

UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA AGRÍCOLA

**MONITORAMENTO DA CULTURA DE CANA-DE-AÇÚCAR  
NO ESTADO DE SÃO PAULO POR MEIO DE IMAGENS  
SPOT VEGETATION E DADOS METEOROLÓGICOS**

**JEFERSON LOBATO FERNANDES**

CAMPINAS  
MAIO DE 2009

UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA AGRÍCOLA

**MONITORAMENTO DA CULTURA DE CANA-DE-AÇÚCAR  
NO ESTADO DE SÃO PAULO POR MEIO DE IMAGENS  
SPOT VEGETATION E DADOS METEOROLÓGICOS**

Dissertação apresentada à Faculdade de Engenharia Agrícola como requisito para obtenção do título de Mestre em Engenharia Agrícola na área de Planejamento e Desenvolvimento Rural Sustentável, com a temática Geotecnologias Aplicadas ao Monitoramento de Áreas Agrícolas e Previsão de Safras.

**JEFERSON LOBATO FERNANDES**

**Orientador: Prof. Dr. Jansle Vieira Rocha**

**Co-orientador: Prof. Dr. Rubens Augusto Camargo Lamparelli**

CAMPINAS  
MAIO DE 2009

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

F391m      Fernandes, Jeferson Lobato  
                 Monitoramento da cultura de cana-de-açúcar no  
                 estado de São Paulo por meio de imagens spot vegetation  
                 e dados meteorológicos / Jeferson Lobato Fernandes. --  
                 Campinas, SP: [s.n.], 2009.

                 Orientadores: Jansle Vieira Rocha, Rubens Augusto  
                 Camargo Lamparelli.

                 Dissertação de Mestrado - Universidade Estadual de  
                 Campinas, Faculdade de Engenharia Agrícola.

                 1. Sensoriamento remoto. 2. Séries temporais. 3.  
                 Estimativa. 4. Produtividade agrícola. I. Rocha, Jansle  
                 Vieira. II. Lamparelli, Rubens Augusto Camargo. III.  
                 Universidade Estadual de Campinas. Faculdade de  
                 Engenharia Agrícola. IV. Título.

Título em Inglês: Monitoring sugar cane crop plantations in São Paulo state using spot  
vegetation images and meteorological data

Palavras-chave em Inglês: Remote sensing, Time series, Estimation, Agricultural  
productivity

Área de concentração: Planejamento e Desenvolvimento Rural Sustentável

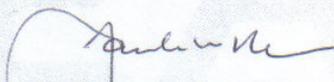
Titulação: Mestre em Engenharia Agrícola

Banca examinadora: Julio Cesar Dalla Mora Esquerdo, Glauco de Souza Rolim

Data da defesa: 29/05/2009

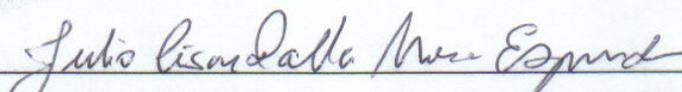
Programa de Pós Graduação: Engenharia Agrícola

Este exemplar corresponde à redação final da **Dissertação de Mestrado** defendida por **Jeferson Lobato Fernandes**, aprovada pela Comissão Julgadora em 29 de maio de 2009, na Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas.



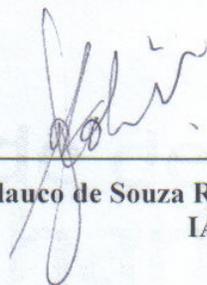
---

**Prof. Dr. Jansle Vieira Rocha – Presidente e Orientador**  
**Feagri/Unicamp**



---

**Dr. Julio Cesar Dalla Mora Esquerdo – Membro Titular**  
**Embrapa/CNPTIA**



---

**Dr. Glaucio de Souza Rolim - Membro Titular**  
**IAC**

## **Dedicatória**

*Dedico este trabalho à minha família,  
em especial aos meus pais, Luiz e Vani,  
que sempre me deram condições e incentivo  
para buscar meus sonhos.*

## **Agradecimentos**

Agradeço,

à Prof. Dra. Daniella Jorge de Moura, em nome do programa de pós-graduação da Faculdade de Engenharia Agrícola da Unicamp, pela possibilidade de realização deste mestrado;

ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, pela bolsa concedida;

ao meu orientador Prof. Dr. Jansle Vieira Rocha, por me aceitar como orientado e acreditar em meu trabalho, pela orientação ao longo do trabalho e por me despertar para outras perspectivas de olhares para a vida;

ao meu co-orientador Prof. Dr. Rubens Augusto Camargo Lamparelli, por ter dado o passo inicial para minha entrada ao “time Labgeo” e também pelas importantes contribuições ao longo do trabalho;

ao Prof. Dr. Nilson Antonio Modesto Arraes, por ter me incentivado a continuar no caminho da pesquisa;

a todos funcionários da Comissão de Pós-Graduação da Faculdade de Engenharia Agrícola da Unicamp, pela competência e empenho que facilitam a vida do aluno;

ao Dr. Júlio César Dalla Mora Esquerdo, amigo que viabilizou o processamento dos dados por meio de suas brilhantes rotinas em IDL;

aos Profs. Drs. Stanley Robson de Medeiros e Luis Henrique Antunes Rodrigues, por me mostrar as infinitas possibilidades de tratamento de banco de dados por meio da mineração de dados, além das valiosas contribuições;

à banca examinadora da qualificação, Drs. Glauco de Souza Rolim e Maurício dos Santos Simões, que avaliaram criteriosamente meu projeto, o que contribuiu bastante para sua melhora;

aos queridos amigos de sala: Michelle Picoli, Gleyce Dantas e Jerry Johann, pelos ótimos momentos de nossa harmoniosa convivência, pelas trocas de idéias e também pelas contribuições ao meu trabalho;

ao funcionário e amigo Agmon Moreira Rocha, que sempre fez de tudo para que os equipamentos do Labgeo estivessem em bom funcionamento;

às amigadas feitas ao longo do mestrado: Tânia, Vânia, Laura, Rojane, Pedro, André, Anderson, Márcio, Hermógenes, Douglas, Adriana, Conan, Carol, Júnior, Karol, Zé Roberto, Marcílio, Iris, Sérgio, Cris, Micheli, outra Tânia;

aos meus amigos da vida e de sempre;

e à Deus.

Levarei para sempre ótimas lembranças, principalmente dos momentos de descontração durante o cafezinho, os deliciosos bolos de quinta-feira, happy-hours, cinegeo e geovolei!

Obrigado!

## SUMÁRIO

<b>LISTA DE FIGURAS.....</b>	<b>ix</b>
<b>LISTA DE TABELAS.....</b>	<b>xii</b>
<b>RESUMO.....</b>	<b>xiii</b>
<b>ABSTRACT .....</b>	<b>xv</b>
<b>1. INTRODUÇÃO .....</b>	<b>1</b>
<b>2. REVISÃO BIBLIOGRÁFICA .....</b>	<b>4</b>
2.1. A importância da cana-de-açúcar.....	4
2.1.1 Fisiologia .....	5
2.2. O sensoriamento remoto aplicado à agricultura.....	10
2.2.1. Resoluções temporal e espacial de sistemas sensores.....	11
2.2.2. Séries multitemporais .....	13
2.2.3. Resposta espectral da vegetação .....	14
2.2.4. Índices de Vegetação .....	15
2.2.5. O sistema SPOT <i>Vegetation</i> .....	16
2.3. Meteorologia .....	17
2.3.1. O modelo ECMWF .....	18
2.4. A previsão de safras no Brasil.....	20
2.4.1. A previsão de safras para cana-de-açúcar .....	21
2.5. O uso de sensoriamento remoto e dados meteorológicos no monitoramento da cana-de-açúcar e previsão de safras .....	22
2.6. Mineração de dados.....	24
<b>3. MATERIAL E MÉTODOS.....</b>	<b>25</b>
3.1. Área e período de estudo.....	25
3.2. Aptidão edafoclimática para a cana-de-açúcar no estado de São Paulo .....	27
3.3. Dados espectrais: sistema SPOT <i>Vegetation</i> .....	28
3.3.1. Suavização dos perfis temporais do NDVI.....	29
3.4. Identificação das áreas canavieiras .....	30
3.5. Dados meteorológicos: modelo do ECMWF .....	34
3.6. Perfis temporais.....	34

3.7. Relação entre NDVI, condições meteorológicas e produtividade.....	36
3.7.1 Geração dos atributos espectrais e meteorológicos a partir das fases de desenvolvimento .....	36
3.7.2 Relações entre atributos meteorológicos, espectrais e a produtividade da cultura ....	36
<b>4. RESULTADOS E DISCUSSÃO .....</b>	<b>38</b>
4.1. Identificação das áreas canavieiras .....	38
4.1.1. Escala estadual .....	38
4.1.2. Escala municipal .....	41
4.1.2.1. Primeira abordagem.....	41
4.1.2.2. Segunda abordagem.....	46
4.2. Relação entre NDVI, dados meteorológicos e produtividade .....	52
4.2.1 Determinação das fases de desenvolvimento.....	52
4.2.2. Geração dos atributos espectrais e meteorológicos a partir das fases de desenvolvimento .....	54
4.2.3. Correlações entre NDVI e variáveis meteorológicas.....	56
4.2.3.1. Determinação do início do ciclo de desenvolvimento.....	56
4.2.4. Classificação da produtividade por meio dos atributos espectrais e meteorológicos .....	59
4.2.4.1. Discretização do atributo classe .....	60
4.2.4.2. Seleção de atributos .....	60
4.2.4.3. Classificação por meio de árvore de decisão.....	62
<b>5. CONCLUSÕES.....</b>	<b>69</b>
<b>6. CONSIDERAÇÕES E SUGESTÕES.....</b>	<b>70</b>
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>72</b>
<b>ANEXOS .....</b>	<b>81</b>

## LISTA DE FIGURAS

Figura 1: Evolução da produção de cana-de-açúcar no Brasil e no Estado de São Paulo, de 1990 a 2006.....	5
Figura 2: Ciclo de uma cultura de cana-de-açúcar com cinco cortes. Fonte: Modificado de Rodrigues (1995). ....	6
Figura 3: Ciclo de desenvolvimento e acúmulo de matéria seca da cana de ano (CA) e cana de ano-e-meio (CAM) e variações climáticas da região centro sul do Brasil. Fonte: Modificado de Casagrande (1991) e Fortes (2003). ....	8
Figura 4: Representação da resposta espectral do dossel de uma cultura ao longo do ciclo de desenvolvimento. Fonte: Adaptado de Guyot (1989).....	15
Figura 5: Mapa da distribuição da precipitação acumulada entre os dias 01 e 10 de janeiro de 2000 no Brasil proveniente de imagem decendial do modelo do ECMWF. ....	19
Figura 6: Fluxograma das principais etapas desenvolvidas.....	25
Figura 7: Mapa da distribuição da produção da cana-de-açúcar no estado de São Paulo em 2006 e localização espacial dos municípios selecionados. Fonte: IBGE (2008).....	27
Figura 8: Mapa de aptidão edafoclimática plena da cultura da cana-de-açúcar, estado de São Paulo. Fonte: ZONEAMENTO DE CULTURAS BIOENERGÉTICAS NO ESTADO DE SÃO PAULO (2008). ....	28
Figura 9: Exemplo do resultado do algoritmo de suavização Hants para um pixel em área canavieira, aplicado em imagens decendiais do SPOT Vegetation. ....	29
Figura 10: Procedimento de suavização das imagens para cada ano-safra. ....	30
Figura 11: Exemplos de resposta temporal do NDVI ao longo de uma safra para os seguintes alvos: área urbana, cultura anual, cana-de-açúcar, pasto e mata. Perfis obtidos em imagens decendiais suavizadas do SPOT <i>Vegetation</i> . ....	31
Figura 12: Comportamento temporal característico do NDVI para áreas canavieiras, obtido em imagens decendiais suavizadas do SPOT Vegetation, e critérios utilizados para identificação destas áreas.....	32
Figura 13: Ilustração do processo de extração dos perfis médios municipais das imagens. ....	35
Figura 14: Gráfico de dispersão, $R^2$ , número do teste que apresentou o melhor resultado e porcentagem de diferença entre área total estimada e oficial de cana colhida no Estado,	

para o melhor resultado de cada safra. Os eixos X e Y dos gráficos representam de área, em unidade de mil hectares.....	39
Figura 15: Mapas de identificação da cultura de cana-de-açúcar no estado de São Paulo para as sete safras analisadas. ....	40
Figura 16: Resultado de identificação de áreas canavieiras em escala municipal para a safra 1999 / 2000 utilizando o teste No 23 da Tabela 3. ....	42
Figura 17: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2000 / 2001 utilizando o teste No 12 da Tabela 3. ....	43
Figura 18: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2001 / 2002 utilizando o teste No 8 da Tabela 3. ....	43
Figura 19: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2002 / 2003 utilizando o teste No 6 da Tabela 3. ....	44
Figura 20: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2003 / 2004 utilizando o teste No 4 da Tabela 3. ....	44
Figura 21: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2004 / 2005 utilizando o teste No 1 da Tabela 3. ....	45
Figura 22: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2005 / 2006 utilizando o teste No 17 da Tabela 3. ....	45
Figura 23: Resultados para cada município referentes à segunda abordagem. ....	49
Figura 24: Mapa indicando para cada município a área relativa de cana-de-açúcar, média entre as sete safras selecionadas. Fonte: IBGE (2008). ....	50
Figura 25: Mapa indicando para cada município o número do teste citado na Tabela 5. ....	51
Figura 26: Perfis médios municipais de NDVI, separados por safra. Sete safras, entre 1999 e 2006. ....	53
Figura 27: Perfis médios de NDVI para cada safra e definição das quatro fases do ciclo de desenvolvimento. ....	54
Figura 28: Correlação entre precipitação acumulada na fase 1 (eixo X, em mm) e decêndio de ocorrência do valor mínimo do NDVI (eixo Y, em decêndios, iniciando no 1º decêndio de agosto). Cada ponto no gráfico é referente a uma safra.....	59

Figura 29: Árvore de decisão para determinação da produtividade obtida aplicando-se aos atributos ndvi2_m, ndvi2_final, ndvi3_s, ndvi3_inic e ndvi_m o algoritmo de classificação J48, com número mínimo de objetos igual a 10. ....	64
Figura 30: Árvore de decisão para determinação da produtividade obtida aplicando-se aos atributos ndvi2_final e ndvi2_m o algoritmo de classificação J48, com número mínimo de objetos igual a 10. ....	66
Figura 31: a) Correlação entre o ndvi final na fase 2 e a produtividade média de cada safra (média entre os 20 municípios). b) Correlação entre o ndvi médio na fase 2 e a produtividade média de cada safra (média entre os 20 municípios).....	67
Figura 32: a) Correlação entre o ndvi final na fase 2 e a produtividade média de cada município - série histórica entre 1999 e 2006. b) Correlação entre o ndvi médio na fase 2 e a produtividade média de cada município - série histórica entre 1999 e 2006.....	68

## LISTA DE TABELAS

Tabela 1: Resolução espacial e temporal de alguns sistemas sensores. ....	12
Tabela 2: Dados de 2006 de produção (toneladas), porcentagem da produção em relação à produção total do Estado, área colhida (hectares), área total do município (hectares) e a área relativa à cana no município (porcentagem), para os 20 municípios selecionados.....	26
Tabela 3: Especificação dos testes realizados para definição da máscara de área colhida de cana-de-açúcar. ....	33
Tabela 4: Testes utilizados para os resultados de identificação de áreas canavieiras em escala municipal, referente à primeira abordagem. ....	42
Tabela 5: Testes utilizados para identificação de áreas canavieiras em escala municipal, referente à segunda abordagem.....	47
Tabela 6: Atributos espectrais e meteorológicos gerados para cada fase de desenvolvimento da cultura e para toda a safra. ....	55
Tabela 7: Limites inferior e superior para as três classes geradas para o atributo produtividade após discretização e classificação. Os valores referem-se à produtividade, em toneladas por hectare.....	60
Tabela 8: Atributos selecionados em cada método de seleção. ....	61
Tabela 9: Resultado da classificação da produtividade, para cada classe, utilizando-se os atributos ndvi2_m, ndvi2_final, ndvi3_s, ndvi3_inic e ndvi_m. ....	63
Tabela 10: Matriz de confusão do classificador J48 para a classificação da produtividade a partir dos atributos ndvi2_m, ndvi2_final, ndvi3_s, ndvi3_inic e ndvi_m.....	63
Tabela 11: : Resultado da classificação da produtividade, para cada classe, utilizando-se os atributos ndvi2_m e ndvi2_final. ....	65
Tabela 12: Matriz de confusão do classificador J48 para a classificação da produtividade a partir dos atributos ndvi2_m, ndvi2_final, ndvi3_s, ndvi3_inic e ndvi_m.....	65
Tabela 13: Produtividade média para cada safra, considerando os 20 municípios selecionados. ....	67

## RESUMO

O atual sistema de previsão de safras para a cultura da cana-de-açúcar utilizado no Brasil depende, em boa parte, de informações subjetivas, baseadas no conhecimento de técnicos do setor e informações da cadeia produtiva. Apesar da experiência dos técnicos envolvidos, este sistema está sujeito a questionamentos sobre a qualidade dos números apresentados e abre margem a ações especulativas por parte de agentes externos. O monitoramento da cultura de cana-de-açúcar por meio de séries multitemporais de índices NDVI do sensor SPOT *Vegetation*, associado a dados meteorológicos provenientes do modelo do ECMWF, viabiliza o acompanhamento do desenvolvimento da cultura e sua correlação com a produtividade, diminuindo assim a subjetividade das estimativas. Este trabalho avaliou o potencial de uso de imagens decendiais do sensor SPOT *Vegetation* para a identificação de áreas canavieiras do estado de São Paulo e a capacidade de inferir sobre os dados de produtividade oficiais registrados em municípios e safras previamente selecionados, baseando-se em dados provenientes dos perfis temporais do NDVI – SPOT *Vegetation* e de variáveis meteorológicas do modelo do ECMWF. Foram selecionados 20 municípios do estado de São Paulo e sete safras compreendidas entre o período de 1999 e 2006. A identificação das áreas canavieiras foi realizada por meio do perfil temporal do NDVI de cada pixel, onde foram selecionados os pixels que ao longo da safra apresentaram comportamento característico da cultura. Para a determinação da produtividade média da safra por meio de dados espectrais e meteorológicos, foram extraídos valores dos perfis temporais do NDVI e das variáveis meteorológicas temperatura mínima, temperatura média, temperatura máxima, precipitação e radiação global, onde, utilizando-se técnicas de mineração de dados, separou-se a produtividade média da safra em 3 classes, baixo-média, média e média-alta. Os resultados de identificação de áreas canavieiras no estado de São Paulo foram avaliados em escala estadual e municipal. Para a escala estadual, o método proposto apresentou resultados satisfatórios. Os índices de determinação ( $R^2$ ), obtidos através da correlação entre a área colhida de cana-de-açúcar oficial e a estimada, por município e safra, considerando os 637 municípios do Estado, variaram entre 0,771 e 0,829. Para a escala municipal não foi possível determinar com exatidão a área de cana-de-açúcar colhida de cada município devido à baixa resolução espacial das imagens. Para o método de classificação da produtividade média municipal proposto, houve maior correlação

com os dados os dados espectrais que com os dados meteorológicos. Valores extraídos do perfil temporal do NDVI entre os meses de dezembro e janeiro permitiram classificar a tendência do resultado de produtividade alcançado no final da safra. Os resultados foram mais efetivos para os casos onde a produtividade foi classificada como sendo média ou acima da média, com acertos de 86,5 e 66,7% respectivamente. Existe potencial no uso de imagens de NDVI do SPOT *Vegetation* para a identificação de áreas canavieiras em escala estadual e para a classificação da produtividade média municipal da safra.

**Palavras-chave:** sensoriamento remoto, séries temporais, produtividade agrícola, estimativa.

## ABSTRACT

The current crop yield forecasting system for sugar cane used in Brazil is dependent on subjective information which are based on personal knowledge of technicians and information from the industry. Despite the experience of technicians, this system is vulnerable once it allows questioning about the quality of the figures presented, what could lead to speculation by market agents. Sugar cane monitoring through multitemporal series of the NDVI data from SPOT Vegetation, associated with meteorological data from the ECMWF model, provides information about the plant development and its correlation with yield, contributing to the decrease in the subjectivity of the forecasts. This study evaluated the potential use of decadal images of the SPOT Vegetation sensor to identify sugar cane areas in São Paulo state and the feasibility to estimate the official productivity registered in municipalities and crops previously selected by the use of NDVI – SPOT Vegetation and ECMWF meteorological data. For this study 20 municipalities of São Paulo state and seven cropping seasons were selected between 1999 and 2006. To identify sugar cane areas the temporal profile of each pixel was analysed during the cropping season, selecting pixels that presented a characteristic temporal NDVI profile of sugar cane. To determine the average yield using spectral and meteorological data, values of the temporal profiles of NDVI and minimum, mean and maximum temperature, precipitation and global radiation were extracted. Through data mining techniques, the average yield was classified in three classes: low-mean, mean and mean-high. Results for the identification of sugar cane areas in São Paulo state were evaluated at regional and local scale. For the regional scale the proposed method showed good results. The  $R^2$  for the correlation between official and estimated sugar cane planted area, considering the 637 municipalities of the state, varied between 0,771 and 0,829. For the local scale, it was not possible to determine the sugar cane area due to the low spatial resolution of the images. For the classification of the average crop yield, by proposed method, spectral data was more relevant than meteorological data. Some values extracted from the temporal profile of NDVI between December and January permitted to classify the yield tendency. Best results were found when the yield was classified as mean or mean-high, with 86,5 and 66,7 percent correctively classified, respectively. The results showed that there is potential for the use of NDVI – SPOT Vegetation images to identify sugar cane areas at regional scale and to classify the average crop yield.

**Keywords:** remote sensing, time series, agricultural productivity, estimates.

## 1. INTRODUÇÃO

A crescente preocupação com o aquecimento global, causado pelo aumento na concentração de gás carbônico da Atmosfera, tem despertado o mundo para a busca por fontes energéticas menos poluidoras. Neste cenário o etanol se destaca por ser uma alternativa viável aos combustíveis fósseis, além de fonte renovável de energia. O Brasil possui grande potencial para a produção de etanol proveniente da cana-de-açúcar, pois possui extensas áreas agricultáveis, clima favorável e conhecimento tecnológico sobre o processo produtivo, sendo atualmente o segundo maior produtor mundial de etanol e o maior produtor mundial de cana-de-açúcar.

A importância desta cultura para a economia brasileira aumenta na medida em que se aumenta a produção. A demanda pelo etanol brasileiro vem aumentando. No mercado interno isto se deve principalmente ao aumento na frota de veículos bicompostíveis. No mercado externo as variações de preço do barril de petróleo e prerrogativas do Protocolo de Kyoto devem estimular a exportação do etanol do Brasil. Dessa forma, a produção de cana-de-açúcar no Brasil, que há tempos se expande, deverá manter esta tendência. No período entre 1990 e 2007 a produção brasileira cresceu cerca de 80%. No estado de São Paulo, atualmente o maior produtor, o aumento foi em torno de 100% para o mesmo período.

O governo necessita de métodos eficazes para o acompanhamento da produção da cana-de-açúcar, gerando estatísticas confiáveis que subsidiem as negociações internacionais e evitem a ação especulativa. No entanto, o sistema atual de previsão de safras utilizado no Brasil é muito dependente de informações subjetivas, obtidas por meio da avaliação de técnicos que percorrem os canaviais e coleta de informações entre agentes da cadeia produtiva. Apesar da experiência dos agentes envolvidos, este sistema está sujeito a questionamentos sobre a qualidade dos números apresentados.

O uso de técnicas de sensoriamento remoto pode contribuir significativamente no monitoramento de áreas agrícolas, permitindo obter metodologias eficazes e de baixo custo, especialmente no Brasil, que possui grande extensão territorial e volume de produção. Por meio de uma série de imagens é possível acompanhar o desenvolvimento da cultura em grandes áreas. O *Normalized Difference Vegetation Index* (NDVI) é um índice de vegetação, gerado a partir de imagens de satélites, comumente utilizado em estudos de áreas agrícolas e

que fornece informações sobre a biomassa da cultura. O *Systeme Pour L'Observation de la Terre (SPOT) Vegetation* é um sensor a bordo do satélite SPOT4 adequado ao monitoramento de coberturas vegetais. As imagens deste sensor de baixa resolução espacial e alta resolução temporal são disponibilizadas gratuitamente, tendo como um de seus produtos imagens de NDVI.

Informações meteorológicas também são de extrema importância quando se pretende monitorar o comportamento das culturas e correlacionar seu desenvolvimento no campo com resultados de produtividade final alcançados. Porém, no Brasil existem dificuldades de utilização de dados meteorológicos, provenientes de estações meteorológicas, para o monitoramento de culturas. Uma alternativa é o uso de dados meteorológicos estimados por modelos, como o modelo global do *European Center for Medium-range Weather Forecast (ECMWF)*, disponíveis através do banco de dados meteorológicos do *Joint Research Centre (JRC)*, centro de pesquisas da Comissão Européia.

O monitoramento da cultura de cana-de-açúcar por meio de séries multitemporais do NDVI – *SPOT Vegetation*, associado a dados meteorológicos provenientes do modelo do ECMWF, viabiliza o acompanhamento do desenvolvimento da cultura e sua correlação com a produtividade.

O objetivo deste trabalho foi avaliar o potencial de uso de imagens decendiais do sensor *SPOT Vegetation* para a identificação de áreas canavieiras do estado de São Paulo e inferir sobre os dados de produtividade oficiais registrados em municípios e safras previamente selecionados, baseando-se em dados provenientes dos perfis temporais do NDVI – *SPOT Vegetation* e de variáveis meteorológicas do modelo do ECMWF.

Para isso, os seguintes objetivos específicos foram definidos:

- avaliar em escala estadual e municipal o resultado da identificação das áreas canavieiras;
- obter o perfil de desenvolvimento da cultura da cana-de-açúcar em 20 importantes municípios produtores do Estado de São Paulo, durante sete safras entre os anos de 1999 e 2006, por meio de séries multitemporais do NDVI provenientes de imagens decendiais captadas pelo *SPOT Vegetation*;
- verificar correlações entre variáveis meteorológicas;
- verificar correlações entre variáveis meteorológicas e espectrais;

- avaliar o potencial de determinação da produtividade média municipal a partir de variáveis meteorológicas e espectrais.

Destaca-se que o método proposto neste trabalho para a identificação das áreas canavieiras a partir dos perfis temporais do NDVI foi inédito e que também se buscou inferir sobre a tendência da produtividade média municipal a partir de insumos gratuitos, como as imagens *SPOT Vegetation* e os dados meteorológicos do ECMWF e usando um software gratuito de mineração de dados.

## **2. REVISÃO BIBLIOGRÁFICA**

### **2.1. A importância da cana-de-açúcar**

Utilizada como matéria-prima na produção de álcool combustível (etanol), a cana-de-açúcar voltou a ocupar lugar de destaque no Brasil e no mundo. No Brasil, os carros bicomustíveis, que representam mais de 75% das vendas, são os grandes responsáveis pelo renascimento do etanol. O etanol é visto mundialmente como um sistema agroindustrial exemplar para a produção de combustíveis renováveis. E tudo isso estimula o setor sucroalcooleiro a aumentar a oferta do produto neste mercado em expansão. Entre os desafios estão o aumento da área plantada e a obtenção de novas variedades mais produtivas (Oliveira e Vasconcelos, 2006).

O setor sucroalcooleiro movimenta cerca de 2% do Produto Interno Bruto (PIB) do Brasil. Segundo dados da FAO referentes ao ano de 2007, o Brasil é o maior produtor mundial de cana-de-açúcar, seguido pela Índia, China, Tailândia e Paquistão (FAOSTAT, 2009). Segundo a CONAB (2008), para a safra 2008/2009 a estimativa nacional de produção de cana-de-açúcar destinada à indústria sucroalcooleira é de 571,4 milhões de toneladas, correspondendo a 13,9% a mais em relação à safra passada. Da produção total de cana-de-açúcar no Brasil, aproximadamente 56,9% são para a produção de etanol. O Estado de São Paulo será responsável por aproximadamente 60% da produção nacional de cana-de-açúcar.

A cana-de-açúcar é uma cultura em expansão no Brasil. Na Figura 1 é mostrada a evolução da produção de cana-de-açúcar no Brasil e no Estado de São Paulo entre 1990 e 2007.

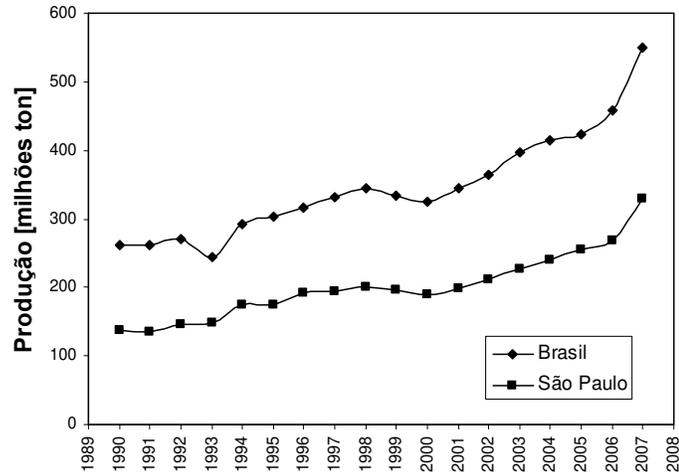


Figura 1: Evolução da produção de cana-de-açúcar no Brasil e no Estado de São Paulo, de 1990 a 2006.

Fonte: IBGE (2008).

De acordo com a CONAB (2008) a área plantada na safra 2008/2009 foi estimada em 7,01 milhões de hectares, 0,7 % superior à safra 2007/2008 e 14% superior à safra 2006/2007.

### 2.1.1 Fisiologia

#### Características gerais

A cana-de-açúcar é uma gramínea pertencente ao gênero *Saccharum*. É cultivada preferencialmente entre as latitudes 15°N e 30°S (Lucchesi, 1995). Não é uma cultura exigente em solos, porém se desenvolve melhor em solos com boa aeração, fertilidade, profundos, argilosos e com boa capacidade de retenção de água (Copersucar, 1988; Benvenuti, 2005).

O sistema radicular varia com a idade do canavial. As raízes da cana da primeira colheita (cana-planta) são mais desenvolvidas que àquelas da cana rebrotada (cana-soca), apesar de a cada rebrota aparecer um sistema radicular novo (Humbert, 1974). A cada corte o sistema radicular se torna mais superficial (Dillewijn, 1952; Castro, 2000).

Rodrigues (1995) afirma que a grande capacidade da cana-de-açúcar para a produção de matéria orgânica está na alta taxa de fotossíntese por unidade de superfície de terreno, que é determinada pelo Índice de Área Foliar (IAF), sendo este um dos mais importantes parâmetros agrônômicos na análise de crescimento. O número de folhas por colmo é afetado

principalmente pelas condições climáticas, aumentando a quantidade em condições climáticas favoráveis (Dillewijn, 1952; Casagrande, 1991).

### Ciclo de desenvolvimento

A cana-de-açúcar é uma cultura semiperene, permitindo cerca de cinco cortes (ciclos de desenvolvimento), para posteriormente ser feita a reforma do canavial. Na Figura 2 é representado um Ciclo de Reforma com cinco cortes.

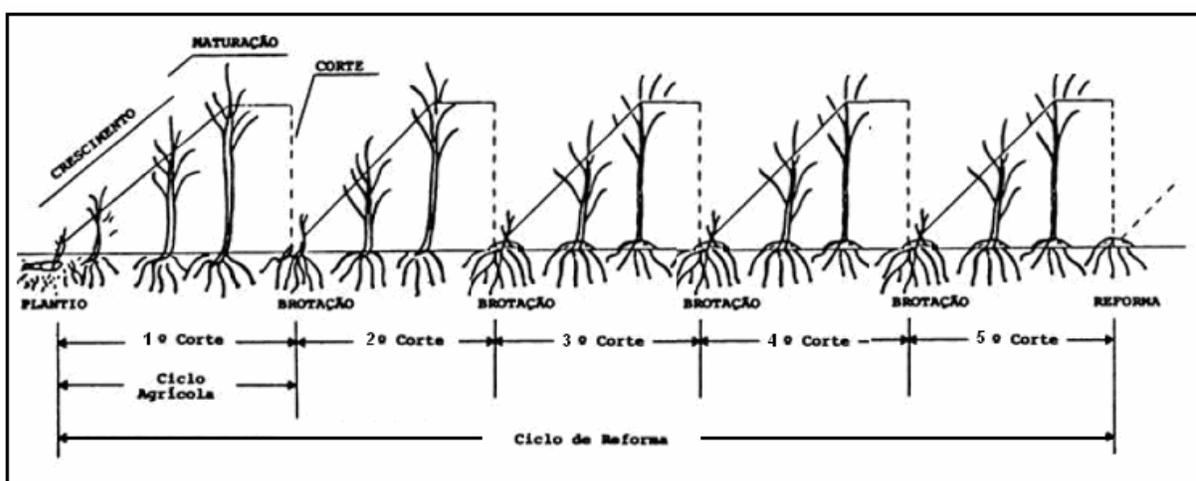


Figura 2: Ciclo de uma cultura de cana-de-açúcar com cinco cortes. Fonte: Modificado de Rodrigues (1995).

O ciclo de desenvolvimento da cana-de-açúcar (chamado de Ciclo Agrícola na Figura 2) pode ser dividido em quatro fases: brotação, perfilhamento, crescimento e maturação. A fase de brotação não será detalhada neste trabalho por ser o estágio mais prematuro da cultura, na qual há predomínio de solo exposto.

O perfilhamento é a formação de outros rebentos no tolete após a brotação das gemas (Dillewijn, 1952). Ocorre até quatro meses após a brotação. O processo de perfilhamento é bastante afetado pelos seguintes fatores: luz, temperatura, água, nutrientes, espaçamento, profundidade do plantio, doenças, pragas e época de plantio (Planalsucar, 1983). Quando atinge o estabelecimento, o autosombreamento induz inibição do perfilhamento e promove a aceleração do crescimento do colmo principal (Rodrigues, 1995).

A fase de crescimento consiste no alongamento dos colmos e acúmulo de matéria seca. Segundo Benvenuti (2005), a distribuição do crescimento da cana-de-açúcar é variável e depende principalmente da época de plantio, que determinará em quais condições climáticas se dará o desenvolvimento da cultura. De acordo com Dillewijn (1952), em um período de estiagem o alongamento dos colmos pode ser reduzidos ou paralisado, resultando em menor volume final. Para a região Centro-Sul do Brasil, uma cana plantada ou cortada em maio passará todo o inverno com crescimento lento, retomando-o quando as condições térmicas e hídricas se tornarem favoráveis. Para uma cana plantada ou cortada em outubro/novembro o crescimento rápido se dará logo nos primeiros meses, para a partir de maio se submeter gradualmente a condições pouco favoráveis. O crescimento em altura continua até a ocorrência de alguma limitação no suprimento de água, ocorrência de baixas temperaturas ou ainda devido ao florescimento, sendo este processo indesejável em culturas comerciais (Rodrigues, 1995).

A fase final corresponde à maturação, na qual, devido a restrições térmicas e hídricas, a taxa de crescimento é pouco intensa e os produtos da fotossíntese são destinados principalmente ao armazenamento da sacarose nos colmos (Benvenuti, 2005). Castro (2000) verificou que a porcentagem de sacarose nos quatro meses antes da colheita aumentou cerca de 200%, enquanto que a massa dos colmos aumentou somente cerca de 5%. Segundo Humbert (1974) nesta fase o número de folhas ativas diminui e o crescimento retarda.

A cana-de-açúcar é denominada cana-planta até a primeira colheita, tendo um período de crescimento entre 12 meses (cana de ano) e 18 meses (cana de ano-e-meio), dependendo da época do plantio e da variedade. Após a primeira colheita ela passa a ser denominada cana-soca, seguida pela ressoça, segunda ressoça, e assim sucessivamente, tendo cada uma dessas fases um período de crescimento de aproximadamente 12 meses (cana de ano). Na Figura 3 é apresentada a dinâmica de cultivo e o acúmulo de matéria seca ao longo do período de crescimento, bem como as variações climáticas para a região Centro-Sul do Brasil.

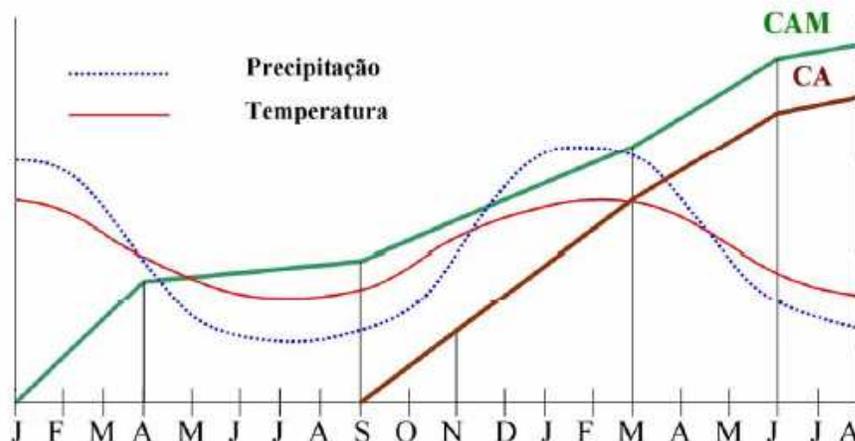


Figura 3: Ciclo de desenvolvimento e acúmulo de matéria seca da cana de ano (CA) e cana de ano-e-meio (CAM) e variações climáticas da região centro sul do Brasil. Fonte: Modificado de Casagrande (1991) e Fortes (2003).

Para a região centro sul do Brasil, a cana de ano, plantada em setembro-outubro, tem sua maior taxa de crescimento de novembro a abril, diminuindo posteriormente devido às condições climáticas adversas do período de inverno, podendo essa ser colhida a partir de julho, variando para cada cultivar. A cana de ano-e-meio, plantada de janeiro ao início de abril, apresenta taxa de crescimento mínima ou nula de maio a setembro em função das condições pouco favoráveis do inverno, como pequena disponibilidade hídrica no solo ou mesmo déficit hídrico, baixas temperaturas e menores intensidades de radiação. Já a fase de maior desenvolvimento da cultura acontece de outubro a abril, com o início das precipitações, aumento da intensidade luminosa e também da temperatura, tendo o pico de crescimento por volta de dezembro a abril (Rodrigues, 1995).

### **Exigências climáticas**

A cana-de-açúcar necessita de luminosidade, calor e água, variando a quantidade de acordo com a fase de desenvolvimento. A seguir seguem alguns comentários sobre os fatores hídricos, térmicos e de radiação.

#### a) condições hídricas:

Em lavouras não irrigadas, como é o caso no Estado de São Paulo, não só a quantidade de precipitação é importante, mas também a distribuição ao longo da safra. Humbert (1974)

relata que excessos durante a estação chuvosa podem reduzir as taxas de crescimento em áreas com drenagem deficiente e também podem diminuir a quantidade de radiação direta disponível para a planta.

A satisfação da necessidade de água desde o plantio/brotação até a fase de máximo crescimento determinará a formação de massa vegetal (Fauconnier e Bassereau, 1975). Benvenuti (2005) afirma que períodos de secas ocasionais irão influenciar decisivamente na produção final da cultura. Em estudo de irrigação de cana-de-açúcar no Estado de São Paulo, Rosenfeld e Leme (1984) concluíram que as maiores reduções na produtividade ocorreram com deficiências hídricas durante os primeiros oito meses do ciclo da cana-planta. Delgado-Rojas e Barbieri (1999) concluíram que no período de maturação a influência do déficit hídrico é muito pequena, tanto para colmo como para açúcar.

b) temperatura do ar:

Segundo Rodrigues (1995) a temperatura do ar exerce grande influência no crescimento dos colmos. A cana-de-açúcar possui uma faixa de temperatura ideal entre 25 e 35°C, sendo que acima ou abaixo desta faixa o crescimento é reduzido. Para a fase de maturação é importante que haja queda na temperatura (Benvenuti, 2005). Fauconnier e Bassereau (1975) relatam que frios noturnos também são importantes para o bom desenvolvimento.

c) radiação:

A quantidade e qualidade da radiação disponível dentro do dossel vegetal afetam a maioria dos processos fisiológicos das plantas (Machado et al., 1985). O metabolismo fotossintético da cana-de-açúcar é C<sub>4</sub>, que caracteriza os vegetais com alto valor de saturação à luz. Segundo Rodrigues (1995), altas intensidades luminosas é fator significativo para altas eficiências fotossintéticas. 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.

## **Variedades**

As características varietais definem a capacidade de brotação de soqueiras, perfilhamento, fechamento de entrelinhas, a altura, o número de colmos por planta, diâmetro do colmo, o comprimento e a largura das folhas, a arquitetura da parte aérea, resistência a pragas e doenças, maturação, produtividade, entre outras características (Rodrigues, 1995; PMGCA, 2008).

Existem variedades de ciclo precoce, médio e tardio, de acordo com a duração da fase de maturação de cada variedade (Rodrigues, 1995). Isso permite que a colheita seja realizada durante um longo período, que normalmente varia de maio (precoce) a novembro (tardio).

De acordo com censo varietal realizado pelo Programa de Melhoramento Genético da Cana-de-Açúcar - PMGCA (2008) em 119 unidades da região centro sul do Brasil, abrangendo uma área total de 2.340.057 ha, 54% são variedades RB, 42% são variedades SP (Centro de Tecnologia Canavieira - CTC), 1% é variedade IAC (Instituto Agrônômico de Campinas) e 3% correspondem a outras variedades.

## **2.2. O sensoriamento remoto aplicado à agricultura**

O sensoriamento remoto pode ser definido como a aquisição de dados físicos de um objeto sem tocá-lo ou haver contato (Lintz e Simonett, 1976). Há mais de 35 anos pesquisadores vêm estudando o potencial de seu uso aplicado à agricultura. Nas décadas de 60 e 70 pesquisadores do *Agricultural Reserch Service (ARS)* do *United States Department of Agriculture (USDA)* investigavam as características morfológicas de culturas e suas propriedades espectrais (Allen et al., 1969; Gausman et al., 1969), assinaturas espectrais para espécies naturais e cultivadas bem como características espectrais de plantas saudáveis e sob condições de stress (Gausman e Allen, 1973; Gausman et al., 1975).

O sensoriamento remoto é uma ferramenta que possibilita monitorar o crescimento e desenvolvimento das culturas e a detecção de condições de stress que limitam a produtividade. Esta tecnologia permite obter informações espaciais e temporais detalhadas sobre o comportamento das culturas nos diferentes ambientes. Diversas pesquisas têm mostrado o potencial de uso do sensoriamento remoto na agricultura, como mapeamento e monitoramento de culturas (Maselli e Rembold, 2001; Pellegrino, 2001; Doraiswamy et al., 2003; Rudorff, 2005; Rocha, 2006; Esquerdo, 2007), manejo de água para irrigação (Millard et al., 1978;

Pinter e Reginato, 1982; Gardner et al., 1992), controle de nutrientes (Scheppers et al., 1996) e ocorrência de pragas e doenças (Hatfield e Pinter Jr, 1993; Apan et al., 2004).

A visão sinótica e a temporalidade do sensoriamento remoto, por possibilitarem gerar mapas de desenvolvimento da cultura e identificar situações de stress durante a safra, permitem diagnosticar a produção de forma mais acurada (Jackson et al., 1986; Pinter Jr., 2003; Apan et al., 2004; Pontes et al. 2005).

Outra característica importante do sensoriamento remoto é que a disponibilidade de imagens para toda a superfície terrestre permite o monitoramento de culturas além das fronteiras nacionais. Doraiswamy et al. (2003) relatam que em 1977 o *Large Area Crop Inventory Experiment* (LACIE), projeto conjunto entre a *National Oceanic and Atmospheric Administration* (NASA), o *United States Department of Agriculture* (USDA) e *National Oceanic and Atmospheric Administration* (NOAA), previram durante a safra uma perda de 30% na produção soviética de trigo, que somente após alguns meses da colheita foi confirmada pelos órgãos oficiais soviéticos como sendo de 10%. Por meio do sensoriamento remoto, em 2006 o projeto Geosafra (2006) buscou estimar a área plantada com arroz na Argentina e Uruguai.

Os órgãos oficiais brasileiros têm interesse em incorporar essa tecnologia em seus métodos de previsão e estimativa de safras (O ESTADO DE SÃO PAULO, 2004; Rudorff et al., 2005).

### **2.2.1. Resoluções temporal e espacial de sistemas sensores**

Na Tabela 1 são apresentadas as resoluções espacial e temporal de alguns sistemas sensores comumente utilizados em pesquisas na agricultura.

Tabela 1: Resolução espacial e temporal de alguns sistemas sensores.

<b>Sensor</b>	<b>Satélite</b>	<b>Resolução Espacial</b>	<b>Resolução Temporal</b>
AVHRR	NOAA	1,1 Km	12 horas
Vegetation	SPOT	1,0 Km	24 horas
MODIS	Terra / Aqua	250 m, 500 m, 1,0 Km	1 ou 2 dias
TM	Landsat 5	30 m	16 dias
CCD	CBERS 2B	20 m	26 dias

A resolução temporal está relacionada com a repetitividade de observação do sensor numa mesma área, ou seja, o intervalo temporal entre revisitas. Para estudos de áreas agrícolas recomenda-se o uso de sensores de alta resolução temporal, a fim de aumentar a periodicidade no monitoramento das culturas e a possibilidade de obter imagens livres de nuvens.

A resolução espacial é definida como a capacidade do sistema sensor em identificar objetos na superfície terrestre, refere-se ao tamanho do pixel da imagem gerada. Assim, quanto menor o tamanho do alvo a ser identificado, maior deve ser a resolução espacial e vice-versa.

Diferentes resoluções espaciais possuem diferentes potenciais de aplicação. Um fator importante para a escolha do sensor, além da questão econômica, é a escala espacial que se pretende realizar o trabalho.

Labus et al., (2002) e Ferencz et al. (2004) avaliaram o uso de sensoriamento remoto para estimativa de produtividade de culturas em diferentes escalas: em nível local (fazenda) e regional (país). Ambos concluíram que os dados de sensores de baixa resolução (NOAA/AVHRR) são adequados ao monitoramento em nível regional, o que não acontece em nível local.

Doraiswamy et al. (2003) utilizaram imagens com diferentes resoluções espaciais (Landsat e NOAA/AVHRR) em um modelo para estimar a produtividade da cultura do trigo no Norte de Dakota-EUA e constataram que, devido à baixa resolução espacial, as imagens AVHRR não foram adequadas para a classificação das áreas cultivadas. Porém, os dados AVHRR possibilitaram detectar stress hídrico que retardou o crescimento da cultura e antecipou a senescência.

Verifica-se assim que, para determinadas aplicações, o uso de dados de sensores de baixa resolução espacial (NOAA/AVHRR e SPOT *Vegetation*) é indicado para estudos em escala regional.

### **2.2.2. Séries multitemporais**

O uso de séries multi-temporais em uma área cultivada permite acompanhar tanto o comportamento da cultura ao longo do desenvolvimento como também comparar diferentes comportamentos de diferentes ciclos da cultura.

A riqueza de informação espectral, espacial e temporal contida em uma seqüência de imagens de sensoriamento remoto oferece uma oportunidade única para o monitoramento e manejo da agricultura (Pinter et al., 2003).

Segundo Labus et al. (2002) a maior parte dos estudos de estimativa de produtividades por meio de índices de vegetação utiliza dados de apenas um subperíodo do ciclo da cultura. Porém, o uso de perfis de desenvolvimento ao longo da safra e a identificação de períodos críticos podem prover a base para o monitoramento e estimativa eficaz da produtividade.

O uso de séries multitemporais de NDVI do sensor AVHRR/NOAA permitiram obter resultados satisfatórios de estimativa da produtividade em escala regional para a cultura do trigo em estudos realizados no Canadá e EUA. Boken e Shayewich (2002) incorporaram o uso de séries multitemporais em modelo de estimativa originalmente baseado em dados de temperatura e precipitação mensais, avaliando os resultados em 5 distritos do Canadá. O uso dos dados espectrais permitiu melhorar índices de determinação ( $R^2$ ) encontrados, passando de 0.13, 0.70, 0.75, 0.50 e 0.00 para 0.79, 0.96, 0.83, 0.95 e 0.39 respectivamente para cada distrito. O uso de séries multitemporais permitiu que Labus et al. (2002) estimassem a produtividade da cultura do trigo nos EUA com resultados de  $R^2$  de até 0.75.

As séries multitemporais também possibilitam a identificação e mapeamento de culturas. Rudorff et al. (2005), utilizando imagens de satélite da série Landsat obtidas em quatro diferentes e específicas épocas, ao longo de um período de doze meses, realizaram com boa confiabilidade o mapeamento das áreas ocupadas com a cultura de cana-de-açúcar no Estado de São Paulo para a safra 2003/04, subdividindo-a em: área de cana disponível para colheita; área de cana em reforma e; área de cana em expansão. A resposta espectral durante a safra

permitiu a diferenciação entre cana de ano e cana de ano-e-meio. Por meio de séries multi-temporais de dados MODIS, Xavier et al. (2006) traçaram os perfis de desenvolvimento da vegetação para o Estado de São Paulo, o que possibilitou identificar e classificar a cultura da cana-de-açúcar em 8 grupos de distinto comportamento temporal do índice de vegetação ao longo do da safra.

### **2.2.3. Resposta espectral da vegetação**

A radiação eletromagnética, ao incidir na planta e interagir com esta, se subdivide em três componentes: a refletida, a absorvida e a transmitida. A resposta espectral da refletância da vegetação refere-se à capacidade da cultura em refletir a radiação eletromagnética nos diferentes comprimentos de onda do espectro.

Os três principais mecanismos biológicos que influenciam a refletância das folhas são: pigmentos, quantidade de água e estrutura celular, além de fatores físicos como o ângulo de incidência da radiação, ângulo de visada e arquitetura da planta. Os mecanismos biológicos respondem de forma distinta de acordo com a disponibilidade de água no ambiente, saúde e idade da planta (Moreira, 2001).

Na Figura 4 é exemplificada a dinâmica da resposta espectral da vegetação para as fases de crescimento e senescência da cultura, nos comprimentos de onda do visível, infravermelho próximo (IVP) e infravermelho médio (IVM). De modo geral, verifica-se que a planta, ao passar da fase de crescimento para a de senescência, aumenta a refletância na faixa do visível e diminui na faixa do IVP devido a alterações nos mecanismos biológicos, se assemelhando assim a resposta espectral do solo. Essa dinâmica possibilita a obtenção dos índices de vegetação, pois, cada banda espectral dos sensores capta a radiação eletromagnética nos diferentes comprimentos de onda.

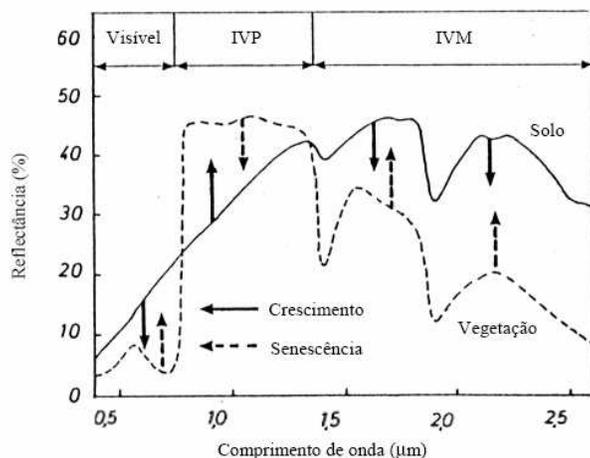


Figura 4: Representação da resposta espectral do dossel de uma cultura ao longo do ciclo de desenvolvimento. Fonte: Adaptado de Guyot (1989).

Guyot (1989) relata que a energia captada pelo satélite proveniente de uma área agrícola é influenciada pelo solo devido às características da cultura, como espaçamento, porte, vigor e fase de desenvolvimento. No início do crescimento da planta o solo exposto aparece entre as plantas jovens. Com o desenvolvimento da cultura a cobertura vegetal passa a predominar, diminuindo assim a influência do solo.

#### 2.2.4. Índices de Vegetação

Segundo Moreira (2001) os índices de vegetação foram criados com o intuito de ressaltar o padrão da variação da vegetação em relação ao solo e a outros alvos da superfície terrestre. Trata-se de transformações lineares do fator de refletância obtido de duas ou mais bandas espectrais do sensor, envolvendo operações matemáticas como soma, diferença, razão ou qualquer outra (Wiegang et al., 1991).

Diversos trabalhos têm reconhecido que o desenvolvimento da planta, condições de stress e o potencial de produção podem ser quantificados por meio de índices de vegetação (Jackson et al., 1986).

Dentre eles, o *Normalized Difference Vegetation Index* (NDVI), proposto por Rouse et al. (1973), é o mais utilizado em trabalhos envolvendo a análise de culturas agrícolas.

O NDVI é calculado da seguinte forma:

$$NDVI = \frac{(IVP - V)}{(IVP + V)} \quad (1)$$

Sendo que:

IVP = refletância na banda do infravermelho próximo;

V = refletância na banda do vermelho.

Pellegrino (2001) e Simões et al. (2005) demonstraram que o uso do NDVI na cana-de-açúcar é interessante quando se busca correlacionar dados espectrais com a produtividade da cultura devido à correlação entre este índice e a quantidade de pigmento clorofila, determinada indiretamente pela variável agrônômica Índice de Área Foliar (IAF) medida no campo.

Lamparelli et al. (2003) obtiveram relação entre o NDVI e a biomassa em área canavieira, verificando o potencial uso do NDVI na estimativa de produtividade para a cana-de-açúcar. Os valores do NDVI variam entre -1 e +1. Em áreas com cobertura vegetal, valores de NDVI próximos a 0 (zero) significam pouca biomassa e valores próximos de +1, maior quantidade de biomassa.

### **2.2.5. O sistema SPOT *Vegetation***

O SPOT *Vegetation* é um sensor a bordo do satélite SPOT4, em operação desde 24 de março de 1998. Este satélite está a uma altura entre 820 e 830 Km. A resolução espacial é de 1,15 Km no nadir, adequadas para monitoramento de grandes áreas com vegetação. A resolução temporal é diária. A largura da imagem é de 2200 Km. Quanto às bandas espectrais de interesse neste trabalho, estas são de 0,61 – 0,68  $\mu\text{m}$  para a banda do vermelho (R) e 0,78 – 0,89  $\mu\text{m}$  para a banda do infravermelho próximo (NIR) (VEGETATION, 2007).

Um dos produtos do SPOT *Vegetation* são as imagens de NDVI, disponibilizadas gratuitamente com intervalo temporal diário, decendial e mensal. Para o uso de imagens diárias, existe uma dificuldade associada à presença de nuvens, principalmente na estação

chuvosa. Nos casos de imagens decendiais e mensais, a probabilidade de se obter uma imagem livre de nuvens aumenta.

### **2.3. Meteorologia**

As condições climatológicas (baseada em valores médios de no mínimo 30 anos dos fenômenos atmosféricos) indicam o tipo de atividade agrícola mais viável em um local, e as condições meteorológicas (baseada nas condições atmosféricas num período específico) determinam o nível de produtividade para aquela atividade, em um certo período, além de interferir na tomada de decisão com relação às diversas práticas agrícolas (Pereira, 2002).

As condições meteorológicas influenciam diretamente no crescimento, desenvolvimento e produtividade das culturas. Os modelos de estimativas de produtividade de culturas utilizam informações meteorológicas, baseado nas respostas das culturas às condições meteorológicas ocorridas nas lavouras (Fontana et al., 2001). De acordo com Rudorff e Batista (1990), flutuações anuais na produtividade, na maioria dos casos, estão relacionadas com as condições meteorológicas ocorridas ao longo da safra, embora outros fatores, como práticas agrícolas e políticas econômicas, também são importantes.

As observações meteorológicas de superfície são realizadas nas estações meteorológicas, que coletam e registram rotineiramente os dados, caracterizando o estado da atmosfera. Segundo Pereira (2002), o Brasil por suas dimensões continentais ainda não apresenta uma rede de estações meteorológicas que atenda às suas necessidades. Há um maior número nas áreas mais desenvolvidas e poucas em áreas mais remotas. Melo e Fontana (2007) relatam que, além da baixa densidade de estações, a defasagem de tempo entre a coleta dos dados e o recebimento destes também dificulta a utilização de dados medidos em estações meteorológicas no monitoramento de culturas no Brasil.

Diversos trabalhos têm mostrado alta correlação entre as condições meteorológicas ocorridas no período de crescimento e desenvolvimento das culturas e a produção agrícola obtida (Berlato et al. 1992; Matzenauer et al., 1995; Delgado-Rojas e Barbieri, 1999; Fontana et al., 2001). Cada cultura possui uma dinâmica de crescimento, existindo subperíodos mais críticos em relação às demandas climáticas. Estresses causados nesses momentos por

alterações nas condições meteorológicas podem comprometer a produtividade da cultura, e assim, a produção final.

O sistema agrícola é complexo, tornando os modelos de simulação complexos devido ao grande número de variáveis envolvidas (Rudorff e Batista, 1990). Uma parte destas variáveis está relacionada a práticas de manejo e condições do solo, sendo estas cada dia melhor controladas por novos sistemas e tecnologias que buscam otimizar o processo produtivo. A outra parte de variáveis está associada ao clima e ao tempo que não são controláveis, exceto em casas de vegetação. Com isso, o monitoramento da relação entre as condições meteorológicas e o desenvolvimento da cultura ganha importância na busca por prever o resultado final de produção.

### **2.3.1. O modelo ECMWF**

No Brasil existem dificuldades de utilização de dados meteorológicos provenientes de estações meteorológicas para o monitoramento de culturas, como o atraso existente entre o momento de coleta e de divulgação do dado e a desuniformidade na densidade de estações distribuídas pelo território brasileiro. Uma alternativa para estas dificuldades é o uso do banco de dados meteorológicos disponibilizados gratuitamente pelo *Joint Research Centre* (JRC), oriundos do modelo do ECMWF (*European Center for Medium-Range Weather Forecast*). São disponibilizados dados simulados de temperatura média, temperatura máxima, temperatura mínima, precipitação (acumulada), evaporação (acumulada, E<sub>0</sub>, solo exposto), evaporação (acumulada, E<sub>0</sub>, sobre a água), evapotranspiração (acumulada, E<sub>T0</sub>, Penman-Monteith) e radiação global (acumulada). Há séries históricas desde 1974, nos intervalos de tempo diário, decendial e mensal. Estes dados são disponibilizados em formato de imagem e apresentam resolução espacial de 0,5 grau de latitude e longitude (JRC, 2007). A utilização dos dados simulados pelo ECMWF apresenta grandes vantagens em relação aos dados provenientes de estações meteorológicas, como rapidez na obtenção dos dados (em tempo quase real) e cobertura para toda a superfície terrestre.

Na Figura 5 é apresentado um mapa de distribuição da precipitação acumulada entre os dias 01 e 10 de janeiro de 2000 no Brasil, proveniente de imagem decendial do modelo do ECMWF.

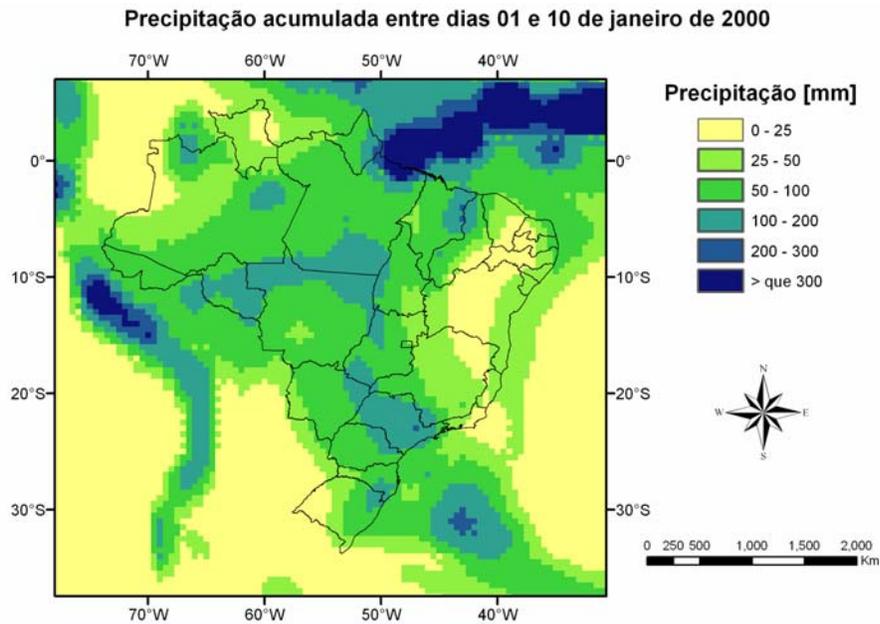


Figura 5: Mapa da distribuição da precipitação acumulada entre os dias 01 e 10 de janeiro de 2000 no Brasil proveniente de imagem decendial do modelo do ECMWF.

Dados decendiais de temperatura e precipitação do modelo do ECMWF foram comparados com dados registrados em estações meteorológicas do estado de São Paulo por Blain et al. (2006) e no estado do Rio Grande do Sul por Melo e Fontana (2006). Para a variável temperatura foram encontradas correlações satisfatórias, com  $R^2$  acima de 0,78 em São Paulo e acima de 0,73 no Rio Grande do Sul. Para a variável precipitação, ambos verificaram grande dispersão entre os dados do modelo e de estações devido à grande variabilidade dos eventos de chuva na região Centro-Sul. Melo e Fontana (2006) obtiveram coeficientes de correlação entre os dados simulados e observados de 0,586 e 0,6 nos períodos decendial e mensal, respectivamente. Analisando qualitativamente, Blain et al. (2006) constataram que praticamente todas as ocorrências de precipitação apontadas pelo modelo do ECMWF também foram registradas nas estações meteorológicas. Deppe et al. (2006) estudaram a correlação entre dados de precipitação decendiais simulados pelo modelo do ECMWF e medidos em estações meteorológicas do Estado do Paraná e verificaram que, apesar da diferença entre valores medidos e estimados, o modelo do ECMWF acompanhou a dinâmica de ocorrência das chuvas ao longo do ano.

Melo e Fontana (2007) avaliaram o desempenho de um modelo agrometeorológico-espectral de estimativa de produtividade da soja no Rio Grande do Sul utilizando dados

meteorológicos de estações meteorológicas e dados simulados pelo ECMWF e concluíram que os dados do ECMWF poderiam ser utilizados como opção para o cálculo da estimativa de produtividade da soja na região maior produtora do Rio Grande do Sul.

#### **2.4. A previsão de safras no Brasil**

O monitoramento das culturas durante a safra e a estimativa de seu potencial de produção são atividades de grande importância, principalmente nos países onde a agricultura representa relevante papel na economia. A previsão de safras é fundamental ao planejamento agrícola do país, pois fornece estatísticas agrícolas que subsidiam a formação dos preços, armazenamento, abastecimento do mercado interno e exportações. De acordo com Esquerdo (2007), previsões de safras confiáveis conduzidas por agentes internos ao país também representam um mecanismo eficaz de proteção da produção nacional, podendo impedir ou reduzir a ação especulativa causada por estimativas geradas por agentes externos, muitas vezes pertencentes a países concorrentes no mercado internacional.

A Companhia Nacional de Abastecimento - CONAB e o Instituto Brasileiro de Geografia e Estatística - IBGE são órgãos oficiais que realizam previsão de produção. Os métodos tradicionais adotados por estes órgãos são considerados subjetivos por se basearem em consulta direta ao setor produtivo, por meio de entrevistas com agentes da cadeia produtiva e levantamentos de campo, além de dados históricos da produção (Picoli et al., 2007; Sugawara et al., 2007; Rizzi e Rudorff, 2007; GEOSAFRAS, 2008). De acordo com Fontana et al. (2000), além de serem subjetivos, os métodos são lentos e onerosos. Luiz e Epiphanyo (2001) e Machado (2003) relatam que outra dificuldade deste método está associada à dificuldade de acesso às áreas cultivadas.

A CONAB vem buscando aprimorar seu método de previsão de safras por meio do projeto Geosafra. Desde 2004 o uso de recursos tecnológicos, como modelos estatísticos, sensoriamento remoto, posicionamento por satélite (GPS) e sistemas de informações geográficas, vem auxiliando as previsões (GEOSAFRAS, 2008).

Pesquisas sobre métodos de aplicação de sensoriamento remoto no monitoramento de safras em escala regional, com baixo custo, poderão contribuir com este aprimoramento.

### **2.4.1. A previsão de safras para cana-de-açúcar**

A previsão de safras para a cana de açúcar pode ser feita em diferentes escalas de acordo com a área de abrangência: a local e a regional. As usinas de cana-de-açúcar realizam a previsão de safra numa escala local, ou seja, buscam estimar a produção nos talhões pertencentes à usina. Já os órgãos oficiais buscam um levantamento regional, a fim de estimar a produção por municípios e estados.

Rudorff (1985) descreveu o método utilizado pela Usina Barra Grande em Lençóis Paulista - SP, o qual possuía duas abordagens: subjetiva e objetiva. A subjetiva era feita por meio da visita dos técnicos da usina às lavouras antes do início da safra, a qual dependia muito da experiência dos técnicos. A outra abordagem, objetiva, realizada pouco antes do início da colheita onde, por meio de amostragem, se calculava o número de colmos por hectare e o peso médio de cada colmo, obtendo assim uma produtividade média estimada para a safra. Por meio do produto da produtividade média pela área plantada estimava-se a produção da safra. Machado (2003) considerou o método para a estimativa de safra adotada pela a Usina São João, em Araras – SP, bastante semelhante ao descrito por Rudorff (1985), diferenciando apenas nas datas de previsão. Machado (2003) observou também que, durante a inspeção, o técnico analisa o comportamento das plantas principalmente nas bordas dos talhões, desconhecendo a proporção e a distribuição em todo o talhão de possíveis padrões diferenciados nas bordas da cultura.

Para a avaliação da safra e estimativa da produção de cana de açúcar em escala regional (municipal), a CONAB baseia-se na pesquisa de campo nos municípios produtores, com aplicação de questionários diretamente aos detentores das informações sobre o sistema agroindustrial sucroalcooleiro, gerando assim os dados que são apresentados em cada levantamento. Em 2007 foram realizados três levantamentos para a cana-de-açúcar. No último levantamento, realizado no período de 5 a 16 de novembro, 47 técnicos da Companhia visitaram 355 usinas de açúcar e destilarias de álcool e 60 entidades de classe em todo o território nacional (CONAB, 2007a, b, c). A CONAB, por meio do projeto Geosafra, utiliza imagens de satélite como ferramenta de auxílio na definição de pontos amostrais para as visitas de campo.

O IBGE realiza a previsão da produção da cana-de-açúcar por meio do Levantamento Sistemático da Produção Agrícola (LSPA), com periodicidade mensal e abrangendo todos os municípios. São levantadas informações sobre a demanda de insumos agrícolas (sementes, fertilizantes, corretivos, etc.) junto aos produtores, associações de classe e cooperativas, utilização de dados históricos de produtividade e levantamentos e observações de campo sobre o comportamento da cultura em face de ocorrências climáticas e/ou fitossanitárias (IBGE, 2002).

Diferentes métodos subjetivos adotados pelos órgãos oficiais geram diferentes resultados. Na safra 2005/2006 a CONAB, em seu último levantamento, estimou uma produção de cana-de-açúcar no Estado de São Paulo de 262.851.200 toneladas. Posteriormente, o IBGE divulgou uma produção de 269.134.237 toneladas e o Instituto de Economia Agrícola (IEA) de 284.916.738 toneladas, para a mesma safra.

Métodos de estimativas de produção auxiliados pelo uso do sensoriamento remoto podem potencialmente contribuir para a diminuição destas discrepâncias.

## **2.5. O uso de sensoriamento remoto e dados meteorológicos no monitoramento da cana-de-açúcar e previsão de safras**

A cultura da 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 uniforme. Porém, embora o uso do sensoriamento remoto na agricultura venha sendo estudado há mais de trinta e cinco anos, a aplicação desta tecnologia na cultura da cana-de-açúcar ainda é relativamente pequena (Apan et al., 2004; Simões et al., 2005; Galvão et al., 2006). Alguns estudos vêm mostrando seu potencial de aplicação na cana-de-açúcar, tais como: identificação de áreas canavieiras e distinção entre cana ano e cana ano-e-meio (Rudorff et al., 2005; Xavier et al., 2006), uso de água na irrigação (Immerzeel et al., 2008), detecção de doenças específicas como ferrugem amarela (Apan et al., 2004) e identificação de diferentes variedades (Fortes e Demattê, 2006).

O CANASAT, em andamento desde 2003, é um projeto do Instituto Nacional de Pesquisa Espaciais (INPE) que utiliza imagens de sensoriamento remoto do sensor Landsat para mapear e quantificar a área plantada com a cultura da cana-de-açúcar no estado de São

Paulo (CANASAT, 2007). Entretanto, segundo Pellegrino (2001), as imagens de baixa resolução espacial e alta resolução temporal são adequadas para o monitoramento da cana-de-açúcar por meio de índices de vegetação, sendo ainda desconhecido o potencial de uso das imagens SPOT *Vegetation* na identificação de áreas canavieiras no estado de São Paulo.

Outra aplicação de grande interesse ao setor público e privado refere-se à estimativa da produtividade da cana-de-açúcar. A possibilidade de determinação do IAF por meio de dados espectrais, como o NDVI (Pellegrino, 2001; Simões et al., 2005b) e a relação entre o IAF e a produtividade da cana-de-açúcar (Simões et al., 2005a) mostram o potencial de utilização de dados espectrais na estimativa de produtividade. Alguns pesquisadores têm buscado estimar a produtividade da cana-de-açúcar por meio de dados espectrais (Rudorff, 1985; Pontes et al., 2005; Picoli, 2006; Sugawara et al., 2007), porém, em escala local (fazendas). Existe a necessidade de se verificar este potencial em escala regional.

O uso de dados meteorológicos pode contribuir significativamente para a estimativa de produtividade na cana-de-açúcar (Muchow et al., 1994; Delgado-Rojas e Barbieri, 1999), principalmente em relação à disponibilidade de água (Delgado-Rojas e Barbieri, 1999). Rudorff e Batista (1990) propuseram um modelo de estimativa de produtividade da cana-de-açúcar no município de Lençóis Paulista, estado de São Paulo, utilizando dados espectrais do sensor MSS dos satélites Landsat 4 e 5 em conjunto com dados agrometeorológicos. Apesar de não terem encontrado resultados satisfatórios, eles concluíram que a associação entre dados espectrais e agrometeorológicos poderia trazer importantes incrementos nas estimativas de produtividade.

Alguns métodos de estimativa de produtividade para a cana-de-açúcar em escala local utilizam um grande número de variáveis relacionadas à cultura e seu manejo, como variedade, ano safra, estágio de corte e aplicação de insumos (Picoli et al., 2007). Isso possibilita um maior controle dos fatores envolvidos no desenvolvimento da cultura. No entanto, podem ser de difícil obtenção e específicas ao manejo aplicado. Para o monitoramento de culturas em escala regional (municípios, estados e países) é necessário que se desenvolvam métodos simples de estimativa de produtividade, nos quais se utilizem variáveis de fácil obtenção e abrangentes territorialmente.

## 2.6. Mineração de dados

O volume de dados digitais armazenados em todas as áreas do conhecimento cresce a cada dia, havendo atualmente ampla disponibilidade de grandes bases de dados. Na agricultura, o uso de imagens de satélite, dados meteorológicos, de manejo da cultura, entre outros, possibilitam avaliar o desenvolvimento das culturas de forma cada vez mais consistente e abrangente. No entanto, se por um lado esse volume de dados possibilita realizar estudos com análises cada vez mais ricas e completas, por outro, aumenta a necessidade do uso de ferramentas computacionais que permitam tratar esses dados e avaliar os resultados em tempo hábil. Neste contexto, o uso de técnicas de mineração de dados e descoberta de conhecimento em banco de dados pode contribuir fortemente na viabilização de estudos e testes de hipóteses, permitindo transformar dados em informação e conhecimento.

Segundo Fayyad et al. (1996), a mineração de dados é um “processo não trivial de identificação de padrões válidos, novos, úteis e implicitamente presentes em grandes volumes de dados”. Existem três tarefas comumente aplicadas em mineração de dados: associação, classificação e agrupamento. E, segundo Fayyad (1997), a classificação é uma das técnicas utilizada em mineração de dados quando se pretende predizer alguma característica baseada em outras características. Dentre os métodos de classificação, a árvore de decisão é interessante por apresentar de forma organizada e de fácil compreensão quais os relacionamentos e hierarquias definidos entre os atributos para a determinação do atributo classe. O uso desta ferramenta como forma de verificar a possibilidade de determinação do atributo classe “produtividade” por meio de atributos provenientes de dados espectrais e meteorológicos é uma opção interessante para avaliar os resultados.

Alguns pesquisadores têm usado a mineração de dados como ferramenta de análise de dados espectrais provenientes de áreas agrícolas, como Karimi et al. (2004) na detecção de stress causado por falta de nutrientes, Shock et al., (2002) e Bajwa et al., (2004) na classificação do uso da terra e Crisler et al. (2002) na análise de dados de produção.

### 3. MATERIAL E MÉTODOS

Na Figura 6 é apresentado o fluxograma das principais etapas realizadas neste trabalho, desde a definição da área e período de estudo até a análise dos resultados. O detalhamento de cada etapa do fluxograma está descrito nos próximos tópicos do trabalho.

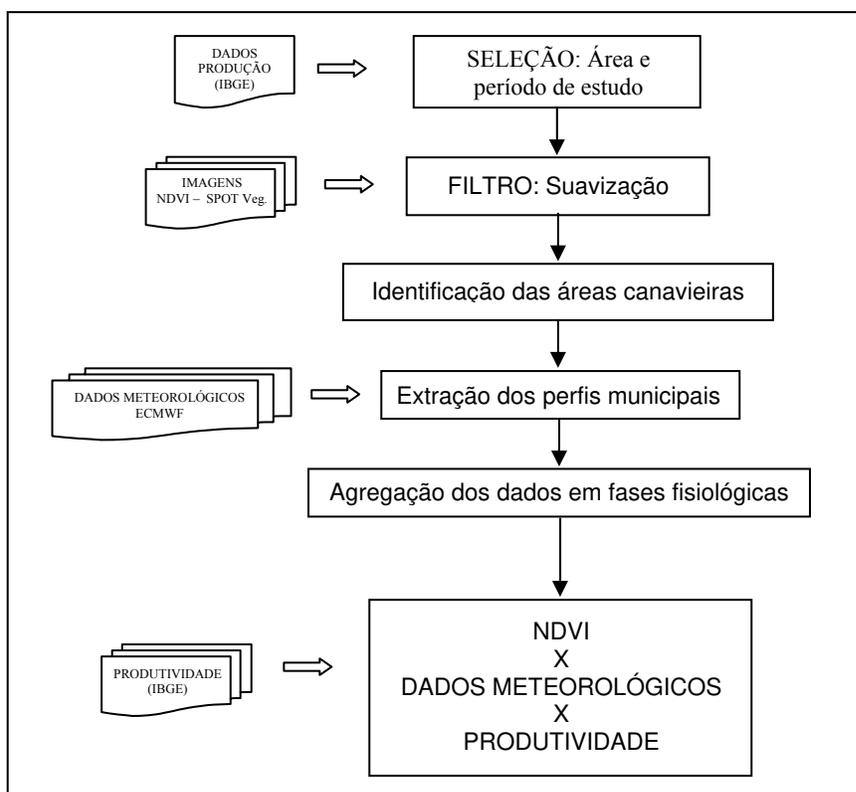


Figura 6: Fluxograma das principais etapas desenvolvidas.

#### 3.1. Área e período de estudo

O estudo foi realizado em 20 municípios do Estado de São Paulo - Brasil, selecionados segundo dois critérios:

- importância produtiva em 2006;
- distribuição espacial no Estado.

A importância produtiva foi verificada por meio dos dados de produção agrícola municipal do IBGE (2008). Na Tabela 2 estão ordenados de forma decrescente os dados de produção (toneladas), porcentagem da produção em relação à produção total do Estado, área

colhida (hectares), área total do município (hectares) e a área relativa de cana no município, referentes à safra 2005 / 2006, para os 20 municípios selecionados.

Tabela 2: Dados de 2006 de produção (toneladas), porcentagem da produção em relação à produção total do Estado, área colhida (hectares), área total do município (hectares) e a área relativa à cana no município (porcentagem), para os 20 municípios selecionados.

2006	Município	Produção (t)	Prod. de SP (%)	Área colhida (ha)	Área município (ha)	Área relativa (%)
1	Morro Agudo	7835267	2,9	86602	138620	62
2	Jaboticabal	3600000	1,3	40000	70670	57
3	Paraguaçu Paulista	3500000	1,3	50000	100110	50
4	Batatais	3272500	1,2	38500	85070	45
5	Barretos	3270300	1,2	36967	156360	24
6	Piracicaba	3200000	1,2	40000	136840	29
7	Valparaíso	3150000	1,2	42000	85870	49
8	Jaú	2962500	1,1	39500	68840	57
9	Lençóis Paulista	2844000	1,1	36000	80380	45
10	Araraquara	2809157	1,0	33900	100600	34
11	Pitangueiras	2786247	1,0	32066	42960	75
12	São Manuel	2610000	1,0	29000	65110	45
13	Penápolis	2560000	1,0	32000	70850	45
14	Guaíra	2475000	0,9	27500	125860	22
15	Guararapes	2465000	0,9	29000	95650	30
16	Araras	2438085	0,9	26383	64340	41
17	Dois Córregos	2400000	0,9	30000	63260	47
18	Sertãozinho	2400000	0,9	30000	40280	74
19	Novo Horizonte	2362800	0,9	29535	93280	32
20	Ribeirão Preto	2254000	0,8	32200	65040	50
TOTAL		61194856	22,7			
TOTAL - Estado de SP		269134237				

Fonte: (IBGE, 2008).

A distribuição espacial dos municípios selecionados foi verificada por meio de um mapa da distribuição da produção da cana-de-açúcar no estado de São Paulo em 2006, onde se buscou contemplar diferentes regiões do Estado. Na Figura 7 é apresentado o mapa da distribuição da produção da cana-de-açúcar no estado de São Paulo em 2006 e os municípios selecionados.

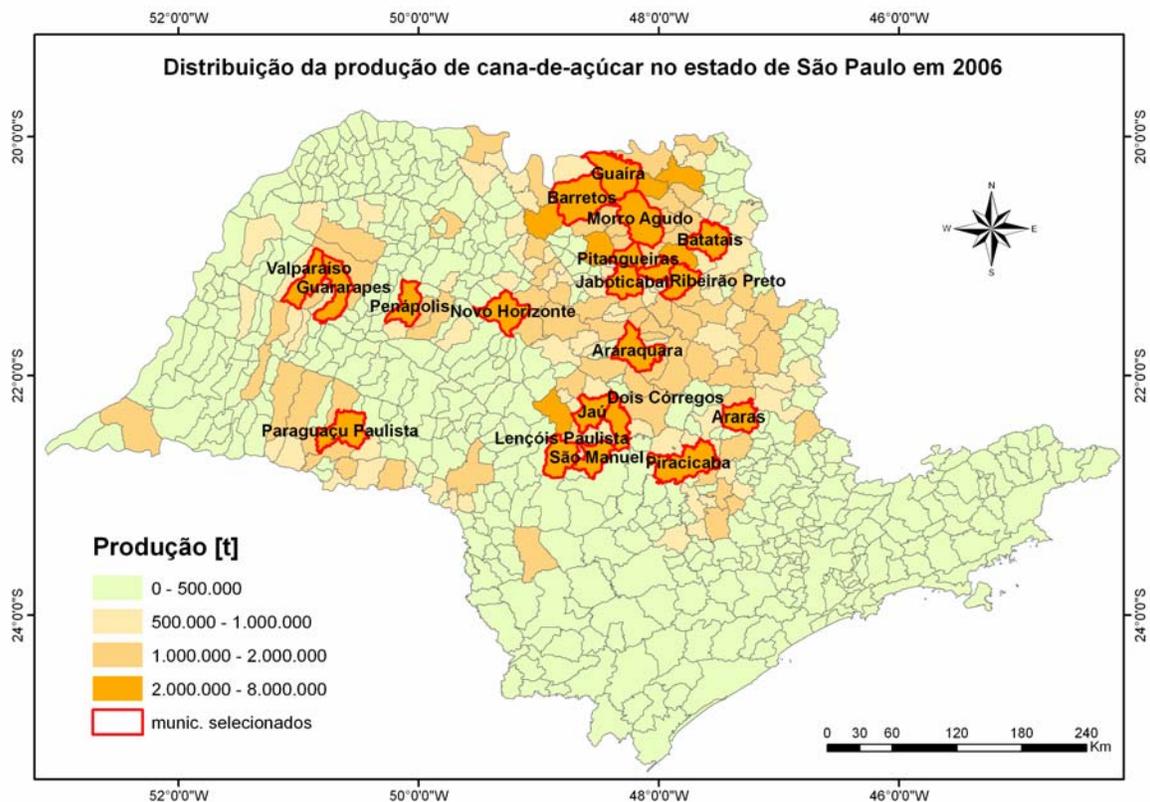


Figura 7: Mapa da distribuição da produção da cana-de-açúcar no estado de São Paulo em 2006 e localização espacial dos municípios selecionados. Fonte: IBGE (2008).

O período de estudo foi escolhido de acordo com a disponibilidade de imagens do SPOT *Vegetation*. Definiu-se o período entre agosto de 1999 e outubro de 2006, considerando sete safras consecutivas.

### 3.2. Aptidão edafoclimática para a cana-de-açúcar no estado de São Paulo

O Governo do Estado de São Paulo realizou um estudo denominado Zoneamento de Culturas Bioenergéticas no Estado de São Paulo. Um dos produtos deste trabalho foi o Mapa de aptidão edafoclimática plena da cultura da cana-de-açúcar, apresentado na Figura 8. Este mapa foi usado para caracterizar edafoclimaticamente a área de estudo do trabalho.

Analisando as Figuras 7 e 8, verifica-se as regiões onde se encontram os municípios selecionados são classificadas como sendo de aptidão edáfica boa ou regular.

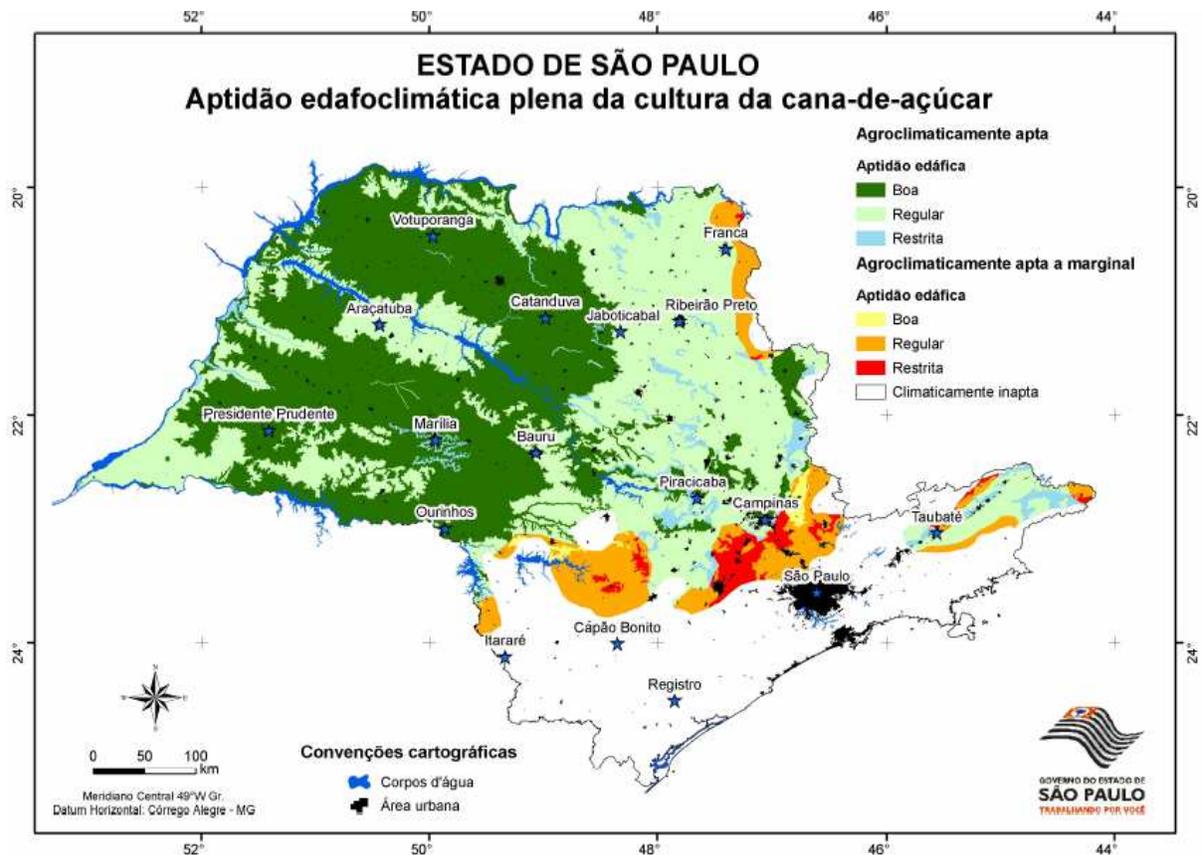


Figura 8: Mapa de aptidão edafoclimática plena da cultura da cana-de-açúcar, estado de São Paulo. Fonte: ZONEAMENTO DE CULTURAS BIOENERGÉTICAS NO ESTADO DE SÃO PAULO (2008).

### 3.3. Dados espectrais: sistema SPOT *Vegetation*

Foi utilizado o produto NDVI com composições decendiais (S10) do SPOT *Vegetation*, que já possuem correções atmosférica e geométrica.

A resolução radiométrica é de 8 bits e os valores digitais originais estavam na escala de níveis de cinza, variando de 0 a 255. A conversão para a escala de valores NDVI, que varia entre -1 e +1, foi feita utilizando a seguinte expressão, disponível em VITO (2007):

$$NDVI = 0,004 * DN - 0,1 \quad (2)$$

Sendo *NDVI* o valor corrigido e *DN* o Número Digital da imagem original.

Foram utilizadas 36 imagens por ano-safra, totalizando 252 para os 7 anos de estudo.

### 3.3.1. Suavização dos perfis temporais do NDVI

Foi aplicado nas séries temporais anuais do NDVI o algoritmo de suavização Hants (*Harmonic Analysis of NDVI Time-Series*) com o objetivo de eliminar variações bruscas nos valores decendiais de NDVI comumente causadas por presença de nuvens, ângulo de visada e também outros fatores. Este algoritmo, proposto por Roerink et al. (2000), considera o comportamento temporal do NDVI, ao longo de uma safra, como sendo harmônico e de baixa frequência. Desta forma, a série é ajustada eliminando as oscilações de alta frequência, tratadas como “ruídos”. O ajuste é baseado no erro mínimo quadrático.

Na Figura 9 é apresentado o resultado da aplicação do algoritmo Hants em um perfil temporal obtido em um pixel de imagens decendiais do SPOT *Vegetation* localizado em uma área canvieira.

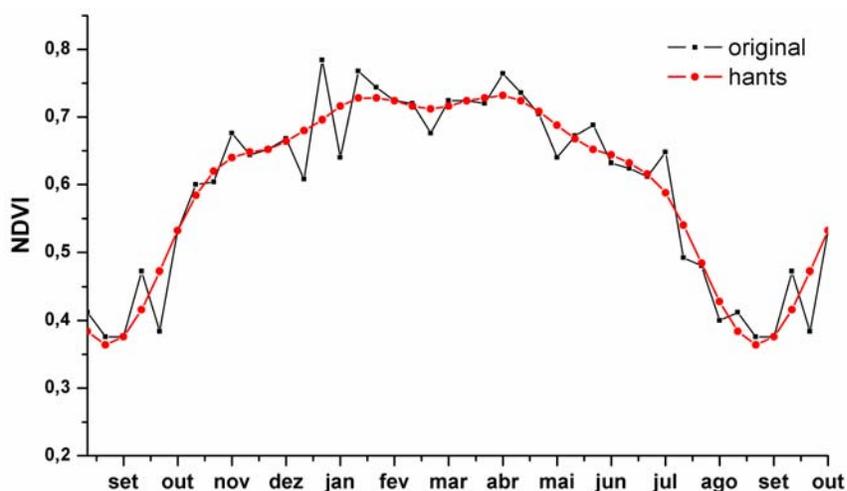


Figura 9: Exemplo do resultado do algoritmo de suavização Hants para um pixel em área canvieira, aplicado em imagens decendiais do SPOT *Vegetation*.

O valor de cada pixel das imagens é ajustado conforme exemplificado na Figura 9.

Na Figura 10 é mostrado o procedimento para aplicação do algoritmo Hants. As quatro etapas descritas na Figura 10 foram realizadas de forma automática, utilizando um sistema computacional na linguagem IDL, desenvolvido no Laboratório de Geoprocessamento da Faculdade de Engenharia Agrícola da Unicamp.

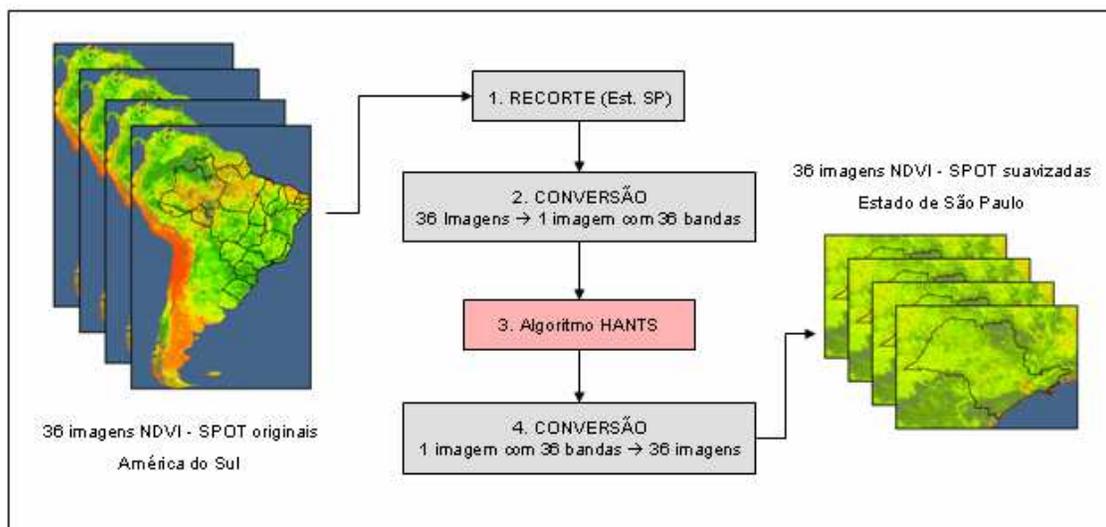


Figura 10: Procedimento de suavização das imagens para cada ano-safra.

O primeiro passo do procedimento de suavização mostrado na Figura 10 foi o recorte das imagens para o estado de São Paulo, com o objetivo de diminuir o tamanho dos arquivos, reduzindo assim o volume ocupado em disco. O segundo passo tratou-se da conversão dos arquivos; os 36 arquivos de imagens decenciais, para cada ano, foram convertidos para apenas um arquivo de 36 bandas. Esta etapa é um pré-requisito para a aplicação do algoritmo Hants a fim de possibilitar a leitura de valores antecedentes e consequentes no perfil temporal. O terceiro passo foi a aplicação do algoritmo Hants na imagem de 36 bandas. Este algoritmo gera outra imagem de 36 bandas com os valores de cada pixel e de cada banda ajustados. No quarto passo, cada banda do arquivo gerado no terceiro passo foi convertida para uma imagem decencial, obtendo assim 36 imagens decenciais suavizadas.

### 3.4. Identificação das áreas canavieiras

A identificação das áreas canavieiras, também chamada de “máscara” da cultura, foi feita utilizando-se um método de classificação baseado na resposta temporal do NDVI de cada pixel ao longo da safra, partindo da premissa de que cada alvo possui um comportamento característico. Primeiramente, foi feita uma avaliação visual dos tipos de perfis predominantes no estado de São Paulo e na Figura 11 são apresentados alguns exemplos de perfis relativos aos seguintes alvos: área urbana, cultura anual, cana-de-açúcar, pasto e mata. Estes perfis

foram obtidos utilizando-se imagens decendiais suavizadas do SPOT *Vegetation* em áreas onde se havia conhecimento prévio dos alvos.

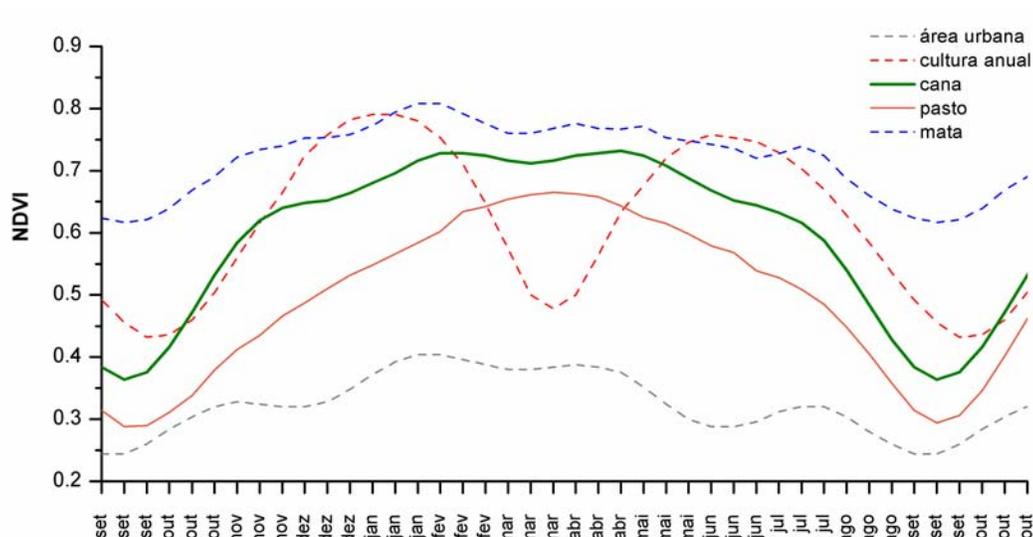


Figura 11: Exemplos de resposta temporal do NDVI ao longo de uma safra para os seguintes alvos: área urbana, cultura anual, cana-de-açúcar, pasto e mata. Perfis obtidos em imagens decendiais suavizadas do SPOT *Vegetation*.

Na Figura 12 é apresentada a resposta temporal característica do NDVI para áreas canavieiras (cana de ano), obtido em imagens decendiais suavizadas do SPOT *Vegetation*. Observa-se que baixos valores de NDVI ocorrem no início e término da safra, enquanto que os altos valores ocorrem na época de grande crescimento vegetativo.

A fim de selecionar nas imagens decendiais suavizadas SPOT *Vegetation* os pixels que apresentassem a resposta temporal característica do NDVI da cultura da cana-de-açúcar, foram definidos cinco critérios de seleção, apresentados na Figura 12. A definição destes cinco critérios, além de selecionar pixels com o perfil característico da cana-de-açúcar, buscou excluir pixels com comportamento de alvos indesejáveis.

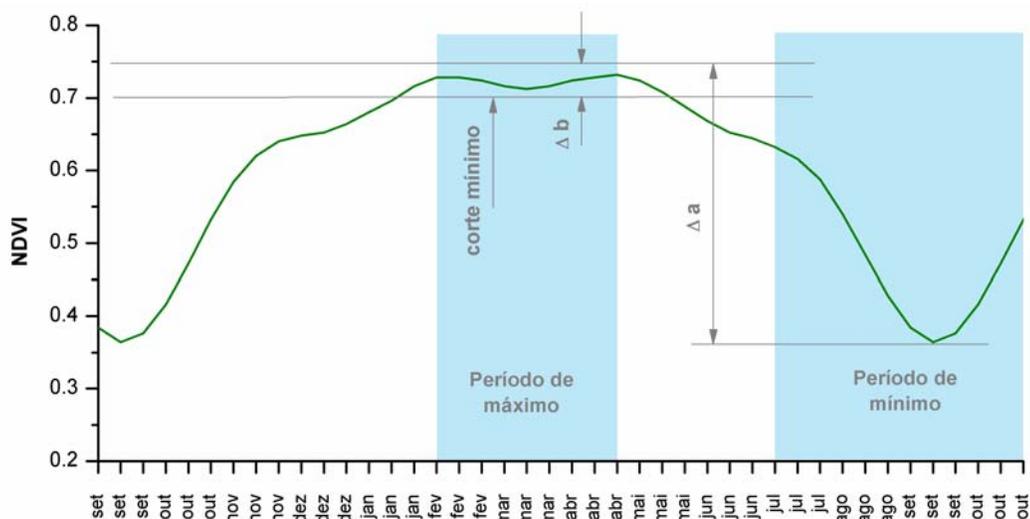


Figura 12: Comportamento temporal característico do NDVI para áreas canavieiras, obtido em imagens decendiais suavizadas do SPOT Vegetation, e critérios utilizados para identificação destas áreas.

- Período de máximo:

Refere-se ao período onde, para a cana de ano, o NDVI atinge os valores máximos, definido entre o início de fevereiro e final de abril.

- Período de mínimo:

Trata-se do período onde, para a cana de ano, variedades precoce, média ou tardia, o NDVI atinge os valores mínimos, definido entre o início de julho e final de outubro.

- Δ a:

Trata-se de uma diferença mínima entre o valor máximo do NDVI, ocorrido no período de máximo, e o valor mínimo do NDVI, ocorrido no período de mínimo.

Esta diferença (delta –  $\Delta$ ) permite identificar os pixels que possuem comportamento de áreas plantadas com culturas anuais ou pastagens, excluindo áreas onde o perfil se mantém relativamente estável ao longo do ano, como áreas de mata nativa, reflorestamentos, culturas perenes, áreas urbanas entre outras.

- Δ b:

Variação máxima permitida no período de máximo.

De modo geral, em imagens suavizadas, o comportamento do perfil temporal do NDVI da cana-de-açúcar apresenta pequenas oscilações no período de máximo. Este parâmetro exclui alvos com grandes oscilações na resposta temporal do NDVI no período de máximo, como

áreas de culturas com duas safras anuais (milho e soja, por exemplo) e algumas áreas de pastagens.

- Corte mínimo:

Valor mínimo a ser alcançado no período de máximo.

Este parâmetro exclui a maioria das áreas de pastagens, pois estas áreas, apesar de possuir comportamento semelhante ao da cana-de-açúcar, geralmente apresentam valores máximos de NDVI suavizado inferiores ao da cana-de-açúcar.

A partir de alguns ensaios preliminares, foram definidos 24 testes provenientes da combinação entre quatro valores para o  $\Delta a$ , três valores para o corte mínimo e dois valores para o  $\Delta b$ , apresentados na Tabela 3.

Tabela 3: Especificação dos testes realizados para definição da máscara de área colhida de cana-de-açúcar.

$\Delta a$ mínimo	Corte mínimo	$\Delta b$ máximo	Nº teste
0,20	0,70	0,02	1
		0,06	2
	0,72	0,02	3
		0,06	4
	0,74	0,02	5
		0,06	6
0,22	0,70	0,02	7
		0,06	8
	0,72	0,02	9
		0,06	10
	0,74	0,02	11
		0,06	12
0,24	0,70	0,02	13
		0,06	14
	0,72	0,02	15
		0,06	16
	0,74	0,02	17
		0,06	18
0,26	0,70	0,02	19
		0,06	20
	0,72	0,02	21
		0,06	22
	0,74	0,02	23
		0,06	24

Os 24 testes descritos na Tabela 3 foram realizados para cada safra.

### Avaliação dos resultados das máscaras

Os resultados das máscaras foram avaliados buscando compará-los com dados de área colhida oficiais, divulgados pelo IBGE. As análises foram feitas para todas as safras e em duas escalas espaciais: estadual e municipal. Para a escala estadual, buscou-se verificar a correlação entre o valor de área de cana-de-açúcar colhida oficial e o valor encontrado pelo método proposto, para cada um dos 637 municípios do estado de São Paulo. Além da correlação, verificou-se a diferença, em porcentagem, entre a área oficial total de cana-de-açúcar colhida no Estado e a área total encontrada pelo método proposto. Para a escala municipal, verificou-se em cada um dos 20 municípios selecionados a diferença entre a área colhida oficial e a área encontrada pelo método proposto.

A extração dos dados de área, baseada na contagem de pixels em cada área de interesse, foi feita de forma automática, por meio de um sistema computacional em IDL, desenvolvido no Laboratório de Geoprocessamento da Faculdade de Engenharia Agrícola da Unicamp, baseado no trabalho de Esquerdo (2007).

### **3.5. Dados meteorológicos: modelo do ECMWF**

Foram utilizadas imagens do modelo do ECMWF para o período entre janeiro de 1999 e dezembro de 2006. Foi escolhido o intervalo temporal decenal, o mesmo definido para as imagens do SPOT *Vegetation*, para as seguintes variáveis meteorológicas:

- precipitação (em mm, valor acumulado em 10 dias);
- radiação global (em wh/m<sup>2</sup>, acumulada em 10 dias);
- temperatura mínima (em °C, média entre 10 dias);
- temperatura média (em °C, média entre 10 dias);
- temperatura máxima (em °C, média entre 10 dias);

### **3.6. Perfis temporais**

Obtidas as áreas canavieiras do Estado para cada safra em estudo, iniciou-se o processo de extração dos dados das imagens e geração dos perfis médios municipais para as variáveis espectral e meteorológicas. Utilizou-se um processo automático de extração de dados de

imagens. Trata-se de um sistema computacional em IDL, desenvolvido no Laboratório de Geoprocessamento da Faculdade de Engenharia Agrícola da Unicamp, baseado no trabalho de Esquerdo (2007).

Este sistema utiliza os seguintes dados de entrada:

- Série de imagens: 36 arquivos para cada safra, contendo todas as imagens decendiais, em formato matricial. O processo de extração dos dados é feito para cada variável separadamente (NDVI, precipitação, radiação global temperatura mínima, média e máxima);
- Vetores municipais: Arquivo contendo os limites territoriais dos municípios de interesse em formato vetorial, obtido no banco de dados do IBGE (2008);
- Áreas de interesse: Máscara da cultura, que se trata de um arquivo em formato ASCII contendo as coordenadas de todos os pixels referentes às áreas de cana-de-açúcar nos municípios de interesse.

A partir destes dados de entrada o sistema constrói o perfil médio para cada município e cada variável. O procedimento é repetido para cada safra. A Figura 13 ilustra o funcionamento do sistema, bem como o formato dos arquivos de entrada e saída.

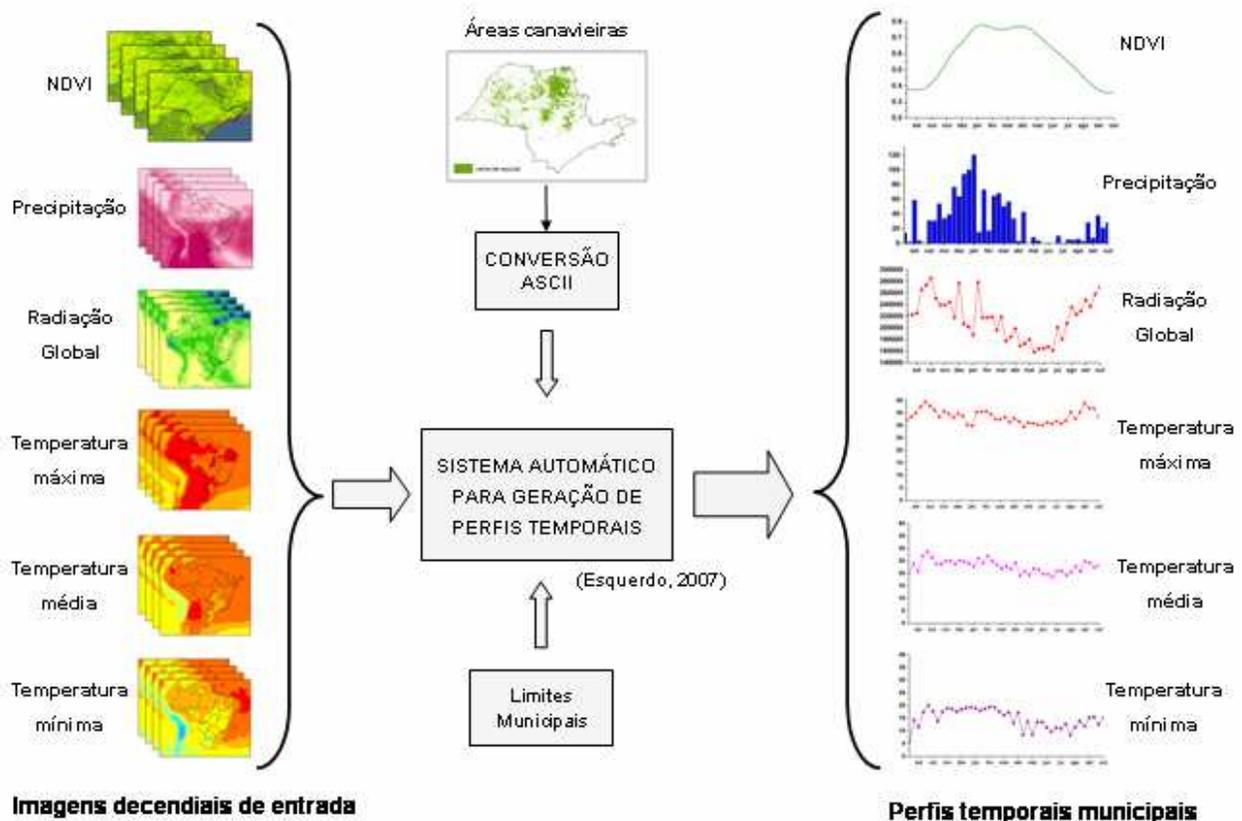


Figura 13: Ilustração do processo de extração dos perfis médios municipais das imagens.

### **3.7. Relação entre NDVI, condições meteorológicas e produtividade**

Para analisar as possíveis relações existentes entre os dados de NDVI, meteorológicos e de produtividade, foi incorporado o conhecimento sobre a fisiologia da cultura, de forma a subdividir a safra em fases de desenvolvimento, analisando assim os dados agregados por fase.

#### **3.7.1 Geração dos atributos espectrais e meteorológicos a partir das fases de desenvolvimento**

Primeiramente foram definidas quatro fases de desenvolvimento: perfilhamento, crescimento acelerado, crescimento lento e maturação. Em seguida, determinadas as fronteiras entre cada fase. Finalmente, os dados espectrais e meteorológicos foram agregados por fase fisiológica para, posteriormente, analisar as possíveis relações entre atributos e entre estes e a produtividade média municipal.

#### **3.7.2 Relações entre atributos meteorológicos, espectrais e a produtividade da cultura**

A análise das relações entre os atributos iniciou-se com a verificação das possíveis correlações entre atributos espectrais e meteorológicos. A verificação foi feita por meio do índice de determinação ( $R^2$ ).

Posteriormente, foram verificadas as relações dos atributos espectrais e meteorológicos com a produtividade municipal oficial, obtida no IBGE (2008). Buscou-se inferir sobre a produtividade da cultura por meio de dados meteorológicos e espectrais. Nesta etapa utilizaram-se técnicas de mineração de dados a fim de selecionar, entre todos os atributos gerados para cada safra e município, os atributos mais relevantes e verificar como estes influenciavam na determinação da produtividade da cultura. Para a aplicação das técnicas de mineração foi usado o software Weka 3.5. A seguir, é detalhado o procedimento para preparação e análise dos dados:

### 1. Geração do banco de dados.

Para cada município e safra foram gerados 51 atributos provenientes dos dados espectrais e meteorológicos ocorridos ao longo do ciclo de desenvolvimento da cultura, citados na Tabela 6. Dessa forma, cada município / safra passou a ser uma instância do banco de dados, tendo a produtividade como atributo classe. Foi então gerado um banco de dados contendo 52 atributos (colunas) e 140 instâncias (linhas).

### 2. Discretização do atributo classe.

O uso de alguns algoritmos de seleção de atributos e de classificação em mineração de dados requer que o atributo classe, no caso a produtividade, esteja discretizado e não seja numérico. Dessa forma, os dados referentes à produtividade para os 20 municípios e sete safras foram discretizados e classificados em três classes por meio de percentis.

### 3. Seleção de atributos.

Foram utilizados quatro métodos para a seleção dos atributos mais relevantes para a determinação do atributo classe:

- 1) teste do Qui-quadrado;
- 2) método Wrapper com algoritmo de árvore de decisão J48;
- 3) método CFS (Correlation Feature Selection);
- 4) combinação dos métodos InfoGAin (Ganho de Informação) e GainRatio (Taxa de Ganho).

Foram considerados como sendo atributos relevantes aqueles selecionados pelo maior número de abordagens.

### 4. Classificação

A partir dos atributos selecionados, foi aplicado o algoritmo de classificação por árvore de decisão J48 a fim de estruturar a hierarquia dos atributos selecionados para a determinação da produtividade.

## **4. RESULTADOS E DISCUSSÃO**

Neste capítulo os resultados foram divididos em duas partes. Na primeira, foram abordados os resultados referentes à identificação das áreas canavieiras, em escala estadual e municipal. Na segunda, apresentados e discutidos os resultados sobre as análises das relações entre dados espectrais, meteorológicos e de produtividade.

### **4.1. Identificação das áreas canavieiras**

#### **4.1.1. Escala estadual**

Foram feitas 24 estimativas de área de cana-de-açúcar colhida para cada safra, obtidas por meio dos 24 testes citados na Tabela 3, abrangendo todo o estado de São Paulo. Em seguida, foi extraída a área estimada em cada um dos 637 municípios do Estado, obtendo-se 24 simulações de mapeamento de cana para cada safra. Então, para cada simulação, foi feita uma análise de correlação entre a área estimada e área oficial de cada município, verificando o coeficiente de determinação ( $R^2$ ). Associado ao coeficiente de determinação, verificou-se também a porcentagem da diferença entre área estimada e oficial total do Estado (% dif). Conforme apresentado no Anexo 1, foram encontrados resultados com valores de % dif ora subestimados ora superestimados.

Os resultados de  $R^2$  e diferença entre área estimada e oficial, para os 24 testes propostos em cada safra, estão apresentados no Anexo 1. Foram selecionados os melhores resultados e, na Figura 14, são apresentados para cada safra os gráficos de dispersão, a correlação entre área estimada e área oficial para os 637 municípios, o número do teste que apresentou o melhor resultado em cada safra e a porcentagem de diferença entre área total estimada e oficial de cana colhida no Estado.

Na Figura 15 são apresentados os mapas de identificação de áreas canavieiras em cada safra analisada, referentes aos resultados citados na Figura 14.

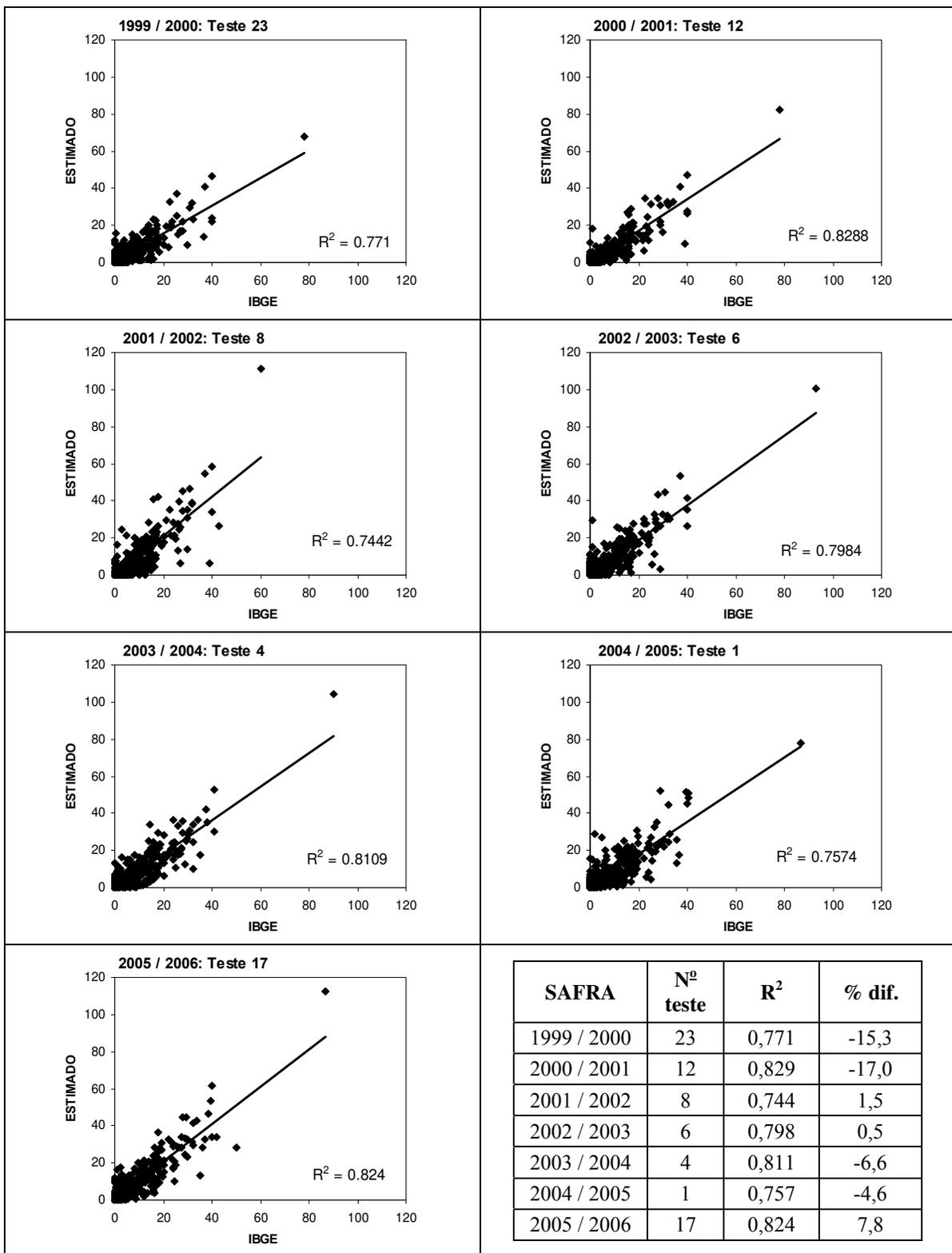


Figura 14: Gráfico de dispersão, R<sup>2</sup>, número do teste que apresentou o melhor resultado e porcentagem de diferença entre área total estimada e oficial de cana colhida no Estado, para o melhor resultado de cada safra. Os eixos X e Y dos gráficos representam de área, em unidade de mil hectares.

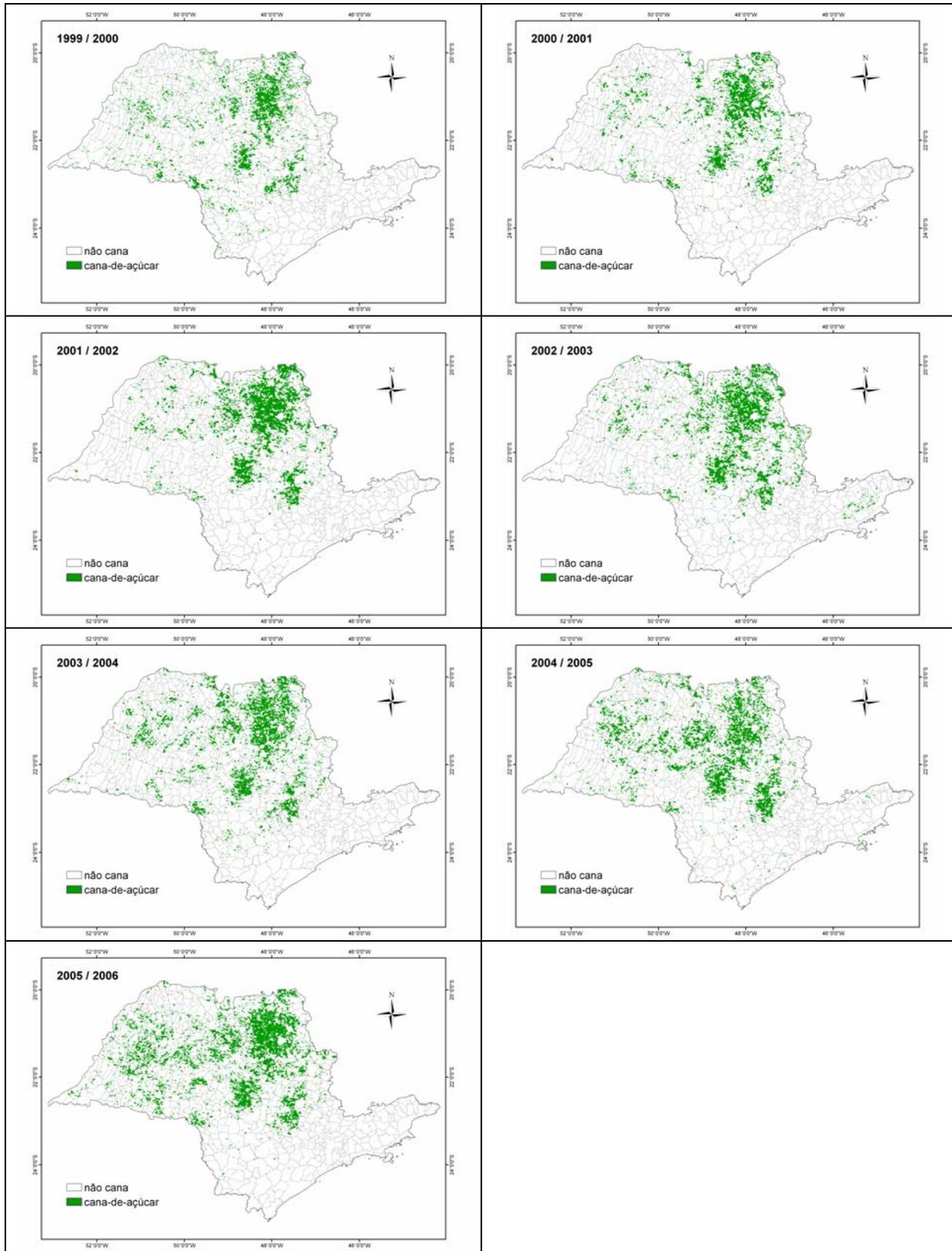


Figura 15: Mapas de identificação da cultura de cana-de-açúcar no estado de São Paulo para as sete safras analisadas.

Verifica-se que as áreas tradicionais para o plantio da cana-de-açúcar no Estado foram mapeadas, sendo possível perceber, ao longo das safras, a expansão das áreas canavieiras para a região noroeste do Estado. Observa-se também que, salvo algumas exceções, áreas onde não há presença de cana-de-açúcar não foram identificadas.

Conforme mostrado na Figura 14, o melhor resultado para cada safra ocorreu usando-se diferentes testes, ou seja, para cada safra foi necessário adequar os parâmetros de cada critério de seleção dos perfis temporais de cana-de-açúcar citados na Figura 12, mostrando haver diferença, entre cada safra, nos perfis temporais do NDVI de pixels de áreas canavieiras. Isso pode estar relacionado a alguns fatores, como mudanças na ocupação da terra ao longo do tempo e diferentes regimes climáticos entre cada safra.

#### **4.1.2. Escala municipal**

Os 24 testes de identificação de áreas canavieiras, realizados para cada safra em escala estadual, também foram analisados em escala municipal, verificando qual foram os resultados de estimativa de área para cada um dos 20 municípios de interesse. Esta análise foi feita com duas diferentes abordagens:

- a) Na primeira, para cada safra, fixou-se o mesmo teste para todos os municípios;
- b) Na segunda, para cada município, utilizou-se o mesmo teste para todas as safras.

##### **4.1.2.1. Primeira abordagem**

Os resultados apresentados nesta abordagem referem-se aos mesmos testes que apresentaram os melhores resultados na escala estadual, ou seja, analisou-se a resposta municipal para os melhores resultados selecionados na escala estadual.

Na Tabela 4 são apresentados os números dos testes aplicados em cada município / safra.

Tabela 4: Testes utilizados para os resultados de identificação de áreas canvieiras em escala municipal, referente à primeira abordagem.

Município	1999/2000	2000/2001	2001/2002	2002/2003	2003/2004	2004/2005	2005/2006
1 Morro Agudo	23	12	8	6	4	1	17
2 Jaboticabal	23	12	8	6	4	1	17
3 Paraguaçu Paulista	23	12	8	6	4	1	17
4 Batatais	23	12	8	6	4	1	17
5 Barretos	23	12	8	6	4	1	17
6 Piracicaba	23	12	8	6	4	1	17
7 Valparaíso	23	12	8	6	4	1	17
8 Jaú	23	12	8	6	4	1	17
9 Lençóis Paulista	23	12	8	6	4	1	17
10 Araraquara	23	12	8	6	4	1	17
11 Pitangueiras	23	12	8	6	4	1	17
12 São Manuel	23	12	8	6	4	1	17
13 Penápolis	23	12	8	6	4	1	17
14 Guaira	23	12	8	6	4	1	17
15 Guararapes	23	12	8	6	4	1	17
16 Araras	23	12	8	6	4	1	17
17 Dois Córregos	23	12	8	6	4	1	17
18 Sertãozinho	23	12	8	6	4	1	17
19 Novo Horizonte	23	12	8	6	4	1	17
20 Ribeirão Preto	23	12	8	6	4	1	17

Nas Figuras 16, 17, 18, 19, 20, 21 e 22 são apresentados os resultados para cada município, separados por safra, por meio de gráficos de colunas, onde são apresentadas as áreas colhidas oficiais (IBGE) e estimadas.

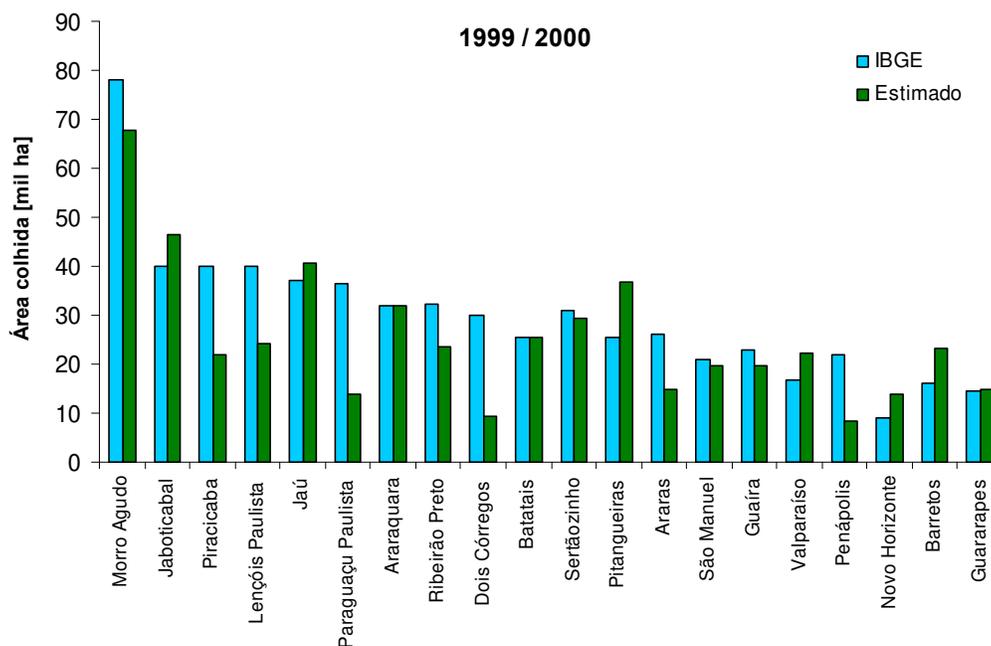


Figura 16: Resultado de identificação de áreas canvieiras em escala municipal para a safra 1999 / 2000 utilizando o teste No 23 da Tabela 3.

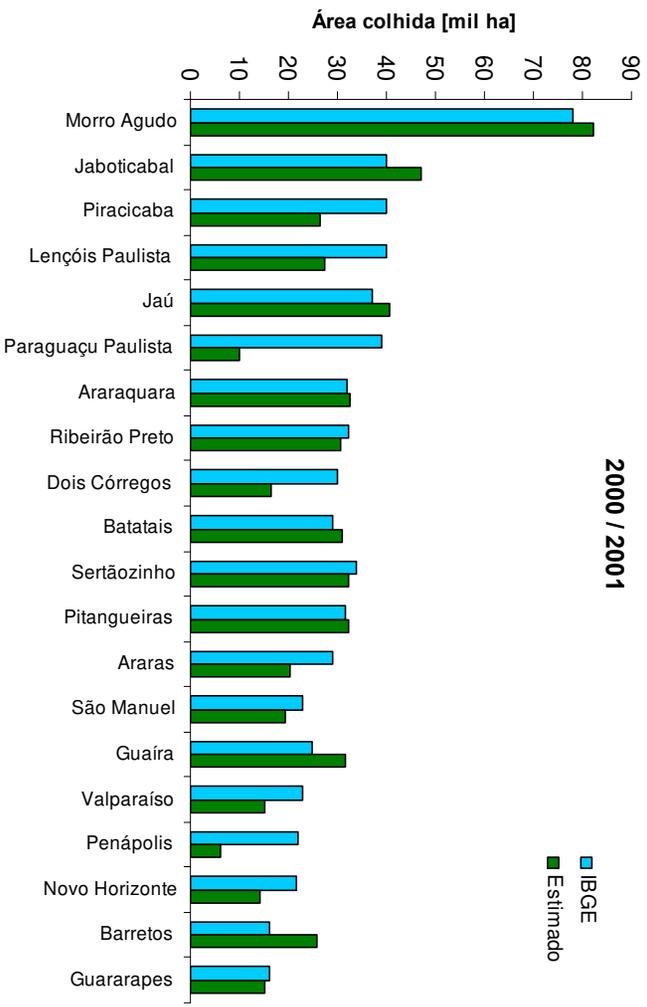


Figura 17: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2000 / 2001 utilizando o teste No 12 da Tabela 3.

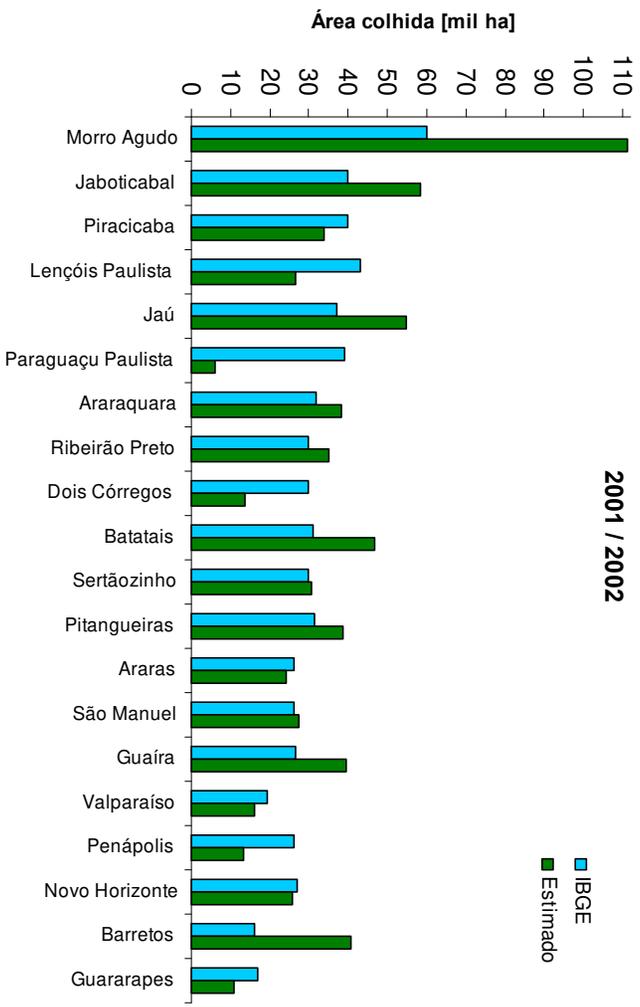


Figura 18: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2001 / 2002 utilizando o teste No 8 da Tabela 3.

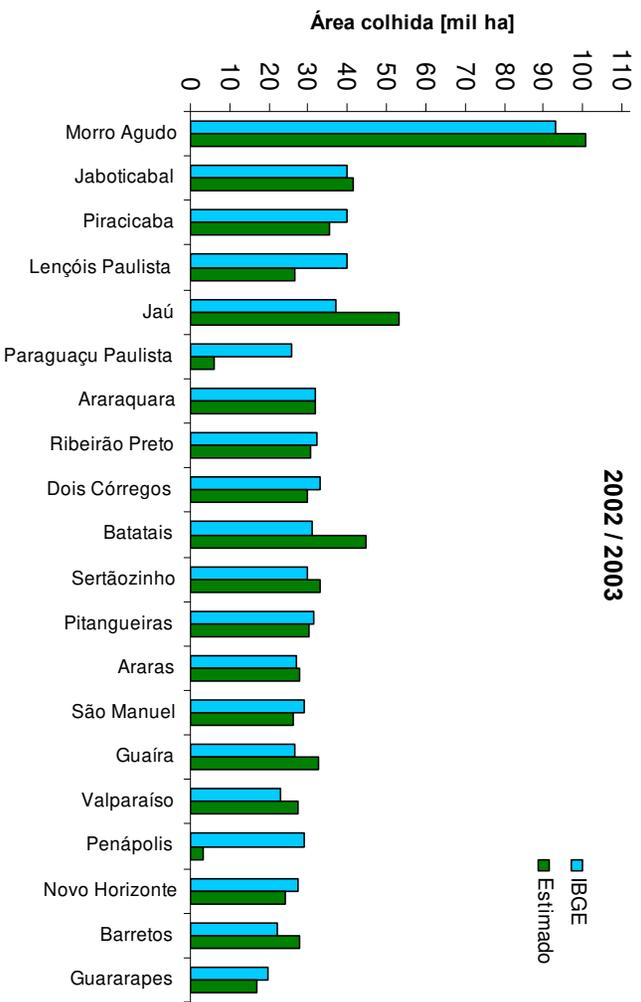


Figura 19: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2002 / 2003 utilizando o teste No 6 da Tabela 3.

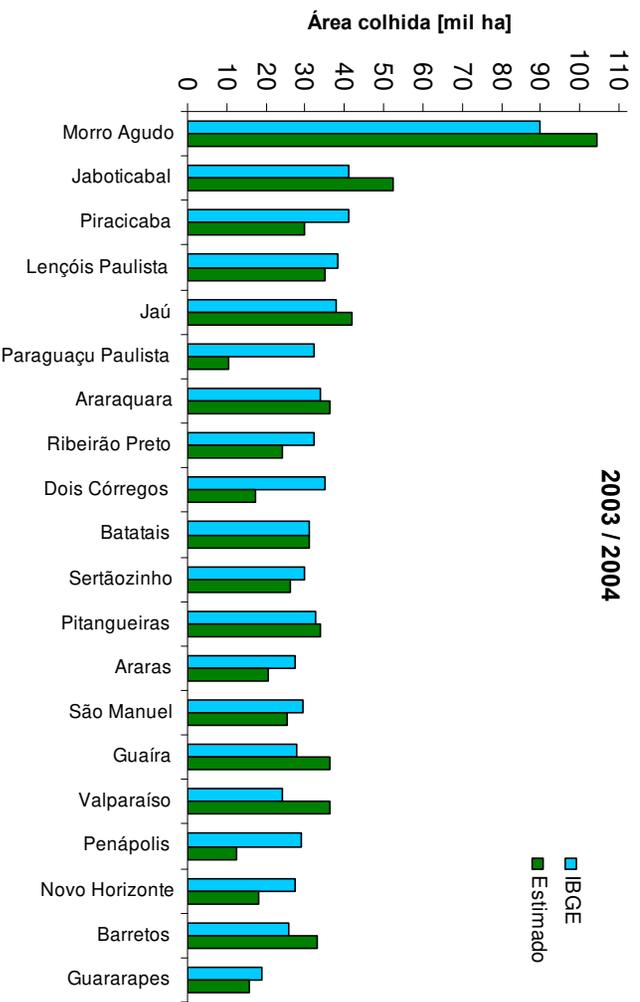


Figura 20: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2003 / 2004 utilizando o teste No 4 da Tabela 3.

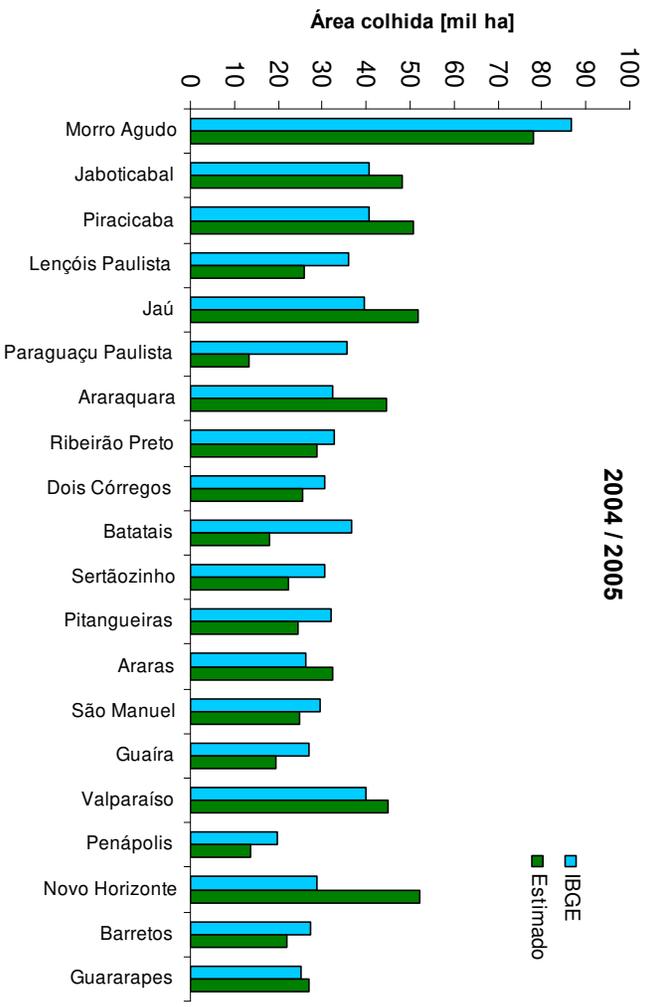


Figura 21: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2004 / 2005 utilizando o teste No 1 da Tabela 3.

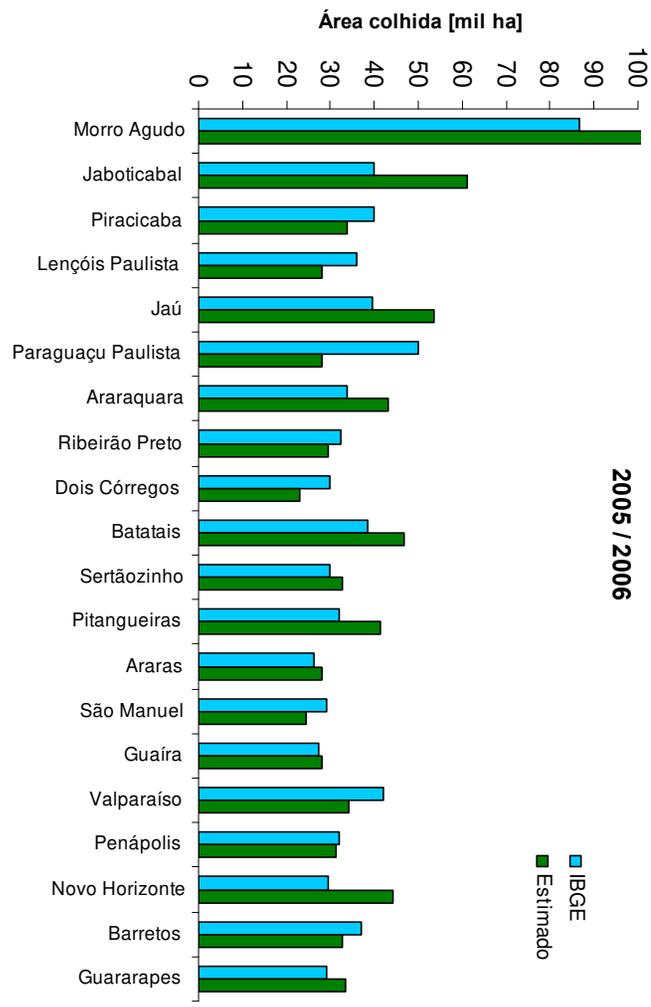


Figura 22: Resultado de identificação de áreas canavieiras em escala municipal para a safra 2005 / 2006 utilizando o teste No 17 da Tabela 3.

Analisando as Figuras 16, 17, 18, 19, 20, 21, e 22, verifica-se que, apesar dos resultados em escala estadual terem apresentado índices de determinação satisfatórios, alguns municípios apresentam diferenças significativas entre área estimada e oficial. Observam-se algumas tendências quanto à diferença entre área estimada e oficial. Os municípios de Ribeirão Preto, Lençóis Paulista, Dois Córregos, São Manuel, Piracicaba, Penápolis e Paraguaçu Paulista mantiveram uma tendência de subestimar os resultados de estimativa da área canavieira em praticamente todas as safras. Já nos municípios Batatais, Jaboticabal, Araraquara e Jaú, ocorreram tendência de superestimar os resultados. Nos demais municípios não houve tendência em subestimar ou superestimar os resultados ao longo das safras. Percebe-se, portanto, haver indícios de que cada município apresenta características próprias quanto ao comportamento do perfil temporal e que estas características se mantêm ao longo do tempo. Esta possibilidade motivou a análise dos resultados por meio da segunda abordagem.

#### **4.1.2.2. Segunda abordagem**

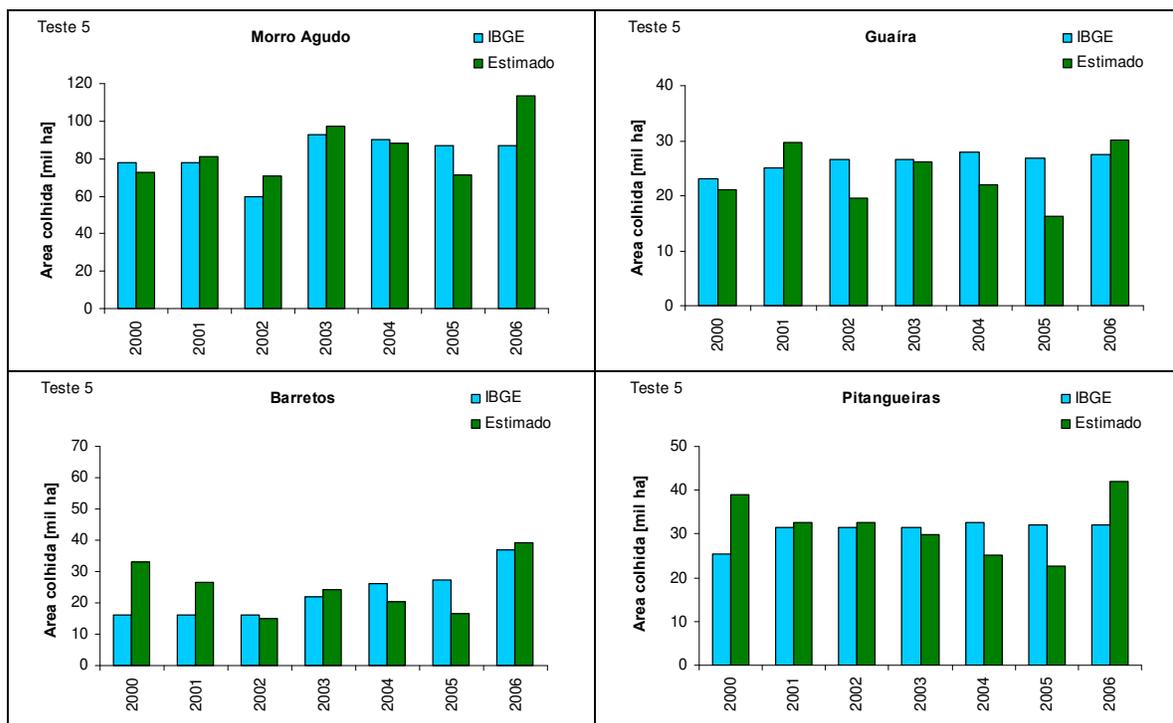
A fim de verificar se, em cada município, o comportamento do perfil temporal do NDVI nas áreas canavieiras seguia algum padrão ao longo das safras, analisou-se os dados fixando-se o mesmo ajuste de seleção dos perfis temporais (teste) para todas as safras.

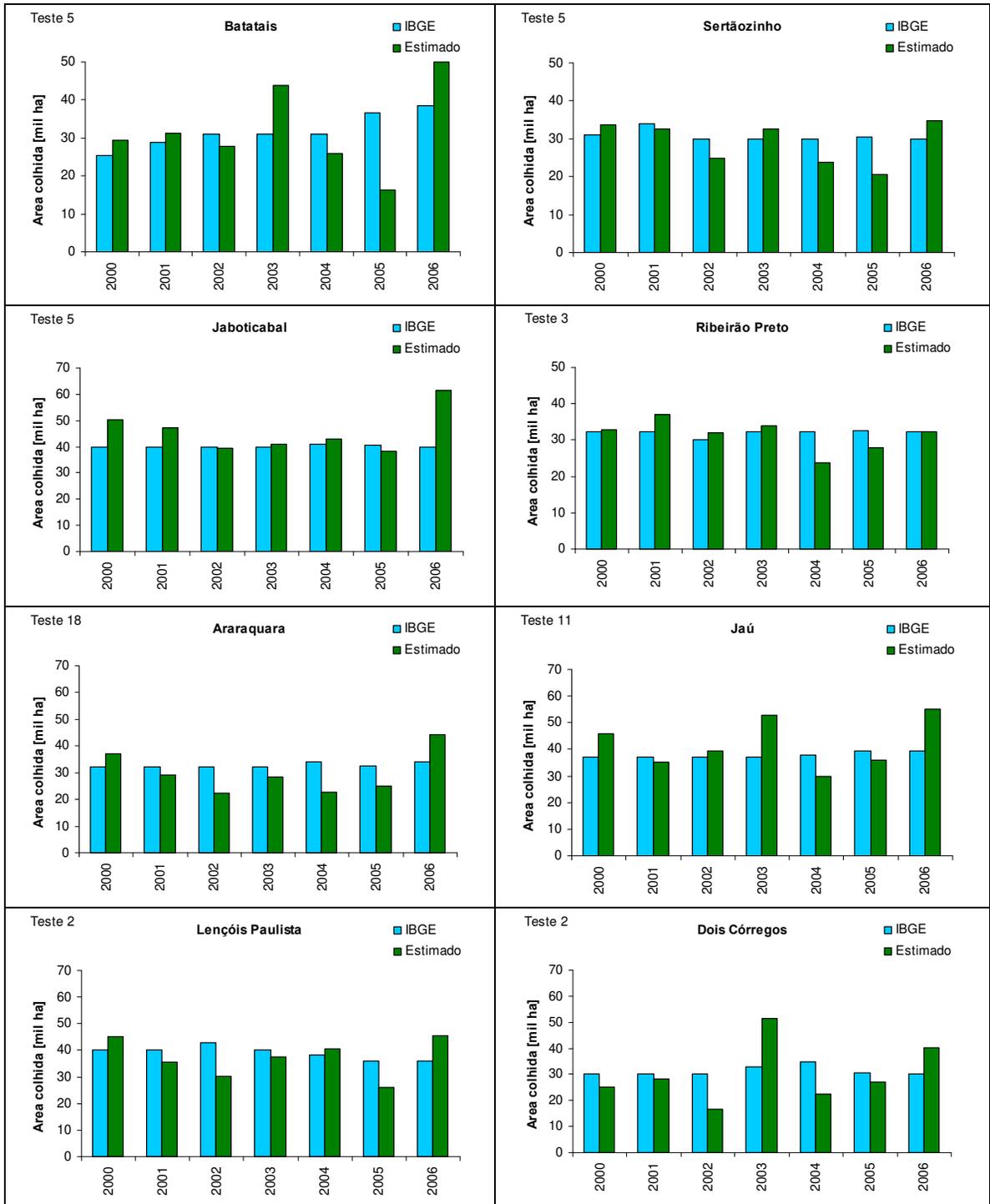
Para cada município foi selecionado o teste que apresentou resultado para as sete safras mais próximos dos dados oficiais, quanto à diferença entre a área oficial e a estimada. Na Tabela 5 são apresentados os testes que apresentaram os melhores resultados em cada município.

Tabela 5: Testes utilizados para identificação de áreas canavieiras em escala municipal, referente à segunda abordagem.

	Município	1999/2000	2000/2001	2001/2002	2002/2003	2003/2004	2004/2005	2005/2006
1	Morro Agudo	5	5	5	5	5	5	5
2	Jaboticabal	5	5	5	5	5	5	5
3	Paraguaçu Paulista	8	8	8	8	8	8	8
4	Batatais	5	5	5	5	5	5	5
5	Barretos	5	5	5	5	5	5	5
6	Piracicaba	9	9	9	9	9	9	9
7	Valparaíso	5	5	5	5	5	5	5
8	Jaú	11	11	11	11	11	11	11
9	Lençóis Paulista	2	2	2	2	2	2	2
10	Araraquara	18	18	18	18	18	18	18
11	Pitangueiras	5	5	5	5	5	5	5
12	São Manuel	3	3	3	3	3	3	3
13	Penápolis	7	7	7	7	7	7	7
14	Guaira	5	5	5	5	5	5	5
15	Guararapes	21	21	21	21	21	21	21
16	Araras	5	5	5	5	5	5	5
17	Dois Córregos	2	2	2	2	2	2	2
18	Sertãozinho	5	5	5	5	5	5	5
19	Novo Horizonte	23	23	23	23	23	23	23
20	Ribeirão Preto	3	3	3	3	3	3	3

Na Figura 23 são apresentados os resultados gráficos de estimativa de área para cada município, referentes aos testes selecionados na Tabela 5, para as sete safras.





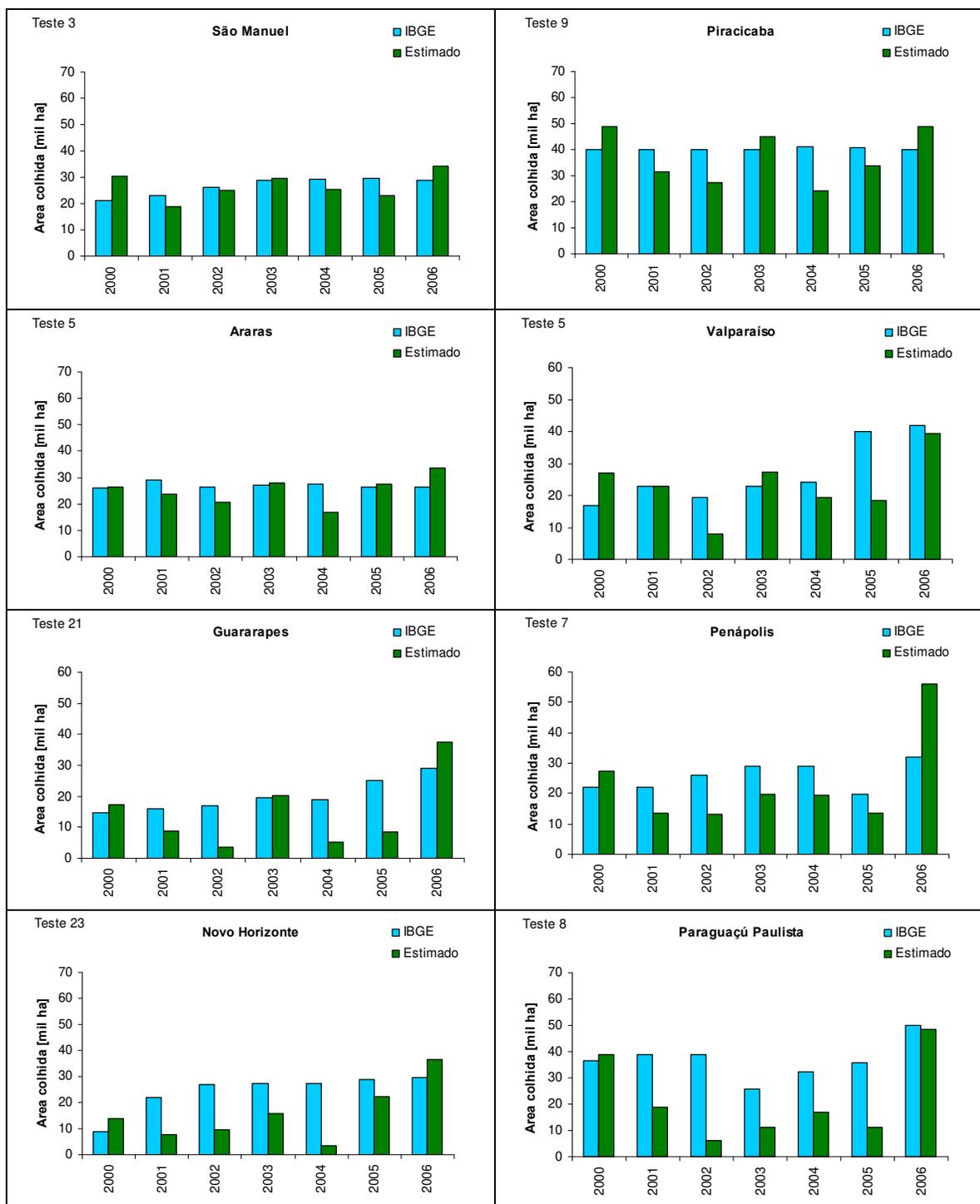


Figura 23: Resultados para cada município referentes à segunda abordagem.

Analisando a Figura 23 verifica-se que os municípios Morro Agudo, Sertãozinho, Jaboticabal, Ribeirão Preto, Lençóis Paulista, São Manuel e Araras, de modo geral, apresentaram resultados de estimativa de área próximos aos dados oficiais. Outros, como

Valparaíso, Guararapes, Penápolis, Novo Horizonte e Paraguaçu Paulista, apresentaram resultados bastante discrepantes.

Constatadas estas diferenças regionais nos resultados, foi analisado se estas estavam associadas à área de cana-de-açúcar relativa do município, partindo da hipótese de que municípios com maior cobertura de área plantada permitiam a obtenção de um maior número de pixels puros, e assim, perfis temporais mais característicos da cultura. Na Figura 24 é apresentado um mapa indicando a área relativa média de cana-de-açúcar em cada município, considerando as sete safras entre 1999 e 2006.

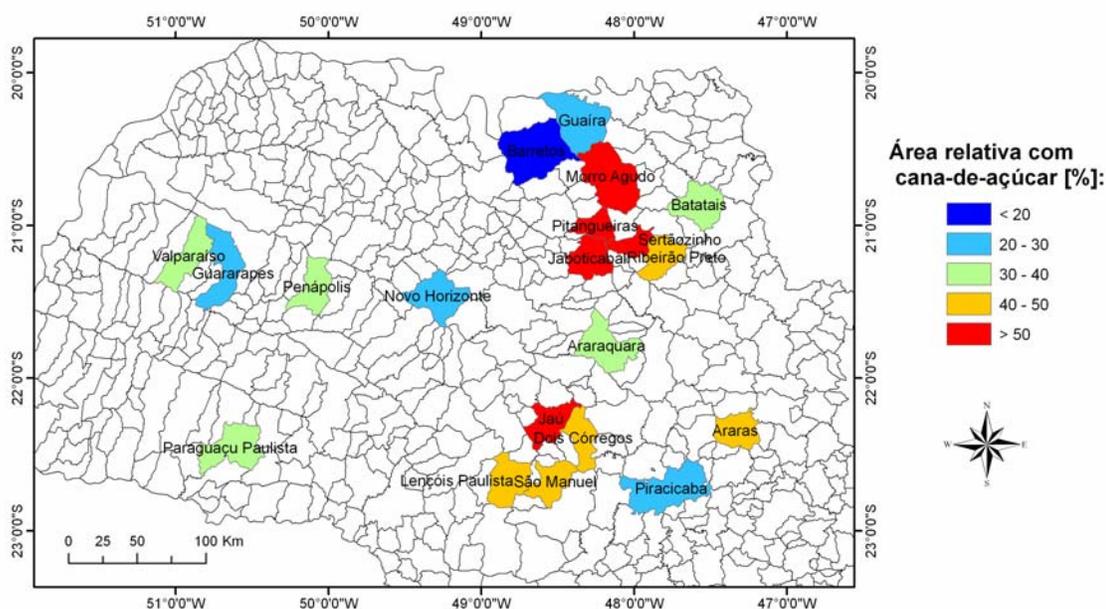


Figura 24: Mapa indicando para cada município a área relativa de cana-de-açúcar, média entre as sete safras selecionadas. Fonte: IBGE (2008).

Analisando a Figura 24 constata-se que a qualidade dos resultados, referente à segunda abordagem, pode estar relacionada à área relativa de cana-de-açúcar no município.

Na Figura 25 é apresentado um mapa indicando o número do teste que apresentou o melhor resultado referente à segunda abordagem.

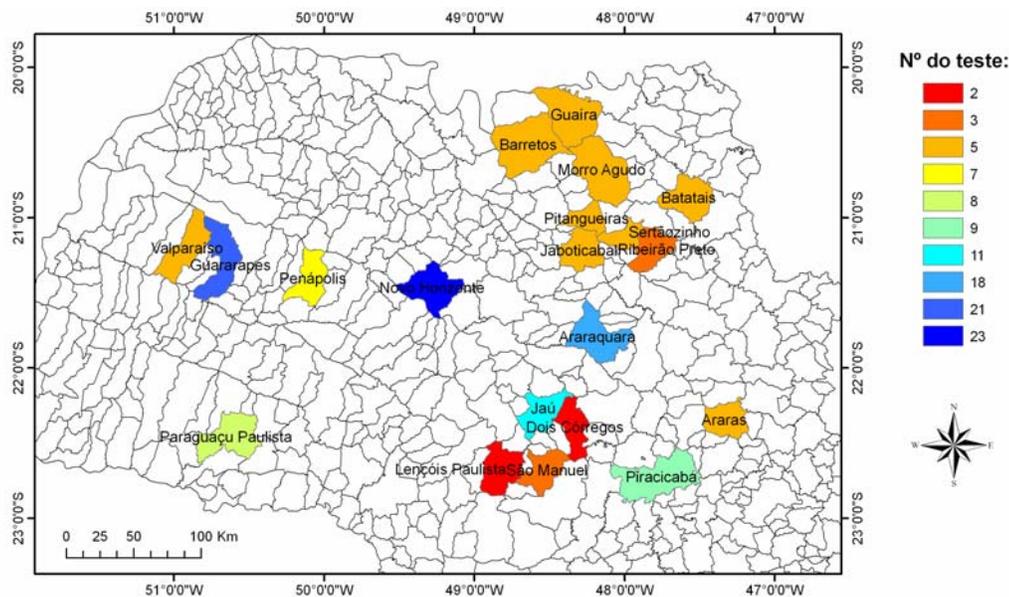


Figura 25: Mapa indicando para cada município o número do teste citado na Tabela 5.

Verifica-se na Figura 25 que há indícios de similaridade entre perfis temporais de municípios vizinhos. Isso pode estar associado ao número de pixels puros (pixels contendo apenas um tipo de alvo) obtidos dentro do município, pois quanto mais puro for o pixel, mais característico será seu perfil temporal. Na região nordeste do Estado o teste nº 5 foi o que predominou quanto à qualidade dos resultados. Mesmo teste significa perfis temporais similares. Assim, por ser uma região densamente ocupada por áreas canaveiras, esta região apresentou perfis temporais bastante característicos e maior qualidade nos resultados.

Já a região noroeste do Estado, por ser uma região de expansão canaveira, apresenta menor probabilidade de ocorrência de pixels puros, o que tornam os perfis temporais desta região menos característicos de área canaveira, se comparados aos perfis temporais da região nordeste.

O método proposto para identificação das áreas canaveiras no estado de São Paulo se mostrou eficaz em algumas áreas e ineficaz em outras. Devido à baixa resolução espacial, melhores resultados de classificação foram alcançados em regiões do Estado que apresentaram áreas plantadas maiores e mais homogêneas. Regiões de expansão canaveira, de forma geral, apresentaram resultados menos satisfatórios. Destaca-se que a expansão canaveira vem se dando em áreas até então de pastagens. A pastagem é uma cobertura vegetal que apresenta perfil temporal de NDVI semelhante ao da cana, característica essa também detectada por

Xavier et al. (2006). Dessa forma, a ocorrência de um plantio menos homogêneo associado a áreas de pastagens no entorno, dificultaram o mapeamento da cana em áreas de expansão no Estado por meio do método proposto.

## **4.2. Relação entre NDVI, dados meteorológicos e produtividade**

### **4.2.1 Determinação das fases de desenvolvimento**

A partir dos perfis médios municipais do NDVI e do conhecimento fisiológico da cultura buscou-se dividir o ciclo de desenvolvimento em quatro fases: perfilhamento, crescimento acelerado, crescimento lento e maturação. Para definir as fronteiras entre cada fase foram primeiramente analisados visualmente os perfis médios municipais separados por safra, apresentados na Figura 26.

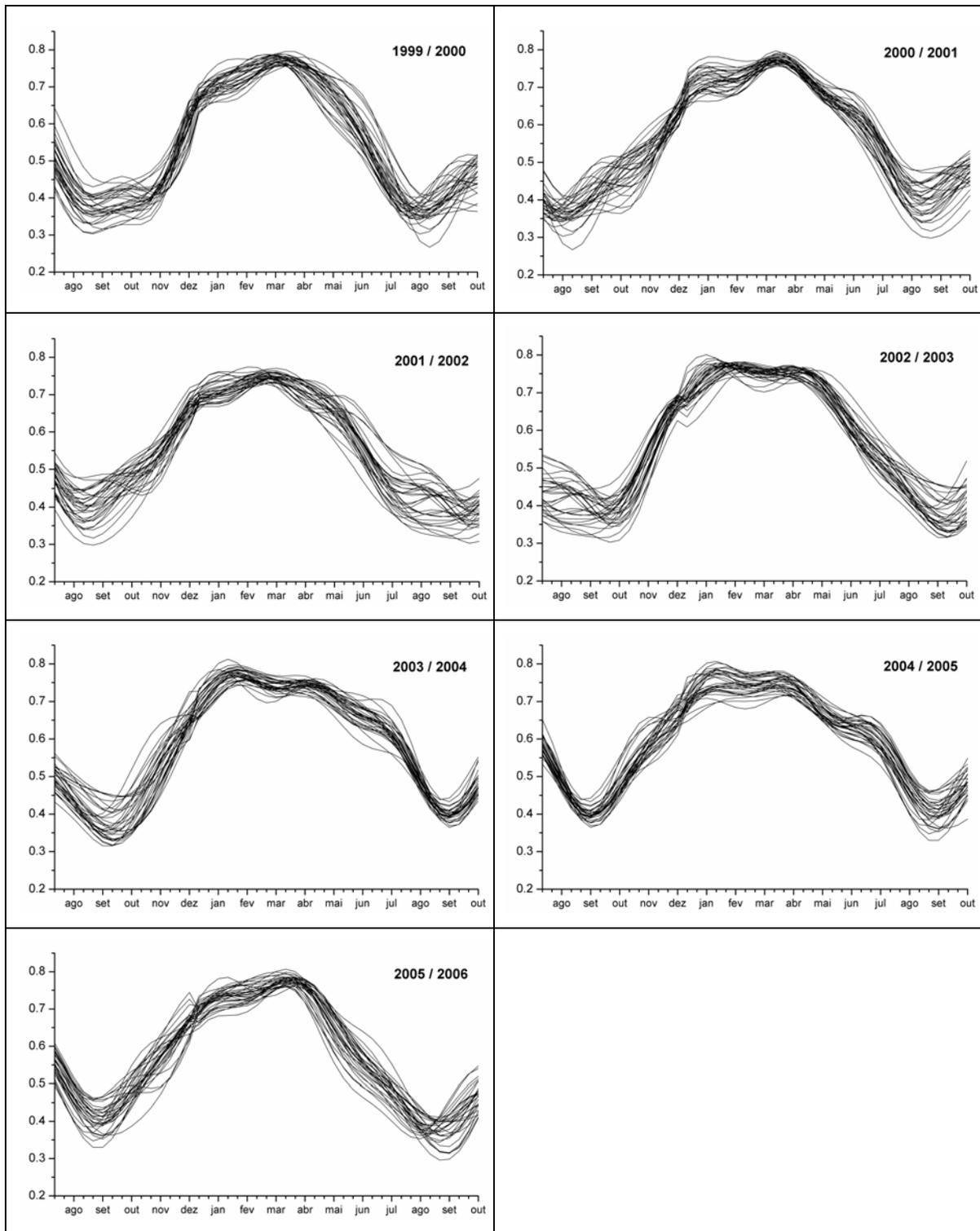


Figura 26: Perfis médios municipais de NDVI, separados por safra. Sete safras, entre 1999 e 2006.

Foi calculado para cada safra o perfil médio entre os 20 municípios, obtendo assim o perfil médio da safra. E, a partir de informações sobre a fisiologia e do comportamento destes perfis médios, foram definidas as fronteiras de cada fase. Os perfis médios da safra, bem como as fronteiras entre as quatro fases, estão apresentados na Figura 27.

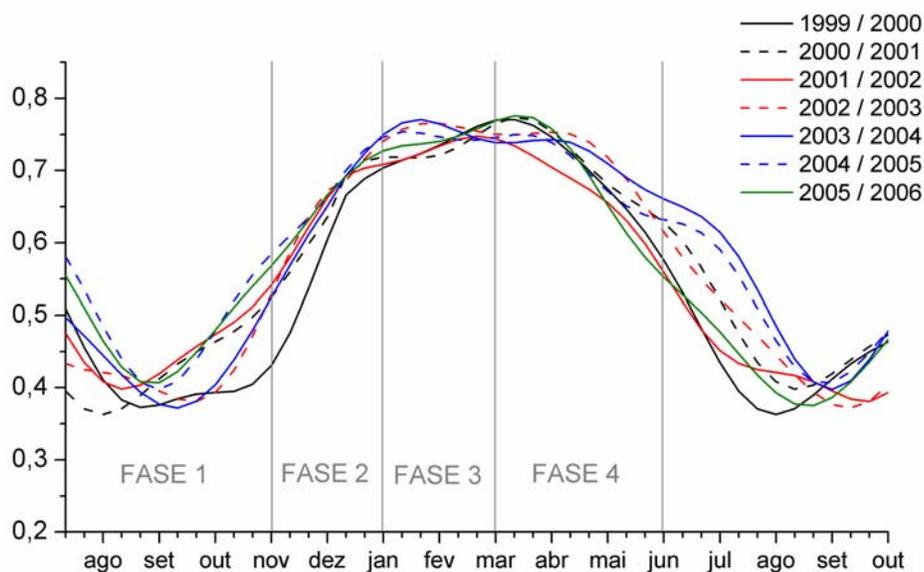


Figura 27: Perfis médios de NDVI para cada safra e definição das quatro fases do ciclo de desenvolvimento.

- fase 1: perfilamento, entre primeiro decêndio de agosto e terceiro decêndio de novembro;
- fase 2: crescimento rápido, entre primeiro decêndio de dezembro e terceiro decêndio de janeiro;
- fase 3: crescimento lento, entre primeiro decêndio de fevereiro e terceiro decêndio de março;
- fase 4: maturação, entre primeiro decêndio de abril e terceiro decêndio de junho.

Ressalta-se que a definição destas fases baseou-se em um comportamento médio da cultura.

#### 4.2.2. Geração dos atributos espectrais e meteorológicos a partir das fases de desenvolvimento

Definidas as fases do ciclo de desenvolvimento, foram gerados atributos para cada fase e para a safra como um todo a partir dos perfis do NDVI e meteorológicos de cada município / safra. Os atributos de cada fase, bem como sua descrição, são mostrados na Tabela 6.

Tabela 6: Atributos espectrais e meteorológicos gerados para cada fase de desenvolvimento da cultura e para toda a safra.

<b>Fase</b>	<b>Nome do Atributo</b>	<b>Descrição</b>	
<b>FASE 1</b>	1	ndvi1_s	Soma dos valores do NDVI da fase 1 (acumulado).
	2	ndvi1_m	Média aritmética entre os valores do NDVI da fase 1.
	3	ndvi1_final	Valor do NDVI no último decêndio da fase 1 (3º decêndio de novembro).
	4	rain1_s	Precipitação acumulada na fase 1.
	5	grs1_s	Radiação global acumulada na fase 1.
	6	tmin1_m	Media aritmética dos valores de temperatura mínima da fase 1.
	7	tmean1_m	Media aritmética dos valores de temperatura média da fase 1.
	8	tmax1_m	Media aritmética dos valores de temperatura máxima da fase 1.
<b>FASE 2</b>	9	ndvi2s	Soma dos valores do NDVI da fase 2 (acumulado).
	10	ndvi2_inic	Valor do NDVI no primeiro decêndio da fase 2 (1º decêndio de dezembro).
	11	ndvi2_m	Média aritmética entre os valores do NDVI da fase 2.
	12	ndvi2_final	Valor do NDVI no último decêndio da fase 2 (3º decêndio de janeiro).
	13	ndvi2_d	Diferença entre o valor do NDVI do último e do primeiro decêndios da fase 2.
	14	rain2s	Precipitação acumulada na fase 2.
	15	grs2s	Radiação global acumulada na fase 2.
	16	tmin2m	Media aritmética entre os valores de temperatura mínima da fase 2.
	17	Tmean2m	Media aritmética entre os valores de temperatura média da fase 2.
	18	tmax2m	Media aritmética entre os valores de temperatura máxima da fase 2.
<b>FASE 3</b>	19	ndvi3s	Soma dos valores do NDVI da fase 3 (acumulado).
	20	ndvi3_inic	Valor do NDVI no primeiro decêndio da fase 3 (1º decêndio de fevereiro).
	21	ndvi3_m	Média aritmética entre os valores do NDVI da fase 3.
	22	ndvi3_final	Valor do NDVI no último decêndio da fase 3 (3º decêndio de março).
	23	ndvi3_d	Diferença entre o valor do NDVI do último e do primeiro decêndios da fase 3.
	24	rain3s	Precipitação acumulada na fase 3.
	25	grs3s	Radiação global acumulada na fase 3.
	26	tmin3m	Media aritmética entre os valores de temperatura mínima da fase 3.
	27	Tmean3m	Media aritmética entre os valores de temperatura média da fase 3.
	28	tmax3m	Media aritmética entre os valores de temperatura máxima da fase 3.
<b>FASE 4</b>	29	ndvi4s	Soma dos valores do NDVI da fase 4 (acumulado).
	30	ndvi4_inic	Valor do NDVI no primeiro decêndio da fase 4 (1º decêndio de abril).
	31	ndvi4_m	Média aritmética entre os valores do NDVI da fase 4.
	32	ndvi4_final	Valor do NDVI no último decêndio da fase 4 (3º decêndio de junho).
	33	ndvi4_d	Diferença entre o valor do NDVI do último e do primeiro decêndios da fase 4.
	34	rain4s	Precipitação acumulada na fase 4.
	35	grs4s	Radiação global acumulada na fase 4.
	36	tmin4m	Media aritmética entre os valores de temperatura mínima da fase 4.
	37	tmean4m	Media aritmética entre os valores de temperatura média da fase 4.
	38	tmax4m	Media aritmética entre os valores de temperatura máxima da fase 4.

<b>TODA SAFRA*</b>	39	ndvis	Soma dos valores do NDVI em todas as fases (acumulado entre agosto e junho).
	40	ndvi_min	Menor valor do NDVI ocorrido na safra.
	41	dec_min	Decêndio em que ocorreu o menor valor do NDVI, a partir do 1º decêndio de agosto.
	42	ndvi_m	Média aritmética entre os valores do NDVI de toda a safra.
	43	ndvi_max	Maior valor do NDVI ocorrido na safra.
	44	dec_max	Decêndio em que ocorreu o maior valor do NDVI, a partir do 1º decêndio de agosto.
	45	ndvi_d	Diferença entre o maior e o menor valor do NDVI ocorridos na safra.
	46	periodo_mM	Número de decêndios (duração) entre a ocorrência do menor e do maior valor do NDVI da safra.
	47	rains	Precipitação acumulada na safra.
	48	grss	Radiação global acumulada na safra.
	49	tminm	Media aritmética entre os valores de temperatura mínima da safra.
	50	tmeanm	Media aritmética entre os valores de temperatura média da safra.
	51	tmaxm	Media aritmética entre os valores de temperatura máxima da safra.

Este procedimento foi aplicado para todos os municípios e safras. A partir destes atributos analisaram-se as possíveis relações entre os mesmos e entre estes e os índices de produtividade municipais. Os dados referentes à produtividade da cana-de-açúcar, para cada município e safra, foram obtidos no banco de dados sobre produção agrícola municipal do IBGE (2008).

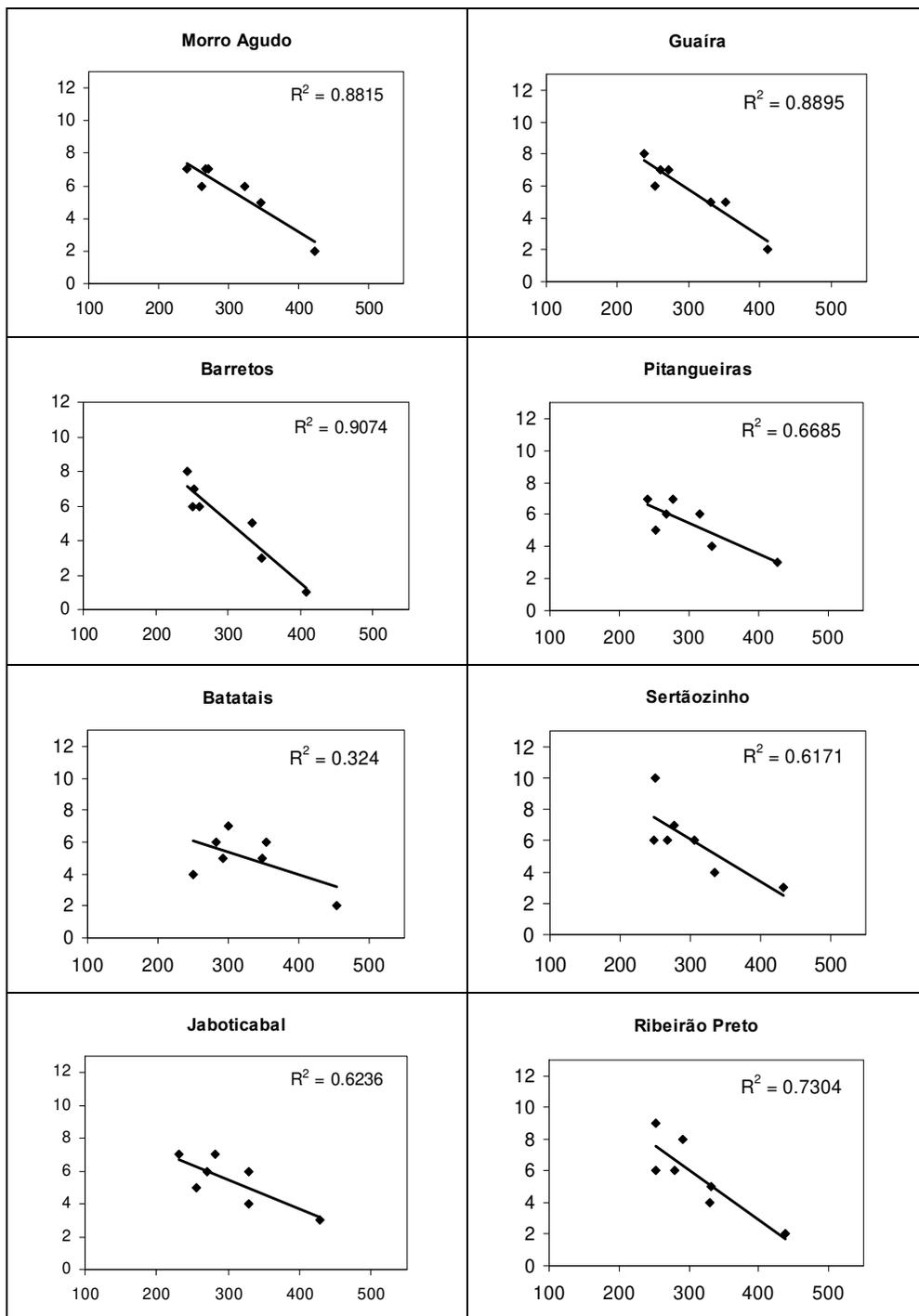
#### **4.2.3. Correlações entre NDVI e variáveis meteorológicas**

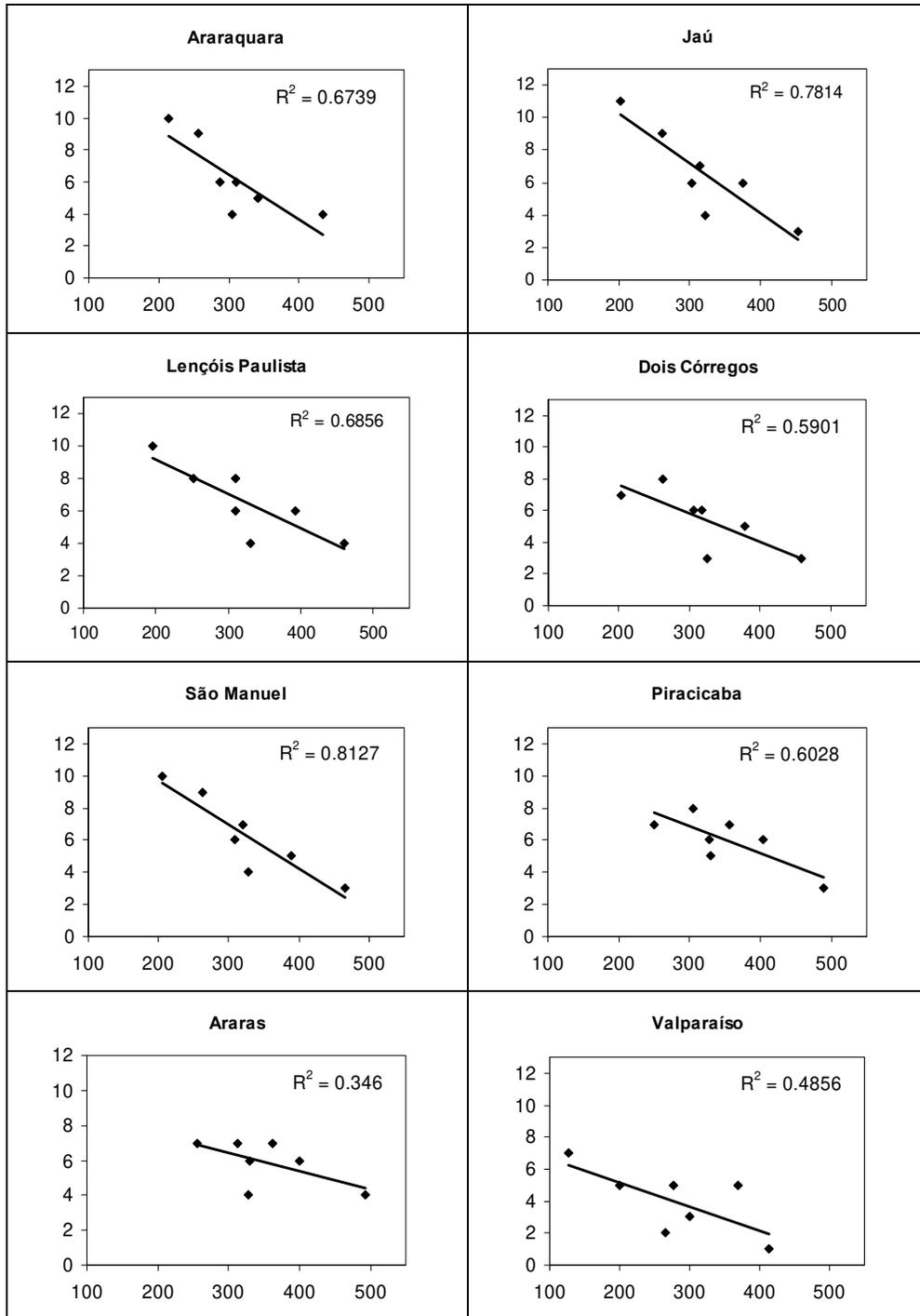
Alguns atributos espectrais podem fornecer informações importantes a respeito do comportamento da cultura. Aqui, foram verificadas possíveis correlações entre atributos espectrais e meteorológicos.

##### **4.2.3.1. Determinação do início do ciclo de desenvolvimento**

O atributo dec\_min (decêndio de ocorrência do valor mínimo do NDVI no perfil temporal) pode ser interpretado como o momento de início do crescimento vegetativo da cultura, ou seja, início da safra. Foi identificada correlação entre atributo espectral dec\_min e o atributo meteorológico rain1\_s.

Na Figura 28 são apresentados os resultados da correlação entre estes dois atributos, por município, e os respectivos índices de determinação.





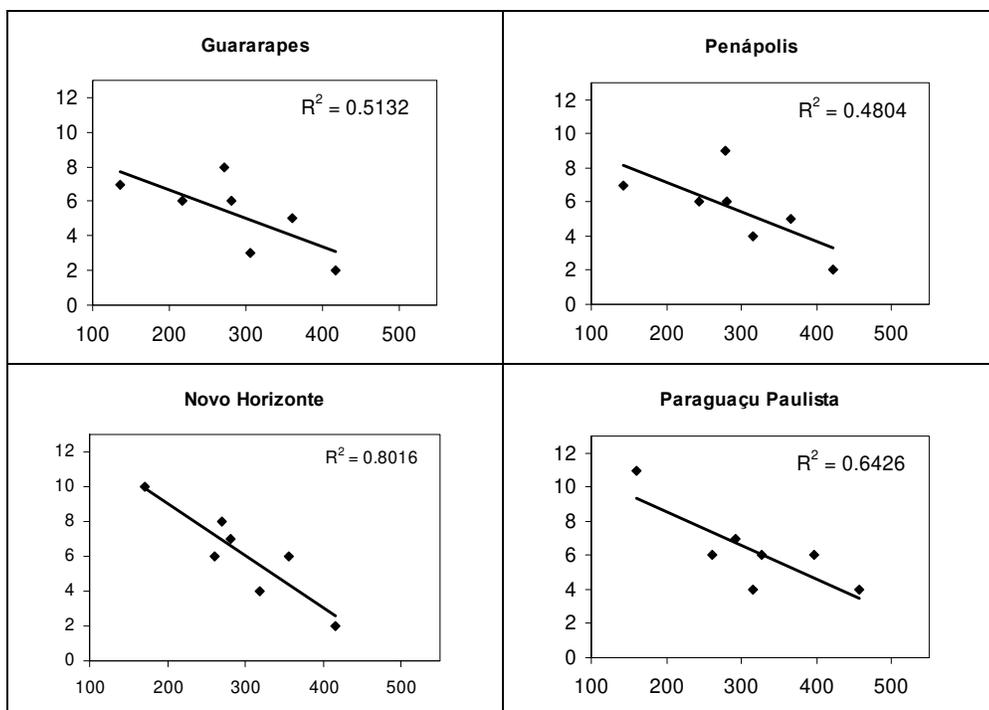


Figura 28: Correlação entre precipitação acumulada na fase 1 (eixo X, em mm) e decêndio de ocorrência do valor mínimo do NDVI (eixo Y, em decêndios, iniciando no 1º decêndio de agosto). Cada ponto no gráfico é referente a uma safra.

Observa-se que o decêndio de ocorrência do valor mínimo do NDVI (*dec\_mim*), interpretado como o início do crescimento vegetativo da cultura, está inversamente correlacionado com a precipitação acumulada na fase de perfilhamento (fase 1), ou seja, quanto maior a precipitação no início da safra, mais cedo iniciará o crescimento da cana-de-açúcar e vice-versa.

Verifica-se que, apesar dos valores de R<sup>2</sup> não ter sido satisfatórios em todos os municípios, a tendência se manteve a mesma, uma correlação inversa entre a quantidade de chuva e o início do crescimento da cultura, mostrando coerência com a realidade no campo.

#### 4.2.4. Classificação da produtividade por meio dos atributos espectrais e meteorológicos

Foram usadas técnicas de mineração de dados para buscar compreender como os atributos espectrais e meteorológicos poderiam contribuir para determinação da produtividade média municipal. Embora o tamanho do banco de dados, no que se refere ao número de instâncias (140), não fosse grande o suficiente para permitir maior eficiência no uso das técnicas, como recomendam especialistas em mineração de dados, optou-se pelo uso destas

técnicas como forma de identificar os atributos mais relevantes e de que forma estes poderiam estar relacionados com os dados de produtividade oficiais. Os resultados encontrados através destas técnicas também foram analisados no Excel, a fim de ampliar a compreensão e verificar a coerência dos resultados.

Para a preparação e análise do banco de dados foram realizados os seguintes passos: Discretização do atributo classe, Seleção de atributos e Classificação.

#### 4.2.4.1. Discretização do atributo classe

Com o intuito de permitir o uso de técnicas de seleção de atributos e de classificação, o atributo classe foi discretizado e classificado. O atributo classe “produtividade” foi discretizado e classificado em 3 classes por meio de percentis, conforme apresentado na Tabela 7.

Tabela 7: Limites inferior e superior para as três classes geradas para o atributo produtividade após discretização e classificação. Os valores referem-se à produtividade, em toneladas por hectare.

<b>Classe</b>	<b>Descrição</b>	<b>Limite inferior</b>	<b>Limite superior</b>
B-M	Baixo-médio	60	73
M	Médio	74	85
M-A	Médio-alto	86	110

#### 4.2.4.2. Seleção de atributos

Após discretização e classificação do atributo classe, foram aplicados ao banco de dados 4 métodos de seleção de atributos, disponíveis no software Weka 3.5. O objetivo foi determinar quais atributos seriam selecionados pelo maior número de métodos. Os resultados completos de cada método estão no Anexo 6. Na Tabela 8 são apresentados os atributos selecionados por cada método de seleção:

- 1) teste do Qui-quadrado;
- 2) método Wrapper com algoritmo de árvore de decisão J48;
- 3) método CFS (Correlation Feature Selection);

4) combinação dos métodos InfoGAin (Ganho de Informação) e GainRatio (Taxa de Ganho).

Tabela 8: Atributos selecionados em cada método de seleção.

Fase	Atributos	Métodos de seleção de atributos			
		1	2	3	4
FASE 1	ndvi1_s				
	ndvi1_m				
	ndvi1_final		X		
	rain1_s				
	grs1_s				
	tmin1_m				
	tmean1_m				
	tmax1_m				
FASE 2	ndvi2s				
	ndvi2_inic				
	<b>ndvi2_m</b>	X		X	X
	<b>ndvi2_final</b>	X		X	X
	ndvi2_d				
	rain2s				
	grs2s				
	tmin2m				
	tmean2m				
	tmax2m				
FASE 3	<b>ndvi3s</b>	X		X	X
	<b>ndvi3_inic</b>	X	X	X	X
	ndvi3_m				
	ndvi3_final				
	ndvi3_d				
	rain3s				
	grs3s				
	tmin3m				
	tmean3m				
	tmax3m				
FASE 4	ndvi4s				
	ndvi4_inic				
	ndvi4_m				
	ndvi4_final				
	ndvi4_d				
	rain4s				
	grs4s				
	tmin4m				
	tmean4m				
	tmax4m				

<b>TODA SAFRA</b>	ndvis				
	ndvi_min				
	dec_min				
	<b>ndvi_m</b>	X		X	X
	ndvi_max		X		
	dec_max	X			X
	ndvi_d				
	periodo_mM				
	rains				
	grss				
	tminm				
	tmeanm				
	tmaxm				

Considerando os resultados dos 4 métodos conjuntamente, concluiu-se que os atributos ndvi2\_m, ndvi2\_final, ndvi3\_s, ndvi3\_inic e ndvi\_m (destacados em cinza na Tabela 8), dentre os atributos gerados, seriam os mais relevantes para a determinação da produtividade municipal. Observa-se que não foram selecionados atributos gerados a partir dos dados meteorológicos.

#### 4.2.4.3. Classificação por meio de árvore de decisão

Após seleção dos atributos ndvi2\_m, ndvi2\_final, ndvi3\_s, ndvi3\_inic e ndvi\_m, foi aplicado o algoritmo de classificação J48, baseado em árvore de decisão. Foi utilizada uma pré-poda, definindo o número mínimo de objetos igual a 10. O resultado completo da classificação é apresentado no Anexo 7.

O classificador classificou corretamente 70% das instâncias, o que correspondeu a 98 entre o total de 140 instâncias.

Na Tabela 9 são apresentadas as quantidades de instâncias pertencentes a cada classe e o número de instâncias classificadas corretamente, bem como a respectiva porcentagem de acerto.

Tabela 9: Resultado da classificação da produtividade, para cada classe, utilizando-se os atributos ndvi2\_m, ndvi2\_final, ndvi3\_s, ndvi3\_inic e ndvi\_m.

<b>Classe</b>	<b>Original</b>	<b>Classificado corretamente</b>	<b>Acerto [%]</b>
<b>B-M</b>	24	15	62,5
<b>M</b>	74	55	74,3
<b>M-A</b>	42	28	66,7

Na Tabela 10 é apresentada a matriz de confusão para a classificação realizada. Observa-se que há certo equilíbrio e coerência nos resultados, uma vez que a confusão concentrou-se mais entre classes vizinhas (B-M e M ou M e M-A) ao invés de classes extremas (B-M e M-A). Esta confusão pode estar associada ao método utilizado de discretização do atributo classe.

Tabela 10: Matriz de confusão do classificador J48 para a classificação da produtividade a partir dos atributos ndvi2\_m, ndvi2\_final, ndvi3\_s, ndvi3\_inic e ndvi\_m.

<b>ORIGINAL</b> <b>CLASSIF.</b>	<b>B-M</b>	<b>M</b>	<b>M-A</b>
<b>B-M</b>	<b>15</b>	9	2
<b>M</b>	6	<b>55</b>	12
<b>M-A</b>	3	10	<b>28</b>

Na Figura 29 é apresentada a árvore de decisão para determinação da produtividade municipal, obtida aplicando-se aos atributos selecionados (ndvi2\_m, ndvi2\_final, ndvi3\_s, ndvi3\_inic e ndvi\_m) o algoritmo de classificação J48, com número mínimo de objetos igual a 10.

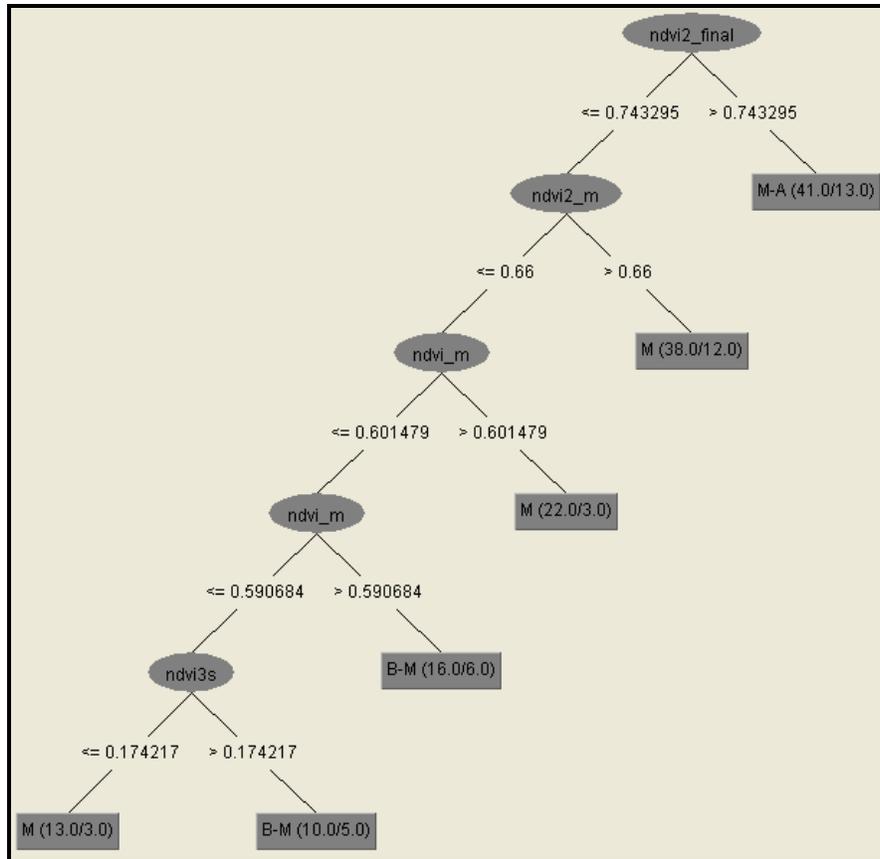


Figura 29: Árvore de decisão para determinação da produtividade obtida aplicando-se aos atributos ndvi2\_m, ndvi2\_final, ndvi3\_s, ndvi3\_inic e ndvi\_m o algoritmo de classificação J48, com número mínimo de objetos igual a 10.

Observa-se que os atributos ndvi2\_final e ndvi2\_m estão no topo da árvore, apresentando maior força na classificação da produtividade. É interessante destacar que se tratam de atributos provenientes da fase 2, referente ao período entre início de dezembro e final de janeiro e, portanto, com antecedência de, no mínimo, dois meses em relação ao início do período de colheita.

Decidiu-se então realizar nova classificação utilizando-se somente os atributos ndvi2\_final e ndvi2\_m. O algoritmo de classificação continuou sendo o J48, utilizando a pré-poda com número mínimo de objetos igual a 10. O resultado completo desta classificação está apresentado no Anexo 8.

O classificador classificou corretamente 67% das instâncias, o que corresponde a 94 entre o total de 140 instâncias. Na Tabela 11 são apresentados os resultados desta nova classificação.

Tabela 11: : Resultado da classificação da produtividade, para cada classe, utilizando-se os atributos ndvi2\_m e ndvi2\_final.

<b>Classe</b>	<b>Original</b>	<b>Classificado corretamente</b>	<b>Acerto [%]</b>
<b>B-M</b>	24	2	8,3
<b>M</b>	74	64	86,5
<b>M-A</b>	42	28	66,7

Na Tabela 12 é apresentada a matriz de confusão para esta nova classificação.

Tabela 12: Matriz de confusão do classificador J48 para a classificação da produtividade a partir dos atributos ndvi2\_m, ndvi2\_final, ndvi3\_s, ndvi3\_inic e ndvi\_m.

<b>ORIGINAL</b> <b>CLASSIF.</b>	<b>B-M</b>	<b>M</b>	<b>M-A</b>
<b>B-M</b>	<b>2</b>	0	0
<b>M</b>	19	<b>64</b>	14
<b>M-A</b>	3	10	<b>28</b>

Analisando as Tabelas 11 e 12 observa-se que houve piora significativa para a classificação da produtividade Baixo-médio (B-M), caindo de 62,5 para 8,3 a porcentagem de acerto desta classe. A classificação da produtividade Médio (M) melhorou, aumentando de 74,3 para 86,5 a porcentagem de acerto.

Comparando as Tabelas 9 e 10 com as Tabelas 11 e 12, verifica-se que ao usar somente os atributos ndvi2\_m e ndvi2\_final o resultado para a classificação da classe M-A se manteve o mesmo.

Na Figura 30 é apresentada a árvore de decisão para determinação da produtividade municipal, obtida utilizando somente os atributos ndvi2\_m e ndvi2\_final.

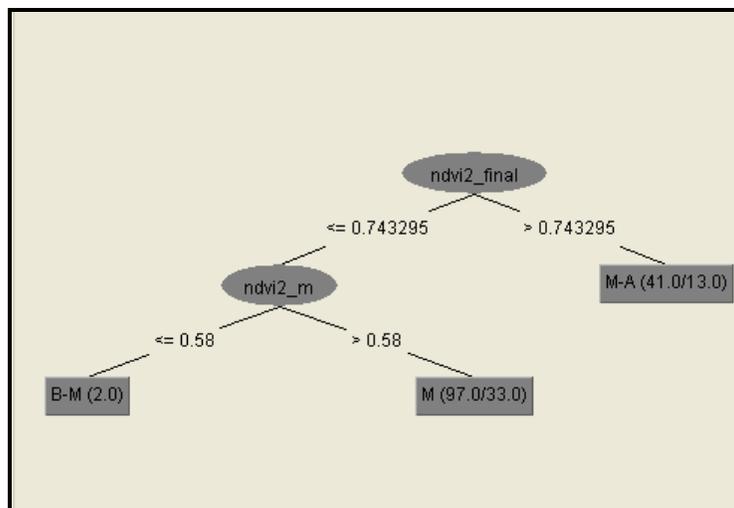


Figura 30: Árvore de decisão para determinação da produtividade obtida aplicando-se aos atributos ndvi2\_final e ndvi2\_m o algoritmo de classificação J48, com número mínimo de objetos igual a 10.

Analisando os resultados das duas classificações conclui-se que, para determinação das classes de produtividade M e M-A, é possível obter resultados satisfatórios utilizando-se somente os atributos ndvi2\_final e ndvi2\_m, o que não acontece com a classe B-M. Ao analisar as árvores de decisão das Figuras 29 e 30, verifica-se que somente o atributo ndvi2\_final foi determinante para a classificação da classe M-A.

É importante destacar que outros fatores, além dos considerados neste trabalho, influenciam nos índices de produtividade alcançados. Atenção especial deve ser dada à classe B-M, referente a baixos índices de produtividade, pois, mesmo com condições de campo favoráveis ao desenvolvimento da cultura, baixas produtividades podem ocorrer se fatores econômicos, por exemplo, não estiverem favoráveis. No entanto, para que ocorram altos índices de produtividade é necessário que as condições de campo estejam favoráveis.

A fim de buscar mais evidências sobre a importância dos atributos ndvi2\_final e ndvi2\_m para a determinação da produtividade municipal e verificar se estes atributos apresentariam a mesma tendência apontada pelas árvores de decisão, foram feitas no Excel duas verificações de correlação para cada atributo.

Na primeira verificação, calculou-se a média do atributo espectral entre os 20 municípios para cada safra e correlacionou-se este com a produtividade média da safra, também calculada considerando os 20 municípios de interesse. Na Tabela 13 é apresentada a produtividade média de cada safra, considerando os 20 municípios.

Tabela 13: Produtividade média para cada safra, considerando os 20 municípios selecionados.

Safra	Produtividade média (ton/ha)
1999 / 2000	77,4
2000 / 2001	79,1
2001 / 2002	81,2
2002 / 2003	81,9
2003 / 2004	82,0
2004 / 2005	84,2
2005 / 2006	82,0

Nas Figuras 31a e 31b são apresentados os gráficos de dispersão e índices de determinação entre os atributos espectrais (ndvi2\_final e ndvi2\_m) e a produtividade média para cada safra. Cada ponto no gráfico se refere a uma safra. Os valores médios de cada safra referem-se à média de cada atributo entre os 20 municípios.

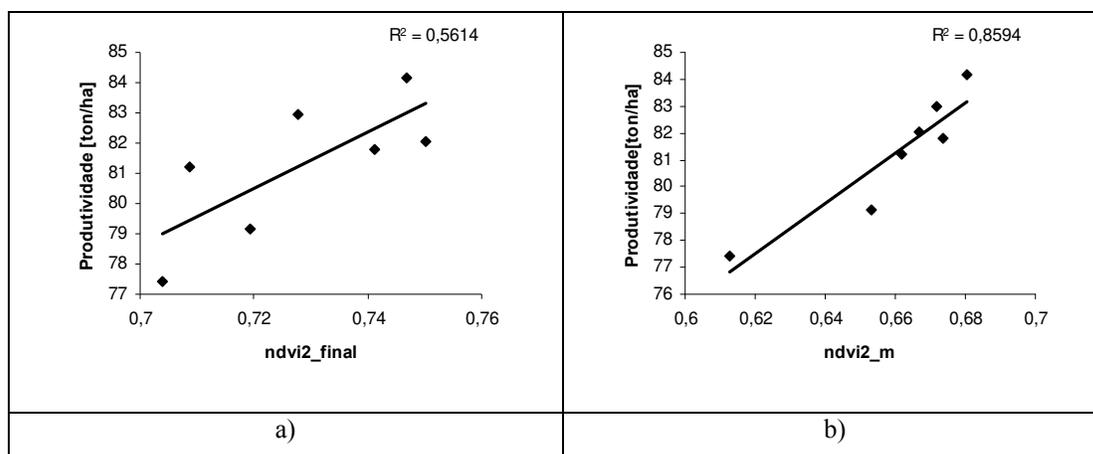


Figura 31: a) Correlação entre o ndvi final na fase 2 e a produtividade média de cada safra (média entre os 20 municípios). b) Correlação entre o ndvi médio na fase 2 e a produtividade média de cada safra (média entre os 20 municípios).

Na segunda verificação, foram analisadas as relações entre os mesmos atributos espectrais (ndvi2\_final e ndvi2\_m) e os dados de produtividade, porém analisando neste caso a média histórica dos atributos espectrais e da produtividade para cada município. Cada ponto dos gráficos se refere ao valor médio de um município. Nas Figuras 32a e 32b são apresentados os gráficos de dispersão e índices de determinação.

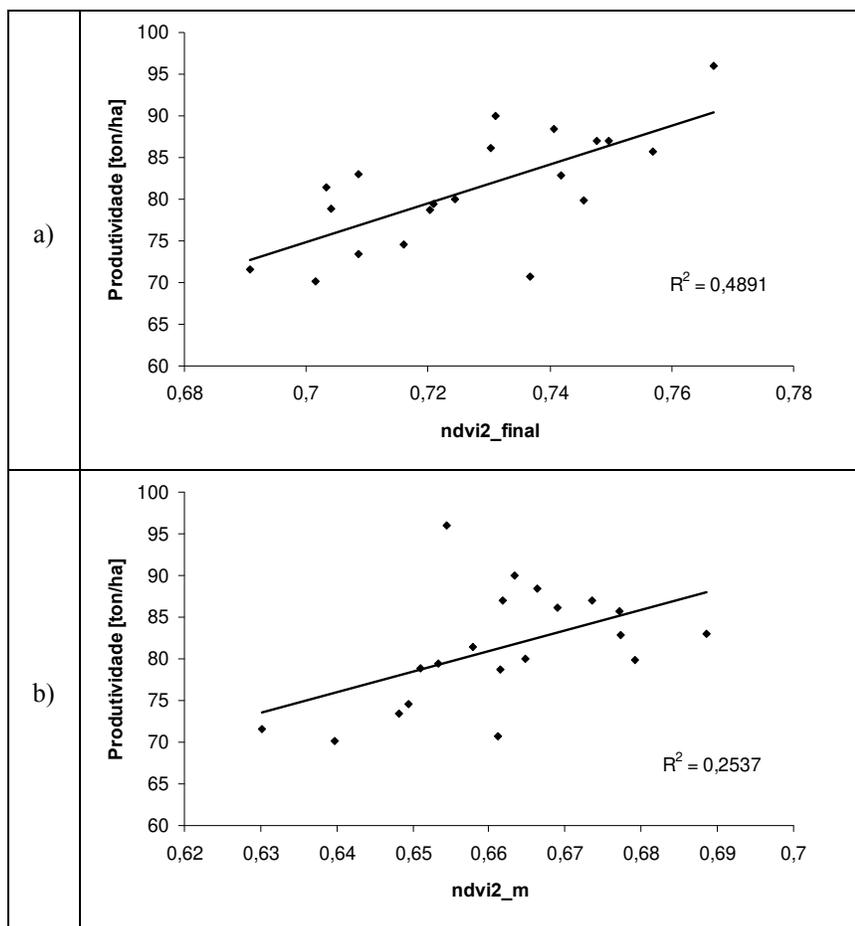


Figura 32: a) Correlação entre o ndvi final na fase 2 e a produtividade média de cada município - série histórica entre 1999 e 2006. b) Correlação entre o ndvi médio na fase 2 e a produtividade média de cada município - série histórica entre 1999 e 2006.

Analisando os resultados apresentados nas Figuras 31 e 32, verifica-se que a tendência apontada na árvore de decisão da Figura 30 para a relação entre os atributos espectrais e a produtividade está coerente com os resultados apresentados nos gráficos de dispersão, ou seja, valores mais altos indicando maior produtividade e vice-versa. Apesar dos índices de determinação das Figuras 31 e 32 serem baixos, é importante destacar que a tendência apontada foi coerente com os resultados apontados na classificação. Baixos índices de determinação já eram esperados, uma vez que a determinação da produtividade não depende apenas dos atributos espectrais apresentados.

## 5. CONCLUSÕES

Foi possível identificar, em escala estadual, importantes áreas canavieiras no estado de São Paulo por meio das imagens decendiais de baixa resolução espacial do sensor SPOT *Vegetation*. A qualidade do resultado variou para diferentes regiões do estado. Regiões que possuíam grandes e homogêneas áreas plantadas possibilitaram a obtenção dos melhores resultados. Já nas regiões de expansão das áreas canavieiras os resultados foram inferiores. Os índices de determinação ( $R^2$ ), obtidos através da correlação entre a área colhida de cana-de-açúcar oficial e a estimada, por município e safra, considerando os 637 municípios do Estado, variaram entre 0,771 e 0,829. No entanto, não foi possível determinar com exatidão a área de cana-de-açúcar colhida de cada município devido à baixa resolução espacial das imagens.

O momento do início do crescimento vegetativo da cultura, tratado aqui como sendo o decêndio no qual ocorre o valor mínimo do NDVI no perfil temporal, está inversamente correlacionado com a quantidade de chuva acumulada na fase considerada de perfilhamento, que se inicia em agosto e se estende até o final de novembro. O resultado de  $R^2$  variou para cada município, desde 0,324 em Batatais até 0,900 em Barretos, sendo que na média, considerando os 20 municípios, o  $R^2$  foi de 0,653. Apenas 4 municípios apresentaram  $R^2$  abaixo de 0,500. Dessa forma, é possível avaliar o momento do início da safra a partir de informações sobre precipitação na fase 1.

Houve relação entre atributos espectrais e a produtividade média municipal. Por meio do perfil médio municipal do NDVI foi possível extrair valores, como o valor médio do NDVI entre os meses de dezembro e janeiro e o valor do NDVI no 3º decêndio de janeiro, que permitiram classificar a tendência do resultado de produtividade alcançado no final da safra. Os atributos espectrais da fase 2 apontaram esta tendência de forma mais efetiva para os casos onde a produtividade foi classificada como sendo média ou acima da média, com acertos de 86,5 e 66,7% respectivamente. Para municípios e safras onde a produtividade foi classificada como abaixo da média, somente com os valores espectrais da fase 2 não foi possível obter bons resultados. Para o método proposto, os dados espectrais tiveram mais relevância que os dados meteorológicos na classificação da produtividade.

Existe potencial no uso de imagens de NDVI do SPOT *Vegetation* para a identificação de áreas canavieiras em escala estadual e para a classificação da produtividade média municipal da safra.

## 6. CONSIDERAÇÕES E SUGESTÕES

### **Identificação das áreas canavieiras do estado de São Paulo**

O ajuste dos parâmetros  $\Delta a$ ,  $\Delta b$  e corte mínimo, mostrados na Figura 12, somente foi possível após a aplicação do filtro Hants para suavização do perfil temporal do NDVI em imagens decendiais do SPOT *Vegetation*.

A análise dos resultados em escala municipal mostrou que cada região do Estado apresenta características próprias, sendo necessário tratá-las de forma diferenciada. Uma sugestão para aprimoramento da técnica de classificação por meio de perfis temporais seria a aplicação da máscara em sub regiões de comportamento homogêneo no Estado.

Constataram-se dois desafios para o mapeamento de culturas: a identificação das áreas de interesse e a exclusão dos alvos indesejáveis. Este método teve como ponto positivo o cumprimento destes dois papéis, embora com qualidade variável de acordo com a região do Estado. Trata-se de um método interessante para separar diferentes usos da terra de acordo com o perfil espectral de cada alvo devido à alta resolução temporal do sensor, que permite obter um perfil espectral detalhado.

### **Relações entre NDVI, dados meteorológicos e produtividade**

A cana-de-açúcar é uma cultura complexa. Diferentes variedades e cortes apresentam diferentes índices de produtividade. Associado a esses fatores, a qualidade do manejo de cada usina e sua abrangência no município influenciam diretamente na produtividade média municipal alcançada. Por fim, fatores econômicos também influenciam na produtividade final, principalmente quando se trata de momentos desfavoráveis da economia que, independente das condições de campo e desenvolvimento da cultura, refletirão em baixos índices de produtividade. Os baixos preços podem levar o produtor a não realizar a colheita de toda a área, conseqüentemente diminuindo a produtividade média da usina e, assim, a produtividade média do município.

Neste trabalho foram usados perfis médios municipais do NDVI, desconsiderando informações sobre a variedade e o corte devido à baixa resolução espacial. Porém, o

conhecimento sobre a proporção da área ocupada por cada variedade e em qual corte se encontra podem fornecer subsídios que indiquem com mais precisão a produtividade esperada.

A proposta deste trabalho foi avaliar o potencial de uso de insumos de fácil aquisição, como imagens do SPOT Vegetation e dados do modelo ECMWF, disponibilizados gratuitamente, na obtenção de um método simples que pudesse subsidiar o atual sistema brasileiro de previsão de safras para a cana-de-açúcar com informações mais consistentes.

Avanços na área de tecnologia da informação e comunicação via *Web*, associados ao conhecimento técnico na agricultura, permitem a cada dia o tratamento e transmissão de grande volume de dados e informações em tempo cada vez menor. O projeto Webmaps é um grande exemplo do potencial desta fusão. Trata-se de um projeto desenvolvido conjuntamente entre o Laboratório de Sistemas de Informação do Instituto de Computação e do Laboratório de Geoprocessamento da Faculdade de Engenharia Agrícola, ambos da Universidade Estadual de Campinas, que busca desenvolver um sistema de informação via *Web* que subsidie o planejamento e monitoramento da agricultura no Brasil. Dessa forma, é necessário que se busque ferramentas que dinamize a análise de grande volume de dados espectrais e climáticos em tempo hábil, principalmente quando se busca prever resultados.

Neste sentido, optou-se por aplicar técnicas de mineração de dados. Atenção especial foi dada à etapa de preparação dos dados, onde se buscou gerar atributos que levassem em consideração a fisiologia da cultura. Embora o banco de dados fosse considerado pequeno quanto ao número de instâncias para a aplicação destas técnicas, este trabalho procurou propor um método de tratamento dos dados, método este que apontou para resultados coerentes. É importante que se busque a aplicação destas técnicas em bancos de dados com maior número de instâncias e com diferentes abordagens quanto à preparação dos dados.

O uso dos dados meteorológicos na etapa de seleção de atributos não pareceu ter sido boa idéia, uma vez que não houve seleção de nenhum destes atributos. Sugere-se que se analisem os dados espectrais e meteorológicos de forma separada, considerando primeiramente apenas os dados espectrais para, num segundo momento usar os dados meteorológicos, e não de forma conjunta como foi feita neste trabalho.

Outra sugestão está relacionada à determinação das fases fisiológicas. Pode-se fixar a duração de cada fase ajustar o início da primeira fase a partir da ocorrência do menor valor do NDVI, e não em data fixa (primeiro decêndio de agosto) como foi feito neste trabalho.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

ALLEN, W.A.; GAUSMAN, H.W.; RICHARDSON, A.J.; THOMAS, J.R. Interaction of isotropic light with a compact plant leaf. **Journal of the Optical Society of America**, v.59 n.10, p.1376-1390, 1969.

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.

BAJWA, S.G.; BAJCSY, P.; GROVES, P.; TIAN, L.F. Hiperspectral image data mining for band selection in agricultural applications. **Transactions of the ASAE**, v.47, n.3, p. 895-907, 2004.

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

BERLATO, M.A.; FONTANA, D.C.; GONÇALVES, H.M. Relação entre o rendimento de grãos de soja e variáveis meteorológicas. **Pesquisa Agropecuária Brasileira**, Brasília, v.27, n.5, p.695-702, 1992.

BLAIN, G.C.; ROLIM, G.S.; CAMARGO, M.B.P.; BRUNINI, O. Adjustment of meteorological observations of MARS-FOOD system in relation of surface data for soybean yield estimation in the state of São Paulo, Brazil. In: 2º International Workshop on Crop Monitoring and Forecasting in South America, 2006. **Proceedings...** Montevideo: South America Scientific Network on Crop Monitoring and Forecasting, 2006. CD-ROM

BOKEN, V.K. and SHAYEWICH, C.F. Improving an operational wheat yield model using phenological phase-based Normalized Difference Vegetation Index. **International Journal of Remote Sensing**, v.23, n.20, p.4155-4168, 2002.

CAMPBELL, James. B. **Introduction to Remote Sensing**. 3rd ed. New York: The Guilford Press, 2002.

CANASAT. Mapeamento da cana via imagens de satélite e observação da Terra. <<http://www.dsr.inpe.br/mapdsr/intro.htm>>. 13 nov. 2007.

CASAGRANDE, A.A. **Tópicos de morfologia e fisiologia da cana-de-açúcar**. Jaboticabal: Funep, Unesp, 1991.

CASTRO, P.R.C. Aplicações da fisiologia vegetal no sistema de produção da cana-de-açúcar. In: Simpósio Internacional de Fisiologia da Cana-de-Açúcar. STAB, Piracicaba, 2000.

CEPAGRI. Centro de Pesquisas Meteorológicas e Climáticas Aplicadas a Agricultura. <<http://www.cpa.unicamp.br/>>. 12 jan. 2008.

CONAB. Companhia Nacional de Abastecimento. **Acompanhamento da Safra Brasileira Cana-de-Açúcar Safra 2007/2008, primeiro levantamento, maio/2007**. Brasília: Conab, 2007a, 12p.

CONAB. Companhia Nacional de Abastecimento. **Acompanhamento da Safra Brasileira Cana-de-Açúcar Safra 2007/2008: segundo levantamento, agosto/2007**. Brasília: Conab, 2007b, 12p.

CONAB. Companhia Nacional de Abastecimento. **Acompanhamento da Safra Brasileira Cana-de-Açúcar Safra 2007/2008: terceiro levantamento, novembro/2007**. Brasília: Conab, 2007c, 13p.

CONAB. Companhia Nacional de Abastecimento. **Acompanhamento da Safra Brasileira Cana-de-Açúcar Safra 2008/2009: terceiro levantamento, dezembro/2008**. Brasília: Conab, 2008, 18p.

COPERSUCAR. Cooperativa dos Produtores de Cana, Açúcar e Álcool do Estado de São Paulo. **Recomendação de adubação para a cultura da cana-de-açúcar**. Piracicaba/SP. Centro de Tecnologia Copersucar, 1988. 7p.

CRISLER, M.T.; STRICKLAND, R.M.; ESS, D.R.; PARSONS, S.D. Data mining methods for use with geo-referenced field crop data. In: World Congress of Computers in Agriculture and Natural Resources, 2006. **Proceedings...** Iguazu Falls: 2002. pp. 265-271.

DELGADO-ROJAS, J.S. e BARBIERI, V. Modelo agrometeorológico de estimativa da produtividade da cana-de-açúcar. **Revista Brasileira de Agrometeorologia**, v.7, n.1, p.67-73, 1999.

DEPPE, F.; MARTINI, L.; LONHMANN, M.; ADAMI, M. Validation studies of ECMWF precipitation data with observed SIMEPAR ground data (meteorological stations). In: 2º International Workshop on Crop Monitoring and Forecasting in South America, 2006. **Proceedings...** Montevideo: South America Scientific Network on Crop Monitoring and Forecasting, 2006. CD-ROM

DILLEWIJN, C.V. **Botany of sugarcane**. Waltham: The Chronica Botanica, 1952.

DORAISWAMY, P.C.; MOULIN, S.; COOK, P.W.; STERN, A. Crop Yield Assessment from Remote Sensing. **Photogrammetric Engineering & Remote Sensing**, v.69, n.6, p.665-674, 2003.

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

FAOSTAT. Food and Agriculture Organization of the United Nations - FAO: <<http://faostat.fao.org/>>, 20 jan 2009.

FAUCONNIER, R. y BASSEREAU, D. **La caña de azúcar, técnicas agrícolas y producciones tropicales**. Barcelona: Editorial Blume, 1975.

FAYYAD, U. M.; PIATESKY-SHAPIO, G.; SMITH, P. **Advances in Knowledge Discovery and Data Mining**. Massachusetts: MIT Press, 1996. 560 p.

FAYYAD, U. Data Mining and Knowledge Discovery in Databases: Implications for Scientific Databases. In: International Conference on 94Th Scientific and Statistical Database Management, 9th, Olympia/Washington/USA, 1997. **Proceedings...** Olympia: IEEE Computer Society, 1997. pp. 2-11.

FERENCZ, C. et al. Crop yield estimation by remote sensing. **International Journal of Remote Sensing**, v.25, n.20, p.4113-4149, 2004.

FONTANA, D.C.; BERLATO, M.A.; LAUSCHNER, M.H.; MELO, R.W. Modelo de estimativa de rendimento de soja no estado do Rio Grande do Sul, **Pesquisa Agropecuária Brasileira**, v.26, n.3, p.399-403, 2001.

FONTANA, D.C.; WEBER, E.; DUCATI, J.; FIGUEIREDO, D.C.; BERLATO, M.A.; BERGAMASCHI, H. A case study for crop monitoring and harvest forecast in south Brazil. In: International Symposium on Remote Sensing of Environment, 28., 2000, Cape Town. **Proceedings...** ICRSE, 2000. p.91-94.

FORTES, C. **Discriminação varietal e estimativa de produtividade agroindustrial de cana-de-açúcar pelo sensor orbital ETM+/Landsat 7**. 131p. Dissertação (Mestrado em Solos e Nutrição de Plantas) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba. 2003.

FORTES, C. and DEMATTÊ, A.M. Discrimination of sugarcane varieties using Landsat 7 ETM+ spectral data. **International Journal of Remote Sensing**, v.27, n.7-8, p.1395-1412, 2006.

GARDNER, B.R.; NIELSEN, D.C.; SHOCK, C.C. Infrared thermometry and the crop water-stress index: 1. History, theory and base-lines. **Journal of Production Agriculture**, v.5, n.4, p.462-466, 1992.

GAUSMAN, H.W.; ALLEN, W.A.; MYERS, V.I.; CARDENAS, R. Reflectance and internal structure of cotton leaves *Gossypium hirsutum* L. **Agronomy Journal**, v.61, n.3, p.374, 1969.

GAUSMAN, H.W. and ALLEN, W.A. Optical parameters of leaves of 30 plant species. **Plant Physiology**, v.52, n.1, p. 57-62, 1973.

GAUSMAN, H.W.; GERBERMANN, A.H.; WIEGANG, C.L. Use of ERTS-1 data to detect chlorotic grain-sorghum. **Photogrammetric Engineering & Remote Sensing**, v.41, n.2, p.177-179, 1975.

GEOSAFRAS. Relatório Final: Tarefas e Produtos Desenvolvidos, Programa de Revitalização da Conab, Grupo UFRGS-FEPAGRO, Produto 9, 2006.

GEOSAFRAS. Projeto Geosafra, Companhia Nacional de Abastecimento. <<http://www.conab.gov.br/conabweb/index.php?PAG=81>>, 20 jan. 2008.

GUSSO, A.; FONSECA, E.L.; FONTANA, D.C. Monitoramento do desenvolvimento da cultura da soja no sul do Brasil usando dados SPOT Vegetation e NOAA/AVHRR. In: Simpósio Brasileiro de Sensoriamento Remoto, 13., 2007. Florianópolis. **Anais...** Florianópolis: INPE, 2007, p. 225-232.

GUYOT, G. **Signatures spectrales des surfaces naturelles**. Collection “Télédétection stellitaire”. Caen: Édition Paradigme, 1989.

HATFIELD, J.L. and PINTER JR, P.J. Remote sensing for crop protection. **Crop Protection**, v.12, n.6, p.403-413, 1993.

HUMBERT, R.P. **El cultivo de la caña de azúcar**. México: Compañía Editorial Continental, 1974.

IBGE. Instituto Brasileiro de Geografia e Estatística. Pesquisas Agropecuárias, Série Relatórios Metodológicos. – Rio de Janeiro, 2ª. ed., v.6, 2002.

IBGE. Instituto Brasileiro de Geografia e Estatística: Banco de Dados Agregados. Sistema IBGE De Recuperação Automática – SIDRA. <<http://www.sidra.ibge.gov.br/>>, 29 jun. 2008.

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.

JACKSON, R.D.; PINTER JR, P.J.; REGINATO, R.J.; IDSO, S.B. Detection and Evaluation of Plant Stresses for Crop Management Decisions. **IEEE Transactions on Geoscience and Remote Sensing**, v.GE-24, n.1, 1986.

JRC. Joint Research Centre. Dados meteorológicos simulados pelo modelo do ECMWF. <<http://agrifish.jrc.it/marsfood/>>, 25 abr 2007.

KARIMI, Y.; PRASHER, S.O.; MCNAIRN, H.; BONNELL, R.B.; DUTILLEUL, P.; GOEL, P.K. Classification accuracy of discriminant analysis, artificial neural networks and decision trees for weed and nitrogen stress detection in corn. **Transactions of the ASAE**, v.48, n.3, p. 1261-1268, 2005.

LABUS, M.P.; NIELSEN, G.A.; LAWRENCE, R.L.; ENGEL, R. Wheat yield estimates using multi-temporal NDVI satellite imagery. **International Journal of Remote Sensing**, v.23, n.20, p. 4169-4180, 2002.

LAMPARELLI, Rubens Augusto Camargo ; ROCHA, J. V. ; MACHADO, Hermógenes Moura ; ZULLO JUNIOR, Jurandir . Using NDVI as a support to sugar cane yield estimates. In: European Conference of Precision Agriculture, 2003, Berlin. **Proceedings...**Berlin: 2003. v. 1. p. 82-86.

LINTZ, J. e SIMONETT, D.S. **Remote Sensing of environment**. Readig, MA: Addison-Wesley, 1976.

LOHMANN, M.; MARTINI, L.; DEPPE, F. Monitoramento da evolução temporal da cultura da soja no estado do Paraná utilizando imagens do sensor MODIS. In Simpósio Brasileiro de Sensoriamento Remoto, 13., 2007. Florianópolis. **Anais...**Florianópolis: INPE, 2007, p.263-269.

LUCCHESI, A.A. Processos fisiológicos da cultura da cana-de-açúcar (*Saccharum spp*). Boletim Técnico ESALQ/SENA, n.7, p.50, 1995.

LUIZ, A.J.B.; EPIPHANIO, J.C.N. Amostragem por pontos em imagens de sensoriamento remoto para estimativa de área plantada por município. In: Simpósio Brasileiro de Sensoriamento Remoto (SBSR), 10., 2001, Foz do Iguaçu. **Anais...** São José dos Campos: INPE, 2001, p.111-118. CD-ROM.

MACHADO, H.M. Determinação da biomassa de cana-de-açúcar considerando a variação espacial de dados espectrais do satélite Landsat 7 – ETM+. 61p. Dissertação (Mestrado em Engenharia Agrícola) - Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas - SP. 2003.

MACHADO, E.C.; PEREIRA, A.R.; CAMARGO, M.B.P.; FAHL, J.I. Relações radiométricas de uma cultura de cana-de-açúcar. **Bragantia**, v.44, n.1, p.229-238, 1985.

MASELLI, F. and REMBOLD, F. Analysis of GAC NDVI Data for Cropland Identification and Yield Forecasting in Mediterranean African Countries. **Photogrammetric Engineering & Remote Sensing**, v. 67,n.5, p.593-602, 2001.

MATZENAUER, R.; BERGAMASCHI, H.; BERLATO, M. A.; RIBOLDI, J. Modelos agrometeorológicos para estimativa do rendimento de milho, em função da disponibilidade hídrica no Estado do Rio Grande do Sul. **Pesquisa Agropecuária Gaúcha**, Porto Alegre, v. 1, n. 2, p. 225-241, 1995.

MELO, R.W. e FONTANA, D.C. Avaliação de dados meteorológicos disponíveis no *website* do *Joint Research Centre*. In: 2º International Workshop on Crop Monitoring and Forecasting in South America, 2006. **Proceedings...** Montevideo: South America Scientific Network on Crop Monitoring and Forecasting, 2006. CD-ROM

MELO, R.W. e FONTANA, D.C. Estimativa do rendimento de soja usando dados do modelo do ECMWF em um modelo agrometeorológico-espectral no estado do Rio Grande do Sul. In: Simpósio Brasileiro de Sensoriamento Remoto, 13., 2007. Florianópolis. **Anais...** Florianópolis: INPE, 2007, p.279-286.

MILLARD, J.P.; JACKSON, R.D.; GOETTELMAN, R.C.; REGINATO, R.J.; IDSO, S.B. Crop water-stress assessment using an airborne thermal scanner. **Photogrammetric Engineering & Remote Sensing**, v.44, p.77-85, 1978.

MOREIRA, Mauricio. A. **Fundamentos do Sensoriamento Remoto e Metodologias de Aplicação**. 1ª ed. São José dos Campos: INPE, 2001.

MUCHOW, R.C.; SPILLMAN, M.F.; WOOD, A.W.; THOMAS, M.R. Radiation Interception and Biomass Accumulation in a Sugarcane Crop Grown under Irrigated Tropical Conditions. **Australian Journal of Agricultural Research**, v.45, p.37-49, 1994.

O ESTADO DE SÃO PAULO. Governo lança projeto de previsão de safra por satélite. <<http://www.estadao.com.br/arquivo/economia/2004/not20040513p19928.htm>>. 25 ago. 2007.

OLIVEIRA, M. e VASCONCELOS, Y. Revolução no canavial. **Pesquisa FAPESP**, n.122, p.62-70, 2006.

PELLEGRINO, G.Q. Utilização de dados espectrais do satélite NOAA14/AVHRR como fonte de dados para modelos matemáticos de estimativa da fitomassa da cana-de-açúcar. 116p. Tese (Doutorado em Engenharia Agrícola) - Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas - SP. 2001.

PEREIRA, Antônio .R.; ANGELOCCI, Luiz R.; SENTELHAS, Paulo C. **Agrometeorologia: Fundamentos e Aplicações Práticas**. Guaíba/RS: Agropecuária, 2002.

PICOLI, M.C.A. Estimativa da produtividade agrícola da cana-de-açúcar utilizando agregados de redes neurais artificiais: estudo de caso Usina Catanduva. 95p. Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos – SP. 2006.

PICOLI, M.C.A.; RUDORFF, B.F.T.; RIZZI, R.; VON ZUBEN, F.J. Estimativa da produtividade agrícola da cana-de-açúcar: estudo de caso Usina Catanduva. In: Simpósio Brasileiro de Sensoriamento Remoto, 13., 2007. Florianópolis. **Anais...**Florianópolis: INPE, 2007, p. 331-333.

PINTER JR, P.J.; HATFIELD, J.L.; SCHEPERS, J.S.; BARNES, E.M.; MORAN, M.S.; DAUGHTRY, C.S.T.; UPCHURCH, D.R. Remote Sensing for Crop Management. **Photogrammetric Engineering & Remote Sensing**, v.69, n.6, p.647-664, 2003.

PINTER JR, P.J. and REGINATO, R.J. A thermal infrared technique for monitoring cotton water-stress and scheduling irrigations. **Transactions of ASAE**, v.25, n.6, p.1651-1655, 1982.

PLANALSUCAR. Nutrição e adubação da cana-de-açúcar no Brasil. Coleção IAA/Planalsucar, n.2, Coord: Orlando F.J., Piracicaba-SP, 1983, 369p.

PMGCA. Programa de Melhoramento Genético da Cana-de-Açúcar.UFSCar. <<http://pmgca.dbv.cca.ufscar.br/index.php>>, 22 jan. 2008.

PONTES, P. P. B; ROCHA, J. V.; LAMPARELLI, R. A. C. Análise temporal de índices de vegetação como subsídio à previsão de safras de cana-de-açúcar. In: Simpósio Brasileiro de Sensoriamento Remoto, 12., 2005. Goiânia. **Anais...**Goiânia: INPE, 2005, p. 217-224.

PRADO, T. B. G.; PELLEGRINO, G. Q.; ZULLO JÚNIOR, J.; LAMPARELLI, R. A. C. Avaliação de modelo matemático-espectral para estimativa da safra de cana-de-açúcar. In: Simpósio Brasileiro de Sensoriamento Remoto, 13., 2007. Florianópolis. **Anais...**Florianópolis: INPE, 2007, p. 351-353.

RIZZI, R.; RUDORFF, B. F. T. Imagens Landsat na estimativa da área plantada com soja em municípios do Rio Grande do Sul. In: Simpósio Brasileiro de Sensoriamento Remoto, 11., 2003. Belo Horizonte. **Anais...**Belo Horizonte: INPE, 2003, p. 231-238.

RIZZI, R. e RUDORFF, B.F.T. Imagens do sensor MODIS associadas a um modelo agrônômico para estimar a produtividade de soja. **Pesquisa Agropecuária Brasileira**, v.42, n.1, p.73-80, 2007.

ROCHA, J.V. **Crop monitoring in South America (MERCOSUR + Bolivia)**. Bulletin no. 7/2006 (August, 2006). Ispra, Italy: JRC (Joint Research Centre), 2006.

RODRIGUES, J.D. Fisiologia da cana-de-açúcar. Unesp Botucatu, Instituto de Biociências, Botucatu, 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.

ROSENFELD, U. e LEME, F.J.A. Produtividade da cana-de-açúcar irrigada por aspersão. Estudo de épocas de irrigação. In: Congresso Nacional da STAB, 3., São Paulo, 1984. **Anais...**, São Paulo: STAB, 1984, p.18-24.

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., Washington, 1973. **Proceedings...** Whashington: NASA, 1974, v.1, p.309-317, 1973.

RUDORFF, B.F.T. et al. Estimativa de área plantada com cana-de-açúcar em municípios do Estado de São Paulo por meio de imagens de satélites e técnicas de geoprocessamento: ano safra 2003/2004, INPE, São José dos Campos, 2004.

RUDORFF, B.F.T. et al. Imagens de satélite no mapeamento e estimativa de área de cana-de-açúcar em São Paulo: ano-safra 2003/04. **Agricultura em São Paulo**, v.52, n.1, p.21-39, 2005.

RUDORFF, B.F.T. and BATISTA, G.T. Yield Estimation of Sugarcane Based on Agrometeorological-Spectral Models. **Remote Sensing of Environment**, v.33, p.183-192, 1990.

RUDORFF, Bernardo Friedrich Theodor. Dados LANDSAT na estimativa da produtividade agrícola da cana-de-açúcar. 114p. Dissertação (Mestrado em Sensoriamento Remoto e Aplicações) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 1985.

SCHEPERS, J.S.; BLACKMER, T.M.; WILHELM, W.W.; RESENDE, M. Transmittance and reflectance measurements of corn leaves from plants with different nitrogen and water supply. **Journal of Plant Physiology**, v.148, n.5, p.523-529, 1996.

SHOCK, B.M.; CARPENTER, G.A.; GOPAL, S.; WOODCOCK, C.E. ARTMAP neural network classification of land use change. In: World Congress of Computers in Agriculture and Natural Resources, 2006. **Proceedings...** Iguazu Falls: 2002. pp. 22-28.

SIMÕES, M.S.; ROCHA, J.V.; LAMPARELLI, R.A.C. Growth indices and productivity in sugarcane. **Scientia Agricola**, v.62, n.1, p.23-30, 2005a.

SIMÕES, M.S.; ROCHA, J.V.; LAMPARELLI, R.A.C. Spectral variables growth analysis and yield of sugarcane. **Scientia Agricola**, v.62, n.3, p.199-207, 2005b.

SUGAWARA, L.M.; RUDORFF, B.F.T.; FREITAS, C.C.; PICOLI, M.C.; ADAMI, M. Estimativa de produtividade de cana-de-açúcar (*Saccharum officinarum* L.) por meio de técnica de análise de regressão linear múltipla. In: Simpósio Brasileiro de Sensoriamento Remoto, 13., 2007. Florianópolis. **Anais...** Florianópolis: INPE, 2007, p.435-442.

VEGETATION. Vegetation Programme.

<<http://spot-vegetation.com/vegetationprogramme/index.htm>> , 22 mar. 2007

VITO. What is the relation between the digital number and the real NDVI? <<http://www.vgt.vito.be/faq/faq.html>>. 23 mai. 2007.

XAVIER, A.C.; RUDORFF, B.F.T.; SHIMABUKURO, Y.E.; BERKA, L.M.S.; MOREIRA, M.A. Multi-temporal analysis of MODIS data to classify sugarcane crop. **International Journal of Remote Sensing**, v.27, n.3-4, p.755-768, 2006.

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.

ZONEAMENTO DE CULTURAS BIOENERGÉTICAS NO ESTADO DE SÃO PAULO: Aptidão Edafoclimática da Cultura da Cana-de-Açúcar. Governo do Estado de São Paulo, 2008. CD-ROM.

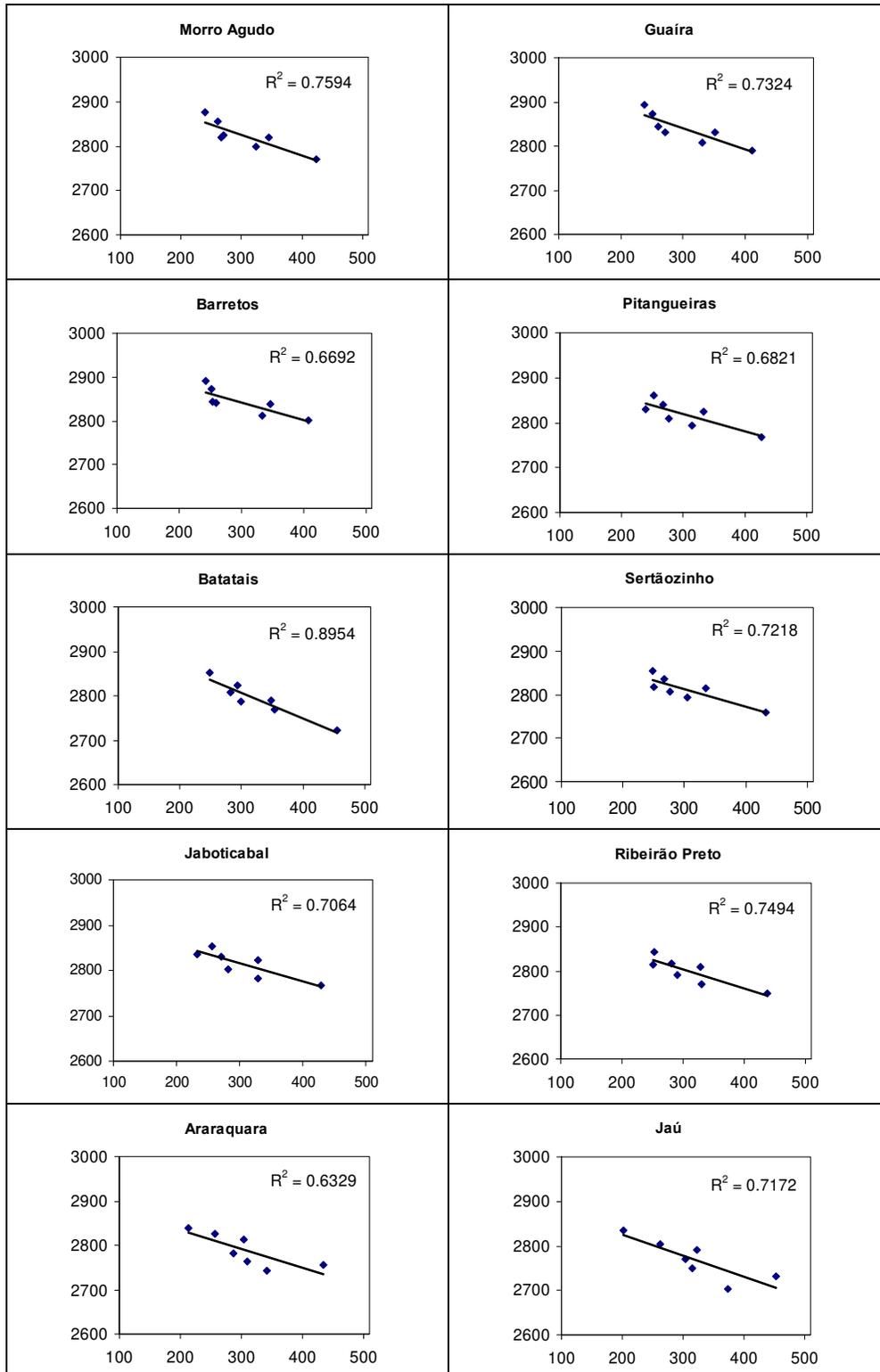
## **ANEXOS**

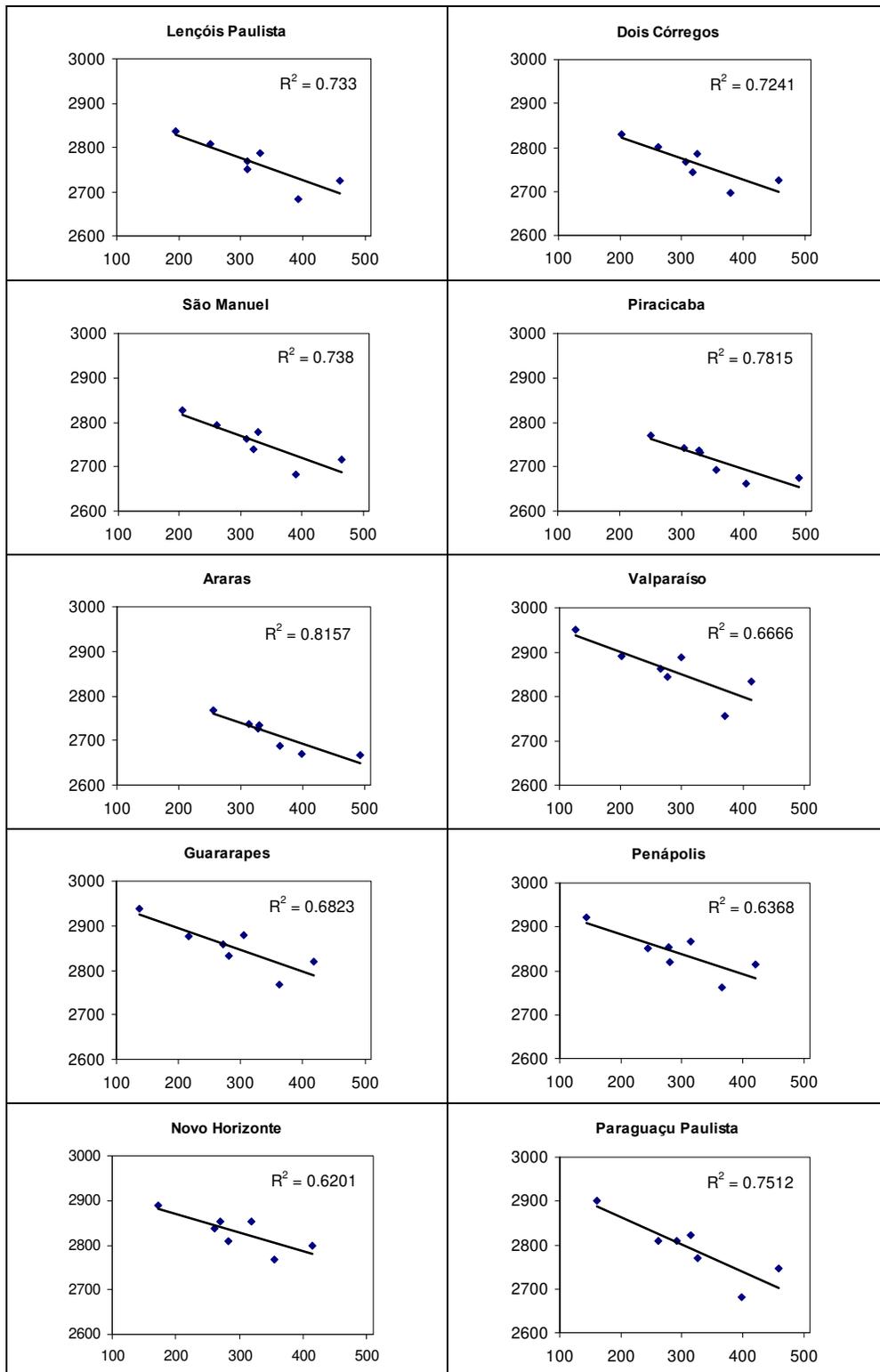
Nº teste	1999 / 2000		2000 / 2001		2001 / 2002		2002 / 2003		2003 / 2004		2004 / 2005		2005 / 2006	
	R <sup>2</sup>	% dif.	R <sup>2</sup>	% dif.	R <sup>2</sup>	% dif.	R <sup>2</sup>	% dif.	R <sup>2</sup>	% dif.	R <sup>2</sup>	% dif.	R <sup>2</sup>	% dif.
1	0.490	177.4	0.728	56.6	0.745	11.5	0.698	97.6	0.791	15.4	<b>0.757</b>	<b>-4.6</b>	0.655	135.6
2	0.431	244.8	0.744	66.7	0.754	18.5	0.711	104.8	0.788	22.4	0.743	-1.1	0.613	157.0
3	0.627	101.2	0.801	21.4	0.731	-14.7	0.773	45.6	0.813	-11.2	0.799	-23.6	0.737	89.4
4	0.555	153.7	0.817	29.4	0.740	-9.9	0.780	50.8	<b>0.811</b>	<b>-6.6</b>	0.802	-21.1	0.706	105.3
5	0.706	33.3	0.822	-11.8	0.693	-42.5	0.795	-2.9	0.789	-38.2	0.789	-43.2	0.790	39.1
6	0.655	68.3	0.818	-9.8	0.701	-39.5	<b>0.798</b>	<b>0.5</b>	0.785	-35.3	0.790	-41.6	0.775	49.8
7	0.541	131.7	0.725	27.8	0.739	-4.5	0.730	68.0	0.769	-10.1	0.724	-24.1	0.699	99.7
8	0.465	192.2	0.740	36.1	<b>0.744</b>	<b>1.5</b>	0.739	74.5	0.762	-4.2	0.728	-21.3	0.658	118.4
9	0.671	73.4	0.791	3.2	0.722	-23.9	0.802	26.9	0.792	-28.7	0.764	-37.6	0.774	64.6
10	0.586	121.1	0.806	9.7	0.728	-19.7	0.806	31.6	0.787	-24.6	0.767	-35.6	0.743	78.7
11	0.742	17.6	0.812	-22.1	0.684	-46.6	0.810	-12.7	0.770	-48.3	0.765	-52.5	0.815	24.1
12	0.687	49.7	<b>0.829</b>	<b>-17.1</b>	0.691	-43.9	0.811	-11.7	0.764	-45.7	0.766	-51.2	0.800	33.9
13	0.610	82.3	0.716	1.8	0.712	-20.9	0.744	39.5	0.719	-31.9	0.676	-42.4	0.739	61.2
14	0.518	133.3	0.729	8.5	0.715	-15.9	0.749	45.1	0.708	-27.0	0.680	-40.3	0.702	76.6
15	0.711	43.2	0.766	-13.6	0.698	-34.2	0.808	8.8	0.750	-44.4	0.714	-51.3	0.798	37.4
16	0.618	84.9	0.781	-8.2	0.703	-30.6	0.810	13.0	0.742	-40.9	0.716	-49.6	0.770	49.5
17	0.764	1.3	0.786	-32.0	0.667	-51.5	0.808	-22.3	0.735	-58.2	0.725	-61.8	<b>0.824</b>	<b>7.8</b>
18	0.705	29.9	0.804	-27.8	0.673	-49.1	0.808	-19.5	0.727	-55.9	0.726	-60.7	0.811	16.5
19	0.686	35.5	0.699	-21.9	0.671	-35.7	0.739	12.2	0.659	-50.6	0.627	-57.9	0.753	25.2
20	0.587	75.2	0.707	-16.6	0.671	-31.6	0.741	16.9	0.644	-46.6	0.628	-56.3	0.723	37.1
21	0.746	12.9	0.729	-30.7	0.663	-44.4	0.793	-8.5	0.697	-58.5	0.659	-63.3	0.795	10.6
22	0.655	46.7	0.741	-26.4	0.667	-41.3	0.793	-4.9	0.687	-55.6	0.660	-62.0	0.772	20.4
23	<b>0.771</b>	<b>-15.3</b>	0.746	-43.1	0.639	-57.3	0.792	-31.9	0.691	-67.9	0.671	-70.2	0.809	-9.5
24	0.717	8.9	0.762	-39.7	0.644	-55.1	0.791	-29.4	0.682	-65.9	0.672	-69.4	0.799	-2.1

**Anexo I**  
 Resultados de R<sup>2</sup> e diferença entre área estimada e oficial, para os 24 testes propostos em cada safra.

## ANEXO 2

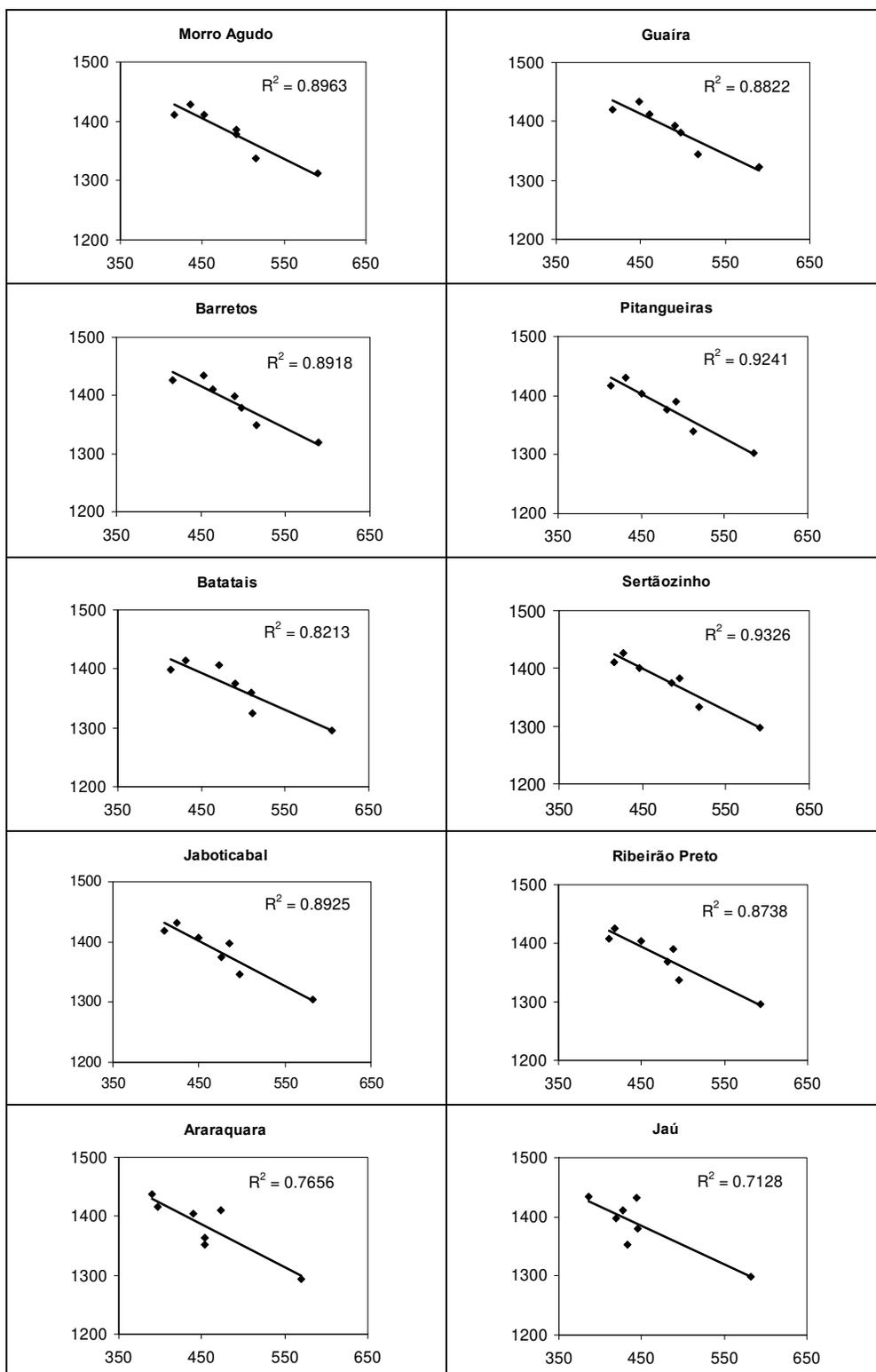
Correlação entre precipitação acumulada na fase 1 (rain1\_s: eixo X, em mm) e radiação global acumulada na fase 1 (grs1\_s: eixo Y, em kwh/m<sup>2</sup>). Gráficos de dispersão e índices de determinação (R<sup>2</sup>) separados por município.

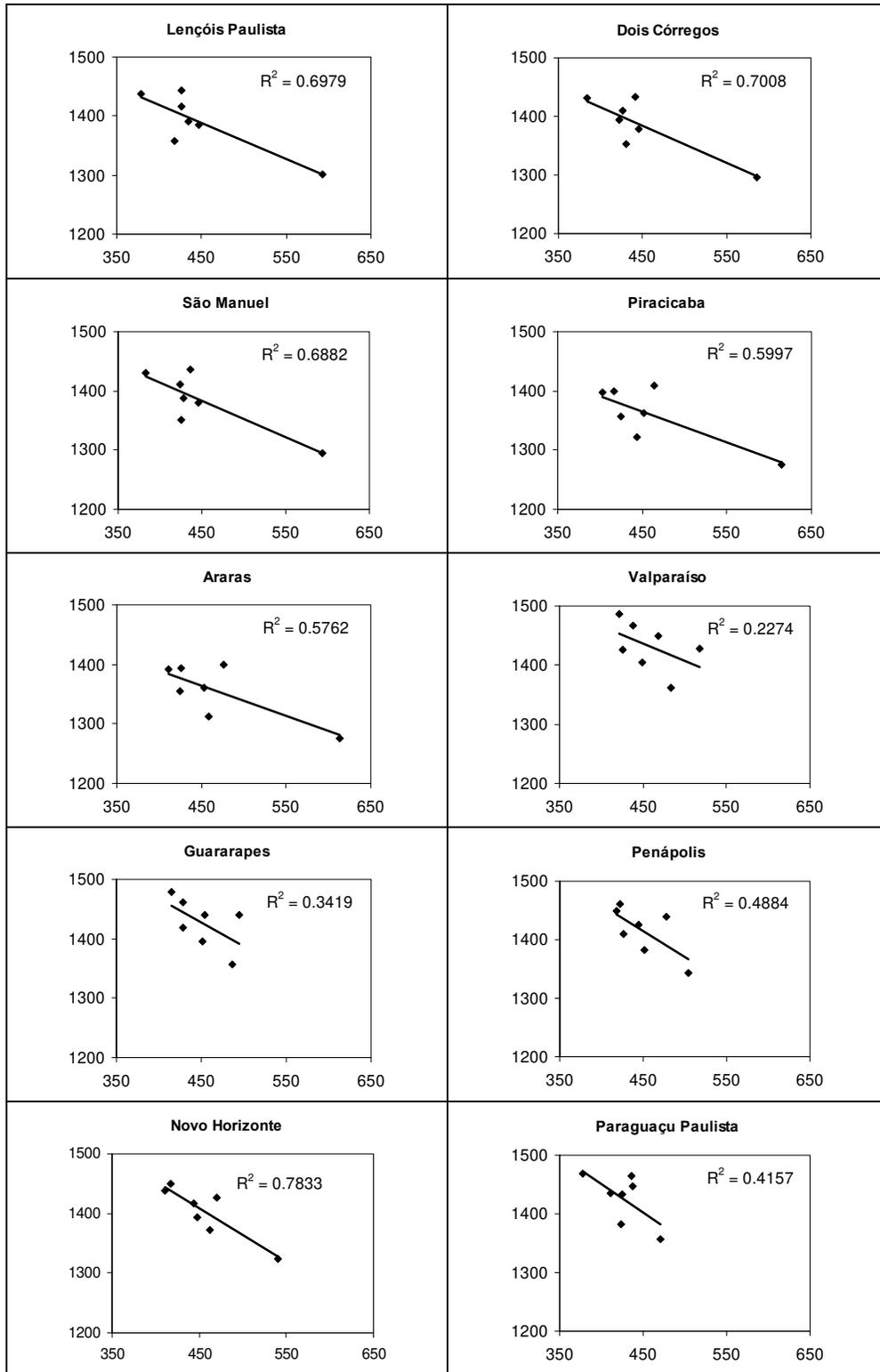




### ANEXO 3

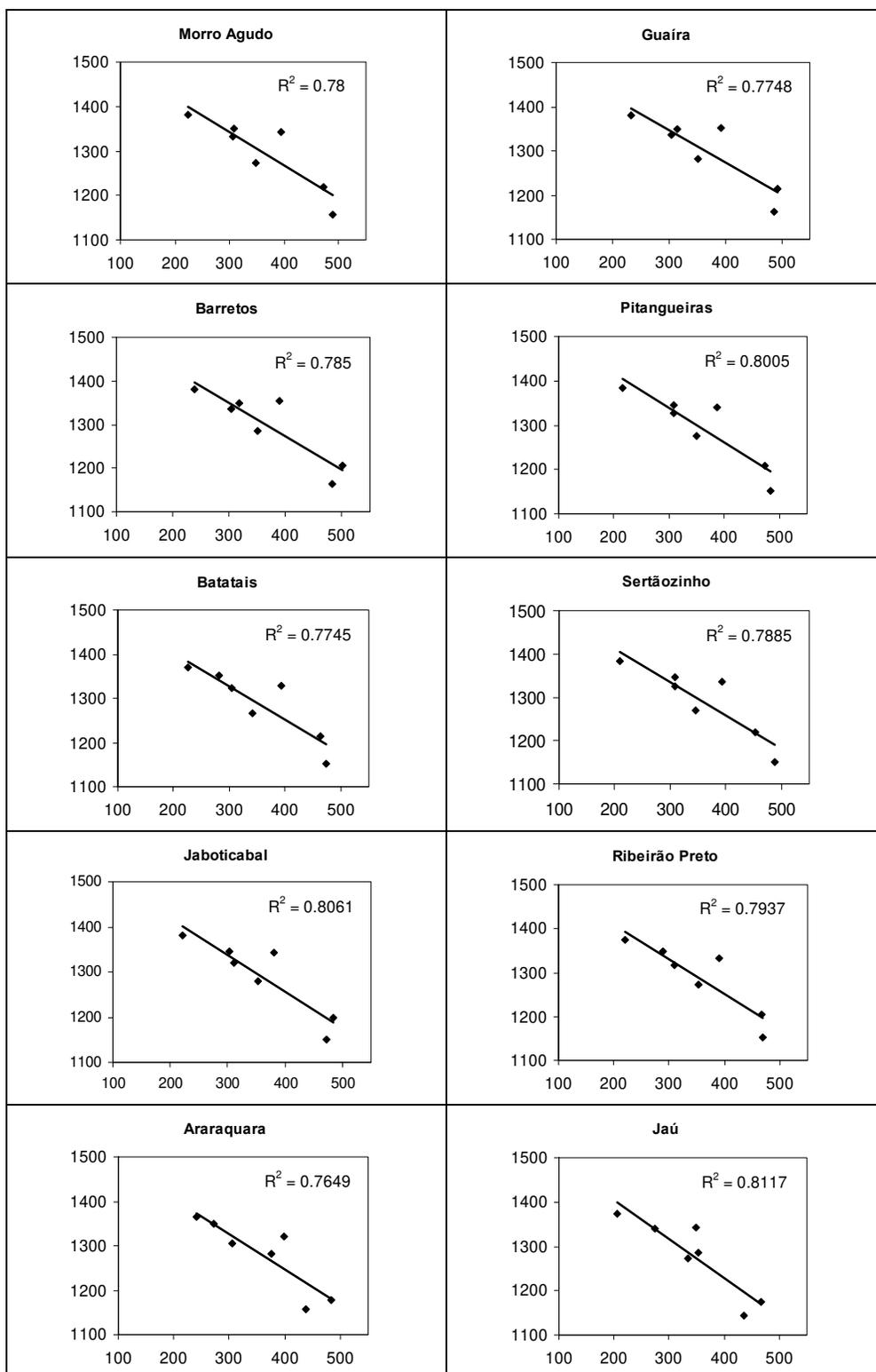
Correlação entre precipitação acumulada na fase 2 (rain2\_s: eixo X, em mm) e radiação global acumulada na fase 2 (grs2\_s: eixo Y, em kwh/m<sup>2</sup>). Gráficos de dispersão e índices de determinação (R<sup>2</sup>) separados por município.

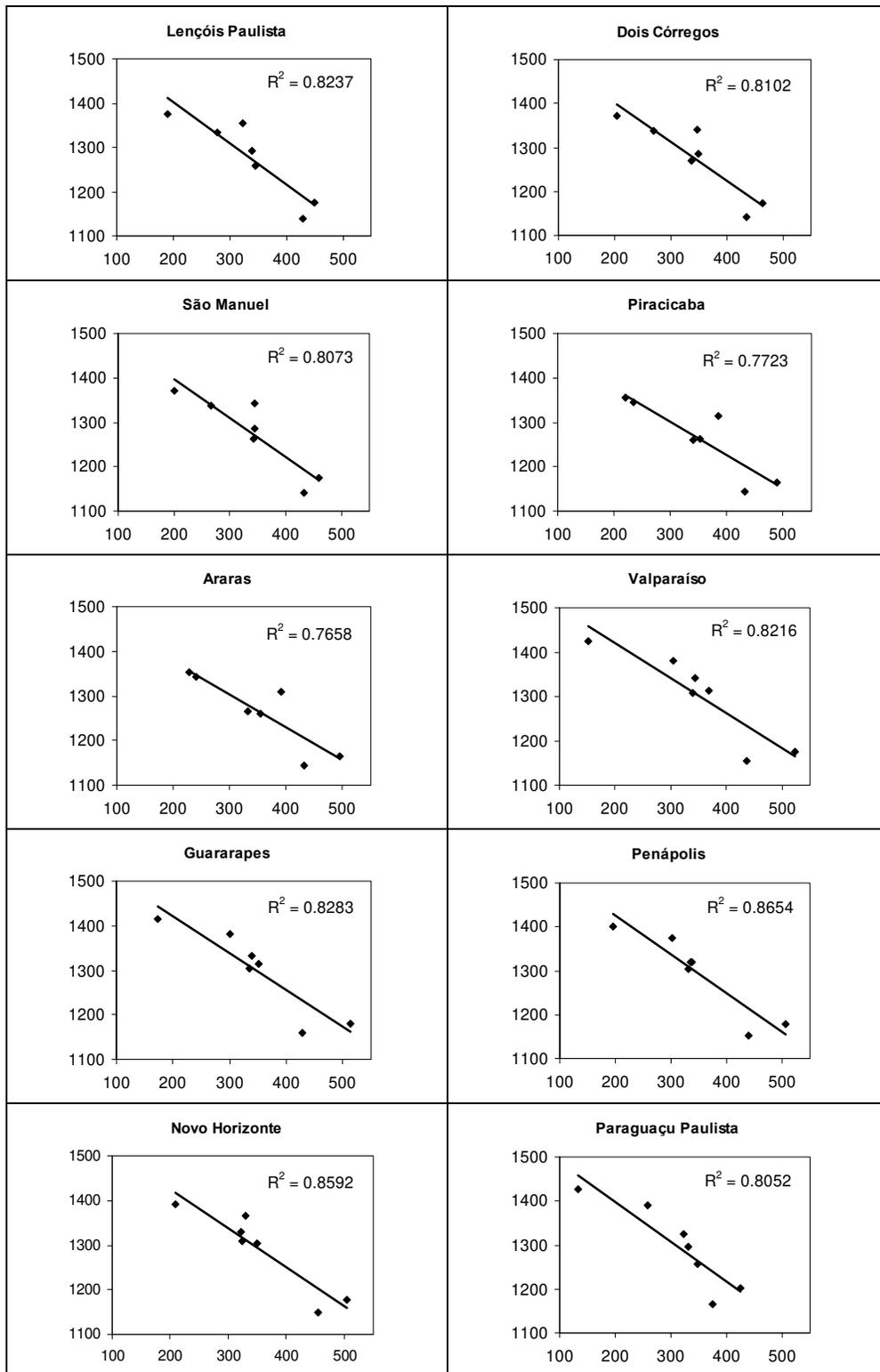




## ANEXO 4

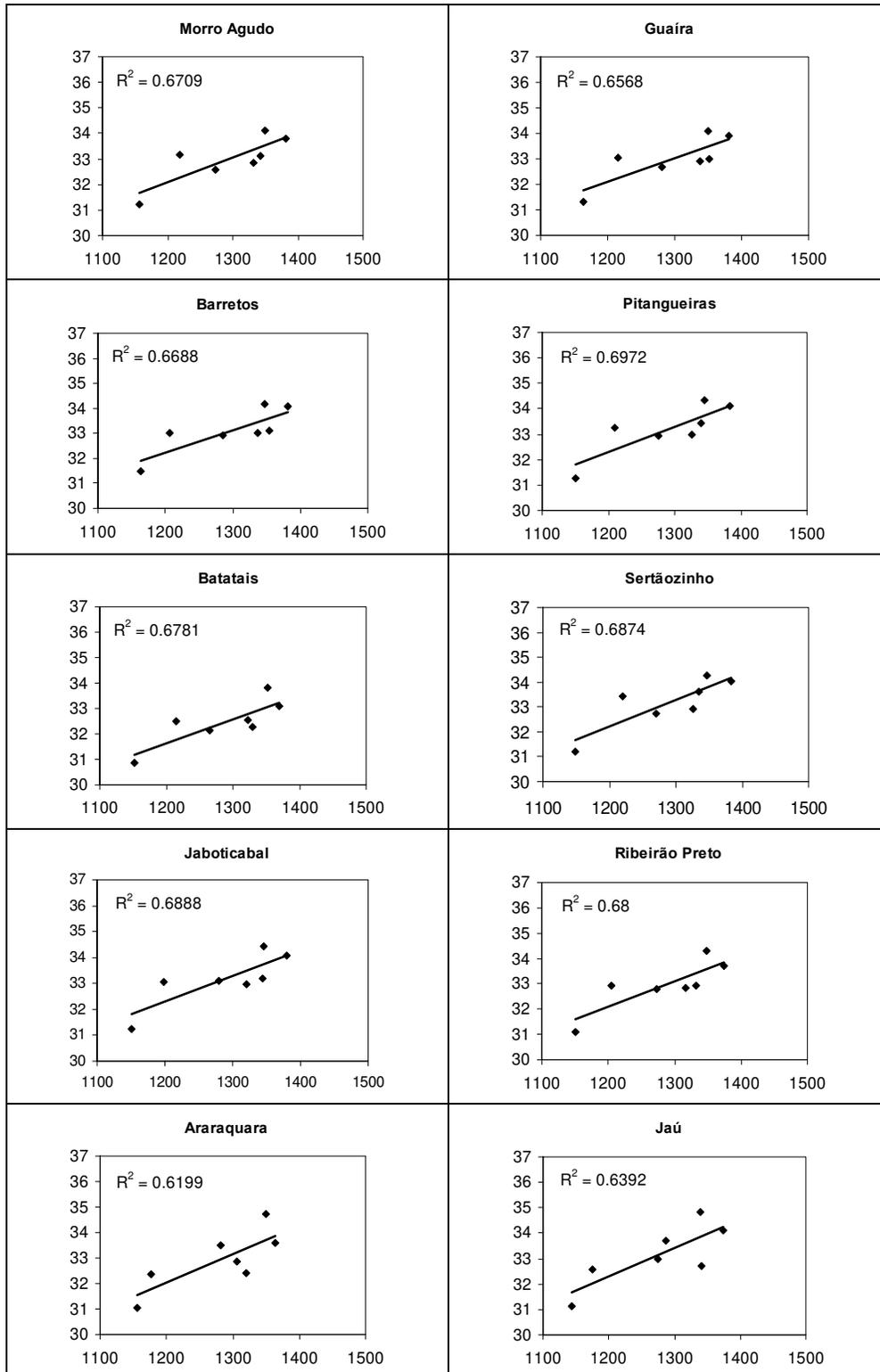
Correlação entre precipitação acumulada na fase 3 (rain3\_s: eixo X, em mm) e radiação global acumulada na fase 3 (grs3\_s: eixo Y, em kwh/m<sup>2</sup>). Gráficos de dispersão e índices de determinação (R<sup>2</sup>) separados por município.

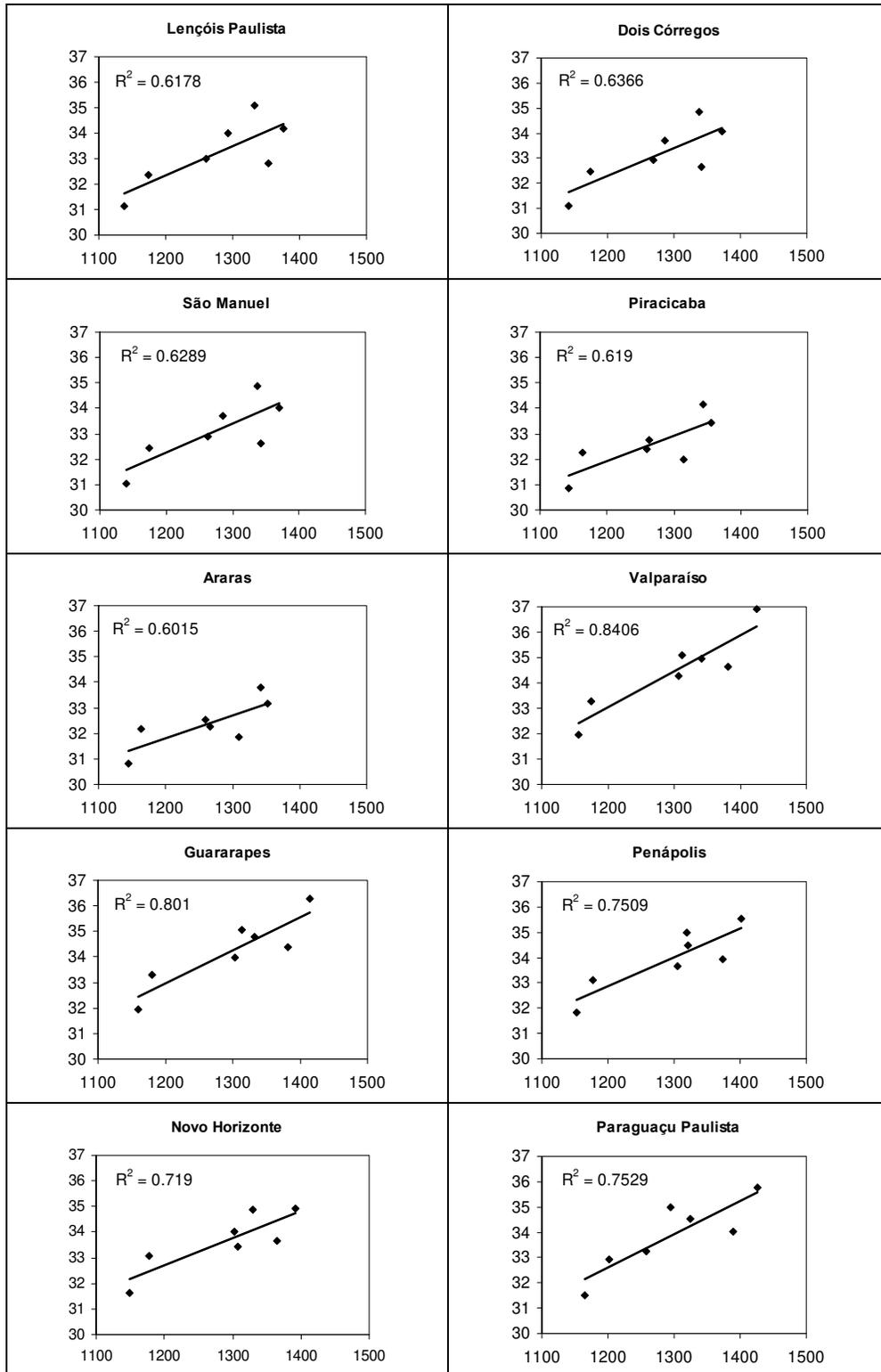




## ANEXO 5

Correlação entre radiação global acumulada na fase 3 (grs3\_s: eixo X, em kwh/m<sup>2</sup>) e média da temperatura máxima na fase 3 (tmax3\_m: eixo Y, em °C). Gráficos de dispersão e índices de determinação (R<sup>2</sup>) separados por município.





## ANEXO 6

Resultados da seleção de atributos retirados do Weka, para cada um dos métodos utilizados, considerando todos os atributos.

### 1) Qui-quadrado:

=== Attribute Selection on all input data ===

Search Method:

Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 54 rendimento):

Chi-squared Ranking Filter

Ranked attributes:

40.494 10 ndvi2\_final  
34.572 18 ndvi3\_inic  
30.544 17 ndvi3s  
22.78 40 ndvi\_m  
19.655 51 ndvi2\_m  
17.963 42 dec\_max  
0 16 tmax2m  
0 14 tmin2m  
0 23 grs3s  
0 15 tmean2m  
0 19 ndvi3\_final  
0 21 ndvi3\_d  
0 20 max-final3  
0 25 tmean3m  
0 24 tmin3m  
0 22 rain3s  
0 4 grs1\_s  
0 6 tmean1\_m  
0 5 tmin1\_m  
0 1 ndvi1\_s  
0 3 rain1\_s  
0 2 ndvi1\_final  
0 11 ndvi2\_d  
0 13 grs2s  
0 12 rain2s  
0 7 tmax1\_m  
0 9 ndvi2\_inic  
0 8 ndvi2s  
0 44 periodo\_mM  
0 46 grss  
0 45 rains  
0 39 dec\_min  
0 43 ndvi\_d  
0 41 ndvi\_max  
0 50 ndvi1\_m  
0 53 ndvi4\_m  
0 52 ndvi3\_m  
0 47 tminm  
0 49 tmaxm  
0 48 tmeanm

0 29 ndvi4\_final  
0 31 ndvi4\_d  
0 30 max-final4  
0 26 tmax3m  
0 28 ndvi4\_inic  
0 27 ndvi4s  
0 36 tmax4m  
0 38 ndvi\_min  
0 37 ndvis  
0 33 grs4s  
0 32 rain4s  
0 35 tmean4m  
0 34 tmin4m

Selected attributes:

10,18,17,40,51,42,16,14,23,15,19,21,20,25,24,22,4,6,5,1,3,2,11,13,12,7,9,8,44,46,45,39,43,41,50,53,52,47,49,48,  
29,31,30,26,28,27,36,38,37,33,32,35,34 : 53

## 2) WrapperSubsetEval (J48):

### Greedy Stepwise

=== Attribute Selection on all input data ===

Search Method:

Greedy Stepwise (forwards).

Start set: no attributes

Merit of best subset found: 0.352

Attribute Subset Evaluator (supervised, Class (nominal): 54 rendimento):

Wrapper Subset Evaluator

Learning scheme: weka.classifiers.trees.J48

Scheme options: -C 0.25 -M 2

Accuracy estimation: classification error

Number of folds for accuracy estimation: 5

Selected attributes: 2,18,41 : 3

ndvi1\_final

ndvi3\_inic

ndvi\_max

## 3) CFS

### Greedy Stepwise

=== Attribute Selection on all input data ===

Search Method:

Greedy Stepwise (forwards).

Start set: no attributes

Merit of best subset found: 0.211

Attribute Subset Evaluator (supervised, Class (nominal): 54 rendimento):

CFS Subset Evaluator

Including locally predictive attributes

Selected attributes: 10,17,18,40,51 : 5

ndvi2\_final  
ndvi3s  
ndvi3\_inic  
ndvi\_m  
ndvi2\_m

#### 4) InfoGain

=== Attribute Selection on all input data ===

Search Method:

Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 54 rendimiento):  
Information Gain Ranking Filter

Ranked attributes:

0.2017 10 ndvi2\_final  
0.1731 18 ndvi3\_inic  
0.1457 17 ndvi3s  
0.1119 40 ndvi\_m  
0.0991 51 ndvi2\_m  
0.0881 42 dec\_max  
0 16 tmax2m  
0 14 tmin2m  
0 23 grs3s  
0 15 tmean2m  
0 19 ndvi3\_final  
0 21 ndvi3\_d  
0 20 max-final3  
0 25 tmean3m  
0 24 tmin3m  
0 22 rain3s  
0 4 grs1\_s  
0 6 tmean1\_m  
0 5 tmin1\_m  
0 1 ndvi1\_s  
0 3 rain1\_s  
0 2 ndvi1\_final  
0 11 ndvi2\_d  
0 13 grs2s  
0 12 rain2s  
0 7 tmax1\_m  
0 9 ndvi2\_inic  
0 8 ndvi2s  
0 44 periodo\_mM  
0 46 grss  
0 45 rains  
0 39 dec\_min  
0 43 ndvi\_d  
0 41 ndvi\_max  
0 50 ndvi1\_m  
0 53 ndvi4\_m

```

0 52 ndvi3_m
0 47 tminm
0 49 tmaxm
0 48 tmeanm
0 29 ndvi4_final
0 31 ndvi4_d
0 30 max-final4
0 26 tmax3m
0 28 ndvi4_inic
0 27 ndvi4s
0 36 tmax4m
0 38 ndvi_min
0 37 ndvis
0 33 grs4s
0 32 rain4s
0 35 tmean4m
0 34 tmin4m

```

Selected attributes:

10,18,17,40,51,42,16,14,23,15,19,21,20,25,24,22,4,6,5,1,3,2,11,13,12,7,9,8,44,46,45,39,43,41,50,53,52,47,49,48,  
29,31,30,26,28,27,36,38,37,33,32,35,34 : 53

## 5) GainRatio

=== Attribute Selection on all input data ===

Search Method:

Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 54 rendimento):

Gain Ratio feature evaluator

Ranked attributes:

```

0.231 10 ndvi2_final
0.226 17 ndvi3s
0.193 18 ndvi3_inic
0.152 40 ndvi_m
0.11 51 ndvi2_m
0.103 42 dec_max
0 25 tmean3m
0 14 tmin2m
0 16 tmax2m
0 15 tmean2m
0 21 ndvi3_d
0 23 grs3s
0 22 rain3s
0 24 tmin3m
0 20 max-final3
0 19 ndvi3_final
0 4 grs1_s
0 6 tmean1_m
0 5 tmin1_m
0 1 ndvi1_s
0 3 rain1_s
0 2 ndvi1_final
0 11 ndvi2_d

```

0 13 grs2s  
 0 12 rain2s  
 0 7 tmax1\_m  
 0 9 ndvi2\_inic  
 0 8 ndvi2s  
 0 44 periodo\_mM  
 0 46 grss  
 0 45 rains  
 0 39 dec\_min  
 0 43 ndvi\_d  
 0 41 ndvi\_max  
 0 50 ndvi1\_m  
 0 53 ndvi4\_m  
 0 52 ndvi3\_m  
 0 47 tminm  
 0 49 tmaxm  
 0 48 tmeanm  
 0 29 ndvi4\_final  
 0 31 ndvi4\_d  
 0 30 max-final4  
 0 26 tmax3m  
 0 28 ndvi4\_inic  
 0 27 ndvi4s  
 0 36 tmax4m  
 0 38 ndvi\_min  
 0 37 ndvis  
 0 33 grs4s  
 0 32 rain4s  
 0 35 tmean4m  
 0 34 tmin4m

Selected attributes:

10,17,18,40,51,42,25,14,16,15,21,23,22,24,20,19,4,6,5,1,3,2,11,13,12,7,9,8,44,46,45,39,43,41,50,53,52,47,49,48,  
 29,31,30,26,28,27,36,38,37,33,32,35,34 : 53

## ANEXO 7

Classificação utilizando os atributos selecionados ndvi2\_final, ndvi2\_m, ndvi3s, ndvi3\_inic e ndvi\_m.

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 10  
Relation: S5\_t0904-weka.filters.unsupervised.attribute.Remove-R1-2-  
weka.filters.unsupervised.attribute.Remove-R1-9,11-16,19-39,41-50,52-53  
Instances: 140  
Attributes: 6  
    ndvi2\_final  
    ndvi3s  
    ndvi3\_inic  
    ndvi\_m  
    ndvi2\_m  
    rendimento

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	98	70	%
Incorrectly Classified Instances	42	30	%
Kappa statistic	0.5039		
Mean absolute error	0.2881		
Root mean squared error	0.3795		
Relative absolute error	71.7111	%	
Root relative squared error	84.777	%	
Total Number of Instances	140		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.625	0.095	0.577	0.625	0.6	0.803	B-M
0.743	0.273	0.753	0.743	0.748	0.77	M
0.667	0.133	0.683	0.667	0.675	0.842	M-A

=== Confusion Matrix ===

a b c <-- classified as  
15 6 3 | a = B-M  
9 55 10 | b = M  
2 12 28 | c = M-A

## ANEXO 8

Classificação usando somente os atributos ndvi2\_final e ndvi2\_m.

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2  
Relation: S5\_t0904-weka.filters.unsupervised.attribute.Remove-R1-11,13-41,43-52,54-55-  
weka.filters.unsupervised.attribute.Remove-R2  
Instances: 140  
Attributes: 3  
    ndvi2\_final  
    ndvi2\_m  
    rendimento

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	94	67.1429 %
Incorrectly Classified Instances	46	32.8571 %
Kappa statistic	0.3954	
Mean absolute error	0.325	
Root mean squared error	0.4031	
Relative absolute error	80.9003 %	
Root relative squared error	90.0451 %	
Total Number of Instances	140	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.083	0	1	0.083	0.154	0.629	B-M
0.865	0.5	0.66	0.865	0.749	0.684	M
0.667	0.133	0.683	0.667	0.675	0.77	M-A

=== Confusion Matrix ===

```
a b c <-- classified as
2 19 3 | a = B-M
0 64 10 | b = M
0 14 28 | c = M-A
```