



Ministério da
Ciência e Tecnologia



INPE-16222-TDI/1543

**CLASSIFICAÇÃO ESPECTRO-TEMPORAL DE
IMAGENS ORBITAIS PARA O MAPEAMENTO DA
COLHEITA DA CANA-DE-AÇÚCAR COM QUEIMA DA
PALHA**

Márcio Pupin de Mello

Dissertação de Mestrado do Curso de Pós-Graduação em Sensoriamento Remoto,
orientada pelos Drs. Bernardo Friedrich Theodor Rudorff e Carlos Antonio Oliveira
Vieira, aprovada em 22 de maio de 2009.

Registro do documento original:

<<http://urlib.net/sid.inpe.br/mtc-m18@80/2009/05.12.17.08>>

INPE
São José dos Campos
2009

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3945-6911/6923

Fax: (012) 3945-6919

E-mail: pubtc@sid.inpe.br

CONSELHO DE EDITORAÇÃO:**Presidente:**

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

Membros:

Dr^a Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Haroldo Fraga de Campos Velho - Centro de Tecnologias Especiais (CTE)

Dr^a Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Dr. Ralf Gielow - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr. Wilson Yamaguti - Coordenação Engenharia e Tecnologia Espacial (ETE)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Jefferson Andrade Ancelmo - Serviço de Informação e Documentação (SID)

Simone A. Del-Ducca Barbedo - Serviço de Informação e Documentação (SID)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Marilúcia Santos Melo Cid - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

EDITORAÇÃO ELETRÔNICA:

Viveca Sant´Ana Lemos - Serviço de Informação e Documentação (SID)



Ministério da
Ciência e Tecnologia



INPE-16222-TDI/1543

**CLASSIFICAÇÃO ESPECTRO-TEMPORAL DE
IMAGENS ORBITAIS PARA O MAPEAMENTO DA
COLHEITA DA CANA-DE-AÇÚCAR COM QUEIMA DA
PALHA**

Márcio Pupin de Mello

Dissertação de Mestrado do Curso de Pós-Graduação em Sensoriamento Remoto,
orientada pelos Drs. Bernardo Friedrich Theodor Rudorff e Carlos Antonio Oliveira
Vieira, aprovada em 22 de maio de 2009.

Registro do documento original:

<<http://urlib.net/sid.inpe.br/mtc-m18@80/2009/05.12.17.08>>

INPE
São José dos Campos
2009

Dados Internacionais de Catalogação na Publicação (CIP)

M489c Mello, Márcio Pupin de.
Classificação espectro-temporal de imagens orbitais para o mapeamento da colheita da cana-de-açúcar com queima da palha / Márcio Pupin de Mello. – São José dos Campos : INPE, 2009.
130p. ; (INPE-16222-TDI/1543)

Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2009.

Orientadores : Drs. Bernardo Friedrich Theodor Rudorff e Carlos Antonio Oliveira Vieira.

1. Classificação automática. 2. Sensoriamento remoto. 3. Cana-de-açúcar. 4. Queima da palha. I. Título.

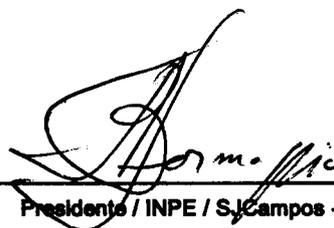
CDU 528.854

Copyright © 2009 do MCT/INPE. Nenhuma parte desta publicação pode ser reproduzida, armazenada em um sistema de recuperação, ou transmitida sob qualquer forma ou por qualquer meio, eletrônico, mecânico, fotográfico, reprográfico, de microfilmagem ou outros, sem a permissão escrita da Editora, com exceção de qualquer material fornecido especificamente no propósito de ser entrado e executado num sistema computacional, para o uso exclusivo do leitor da obra.

Copyright © 2009 by MCT/INPE. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, microfilming or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use of the reader of the work.

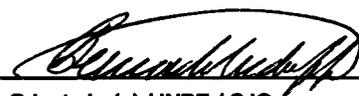
Aprovado (a) pela Banca Examinadora
em cumprimento ao requisito exigido para
obtenção do Título de Mestre em
Sensoriamento Remoto

Dr. Antonio Roberto Formaggio



Presidente / INPE / SJCampos - SP

Dr. Bernardo Friedrich Theodor Rudorff



Orientador(a) / INPE / SJCampos - SP

Dr. Carlos Antonio Oliveira Vieira



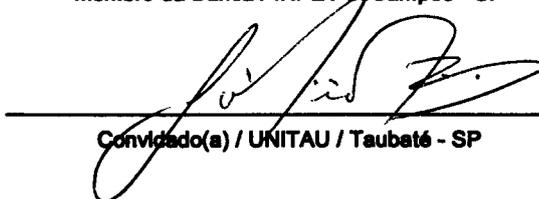
Orientador(a) / UFV / Viçosa - MG

Dr. Camilo Daleles Rennó



Membro da Banca / INPE / SJCampos - SP

Dr. Getúlio Teixeira Batista



Convidado(a) / UNITAU / Taubaté - SP

Aluno (a): Márcio Pupin de Mello

São José dos Campos, 22 de maio de 2009

“Serei breve, mas não tão breve que a eternidade escape do coração.”

Lindolf Bell

*À minha mãe, meu Anjo, minha Angela,
por ser mãe, colo, ombro, vida, minha vida...*

AGRADECIMENTOS

A Deus.

Aos meus orientadores Dr. Bernardo F. T. Rudorff e Dr. Carlos A. O. Vieira, pelos conhecimentos transmitidos, por acreditar e confiar em mim e, principalmente, pela amizade construída e consolidada.

Ao Dr. Morton J. Canty, do *Institute for Chemistry and Dynamics of the Geosphere*, no *Jülich Research Center* – Alemanha, pela ajuda com a normalização radiométrica IR-MAD.

Aos Drs. Luiz A. Peternelli, Flávio J. Ponzoni e Mauro A. H. Antunes e ao M.Sc. Daniel A. Aguiar, pelos esclarecimentos prestados e pela solicitude.

Aos Drs. Getulio T. Batista, Camilo D. Rennó e Antônio R. Formaggio, pela excelente discussão e pelas contribuições na banca de defesa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

Ao INPE pela oportunidade de desfrutar de tão brilhante instituição.

À minha família, por todo o apoio e formação.

Aos amigos das repúblicas de São José dos Campos - SP e Viçosa - MG, pelos bons momentos vividos. Ah, e ao Azul por cozinhar tão bem!

Aos grandes amigos e amigas que fiz no INPE.

A todos que, de alguma forma, contribuíram.

Aos brasileiros, que com o dinheiro de seus impostos financiaram toda minha formação, sem muitas vezes poder sequer escolher o mesmo destino a seus filhos.

E por fim, à Bianka, meu amor, por ser meu tudo, minha família. Eu te amo!

RESUMO

Visando antecipar a extinção da queima da palha da cana-de-açúcar na etapa de pré-colheita de 2021 para 2014, a Secretaria do Meio Ambiente do Estado de São Paulo (SMA-SP) e a União da Indústria de Cana-de-Açúcar (UNICA) assinaram, em junho de 2007, um protocolo de intenções. Dessa forma, o monitoramento das lavouras de cana-de-açúcar com e sem a prática da queima da palha torna-se importante para avaliar a eficácia do protocolo. O Instituto Nacional de Pesquisas Espaciais (INPE), em parceria com a SMA-SP e a UNICA, vêm monitorando desde 2006, através de imagens de satélites de média resolução espacial (Landsat e CBERS), a colheita da cana-de-açúcar no Estado de São Paulo. Entretanto, o processo atual de mapeamento é fortemente apoiado na interpretação visual, que fornece resultados precisos, mas demanda um tempo razoavelmente grande considerando a extensão do cultivo da cana no Estado. Nesse contexto, o presente trabalho teve como objetivo avaliar o uso de classificações espectro-temporais de imagens orbitais para mapear áreas de colheita da cana-de-açúcar, com e sem queima da palha. Ao todo foram avaliadas sete classificações, sendo: quatro diferentes variantes do método de superfícies de resposta espectro-temporal (STRS: *Spectral-Temporal Response Surface*); uma classificação utilizando um Modelo Linear de Mistura Espectral (MLME); uma classificação multitemporal com base nas bandas espectrais de todas as imagens da série temporal; e por último, uma classificação de única data, baseada na última imagem da série temporal. Os resultados mostraram que as classificações STRS, a classificação baseada no MLME, assim como a classificação multitemporal utilizando as bandas espectrais de todas as imagens da série temporal, foram eficientes em mapear com precisão as áreas de colheita de cana-de-açúcar, obtendo, em geral, valores de Exatidão Global (EG) acima de 90%. Já a classificação de única data, com EG menor que 70%, apenas corroborou com a literatura, reafirmando a necessidade da utilização de técnicas multitemporais quando de aplicações em monitoramento agrícola. De maneira geral, os resultados mostraram que é possível utilizar métodos de classificação automática de imagens multitemporais para identificar áreas de colheita da cana-de-açúcar, com e sem queima da palha e, assim, automatizar grande parte do processo de interpretação visual.

SPECTRAL-TEMPORAL CLASSIFICATION OF ORBITAL IMAGES TO MAP SUGARCANE WITH PRE-HARVEST STRAW BURNING

ABSTRACT

In order to reduce the time established by law to cease the pre-harvest burning of sugarcane from 2021 to 2014, the Environmental Secretary of São Paulo State (SMA-SP) and the Sugarcane Industry Association (UNICA) signed, in June 2007, an Agro-Environmental protocol of intentions. Therefore, the monitoring of the sugarcane crop, with and without the practice of pre-harvest straw burning, becomes important to assess the effectiveness of the protocol. The National Institute for Space Research (INPE), in partnership with SMA-SP and UNICA, are monitoring the sugarcane harvest procedure in São Paulo State since 2006 using remote sensing images of medium spatial resolution (Landsat and CBERS). The current mapping process is strongly supported by visual interpretation, which provides accurate results, but demands a large amount of time, considering the extension of the cultivated sugarcane area in São Paulo State. In this context, this study has the objective to evaluate the use of spectral-temporal classifications of satellite images to map sugarcane fields with and without pre-harvest straw burning. Seven classification procedures were evaluated: four classifications with different methods of the Spectral-Temporal Response Surface (STRS); one classification with the Linear Spectral Mixture Model (MLME); one classification with the entire time series of Landsat images; and one single date classification with the last image of the time series. The results showed that the STRS classifications, the MLME classification, and the entire time series of Landsat images classification are all effective to precisely map sugarcane fields with and without pre-harvest straw burning, providing, in general, overall indexes greater than 90%. The single date classification showed an overall index less than 70% which was not satisfactory, corroborating with the fact that multitemporal images are paramount for agriculture monitoring applications. Finally, results showed that it is possible to use automated image classification procedures to identify sugarcane harvested fields with and without pre-harvest straw burning and, therefore, substitute great part of the visual interpretation process.

SUMÁRIO

Pág.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SIGLAS E ABREVIATURAS

1. INTRODUÇÃO	25
1.1 Objetivo Geral	27
1.2 Objetivos Específicos	27
2. FUNDAMENTAÇÃO TEÓRICA	29
2.1 Cana-de-açúcar	29
2.1.1 Colheita da cana-de-açúcar	30
2.1.2 Calendário agrícola da cana-de-açúcar	31
2.2 Landsat-5/TM	32
2.3 Modelo Linear de Mistura Espectral	33
2.4 Classificação multitemporal STRS	34
3 MATERIAIS E MÉTODOS	37
3.1 Área de estudo	37
3.2 Imagens utilizadas	37
3.3 Pré-processamento das imagens	39
3.3.1 Correções geométricas	39
3.3.2 Correções radiométricas	39
3.4 Máscara utilizada, mapa de referência e amostras de treinamento	41
3.5 Classificações testadas	43
3.6 Classificação por Superfície de Resposta Espectro-Temporal	43
3.6.1 O comportamento espectro-temporal da colheita da cana-de-açúcar	44
3.6.2 Os pontos de controle da interpolação de superfícies STRS	45
3.6.3 Nova metodologia de distribuição dos pontos de controle	46
3.6.4 Re-escalamento nos pontos de controle	48
3.6.5 Interpolação das superfícies analíticas	50
3.6.6 Re-escalamento dos coeficientes ajustados	51
3.6.7 Amostras de treinamento	52
3.7 Modelo Linear de Mistura Espectral utilizado	55
3.7.1 O modelo	55
3.7.2 Procedimento da classificação multitemporal usando o MLME	55

3.8 Classificações 6 e 7	56
3.9 Avaliação das classificações	57
4. RESULTADOS E DISCUSSÃO	59
4.1 Comportamento espectro-temporal da colheita da cana-de-açúcar.....	59
4.2 Classificações	65
4.2.1 Classificação 1 (C1)	65
4.2.2 Classificação 2 (C2)	72
4.2.3 Classificação 3 (C3)	77
4.2.4 Classificação 4 (C4)	83
4.2.5 Classificação 5 (C5)	90
4.2.6 Classificação 6 (C6)	94
4.2.7 Classificação 7 (C7)	97
4.3 Comparação entre as sete classificações	100
5. CONCLUSÕES E RECOMENDAÇÕES	103
5.1 Conclusões	103
5.2 Recomendações.....	103
REFERÊNCIAS BIBLIOGRÁFICAS	105
APÊNDICE A – CORREÇÕES RADIOMÉTRICAS	115
A.1 Algoritmo em LEGAL (converte número digital em reflectância aparente)	115
A.2 Dados de entrada do algoritmo 6S (um exemplo).....	116
A.3 Código IDL (adaptado) normalização radiométrica IR-MAD	116
APÊNDICE B – CÓDIGOS IMPLEMENTADOS EM LINGUAGEM R.....	123
B.1 Algoritmo para gerar os coeficientes ajustados de toda a imagem	123
B.2 Função para interpolar os coeficientes utilizando arquivos *.TSA	125
B.3 Função que calcula índices de exatidão, suas variâncias e testes Z	128

LISTA DE FIGURAS

2.1 – Plantação de cana-de-açúcar no início do crescimento da cultura, com a presença do solo exposto e alguns vestígios de palha.	29
2.2 – a) Colheita da cana-de-açúcar crua (sem o uso do fogo) de maneira mecanizada; e b) Procedimento de queima na etapa de pré-colheita da cana-de-açúcar.	31
3.1 – Imagem em composição falsa-cor (R4 G5 B3) da área de estudo, localizada no Estado de São Paulo. Imagem adquirida em 06 de julho de 2007 pelo sensor Landsat-5/TM. Coordenadas no sistema Geográfico/SAD69. Destaque para como se apresentam visualmente as três classes temáticas investigadas nesse estudo: <i>Cana Crua</i> , <i>Cana Queima</i> e <i>Cana-em-Pé</i>	37
3.2 – Cronologia das seis imagens utilizadas do ano de 2007, dentro do período da colheita da cana-de-açúcar no Estado de São Paulo, que abrange os meses de abril a novembro.	38
3.3 – Seis imagens utilizadas do ano de 2007, em composição R4 G5 B3. Elas são datadas de: a) 03 de maio; b) 20 de junho; c) 06 de julho; d) 07 de agosto; e) 23 de agosto; e f) 08 de setembro.	38
3.4 – Mapa de referência, elaborado por interpretação visual, contendo as classes <i>Cana Crua</i> (CC), <i>Cana Queima</i> (CQ) e <i>Cana-em-Pé</i> (CP). As áreas em cor preta correspondem às “eliminadas” pela aplicação da máscara.	41
3.5 – Mapa de Treinamento, contendo os polígonos delimitados por edição vetorial sobre as imagens, utilizado na seleção de amostras de treinamento das classificações. As classes temáticas são: <i>Cana Crua</i> colhida na data 1 (CC1), colhida na data 2 (CC2), CC3, CC4, CC5, CC6; <i>Cana Queima</i> na data 1 (CQ1), CQ2, CQ3, CQ4, CQ6; e <i>Cana-em-Pé</i> (CP).	42
3.6 – Esquema ilustrativo da metodologia empregada no estudo do comportamento espectro-temporal da colheita da cana-de-açúcar.	45
3.7 – Exemplos de superfícies espectro-temporais da classe <i>Cana Queima</i> : a) colhida na primeira data (03 de maio) e b) colhida na última data (08 de setembro).	53
3.8 – Esquema geral ilustrando a metodologia de classificação STRS.	54
3.9 – Ilustração dos procedimentos da Classificação 5 (C5) utilizando o MLME.	56
4.1 – Gráfico exibindo o comportamento espectro-temporal das classes (a) <i>Cana Crua</i> e (b) <i>Cana Queima</i> . A primeira data, representada pela linha vermelha, corresponde ao período antes da colheita, ou seja, quando a cana estava em pé.	61
4.2 – Superfícies ajustadas para a classe <i>Cana Crua</i> : a) utilizando o método tradicional, em que os valores dos comprimentos de onda são re-escalonados proporcionalmente entre zero e um; e b) utilizando o novo método proposto, em que os valores de comprimento de onda são dispostos equidistantemente entre zero e um.	62
4.3 – Superfícies ajustadas utilizando a distribuição tradicional, proporcionalmente re-escalonada, para o eixo de comprimento de onda, das três classes: a) <i>Cana Crua</i> ; b) <i>Cana Queima</i> ; e c) <i>Cana-em-Pé</i>	62
4.4 – Superfícies ajustadas para as três classes: a) <i>Cana Crua</i> ; b) <i>Cana Queima</i> ; e c) <i>Cana-em-Pé</i> , utilizando o novo método proposto em que os valores de comprimento de onda são dispostos equidistantemente no intervalo [0,1].	63
4.5 – Superfícies interpoladas utilizando o método tradicional de distribuição dos pontos de controle no eixo do comprimento de onda e após o re-escalonamento dos coeficientes ajustados.	64
4.6 – Superfícies interpoladas utilizando o novo método de distribuição dos pontos de controle no eixo do comprimento de onda e após o re-escalonamento dos coeficientes ajustados.	64

4.7 – Amostras de superfícies finais, utilizadas na Classificação 1 (C1), representativas das classes: a) Cana Crua (colhida na data 1 – CC1); b) Cana Crua (colhida na data 2 – CC2); c) CC3; d) CC4; e) CC5; f) CC6; g) Cana Queima (colhida na data 1 – CQ1); h) CQ2; i) CQ3; j) CQ4; k) CQ6; e l) Cana-em-Pé (CP). Os eixos <i>x</i> e <i>y</i> , representando o tempo (dado em dia do ano) e o comprimento de onda (dado em micrometros), respectivamente, estão no intervalo [0,1] e, por isso, suas escalas foram omitidas.....	66
4.8 – Comportamento médio dos coeficientes ajustados para as 12 classes treinadas na C1.	67
4.9 – a) Mapa detalhado contendo as 12 classes da Classificação 1: Cana Crua representadas em verde (variando de verde mais claro para <i>CC1</i> a verde mais escuro para <i>CC6</i>); Cana Queima representadas em azul (azul claro para <i>CQ1</i> até azul escuro para <i>CQ6</i>); e Cana-em-Pé, em vermelho; e b) Mapa geral, onde as 12 classes foram agrupadas nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).....	68
4.10 – Mapa de discordâncias entre o mapa geral da Classificação 1 e o mapa de referência. Os pontos do mapa representam os <i>pixels</i> classificados em classes diferentes da sua classe de referência.....	70
4.11 – Amostras de superfícies finais, utilizadas na Classificação 2 (C2), representativas das classes: a) Cana Crua (colhida na data 1 – CC1); b) CC2; c) CC3; d) CC4; e) CC5; f) CC6; g) Cana Queima (colhida na data 1 – CQ1); h) CQ2; i) CQ3; j) CQ4; k) CQ6; e l) Cana-em-Pé (CP).	72
4.12 – Comportamento médio dos coeficientes ajustados (das 200 amostras de treinamento) para as 12 classes treinadas na Classificação 2.....	74
4.13 – a) Mapa detalhado contendo as 12 classes da Classificação 2: Cana Crua representadas em verde (variando de verde mais claro para <i>CC1</i> a verde mais escuro para <i>CC6</i>); Cana Queima representadas em azul (azul claro para <i>CQ1</i> até azul escuro para <i>CQ6</i>); e Cana-em-Pé, em vermelho; e b) Mapa geral, onde as 12 classes foram agrupadas nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).....	75
4.14 – Mapa de discordâncias entre o mapa geral da Classificação 2 e o mapa de referência.	76
4.15 – Amostras de superfícies finais, utilizadas na Classificação 3 (C3), representativas das classes: a) Cana Crua (colhida na data 1 – CC1); b) CC2; c) CC3; d) CC4; e) CC5; f) CC6; g) Cana Queima (data 1 – CQ1); h) CQ2; i) CQ3; j) CQ4; k) CQ6; e l) Cana-em-Pé (CP).....	78
4.16 – Comportamento médio dos coeficientes ajustados para as 12 classes treinadas na C3.	79
4.17 – a) Mapa detalhado contendo as 12 classes da Classificação 3: Cana Crua representadas em verde (variando de verde mais claro para <i>CC1</i> a verde mais escuro para <i>CC6</i>); Cana Queima representadas em azul (azul claro para <i>CQ1</i> até azul escuro para <i>CQ6</i>); e Cana-em-Pé, em vermelho; e b) Mapa geral, onde as 12 classes foram agrupadas nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).....	81
4.18 – Mapa de discordâncias entre o mapa geral da Classificação 3 e o mapa de referência.	82
4.19 – Amostras de superfícies finais, utilizadas na primeira etapa da Classificação 4 (C4-1), representativas das classes: a) Cana Crua (colhida na data 1 – CC1); b) CC2; c) CC3; d) CC4; e) Cana Queima (colhida na data 1 – CQ1); f) CQ2; g) CQ3; h) CQ4; e i) Cana-em-Pé (CP). (*) <i>pode ser colhida nas datas posteriores</i>	84
4.20 – Amostras de superfícies finais, utilizadas na segunda etapa da Classificação 4 (C4-2), representativas das classes: a) CC1; b) CC2; c) CC3; d) CC4; e) CC5; f) CQ1; g) CQ2; h) CQ3; i) CQ4 e j) CP. (*) <i>pode ser colhida nas datas posteriores</i>	85
4.21 – a) Mapa detalhado das três etapas da Classificação 4: a) C4-1, b) C4-2, e c) C4-3; e d) mapa geral onde os resultados das três etapas foram agrupados nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP). (*) <i>pode ser colhida em datas futuras</i>	86
4.22 – Mapa de discordâncias entre o mapa geral da Classificação 4 e o mapa de referência.	88

4.23 – Investigação de dois talhões classificados de maneira discordante à referência nas classificações apresentadas até o momento (C1, C2, C3 e C4) Imagens em composição R4 G5 B3.	89
4.24 – Comportamento espectral dos <i>endmembers</i> selecionados para o MLME aplicado à segunda imagem da série temporal.	91
4.25 – Imagens-fração resultantes da aplicação do MLME na segunda imagem (data 2) da série temporal. Frações: a) Vegetação (V), b) Palha (P), e c) Cinza/Sombra (C).....	91
4.26 – a) Mapa detalhado contendo as 12 classes da Classificação 5: Cana Crua representadas em verde; Cana Queima representadas em azul; e Cana-em-Pé, em vermelho; e b) Mapa geral, onde as 12 classes foram agrupadas nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).....	92
4.27 – Mapa de discordâncias entre o mapa geral da Classificação 5 e o mapa de referência.	93
4.28 – a) Mapa detalhado contendo as 12 classes da Classificação 6: Cana Crua representadas em verde (variando de verde mais claro para <i>CC1</i> a verde mais escuro para <i>CC6</i>); Cana Queima representadas em azul (azul claro para <i>CQ1</i> até azul escuro para <i>CQ6</i>); e Cana-em-Pé, em vermelho; e b) Mapa geral, onde as 12 classes foram agrupadas nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).....	95
4.29 – Mapa de discordâncias entre o mapa geral da Classificação 6 e o mapa de referência.	96
4.30 – Imagem de 08/09/2007 utilizada na Classificação 7. Composição falsa-cor R4 G5 B3.....	97
4.31 – Mapa geral da Classificação 7 contendo as três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).....	98
4.32 – Mapa de discordâncias entre o mapa geral da Classificação 7 e o mapa de referência.	99

LISTA DE TABELAS

2.1 – Calendário agrícola da cana-de-açúcar no Estado de São Paulo, em que A corresponde à “cana de ano” e AM corresponde a “cana de ano e meio”.....	31
2.2 – Características das bandas espectrais do sensor TM.....	32
3.1 – Resumo da descrição das classificações investigadas neste estudo.....	43
3.2 – Exemplo ilustrativo de valores das coordenadas dos pontos de controle de um <i>pixel</i> representado na metodologia STRS.....	46
3.3 – Exemplo ilustrativo de valores das coordenadas dos pontos de controle de um <i>pixel</i> representado na metodologia STRS, utilizando o novo método de distribuição no eixo do comprimento de onda (eixo y).....	47
3.4 – Exemplo ilustrativo de valores das coordenadas dos pontos de controle de um <i>pixel</i> representado na metodologia STRS, após o re-escalonamento.....	49
4.1 – Valores médios de reflectância de superfície das classes <i>Cana Crua</i> , <i>Cana Queima</i> e <i>Cana-em-Pé</i> , para cada data, nas diferentes bandas do sensor TM, em que o grifo corresponde à data da imagem na qual a colheita foi identificada.....	59
4.2 – Matriz de confusão da Classificação 1, tendo como referência o mapa de interpretação visual; e valores dos índices <i>EG</i> e κ	69
4.3 – Matriz de confusão da Classificação 2, tendo como referência o mapa de interpretação visual; e valores dos índices <i>EG</i> e κ	76
4.4 – Valores médios do coeficiente a_8 , em cada uma das 12 classes, para C1, C2 e C3.....	80
4.5 – Matriz de confusão da Classificação 3 e valores dos índices <i>EG</i> e κ	82
4.6 – Matriz de confusão da Classificação 4 e valores dos índices <i>EG</i> e κ	87
4.7 – Matriz de confusão da Classificação 5 e valores dos índices <i>EG</i> e κ	93
4.8 – Matriz de confusão da Classificação 6 e valores dos índices <i>EG</i> e κ	95
4.9 – Matriz de confusão da Classificação 7 e valores dos índices <i>EG</i> e κ	99
4.10 – Resultado da comparação entre as classificações pelo teste <i>Z</i> sobre o índice <i>Kappa</i> : Valores de <i>Kappa</i> (κ); variância do <i>Kappa</i> (V_κ); na triangular-superior estão os valores da estatística <i>Z</i> (Z_{calc}); na triangular-inferior os valores de <i>p-valor</i> dos testes (simétricos e associados ao Z_{calc}). Na diagonal, os valores de Z_{calc} estão na parte de cima, enquanto que o <i>p-valor</i> associado ao Z_{calc} está na parte inferior da célula. As células em fonte vermelha destacam os resultados significativos a 5% de significância.....	101

LISTA DE SIGLAS E ABREVIATURAS

3D	Três Dimensões
AOD	<i>Aerosol Optical Depth</i>
C1	Classificação 1 (STRS tradicional)
C2	Classificação 2 (STRS com pontos do eixo y equidistantes)
C3	Classificação 3 (STRS com pontos do eixo y equidistantes, sem o re-escalonamento dos coeficientes ajustados)
C4	Classificação 4 (STRS progressivo, realizada em três etapas de classificações STRS)
C5	Classificação 5 (Classificação multitemporal utilizando MLME)
C6	Classificação 6 (Classificação multitemporal utilizando as bandas de todas as imagens da série temporal)
C7	Classificação 7 (Classificação com uma única data)
CAPES	Comissão de Aperfeiçoamento de Pessoal de Nível Superior
CBERS	<i>China-Brazil Earth Resources Satellite</i>
CC	Cana Crua
COL	<i>Collocation</i>
CP	Cana-em-Pé
CQ	Cana Queima
CTC	Centro de Tecnologia Canavieira
EG	índice de Exatidão Global
ENVI	<i>Environment for Visualizing Images</i>
ERS	<i>Earth Resources Satellite</i>
GO	Goiás
h	horas
ha	hectares
IBGE	Instituto Brasileiro de Geografia e Estatística
IDL	<i>Interactive Data Language</i>
IFOV	<i>Instantaneous Field-Of-View</i>
INPE	Instituto Nacional de Pesquisas Espaciais
IR-MAD	<i>Iteratively Re-weighted Multivariate Alteration Detection</i>
IVP	Infra-Vermelho Próximo
κ	índice <i>Kappa</i>
Km	Quilômetros
MaxVer	Máxima Verossimilhança
MG	Minas Gerais

MLME	Modelo Linear de Mistura Espectral
MMQ	Método dos Mínimos Quadrados
MODIS	<i>Moderate Resolution Imaging Spectroradiometer</i>
MS	Mato Grosso do Sul
MSS	<i>Multispectral Scanner System</i>
MT	Mato Grosso
PA	Progressão Aritmética
PIB	Produto Interno Bruto
PR	Paraná
PTS	<i>Polynomial Trend Surface</i>
RMS	<i>Root Mean Square</i>
RNA	Redes Neurais Artificiais
S	<i>South</i>
SAA-SP	Secretaria de Agricultura e Abastecimento do estado de São Paulo
SAD69	<i>South American Datum of 1969</i>
SMA-SP	Secretaria de Meio Ambiente do estado de São Paulo
SP	São Paulo
SPRING	Sistema de Processamento de Imagens Georeferenciadas
SR	Sensoriamento Remoto
STRS	<i>Spectral-Temporal Response Surface</i>
TM	<i>Thematic Mapper</i>
UFV	Universidade Federal de Viçosa
UNICA	União da Indústria de Cana-de-Açúcar
USGS	<i>United States Geological Service</i>
W	<i>West</i>
WRS	<i>Worldwide Reference System</i>
www	<i>World Wide Web</i>

1. INTRODUÇÃO

As tecnologias de Sensoriamento Remoto (SR) e geoinformação têm sido utilizadas em diversas áreas do conhecimento, particularmente nas questões relacionadas ao monitoramento dos recursos terrestres (JENSEN, 2006). A intensificação da atividade agrícola, pressionada pelo aumento da produção de alimentos, e mais recentemente, pela demanda por biocombustíveis, precisa ser realizada de forma equilibrada a fim de não comprometer ainda mais a degradação do meio ambiente (GOLDEMBERG et al., 2008; ZUURBIER e VAN DE VOOREN, 2008; ESCOBAR et al., 2009). Portanto, é primordial investir em metodologias que aperfeiçoem os processos de obtenção de informações precisas, em escala e em tempo hábil para o adequado monitoramento do meio ambiente e da atividade agrícola.

No contexto das culturas agrícolas cultivadas no Brasil, a cana-de-açúcar exerce um papel fundamental na economia do país. O Brasil, além de maior produtor de cana-de-açúcar do mundo, é ainda o país que apresenta o maior crescimento na produção dessa cultura (NEVES e CONEJERO, 2007). Atualmente, a área cultivada com cana no Brasil é da ordem de 9,4 milhões de hectares e cresceu 94,3% entre as safras 2000/01 e 2008/09, segundo dados do Instituto Brasileiro de Geografia e Estatística – IBGE (2000, 2008). As duas principais regiões produtoras de cana no Brasil são a Centro-Sul (SP, MG, PR, GO, MT e MS) e a Nordeste, que na safra 2008/09 cultivaram 7,8 e 1,3 milhões de hectares, respectivamente (IBGE, 2008).

No cenário nacional, o Estado de São Paulo sobressai na produção da cana-de-açúcar, com 5,3 milhões de hectares cultivados na safra 2008/09, sendo responsável pela produção de quase 56% de toda a cana do Brasil (IBGE, 2008). Entretanto, apesar do elevado nível tecnológico adotado no cultivo da cana, a colheita ainda é efetuada, em grande parte, de forma manual, requerendo que seja feita a queima da palha na pré-colheita, o que contribui com o aumento da poluição atmosférica e consequentes prejuízos à saúde pública (KIRCHHOFF et al., 1991; CANÇADO et al., 2006; URIARTE et al., 2009).

Visando antecipar a extinção da prática da queima na pré-colheita, que por força da lei deve cessar em 2031, a Secretaria do Meio Ambiente do Estado de São Paulo (SMA-SP), Secretaria da Agricultura e Abastecimento (SAA-SP) e a União da Indústria de Cana-de-Açúcar (UNICA) assinaram, em junho de 2007, um protocolo de intenções, chamado de “Protocolo Verde”, para extinguir as queimadas em áreas mecanizáveis (com declividade $\leq 12\%$) até 2014. O protocolo também estabelece metas para que pelo menos 70% das áreas cultivadas sejam colhidas sem o uso do fogo até 2010 e que as lavouras de cana que entram em produção pela primeira vez na safra 2009/10, não podem mais ser colhidas com queima (SMA-SP, 2009; UNICA, 2009). Nesse contexto, o monitoramento das lavouras de cana-de-açúcar com e sem a prática da queima da palha na pré-colheita torna-se importante, tanto para avaliar a eficácia do protocolo, quanto para fiscalizar o andamento da queima da palha em lavouras autorizadas pela SMA-SP (AGUIAR et al., 2008).

O sensoriamento remoto orbital permite a obtenção de imagens, sistemática e repetitivamente, em faixas do espectro eletromagnético, possibilitando diferenciar os alvos da superfície terrestre em função da sua resposta espectral (MOREIRA, 2005; JENSEN, 2006); portanto, elas são peça fundamental para monitorar o procedimento de colheita da cana-de-açúcar ao longo da safra, a fim de identificar as lavouras colhidas com e sem a queima da palha, uma vez que, visualmente, essas lavouras aparecem de forma bem diferenciada nas imagens (AGUIAR, 2007). O Instituto Nacional de Pesquisas Espaciais (INPE), em parceria com a SMA-SP, UNICA e o Centro de Tecnologia Canavieira (CTC), vem monitorando desde 2006, através de imagens de satélites de média resolução espacial (Landsat-5 e CBERS), a colheita da cana-de-açúcar no Estado de São Paulo (AGUIAR et al., 2007; AGUIAR et al., 2009). Entretanto, a análise das imagens para fins de monitoramento da colheita da cana é fortemente apoiada na técnica de interpretação visual, que classifica com precisão as áreas de cana colhidas com e sem a queima da palha, mas requer um considerável esforço manual dada a extensão do cultivo da cana no Estado.

Um método inovador para classificação automática de culturas agrícolas em imagens multiespectrais e multitemporais, denominado superfície de resposta espectro-temporal

(STRS: *Spectral Temporal Response Surface*), foi desenvolvido por Vieira (2000). Esse método leva em conta a assinatura espectral, no tempo, de cada elemento (*pixel*) da cena (VIEIRA et al., 2002). No Brasil, o STRS foi aplicado para estimar a área plantada com a cultura da soja nos estados do Rio Grande do Sul (RUDORFF et al., 2007) e do Mato Grosso (EPIPHANIO, 2007), assim como para discriminar entre diversas culturas em alguns municípios do Estado de São Paulo (BRAGA et al., 2007); todos esses trabalhos utilizando o sensor MODIS (*Moderate Resolution Imaging Spectroradiometer*). Uma vez que a metodologia STRS contempla o componente temporal (análise multitemporal), pode-se intuir que este método possa ser aplicado para mapear um fenômeno que acontece em um determinado instante no tempo. Assim, a classificação STRS poderia indicar o momento da colheita de uma determinada cultura e, no caso da cana-de-açúcar, classificar o tipo da colheita – com e sem queima – em função da resposta espectro-temporal diferenciada dessas classes, como mostrado em Mello et al. (2009a).

No contexto apresentado, o presente trabalho teve como hipótese que métodos de classificação automática de imagens multitemporais e multiespectrais, em especial o método STRS, podem ser aplicados para identificar a colheita da cana-de-açúcar com e sem queima da palha, fornecendo resultados compatíveis com a interpretação visual.

1.1 Objetivo Geral

O objetivo deste trabalho foi avaliar o uso de classificações espectro-temporais de imagens orbitais do Landsat-5/TM, em especial a classificação STRS, para mapear áreas de colheita da cana-de-açúcar com e sem queima da palha.

1.2 Objetivos Específicos

O trabalho teve como objetivos específicos:

- a) Adaptar o classificador STRS para mapear áreas de colheita da cana-de-açúcar com e sem a queima da palha;

- b) Explorar alternativas da metodologia STRS que melhor se adaptem ao processo de mapeamento da colheita da cana-de-açúcar;
- c) Utilizar outras técnicas de classificação espectro-temporal como, por exemplo, utilizando o modelo de mistura espectral, e comparar os resultados com os obtidos pela técnica STRS.

2. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, abordar-se-ão considerações fundamentadas sobre os aspectos da cultura da cana-de-açúcar, sua colheita e calendário agrícola. Em sequência será feita uma breve descrição das principais características do sensor Thematic Mapper (TM) a bordo do satélite Landsat-5. Posteriormente, explicar-se-ão aspectos sobre MLME (Modelo Linear de Mistura Espectral) e sobre a classificação STRS (*Spectral-Temporal Response Surface*). O capítulo termina com um rápido comentário sobre avaliação de qualidade de classificação de imagens.

2.1 Cana-de-açúcar

A cana-de-açúcar (Figura 2.1) exerce um importante papel tanto na economia do Brasil quanto nas questões ambientais relacionadas com a redução da emissão de gases do efeito estufa, em virtude do grande potencial que esta cultura apresenta na geração de energia renovável. De acordo com a Empresa de Pesquisa Energética – EPE (2008a) a previsão é que a demanda de etanol no Brasil cresça 150% nos próximos 10 anos, passando a produzir 63,9 bilhões de litros até 2017. Ainda segundo a EPE (2008b) a energia gerada pela cana-de-açúcar por meio do etanol e do bagaço chegou a 16% na composição das fontes primárias de energia utilizadas no país até o mês de maio de 2008, ocupando assim a segunda posição, atrás apenas do petróleo.



Figura 2.1 – Plantação de cana-de-açúcar no início do crescimento da cultura, com a presença do solo exposto e alguns vestígios de palha.

A produção da cana-de-açúcar, no que tange ao setor pós-colheita de industrialização, está focada em dois aspectos: a produção de álcool e seus derivados combustíveis (HIRA e OLIVEIRA, 2009; LUO et al., 2009); e a produção de açúcar. Segundo dados apresentados por Neves e Conejero (2007), 60% da safra é destinada a produção de álcool e todos os seus produtos, e os 40% restantes são destinados a produção de açúcar. De acordo com Goldemberg et al. (2008), o etanol possui inúmeras vantagens como substituto da gasolina, como por exemplo, menores taxas de emissão de dióxido de carbono, pela ocasião de sua queima.

2.1.1 Colheita da cana-de-açúcar

Mesmo com todo o desenvolvimento tecnológico, a produção de cana-de-açúcar ainda encontra alguns problemas de ordem socioeconômica e ambiental, como a queima da palha na etapa de pré-colheita que, dentre outros problemas, emite gases poluentes diminuindo a qualidade do ar e causando prejuízos à saúde pública (ROSEIRO e TAKAYANAGUI, 2004; LE BLOND et al., 2008), além de eliminar as palhas e matérias orgânicas que se decomporiam no solo e o enriqueceriam para a safra seguinte (SOUZA et al., 2005) ou seriam aproveitadas para a co-geração de energia juntamente com o bagaço (ALONSO PIPPO et al., 2007; DAWSON e BOOPATHY, 2007).

Existem basicamente dois tipos de colheita da cana: manual e mecanizada (SOUZA et al., 2005). Em ambos os casos, a colheita pode ocorrer com a queima da palha (Figura 2.2b) ou sem a queima, colhendo a “cana crua” (Figura 2.2a). Entretanto, a colheita manual, quando feita sem o uso do fogo, é difícil de ser executada, uma vez que a queima da palha serve tanto para limpar o local, eliminando a grande biomassa de folhas, quanto para afugentar os animais peçonhentos. Já a colheita mecanizada é perfeitamente viável com a “cana crua” (RODRIGUES e SAAB, 2007). Contudo, existem alguns requisitos que devem ser atendidos, como inclinação do terreno menor que 12%, por exemplo, para que um talhão de cana possa ser considerado mecanizável (BRAUNBECK e OLIVEIRA, 2006).



Figura 2.2 – a) Colheita da cana-de-açúcar crua (sem o uso do fogo) de maneira mecanizada; e b) Procedimento de queima na etapa de pré-colheita da cana-de-açúcar.

2.1.2 Calendário agrícola da cana-de-açúcar

É importante, quando do uso de imagens do sensoriamento remoto para monitoramento de culturas agrícolas, que se conheça o calendário agrícola da cultura de interesse (VIEIRA et al., 2002; SANCHES et al., 2005). Segundo Rudorff et al. (2004), o calendário agrícola da cana-de-açúcar no estado de São Paulo, de modo geral, segue como na Tabela 2.1. Nessa tabela, o termo *A* (cana de ano) corresponde à cana cujo plantio vai de agosto a outubro, quando se inicia o período das precipitações de primavera. As variedades de cana de ano exigem solos mais férteis e devem estar prontas para a colheita do meio para o fim da safra; já o termo *AM* (cana de ano e meio) corresponde à cana plantada entre outubro a abril. Normalmente essa modalidade possui boa produtividade, com tempo de crescimento de 15 a 18 meses e variedades precoces, médias e tardias.

Tabela 2.1 – Calendário agrícola da cana-de-açúcar no Estado de São Paulo, em que *A* corresponde à “cana de ano” e *AM* corresponde a “cana de ano e meio”.

FASE	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Plantio	AM	AM	AM	AM				A	A	AM A	AM	AM
Colheita												

Fonte: Adaptada de Rudorff et al. (2004)

De acordo com Sanches et al. (2005) e Rizzi et al. (2009), o conhecimento do calendário agrícola é de fundamental importância para aplicações do sensoriamento remoto, visto que é o calendário agrícola que irá definir as melhores datas de aquisição de imagens e permitirá inferir sobre toda a dinâmica espectro-temporal das culturas imageadas pelos sensores orbitais.

2.2 Landsat-5/TM

Projetado para uma vida útil de três anos, o satélite Landsat-5 é o satélite de sensoriamento remoto mais antigo em operação. Ele foi lançado no dia 1º de março de 1984 e completou 25 anos de operação em 2009, totalizando mais de 700 mil imagens da superfície da Terra, segundo a *United States Geological Survey – USGS* (2009). Em outubro de 2007 o Landsat-5 apresentou alguns problemas com uma de suas baterias e saiu temporariamente de operação, retomando o funcionamento normal no final do mês de fevereiro de 2008 (USGS, 2008).

O Landsat-5 tem órbita circular, heliosíncrona e quase-polar, de 98,20° de inclinação, com altitude média de 705 Km. Isso confere ao satélite um período de revisita de 16 dias, sendo que o horário de passagem é sempre às 9:45h do horário local. A bordo do Landsat-5 estão os sensores Multispectral Scanner System (MSS) e Thematic Mapper (TM), ambos com faixa imageadora (*swath*) de 185 Km. O sensor TM possui resolução radiométrica de 8 bits e é o principal instrumento do satélite. As características do TM (FREDEN e GORDON JR, 1983) são descritas na Tabela 2.2.

Tabela 2.2 – Características das bandas espectrais do sensor TM.

# Banda	Região do Espectro	Limite		Resolução Espacial (m)
		Inferior (µm)	Superior (µm)	
1	Azul	0,45	0,52	30
2	Verde	0,52	0,60	30
3	Vermelho	0,63	0,69	30
4	Infravermelho próximo	0,76	0,90	30
5	Infravermelho de ondas curtas	1,55	1,75	30
6	Infravermelho termal	10,40	12,50	120
7	Infravermelho de ondas curtas	2,08	2,35	30

As bandas espectrais de um sensor carregam informações importantes para a discriminação dos alvos na superfície da Terra. A informação contida em cada elemento

fundamental (*pixel*) da imagem é resultado da resposta espectral integrada dos elementos que estão contidos no campo de visada instantânea (IFOV: *Instantaneous Field-Of-View*) de cada detector que compõe o sensor (JENSEN, 2006). Assim, um único *pixel* em uma imagem pode conter uma proporção de vários componentes como, por exemplo, vegetação, solo e sombra (KESHAVA e MUSTARD, 2002).

2.3 Modelo Linear de Mistura Espectral

Com o objetivo de quantificar as proporções de cada componente dentro de cada *pixel*, Shimabukuro e Smith (1991) propuseram uma modelagem linear por ajuste de mínimos quadrados com restrições. O objetivo desse método é estimar a proporção dos componentes presentes em cada *pixel* utilizando o chamado Modelo Linear de Mistura Espectral (MLME), e informações espectrais de amostras de *pixels* que contenham apenas um dos componentes a ser estimado. São os chamados *pixels* puros ou *endmembers* (TOMPKINS et al., 1997). O modelo, ajustado para cada *pixel* da imagem, na *i*-ésima banda, é como apresentado por:

$$r_i = \sum_{j=1}^n (a_{ij}x_j) + e_i \quad (2.1)$$

Em que:

- r_i : resposta espectral da *i*-ésima banda para o *pixel* ($i = 1, 2, \dots, m$ bandas);
- a_{ij} : valor na *i*-ésima banda para *endmember* do *j*-ésimo componente ($j = 1, \dots, n$);
- x_j : *j*-ésimo componente (a ser estimado pelo modelo) para o *pixel*;
- e_i : erro associado à observação da *i*-ésima banda para o *pixel*.

Ao modelo são impostas duas restrições: a primeira é que a soma das proporções estimadas por cada componente deve ser igual a uma unidade; a segunda restrição é que cada proporção estimada, em cada *pixel* não pode ser negativa. Assim, com um número de bandas maior que o número de componentes a estimar o sistema é superdeterminado e pode ser resolvido (entre diversos métodos) pelo método que minimiza os desvios (quadrado dos desvios), conhecido por Método dos Mínimos Quadrados (MMQ). Toda a metodologia é descrita detalhadamente em Shimabukuro e Smith (1991).

Como resultado da aplicação da técnica de MLME tem-se as imagens-fração, que representam a proporção estimada de cada componente do modelo. Em alguns trabalhos recentes, como o de Mello et al. (2008) e, posteriormente, Machado et al. (2008), as imagens-fração resultantes da aplicação do MLME foram utilizadas como elementos do vetor de atributos em uma classificação de máxima verossimilhança, apresentando resultados satisfatórios.

2.4 Classificação multitemporal STRS

No uso do sensoriamento remoto em aplicações no monitoramento de culturas agrícolas é fundamental uma abordagem que contemple informações captadas pelos sensores durante o período de tempo que se pretende monitorar. Vieira et al. (2000) utilizou imagens que abrangiam o período de plantio e desenvolvimento de diversas culturas, no interior da Inglaterra, como batata, trigo, cebola e beterraba, entre outras. A justificativa é que cada cultura apresenta um padrão espectral distinto ao longo do ciclo de desenvolvimento, de maneira que é possível, por meio de uma abordagem espectro-temporal, discriminar entre as diferentes culturas agrícolas abordadas.

Dentre os vários métodos de análise multitemporal em SR, o método de classificação por superfícies de resposta espectro-temporal (STRS, do inglês: *Spectral-Temporal Response Surface*), desenvolvido por Vieira (2000), mostrou-se bastante eficiente em mapear com precisão os campos de culturas agrícolas, como pode ser visto em Vieira et al. (2000). A idéia básica é que um determinado alvo agrícola, ao invés de possuir uma única curva espectral, possui curvas espectrais distintas para diferentes tempos de desenvolvimento da cultura. Assim, o alvo que possui uma curva espectral para uma única data, terá então uma superfície espectro-temporal, quando consideradas várias datas de imageamento.

Neste tipo de representação, cada elemento (*pixel*) da imagem é modelado em três dimensões, onde o eixo *x* representa a variação temporal, o eixo *y* a variação espectral e o eixo *z* a reflectância. Através desse sistema 3D, são interpoladas as superfícies que descrevem a dinâmica espectro-temporal dos alvos. Vieira (2000) descreve dois métodos de interpolação das superfícies espectro-temporais: o *Collocation* (COL) e o

Polynomial Trend Surface (PTS). O primeiro ajusta a superfície de forma rígida, tangenciando os pontos de controle. Já o segundo ajusta uma superfície mais suavizada, que descreve a tendência do comportamento espectro-temporal dos alvos, possuindo assim a vantagem de minimizar problemas como, por exemplo, *pixels* com nuvens (VIEIRA et al., 2000).

As superfícies interpoladas serão utilizadas pelo classificador para rotular cada elemento da imagem a uma classe previamente treinada. O processo de classificação se reduz a comparação da forma das superfícies analíticas, que pode ser feita de forma direta, o que demanda muito recurso de processamento computacional. Contudo, como os coeficientes ajustados para cada superfície carregam informações intrínsecas a respeito da forma das mesmas (por exemplo: o coeficiente a_0 representa a translação da superfície em relação à origem¹), a idéia é utilizar esses coeficientes como elementos do vetor de atributos no processo de classificação (VIEIRA, 2000).

Vieira et al. (2000) mostraram que pode-se atingir bons resultados com o uso dos dois interpoladores (COL e PTS) anteriormente citados e utilizando classificadores multitemporais pelos métodos de Máxima Verossimilhança (MaxVer) e Redes Neurais Artificiais (RNA). Em outro trabalho, Vieira et al. (2003) observaram que em seus resultados, o índice *Kappa* de uma classificação pelo método MaxVer aumentou significativamente de 0,68, conseguido com o uso de uma única imagem, para 0,85 usando classificação STRS. No mesmo trabalho, é feita a recomendação de que o classificador MaxVer é o mais indicado para aplicações em STRS, uma vez que encontrou excelentes resultados, assim como a RNA, mas não demandou tanto tempo computacional. Ambos os processos de classificação (MaxVer e RNA) são descritos por Mather (2004).

¹ O modelo utilizado na interpolação das superfícies STRS por ser visto na Equação 3.4 (seção 3.6.5).

3 MATERIAIS E MÉTODOS

3.1 Área de estudo

Para a aplicação das metodologias e testes propostos, foi escolhida uma área com 225 km² (correspondendo a um recorte de 500 x 500 *pixels* em uma imagem de resolução espacial de 30 x 30 m) localizada no Estado de São Paulo (Figura 3.1), aproximadamente entre os municípios de Ribeirão Preto e São Carlos. Trata-se de uma região onde predomina o cultivo da cana-de-açúcar.

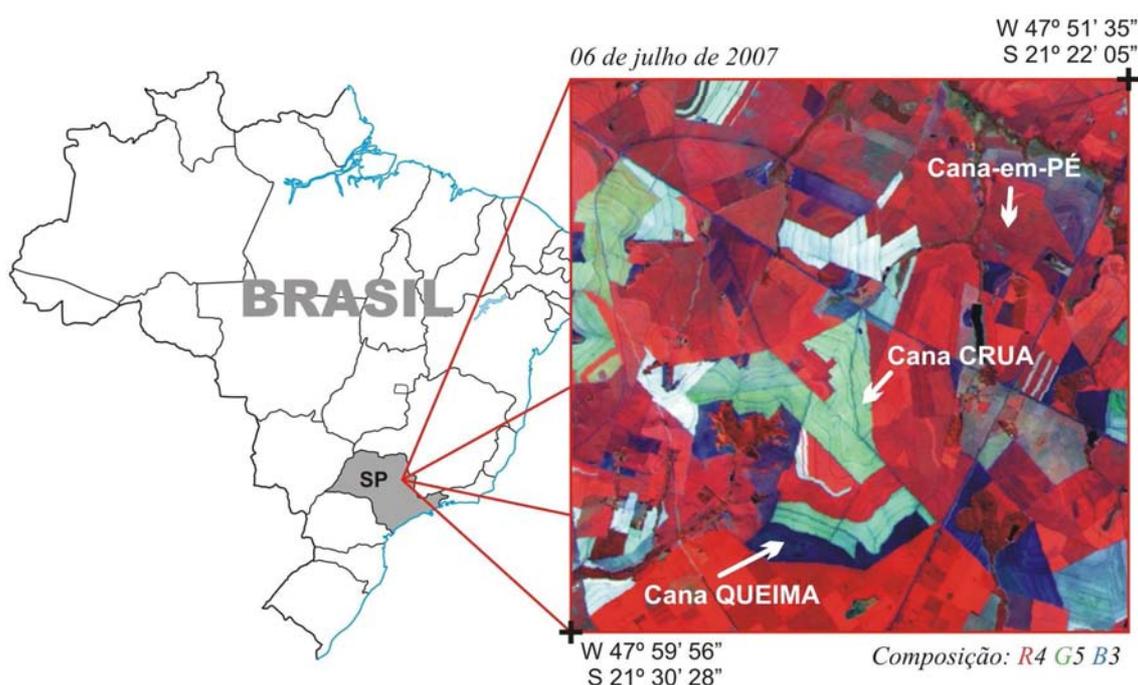


Figura 3.1 – Imagem em composição falsa-cor (R4 G5 B3) da área de estudo, localizada no Estado de São Paulo. Imagem adquirida em 06 de julho de 2007 pelo sensor Landsat-5/TM. Coordenadas no sistema Geográfico/SAD69. Destaque para como se apresentam visualmente as três classes temáticas investigadas nesse estudo: *Cana Crua*, *Cana Queima* e *Cana-em-Pé*.

3.2 Imagens utilizadas

Foram adquiridas seis imagens, obtidas pelo sensor Landsat-5/TM com boas condições de visibilidade e livres de nuvens, da área de estudo em questão. A Figura 3.2 mostra um esquema com a cronologia das seis imagens, dentro do período da colheita da cana no Estado de São Paulo, que habitualmente abrange os meses de abril a novembro. Uma interrupção temporária no funcionamento do Landsat5 (USGS, 2008) impossibilitou a

aquisição de imagens do fim do período de colheita do ano de 2007 (meses de outubro e novembro).

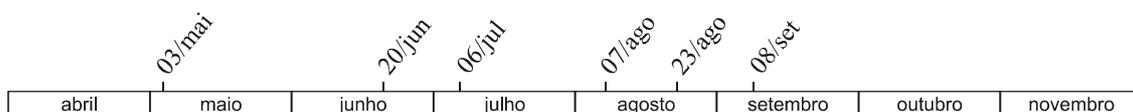


Figura 3.2 – Cronologia das seis imagens utilizadas do ano de 2007, dentro do período da colheita da cana-de-açúcar no Estado de São Paulo, que abrange os meses de abril a novembro.

As imagens correspondem, no sistema WRS (*Worldwide Reference System*) do Landsat, à órbita/ponto 220/075. Elas foram escolhidas por serem as imagens livres de nuvens disponíveis no período da colheita da cana no Estado de São Paulo. Todas foram adquiridas sem custo pelo catálogo de imagens do INPE hospedado na *web* no sítio <http://www.dgi.inpe.br/CDSR/>. As seis imagens supracitadas são mostradas na Figura 3.3.

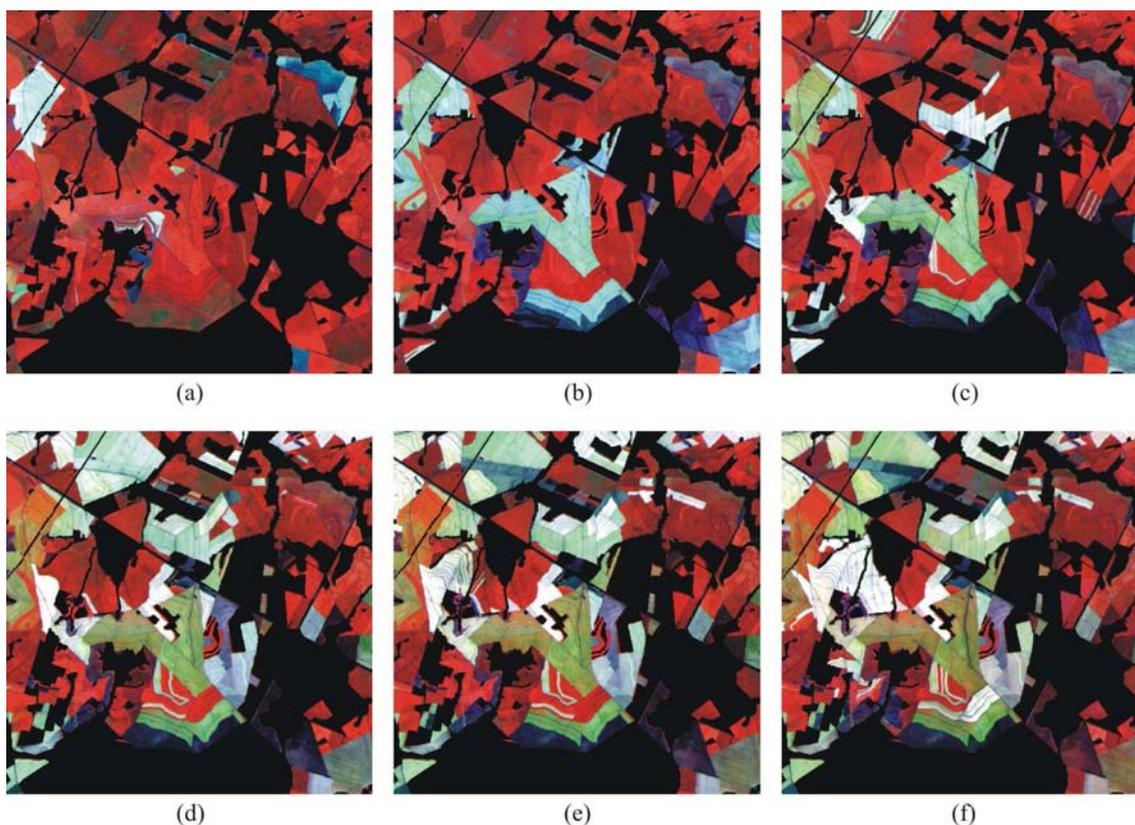


Figura 3.3 – Seis imagens utilizadas do ano de 2007, em composição R4 G5 B3. Elas são datadas de: a) 03 de maio; b) 20 de junho; c) 06 de julho; d) 07 de agosto; e) 23 de agosto; e f) 08 de setembro.

3.3 Pré-processamento das imagens

3.3.1 Correções geométricas

As seis imagens adquiridas foram registradas, com o uso do aplicativo ENVI 4.3 (RSI, 2006), utilizando-se como base uma imagem previamente georrefenciada, que também tinha sido usada como base no registro das imagens que compunham o banco de dados do projeto Canasat (2009) de mapeamento da cana no ano safra 2007/08 realizado pelo INPE (SUGAWARA et al., 2008).

O registro utilizou a metodologia de pontos homólogos nas duas imagens: na imagem base e na imagem a ser registrada. Polinômios de primeiro e segundo graus foram utilizados na modelagem das equações de correspondência. Assim como recomendado por Dai e Khorram (1998), em todos os casos o RMS médio foi menor que 0,5 *pixel*.

3.3.2 Correções radiométricas

Como este estudo aborda modelagens multitemporais que consideram, essencialmente, o comportamento espectral dos alvos ao longo do tempo, é de extrema importância que sejam aplicadas correções radiométricas adequadas nas imagens (SONG et al., 2001). Segundo sugere Schroeder et al. (2006), no caso de séries temporais de imagens, os melhores resultados são obtidos da seguinte maneira: procede-se à correção radiométrica completa (que converte a imagem de valor digital para reflectância de superfície) em uma das imagens da série temporal, preferencialmente uma imagem que se posicione medianamente em relação às datas de todo o conjunto de imagens; posteriormente, utilizando-se a imagem corrigida, agora chamada de imagem base da correção radiométrica, procede-se à normalização radiométrica das demais imagens da série temporal.

A imagem datada de 06 de julho de 2007 foi escolhida para ser corrigida radiometricamente. Seus valores foram convertidos para reflectância aparente valendo-se dos parâmetros de calibração propostos por Chander et al. (2007), utilizando um algoritmo programado em linguagem LEGAL¹ no aplicativo SPRING, versão 4.3.3

¹ Linguagem Espacial para Geoprocessamento Algébrico

(CÂMARA et al., 1996). Posteriormente, os valores foram convertidos para reflectância de superfície utilizando o algoritmo *Second Simulation of the Satellite Signal in the Solar Spectrum – 6S* (VERMOTE et al., 1997), adaptado por Antunes et al. (2003).

Na correção 6S, foi utilizado o valor de visibilidade de 50 Km. Esse valor foi estipulado empiricamente com base no valor de profundidade óptica de aerossol à 550nm (AOD, do inglês: *Aerosol Optical Depth*) obtida pelo produto MOD04¹, uma vez que a visibilidade e a AOD estão diretamente relacionadas (PINTO, 2001). Algumas metodologias, como as apresentadas em Kaufman (1988), Fallah-Adl et al. (1997) e Liang et al. (1997), propõem usar os *pixels* escuros de uma cena para estimar a AOD, segundo a pressuposição de que a banda 7 do sensor TM não seria significativamente afetada pela atmosfera. No entanto, como mostrado em Schroeder et al. (2006), essa metodologia tende a superestimar substancialmente os valores de AOD.

De posse da imagem de 06 de julho de 2007, corrigida radiometricamente para valores de reflectância de superfície, procedeu-se à normalização radiométrica das demais imagens do banco de dados, utilizando a previamente corrigida como base. Após o procedimento, todas as imagens da série temporal representavam valores de reflectância de superfície (SCHROEDER et al., 2006).

O procedimento de normalização radiométrica adotado foi o sugerido por Schroeder et al. (2006). A normalização leva em consideração os *pixels* considerados pseudo-invariantes na imagem (SCHOTT et al., 1988) e foi realizada segundo a técnica proposta em Nielsen et al. (1998) e Nielsen et al. (2002), e aprimorada por Nielsen (2007). Como descrita em Canty et al. (2004) e posteriormente em Canty e Nielsen (2008), a técnica, implementada em algoritmo de linguagem IDL², do pacote de *softwares* ENVI 4.3 (RSI, 2006), é conhecida pelo nome *Iteratively Re-weighted Multivariate Alteration Detection* (IR-MAD) e utiliza análise de correlação canônica

¹ A descrição dos produtos do sensor MODIS pode ser encontrada em Correia et al. (2007).

² O algoritmo original (disponível em: <<http://mcanty.homepage.t-online.de/software.html>>, acesso em maio de 2008) programado em IDL, foi modificado com a ajuda do Dr. Morton J. Canty para obter como resultado final a imagem normalizada no formato de arquivo “Byte” (de 8bits) e não “Floating Point” como originalmente implementado. O algoritmo adaptado (*radcal_run.pro*) encontra-se no Apêndice A.

para encontrar os *pixels* pseudo-invariantes de forma automática. O parâmetro de entrada chamado de “*penalization*” foi definido como nulo, assim como sugerido por Nielsen e Canty (2005).

3.4 Máscara utilizada, mapa de referência e amostras de treinamento

Uma das etapas do projeto Canasat de monitoramento da cana envolve a identificação, por interpretação visual de imagens Landsat-5/TM, CBERS2 e CBERS-2B/CCD, das áreas de cana para a colheita, gerando assim uma máscara para eliminar as áreas que não devem ser abordadas, como lagos e estradas (RUDORFF et al., 2005). Essa máscara também foi utilizada na presente pesquisa com o objetivo de eliminar da imagem as áreas que não eram de interesse.

Visando a comparação direta dos resultados obtidos pelas metodologias aqui propostas, um mapa de referência (Figura 3.4) foi criado pela equipe do projeto Canasat, utilizando a análise de interpretação visual das mesmas seis imagens da série temporal. O mapa de referência assim criado permite, então, uma comparação direta entre as metodologias tradicionais de interpretação visual, atualmente empregadas pelo programa do INPE, e as novas metodologias automatizadas propostas na presente pesquisa.

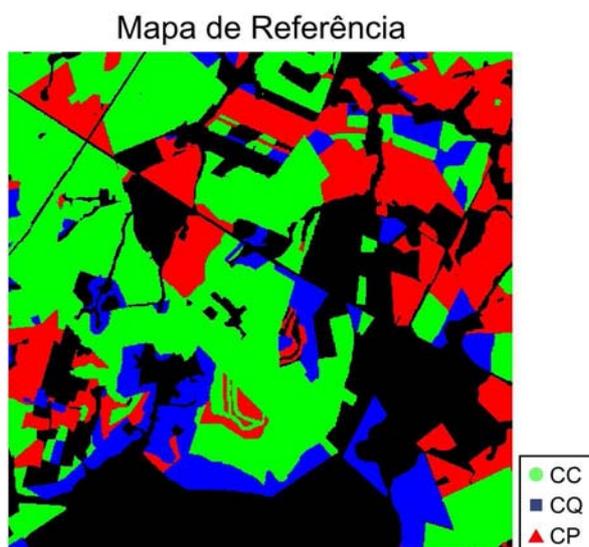


Figura 3.4 – Mapa de referência, elaborado por interpretação visual, contendo as classes Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP). As áreas em cor preta correspondem às “eliminadas” pela aplicação da máscara.

É importante lembrar que o mapa de referência, por se tratar de uma interpretação visual, não traduz exatamente a realidade de campo. Assim, quando uma discordância entre a classificação e a referência for tratada como “erro”, é fundamental o entendimento da subjetividade do termo. Visando minimizar os possíveis erros de interpretação visual no treinamento do classificador, um mapa contendo as classes de treinamento necessárias às classificações foi cuidadosamente criado utilizando edição vetorial sobre as imagens. Nesse mapa, chamado de *Mapa de Treinamento* (Figura 3.5), foram selecionados, cuidadosamente, polígonos localizados no meio dos talhões, de modo a evitar os *pixels* de bordas.

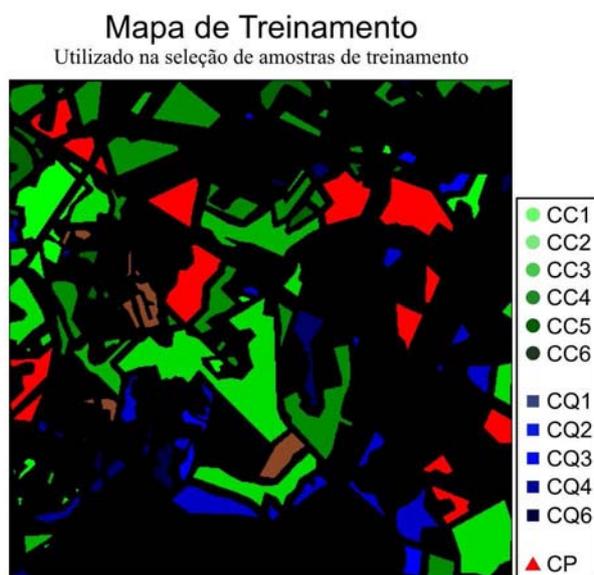


Figura 3.5 – Mapa de Treinamento, contendo os polígonos delimitados por edição vetorial sobre as imagens, utilizado na seleção de amostras de treinamento das classificações. As classes temáticas são: Cana Crua colhida na data 1 (CC1), colhida na data 2 (CC2), CC3, CC4, CC5, CC6; Cana Queima na data 1 (CQ1), CQ2, CQ3, CQ4, CQ6; e Cana-em-Pé (CP).

Uma vez que um *pixel* identificado como Cana Crua (cana colhida sem o uso do fogo) na primeira data, por exemplo, tem o comportamento espectro-temporal totalmente diferente (a resposta espectral característica que permite identificar a colheita é deslocada no eixo do tempo) de um *pixel* de Cana Crua cuja colheita foi identificada na última data analisada, por exemplo, foi necessário treinar uma classe para cada tipo de colheita (Crua ou Queima) em cada data. Assim, as classificações procederam com 12 classes, sendo elas: Cana Crua colhida na data 1 (CC1), Cana Crua colhida na data 2

(CC2), CC3, CC4, CC5, CC6; Cana Queima (cana colhida com a queima da palha) na data 1 (CQ1), CQ2, CQ3, CQ4, CQ6 (a CQ5 não apresentou quantidade de *pixels* suficientes para que pudesse ser utilizada no treinamento do classificador); e Cana-em-Pé (CP), representando a cana que não foi colhida.

3.5 Classificações testadas

Neste trabalho foram testadas sete classificações automáticas de imagens, sendo seis multitemporais e uma classificação de única data. Dessas sete classificações: quatro envolvem a classificação pelo método STRS; uma utiliza aplicação de MLME; e por fim, as duas últimas são classificações convencionais utilizando as bandas das imagens da série temporal. Essas sete classificações foram numeradas de acordo com o especificado na Tabela 3.1.

Tabela 3.1 – Resumo da descrição das classificações investigadas neste estudo.

Classificação	Descrição
C1	STRS tradicional
C2	STRS com pontos do eixo y equidistantes
C3	STRS com pontos do eixo y equidistantes, sem o re-escalamento dos coeficientes ajustados
C4	STRS progressivo, realizada em três etapas de classificações STRS
C5	Classificação multitemporal utilizando MLME
C6	Classificação multitemporal utilizando as bandas de todas as imagens da série temporal
C7	Classificação utilizando uma única data (08/09/2007)

3.6 Classificação por Superfície de Resposta Espectro-Temporal

A classificação por Superfície de Resposta Espectro-Temporal (STRS: do inglês, *Spectral Temporal Response Surface*) foi desenvolvida por Vieira (2000) e permite considerar a variação temporal das respostas espectrais das culturas agrícolas durante seu ciclo fenológico (VIEIRA et al., 2000). A metodologia de classificação (*pixel-a-pixel*) STRS baseia-se na assinatura espectral do *pixel* no tempo. O procedimento caracteriza-se por usar, para cada *pixel* da imagem, pontos de controle em três eixos cartesianos. Estes eixos correspondem, respectivamente: a) ao tempo, de

acordo com as datas das imagens, dado em dia do ano¹; b) ao comprimento de onda médio das bandas espectrais do sensor, dado em micrometros; e c) aos valores de reflectância espectral do *pixel* nas respectivas datas das imagens e bandas espectrais.

3.6.1 O comportamento espectro-temporal da colheita da cana-de-açúcar

Um estudo preliminar foi realizado visando investigar o comportamento espectro-temporal de lavouras de cana-de-açúcar colhidas com e sem a queima da palha na pré-colheita. O objetivo desse estudo foi investigar a viabilidade da aplicação de técnicas automatizadas de classificação multitemporal de imagens utilizando superfícies de resposta espectro-temporais para o monitoramento da colheita da cana (MELLO et al., 2009a).

De posse das imagens pré-processadas, procedeu-se ao estudo do comportamento espectro-temporal das três classes de interesse: Cana Crua (CC), que corresponde à cana colhida sem o uso do fogo; Cana Queima (CQ), que representa a cana colhida com a queima da palha; e Cana-em-Pé (CP), representando a cana que não foi colhida. Uma vez que o principal objetivo desse estudo pré-liminar do comportamento espectro-temporal da colheita da cana-de-açúcar foi investigar como se comportam os alvos quando da colheita com e sem o uso do fogo, optou-se por escolher a colheita identificada na segunda data da serie temporal, ou seja, na imagem de 20 de junho de 2007 (dia do ano 171). Dessa forma, é possível verificar o comportamento da curva espectral dos alvos antes da colheita (na primeira data) e como seria a dinâmica espectral pós-colheita (visualizada nas datas subsequentes).

Para a realização desse estudo utilizou-se amostras de 5.000 *pixels* para cada uma das três classes supracitadas, assumindo que esse seria um número suficiente para representar o fenômeno estudado. Esses *pixels* foram coletados com base no mapa de treinamento (Figura 3.5, apresentada na seção 3.4). A análise do comportamento espectro-temporal, para cada classe de interesse, foi feito com base nas amostras coletadas estudando-se as médias das respostas espectrais em cada banda. Estas médias

¹ Dia contado de forma contínua, a partir do primeiro dia do ano, sendo que 1º de janeiro corresponde ao dia do ano número 1 e o dia 31 de dezembro corresponde ao dia do ano número 365 (exceto em casos de ano bissexto). Muitas vezes o dia do ano é chamado equivocadamente de dia Juliano.

foram utilizadas na interpolação de superfícies de tendência polinomiais levando-se em conta a assinatura espectral média da classe no tempo. A Figura 3.6 ilustra a metodologia utilizada nessa etapa de investigação.

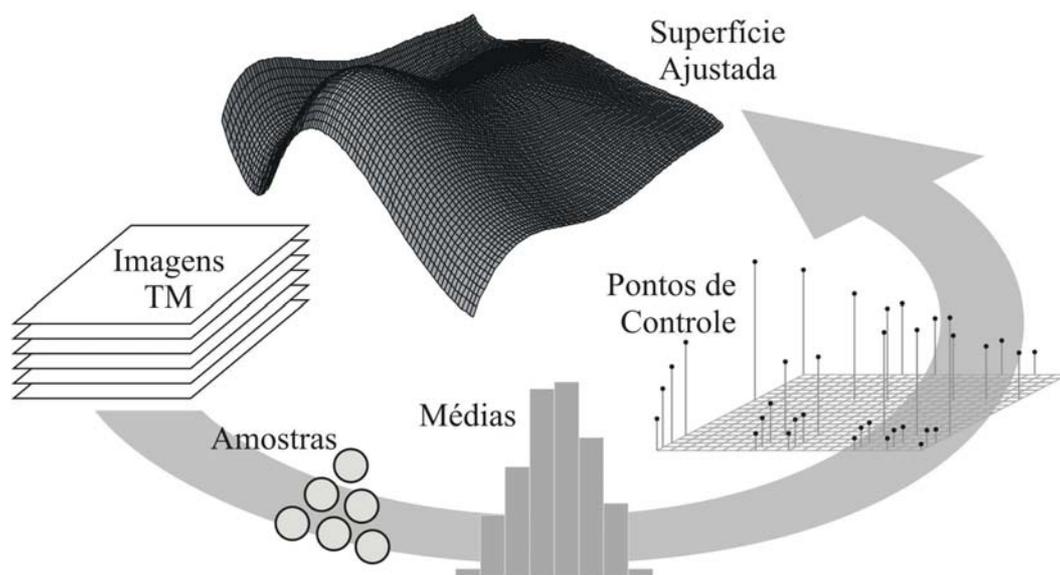


Figura 3.6 – Esquema ilustrativo da metodologia empregada no estudo do comportamento espectral-temporal da colheita da cana-de-açúcar.

Um dos resultados dessa investigação sugere que a distribuição, ao longo dos eixos, dos pontos de controle utilizados na interpolação das superfícies pode não ser adequada. Dessa forma, os resultados conduzem ao fato de que distribuir os pontos de controle de maneira diferente, usando intervalos equidistantes para o eixo do comprimento de onda, poderia apresentar melhores resultados. Essas variações foram testadas visando à identificação da melhor maneira possível de se obter bons resultados com a metodologia STRS no monitoramento da colheita da cana utilizando imagens de satélites.

3.6.2 Os pontos de controle da interpolação de superfícies STRS

Apenas para o melhor entendimento da metodologia, explicar-se-á o processo de classificação STRS com base nas seis imagens e seis bandas do sensor TM (todas exceto a banda 6) utilizadas nesse estudo. Assim, com as seis datas de aquisição das imagens pelo Landsat-5, nas seis bandas do sensor TM, tem-se 36 pontos de controle (6 datas x 6 bandas) no sistema cartesiano (x, y, z) , para cada *pixel*. Os pontos de controle, para cada *pixel*, são como mostrados no (*pixel* de) exemplo da Tabela 3.2.

Tabela 3.2 – Exemplo ilustrativo de valores das coordenadas dos pontos de controle de um *pixel* representado na metodologia STRS.

<i>x</i> (dia do ano)	<i>y</i> (μm)	<i>z</i> (reflectância de superfície)
123	0,485	0,078431
123	0,570	0,137255
123	0,659	0,172549
123	0,841	0,188235
123	1,676	0,298039
123	2,222	0,231373
171	0,485	0,050980
171	0,570	0,090196
171	0,659	0,109804
171	0,841	0,196078
171	1,676	0,231373
171	2,222	0,164706
187	0,485	0,050980
187	0,570	0,086275
187	0,659	0,098039
187	0,841	0,200000
187	1,676	0,207843
187	2,222	0,145098
219	0,485	0,043137
219	0,570	0,086275
219	0,659	0,094118
219	0,841	0,223529
219	1,676	0,211765
219	2,222	0,141176
235	0,485	0,043137
235	0,570	0,074510
235	0,659	0,086275
235	0,841	0,207843
235	1,676	0,200000
235	2,222	0,129412
251	0,485	0,035294
251	0,570	0,070588
251	0,659	0,082353
251	0,841	0,200000
251	1,676	0,180392
251	2,222	0,129412

3.6.3 Nova metodologia de distribuição dos pontos de controle

Um dos resultados do estudo preliminar do comportamento espectro-temporal da colheita da cana-de-açúcar, apresentado no trabalho de Mello et al. (2009a), sugere uma nova metodologia de distribuição dos pontos de controle ao longo do eixo do comprimento de onda, utilizado na metodologia STRS. A metodologia supõe que a disposição das bandas do sensor TM não permite uma boa distribuição dos pontos de controle utilizados na interpolação das superfícies STRS. Isso porque, pela própria concepção do sensor, elas foram mais concentradas na região do visível e infravermelho próximo, o que faz com que os pontos de controle se concentrem numa certa região, no que diz respeito ao eixo do comprimento de onda, “prejudicando” assim o posterior ajuste da superfície de resposta.

O pressuposto da nova metodologia é que, quanto melhor distribuídos por todo o eixo estiverem os pontos, melhor seria o ajuste da superfície. A partir daí, surge a proposição de se usar intervalos equidistantes. A metodologia foi elaborada de modo a gerar valores no eixo do comprimento de onda (eixo y) com distribuição equidistante no intervalo fechado entre zero e um ($[0,1]$). Para tal é utilizada uma Progressão Aritmética (PA) em que a razão (r) é calculada de acordo com o número de bandas consideradas, como mostrado em:

$$r = \frac{(\lim_{sup} - \lim_{inf})}{(n_{bandas} - 1)} \quad (3.1)$$

Em que: r é a razão da PA; \lim_{sup} e \lim_{inf} são os limites superior e inferior do intervalo, respectivamente, igual a 1 e 0; e n_{bandas} é o número de bandas espectrais utilizadas. Assim, para o caso de seis bandas e intervalo $[0,1]$ (escolhido de forma a coincidir com o intervalo do re-escalamento, como será mostrado em sequência), a razão seria 0,2 e teríamos os seis valores distribuídos nos níveis: 0,0; 0,2; 0,4; 0,6; 0,8; e 1,0. Dessa maneira, o exemplo ilustrando os pontos de controle exibidos na Tabela 3.2, ficaria agora como descrito na Tabela 3.3.

Tabela 3.3 – Exemplo ilustrativo de valores das coordenadas dos pontos de controle de um *pixel* representado na metodologia STRS, utilizando o novo método de distribuição no eixo do comprimento de onda (eixo y).

x (dia do ano)	y (μm)	z (reflectância de superfície)
123	0,0	0,078431
123	0,2	0,137255
123	0,4	0,172549
123	0,6	0,188235
123	0,8	0,298039
123	1,0	0,231373
171	0,0	0,050980
171	0,2	0,090196
171	0,4	0,109804
171	0,6	0,196078
171	0,8	0,231373
171	1,0	0,164706
187	0,0	0,050980
187	0,2	0,086275
187	0,4	0,098039
187	0,6	0,200000
187	0,8	0,207843
187	1,0	0,145098

(continua)

Tabela 3.3 – Conclusão.

219	0,0	0,043137
219	0,2	0,086275
219	0,4	0,094118
219	0,6	0,223529
219	0,8	0,211765
219	1,0	0,141176
235	0,0	0,043137
235	0,2	0,074510
235	0,4	0,086275
235	0,6	0,207843
235	0,8	0,200000
235	1,0	0,129412
251	0,0	0,035294
251	0,2	0,070588
251	0,4	0,082353
251	0,6	0,200000
251	0,8	0,180392
251	1,0	0,129412

É importante comentar que esse novo método de distribuição proposto descaracteriza o significado físico do eixo de comprimento de onda, ou seja, os valores desse eixo deixam de representar (em qualquer escala) os reais valores dos comprimentos de onda das bandas do sensor utilizado, mesmo que neste trabalho, por convenção, esse eixo continue sendo chamado de “eixo de comprimento de onda”.

3.6.4 Re-escalamento nos pontos de controle

Segundo recomenda Vieira (2000), pelo fato de assumirem valores muito discrepantes entre si, tanto as informações dos eixos cartesianos dos pontos de controle, quanto os coeficientes ajustados pelo modelo de superfície de tendência (como será mostrado adiante) devem ser re-escalados em um intervalo fechado entre zero e um ([0,1]). O re-escalamento para determinado intervalo pode ser realizado com o uso da seguinte equação:

$$X' = \left[\frac{(X - \min)}{(\max - \min)} \right] \cdot (\lim_{\sup} - \lim_{\inf}) + \lim_{\inf} \quad (3.2)$$

Em que X' é o novo valor re-escalado para o intervalo desejado; X é o valor a ser re-escalado; \min e \max representam o menor e o maior valor, respectivamente, dos dados referentes à normalização de X' ; e \lim_{\sup} e \lim_{\inf} representam os limites superior e

inferior, respectivamente, do intervalo (fechado) desejado. Para o intervalo fechado entre zero e um ($[0,1]$) a Equação 3.2 se reduz a:

$$X' = \frac{(X - \min)}{(\max - \min)} \quad (3.3)$$

Assim, no que diz respeito aos pontos de controle, cada uma das colunas correspondentes aos eixos x e y terão seus valores entre 0 e 1. Já na coluna z , que representa os valores de reflectância de superfície, conceitualmente já no intervalo $[0,1]$, não há necessidade do re-escalamento. Dessa forma, apenas a título de exemplo, os dados dos pontos de controle explicitados na Tabela 3.2, ficariam, após o re-escalamento, como mostrados na Tabela 3.4.

Tabela 3.4 – Exemplo ilustrativo de valores das coordenadas dos pontos de controle de um *pixel* representado na metodologia STRS, após o re-escalamento.

x (dia ano ano)	y (μm)	z (reflectância de superfície)
0,000	0,000000	0,078431
0,000	0,048935	0,137255
0,000	0,100173	0,172549
0,000	0,204951	0,188235
0,000	0,685665	0,298039
0,000	1,000000	0,231373
0,375	0,000000	0,050980
0,375	0,048935	0,090196
0,375	0,100173	0,109804
0,375	0,204951	0,196078
0,375	0,685665	0,231373
0,375	1,000000	0,164706
0,500	0,000000	0,050980
0,500	0,048935	0,086275
0,500	0,100173	0,098039
0,500	0,204951	0,200000
0,500	0,685665	0,207843
0,500	1,000000	0,145098
0,750	0,000000	0,043137
0,750	0,048935	0,086275
0,750	0,100173	0,094118
0,750	0,204951	0,223529
0,750	0,685665	0,211765
0,750	1,000000	0,141176
0,875	0,000000	0,043137
0,875	0,048935	0,074510
0,875	0,100173	0,086275
0,875	0,204951	0,207843
0,875	0,685665	0,200000
0,875	1,000000	0,129412
1,000	0,000000	0,035294
1,000	0,048935	0,070588
1,000	0,100173	0,082353
1,000	0,204951	0,200000
1,000	0,685665	0,180392
1,000	1,000000	0,129412

3.6.5 Interpolação das superfícies analíticas

Os pontos de controle são utilizados na interpolação das superfícies analíticas. Existem diversas maneiras e procedimentos para a interpolação de superfícies, como pode ser visto em Watson (1992). Vieira et al. (2000) recomendam a utilização da metodologia de interpolação de superfícies conhecida por Superfícies de Tendência Polinomial (PTS: *Polynomial Trend Surface*) em concomitância com o classificador de Máxima Verossimilhança, quando levados em consideração o tempo e complexidade computacional para produzir resultados confiáveis.

O PTS tem a característica de suavizar as superfícies interpoladas, minimizando os efeitos de *pixels* saturados ou com nuvens, por exemplo. O procedimento utiliza os dados dos pontos de controle, de cada *pixel*, para interpolar uma superfície de tendência de grau d . Essa modelagem é feita considerando os eixos do tempo (x) e comprimento de onda (y) como variáveis independentes, e o eixo das reflectâncias (z) como variável dependente, segundo o modelo expresso em:

$$\hat{z} = f(x, y) = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 + a_6x^3 + a_7x^2y + \dots + a_{\frac{(d+1)(d+2)-2}{2}}y^d \quad (3.4)$$

No caso do modelo de grau cinco, são 21 coeficientes ajustados (denotados por a_i – que vão do a_0 ao a_{20}). O número de coeficientes a ajustar no modelo (N_{coef}) é dado em função do grau (d) polinomial do mesmo, segundo

$$N_{coef} = \frac{(d+1)(d+2)}{2} \quad (3.5)$$

Em aplicações recentes no Brasil (BRAGA et al., 2007; MELLO et al., 2009b), o modelo tem sido ajustado por um algoritmo implementado por Mello e Vieira (2006) no *software* R (R DEVELOPMENT CORE TEAM, 2009), utilizando solução pelo Método dos Mínimos Quadrados (MMQ), que usa inversão de matriz em seu modelo-solução, como apresentado a seguir:

$$Z = O.A + \varepsilon \quad (3.6)$$

$$\begin{aligned} O'.Z &= O'.O.A \\ (O'.O)^{-1} \cdot (O'.Z) &= (O'.O)^{-1} \cdot (O'.O).A \\ (O'.O)^{-1} \cdot (O'.Z) &= I.A \end{aligned} \quad (3.7)$$

$$A = (O'.O)^{-1} \cdot (O'.Z) \quad (3.8)$$

Em que:

Z : vetor das observações da variável dependente (eixo z);

O : matriz das observações das variáveis independentes (eixos x e y);

A : vetor dos coeficientes (a_i) a serem ajustados (incógnitas);

ε : erro aleatório; $\varepsilon = I.\xi$, em ξ que segue distribuição normal: $\xi \sim N(0, \sigma^2)$.

I : matriz identidade.

Entretanto, este método impõe uma limitação algébrica ao sistema¹. Cada superfície a ser ajustada possui X níveis no fator tempo (eixo x) e Y níveis no fator comprimento de onda (eixo y), que atuam como variáveis independentes no modelo a ser ajustado. Como o processo envolve inversão de matriz, o grau do polinômio a ser ajustado fica restrito ao posto da matriz que representa as variáveis independentes no modelo solução, chamada matriz de delineamento (Matriz O), ou seja, ao número de níveis menos uma unidade, do fator com o menor número de níveis. Como, no caso desse estudo, os eixos x (tempo) e y (espectro – comprimento de onda) possuem seis níveis cada, o grau do polinômio fica restrito ao valor 5.

3.6.6 Re-escalonamento dos coeficientes ajustados

Como sugere Vieira (2000), os coeficientes ajustados podem assumir intervalos bastante distintos entre si, o que prejudicaria a classificação. A solução seria re-escalonar no intervalo $[0,1]$ também os coeficientes ajustados na interpolação das superfícies

¹ Um procedimento diferente do implementado por Mello e Vieira (2006) foi utilizado por Vieira (2000), em que o ajuste das superfícies espectro-temporais é feito por solução de mínimos quadrados utilizando a ortogonalização Gram-Schmidt, como descrito por Björck (1967) e apresentado em Mather (1976). Nesse caso, a única restrição é que o número de coeficientes a ajustar (Equação 3.5) seja menor que o número de pontos de controle, em cada *pixel*.

analíticas. Para o re-escalamento dos coeficientes ajustados, utiliza-se a mesma equação que é utilizada para o re-escalamento dos pontos de controle (Equação 3.3).

Para cada *pixel*, tem-se interpolada uma superfície espectro-temporal. Assim, cada superfície interpolada possui um coeficiente a_0 , um coeficiente a_1 , e assim por diante, até o coeficiente a_{20} (no caso dos polinômios de grau cinco). O re-escalamento leva em consideração cada coeficiente do polinômio, buscando, entre todos os coeficientes a_0 de todas as superfícies ajustadas, qual o menor e qual o maior valor, e utilizando-os na Equação 3.3 de modo a re-escalonar os valores dos coeficientes a_0 de todas as superfícies ajustadas para o intervalo [0,1]. Posteriormente, procede-se da mesma forma para o coeficiente a_1 . E assim sucessivamente até re-escalonar todos os coeficientes.

3.6.7 Amostras de treinamento

As amostras são coletadas com base nas classes de interesse: Cana Crua (CC), que corresponde à cana colhida sem o uso do fogo; Cana Queima (CQ), correspondendo à cana colhida com a queima da palha na pré-colheita; e Cana-em-Pé (CP), que representa a cana que não foi colhida.

Uma vez que o classificador compara a forma da superfície ajustada para um determinado *pixel* da imagem com a forma das superfícies para as quais ele foi treinado, foi necessário treinar 12 classes, sendo seis classes de Cana Crua (CC1, CC2, CC3, CC4, CC5 e CC6, para Cana Crua colhida nas datas 1, 2, 3, 4, 5, e 6, respectivamente), cinco classes de Cana Queima (CQ1, CQ2, CQ3, CQ4 e CC6, para Cana Queima colhida nas datas 1, 2, 3, 4 e 6, respectivamente – a Cana Queima colhida na quinta data não tinha uma quantidade de *pixels* suficiente para a amostragem) e uma classe de Cana-em-Pé (CP). Isso foi necessário porque, como mostra a Figura 3.7, um *pixel* de Cana Queima colhida na primeira imagem (03 de maio), por exemplo, tem a superfície espectro-temporal totalmente deslocada no eixo do tempo em relação a um *pixel* de Cana Queima colhida na última imagem (08 de setembro), o que dificultaria a análise e implicaria em erros de classificação.

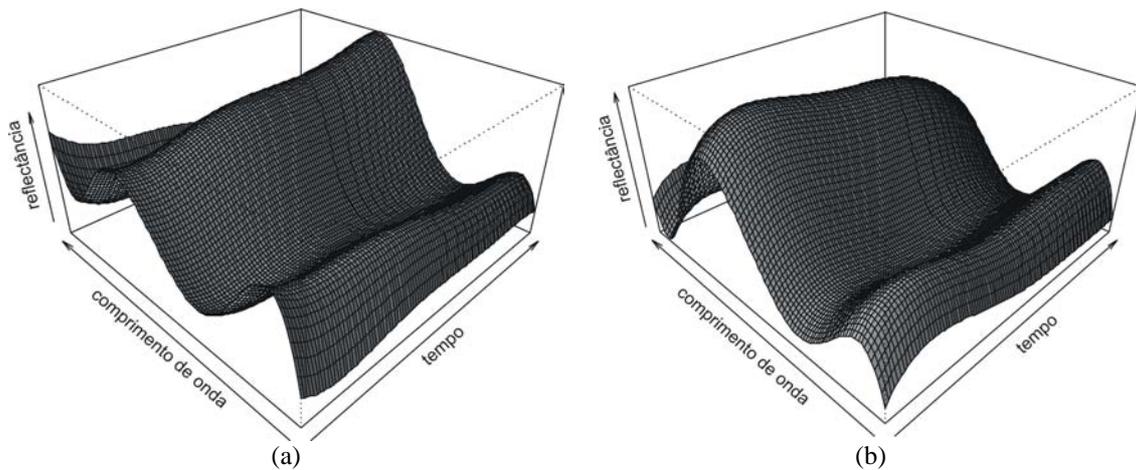


Figura 3.7 – Exemplos de superfícies espectro-temporais da classe Cana Queima: a) colhida na primeira data (03 de maio) e b) colhida na última data (08 de setembro).

Portanto, as superfícies espectro-temporais são ajustadas para as classes Cana Crua e Cana Queima em cada uma das seis imagens utilizadas (exceto a classe Cana Queima da imagem 5), além da classe Cana-em-Pé. No final da classificação, as seis classes de Cana Crua foram agrupadas, assim como as cinco classes de Cana Queima, resultando em um mapa geral com as classes: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).

Nas classificações envolvendo a metodologia STRS (C1, C2, C3 e C4) foram amostrados, de maneira aleatória e independente, 300 *pixels* por classe, totalizando 3600 *pixels* em 12 classes. Esses 300 *pixels* em cada classe foram separados em 200 para o treinamento e 100 para o teste de validação. Esse teste é feito utilizando uma matriz de confusão onde os 100 *pixels* de teste (que possuem suas classes de referência definidas, uma vez que foram selecionados pelo usuário) são classificados com base no treinamento realizado com os 200 *pixels*. A partir dessa matriz são derivados os vários índices de exatidão, que permitem avaliar a qualidade do treinamento da mesma maneira que se avalia uma classificação.

De posse das amostras devidamente treinadas, utilizou-se, para a classificação, o classificador de Máxima Verossimilhança (MaxVer) adaptado por Vieira (2000) e implementado em linguagem C (aplicativo denominado “Max_like_cof.exe”), que é executado por meio da interface criada no aplicativo “Transform.exe”, também

desenvolvido por Vieira (2000). O classificador compara as superfícies com base nos coeficientes ajustados das mesmas, atribuindo um determinado rótulo ao *pixel* cuja superfície ajustada seja mais verossímil a uma das superfícies treinadas. A Figura 3.8 ilustra, de forma geral, a metodologia STRS.

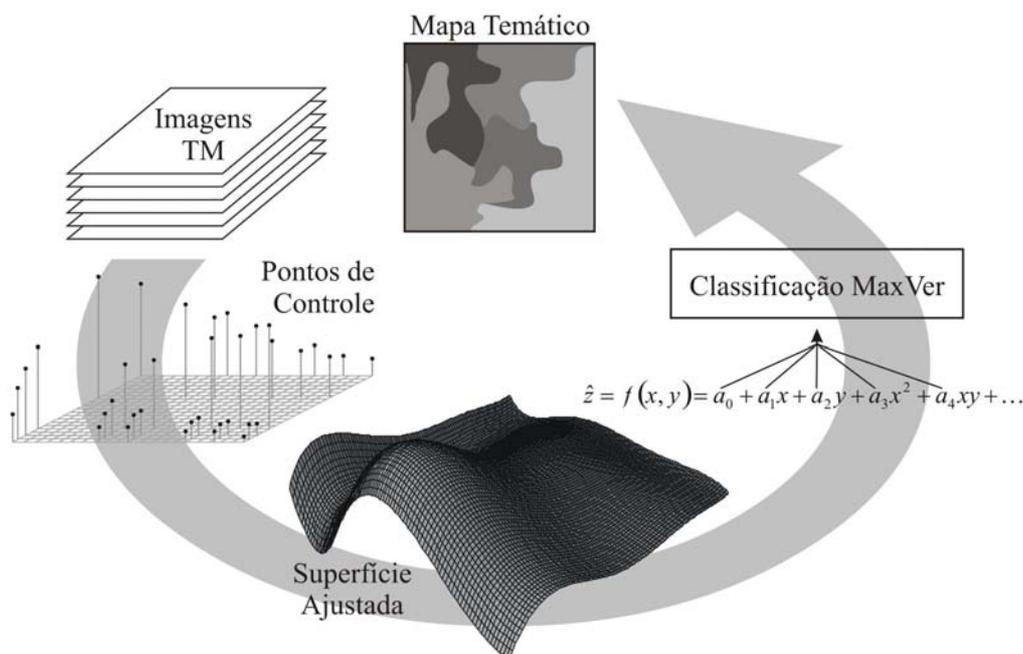


Figura 3.8 – Esquema geral ilustrando a metodologia de classificação STRS.

A metodologia ilustrada na Figura 3.8 consiste, em linhas gerais, das etapas: preparação das imagens (pré-processamento); obtenção dos pontos de controle que serão utilizados na interpolação das superfícies; interpolação das superfícies e obtenção dos coeficientes ajustados; utilização dos coeficientes como elementos do vetor de atributos do classificador MaxVer; e por fim, o resultado da classificação é um mapa temático, chamado de *mapa detalhado* contendo as 12 classes treinadas. Visando chegar a um mapa com as três classes de interesse, as classes semelhantes foram agrupadas, ou seja, as classes CC1, CC2, CC3, CC4, CC5 e CC6 foram agrupadas na classe Cana Crua (CC), assim como as classes CQ1, CQ2, CQ3, CQ4 e CQ6 foram agrupadas na classe Cana Queima (CQ) e, por fim, a Classe Cana-em-Pé (CP) manteve-se. Esse mapa resultante contendo apenas três classes foi chamado de *mapa geral*.

3.7 Modelo Linear de Mistura Espectral utilizado

Além das abordagens de classificações STRS, um procedimento de classificação envolvendo a técnica de Modelo Linear de Mistura Espectral também foi investigado.

3.7.1 O modelo

A Classificação 5 (C5) foi realizada utilizando a metodologia de Modelo Linear de Mistura Espectral (MLME). Esses modelos pressupõem que os *pixels* de uma imagem são compostos por uma relação linear de mistura espectral de alguns componentes presentes dentro da área de resolução do sensor. No caso desse estudo, com a utilização da máscara nas imagens, os elementos predominantes na cena são basicamente *Vegetação*, *Palha* (em decorrência da colheita da cana) e traços da queima da palha, em decorrência da colheita com o uso do fogo, que será denominado de *Cinza/Sombra* (isso porque a componente sombra e a componente cinza possuem padrão espectral muito semelhante). Assim, o MLME utilizado contém os componentes: *Vegetação* (V), *Palha* (P) e *Cinza/Sombra* (C); e é como descrito por:

$$r_i = a_{1i} \cdot V + a_{2i} \cdot P + a_{3i} \cdot C + e_i \quad (3.9)$$

Em que:

r_i : resposta espectral do *pixel* na i -ésima banda;

a_{1i} : resposta do *endmember* para a componente V na i -ésima banda;

a_{2i} : resposta do *endmember* para a componente P na i -ésima banda;

a_{3i} : resposta do *endmember* para a componente C na i -ésima banda;

V: proporção estimada para a componente *Vegetação* no *pixel*;

P: proporção estimada para a componente *Palha* no *pixel*;

C: proporção estimada para a componente *Cinza/Sombra* no *pixel*;

e_i : erro aleatório associado à i -ésima banda espectral no *pixel*.

3.7.2 Procedimento da classificação multitemporal usando o MLME

Na aplicação do modelo, em cada data, foram utilizadas as seis bandas espectrais. Como resultado tem-se três imagens-fração em cada data, totalizando 18 imagens-fração ao

longo da série temporal. A metodologia utilizada na C5 é como mostrada na Figura 3.9: cada imagem (data) gera três imagens-fração (os *endmembers* são selecionados na própria imagem); em seguida todos os seis conjuntos de imagens-fração (três por data) são utilizados como elementos do vetor de atributos na classificação MaxVer, gerando um mapa temático.

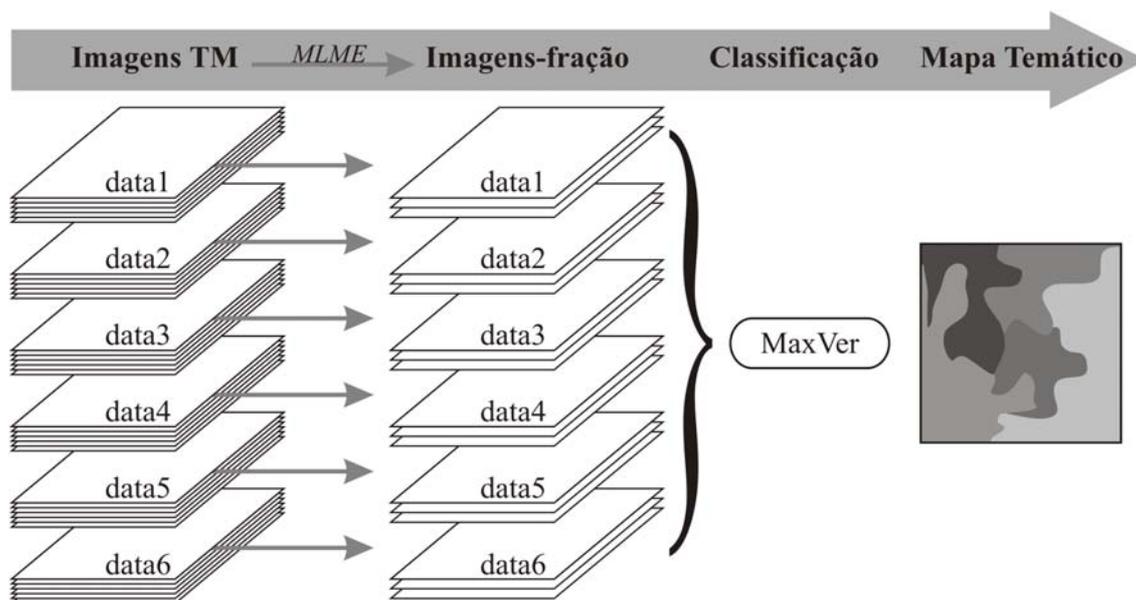


Figura 3.9 – Ilustração dos procedimentos da Classificação 5 (C5) utilizando o MLME.

Para a classificação MaxVer com as imagens-fração resultantes do MLME, foram sorteadas, de maneira independente e aleatória, amostras de 200 *pixels* por classe (foram treinadas as mesmas 12 classes descritas anteriormente).

3.8 Classificações 6 e 7

A Classificação 6 (C6) deu-se da seguinte maneira: o vetor de atributos da classificação MaxVer foi composto por todas as seis bandas espectrais utilizadas e todas as seis datas de imagens analisadas. Na C6, assim como na C5, foram utilizadas amostras de 200 *pixels* por classe, nas 12 classes treinadas.

Já a Classificação 7 (C7) foi realizada utilizando-se apenas uma única imagem (data). A data utilizada foi a última da série temporal (08 de setembro). Na C7 foram treinadas

apenas três classes: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP), sendo também utilizados 200 *pixels* por classe.

3.9 Avaliação das classificações

As sete classificações testadas foram avaliadas usando como referência o mapa obtido por interpretação visual das mesmas seis imagens utilizadas nas classificações. O objetivo foi realizar uma comparação direta entre a interpretação visual e os métodos automatizados testados nesta pesquisa.

A análise da qualidade das classificações procedeu-se de duas formas:

- a) quantitativamente, através dos índices de exatidão global (*EG*) e *Kappa* (κ);
- b) qualitativamente, usando a idéia de imagem binária de erros, como apresentado em Vieira e Mather (2000). Utilizando-se de operações matemáticas, identifica-se a qual *pixel* do mapa temático gerado pela classificação não foi atribuído o mesmo rótulo (classe) do mapa de referência obtido por interpretação visual. O resultado é um mapa de discordâncias, onde se podem localizar, espacialmente na imagem, os *pixels* em que o classificador se diferenciou do intérprete visual, permitindo, além de uma aferição mais detalhada das divergências, uma rápida percepção do quão custoso seria uma posterior correção do mapa classificado pela metodologia automática.

Um teste estatístico *Z* foi realizado para avaliar os valores dos índices *Kappa* derivado das classificações, individualmente e aos pares (CONGALTON e MEAD, 1983). O teste individual verifica se a classificação foi significativamente melhor do que uma classificação casual e é realizado por meio da equação:

$$Z_{calc} = \frac{k}{\sqrt{\sigma_k^2}} \quad (3.10)$$

A Equação 3.10 (em que k representa o valor da estimativa do índice *Kappa* e, dentro do radical, no denominador, está a estimativa de sua variância¹) testa as hipóteses: H_0 : a classificação se deu devido ao acaso; H_a : não H_0 , ou seja, não há evidências para se afirmar que a classificação não foi significativamente melhor que uma classificação casual.

Já o teste descrito pela Equação 3.11 permite a comparação (aos pares) de duas classificações distintas, testando as hipóteses: H_0 : $\kappa_1 = \kappa_2$; e H_a : $\kappa_1 \neq \kappa_2$. Em que k_1 e k_2 representam os valores das estimativas dos índices *Kappa* das classificações 1 e 2, respectivamente (as estimativas de suas variâncias seguem os mesmos padrões de nomenclatura). Assim, o teste que permite a comparação entre duas classificações é dado por:

$$Z_{calc} = \frac{|k_1 - k_2|}{\sqrt{\sigma_{k_1}^2 + \sigma_{k_2}^2}} \quad (3.11)$$

O teste Z considera as pressuposições de: normalidade; que as classes sejam mutuamente exclusivas no mapa; e que as amostragens devam ser realizadas de maneira aleatória e independente. Para que se rejeite H_0 , nos dois testes supra descritos, o valor de Z calculado (Z_{calc}) deve ser maior ou igual ao valor de referência de Z (Z_{tab}). Esse valor de referência, para um nível de significância de 5%, é aproximadamente igual a 1,96.

¹ As equações para o cálculo das estimativas do índice *Kappa* e da variância do *Kappa* podem ser encontradas em Hudson e Ramm (1987).

4. RESULTADOS E DISCUSSÃO

Primeiramente, será abordado o estudo do comportamento espectro-temporal da colheita da cana-de-açúcar. Em sequência, abordar-se-ão cada uma das sete classificações avaliadas. Por fim, será apresentada a análise comparativa de todas as classificações testadas.

4.1 Comportamento espectro-temporal da colheita da cana-de-açúcar

Visando estudar a viabilidade da implantação de técnicas de classificação multitemporal de imagens de satélite no monitoramento da colheita da cana-de-açúcar, especialmente a classificação por STRS, realizou-se um estudo do comportamento espectro-temporal das classes de colheita da cana. De posse das amostras (5.000 *pixels* por classe) das classes de *Cana Crua* (CC), *Cana Queima* (CQ) e *Cana-em-Pé* (CP), foi gerada a Tabela 4.1 com as médias de valores de reflectâncias de superfície em cada uma das seis bandas do sensor TM (todas, exceto a banda 6), para cada uma das seis imagens da série temporal, visando à análise comparativa do comportamento espectro-temporal entre essas classes. As classes de colheita da cana referem-se à colheita identificada na segunda imagem da série temporal (imagem de 20/06/2007).

Tabela 4.1 – Valores médios de reflectância de superfície das classes *Cana Crua*, *Cana Queima* e *Cana-em-Pé*, para cada data, nas diferentes bandas do sensor TM, em que o grifo corresponde à data da imagem na qual a colheita foi identificada.

Cana Crua (CC)						
Data da Imagem (dia do ano)	Banda					
	1 ^a	2 ^b	3 ^c	4 ^d	5 ^e	7 ^f
03/05/2007 (123)	,026	,061	,050	,267	,141	,063
20/06/2007 (171)	,058	,103	,147	,216	,338	,245
06/07/2007 (187)	,053	,091	,125	,215	,320	,229
07/08/2007 (219)	,045	,080	,110	,216	,294	,212
23/08/2007 (235)	,041	,074	,095	,222	,273	,192
08/09/2007 (251)	,034	,065	,077	,226	,243	,164
Cana Queima (CQ)						
Data da Imagem (dia do ano)	Banda					
	1 ^a	2 ^b	3 ^c	4 ^d	5 ^e	7 ^f
03/05/2007 (123)	,025	,055	,048	,243	,136	,063
20/06/2007 (171)	,019	,043	,071	,100	,113	,106
06/07/2007 (187)	,018	,042	,068	,098	,104	,099
07/08/2007 (219)	,019	,043	,069	,109	,109	,100
23/08/2007 (235)	,019	,043	,066	,117	,109	,095
08/09/2007 (251)	,018	,041	,055	,131	,106	,084

(continua)

Tabela 4.1 – Conclusão.

Cana-em-Pé (CP)						
Data da Imagem (dia do ano)	Banda					
	1 ^a	2 ^b	3 ^c	4 ^d	5 ^e	7 ^f
03/05/2007 (123)	,023	,052	,047	,218	,135	,065
20/06/2007 (171)	,023	,052	,045	,237	,133	,059
06/07/2007 (187)	,024	,052	,046	,235	,132	,061
07/08/2007 (219)	,024	,053	,048	,234	,135	,061
23/08/2007 (235)	,024	,053	,048	,240	,131	,059
08/09/2007 (251)	,023	,051	,047	,246	,128	,057

^a azul; ^b verde; ^c vermelho; ^d infravermelho próximo; ^e e ^f infravermelho de ondas curtas (ou médio).

Analisando os valores da Tabela 4.1 pode-se notar que, para as duas classes de colheita da cana (CC e CQ), é possível ver uma alteração no valor médio da reflectância quando ocorre a colheita. O que não acontece na classe CP, que mantém sua resposta espectral praticamente constante ao longo do tempo. Isso já seria suficiente para considerar a potencialidade de uma classificação multitemporal na identificação da colheita da cultura.

Em geral, para a classe CC, os valores de reflectância aumentaram consideravelmente em todas as bandas, exceto na banda 4, que corresponde à região do infravermelho próximo, com comprimento de onda entre 0,76 e 0,90 μm . Isso acontece pois a palha seca da cana e o solo refletem mais em todos os comprimentos de onda das bandas do TM, exceto na banda 4 em que a vegetação tem reflectância mais alta. Já para a classe CQ, os valores de reflectância tenderam a diminuir, exceto nas bandas 3 e 7, que correspondem às faixas espectrais do vermelho (0,63 a 0,69 μm) e infravermelho médio ou de ondas curtas (2,08 a 2,35 μm), respectivamente. Esse comportamento se deve ao fato do solo ter reflectância muito menor com a ausência da palha e presença da fuligem. A Figura 4.1 apresenta o comportamento espectro-temporal das classes de colheita da cana.

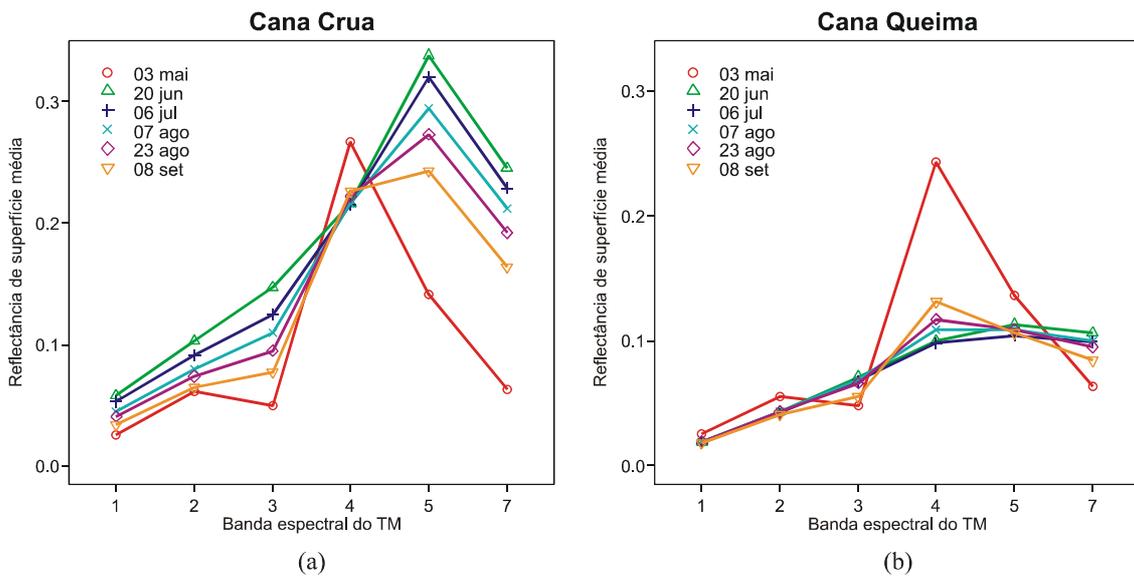


Figura 4.1 – Gráfico exibindo o comportamento espectro-temporal das classes (a) *Cana Crua* e (b) *Cana Queima*. A primeira data, representada pela linha vermelha, corresponde ao período antes da colheita, ou seja, quando a cana estava em pé.

Durante a etapa de interpolação das superfícies pela metodologia STRS, notou-se que a superfície ajustada para a *Cana Crua* não aparentava descrever corretamente a tendência dos pontos de controle, como pode ser visto na Figura 4.2a. Provavelmente essa falta de ajuste da superfície se deve à distribuição dos pontos de controle em relação ao eixo de comprimento de onda (eixo y), onde as bandas do visível e do infravermelho próximo se encontram muito concentradas em relação à faixa do espectro abrangida pelas bandas do TM.

Dessa forma, testou-se, neste trabalho, um método alternativo de distribuição dos pontos de controle, de maneira que, ao invés de re-escalonar os valores do eixo de comprimento de onda, esses valores fossem gerados de modo que os seis níveis (representando as seis bandas do TM utilizadas) ficassem localizados equidistantemente no intervalo fechado entre zero e um ($[0,1]$). A superfície ajustada utilizando esse novo método de organizar o eixo de comprimento de onda é mostrada na Figura 4.2b.

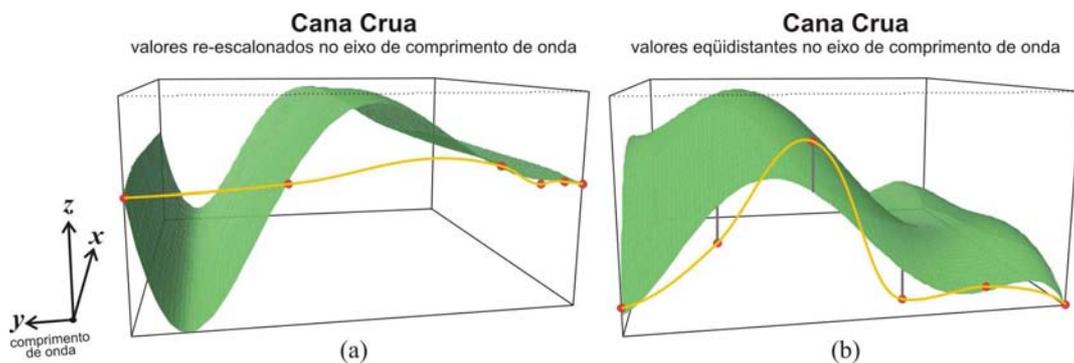


Figura 4.2 – Superfícies ajustadas para a classe *Cana Crua*: a) utilizando o método tradicional, em que os valores dos comprimentos de onda são re-escalados proporcionalmente entre zero e um; e b) utilizando o novo método proposto, em que os valores de comprimento de onda são dispostos equidistantemente entre zero e um.

Observando a Figura 4.2, nota-se que a linha de tendência dos pontos da superfície ajustada pelo método convencional (Figura 4.2a), apesar de praticamente tangenciar os pontos de controle, exagera demasiadamente em curvaturas onde a densidade de pontos é menor, que corresponde aos maiores comprimentos de onda (bandas 5 e 7 do TM). Já a superfície que utiliza a nova metodologia, em que os valores do eixo de comprimento de onda são equidistantes (Figura 4.2b), descreve mais fielmente a tendência mostrada pelos pontos de controle, apesar de não tangenciar todos os pontos. Entretanto, o fato de não tangenciar os pontos não caracteriza um problema, pois, como o próprio nome (*Polynomial Trend Surface*) diz, o método interpola uma superfície que descreve a tendência dos pontos de controle.

A Figura 4.3 mostra as três superfícies interpoladas pela metodologia convencional STRS, com o eixo de comprimento de onda re-escalado proporcionalmente entre zero e um.

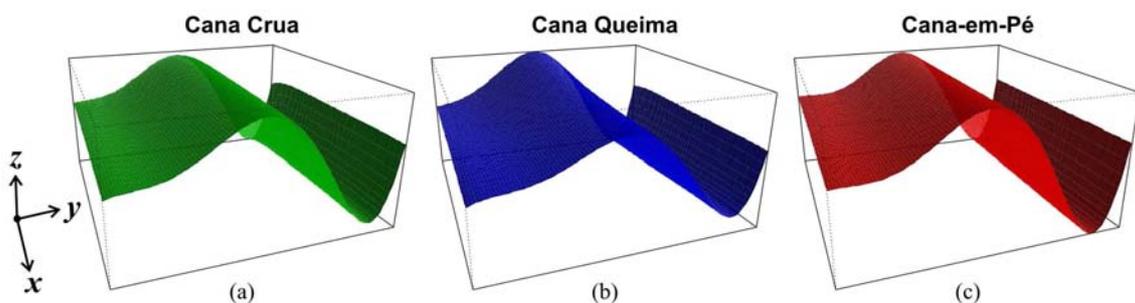


Figura 4.3 – Superfícies ajustadas utilizando a distribuição tradicional, proporcionalmente re-escalada, para o eixo de comprimento de onda, das três classes: a) *Cana Crua*; b) *Cana Queima*; e c) *Cana-em-Pé*.

Visualmente, as três superfícies ajustadas pelo método tradicional (Figura 4.3) têm formas muito semelhantes. Também é possível perceber que nas classes CQ e CP as superfícies (Figuras 4.3b e 4.3c, respectivamente) apresentaram o mesmo problema citado para o ajuste referente à classe CC, quanto ao exagero demasiado nas curvaturas, especialmente na concavidade existente entre as bandas 5 e 7.

As mesmas superfícies das três classes foram novamente interpoladas, agora utilizando a nova metodologia proposta, em que os valores ficam dispostos equidistantemente no eixo do comprimento de onda. O resultado é apresentado na Figura 4.4.

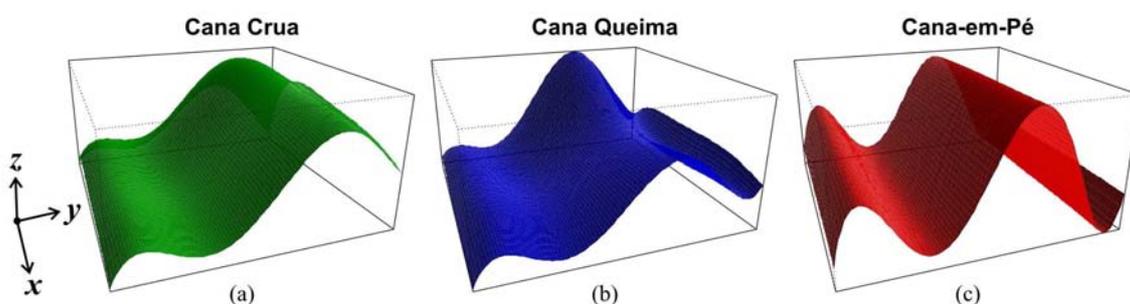


Figura 4.4 – Superfícies ajustadas para as três classes: a) *Cana Crua*; b) *Cana Queima*; e c) *Cana-em-Pé*, utilizando o novo método proposto em que os valores de comprimento de onda são dispostos equidistantemente no intervalo [0,1].

Foi possível perceber que para a classe CP (Figura 4.4c) a superfície ajustada também apresentou o problema da criação de concavidade exagerada entre as bandas 5 e 7. Entretanto, foi possível perceber que a nova metodologia proposta resultou em superfícies visualmente bem mais distintas que a metodologia tradicional. Isso pode ser um indicativo de que a nova metodologia possibilite melhorar os resultados da classificação da colheita da cana utilizando o procedimento STRS.

Contudo, é importante ressaltar que, depois de interpoladas as superfícies, a metodologia sugere o re-escalamento dos coeficientes ajustados. Isso, independentemente do método (seja ele o tradicional ou o novo) de disposição dos valores no eixo do comprimento de onda, certamente altera a forma das superfícies. Assim, outra avaliação pertinente deu-se quanto ao re-escalamento dos coeficientes ajustados. Visando identificar o efeito que esse re-escalamento pudesse ter, foram realizados os re-escalamentos das superfícies das três classes nos dois casos

apresentados, ou seja, utilizando o método tradicional e o novo método de disposição dos valores no eixo do comprimento de onda. O resultado está apresentado nas Figuras 4.5 e 4.6.

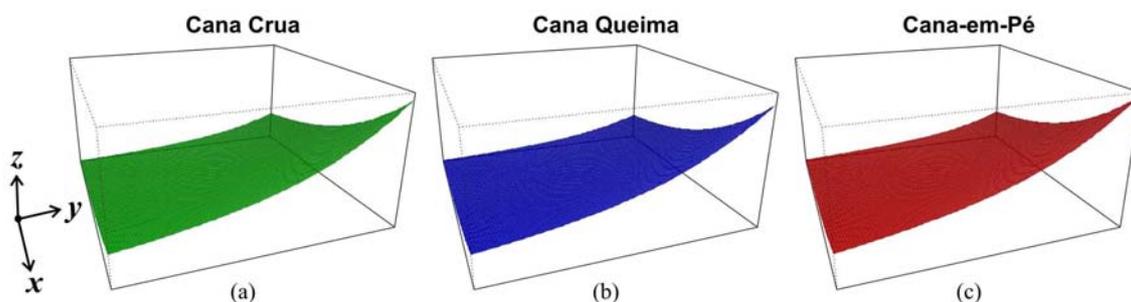


Figura 4.5 – Superfícies interpoladas utilizando o método tradicional de distribuição dos pontos de controle no eixo do comprimento de onda e após o re-escalamento dos coeficientes ajustados.

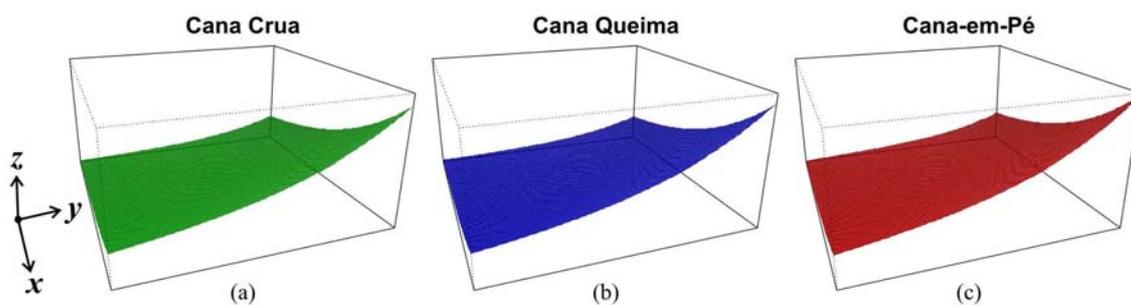


Figura 4.6 – Superfícies interpoladas utilizando o novo método de distribuição dos pontos de controle no eixo do comprimento de onda e após o re-escalamento dos coeficientes ajustados.

Como pode ser visto na Figura 4.5, para o método tradicional de distribuição do eixo y , após o re-escalamento dos coeficientes ajustados, apesar de formas completamente diferentes das que tinham as superfícies antes do re-escalamento dos coeficientes (Figura 4.3), as superfícies das três classes estudadas mantiveram as similaridades visuais entre si. Contudo, o resultado mais interessante vem do re-escalamento dos coeficientes ajustados quando as superfícies foram geradas utilizando o novo método de distribuição dos valores no eixo do comprimento de onda. As superfícies, antes muito distintas visualmente (Figura 4.4), passaram a ser bastante similares (Figura 4.6), inclusive com as superfícies da Figura 4.5. Isso sugere que o re-escalamento dos coeficientes ajustados pode ser prejudicial ao processo de distinção (realizado pelo

classificador), mesmo quando utilizado o novo método de distribuição dos valores de comprimento de onda para os pontos de controle.

Entretanto, é importante lembrar que, embora as diferenças visuais entre as três classes sejam quase imperceptíveis nas Figuras 4.3, 4.5 e 4.6, não significa que isso implique em uma dificuldade para o classificador de máxima verossimilhança (MaxVer), uma vez que ele utiliza relações matemáticas probabilísticas para a distinção das classes, e por isso comporta-se de maneira bem diferente do sentido humano da visão. Assim, é necessário testar as diferentes variações da metodologia STRS, buscando uma maneira de aperfeiçoar os procedimentos buscando sempre a melhoria na qualidade dos resultados.

4.2 Classificações

Ao total, foram realizadas sete diferentes classificações (veja Tabela 3.2), nomeadas de Classificação 1 (C1) à Classificação 7 (C7). Os resultados decorrentes de cada uma das classificações serão descritos subsequencialmente na ordem numérica crescente das mesmas. As quatro primeiras discorrem sobre os métodos e variações da classificação STRS. A quinta classificação trata do uso da técnica de MLME. Já as duas últimas são classificações simples, uma multitemporal e outra de única data, utilizando apenas as bandas espectrais das imagens.

4.2.1 Classificação 1 (C1)

Trata-se da classificação STRS convencional, assim como apresentada por Vieira, (2000). Nela, os valores das coordenadas dos pontos de controle em cada eixo que compõe o espaço cartesiano são re-escalados para o intervalo [0,1], assim como os coeficientes ajustados das superfícies também o são. Foram utilizadas seis imagens (datas) e seis bandas do sensor Landsat 5-TM (todas, exceto a banda 6). Para cada uma das 12 classes treinadas foram coletadas amostras de 300 *pixels*, separadas em 200 para o treinamento e 100 para a validação do treinamento.

4.2.1.1 Superfícies

A Figura 4.7 apresenta o aspecto final de 12 superfícies ajustadas para as 12 diferentes classes: Cana Crua (CC1, CC2, CC3, CC4, CC5 e CC6), Cana Queima (CQ1, CQ2, CQ3, CQ4 e CQ6 – lembrando que a classe Cana Queima imagem 5 (CQ5) não apresentou número suficiente de *pixels* para a amostragem) e Cana-em-Pé (CP).

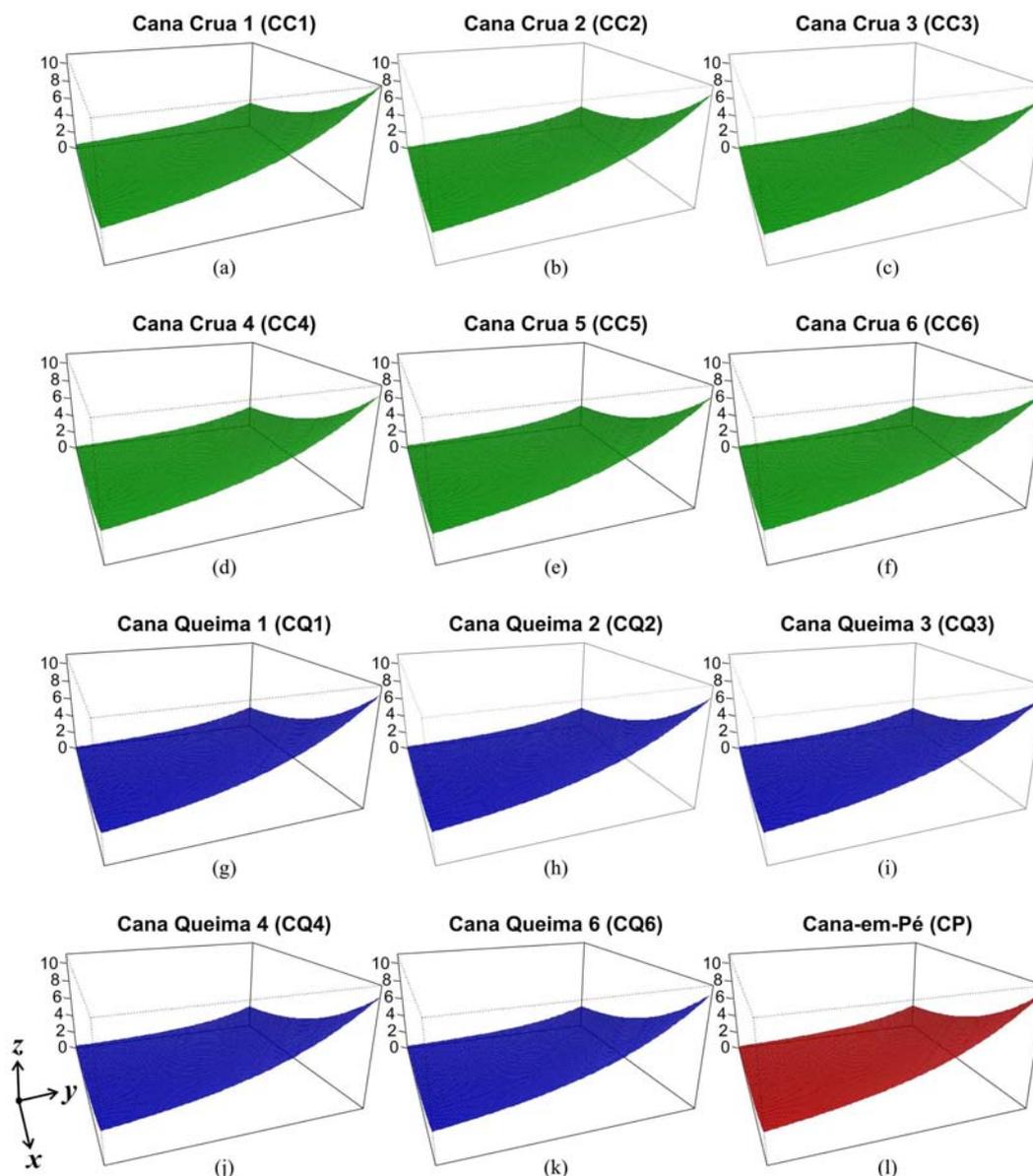


Figura 4.7 – Amostras de superfícies finais, utilizadas na Classificação 1 (C1), representativas das classes: a) Cana Crua (colhida na data 1 – CC1); b) Cana Crua (colhida na data 2 – CC2); c) CC3; d) CC4; e) CC5; f) CC6; g) Cana Queima (colhida na data 1 – CQ1); h) CQ2; i) CQ3; j) CQ4; k) CQ6; e l) Cana-em-Pé (CP). Os eixos x e y , representando o tempo (dado em dia do ano) e o comprimento de onda (dado em micrometros), respectivamente, estão no intervalo $[0,1]$ e, por isso, suas escalas foram omitidas.

Da mesma maneira apresentada anteriormente no estudo preliminar do comportamento espectro-temporal da colheita da cana-de-açúcar, a Figura 4.7 mostra a similaridade visual existente entre as superfícies. Isso ocorre, principalmente devido ao re-escalamento dos coeficientes ajustados. Os eixos x e y , representando o tempo (dado em dia do ano) e o comprimento de onda (dado em micrometros), respectivamente, estão no intervalo $[0,1]$ e, por isso, suas escalas foram omitidas na Figura 4.7.

Outra investigação importante a se fazer é sobre os valores dos 21 coeficientes ajustados para cada uma das 12 classes treinadas. A Figura 4.8 mostra os valores de médias (das 200 amostras de treinamento) para os coeficientes ajustados e re-escalados em cada uma das 12 classes. Uma vez que o classificador utilizado foi o de Máxima Verossimilhança, então, no gráfico de dispersão da Figura 4.8, para cada coeficiente ajustado, quanto mais dispersos estiverem os pontos, melhor será a discriminação entre as classes (outro fator que também influencia na separabilidade pode ser quantificado por uma medida de dispersão, como por exemplo, a variância de cada coeficiente em cada classe, que, quanto menor, melhor também é a distinção entre as classes).

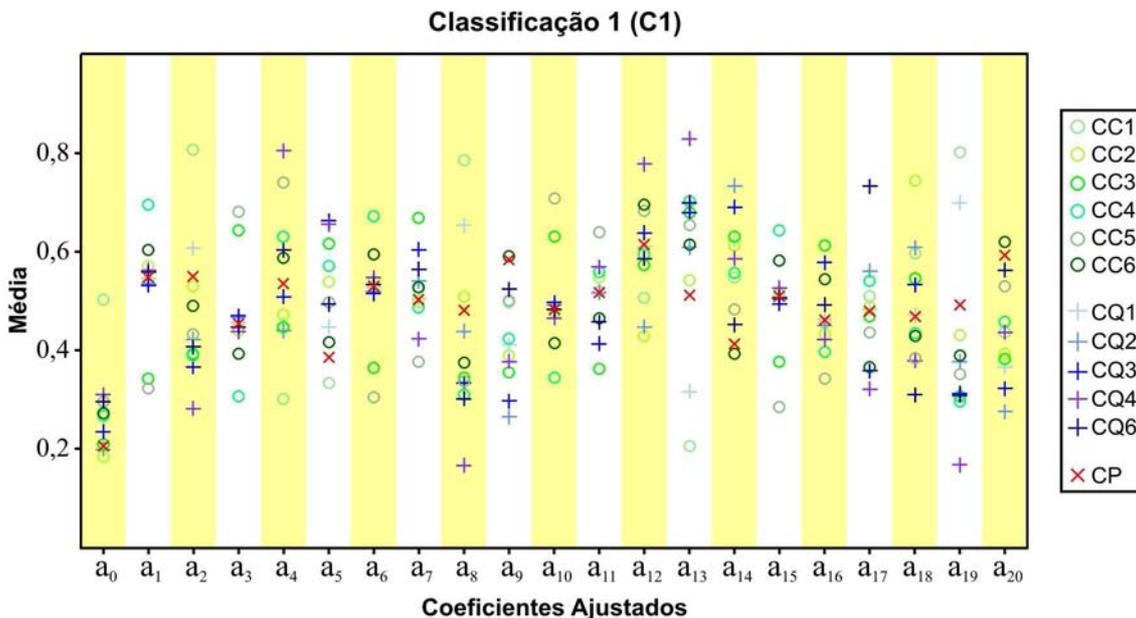


Figura 4.8 – Comportamento médio dos coeficientes ajustados para as 12 classes treinadas na C1.

Analisando a Figura 4.8 foi possível notar que alguns coeficientes apresentam maior separabilidade entre as classes, como, por exemplo, o coeficiente a_8 , enquanto em

outros as classes são menos discrepantes entre si, como no a_9 . Na Figura 4.8, as classes semelhantes (como as classes CC1, CC2, CC3, ..., CC6, por exemplo) possuem a mesma simbologia e cor, variando apenas em tonalidades. É importante lembrar também que não há problemas se houver confusão entre essas classes semelhantes, uma vez que todas serão agrupadas para gerar o “mapa geral” com as três classes de interesse.

4.2.1.2 Mapa temático

A classificação de máxima verossimilhança resultou em um mapa temático chamado de “mapa detalhado” (Figura 4.9a), onde estão presentes as 12 classes anteriormente definidas (CC1 a CC6 para Cana Crua, CQ1 a CQ6 (exceto a CQ5) para Cana Queima, e Cana-em-Pé). As classes CC1 a CC6 são representadas na Figura 4.9a em tons de verde, variando entre verde claro (CC1) a verde escuro (CC6), assim como as classes de CQ1 a CQ6 são representadas em tons de azul claro (CQ1) a azul escuro (CQ6), enquanto que a Cana-em-Pé é representada em vermelho. As classes semelhantes do mapa detalhado foram agrupadas gerando o mapa geral (Figura 4.9b) com as classes Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).

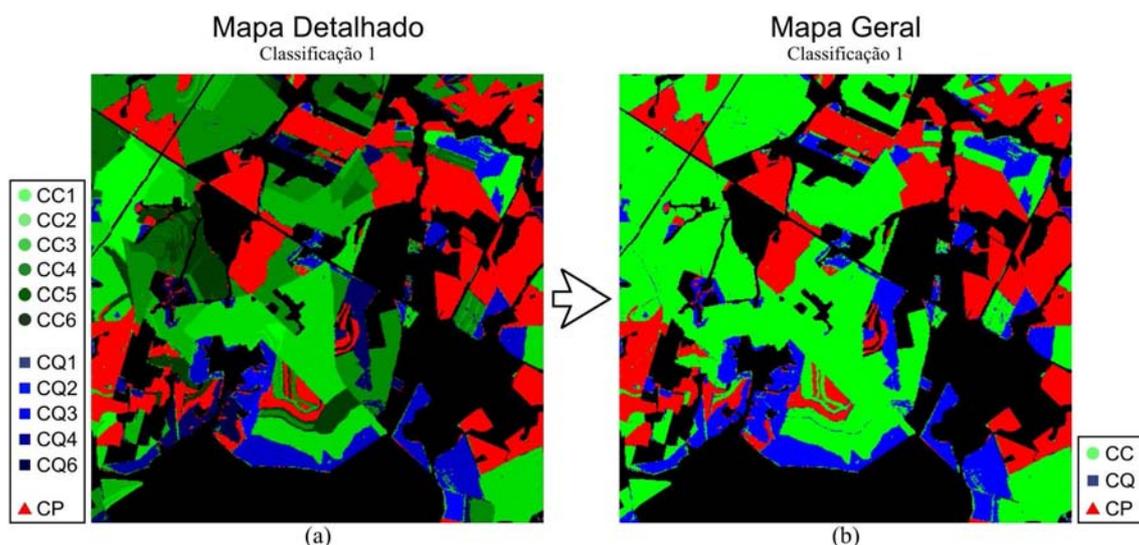


Figura 4.9 – a) Mapa detalhado contendo as 12 classes da Classificação 1: Cana Crua representadas em verde (variando de verde mais claro para CC1 a verde mais escuro para CC6); Cana Queima representadas em azul (azul claro para CQ1 até azul escuro para CQ6); e Cana-em-Pé, em vermelho; e b) Mapa geral, onde as 12 classes foram agrupadas nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).

O mapa temático final (Mapa Geral) apresentou, em geral, padrões uniformes, com pouquíssimos *pixels* isolados. Isso pode ser confirmado com o simples fato de que, apesar de tratar-se de uma classificação *pixel-a-pixel*, vê-se claramente talhões bem definidos.

4.2.1.3 Análise dos resultados

Das 300 amostras selecionadas para cada classe, 200 foram utilizadas no treinamento e 100 na validação desse treinamento. Dos 1.200 *pixels* testados na validação do treinamento (100 *pixels* em cada uma das 12 classes), o classificador errou apenas em 26, resultando em uma Exatidão Global (*EG*) de 97,8%. Como as classes semelhantes serão agrupadas posteriormente, então o possível erro entre elas pode ser desconsiderado, ou seja, se o classificador atribuir a um *pixel* o rótulo CC1, sendo que ele pertence à classe CC2, esse erro deixará de existir quando as classes se agruparem para gerar o mapa geral. Assim, desconsiderando os erros entre classes semelhantes, a *EG* do treinamento sobe para 98,4% (19 *pixels* erroneamente classificados).

A análise quantitativa da precisão do mapa geral resultante da Classificação 1 foi realizada tomando como referência o mapa obtido por interpretação visual. Foi selecionado aleatória e independentemente um total de 1.500 *pixels*, sendo 500 para cada uma das três classes estudadas. Dessa avaliação resulta a matriz de confusão apresentada na Tabela 4.2, que também apresenta os valores dos índices de Exatidão Global (*EG*) e *Kappa* (κ).

Tabela 4.2 – Matriz de confusão da Classificação 1, tendo como referência o mapa de interpretação visual; e valores dos índices *EG* e κ

C1	Referência			TOTAL	Índices (C1)	
	CC	CQ	CP		<i>EG</i>	<i>Kappa</i>
CC	475	54	14	543	91,6%	
CQ	12	424	11	447		0,874
CP	13	22	475	510		
TOTAL	500	500	500	1.500		

Segundo Foody (2002), uma boa classificação deve apresentar um índice *EG* maior que 85%, o que é o caso desta classificação, que apresentou um *EG* de 91,6%. Também o índice κ de 0,874 é classificado como excelente por Landis e Koch (1977). A maior

confusão aconteceu em erros de omissão à classe CQ, quando o classificador classificou de maneira divergente à referência 10,8% (54/500) e 4,4% (22/500) do total de *pixels* da classe CQ como pertencentes às classes CC e CP, respectivamente.

Em complemento à análise quantitativa, gerou-se um mapa de discordâncias (Figura 4.10). Trata-se de um mapa da área de estudo contendo os *pixels* que o classificador rotulou com uma classe diferente em relação ao mapa de referência.

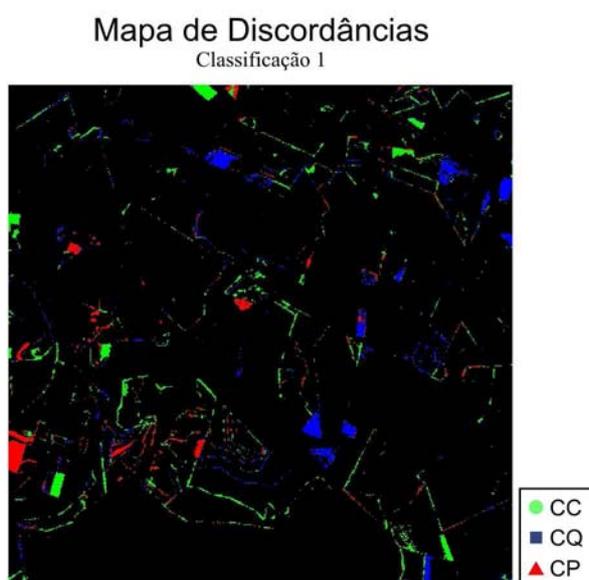


Figura 4.10 – Mapa de discordâncias entre o mapa geral da Classificação 1 e o mapa de referência. Os pontos do mapa representam os *pixels* classificados em classes diferentes da sua classe de referência.

Do total de 10.869 *pixels* classificados de maneira divergente pela C1, 42,4% (4.609) foram erroneamente classificados como Cana Crua, 33,7% (3.658) foram equivocadamente comissionados à classe Cana Queima, e 23,9% (2.602) dos *pixels* foram rotulados como Cana-em-Pé, quando na referência pertenciam a outra classe. Pode-se perceber, inspecionando o mapa de discordâncias (Figura 4.10), que as divergências concentram-se nas bordas de talhões, o que já era esperado, uma vez que, além da mistura espectral inerente desses *pixels*, existem as imperfeições do processo de registro das imagens. Uma discussão importante é que, em geral, as bordas estão associadas a erros de atribuição à classe CC, justificando em parte o maior percentual de divergências supracitado de 42,4% relacionado a essa classe.

Outros erros em talhões de cana colhida (CC e CQ) são decorrentes do seguinte fato: por convenção adotada no projeto Canasat, as classes de CC ou CQ são rotuladas de acordo com a primeira percepção visual diferente da cana em pé. Por exemplo, quando o intérprete identifica que a cana de um talhão foi colhida crua, ele mapeia como CC independente se na imagem subsequente a palha da cana foi queimada, pois este talhão já foi rotulado e não passa mais por uma nova avaliação. Isso é uma dificuldade para o classificador MaxVer, pois ele tem que verificar, por verossimilhança, a qual padrão espectro-temporal aquele talhão (*pixels* do talhão) mais se assemelha. Se, após a colheita da cana-crua, o talhão passa por uma gradagem ou a palha é queimada, o classificador provavelmente vai classificar aquele talhão como CQ.

A classe CC foi principalmente confundida com a CQ. Essas divergências relacionadas à atribuição indevida de *pixels* à classe CC acontecem, em geral, quando o talhão foi colhido logo após uma passagem do satélite, fazendo com que a defasagem entre a data da colheita e a data da próxima imagem seja suficientemente grande para descaracterizar a cicatriz do modo de colheita. Esta dificuldade também ocorre na interpretação visual fazendo com que o dado de referência esteja sujeito a erros. Por fim, algumas discordâncias relacionadas aos *pixels* atribuídos à classe CP podem ocorrer para talhões colhidos bem no início da safra (seja com ou sem queima) e que têm uma rebrota rápida, fazendo com que o classificador “entenda” que o padrão espectro-temporal destes talhões se assemelhe aos de CP.

Assim como apresentado por Vieira et al. (2000) em aplicações de mapeamento de culturas agrícolas na Inglaterra, os resultados aqui apresentados foram excelentes, alcançando precisões compatíveis para tornar potencialmente válidos os procedimentos de classificação STRS aplicado ao mapeamento de áreas de colheita da cana-de-açúcar com e sem a queima da palha. É importante mencionar que a resolução espacial adequada, associada com as características da área mapeada (relevo levemente ondulado e talhões bem definidos) foi um fator importante para esse resultado. Rudorff et al. (2007) encontraram resultados não tão bons no mapeamento da soja no Rio Grande do Sul, especialmente em áreas de minifúndio, por utilizar imagens de resolução moderada (sensor MODIS – 250 x 250 m). Da mesma forma, Braga et al. (2007)

obtiveram resultados com índices *Kappa* menores que 0,6 no mapeamento de diversas culturas utilizando imagens também de resolução moderada (MODIS – 250 x 250 m).

4.2.2 Classificação 2 (C2)

A classificação 2 (C2) foi realizada visando testar a influência da distribuição dos valores dos pontos de controle pelo eixo do comprimento de onda. Ela diferencia-se da anterior (C1) apenas na maneira de dispor os valores dos pontos de controle utilizados na interpolação das superfícies espectro-temporais. Na C2 foi utilizado o novo método de distribuição dos valores no eixo do comprimento de onda (eixo *y*), onde esses valores são dispostos equidistantemente no intervalo [0,1]. Nessa classificação, os valores dos coeficientes ajustados são, assim como na C1, também re-escalados no intervalo [0,1] e também foram usadas seis imagens (datas) e seis bandas do sensor Landsat5-TM (todas, exceto a banda 6). Para cada uma das 12 classes treinadas foram coletadas amostras de 300 *pixels*, separadas em 200 para o treinamento e 100 para a validação do treinamento.

4.2.2.1 Superfícies

A Figura 4.11 apresenta o aspecto final de 12 superfícies ajustadas para as 12 diferentes classes: Cana Crua (CC1, CC2, CC3, CC4, CC5 e CC6), Cana Queima (CQ1, CQ2, CQ3, CQ4 e CQ6 – lembrando que a classe Cana Queima imagem 5 (CQ5) não apresentou quantidade de *pixels* suficiente para a amostragem) e Cana-em-Pé (CP), respectivamente.

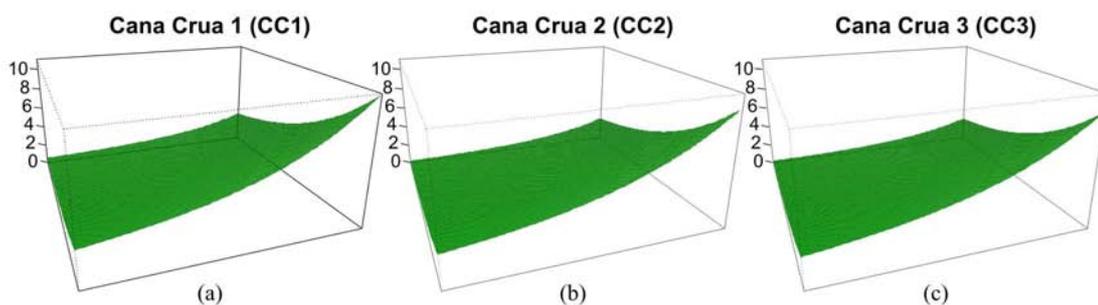


Figura 4.11 – Amostras de superfícies finais, utilizadas na Classificação 2 (C2), representativas das classes: a) Cana Crua (colhida na data 1 – CC1); b) CC2; c) CC3; d) CC4; e) CC5; f) CC6; g) Cana Queima (colhida na data 1 – CQ1); h) CQ2; i) CQ3; j) CQ4; k) CQ6; e l) Cana-em-Pé (CP).

(continua)

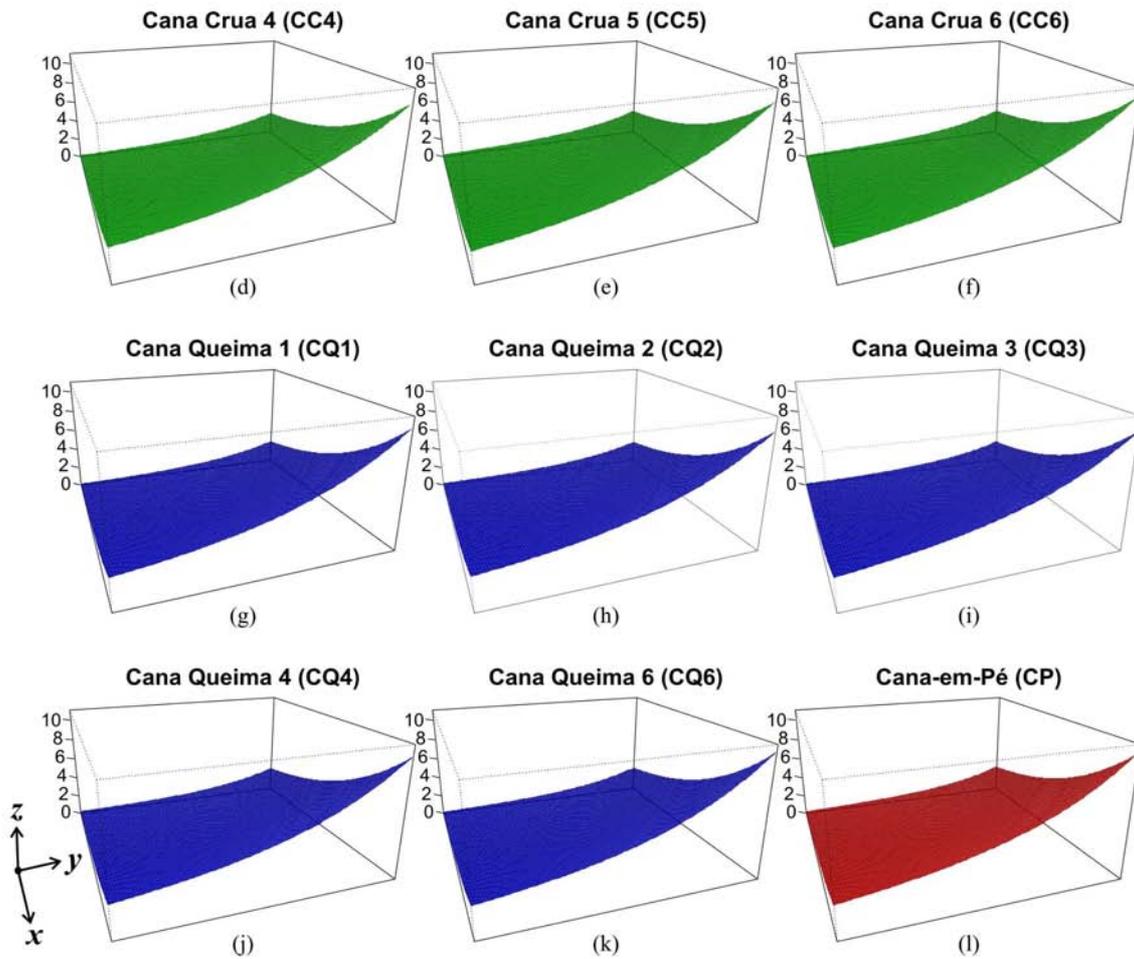


Figura 4.11 – Conclusão.

Assim como na C1, a C2 apresenta similaridade visual entre as superfícies (Figura 4.11), reforçando o fato de que o re-escalamento dos coeficientes ajustados aumenta a similaridade (visual) entre as superfícies.

Na investigação sobre os valores dos 21 coeficientes ajustados para cada uma das 12 classes treinadas Figura 4.12 percebe-se um padrão muito semelhante ao da C1, indicando que toda a melhora na forma das superfícies de cada classe, gerada pelo novo método de distribuição dos pontos de controle, provavelmente é perdido com o re-escalamento dos coeficientes ajustados.

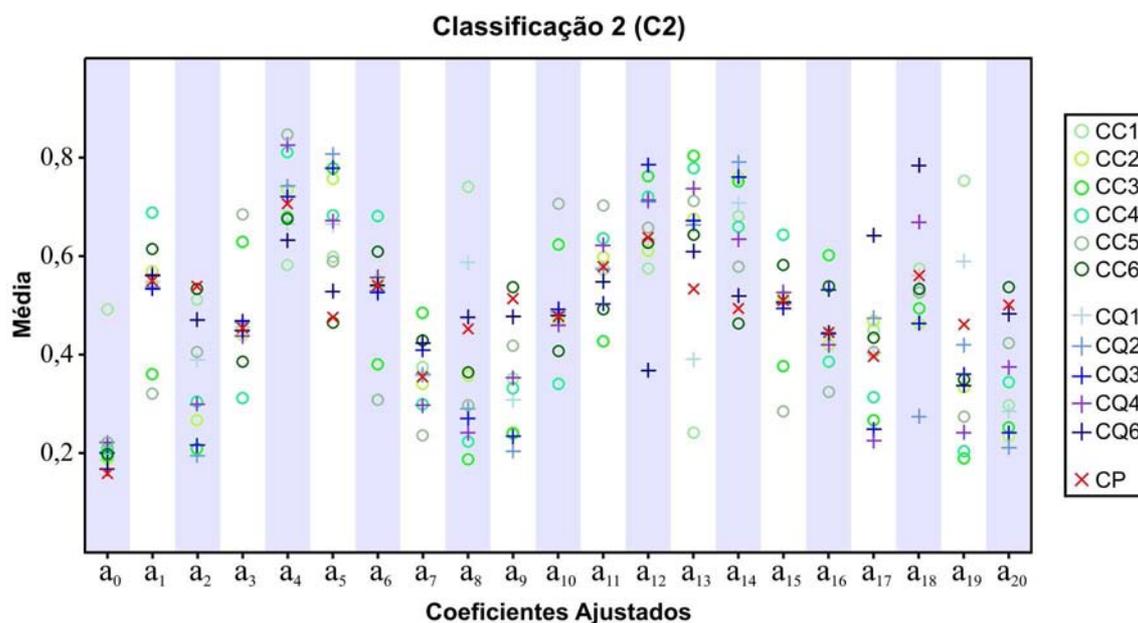


Figura 4.12 – Comportamento médio dos coeficientes ajustados (das 200 amostras de treinamento) para as 12 classes treinadas na Classificação 2.

Na Figura 4.12, as classes semelhantes (como as classes CC1, CC2, CC3, ..., CC6, por exemplo) possuem a mesma simbologia e cor, variando apenas em tonalidades. É importante lembrar novamente que não há problemas se houver confusão entre essas classes semelhantes, uma vez que todas serão agrupadas para gerar o “mapa geral” com as três classes de interesse.

4.2.2.2 Mapa temático

A classificação MaxVer permitiu a geração do “mapa detalhado” da C2 (Figura 4.13a), contendo as classes CC1 a CC6 para Cana Crua, CQ1 a CQ6 (exceto a CQ5) para Cana Queima, e Cana-em-Pé. As classes semelhantes do mapa detalhado foram agrupadas gerando o mapa geral (Figura 4.13b) da C2 com as classes Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP) que, assim como o mapa geral da C1, apresentou padrões bem definidos para os talhões.

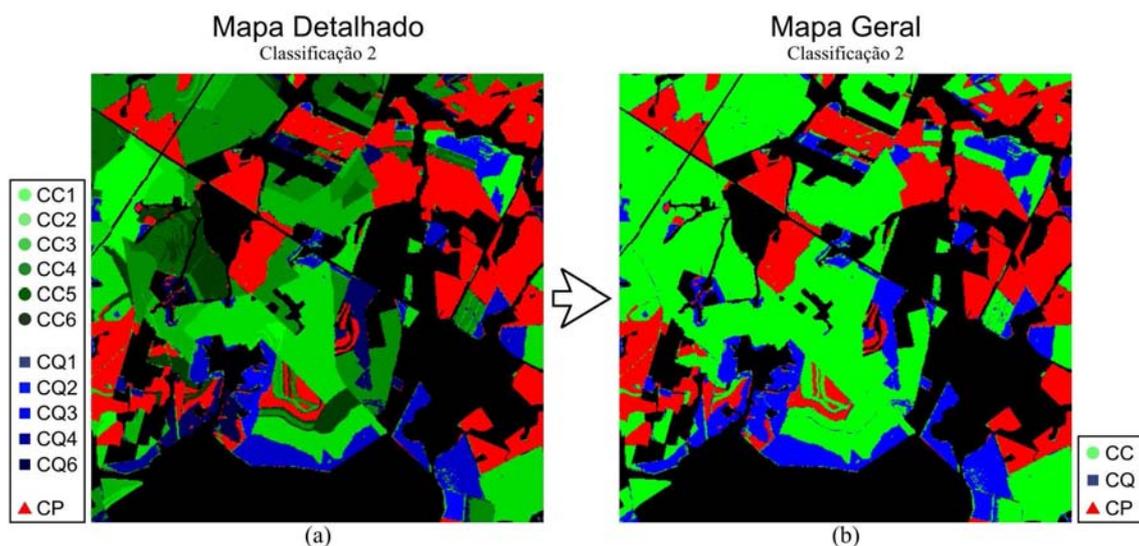


Figura 4.13 – a) Mapa detalhado contendo as 12 classes da Classificação 2: Cana Crua representadas em verde (variando de verde mais claro para *CC1* a verde mais escuro para *CC6*); Cana Queima representadas em azul (azul claro para *CQ1* até azul escuro para *CQ6*); e Cana-em-Pé, em vermelho; e b) Mapa geral, onde as 12 classes foram agrupadas nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).

4.2.2.3 Análise dos resultados

Na C2, assim como na C1, foram amostrados 300 *pixels* de cada classe, sendo 200 para o treinamento e 100 para a validação do treinamento. Dos 1.200 *pixels* testados (100 *pixels* em cada uma das 12 classes) o classificador errou em apenas 30, resultando em uma Exatidão Global (*EG*) de 97,5%. Desconsiderando-se os erros entre classes semelhantes, esse número de *pixels* errados diminuiu para 16, aumentando a *EG* do treinamento para 98,7%. Percebe-se que a diferença entre a precisão do treinamento (representado pela *EG*) da C2 em relação a C1 foi de apenas 0,3%.

Assim como para a C1, a análise quantitativa da precisão do mapa geral resultante da C2 foi realizada tomando como referência o mapa obtido por interpretação visual. Foram sorteadas, de maneira aleatória e independente, 1.500 amostras da C2, sendo 500 para cada uma das três classes, gerando assim a matriz de confusão da C2 (Tabela 4.3), juntamente com os valores dos índices de exatidão derivados dessa matriz.

Tabela 4.3 – Matriz de confusão da Classificação 2, tendo como referência o mapa de interpretação visual; e valores dos índices *EG* e κ :

C2	Referência			TOTAL	Índices (C2)	
	CC	CQ	CP		<i>EG</i>	<i>Kappa</i>
CC	478	64	18	560	90,6%	
CQ	15	416	17	448		0,859
CP	7	20	465	492		
TOTAL	500	500	500	1.500		

Percebe-se que as diferenças nos valores dos índices são pequenas em relação à C1. A maior confusão também aconteceu em erros de omissão à classe CQ, quando o classificador classificou de maneira divergente à referência 12,8% (64/500) e 4,0% (20/500) do total de *pixels* da classe CQ como pertencentes às classes CC e CP, respectivamente.

A análise qualitativa dos resultados é feita pelo estudo do mapa de discordâncias da C2 (Figura 4.14), onde são representados os *pixels* que o classificador rotulou de maneira diferente do intérprete (mapa de referência).

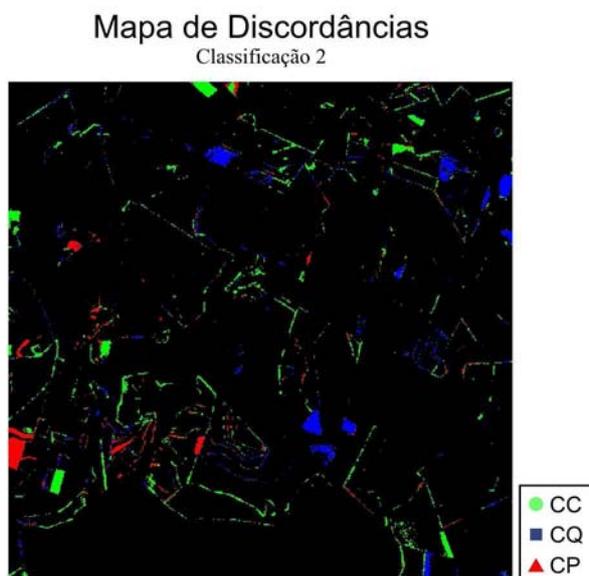


Figura 4.14 – Mapa de discordâncias entre o mapa geral da Classificação 2 e o mapa de referência.

Do total de 10.648 *pixels* classificados de maneira divergente pela C2, 48,0% (5.114) foram classificados como Cana Crua, 32,1% (3.421) como Cana Queima, e 19,9% (2.113) como Cana-em-Pé, quando na referência pertenciam a outra classe. Também na C2 as divergências se concentram nas bordas, sendo que as diferenças para o mapa de

discordâncias da C1 são praticamente insignificantes e sutis, de maneira que a C2 tendeu a “errar” um pouco menos nos talhões e um pouco mais nas bordas, em relação a C2. Assim, as mesmas discussões feitas a respeito da C1 também valem para a C2.

De maneira geral a C2 apresentou resultados bastante similares à C1, indicando que o novo método de distribuição dos pontos de controle no eixo do comprimento de onda, quando utilizado juntamente com o re-escalamento dos coeficientes ajustados, não produz resultados expressivamente diferentes em relação à metodologia tradicional STRS. Mello et al. (2009a) recomendaram avaliar o método de distribuição equidistante dos pontos, suspeitando que esse método melhoraria os resultados da classificação. Entretanto, essa melhora não foi observada no presente trabalho.

4.2.3 Classificação 3 (C3)

Com o intuito de investigar os efeitos do re-escalamento dos coeficientes ajustados sobre os resultados, procedeu-se à Classificação 3 (C3). Essa classificação, assim como a C2, diferencia-se da C1 na disposição dos pontos de controle utilizados na interpolação das superfícies espectro-temporais. A C3 utiliza o novo método de distribuição dos valores dos pontos de controle no eixo do comprimento de onda, em que esses valores são dispostos equidistantemente no intervalo [0,1]. Entretanto, na C3, os coeficientes ajustados na interpolação das superfícies não são re-escalados, sendo esta a grande diferença entre a C3 e as classificações anteriormente apresentadas (C1 e C2).

Na C3, assim como nas classificações anteriores, foram utilizadas seis imagens (datas) e seis bandas espectrais do sensor Landsat 5-TM. Para cada uma das 12 classes treinadas foram coletadas amostras de 300 *pixels*, separadas em 200 para o treinamento e 100 para a validação do treinamento.

4.2.3.1 Superfícies

A Figura 4.15 apresenta o aspecto final de 12 superfícies ajustadas para as 12 diferentes classes: Cana Crua (CC1, CC2, CC3, CC4, CC5 e CC6), Cana Queima (CQ1, CQ2, CQ3, CQ4 e CQ6 – lembrando mais uma vez que a classe Cana Queima imagem 5

(CQ5) não apresentou quantidade de *pixels* suficiente para a amostragem) e Cana-em-Pé (CP), respectivamente.

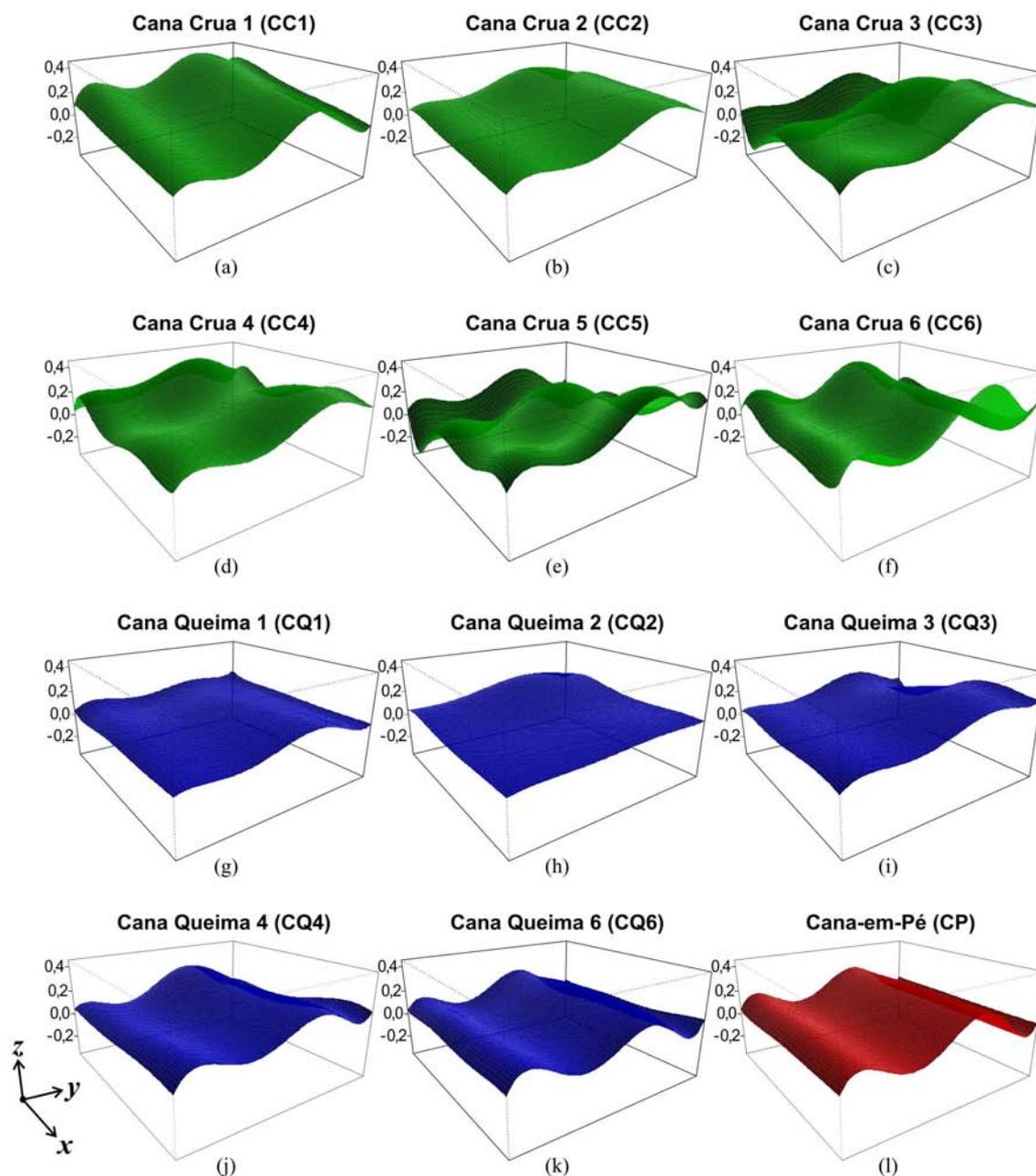


Figura 4.15 – Amostras de superfícies finais, utilizadas na Classificação 3 (C3), representativas das classes: a) Cana Crua (colhida na data 1 – CC1); b) CC2; c) CC3; d) CC4; e) CC5; f) CC6; g) Cana Queima (data 1 – CQ1); h) CQ2; i) CQ3; j) CQ4; k) CQ6; e l) Cana-em-Pé (CP).

Muito diferentemente das classificações anteriores (C1 e C2) a C3 apresentou superfícies bem distintas entre si para as 12 classes. Isso é um indício de que o re-

escalonamento dos coeficientes ajustados, de certa forma, uniformiza as superfícies. É importante observar, contudo, a escala do eixo das reflectâncias. Nas classificações C1 e C2 esse eixo estava no intervalo [0;1] e na C3 o intervalo foi bem menor [-0,2;0,4]. Isso induz a pensar que, se observadas sobre a mesma escala da C3, as superfícies da C1 e C2 podem apresentar também grandes diferenças, o que explicaria os bons resultados de precisão das classificações C1 e C2 mesmo quando suas superfícies são visualmente muito semelhantes entre si.

Investigando o comportamento médio dos 21 coeficientes ajustados para as 12 classes treinadas (Figura 4.16), percebe-se que o intervalo [-100;100], dessa vez, é muito maior do que quando ocorre o re-escalonamento dos coeficientes ajustados.

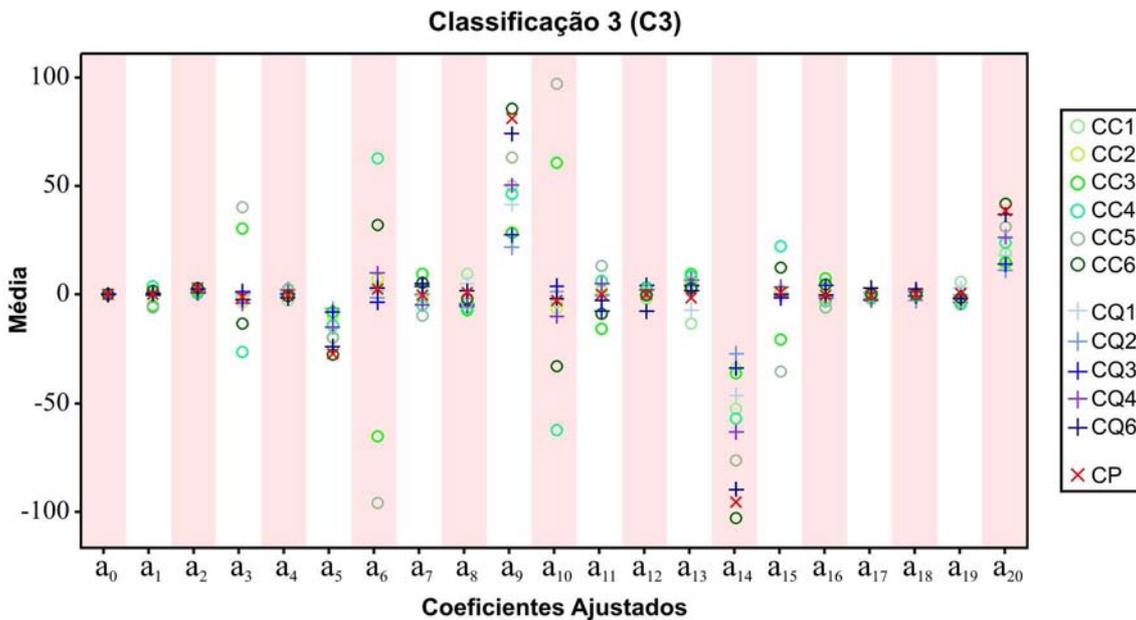


Figura 4.16 – Comportamento médio dos coeficientes ajustados para as 12 classes treinadas na C3.

É notável a diferença que o re-escalonamento provoca, quando comparamos os diagramas de dispersão das duas classificações anteriores (Figuras 4.8 e 4.12) com a Classificação 3 (Figura 4.16). Nesse caso (Figura 4.16), a escala do gráfico da C3 induz a pensar que alguns coeficientes, como por exemplo o a_8 , são mais díspares na C1 e C2 do que na C3. Só para se ter uma idéia do quanto a escala gráfica pode confundir a interpretação visual, os valores do coeficiente a_8 das classificações C1, C2 e C3, para cada classe, são apresentados na Tabela 4.4.

Tabela 4.4 – Valores médios do coeficiente a_8 , em cada uma das 12 classes, para C1, C2 e C3.

Classe	C1	C2	C3
CC1	0,7855929	0,7393831	9,7086647
CC2	0,5075959	0,3573736	-1,9840430
CC3	0,3440935	0,1889533	-7,1391203
CC4	0,3103660	0,2243333	-6,0561948
CC5	0,3372772	0,2987647	-3,7779669
CC6	0,3759508	0,3635018	-1,7964675
CQ1	0,6533906	0,5869793	5,0438259
CQ2	0,4389054	0,2900147	-4,0457911
CQ3	0,3341223	0,2711034	-4,6246353
CQ4	0,1674628	0,2415494	-5,5292371
CQ6	0,3020798	0,4762936	1,6559117
CP	0,4802837	0,4516875	0,9027585

De posse da Tabela 4.4, é possível perceber que os valores de a_8 para C3 não são menos díspares do que para C1 ou C2. É importante lembrar que, uma vez que as classes semelhantes serão agrupadas no mapa final (mapa geral), não existe nenhum problema no fato de, por apresentarem valores semelhantes para o coeficiente a_8 , haver uma possível confusão entre as classes CQ2 (-4,0457911) e CQ3 (-4,6246353) para a C3, por exemplo.

4.2.3.2 Mapa temático

O mapa detalhado resultante da classificação C3 (Figura 4.17a) contém as mesmas 12 classes das duas classificações anteriores (C1 e C2). Esse mapa é praticamente idêntico ao mapa geral da C2, com exceção de uns raríssimos *pixels*. Isso é um forte indício de que o re-escalamento dos coeficientes (única diferença entre as classificações C2 e C3) provavelmente não afeta a capacidade de distinção entre as classes pelo classificador MaxVer. É provável que, como o re-escalamento altera apenas a escala dos coeficientes ajustados, ele não interfira na habilidade do MaxVer em distinguir as classes, pois preserva a mesma inter-relação dos elementos do vetor de atributos do MaxVer (apenas com diferenças de escala). Os raríssimos *pixels* que apresentaram diferença entre a C3 e a C2 o fizeram, provavelmente, apenas por questão de arredondamento.

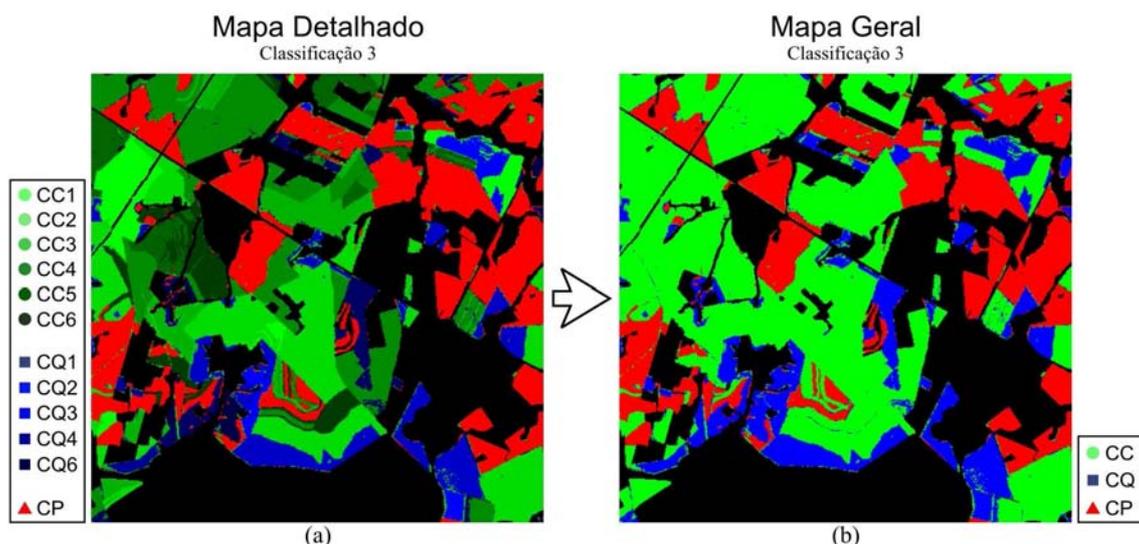


Figura 4.17 – a) Mapa detalhado contendo as 12 classes da Classificação 3: Cana Crua representadas em verde (variando de verde mais claro para *CC1* a verde mais escuro para *CC6*); Cana Queima representadas em azul (azul claro para *CQ1* até azul escuro para *CQ6*); e Cana-em-Pé, em vermelho; e b) Mapa geral, onde as 12 classes foram agrupadas nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).

4.2.3.3 Análise dos resultados

Também para a C3 foram utilizadas 300 amostras para cada classe, sendo 200 para o treinamento e 100 para a validação do treinamento. Dos 1.200 *pixels* testados na validação (100 *pixels* em cada uma das 12 classes), o classificador errou em apenas 29, resultando em uma Exatidão Global (*EG*) de 97,6%. Desconsiderando-se os erros entre classes semelhantes, esse número de *pixels* errados diminuiu para 15, aumentando a *EG* do treinamento para 98,8%. Esse resultado mostra que a diferença entre as precisões dos treinamentos (representados pela *EG*) das classificações C1, C2 e C3 foi muito pequeno.

Procedendo como para as classificações anteriores, a análise quantitativa da precisão do mapa geral resultante da Classificação 3 foi realizada com base no mapa de referência. Foram sorteadas 1.500 amostras, sendo 500 para cada uma das três classes. A Tabela 4.5 apresenta a matriz de confusão para a C3, assim como os índices de exatidão derivados da mesma.

Tabela 4.5 – Matriz de confusão da Classificação 3 e valores dos índices *EG* e κ

C3	Referência			TOTAL	Índices (C3)	
	CC	CQ	CP		<i>EG</i>	<i>Kappa</i>
CC	480	66	24	570	91,0%	
CQ	9	420	11	440		0,865
CP	11	14	465	490		
<i>TOTAL</i>	<i>500</i>	<i>500</i>	<i>500</i>	<i>1.500</i>		

As diferenças dos valores dos índices são mínimas em relação às classificações anteriores, principalmente em relação à C2. Na C3, assim como na C1 e C2, a maior confusão manteve-se na omissão à classe CQ.

A análise qualitativa da precisão da classificação, assim como para as classificações anteriores, foi realizada com o mapa de discordâncias da C3 (Figura 4.18).

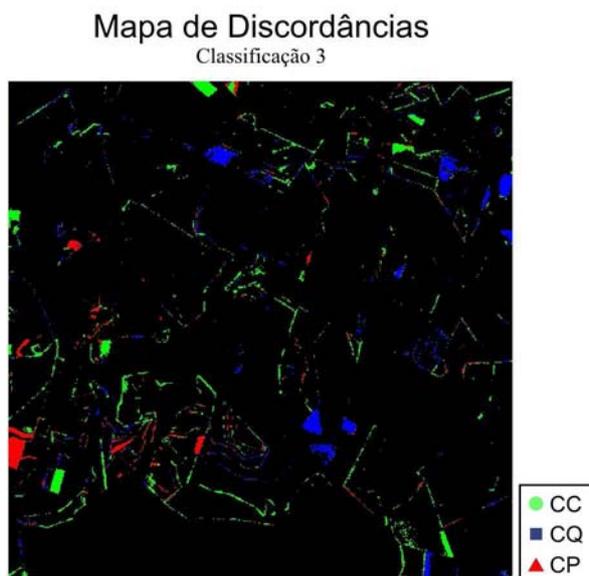


Figura 4.18 – Mapa de discordâncias entre o mapa geral da Classificação 3 e o mapa de referência.

Assim como para o mapa geral, o mapa de divergências da C3 foi praticamente idêntico ao da C2. A C3 classificou de maneira divergente um total de 10.642 *pixels*, o que representa apenas oito a menos que a classificação C2. Desse total, 47,5% (5.050) foram classificados como Cana Crua, 32,5% (3.467) como Cana Queima, e 20,0% (2.125) como Cana-em-Pé, quando na referência pertenciam a outra classe.

Os resultados da C3 indicaram que, mesmo quando os coeficientes ajustados das superfícies espectro-temporais não são re-escalados, o novo método de distribuição

dos pontos de controle no eixo do comprimento de onda não apresenta resultados significativamente diferentes da metodologia tradicional STRS. Analisando os resultados até aqui expostos, percebe-se que as variações testadas da metodologia STRS não apresentaram diferenças expressivas nos resultados, quanto a precisão dos mapas temáticos gerados pela classificação. A nova metodologia de disposição dos valores no eixo do comprimento de onda faz com que esse eixo perca seu real significado físico, deixando de representar (em qualquer escala que seja) os reais valores de comprimento de onda das bandas do sensor utilizado. Isso caracteriza uma desvantagem em relação à metodologia tradicional, que segundo Vieira (2000), permite ser utilizada com diversos sensores ao mesmo tempo (multi-sensor). Além disso, os métodos aplicados nas classificações C2 e C3 diminuem a robustez do classificador, uma vez que introduzem distorções no sistema. Isso poderia ser verificado caso a mesma metodologia fosse aplicada em outra safra (ano) ou utilizando outro sensor.

4.2.4 Classificação 4 (C4)

Como exposto anteriormente, parte das discordâncias entre o mapa geral, gerado pelas classificações, e o mapa de referência, pode ser explicada com o seguinte fato: na interpretação visual, as classes de CC ou CQ são rotuladas de acordo com a primeira percepção visual do intérprete que seja diferente da cana em pé; além disso, ainda acontecem casos em que o talhão é colhido logo após uma passagem do satélite, fazendo com que a defasagem entre a data da colheita e a data da próxima imagem seja suficientemente grande para descaracterizar a cicatriz do modo de colheita.

A Classificação 4 (C4) foi realizada na tentativa de minimizar o problema exposto. Trata-se da aplicação de três classificações STRS convencionais (método tradicional de disposição dos valores de comprimento de onda e re-escalamento dos coeficientes ajustados) realizadas em etapas: A primeira (C4-1) utilizando as quatro primeiras imagens adquiridas, no início da colheita; a segunda (C4-2) utilizando as quatro primeiras, mais uma nova imagem, e não re-classificando os *pixels* já rotulados como cana colhida (CC e CQ) na classificação anterior; e, por fim, a terceira (C4-3) utilizando as seis imagens e não re-classificando os *pixels* já rotulados como cana colhida nas duas classificações anteriores.

Na C4-1 foram treinadas 9 classes, sendo elas: CC1, CC2, CC3, CC4, CQ1, CQ2, CQ3, CQ4, e CP (nessa etapa, essa classe representa a cana cuja colheita não foi identificada nas imagens analisadas, mesmo que o talhão seja colhido nas imagens subsequentes). Na C4-2 foram treinadas 10 classes, sendo as mesmas nove classes da C4-1 mais a classe CC5 (lembrando mais uma vez que a classe CQ5 não aparece nas classificações por não apresentar quantidade de *pixels* suficiente para a amostragem). E, por fim, na C4-3 foram treinadas as mesmas 12 classes das classificações C1, C2 e C3.

Em todas as etapas da C4 (C4-1, C4-2 e C4-3), foram utilizadas seis bandas espectrais do sensor Landsat-5/TM e, para cada uma das classes em cada etapa, foram coletadas, aleatória e independentemente, amostras de 300 *pixels*, separadas em 200 para o treinamento e 100 para a validação.

4.2.4.1 Superfícies

A Figura 4.19 apresenta o aspecto final de nove superfícies ajustadas para as nove classes treinadas na etapa C4-1.

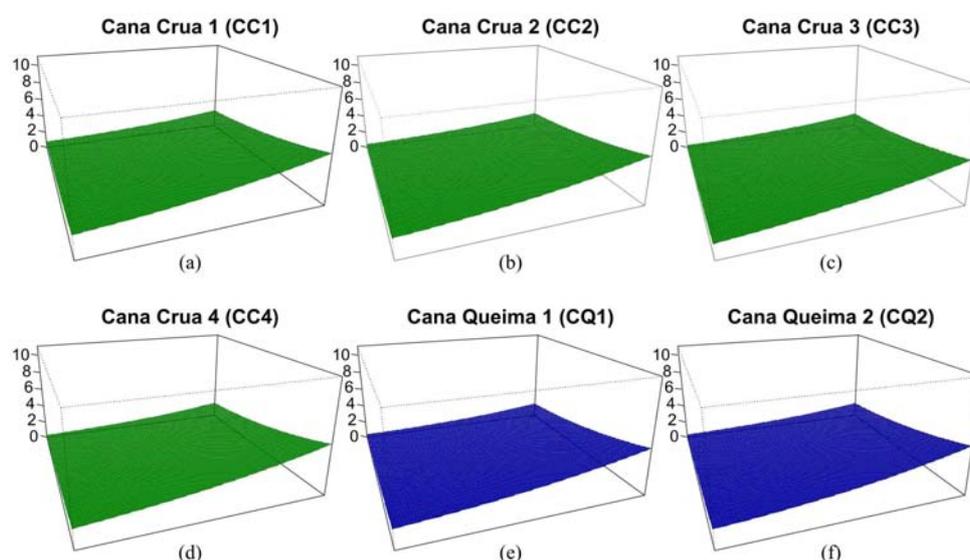


Figura 4.19 – Amostras de superfícies finais, utilizadas na primeira etapa da Classificação 4 (C4-1), representativas das classes: a) Cana Crua (colhida na data 1 – CC1); b) CC2; c) CC3; d) CC4; e) Cana Queima (colhida na data 1 – CQ1); f) CQ2; g) CQ3; h) CQ4; e i) Cana-em-Pé (CP). (*) *pode ser colhida nas datas posteriores.*

(continua)

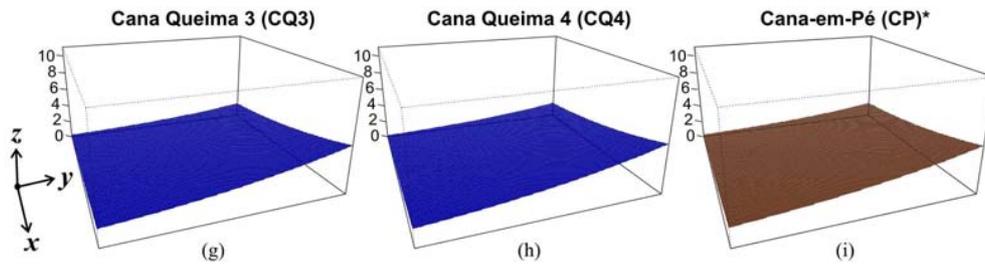


Figura 4.19 – Conclusão.

A Figura 4.20 apresenta o aspecto final de 10 superfícies ajustadas na etapa C4-2.

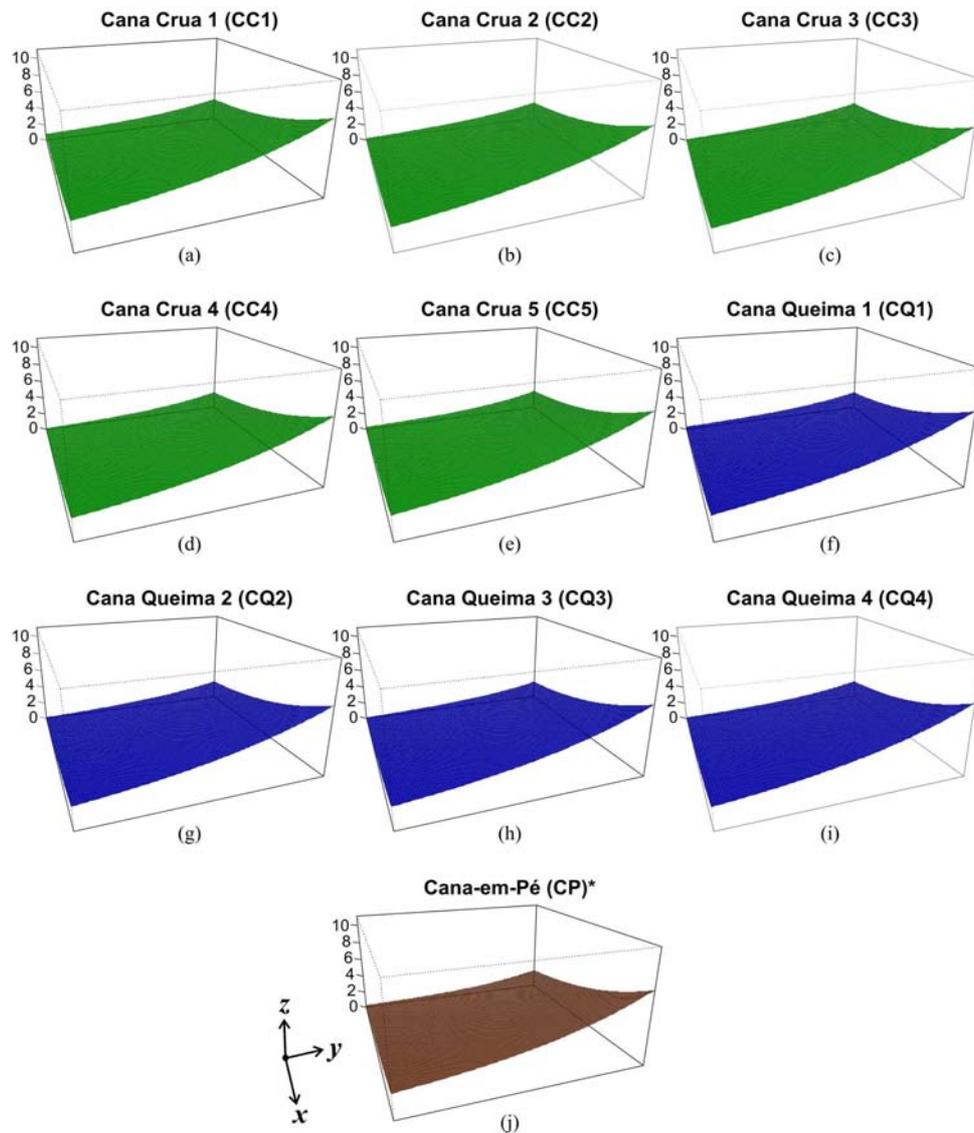


Figura 4.20 – Amostras de superfícies finais, utilizadas na segunda etapa da Classificação 4 (C4-2), representativas das classes: a) CC1; b) CC2; c) CC3; d) CC4; e) CC5; f) CQ1; g) CQ2; h) CQ3; i) CQ4 e j) CP. (*) pode ser colhida nas datas posteriores.

Percebe-se que a forma das superfícies é praticamente a mesma apresentada na C1 e C2 (Figuras 4.7 e 4.11, respectivamente), com a diferença de que na C4-1 (Figura 4.19) elas possuem valores mais baixos para a reflectância quando x e y são máximos. Na C4-2 (Figura 4.20) as superfícies mantêm o padrão, elevando-se mais um pouco nas proximidades da região $x = 1$ e $y = 1$. Por fim na C4-3 o aspecto das superfícies é idêntico ao apresentado na C1 (Figura 4.7). Esse fato provavelmente está ligado ao grau do polinômio utilizado. Como na C4-1 foram utilizadas apenas quatro imagens, o grau do polinômio se restringe ao valor 3. Assim também, na C4-2, etapa em que as superfícies foram interpoladas utilizando polinômios de quarto grau, uma vez que o processo se deu com cinco imagens.

4.2.4.2 Mapa temático

Cada etapa da C4 gerou um mapa detalhado, que foram agrupados resultando no mapa geral da C4 (Figura 4.21d) contendo as três classes de interesse.

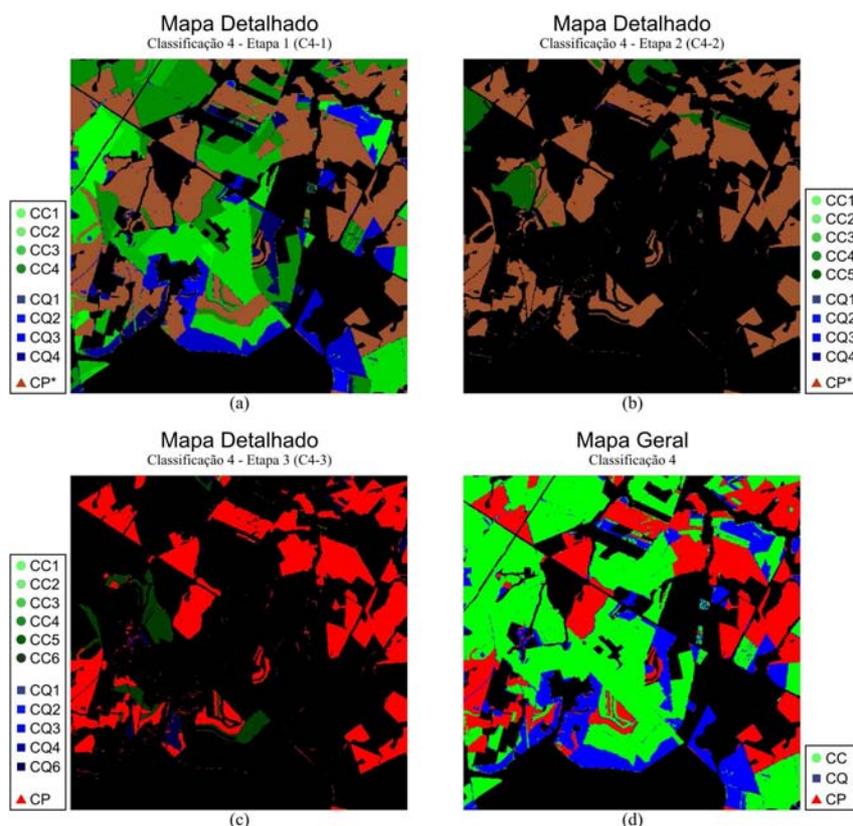


Figura 4.21 – a) Mapa detalhado das três etapas da Classificação 4: a) C4-1, b) C4-2, e c) C4-3; e d) mapa geral onde os resultados das três etapas foram agrupados nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP). (*) *pode ser colhida em datas futuras.*

O mapa geral da C4, grosso modo, apresenta semelhança aos mapas resultantes das classificações anteriormente vistas (C1, C2 e C3), com feições bem delimitadas e poucos *pixels* classificados de maneira isolada.

4.2.4.3 Análise dos resultados

Em todas as etapas da C4 também foram coletadas, de maneira aleatória e independente, 300 amostras para cada classe, sendo 200 para o treinamento e 100 para a validação do treinamento. Na C4-1, os valores das *EGs* foram de 97,9% (19/900) e 98,8% (11/900), considerando e não considerando os *pixels* com erro em classes semelhantes, respectivamente. Para a C4-2 os valores das *EGs* foram 96,8% (32/1.000) e 97,4% (26/1.000), respectivamente. E, por fim, para a última etapa (C4-3) os valores das *EGs* foram 97,8% (26/1.200) e 98,4% (19/1.200). Os resultados dos testes sobre os treinamentos mostram que os mesmos procederam-se válidos.

Assim como feito para as classificações anteriores, a análise quantitativa da precisão do mapa geral resultante da Classificação 4 foi realizada com base no mapa de referência. Foram selecionadas, de forma aleatória e independente, 1.500 amostras, sendo 500 para cada uma das três classes. A Tabela 4.6 apresenta a matriz de confusão para a C4 e os valores dos índices de exatidão derivados da mesma.

Tabela 4.6 – Matriz de confusão da Classificação 4 e valores dos índices *EG* e κ

C4	Referência			TOTAL	Índices (C4)	
	CC	CQ	CP		<i>EG</i>	<i>Kappa</i>
CC	474	21	17	512		
CQ	18	465	20	503		
CP	8	14	463	485		
<i>TOTAL</i>	500	500	500	1.500		

Os valores dos índices de exatidão, em geral, são sutilmente maiores que os das classificações anteriores (C1, C2 e C3), com o valor do *Kappa* maior que 0,9. Dando continuidade à análise de qualidade da C4 gerou-se o mapa de discordâncias da C4 (Figura 4.22).

Mapa de Discordâncias Classificação 4

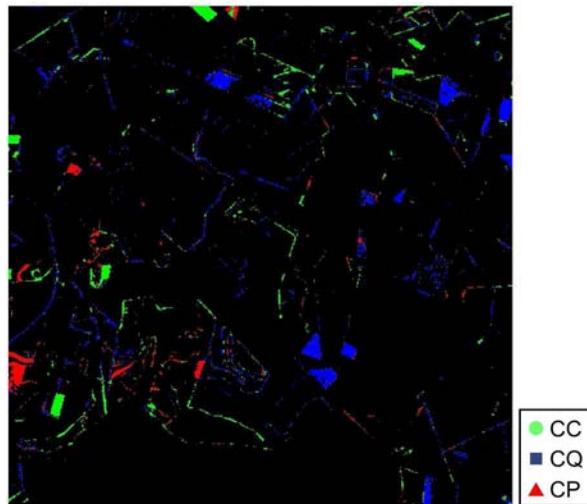


Figura 4.22 – Mapa de discordâncias entre o mapa geral da Classificação 4 e o mapa de referência.

Analisando o mapa de discordâncias da C4 (Figura 4.22) percebe-se um fato interessante: as bordas, normalmente classificadas equivocadamente como CC nas classificações anteriores, o foram predominantemente como CQ na C4, justificando o fato de, dessa vez, a CQ ser a classe com maior quantidade de discordâncias, totalizando 51,6% (5.724 em 11.096 *pixels*) das mesmas. A classe CC, por sua vez, contribuiu com 30,1% (3.338 *pixels*), enquanto a CP representou apenas 18,3% (2.034 *pixels*) das discordâncias.

Com base na inspeção visual dos mapas de discordâncias das quatro classificações até o momento apresentadas, é possível proceder algumas investigações acerca de alguns talhões discordantes entre as classificações e a referência (interpretação visual). Podem ser observados, no canto inferior esquerdo dos mapas de discordâncias (Figura 4.22, por exemplo), dois polígonos bem definidos: um polígono vermelho, representando um talhão que foi classificado como CP, de maneira divergente à referência, que indica que aquela área é da classe CC; e um polígono verde, representando um talhão erroneamente classificado como CC, quando na referência ele é mostrado como CQ.

Uma investigação desses dois talhões foi realizada visando expor duas idéias: a primeira (Análise 1) é que a cicatriz do modo de colheita pode ser descaracterizada pelos intervalos de tempo decorridos entre as passagens do satélite; e a segunda (Análise 2)

mostra que, apesar de utilizado como referência, o mapa de interpretação visual não pode ser tomado como verdade absoluta, e pode conter alguns erros. Veja a Figura 4.23.

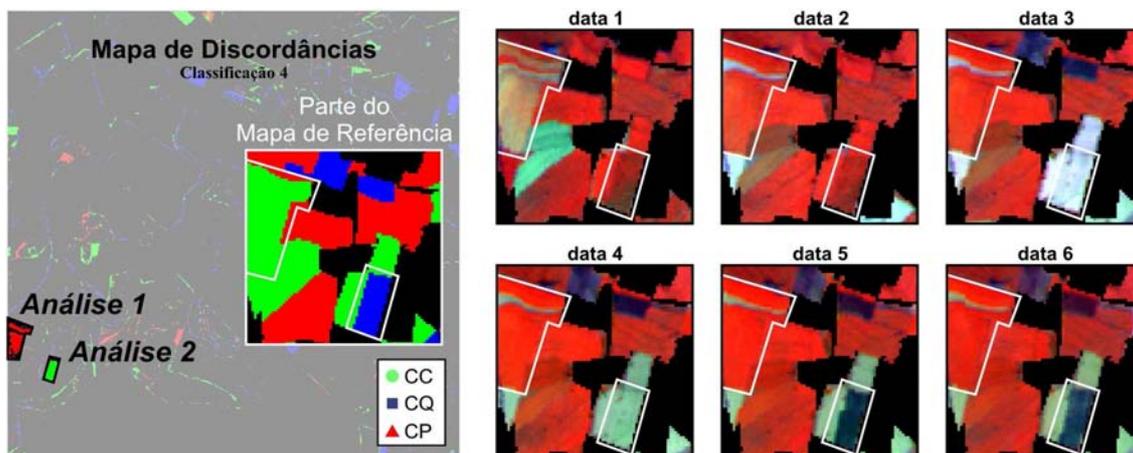


Figura 4.23 – Investigação de dois talhões classificados de maneira discordante à referência nas classificações apresentadas até o momento (C1, C2, C3 e C4) Imagens em composição R4 G5 B3.

Analisando a Figura 4.23 é possível fazer investigações em cada análise proposta nos dois talhões supracitados, que foram classificados de forma discordante à referência. Observando as imagens montadas em composição falsa-cor (R4 G5 B3) nas diversas datas, pode-se perceber que o talhão da Análise 1 é identificado visualmente como colhido sem queima na primeira data. Entretanto, a mesma imagem já indica que há uma rápida rebrota, que pode ser confirmada inspecionando a data 2. Em todas as imagens subsequentes o talhão se comporta espectralmente como cana em pé. Dessa forma o classificador “entende” que o talhão é mais verossímil a um talhão de CP. Essa análise demonstra claramente uma limitação da classificação automática em relação à interpretação visual.

Já no talhão da Análise 2 (Figura 4.23) percebe-se que há a colheita crua na data 3, o que faria com que o intérprete classificasse esse talhão como CC, uma vez que, como convencional no projeto Canasat, deve-se rotular um talhão colhido com a classe referente ao tipo de colheita de acordo com a primeira percepção diferente da cana-em-pé. Entretanto, o mapa de referência mostra que o intérprete cometeu um erro, assinalando esse talhão à classe CQ. Isso aconteceu, provavelmente, pelo fato de o intérprete ter visto o mesmo talhão com aspecto muito semelhante a talhões colhidos

com queima nas datas 5 e 6. Contudo, esse assunto é totalmente discutível do ponto de vista que, se o objetivo do mapeamento é identificar fontes de emissão de gases poluentes, por exemplo, então o importante é identificar a queima, independente se ela acontece antes ou depois da colheita. O que tornaria conceitual, e não operacional, o problema supracitado, e dessa forma, o intérprete é quem teria obtido o resultado correto.

4.2.5 Classificação 5 (C5)

A Classificação 5 (C5) é uma classificação multitemporal com base na aplicação da técnica de Modelo Linear de Mistura Espectral (MLME). Um modelo de mistura, como descrito na Equação 3.9, é aplicado em cada imagem da série temporal, gerando, para cada uma dessas imagens, três imagens-fração que representam as proporções dos componentes do modelo em cada *pixel*. Essas imagens-fração são, então, utilizadas como elementos do vetor de atributos do classificador MaxVer.

4.2.5.1 Aplicação do MLME

Para cada uma das seis imagens da série temporal, procedeu-se um MLME específico, que tinha como componentes do modelo as frações: Vegetação (V), Palha (P) e Cinza/Sombra (C). Ao final do processo, cada componente dá origem a uma imagem-fração. Assim, utilizando as seis imagens da série temporal, obteve-se, ao fim do processo, 18 imagens-fração, sendo três (fração V, fração P e fração C) para cada data.

Tomando como exemplo a data 2 (20/06/2007), demonstrar-se-á, a seguir, como procedeu a aplicação do MLME:

De posse da imagem da data 2 (6 bandas espectrais), efetuou-se a identificação dos *pixels* puros, conhecidos como *endmembers*, de cada uma das três componentes do modelo por meio da identificação em tela do aplicativo SPRING. A Figura 4.24 mostra o comportamento espectral dos *endmembers* selecionados para o MLME aplicado à data 2.

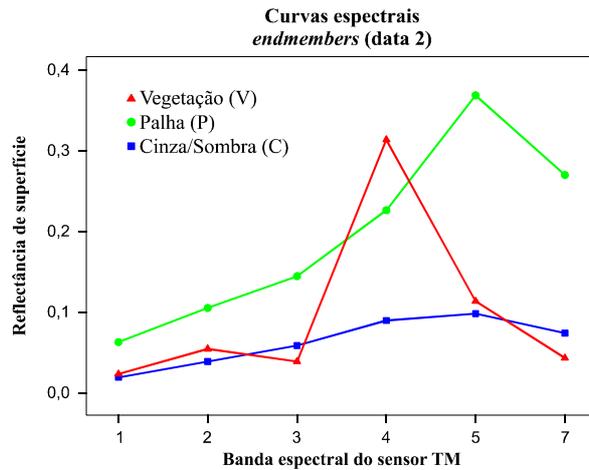


Figura 4.24 – Comportamento espectral dos *endmembers* selecionados para o MLME aplicado à segunda imagem da série temporal.

É possível perceber (Figura 4.24) que o comportamento espectral do *endmember* P é bastante distinto em relação aos demais. Entretanto, nota-se, inspecionando a Figura 4.24, que a diferença entre os *endmembers* de V e C só é consideravelmente grande na banda 4, que corresponde ao infra-vermelho próximo. A alta resposta da componente V na banda 4 está relacionada com a estrutura foliar da planta. Segundo Ponzoni e Shimabukuro (2007), os valores dos fatores de reflectância serão maiores quanto mais lacunosa for a estrutura interna foliar.

A aplicação do MLME para a data 2 teve como resultado três imagens-fração (Figura 4.25), cada uma correspondendo a um componente do modelo: V, P e C (Figura 4.25a, 4.25b e 4.25c, respectivamente).

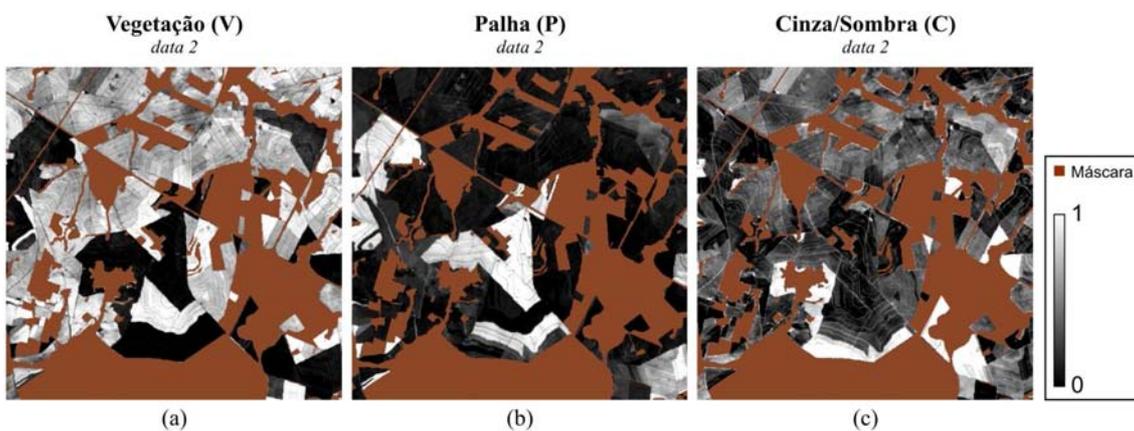


Figura 4.25 – Imagens-fração resultantes da aplicação do MLME na segunda imagem (data 2) da série temporal. Frações: a) Vegetação (V), b) Palha (P), e c) Cinza/Sombra (C).

Analisando a Figura 4.25 percebe-se que o MLME é útil em distinguir as feições, delimitando os talhões correspondentes a cada classe que compõe o modelo, apesar das frações V e C não se delimitarem tão bem quanto a P. Essa distinção não tão grande entre os componentes V e C ocorre pela própria estrutura do dossel da plantação de cana que, devido à disposição das folhas, produzem sombra e fazem com que a componente Palha/Sombra responda de maneira significativa em alguns locais onde predomina a presença de Vegetação.

De modo geral, a aplicação do MLME é útil no sentido de sintetizar os dados, reduzindo sua dimensionalidade, de seis bandas espectrais para três frações, em cada imagem da série temporal. Essa redução no vetor de atributos de 36 (6 datas x 6 bandas) para 18 (6 datas x 3 imagens-fração) elementos é válida na aplicação da classificação MaxVer no sentido de minimizar o problema conhecido como “maldição da dimensionalidade” ou “fenômeno de Hughes” (HUGHES, 1968; ABEND et al., 1969).

4.2.5.2 Mapa temático

A classificação MaxVer deu-se com as imagens-fração das seis datas como elementos do vetor de atributos. Por fim, o resultado foi um mapa temático específico (Figura 4.26a), contendo as 12 classes treinadas (da mesma maneira que para as classificações anteriores). Na C5 o mapa geral (Figura 4.26b) também foi obtido pelo agrupamento das classes semelhantes.

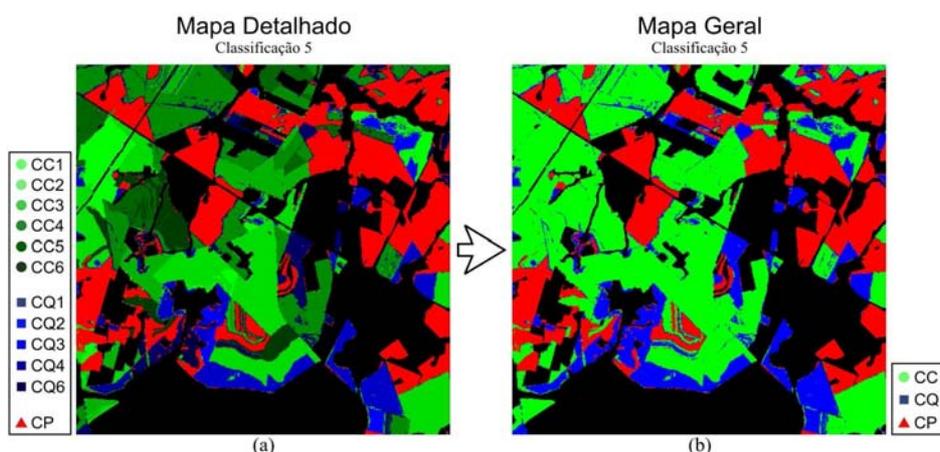


Figura 4.26 – a) Mapa detalhado contendo as 12 classes da Classificação 5: Cana Crua representadas em verde; Cana Queima representadas em azul; e Cana-em-Pé, em vermelho; e b) Mapa geral, onde as 12 classes foram agrupadas nas três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).

4.2.5.3 Análise dos resultados

Foram coletadas, assim como nas classificações anteriores, 1.500 amostras para a avaliação da C5, sendo 500 para cada uma das três classes. A Tabela 4.7 apresenta a matriz de confusão obtida para a C5 e os valores dos índices de exatidão desta classificação.

Tabela 4.7 – Matriz de confusão da Classificação 5 e valores dos índices *EG* e κ

C5	Referência			TOTAL	Índices (C5)	
	CC	CQ	CP		<i>EG</i>	<i>Kappa</i>
CC	453	56	5	514	89,6%	
CQ	37	414	18	469		0,844
CP	10	30	477	517		
<i>TOTAL</i>	<i>500</i>	<i>500</i>	<i>500</i>	<i>1.500</i>		

Os índices de precisão da C5 indicam que essa classificação foi bastante eficiente. Assim como para as classificações STRS (C1, C2, C3 e C4) o maior erro se concentrou na confusão entre as classes CC e CQ.

Cruzando as informações do mapa geral da C5 e do mapa de referência gerou-se o mapa de discordâncias da C5 (Figura 4.27), que posiciona espacialmente os *pixels* que o classificador classificou de maneira diferente à rotulada na referência.

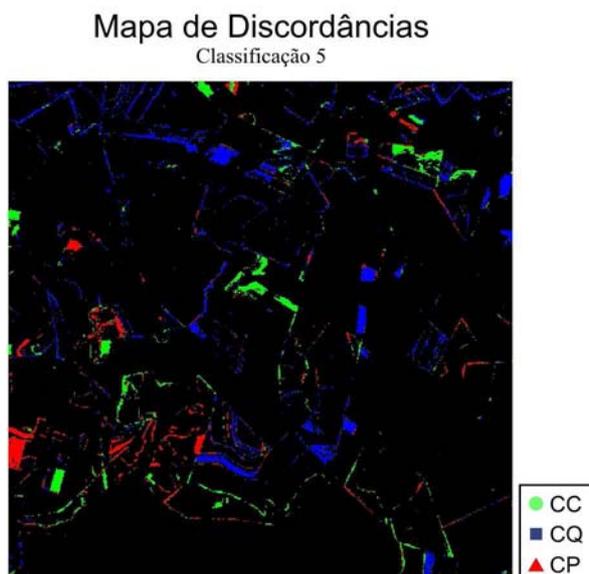


Figura 4.27 – Mapa de discordâncias entre o mapa geral da Classificação 5 e o mapa de referência.

O mapa de discordâncias da C5 (Figura 4.27) apresentou um total de 14.286 *pixels*. Sendo: 26,7% (3.809) da classe CC; 49,1% (7.014) da CQ; e 24,2% (3.463) da CP. A maior confusão ocorreu com os *pixels* que foram classificados como CQ e pertenciam à classe CC na referência (5.442).

Os resultados mostraram-se satisfatórios para a C5. Entretanto, os procedimentos de criação e aplicação dos MLMs, apesar de já implementados em aplicativos como, por exemplo, o SPRING, exigem conhecimentos específicos sobre o comportamento espectral dos *endmembers*, e demandam um tempo consideravelmente razoável por parte do usuário. Entretanto, uma vez que se usem sempre imagens corrigidas radiometricamente, pode-se padronizar os valores espectrais dos *endmembers* para toda imagem de um mesmo sensor (TOMPKINS et al., 1997).

4.2.6 Classificação 6 (C6)

Nessa classificação (C6), o vetor de atributos é composto por todas as seis bandas de todas as seis imagens da série temporal, sem nenhuma metodologia adicional. Todas as 36 bandas (6 bandas em cada uma das 6 imagens) foram utilizadas na classificação MaxVer, que se deu no aplicativo ENVI, versão 4.3 (RSI, 2006).

4.2.6.1 Mapa temático

Para a C6 foram treinadas, utilizando também 200 *pixels* por classe, as mesmas 12 classes das classificações anteriores. O resultado pode ser visto no mapa específico (Figura 4.28a), contendo as 12 classes, que após o agrupamento das classes semelhantes, resultou no mapa geral (Figura 4.28b) da C6.

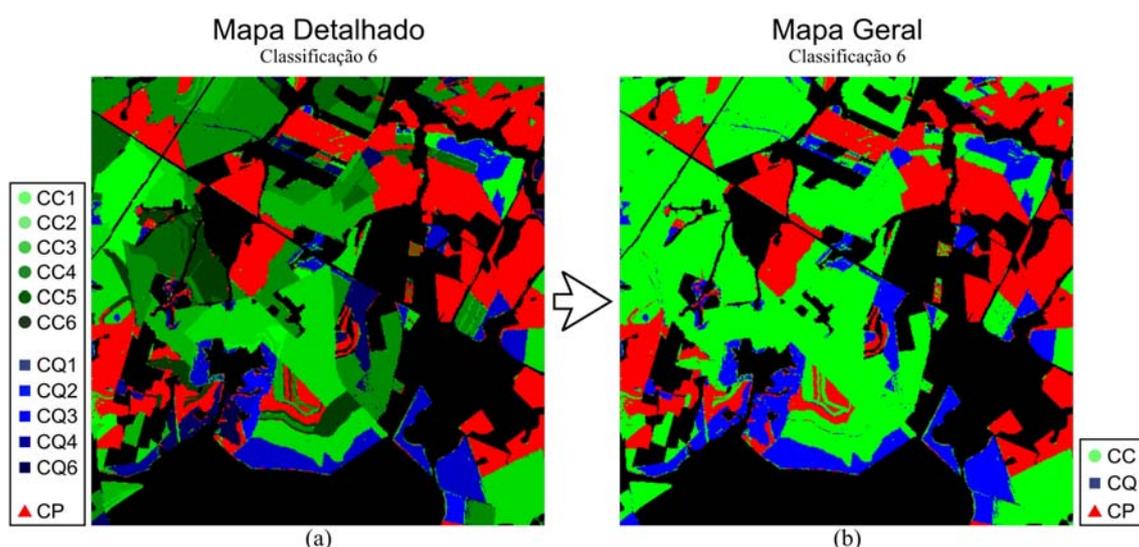


Figura 4.28 – a) Mapa detalhado contendo as 12 classes da Classificação 6: Cane Crua representadas em verde (variando de verde mais claro para *CC1* a verde mais escuro para *CC6*); Cane Queima representadas em azul (azul claro para *CQ1* até azul escuro para *CQ6*); e Cane-em-Pé, em vermelho; e b) Mapa geral, onde as 12 classes foram agrupadas nas três classes de interesse: Cane Crua (CC), Cane Queima (CQ) e Cane-em-Pé (CP).

Aparentemente o mapa geral da C6 (Figura 4.28b) apresentou menores quantidades de *pixels* dispersos em bordas, especialmente da classe CQ, do que na classificação utilizando MLME anteriormente apresentada (C5). Foi possível perceber que, assim como em todas as classificações anteriores (STRS: C1, C2, C3 e C4; e MLME: C5), o mapa temático apresenta feições bem delimitadas dos talhões.

4.2.6.2 Análise dos resultados

Assim como para as classificações anteriores, a análise quantitativa da precisão do mapa geral resultante da Classificação 6 foi realizada tomando como referência o mapa obtido por interpretação visual. Foram utilizadas 1.500 amostras, obtidas de forma aleatória e independente, sendo 500 para cada uma das três classes. O resultado pode ser visto na matriz de confusão para a C6 (Tabela 4.8), juntamente com os valores dos índices de precisão derivados dessa matriz.

Tabela 4.8 – Matriz de confusão da Classificação 6 e valores dos índices *EG* e κ .

C6	Referência			TOTAL	Índices (C6)	
	CC	CQ	CP		<i>EG</i>	<i>Kappa</i>
CC	482	63	19	564	90,9%	0,864
CQ	8	406	5	419		
CP	10	31	476	517		
TOTAL	500	500	500	1.500		

Os valores dos índices encontrados são sutilmente maiores que os da C5. O uso de um vetor de atributos tão grande como o da C6, com 36 elementos (6 imagens com 6 bandas cada imagem), aparentemente não influenciou negativamente nos resultados. Ao realizar uma classificação idêntica a C6, porém usando apenas as bandas 3, 4 e 5, o valor do índice *EG* sofre uma pequena, porém insignificante, redução (90,7%). Se a classificação for feita utilizando apenas as bandas 3 e 4, o índice *EG* reduz um pouco mais e atinge o valor 89,1%. Só quando a classificação é feita com o uso apenas da banda 4, é que o índice *EG* reduz significativamente (73,3%)¹.

Uma idéia da dispersão das discordâncias entre a C6 e o mapa de referência pode ser obtida com inspeção ao mapa de discordâncias da C6 (Figura 4.29).

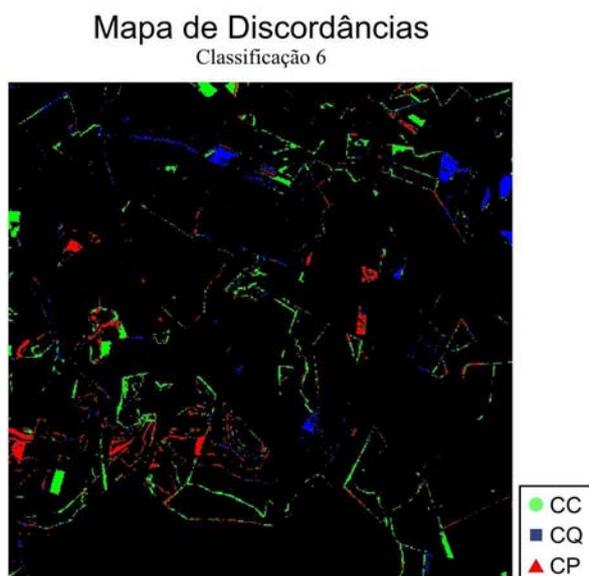


Figura 4.29 – Mapa de discordâncias entre o mapa geral da Classificação 6 e o mapa de referência.

O mapa de discordâncias da C6 (Figura 4.29) foi o que apresentou menor quantidade de *pixels*, totalizando 10.428 *pixels* classificados de maneira discordante à referência. Desse total, 48,6% (5.072) foram indevidamente atribuídos à classe CC, 22,4% (2.332) à CQ, e 29,0% à classe CP.

¹ A escolha da banda 4 se deu com base no estudo do comportamento espectro-temporal das classes de colheita da cana-de-açúcar, previamente apresentado na seção 4.1 (ver Tabela 4.1 e Figura 4.1), por ser a banda que apresenta valores médios de reflectância mais distintos entre as três classes: Cana Crua, Cana Queima e Cana-em-Pé. A escolha das bandas 3 e 4, assim como a escolha das bandas 3, 4 e 5, para as classificações com duas e três bandas, respectivamente, também se deu de acordo com o estudo apresentado na seção 4.1.

Os resultados da Classificação 6 mostram-se excelentes, equivalendo-se aos resultados apresentados nas classificações STRS e na classificação MLME. Contudo, a C6 não possui a robustez (principalmente em relação à classificação STRS tradicional – C1) das classificações anteriormente apresentadas. Essa metodologia utilizada na C6, caso as imagens apresentassem *pixels* com nuvens ou ruídos, provavelmente não obteria resultados tão expressivos em termos de precisão.

4.2.7 Classificação 7 (C7)

Apenas com o intuito de reafirmar a superioridade das classificações multitemporais sobre as classificações com uma única data, para estudos como este, realizou-se também uma classificação MaxVer simples, utilizando apenas as seis bandas da última imagem da série temporal, ou seja, a imagem de 08/09/2007 (data 6) (Figura 4.30). A última data foi escolhida por ser a data que, teoricamente, possui a informação mais completa sobre a colheita realizada no ano de estudo (2007).

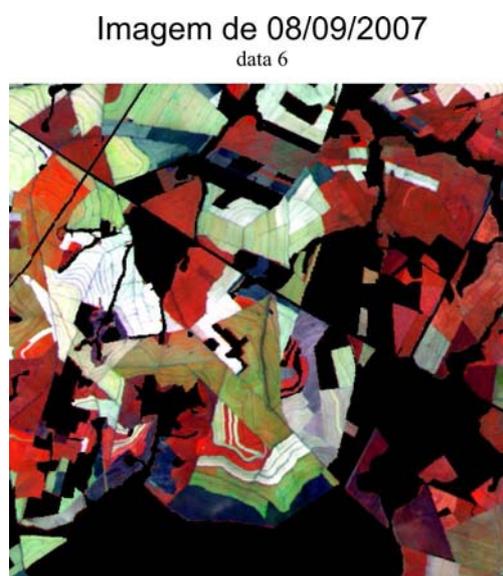


Figura 4.30 – Imagem de 08/09/2007 utilizada na Classificação 7. Composição falsa-cor R4 G5 B3.

4.2.7.1 Mapa temático

No treinamento das classes da C7, foram amostrados 200 *pixels* por classe. Entretanto, como se trata de uma única imagem, só foram utilizadas as amostras de: Cana Crua colhida na data 6 para representar a CC; Cana Queima colhida na data 6, para

representar a CC, e a classe CP. O resultado foi um mapa temático com as três classes de interesse (Figura 4.31).

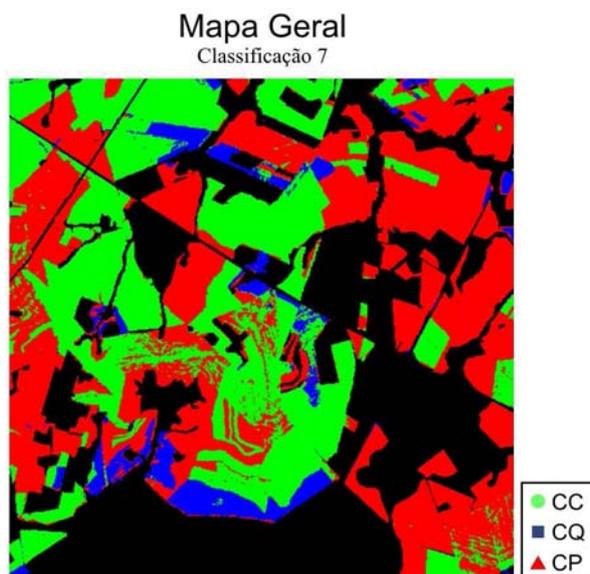


Figura 4.31 – Mapa geral da Classificação 7 contendo as três classes de interesse: Cana Crua (CC), Cana Queima (CQ) e Cana-em-Pé (CP).

É possível notar que o mapa da C7 (Figura 4.31) apresenta maior quantidade de *pixels* na classe CP. Isso pode ser justificado pelo fato de que, ao usar apenas a última imagem da série temporal, vários talhões colhidos mais no início do período de safra já foram replantados ou ocorreu rebrota, apresentando assim, comportamento espectral de cana-em-pé. Outro fato interessante de se mencionar é que, nesse tipo de classificação seria impossível classificar um talhão como CC se o mesmo tiver sido colhido sem o fogo, mas tivesse tido sua palha queimada após a colheita, por exemplo.

4.2.7.2 Análise dos resultados

Assim como para as seis classificações anteriormente apresentadas, a análise de precisão da Classificação 7 foi realizada tomando como referência o mapa obtido por interpretação visual. Foram utilizadas 1.500 amostras, sendo 500 para cada uma das três classes. O resultado é apresentado na matriz de confusão para a C7 (Tabela 4.9), juntamente com os valores dos índices de exatidão derivados da mesma.

Tabela 4.9 – Matriz de confusão da Classificação 7 e valores dos índices *EG* e κ

C7	Referência			TOTAL	Índices (C7)	
	CC	CQ	CP		<i>EG</i>	<i>Kappa</i>
CC	368	37	3	408	69,3%	
CQ	8	182	7	197		0,540
CP	124	281	490	895		
TOTAL	500	500	500	1.500		

Inspecionando a Tabela 4.9, percebe-se que os erros, em sua maioria absoluta, aconteceram como inclusão à classe CP, devido ao fato de que talhões colhidos no início do período da safra (nos meses de abril a julho) e re-plantados ou com rebrota apresentam padrão espectral semelhante à CP. O erro de inclusão à classe CQ associado com a omissão à CC acontece, em geral, pela limitação do classificador em identificar a colheita sem o uso do fogo em talhões onde a palha foi queimada ou ocorreu uma gradagem após a colheita. O mapa de discordâncias da C7 (Figura 4.32) ilustra espacialmente das divergências entre a C7 e o mapa de referência.

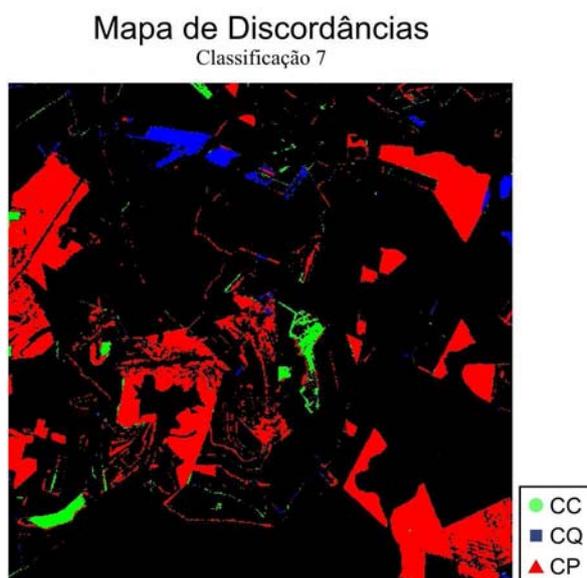


Figura 4.32 – Mapa de discordâncias entre o mapa geral da Classificação 7 e o mapa de referência.

O mapa de discordâncias da C7 (Figura 4.32) só fez confirmar o exposto, de que as classificações com única data são inviáveis de ser utilizadas em aplicações de sensoriamento remoto em mapeamentos agrícolas. Do total de 41.311 *pixels* classificados de maneira discordante à referência, 86,4% (35.698) foram indevidamente atribuídos à classe CP e 6,8% (2.822 e 2.791, respectivamente) às classes CQ e CC.

Em geral a C7 esteve bem aquém dos resultados das outras classificações desse estudo. Isso já era esperado, uma vez que é praticamente impossível monitorar, com apenas uma imagem (uma única data) um fenômeno que ocorre no tempo, como é o caso da colheita da cana-de-açúcar.

4.3 Comparação entre as sete classificações

Um total de sete classificações de máxima verossimilhança foi realizado nesse estudo visando avaliar sobre diferentes técnicas de classificação automática de imagens digitais. As quatro primeiras classificações (C1, C2, C3 e C4) utilizaram o procedimento de classificação por superfícies de resposta espectro-temporais (STRS). A Classificação 5 (C5) fez uso da técnica conhecida por Modelo Linear de Mistura Espectral (MLME). A C6 foi realizada utilizando todas as bandas de todas as imagens da série temporal como elementos do vetor de atributos do classificador. E por fim, a C7 foi realizada como uma classificação convencional, utilizando apenas uma única imagem: a última da série temporal, adquirida em 08/09/2007. Diante de tantos resultados, é extremamente válido realizar comparações diretas entre as classificações, na tentativa de identificar quais classificações se equivalem em termos de precisão.

Uma maneira de avaliar comparativamente o resultado das classificações é utilizando o teste estatístico Z , baseado nas Equações 3.10 e 3.11 (apresentadas na seção 3.9). O resultado do teste aplicado ao índice $Kappa$ é apresentado na Tabela 4.10. Na diagonal, valores significativos representam uma classificação significativamente melhor que uma classificação casual. Os valores fora da diagonal comparam os resultados das classificações, par a par, em que valores significativos indicam diferença estatística entre as duas classificações testadas.

Tabela 4.10 – Resultado da comparação entre as classificações pelo teste Z sobre o índice *Kappa*: Valores de *Kappa* (κ); variância do *Kappa* (V_{κ}); na triangular-superior estão os valores da estatística Z (Z_{calc}); na triangular-inferior os valores de *p-valor* dos testes (simétricos e associados ao Z_{calc}). Na diagonal, os valores de Z_{calc} estão na parte de cima, enquanto que o *p-valor* associado ao Z_{calc} está na parte inferior da célula. As células em fonte vermelha destacam os resultados significativos a 5% de significância.

	C1	C2	C3	C4	C5	C6	C7
κ	0,874	0,859	0,865	0,902	0,844	0,864	0,540
$V_{\kappa} \times 10^{-4}$	1,149787	1,271485	1,222098	0,915801	1,394475	1,226165	2,668623
C1	81,508475 ≈ 0	0,963983	0,584380	1,948212	1,880790	0,648756	17,092485
C2	0,335054	76,179408 ≈ 0	0,379961	2,907473	0,918680	0,316376	16,070769
C3	0,558965	0,703974	78,246161 ≈ 0	2,530510	1,298234	0,063910	16,476627
C4	0,051390	0,003644	0,011390	94,255339 ≈ 0	3,815892	2,596434	19,120485
C5	0,060000	0,358263	0,194207	0,000136	71,472193 ≈ 0	1,235453	15,081515
C6	0,516496	0,751717	0,949042	0,009420	0,216662	78,02598 ≈ 0	16,41735
C7	≈ 0	≈ 0	33,05599 ≈ 0				

Os valores apresentados na diagonal da Tabela 4.10 mostram que todas as sete classificações avaliadas nesta pesquisa apresentaram resultados significativamente melhores que uma classificação casual. Os valores fora da diagonal representam o resultado dos testes feito aos pares de classificações. Esse resultado, a 5% de significância, indica não haver diferenças significativas entre a classificação STRS tradicional (C1) e as demais classificações multitemporais (C2, C3, C4, C5 e C6). Os testes ainda indicam que a classificação STRS em etapas (C4) foi estatisticamente melhor que as demais classificações, exceto quando comparada à Classificação 1. Já a classificação C5 (que utiliza a técnica de MLME) apresentou-se significativamente igual à C6 e às classificações SRTS, exceto à C4, que foi superior a C5. Por fim, como esperado, a classificação de uma única imagem (C7) apresentou-se estatisticamente inferior a todas as outras (multitemporais) para a discriminação das classes de colheita da cana (CC e CQ) e cana-em-pé (CP).

Dados os altos valores de precisão encontrados nas avaliações das classificações, de modo geral, é intuitivo pensar sobre alguns fatores. O primeiro deles é a quantidade de classes estudadas e seus comportamentos espectro-temporais. Os fenômenos de colheita

da cana sejam eles sem o uso do fogo ou com a queima da palha na pré-colheita, associados respectivamente às classes temáticas Cana Crua (CC) e Cana Queima (CQ), assim como a cana que não foi colhida até a data final da série temporal analisada, representada na classe Cana-em-Pé (CP), apresentam comportamentos espectro-temporais muito distintos entre si.

Outro fato importante a considerar, para que fossem obtidas elevadas precisões nas classificações, foi a cuidadosa seleção das amostras de treinamento, além da importante consideração sobre detalhar as classes de colheita da cana em cada data de imagem utilizada. Só para se ter uma idéia de como esse fato pode afetar a classificação, efetuando-se uma classificação de máxima verossimilhança simples, utilizando apenas a última imagem da série temporal (data 6), porém utilizando as mesmas 12 classes treinadas para as classificações multitemporais (C1 a C6)¹, obteve-se como resultado um mapa temático cujos índices de Exatidão Global (EG) e *Kappa* (κ) foram de 85,9% e 0,789, respectivamente. Além disso, o fato da área de estudo ser relativamente pequena e apresentar talhões com padrões bem definidos colaborou com as elevadas precisões encontradas nas classificações.

Valores de *Kappa* acima de 0,8, como nas classificações multitemporais (C1 a C6), são considerados excelentes segundo Landis e Koch (1977). Ainda com relação a qualificar as classificações por meio dos valores dos índices, de acordo com Foody (2002), é desejável que uma classificação apresente um índice de Exatidão Global acima de 85%, como também é o caso das classificações multitemporais apresentadas nesta pesquisa.

¹ Desconsiderando-se o fato de que, apenas com a última imagem da série, o intérprete não teria como coletar amostras para as classes de colheita da cana nas primeiras datas da série temporal, como por exemplo, as classes: Cana Crua colhida na data 1 (CC1), CC2, ..., ou mesmo Cana Queima colhida na data 1(CQ1), CQ2, ...

5. CONCLUSÕES E RECOMENDAÇÕES

5.1 Conclusões

Os resultados deste trabalho indicaram que é possível utilizar técnicas de classificação espectro-temporal de imagens digitais para automatizar grande parte do processo de interpretação visual, obtendo elevados graus de precisão e confiabilidade no mapeamento de áreas de colheita de cana-de-açúcar (com e sem a queima da palha) por imagens de satélites de sensoriamento remoto orbital.

A metodologia de classificação por superfícies de resposta espectro-temporais (STRS), mostrou-se adequada para mapear as áreas de colheita da cana-de-açúcar conforme comprovado pelos excelentes índices de Exatidão Global, maiores que 90% (FOODY, 2002) e por valores de *Kappa* maiores que 0,85 (LANDIS e KOCH, 1976).

As classificações alternativas da metodologia STRS (C2, C3 e C4) não apresentaram ganho de qualidade nos resultados, quando comparadas com a metodologia tradicional STRS proposta por Vieira (2000). Dentre as classificações alternativas, aquela feita por etapas (C4), foi a que apresentou melhor resultado.

Dentre as outras classificações testadas (C5, C6 e C7) a C5 e C6 apresentaram os melhores resultados. No entanto, nenhuma destas classificações teve desempenho superior às classificações STRS (C1 a C4), apresentando, inclusive, resultados inferiores à C4. Contudo, o resultado favorável da classificação C6 deve ser visto com cuidado, pois essa metodologia de classificação não possui a robustez das demais classificações multitemporais, especialmente em relação às classificações STRS (C1 a C4), podendo apresentar resultados bem distintos se utilizada em outros anos-safra. Já a classificação de uma só data (C7), sempre teve desempenho inferior às classificações multitemporais, indicando que a característica multitemporal das imagens de sensoriamento remoto, aplicado ao monitoramento agrícola, é fundamental.

5.2 Recomendações

Em virtude dos estudos realizados e da experiência adquirida recomenda-se:

- a) Testar as técnicas empregadas neste estudo, também em imagens de vários anos-safra, de modo a verificar a robustez das classificações e também visando criar uma biblioteca espectral para ser utilizada nas classificações posteriores, sem a necessidade da coleta de amostras de treinamento.
- b) Testar maneiras de reduzir o número de coeficientes utilizados nas classificações STRS, seja pela redução do grau do polinômio ou por procedimentos de seleção dos melhores coeficientes, como mostrado em Vieira et al. (2000).
- c) Avaliar a aplicação da metodologia de classificação STRS (C1) em grandes áreas e utilizando outros sensores.
- d) Implementar o classificador multitemporal por superfícies de resposta espectro-temporais (STRS) no aplicativo SPRING, considerando os procedimentos necessários de engenharia de *software* para minimizar ao máximo o tempo computacional demandado no processamento.

REFERÊNCIAS BIBLIOGRÁFICAS

ABEND, K.; HARLEY, T.; CHANDRASEKARAN, B.; HUGHES, G. Comments on 'On the mean accuracy of statistical pattern recognizers' by Hughes, G. F. **IEEE Transactions on Information Theory**, v. 15, n. 3, p. 420-423, 1969.

AGUIAR, D. A. **Monitoramento da área colhida de cana-de-açúcar por meio de imagens do sensor MODIS**. 2007. 103 p. (INPE-14803-TDI/1246). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos-SP, 2007.

AGUIAR, D. A.; ADAMI, M.; RUDORFF, B. F. T.; SÁNCHEZ, G. A.; BARROS, M. A.; SUGAWARA, L. M.; SHIMABUKURO, Y. E.; MOREIRA, M. A. **Mapeamento da colheita da cana-de-açúcar no estado de São Paulo - ano safra 2006/2007**. São José dos Campos-SP: INPE, 2007. 48 p. (INPE-14788-RPE/811). Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m17@80/2007/12.10.12.56>>. Acesso em: 17 ago. 2008.

AGUIAR, D. A.; RUDORFF, B. F. T.; SILVA, W. F. Monitoramento do modo de colheita da cana-de-açúcar no Estado de São Paulo - Brasil por meio de imagens de sensores orbitais em dois anos-safra. In: Simposio Latinoamericano de Percepción Remota y Sistemas de Información Espacial (SELPER), 13., 2008, Havana, Cuba. **Anais...** Havana: SELPER, 2008.

AGUIAR, D. A.; RUDORFF, B. F. T.; SILVA, W. F.; CARVALHO, M. A.; AULICINO, T. L. I. N.; BRANDÃO, D.; GOLTZ, E.; ADAMI, M.; SUGAWARA, L. M. **Mapeamento da colheita da cana-de-açúcar no estado de São Paulo - ano safra 2007/2008**. São José dos Campos-SP: INPE, 2009. 65 p. (INPE-15724-RPQ/821).

ALONSO PIPPO, W.; GARZONE, P.; CORNACCHIA, G. Agro-industry sugarcane residues disposal: the trends of their conversion into energy carriers in Cuba. **Waste Management**, v. 27, n. 7, p. 869-885, 2007.

ANTUNES, M. A. H.; FREIRE, R. M. B.; BOTELHO, A. S.; TONIOLLI, L. H. Correções atmosféricas de imagens de satélites utilizando o modelo 6S. In: Congresso Brasileiro de Cartografia, 21., 2003, Belo Horizonte, MG. **Anais...** Rio de Janeiro: Sociedade Brasileira de Cartografia, 2003. Disponível em: <http://www.cartografia.org.br/xxi_cbc/174-SR21.pdf>. Acesso em: 27 nov. 2007.

BJÖRCK, Å. Solving linear least squares problems by Gram-Schmidt orthogonalization. **BIT Numerical Mathematics**, v. 7, n. 1, p. 1-21, 1967.

BRAGA, A. L.; VIEIRA, C. A. O.; ASSIS, L. C. Identificação de extensas áreas de culturas agrícolas empregando uma abordagem espectro-temporal utilizando imagens MODIS. In: Simpósio Brasileiro de Sensoriamento Remoto, 13., 2007, Florianópolis, SC. **Anais...** São José dos Campos: INPE, 2007. p. 5603-5610. Disponível em: <<http://marte.dpi.inpe.br/col/dpi.inpe.br/sbsr@80/2006/11.14.17.06.47/doc/5603-5610.pdf>>. Acesso em: 02 nov. 2007.

BRAUNBECK, O. A.; OLIVEIRA, J. T. A. Colheita de cana-de-açúcar com auxílio mecânico. **Engenharia Agrícola**, v. 26, n. 1, p. 300-308, 2006. Disponível em: <www.scielo.br/pdf/eagri/v26n1/30121.pdf>. Acesso em: 23 fev. 2008.

CÂMARA, G.; SOUZA, R. C. M.; FREITAS, U. M.; GARRIDO, J.; II, F. M. SPRING: integrating remote sensing and GIS by object-oriented data modelling. **Computers & Graphics**, v. 20, n. 3, p. 395-403, 1996.

CANASAT. **Mapeamento de cana via satélites de observação da Terra**. 2009. Disponível em: <<http://www.dsr.inpe.br/mapdsr/>>. Acesso em: 07 abr. 2009.

CANÇADO, J. E.; SALDIVA, P. H.; PEREIRA, L. A.; LARA, L. B.; ARTAXO, P.; MARTINELLI, L. A.; ARBEX, M. A.; ZANOBETTI, A.; BRAGA, A. L. The impact of sugar cane-burning emissions on the respiratory system of children and the elderly. **Environ Health Perspect**, v. 114, n. 5, p. 725-729, 2006.

CANTY, M. J.; NIELSEN, A. A. Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. **Remote Sensing of Environment**, v. 112, n. 3, p. 1025-1036, 2008.

CANTY, M. J.; NIELSEN, A. A.; SCHMIDT, M. Automatic radiometric normalization of multitemporal satellite imagery. **Remote Sensing of Environment**, v. 91, n. 3-4, p. 441-451, 2004.

CHANDER, G.; MARKHAM, B. L.; BARSÍ, J. A. Revised Landsat-5 Thematic Mapper radiometric calibration. **IEEE Geoscience and Remote Sensing Letters**, v. 4, n. 3, p. 490-494, 2007.

CONGALTON, R. G.; MEAD, R. A. A quantitative method to test for consistency and correctness in photointerpretation. **Photogrammetric Engineering & Remote Sensing**, v. 49, n. 1, p. 69-74, 1983.

CORREIA, A.; CASTANHO, A. D. A.; MARTINS, J. V.; LONGO, K. M.; YAMASOE, M. A.; ARTAXO, P. Inferência de aerossóis. In: RUDORFF, B. F. T.; SHIMABUKURO, Y. E.; CEBALLOS, J. C. (Org.). **O sensor MODIS e suas aplicações ambientais no Brasil**. São José dos Campos-SP: Editora Parêntese, 2007. cap. 21, p. 297-315.

DAI, X.; KHORRAM, S. The effects of image misregistration on the accuracy of remotely sensed change detection. **IEEE Transactions on Geoscience and Remote Sensing**, v. 36, n. 5, p. 1566-1577, 1998.

DAWSON, L.; BOOPATHY, R. Use of post-harvest sugarcane residue for ethanol production. **Bioresource Technology**, v. 98, n. 9, p. 1695-1699, 2007.

EMPRESA DE PESQUISA ENERGÉTICA (EPE). **Informe à imprensa** – PDE 2008-2017 - etanol: demanda de etanol no país crescerá 150% nos próximos dez anos. 2008a. Disponível em: <http://www.epe.gov.br/PressReleases/20080924_1.pdf>. Acesso em: 10 nov. 2008.

_____. **Informe à imprensa** – resultados preliminares BEN 2008: cana-de-açúcar já é a segunda fonte primária de energia no Brasil. 2008b. Disponível em: <http://www.epe.gov.br/PressReleases/20080508_1.pdf>. Acesso em: 10 nov. 2008.

EIPHANIO, R. D. V. **Avaliação da potencialidade das imagens MODIS na estimativa da área de soja no estado do Mato Grosso**. 2007. 103 p. (INPE-15190-TDI/1299). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos-SP, 2007. Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m17@80/2007/07.18.12.27>>. Acesso em: 25 abr. 2008.

ESCOBAR, J. C.; LORA, E. S.; VENTURINI, O. J.; YÁÑEZ, E. E.; CASTILLO, E. F.; ALMAZAN, O. Biofuels: environment, technology and food security. **Renewable and Sustainable Energy Reviews**, v. 13, n. 6-7, p. 1275-1287, 2009.

FALLAH-ADL, H.; JÁJÁ, J.; LIANG, S. Fast algorithms for estimating aerosol optical depth and correcting Thematic Mapper imagery. **The Journal of Supercomputing**, v. 10, n. 4, p. 315-329, 1997.

FOODY, G. M. Status of land cover classification accuracy assessment. **Remote Sensing of Environment**, v. 80, n. 1, p. 185-201, 2002.

FREDEN, S. C.; GORDON JR, F. Landsat Satellites. In: COWELL, R. N. (Org.). **Manual of Remote Sensing**. Bethesda, MD, USA: American Society for Photogrammetry and Remote Sensing, 1983. cap. 12, p. 517-570.

GOLDEMBERG, J.; COELHO, S. T.; GUARDABASSI, P. The sustainability of ethanol production from sugarcane. **Energy Policy**, v. 36, n. 6, p. 2086-2097, 2008.

HIRA, A.; OLIVEIRA, L. G. No substitute for oil? How Brazil developed its ethanol industry. **Energy Policy**, v. 37, n. 6, p. 2450-2456, 2009.

HUDSON, W. D.; RAMM, C. W. Correct formulation of the kappa coefficient of agreement. **Photogrammetric Engineering & Remote Sensing**, v. 53, n. 4, p. 421-422, 1987.

HUGHES, G. On the mean accuracy of statistical pattern recognizers. **IEEE Transactions on Information Theory**, v. 14, n. 1, p. 55-63, 1968.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Levantamento sistemático da produção agrícola** - dezembro de 2000. Rio de Janeiro-RJ: IBGE, 2000. 76 p. (ISSN 0103-443X). Disponível em: <[ftp://ftp.ibge.gov.br/Producao_Agricola/Levantamento_Sistematico_da_Producao_Agricola_\[mensal\]/Fasciculo/2000/lspa_200012.zip](ftp://ftp.ibge.gov.br/Producao_Agricola/Levantamento_Sistematico_da_Producao_Agricola_[mensal]/Fasciculo/2000/lspa_200012.zip)>. Acesso em: 05 abr. 2009.

_____. **Levantamento sistemático da produção agrícola** - dezembro de 2008. Rio de Janeiro-RJ: IBGE, 2008. 80 p. (ISSN 0103-443X). Disponível em: <[ftp://ftp.ibge.gov.br/Producao_Agricola/Levantamento_Sistematico_da_Producao_Agricola_\[mensal\]/Fasciculo/2008/lspa_200812caderno.zip](ftp://ftp.ibge.gov.br/Producao_Agricola/Levantamento_Sistematico_da_Producao_Agricola_[mensal]/Fasciculo/2008/lspa_200812caderno.zip)>. Acesso em: 18 ago. 2008.

JENSEN, J. R. **Remote sensing of the environment: an Earth Resource perspective**. 2 ed. Upper Saddle River, NJ, USA: Pearson Prentice Hall, 2006. 608 p. (Prentice Hall series in geographic information science).

KAUFMAN, Y. J. Atmospheric effect on spectral signature-measurements and corrections. **IEEE Transactions on Geoscience and Remote Sensing**, v. 26, n. 4, p. 441-450, 1988.

KESHAHA, N.; MUSTARD, J. F. Spectral unmixing. **IEEE Signal Processing Magazine**, v. 19, n. 1, p. 44-57, 2002.

KIRCHHOFF, V. W. J. H.; MARINHO, E. V. A.; DIAS, P. L. S.; PEREIRA, E. B.; CALHEIROS, R.; ANDRÉ, R.; VOLPE, C. Enhancements of CO and O3 from burnings in sugar cane fields **Journal of Atmospheric Chemistry**, v. 12, n. 1, p. 87-102, 1991.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, n. 1, p. 159-174, 1977.

LE BLOND, J. S.; WILLIAMSON, B. J.; HORWELL, C. J.; MONRO, A. K.; KIRK, C. A.; OPPENHEIMER, C. Production of potentially hazardous respirable silica airborne particulate from the burning of sugarcane. **Atmospheric Environment**, v. 42, n. 22, p. 5558-5568, 2008.

LIANG, S.; FALLAH-ADL, H.; KALLURI, S.; JÁJÁ, J.; KAUFMAN, Y. J.; TOWNSHEND, J. R. G. An operational atmospheric correction algorithm for Landsat Thematic Mapper imagery over the land. **Journal of Geophysical Research**, v. 102, n. D14, p. 17173-17186, 1997.

LUO, L.; VAN DER VOET, E.; HUPPES, G. Life cycle assessment and life cycle costing of bioethanol from sugarcane in Brazil. **Renewable and Sustainable Energy Reviews**, v. 13, n. 6-7, p. 1613-1619, 2009.

MACHADO, K. J.; SOARES, E. G. S.; CARVALHO, L. T.; MELLO, M. P.; VIEIRA, C. A. O. Avaliação do uso do modelo linear de mistura espectral na classificação de imagem ALOS. In: Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação, 2., 2008, Recife, PE. **Anais...** Recife: UFPE, 2008. CD-ROM. Disponível em: <<http://mtc-m18.sid.inpe.br:80/rep-/sid.inpe.br/mtc-m18@80/2008/10.10.21.06>>. Acesso em: 11 out. 2008.

MATHER, P. M. **Computational methods of multivariate analysis in physical geography**. Londres, Inglaterra: John Wiley & Sons, 1976. 532 p.

_____. **Computer processing of remotely-sensed images: an introduction.** 3 ed. Chichester, Inglaterra: John Wiley & Sons, 2004. 324 p.

MELLO, M. P.; SILVA, G. B. S.; PEREIRA, G.; PRADO, B. R.; SHIMABUKURO, Y. E.; RUDORFF, B. F. T.; VIEIRA, C. A. O.; PETERNELLI, L. A. Avaliação do uso do modelo linear de mistura espectral na discriminação de fitofisionomias do Cerrado. In: Simpósio Nacional do Cerrado - Simpósio Internacional de Savanas Tropicais, 9-2., 2008, Brasília, DF. **Anais...** Planaltina, DF: EMBRAPA-CERRADOS, 2008. (CD-ROM). Disponível em: <<http://mtc-m18.sid.inpe.br:80/rep-/sid.inpe.br/mtc-m18@80/2008/10.17.00.51>>. Acesso em: 16 out. 2008.

MELLO, M. P.; VIEIRA, C. A. O. Modelagem de um interpolador “Polynomial Trend Surface” usando o software livre R para gerar superfícies de assinaturas espectrais multitemporais de culturas agrícolas. In: Simpósio de Iniciação Científica da Universidade Federal de Viçosa, 15., 2006, Viçosa, MG. **Resumos...** Viçosa: UFV, 2006. (897) (CD-ROM). Disponível em: <<http://mtc-m17.sid.inpe.br:80/rep-/sid.inpe.br/mtc-m17@80/2008/04.29.02.11>>. Acesso em: 29 abr. 2008.

MELLO, M. P.; RUDORFF, B. F. T.; VIEIRA, C. A. O.; AGUIAR, D. A. Comportamento espectro-temporal da cana-de-açúcar: uma nova proposta na distribuição espacial dos pontos de controle utilizados na interpolação de superfícies de resposta espectro-temporais. In: Simpósio Brasileiro de Sensoriamento Remoto, 14., 2009a, Natal, RN. **Anais...** São José dos Campos: INPE, 2009. p. 271-278. Disponível em: <<http://sbsr.dpi.inpe.br/col/dpi.inpe.br/sbsr@80/2008/11.17.19.02/doc/271-278.pdf>>. Acesso em: 22 abr. 2009.

MELLO, M. P.; VIEIRA, C. A. O.; AGUIAR, D. A.; RUDORFF, B. F. T. Classificação da colheita da cana-de-açúcar por meio de imagens de satélite utilizando superfícies de resposta espectro-temporais. In: Simpósio Brasileiro de Sensoriamento Remoto, 14., 2009b, Natal, RN. **Anais...** São José dos Campos: INPE, 2009. p. 279-286. Disponível em: <<http://sbsr.dpi.inpe.br/col/dpi.inpe.br/sbsr@80/2008/11.17.05.02.44/doc/279-286.pdf>>. Acesso em: 22 abr. 2009.

MOREIRA, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação.** 3 ed. Viçosa, MG: Editora UFV, 2005. 421 p.

NEVES, M. F.; CONEJERO, M. A. Sistema agroindustrial da cana: cenários e agenda estratégica. **Economia Aplicada**, v. 11, n. 4, p. 587-604, 2007. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-80502007000400007&nrm=iso>. Acesso em: 22 fev. 2008.

NIELSEN, A. A. The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data. **IEEE Transactions on Image Processing**, v. 16, n. 2, p. 463-478, 2007.

NIELSEN, A. A.; CANTY, M. J. Multi- and hyperspectral remote sensing change detection with generalized difference images by the IR-MAD method. In: International Workshop on the Analysis of Multi-Temporal Remote Sensing Images (MultiTemp), 3., 2005, Biloxi, Mississippi, USA. **Proceedings...** USA: IEEE International, 2005. p. 169-173.

NIELSEN, A. A.; CONRADSEN, K.; ANDERSEN, O. B. A change oriented extension of EOF analysis applied to the 1996-1997 AVHRR sea surface temperature data. **Physics and Chemistry of the Earth**, v. 27, n. 32-34, p. 1379-1386, 2002.

NIELSEN, A. A.; CONRADSEN, K.; SIMPSON, J. J. Multivariate Alteration Detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: new approaches to change detection studies. **Remote Sensing of Environment**, v. 64, n. 1, p. 1-19, 1998.

PINTO, M. L. A. **Relação empírica da visibilidade com profundidade óptica, concentração de aerossóis e focos de queimadas em Alta Floresta e Cuiabá, em 1993 e 1994.** 2001. 141 p. (INPE-10097-TDI/893). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos-SP, 2001. Disponível em: <<http://urlib.net/sid.inpe.br/jeferson/2003/09.18.08.48>>. Acesso em: 25 jun. 2008.

PONZONI, F. J.; SHIMABUKURO, Y. E. **Sensoriamento remoto no estudo da vegetação.** 1 ed. São José dos Campos, SP: Parêntese Editora, 2007. 144 p.

R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing.** ver. 2.9.0. Vienna, Austria: R Foundation for Statistical Computing. 2009. Disponível em: <www.R-project.org>. Acesso em: 18 abr. 2009.

RIZZI, R.; RISSO, J.; EPIPHANIO, R. D. V.; RUDORFF, B. F. T.; FORMAGGIO, A. R.; SHIMABUKURO, Y. E.; FERNANDES, S. L. Estimativa da área de soja no Mato Grosso por meio de imagens MODIS. In: Simpósio Brasileiro de Sensoriamento Remoto, 14., 2009, Natal, RN. **Anais...** São José dos Campos: INPE, 2009. p. 387-394. Disponível em: <<http://sbsr.dpi.inpe.br/col/dpi.inpe.br/sbsr@80/2008/11.16.18.50.57/doc/387-394.pdf>>. Acesso em: 28 abr. 2009.

RODRIGUES, E. B.; SAAB, O. J. G. A. Avaliação técnico-econômica da colheita manual e mecanizada da cana-de-açúcar (*saccharum spp*) na região de Bandeirantes-PR. **Semina. Ciências Agrárias (Online)**, v. 28, n. 4, p. 581-588, 2007. Disponível em: <www.uel.br/proppg/semina/pdf/semina_28_4_19_5.pdf>. Acesso em: 17 fev. 2008.

ROSEIRO, M. N. V.; TAKAYANAGUI, A. M. M. Meio ambiente e poluição atmosférica: o caso da cana-de-açúcar. **Revista Saúde**, v. 30, n. 1-2, p. 76-83, 2004.

RSI. **Environment for Visualizing Images - ENVI.** ver. 4.3. Boulder, CO, USA: ITT Industries. 2006.

RUDORFF, B. F. T.; BERKA, L. M. S.; MOREIRA, M. A.; DUARTE, V.; XAVIER, A. C.; ROSA, V. G. C.; SHIMABUKURO, Y. E. Imagens de satélite no mapeamento e estimativa de área de cana-de-açúcar em São Paulo: ano-safra 2003/2004. **Agricultura em São Paulo**, v. 52, n. 1, p. 21-39, 2005. Disponível em:

<<http://www.iea.sp.gov.br/out/publicacoes/pdf/asp-1-05-2.pdf>>. Acesso em: 03 jan. 2008.

RUDORFF, B. F. T.; BERKA, L. M. S.; XAVIER, A. C.; MOREIRA, M. A.; DUARTE, V.; ROSA, V. G. C.; SHIMABUKURO, Y. E. **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**. São José dos Campos-SP: INPE, 2004. 47 p. (INPE-10791-RPQ/759). Disponível em:

<<http://www.dsr.inpe.br/mapdsr/data/artigos/2003.pdf>>. Acesso em: 15 set. 2008.

RUDORFF, C. M.; RIZZI, R.; RUDORFF, B. F. T.; SUGAWARA, L. M.; VIEIRA, C. A. O. Superfícies de resposta espectro-temporal de imagens do sensor MODIS para classificação de área de soja no Estado do Rio Grande do Sul. **Ciência Rural**, v. 37, n. 1, p. 118-125, 2007.

SANCHES, I. D. A.; EPIPHANIO, J. C. N.; FORMAGGIO, A. R. Culturas agrícolas em imagens multitemporais do satélite Landsat1. **Agricultura em São Paulo**, v. 52, n. 1, p. 83-96, 2005.

SCHOTT, J. R.; SALVAGGIO, C.; VOLCHOK, W. J. Radiometric scene normalization using pseudoinvariant features. **Remote Sensing of Environment**, v. 26, n. 1, p. 1-16, 1988.

SCHROEDER, T. A.; COHEN, W. B.; SONG, C.; CANTY, M. J.; YANG, Z. Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon. **Remote Sensing of Environment**, v. 103, n. 1, p. 16-26, 2006.

SECRETARIA DE MEIO AMBIENTE DO ESTADO DE SÃO PAULO (SMA-SP). **Eliminação gradativa da queima da palha da cana-de-açúcar**. 2009. Disponível em: <<http://www.sigam.ambiente.sp.gov.br/sigam2/default.aspx?idPagina=123>>. Acesso em: 15 mai. 2009.

SHIMABUKURO, Y. E.; SMITH, J. A. The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 29, n. 1, p. 16-20, 1991.

SONG, C.; WOODCOCK, C. E.; SETO, K. C.; LENNEY, M. P.; MACOMBER, S. A. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? **Remote Sensing of Environment**, v. 75, n. 2, p. 230-244, 2001.

SOUZA, Z. M.; PRADO, R. M.; PAIXÃO, A. C. S.; CESARIN, L. G. Sistemas de colheita e manejo da palhada de cana-de-açúcar. **Pesquisa Agropecuária Brasileira**, v. 40, n. 3, p. 271-278, 2005.

SUGAWARA, L. M.; RUDORFF, B. F. T.; VIEIRA, R. M. D. S. P.; AFONSO, A. G.; AULICINO, T. L. I. N.; CARVALHO, M. A. D.; MOREIRA, M. A.; DUARTE, V.; SILVA, W. F. D.; AGUIAR, D. A. D. **Imagens de satélites na estimativas de área plantada com cana na safra 2005/2006 - região centro-sul**. São José dos Campos-SP: INPE, 2008. 75 p. (INPE-15254-RPQ/815). Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m18@80/2008/07.10.19.59>>. Acesso em: 17 set. 2008.

TOMPKINS, S.; MUSTARD, J. F.; PIETERS, C. M.; FORSYTH, D. W. Optimization of endmembers for spectral mixture analysis. **Remote Sensing of Environment**, v. 59, n. 3, p. 472-489, 1997.

UNIÃO DAS INDÚSTRIAS DE CANA-DE-AÇÚCAR (UNICA). **Cresce a mecanização na colheita de cana-de-açúcar em São Paulo**. 2009. Disponível em: <<http://unica.com.br/noticias/show.asp?nwsCode=9C316670-3A8A-4908-B41A-CEDC6E8B1088>>. Acesso em: 12 mai. 2009.

UNITED STATES GEOLOGICAL SURVEY (USGS). **Landsat 5 satellite back in action**. 2008. Disponível em: <<http://www.usgs.gov/newsroom/article.asp?ID=1881>>. Acesso em: 5 mar. 2009.

_____. **Designed for 3 years, 'workhorse' satellite completes 25**. 2009. Disponível em: <http://www.usgs.gov/newsroom/article.asp?ID=2154&from=rss_home>. Acesso em: 5 mar. 2009.

URIARTE, M.; YACKULIC, C. B.; COOPER, T.; FLYNN, D.; CORTES, M.; CRK, T.; CULLMAN, G.; MCGINTY, M.; SIRCELY, J. Expansion of sugarcane production in São Paulo, Brazil: Implications for fire occurrence and respiratory health. **Agriculture, Ecosystems & Environment**, v. 132, n. 1-2, p. 48-56, 2009.

VERMOTE, E. F.; TANRE, D.; DEUZE, J. L.; HERMAN, M. A.; MORCETTE, J. J. A. Second simulation of the satellite signal in the solar spectrum, 6S: an overview. **IEEE Transactions on Geoscience and Remote Sensing**, v. 35, n. 3, p. 675-686, 1997.

VIEIRA, C. A. O. **Accuracy of remotely sensing classification of agricultural crops: a comparative study**. 2000. 323 p. Tese (Ph.D. em Geografia Física) - University of Nottingham, Nottingham, Inglaterra, 2000.

VIEIRA, C. A. O.; MATHER, P. M. Visualisation of measures of classifier reliability and error in remote sensing. In: ICG ACCURACY, 4., 2000, Amsterdam. **Anais...** Netherlands: Delft University Press, 2000. p. 701-708.

VIEIRA, C. A. O.; MATHER, P. M.; APLIN, P. Multitemporal classification of agricultural crops using the spectral-temporal response surface. In: BRUZZONE, L.; SMITS, P. (Org.). **Analysis of multi-temporal remote sensing images**. Londres, Inglaterra: World Scientific Publishing Co., v. 2, 2002. cap. 32, p. 290-297.

_____. Agricultural crop classification using the spectral-temporal response surface. In: Simpósio Brasileiro de Sensoriamento Remoto, 11., 2003, Belo Horizonte, MG. **Anais...** São José dos Campos: INPE, 2003. p. 255-262. Disponível em: <http://marte.dpi.inpe.br/col/ltid.inpe.br/sbsr/2002/11.13.10.49/doc/01_093.pdf>. Acesso em: 19 set. 2007.

VIEIRA, C. A. O.; MATHER, P. M.; MCCULLAGH, M. The Spectral-Temporal Response Surface and its use in the multi-sensor, multitemporal classification of agricultural crops. In: Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS), 33., 2000, Amsterdam. **Proceedings...** Amsterdam, Netherlands: International Archive Photogrammetry and Remote Sensing, 2000. p. 16-23.

WATSON, D. F. **Contouring**: a guide to the analysis and display of spatial data - with programs on diskette. 1 ed. Oxford, England: Pergamon Press, v. 10, 1992. 321 p. (Computer Methods in the Geosciences).

ZUURBIER, P.; VAN DE VOOREN, J., (Eds.). **Sugarcane ethanol**: contributions to climate change mitigation and the environment. Wageningen, Holanda: Wageningen Academic Publishers, 2008. 256 p.

APÊNDICE A – CORREÇÕES RADIOMÉTRICAS

A.1 Algoritmo em LEGAL (converte número digital em reflectância aparente)

```
{
//transforma número digital em Radiância Aparente p/imagens LANDSAT5
//Idealizado pelo Dr. Flavio J. Ponzoni
//criado pelo M.Sc. Gabriel Pereira
//e editado por Márcio Pupin de Mello

//Declaração da categoria contendo as imagens a serem transformadas
Image Im1, Im2, Im3, Im4, Im5, Im6 ("IM_BRT_20070706");

//Declaração da categoria que conterà os dados transformados
//Criar os PIs abaixo (os nomes q estao entre aspas - imagem e numerico)
Image Im13, Im14, Im15, Im16, Im17, Im18 ("IM_Rap_20070706");
Digital Im7, Im8, Im9, Im10, Im11, Im12 ("IM_RapNum_20070706");

//Atribuir as informações a cada variável
Im1=Recupere (Nome="20070706_1");
Im2=Recupere (Nome="20070706_2");
Im3=Recupere (Nome="20070706_3");
Im4=Recupere (Nome="20070706_4");
Im5=Recupere (Nome="20070706_5");
Im6=Recupere (Nome="20070706_7");

// PIs de saída que serão criados
Im7=Novo (Nome="Rfap_1", ResX=30, ResY=30, Escala=50000, Min=0, Max=255);
Im8=Novo (Nome="Rfap_2", ResX=30, ResY=30, Escala=50000, Min=0, Max=255);
Im9=Novo (Nome="Rfap_3", ResX=30, ResY=30, Escala=50000, Min=0, Max=255);
Im10=Novo (Nome="Rfap_4", ResX=30, ResY=30, Escala=50000, Min=0, Max=255);
Im11=Novo (Nome="Rfap_5", ResX=30, ResY=30, Escala=50000, Min=0, Max=255);
Im12=Novo (Nome="Rfap_7", ResX=30, ResY=30, Escala=50000, Min=0, Max=255);

// PIs numéricos de saída que serão criados
Im13=Novo (Nome="Ra_1", ResX=30, ResY=30, Nbits=8);
Im14=Novo (Nome="Ra_2", ResX=30, ResY=30, Nbits=8);
Im15=Novo (Nome="Ra_3", ResX=30, ResY=30, Nbits=8);
Im16=Novo (Nome="Ra_4", ResX=30, ResY=30, Nbits=8);
Im17=Novo (Nome="Ra_5", ResX=30, ResY=30, Nbits=8);
Im18=Novo (Nome="Ra_7", ResX=30, ResY=30, Nbits=8);

//Os cossenos devem ser em radianos!
//Valor de 1 radiano (não alterar)
rad = 57.302941550999617980389660002547;

//valores dos parâmetros
zen=(53.80012123237)/(rad);
pi=3.141592654;
dmax=255;
dist=0.96659706218;

//Parâmetros de calibração do Landsat5-TM
Lmin1=-0.152;
Lmin2=-0.284;
Lmin3=-0.117;
Lmin4=-0.151;
Lmin5=-0.037;
Lmin7=-0.015;

Lmax1=19.30;
Lmax2=36.50;
Lmax3=26.40;
Lmax4=22.10;
Lmax5=3.020;
Lmax7=1.650;

esun1=195.7;
esun2=182.6;
```

```

esun3=155.4;
esun4=103.6;
esun5=21.5;
esun7=8.07;

//Contas feitas pelo modelo
Im7= (((Lmin1+(Lmax1-Lmin1)*(Digital (Im1/dmax)))*pi*dist^2)/(esun1*cos(zen)))*255;
Im7= Im7 <= 0 ? 0 : Im7;
Im13= Imagem (Im7);
Im8= (((Lmin2+(Lmax2-Lmin2)*(Digital (Im2/dmax)))*pi*dist^2)/(esun2*cos(zen)))*255;
Im8= Im8 <= 0 ? 0 : Im8;
Im14= Imagem (Im8);
Im9= (((Lmin3+(Lmax3-Lmin3)*(Digital (Im3/dmax)))*pi*dist^2)/(esun3*cos(zen)))*255;
Im9= Im9 <= 0 ? 0 : Im9;
Im15= Imagem (Im9);
Im10= (((Lmin4+(Lmax4-Lmin4)*(Digital (Im4/dmax)))*pi*dist^2)/(esun4*cos(zen)))*255;
Im10= Im10 <= 0 ? 0 : Im10;
Im16= Imagem (Im10);
Im11= (((Lmin5+(Lmax5-Lmin5)*(Digital (Im5/dmax)))*pi*dist^2)/(esun5*cos(zen)))*255;
Im11= Im11 <= 0 ? 0 : Im11;
Im17= Imagem (Im11);
Im12= (((Lmin7+(Lmax7-Lmin7)*(Digital (Im6/dmax)))*pi*dist^2)/(esun7*cos(zen)))*255;
Im12= Im12 <= 0 ? 0 : Im12;
Im18= Imagem (Im12);

}

```

A.2 Dados de entrada do algoritmo 6S (um exemplo)

7	(Landsat TM geometrical conditions)
07 06 13.0644 -47.35 -21.567	(month, day, hh.ddd (hour in univ. time), long., lat.)
1	(tropical atmospheric mode)
1	(continental aerosols model)
50	(visibility in km - aerosol model concentration)
-.650	(target at km above sea level)
-1000	(sensor on board of satellite)
25	(Landsat5-TM b1=25 b2=26 b3=27 b4=28 b5=29 b7=30)
-1	(Apparent reflectance)
250000	(number of pixels of the image = number of bytes)

A.3 Código IDL (adaptado) normalização radiométrica IR-MAD

```

;+
; NAME:
;   RADCAL_RUN
; PURPOSE:
;   Radiometric calibration using MAD
;   Ref: M. Canty et al. Remote Sensing of
;   Environment 91(3,4) (2004) 441-451
;   Reference and target images must have
;   equal spatial and spectral dimensions,
;   at least 2 spectral components, and be
;   registered to one another.
;   Once the regression coefficients have been
;   determined, they can be used to
;   calibrate another file, for example a full
;   scene, which need not be registered
;   to the reference image.
; AUTHOR
;   Mort Canty (2006) m.canty@fz-juelich.de
;   Juelich Research Center
; ADAPTATION
;   A simple adaptation was made in July/2008
;   by Marcio Pupin de Mello
;   (from Brazilian National Institute for Space Research - INPE)
;   under Dr. Canty's supervision
;   for the set up the output normalized image in "byte" (8bits) format
; CALLING SEQUENCE:

```

```

;       Radcal_Run
; ARGUMENTS:
;       Event (if used as a plug-in menu item)
; KEYWORDS:
;       None
; DEPENDENCIES:
;       ENVI
;       ORTHO_REGRESS
;       WISHART
;       PROGRESSBAR_DEFINE (FSC_COLOR)
;-----
; This program is free software; you can redistribute it and/or modify
; it under the terms of the GNU General Public License as published by
; the Free Software Foundation; either version 2 of the License, or
; (at your option) any later version.
;
; This program is distributed in the hope that it will be useful,
; but WITHOUT ANY WARRANTY; without even the implied warranty of
; MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
; GNU General Public License for more details.
;
; You should have received a copy of the GNU General Public License
; along with this program; if not, write to the Free Software
; Foundation, Inc., 59 Temple Place, Suite 330, Boston, MA 02111-1307 USA
;-----

PRO radcal_run_define_buttons, buttonInfo
  ENVI_DEFINE_MENU_BUTTON, buttonInfo, $
  VALUE = 'Radiometric Normalization', $
  REF_VALUE = 'Change Detection Statistics', $
  EVENT_PRO = 'radcal_run', $
  UVALUE = 'RADCAL', $
  POSITION = 'after', $
  /SEPARATOR
END

pro radcal_run, event

print, '-----'
print, 'Radiometric calibration with MAD'
print, systemtime(0)

catch, theError
if theError ne 0 then begin
  void = Dialog_Message(!Error_State.Msg, /error)
  return
endif

envi_select, title='Choose (subset of) reference image', fid=fid1, dims=dims1,pos=pos1
  if (fid1 eq -1) then begin
    print, 'cancelled'
    return
  endif
envi_file_query, fid1, fname=fname1, ns=ns, nl=nl, nb=nb
print, 'Reference: ', fname1

envi_select, title='Choose (subset of) target image', fid=fid2, dims=dims2,pos=pos2
  if (fid2 eq -1) then begin
    print, 'cancelled'
    return
  endif
envi_file_query, fid2, fname=fname2, bnames=bnames2, ns=ns, nl=nl, nb=nb,
  xstart=xstart2, ystart=ystart2
print, 'Target:      ', fname2
print, '-----'
bnames = 'renorm('+bnames2[pos2]+' )'

if ((dims1[2]-dims1[1]) ne (dims2[2]-dims2[1])) or $
  ((dims1[4]-dims1[3]) ne (dims2[4]-dims2[3])) or $
  (n_elements(pos1) ne n_elements(pos2)) or $

```

```

; (min(pos1 eq pos2) eq 0) or $
  (n_elements(pos1) lt 2) then begin
    print, 'Dimensions of the two images are different. Aborting.'
    Message, 'Dimensions are different. Aborting.'
    return
  endif

envi_select, title='Choose (spatial subset of) chi_square image', fid=fid3,
  dims=dims3,pos=pos3,/band_only
  if (fid3 eq -1) then begin
    print, 'cancelled'
    return
  endif
envi_file_query, fid3, fname=fname3, ns=ns, nl=nl, nb=nb3
print, 'chi-square image: ',fname3

if ((dims1[2]-dims1[1]) ne (dims3[2]-dims3[1])) or $
  ((dims1[4]-dims1[3]) ne (dims3[4]-dims3[3])) then begin
  print, 'Dimensions of the two images are different. Aborting.'
  Message, 'Dimensions are different. Aborting.'
endif

; read and plot chi-square histogram out to 5 times 99.9 percentile
p999=chisqr_cvf(0.001,nb3-1)
chi_sqr = envi_get_data(fid=fid3,dims=dims3,pos=pos3)
ncp = 1.0 - chisqr_pdf(chi_sqr,nb3-1)
X = randomu(seed,nb3-1,n_elements(chi_sqr),/normal)
X2 = total(X^2,1)
hst = histogram(X2, min=0.0, max=5*p999, nbins=500)
hst1 = histogram(chi_sqr, min=0.0, max=5*p999, nbins=500)
envi_plot_data,findgen(500)*5*p999/500,[[hst],[hst1]], $
  title='Chi-Sqr Image',$
  ytitle='Frequency', $
  xtitle='chi_sqr'

; select no-change
base = widget_auto_base(title='No-change threshold')
wg = widget_slider(base, title='Select minimum no-change probability (%)', min=50,
  max=999, scale=1, $
  value=950, dt=1, uvalue='slide', xs=[300,7],/auto)
result = auto_wid_mng(base)
if (result.accept eq 0) then begin
  print, 'cancelled'
  return
endif
prct = float(result.slide)/1000.0
mask_nochange = where(ncp ge prct)

; output destination
base = widget_auto_base(title='Save normalized image')
sb = widget_base(base, /row, /frame)
wp = widget_outfm(sb, uvalue='outf', /auto)
result = auto_wid_mng(base)

widget_control,/hourglass

num_columns = dims1[2]-dims1[1]+1
num_rows = dims1[4]-dims1[3]+1
num_pixels = (num_columns*num_rows)
num_bands = n_elements(pos1)

; tie point
map_info2 = envi_get_map_info(fid=fid2)
envi_convert_file_coordinates, fid2, dims2[1], dims2[3], e, n, /to_map
map_info2.mc[2:3]= [e,n]

image3=fltarr(num_bands,num_pixels)

mrs=fltarr(num_bands) ; means of reference image

```

```

mts=fltarr(num_bands) ; means of target image
mns=fltarr(num_bands) ; means of normalized image
vrs=fltarr(num_bands) ; variances of reference image
vts=fltarr(num_bands) ; variances of target image
vns=fltarr(num_bands) ; variances of normalized image
df=intarr(num_bands) ; degrees of freedom
aa= fltarr(num_bands) ; slope of orthogonal regression curve
xm = fltarr(num_bands); mean of X
ym = fltarr(num_bands); mean of Y
table = fltarr(9,num_bands) ; table of fit parameters
t_stat = fltarr(num_bands)
t_signif = fltarr(num_bands)
f_stat = fltarr(num_bands)
f_signif = fltarr(num_bands)

; holdout 1/3 for test purposes
indices1 = where(indgen(n_elements(mask_nochange)) mod 3,complement=indices2)
mask_test = mask_nochange[indices2]
mask_train = mask_nochange[indices1]
n_test = n_elements(mask_test)
n_train = n_elements(mask_train)
n_nochange = n_elements(mask_nochange)

; export regression points to roi
envi_delete_rois, envi_get_roi_ids()
roi_id = envi_create_roi(color=4, name='RadCal regression pixels', ns=num_columns,
                        nl=num_rows)
xpts = mask_train mod num_columns
ypts = (mask_train - (mask_train mod num_columns))/num_columns
envi_define_roi, roi_id, /point, xpts=xpts, ypts=ypts

print, 'orthogonal regression using ',n_train,' no-change pixels ...'

for i=0,num_bands-1 do begin
  image1 = envi_get_data(fid=fid1,dims=dims1,pos=pos1[i])
  Y = (image1)[mask_train]
  image2 = envi_get_data(fid=fid2,dims=dims2,pos=pos2[i])
  X = (image2)[mask_train]
  window,10+i,xsize=500,ysize=400,xpos=100+10*i,ypos=100+10*i,title="Regression on Band
    "+strtrim(pos1[i]+1,2)
  wset,10+i
  plot, X, Y, pSym=3,title='Band '+strtrim(pos1[i]+1,2),xtitle="X",ytitle="Y", $
    color=0,background='FFFFFF'XL,xrange=[0,max(X)],yrange=[0,max(Y)]
; orthogonal regression on ax+c
  ortho_regress, transpose(X), transpose(Y), ai, xmi, ymi, sigma_aa, sigma_bb, sigma,
    rank=rank
  aa[i]=ai
  xm[i]=xmi
  ym[i]=ymi
  x0 = min(X) < 0.0
  x1 = max(X)
  Oplot,[x0,x1],[ym[i]-aa[i]*xm[i]+x0*aa[i],ym[i]-aa[i]*xm[i]+x1*aa[i]],color=0
  table[0,i] = i+1 ; band
  table[1,i] = ym[i]-aa[i]*xm[i]; intercept
  table[2,i] = sigma_aa ; standard error in intercept
  table[3,i] = aa[i] ; slope
  table[4,i] = sigma_bb ; standard error in slope
  table[5,i] = correlate(X,Y) ; correlation
  table[6,i] = sigma ; RMSE
  table[7,i] = rank[0] ; Spearman rank-order correlation
  table[8,i] = rank[1] ; significance of differene from zero
; renormalize
  image3[i,*]=ym[i]-aa[i]*xm[i]+image2*aa[i]
endfor
print, 'band, a, sigma_a, b, sigma_b, r, rmse, R, s'
print, table

; determine means and standard deviations
; using holdout test pixels only
reference = fltarr(num_bands,n_test)

```

```

target      = fltarr(num_bands,n_test)
normalized = fltarr(num_bands,n_test)
for i=0,num_bands-1 do begin
    reference[i,*] = (envi_get_data(fid=fid1,dims=dims1,pos=pos1[i]))[mask_test]
    target[i,*]    = (envi_get_data(fid=fid2,dims=dims2,pos=pos2[i]))[mask_test]
    normalized[i,*] = (image3[i,*])[mask_test]
    mrs[i]=mean(reference[i,*])
    mts[i]=mean(target[i,*])
    mns[i]=mean(normalized[i,*])
    vrs[i]=variance(reference[i,*])
    vts[i]=variance(target[i,*])
    vns[i]=variance(normalized[i,*])
    tm = tm_test(reference[i,*],normalized[i,*],/paired)
    t_stat[i] = tm[0]
    t_signif[i] = tm[1]
    fv = fv_test(reference[i,*],normalized[i,*])
    f_stat[i] = fv[0]
    f_signif[i] = fv[1]
endfor
; determine covariance matrices and do Wishart test
Sr = correlate(reference,/double,/covariance)
Sn = correlate(normalized,/double,/covariance)
w_signif = wishart(mrs,mns,Sr,Sn,n_test,statistic=Q)
print,"Comparison Statistics using",n_test," test pixels"
print,"Means"
print,"target      ",mts
print,"reference    ",mrs
print,"normalized    ",mns
print,"t-statistic   ",t_stat
print,"p-value      ",t_signif
print,"Variances"
print,"target      ",vts
print,"reference    ",vrs
print,"normalized    ",vns
print,"F-statistic   ",f_stat
print,"p-value      ",f_signif
print,"Wishart      ",Q
print,"p-value     ",W_signif

envi_info_wid,['RADCAL statistics: '+strtrim(n_train,2)+' training and
'+strtrim(n_test,2)+' test pixels'], $
    [' band      intercept sigma      slope      sigma      r      RMSE
R          s'], $
    string( table, format='(9F10.4)'), $
'Means', ['target
'+string(mts,format='('+strtrim(num_bands)+'F10.4)')], $
    ['reference
'+string(mrs,format='('+strtrim(num_bands)+'F10.4)')], $
    ['normalized
'+string(mns,format='('+strtrim(num_bands)+'F10.4)')], $
    ['t-stat
'+string(t_stat,format='('+strtrim(num_bands)+'F10.4)')], $
    ['p-value
'+string(t_signif,format='('+strtrim(num_bands)+'F10.4)')], $
'Variances',['target
'+string(vts,format='('+strtrim(num_bands)+'F12.4)')], $
    ['reference
'+string(vrs,format='('+strtrim(num_bands)+'F12.4)')], $
    ['normalized
'+string(vns,format='('+strtrim(num_bands)+'F12.4)')], $
    ['F-stat
'+string(f_stat,format='('+strtrim(num_bands)+'F12.4)')], $
    ['p-value
'+string(f_signif,format='('+strtrim(num_bands)+'F12.4)')], $
    ['Wishart      '+string(Q,format='(F10.4)')], $
    ['p-value      '+string(W_signif,format='(F10.4)')]], $
    title='RADCAL statistics'

; save to memory or disk
out_array = bytarr(num_columns,num_rows,num_bands)

```

```

for i = 0,num_bands-1 do begin
    temp = reform(image3[i,*],num_columns,num_rows,/overwrite) > 0.0
    out_array[*,* ,i] = temp < 255.0
endfor
if (result.accept eq 0) then begin
    print, 'output cancelled'
end else if (result.outf.in_memory eq 1) then begin
    envi_enter_data, out_array,
        bnames=bnames,map_info=map_info2,xstart=xstart2+dims2[1],ystart=ystart2+dims2[3]
    print, 'result written to memory'
end else begin
    openw, unit, result.outf.name, /get_lun
    band_names=strarr(num_bands)
    for i=0,num_bands-1 do writeu, unit, out_array[*,* ,i]
    envi_setup_head ,fname=result.outf.name, ns=num_columns, $
        map_info=map_info2, $
        xstart=xstart2+dims2[1], ystart=ystart2+dims2[3], $
        nl=num_rows, nb=num_bands, $
        data_type=1, interleave=0, /write, $
        bnames=bnames, $
        descrip='RadCal normalized'
    print, 'file created ', result.outf.name
    close, unit
endelse

; process another file (e.g. a full scene), writing to disk only

answer=dialog_message('Do you want to use present results to calibrate another file?',
    /question)
if answer eq 'Yes' then begin
    envi_select, title='Choose file to be calibrated', fid=fid, dims=dims,pos=pos
    if (fid eq -1) then begin
        print,'cancelled'
        return
    endif
    if n_elements(pos) ne num_bands then begin
        print, 'incorrect spectral dimension'
        Message, 'Spectral dimensions are different. Aborting.'
        return
    endif
    envi_file_query, fid, fname=fname, bnames=bnames, interleave=interleave, $
        xstart=xstart, ystart=ystart
    map_info = envi_get_map_info(fid=fid)
    bnames = 'renorm('+bnames+')'
    print,'calibrating file '+fname
    if n_elements(pos1) ne n_elements(pos2) then begin
        print, 'Spectral dimension of file does not match. Aborting.'
        Message, 'Dimensions are different. Aborting.'
        return
    endif
    num_columns = dims[2]-dims[1]+1
    num_rows = dims[4]-dims[3]+1
    base = widget_auto_base(title='Choose output file name')
    sb = widget_base(base, /row, /frame)
    wp = widget_outf(sb, uvalue='outf', /auto)
    result = auto_wid_mng(base)
    if (result.accept eq 0) then begin
        print, 'output cancelled'
        return
    endif
    outfilename = result.outf
    openw, unit, outfilename, /get_lun
    progressbar = Obj_New('progressbar', Color='blue', Text='0',$
        title='processing tile ...',xsize=250,ysize=20)
    progressbar->start
; get tile id
    tile_id = envi_init_tile(fid,pos,num_tiles=num_tiles, $
        interleave=interleave,xs=dims[1],xe=dims[2],ys=dims[3],ye=dims[4])
; start tiling

```

```

yml = fltarr(num_bands,num_columns)
xml = fltarr(num_bands,num_columns)
aal = fltarr(num_bands,num_columns)
for i=0,num_columns-1 do begin
    yml[*,i] = ym
    xml[*,i] = xm
    aal[*,i] = aa
endfor
for tile_index=0L,num_tiles-1 do begin
    tile = envi_get_tile(tile_id,tile_index,band_index=k)
    case interleave of
        0: imagel = ym[k] - aa[k]*xm[k] + aa[k]*tile
        1: imagel = transpose(yml-aal+xml) + transpose(aal)*tile
        2: imagel = yml - aal+xml + aal*tile
    endcase
    if progressbar->CheckCancel() then begin
        print,'Calibration aborted'
        free_lun, unit
        progressbar->Destroy
        return
    endif
    progressbar->Update,tile_index*100/num_tiles,text=strtrim(tile_index,2)
    writeu,unit,imagel
endfor
; tidy up
free_lun, unit
envi_tile_done, tile_id
progressbar->Destroy
envi_setup_head,fname=outfilename, ns=num_columns, $
    nl=num_rows, nb=num_bands, $
    data_type=1, $
    file_type=0,map_info=map_info, $
    interleave=interleave, /write, $
    xstart=xstart+dims[1], $
    ystart=ystart+dims[3], $
    bnames=bnames
print, 'calibrated file created ', outfilename
endif
end

```

APÊNDICE B – CÓDIGOS IMPLEMENTADOS EM LINGUAGEM R

B.1 Algoritmo para gerar os coeficientes ajustados de toda a imagem

```
#####  
#           Implementado por: Márcio Pupin de Mello           #  
#           #           #           #           #           #  
#           Mestrado em Sensoriamento Remoto (INPE) 2007-2009 #  
#           #           #           #           #           #  
#           email: marciopupin@yahoo.com.br                 #  
#           #           #           #           #           #  
#           data: abril/2009                               #  
#####  
  
#####  
#Algoritmo que:                                           #  
#  monta a estrutura de pontos de controle                 #  
#  Calcula as superfícies de todos os pixels da imagem toda #  
#  Grava os valores dos coeficientes ajustados num arquivo #  
#  de saída chamado "imagens0.cof"                         #  
#####  
  
#Informa na tela a data e hora do início do procedimento de processamento  
cat(paste("\nInício do procedimento em",Sys.time()))  
  
#####  
#dados necessarios  
#####  
  
#diretório de trabalho  
#esse dado deve ser informado com exatidão por cada usuário  
setwd("F:##MESTRADO/CLASSIFICACOES/C1/STRS")  
  
#necessário para não usar notação científica nos números  
options(scipen=10000)  
  
#obtendo os nomes dos arquivos de imagens  
#as imagens devem ser nomeadas com um padrão:  
#para a data 03/05/2007, por exemplo,  
#a imagem deve estar em formato generic binary  
#e ter o nome "20070503" sem nenhuma extensão de arquivo  
#as imagens devem estar dentro de uma subpasta chamada "imagens_gen_bin_BIP"  
arquivos.de.imagens<-sort(dir("./imagens_gen_bin_BIP"))  
  
#informando ao programa as datas das imagens (toda devem ser do mesmo ano)  
datas<-as.Date(arquivos.de.imagens,format="%Y%m%d")  
  
#informa de qual linha da imagem deve-se iniciar o processamento  
#esse objeto é útil quando deseja-se dividir o processamento da imagem  
linha.inicial<-1  
  
#obtendo informações necessárias  
nimagens<-length(arquivos.de.imagens)  
nlinhas<-500  
ncolunas<-500  
nbandas<-6  
valor.maximo<-255 #utilizando imagens de 8bits  
  
#o grau do polinômio é dado pelo maior grau possível  
grau.do.pts<-min(c(nimagens,nbandas))-1  
  
#####  
#algumas funções  
#####  
  
#####  
#função que faz o re-escalonamento de conjuntos de dados
```

```

reescala<-function(obj,valmin=0,valmax=1)
{
  n<-length(obj)
  obj2<-rep(9999,n)
  for (i in 1:n)
    obj2[i]<-((obj[i]-min(obj))/(max(obj)-min(obj)))*(valmax-valmin)+valmin
  return(obj2)
}

#####
#funcao que gera o modelo matematico de grau d para ser usado
fpol<-function(grau)
{
  nt<-(((grau+1)*(grau+2))/2)-1
  termos1<-rep("vazio",nt)
  cont<-1
  for (j in 1:grau)
    for (i in j:0)
      {
        termos1[cont]<-paste("I(V1^",i,"*V2^",(j-i),")",sep="")
        cont<-cont+1
      }
  f1<-as.formula(paste("V3~",paste(termos1,collapse="+")))
  return(f1)
}

#####
#criando as conexões para os arquivos
#####

#esse procedimento é realizado porque, dependendo do tamanho
#das imagens não há espaço de memória suficiente para elas
#assim, as conexões criam acessos às imagens direto do HD

for (img in 1:nimagens)
{
  comando<-paste("conexao",img,"<-
    file('./imagens_gen_bin_BIP/",arquivos.de.imagens[img],"',open='rb')",sep="
    ")
  eval(parse(text = comando))
}

#Caso precise continuar a análise de uma certa linha qualquer...
for (img in 1:nimagens)
{
  comando<-paste("seek(conexao",img,"where=",((linha.inicial-
    1)*nbandas*ncolunas),"origin='start')",sep="")
  eval(parse(text = comando))
}

#####
#montando a estrutura dos pontos de controle do pixel
#####

#apenas para escrever algumas informações no arquivo a ser gerado
#informação do número do pixel analisado e do número de coeficientes ajustados para o
mesmo
auxiliar.PTS<-c(0,((grau.do.pts+1)*(grau.do.pts+2))/2)

#monta a equacao usada no modelo de interpolação de grau igual a "grau.do.pts"
equacao<-fpol(grau.do.pts)

#colunal = coluna do tempo
colunal<-as.numeric(as.vector(gl(nimagens,nbandas,labels=round(
  reescala(as.numeric(julian(as.Date(arquivos.de.imagens,format="%Y%m%d"))),
  6))))))

#coluna2 = coluna do comprimento de onda (OBS: dados para o sensor Landsat-TM)

```

```

coluna2<-round(rep(c(0,.043353,.101156,.199422,.670520,1),nimagens),6)

#caso utilize a nova metodologia de disposiçao de valores no eixo do comprimento de
ondas
#coluna2<-round(rep(seq(0,1,((1-0)/(nbandas-1))),nimagens),6)

#criando a estrutura da coluna das reflectancias (os valores serao lidos das imagens)
coluna3<-rep(0,nimagens*nbandas)

#procedimento completo em loop que le os valores da coluna 3 e interpola as superficies
for (linha in linha.inicial:nlinhas)
{
  for (coluna in 1:ncolunas)
  {

    #le os valores da coluna 3 (arredondando para 6 casas decimais)
    for (img in 1:nimagens)
    {
      comando<-paste("coluna3[",((img-1)*nbandas)+1,":",((img-1)*nbandas)+nbandas,")<-
        round(readBin(conexao",img,",integer(),n=",nbandas,",size=1,signed=F)/
          valor.maximo,6)",sep="")
      eval(parse(text = comando))
    }

    #monta o conjunto de dados de pontos de controle do pixel a ser processado
    pts.de.controle<-data.frame(V1=coluna1,V2=coluna2,V3=coluna3)
    numero.do.pixel<-((linha-1)*ncolunas)+coluna
    auxiliar.PTS[1]<-numero.do.pixel

    #relativo ao Polinomial Trend Surface (PTS)
    coeficientes.ajustados<-
      round(as.numeric(lm(equacao,data=pts.de.controle)$coefficients),6)

    #escrever o arquivo de saida PTS (de nome "imagems0.cof")

      write.table(t(auxiliar.PTS),"imagems0.cof",col.names=F,row.names=F,quote=F
        ,append=T)
    write.table(t(coeficientes.ajustados),"imagems0.cof",col.names=F,
      row.names=F,quote=F,append=T)

  }

  #informa ao usuario em qual a ultima linha da imagem processada
  cat(paste("\nProcessada a linha",linha,"em",Sys.time()))

  #permite que as informacoes da linha atual seja exibidas mesmo quando a opcao "saida
  com buffer" estiver ativa
  flush.console()

  #libera a memoria livre
  gc()

}

#Informa na tela a data e hora do fim do procedimento de processamento
cat(paste("\n\nFim do procedimento em",Sys.time()))

```

B.2 Função para interpolar os coeficientes utilizando arquivos *.TSA

```

#####
#          Implementado por: Márcio Pupin de Mello          #
#                                                                 #
#          Mestrado em Sensoriamento Remoto (INPE) 2007-2009  #
#                                                                 #
#          email: marciopupin@yahoo.com.br                  #
#                                                                 #
#          data: abril/2009                                  #
#####

```

```
#####
#função completa
#####

geraPTS<-function()
{
  #algumas informações

  #nc é o número de classes
  #np é o número de pixels
  #d é o grau do polinômio
  #cam é o caminho do diretório de trabalho
  #observar a barra que não é invertida nem no R para Windows!!!
  #não colocar "/" após a última sub-pasta!
  #nom á a parte do nome comum aos arquivos TSA
  #observar que a parte "ls" não é dado de entrada necessário!
  #naotreina é uma variável booleana que indica nc=0, ou seja, não é treinamento.
  #nesse caso, trata-se do processamento de toda a imagem lendo os arquivos TSA

  #apenas para nunca escrever em notação científica
  options(scipen=1000)

  #lendo informações úteis

  #sobre o diretório de trabalho (caminho)
  cat("\nO diretório de trabalho atual é:\n",getwd(),"\n\n")
  resp<-"0" #default
  resp<-readline("Deseja alterá-lo? (0-NÃO ; 1-SIM): ")
  if(resp=="1")
  {cam<-readline("Digite o novo caminho (Ex.: C:/novo/caminho): ")
  setwd(cam)
  if(getwd()==cam) cat("\nNovo diretório de trabalho:\n",getwd(),"\n\n")}

  nom<-readline("Nome dos TSA (Ex.: para treinamentols1_1.TSA; etc... digite:
  treinamento / default: treina): ")
  if (nom=="") nom<-"treina" #default

  cat("\n")#apenas passa p/ linha de baixo

  d<-as.numeric(readline("          Graus do polinômio: "))
  if (d<1||d>20) stop ("Grau não aceito")

  nc<-as.numeric(readline("Número de Classes (0 - imagem): "))
  naotreina<-F
  if (nc==0)
  {nc<-1
  naotreina<-T} else if (nc<0) stop ("O número de classes deve ser no mínimo 0")

  np<-as.numeric(readline("Número de Pixels por Classe (se estiver processando a
  IMAGEM entre com o numero de pixels total da imagem): "))
  if (np<1) stop ("O número de pixels por classe deve ser no mínimo 1")

  cat("\n")#apenas passa p/ linha de baixo

  if (naotreina==F)
  {
  metodo<-as.numeric(readline("Qual método para dispor o eixo de comprimento de
  onda? (1) Re-escalonado (2) Equidistante : "))
  cat("\n")#apenas passa p/ linha de baixo

  nsup<-as.numeric(readline("Quantas superfícies por classe deseja gerar? (0 para
  nenhuma): "))
  if (nsup>np) stop (paste("O número máximo de superfícies a gerar deve ser menor
  ou igual a",np))} else Uni

  #ncoef é o número de coeficientes gerados por cada pixel (superfície)
  ncoef<-((d+1)*(d+2))/2

```

```

#função que gera o modelo matemático de grau d p/ ser usada no comando lm()
fpol<-function(grau)
{
  nt<-(((grau+1)*(grau+2))/2)-1
  termos1<-rep("vazio",nt)
  termos2<-termos1
  cont<-1
  for (j in 1:grau)
    for (i in j:0)
      {
        termos1[cont]<-paste("I(V1^",i,"*V2^",j-i,")",sep="")
        termos2[cont]<-paste("pts$coefficients[[",cont+1,"]] *x^",i,"*y^",j-
          i,sep="")
        cont<-cont+1
      }
  f1<-as.formula(paste("V3~",paste(termos1,collapse="+")))
  f2<-paste("pts$coefficients[[1]]+",paste(termos2,collapse="+"))
  f<-list(f1,f2)
  return(f)
}

#objeto que recebe os coeficientes temporariamente
aux<-c(0,0)

#objeto auxiliar
parar.por.NA<- "qqcoisa"

#coração da função
for (classe in 1:nc)
  for (p in 1:np)
    {
      #le os arquivos de dados
      if (naotreina==T) dados<-read.table(paste(nom,"ls",classe-1,
        "_",p,".TSA",sep=""),h=F,skip=1) else
        dados<-read.table(paste(nom,"ls",classe,"_",p,".TSA",sep=""),h=F,skip=1)

      #infere sobre o método da distribuição dos níveis no eixo do comprimento de
      onda (método 2 = valores equidistantes)
      if (metodo==2)
        {
          nbandas<-length(levels(as.factor(dados[,2])))
          ndatas<-length(levels(as.factor(dados[,1])))
          dados[,2]<-round(rep(seq(0,1,1/(nbandas)-1),ndatas),6)
        }

      #informa sobre a possibilidade de coeficientes não ajustados e sobre proceder
      ao ajuste
      if (classe==1 && p==1 && naotreina==F)
        {
          if (ncoef>=(length(dados[[1]])-1)) print("Não é possível gerar a PTS com
            esse grau\nArquivos de entrada com dados insuficientes.")
          if
            ((length(levels(factor(dados[[1]]))<=d)||length(levels(factor(dados[[2]]
              ))<=d)||length(levels(factor(dados[[3]]))<=d)) parar.por.NA<-
              as.numeric(readline("Possivelmente alguns coeficientes não serão ajustados
                (aparecerá NA no resultado)... deseja continuar? (0-NÃO ; 1-SIM):"))
            if (parar.por.NA=="0") stop ("Abortado pelo usuário")
          }
        }

      #gera os coeficientes do polinômio ajustado pelos mínimos quadrados
      funcao<-fpol(d)
      pts<-lm(funcao[[1]],data=dados)

      #escrever o arquivo de saída
      ctmp<-round(as.numeric(pts$coefficients),6)
      aux[1]<-p
      aux[2]<-ncoef
      if (naotreina==T)

```

```

        {write.table(t(aux),paste(nom,"ls",classe-1,".cof",sep=""),sep=" ",
col.names=F,row.names=F,quote=F,append=T)
write.table(t(ctmp),paste(nom,"ls",classe-1,".cof",sep=""),sep=" ",
col.names=F,row.names=F,quote=F,append=T)
} else {
write.table(t(aux),paste(nom,"ls",classe,".cof",sep=""),sep=" ",
col.names=F,row.names=F,quote=F,append=T)
write.table(t(ctmp),paste(nom,"ls",classe,".cof",sep=""),sep=" ",
col.names=F,row.names=F,quote=F,append=T)
}

#gera as superfícies quando for pertinente
if(nsup>0 && p<=nsup)
{
funcao.ajustada<-function(x,y) {eval(parse(text = funcao[[2]]))}
superficie<-round(outer(seq(min(dados$V1),max(dados$V1)-
min(dados$V1))/100),seq(min(dados$V2),max(dados$V2)-
min(dados$V2))/100),funcao.ajustada),6)
write.table(superficie,paste(nom,"ls",classe,"_",p,"___R.dat",sep=""),
col.names=F,row.names=F,quote=F)
}

#informa em que passo esta a operação caso esteja processando a imagem toda
if ((naotreina==T) && (.Platform$OS.type=="windows") && (p%1000==0))
{cat(paste("\nArquivo ",nom,"ls",classe-1,"_",p,".TSA
processado...",sep=""))
flush.console()}

}

cat("\nOs arquivos de saída foram gerados com êxito!\n")
}

```

B.3 Função que calcula índices de exatidão, suas variâncias e testes Z

```

#####
# Implementado por: Márcio Pupin de Mello #
# #
# Mestrado em Sensoriamento Remoto (INPE) 2007-2009 #
# #
# email: marciopupin@yahoo.com.br #
# #
# data: abril/2009 #
#####

#####
#Algoritmo que: #
# calcula índices (e suas respectivas variâncias): #
# - Exatidão Global #
# - Kappa #
# - Tau #
# Procedo o teste Z para avaliar as classificações: #
# - individualmente #
# - aos pares #
#####

#####
#função que calcula índices com base na matriz de confusão
#####

calcula.indices<-function()
{

#configuração para nunca utilizar notação científica para valores
options(scipen=1000)

#quantas classes tem o mapa temático?

```

```

nc<-as.numeric(readline("\nQuantas classes possui o mapa temático a ser avaliado?:
"))
m1<-matrix(rep(-1,nc^2),ncol=nc)

#entrando com os dados da matriz de confusão
for (i in 1:nc)
  for (j in 1:nc)
    m1[i,j]<-as.numeric(readline(paste("\nEntre com o elemento (",i,",",j,"):
",sep="")))

m2<-cbind(m1,margin.table(m1,1));m2<-rbind(m2,margin.table(m2,2))

#calculos dos tetas
N<- sum(m1)
tetal<- sum(diag(m1)) / N
teta2<- sum(m2[,nc+1][1:nc]*m2[nc+1,][1:nc]) / (N^2)
teta3<-sum(diag(m1)*(m2[,nc+1][1:nc]+m2[nc+1,][1:nc])) / (N^2)
teta4<-0
  for (i in 1:nc)
    for (j in 1:nc)
      teta4<-teta4 + (m1[i,j] * ( (sum(m1[j,]) + sum(m1[,i])) ^2));
      teta4<-teta4 / (N^3)

#calculos das estatisticas
exatidao.global<-tetal
var.exatidao.global<-(tetal*(1-tetal))/N
kappa<-(tetal-teta2)/(1-teta2)
var.kappa<- (1/N) * ( ((tetal*(1-tetal))/((1-teta2)^2)) + ((2*(1-
tetal)*((2*tetal*teta2)-teta3))/((1-teta2)^3)) + (((1-tetal)^2)*(teta4-
(4*(teta2^2)))/((1-teta2)^4)) )
tau<-(tetal-(1/nc))/(1-(1/nc))
var.tau<-(tetal/N)*((1-tetal)/((1-(1/nc))^2))

cat("\n\nProcessado!\n\n")

resultado<-
  list(matriz.de.confusao=m1,exatidao.global=exatidao.global,var.exatidao.glo
bal=var.exatidao.global,kappa=kappa,var.kappa=var.kappa,tau=tau,var.tau=var
.tau)
return(resultado)
}

#####
#função que procede o teste Z
#####

calcula.Z<-function()
{

#configuração para nunca utilizar notação científica para valores
options(scipen=1000)

#sobre o numero de classificações avaliadas
nclassif<-as.numeric(readline("\nQuantas classificações serão computadas?: "))

#pergunta sobre digitar os valores do índice e suas variâncias ou ler de um objeto da
função "calcula.indices()"
ler.existente.ou.digitar<-as.numeric(readline("\n\nDeseja digitar os valores ou ler
de objeto(s) da funcao 'calcula.indices'? \n (1) Ler de objetos(s)\n (2)
Digitar os valores\n\n Opção: "))

#pergunta sobre qual índice deseja utilizar
qual.indice<-as.numeric(readline("\n\nQual índice de exatidão deseja utilizar?\n
(1) Exatidão Global\n (2) Kappa\n (3) Tau\n\n Opção: "))
if(qual.indice==1)
  dados<-data.frame(nome=rep("vazio",nclassif),exat.global=rep(-1,
nclassif),var.exat.global=rep(-1,nclassif)) else
  if (qual.indice==2)

```

```

        dados<-data.frame(nome=rep("vazio",nclassif),kappa=rep(-1,
            nclassif),var.kappa=rep(-1,nclassif)) else
            if (qual.indice==3)
                dados<-data.frame(nome=rep("vazio",nclassif),tau=rep(-1,
                    nclassif),var.tau=rep(-1,nclassif))
dados$nome<-as.vector(dados$nome)

#procede os testes Z
for (i in 1:nclassif)
    {
        if (ler.existente.ou.digitar==1)
            {
                lixo<-readline(paste("\nNome do objeto ",i," : ",sep=""))
                dados[i,1]<-lixo
                comando1<-paste("dados[i,2]<-",lixo,"[[" ,((qual.indice-1)*2)+2,"]]",sep="")
                eval(parse(text=comando1))
                comando2<-paste("dados[i,3]<-",lixo,"[[" ,((qual.indice-1)*2)+3,"]]",sep="")
                eval(parse(text=comando2))
            } else
                {
                    dados[i,1]<-readline(paste("\nNome da classificacao ",i," : ",sep=""))
                    dados[i,2]<-as.numeric(readline(paste("Valor do Kappa",i," : ",sep="")))
                    dados[i,3]<-as.numeric(readline(paste("Valor da variância do Kappa",i," : ",
                        sep="")))
                }
    }

m1<-matrix(rep(-1,nclassif^2),ncol=nclassif,dimnames=list(dados[,1],dados[,1]))

for (i in 1:nclassif)
    for (j in 1:nclassif)
        if (i==j) m1[i,j]<-dados[i,2]/sqrt(dados[i,3]) else m1[i,j]<-
            abs(dados[i,2]-dados[j,2])/sqrt(dados[i,3]+dados[j,3])
m1<-round(m1,6)
m1[upper.tri(m1,diag=F)]<-NA

m2<-matrix(rep(-1,nclassif^2),ncol=nclassif,dimnames=list(dados[,1],dados[,1]))
for (i in 1:nclassif)
    for (j in 1:nclassif)
        m2[i,j]<-(1-pnorm(m1[i,j]))*2
m2<-round(m2,6)
m2[upper.tri(m1,diag=F)]<-NA

#objeto que contem o resultado dos procedimentos
resultado<-list(classificacoes=dados,estatisticaZ=m1,valor.p=m2)
return(resultado)
}

```