos neurônios acionados com maior frequência para diferentes ordenações dos dados, de forma a maximizar a generalização das categorias de reconhecimento e minimizar os erros de predição da rede.

O módulo Inter-ART compara cada padrão de entrada com os neurônios da rede por meio do teste de ressonância, para verificar se é suficientemente semelhante a um dos padrões existentes e, assim, escolher o neurônio vencedor. A seguir são descritos resumidamente os passos do algoritmo, baseado em MALANGE (2010).

- 1 - Redimensionamento do vetor de entrada de ART_a para o intervalo $[0, 1]$.
- 2 - Codificação complementar dos vetores de entrada de ART_a e ART_b para preservar a amplitude da informação.
- 3 - Iniciar os pesos e parâmetros de ART_a , ART_b e Inter-ART. As unidades desativadas têm pesos com valor igual a um; parâmetro escolha (α), taxa de treinamento (β) e parâmetro de vigilância (ρ_a , ρ_b e ρ_{ab}) com valores no intervalo $[0, 1]$.
- 4 - Escolha da categoria de reconhecimento de ART_a e ART_b . A escolha da categoria é realizada quando no máximo um neurônio torna-se ativo. Se mais de um neurônio está ativo, é escolhido aquele com o maior valor do índice de ordenação.
- 5 - Teste de vigilância de ART_a e ART_b . A ressonância ocorre se é satisfeito o critério de vigilância. Se o critério de vigilância falhar para a categoria escolhida, ocorre a reinicialização e um novo índice é escolhido, retornando ao passo 4. O processo de busca se repete até que o índice escolhido satisfaça o teste de vigilância.
- 6 - O módulo Inter-ART controla o treinamento de uma mapa associativo de categorias de reconhecimento de ART_a para categorias de reconhecimento de ART_b e, também, realiza o teste de ressonância (*match tracking*) entre ART_a e ART_b para verificar se houve associação da

categoria de reconhecimento de entrada para a de saída. Se não houver, procura outro índice que satisfaça o teste.

7 - Adaptação dos pesos. O vetor da camada F_2 de ART_a , ART_b e Inter- ART são atualizados com os novos pesos.

8 - Repetir passos 4 a 7 para todos os pares de entrada a serem treinados.

Segundo LI (2007), o redimensionamento do vetor de entrada pode ser visto com uma analogia à abordagem de “fuzzificação” utilizando uma função de pertinência Gaussiana, visando padronizar um vetor n-dimensional $\mathbf{x} \in \mathfrak{R}^n = [x_1, x_2, \dots, x_n]$ para o intervalo entre zero e um, por meio da Equação 3.10.

$$x_{i\text{red}} = \begin{cases} \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, & \text{se } x_i \neq 0. \\ 0, & \text{se } x_i = 0. \end{cases} \quad (3.10)$$

onde, x_i é o dado de entrada, x_{\min} é valor mínimo e x_{\max} é o valor máximo do vetor de entrada, respectivamente.

A vantagem é que esse procedimento de conversão do vetor de entrada, além de conservar as propriedades de continuidade, não altera os padrões da forma e distribuição dos dados, mantendo uma correlação perfeita entre os dados originais e os redimensionados.

Um neurônio em F_2 será escolhido como o vencedor se é semelhante o suficiente com o padrão de entrada, que é determinado usando uma operação de interseção fuzzy, conforme a Equação 3.11:

$$\text{vencedor} = \arg \max_j \left(\frac{|x_i \wedge w_{ji}|}{\alpha + |w_{ji}|} \right) = \arg \max_j \frac{\sum_{i=1}^n \min(x_i, w_{ji})}{\alpha + \sum_{i=1}^n w_{ji}} \quad (3.11)$$

onde, x_i é o padrão de entrada, w_{ji} é o peso da conexão entre a camada F_1 e F_2 e a expressão entre parênteses é a função de escolha, sendo α o parâmetro de escolha.

O vencedor selecionado está sujeito ao teste de vigilância pela Equação 3.12:

$$\frac{\sum_{i=1}^n \min(x_i, w_{ji})}{\sum_{i=1}^n x_i} \geq \rho \quad (3.12)$$

onde, ρ é o parâmetro de vigilância.

Se o teste de vigilância falhar, este vencedor é considerado inválido, ocorre então a reinicialização e a busca é repetida até que um novo vencedor satisfatório seja encontrado, ou seja, se ocorre a ressonância. Se nenhum vencedor é selecionado, um novo neurônio será gerado para acomodar este padrão de entrada. Quando um vencedor é selecionado, os pesos da conexão com os neurônios de F_1 são atualizados, conforme a Equação 3.13:

$$w_{ji}^{t+1} = \beta(x_i \wedge w_{ji}^t) + (1 - \beta)w_{ji}^t \quad (3.13)$$

onde, β é a taxa de treinamento entre as camadas F_1 e F_2 .

Esse algoritmo é interessante na medida em que pode acomodar novos padrões sem esquecer os antigos, dado pelos diferentes valores utilizados da taxa de treinamento. Uma maior taxa de treinamento faz o modelo reter a memória mais recente.

O treinamento supervisionado do Fuzzy ARTMAP é realizado pelo módulo Inter-ART que executa o teste de ressonância. Quando a relação da taxa do Inter-ART for maior ou igual ao parâmetro de vigilância, o vetor de pesos w_{j2} de um neurônio selecionado entre F_2 e Inter-ART será atualizado pela Equação 3.14:

$$w_{j2_i}^{t+1} = \beta_2(o_i \wedge w_{j2_i}^t) + (o_i - \beta_2)w_{j2_i}^t \quad (3.14)$$

onde, o_i é o vetor de saída desejada de ART_b que corresponde a categoria ativa de ART_a e β_2 é a taxa de treinamento entre F_2 e Inter-ART.

Como a variabilidade espectral intraclasse pode prejudicar demasiadamente o processo de aprendizagem para o reconhecimento de padrões, neste trabalho foi utilizado o algoritmo de classificação sub-pixel da rede neural Fuzzy ARTMAP baseado na medida de Compromisso CMT para expressar o grau de compromisso de um pixel em relação a uma determinada classe.

De acordo com LI (2008), os neurônios na camada F_2 crescem dinamicamente e todos têm a oportunidade de serem acionados, não existindo neurônios redundantes. Assim sendo, cada neurônio em F_2 deve representar pelo menos um agrupamento, onde grupos com alta variabilidade serão associados com mais neurônios.

Com base nisso, o grau de compromisso que o modelo Fuzzy ARTMAP pode atribuir a um padrão de entrada pertencente a uma classe, pode ser determinado calculando-se a proporção de compromisso da classe no neurônio selecionado, por meio da Equação 3.15:

$$C_i = \frac{P_i(j)}{\sum_{i=1}^m P_i(j)} \quad (3.15)$$

onde, $P_i(j)$ é a proporção nas amostras de treinamento da classe i ($i = 1, 2, \dots, m$) do neurônio selecionado j ($j = 1, 2, \dots, n$), sendo calculada pela Equação 3.16:

$$P_i(j) = \frac{f_i(j)}{n_i} \quad (3.16)$$

onde, $f_i(j)$ é a frequência do neurônio selecionado j pelos pixels rotulados como classe i e n_i é o número total de amostras da classe i no conjunto de treinamento.

A rede neural Fuzzy ARTMAP com o algoritmo de classificação sub-pixel baseado na medida de Compromisso CMT utilizado neste trabalho foi desenvolvida por LI (2007) e faz parte do conjunto de classificadores do IDRISI.

3.2.5. Estimativa sub-pixel de área plantada de cana-de-açúcar

A partir das imagens de entrada e das amostras de treinamento, a rede neural Fuzzy ARTMAP com o algoritmo de classificação sub-pixel baseado na medida de Compromisso CMT gera um conjunto de imagens de adesão, uma para cada classe, com os graus de compromisso dos

pixels medidos entre zero e um, que permite estimar as proporções de cobertura do solo misturadas dentro de cada pixel.

Baseado nisso, a partir da imagem de adesão correspondente à classe Cana, pode-se obter a porção de área de cada pixel a partir das medidas de Compromisso. Dessa forma, foi calculada a estimativa de área plantada de cana-de-açúcar por meio de uma operação matricial de estatística zonal de soma dos pixels internos aos limites vetoriais de cada município, realizada no QGIS com a Equação 3.17:

$$A_{cana} = A_{pixel} * \sum_{i=1}^n CMT_{cana_i} \quad (3.17)$$

onde,

A_{cana} = área de cana-de-açúcar em hectares (ha);

A_{pixel} = área do pixel MODIS igual a 6,25 ha (250 m x 250 m = 62.500 m²);

CMT_{cana_i} = medida de Compromisso CMT da classe Cana por pixel;

n = número de pixels por município.

As estimativas de área plantada de cana-de-açúcar foram obtidas para os 645 municípios do Estado de São Paulo, onde o somatório da área plantada em nível municipal foi consolidado pelas 15 mesorregiões e totalizado para o nível estadual. A classificação Fuzzy ARTMAP foi executada ao longo das safras 2004/2005 a 2011/2012 para realizar o monitoramento sistemático do cultivo da cana-de-açúcar no Estado de São Paulo.

3.2.6. Análises Estatísticas

Os resultados dos modelos da classificação Fuzzy ARTMAP foram submetidos a uma série de análises estatísticas para avaliação e validação da acurácia, bem como para a determinação do nível dos erros entre as estimativas de área plantada.

A acurácia dos modelos ART gerados pela classificação Fuzzy ARTMAP para as oito safras foi avaliada a partir das amostras de teste delimitadas por amostragem com a técnica aleatória estratificada proporcional nas classes Cana e Não-Cana, tomando-se como referência os mapas temáticos do Canasat, dando origem à matriz de erros apresenta na Tabela 3.1.

Tabela 3.1: Matriz de erros.

Matriz de Erros						
		Referência		Soma Amostra	Total População	
		Cana j = 1	Não-Cana j = 2			
Classificação	Cana	i = 1	n_{11} (VP)	n_{12} (FP)	Σn_{1j}	N_1
	Não-Cana	i = 2	n_{21} (FN)	n_{22} (VN)	Σn_{2j}	N_2

Sendo que:

n_{ij} = amostras da classe de referência j ocorrendo na classe i;

Σn_{ij} = soma das amostras da classe de referência j ocorrendo na classe i;

N_i = total da população da classe i;

VP = Verdadeiros Positivos: amostras que pertencem à classe Cana e foram corretamente classificados como pertencentes à classe Cana;

FN = Falsos Negativos: amostras que pertencem à classe Cana e foram incorretamente classificados como pertencentes à classe Não-Cana;

FP = Falsos Positivos: amostras que pertencem à classe Não-Cana e foram incorretamente classificados como pertencentes à classe Cana;

VN = Verdadeiros Negativos: amostras que pertencem à classe Não-Cana e foram corretamente classificados como pertencentes à classe Não-Cana.

A sensibilidade do modelo de classificação é representada pela fração das amostras classificadas corretamente como classe Cana (VP) entre as amostras da classe de referência Cana (VP + FN).

A especificidade do modelo de classificação é representada pela fração das amostras classificadas corretamente como classe Não-Cana (VN) entre as amostras da classe de referência Não-Cana (VN + FP).

A partir da matriz de erros é obtida a matriz de proporções que expressa a proporção das amostras estimadas nas classes de referência, conforme apresentada na Tabela 3.2.

Tabela 3.2: Matriz de proporções.

		Matriz de Proporções			
		Referência		Total	
Classificação	Cana	j = 1	Não-Cana		
	Não-Cana	j = 2	p_{11}	p_{12}	Σp_{1j}
			p_{21}	p_{22}	Σp_{2j}
	Total	Σp_{i1}	Σp_{i2}	1,0	

Sendo que:

p_{ij} = proporção de amostras da classe de referência j representada na classe i;

Σp_{ij} = soma da proporção de amostras da classe de referência j ocorrendo na classe i,

onde p_{ij} é calculado pela Equação 3.18:

$$p_{ij} = \left(\frac{n_{ij}}{\sum_{j=1}^J n_{ij}} \right) * \left(\frac{N_i}{\sum_{j=1}^J N_i} \right) \quad (3.18)$$

Com base em PONTIUS e MILLONES (2011), quando a matriz de erros (Tabela 3.1) deriva de amostragem estratificada, que é o caso neste trabalho porque as classes Cana e Não-Cana são naturalmente desbalanceadas, a matriz de proporções (Tabela 3.2) apresenta estimativas imparciais e não tendenciosas das medidas estatísticas descritas a seguir, para toda a população, ou seja, para toda a região de estudo.

A Exatidão Global corresponde à concordância das proporções classificadas corretamente presentes na diagonal principal da matriz. O Índice Kappa também mede a concordância, mas incorpora os elementos das linhas e das colunas marginais referentes aos erros de classificação, além dos que se situam na diagonal principal. Porém, deve ser usado com cautela porque pode fornecer resultados enganosos, já que calcula a acurácia baseado na aleatoriedade, o que não é razoável para a construção de mapas. Em função disso, sugere-se analisar a Discordância Global formada pela componente de Quantidade, que é a classificação de proporções incorretas de pixels nas classes, e pela componente de Alocação, que se refere à distribuição espacial incorreta de pixels nas classes, as quais fornecem informações adicionais que ajudam a explicar os erros. O Erro de Comissão corresponde às proporções classificadas em

uma classe quando na realidade pertencem à outra classe. O Erro de Omissão corresponde às proporções de uma classe que foram atribuídas a outra classe.

A acurácia dos modelos CMT da classe Cana, gerados pela classificação Fuzzy ARTMAP para as safras 2004/2005 a 2011/2012, foi avaliada pelo coeficiente de correlação de Pearson (r), para analisar a relação entre as proporções de cana-de-açúcar obtidas a partir do Canasat e as medidas de Compromisso, para cada pixel da área de estudo, por meio do módulo de estatística do IDRISI.

Para validação dos modelos CMT da classe Cana, gerados pela classificação Fuzzy ARTMAP, foi empregada a análise da curva ROC baseado em PONTIUS e PARMENTIER (2014), que traça a taxa de Verdadeiros Positivos (VP) em relação à taxa de Falsos Positivos (FP) obtidas da matriz de erros (Tabela 3.1), para diferentes limiares, a partir da comparação da imagem das medidas de Compromisso com a imagem de referência do Canasat.

O valor ROC é a área sob a curva (AUC) que resulta conectando-se os pontos obtidos para cada limiar. AUC igual a um refere-se à concordância espacial perfeita entre o mapa das medidas de Compromisso e o mapa de referência do Canasat, enquanto que AUC igual 0,5 indica a tendência das medidas de Compromisso ser distribuídas aleatoriamente.

As medidas estatísticas e a curva ROC foram calculadas por meio de planilhas eletrônicas Excel, desenvolvidas por PONTIUS (2014), que oferecem uma variedade de análises e resumos estatísticos associados aos módulos de avaliação e validação da acurácia de classificação do IDRISI.

A análise estatística dos resultados consistiu na comparação pareada entre as estimativas de área plantada de cana-de-açúcar oficial do IBGE, de referência do Canasat e as obtidas pelo modelo CMT da classe Cana gerado pelo classificador Fuzzy ARTMAP, para as oito safras.

Em função disso, todos os conjuntos de dados, em nível municipal e por mesorregião, foram submetidos ao Teste de Shapiro-Wilk a 5% de significância (nível de confiança de 95%) para testar a hipótese de normalidade, onde a hipótese nula (H_0) indica que os dados possuem normalidade e a hipótese alternativa (H_1) indica que os dados não possuem normalidade.

As estimativas de área plantada de cana-de-açúcar foram comparadas por meio de análise de correlação. Com base na suposição de normalidade dos dados, a acurácia dos ajustes foi avaliada pelo coeficiente de correlação de Pearson (r). Caso contrário, optou-se então por utilizar o coeficiente de correlação não paramétrica de Spearman (r_s). Os coeficientes de

correlação foram submetidos ao teste de aleatoriedade a 5% de significância (nível de confiança de 95%) para testar a hipótese de associação entre os dados, onde a hipótese nula (H0) indica que não há correlação entre os dados e a hipótese alternativa (H1) indica que há correlação entre os dados.

Para melhor avaliar o desempenho dos ajustes foi utilizado o índice de concordância refinado de Willmott (d_r), proposto por WILLMOTT et al. (2012) em reformulação ao índice desenvolvido por WILLMOTT et al. (1985), para medir as distâncias dos valores na dispersão dos dados correlacionados em torno da reta 1:1, ou seja, indica o grau de afastamento dos valores estimados em relação aos observados, variando de -1,0 para nenhuma concordância, a +1,0 para a concordância perfeita, calculado pela Equação 3.19.

Para complementar a análise estatística dos resultados foram utilizadas medidas acerca de desvios entre os dados, em nível municipal para as safras 2004/2005 a 2011/2012, para fornecer uma melhor percepção em termos de precisão e exatidão dos ajustes entre as estimativas de área plantada, conforme descritas a seguir com base em PIMENTEL-GOMES (2000).

Erro Relativo (ER): mede a porcentagem da diferença dos valores estimados em relação aos valores observados, calculado pela Equação 3.20.

Erro Médio (*Mean Error*) (ME): mede o afastamento médio dos valores estimados em relação aos valores observados, constituindo na média dos erros, calculado pela Equação 3.21. Tende a ser pequeno na medida em que a soma dos erros negativos e positivos se compensam.

Erro Médio Absoluto (*Mean Absolute Error*) (MAE): mede também a magnitude do afastamento médio dos valores estimados em relação aos valores observados, mas não considera se um erro foi superestimado ou subestimado, calculado pela Equação 3.22.

Raiz do Erro Médio Quadrático (*Root Mean Square Error*) (RMSE): mede a dimensão dos erros entre os valores estimados e os valores observados, calculado pela Equação 3.23.

$$d_r = 1 - \frac{\sum_{i=1}^n |E_i - O_i|}{2 * \sum_{i=1}^n |O_i - \bar{O}|} \quad (3.19)$$

$$ER = \frac{(E_i - O_i)}{O_i} * 100 \quad (3.20)$$

$$ME = \frac{1}{n} * \sum_{i=1}^n (E_i - O_i) \quad (3.21)$$

$$MAE = \frac{1}{n} * \sum_{i=1}^n |E_i - O_i| \quad (3.22)$$

$$RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^n (E_i - O_i)^2} \quad (3.23)$$

onde,

E_i = área plantada de cana-de-açúcar estimada do município;

O_i = área plantada de cana-de-açúcar observada do município;

\bar{O} = área plantada média de cana-de-açúcar observada nos municípios;

n = número de municípios.

O cálculo das medidas de correlação, concordância e de desvios entre as estimativas de área plantada foi todo automatizado em planilhas eletrônicas Excel, com a integração do software Action para realizar iterativamente os testes estatísticos.

Os dados de área plantada gerados a partir do monitoramento sistemático do cultivo da cana-de-açúcar foram cruzados com os dados do ZAE no ArcGIS, para analisar a dinâmica e a tendência da expansão de área plantada de cana-de-açúcar no Estado de São Paulo.

3.2.7. Padronização do ano-safra

O ano-safra corresponde ao período em que se realizam as operações culturais necessárias à produção agrícola da cana-de-açúcar. O ano agrícola compreende o ciclo fenológico da cana-de-açúcar, que se inicia em outubro do ano anterior e termina em setembro do ano da colheita, referente ao período do monitoramento com as imagens MODIS para obtenção das estimativas de área plantada.

A disponibilização dos dados do Canasat considera o ano industrial da UNICA, em que o período começa em abril do ano da colheita e termina em março do ano seguinte. Como a colheita da cana-de-açúcar no Estado de São Paulo ocorre no período de abril a dezembro e os dados de área plantada do IBGE referem-se à área destinada à colheita no ano, a disponibilização dos dados pelo SIDRA considera o ano civil. Baseado nisso, a Tabela 3.3 apresenta os anos-safra das respectivas fontes de dados.

Tabela 3.3: Anos-safra das fontes de dados.

Safras	MODIS	Canasat	IBGE
1	2003/2004	2004/2005	2004
2	2004/2005	2005/2006	2005
3	2005/2006	2006/2007	2006
4	2006/2007	2007/2008	2007
5	2007/2008	2008/2009	2008
6	2008/2009	2009/2010	2009
7	2009/2010	2010/2011	2010
8	2010/2011	2011/2012	2011

Para fins de padronização do ano-safra foi adotada como referência a nomenclatura do Canasat, identificada como “safra” em todo o trabalho.

4. RESULTADOS E DISCUSSÕES

4.1. Reconhecimento do padrão espectro-temporal da cana-de-açúcar

A Figura 4.1 apresenta os perfis temporais do EVI2 do MODIS gerados no sistema *Series View* do INPE para cada uma das 10 áreas de cultivo da cana-de-açúcar de referência.

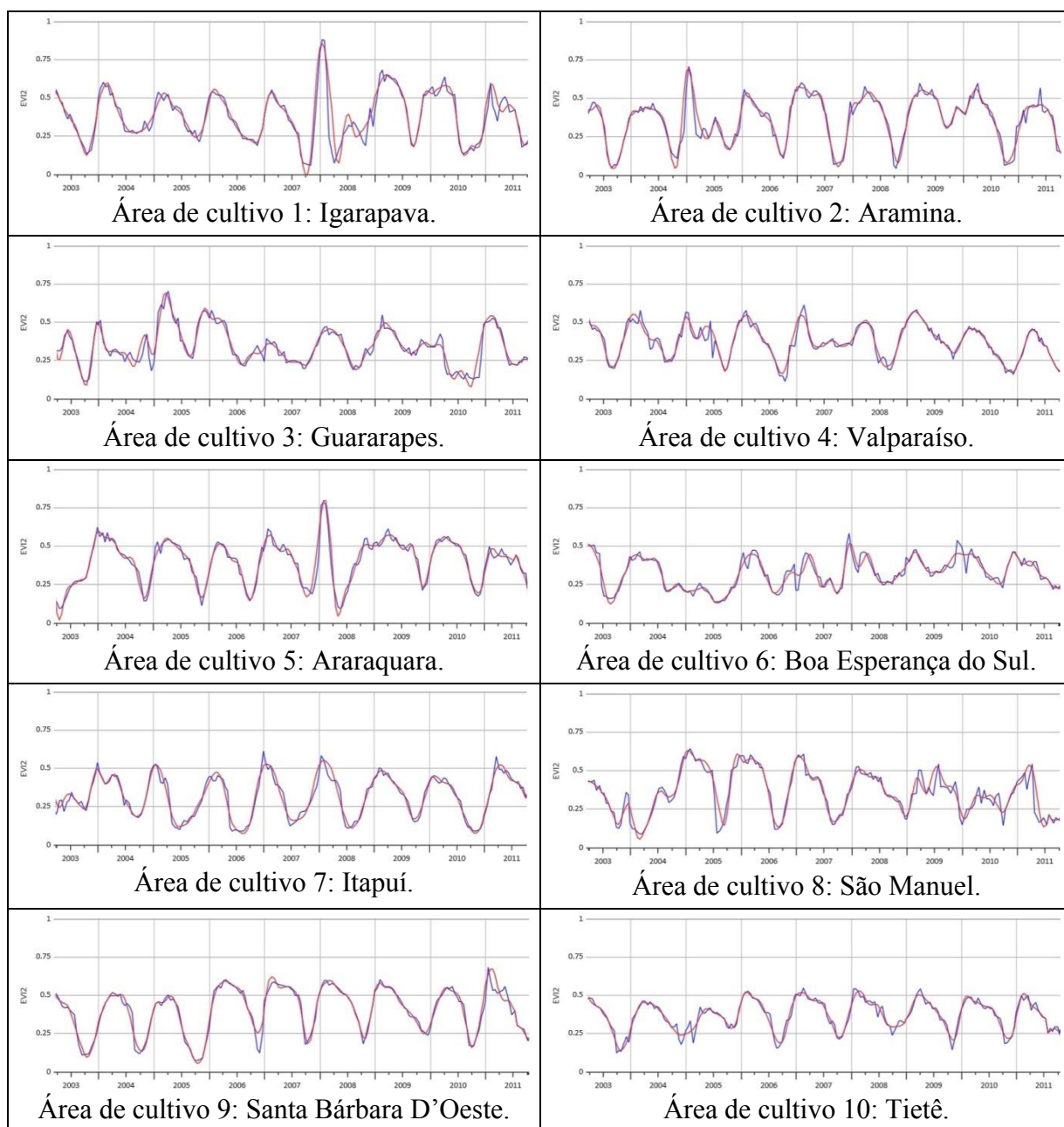


Figura 4.1: Perfis temporais do EVI2 nas áreas de cultivo da cana-de-açúcar de referência.

A curva em azul na Figura 4.1 representa os dados originais do EVI2 com a presença de ruídos observados em um pixel das imagens MODIS ao longo do tempo. A curva em vermelho é resultado da suavização utilizando a Transformada de Wavelet Daubechies 8.

De forma geral, a fenologia das áreas de cultivo da cana-de-açúcar apresentam grandes variações do EVI2, relacionada à alta dinâmica espectral da cana-de-açúcar durante o ano, onde a cultura é caracterizada por picos na série temporal, com predomínio do ciclo da cana de ano. Entre os meses de agosto e novembro a cana-de-açúcar está na fase de perfilhamento e os valores do EVI2 são baixos, aumentando fortemente a partir de dezembro na fase de crescimento, com o período de maior vigor vegetativo ocorrendo em março, decrescendo a partir de abril na fase de maturação quando se inicia a colheita.

A área 1 em Igarapava e a área 5 em Araraquara apresentam padrões semelhantes, onde se podem observar quatro ciclos de desenvolvimento de 2003 a 2007, iniciado com o plantio da cana de ano-e-meio e a rebrota da cana de ano a cada corte. A reforma do canavial ocorreu no final de 2007, identificada por um pico acentuado do EVI2 provavelmente devido ao cultivo de uma forrageira, seguida por mais três ciclos de desenvolvimento, da cana de ano-e-meio em 2008 e a rebrota da cana de ano a cada corte em 2010 e 2011.

Na área 3 em Guararapes e na área 4 em Valparaíso podem-se ver três ciclos de desenvolvimento da cana de ano de 2004 a 2006, a reforma do canavial no início de 2007, seguida por um ciclo da cana de ano-e-meio em 2008 e a rebrota da cana de ano a cada corte de 2009 a 2011. Na área 7 em Itapuí, na área 9 em Santa Bárbara D'Oeste e na área 10 em Tietê podem-se perceber oito ciclos de desenvolvimento da cana de ano de 2003 a 2011. Na área 2 em Aramina pode-se observar um ciclo da cana de ano em 2004 com a reforma do canavial no início de 2005, seguida por seis ciclos da cana de ano. Na área 8 em São Manuel pode-se notar a reforma do canavial no final de 2003, seguida por sete ciclos de desenvolvimento, da cana de ano-e-meio em 2004 e a rebrota da cana de ano a cada corte entre 2006 e 2011. A área 6 em Boa Esperança do Sul apresenta uma alta variabilidade do EVI2 causada por uma grande mistura espectral inerente da fisiologia da cana-de-açúcar, onde é possível observar um ciclo da cana de ano em 2004 com uma provável reforma do canavial em 2005, seguida por seis ciclos da cana de ano, que não estão bem caracterizados pelo perfil temporal.

A Transformada de Wavelet Daubechies 8 aplicada à série temporal do EVI2 mostrou ser uma técnica robusta, pois conseguiu eliminar os ruídos, propiciando, assim, uma melhor

captura das tendências dos ciclos de desenvolvimento da cana-de-açúcar em toda a série temporal.

Esses resultados mostraram o potencial do monitoramento temporal na detecção de padrões de fenologia da cana-de-açúcar, cujo entendimento é importante para se verificar as variações interanuais.

4.2. Decomposição das séries temporais MODIS

As séries temporais de EVI e NDVI do MODIS/TERRA para o Estado de São Paulo foram utilizadas para o monitoramento sistemático da cana-de-açúcar ao longo das safras 2004/2005 a 2011/2012. Para isso foi considerado o período que se inicia em outubro do ano anterior e termina em setembro do ano da colheita, ilustrado na Figura 4.2, que engloba, de forma geral, o ciclo fenológico característico da cana-de-açúcar de duração de 12 meses predominante na área de estudo, conforme foi verificado por meio dos perfis temporais do EVI2. Deste modo, as imagens foram obtidas do período de outubro de 2003 a setembro de 2011, sendo 23 composições de cada índice de vegetação para cada ano, totalizando 368 imagens.

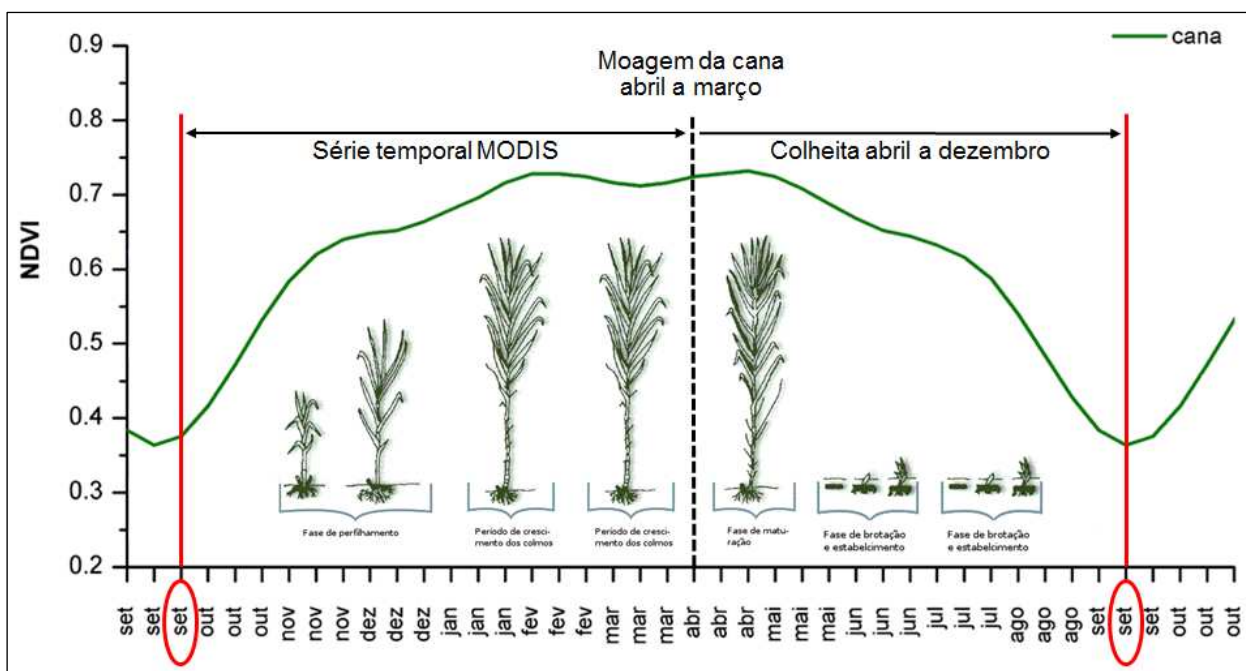


Figura 4.2: Ciclo fenológico característico da cana-de-açúcar no Estado de São Paulo.
Fonte: Adaptado de GASCHO e SHIH (1983) e FERNANDES (2009).

As séries temporais de EVI e NDVI foram decompostas pelo HANTS, separadamente para o período de seis meses, de meados de setembro a meados de abril de cada safra, sendo 14 imagens de cada índice de vegetação que cobrem o ciclo fenológico da cana-de-açúcar, desde a fase de perfilhamento até atingir o pico vegetativo na fase de crescimento, com cobertura total do solo, indo para a fase de maturação quando é iniciada a colheita.

Os termos harmônicos gerados foram a amplitude e a fase para as frequências 0, 1, 2 e 3 que correspondem a média do EVI e NDVI de cada série temporal e aos ciclos de seis meses, três meses e dois meses, respectivamente. De forma geral, esses três harmônicos são suficientes para explicar a maioria da variância das séries temporais de imagens.

A transformação HLS foi utilizada para ilustrar o ciclo fenológico da cana-de-açúcar no período de outubro a março em cada safra, baseando-se fundamentalmente nos termos harmônicos zero e primeiro que representam a média e a amplitude do ciclo de seis meses dos índices de vegetação. Para cada pixel das séries temporais decompostas de EVI e NDVI, a fase do ciclo de seis meses foi atribuída à componente matiz, a amplitude do ciclo de seis meses à componente luminosidade e a média do ciclo de seis meses à componente saturação. Na representação HLS resultante foi aplicado o realce de contraste linear, visando destacar pequenas diferenças para facilitar a interpretação visual.

As Figuras 4.3 e 4.4 mostram as representações HLS dos termos harmônicos de EVI e NDVI para o Estado de São Paulo, respectivamente, do período de outubro a março, das oito safras. Para auxiliar na compreensão das representações HLS, a legenda de cores das Figuras 4.3 e 4.4 estabelece a correspondência com a ocorrência do pico vegetativo nas diferentes safras, o que dá uma impressão rápida das mudanças nas regiões de cultivo da cana-de-açúcar no Estado de São Paulo.

Neste tipo de representação HLS, uma alta intensidade refere-se a um EVI ou NDVI constantemente elevado durante o período de seis meses, como em regiões de produção agrícola. As áreas de coloração mais clara possuem um forte efeito do ciclo de seis meses. Como a matiz indica o tempo de máximo EVI ou NDVI, a ordem das cores vermelho - amarelo - verde - ciano - azul - magenta corresponde aproximadamente à temporização da fase nos meses de outubro - novembro - dezembro - janeiro - fevereiro - março, respectivamente.

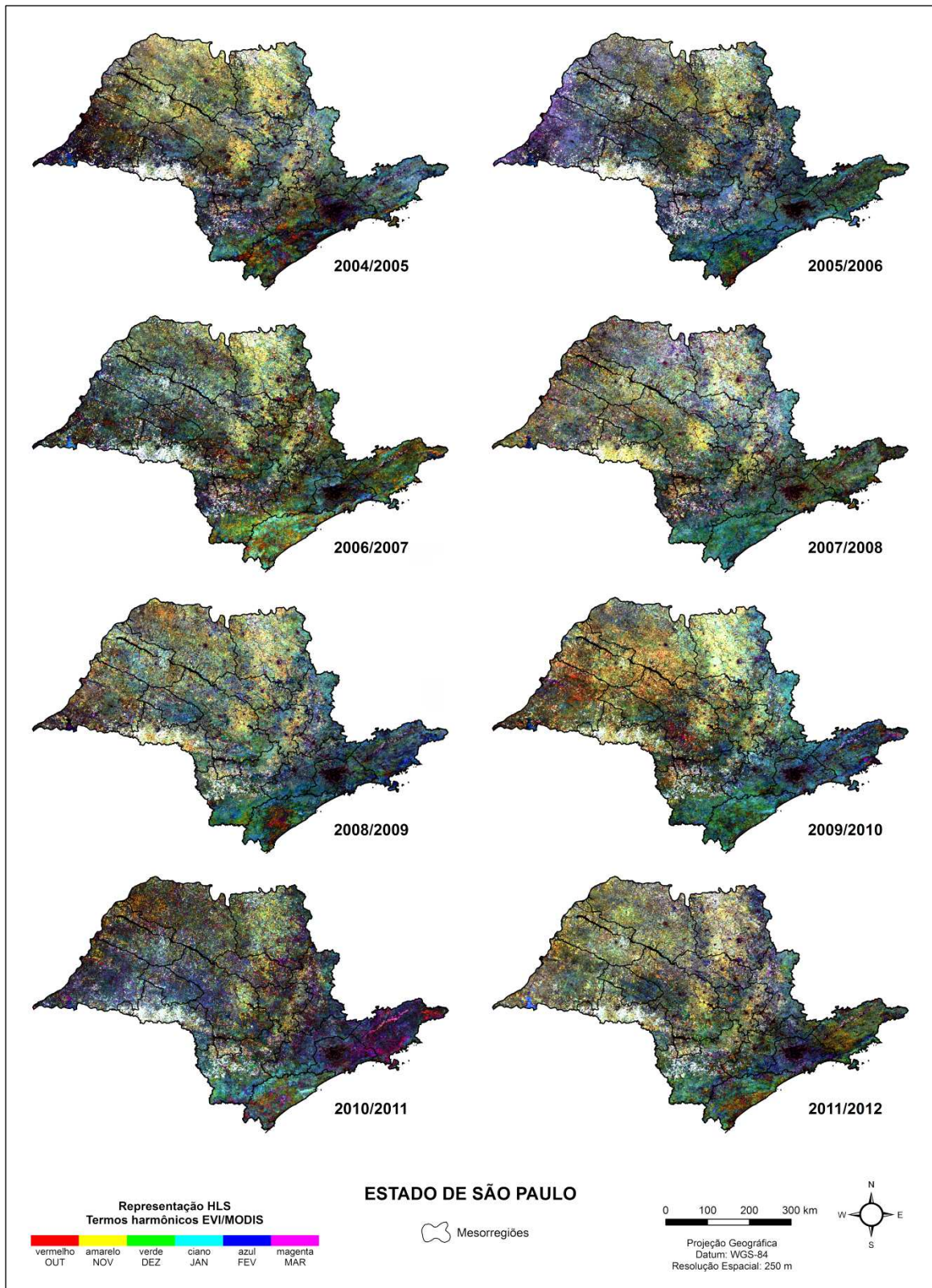


Figura 4.3: Representação HLS dos termos harmônicos de EVI do período de outubro a março das oito safras.

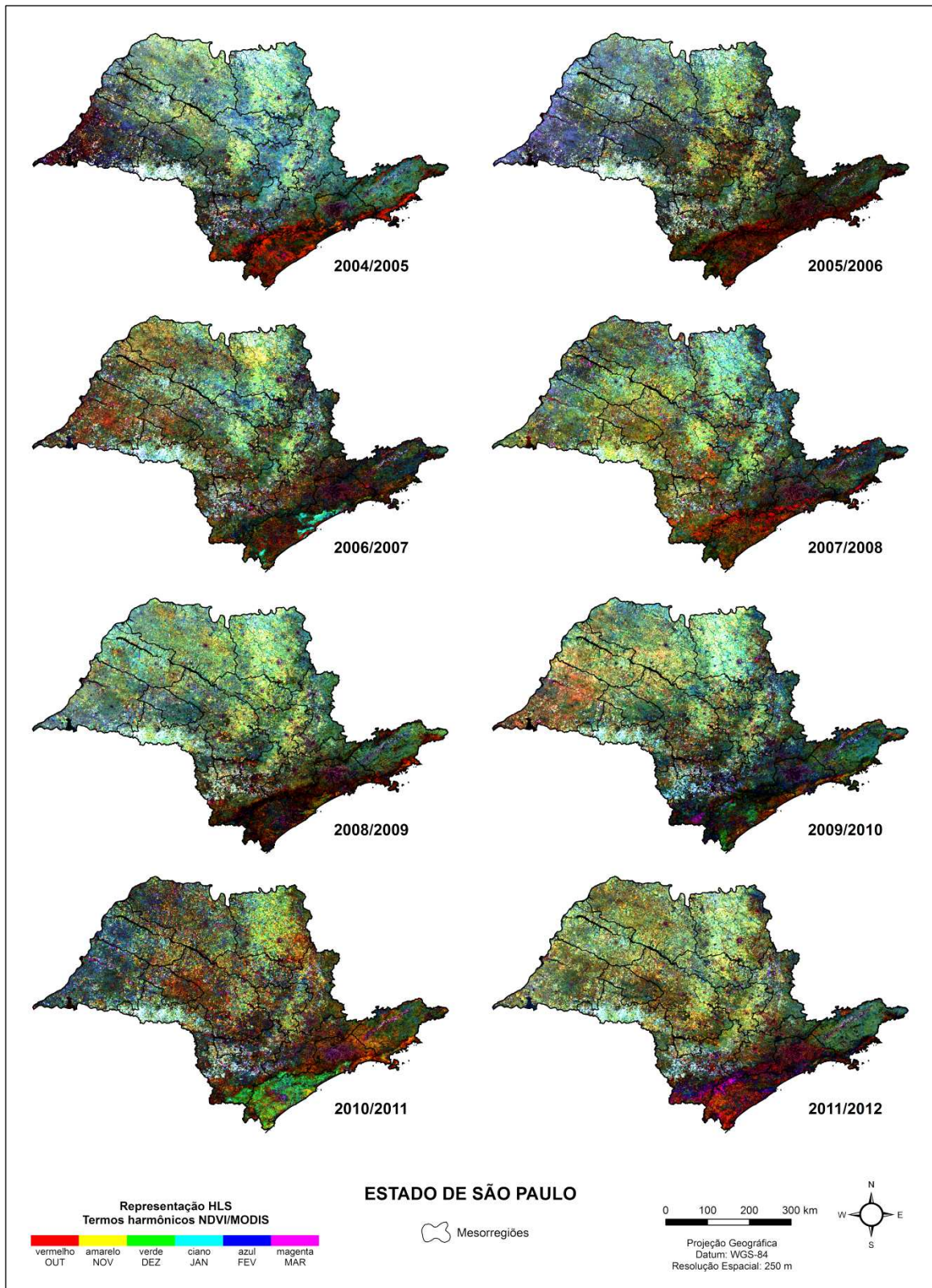


Figura 4.4: Representação HLS dos termos harmônicos de NDVI do período de outubro a março das oito safras.

A intensidade brilhante de cinza claro a branco refere-se a um EVI ou NDVI elevado durante o ciclo vegetativo que se manifesta em todos os harmônicos, como nas mesorregiões de Ribeirão Preto (Figura 4.5) e Bauru, onde prevalece a intensificação do cultivo da cana-de-açúcar.

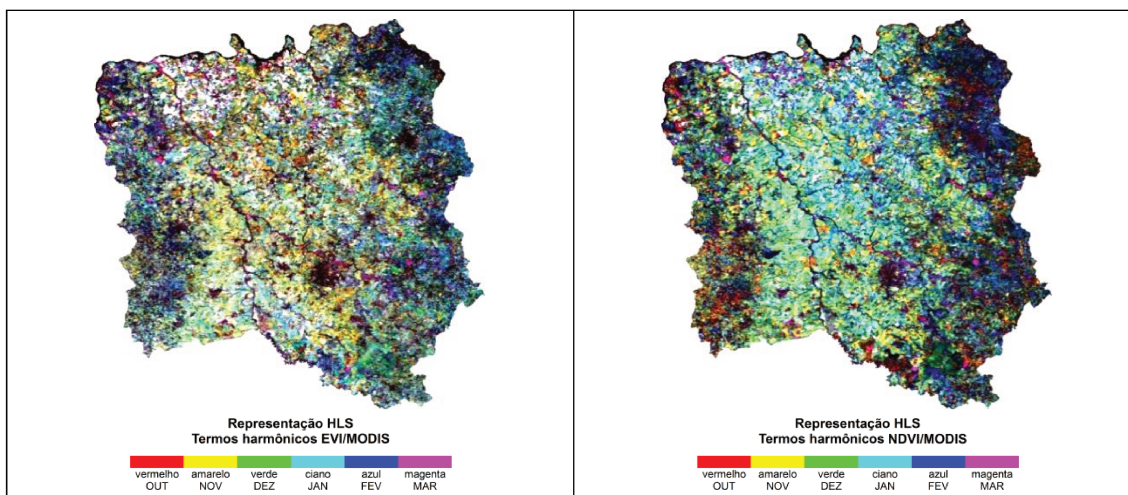


Figura 4.5: Representações HLS dos termos harmônicos de EVI e NDVI para a mesorregião de Ribeirão Preto de outubro a março da safra 2007/2008.

O cultivo da cana-de-açúcar também pode ser observado predominantemente em tons de azul, com pico vegetativo ocorrendo em fevereiro nas áreas de expansão localizadas nas mesorregiões de São José do Rio Preto (Figura 4.6) e Araçatuba.

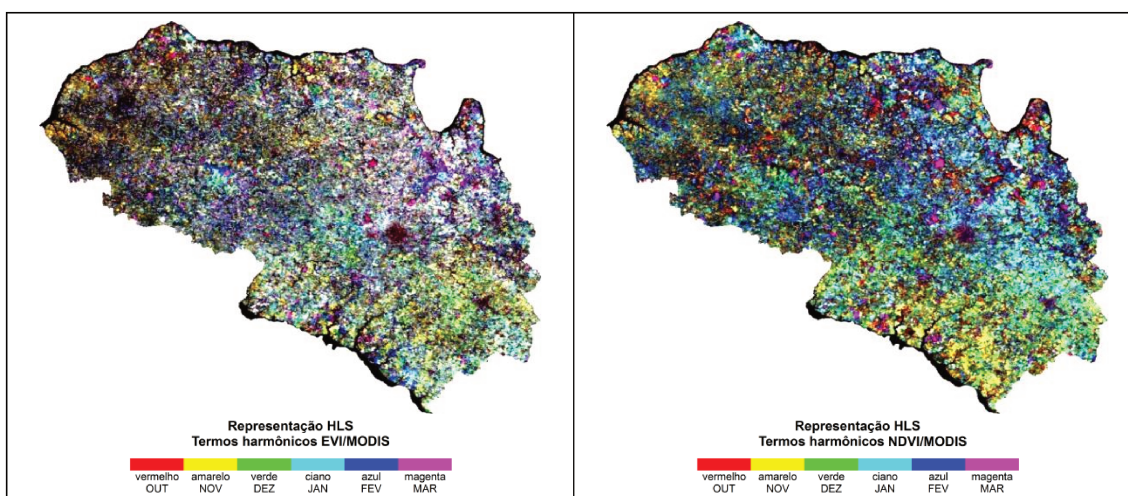


Figura 4.6: Representações HLS dos termos harmônicos de EVI e NDVI para a mesorregião de São José do Rio Preto de outubro a março da safra 2007/2008.

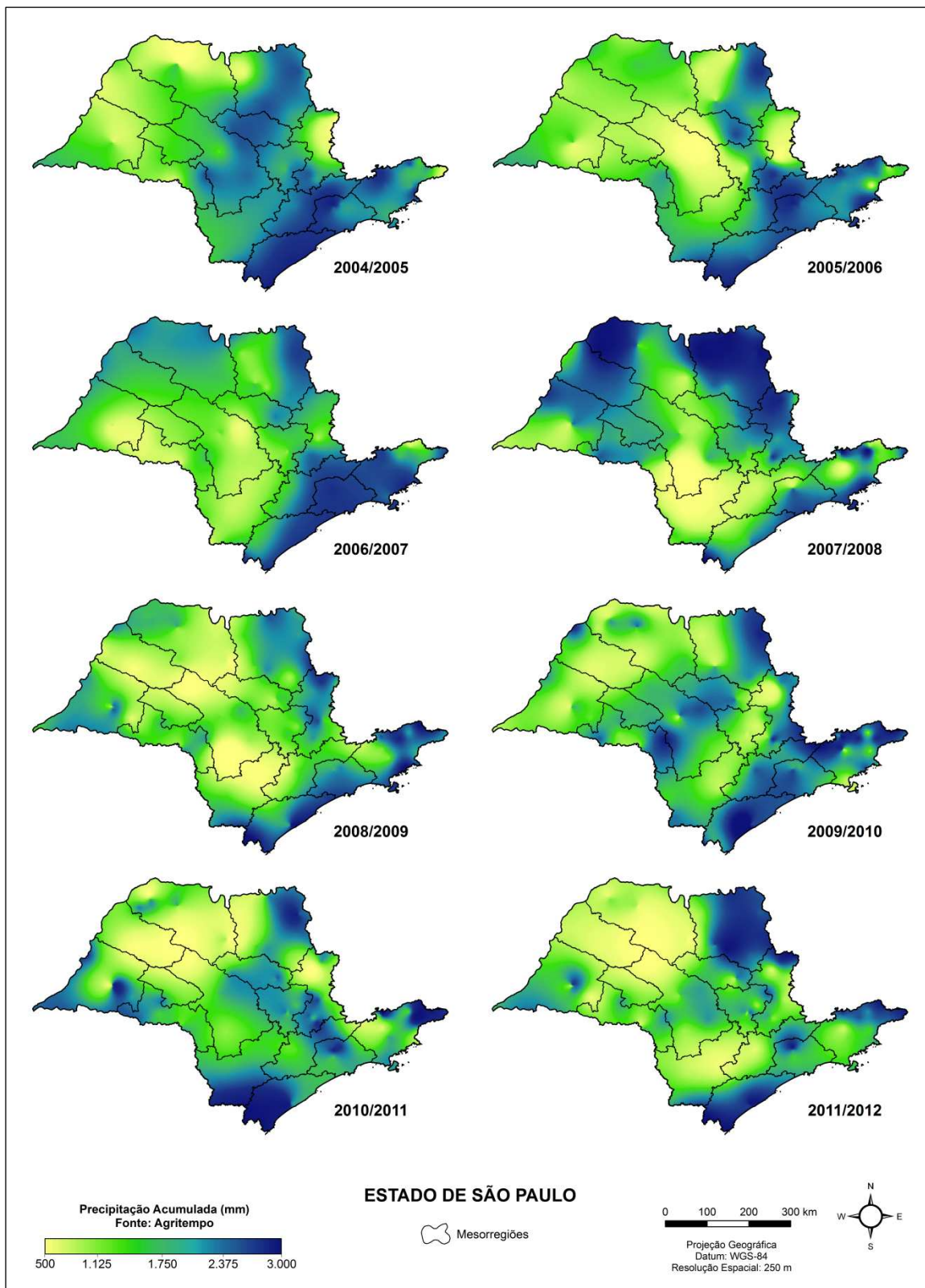


Figura 4.7: Mapas de precipitação acumulada.

A Figura 4.7 apresenta os mapas de precipitação acumulada do Estado de São Paulo, do período de outubro a setembro das safras 2004/2005 a 2011/2012, onde pode ser visualizada de forma geral a espacialização das chuvas. Na análise comparativa realizada a seguir, para caracterização da média histórica da precipitação acumulada utilizou-se a última normal climatológica de 30 anos, referente ao período de 1961 a 1990.

As safras 2004/2005, 2006/2007, 2007/2008 e 2008/2009 foram caracterizadas como sendo de acumulados de chuvas próximos às médias históricas, ou seja, de condições bastante favoráveis ao desenvolvimento da cana-de-açúcar no Estado de São Paulo, em termos de quantidade acumulada e distribuição das chuvas ao longo do ciclo vegetativo. Por essa razão que nas representações HLS do Estado de São Paulo desses respectivos anos (Figuras 4.3 e 4.4), podem-se observar áreas de cultivo da cana-de-açúcar onde predominam tons de azul e de cinza claro a branco, com pico vegetativo ocorrendo normalmente em fevereiro durante a fase de crescimento.

As safras 2010/2011 e 2011/2012 foram caracterizadas pelo acúmulo de chuva abaixo da média histórica e pela má distribuição das chuvas, acarretando baixas condições de umidade e elevação das temperaturas, que são desfavoráveis para o crescimento da cana-de-açúcar, podendo antecipar a maturação e por consequência a colheita.

Já safras 2005/2006 e 2009/2010 foram marcadas pelos acumulados de chuva bem acima da média histórica distribuída ao longo do ciclo vegetativo. O excesso de umidade causado pelo grande volume de chuvas prejudica muito a qualidade da cana-de-açúcar em termos do armazenamento de sacarose nos colmos e pode acarretar no atraso da colheita, fazendo com que cana madura permaneça no campo para ser colhida na próxima safra, o que é conhecida por cana bisada. Em função das condições climáticas dessas safras, nas representações HLS desses referidos anos (Figuras 4.3 e 4.4), prevalecem os tons de ciano, indicando que o pico vegetativo foi adiantado para janeiro.

Esses resultados mostram o potencial da análise harmônica no estudo do comportamento cíclico dos índices de vegetação, cujo entendimento é importante para evidenciar alterações no desenvolvimento da cana-de-açúcar e revelar padrões da dinâmica temporal para a tarefa de classificação de imagens.

As representações HLS dos termos harmônicos das séries temporais MODIS indicam que o cultivo da cana-de-açúcar no Estado de São Paulo possui padrões espaciais coerentes com o

ciclo vegetativo, consistente com a variabilidade das precipitações sazonais que afetam diretamente o momento de máximo EVI e NDVI.

4.3. Extração de métricas fenológicas sazonais da cana-de-açúcar

As séries temporais de EVI e NDVI foram processadas com o TIMESAT, sendo 253 imagens de cada índice de vegetação que cobrem o período de outubro de 2001 a setembro de 2012, para que os parâmetros fenológicos fossem obtidos para as oito safras.

As 11 imagens das métricas fenológicas sazonais de EVI e NDVI foram extraídas utilizando uma janela temporal suficientemente grande que considerou seis imagens antes de outubro e seis imagens depois de setembro para cada safra, de modo a abranger qualquer variação anterior ou posterior do ciclo fenológico da cana-de-açúcar. A partir disso foi calculada a média das 11 imagens das métricas fenológicas sazonais, representando, assim, imagens sínteses de toda a série temporal de EVI e NDVI.

Para ilustração, as Figuras 4.8 e 4.9 mostram as imagens sínteses das métricas fenológicas sazonais nível base e taxa de senescência da série temporal de EVI e NDVI para o Estado de São Paulo, respectivamente.

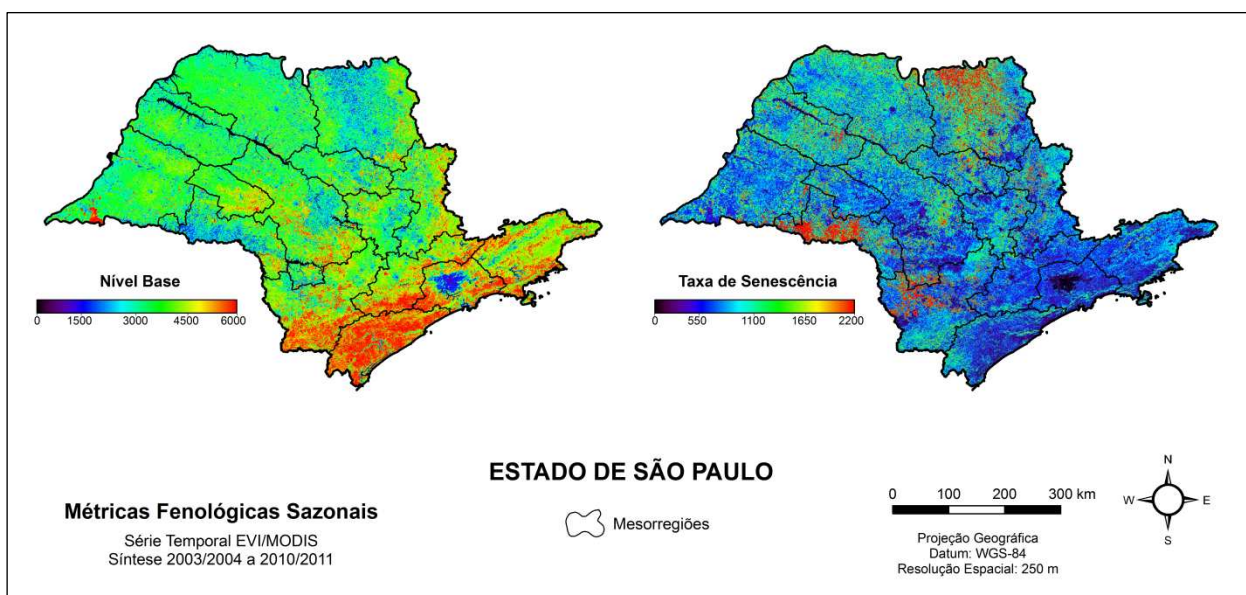


Figura 4.8: Métricas fenológicas sazonais nível base e taxa de senescência sínteses da série temporal de EVI.

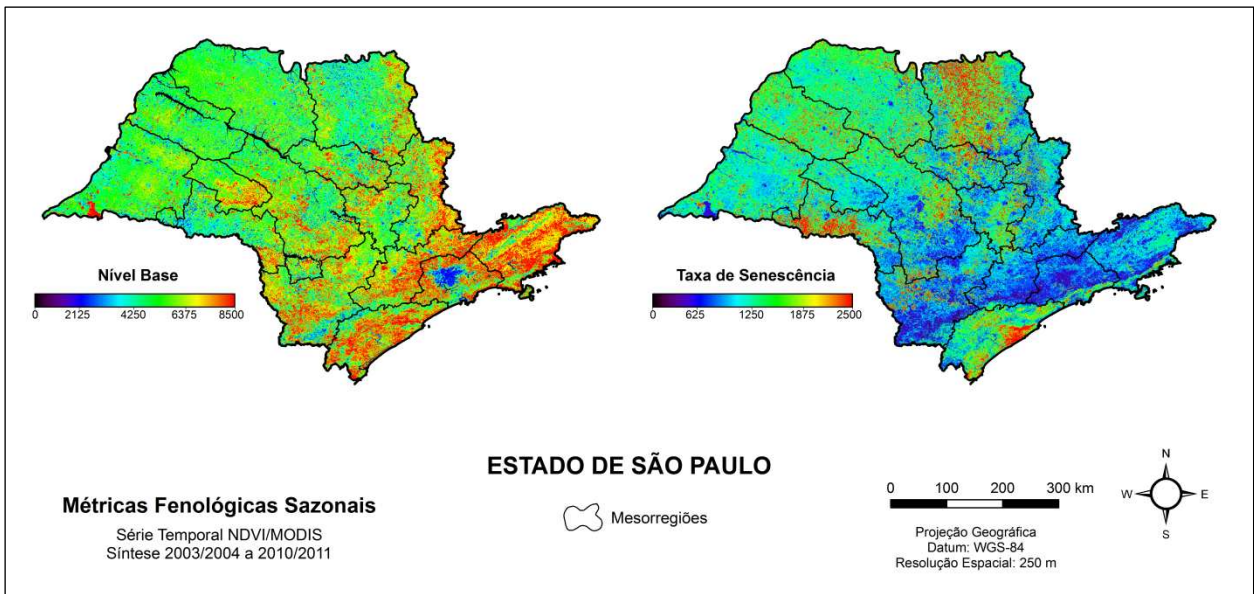


Figura 4.9: Métricas fenológicas sazonais nível base e taxa de senescência sínteses da série temporal de NDVI.

Pela análise visual das Figuras 4.8 e 4.9, pode-se observar que áreas de cultivo da cana-de-açúcar representadas em tons azulados geralmente possuem valores baixos da métrica fenológica sazonal nível base. Já o contrário acontece com a métrica fenológica sazonal taxa de senescência, que possuem valores elevados em áreas de cultivo da cana-de-açúcar representadas em tons avermelhados, o que está relacionado à forte redução da atividade fotossintética durante a fase de maturação.

Para descobrir quais métricas fenológicas geravam os modelos de classificação mais preditivos de acordo com a relevância de cada uma delas, foi realizada uma análise combinatória de importância relativa com as variáveis de entrada, tomadas uma a uma, duas a duas, três a três, etc., até as 11 métricas fenológicas de EVI e NDVI.

Com isso, verificou-se que as métricas fenológicas nível base e taxa de senescência juntas propiciaram o maior aumento na acurácia do modelo de classificação, que ficou em torno de 7,7%, mostrando serem preditores úteis no mapeamento da cana-de-açúcar. As demais métricas fenológicas foram descartadas porque não trouxeram um ganho relevante nesse nível de desempenho para a classificação.

4.4. Conversão de escala do Canasat

Os dados do Canasat do Estado de São Paulo foram processados para geração dos mapas temáticos de referência para as safras 2004/2005 a 2011/2012, por meio da conversão de escala para compatibilizar as resoluções espaciais dos diferentes sensores utilizados neste trabalho.

Primeiramente cada vetor do Canasat foi convertido da projeção policônica SAD-69 para geográfica lat/lon WGS-84 e transformado em imagem na resolução espacial de 30 m, escalonada para 0 e 255 ND. Após isso, com base em FOODY e COX (1994), cada imagem do Canasat foi degradada espacialmente com um filtro de convolução passa-baixa de dimensões 9 x 9 pixels, valor próximo à razão entre o tamanho do pixel MODIS (250 m) e o tamanho do pixel Landsat (30 m), fazendo com que a resolução espacial ficasse em 9 x 30 m igual a 270 m.

Em seguida, essa imagem foi reamostrada pelo método do vizinho mais próximo para o tamanho de 0,002245°, compatível com o pixel MODIS de 250 m. A Figura 4.10 mostra os efeitos da filtragem e da reamostragem na conversão de escala do Canasat.

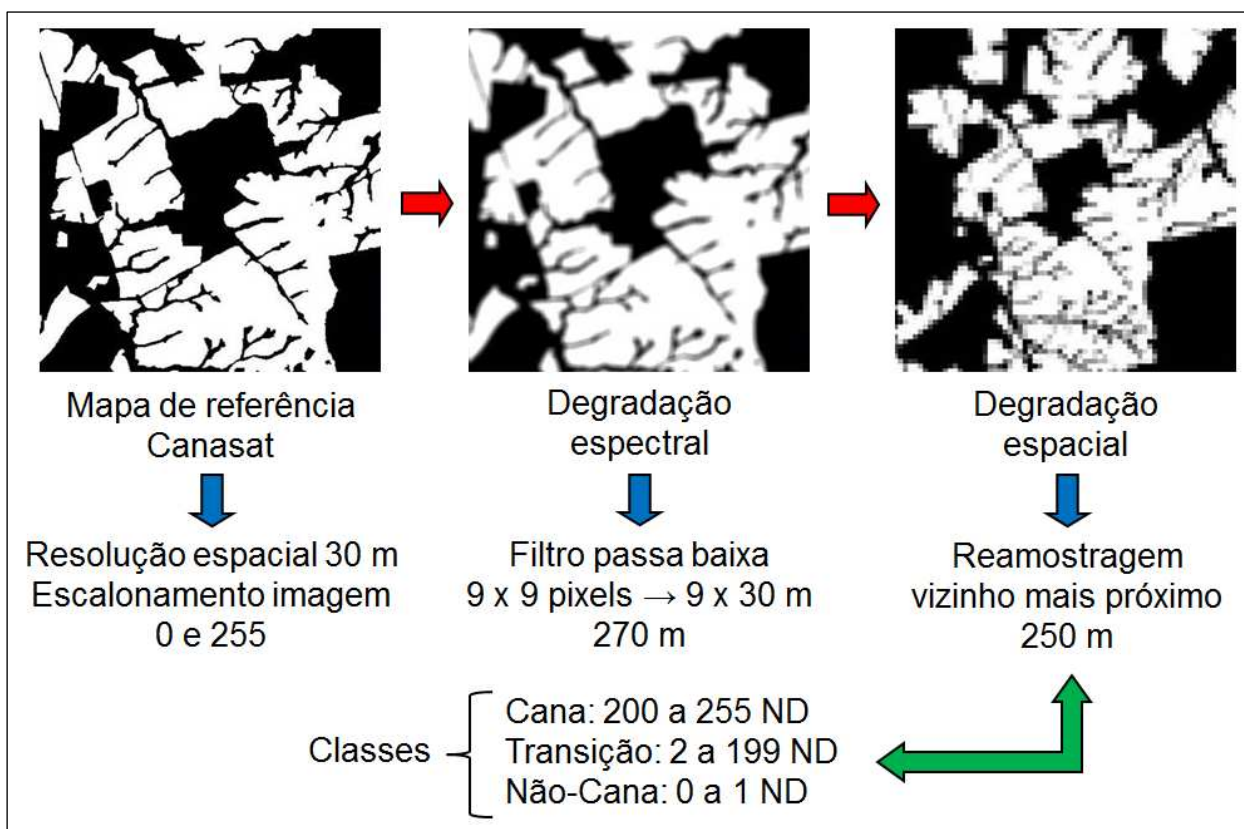


Figura 4.10: Conversão de escala do Canasat.

O resultado é uma imagem com valores entre 0 a 255 ND, onde o valor 255 representa os pixels mais homogêneos em áreas de cultivo da cana-de-açúcar e, conseqüentemente, de menor mistura espectral. Para delimitação das áreas representativas de cana-de-açúcar foram definidos os limiares de 200 a 255 ND para a classe Cana, de 2 a 199 ND para a classe Transição e 0 a 1 ND para a classe Não-Cana. Dessa forma, os pixels de borda são suprimidos da classe Cana e o universo amostral fica formado por todos os pixels que estavam totalmente contidos pelos polígonos do Canasat.

A Figura 4.11 representa o mapa de referência do Canasat convertido para a resolução espacial de 250 m referente à safra 2011/2012. Este procedimento foi realizado para as demais safras.

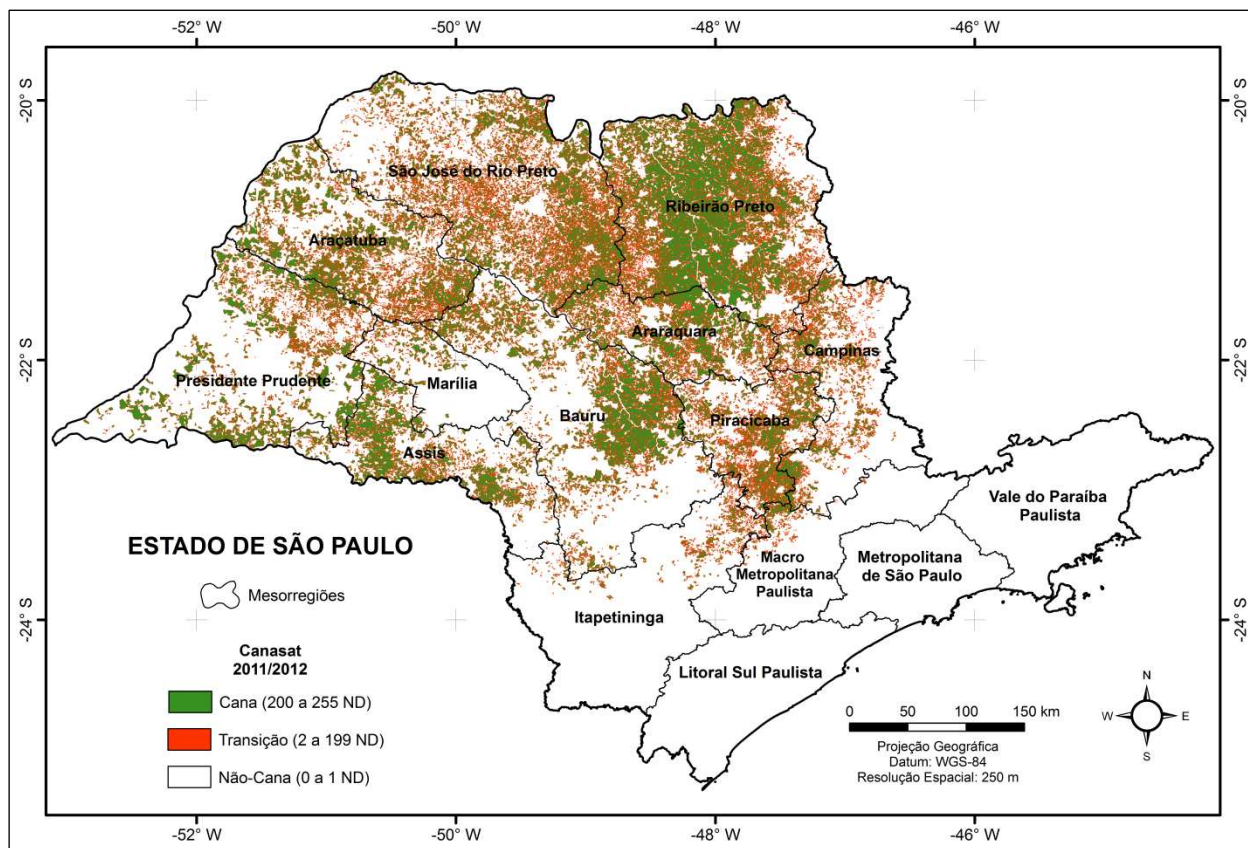


Figura 4.11: Mapa de referência do Canasat da safra 2011/2012.

Os mapas temáticos de referência gerados a partir do Canasat servem como uma base bastante confiável para seleção das amostras de treinamento de pixels homogêneos das classes Cana e Não-Cana a serem utilizadas pelo classificador e, também, para seleção das amostras de teste independentes para avaliação da acurácia do modelo de classificação ART.

Além disso, a partir do reescalonamento dos valores destes mapas temáticos de referência do Canasat para o intervalo de zero a um, são obtidos os mapas de proporções de cana-de-açúcar para cada pixel da área de estudo, que foram utilizados para a análise de correlação estatística com o modelo de classificação CMT. A Figura 4.12 apresenta o mapa de proporções de cana-de-açúcar na resolução de 250 m referente à safra 2011/2012 para o Estado de São Paulo.

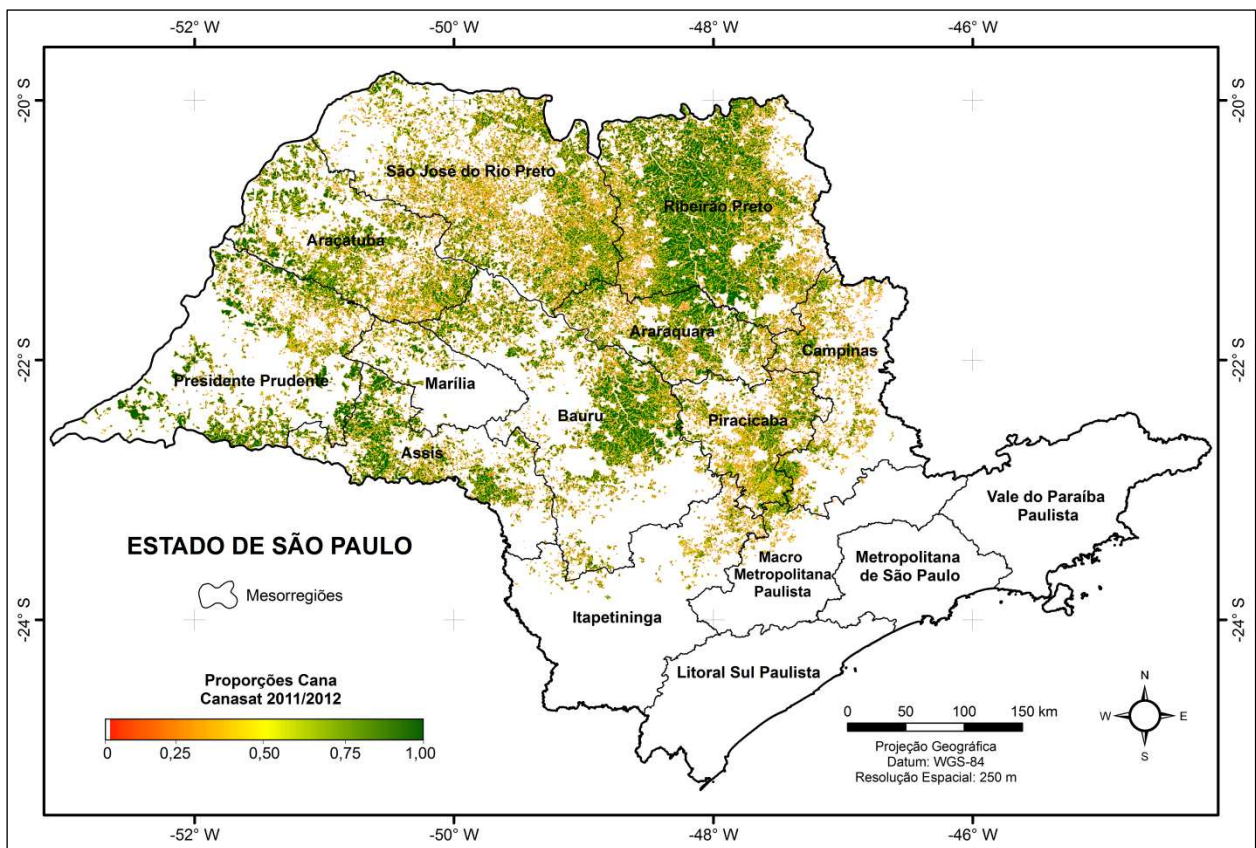


Figura 4.12: Mapa de proporções de cana-de-açúcar da safra 2011/2012.

O resultado agora é uma imagem com valores entre zero e um, onde o valor um representa os pixels mais homogêneos em áreas de cultivo da cana-de-açúcar. O gradiente de variação das cores em tons de verde, para os tons de amarelo e tons de vermelho, representa a transição da classe Cana, com altas proporções de cana-de-açúcar dentro do pixel para a classe Não-Cana, sem a presença de áreas de cultivo.

4.5. Seleção do conjunto de amostras

Para obter um erro de classificação aceitável em torno de 5%, CANTY (2010) com base em VAN NIEL et al. (2005) sugere utilizar como heurística a regra de ouro para o tamanho mínimo do conjunto de amostras, calculado pela Equação 4.1:

$$n \approx 30 \times N \times K \quad (4.1)$$

onde,

n = total de amostras de treinamento;

N = dimensão dos dados;

K = número de classes.

Sendo:

$N = 18$ imagens, sendo os sete termos harmônicos (média mais amplitude e fase para as frequências 1, 2 e 3 do EVI e NDVI) e as duas métricas fenológicas sazonais sínteses (nível base e taxa de senescência do EVI e NDVI);

$K = 2$ classes (Cana e Não-Cana).

Então:

$$n \approx 30 \times 18 \times 2 = 1.080 \times 3/2 \text{ (referente às amostras de teste)} = 1.620.$$

Portanto, o tamanho mínimo definido do conjunto de amostras deve ser de 1.620 pixels, sendo 1.080 pixels de amostras de treinamento (2/3) e 540 pixels de amostras de teste (1/3). Para produzir o conjunto de amostras em forma de polígonos homogêneos para auxiliar no processo de aprendizagem do classificador, foi realizado o procedimento ilustrado na Figura 4.13, baseado no mapa de referência do Canasat na resolução espacial de 250 m da safra 2011/2012.

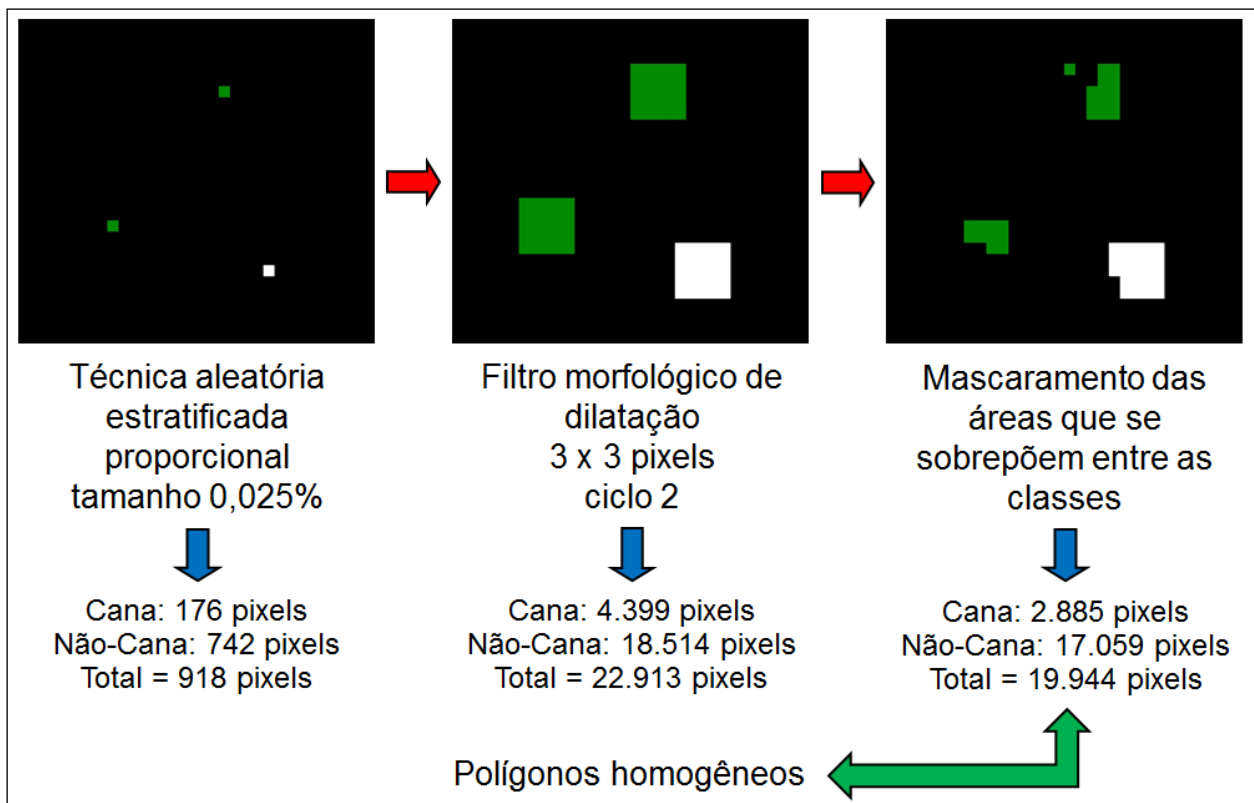


Figura 4.13: Geração de polígonos homogêneos de treinamento da safra 2011/2012.

Por naturalmente haver diferentes proporções entre as classes Cana e Não-Cana presentes nos mapas de referência do Canasat, a amostragem foi realizada com a técnica aleatória estratificada proporcional, com tamanho mínimo definido em 0,025%. Esta técnica envolve a divisão das classes em subgrupos homogêneos e, em seguida, seleciona uma amostra aleatória simples em cada subgrupo, gerando amostras que são diretamente relacionadas ao tamanho das classes. Esse processamento selecionou 176 pixels da classe Cana e 742 pixels da classe Não-Cana, totalizando 918 pixels.

Após isso foi aplicado o filtro morfológico de dilatação de dimensões 3 x 3 pixels com dois ciclos de iteração para expandir as áreas em torno do elemento central, produzindo 4.399 pixels da classe Cana e 18.514 pixels da classe Não-Cana, num total de 22.913 pixels. Ao final foi feito o mascaramento das áreas expandidas que se sobrepuseram entre as classes, que gerou 2.885 pixels da classe Cana e 17.059 pixels da classe Não-Cana, totalizando 19.944 amostras de treinamento, conforme apresentado na Figura 4.14.

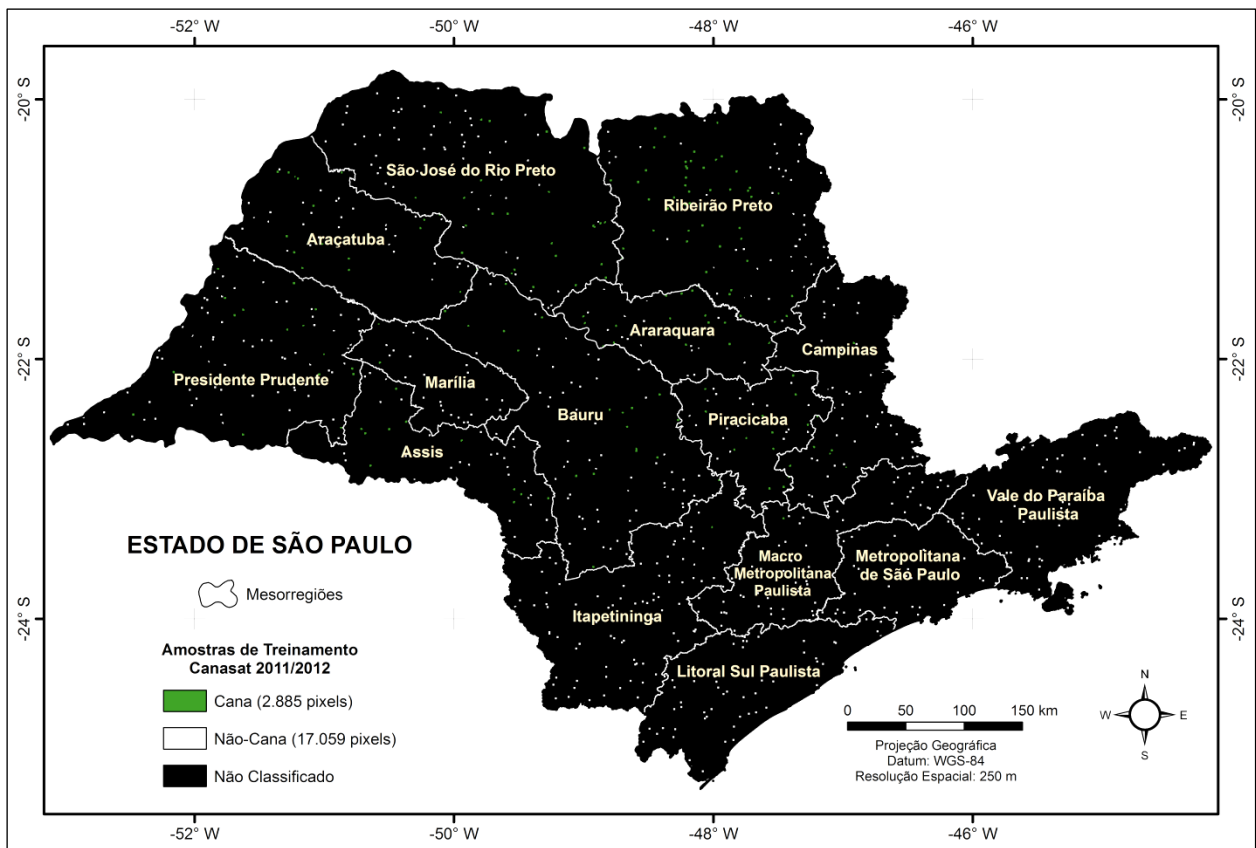


Figura 4.14: Amostras de treinamento da safra 2011/2012.

Um procedimento análogo foi realizado para geração das amostras de teste que devem ser independentes das amostras de treinamento, ou seja, a amostragem deve ocorrer sem repetição. Para isso, primeiramente foi feito o mascaramento dos pixels das amostras de treinamento no mapa de referência do Canasat da safra 2011/2012. A seguir foi realizada a amostragem aleatória estratificada proporcional, com tamanho mínimo definido em 0,0125%, que selecionou 88 pixels da classe Cana e 369 pixels da classe Não-Cana, totalizando 457 pixels.

Após isso foi aplicado o filtro morfológico de dilatação de dimensões 3 x 3 pixels com dois ciclos de iteração, produzindo 2.200 pixels da classe Cana e 9.214 pixels da classe Não-Cana, num total de 11.414 pixels. Ao final foi feito o mascaramento das áreas expandidas que se sobrepuseram entre as classes, que gerou 1.494 pixels da classe Cana e 8.401 pixels da classe Não-Cana, totalizando 9.895 amostras de teste, conforme mostrado na Figura 4.15.

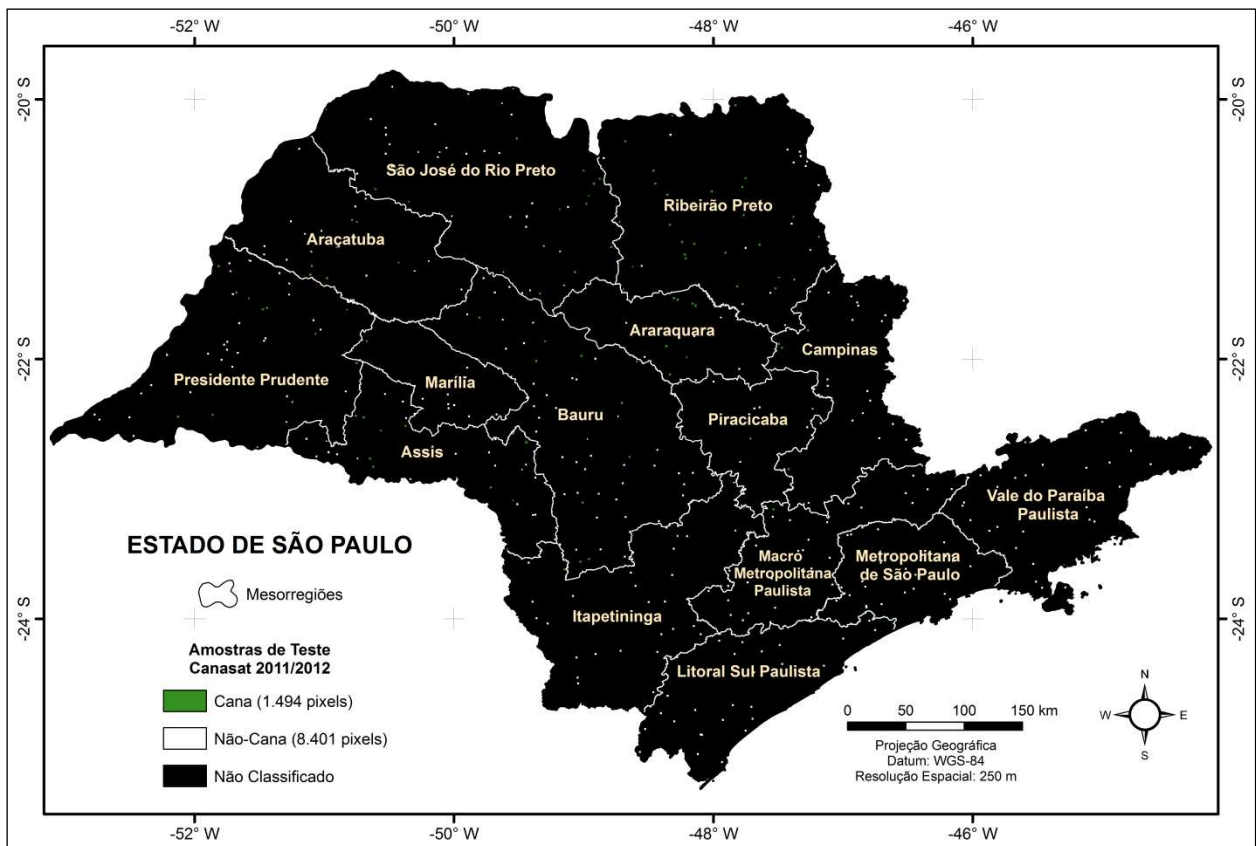


Figura 4.15: Amostras de teste da safra 2011/2012.

As amostras de treinamento (19.944 pixels) e as amostras de teste (9.895 pixels) correspondem, como planejado, a aproximadamente 2/3 e 1/3 do conjunto total de amostras (29.839 pixels), respectivamente. Além disso, atendem perfeitamente a regra do tamanho mínimo definida inicialmente.

Esse procedimento de amostragem foi repetido para os demais mapas de referência do Canasat na resolução espacial de 250 m, definindo os mesmos tamanhos mínimos para as amostras de treinamento e de teste da safra 2011/2012, já que as proporções entre as classes Cana e Não-Cana se alteram de uma safra para a outra no decorrer do tempo. Dessa forma, procurou-se manter os níveis numéricos similares dos conjuntos de amostras para todas as safras, como apresentado na Tabela 4.1.

Tabela 4.1: Conjuntos de amostras.

Canasat 250 m Safras	Amostras (pixels)					
	Treinamento			Teste		
	Cana	Não-Cana	Total	Cana	Não-Cana	Total
2004/2005	2.918	17.737	20.655	1.501	8.813	10.314
2005/2006	2.995	17.673	20.668	1.565	8.822	10.387
2006/2007	2.973	17.735	20.708	1.561	8.788	10.349
2007/2008	3.031	17.439	20.470	1.496	8.667	10.163
2008/2009	2.808	17.396	20.204	1.359	8.546	9.905
2009/2010	2.678	17.189	19.867	1.404	8.577	9.981
2010/2011	2.759	17.348	20.107	1.457	8.562	10.019
2011/2012	2.885	17.059	19.944	1.494	8.401	9.895
MÉDIA	2.881	17.447	20.328	1.480	8.647	10.127

4.6. Classificação Fuzzy ARTMAP

O conjunto de 18 imagens, sendo os sete termos harmônicos (média mais amplitude e fase para as frequências 1, 2 e 3 do EVI e NDVI) e as duas métricas fenológicas sazonais sínteses (nível base e taxa de senescência do EVI e NDVI), foram utilizadas como variáveis de entrada do classificador Fuzzy ARTMAP, para cada safra. Além do modelo de classificação ART, também é gerado um conjunto de imagens baseado na medida CMT, que expressa o grau de compromisso dos pixels em relação às classes descritas pelas amostras de treinamento, neste caso para as classes Cana e Não-Cana, já que o interesse principal é mapear a cultura da cana-de-açúcar para estimar a área plantada.

Os ajustes de todos os parâmetros do classificador Fuzzy ARTMAP foram conduzidos com base na safra 2011/2012, para depois serem reproduzidos para as demais safras. Esta última safra analisada é a mais representativa das condições de cultivo da cana-de-açúcar em termos de área plantada. Além disso, o procedimento de amostragem adotado anteriormente gerou um conjunto de amostras bem maior ao definido pela regra do tamanho mínimo, no sentido de representar a variabilidade espectral da cana-de-açúcar ao longo das oito safras.

O treinamento do Fuzzy ARTMAP no modo supervisionado foi feito com as 19.944 amostras, sendo 2.885 pixels da classe Cana e 17.059 pixels da classe Não-Cana (Tabela 4.1). A parametrização utilizada para ART_a foi o parâmetro de escolha igual a 0,01, a taxa de treinamento igual a 0,93 e o parâmetro de vigilância igual 0,94. Para ART_b utilizou-se a taxa de treinamento igual a 1,00 e o parâmetro de vigilância igual 1,00. Esses ajustes foram estabelecidos

por simulações de ensaio, como forma de alcançar a melhor a generalização e obter a maior acurácia do modelo de classificação. A camada de mapeamento que conecta ART_a a ART_b possui a mesma dimensão do número de classes de saída, neste caso dois referente às classes Cana e Não-Cana. A Figura 4.16 ilustra a interface do classificador Fuzzy ARTMAP no IDRISI.

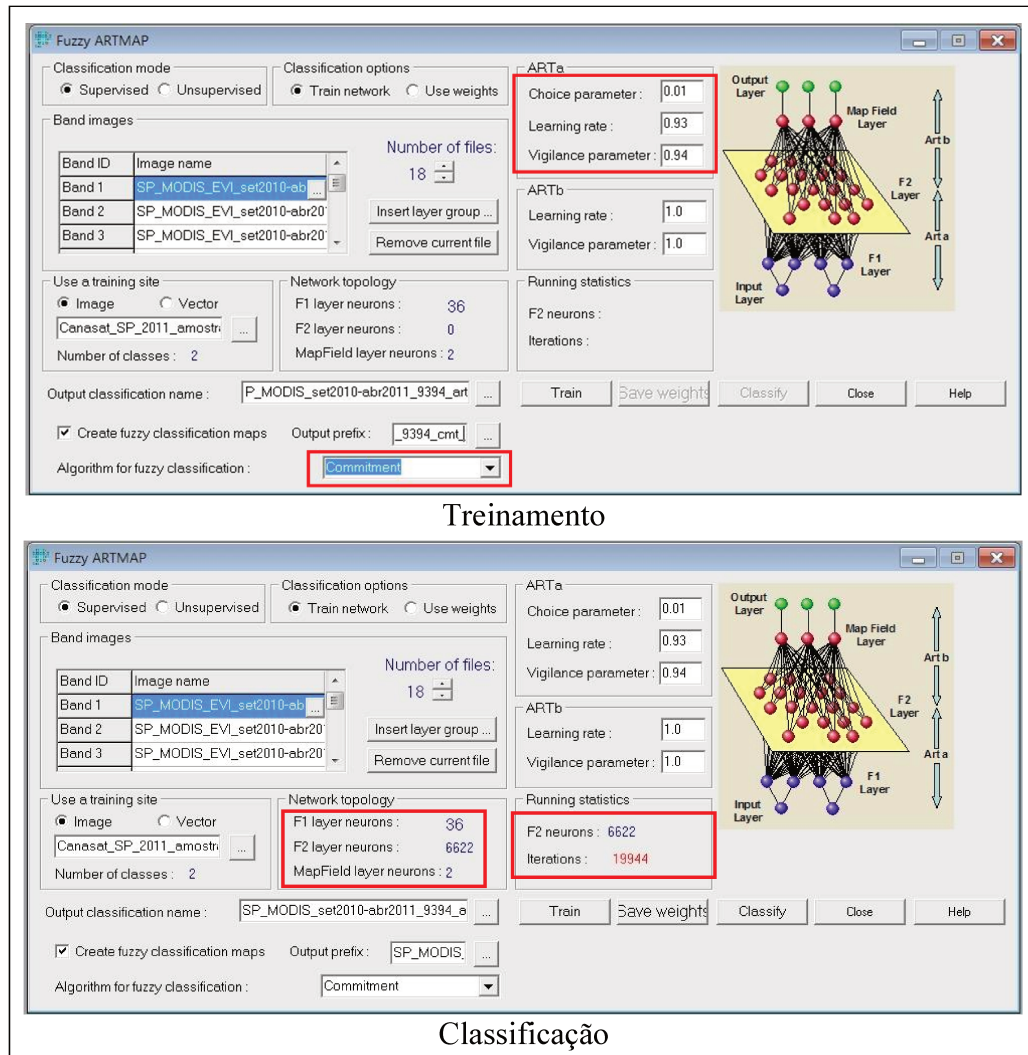


Figura 4.16: Interface do Fuzzy ARTMAP no IDRISI.

A fase de treinamento do Fuzzy ARTMAP foi realizada rapidamente em aproximadamente cinco minutos. Esse tempo muito baixo está relacionado à utilização do operador de intersecção fuzzy que é uma generalização no sentido de melhorar o desempenho da rede no que se refere, principalmente, ao tempo de execução do treinamento.

Em relação à topologia da rede, a camada F_1 contém o dobro de neurônios da dimensão das variáveis de entrada para preservar a amplitude das informações, neste caso 36, igual a 2×18 variáveis de entrada. O número de neurônios da camada F_2 cresce dinamicamente durante o processo de aprendizagem e está diretamente relacionado à parametrização de ART_a . Quanto maior a taxa de treinamento, mais neurônios na camada F_2 e mais informação de novos padrões é incorporada à rede. Quanto maior o parâmetro de vigilância, também mais neurônios são gerados na camada F_2 , o que produz imagens com pixels mais homogêneos, tendo valores mais elevados da medida CMT. O parâmetro de escolha determina o vencedor da classe de saída para determinado pixel e está sujeito ao teste de ressonância baseado no parâmetro de vigilância. A fase de classificação foi bastante lenta, devido às dimensões das variáveis de entrada, levando em torno de oito horas de processamento.

A Figura 4.17 apresenta o modelo ART resultado da classificação do Fuzzy ARTMAP para a safra 2011/2012 do Estado de São Paulo.

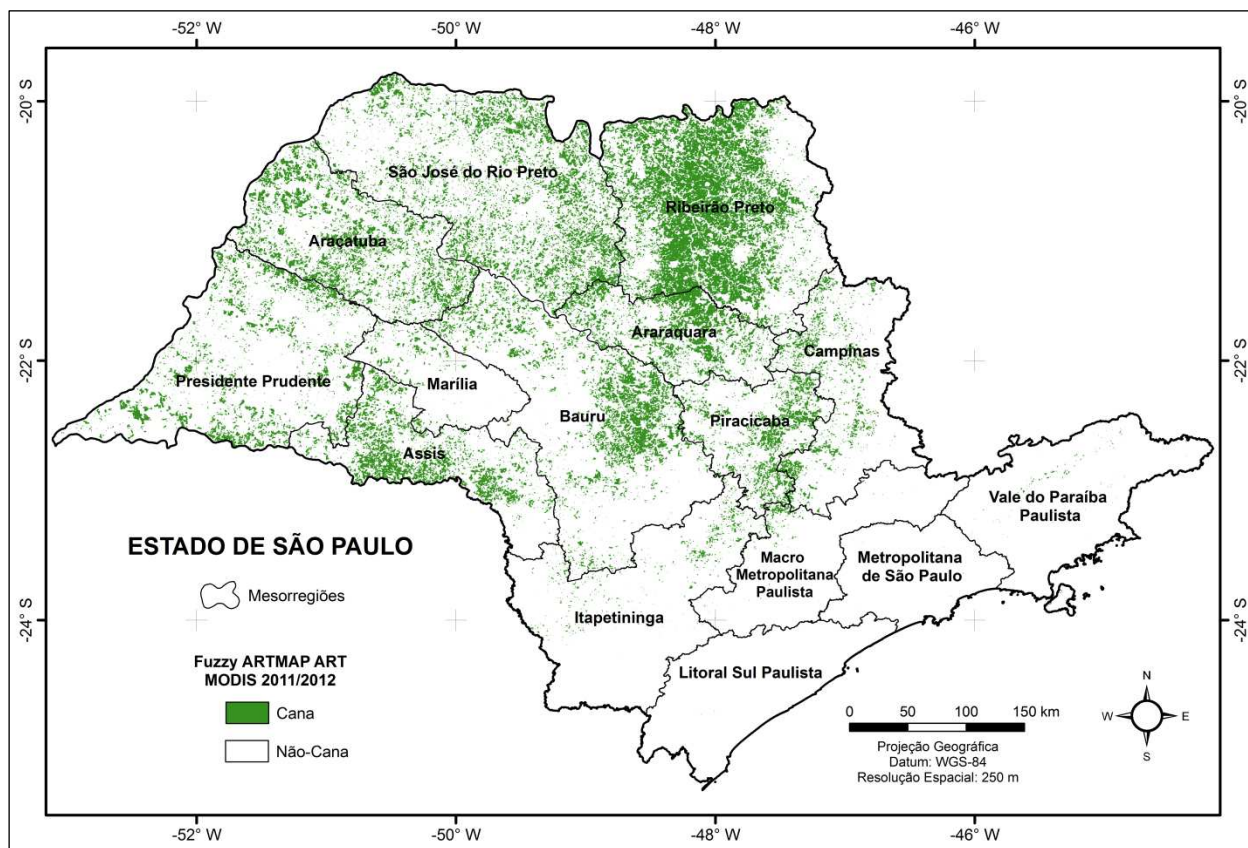


Figura 4.17: Modelo de classificação ART da safra 2011/2012.

Para avaliação da acurácia do modelo de classificação ART foram utilizadas as 9.895 amostras de teste, sendo 1.494 pixels da classe Cana e 8.401 pixels da classe Não-Cana (Tabela 4.1), gerando a matriz de erros e a matriz de proporções das quais foram obtidas as medidas estatísticas, conforme apresentado na Figura 4.18.



Figura 4.18: Medidas de acurácia do modelo de classificação ART da safra 2011/2012.

A classificação alcançou uma exatidão global de 95,15%, com índice Kappa de 0,851. Dessa forma, a discordância total foi de 4,85%, sendo 2,32% da discordância de quantidade e 2,53% da discordância de alocação. Para a classe Cana, o erro de omissão foi de 16,62% e o erro de comissão foi de 6,58%. Para a classe Não-Cana, o erro de omissão foi de 1,61% e o erro de comissão foi de 4,43%.

Ao se analisar as medidas obtidas da matriz de erros, pode-se dizer que a classificação Fuzzy ARTMAP que deu origem ao modelo de classificação ART da safra 2011/2012 foi caracterizada por uma qualidade temática muito boa, consistente com os altos valores de

exatidão, os baixos valores de discordância e, também, pelos relativamente baixos valores de omissão e comissão.

Apesar das amostras de treinamento terem sido bem delineadas, esse nível mais alto do erro de omissão da classe Cana está associado à variabilidade espectral inerente da cana-de-açúcar ao longo do seu ciclo vegetativo, o que provoca uma maior dificuldade no reconhecimento do padrão da cultura pelo classificador.

A contribuição da componente de quantidade na discordância total implica em maiores proporções incorretas de pixels designados nas classes. Isso também é refletido pelo erro de omissão da classe Cana, o que pode indicar uma tendência à subestimação de áreas de cultivo da cana-de-açúcar.

Dentre as várias simulações de ajuste do classificador Fuzzy ARTMAP, a combinação ótima para ART_a com a taxa de treinamento igual a 0,93 e o parâmetro de vigilância igual 0,94 foi a melhor generalização do modelo ART obtendo a maior acurácia.

A análise de correlação foi utilizada para avaliar a relação entre as proporções de cana-de-açúcar obtidas a partir do Canasat e as medidas CMT geradas pelo Fuzzy ARTMAP, para cada pixel da área de estudo. A Figura 4.19a apresenta o modelo CMT Cana resultado da classificação do Fuzzy ARTMAP para a safra 2011/2012. O coeficiente de correlação de Pearson (r) foi de 0,648, mostrando que o ajuste teve uma correlação positiva moderada, estatisticamente significativa a 5%.

Porém, na inspeção visual da imagem da classe Cana constatou-se que faltavam características espaciais às medidas CMT, pois o classificador distinguiu bem os pixels mais homogêneos em áreas de cultivo da cana-de-açúcar (valores mais altos) e sem a presença de áreas de cultivo (valores mais baixos), mas nem tanto os pixels de transição entre as duas classes. A partir disso, procedeu-se a pós-classificação da imagem referente à classe Cana com o propósito de aumentar o nível de correlação.

Segundo PALUBINSKAS et al. (1995), a vizinhança em torno de cada pixel pode ser utilizada para extrair informações adicionais sobre o relacionamento espacial entre os pixels de uma imagem, o que tem a capacidade de aumentar a acurácia da classificação numa abordagem de pós-classificação considerando a análise de contexto. Baseado nisso, foi então aplicada a soma de 3 x 3 pixels vizinhos mais próximos das medidas CMT, normalizada para o intervalo de zero a um, como mostrado na Figura 4.19b.

O resultado desse pós-processamento mostra a incorporação de muitas informações espaciais de contexto as medidas CMT. Dessa forma, pode-se observar em regiões da imagem com pixels mais homogêneos da cana-de-açúcar que os valores das medidas CMT continuaram elevados (próximos a um), por consequência da soma normalizada da vizinhança, criando um gradiente de transição entre as classes.

No entanto, destacou demais os valores mais baixos das medidas CMT em regiões sem cultivo da cana-de-açúcar, o que foi contornado com o mascaramento espacial pela imagem do modelo de classificação ART (Figura 4.17) gerada pelo classificador Fuzzy ARTMAP, já que alcançou um bom nível de acurácia, sobrepondo os pixels das medidas CMT em áreas sem cultivo da cana-de-açúcar. A Figura 4.19c mostra o resultado final do modelo de classificação CMT Cana da safra 2011/2012 para o Estado de São Paulo.

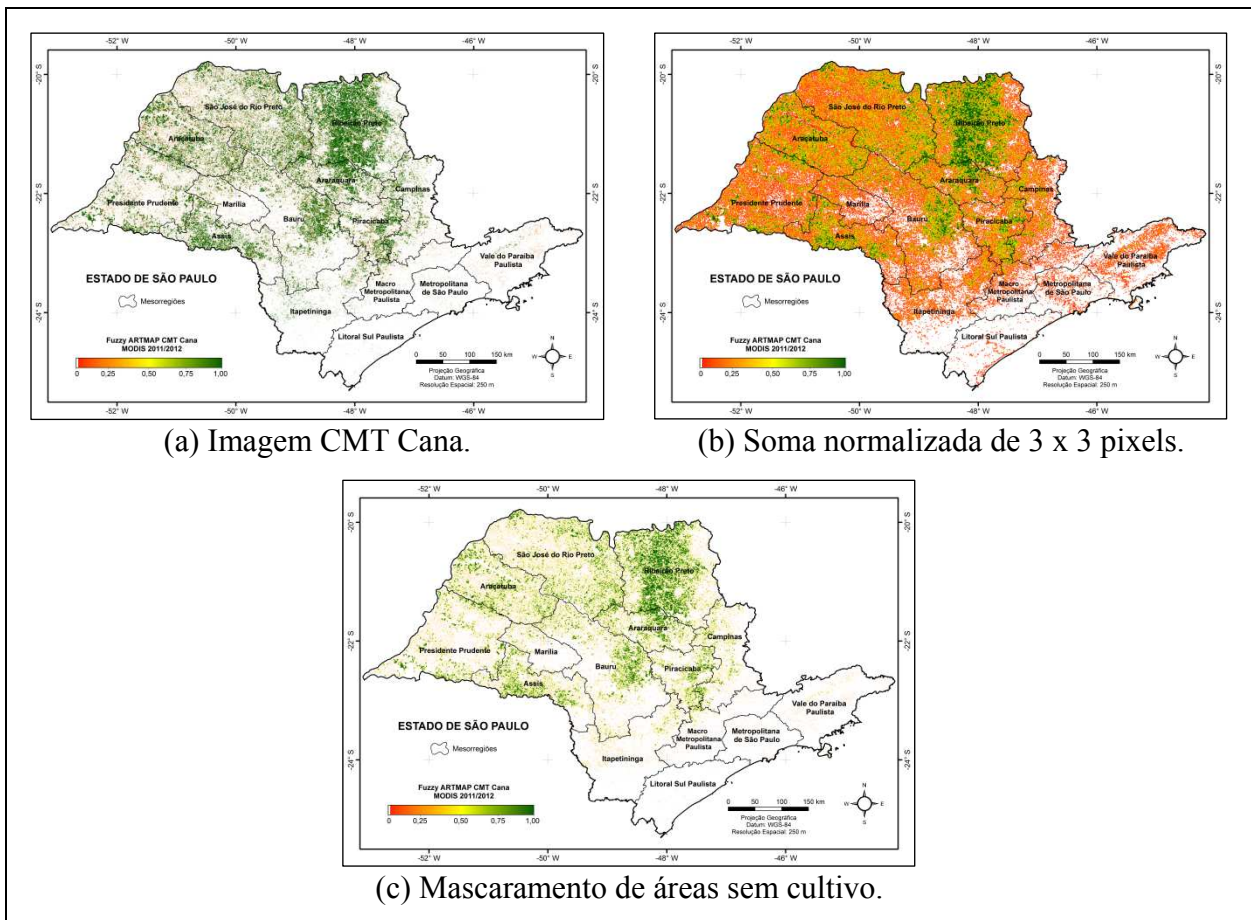


Figura 4.19: Pós-processamento do modelo de classificação CMT Cana da safra 2011/2012.

Cabe ressaltar que esse processamento de pós-classificação para incorporação de informações espaciais de contexto de vizinhança é plausível, no sentido de que a estimativa de área plantada de cana-de-açúcar baseada nas medidas CMT será calculada a partir do somatório dos pixels internos aos limites geográficos de cada município, consolidada pelas mesorregiões do Estado de São Paulo, onde cada uma congrega diversos municípios.

Para manter a correspondência da mesma base geográfica na análise de correlação, uma abordagem análoga teve que ser aplicada na imagem proporções Cana, mostrada na Figura 4.20a. Assim sendo, foi realizada a soma de 3 x 3 pixels vizinhos mais próximos das proporções Cana, normalizada para o intervalo de zero a um, apresentada na Figura 4.20b. Em seguida, também foi aplicado o mascaramento espacial pela imagem do modelo ART (Figura 4.17), para sobrepor os pixels das proporções Cana em áreas sem cultivo da cana-de-açúcar. A Figura 4.20c mostra o resultado final da imagem proporções Cana da safra 2011/2012 para o Estado de São Paulo.

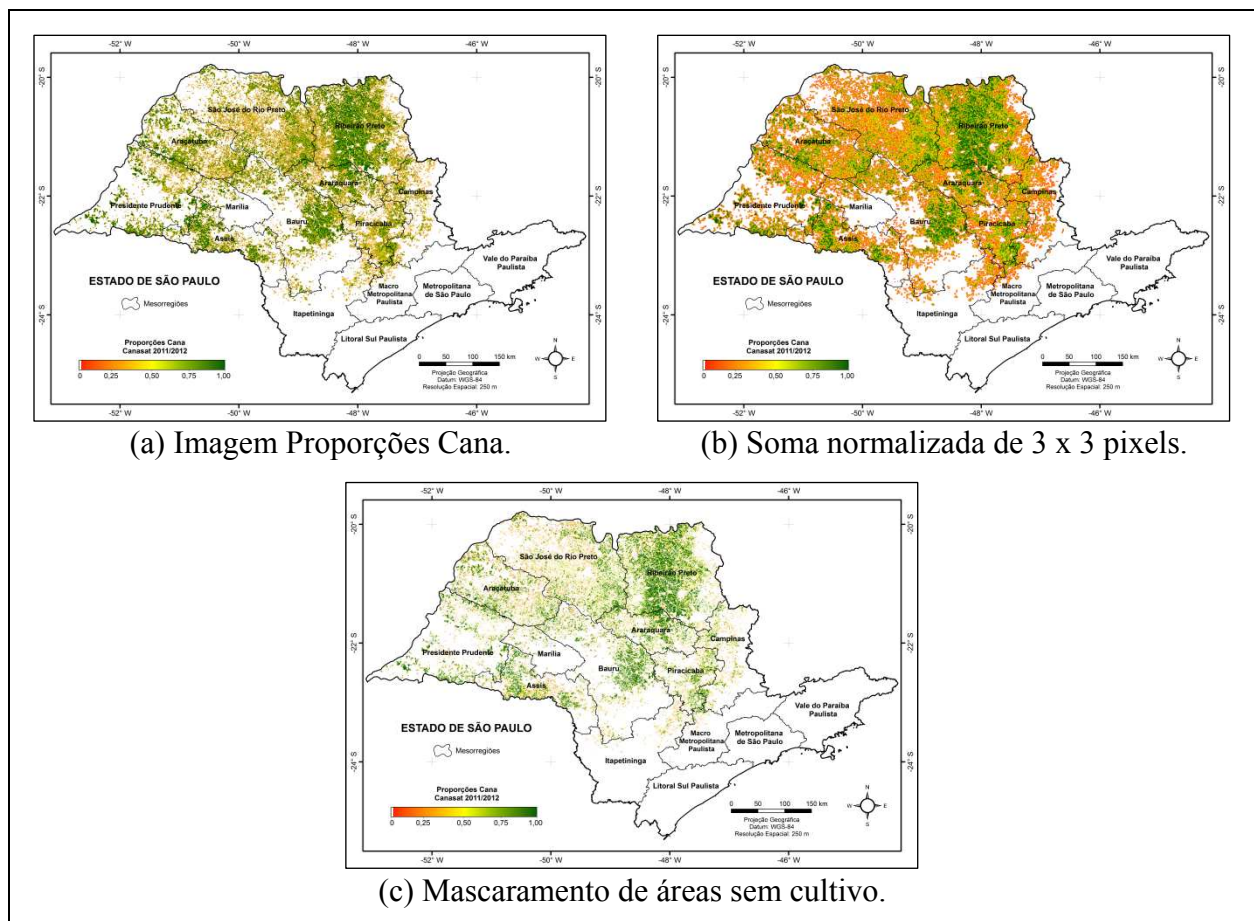


Figura 4.20: Pós-processamento da imagem proporções Cana da safra 2011/2012.

Após esse procedimento, o coeficiente de correlação de Pearson (r) aumentou substancialmente para 0,875 com a soma normalizada de 3 x 3 pixels vizinhos mais próximos e, com o mascaramento das áreas sem cultivo da cana-de-açúcar subiu para 0,943, o que caracteriza um ajuste com correlação positiva forte, estatisticamente significativa a 5%, conforme representado pelo gráfico Figura 4.21.

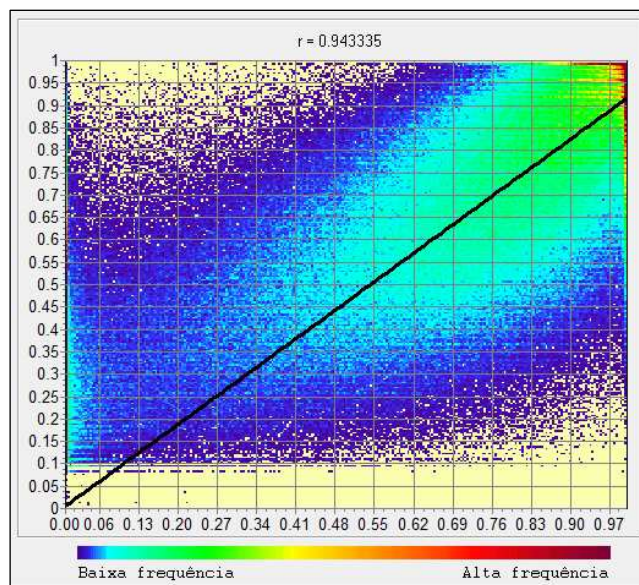


Figura 4.21: Relação entre as proporções Cana e as medidas CMT Cana após o procedimento de pós-processamento.

Para validação do modelo CMT Cana gerado pela classificação Fuzzy ARTMAP, foi empregada a análise da curva ROC. Para isso, a imagem das medidas CMT da safra 2011/2012 (Figura 4.22a), que expressa o grau de compromisso dos pixels a classe Cana, foi comparada com a imagem booleana de referência Cana e Não-Cana da safra 2011/2012 (Figura 4.22b), que aponta a localização geográfica das áreas de cultivo da cana-de-açúcar.

A curva ROC traça a taxa de verdadeiros pela taxa de falsos positivos do modelo de classificação CMT Cana, calculadas para cada limiar definido. A Figura 4.22c exibe o gráfico ROC considerando 100 limiares de intervalos iguais para calcular a área sob a curva (AUC). A partir da análise da curva ROC, pode-se observar que à medida que se aumenta o limiar de corte, o poder discriminante do modelo CMT Cana também aumenta, mantendo uma boa relação entre a sensibilidade e a especificidade do classificador, o que confere uma maior área sob a curva e, consequentemente, obtém-se um melhor desempenho.

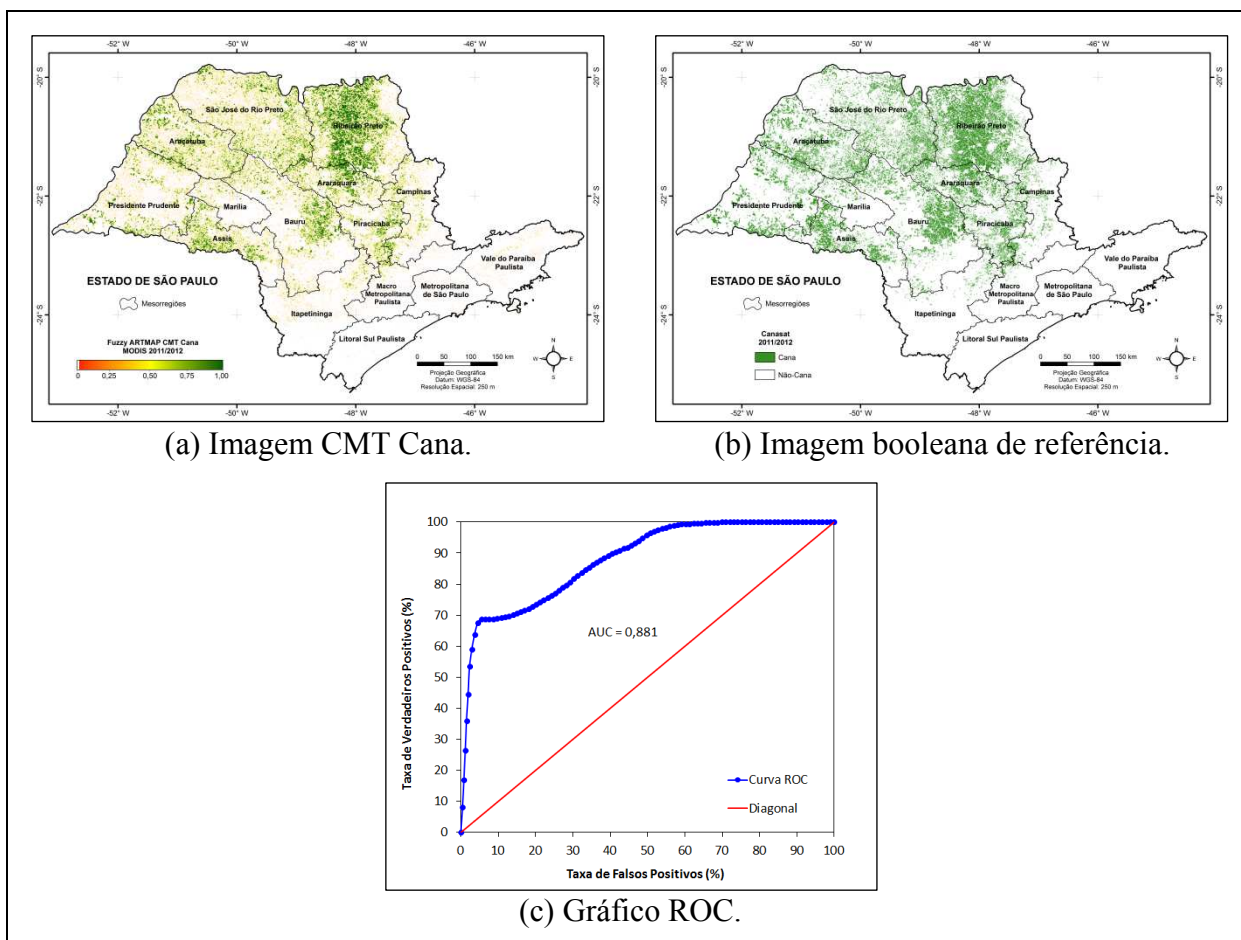


Figura 4.22: Curva ROC entre o modelo de classificação CMT Cane e a imagem booleana de referência da safra 2011/2012.

A AUC do modelo de classificação CMT Cane foi de 0,881 que é um valor extremamente alto, caso as medidas CMT tivessem a tendência de serem distribuídas uniformemente sobre a diagonal ($AUC = 0,50$) (PONTIUS e PARMENTIER, 2014), demonstrando que é válido mapear a cultura da cana-de-açúcar baseado nas medidas CMT.

Os procedimentos para avaliação da acurácia do modelo de classificação ART, bem como a análise de correlação e validação pela curva ROC do modelo CMT Cane, gerados pelo classificador Fuzzy ARTMAP, foram reproduzidos para as demais safras em estudo.

A Figura 4.23 mostra os modelos de classificação ART e a Figura 4.24 os modelos de classificação CMT Cane das safras 2004/2005 a 2011/2012 para o Estado de São Paulo.

A Figura 4.25 apresenta as imagens proporções Cane e a Figura 4.26 as imagens booleanas de referência Cane e Não-Cane das safras 2004/2005 a 2011/2012 para o Estado de São Paulo.

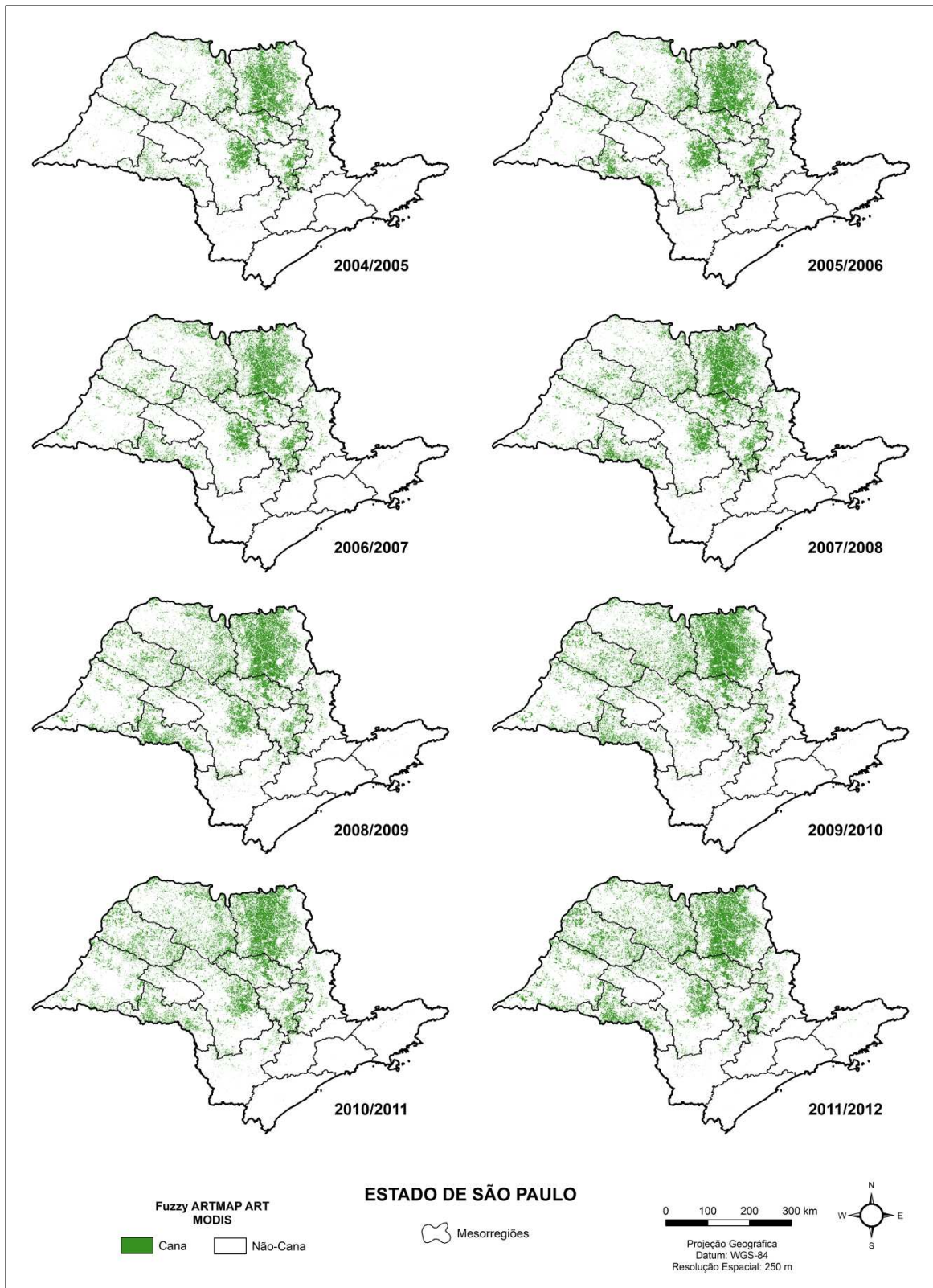


Figura 4.23: Modelos de classificação ART.

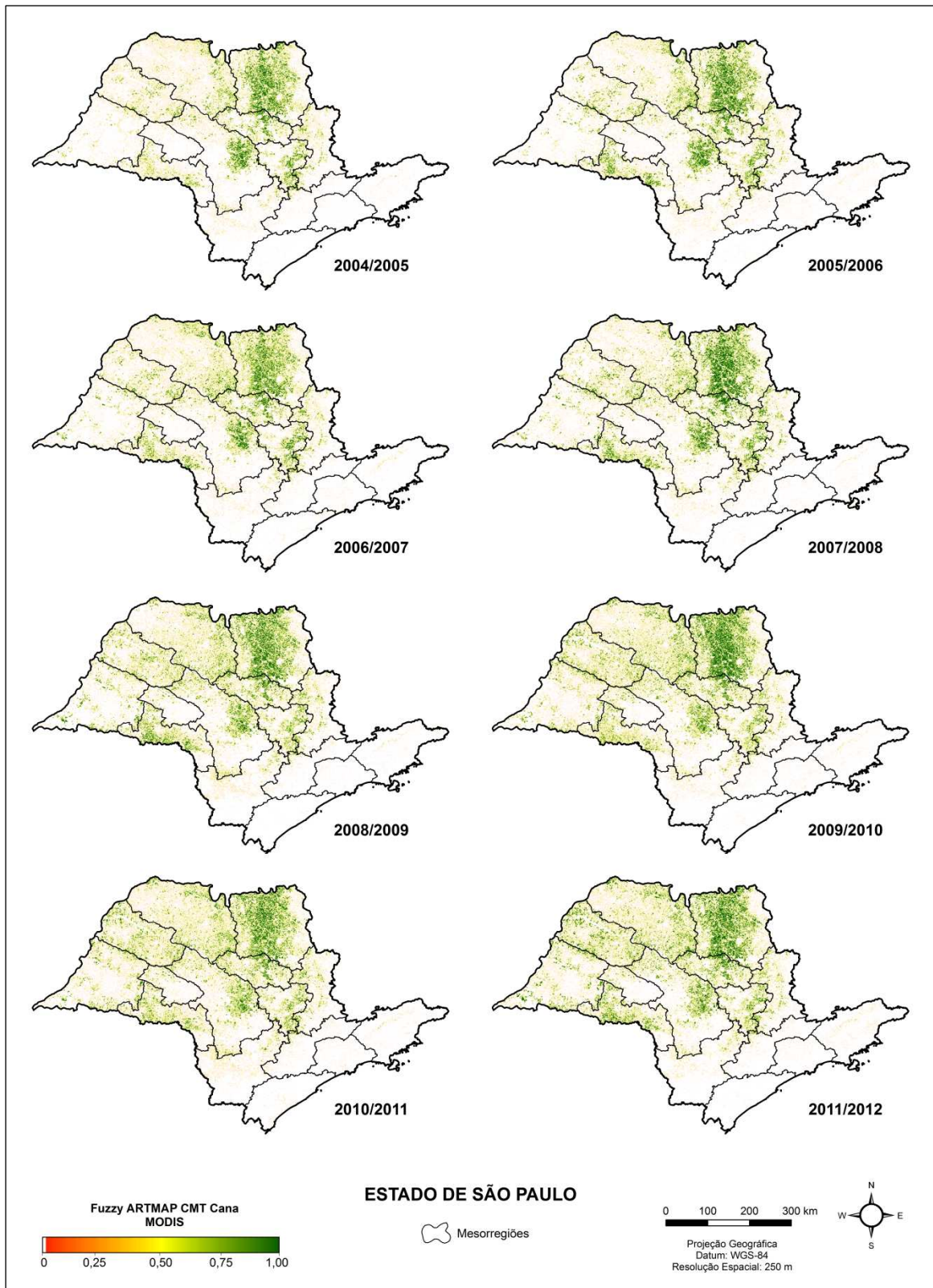


Figura 4.24: Modelos de classificação CMT Cana.

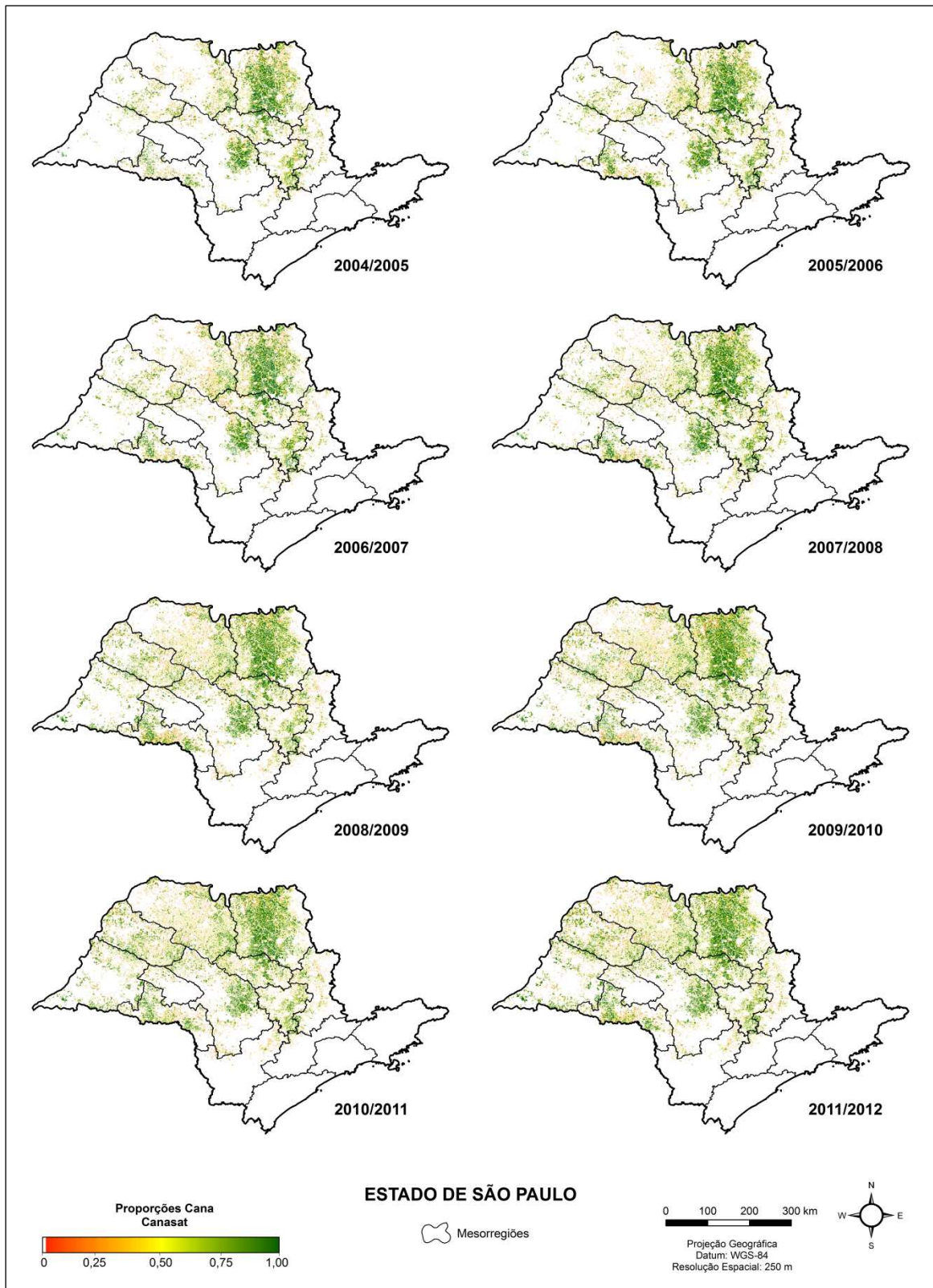


Figura 4.25: Imagens de proporções Cana.

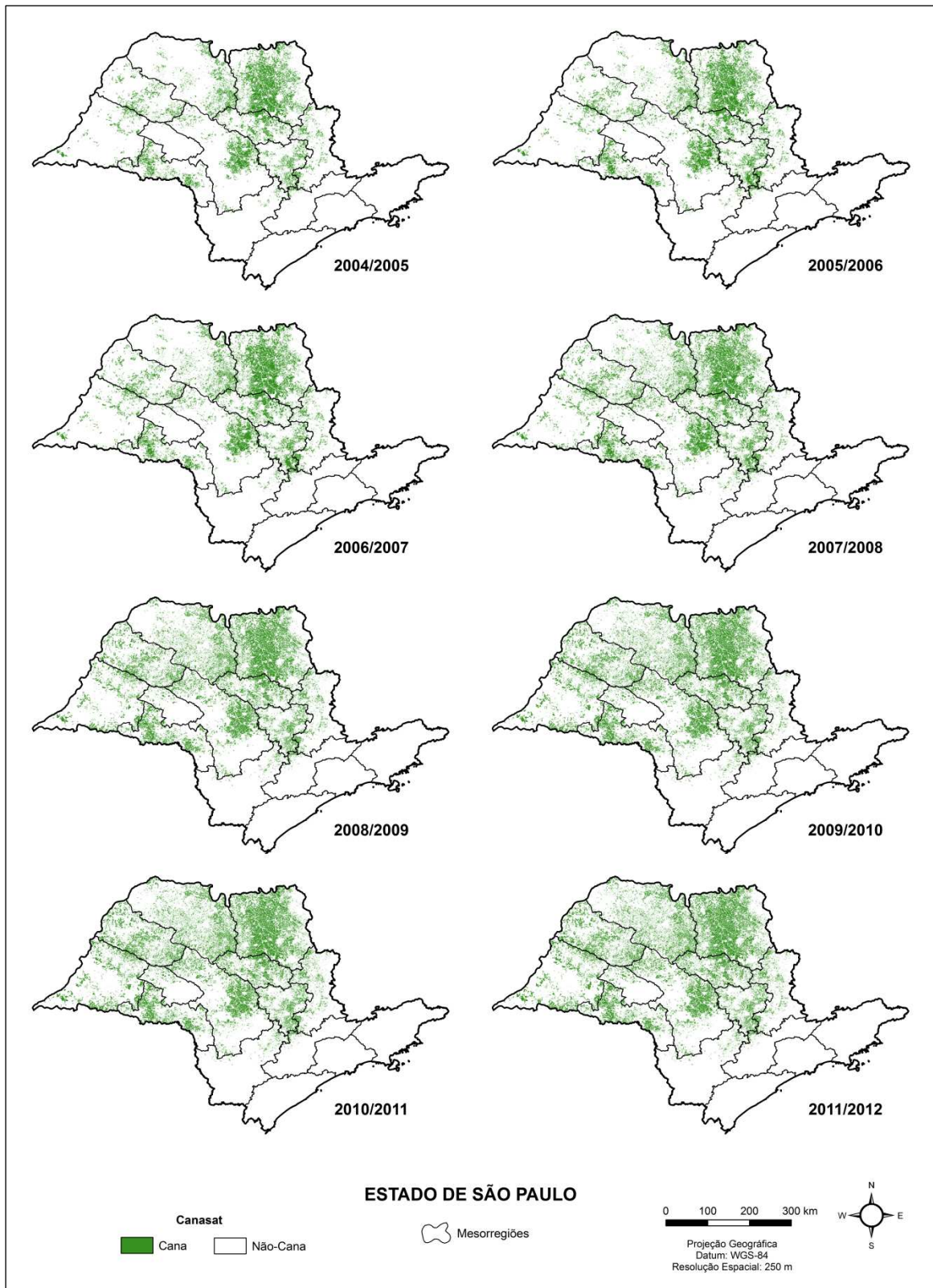


Figura 4.26: Imagens de referência Cana e Não-Cana.

As Tabelas 4.2 e 4.3 resumizam as medidas estatísticas para avaliação da acurácia dos modelos de classificação ART das safras 2004/2005 a 2011/2012, com base nas respectivas amostras de teste apresentadas na Tabela 4.1.

Tabela 4.2: Exatidão global, índice Kappa e discordância dos modelos de classificação ART.

Modelos ART Safras	Medidas Estatísticas			
	Exatidão Global (%)	Índice Kappa	Discordância	
			Quantidade (%)	Alocação (%)
2004/2005	95,80	0,808	3,18	1,02
2005/2006	95,77	0,816	2,90	1,33
2006/2007	95,09	0,806	3,83	1,08
2007/2008	95,41	0,832	2,74	1,85
2008/2009	95,51	0,848	1,86	2,63
2009/2010	95,39	0,854	2,24	2,37
2010/2011	95,14	0,849	2,71	2,15
2011/2012	95,15	0,851	2,32	2,53
MÉDIA	95,41	0,833	2,72	1,87

Tabela 4.3: Erros de omissão e comissão das classes Cana e Não-Cana dos modelos de classificação ART.

Modelos ART Safras	Medidas Estatísticas			
	Classe Cana		Classe Não-Cana	
	Erro de Omissão (%)	Erro de Comissão (%)	Erro de Omissão (%)	Erro de Comissão (%)
2004/2005	26,24	4,67	0,59	4,14
2005/2006	24,35	5,67	0,78	4,04
2006/2007	26,15	4,18	0,65	5,02
2007/2008	20,67	6,18	1,12	4,31
2008/2009	16,73	7,68	1,62	3,83
2009/2010	16,44	6,38	1,50	4,21
2010/2011	17,56	5,71	1,37	4,67
2011/2012	16,62	6,58	1,61	4,43
MÉDIA	20,60	5,88	1,16	4,33

As classificações alcançaram uma exatidão global média de 95,41%, com índice Kappa médio de 0,833. Conseqüentemente, a discordância total média foi de 4,59%, sendo 2,72% da discordância de quantidade e 1,87% da discordância de alocação. Para a classe Cana, o erro de omissão médio foi de 20,60% e o erro de comissão foi médio de 5,88%. Para a classe Não-Cana, o erro de omissão médio foi de 1,16% e o erro de comissão médio foi de 4,33%.

Com base nessas medidas estatísticas, pode-se dizer, de modo geral, que a classificação Fuzzy ARTMAP das oito safras produziu modelos ART de ótima qualidade temática, consistente com os altos valores de exatidão, os baixos valores de discordância e, também, pelos relativamente baixos valores de omissão e comissão.

A contribuição da componente de quantidade na discordância total média foi maior que a componente de alocação, implicando em maiores proporções incorretas de pixels designados nas classes. Isso pode ser observado pelo nível mais alto do erro de omissão da classe Cana, apontando para uma propensão à subestimação de áreas de cultivo da cana-de-açúcar ao longo das safras.

A Tabela 4.4 resume as medidas estatísticas para análise de correlação e validação dos modelos de classificação ART Cana das safras 2004/2005 a 2011/2012.

Tabela 4.4: Correlação e curva ROC dos modelos de classificação ART Cana.

Modelos ART Cana Safras	Medidas Estatísticas	
	r	AUC
2004/2005	0,924	0,881
2005/2006	0,935	0,896
2006/2007	0,917	0,876
2007/2008	0,944	0,885
2008/2009	0,940	0,881
2009/2010	0,936	0,870
2010/2011	0,935	0,871
2011/2012	0,943	0,881
MÉDIA	0,934	0,880

O coeficiente de correlação de Pearson (r) médio da relação entre as proporções Cana e os modelos CMT Cana gerados pelo Fuzzy ARTMAP foi de 0,934, conseguido com a abordagem de pós-processamento, considerando o contexto de vizinhança da soma normalizada de 3 x 3 pixels vizinhos mais próximos e com mascaramento das áreas sem cultivo da cana-de-açúcar. Esse valor representa um ajuste com correlação positiva, forte e estatisticamente significativa a 5%.

A AUC média dos modelos CMT Cana foi de 0,880 que é um valor notavelmente superior aos modelos aleatórios, comprovando que realmente é plausível mapear a cultura da cana-de-açúcar baseado nas medidas CMT. LI e FOX (2012) obtiveram valores de AUC entre

0,808 a 0,825 na classificação sub-pixel baseada em tipicidades Mahalanobis de dados NDVI/MODIS para mapear a distribuição de crescimento de seringueira no sudeste da Ásia.

Portanto, na classificação Fuzzy ARTMAP das oito safras, foi possível aliar a geração de modelos ART de maior exatidão e com menor discordância, com modelos CMT Cana que alcançaram uma forte correlação linear e com melhor desempenho, o que é necessário para propiciar a obtenção de estimativas de área plantada de cana-de-açúcar confiáveis.

4.7. Estimativa sub-pixel de área plantada de cana-de-açúcar

As medidas CMT expressam o grau de compromisso dos pixels à classe Cana. Com a finalidade de eliminar áreas sem cultivo da cana-de-açúcar, primeiramente foi aplicado o mascaramento espacial da imagem do modelo ART sobre o modelo CMT Cana, ambos gerados pelo classificador Fuzzy ARTMAP, com base na safra 2011/2012.

Após isso, a área plantada de cana-de-açúcar foi calculada a partir da multiplicação da área do pixel MODIS pelo somatório das medidas CMT Cana dos pixels dentro dos limites geográficos internos de cada município, por meio da Equação 3.17. Em seguida foi realizada a consolidação por mesorregiões, onde cada uma congrega diversos municípios. Desse modo, foram então obtidas as estimativas de área plantada de cana-de-açúcar para cada um dos 645 municípios do Estado de São Paulo, consolidadas pelas 15 mesorregiões e totalizadas para o nível estadual.

Esse procedimento foi reproduzido para as demais safras em estudo, baseado nos modelos CMT Cana com o mascaramento espacial pelos respectivos modelos ART, gerados pelo classificador Fuzzy ARTMAP, para obtenção das estimativas de área plantada de cana-de-açúcar.

As Tabelas 4.5 a 4.12 reúnem as estimativas de área plantada de cana-de-açúcar oficial do IBGE, a de referência do Canasat e a obtida do modelo CMT Cana gerado pelo classificador Fuzzy ARTMAP, consolidadas pelas 15 mesorregiões e totalizadas para o Estado de São Paulo, das safras 2004/2005 a 2011/2012, respectivamente.

Tabela 4.5: Estimativas de área plantada de cana-de-açúcar da safra 2004/2005.

Mesorregiões Estado de São Paulo	Número de Municípios	Área Plantada de Cana-de-açúcar (ha)		
		Safra 2004/2005		
		IBGE	Canasat	CMT
Araçatuba	36	223.053	233.460	251.218
Araraquara	21	265.657	271.134	240.288
Assis	35	203.343	247.388	194.280
Bauru	56	351.480	390.937	412.477
Campinas	49	197.127	167.985	176.729
Itapetininga	36	35.528	31.453	82.750
Litoral Sul Paulista	17	0	0	7.166
Macro Metropolitana Paulista	36	19.434	20.855	33.765
Marília	20	7.760	9.066	23.088
Metropolitana de São Paulo	45	0	6.364	7.412
Piracicaba	26	253.035	249.146	215.682
Presidente Prudente	54	123.182	131.929	127.119
Ribeirão Preto	66	958.323	1.035.596	1.007.969
São José do Rio Preto	109	312.417	370.074	476.213
Vale do Paraíba Paulista	39	1.465	0	21.154
TOTAL	645	2.951.804	3.165.387	3.277.309

Tabela 4.6: Estimativas de área plantada de cana-de-açúcar da safra 2005/2006.

Mesorregiões Estado de São Paulo	Número de Municípios	Área Plantada de Cana-de-açúcar (ha)		
		Safra 2005/2006		
		IBGE	Canasat	CMT
Araçatuba	36	226.406	248.018	261.512
Araraquara	21	268.876	282.395	273.675
Assis	35	215.274	254.842	283.412
Bauru	56	351.347	413.259	451.647
Campinas	49	204.467	174.295	186.581
Itapetininga	36	40.941	38.199	98.361
Litoral Sul Paulista	17	157	0	6.428
Macro Metropolitana Paulista	36	20.274	23.399	43.633
Marília	20	11.758	13.978	34.275
Metropolitana de São Paulo	45	0	6.747	12.712
Piracicaba	26	248.148	271.356	217.947
Presidente Prudente	54	146.980	151.081	157.371
Ribeirão Preto	66	987.787	1.080.202	1.076.774
São José do Rio Preto	109	360.488	406.933	518.694
Vale do Paraíba Paulista	39	1.849	0	28.633
TOTAL	645	3.084.752	3.364.704	3.651.654

Tabela 4.7: Estimativas de área plantada de cana-de-açúcar da safra 2006/2007.

Mesorregiões Estado de São Paulo	Número de Municípios	Área Plantada de Cana-de-açúcar (ha)		
		Safra 2006/2007		
		IBGE	Canasat	CMT
Araçatuba	36	289.699	279.294	293.654
Araraquara	21	261.351	303.916	284.876
Assis	35	278.016	271.060	270.760
Bauru	56	397.684	440.739	428.768
Campinas	49	219.431	184.147	182.205
Itapetininga	36	47.848	41.100	83.261
Litoral Sul Paulista	17	13	0	15.460
Macro Metropolitana Paulista	36	21.559	23.902	33.284
Marília	20	16.621	15.891	36.036
Metropolitana de São Paulo	45	0	6.988	12.917
Piracicaba	26	251.994	278.264	227.931
Presidente Prudente	54	207.471	184.341	198.803
Ribeirão Preto	66	1.066.831	1.147.672	1.047.934
São José do Rio Preto	109	437.899	483.841	587.483
Vale do Paraíba Paulista	39	1.848	0	26.339
TOTAL	645	3.498.265	3.661.155	3.729.710

Tabela 4.8: Estimativas de área plantada de cana-de-açúcar da safra 2007/2008.

Mesorregiões	Número de Municípios	Área Plantada de Cana-de-açúcar (ha)		
		Safra 2007/2008		
		IBGE	Canasat	CMT
Araçatuba	36	345.733	373.717	310.874
Araraquara	21	297.937	327.405	297.183
Assis	35	271.513	318.515	294.086
Bauru	56	437.514	529.180	458.720
Campinas	49	226.408	204.571	209.321
Itapetininga	36	47.367	57.682	107.721
Litoral Sul Paulista	17	0	0	8.114
Macro Metropolitana Paulista	36	19.594	26.559	34.808
Marília	20	23.585	32.403	41.373
Metropolitana de São Paulo	45	0	8.056	10.203
Piracicaba	26	275.798	293.472	235.483
Presidente Prudente	54	223.876	248.733	237.524
Ribeirão Preto	66	1.174.672	1.225.322	1.174.115
São José do Rio Preto	109	544.891	604.307	542.314
Vale do Paraíba Paulista	39	1.526	0	26.949
TOTAL	645	3.890.414	4.249.922	3.988.788

Tabela 4.9: Estimativas de área plantada de cana-de-açúcar da safra 2008/2009.

Mesorregiões Estado de São Paulo	Número de Municípios	Área Plantada de Cana-de-açúcar (ha)		
		Safra 2008/2009		
		IBGE	Canasat	CMT
Araçatuba	36	438.875	478.512	393.159
Araraquara	21	355.253	361.782	306.735
Assis	35	329.018	349.773	355.518
Bauru	56	493.145	596.829	513.946
Campinas	49	239.757	217.165	198.157
Itapetininga	36	64.167	62.644	126.695
Litoral Sul Paulista	17	157	0	5.035
Macro Metropolitana Paulista	36	21.779	28.072	35.402
Marília	20	47.617	48.822	58.032
Metropolitana de São Paulo	45	0	9.079	8.658
Piracicaba	26	311.312	300.631	218.566
Presidente Prudente	54	311.275	341.151	297.389
Ribeirão Preto	66	1.261.347	1.326.217	1.205.171
São José do Rio Preto	109	666.079	753.263	687.110
Vale do Paraíba Paulista	39	1.728	0	23.373
TOTAL	645	4.541.509	4.873.940	4.432.945

Tabela 4.10: Estimativas de área plantada de cana-de-açúcar da safra 2009/2010.

Mesorregiões Estado de São Paulo	Número de Municípios	Área Plantada de Cana-de-açúcar (ha)		
		Safra 2009/2010		
		IBGE	Canasat	CMT
Araçatuba	36	490.197	532.252	444.465
Araraquara	21	364.033	375.774	302.653
Assis	35	315.436	366.226	307.921
Bauru	56	587.662	640.457	487.256
Campinas	49	235.522	241.915	183.461
Itapetininga	36	67.961	71.233	93.529
Litoral Sul Paulista	17	182	0	7.030
Macro Metropolitana Paulista	36	22.461	28.851	33.590
Marília	20	54.130	62.452	78.299
Metropolitana de São Paulo	45	0	0	12.156
Piracicaba	26	322.650	311.408	213.570
Presidente Prudente	54	367.824	423.180	342.967
Ribeirão Preto	66	1.354.770	1.362.770	1.316.544
São José do Rio Preto	109	792.640	825.970	780.421
Vale do Paraíba Paulista	39	1.609	0	23.949
TOTAL	645	4.977.077	5.242.488	4.627.810

Tabela 4.11: Estimativas de área plantada de cana-de-açúcar da safra 2010/2011.

Mesorregiões Estado de São Paulo	Número de Municípios	Área Plantada de Cana-de-açúcar (ha)		
		Safra 2010/2011		
		IBGE	Canasat	CMT
Araçatuba	36	535.229	547.078	450.457
Araraquara	21	327.602	378.948	302.369
Assis	35	329.625	361.931	315.206
Bauru	56	607.248	641.269	525.852
Campinas	49	224.669	239.994	196.128
Itapetininga	36	61.699	70.593	143.846
Litoral Sul Paulista	17	9	0	17.054
Macro Metropolitana Paulista	36	22.469	28.903	43.045
Marília	20	49.454	65.219	75.384
Metropolitana de São Paulo	45	0	0	14.411
Piracicaba	26	317.969	307.965	232.934
Presidente Prudente	54	421.485	444.535	369.995
Ribeirão Preto	66	1.300.950	1.364.588	1.173.847
São José do Rio Preto	109	870.861	852.319	745.033
Vale do Paraíba Paulista	39	1.936	0	31.152
TOTAL	645	5.071.205	5.303.342	4.636.712

Tabela 4.12: Estimativas de área plantada de cana-de-açúcar da safra 2011/2012.

Mesorregiões	Número de Municípios	Área Plantada de Cana-de-açúcar (ha)		
		Safra 2011/2012		
		IBGE	Canasat	CMT
Araçatuba	36	545.853	557.177	507.823
Araraquara	21	376.591	390.588	315.975
Assis	35	346.194	367.004	340.851
Bauru	56	598.601	640.228	509.228
Campinas	49	240.800	240.181	199.776
Itapetininga	36	67.919	70.595	103.572
Litoral Sul Paulista	17	9	0	5.191
Macro Metropolitana Paulista	36	24.814	29.175	36.660
Marília	20	67.482	71.145	65.223
Metropolitana de São Paulo	45	0	0	11.276
Piracicaba	26	310.463	306.459	224.198
Presidente Prudente	54	455.276	462.817	418.459
Ribeirão Preto	66	1.333.636	1.376.400	1.220.752
São José do Rio Preto	109	847.958	889.054	805.996
Vale do Paraíba Paulista	39	895	0	35.763
TOTAL	645	5.216.491	5.400.823	4.800.742

4.8. Comparação entre as estimativas de área plantada de cana-de-açúcar

A análise estatística dos dados iniciou-se com a verificação de normalidade das estimativas de área plantada de cana-de-açúcar oficial do IBGE, a de referência do Canasat e a do modelo de classificação Fuzzy ARTMAP CMT das oito safras.

Para isso, foi aplicado o Teste de Shapiro-Wilk a 5% de significância, onde todos os 408 conjuntos de dados referentes a todas as safras, tanto em nível municipal como por mesorregião, tiveram a hipótese de normalidade rejeitada ($p\text{-valor} < 0,05$), evidenciando que os dados não seguem uma distribuição normal.

Esse resultado está relacionado ao fato de que todos os 645 municípios e as 15 mesorregiões do Estado de São Paulo foram considerados no estudo, independentemente se representavam ou não expressivas áreas de cultivo da cana-de-açúcar.

Como a validade do coeficiente de correlação estatística de Pearson (r) pressupõe a normalidade dos conjuntos de dados, optou-se então por utilizar o coeficiente de correlação não paramétrica de Spearman (r_s) para proceder a comparação pareada entre as estimativas de área plantada de cana-de-açúcar e, dessa forma, não incorrer em interpretações errôneas dos resultados.

As Figuras 4.27, 4.28 e 4.29 apresentam os gráficos da relação entre as estimativas de área plantada de cana-de-açúcar oficial do IBGE e de referência do Canasat, de referência do Canasat e do modelo de classificação Fuzzy ARTMAP CMT Cana e, por último, a oficial do IBGE e a do modelo de classificação Fuzzy ARTMAP CMT Cana, respectivamente, dos 645 municípios do Estado de São Paulo, das safras 2004/2005 a 2011/2012.

Assim sendo, para comparação estatística entre as estimativas de área plantada de cana-de-açúcar foi utilizada a análise de correlação e de concordância. Para complementar as análises estatísticas foi discutida a informação acerca das medidas de desvio entre os dados, que é fundamental para fornecer uma melhor percepção em termos de precisão e exatidão da relação entre as estimativas de área plantada.

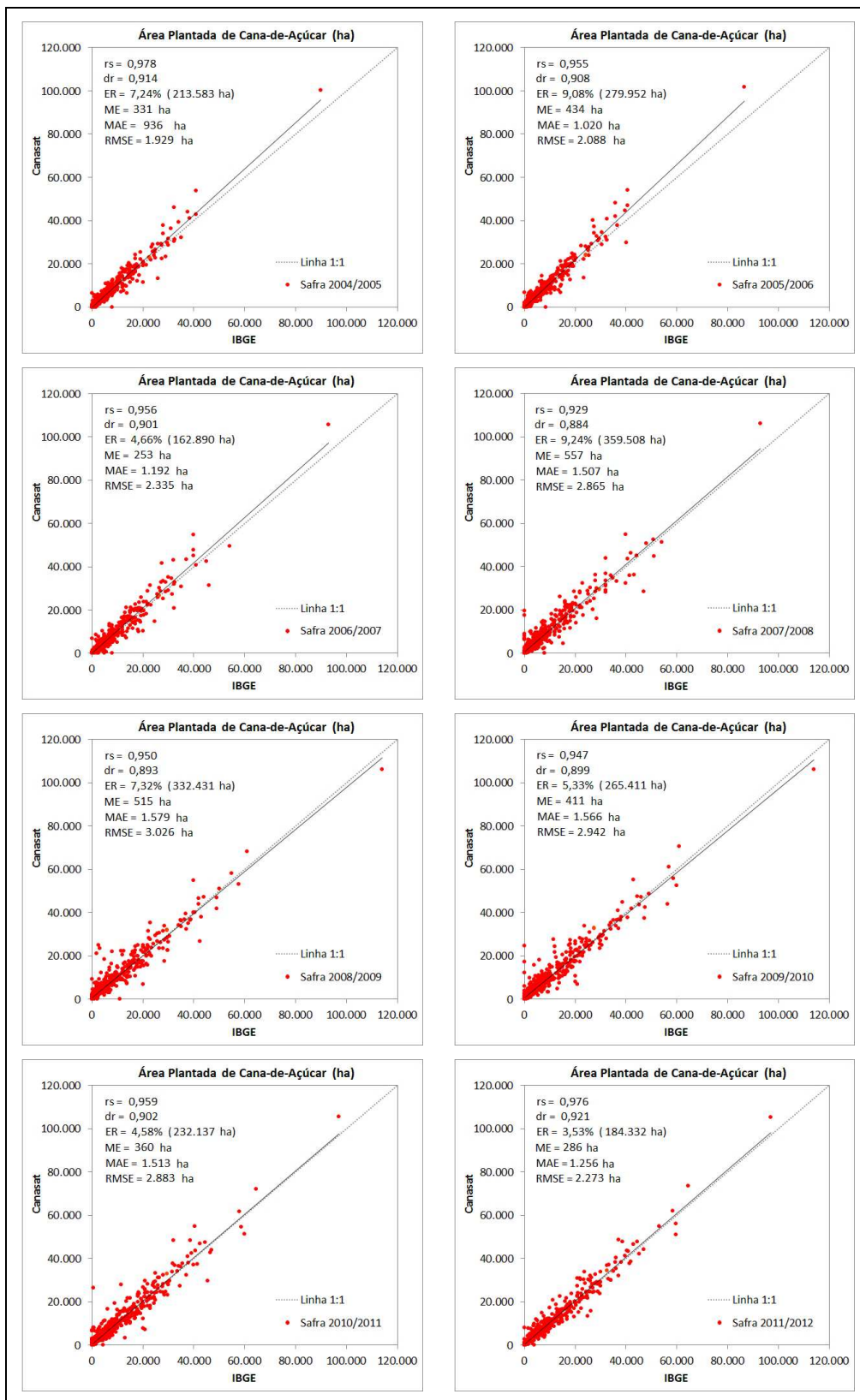


Figura 4.27: Relação entre a área plantada de cana-de-açúcar do IBGE e do Canasat em nível municipal nas oito safras.

Os ajustes lineares realizados entre os dados do IBGE e do Canasat em nível municipal, tiveram a hipótese de aleatoriedade rejeitada ($p\text{-valor} < 0,05$), isto é, a correlação foi estatisticamente significativa a 5% para todas as safras em estudo. A discussão desses resultados é baseada nas informações mostradas nos gráficos da Figura 4.27.

Os valores do coeficiente de correlação de Spearman (r_s) variaram entre 0,929 (safra 2007/2008) a 0,976 (safra 2011/2012), indicando que as estimativas de área plantada de cana-de-açúcar são fortemente correlacionadas e possuem a mesma tendência, movendo-se em perfeita proporção na mesma direção. Esses resultados comprovam a grande precisão do ajuste na relação entre os dados.

Os valores do índice de concordância de Willmott (d_r) variaram entre 0,884 (safra 2007/2008) e 0,921 (safra 2011/2012), confirmando o ótimo desempenho do ajuste, que obteve uma alta exatidão entre as estimativas de área plantada de cana-de-açúcar do IBGE e do Canasat.

O menor erro relativo entre os dados do IBGE e os do Canasat foi de 3,53% (safra 2011/2012) e o maior foi de 9,24% (safra 2007/2008), o que representa uma diferença de 184.332 ha e 359.508 ha da área total de cana-de-açúcar do Estado de São Paulo, respectivamente. Essa variação comprova o que foi indicado pelo coeficiente de correlação de Spearman (r_s) e pelo índice de concordância de Willmott (d_r), onde os menores valores foram obtidos para a safra 2007/2008 e os maiores valores para a safra 2011/2012.

Pelo erro médio (ME), constatou-se que as estimativas de área do Canasat foram superestimadas em relação às do IBGE entre 253 ha (safra 2006/2007) a 557 ha (safra 2007/2008), na média municipal. O erro médio absoluto (MAE) mostrou que a magnitude média municipal dos erros entre as estimativas de área variou de 936 ha (safra 2004/2005) a 1.579 ha (safra 2008/2009). A raiz do erro médio quadrático (RMSE) revelou que a dimensão do erro teve uma variação média municipal entre 1.929 ha (safra 2004/2005) e 3.026 ha (safra 2008/2009).

Portanto, os ajustes obtiveram um excelente desempenho em termos de precisão e exatidão, com uma superestimação média de 6,37% das estimativas de área plantada de cana-de-açúcar do Canasat em relação às do IBGE.

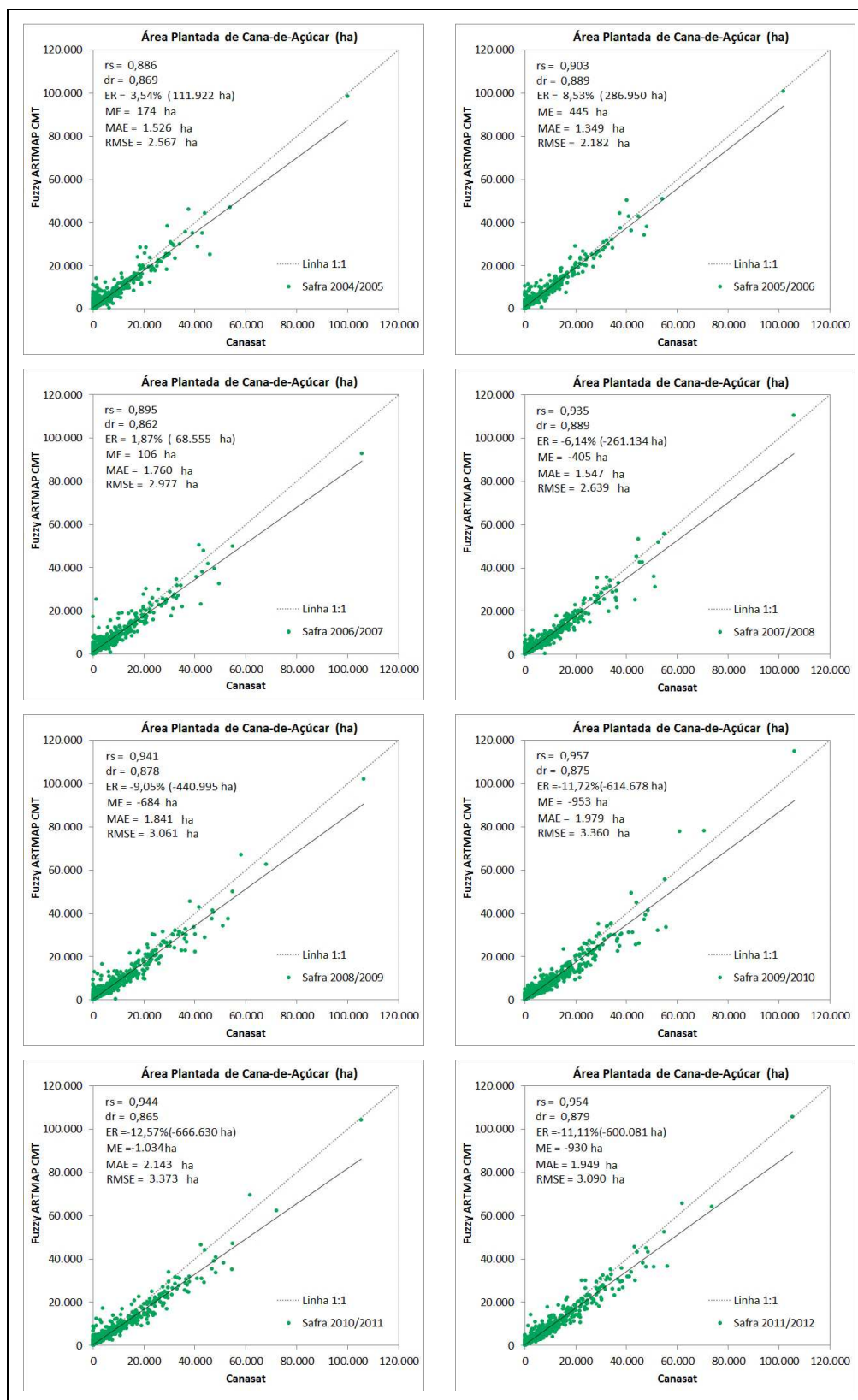


Figura 4.28: Relação entre a área plantada de cana-de-açúcar do Canasat e do Fuzzy ARTMAP CMT em nível municipal nas oito safras.

Os ajustes lineares realizados entre os dados do Canasat e do Fuzzy ARTMAP CMT em nível municipal, tiveram a hipótese de aleatoriedade rejeitada (p -valor $< 0,05$), isto é, a correlação foi estatisticamente significativa a 5% para todas as safras em estudo. A discussão desses resultados é baseada nas informações mostradas nos gráficos da Figura 4.28.

Os valores do coeficiente de correlação de Spearman (r_s) variaram entre 0,886 (safra 2004/2005) a 0,957 (safra 2009/2010), indicando que as estimativas de área plantada de cana-de-açúcar são fortemente correlacionadas e possuem a mesma tendência, movendo-se em perfeita proporção na mesma direção. Esses resultados também comprovam a grande precisão do ajuste na relação entre os dados.

Os valores do índice de concordância de Willmott (d_r) variaram entre 0,862 (safra 2006/2007) e 0,889 (safra 2005/2006), confirmando o ótimo desempenho do ajuste, que obteve uma alta exatidão entre as estimativas de área plantada de cana-de-açúcar do Canasat e do Fuzzy ARTMAP CMT.

O menor erro relativo entre os dados do Canasat e do Fuzzy ARTMAP CMT foi de 1,87% (safra 2006/2007) e o maior foi de -12,57% (safra 2010/2011), o que representa uma diferença de 68.555 ha e -666.630 ha da área total de cana-de-açúcar do Estado de São Paulo, respectivamente. Esse grau de variação aponta um afastamento um pouco maior entre os dados.

Pelo erro médio (ME), verificou-se que as estimativas de área do Fuzzy ARTMAP CMT foram superestimadas em relação às do Canasat de 174 ha (safra 2004/2005) a 445 ha (safra 2005/2006) e subestimadas de 405 ha (safra 2007/2008) a 1.034 ha (safra 2010/2011), na média municipal. O erro médio absoluto (MAE) mostrou que a magnitude média municipal dos erros entre as estimativas de área variou de 1.349 ha (safra 2005/2006) a 2.143 ha (safra 2010/2011). A raiz do erro médio quadrático (RMSE) revelou que a dimensão do erro teve uma variação média municipal entre 2.182 ha (safra 2005/2006) e 3.173 ha (safra 2010/2011).

Dessa forma, os ajustes obtiveram um ótimo desempenho em termos de precisão e exatidão, porém indicando uma tendência de subestimação das estimativas de área plantada de cana-de-açúcar do Fuzzy ARTMAP CMT em relação às do Canasat.

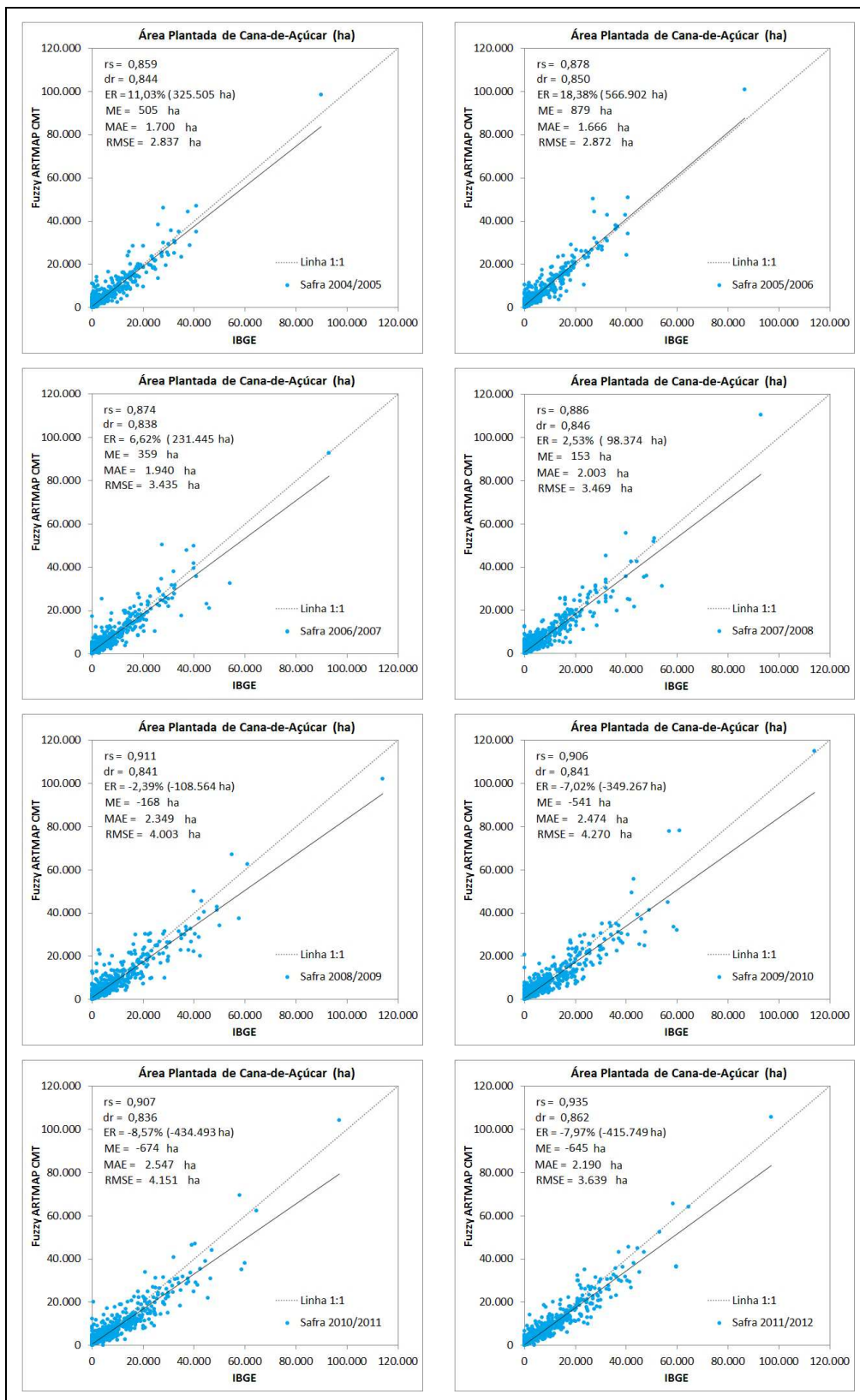


Figura 4.29: Relação entre a área plantada de cana-de-açúcar do IBGE e do modelo Fuzzy ARTMAP CMT em nível municipal nas oito safras.

Os ajustes lineares realizados entre os dados do IBGE e do Fuzzy ARTMAP CMT em nível municipal, tiveram a hipótese de aleatoriedade rejeitada (p -valor $< 0,05$), isto é, a correlação foi estatisticamente significativa a 5% para todas as safras em estudo. A discussão desses resultados é baseada nas informações mostradas nos gráficos da Figura 4.29.

Os valores do coeficiente de correlação de Spearman (r_s) variaram entre 0,859 (safra 2004/2005) a 0,935 (safra 2011/2012), indicando que as estimativas de área plantada de cana-de-açúcar são fortemente correlacionadas e possuem a mesma tendência, movendo-se em perfeita proporção na mesma direção. Esses resultados comprovam a boa precisão do ajuste na relação entre os dados.

Os valores do índice de concordância de Willmott (d_r) variaram entre 0,836 (safra 2010/2011) e 0,862 (safra 2011/2012), confirmando o bom desempenho do ajuste, que obteve uma alta exatidão entre as estimativas de área plantada de cana-de-açúcar do IBGE e do Fuzzy ARTMAP CMT.

O menor erro relativo entre os dados do IBGE e do Fuzzy ARTMAP CMT foi de -2,39% (safra 2008/2009) e o maior foi de 18,38% (safra 2005/2006), o que representa uma diferença de -108.564 ha e 566.902 ha da área total de cana-de-açúcar do Estado de São Paulo, respectivamente. Esse grau de variação mostra uma dispersão maior entre os dados.

Pelo erro médio (ME), verificou-se que as estimativas de área do Fuzzy ARTMAP CMT foram superestimadas em relação às do IBGE de 153 ha (safra 2007/2008) a 879 ha (safra 2005/2006) e subestimadas de 168 ha (safra 2008/2009) a 674 ha (safra 2010/2011), na média municipal. O erro médio absoluto (MAE) mostrou que a magnitude média municipal dos erros entre as estimativas de área variou de 1.666 ha (safra 2005/2006) a 2.547 ha (safra 2010/2011). A raiz do erro médio quadrático (RMSE) revelou que a dimensão do erro teve uma variação média municipal entre 2.837 ha (safra 2004/2005) e 4.270 ha (safra 2009/2010).

Desse modo, todos os ajustes obtiveram um bom desempenho em termos de precisão e exatidão, porém apontando uma tendência de superestimação das estimativas de área plantada de cana-de-açúcar do Fuzzy ARTMAP CMT em relação às do IBGE.

De maneira geral, pode-se dizer que os ajustes obtidos pela comparação pareada entre as estimativas de área plantada de cana-de-açúcar, em nível municipal, alcançaram desempenho similar para todas as safras em estudo no Estado de São Paulo.

As estimativas de área plantada de cana-de-açúcar do Canasat aproximaram-se mais do IBGE, porque obtiveram os maiores indicadores de correlação e concordância (r_s de 0,929 a 0,976; d_r de 0,884 a 0,921), com os menores desvios entre os dados (ER de 3,53% a 9,24%). Já as estimativas de área plantada de cana-de-açúcar do Fuzzy ARTMAP CMT aproximaram-se mais do Canasat, determinado pelos altos indicadores de correlação e de concordância (r_s 0,886 a 0,957; d_r 0,862 a 0,889), além de também estarem bem balizadas com as do IBGE (r_s 0,859 a 0,935; d_r 0,836 a 0,862), porém com desvios um pouco menores em relação aos dados do Canasat (ER de 1,87% a -12,57%) do que os dados do IBGE (ER -2,39% a 18,38%).

Cabe destacar que as estimativas de área plantada de cana-de-açúcar do Canasat são baseadas no mapeamento de imagens tipo Landsat (30 m) por talhão, enquanto que as do modelo Fuzzy ARTMAP CMT são obtidas a partir da classificação sub-pixel de imagens MODIS (250 m), sendo assim mais um indicativo de boa precisão.

Ressalta-se também as grandes vantagens da metodologia para obtenção das estimativas de área plantada de cana-de-açúcar do modelo Fuzzy ARTMAP CMT, que além de ser realizada de forma objetiva, com baixo custo e bastante rapidez, propicia o conhecimento da distribuição espacial do cultivo, o que não é possível quando se utiliza a metodologia do IBGE.

A Tabela 4.13 apresenta a correlação e a concordância médias das oito safras por mesorregiões do Estado de São Paulo, na comparação pareada entre as estimativas de área plantada de cana-de-açúcar oficial do IBGE e de referência do Canasat, de referência do Canasat e do Fuzzy ARTMAP CMT Cana e, também, a oficial do IBGE e a do Fuzzy ARTMAP CMT Cana, respectivamente. As mesorregiões Litoral Sul Paulista, Metropolitana de São Paulo e Vale do Paraíba Paulista não constam da análise porque o cultivo da cana-de-açúcar nos seus respectivos municípios é inexistente ou pouco expressivo, o que impossibilita o cálculo das medidas estatísticas. A discussão dos resultados por mesorregiões a seguir é baseada nas informações da Tabela 4.13.

Os ajustes lineares realizados entre os dados do IBGE e do Canasat por mesorregiões, tiveram a hipótese de aleatoriedade rejeitada (p -valor $< 0,05$), isto é, a correlação foi estatisticamente significativa a 5% para todas as safras em estudo.

Os valores médios do coeficiente de correlação de Spearman (r_s) variaram entre 0,623 (Macro Metropolitana Paulista) a 0,975 (Ribeirão Preto), indicando que as estimativas de área plantada de cana-de-açúcar possuem uma correlação positiva, de moderada a forte.

Tabela 4.13: Medidas estatísticas médias das oito safras.

Mesorregiões Estado de São Paulo	Número de Municípios	Medidas Estatísticas		
		IBGE X Canasat	Canasat X CMT	IBGE X CMT
Araçatuba	36			
Correlação de Spearman (r_s)		0,927	0,884	0,832
Concordância de Willmott (d_r)		0,867	0,833	0,780
Araraquara	21			
Correlação de Spearman (r_s)		0,915	0,969	0,901
Concordância de Willmott (d_r)		0,823	0,844	0,775
Assis	35			
Correlação de Spearman (r_s)		0,966	0,944	0,915
Concordância de Willmott (d_r)		0,871	0,841	0,793
Bauru	56			
Correlação de Spearman (r_s)		0,915	0,944	0,845
Concordância de Willmott (d_r)		0,877	0,856	0,807
Campinas	49			
Correlação de Spearman (r_s)		0,887	0,925	0,904
Concordância de Willmott (d_r)		0,845	0,876	0,830
Itapetininga	36			
Correlação de Spearman (r_s)		0,870	0,711	0,617
Concordância de Willmott (d_r)		0,831	0,531	0,496
Macro Metropolitana Paulista	36			
Correlação de Spearman (r_s)		0,623	0,555	0,401
Concordância de Willmott (d_r)		0,850	0,731	0,704
Marília	20			
Correlação de Spearman (r_s)		0,902	0,854	0,805
Concordância de Willmott (d_r)		0,809	0,672	0,564
Piracicaba	26			
Correlação de Spearman (r_s)		0,963	0,948	0,920
Concordância de Willmott (d_r)		0,871	0,816	0,813
Presidente Prudente	54			
Correlação de Spearman (r_s)		0,897	0,814	0,722
Concordância de Willmott (d_r)		0,809	0,764	0,704
Ribeirão Preto	66			
Correlação de Spearman (r_s)		0,975	0,978	0,955
Concordância de Willmott (d_r)		0,895	0,893	0,862
São José do Rio Preto	109			
Correlação de Spearman (r_s)		0,949	0,881	0,841
Concordância de Willmott (d_r)		0,867	0,796	0,746

Os valores médios do índice de concordância de Willmott (d_r) variaram entre 0,809 (Marília) e 0,895 (Ribeirão Preto), ou seja, um ajuste com alta exatidão entre as estimativas de área plantada de cana-de-açúcar do IBGE e do Canasat.

Os ajustes lineares realizados entre os dados do Canasat e do Fuzzy ARTMAP CMT por mesorregiões, tiveram a hipótese de aleatoriedade rejeitada (p -valor $< 0,05$), isto é, a correlação foi estatisticamente significativa a 5% para todas as safras em estudo.

Os valores médios do coeficiente de correlação de Spearman (r_s) variaram entre 0,555 (Macro Metropolitana Paulista) a 0,978 (Ribeirão Preto), indicando que as estimativas de área plantada de cana-de-açúcar possuem uma correlação positiva, de moderada a forte.

Os valores médios do índice de concordância de Willmott (d_r) variaram entre 0,531 (Itapetininga) e 0,893 (Ribeirão Preto), ou seja, um ajuste com média a alta exatidão entre as estimativas de área plantada de cana-de-açúcar do Canasat e do Fuzzy ARTMAP CMT.

Os ajustes lineares realizados entre os dados do IBGE e do Fuzzy ARTMAP CMT por mesorregiões, tiveram a hipótese de aleatoriedade rejeitada (p -valor $< 0,05$), isto é, a correlação foi estatisticamente significativa a 5% para todas as safras em estudo.

Os valores médios do coeficiente de correlação de Spearman (r_s) variaram entre 0,401 (Macro Metropolitana Paulista) a 0,955 (Ribeirão Preto), indicando que as estimativas de área plantada de cana-de-açúcar possuem uma correlação positiva, de moderada a forte.

Os valores médios do índice de concordância de Willmott (d_r) variaram entre 0,496 (Itapetininga) e 0,862 (Ribeirão Preto), ou seja, um ajuste com média a alta exatidão entre as estimativas de área plantada de cana-de-açúcar do IBGE e do Fuzzy ARTMAP CMT.

De maneira geral, pode-se dizer que os ajustes obtidos pela comparação pareada entre as estimativas de área plantada de cana-de-açúcar alcançaram um desempenho médio equivalente em termos de precisão e exatidão, o que pode estar relacionado à distribuição espacial do cultivo nas mesorregiões, influenciando o reconhecimento do padrão da cana-de-açúcar pelo classificador Fuzzy ARTMAP.

As estimativas de área plantada de cana-de-açúcar do Canasat novamente se aproximaram mais do IBGE, porque obtiveram indicadores de correlação e de concordância mais altos. Já as estimativas de área plantada de cana-de-açúcar do Fuzzy ARTMAP CMT se aproximaram mais do Canasat, determinado pelos indicadores de correlação e de concordância em níveis semelhantes, ainda assim também são balizadas com as do IBGE.

Com base em PISSINATO (2014), que analisou detalhadamente o panorama da evolução histórica da cultura da cana-de-açúcar no Estado de São Paulo de 1950 até 2010, as mesorregiões podem ser caracterizadas de acordo com a relação entre a área plantada e o percentual de participação no total das terras agrícolas.

As mesorregiões de Araçatuba, Araraquara, Bauru, Piracicaba, Presidente Prudente, Ribeirão Preto e São José do Rio Preto são consideradas muito expressivas no cultivo da cana-de-açúcar, com participação superior a 70% de área plantada no total das terras agrícolas, sendo Ribeirão Preto e Piracicaba regiões tradicionais de intensificação do cultivo e, as demais regiões apresentando forte expansão. As mesorregiões de Assis, Campinas e Marília são consideradas de cultivo expressivo da cana-de-açúcar, com participação superior a 40% de área plantada no total das terras agrícolas, sendo regiões de expansão do cultivo. As mesorregiões de Itapetininga e Macro Metropolitana Paulista são consideradas de cultivo regular da cana-de-açúcar, com participação de até 20% de área plantada no total das terras agrícolas, sendo uma cultura pouco relevante nessas regiões.

A partir disso, realmente pode-se comprovar que o desempenho do classificador Fuzzy ARTMAP está intrinsecamente relacionado à caracterização da distribuição geográfica do cultivo da cana-de-açúcar nas mesorregiões do Estado de São Paulo. As mesorregiões de cultivo muito expressivo da cana-de-açúcar alcançaram os melhores ajustes, como Ribeirão Preto e Piracicaba que possuem grandes áreas de plantio mais homogêneas. As mesorregiões de Araçatuba, Araraquara, Bauru, Presidente Prudente e São José do Rio Preto também apresentaram ótimos ajustes, mas como são regiões de grande expansão principalmente sobre pastagens, possuem áreas bastante heterogêneas. As mesorregiões de Assis, Campinas e Marília tiveram bons ajustes, relacionado ao cultivo expressivo da cana-de-açúcar, mas com áreas de expansão menos significativas que não interferem na importância de outras culturas agrícolas. Os piores ajustes foram encontrados para as mesorregiões de Itapetininga e Macro Metropolitana Paulista, onde não predomina o cultivo da cana-de-açúcar.

As Figuras 4.30, 4.31 e 4.32 apresentam os gráficos da relação entre as estimativas de área plantada de cana-de-açúcar oficial do IBGE e de referência do Canasat, de referência do Canasat e do Fuzzy ARTMAP CMT Cana e, por último, a oficial do IBGE e a do Fuzzy ARTMAP CMT Cana, respectivamente, das 15 mesorregiões consolidadas para o Estado de São Paulo, das safras 2004/2005 a 2011/2012.

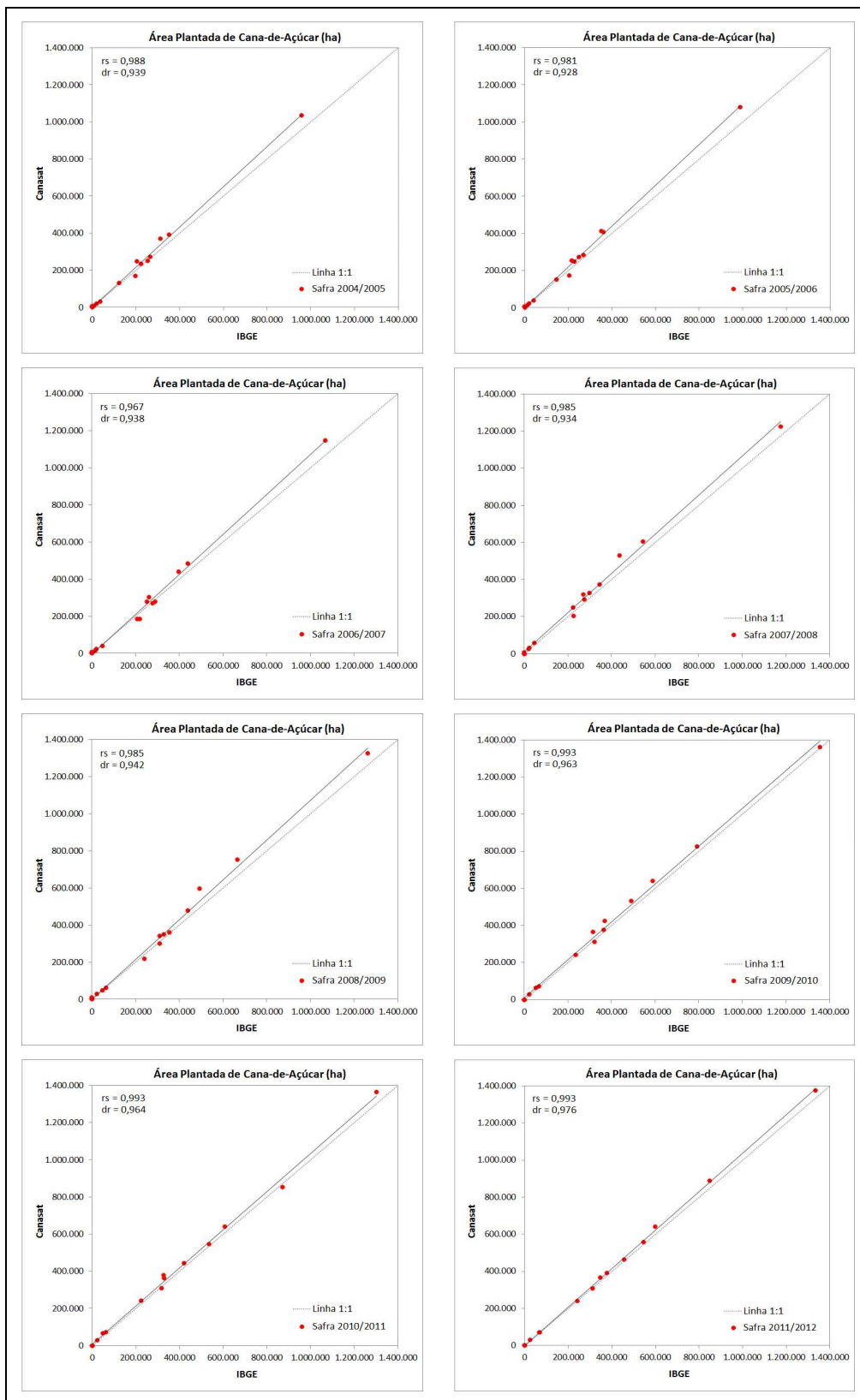


Figura 4.30: Relação entre a área plantada de cana-de-açúcar do IBGE e do Canasat em nível de mesorregiões das oito safras.

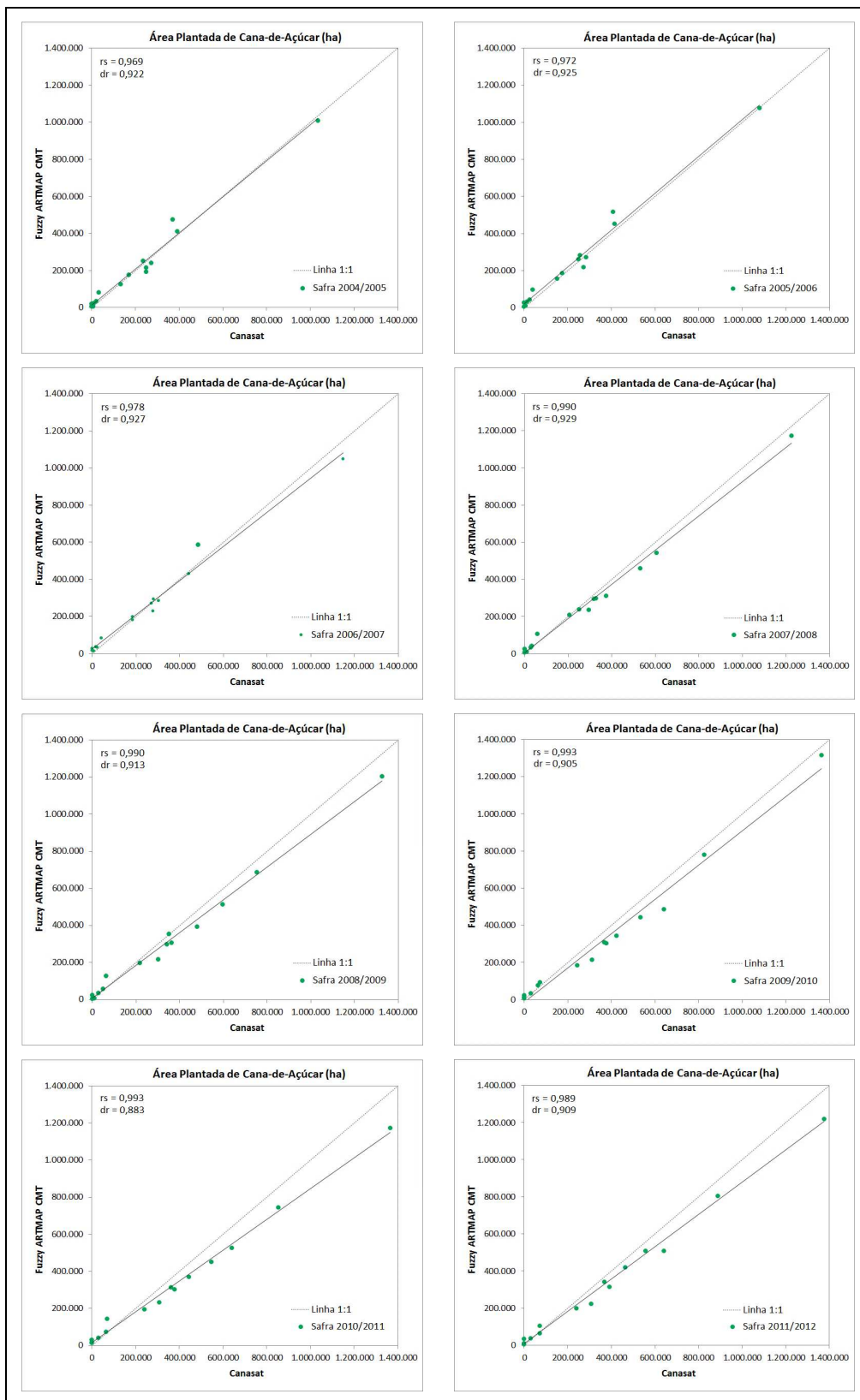


Figura 4.31: Relação entre a área plantada de cana-de-açúcar do Canasat e do Fuzzy ARTMAP CMT em nível de mesorregiões das oito safras.

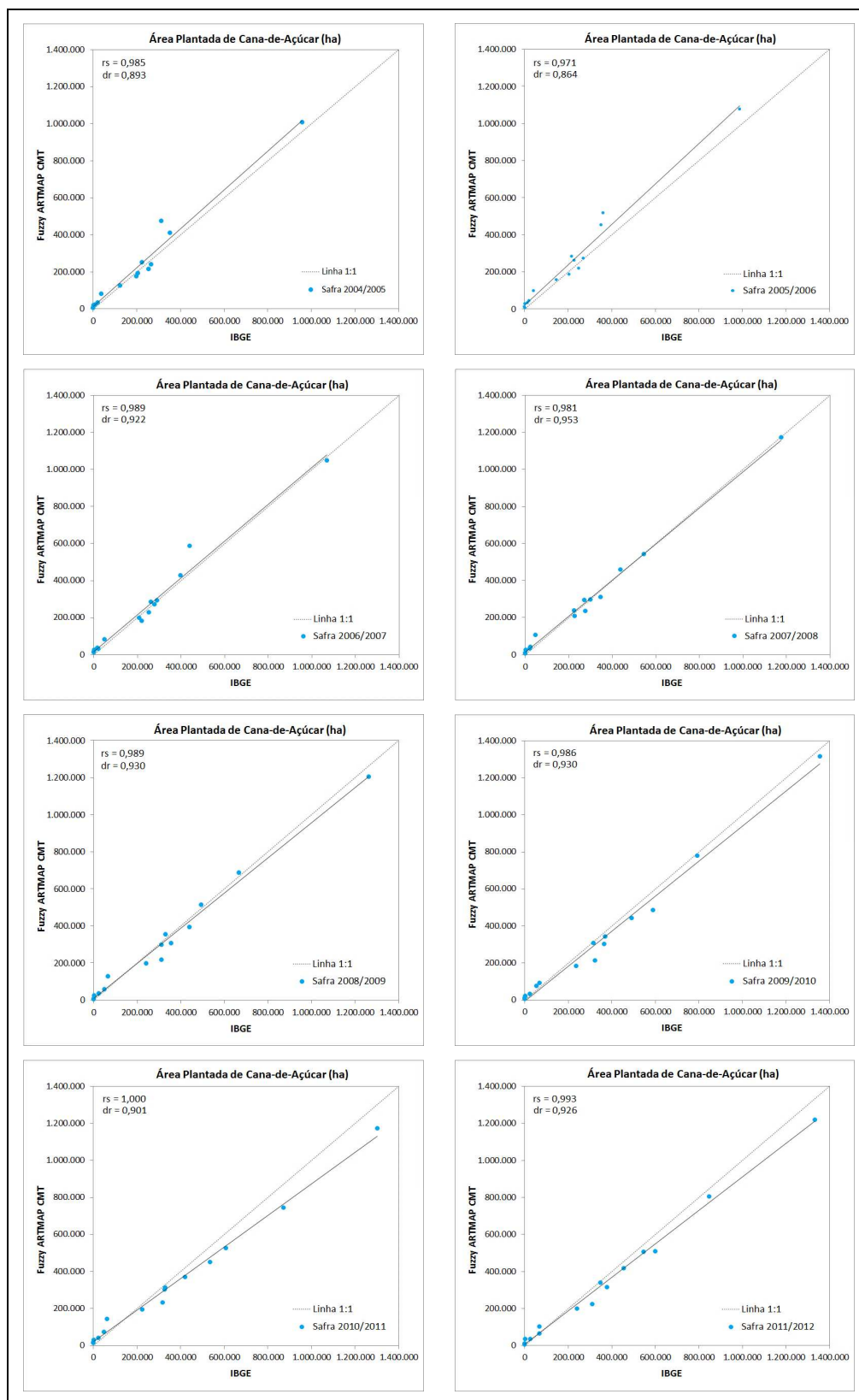


Figura 4.32: Relação entre a área plantada de cana-de-açúcar do IBGE e do modelo Fuzzy ARTMAP CMT em nível de mesorregiões das oito safras.

Os ajustes lineares realizados entre os dados do IBGE e do Canasat em nível de mesorregiões, tiveram a hipótese de aleatoriedade rejeitada (p -valor $< 0,05$), isto é, a correlação foi estatisticamente significativa a 5% para todas as safras em estudo. A discussão desses resultados é baseada nas informações mostradas nos gráficos da Figura 4.30.

Os valores do coeficiente de correlação de Spearman (r_s) variaram entre 0,967 (safra 2006/2007) a 0,993 (safras 2009/2010, 2010/2011 e 2011/2012), indicando que as estimativas de área plantada de cana-de-açúcar são fortemente correlacionadas e possuem a mesma tendência, movendo-se em perfeita proporção na mesma direção. Esses resultados comprovam a excelente precisão do ajuste na relação entre os dados.

Os valores do índice de concordância de Willmott (d_r) variaram entre 0,928 (safra 2005/2006) e 0,976 (safra 2011/2012), confirmando o excelente desempenho do ajuste, que obteve uma altíssima exatidão entre as estimativas de área plantada de cana-de-açúcar do IBGE e do Canasat.

Os ajustes lineares realizados entre os dados do Canasat e do Fuzzy ARTMAP CMT em nível de mesorregiões, tiveram a hipótese de aleatoriedade rejeitada (p -valor $< 0,05$), isto é, a correlação foi estatisticamente significativa a 5% para todas as safras em estudo. A discussão desses resultados é baseada nas informações mostradas nos gráficos da Figura 4.31.

Os valores do coeficiente de correlação de Spearman (r_s) variaram entre 0,969 (safra 2004/2005) a 0,993 (safras 2009/2010 e 2010/2011), indicando que as estimativas de área plantada de cana-de-açúcar são fortemente correlacionadas e possuem a mesma tendência, movendo-se em perfeita proporção na mesma direção. Esses resultados também comprovam a excelente precisão do ajuste na relação entre os dados.

Os valores do índice de concordância de Willmott (d_r) variaram entre 0,883 (safra 2010/2011) e 0,929 (safra 2007/2008), confirmando o ótimo desempenho do ajuste, que obteve uma alta exatidão entre as estimativas de área plantada de cana-de-açúcar do Canasat e do Fuzzy ARTMAP CMT.

Os ajustes lineares realizados entre os dados do IBGE e do Fuzzy ARTMAP CMT em nível de mesorregiões, tiveram a hipótese de aleatoriedade rejeitada (p -valor $< 0,05$), isto é, a correlação foi estatisticamente significativa a 5% para todas as safras em estudo. A discussão desses resultados é baseada nas informações mostradas nos gráficos da Figura 4.32.

Os valores do coeficiente de correlação de Spearman (r_s) variaram entre 0,971 (safra 2005/2006) a 1,000 (safra 2011/2012), indicando que as estimativas de área plantada de cana-de-açúcar são fortemente correlacionadas e possuem a mesma tendência, movendo-se em perfeita proporção na mesma direção. Novamente, esses resultados comprovam a excelente precisão do ajuste na relação entre os dados.

Os valores do índice de concordância de Willmott (d_r) variaram entre 0,864 (safra 2005/2006) e 0,953 (safra 2007/2008), confirmando o ótimo desempenho do ajuste, que obteve uma alta exatidão entre as estimativas de área plantada de cana-de-açúcar do IBGE e do Fuzzy ARTMAP CMT.

Assim sendo, pode-se dizer que os ajustes obtidos pela comparação pareada entre as estimativas de área plantada de cana-de-açúcar, na consolidação das 15 mesorregiões no Estado de São Paulo, obtiveram excelentes desempenhos, alcançando níveis muito semelhantes em termos de precisão e exatidão para todas as safras em estudo. Esse fato é indicado por valores extremamente altos de correlação e concordância entre as estimativas de área do Canasat e do IBGE (r_s de 0,967 a 0,993; d_r de 0,928 a 0,976), do Fuzzy ARTMAP CMT e do Canasat (r_s 0,969 a 0,993; d_r 0,883 a 0,929) e do Fuzzy ARTMAP CMT e do IBGE (r_s de 0,971 a 1,000; d_r de 0,864 a 0,953).

A metodologia de estimativa de área plantada baseada nas medidas de Compromisso CMT do classificador Fuzzy ARTMAP foi eficiente para o mapeamento da cana-de-açúcar em nível municipal, demonstrando grande potencial para a análise sub-pixel de séries temporais de índices de vegetação MODIS. Tal fato se torna ainda mais evidente quando se fez a consolidação para o nível de mesorregiões, onde os ajustes alcançaram um desempenho ainda superior, validados pelos resultados inovadores conseguidos a partir do monitoramento sistemático do cultivo da cana-de-açúcar ao longo das oito safras.

4.9. Cruzamento dos dados de área plantada de cana-de-açúcar com o ZAE

Os dados de área plantada gerados pelo modelo de classificação Fuzzy ARTMAP CMT obtidos a partir do monitoramento sistemático do cultivo da cana-de-açúcar foram cruzados com os dados do ZAE, permitindo analisar os resultados da dinâmica e da tendência da expansão de área da cana-de-açúcar no Estado de São Paulo.

Como na elaboração do ZAE foram excluídas áreas cultivadas com cana-de-açúcar na safra 2007/2008, torna-se válido somente da safra 2008/2009 em diante. Em função disso, a análise foi realizada para as safras 2008/2009 a 2011/2012, considerando a expansão como a diferença de área plantada de uma safra para a outra, conforme apresentado na Tabela 4.14.

Tabela 4.14: Expansão da área plantada de cana-de-açúcar em relação ao ZAE.

Safras	Área Plantada Cana (ha)	Apta (%)		Aptidão (%)			Uso da Terra (%)		
		Sim	Não	Alta	Média	Baixa	Ac	Ag	Ap
2008/2009	4.432.945	89,09	10,91	74,86	24,50	0,63	64,64	3,82	31,55
2009/2010	4.627.810	88,50	11,50	70,63	28,79	0,58	55,42	4,02	40,56
2010/2011	4.636.712	88,91	11,09	72,27	27,61	0,12	61,19	3,23	35,58
2011/2012	4.800.742	86,37	13,63	76,62	23,21	0,17	63,85	3,45	32,70
MÉDIA	4.143.209	88,22	11,78	73,60	26,03	0,37	61,27	3,63	35,10

Legenda do ZAE: Ac - Agricultura, Ag - Agropecuária, Ap - Pastagem.

De acordo com a Tabela 4.14, pode-se observar que a expansão do cultivo da cana-de-açúcar em regiões indicadas como aptas foi em média de 88,22%, sendo 73,60% em regiões com aptidão agrícola alta, 26,03% em regiões com aptidão agrícola média e 0,37% em regiões com aptidão agrícola baixa.

A dinâmica do cultivo da cana-de-açúcar mostrou que a expansão ocorreu em média de 61,27% em regiões com uso em agricultura, 3,63% em regiões com uso em agropecuária e 35,10% em regiões com uso em pastagem, mostrando a tendência atual da cana-de-açúcar de incorporar mais áreas para produção agrícola do que pastagens.

Cabe ressaltar que os mapas de uso da terra utilizados pelo ZAE são do ano de 2002 e precisariam ser atualizados para representar melhor as mudanças recentes da expansão da cana-de-açúcar no Estado de São Paulo.

Mesmo assim, esses resultados são corroborados por PISSINATO (2014), em que o avanço da cana-de-açúcar em áreas aptas era tido como potencial, especificamente para estas safras analisadas. Nesse período ocorreu um crescimento da produção por contribuição por área no Estado de São Paulo, superior à contribuição por rendimento, sugerindo um aumento na produção extensiva canavieira em regiões ocupadas por outras culturas agrícolas, maior do que em regiões de pastagens, demonstrando que o recente avanço em áreas aptas já utilizadas para produção agrícola é válido. No entanto, apesar da predominância na incorporação de áreas

agrícolas pela cana-de-açúcar, não houve queda na produção de outras culturas, devido ao aumento da produtividade da terra.

4.10. Considerações finais e sugestões

Os resultados deste trabalho mostraram o potencial da análise harmônica no estudo do comportamento cíclico dos índices de vegetação, cujo entendimento foi importante para evidenciar alterações no desenvolvimento da cana-de-açúcar e revelar padrões da dinâmica temporal para a tarefa de classificação de imagens.

As representações HLS dos termos harmônicos das séries temporais de índices de vegetação MODIS indicaram que o cultivo da cana-de-açúcar no Estado de São Paulo possui padrões espaciais coerentes com o ciclo vegetativo, consistente com a variabilidade das precipitações sazonais que afetam diretamente o momento de máximo EVI e NDVI.

A Transformada de Wavelet mostrou ser uma técnica robusta, pois conseguiu eliminar os ruídos, propiciando, assim, uma melhor captura das tendências dos ciclos de desenvolvimento da cana-de-açúcar em toda a série temporal do EVI2, para a detecção de padrões do ciclo fenológico da cana de ano predominante na área de estudo, cuja caracterização possibilitou observar as variações interanuais.

As métricas fenológicas sazonais sínteses, nível base e taxa de senescência extraídas das séries temporais de EVI e NDVI, foram preditores úteis no mapeamento da cana-de-açúcar, pois propiciaram um aumento de acurácia no modelo, auxiliando na separabilidade entre as classes de cobertura do solo.

Os mapas temáticos de referência do Canasat mostraram-se fundamentais e bastante confiáveis na delimitação das amostras de treinamento a serem utilizadas pelo classificador e, também, das amostras de teste para avaliação da acurácia do modelo de classificação.

A classificação Fuzzy ARTMAP das safras 2004/2005 a 2011/2012 produziu modelos ART de qualidade temática muito boa, consistente com os altos valores de exatidão e os baixos valores de discordância. A combinação ótima dos parâmetros de ART_a com a taxa de treinamento igual a 0,93 e o parâmetro de vigilância igual 0,94 foi a melhor generalização dos modelos ART obtendo a maior acurácia.

Os modelos CMT Cana gerados pela classificação Fuzzy ARTMAP das oito safras alcançaram uma forte correlação linear com as proporções Cana de referência, conseguido com a abordagem de pós-processamento, considerando o contexto de vizinhança da soma normalizada de 3 x 3 pixels vizinhos mais próximos e com mascaramento das áreas sem cultivo da cana-de-açúcar. Além disso, a área sob a curva ROC dos modelos CMT Cana foi notavelmente superior aos modelos aleatórios, comprovando que realmente é plausível mapear a cultura da cana-de-açúcar baseado nas medidas de Compromisso.

Aliar uma boa acurácia dos modelos de classificação ART com uma alta precisão dos modelos de classificação CMT Cana consistiu numa condição imprescindível para propiciar a obtenção de estimativas de área plantada de cana-de-açúcar confiáveis.

Os ajustes obtidos entre as estimativas de área plantada de cana-de-açúcar alcançaram um desempenho superior com alta precisão e exatidão em nível municipal. No entanto, verificou-se que o desempenho do classificador Fuzzy ARTMAP está intrinsecamente relacionado à caracterização da distribuição geográfica do cultivo da cana-de-açúcar nas mesorregiões do Estado de São Paulo. As mesorregiões de cultivo muito expressivo alcançaram os melhores ajustes. Outras mesorregiões de cultivo muito expressivo, mas com grande expansão apresentaram ótimos ajustes. As mesorregiões de cultivo expressivo com expansão menos significativa tiveram bons ajustes. Os piores ajustes foram encontrados para as mesorregiões onde não predomina o cultivo da cana-de-açúcar.

Na consolidação para o nível de mesorregiões, os ajustes obtidos entre as estimativas de área plantada de cana-de-açúcar alcançaram um desempenho ainda superior, indicado por valores extremamente altos de correlação e concordância.

A metodologia de estimativa de área plantada baseada nas medidas de Compromisso CMT do classificador Fuzzy ARTMAP mostrou ser viável, tendo como grandes vantagens ser realizada de forma objetiva, com baixo custo e bastante rapidez, além de propiciar o conhecimento da distribuição espacial do cultivo da cana-de-açúcar. Portanto, pode ser empregada com o propósito de auxiliar os órgãos oficiais responsáveis pelas estatísticas de produção agrícola no Brasil.

Em estudos futuros, para melhorar os resultados do classificador Fuzzy ARTMAP, sugere-se utilizar amostras de treinamento selecionadas por classes mais específicas de uso da terra, com o propósito de refletir melhor a variabilidade espectral entre as classes e, assim, auxiliar no processo de aprendizagem para o reconhecimento de padrões.

Além disso, seria interessante explorar outros modelos de classificação sub-pixel, como por exemplo, os algoritmos de redes neurais SOM (*Self-Organizing Map*) e também SVM (*Support Vector Machine*).

5. CONCLUSÕES

A metodologia de estimativa de área plantada baseada nas medidas de Compromisso CMT do classificador Fuzzy ARTMAP foi eficiente para o mapeamento da cultura da cana-de-açúcar, demonstrando grande potencial para a análise sub-pixel de séries temporais de índices de vegetação MODIS, pois possibilitou realizar o monitoramento sistemático do cultivo da cana-de-açúcar considerando as safras 2004/2005 a 2011/2012 analisadas no Estado de São Paulo, comprovando a hipótese proposta no trabalho.

As estimativas de área plantada de cana-de-açúcar foram obtidas até o final de março que é o período de maior vegetativo da cultura, com antecedência em relação à colheita, ao contrário dos levantamentos oficiais que se estendem até o final da safra, além de utilizarem dados subjetivos vindos do campo.

Tanto em nível municipal como por mesorregião, as estimativas de área plantada de cana-de-açúcar do Canasat se aproximaram mais do IBGE, porque obtiveram os maiores indicadores de correlação e concordância. Já as estimativas de área plantada de cana-de-açúcar do Fuzzy ARTMAP CMT se aproximaram mais do Canasat, além de também estarem bem balizadas com as do IBGE, determinado pelos altos indicadores de correlação e de concordância. O nível dos desvios entre as estimativas de área, em termos de magnitude e dimensão, mostrou uma variação média municipal menor em relação aos dados do Canasat do que os dados do IBGE.

Na comparação da área plantada de cana-de-açúcar das safras de 2008/2009 a 2011/2012 com os dados do ZAE para o Estado de São Paulo, observou-se que a expansão do cultivo da cana-de-açúcar predomina essencialmente em regiões aptas, com a tendência recente de incorporar mais áreas para produção agrícola do que pastagens.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ADAMI, M. **Estimativa da data de plantio da soja por meio de séries temporais de imagens MODIS**. 163p. Tese (Doutorado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2010.

ADAMI, M.; MELLO, M. P.; AGUIAR, D. A.; RUDORFF, B. F. T.; SOUZA, A. F. A Web platform development to perform thematic accuracy assessment of sugarcane mapping in South-Central Brazil. **Remote Sensing**, v. 4, n. 10, p. 3201-3214, 2012.

AGUIAR, D. A.; RUDORFF, B. F. T.; SILVA, W. F.; ADAMI, M.; MELLO, M. P. Remote sensing images in support of environmental protocol: monitoring the sugarcane harvest in São Paulo State, Brazil. **Remote Sensing**, v. 3, n. 12, p. 2682-2703, 2011.

ALCANTARA, C.; KUEMMERLE, T.; PRISHCHEPOV, A. V.; RADELOFF, V. C. Mapping abandoned agriculture with multi-temporal MODIS satellite data. **Remote Sensing of Environment**, v. 124, p. 334-347, 2012.

ANJOS, I. A. dos.; FIGUEIREDO, P. A. M. de. **Aspectos fitotécnicos do plantio**. In: Cana-de-açúcar. - DINARDO-MIRANDA, L. L.; VASCONCELOS, A. C. M. de; LANDELL, M. G. de A. (Eds.) Campinas: Instituto Agrônômico, 1ª. edição, 882p., 2010.

ANTUNES, J. F. G. **Aplicação de lógica fuzzy para estimativa de área plantada da cultura de soja utilizando imagens AVHRR-NOAA**. 91p. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Estadual de Campinas, 2005.

APAN, A.; HELD, A.; PHINN, S.; MARKLEY, J. Detecting sugarcane 'orange rust' disease using EO-1 Hyperion hyperspectral imagery. **International Journal of Remote Sensing**, v. 25, n. 2, p. 489-498, 2004.

ARAÚJO, M. H. S.; CRUZ, C. B. M.; VICENS, R. S. **Levantamento da cobertura vegetal nativa do bioma Mata Atlântica: relatório final - edital PROBIO 03/2004**. Rio de Janeiro, 84p., 2007.

BARALDI, A.; BINAGHI, E.; BLONDA, P.; BRIVIO, P. A.; RAMPINI, A. Comparison of the multilayer perceptron with neuro-fuzzy techniques in the estimation of cover class mixture in remotely sensed data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 39, n. 5, p. 994-1005, 2001.

BENVENUTI, F. A. **Relação de índices espectrais de vegetação com a produtividade da cana-de-açúcar e atributos edáficos.** 120p. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Estadual de Campinas, Campinas, SP, 2005.

BERNARD, A. C.; WILKINSON, G. G.; KANELLOPOLOUS, I. Training strategies for neural network soft classification of remotely-sensed imagery. **International Journal of Remote Sensing**, v. 18, n. 8, p. 1851-1856, 1997.

BEZDEK, J. C.; PAL, S. K. **Fuzzy models for pattern recognition.** New York: IEEE Press, 1992.

BISHOP, C. M. **Neural networks for pattern recognition.** Oxford, New York: Oxford University Press, 1 edition, 482p., 1995.

BRASIL. Secretaria de Produção e Agroenergia. **Plano nacional de agroenergia 2006-2011.** Brasília: Embrapa Informação Tecnológica, 2ª. edição revisada, 110 p., 2006.

BROWN, J. C.; JEPSON, W. E.; KASTENS, J. H.; WARDLOW, B. D. Multitemporal, moderate-spatial-resolution remote sensing of modern agricultural production and land modification in the Brazilian Amazon. **GIScience & Remote Sensing**, v. 44, n. 2, p. 117-148, 2007.

BROWN, J. C.; KASTENS, J. H.; COUTINHO, A. C.; VICTORIA, D. de C.; BISHOP, C. R. Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data. **Remote Sensing of Environment**, v. 130, p. 39-50, 2013.

CAMARGO, A. M. M. P. de; CASER, D. V.; CAMARGO, F. P. de; OLIVETTE, M. P. de A.; SACHS, R. C. C.; TORQUATO, S. A. Dinâmica e tendência da expansão da cana-de-açúcar sobre as demais atividades agropecuárias, Estado de São Paulo, 2001-2006. **Informações Econômicas**, São Paulo, v. 38, n. 3, p. 47-66, 2008.

CANTY, M. J. **Image analysis, classification and change detection in remote sensing: with algorithms for ENVI/IDL.** New York: CRC Press, Taylor & Francis, 2 edition, 471p., 2010.

CARPENTER, G. A. Neural network models for pattern recognition and associative memory. **Neural Networks**, v. 2, n. 4, p. 243-257, 1989.

CARPENTER G. A.; CROSSBERG, S.; REYNOLDS, J. H. ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing. **Neural Networks**, v. 4, n. 5, p. 565-588, 1991.

CARPENTER, G. A., CROSSBERG, S. ; MARKUZON, N.; REYNOLDS, J. H.; ROSEN, D. B. Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. **IEEE Transactions on Neural Networks**, v. 3, n. 5, p. 698-713, 1992.

CASAGRANDE, A. A.; VASCONCELOS, A. C. M. de. **Fisiologia da parte aérea**. In: Cana-de-açúcar. - DINARDO-MIRANDA, L. L.; VASCONCELOS, A. C. M. de; LANDELL, M. G. de A. (Eds.) Campinas: Instituto Agrônômico, 1ª. edição, 882p., 2010.

CHEAVEGATTI-GIANOTTO, A.; ABREU, H. M. C. de; ARRUDA, P.; BESPALHOK FILHO, J. C.; BURNQUIST, W. L.; CRESTE, S.; CIERO, L. di; FERRO, J. A.; FIGUEIRA, A. V. de O.; FILGUEIRAS, T. de S.; GROSSI-DE-SÁ, M. de F.; GUZZO, E. C.; HOFFMANN, H. P.; LANDELL, M. G. de A.; MACEDO, N.; MATSUOKA, S.; REINACH, F. de C.; ROMANO, E.; SILVA, W. J. da; SILVA FILHO, M. de C.; ULIAN, E. C. Sugarcane (*Saccharum X officinarum*): A reference study for the regulation of genetically modified cultivars in Brazil. **Tropical Plant Biology**, v. 4, n. 1, p. 62-89, 2011.

CLARK LABS. **Products**. Disponível em <<http://www.clarklabs.org/index.cfm>>. Acesso em 6 jul. 2014.

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

COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). **Acompanhamento da safra brasileira de cana-de-açúcar: safra 2013/14, quarto levantamento, abr. 2014**. Brasília, p. 1-14, 2014.

COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). **GeoSafras**. Disponível em <http://www.conab.gov.br/conabweb/geotecnologia/html_geosafras/geosafras.html>. Acesso em 16 ago. 2010.

DAUBECHIES, I. **Ten lectures on wavelets**. Philadelphia: Rutgers University and AT&T Bell Laboratories, 9 edition, 357p., 2006.

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

DE WIT, A. J. W.; SU, B. Deriving phenological indicators from SPOT-VGT data using the HANTS algorithm. In: International SPOT-VEGETATION Users Conference, 2nd., **Proceedings...** Belgium: Antwerp, p. 195-201, 2005.

EASTMAN, J. R.; R. M. LANEY. Bayesian soft classification for sub-pixel analysis: a critical evaluation. **Photogrammetric Engineering and Remote Sensing**, v. 68, n. 11, p. 1149-1154, 2002.

EKLUNDH, L.; JÖNSSON, P. **TIMESAT 3.1: software manual**. Lund University and Malmö University, Sweden, 82p., 2012.

EMBRAPA INFORMÁTICA AGROPECUÁRIA. **Banco de produtos MODIS na base estadual brasileira**. Disponível em <<http://www.modis.cnptia.embrapa.br/geonetwork/srv/pt/main.home>>. Acesso em 6 jul. 2014.

EMBRAPA INFORMÁTICA AGROPECUÁRIA; CEPAGRI/UNICAMP. **Agritempo**. Disponível em <<http://www.agritempo.gov.br/agritempo/index.jsp>>. Acesso em 6 jul. 2014.

ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE (ESRI). **Products**. Disponível em <<http://www.esri.com/>>. Acesso em 6 jul. 2014.

ESQUERDO, J. C. D. M. **Utilização de perfis multi-temporais do NDVI/AVHRR no acompanhamento da safra de soja no oeste do Paraná**. 186p. Tese (Doutorado em Engenharia Agrícola) - Universidade Estadual de Campinas, Campinas, SP, 2007.

ESQUERDO, J. C. D. M.; ANTUNES, J. F. G.; ANDRADE, J. C. de. **Desenvolvimento do banco de produtos MODIS na base estadual brasileira**. (Comunicado técnico, 100) - Campinas: Embrapa Informática Agropecuária, 7p., 2010.

ESTATCAMP. **Portal Action**. Disponível em <<http://www.portalaction.com.br/>>. Acesso em 6 jul. 2014.

EXELIS. VISUAL INFORMATION SOLUTIONS. **Products & Services**. Disponível em <<http://www.exelisvis.com/ProductsServices.aspx>>. Acesso em 6 jul. 2014.

FERNANDES, E. **Monitoramento da cultura de cana-de-açúcar no Estado de São Paulo por meio de imagens SPOT Vegetation e dados meteorológicos**. 114p. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Estadual de Campinas, Campinas - SP, 2009.

FERNANDES, J. L.; ROCHA, J. V. ; LAMPARELLI, R. A. C. Sugarcane yield estimates using time series analysis of SPOT Vegetation images. **Scientia Agrícola**, v. 68, n. 2, p. 139-146, 2011.

FOLEY, J. D.; VAN DAM, A. **Fundamentals of interactive computer graphics**. Addison-Wesley Publishing Company, 664 p., 1984.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). **Multiple frame agricultural surveys: current survey based on area and list sampling methods.** Roma: FAO, v.1, 1996.

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. Supervised image classification by MLP and RBF neural networks with and without an exhaustively defined set of classes. **International Journal of Remote Sensing**, v. 25, n. 15, p. 3091-3104, 2004.

FOODY, G. M.; CAMPBELL, N. A.; TRODD, N. M; WOOD, T. F. Derivation and applications of probabilistic measures of class membership from the maximum likelihood classification. **Photogrammetric Engineering and Remote Sensing**, v. 58, n. 9, p. 1335-1341, 1992.

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.

FORTES, C.; DEMATTÊ, A. M. Discrimination of sugarcane varieties using Landsat 7 ETM+ spectral data. **International Journal of Remote Sensing**, v. 27, n. 7, p. 1395-1412, 2006.

FREITAS, R. M.; ARAI, E.; ADAMI, M.; SOUZA, A. F.; SATO, F. Y.; SHIMABUKURO, Y. E.; ROSA, R. R.; ANDERSON, L. O.; RUDORFF, B. F. T. Virtual laboratory of remote sensing time series: visualization of MODIS EVI2 data set over South America. **Journal of Computational Interdisciplinary Sciences**, v. 2, n. 1, p. 57-68, 2011.

FREITAS, R. M.; SHIMABUKURO, Y. E. Combining wavelets and linear spectral mixture model for MODIS satellite sensor time-series analysis. **Journal of Computational Interdisciplinary Sciences**, v. 1, n. 1, p. 51-56. 2008.

GALFORD, G. L.; MUSTARD, J. F.; MELILLO, J.; GENDRIN, A.; CERRI, C. C., CERRI, C. E. P. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. **Remote Sensing of Environment**, v. 112, n. 2, p. 576-587, 2008.

GASCHO, G. J.; SHIH, S. F. **Sugarcane.** In: Crop-water relations. - TEARE, I. D.; PEET, M.M. (Eds.) New York: John Wiley, 547p., 1983.

GENDRIN, A.; LANGEVIN, Y.; BIBRING, J. P.; FORNI, O. A new method to investigate hyperspectral image cubes: an application of the wavelet transform. **Journal of Geophysical Research**, v. 111, n. E10, E10004, 21p., 2006.

GIACCO, F; THIEL, C.; PUGLIESE, L; SCARPETTA, S.; MARINARO, M. Uncertainty analysis for the classification of multi-spectral satellite images using SVMs and SOMs. **IEEE Transactions on Geoscience and Remote Sensing**, v. 48, n. 10, p. 3769-3779, 2010.

GROSSBERG, S. Adaptive pattern classification and universal recoding, I: parallel development and coding of neural feature detectors. **Biological Cybernetics**, v. 23, n. 3, p. 121-134, 1976.

HALL-BEYER, M. IHS Co-representation of circular and non-circular variables using harmonic analysis parameters. **Canadian Journal of Remote Sensing**, v. 33, n. 5, p. 416-421, 2007.

HAYKIN, S. **Redes Neurais: princípios e práticas**. Porto Alegre: Bookman, 2ª. edição, 900p., 2001.

HIRD, J. N.; MCDERMID, G. J. Noise reduction of NDVI time series: an empirical comparison of selected techniques. **Remote Sensing of Environment**, v. 113, n. 1, p.248-258, 2009.

HOFFER, R. M. Biological and physical considerations in applying computer-aided analysis techniques to remote sensor data. In: SWAIN, P. H.; DAVIS, S. M. (Eds.) **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.

HUETE, A.; JUSTICE, C.; LIU, H. Development of vegetation and soil indices for MODIS-EOS. **Remote Sensing of Environment**, v. 49, n. 3, p. 224-234, 1994.

IMMERZEEL, W. W.; GAUR, A.; ZWART, S. J. Integrating remote sensing and a process-based hydrological model to evaluate water use and productivity in a south Indian catchment. **Agricultural Water Management**, v. 95, n. 1, p.11-24, 2008.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Banco de Dados Agregados. Sistema IBGE de Recuperação Automática - SIDRA**. Disponível em <<http://www.sidra.ibge.gov.br/>>. Acesso em 28 jun. 2012.

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

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **Monitoramento da Cana-de-açúcar via imagens de satélite**. Disponível em <<http://www.dsr.inpe.br/laf/canasat/>>. Acesso em 28 jun. 2012.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **Series view**. Disponível em <<https://www.dsr.inpe.br/laf/series/>>. Acesso em 6 jul. 2014.

IPPOLITI-RAMILO, G.; EPIPHANIO, J. C. N.; SHIMABUKURO, Y. E. Landsat-5 Thematic Mapper data for pre-planting crop area evaluation in tropical countries. **International Journal of Remote Sensing**, v. 24, n. 7, p. 1521-1534, 2003.

JACKSON, R. D.; HUETE, A. R. Interpreting vegetation indices. **Preventive Veterinary Medicine**, v. 11, n. 3-4, 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.

JAKUBAUSKAS, M. E.; LEGATES, D. R.; KASTENS, J. H. Harmonic analysis of time-series AVHRR NDVI data. **Photogrammetric Engineering and Remote Sensing**, v. 67, n. 4, p. 461-470, 2001.

JENSEN, J. R. **Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres**. São José dos Campos, SP: Parêntese, 2^a. edição traduzida, 598p, 2009.

JIANG, Z.; HUETE, A. R.; DIDAN, K.; MIURA, T. Development of a two-band enhanced vegetation index without a blue band. **Remote Sensing of Environment**, v.112, n.10, p.3833-3845, 2008.

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

JÖNSSON, P. EKLUNDH, L. TIMESAT - a program for analysing time-series of satellite sensor data. **Computers & Geosciences**, v. 30, n. 8, p. 833-845, 2004.

KAVZOGLU, T.; MATHER, P. M. The use of backpropagation artificial neural networks in land cover classification. **International Journal of Remote Sensing**, v. 24, n. 13, p. 4907-4938, 2003.

KRÜGER, C. A. M. B.; FONTANA, D. C.; MELO, R. W. Estimativa do rendimento de grãos da soja no Rio Grande do Sul usando um modelo agrometeorológico-espectral regionalizado. **Revista Brasileira de Agrometeorologia**, Piracicaba - SP, v. 15, n. 3, p. 210-219, 2007.

LAND PROCESSES DISTRIBUTED ACTIVE ARCHIVE CENTER (LP DAAC). **Data Pool**. Disponível em <https://lpdaac.usgs.gov/data_access/data_pool>. Acesso em 6 jul. 2014.

LANDELL, M. G. de A.; BRESSIANI, J. A. **Melhoramento genético e manejo varietal**. In: Cana-de-açúcar. - DINARDO-MIRANDA, L. L.; VASCONCELOS, A. C. M. de; LANDELL, M. G. de A. (Eds.) Campinas: Instituto Agrônômico, 1ª. edição, 882p., 2010.

LATORRE, M. L.; SHIMABUKURO, Y. E.; ANDERSON, L. O. **Sensor MODIS: produtos para ecossistemas terrestres - MODLAND**. In: O Sensor Modis e suas aplicações ambientais no Brasil - SHIMABUKURO, Y. E.; RUDORFF, B. F. T.; CEBALLOS, J. C. (Coords.) São José dos Campos: Editora Parêntese, SP, Brasil, 2007.

LI, Z. **Development of soft classification algorithms for neural network models in the use of remotely sensed image classification**. 98p. Thesis (Doctor of Philosophy in the Graduate School of Geography) - Clark University, Worcester, 2007.

LI, Z. Fuzzy ARTMAP based neurocomputational spatial uncertainty measures. **Photogrammetric Engineering and Remote Sensing**, v. 74, n. 12, p. 1573-1584, 2008.

LI, Z.; EASTMAN, J. R. Commitment and typicality measures for the Self-Organizing Map. **International Journal of Remote Sensing**, v. 31, n. 16, p. 4265-4280, 2010.

LI, Z.; EASTMAN, J. R. The nature and classification of unlabelled neurons in the use of Kohonen's Self-Organizing Map for supervised classification. **Transactions in GIS**, v. 10, n. 4, p. 599-613, 2006.

LI, Z.; FOX, J. Mapping rubber tree growth in mainland Southeast Asia using time-series MODIS 250 m NDVI and statistical data. **Applied Geography**, v. 32, n. 2, p. 420-432, 2012.

LI, Z.; HUFFMAN, T.; ZHANG, A.; ZHOU, F.; MCCONKEY, B. Spatially locating soil classes within complex soil polygons - mapping soil capability for agriculture in Saskatchewan Canada. **Agriculture, Ecosystems and Environment**, v. 152, n. 5, p. 59-67, 2012.

LIU, W.; SETO, K.C.; WU, E.Y.; GOPAL, S.; WOODCOCK, C. E. ART-MMAP: a neural network approach to subpixel classification. **IEEE Transactions on Geoscience and Remote Sensing**, v. 42, n. 9, p. 1976-1983, 2004.

LUND UNIVERSITY; MALMÖ UNIVERSITY. **TIMESAT**. Disponível em <<http://www.nateko.lu.se/TIMESAT/timesat.asp>>. Acesso em 6 jul. 2014.

MACEDO, I. C. **A energia da cana-de-açúcar: doze estudos sobre a agroindústria da cana-de-açúcar no Brasil e a sua sustentabilidade.** São Paulo: Berlendis & Vertecchia - UNICA, 2ª. edição, 245p., 2007.

MACHADO, H.; LAMPARELLI, R. A. C. Desenvolvimento metodológico para amostragem em cana-de-açúcar, considerando a resposta espectral da biomassa em imagens orbitais. **Revista Brasileira de Cartografia**, v. 59, n. 2, 2007.

MALANGE, F. C. V. **Rede neuro-fuzzy-wavelet para detecção e classificação de anomalias de tensão em sistemas elétricos de potência.** 128p. Tese (Doutorado em Engenharia Elétrica) - Universidade Estadual Paulista, Ilha Solteira, SP, 2010.

MANNAN, B.; ROY, J. Fuzzy ARTMAP supervised classification of multispectral remotely-sensed images. **International Journal of Remote Sensing**, v. 19, p. 767-774, 1998.

MANZATTO, C. V.; ASSAD, E. D.; BACA, J. F. M.; ZARONI, M. J.; PEREIRA, S. E. M. (Orgs.) **Zoneamento agroecológico da cana-de-açúcar: expandir a produção, preservar a vida, garantir o futuro.** (Documentos, 110) - Rio de Janeiro: Embrapa Solos, 55p., 2009.

MARTÍNEZ, B.; GILABERT, M. A. Vegetation dynamics from NDVI time series analysis using the wavelet transform. **Remote Sensing of Environment**, v. 113, n. 9, p. 1823-1842, 2009.

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

MORETTIN, P. A. **Ondas e ondaletas: Da análise de Fourier à análise de ondaletas.** São Paulo: Editora da Universidade de São Paulo, 276p., 1999.

MORETTIN, P. A.; TOLOI, C.M.C. **Análise de séries temporais.** São Paulo: Editora Edgard Blücher, 2ª. edição, 564p., 2006.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA). **MODIS LAND.** Disponível em <<http://modis-land.gsfc.nasa.gov/>>. Acesso em 6. jul. 2014.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. EARTH OBSERVATORY (NASA/EOS). **Normalized Difference Vegetation Index (NDVI).** Disponível em <http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_2.php>. Acesso em 6 jul. 2014.

NEVES, M. F; TROMBIN, V. G. **A dimensão do setor sucroenergético: mapeamento e quantificação da safra 2013/14.** Ribeirão Preto - SP: Markestrat, Fundace, FEA-RP/USP, 2014.

OPEN SOURCE GEOSPATIAL FOUNDATION (OSGEO). **GDAL: Geospatial Data Abstraction Library**. Disponível em <<http://www.gdal.org/>>. Acesso em 28 jun. 2012.

OPEN SOURCE GEOSPATIAL FOUNDATION (OSGEO). **Geonetwork Opensource**. Disponível em <<http://geonetwork-opensource.org/index.html>>. Acesso em 10 nov. 2010.

OPEN SOURCE GEOSPATIAL FOUNDATION (OSGEO). **QGIS: a free and open source Geographic Information System**. Disponível em <<http://qgis.org/en/site/about/index.html>>. Acesso em 6 jul. 2014.

PALUBINSKAS, G.; LUCAS, R. M.; FOODY, G. M.; CURRAN, P. J. An evaluation of fuzzy and texture-based classification approaches for mapping regenerating tropical forest classes from the Landsat-TM data. **International Journal of Remote Sensing**, v. 16, n. 4, p. 747-759, 1995.

PEDRYCZ, W.; GOMIDE, F. **An introduction to fuzzy sets: analysis and design**. Cambridge: MIT Press, 465p., 1998.

PICOLI, M. C. A.; LAMPARELLI, R. A. C.; SANO, E. E; ROCHA, J. V. Imagens multipolarizadas do sensor Palsar/Alos na discriminação das fases fenológicas da cana-de-açúcar. **Pesquisa Agropecuária Brasileira**, v. 47, n. 9, set., p. 1307-1316, 2012.

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

PINO, F. A. Estimativa subjetiva de safras agrícolas. **Informações Econômicas**, São Paulo, v. 31, n. 6, p. 55-58, 2001.

PISSINATO, B. **A cultura de cana-de-açúcar no Estado de São Paulo entre 1950 e 2010: evolução histórica da área e da produtividade**. 161p. Tese (Doutorado em Economia Aplicada) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, 2014.

PONTIUS, R. G. **Gil Pontius' home page**. Clark University. Disponível em <<http://www.clarku.edu/~rpontius/>>. Acesso em 6 jul. 2014.

PONTIUS, R. G.; MILLONES, M. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. **International Journal of Remote Sensing**, v. 32, n. 15, p. 4407-4429, 2011.

PONTIUS, R. G.; PARMENTIER, B. Recommendations for using the Relative Operating Characteristic (ROC). **Landscape Ecology**, v. 29, n. 3: p. 367-382, 2014.

PONZONI, F. J.; SHIMABUKURO, Y. E.; KUPLICH, T. M. **Sensoriamento remoto da vegetação**. São Paulo: Oficina de Textos, 2a. edição, 176p., 2012.

QIU, F.; JENSEN, J. R. Opening the black box of neural networks for remote sensing image classification. **International Journal of Remote Sensing**, v. 205, n. 9, p 1749-1768, 2004.

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

RODRIGUES, J. D. **Fisiologia da cana-de-açúcar**. Unesp, Instituto de Biociências, Botucatu - SP, 99p., 1995.

ROERINK, G. J.; MENENTI, M.; VERHOEF, W. Reconstructing cloudfree NDVI composites using Fourier analysis of time series. **International Journal of Remote Sensing**, v. 21, n. 9, p.1911-1917, 2000.

ROGAN, J.; FRANKLIN, J.; STOW, D.; MILLER, J.; ROBERTS, D. A.; WOODCOCK, C. Mapping land cover modifications over large areas: a comparison of machine learning techniques. **Remote Sensing of Environment**, v. 112, n. 5, p. 2272-2283, 2008.

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.

RUDORFF, B. F. T.; AGUIAR, D. A.; SILVA, W. F.; SUGAWARA, L. M.; ADAMI, M.; MOREIRA, M. A. Studies on the rapid expansion of sugarcane for ethanol production in São Paulo State (Brazil) using Landsat data. **Remote Sensing**, v. 2, n. 4, p. 1057-1076, 2010.

SAKAMOTO, T.; YOKOZAWA, M.; TORITANI, H.; SHIBAYAMA, M.; ISHITSUKA, N.; OHNO, H. A crop phenology detection method using time-series MODIS data. **Remote Sensing of Environment**, v. 96, n. 3-4, p. 366-374, 2005.

SANO, E. E.; BRITO, J. L. S.; FERREIRA, L. G.; ROSA, R. **Mapeamento da cobertura vegetal do bioma Cerrado: relatório final - edital PROBIO 02/2004**. Brasília, 93p., 2007.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. **Analytical Chemistry**, v. 36, n. 8, p. 1627-1639, 1964.

SCARPARI, M. S.; BEAUCLAIR, E. G. F. de. **Anatomia e botânica**. In: Cana-de-açúcar. - DINARDO-MIRANDA, L. L.; VASCONCELOS, A. C. M. de; LANDELL, M. G. de A. (Eds.) Campinas: Instituto Agrônômico, 1ª. edição, 882p., 2010.

SENF, C.; PFLUGMACHER, D.; VAN DER LINDEN, S.; HOSTERT, P. Mapping rubber plantations and natural forests in Xishuangbanna (Southwest China) using multi-spectral phenological metrics from MODIS time series, **Remote Sensing**, v. 5, n. 6, p. 2795-2812, 2013.

SIMÕES, M. dos S.; ROCHA, J. V. LAMPARELLI, R. A. C. Orbital spectral variables, growth analysis and sugarcane yield. **Scientia Agricola**, Piracicaba, v. 66, n. 4, p.451-461, 2009.

SOARES, J. V.; BATISTA, G. T; SHIMABUKURO. Y. E. **Sensor MODIS: histórico e descrição**. In: O Sensor Modis e suas aplicações ambientais no Brasil - SHIMABUKURO, Y. E.; RUDORFF, B. F. T.; CEBALLOS, J. C. (Coords.) São José dos Campos: Editora Parêntese, SP, Brasil, 2007.

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

UNIÃO DA AGROINDÚSTRIA CANAVIEIRA DE SÃO PAULO (UNICA). **UNICADATA**. Disponível em < <http://www.unicadata.com.br/>>. Acesso em 30 jun. 2014.

VAN NIEL, T. G.; MCVICAR, T. R.; DATT, B. On the relationship between training sample size and data dimensionality: Monte Carlo analysis of broadband multi-temporal classification. **Remote Sensing of Environment**, v. 98, n.4, p. 468-480, 2005.

VERBESSELT, J.; HYNDMAN, R.; NEWNHAM, G.; CULVENOR, D. Detecting trend and seasonal changes in satellite image time series. **Remote Sensing of Environment**, v. 114, n. 1, p. 106-115, 2010.

VICENTE, L. E.; GOMES, D.; VICTORIA, D. de C.; GARÇOM, E. A. M.; BOLFE, E. L.; ANDRADE, R. G.; SILVA, G. B. S. Séries temporais de NDVI do sensor SPOT Vegetation e algoritmo SAM aplicados ao mapeamento de cana-de-açúcar. **Pesquisa Agropecuária Brasileira**, v. 47, n. 9, set., p.1337-1345, 2012.

XAVIER, A. C.; RUDORFF, B. F. T.; SHIMABUKURO, Y. E.; BERKA, L. M. S. B.; MOREIRA, M. A. Multi-temporal analysis of MODIS data to classify sugarcane crop. **International Journal of Remote Sensing**, v. 27, n. 4, p. 755-768, 2006.

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

WARDLOW, B. D.; EGBERT, S. L.; KASTENS, J. H. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. **Remote Sensing of Environment**, v.108, n. 3, p. 290-310, 2007.

WASILEWSKI, F. **PyWavelets: wavelet properties browser**. Disponível em <<http://Wavelets.pybytes.com>>. Acesso em 6 jul. 2014.

WASSERMAN, P. D. **Neural computing: theory and practice**. New York, Van Nostrand Reinhold, 230p., 1989.

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.

WILLMOTT, C. J.; ACKLESON, S. G.; DAVIS, R. E. Statistics for the evaluation and comparison of models. **Journal of Geophysical Research**, v. 90, n. C5, p. 8995-9005, 1985.

WILLMOTT, C. J.; ROBESON, S. M.; MATSUURA, K. A refined index of model performance. **International Journal of Climatology**, v. 32, n. 13, p. 2088-2094, 2012.

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.

ZHANG, J; FOODY, G. M. Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: statistical and artificial neural network approaches. **International Journal of Remote Sensing**, 22: 615-628, 2001.

ZHANG, X.; SUN, R.; ZHANG, B.; TONG, Q. Land cover classification of the North China Plain using MODIS_EVI time series. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 63, n. 4, p. 476-484, 2008.