



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21b/2016/03.01.17.20-TDI

**ANÁLISE DE IMAGENS ORIENTADA A OBJETOS E
AMOSTRAGEM ESTATÍSTICA NO MONITORAMENTO
DE CANA-DE-AÇÚCAR, MILHO E SOJA NO ESTADO
DE SÃO PAULO**

Bruno Schultz

Tese de Doutorado do Curso de
Pós-Graduação em Sensoriamento
Remoto, orientada pelos Drs.
Antonio Roberto Formaggio,
e Alfredo José Barreto Luiz,
aprovada em 07 de março de 2016.

URL do documento original:

[<http://urlib.net/8JMKD3MGP3W34P/3L9BUQ5>](http://urlib.net/8JMKD3MGP3W34P/3L9BUQ5)

INPE
São José dos Campos
2016

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) 3208-6923/6921

Fax: (012) 3208-6919

E-mail: pubtc@inpe.br

**COMISSÃO DO CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO
DA PRODUÇÃO INTELECTUAL DO INPE (DE/DIR-544):****Presidente:**

Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação (CPG)

Membros:

Dr. Plínio Carlos Alvalá - Centro de Ciência do Sistema Terrestre (CST)

Dr. André de Castro Milone - Coordenação de Ciências Espaciais e Atmosféricas (CEA)

Dra. Carina de Barros Melo - Coordenação de Laboratórios Associados (CTE)

Dr. Evandro Marconi Rocco - Coordenação de Engenharia e Tecnologia Espacial (ETE)

Dr. Hermann Johann Heinrich Kux - Coordenação de Observação da Terra (OBT)

Dr. Marley Cavalcante de Lima Moscati - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Silvia Castro Marcelino - Serviço de Informação e Documentação (SID)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon

Clayton Martins Pereira - Serviço de Informação e Documentação (SID)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

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

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

EDITORAÇÃO ELETRÔNICA:

Marcelo de Castro Pazos - Serviço de Informação e Documentação (SID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SID)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21b/2016/03.01.17.20-TDI

**ANÁLISE DE IMAGENS ORIENTADA A OBJETOS E
AMOSTRAGEM ESTATÍSTICA NO MONITORAMENTO
DE CANA-DE-AÇÚCAR, MILHO E SOJA NO ESTADO
DE SÃO PAULO**

Bruno Schultz

Tese de Doutorado do Curso de
Pós-Graduação em Sensoriamento
Remoto, orientada pelos Drs.
Antonio Roberto Formaggio,
e Alfredo José Barreto Luiz,
aprovada em 07 de março de 2016.

URL do documento original:

[<http://urlib.net/8JMKD3MGP3W34P/3L9BUQ5>](http://urlib.net/8JMKD3MGP3W34P/3L9BUQ5)

INPE
São José dos Campos
2016

Dados Internacionais de Catalogação na Publicação (CIP)

Schultz, Bruno.

Sc82a Análise de imagens orientada a objetos e amostragem estatística no monitoramento de cana-de-açúcar, milho e soja no estado de São Paulo / Bruno Schultz. – São José dos Campos : INPE, 2016.

xxxii + 210 p. ; (sid.inpe.br/mtc-m21b/2016/03.01.17.20-TDI)

Tese (Doutorado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2016.

Orientadores : Drs. Antonio Roberto Formaggio, e Alfredo José Barreto Luiz.

1. Estimativa agrícola. 2. Classificação visual. 3. Classificação automática. I.Título.

CDU 528.8:633(815.6)



Esta obra foi licenciada sob uma [Licença Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](#).

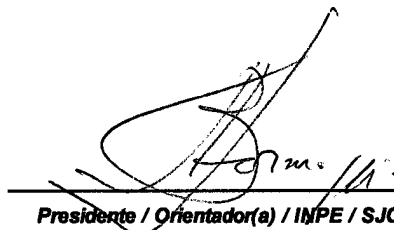
This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](#).

Aluno (a): **Bruno Schultz**

Título: " ANÁLISE DE IMAGENS ORIENTADA A OBJETOS E AMOSTRAGEM
ESTATÍSTICA NO MONITORAMENTO DE CANA-DE-AÇÚCAR, MILHO E SOJA NO
ESTADO DE SÃO PAULO".

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

Dr. Antonio Roberto Formaggio



Presidente / Orientador(a) / INPE / SJCampos - SP

Dr. Alfredo José Barreto Luiz



Orientador(a) / EMBRAPA / Jaguariúna - SP

Dra. Ieda Del' Arco Sanches



Membro da Banca / INPE / São José dos Campos - SP

Dra. Luciana Miura Sugawara



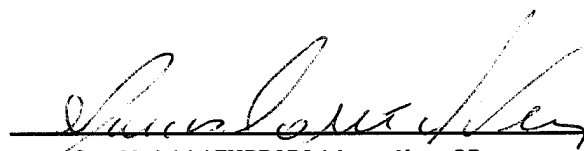
Membro da Banca / INPE / Cachoeira Paulista - SP

Dr. Júlio César Dalla Mora Esquerdo



Convidado(a) / EMBRAPA / Campinas - SP

Dr. Marcos Corrêa Neves



Convidado(a) / EMBRAPA / Jaguariúna - SP

Este trabalho foi aprovado por:

() maioria simples

☒ unanimidade

São José dos Campos, 07 de Março de 2016

“A imaginação é mais importante que a ciência, porque a ciência é limitada, ao passo que a imaginação abrange o mundo inteiro”.

Albert Einstein

A minha mulher Lígia Rampani Duarte e minha sogra Marta Solange Rampani pelo apoio incondicional dado nestes quatro anos de doutorado.

AGRADECIMENTOS

A meus pais Marli Schultz e Edson Luiz Schultz (*in memoriam*), minha irmã Sabrina Schultz e meus tios Élcio Schultz e Tânia Stafin Schultz pela educação oferecida para meu amadurecimento como pessoa e profissional, além de todo suporte psicológico dado em minha jornada acadêmica.

A meus sogros Marta Solange Rampani e Fernando Sergio Duarte e seus respectivos cônjuges Ricardo e Valéria, e toda família Rampani & Duarte que me ofereceram um lar e zelo familiar nestes quatro anos de estadia em São José dos Campos, hoje minha casa.

A minha mulher Lígia Rampani Duarte que me trata com muito amor, carinho, calma e se tornou minha guia para obter meus objetivos, nesta vida.

Ao Dr. Antonio Roberto Formaggio pelo grande profissional que é, e por todos os ensinamentos prestados no programa de doutoramento do INPE.

Ao Dr. Alfredo José Barreto Luiz por ter feito parte de minha carreira profissional e ter contribuído muito para que eu concluísse meu doutorado.

À doutora Ieda Del' Arco Sanches por me escutar e discutir temas não apenas relacionados à ciência.

Ao Dr. Clement Atzberger por me propiciar uma excelente viagem ao exterior que objetivou minha inserção no mundo científico internacional.

Ao Pesquisador Luiz Maurano pela ajuda prestada para concluir uma parte importante de minha pesquisa.

Aos amigos "Inpianos": Msc. Bruno Oliveira, MSc. Cesare Neto, Msc. Gabriel Bertani, Msc. Guilherme Mataveli, Msc. Isaque Eberhardt, Msc. Renato Ferreira, Msc. Sulimar Munira e Msc. Thiago Frank pela amizade e por me ajudarem a chegar ao final do processo de doutorado no INPE.

Aos amigos de BOKU/Viena: Dr. Anja Klisch, MSc. Javier Portillo, Msc. Markus Immitzer, Msc. Kathrin Einzmann e MSc. Wai Tim pela paciência e pelo afeto que tiveram comigo e com minha esposa, em nossa estadia em Viena.

Aos colegas: Dr. André Lima, Msc. Elizabeth Goltz, Dr. Juliana dos Santos, Dr. Kleber Trabaquini, Dr. Marcos Adami, Dr. Márcio Mello, MSc. Moises Salgado e MSc. Willian Foschiera, que "passaram" pelo INPE e contribuíram para minha carreira profissional.

A todos os professores e servidores do INPE, em especial ao professor Dr. Flávio Ponzoni que teve a bondade em me receber no INPE, na primeira vez que estive visitando as instalações do SERE.

Ao Conselho Nacional de Pesquisa (CNPq) pelo suporte financeiro concedido em 3,5 anos de doutorado no Brasil e meio ano de doutorado Sanduiche na Áustria.

A toda comunidade do sensoriamento remoto agrícola que publicou seus trabalhos, com o objetivo de facilitar o desenvolvimento da ciência.

RESUMO

A superfície do estado de São Paulo é recoberta por diversos tipos de usos e coberturas da terra, distribuídos em diferentes porções geográficas. Entre esses usos, a cana-de-açúcar, o milho e a soja ocupam a maior parte das áreas agrícolas do estado. Apesar de essas culturas possuírem grande importância para o cenário agrícola, elas ainda não são monitoradas sistematicamente em tempo quase-real. Como os dados de estimativas agrícolas possuem grande importância nas tomadas de decisão, a presente pesquisa teve como meta avaliar a possibilidade de se obter o valor de área ocupada pelas culturas de cana-de-açúcar, milho e soja no estado de São Paulo, em tempo-quase real (período bi-mensal), utilizando sensoriamento remoto (SR), classificação visual e automática de imagens multitemporais OLI. Três diferentes áreas-teste foram avaliadas: o Estado de São Paulo (**A1**), o município de Cândido Mota/SP (**A2**) e 15 municípios contínuos e localizados no extremo Sul de São Paulo (**A3**). Os objetivos específicos de cada área teste foram: (A1) informar o valor da área ocupada com cana-de-açúcar, milho e soja para o estado de São Paulo, em períodos de dois meses do ano safra 2014/2015; (A2) validar informações de campo, desenvolver um protocolo de classificação visual e avaliá-lo; e (A3) avaliar o desempenho da classificação automática. Para A1, 5.000 pixels aleatórios foram estratificados em cada mesorregião do estado, em cada período **P1** (Nov/Dez), **P2** (Jan/Fev), **P3** (Mar/Abr) e **P4** (Mai/Jun) do ano safra e classificados em laboratório, utilizando protocolo de classificação visual. Além disso, um banco de dados composto por 432 imagens multitemporais OLI foi montado e dividido em 24 mosaicos (n), salvos em composição colorida R-b5 G-b6 B-b4. Cada imagem do mosaico (composto por 18 cenas) foi padronizada através de um algoritmo nominado Aparência Equalizada (ApEq). A ApEq foi executada para que não ocorressem grandes alterações na cor resultante dos alvos pseudo-invariantes presentes nas diferentes imagens dos mosaicos. A análise visual dos resultados da ApEq foi discutida sobre o mosaico de nº 13. Em A2, 200 pixels OLI foram sorteados sobre o município de Cândido Mota/SP e validados em campo no período de 13 a 17 de abril/2015. Informações sobre as áreas ocupadas com cana-de-açúcar, milho e soja foram coletadas, e posteriormente utilizadas para avaliar o protocolo de classificação visual através das estatísticas geradas pela matriz de confusão (Exatidão Global). O protocolo de classificação visual foi aplicado por oito intérpretes, divididos em dois grupos - **G1**: intérpretes com maior experiência na classificação visual de imagens de SR e **G2**: intérpretes com menor experiência. O comportamento espectro-temporal (CET) e a presença de vegetação fotossinteticamente ativa (VFA) das três culturas estudadas foram analisados com uso das imagens multitemporais OLI. Os 200 pixels de A2 também foram usados para treinamento de modelos de classificação automática por análise de imagem orientada a objetos OBIA/*Random Forest* e segmentação via Segmentador Multi-resolução (SM-R). Foram gerados 1150 modelos de classificação, formados pela variação dos parâmetros fator de escala (Fe), forma (Fm) e compacidade (Cp) do algoritmo MR-S aplicado sobre

quatro imagem multitemporais OLI 222/076, datadas de agosto/2014, outubro/2014, janeiro/2015 e maio/2015. Os resultados dos modelos de classificação foram testados pela amostragem *out-of-bag* (OOB), e aquele que obteve maior valor de EG foi replicado para gerar o mapa temático com três classes (cana-de-açúcar, milho e outros), em A3. Outros 200 pixels foram sorteados aleatoriamente sobre A3 e validados por classificação visual, utilizando o protocolo de classificação da presente pesquisa. A ApEq mostrou-se útil no processo de classificação visual, uma vez que o efeito de sobreposição entre órbitas adjacentes foi minimizado pela equalização dos histogramas das imagens, além de proporcionar uma padronização nas cores da VFA das culturas da cana-de-açúcar, milho e soja, nas imagens OLI, em composição colorida R-b5 G-b6 B-b4 de diferentes orbitas/pontos do estado. O protocolo de classificação visual mostrou que a EG média da classificação visual de G2 (94,50 %) foi equivalente ao G1. O modelo de classificação OBIA/*Random Forest* obteve resultados de exatidão global OOB igual a 99 % quando os parâmetros fe (25), fm (30) e cp (10) foram aplicados no segmentador MS-R. Quando o modelo de classificação foi replicado para A3, o resultado de EG foi 86,50 %. No mapa temático produzido em A3 foram obtidos resultados de área ocupada com milho, semelhantes aos dados da Produção Agrícola Municipal (PAM) (mapa temático produzido em A3 subestimou em 4 % as áreas de milho). Em relação à cana-de-açúcar, a subestimativa foi de 11%, em comparação com os dados da PAM-IBGE. De acordo com estes resultados, a classificação visual pode ser utilizada para estimativa agrícola de cana-de-açúcar, milho e soja em São Paulo, bimensalmente. O protocolo de classificação visual em laboratório é viável, o que reduz custos com visitas em campo. Já a classificação automática ainda não obteve resultados de classificação equiparáveis à classificação visual.

Palavras-chave: Estimativa agrícola. Classificação visual. Classificação automática.

OBJECT-BASED IMAGE ANALYSIS AND STATISTICAL SAMPLING APPLIED TO MONITORING SUGARCANE, SOYBEAN AND MAIZE OVER SÃO PAULO STATE

ABSTRACT

The São Paulo State's surface is covered with various land use and land covers (LULC), which are distributed in different geographical portions. Among these LULC, sugarcane, corn and soybeans occupy the main agricultural areas of the state. Even those crops being considered of great importance to the agricultural scene, they are not systematically monitored in near-time real. As the agricultural estimate data has been of great importance for decision-making, the present study was aimed at evaluating the possibility of obtaining the value of occupied area with sugarcane, corn and soybeans, in near-real time (bi-monthly period) in the state São Paulo, using remote sensing (RS), visual and automatic classification of multitemporal OLI images. Three different tests areas were evaluated: the São Paulo State (**A1**), the Cândido Mota/SP municipality (**A2**) and 15 continuous municipalities located in the extreme South of São Paulo (**A3**). The specific objectives of each test area were: (A1) to inform the value of the occupied area with sugarcane, corn and soybeans for São Paulo State, in two months periods of the crop year 2014/2015; (A2) to validate field information, to develop a visual classification protocol and to evaluate it; and (A3) to evaluate the performance of automatic classification. For A1, 5,000 random pixels were stratified into each mesoregion of the state, in each period **P1** (Nov / Dec), **P2** (Jan / Feb), **P3** (Mar / Apr) and **P4** (May / June) of the crop year and classified in laboratory using visual classification protocol. In addition, a database composed of 432 multitemporal images OLI was assembled and divided into 24 tiles (n), saved in R-b5 G-B6 B-B4 color composition and each tile (18 in total) has been standardized via algorithm called Appearance Equalized (ApEq). The ApEq was performed so that no major changes occurred in the resulting color of pseudo-invariant targets present in different images of the tile. Visual analysis of ApEq results was discussed on the mosaic nº 13. In A2, 200 pixels OLI were randomized on the Cândido Mota/SP city and validated in the field during the period 13th-17th April/2015. Information about occupied areas with sugarcane and corn, and also the area that was occupied with soybeans were collected. These data were used to assess the visual classification protocol by statistics generated from the confusion matrix (Overall Accuracy). The visual classification protocol was applied for eight performers divided into two groups **G1**: interpreters with experience in visual classification RS images and **G2**: interpreters with less experience. The temporal spectrum behavior (TEB) and the presence of photosynthetically active vegetation (PAV) of the three studied crops were analyzed using multitemporal OLI images. The 200 pixels from A2 were also used for training automatic classification models for Object Based Image Analysis OBIA/Random Forest and segmentation by

Multi-Resolution segmentation (MR-S). Were generated 1150 classification models formed by varying the MR-S algorithm parameters scale factor (Sf), shape (Sp) and compactness (Cp) applied to four image multitemporal OLI 222/076, dated at August/2014 October/2014 January/2015 and May/2015. The results of classification models have been tested by out-of-bag (OOB) sampling and the best model (highest OC value) was replicated to generate thematic map with three classes (sugarcane, maize and others), in A3 . Another 200 pixels were randomly selected on A3 and validated by visual classification using the classification protocol of this study. The ApEq proved useful in visual classification process, since the effect of overlapping between adjacent images was minimized by the equalizer, and provide standardization in the colors of sugarcane, corn and soybeans PAVs, in R-b5 G-b6 B-b4 colorful composite from different path/rows. The visual classification protocol showed that the EG average from G2 (94.50%) was equivalent to G1. The OBIA/Random Forest classification model achieved an OOB OA equal 99% when Sf (25), Sp (30) and Cp (10) were invested in segmenting MS-R. When the classification model has been replicated to A3, the result of EG was 86.50%. The thematic map produced in A3 were obtained corn occupied area values similar to the area planted with maize from Municipal Agricultural Production (MAP) results (underestimated by 4%). While for sugarcane, the underestimation was 11% compared to the data from MAP-IBGE. According to these results, the visual classification can be used for crop estimation of sugarcane, corn and soybeans in Sao Paulo, bimonthly. The visual classification lab protocol is feasible, which reduces costs in field visits. Automatic classification has not yet obtained classification results comparable to the visual classification.

Keywords: Crop estimate. Visual classification. Automatic classification.

LISTA DE FIGURAS

	<u>Pág.</u>
Figura 2.1 – Proporção da área plantada (31 culturas temporárias e semi-perenes) e área disponível para colheita (35 culturas perenes) das culturas informadas pela Produção Agrícola Municipal (PAM), para SP.	8
Figura 2.2 - Tamanho das propriedades de uso temporário informado pelo Censo de 2006 para os estados do Mato Grosso, Paraná, Rio Grande do Sul e São Paulo.....	13
Figura 2.3 - Variação da exatidão do mapeamento de uma classe qualquer em função o uso de diferentes tamanhos e graus de pureza dos pixels.	15
Figura 2.4 - Exemplos de composições coloridas das imagens de sensoriamento remoto.....	24
Figura 2.5 – Exemplos de cores resultantes da mistura do vermelho, verde e azul no universo de cores RGB.....	25
Figura 2.6 – Comportamento espectral de uma folha sadia em diferentes intervalos de comprimento de onda. Os intervalos B1 a B7 referem-se aos comprimentos de onda do sensor <i>Operator Land Imager</i> (OLI/Landsat-8).	26
Figura 2.7 – Ilustração dos componentes de um pixel de uma área citrícola, visada ao nadir, destacando as árvores cítricas e as sombras nos dosséis.	28
Figura 2.8 – Geometria de iluminação/sombreamento de um elemento de terreno.....	31
Figura 2.9 – (a) Variação da quantidade de sombra produzida por parte de um talhão de cana-de-açúcar em função do ângulo de declinação solar (δ) e orientação do plantio. (b) Efeito do sombreamento de um dossel agrícola devido à variação do alinhamento dos raios solares.	32
Figura 2.10 – Calendários agrícolas utilizados para o plantio e colheita da cana-de-açúcar e milho no estado de São Paulo.....	37
Figura 2.11 – Calendários agrícolas utilizados para o plantio e colheita da soja, amendoim e mandioca no estado de São Paulo.	40
Figura 2.12 – Comparação entre a forma visual dos talhões agrícolas de uma área do <i>Corn Belt</i> (EUA) e de uma área com soja e cana-de-açúcar no município de Ipuã/SP (Brasil).	43
Figura 2.13 – Comparação do erro de classificação usando área teste (linha azul) e técnica OOB com diferentes segmentações (<i>grid size G</i>). 49	

Figura 2.14 – Ilustração do comportamento das exatidões globais dos modelos <i>self-guided segmentation</i> testados para encontrar a melhor segmentação da área de estudo.	50
Figura 2.15 – Comparação dos valores de área modelo (classificação automática) e observado (censo agrícola americano) por municípios e tipo de cultura (milho, soja e trigo) em 358 municípios do estado de Nebraska-EUA.	52
Figura 2.16 – Histogramas de uma imagem OLI, em composição colorida R-b5 G-b6 e B-b4 oriunda do mês de janeiro, em 8 <i>bits</i>	53
Figura 2.17 – Mosaicos de imagens <i>Landsat-like</i> montados sobre a região que compreende a florestal amazônica brasileira, para o ano de 2014. Cada retângulo em amarelo representa uma órbita/ponto Landsat.	55
Figura 2.18 – Passos utilizados para o pré-processamento e processamento de imagens Landsat-like.	59
Figura 2.19 – Histogramas hipotéticos da frequência (Freq.) dos níveis digitais (ND) de uma imagem de sensoriamento remoto que mostram os limites de máximo (máx.) e mínimo (mín.) dos valores de ND. Em (a) o histograma possui baixa amplitude entre os valores de máx. e mín., e em (b) a amplitude é maior.	60
Figura 2.20 – Uma árvore de decisão hipotética utilizada para classificação de A, B e C. Onde V são os diferentes atributos utilizados para separação das classes.	66
Figura 2.21 – Funcionamento do método Random Forest. Os classificadores são representados por hipotéticas árvores de decisão.	67
Figura 3.1 – Fluxograma das atividades desenvolvidas na presente pesquisa.	71
Figura 3.2 – Região de estudo, que compreende o estado de São Paulo e suas 15 mesorregiões.	72
Figura 3.3 – Algumas características relacionadas ao território do estado de São Paulo. (a) Classes de declividade conforme Marques (1971), (b) classes climáticas conforme Alvares <i>et al.</i> (2014), ramais logísticos e porcentagem de cobertura florestal conforme Hansen <i>et al.</i> , (2013).	73
Figura 3.4 – Área plantada com (a) cana-de-açúcar, (b) milho, (c) soja, (d) feijão, (e) amendoim e (f) mandioca, no estado de São Paulo conforme dados da pesquisa agrícola municipal para o ano safra 2013/2014, agregados nas quinze mesorregiões do estado de São Paulo.	74
Figura 3.5 – Uso e ocupação do solo classificado em dados Landsat-like para o estado de São Paulo (CANASAT, 2014; CAFESAT, 2014; FAVARO, 2013; LANDAU <i>et al.</i> , 2013).	75

Figura 3.6 – Formação do banco de dados com dados vetoriais em ambiente SIG, a partir do painel amostral e da intensidade amostral de pixels definida por mesorregião do estado e para os quatro períodos da estimativa do ano safra 2014/2015.	78
Figura 3.7 – Figura ilustrativa que mostra a aquisição de imagens do sensor OLI para os quatro períodos da estimativa e regra de classificação visual (seta em laranja) utilizada para interpretação visual dos 5.000 pixels, por período do ano safra $x / (x+1)$	79
Figura 3.8 – Formação do banco de dados com dados matriciais em ambiente SIG, a partir das 18 imagens OLI em composição colorida R-b5 G-b6 B-b4 sobre o estado de São Paulo, remoção dos valores de <i>Nodata</i> , aplicação da ApEq e recorte do mosaico equalizado, para a abrangência da área do estado de São Paulo.	89
Figura 3.9 – Passos descritos para aplicação da ApEq em imagens multitemporais OLI.	91
Figura 3.10 – Localização da área de estudo 2 (Município de Cândido Mota / número 15 da lista de municípios) e da área de estudo 3 (quinze municípios do estado de São Paulo recobertos pela imagem 222/076 do sensor OLI.	97
Figura 3.11 – (a) Cor do solo presente na região de estudo 2 e 3 e (b) erosões causadas em áreas agrícolas com declividade suave ondulada. ...	97
Figura 3.12 – Principais culturas agrícolas/silvícolas fotografadas no trabalho de campo 2, datado no dia 13-17/abril/2015 no município de Cândido Mota.	98
Figura 3.13 – Ilustração do visor de navegação sobre imagens <i>Google Earth</i> , em ambiente SIG <i>Global Mapper</i> para validação de pixels no campo.	104
Figura 3.14 – Principais dados utilizados para montagem do protocolo de classificação visual 1.	105
Figura 3.15 – Quadrados de 2,5 x 2,5 cm que mostram as imagens OLI e demais produtos usados no protocolo de classificação visual 1 em duas diferentes escaladas de detalhamento, em big Picture e tiny Picture.	107
Figura 3.16 – Fluxograma das atividades desenvolvidas no processo de avaliação da classificação visual dos 200 pixels validados no campo e classificados pelos intérpretes do G1 e G2 utilizando o protocolo 1. A classificação visual decorrente do uso do protocolo 2 (destaque em laranja) foi avaliado apenas pelos intérpretes do G2	111
Figura 3.17 – Fluxograma da metodologia aplicada na classificação automática dos pixels nas imagens multitemporais OLI.	112

- Figura 3.18** – Informação da localização dos dados de referência (treinamento) e validação (área teste e pixels aleatórios, foto-interpretados). ... 113
- Figura 4.1** – As barras ilustram a área ocupada com cultura verde de cana-de-açúcar, milho e soja pelo método MoBARS, por período: Período 1 (Nov/Dez), Período 2 (Jan/Fev), Período 3 (Mar/Abr) e Período 4 (Mai/Jun), no ano safra 2014/2015..... 124
- Figura 4.2** – Imagem base OLI localizada na órbita/ponto 222/076 e equalizada visualmente por intérprete treinado. 129
- Figura 4.3** – As 18 imagens OLI que compreendem o mosaico de Nº 13 usado nesta pesquisa, em aparência equalizada. 133
- Figura 4.4** – Área de sobre posição das imagens OLI localizadas no extremo Sul do estado de São Paulo. As linhas em vermelho indicam as linhas de sobreposição entre imagens adjacentes..... 134
- Figura 4.5** – Área de sobre posição das imagens OLI localizadas ao nordeste do estado de São Paulo. As linhas em vermelho indicam as linhas de sobreposição entre imagens adjacentes. 135
- Figura 4.6** – Resultado das cores da imagem 221/076 do dia 24/janeiro/2015 em aparência equalizada. Em destaque as cores resultantes para a cana-de-açúcar (1), soja (2) e floresta nativa (3) validados em campo, no dia 13-17/abril/2015. 136
- Figura 4.7** – Resultado das cores da imagem 220/075 do dia 17/janeiro/2015 em aparência equalizada. Em destaque as cores resultantes para a soja (1), cana-de-açúcar (2) e milho (3) validados em campo, no dia 13-17/abril/2015. 137
- Figura 4.8** – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Talhões de cana-de-açúcar com: (a) 1,5 m; (b) 2,0 m; (c) 2,5 m; (d) 3,0 m; (e) 3,5 m e (f) 4,0 m. 140
- Figura 4.9** – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Talhões de milho com: (a) 0,3 m; (b) 1,2 m; (c) 1,5 m; (d) 2,0 m; (e) 2,5 m e (f) 2,5 m e pendoadado..... 141
- Figura 4.10** – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. (a) Pivô central recoberto com palhada de soja. (b; c; d) Talhão com palhada e brotação da soja. (e; f) Talhão com palhada e milho de 30 cm. 142
- Figura 4.11** – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. (a; b; c) Talhões de mandioca com 2,0; 1,5 e 1,0 m, respectivamente. (d) Talhão de sorgo sacarino, considerado pixel

“extra”. (e) Remanescente de floresta natural. (f) Área de capoeira.
..... 145

Figura 4.12 – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. (a) Mata ciliar. (b) Divisa entre remanescente florestal e talhão de milho. (c) Pastagem. (d) Talhão de cana-de-açúcar com falhas. (e) área de solo exposto; (f) talhão de milho seco em colheita..... 146

Figura 4.13 – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. (a) talhão de cana-de-açúcar recém-plantada. (b; c) talhões de cana-de-açúcar com plantas de 20 e 50 cm de altura, respectivamente. (d; e; f) Talhões de milho com plantas de 10, 20 e 50 cm de altura, respectivamente 147

Figura 4.14 – Recortes de imagens OLI/Landsat-8 na escala 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Os pixels de 30 x 30 m se localizam na parte central do recorte. A variação das cores dos pixels é resultante da quantidade de VFA da cana-de-açúcar presente em cada uma das 14 datas do ano-safra 2014/2015 e também da ausência de VFA correspondente a palhada gerada pela colheita mecanizada. 150

Figura 4.15 – Recortes de imagens OLI/Landsat-8 na escala 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Os pixels de 30 x 30 m se localizam na parte central do recorte. A variação das cores dos pixels é resultante da quantidade de VFA da cana-de-açúcar presente em cada uma das 14 datas do ano-safra 2014/2015 e também da ausência de VFA gerada pela exposição do solo..... 151

Figura 4.16 – Recortes de imagens OLI/Landsat-8 na escala 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Os pixels de 30 x 30 m se localizam na parte central do recorte. A variação das cores dos pixels é resultante da quantidade de VFA de soja e milho presente em cada uma das 14 datas do ano-safra 2014/2015 e da ausência da VFA devido ao milho seco, a palhada e ao solo exposto. 152

Figura 4.17 – Recortes de imagens OLI/Landsat-8 na escala 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Os pixels de 30 x 30 m se localizam na parte central do recorte. A variação das cores dos pixels é resultante da quantidade de VFA de mandioca, sorgo sacarino, floresta, capoeira e pasto em cada uma das 14 datas do ano-safra 2014/2015 e da ausência de VFA devido a palhada e ao solo exposto. 153

Figura 4.18 – Roteiro de classificação para os pixels representantes das classes de cana-de-açúcar, milho, soja e outros (transição e AVFA) na composição colorida R-b5 R-b6 B-b4.....	156
Figura 4.19 – Interface de classificação desenvolvida em Laboratório SIG. .	158
Figura 4.20 – Interface de classificação desenvolvida em Laboratório Web (... continua).....	159
Figura 4.21 – Interface de classificação desenvolvida em Laboratório Web.	160
Figura 4.22 – Lista dos pixels classificados pelos oito intérpretes (Grupo 1 – 1, 2, 3 e 4; Grupo 2 – 1, 2, 3 e 4) que aplicaram o Protocolo de classificação visual 1 para interpretação de pixels das classes outros (verde), cana-de-açúcar (vermelho) e milho (amarelo). A coluna Rf. na Figura, refere-se aos dados coletados em campo.	166
Figura 4.23 – Recortes de imagens multitemporais OLI em ApEq com duas escalas de observação, 1:1.000 (<i>tyni picture</i>) e 1:30.000 (<i>big picture</i>).	167
Figura 4.24 – Recortes de imagens multitemporais OLI em ApEq com duas escalas de observação, 1:1.000 (<i>tyni picture</i>) e 1:30.000 (<i>big picture</i>).	168
Figura 4.25 – Recortes de imagens multitemporais OLI em ApEq com duas escalas de observação, 1:1.000 (<i>tyni picture</i>) e 1:30.000 (<i>big Picture</i>).	168
Figura 4.26 – Distribuição da exatidão global OOB dos 1150 modelos de classificação RF baseados em diferentes segmentações.	174
Figura 4.27 – Exatidão Global (OOB) para todas as combinações dos parâmetros fator de escala, forma e compacidade, testados com os 200 pixels validados em campo.	176
Figura 4.28 – Resultado da classificação obtida sobre a área de estudo usando a segmentação ótima para o modelo de classificação identificado pela validação OOB (e.g., fator de escala =25, forma = 30 e compacidade = 10).	177
Figura 4.29 – Erros de classificação gerados por OBIA na classificação dos 200 pixels de referência foto interpretados.	179
Figura 4.30 – Erros de classificação gerados por OBIA na classificação dos 200 pixels de referência classificados visualmente.	182
Figura 4.31 – Área teste de 22.500 ha classificada sobre o município de Tarumã-SP. (a) Dados de referência foto-interpretados, (b) mapa temático, (c) divergências entre o mapa de referência e temático e (d) sensibilidade do modelo de classificação.	183

LISTA DE TABELAS

	<u>Pág.</u>
Tabela 2.1 – Variáveis agrícolas definidas e utilizadas para as culturas da cana-de-açúcar, amendoim, mandioca, milho e soja, no estado de São Paulo.....	29
Tabela 2.2 – Resultados literários que mostram os principais atributos utilizados nos modelos de conhecimento para classificação automática de culturas agrícolas no Brasil (1 – 4) e nos EUA (5 – 6).....	47
Tabela 2.3 – Características básicas do sensor OLI.....	54
Tabela 2.4 – Características básicas do sensor MSI	56
Tabela 2.5 – Respectivas fontes de dados para construir as linhas para o cenário agrícola brasileiro. Estes dados sevem de informação secundária no momento da classificação visual do pixel na imagem <i>Landsat-like</i>	57
Tabela 3.1 – Exemplo na organização da lista de pixels utilizada por mesorregião e por período para sorteio da amostra	79
Tabela 3.2 – Área total do território e áreas colhidas (ha) de cana-de-açúcar, milho e soja, e número teórico de pixels de 30 x 30 m, nas mesorregiões e no estado de São Paulo, em 2013 (fonte: IBGE, 2015).	81
Tabela 3.3 – Proporção de área (p_{ih}) e respectivo desvio padrão (s_{ih}) para as culturas de cana, soja e milho, e peso (W_h) de cada estrato no Estado de São Paulo, 2013.....	83
Tabela 3.4 – Proporção de área (p_{ih}) e respectivo desvio padrão (s_{ih}) para as culturas de cana, soja e milho, e peso (W_h) de cada estrato no Estado de São Paulo, 2013.....	84
Tabela 3.5 – Proporção de área (p_{ih}) e respectivo desvio padrão (s_{ih}) para as culturas de cana, soja e milho, e peso (W_h) de cada estrato no Estado de São Paulo, 2013.....	86
Tabela 3.6 – Proporção de área (p_{ih}) e respectivo desvio padrão (s_{ih}) para as culturas de cana, soja e milho, e peso (W_h) de cada estrato no Estado de São Paulo, 2013.....	86
Tabela 3.7 – Datas das imagens OLI utilizadas nos quatros período da estimativa agrícola para cana-de-açúcar, milho e soja no ano safra 2014/2015, em São Paulo. Cada linha da Tabela refere-se a uma visita do sensor OLI sobre o estado de São Paulo.....	88
Tabela 3.8 – Exemplo na organização da lista das métricas obtidas dos valores de reflectância da superfície (σ) das bandas do b4, b5 e b6 para os	

pixels de floresta filtrados pelo passo 1 da ApEq, na imagens: Base (IB), 218/76 e 223/76.....	92
Tabela 3.9 – Exemplo na organização da lista dos valores de Fm das bandas do b4, b5 e b6 para as imagens IB = 218/76 e IC = 223/76.	93
Tabela 3.10 – Exemplo na organização da lista das métricas obtidas dos valores de Fm das bandas do b4, b5 e b6 para as imagens IB = 218/76 e IC= 223/76.....	94
Tabela 3.11 – Imagens OLI e <i>RapidEye</i> usadas no estudo.....	99
Tabela 3.12 – Formação acadêmica dos intérpretes visuais do G2	110
Tabela 3.13 – Sumário descritivo das informações dos dados de referência e validação coletados na área de estudo 2 (município de Cândido Mota) e na área de estudo 3 (quinze municípios do estado de São Paulo recobertos pela imagem OLI 222/076).	113
Tabela 4.1 – Resumo do número de pixels classificados como CN (cana-de-açúcar), ML (milho), SJ (soja), OT (outros) e NV (nuvem) nas 15 mesorregiões do estado de São Paulo (estratos) e nos quatro períodos da estimativa agrícola do ano safra 2014/2015, dado em ordem decrescente do total de pixels sorteados por estrato.	121
Tabela 4.2 – Resumo do número de pixels classificados visualmente como cana-de-açúcar, milho, soja, outros e nuvem no estado de São Paulo, para os quatro períodos da estimativa agrícola do ano safra 2014/2015.	123
Tabela 4.3 – Coeficientes de variação calculados pelo painel amostral MoBARS para as culturas agrícolas da cana-de-açúcar, milho e soja, nos quatro períodos da estimativa agrícola do ano safra 2014/2015	123
Tabela 4.4 – Sumário descritivo das informações dos dados dos valores de reflectância da superfície (σ), fator multiplicativo (FM) e valores de máximo (máx.) e mínimo (mín.) das bandas b4, b5 e b6 do sensor OLI, para as imagens correspondentes ao mosaico N° 13.	131
Tabela 4.5 – Matrizes de erro encontradas para quatro intérpretes do Grupo 1 que aplicaram o protocolo de classificação 1 para classificação visual de 200 pixels, nas classes cana-de-açúcar, milho, soja e outros, em imagens multitemporais OLI.....	164
Tabela 4.6 – Matrizes de erro encontradas para quatro intérpretes do Grupo 2 que aplicaram o protocolo de classificação 1 para classificação visual de 200 pixels, nas classes cana-de-açúcar, milho, soja e outros, em imagens multitemporais OLI.....	165
Tabela 4.7 – Matrizes de erro encontradas para quatro intérpretes do Grupo 2 que aplicaram o protocolo de classificação 2 para classificação visual de 200 pixels, nas classes cana-de-açúcar, milho, soja e outros, em imagens multitemporais OLI.....	170

Tabela 4.8 – Testes para diferença estatística entre os resultados dos Kappas encontrados pelos intérpretes dos dois grupos com melhores resultados de exatidão global.....	172
Tabela 4.9 – Resultados mínimos, medianos e máximos de exatidão global (EG) e coeficiente Kappa para 1150 segmentações e classificações (validação de 200 pixels pela técnica <i>out-of-bag</i>)	174
Tabela 4.10 – Estatística descritiva da exatidão global (EG) e coeficiente Kappa para 1150 segmentações e classificações (validação de 200 pixels pela técnica <i>out-of-bag</i>).....	174
Tabela 4.11 – Matriz de confusão formada com os dados de referência e os dados classificados automaticamente pelo melhor modelo de classificação RF selecionado.	178
Tabela 4.12 – Área ocupada com a cultura da cana-de-açúcar e milho obtida por meio da classificação automática pelo melhor modelo RF em comparação à fornecida pelo IBGE, para a área de estudo.....	180

LISTA DE SIGLAS E ABREVIATURAS

Pág.

GEOGLAM	<i>Global Agricultural Monitoring</i>
G20	20 Países mais ricos do mundo
IBGE	Instituto Brasileiro de Geografia e Estatística
INPE	Instituto Nacional de Pesquisas Espaciais
PREVS	Pesquisa Objetiva de Previsão de Safras
Geosafras	Geotecnologias aplicadas ao Monitoramento Agrícola
BOKU	<i>University of Natural Resources and Life Sciences</i>
AT	Áustria
SR	Sensoriamento remoto
SIG	Sistema de informação geográfica
MoBARS	<i>Monitoring Brazilian Agriculture by Remote Sensing</i>
OBIA	<i>Object-based image analyses</i>
PAM	Produção agrícola municipal
PIB	Produto interno bruto
SP	São Paulo
FAO	<i>Food and agriculture organization</i>
CEPEA	Centro de estudos avançados em economia aplicada
FAOSTAT	<i>Statistics</i> FAO
ONUA	Organização das nações unidas
MODIS	<i>Moderate-Resolution Imaging Spectroradiometer</i>
PR	Paraná
RS	Rio Grande do Sul
MT	Mato Grosso
CANASAT	Monitoramento da Cana-de-açúcar via Imagens de Satélite
UNICA	União das Indústrias de Cana-de-açúcar
CV	Coeficiente de variação
NASS	<i>National Agricultural Statistics Service</i>

VIS	Visível
IVP	Infravermelho próximo
IVC	Infravermelho de ondas curtas
REM	Radiação Eletromagnética
RGB	Ambiente tridimensional de cores RGB (<i>Red Green Blue</i>)
VFA	Vegetação fotossinteticamente ativa
OLI	<i>Operator Land Imager</i>
IVs	Índices de vegetação
IAF	Índice de área foliar
USGS	<i>United State Geological Survey</i>
PRODES	Monitoramento do Desmatamento na Amazônia Legal por Satélite
TM	<i>Thematic Mapper</i>
ETM ⁺	<i>Enhanced Thematic Mapper Plus</i>
CET	Comportamento espectro-temporal
MAPA	Ministério Agricultura, Pesca e Abastecimento
IEA	Instituto de Economia Aplicada
GM	Grupo de maturação fisiológica
SBM	Sociedade Brasileira de Mandioca
EUA	Estados Unidos da América
SM-R	Segmentador multi-resolução
NDVI	Índice de vegetação da diferença normalizada
IVV	Índice de vegetação do verde
GNDVI	Índice de Vegetação da diferença normalizada do verde
NDSVI	Índice de vegetação da diferença normalizada da senescência
MLME	Modelo linear de mistura espectral
OOB	<i>Out-of-bag</i>
RF	<i>Random Forest</i>
IK	Índice Kappa
EG	Exatidão global
EU	Exatidão do usuário

EP	Exatidão do Produtor
Var.K	Variância do Kappa
BE	Bélgica
NASA	<i>National Aeronautics and Space Administration</i>
BQA	<i>Quality Assessment Band</i>
MSI	<i>Multispectral Imager</i>
MMA	Ministério do Meio Ambiente
GLCM	<i>Gray-Level Co-Occurrence Matrix</i>
GFC	<i>Global Forest Change</i>
CDR	<i>Climate Data Records</i>
ND	Nível Digital
LL	<i>Landsat-like</i>
API's	Alvos peseudo-invariantes
CRB	Crescimento de região de Bins
Sm	Similaridade
Ar	Área
Fe	Fator de escala
Fm	Forma
Cp	Compacidade
WEKA	<i>Waikato Environment Knowledge Analyisi</i>
AD	Árvore de decisão
WRS-2	<i>Worldwide Reference System 2</i>
CPU	<i>Central Processing Unit</i>
GB	<i>Gigabyte</i>
RAM	<i>Random Access Memory</i>
TB	<i>Terabyte</i>
HD	<i>Hard disk</i>
UTM	Universal Transversa de Mercator
ID	Identidade
CN	Cana-de-açúcar
SJ	Soja

ML	Milho
OT	Outros
NV	Nuvem
AVFA	Ausência de vegetação fotossinteticamente ativa
TERRACLASS	Mapeamento do Uso e Cobertura Vegetal do Cerrado

SUMÁRIO

	<u>Pág.</u>
1. INTRODUÇÃO	1
1.1. Contextualização	1
1.2. Justificativa	3
1.3. Motivação	3
1.4. Hipóteses	3
1.5. Objetivos	4
1.5.1. Geral	4
1.5.2. Específicos	4
2. FUNDAMENTAÇÃO TEÓRICA	7
2.1. A agricultura no estado de São Paulo	7
2.2. Agrodinâmica do estado de São Paulo	9
2.3. O uso do sensoriamento remoto nas estimativas agrícolas	10
2.3.1. Confecção de mapas com cultivos agrícolas	11
2.3.2. Resolução espacial dos dados MODIS e o tamanho das áreas agrícolas em São Paulo	12
2.3.2.1. Resolução espacial dos dados <i>Landsat-like</i> e o problema das nuvens	15
2.3.3. Amostragem	18
2.3.4. Estimativas agrícolas via painel amostral	19
2.3.4.1. O programa PREVS	20
2.3.4.2. O programa Geosafras	20
2.3.4.3. O programa MoBARS	22
2.4. Classificação de imagens de sensoriamento remoto	23
2.4.1. VFA em escala foliar	25
2.4.2. VFA em escala de dossel	27
2.4.2.1. VFA decorrente do tipo de dossel agrícola	29
2.4.2.2. VFA em decorrência do espaçamento das plantas	30
2.4.2.3. VFA em decorrência da geometria de iluminação do sensor	30
2.4.3. Classificação visual de imagens de sensoriamento remoto	33
2.4.3.1. Feições interpretáveis nas imagens	33
2.4.3.2. Imagens multitemporais	34

2.4.3.3.	Calendário agrícola	35
2.4.3.4.	O calendário agrícola da cana-de-açúcar e milho	36
2.4.3.5.	O calendário agrícola da soja, amendoim e mandioca.....	39
2.4.4.	Classificação digital de imagens de sensoriamento remoto	41
2.4.4.1.	<i>Object-based Image Analysis</i> (OBIA) na agricultura	44
2.4.5.	<i>Sensor Operator Land Imager</i> (OLI).....	52
2.4.5.1.	Resolução radiométrica e espectral do OLI	53
2.4.5.2.	Resolução espacial e temporal do OLI.....	54
2.4.6.	<i>Sensor Multispectral Imager</i> (MSI)	55
2.4.7.	Linhas de base ou <i>baselines</i>	56
2.4.8.	Pré-processamento de dados OLI.....	57
2.4.8.1.	Equalização ou realce	59
2.4.8.2.	Normalização ou uniformização	61
2.4.8.3.	Aparência Equalizada (ApEq)	62
2.4.9.	Processamento de dados OLI – classificação automática	62
2.4.9.1.	Segmentadores de imagens de sensoriamento remoto	62
2.4.9.2.	Algoritmos de classificação automática baseados em técnica de árvore de decisão	64
3.	MATERIAL E MÉTODOS	69
3.1.	Área de estudo 1 – Estado de São Paulo	71
3.1.1.	Material	75
3.1.2.	Métodos	77
3.1.2.1.	Formação do banco de dados: vetores	77
3.1.2.2.	Painel amostral: definição da intensidade amostral	80
3.1.2.3.	Painel amostral: extrapolação	85
3.1.2.4.	Formação do banco de dados matricial	87
3.1.2.4.1.	Aparência Equalizada (ApEq)	89
3.1.2.4.2.	Construção dos mosaicos de imagens em ApEq	95
3.2.	Área de estudo 2 e 3 –Cândido Mota e municípios vizinhos.....	96
3.2.1.	Material	98
3.2.2.	Métodos	100
3.2.2.1.	Trabalho de campo 1: microrregião de Assis	101
3.2.2.2.	Trabalho de campo 2: município de Cândido Mota	102
3.2.2.3.	Montagem do protocolo de classificação visual de pixels	104

3.2.2.3.1.	Interface de classificação 1: Laboratório SIG	105
3.2.2.3.2.	Interface de classificação 2: Laboratório <i>Web</i>	107
3.2.2.4.	Avaliação da classificação visual	109
3.2.2.5.	Classificação automática.....	111
3.2.2.5.1.	Confecção de dados de validação	112
3.2.2.5.2.	Segmentação das imagens multitemporais e extração de atributos	114
3.2.2.5.3.	Formação do modelo de conhecimento e classificação	115
3.2.2.5.4.	Desempenho do modelo de classificação automática.....	117
4.	RESULTADOS E DISCUSSÃO	119
4.1.	Estimativa agrícola de cana-de-açúcar, milho e soja em São Paulo, no ano safra 2014/2015	119
4.1.1.	Coeficientes de variação para cana-de-açúcar, milho e soja	120
4.1.2.	Resultados de área ocupada com cana-de-açúcar, milho e soja no estado de São Paulo, para o ano safra 2014/2015	124
4.2.	Aparência Equalizada em imagens OLI	127
4.2.1.	Imagem Base usada no processo da ApEq	128
4.2.2.	Variação dos valores de reflectância da superfície dos pixels pseudo-invariantes	130
4.2.3.	Mosaico composto por dezoito imagens em ApEq.....	132
4.3.	Definição do protocolo de classificação visual de imagens OLI ..	138
4.3.1.	Comportamento espectro-visual.....	139
4.3.2.	Correlação espectro-temporal visual	148
4.3.3.	Roteiro de classificação visual	155
4.3.4.	Protocolo de classificação visual.....	156
4.3.5.	Avaliação do uso dos protocolos de classificação visual	162
4.4.	Classificação automática.....	173
4.4.1.	Exatidão dos modelos de classificação	173
4.4.2.	Avaliação da robustez do modelo de classificação	176
4.4.3.	Avaliação do mapa temático	180
5.	CONCLUSÕES	184
	REFERÊNCIAS BIBLIOGRÁFICAS	187
	ANEXO A.....	207
	ANEXO B.....	Erro! Indicador não definido.
	ANEXO C.....	208

ANEXO D.....	210
GLOSSÁRIO.....	Erro! Indicador não definido.

1. INTRODUÇÃO

1.1. Contextualização

O Brasil, considerado um dos principais produtores de *commodities* agrícolas do mundo (FAOSTAT, 2015), ainda está convertendo áreas de florestas tropicais e de cerrado em áreas utilizadas na atividade agropecuária (HANSEN *et al.*, 2013; LAPOLA *et al.*, 2013).

O grupo composto pelos países mais importantes economicamente no mundo (G20) fundou o *Global Agricultural Monitoring* (GEOGLAM). O GEOGLAM tem por objetivo principal o levantamento objetivo das variáveis “produtividade” e “área” das principais culturas agrícolas que constituem a dieta da população humana. Na procura por ferramentas úteis para atingir tal objetivo, o GEOGLAM definiu que os dados de sensoriamento remoto deveriam ser utilizados nos métodos de levantamento de safras (ATZBERGER, 2013; BECKER-RESHEF *et al.*, 2010).

Os reflexos positivos das iniciativas adotadas pelo GEOGLAM irão assegurar menores índices de passivos ambientais decorrentes do processo de desmatamento, o controle da volatilização dos preços de bens de consumo e a otimização da distribuição dos alimentos (BECKER-RESHEF *et al.*, 2010; ATZBERGER, 2012). No entanto, para que isto ocorra, caberá, ao Brasil e a cada país membro do G20 realizar todo processo de previsão de safras e informar o seus dados ao GEOGLAM em tempo hábil, para agilizar as principais tomadas de decisão (ATZBERGER, 2012; EBERHARDT, 2014).

No Brasil, os órgãos oficiais responsáveis pelo fornecimento das estimativas agrícolas são o Instituto Brasileiro de Geografia e Estatística (IBGE) e a Companhia Nacional de Abastecimento (CONAB). Os dados fornecidos por estes órgãos apresentam um método bem fundamentado de levantamento, no entanto são dados considerados “subjetivos”, devido à fonte deste tipo de estimativas agrícolas provir de informações relatadas pelos agentes municipais

ligados ao setor agrícola local e não possibilitar a geração dos erros relacionados às variáveis estimadas (PINO, 1999).

Portanto, para que as estatísticas agrícolas possam oferecer dados objetivos, rápidos e confiáveis é demandada a incorporação de inovações tecnológicas aos procedimentos operacionais já em uso.

A partir de 1972, devido ao fornecimento das primeiras imagens do satélite Landsat para o Brasil, coube aos profissionais da área de sensoriamento remoto agrícola do Instituto Nacional de Pesquisas Espaciais (INPE) iniciarem a avaliação desta nova tecnologia, como fomento para as estimativas agrícolas, e assim, algumas iniciativas foram tomadas pelo INPE em parceria com outras instituições (IBGE e CONAB) visando o uso de sensoriamento remoto para o levantamento objetivo de estatísticas agrícolas.

Uma importante iniciativa, desenvolvida em 1986, foi a Pesquisa Objetiva de Previsão de Safras (PREVS) (FIGUEIREDO, 2005; RIZZI, 2005), tendo sido uma segunda, em 2002, o programa denominado Geotecnologias aplicadas ao Monitoramento Agrícola (GeoSafras). Nesta esteira, a terceira iniciativa, datada em 2013, foi o projeto *Monitoring Brazil Agriculture by Remote Sensing* (MoBARS) (EBERHARDT, 2014).

A PREVS e o GeoSafras tiveram grande importância na formação do MoBARS, e a grande contribuição destas duas iniciativas foi a incorporação dos dados remotos aos métodos de levantamento de estimativa agrícola. No entanto, a indispensabilidade de realizar trabalhos de campo para tais metodologias de levantamentos objetivos inviabilizou técnica e economicamente os projetos PREVS e GeoSafras (ADAMI, 2003; CRUSCO, 2007).

Após cerca de 30 anos de experiências com os resultados obtidos pelos projetos PREVS e GeoSafras, atualmente, o projeto MoBARS visa incentivar a sinergia entre as mais modernas geotecnologias para a geração de estatísticas agrícolas confiáveis, em tempo hábil e com custos mínimos.

O Método MoBARS em desenvolvimento pelo INPE em parceria com a *University of Natural Resources and Life Sciences (BOKU)* de Viena-AT tem como finalidade monitorar em Sistema de Informações Geográficas (SIG) a “área ocupada” com três das principais culturas agrícolas brasileiras (cana-de-açúcar, milho e soja), tendo como área de estudo os estados de São Paulo, Paraná e Rio Grande do Sul, em períodos bimensais do ano safra o que torna em tempo “quase-real” a divulgação dos produtos, utilizando para esta finalidade geotecnologias modernas e amostragens estatísticas.

1.2. Justificativa

O Brasil ainda não possui um sistema rotineiro que informe com precisão estatística a área que as três principais culturas agrícolas plantadas (cana-de-açúcar, milho e soja) ocupam, em tempo quase real. Algumas iniciativas estão sendo tomadas, como o programa MoBARS.

No entanto, dentre os principais avanços buscados pelo programa MoBARS para se estimar a área ocupada com culturas agrícolas via imagens de sensoriamento remoto (SR), estão a questão da minimização da dependência de dados de campo e a geração de estimativas em tempo quase-real.

Para isso, um protocolo de classificação visual ou automático de imagens orbitais deve ser definido e avaliado. E para obter tais resultados, o conhecimento técnico-científico, sobre classificação de imagens orbitais, baseado em trabalhos literários e validações de campo, deve ser formado.

1.3. Motivação

A necessidade de informar dados agrícolas em tempo quase real e via imagens orbitais faz com que esta pesquisa seja útil como de levantamento de safras, em nível nacional.

1.4. Hipóteses

- a) É possível, em tempo quase-real, com o uso de amostragem estatística de dados multitemporais de sensoriamento remoto e de

classificação visual, estimar as áreas ocupadas com as culturas da cana-de-açúcar, da soja e do milho em pleno vigor vegetativo no estado de São Paulo;

- b) A classificação automática de imagens multiespectrais e multitemporais por meio da análise de imagem orientada a objetos (OBIA), aplicada para classificação de pixels amostrais de cana-de-açúcar, milho e soja em pleno vigor vegetativo no estado de São Paulo, tem desempenho comparável ao da classificação visual.

1.5. Objetivos

1.5.1. Geral

Desenvolver um método que permita estimar por amostragem estatística as áreas ocupadas com cana-de-açúcar, soja e milho em pleno vigor vegetativo, no estado de São Paulo, em tempo quase-real, em quatro períodos bimensais de um ano agrícola, a partir de dados Landsat-8, por classificação visual e ainda estudar a possibilidade de automatizar o processo de obtenção das estimativas.

1.5.2. Específicos

A presente pesquisa teve os seguintes objetivos específicos:

- 1.** Definir a melhor forma de estratificação e de alocação de pontos amostrais no estado de São Paulo, para reduzir os coeficientes de variação das estimativas de áreas plantadas com cana-de-açúcar, milho e soja.
- 2.** Estabelecer o procedimento de obtenção de um mosaico bimensal de imagens sobre o estado de São Paulo, que permita a interpretação visual dos pontos amostrais.
- 3.** Desenvolver um protocolo de classificação, que englobe o ano safra 2014/2015, para a interpretação dos pixels amostrais nos mosaicos estaduais,

aplicá-lo em quatro períodos bimensais (dezembro/janeiro, fevereiro/março, abril/maio, junho/julho) e gerar as estimativas correspondentes.

4. Avaliar o desempenho do protocolo de classificação desenvolvido para as imagens de sensoriamento remoto tendo como referência dados validados em campo.

5. Definir o método de segmentação mais adequado à situação agrícola da área de estudo utilizada para validação.

6. Definir o método de classificação automática que tenha melhor desempenho para classificação dos pixels validados em campo.

7. Avaliar os resultados da classificação automática com o uso dos dados validados em campo.

8. Comparar os resultados da classificação visual com os dados da classificação automática.

9. Comparar os dados de área ocupada com cana-de-açúcar e milho em 15 municípios localizados no extremo Sul do estado de São Paulo, estimados pela metodologia automática proposta, com os dados oficiais fornecidos pela Produção Agrícola Municipal (PAM) do IBGE.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. A agricultura no estado de São Paulo

O estado de São Paulo mesmo sendo considerado um estado com alto grau de industrialização arrecadou em 2013, R\$ 213 bilhões no setor do agronegócio. O PIB do agronegócio do estado de São Paulo equivaleu a 15 % do PIB do estado e 4,4 % do PIB do País (CEPEA, 2013).

Estes resultados foram alcançados principalmente devido à vastidão do seu território, ao clima, solo e relevo que favorecem a prática da atividade agrícola, à regionalização e à tradição de cultivo praticada nos municípios do estado, às políticas agrícolas favoráveis ao desenvolvimento de infraestruturas; e a disponibilidade de créditos agrícolas mais baratos aos proprietários rurais, para financiamentos agrícolas (SPAROVEK *et al.*, 2010; LAPOVA *et al.*, 2014).

O estado de São Paulo possui 24,8 milhões de hectares ocupados com as mais diversas culturas agrícolas, sejam elas perenes, semi-perenes ou temporárias. A Figura 2.1 ilustra as principais culturas cultivadas no estado de São Paulo, de acordo com as suas proporções em área plantada para o estado e calculadas pela PAM para os anos de 1990-2013.

Entre as atividades agrícolas desenvolvidas em SP, o grande destaque vai para a produção de cana-de-açúcar, que utiliza perto de 22% da área do estado, depois vêm culturas temporárias como o milho (3%) e a soja (2%), perenes como laranja (2%) e café (1%); em seguida, demais culturas temporárias como o feijão e amendoim com 0,4% cada, e a mandioca, que ocupa 0,2% da área do estado (Figura 2.1).

Figura 2.1 – Proporção da área plantada (31 culturas temporárias e semi-perenes) e área disponível para colheita (35 culturas perenes) das culturas informadas pela Produção Agrícola Municipal (PAM), para SP.



Fonte: adaptado do IBGE, (2015).

A cana-de-açúcar lidera a geração de renda em atividades agrícolas no estado, com 58%; entre os mais importantes, vêm a seguir a laranja, com 6% e, depois, a soja e o milho, com 5% cada. As demais culturas possuem importância secundária em geração de renda no mercado agrícola (CEPEA, 2013).

A *Food Agriculture Organization* (FAO) faz todo ano um balanço sobre a quantidade (toneladas) de produtos agrícolas processados e consumidos pela população humana. De acordo com a FAO o óleo de soja e de milho (71 milhões de ton.) e o bagaço e o melaço da cana-de-açúcar (56 milhões de ton.) são os produtos agrícolas mais exportados pelo Brasil (FAOSTAT, 2015). Estes estão entre os mais usados na dieta humana (FOOD BALANCE SHEET, 2015) e estão entre os produtos que mais geram renda ao setor agrícola do estado de São Paulo (CEPEA, 2013).

2.2. Agrodinâmica do estado de São Paulo

Dadas as informações sobre a área ocupada com as principais culturas agrícolas em São Paulo e a importância da cana-de-açúcar, soja e milho no cenário mundial, é importante destacar também que a ocupação do território do estado foi se moldando gradativamente no que é hoje, isto é, ocorreu “dinâmica” da agricultura ao longo dos seus anos (GARAGORRY e CHAIB FILHO, 2008).

O termo “agrodinâmica” foi designado por Garagorry e Chaib Filho (2008) para medir o deslocamento das culturas agrícolas em determinado limite geográfico, por meio de técnicas matemáticas e físicas. Como exemplo disso, a cana-de-açúcar, avaliada nos 23 anos de dados da PAM, foi a cultura que mais sofreu dinâmica sobre o estado de São Paulo (SCHULTZ *et al.*, 2014).

Em 2003, o plantio de cana-de-açúcar no estado foi impulsionado pelas tratativas adotadas pela Organização das Nações Unidas (ONU) para redução da emissão dos gases de efeito estufa causada pelo consumo de automotivo de combustíveis não renováveis (LUCON e GOLDEMBERG, 2010). Assim, uma parte da queima de combustíveis fósseis utilizada diariamente pelos carros movidos à gasolina e óleo diesel, seria substituída pelo uso do etanol de “cana” nos carros *flex* (WALTER *et al.*, 2011).

O resultado da aderência do Brasil a esta tratativa foi que em 23 anos de agrodinâmica (1990-2013), 15% da área de SP encontravam-se já ocupadas com cana-de-açúcar (RUDORFF *et al.*, 2010). Vale ressaltar ainda que, neste período, 25 % dos canaviais expandiram sobre áreas que praticavam a agricultura de verão (plantio de milho e soja, por exemplo). No entanto, a maior proporção da expansão (69,7 %) da cana-de-açúcar foi sobre as áreas de pastagens (ADAMI *et al.*, 2012).

A agrodinâmica das culturas agrícolas sobre o estado de São Paulo pode ainda ser influenciada pela Agenda de Desenvolvimento Pós-2015 e também pelo

Novo Código Florestal Brasileiro. É de se esperar, que estas duas diretrizes norteiem novos rumos para a atividade agrícola brasileira.

2.3. O uso do sensoriamento remoto nas estimativas agrícolas

A área total do estado de São Paulo é de 24,8 milhões de hectares, dos quais 16,9 milhões eram ocupados por estabelecimentos agropecuários segundo o Censo Agropecuário de 2006 do IBGE. A atividade agropecuária do estado, apesar da concentração em alguns produtos, é diversificada e dinâmica (IBGE, 2015).

A vastidão do território, a grande quantidade de culturas plantadas e a alternância da escolha da cultura a ser plantada no ano, são fatores que dificultam o monitoramento objetivo da agricultura em extensas áreas (ADAMI e RUDORFF *et al.*, 2012; IPPOLITI-RAMILO *et al.*, 2003; MATTON *et al.*, 2015; WALDNER *et al.*, 2015).

Para tal finalidade, as imagens de sensoriamento remoto são atualmente a única ferramenta que pode ser considerada sem fronteiras, com o mesmo padrão de uso para as diferentes nações do mundo (LUIZ, 2003) e com um considerado histórico de uso como fomento nas estimativas agrícolas, em determinada região geográfica (ATZBERGER, 2013; GALFORD *et al.*, 2008; MOSLEH *et al.*, 2015; RUDORFF *et al.*, 2010).

O pixel, considerado elemento básico de uma imagem de SR, pode ser uma ferramenta útil para interpretação de acontecimentos históricos em determinada área (ARVOR *et al.*, 2010; GALFORD *et al.* 2008; ADAMI *et al.* 2012b). A história “contada” pelo pixel, por meio de periódicos valores “tonais” ou “digitais”, pode ser processada e aplicada para gerar estatísticas agrícolas.

No Brasil, as estatísticas agrícolas objetivas podem ser geradas por meio de mapas de cultivos agrícolas (EBERHARDT *et al.*, 2015; SCHULTZ *et al.*, 2015), ou por meio de estatística probabilística (ADAMI *et al.*, 2010; LUIZ *et al.*, 2012; LUIZ *et al.*, 2015). Estas duas abordagens possuem suas desvantagens, relacionadas aos seus resultados finais produzidos.

A utilidade do pixel, para confecção de mapas depende das resoluções dos sistemas sensores usados para tal processo (MATTON *et al.*, 2015; MOSLEH *et al.*, 2015; WALDNER *et al.*, 2015) e para a amostragem probabilística, sendo que a desvantagem do método está relacionada à inexistência da espacialização da cultura agrícola em nível de talhão agrícola (por exemplo) (EBERHARDT *et al.* 2015).

2.3.1. Confecção de mapas com cultivos agrícolas

Os mapas de cultivos agrícolas são confeccionados em maior proporção por técnicas automáticas de classificação de imagens. Os produtos são produzidos a partir da oferta de imagens de sensoriamento remoto, de sistemas computacionais atuais existentes e de diferentes técnicas para processamento dos dados. Exemplos desta assertiva foram: o mapeamento do tipo da colheita de cana-de-açúcar praticada em São Paulo (AGUIAR *et al.*, 2011) e validação de áreas classificadas com cana-de-açúcar (ADAMI *et al.*, 2012). Aguiar *et al.* (2011) e Adami *et al.* (2012) utilizaram dados multi-sensor para produção e publicação de produtos baseados em imagens de sensoriamento remoto, para o estado de São Paulo.

Os sensores mais comuns utilizados pelos usuários de dados de sensoriamento remoto nas condições agrícolas do Brasil são os dados do sensor *Moderate-Resolution Imaging Spectroradiometer* (MODIS) e os sensores chamados *Landsat-like* (Schultz *et al.* 2015a). A diferença entre estes dois tipos de dados é dada pelas quatro possíveis resoluções dos seus sistemas sensores, que estão atreladas:

- a) à temporalidade da informação coletada (resolução temporal, dias);
- b) à dimensão espacial do pixel formado graficamente (resolução espacial, metros);
- c) ao número e espessura dos canais espectrais do sensor (resolução espectral, μm);

d) e ao nível/valor do sinal digital captado pelo detector (resolução radiométrica, bits).

Para uso na agricultura, ainda busca-se um sensor capaz de fornecer pixels de qualidade, que tenham alta repetitividade e suportem uma resolução espacial fina, isto é, que atenda ao tamanho médio das propriedades agrícolas. No entanto, nem o MODIS, nem os sensores *Landsat-like* possuem ainda tais habilidades para monitoramento de mais de uma cultura ao mesmo tempo (EBERHARDT *et al.* 2015; SCHULTZ *et al.*, 2015a).

A relação entre resolução espacial e temporal segue caminhos inversos no SR, enquanto que os dados diários possuem pixels em média com 250 m de resolução espacial (e.g., MODIS), os dados *Landsat-like* (± 30 m) possuem apenas 16 dias de repetitividade (JUSTICE *et al.*, 2002; POWELL *et al.*, 2007).

Além disso, os dados MODIS são registrados em grandes cenas (São Paulo, Paraná, Santa Catarina e Rio Grande do Sul em uma mesma imagem) e as cenas *Landsat-like* abrangem pequenas áreas (e.g., três cenas para recobrir a microrregião de Assis/SP), fator este preponderante para uso do MODIS na geração de mapas de cultivos agrícolas em grandes áreas para o Brasil (EBERHARDT *et al.*, 2015).

2.3.2. Resolução espacial dos dados MODIS e o tamanho das áreas agrícolas em São Paulo

Autores como Rudorff *et al.* (2007), Yi *et al.* (2007), Lamparelli *et al.* (2008); Galford *et al.* (2008). Arvor *et al.* (2011); Bernardes *et al.* (2011), Gusso *et al.*, (2012), Johann *et al.* (2012) e Brown *et al.* (2013) utilizaram-se de dados MODIS para produção de mapas de cultivo das culturas de ciclo curto, como o milho e a soja, no Paraná (PR), Rio Grande do Sul (RS) e Mato Grosso (MT).

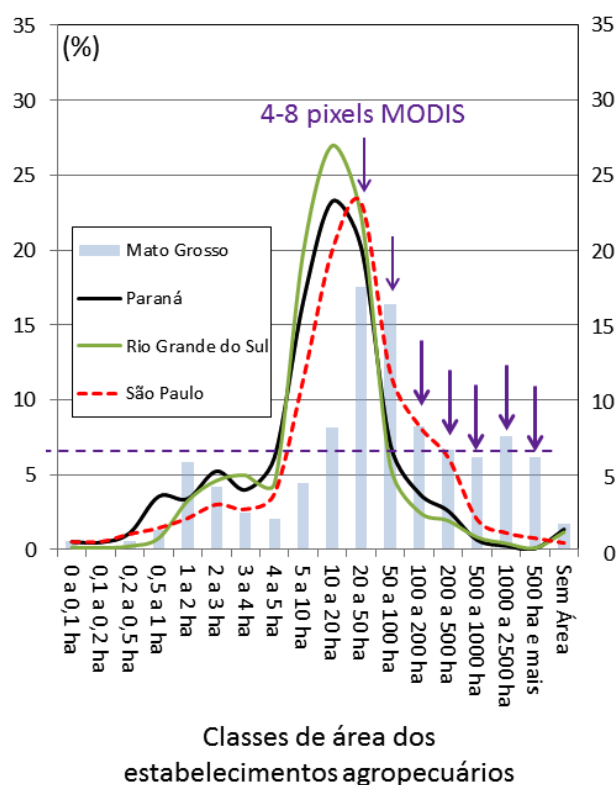
Nas condições agrícolas dos estados do RS e PR, Rudorff *et al.* (2007), Yi *et al.* (2007) e Lamparelli *et al.* (2008) comentaram que o principal entrave do uso de pixels MODIS (resolução espacial de 6,25 ha) utilizados para se estimar a

área de cultivos agrícolas foi omitir áreas de minifúndios, no mapa final produzido.

No entanto, para a agricultura praticada no estado do Mato Grosso, por exemplo, os dados MODIS possuem considerada relevância para mapeamento de cultivos com soja (GALFORD *et al.*, 2008; ARVOR *et al.* 2011; BERNARDES *et al.* 2011; GUSSO *et al.* 2012; BROWN *et al.* 2013).

De fato, no estado do MT, o uso de imagens MODIS para mapeamento de cultivos agrícolas terá sim um desempenho melhor se comparado aos estados do PR, RS e também de SP. O melhor desempenho da confecção do mapa é dado pela quantidade das áreas agrícolas do MT serem muito maiores que as áreas agrícolas dos demais estados citados (Figura 2.2).

Figura 2.2 - Tamanho das propriedades de uso temporário informado pelo Censo de 2006 para os estados do Mato Grosso, Paraná, Rio Grande do Sul e São Paulo.



O eixo y é dado pela porcentagem do número de propriedades por classe de área. Fonte: adaptado do IBGE, (2015).

Para as condições agrícolas de SP, na classe de 20 a 50 ha de área (4 a 8 pixels MODIS), são encontradas mais de 20 % das propriedades agrícolas do estado (Figura 2.2); e as demais classes possuem baixa representatividade.

Em paisagens heterogêneas, como é o caso do estado de São Paulo, as “máscaras” de plantio formadas pelo uso do pixel de 6,25 ha, do MODIS, fica comprometida devido à influência de pixels com mistura espectral, como comentado no trabalho de Löw e Duveiller, (2014).

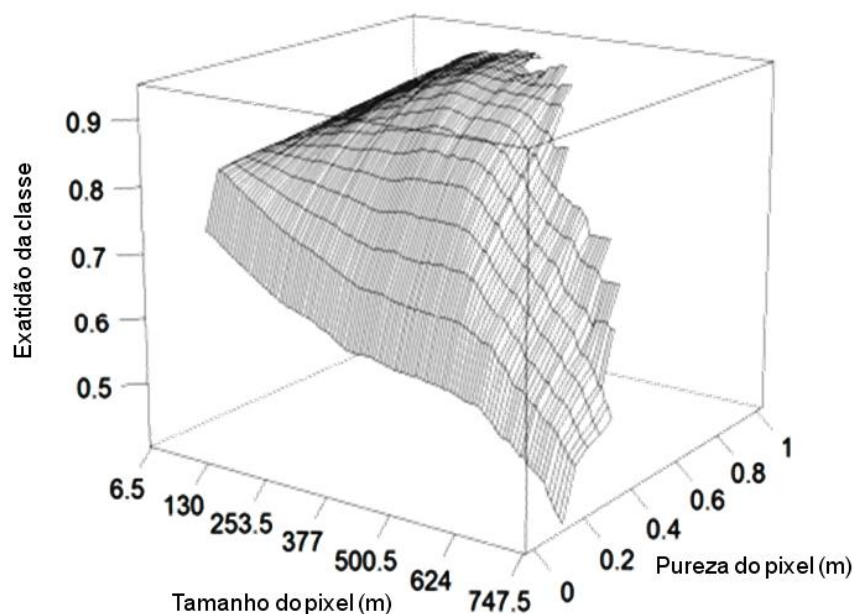
Luiz e Epiphanyo (2001) demonstraram que, devido ao efeito de borda, talhões agrícolas menores que 10 ha apresentam mais de 1/3 dos pixels impuros, enquanto que os talhões de 20 ha, com maior frequência em São Paulo e Mato Grosso (Figure 2.2) possuem mais de 1/4 dos pixels impuros.

Dependendo do grau de mistura (ou pureza) dos pixels que fazem parte de um talhão agrícola, pode ocorrer grandes diferenças na precisão da classificação automática (HSIEH *et al.*, 2001), o que significa que, pixels caracterizadas por diferentes graus de pureza podem não ser tão confiáveis para discriminar as classes sob investigação (OLOFSSON *et al.*, 2013; CONGALTON *et al.*, 2014;).

Para o estado do MT, em todas as classes superiores a 20 a 50 ha de área, é possível encontrar mais de 50 % das propriedades temporárias do estado (Figura 2.2). O número de talhões que suportariam mais de 4 a 8 pixels é muito maior se comparado aos números de propriedades no estado de SP.

Assim, devido à quantidade de pixels puros ser maior que impuros no momento do mapeamento de culturas agrícolas no MT, um maior grau de pureza do pixel melhora a sua capacidade de descrever as propriedades da superfície desejada, como comentado por Löw e Duveiller, (2014). A Figura 2.3 ilustra o comportamento da exatidão de uma cultura agrícola qualquer mapeada em função da variação do tamanho do pixel (*pixel size*) e pureza do pixel (*pixel purity*).

Figura 2.3 - Variação da exatidão do mapeamento de uma classe qualquer em função o uso de diferentes tamanhos e graus de pureza dos pixels.



Fonte: adaptado de Löw e Cuveiller (2014).

Na Figura 2.3 percebe-se que um aumento do tamanho do pixel e diminuição da sua pureza, causa a diminuição gradativa da exatidão do mapeamento (o caso das áreas agrícolas de São Paulo). No entanto, se o pixel sofre um incremento e o grau de pureza não diminui, a exatidão do mapeamento segue um caminho inverso, isto é, aumenta de acordo com o aumento do tamanho do pixel (o caso das áreas agrícolas do Mato Grosso).

2.3.2.1. Resolução espacial dos dados *Landsat-like* e o problema das nuvens

Para controlar o problema decorrente da influência dada pela resolução espacial no mapeamento de culturas temporárias e semi-perenes em São Paulo, a resolução espacial muito mais fina dos dados *Landsat-like*, se torna viável para mapeamento das principais culturas agrícolas (SUGAWARA *et al.*, 2008; MOREIRA *et al.*, 2008; MELLO *et al.*, 2010; RUDORFF. *et al.*, 2010; VIEIRA *et al.*, 2012; SCHULTZ *et al.*, 2015a; SCHULTZ *et al.*, dados não publicados). No entanto, estes dados possuem desvantagens se comparados aos dados MODIS.

Para o caso específico do Brasil, a principal limitação do uso de dados Landsat-like é dada pela interferência de nuvens sobre pixels da imagem. Rizzi e Rudorff (2005), Sano *et al.*, (2007) e Sugawara *et al.* (2008) comentaram que por ser um país tropical e adotar agricultura de verão, o Brasil é mais afetado pela interferência de nuvens sobre o conjunto de dados Landsat-like.

Além disso, Eberhardt *et al.* (dados não publicados) comentaram que a sazonalidade influencia com maior intensidade a presença de nuvem sobre os dados de sensoriamento remoto que recobrem o estado de SP e menos os demais estados localizados em latitudes inferiores a Ele.

De fato, em São Paulo, trabalhos que relacionaram mapeamento de cultivos agrícola de culturas temporárias aos dados *Landsat-like* abrangeram pequenas áreas teste, como exemplo disso ver Mello *et al.*, (2010), Viera *et al.*, (2012), Schultz *et al.*, (2015a) e Schultz *et al.*, (dados não publicados). A resolução espacial deste tipo de dado foi adequada para a classificação automática destas áreas, no entanto a nuvem foi o principal limitante para que o método fosse aplicado em grandes áreas.

Para as culturas perenes (café) e semi-perenes (cana-de-açúcar), que permanecem no campo por muito mais tempo que as culturas temporárias, o problema da nuvem tem menor interferência no mapeamento (MOREIRA *et al.*, 2008; RUDORFF *et al.*, 2010; MENDONÇA *et al.*, 2011). Um exemplo disso era o fomento produzido pelo projeto Monitoramento da Cana-de-açúcar via Imagens de Satélite (CANASAT) do INPE anualmente, para a União das Indústrias de Cana-de-açúcar (UNICA) (RUDORFF *et al.*, 2010; AGUIAR *et al.*, 2011).

Os gestores do projeto CANASAT, percebendo que os dados *Landsat-like* possuíam resolução espacial e temporal adequadas para mapeamento de cana-de-açúcar em São Paulo, desenvolveram um sistema operacional para mapeamento da área disponível para colheita da cana-de-açúcar (RUDORFF *et al.*, 2010). Os dados eram obtidos em todos os anos-safras (Junho / ano x até Julho / ano x+1) e publicados todo mês de abril de cada ano-safra.

O método proposto por Rudorff *et al.* (2010) se mostrou usual para o mapeamento da cana-de-açúcar, porque a cana-de-açúcar permanece 480 a 540 dias no campo, período este, que permite a utilização de 30 a 33 imagens multitemporais de 16 dias úteis para avaliação do crescimento desta cultura. Desta forma, o mapeamento visual da cana-de-açúcar era favorecido, por propiciar maior número de imagens multitemporais sobre seu calendário agrícola e também, por oferecer um produto anual e não em tempo quase real.

O principal avanço do método proposto pelo CANASAT para a agricultura foi a aplicação de dados orbitais multitemporais *Landsat-like* como fonte de informação espectro-temporal (SANCHES *et al.* 2005; CRUSCO, 2007), utilizada na classificação visual do tipo de “cana” (bisada, reforma ou expansão) e do tipo de colheita da “cana” (crua ou queimada) em São Paulo.

O interprete, ao se deparar com um talhão de uma cultura de calendário agrícola longo que foi colhido no inverno, supostamente poderá classificar aquele talhão como cana-de-açúcar. Por outro lado, se o intérprete não detém imagens livres de nuvens nos meses do cultivo de culturas de ciclo curto como milho e soja, isto é, janeiro/fevereiro (soja e milho de primeira safra) e março/abril (milho de segunda safra), essas culturas dificilmente serão caracterizadas nas imagens de SR (RIZZI e RUDORFF, 2005; SUGAWARA *et al.*, 2008).

Para o mapeamento de soja, por exemplo, Sugawara *et al.* (2008) comentaram que o uso de imagens multi-temporais de 16 dias, não seria suficiente para mapear a cultura no estado do Paraná. Além de uma baixa quantidade de imagens oferecidas em um calendário agrícola curto (oito imagens em 130 dias), a maior proporção de imagens com nuvens é derivada dos meses de crescimento vegetativo da soja (RIZZI e RUDORFF, 2005; EBERHARDT *et al.*, 2016).

A nuvem tem baixa transmitância no comprimento da luz visível e as lavouras de soja, que não estavam visíveis na data da imagem deste período, são ocultadas pelas nuvens nestes comprimentos de onda e assim, não serão

computadas como lavouras plantadas com soja (SUGAWARA *et al.*, 2008). Além disso, são necessárias ao menos duas imagens *Landsat-like*, do período vegetativo da cultura da soja, para mapeamento de suas lavouras (RIZZI e RUDORFF, 2005).

Tanto os dados MODIS quanto os dados *Landsat-like* possuem suas particularidades que dificultam seu uso para mapeamento de áreas com cana-de-açúcar, milho e soja em São Paulo.

Uma alternativa é proposta por autores como Adami *et al.*, (2007), Adami *et al.* (2010), Luiz *et al.* (2012) e Luiz *et al.* (2015). Estes autores mostraram a viabilidade do uso de amostragem probabilística realizada em grandes áreas, para levantamento de variáveis agrícolas. Luiz *et al.* (2015) reforçam que a amostragem é o mais aceitável método para levantamento de previsão de safras de mais de uma cultura conjuntamente, em imagens *Landsat-like* e em tempo quase-real.

2.3.3. Amostragem

Ao considerar a imagem de sensoriamento remoto como uma população de n pixels, cada qual com posições fixas no espaço e sortear uma parte deles, estamos realizando um processo de amostragem da população. A amostragem visa entender o comportamento da população, reduzindo os custos e mantendo as suas características, pelos processos de amostragem.

Os processos de amostragem mais comuns são:

(a) Amostra aleatória simples: A seleção das mostras parte do pressuposto de que todas as combinações possíveis de amostras têm igual chance de serem selecionadas (LUIZ *et al.*, 2012). As culturas agrícolas não estão igualmente distribuídas na superfície do estado de São Paulo (LUIZ *et al.*, 2015).

(b) Amostra estratificada: consiste em dividir a população amostrada em algumas classes ou estratos (sub-populações homogêneas internamente e distintas entre si), de modo que dentro da classe se reduza a variabilidade da

variável de interesse (ADAMI *et al.*, 2010). Os principais métodos utilizados para estimativa agrícola no Brasil que utilizaram imagens de sensoriamento remoto tinham por base, em seus painéis amostrais, a amostragem estratificada (EPIPHANIO *et al.*, 2000; LUIZ *et al.*, 2015).

A partir da definição do uso da amostragem estratificada como processo prático para levantamento de safras no Brasil, o próximo passo que segue é mostrar ao leitor, quais os programas brasileiros que já utilizaram estimativa probabilística para levantamento de variáveis agrícola, no País.

2.3.4. Estimativas agrícolas via painel amostral

Como visto anteriormente, a amostragem é uma ferramenta útil para levantamento de uma variável agrícola (ADAMI *et al.*, 2007; ADAMI *et al.*, 2010; LUIZ *et al.*, 2012; LUIZ *et al.*, 2015). Os valores da variável levantada são informados de forma rápida, objetiva, mais barata que o censo e com referência sobre o erro inerente ao processo de medição (coeficiente de variação CV%) (COCHRAN, 1977).

Quando se tem muitas variáveis (culturas agrícolas) a serem avaliadas, a estratificação dificilmente conseguirá obter baixos valores de CV% das estimativas. Para minimizar este problema, o ótimo seria trabalhar separadamente com cada uma delas (exemplo, projeto GEOSAFRAS). No entanto, existem formas estatísticas de se estratificar mais de uma cultura ao mesmo tempo, sem comprometer o valor do CV% da estimativa; como executado pelo método MoBARS (EBERHARDT, 2014; TRABAQUINI *et al.*, 2015; LUIZ *et al.*, 2015).

Nos parágrafos abaixo serão descritos os resultados alcançados pelas estatísticas agrícola geradas pelos projetos PREVS, GeoSafras e os resultados iniciais já alcançados para o MoBARS, no Brasil.

2.3.4.1. O programa PREVS

O programa PREVS foi o primeiro método operacional baseado em estatísticas agrícolas e sensoriamento remoto aplicado no Brasil (Eberhardt, 2014). O PREVS baseava-se nas metodologias propostas pela *National Agricultural Statistics Service* (NASS) em que amostragem probabilística de segmentos era estratificada e aplicada em imagens de sensores remotos sub-orbitais e orbitais (MULLER *et al.*, 1988).

Em campo, os consultores atualizavam as áreas agrícolas nos segmentos confeccionados sobre as imagens e posteriormente no laboratório, a expansão era realizada por estrato (EPIPHANIO *et al.* 2002). O somatório da área estimada com a cultura nos diferentes estratos fornecia a área total da cultura de interesse na região de estudo (ADAMI *et al.*, 2004).

Adami *et al.* (2007) testaram o método proposto pelo PREVS baseando a sua estratificação, no valor de área obtido pela classificação semiautomática de imagens Landsat. Os autores obtiveram valores de CV% que variaram de 6,59% (para soja) a 26,05% (para o milho). Neste método, a área de milho obtida pelo resultado do painel amostral foi subestimada em 53% da área obtida pelos dados oficiais.

É possível perceber que no método PREVS mais de uma cultura foi avaliada ao mesmo tempo (soja e milho) pela amostra probabilística. No entanto, o valor do CV% para o milho foi muito alto se comparado ao da soja. Estes resultados implicam numa baixa confiabilidade no dado produzido. Devido a isso e demais problemas operacionais, o método PREVS não é mais utilizado em escala operacional.

2.3.4.2. O programa Geosafras

O programa GeoSafras associou técnicas de amostragem estatística estratificada com características das imagens obtidas por sensoriamento remoto orbital (FIGUEIREDO, 2005; ADAMI *et al.*, 2010; LUIZ *et al.*, 2012). A base do painel amostral utilizado no programa Geosafras foi a obtenção da

intensidade amostral por estrato e expansão (EPIPHANIO *et al.*, 2000; LUIZ *et al.*, 2012).

Para definir a intensidade amostral, a fonte de dados fornecida para estratificação deste método eram as estimativas municipais geradas pelo IBGE (LUIZ *et al.*, 2012). A partir destes dados, os municípios eram então agrupados por algoritmos de agrupamento e os estratos formados. A distribuição proporcional determinava a quantidade de pixels sorteados em cada estrato (CROCHAN, 1977). O número total de pixels classificados com a cultura agrícola foi então extrapolado para a área de estudo, pelo princípio da expansão direta descrita por Epiphanio *et al.*, (2000).

Adami *et al.* (2010) e Luiz *et al.* (2012) avaliaram o método proposto pelo GeoSafras e obtiveram CVs% iguais a 7 e 3,4%, respectivamente, na estimativa de soja no Rio Grande do Sul. Adami *et al.* (2010) utilizaram (*n*) de 1.990 pixels amostrais enquanto que Luiz *et al.* (2012) dispuseram de (*n*) igual a 3.000 pixels. É postulado que, quanto maior o número de elementos amostrados de uma população, menor tende a ser o CV% obtido para a amostra (LUIZ, FORMAGGIO; EPIPHANIO, 2010).

Mesmo o painel amostral delineado para o GeoSafras tendo obtido CVs% menores que o painel amostral da PREVS, o uso deste método se restringiu apenas para uma única cultura de interesse, pois a estratificação baseava-se na área desta cultura. Além deste problema relatado ao número único de culturas avaliadas estatisticamente, o uso do Geosafras em escala operacional também foi inviabilizado economicamente, principalmente devido aos gastos gerados no processo de validação de pixels no campo (EBERHARDT, 2014).

2.3.4.3. O programa MoBARS

No método proposto por Luiz *et al.* (2015) e aqui nesta presente pesquisa chamado de programa MoBARS, as mesorregiões¹ são usadas como estratos. Estes autores comentaram que as mesorregiões possuem extensões de área relativamente semelhantes; e, também o seu uso permite a alocação de estratos que permitem a avaliação de mais de uma variável ao mesmo tempo.

Neste método, cada um dos estratos (mesorregiões) representava um peso ou “importância para o estado”, referente à proporção de sua área para área total do estado. O painel amostral foi construído com os dados oficiais do IBGE de área plantada com as culturas da cana-de-açúcar, do milho e da soja para o ano de 2011. A partir daí, o método era semelhante ao Geosafras, em que o sorteio aleatório estratificado era realizado sobre os pixels das imagens Landsat, num total de 3.000 pixels por estimativa bi-mensal.

Definido o painel amostral em virtude da subdivisão do número de pixels sorteados em cada estrato, os CVs% alcançados para as estimativas de cana-de-açúcar, milho e soja no Paraná foram de 7,67 %, 3,19 % e 2,49 %, respectivamente (TRABAQUINI *et al.*, 2015) e para o Rio Grande do Sul 7,94 % e 3,53 % na estimativa de milho e soja, respectivamente (EBERHARDT, 2014).

Estes valores de CVs% encontrados para mais de uma cultura ao mesmo tempo foram condizentes com os valores de CV% anteriormente encontrados pelo método Geosafras, para apenas uma cultura. Além disso, o levantamento dos pixels amostrais é realizado no laboratório, praticamente não necessitando trabalho de campo, do qual se dependia fortemente na metodologia Geosafras.

¹ Mesorregião é uma subdivisão dos estados brasileiros que congrega diversos municípios de uma área geográfica com similaridades econômicas e sociais, que por sua vez, são subdivididas em microrregiões. Foi criada pelo IBGE e é utilizada para fins estatísticos e não constitui, portanto, uma entidade política ou administrativa.^[1]

[1] Divisão do Brasil em Mesorregiões e Microrregiões Geográficas, IBGE - Rio de Janeiro, 1990, p. 26

O levantamento dos pixels em laboratório exige conhecimento de classificação visual de automática de imagens de sensoriamento remoto. Desta forma, estes conceitos serão descritos nos tópicos abaixo.

2.4. Classificação de imagens de sensoriamento remoto

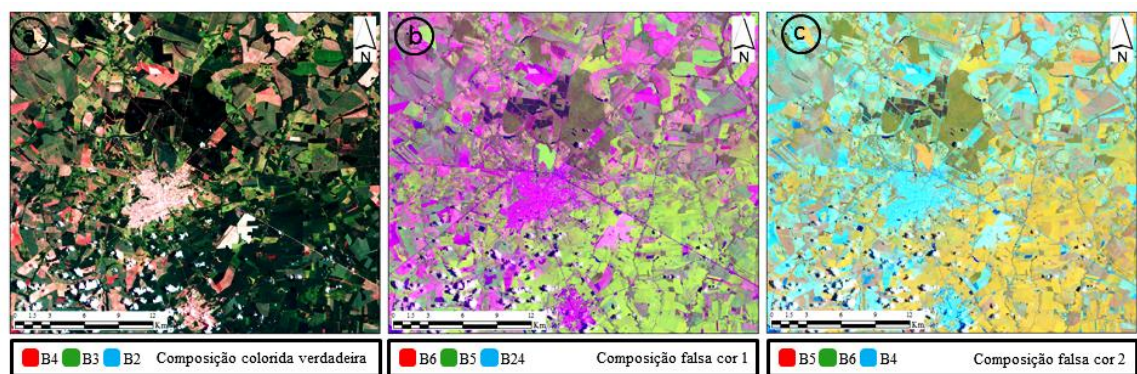
A classificação de uma imagem de sensoriamento remoto visa formar regiões mais homogêneas possíveis em uma cena (classes), utilizando para tal finalidade o conhecimento visual do intérprete (classificação visual) ou o desempenho de algoritmos (classificação digital) (GAROFALO *et al.*, 2015). Neste sentido, os intérpretes e algoritmos são chamados de classificadores de imagem.

Para chegar a tais resultados, os classificadores dependem de dados obtidos das imagens de sensoriamento remoto (MCROBERTS, 2011). Normalmente, estes dados são fornecidos em níveis de cinza, na forma de bandas espectrais (JENSEN, 2011). Em sensores multiespectrais, cada uma das bandas espectrais possui um intervalo, que compreende comprimentos de onda da região do espectro visível (VIS), infravermelho próximo (IVP) e infravermelho de ondas curtas (IVC) da radiação eletromagnética (REM) (NOVO, 2009).

A combinação de apenas três destas bandas espectrais, em ambiente tridimensional de cores RGB, origina as imagens em composição colorida (GONZALES e WOODS, 2007; MOREIRA, 2010; FLORENZANO, 2011)

A composição colorida das bandas multiespectrais (Figura 2.4). é amplamente utilizada para o procedimento de fotointerpretação de alvos agrícolas em imagens orbitais (SANCHES *et al.* 2005; CRUSCO, 2008; MOREIRA, 2010; FLORENZANO, 2011). Por outro lado, na classificação automática, as bandas espectrais são subsídios para formação dos atributos utilizados como dados de entrada para formação de modelos de conhecimento e reconhecimento automático de padrões, nas imagens (BREIMAN, 2001).

Figura 2.4 - Exemplos de composições coloridas das imagens de sensoriamento remoto.



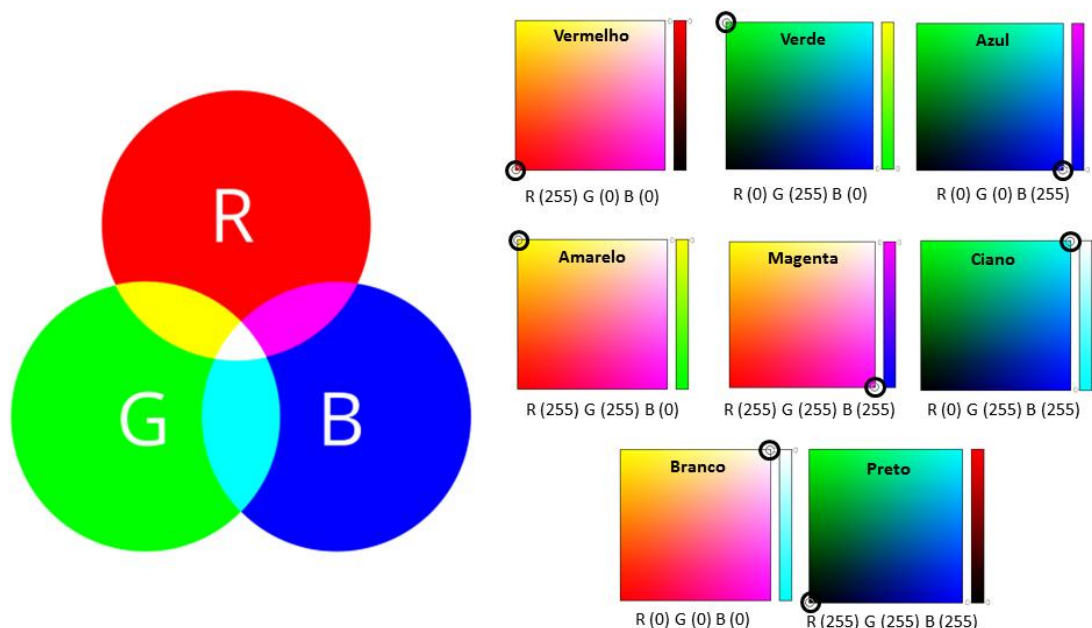
Onde: B2 (intervalo espectral do azul), B3 (do verde), B4 (do vermelho), B5 (do infravermelho próximo) e B6 (do infravermelho de ondas curtas).

Fonte: adaptado de Florenzano, (2011).

Para que um classificador possa reconhecer e separar um pixel “agrícola” de demais usos e ocupação do solo presentes na imagem, a tomada de decisão deve partir da presença ou não de vegetação fotossinteticamente ativa (VFA) presente no pixel, na data da imagem. A VFA é usada como parâmetro para informar se a cultura que ocupa o pixel, no dado momento da passagem do sensor sobre a área agrícola, ainda encontra-se verde e em pé (EBERHARDT *et al.*, 2015).

O processo de reconhecimento da VFA no pixel é feito visualmente, pela cor resultante entre os valores radiométricos das bandas espectrais, em universo RGB (Figura 2.5); ou automática via algoritmo digital. Ambos os métodos, automático ou visual, dependem dos valores radiométricos das bandas espectrais. A resolução radiométrica empregada na classificação visual, normalmente é 8 *bits*, enquanto que, na automática, a resolução radiométrica deve ser a original ao do dado processado (OLI = 16 *bits*).

Figura 2.5 – Exemplos de cores resultantes da mistura do vermelho, verde e azul no universo de cores RGB.



Fonte: adaptado de Colorzilla, (2015).

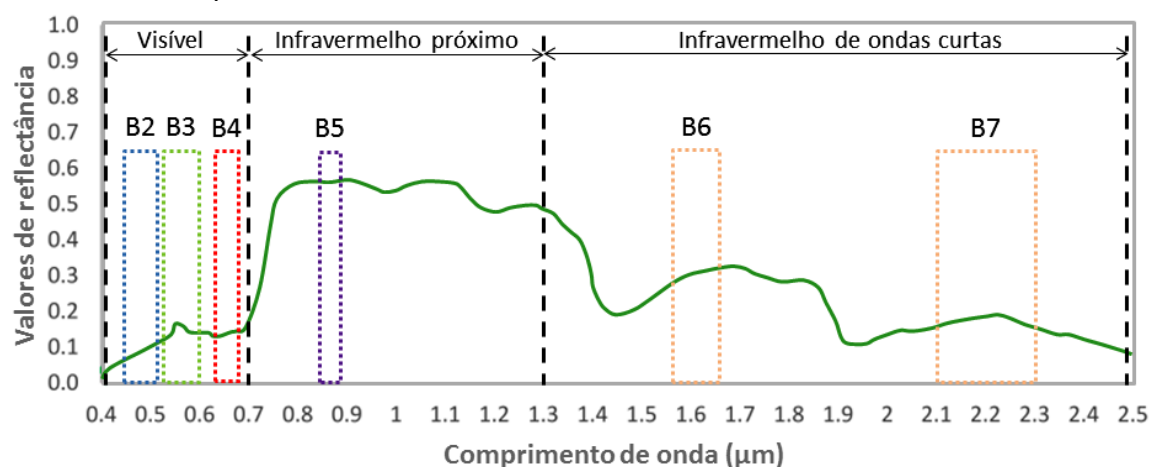
Cada cultura agrícola possui suas particularidades relacionadas a morfologia, fisiologia, tipo de dossel, manejo, estádios fenológicos, etc, que, ao serem captados radiometricamente pelo sensoriamento remoto, resultam em diferentes valores tonais (PONZONI e SHUMABUKURO, 2007). No processo de fotointerpretação, estas diferenças são de suma importância no momento da separação de cada cultura, na imagem.

Logo abaixo, serão apresentados alguns fatores que influenciam na resposta espectral da VFA em duas escalas de sensoriamento remoto: escala foliar (laboratório) e em escala de dossel (orbital).

2.4.1. VFA em escala foliar

A VFA possui comportamento espectral peculiar dentre os demais alvos da natureza (CLARCK *et al.*, 2007). Esta assertiva pode ser observada na Figura 2.6.

Figura 2.6 – Comportamento espectral de uma folha sadia em diferentes intervalos de comprimento de onda. Os intervalos B1 a B7 referem-se aos comprimentos de onda do sensor *Operator Land Imager* (OLI/Landsat-8).



Fonte: adaptado de Clarck *et al.*, (2007).

Na região do visível (B2, B3 e B4), a radiação eletromagnética (REM) incidente sobre a VFA é prevalentemente absorvida em B2 (azul) e em B4 (vermelho) e refletida em B3. A absorção é causada pela fotossíntese, que ativa os pigmentos das plantas (clorofila a, b e carotenoides) nesta faixa (B2 e B3) do espectro eletromagnético, para produção de fotossíntese (BREUNIG, 2011).

Na região do espectro eletromagnético que compreende B5, a reflectância da VFA está relacionada com a estrutura interna das folhas, número de camadas de folhas e quantidade de biomassa. A maior quantidade de biomassa ou de camadas de folhas de VFA produz um aumento não linear da reflectância em B5, e isto ocorre, principalmente, devido a interação da radiação com a estrutura da célula da superfície da folha (PONZONI e SHIMABUKURO, 2007).

Na região do espectro que compreende B6 e B7 a REM é fortemente influenciada pela quantidade de água no lúmen da célula da folha. Isto é, quanto mais água disponível na célula da planta, mais energia será absorvida na faixa do espectro que corresponde a B6 e B7 (VALERIANO, 2003; PONZONI e SHIMABUKURO, 2007).

Em resumo, os comprimentos de onda posicionados em B4, B5, B6 e B7 estão mais relacionados com os aspectos biofísicos da vegetação (SANCHES *et al.*,

2005). A ausência de VFA é comprovada por maiores valores de reflectância em B4, e o contrário ocorre quando são perceptíveis altos valores de B5. A sanidade da planta também pode ser descrita por altos valores de B5 e B6, isto é uma planta sadia absorve a REM nestes comprimentos de onda.

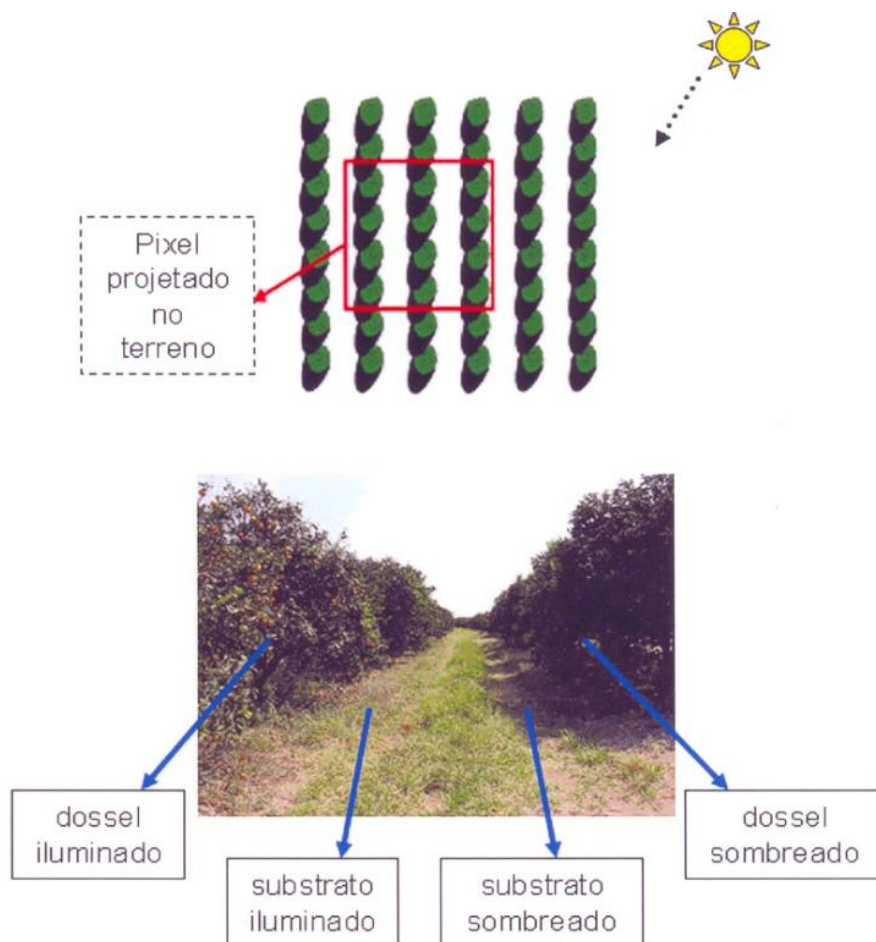
As relações matemáticas destas porções da REM originam os índices espectrais de vegetação (IVs). A principal função dos IVs é relacionar a taxa fotossinteticamente ativa das plantas, com sua biomassa (JACKSON e HUETE, 1991), além disso, eles tendem também a aumentar a quantidade de informação relacionada a VFA e diminuir os efeitos do substrato (BREUNIG, 2011).

Os IVs de bandas largas são calculados a partir de imagens multiespectrais de sensores como os *Landsat-like*. Por este motivo, os IVs serão descritos na seção 2.4.4.1.2, na página 46.

2.4.2. VFA em escala de dossel

O dossel representa o conjunto de elementos da vegetação acima do solo, somando folhas, hastes, caule, frutos, sementes, flores, etc . A resposta de VFA de um talhão agrícola em um pixel projetado no terreno é afetada pela proporção destes elementos presentes no dossel (COLOMBO *et al.*, 2003; SANCHES *et al.*, 2008) e fatores relacionados à geometria de iluminação do sensor (BREUNIG, 2011). A Figura 2.7 ilustra esta assertiva.

Figura 2.7 – Ilustração dos componentes de um pixel de uma área citrícola, visada ao nadir, destacando as árvores cítricas e as sombras nos dosséis.



Fonte: Sanches *et al.*, (2008).

Demais elementos presentes nos dosséis como a presença de substrato (*background*), idade do fechamento da copa, orientação do plantio, agregação e arranjo das folhas e galhos, idade da planta e concentração de clorofila também afetam a resposta da VFA nas imagens (COLOMBO *et al.*, 2003)

Cada cultura agrícola plantada em São Paulo possui suas peculiaridades relacionadas às variáveis suprarreferidas. Como exemplo disso, a Tabela 2.1, mostra as variáveis densidade de plantas por hectare, tipo do dossel das culturas e índice de área foliar (IAF) calculado em pleno vigor vegetativo, para as principais culturas agrícolas semi-perenes e temporárias plantadas em São Paulo.

Tabela 2.1 – Variáveis agrícolas definidas e utilizadas para as culturas da cana-de-açúcar, amendoim, mandioca, milho e soja, no estado de São Paulo

Cultura	Adensamento (Plantas / ha)	Espaçamento entre linhas (m)	IAF*	Tipo do dossel	Referência
Cana-de-açúcar	100 mil - 120 mil	1,50 a 1,80	4	Erectófila ^a	Embrapa, (2015)
Amendoim	140 mil - 150 mil	0,3 a 0,5	7,5	Planófila ^b	Abicab, (2015)
Mandioca	170 mil - 200 mil	1	8	Planófila	Embrapa, (2015)
Milho	400 mil - 800 mil	até 0,45	8	Erectófila	Embrapa, (2015)
Soja	320 mil - 400 mil	0,4 a 0,5	2 a 3	Planófila	Embrapa, (2015)

*Índice de área foliar resultante no período de pleno vigor vegetativo da cultura. ^a predomínio de folhas verticais, que tendem a aprisionar os fótons incidentes, reduzindo a reflectância, ^b as folhas tendem a se distribuir horizontalmente, aumentando a sua reflectância em relação a dosséis erectófilos (KIMES et al., 1994).

É importante compreender como este conjunto de variáveis influencia na resposta espectral da VFA das culturas agrícolas para permitir a correta interpretação delas em imagens coletadas remotamente (VALERIANO, 2003; SANCHES *et al.*, 2005; PONZONI e SHIMABUKURO, 2007).

2.4.2.1. VFA decorrente do tipo de dossel agrícola

A soja é plantada de forma adensada e é considerada uma cultura planófila com movimentos heliotrópicos que deixam as folhas mais perpendiculares à REM no início da manhã e no final da tarde (BREUNIG, 2011). O amendoim e a mandioca são normalmente plantas planófilas e plantadas de forma menos adensada que a soja.

Nos dosséis planófilos as folhas tendem a se distribuir predominantemente na horizontal, aumentando a sua reflectância em relação a dosséis erectófilos. Em geral, as plantas erectófilas, como a cana-de-açúcar e o milho apresentam máxima reflectância na direção do retroespalhamento e mínima próxima ao

nadir. Nos dosséis planófilos esse efeito é menos pronunciado (BREUNIG, 2011; BREUNIG *et al.*, 2011).

2.4.2.2. VFA em decorrência do espaçamento das plantas

Em dosséis mais esparsos (menor densidade de plantas por hectare), o solo apresenta forte influência sobre a radiância detectada pelos sensores. Algumas vezes, devido à presença de restos de vegetação não-fotossinteticamente ativos (p.ex. palha do plantio direto ou colheita da cana-de-açúcar) sobre o solo, a contribuição do solo pode ser reduzida. Desta forma, dosséis abertos podem apresentar comportamento espectral semelhante ao da vegetação seca/senescente (BREUNIG *et al.*, 2012).

No comprimento de onda do visível o sombreamento provoca forte redução na REM que chega ao sensor e no IVP, a redução é menos acentuada (BREUNIG, 2011; BREUNIG *et al.*, 2012). Contudo, para alguns tipos de dosséis erectófilos (p.ex. cana-de-açúcar e milho), a reflectância de hastes e espigas pode provocar diminuição da reflectância no visível com a diminuição do ângulo de declinação solar (APARICIO *et al.*, 2004).

Em dosséis mais fechados (maior quantidade de plantas por hectare), são mínimas as contribuições dadas pelo solo e sombra. A maior quantidade de VFA sobre o talhão tende a aumentar a quantidade de energia captada pelo sensor na região do IVP (EBERHARDT, 2014).

2.4.2.3. VFA em decorrência da geometria de iluminação do sensor

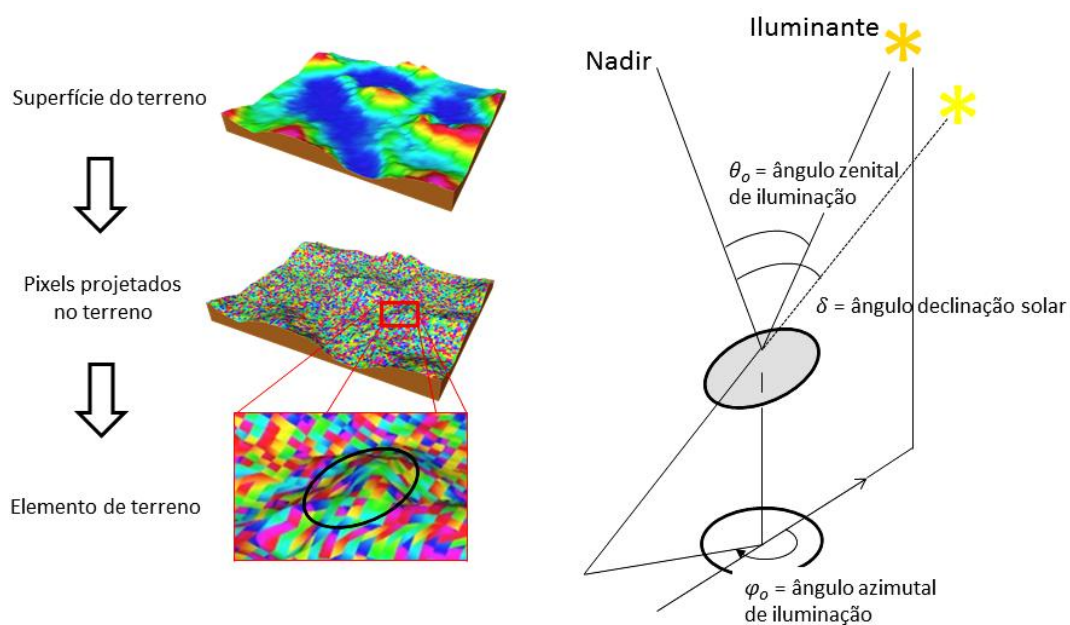
O dossel agrícola não é considerado um alvo lambertiano e as variações da reflectância destes alvos estão associadas à geometria de iluminação (ângulos zenital (θ_i) e azimutal (φ_i) de iluminação) e visada (ângulos zenital (θ_o) e azimutal (φ_o) de visada), são definidos por efeitos bidirecionais (BREUNIG, 2011).

Os dados Landsat são formados em órbita heliosincrônica, em que o satélite sobrevoa a linha do equador sempre no mesmo horário, com o sensor na

posição nadir (JENSEN, 2012; NOVO, 2009). Desta forma, variações da geometria de visada não são consideradas para este tipo de dado (USGS, 2015).

No entanto, os valores radiométricos dos dosséis agrícola projetado no pixel destas imagens, estão em função da geometria de iluminação e topografia do terreno (VALERIANO, 2011). A influência da geometria de iluminação e do fator topográfico sobre estes dados está ilustrada na Figura 2.8. Dosséis agrícolas ou florestais projetados em pixels localizados em topos de morro terão mais brilho que o mesmo dossel localizado na base do morro (VALERIANO, 2011; TELES *et al.*, 2015).

Figura 2.8 – Geometria de iluminação/sombreamento de um elemento de terreno.

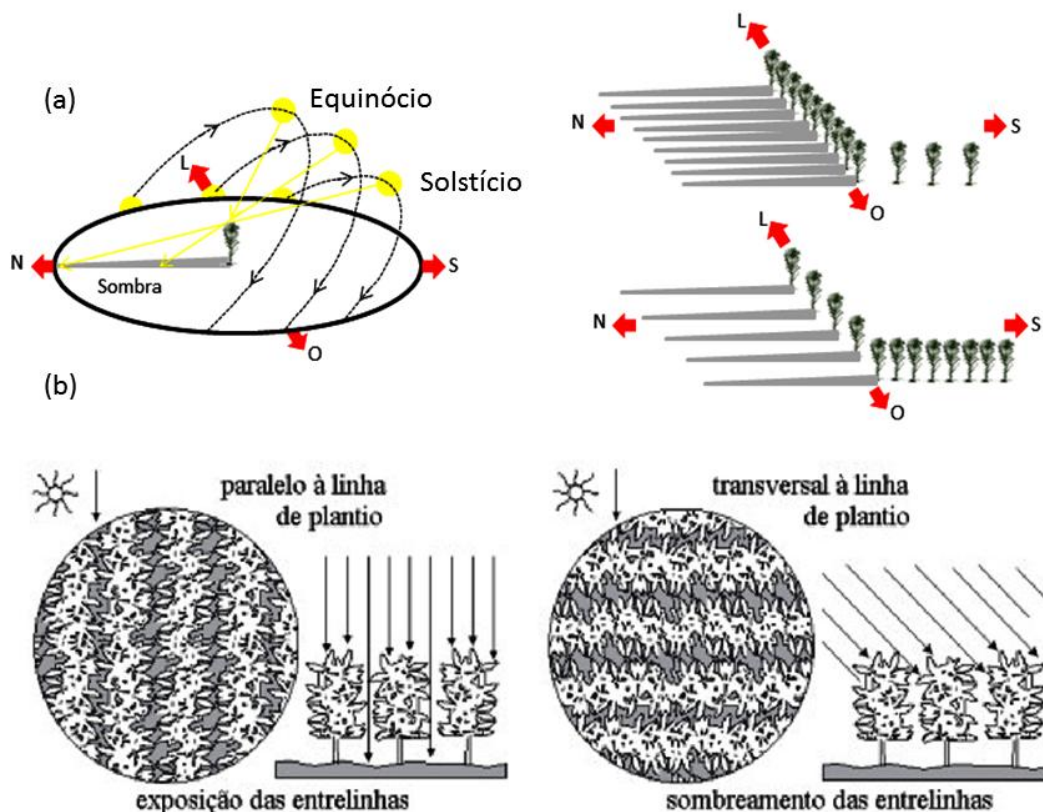


Fonte: Adaptado de Valeriano, (2011).

Ademais, o uso de imagens multitemporais usadas para classificação de culturas agrícolas necessita também, do entendimento sobre a variação do ângulo de declinação (δ) solar sobre os dados Landsat. As imagens multitemporais são tomadas dentro do calendário agrícola das culturas, e apresentam variação angular de $+ 23^\circ 27'$ (no solstício de verão), $0^\circ 0'$ (nos equinócios) e $- 23^\circ 27'$ (no solstício de inverno).

Como exemplo, em imagens obtidas com baixos ângulos solares (Figura 2.9.a), decorrentes dos meses de inverno, a sombra afeta menos os valores radiométricos dos dosséis em plantios orientados no sentido norte/sul (Figura 2.9.a), do que nos dosséis de plantios orientados no sentido leste/oeste (Figura 2.9.b).

Figura 2.9 – (a) Variação da quantidade de sombra produzida por parte de um talhão de cana-de-açúcar em função do ângulo de declinação solar (δ) e orientação do plantio. (b) Efeito do sombreamento de um dossel agrícola devido à variação do alinhamento dos raios solares.



Fonte: adaptado de Lillesand et al., (2007) e Valeriano, (2003).

Além disso, nos equinócios, os raios solares estão paralelos às entrelinhas de plantio, e assim as linhas serão mais expostas. Nestes meses os dosséis terão mais efeito de substrato sobre os valores radiométricos dos pixels (Figura 2.9.b). No solstício, os raios solares estarão transversais às linhas de plantio, desta forma, a sombra terá maior efeito sobre os valores radiométricos (Figura 2.9.b) (VALERIANO, 2003).

2.4.3. Classificação visual de imagens de sensoriamento remoto

2.4.3.1. Feições interpretáveis nas imagens

A cor é a primeira feição das imagens de sensoriamento remoto que pode ser utilizada no momento da classificação de culturas agrícolas (SANCHES et al., 2005). Usualmente, em ambiente SIG é utilizada a composição falsa cor para caracterizar diferentes usos e ocupação do solo em imagens *Landsat-like* (FLORENZANO, 2011).

Na composição RGB falsa cor, as bandas do comprimento de onda do b-5 (IVP) e b-6 (IVC) são utilizadas nos canais RGB e configuram diferentes cores para as imagens.

Por exemplo, ao trabalhar com florestas, o PRODES utiliza a composição R-b4 G-b5 B-b6, tornando a VFA verde e o solo exposto magenta. Esta composição colorida falsa cor possibilita a fácil separação entre áreas com desflorestamento (magenta) e áreas de floresta natural (verde) (PROJETO PRODES, 2013).

Quando a composição colorida R-b5 G-b6 B-b4 é aplicada às imagens, a VFA se torna vermelha e o solo exposto, ciano. O vermelho é uma cor de fácil compreensão e suas variações são melhores detectadas pelo olho humano (SANCHES et al., 2005). Desta forma, qualquer variação do matiz vermelho vai depender especificamente da quantidade de energia emitida ou absorvida pela vegetação em estudo, possibilitando a melhor separação entre elas, em uma única imagem.

Sanches et al. (2005) e Rizzi e Rudorff (2005) mencionaram que a composição R-b4 G-b5 B-b3 dos sensores *Thematic Mapper* (TM) e *Enhanced Thematic Mapper Plus* (ETM⁺) correspondente a composição R-b5 G-b6 B-b4 do OLI, melhor representou as diferenças entre as culturas agrícolas como a cana-de-açúcar e milho (magenta) e soja (amarela).

Além disso, Sanches et al. (2005) comentaram também que nesta mesma composição, áreas de solo preparado têm cor escura (roxo a preto), áreas com palhada adquirem matiz ciano e parcelas agrícolas com plantas em senescência adquirem diferentes matizes (verde escuro para o milho e rosa claro para a soja).

Além da cor ou matiz, Florenzano (2011) e Moreira (2012) mencionaram que o foto-intérprete dispõe de outras informações importantes no momento da classificação visual de alvos em imagens de SR. A tonalidade, a textura, o tamanho, a forma, o padrão e a localização geográfica, são outros elementos de interpretação de imagens, indispensáveis para formulação de um aceitável protocolo de classificação.

Estes elementos são característicos para interpretação de culturas agrícolas em uma única imagem *Landsat-like*. Isto é, na prática, se o intérprete tiver disponível apenas uma imagem OLI R-b5 G-b6 B-b4 do mês de maio, por exemplo, e se deparar com um talhão agrícola de cor vermelha, a dúvida surgirá. Este talhão poderá ser milho safrinha ou cana-açúcar (RUDORFF e SUGAWARA, 2007; RUDORFF *et al.*, 2010). Da mesma forma, apenas a imagem de janeiro não é suficiente para separação de soja e mandioca, nesta mesma composição colorida (SCHULTZ *et al.*, 2015b).

Quando é possível utilizar mais uma imagem multitemporal de uma mesma cena, a fenologia das culturas pode ser relacionada às datas da imagem e assim auxiliar o intérprete no momento da classificação visual dos alvos na imagem base (CRUSCO, 2007).

2.4.3.2. Imagens multitemporais

O plantio de cana-de-açúcar, milho e soja, assim como o manejo do solo praticado em determinada área tradicionalmente agrícola, são relativamente constantes durante alguns anos. Esta característica foi denominada de tradição agrícola (ORTIZ *et al.*, 1997).

Em alguns municípios é possível relacionar um padrão no comportamento do manejo destas culturas, com imagens de sensoriamento remoto, utilizando para isto, a resposta da VFA no tempo (IPPOLITI-RAMILO *et al.*, 2003).

Sanches *et al.* (2005) nomeiam este padrão de comportamento espectro-temporal (CET) das culturas. De acordo com o CET, o intérprete pode utilizar imagens multitemporais de épocas conhecidas dos seus calendários agrícolas como fonte de informação, para separação entre elas.

Um exemplo disso é dado por Rudorff e Sugawara (2007) e Rudorff *et al.* (2010). Estes autores concluíram que imagens provenientes dos meses de março e abril, em composição colorida correspondente a composição colorida falsa cor R-b5 G-b6 B-b4 do OLI, são indispensáveis na separação visual de cana-de-açúcar de culturas agrícolas temporárias como, o milho e soja.

Desta forma, os CETs de três culturas (cana-de-açúcar, milho e soja) foram separados utilizando para isso, apenas duas imagens dos seus calendários agrícolas. O entendimento do calendário agrícola das culturas de maior expressividade em área para o estado de São Paulo é assim, outra característica que orienta o intérprete no momento da classificação visual. O CET das culturas é considerado aqui, outra feição interpretável da imagem.

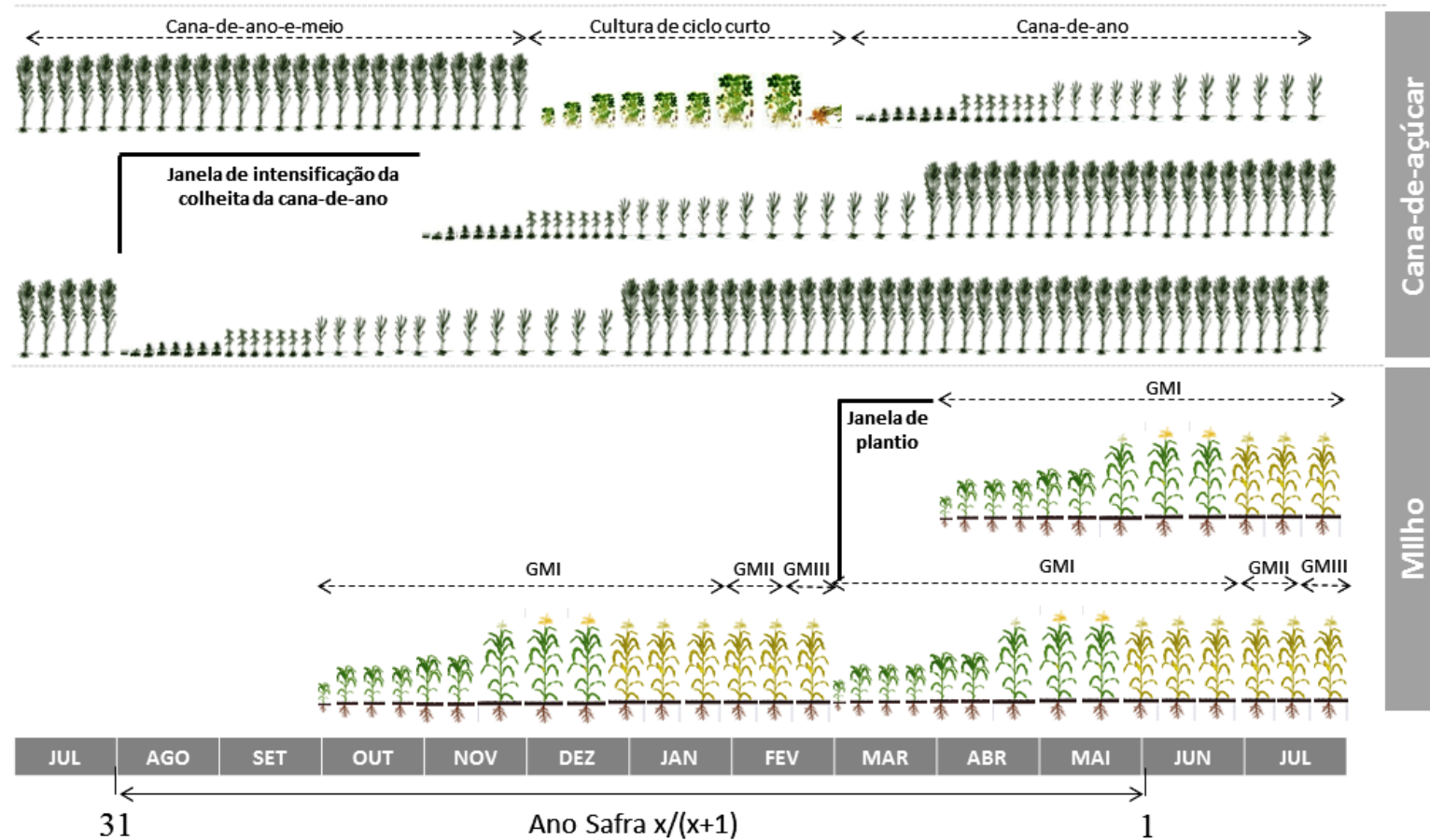
2.4.3.3. Calendário agrícola

De acordo com os dados da PAM de 2013 (IBGE, 2015), a principal cultura plantada em São Paulo é a cana-de-açúcar, e as cinco principais culturas temporárias são o milho, soja, feijão, amendoim e mandioca. No entanto, de acordo com Schultz *et al.* (2015a), a mandioca e o amendoim são culturas que possuem comportamento espectro-temporal semelhantes à soja. Desta forma, o calendário agrícola da cana-de-açúcar, milho, soja, amendoim e mandioca foram apresentados nesta pesquisa.

2.4.3.4. O calendário agrícola da cana-de-açúcar e milho

O ciclo da cana-de-açúcar é semi-perene e sua colheita pode ser realizada de 360 dias (cana-de-açúcar de ano) a 540 dias (cana-de-açúcar de ano-e-meio) após o plantio (Figura 2.10). A cana-de-ano normalmente permanece como rebrota de 5 a 7 anos de plantio, e posterior a este período as socas de cana-de-açúcar são reformadas (Figura 2.10).

Figura 2.10 – Calendários agrícolas utilizados para o plantio e colheita da cana-de-açúcar e milho no estado de São Paulo.



Fonte: adaptado de MAPA, (2005), Vieira *et al.*, (2012) e Schultz *et al.*, (2015a).

O manejo aplicado para a cana-de-ano-e-meio é diferente do usado para a cana-de-ano. Normalmente depois de realizada a colheita nos talhões de cana-de-ano-e-meio, eles são reformados com culturas de ciclo curto, como soja e amendoim (Figura 2.10) (RUDORFF *et al.*, 2010; RUDORFF e SUGAWARA, 2007). Posteriormente a colheita das culturas de ciclo curto, o agricultor pode optar em manejar, novamente, o talhão com cana-de-ano-e-meio ou de ano.

A colheita da cana-de-açúcar é praticada nos meses mais secos do ano (AGUIAR *et al.*, 2011, 2008; MELLO *et al.*, 2010; RUDORFF *et al.*, 2010) e a intensificação ocorre de agosto a outubro (AGUIAR *et al.*, 2008) (Figura 2.10). Antes da aderência do Brasil ao uso das práticas exigidas pelo Protocolo do Etano Verde (AGUIAR *et al.*, 2011), esta época de colheita era propícia para se realizar a queima dos talhões (prática da colheita com queima) e, além disso, o fator sazonal influenciava a perda de água pelos colmos e favorecia o aumento da sacarose nas plantas (AGUIAR *et al.*, 2008).

O milho é uma cultura de ciclo curto plantado em São Paulo como primeira safra (outubro a dezembro), milho de segunda safra (março a abril); e milho irrigado, que pode ser plantado em épocas distintas às demais (Figura 2.10). As datas de plantio permitidas para o milho estão descritas nas normativas do Ministério da Agricultura, Pesca e Abastecimento (MAPA) (MAPA, 2015).

A área total de milho plantada em SP foi 434 mil ha em 2013. No entanto, 74 % desta área foram plantados como milho de segunda safra (safrinha), 15 % como milho de primeira safra e 11 % como milho irrigado (IEA, 2015). É possível perceber que, o milho “safrinha” é o principal manejo adotado para o plantio de milho, no estado de São Paulo.

As cultivares de milho variam de acordo com o seu grau de maturação fisiológica (GM), e são classificadas em três grandes grupos. As cultivares GM I, levam menos de 90 dias para serem colhidas e as cultivares GMII possuem um ciclo um pouco mais longo, com 120 dias. Já as cultivares GM III são as cultivares de ciclo mais longo, e seu plantio dura de 120 a 145 dias em média (Figura 2.10).

2.4.3.5. O calendário agrícola da soja, amendoim e mandioca

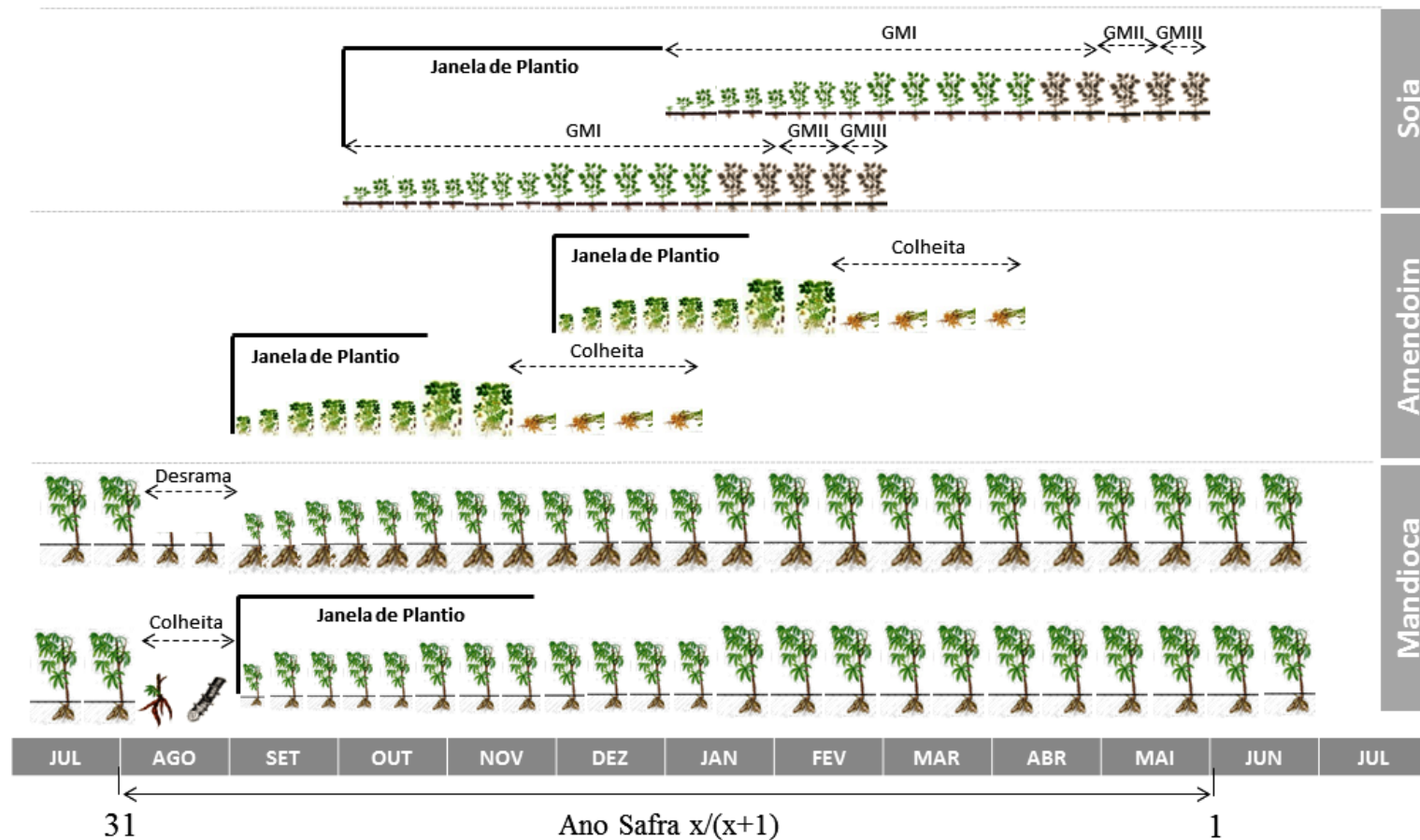
A soja é uma leguminosa e o seu calendário agrícola é determinado pelo zoneamento agrícola dessa cultura, descrito nas portarias do MAPA (MAPA, 2015). De acordo com os dados de riscos climáticos para a cultura da soja, em São Paulo são plantadas soja de ciclo curto GMI (< 120 dias), ciclo médio GMII (121-130 dias) e ciclo longo GMIII (>131 dias) (Figura 2.11) (ADAMI, 2010).

Outra leguminosa plantada em SP é o amendoim. O amendoim é uma planta de crescimento rasteiro ou ereto que leva 90 a 140 dias para completar seu ciclo de crescimento (Figura 2.11). Em São Paulo, os melhores meses para o plantio de amendoim são setembro/outubro, na safra de verão, e dezembro/janeiro quando o amendoim é utilizado na reforma de talhões com cana-de-açúcar (RUDORFF *et al.*, 2010).

Outra cultura agrícola plantada em São Paulo é a mandioca. A mandioca, em São Paulo, normalmente é plantada nos meses de setembro a novembro, e as plantas podem permanecer de 240 dias (mandioca precoce) a 730 dias (mandioca de dois anos), plantadas. Um ano após o plantio, o produtor também pode optar pela desrama das plantas, para favorecer a incorporação de carboidratos pela raiz (Figura 2.11).

Nos municípios do Sul do estado de São Paulo, normalmente são encontrados talhões de mandioca próximos às fecularias. A data de plantio da mandioca, o ciclo da cultura e a data da desrama são fatores que interferem no desenvolvimento da parte aérea da planta e nos valores de produtividade tuberosa (FAGUNDES *et al.*, 2010; AGUIAR. *et al.*, 2011).

Figura 2.11 – Calendários agrícolas utilizados para o plantio e colheita da soja, amendoim e mandioca no estado de São Paulo.



Fonte: adaptado de CONAB (2015), Embrapa (2015) e Schultz et al. (2015a).

Fagundes *et al.* (2010) constataram que a mandioca de ano, plantada em setembro, produziu mais que demais plantios realizados em meses recorrentes. Aguiar *et al.* (2011) comentaram que quando podada no mês de junho, a mandioca de dois anos produziu mais massa fresca que a mandioca podada em outubro. Ao se basear nestas épocas de manejo, a Figura 2.11 ilustra o calendário agrícola da mandioca para São Paulo.

Na produção de mandioca há predominância da pequena propriedade, portanto, é realizada, por pequenos produtores. É conhecida como a “poupança” do sítio. Segundo (SBM, 2015), 68% da raiz processada pela indústria provém de produtores com menos de 12 hectares; outros 20% são fornecidos por produtores que cultivam entre 12 e 70 hectares; e apenas 12% vem de lavouras de mais de 70 ha..

2.4.4. Classificação digital de imagens de sensoriamento remoto

Como falado anteriormente, na classificação visual, os intérpretes extraem feições de interesse das cenas por meio de composições coloridas RGB entre as bandas espectrais de uma única imagem, ou de várias imagens que abrangem o calendário agrícola das culturas avaliadas.

Na classificação digital ou também chamada de automática, os classificadores extraem feições por meio do reconhecimento automático de padrões e de objetos homogêneos, representados em uma única classe (VIERA *et al.*, 2012; GAROFALO *et al.*, 2015)

A classificação digital pode ser pixel a pixel ou por objeto/região. Na classificação pixel a pixel, os valores digitais dos pixels são utilizados como unidade mínima da classificação. Para classificação segmento/região, os pixels são agrupados em regiões com pixels mais homogêneas entre si (*Object Based Image Analysis – OBIA*), tornando os segmentos como unidades mínimas de classificação (MYINT *et al.*, 2009; BLASCHKE, 2010; STEFANSKI *et al.*, 2013; PEÑA *et al.*, 2014; SCHUTZ *et al.*, 2015a).

A abordagem de classificação de imagens orientada a objetos possui vantagens se comparada à classificação por pixel, são elas:

a) reduzir os ruídos inerentes ao processo de classificação pixel a pixel, como o efeito sal-e-pimenta no mapa temático (MYINT *et al.*, 2009; ZHANG e XIE, 2012; SCHULTZ *et al.* 2015a);

b) melhorar a relação entre os valores dos atributos de cada segmento com as variáveis biofísicas da vegetação (BLASCHKE, 2010; GALLETTI e MYINT, 2014; PEÑA *et al.*, 2014);

c) reduzir o tempo de processamento da classificação, devido à redução do número de objetos a serem classificados (BLASCHKE, 2010; TAŞDEMİR *et al.*, 2012).

Autores como Peña-Barrgán *et al.* (2010), Vieira *et al.* (2012), Long *et al.* (2013), Stefanski *et al.* (2013), Stefanski *et al.* (2014), Peña *et al.* (2014), Schultz *et al.*, (2015) e Schultz *et al.* (dados não publicados) comprovaram em seus trabalhos que OBIA é usual para classificação automática de culturas agrícolas em imagens *Landsat-like*.

Além disso, OBIA teve melhores resultados de mapeamento se comparada com a abordagem pixel a pixel utilizada na classificação de culturas agrícolas em imagens *Landsat-like* ao utilizar algoritmos baseados em árvore de decisão (MYINT *et al.*, 2009; DURO *et al.*, 2012; LONG *et al.*, 2013).

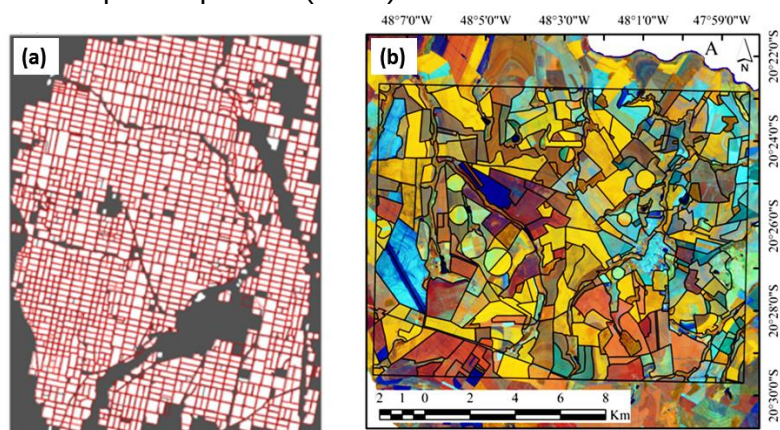
Na agricultura, os segmentos originados sobre uma imagem de sensoriamento remoto podem até modelar o comportamento espacial de um talhão agrícola. Neste caso, a classificação é gerada por talhão (*per-field classification*) e não por segmento (FORSTER *et al.*, 2010).

Em trabalhos realizados por Long *et al.* (2013) e Yan e Roy, (2014) foi possível aplicar a classificação em nível de parcela de campo (*per-field classification*²). No entanto, Long *et al.* (2013) utilizaram dado vetorial cadastral dos talhões agrícolas e para Yan e Roy (2014) a dinâmica espaço-temporal das culturas agrícolas na região do *Corn Belt* nos EUA permitiu a execução de tal procedimento.

O Brasil não detém de dados vetoriais ao nível de parcela agrícola e, além disso, as áreas agrícolas não respeitam um padrão de ocupação espacial (Figura 2.12). A única lei que define a orientação de uma propriedade rural no Brasil é a obrigação do terreno fazer divisa, em uma de suas arestas, com uma estrada municipal ou ferrovia (Lei nº 10.406/2002). Demais orientações estarão em função do relevo do terreno, da posição do sol, do tipo de cultura plantada, etc.

A Figura 2.12 compara as parcelas agrícolas encontradas na região do *Corn Belt* nos EUA e as parcelas agrícolas encontradas na região de Assis /SP.

Figura 2.12 – Comparação entre a forma visual dos talhões agrícolas de uma área do *Corn Belt* (EUA) e de uma área com soja e cana-de-açúcar no município de Ipuã/SP (Brasil).



Fonte: Adaptado de Yan e Roy (2014) e Schultz *et al.*, (dados não publicados).

² Dados originados da remoção das variações presentes em um mesmo talhão agrícola, decorrentes do processo da segmentação

^[3] Forster et al. (2010).

A definição de segmentos que correspondem aos objetos de campo, ainda é um desafio no método OBIA (BLASCHKE, 2010).

2.4.4.1. *Object-based Image Analysis (OBIA)* na agricultura

Em OBIA, quatro passos são fundamentais e realizados anteriormente ao processo de classificação da imagem de sensoriamento remoto (VIERA *et al.*, 2012; STEFANSKIN *et al.*, 2013; SCHULTZ *et al.*, 2015a; Schultz *et al.*, dados não publicados):

- (a) segmentação da imagem de sensoriamento remoto;
- (b) extração de atributos;
- (c) formação do modelo de conhecimento;
- (d) e classificação.

No presente trabalho, adicionamos entre o processo (c) e (d) um item a mais no procedimento OBIA. Este novo processo ficou conhecido como “avaliação das segmentações produzidas”. Este método foi descrito por Stefanski *et al.* (2013) e Schultz *et al.* (2015a), e foi utilizado para eliminar a subjetividade gerada pela fase da segmentação das imagens, no momento da formação do modelo de classificação e classificação final (SCHULTZ *et al.*, 2015a).

2.4.4.1.1. Segmentação das imagens

No processo de segmentação, os pixels das imagens passam por uma fase de pré-classificação, isto é, ocorre uma subdivisão de uma imagem em regiões homogêneas independentes, considerando algumas características intrínsecas, como, por exemplo, os níveis de cinza dos pixels (SANCHES, 2004; ESPÍNDOLA, 2006).

Para que os pixels sejam agrupados em regiões homogêneas, o processo de segmentação de uma imagem é feito por meio de segmentadores automáticos. Os segmentadores exigem a definição *a priori* de parâmetros de segmentação,

informados manualmente pelo usuário, em cada tipo de algoritmo (OLIVEIRA *et al.*, 2003; STEFANSKI *et al.*, 2013; GAROFALO *et al.*, 2014; SCHULTZ *et al.*, 2015b; SCHULTZ *et al.*, dados não publicados).

O conjunto de segmentações gerado pelo uso destes diferentes parâmetros pode resultar em *enésimas* segmentações (OLIVEIRA *et al.*, 2003; STEFANSKI *et al.*, 2013). Porém, apenas uma delas será utilizada no processo de classificação da imagem (SCHULTZ *et al.*, 2015a, SCHULTZ *et al.*, dados não publicados).

A segmentação automática tenta reproduzir as formas dos objetos do campo nas imagens de sensoriamento remoto (OLIVEIRA *et al.*, 2003). No entanto, os algoritmos segmentadores que possuem tal habilidade, ainda têm certas limitações de uso.

Por exemplo, uma grande quantidade de segmentações pode ser obtida pelas variações do valor dos quatro parâmetros do algoritmo chamado Segmentador Multi-resolução SM-R (STEFANSKI *et al.*, 2014; SCHULTZ *et al.*, 2015a). Assim, as segmentações devem ser testadas/supervisionadas (STEFANSKI *et al.*, 2014), para haver relação entre os segmentos confeccionados automaticamente e os objetos de campo presentes na cena de sensoriamento remoto (OLIVEIRA *et al.*, 2003).

No tópico 2.4.4.1.4 da página 48 serão comentadas formas de se obter segmentos homogêneos e condizentes com a realidade dos objetos de campo. Os tópicos a baixo seguem uma ordem lógica no processo de OBIA.

2.4.4.1.2. Extração de atributos das imagens de sensoriamento remoto

Os objetos formados, decorrentes do agrupamento dos pixels semelhantes entre si, terão características próprias, chamadas de *atributos*. Estes atributos são formados por regras matemáticas, que incluem todos os valores espectrais dos pixels internos a um segmento. Um exemplo disso é a *média da banda do IVP*, dada pela média geral de todos os pixels que pertencem a um mesmo segmento.

Assim, um objeto possui, em comparação a um pixel, além de valores espectrais, numerosos outros atributos, incluindo a forma, a textura, a morfologia e o contexto, dentre outros, que são expressas pelas variáveis contínuas; e podem ser empregados em análise de imagem (NAVULUR, 2007; VIEIRA *et al.*, 2012) para formar *variáveis categóricas*.

Os atributos são extraídos por *softwares* como *eCognition* (Trimble Geospatial, 2009), *InterImage* (COSTA *et al.*, 2008), *Software R* (R DEVELOPMENT CORE TEAM, 2015) e *Orfeo Tollbox* (OTB DEVELOPMENT TEAM, 2015), os quais possuem a capacidade de extrair estas características, por já terem incluído em suas plataformas, algoritmos especializados para cálculo de métricas.

Diferentes fontes de dados de sensoriamento remoto podem ser adicionadas a estas plataformas e vários atributos podem ser associados a cada um dos segmentos formados pelo processo de segmentação. No entanto, uma grande quantidade de atributos pode, em certos classificadores, causar sobreajuste (*overfitting*) no momento de gerar o modelo de conhecimento; e assim causar a redução dos resultados de classificação (BREIMAN, 2001; LONG *et al.*, 2013).

Quanto mais relacionados estes atributos estão à classe em estudo, melhores serão os resultados de classificação daquela classe avaliada (IMMITZER *et al.*, 2012; SCHULTZ *et al.*, dados não publicados), e quanto menos correlacionadas as classes estão, poucos atributos são necessários para separação entre elas (LONG *et al.*, 2013).

Técnicas de seleção de atributos podem ser aplicadas em determinados algoritmos de classificação automática supervisionada (IMMITZER *et al.*, 2012) para reduzir problemas de sobreajuste e selecionar aqueles atributos que mais condizem com as classes de interesse. Outra forma pode ser feita, analisando os resultados já obtidos anteriormente pela literatura, ao utilizar atributos derivados de imagens *Landsat-like*, na separação automática de culturas agrícolas (Tabela 2.2).

Tabela 2.2 – Resultados literários que mostram os principais atributos utilizados nos modelos de conhecimento para classificação automática de culturas agrícolas no Brasil (1 – 4) e nos EUA (5 – 6).

Cultura classificada	Atributos importantes	Referência
Cana-de-açúcar	Textura de Haralick de imagens multitemporais ^a	1- Vieira <i>et al.</i> (2012)
Cana-de-açúcar Culturas de verão	Frações do MLME ^b	2- Atzberger <i>et al.</i> , (2014)
Soja Cana-de-açúcar	NDVI de imagens multi-temporais ^c	3- Schultz <i>et al.</i> 2016
Amendoim Mandioca Soja Cana-de-açúcar	Media e desvio padrão das bandas b-4, b-5 e b6 de imagens multitemporais	4- Schultz <i>et al.</i> , (2015)
Culturas de verão	[IVV, NDVI, GNDVI, NDVI b-6, NDVI b-7 e NDSVI] ^d	5 - Peña-Barragán <i>et al.</i> , (2011) e Peña <i>et al.</i> , (2014)
Cereais Ervilha	NDVI e média das bandas b-6 e b-7 de imagens multi-temporais	6 - Long <i>et al.</i> , (2013)

Onde: ^a Haralick *et al.* (1973), ^b modelo linear de mistura espectral (MLME) (SHIMABUKURO e SMITH, 1991), ^c índice de vegetação da diferença normalizada (NDVI) (ROSE *et al.*, 1974), ^d índice de vegetação do verde (IVV) , índice de vegetação da diferença normalizada do verde (GNDVI) (GITELSON *et al.*, 2002), NDVI da banda b-6 (NDVI b-6) e NDVI da banda b-7 (NDVI b-7) (MCNAIER *et al.* 1993) e índice de vegetação da diferença normalizada da senescência (NDSVI) (QI *et al.*, 2002).

Trabalhos de Vieira *et al.* (2012), Atzberger *et al.* (2014), Schultz *et al.* (2015a) e Schultz *et al.* (dados não publicados) retrataram o uso de alguns atributos, derivados das bandas espectrais de sensores *Landsat-like*, que foram utilizados na separação de culturas agrícolas no estado de São Paulo, Brasil. Por outro lado, Peña-Barragán *et al.* (2010), Long *et al.* (2013) e Peña *et al.* (2014) relataram alguns atributos importantes, para classificação de culturas agrícolas no cenário agrícola dos EUA.

Ademais, o uso de atributos obtidos em imagens multitemporais, procedentes de datas estratégicas dos calendários agrícolas das culturas, pode aumentar de 12 a 15 % a exatidão do mapeamento do mapa final produzido (EL-MAGD e TANTON, 2003; LONG *et al.*, 2013; SCHULTZ *et al.*, dados não publicados).

2.4.4.1.3. Formação do modelo de conhecimento

Para formação do modelo de conhecimento supervisionado o usuário necessariamente terá disponível um conjunto de dados de treinamento/calibração. Estes dados são extraídos da imagem. no momento da seleção de segmentos amostrais e seus respectivos atributos relacionados a eles (VIEIRA *et al.*, 2012).

Em determinados algoritmos como o *Random Forest*, uma parte destes dados de treinamento é utilizada para formar um modelo de conhecimento e a outra parte é utilizada para a validação do modelo. Neste caso, apenas o modelo está sendo validado (SCHULTZ *et al.*, 2015a). Este tipo de validação *out-of-bag* (OOB) é aplicado por técnicas de amostragem com reposição, também chamada de *bootstrap* e descrita por Breiman, (2001).

Ao replicar o modelo de conhecimento para outra área, fonte dos dados de treinamento, o usuário estará formando um *Wall-to-Wall Land Cover Map* (Mapa Temático). O mapa temático produzido só será validado, se as amostras utilizadas para validação foram independentes daquelas utilizadas para treinamento do modelo (FOODY, 2004).

A validação do modelo tem por objetivo analisar a coerência do resultado gerado pelo aprendizado do modelo, com a realidade informada pelas amostras de treinamento (STEFANSKI *et al.*, 2013); e a validação do mapa é quantificada, para obter a coerência entre os dados produzidos (classes) com a realidade de campo (POWELL. *et al.*, 2004).

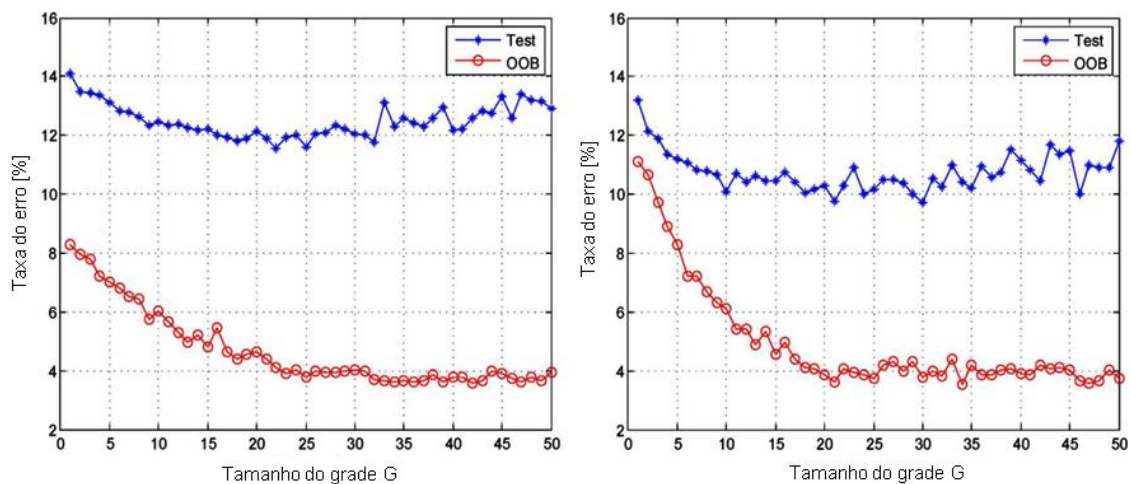
A coerência dos resultados gerados pelo modelo e pelo mapa pode ser obtida pela matriz de confusão. Os principais parâmetros utilizados para quantificar a matriz são: Índice Kappa de Cohen (IK), Exatidão Global (EG), a exatidão do usuário (EU), a exatidão do produtor (EP) e variância do Kappa (*var.K*) (CONGALTON e GREEN, 2009).

2.4.4.1.4. Avaliação das segmentações resultantes

Para cumprir tal objetivo, de selecionar a segmentação mais condizente com os objetos de campo, autores como Peña-Barragan *et al.* (2011), Fernandes *et al.* (2012), Vieira *et al.* (2012), Stefanski *et al.* 2013; Penã *et al.* (2014), Stefanski *et al.* (2013), Schultz *et al.* (2015a) comentaram que diferentes métodos podem ser utilizados.

O método *super-pixel classification* (SPc) (STEFANSKI *et al.* 2013) é o mais indicado entre eles para avaliar a melhor segmentação de um conjunto de segmentações, via supervisão (Figura 2.13).

Figura 2.13 – Comparação do erro de classificação usando área teste (linha azul) e técnica OOB com diferentes segmentações (*grid size G*).



Esquerda: baseado em Luxemburgo-BE; Direita: baseado na Alemanha.

Fonte: Stefanski *et al.*, (2013).

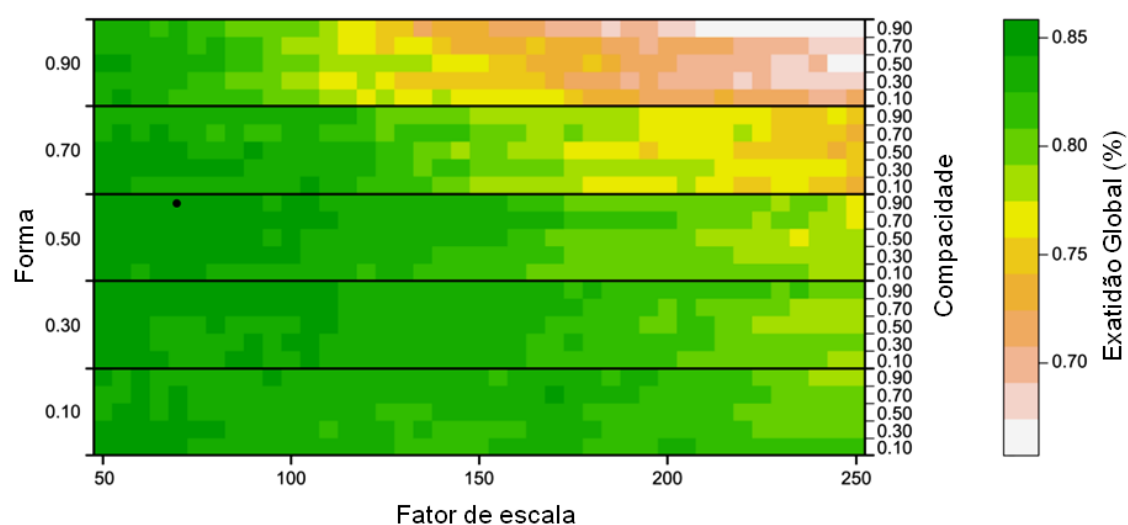
Nele, em cada segmentação gerada, é formado um modelo de conhecimento validado de duas diferentes formas: pela técnica OOB (linha vermelha) e por comparação com áreas testes, foto-interpretadas (linha azul).

Ainda na Figura 2.13, cada um dos modelos de conhecimento produz dois resultados de exatidão. Aquele modelo, representado no eixo x pelo nome *Grid Size G*, que produziu menor razão do erro (%) é aplicado para os dados independentes dos dados de treinamento.

A mesma técnica descrita por Stefanski *et al.* (2013) foi modelada para as condições agrícolas brasileiras utilizando o algoritmo Segmentador Multi-resolução (SM-R), e chamada de *self-guided segmentation* (SCHULTZ *et al.*, 2005a). No trabalho publicado por Schultz *et al.* (2015a), para cada combinação gerada dos parâmetros deste segmentador, que no caso são três: fator de escala (fe), forma (fm) e compacidade (cp), um resultado de exatidão global do modelo foi gerado, isto é, foram avaliadas 1025 diferentes modelos de classificação.

A Figura 2.14 mostra o resultado da exatidão global (coerência entre os dados de referência e classificação) de 677 pixels usados no treinamento do modelo de conhecimento *Random Forest*. Para este algoritmo, cada um dos modelos de conhecimento formado foi validado pela técnica OOB.

Figura 2.14 – Ilustração do comportamento das exatidões globais dos modelos *self-guided segmentation* testados para encontrar a melhor segmentação da área de estudo.



No eixo x, o parâmetro fator de escala e no eixo y, os parâmetros forma e compacidade. Os valores de exatidão global foram ilustrados nas escalas de cores.

Fonte: adaptado de Schultz *et al.*, (2015a)

Na Figura 2.14, entre 1025 modelos de classificação derivados da variação dos parâmetros Fe, Fm e Cp do algoritmo SM-R, o melhor resultado da exatidão global foi obtido quando o modelo operou com valores iguais a (70, 50 e 90), respectivamente para estes parâmetros (ponto em cor preta na Figura

2.14). Esta segmentação é mais condizente com a realidade de campo, dado o valor de exatidão global do modelo.

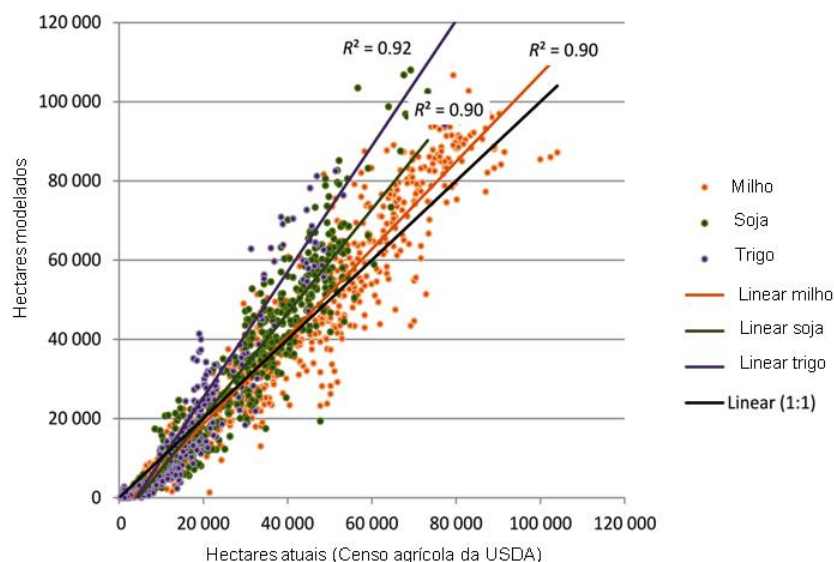
2.4.4.1.5. Replicação do modelo de conhecimento – classificação

No sensoriamento remoto, o termo *Wall-to-Wall Land Cover Map* expressa o significado da extrapolação do modelo de conhecimento formado, para toda área de estudo a ser classificada (GEBHARDT *et al.*, 2014). O mapa temático deve ser validado por amostras independentes daquelas utilizadas no processo de treinamento do modelo (FOODY, 2004). No entanto, a obtenção de dados de validação nem sempre é trabalho de fácil execução (POWELL *et al.*, 2004).

A validação de mapas temáticos agrícolas pode ser feita pela validação do mapa classificado com a referência de campo ou áreas foto-intepretadas, (POWELL *et al.*, 2004; SCHULTZ *et al.*, 2015a). Como já comentado anteriormente, o resultado obtido pelo processo de validação dos dados classificados com os dados de referência gera a matriz de confusão (CONGALTON e GREEN, 2009; SCHULTZ *et al.*, 2015b).

Outra forma pode ser realizada, avaliando o resultado de área do mapa classificado, em nível de município, comparado ao dado de área plantada pelas estimativas agrícolas oficiais, também ao nível de município (RIZZI e RUDORFF, 2005; HOWARD *et al.*, 2012; GUSSO *et al.*, 2012; JOHAN *et al.*, 2012). O resultado desta avaliação gera um modelo de regressão linear simples e o seu coeficiente de correlação linear (R^2) (Figura 2.15).

Figura 2.15 – Comparação dos valores de área modelo (classificação automática) e observado (censo agrícola americano) por municípios e tipo de cultura (milho, soja e trigo) em 358 municípios do estado de Nebraska-EUA.



Trigo, $R^2 = 0,92$; soja, $R^2 = 0,90$; e milho, $R^2 = 0,90$.

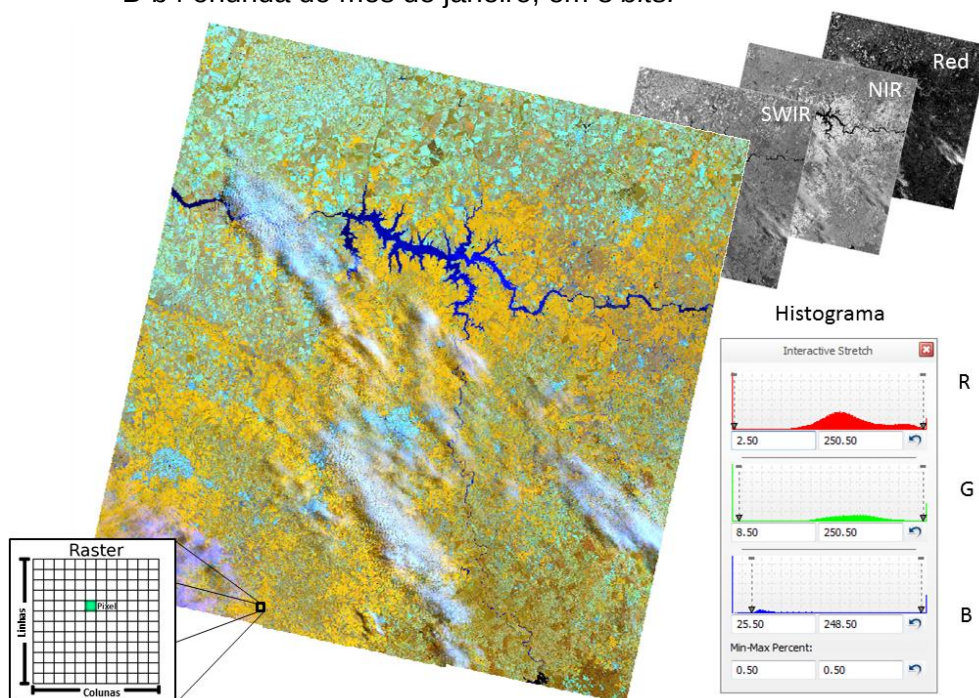
Fonte: Howard *et al.*, (2012).

Na Figura 2.15, existiu uma alta correlação entre os dados classificados no mapa ($R^2 > 0,90$) e os dados oficiais dados pelo censo. Johan *et al.* (2012) obtiveram R^2 superior à 84 % no mapeamento de culturas de verão (soja e milho), nas condições agrícolas do estado do Paraná. Gusso *et al.* (2012) tiveram melhores resultados ($R^2 = 0,91$) que Johan *et al.* (2012), para o mapeamento de soja, no Rio Grande do Sul. Wardlow e Egbert (2008) encontraram ($R^2 = 0,76$), ao considerar todas as culturas de verão (milho, sorgo e soja) nas condições agrícolas dos EUA. Dados de sensoriamento remoto utilizados na pesquisa

2.4.5. Sensor Operator Land Imager (OLI)

O sensor OLI do satélite Landsat-8 vem operando em órbita desde o dia 11 de fevereiro de 2013. As primeiras imagens deste sensor formadas sobre o estado de SP foram datadas no mês de Março/2013. A Figura 2.16 ilustra o aspecto de uma imagem OLI em ambiente SIG em composição falsa cor R-b5 G-b6 B-b4.

Figura 2.16 – Histogramas de uma imagem OLI, em composição colorida R-b5 G-b6 e B-b4 oriunda do mês de janeiro, em 8 *bits*.



Fonte: Adaptado de Florenzano, (2011).

As imagens OLI estão vinculadas ao sistema de referência geográfica *Worldwide Reference System-2* (NASA, 2009), sendo que esses dados estão disponíveis nos websites <http://glovis.usgs.gov/> e <http://earthexplores.usgs.gov/>.

2.4.5.1. Resolução radiométrica e espectral do OLI

As primeiras imagens do OLI fornecidas pela *United State Geological Survey* (USGS) foram oferecidas ao usuário em níveis de cinza (16 *bits*), e no início do ano de 2015, era possível fazer *download* das imagens, já em reflectância da superfície (também 16 *bits*).

A Tabela 2.3 mostra as principais bandas espectrais do sensor OLI utilizadas na presente pesquisa.

Tabela 2.3 – Características básicas do sensor OLI

Nome do Sensor	Operator Land Imager (OLI)
Satélite (númeror)	1
Passagem sobre o Equador (hora)	±11:00
Resolução Radiométrica (<i>Bits</i>)	16
Distância de amostragem do solo (m)	30
Bandas espectrais (nm)	
Aerosols (b1)	430-450
Azul (b2)	450-520
Verde (b3)	530-590
Vermelho (b4)	640-670
IVP (b5)	850-880
IVC (b6 e b7)	1570-1610
PAN (b8)	2110-2290
Cirrus (b9)	500-680
	1360-1380

Além das bandas espectrais mostradas na Tabela 2.3, a USGS fornece também, para cada uma das imagens OLI, um dado chamado BQA (*Quality Assessment Band*). O dado BQA informa a qualidade que cada pixel da imagem possui, em função das interferências no valor radiométrico captado, oriundas da existência de nuvens, sombra, neve e terra sobre pixel (USGS, 2015).

Ao utilizar b1, b2, b9 e o produto BQA, a USGS gera também um dado chamado de *cf mask* (*cloud function of mask*). Este dado é uma máscara em forma matricial, que informa a posição dos pixels recobertos com nuvens e sombra no momento da passagem do sensor; e também informa quais pixels vizinhos, aos pixels mascarados, que tiveram seus valores radiométricos alterados devido a presença de sombra das nuvens (USGS, 2015).

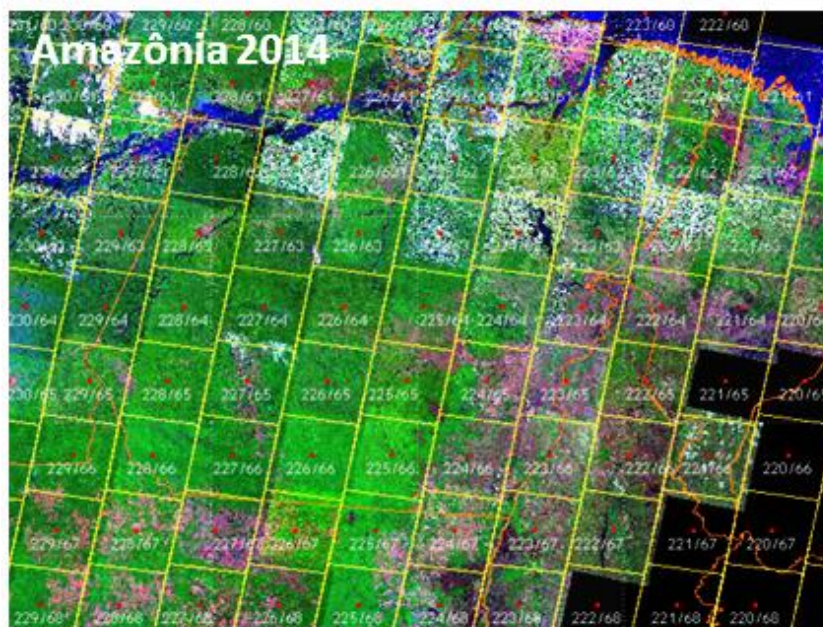
2.4.5.2. Resolução espacial e temporal do OLI

O sensor OLI adquire suas imagens em resolução espacial de 30 m e com resolução temporal de 16 dias.

Normalmente, o tamanho aproximado da cena OLI é de 170 km ao norte/sul por 183 km a leste/oeste. Desta forma, uma imagem Landsat-8 equivale a 3.111 ha e possui mais de 3.456.667 pixels de 30 metros. Por exemplo, são necessárias 18 cenas Landsat, em cinco diferentes órbitas, para recobrir o

estado de São Paulo. A Figura 2.17 ilustra um mosaico de imagens *Landsat-like* que recobre parte nordeste do território nacional.

Figura 2.17 – Mosaicos de imagens *Landsat-like* montados sobre a região que compreende a florestal amazônica brasileira, para o ano de 2014. Cada retângulo em amarelo representa uma órbita/ponto Landsat.



Fonte: adaptado de Projeto PRODES (2015).

Este mosaico apresentado na Figura 2.17 foi utilizado na presente pesquisa apenas para mostrar como várias imagens *Landsat-like* se comportam lado-a-lado, para formar um mosaico de diferentes imagens de resoluções temporal, radiométrica, espectral e espacial, semelhantes.

2.4.6. Sensor *Multispectral Imager* (MSI)

O sensor MSI está a bordo dos cinco satélites *RapidEye* (NAUGHTON *et al.*, 2011). Estes satélites foram lançados em órbita heliossíncrona no dia 29/08/2008. As principais características deste sensor estão descritas na Tabela 2.4.

Tabela 2.4 – Características básicas do sensor MSI

Nome do Sensor	Multispectral Imager (MSI)
Satélite (número)	5
Passagem sobre o Equador (hora)	±11:00
Resolução Radiométrica (<i>Bits</i>)	12
Distância de amostragem do solo (m)	5
Bandas espectrais (nm)	
Azul (B2)	440-510
Verde (B3)	520-590
Vermelho (B4)	630-685
Red-Edge (B5)	690-730

Os dados deste sensor são fornecidos por encomenda. O Ministério do Meio Ambiente (MMA) fez uma compra dos dados MSI para o ano de 2011, que abrangeu um recobrimento sistemático de todo o território brasileiro. Os dados podem ser baixados no catálogo do INPE, pelo *link* (<http://cbers2.dpi.inpe.br/>).

2.4.7. Linhas de base ou *baselines*

Uma linha de base ou *baseline* consiste no melhor mapa de uso e ocupação do solo elaborado e possível para área de estudo avaliada (WALDNER *et al.*, 2015; MATON *et al.*, 2015). Estas linhas de base podem ser utilizadas como dados auxiliares para classificação automática ou validação, como comentado por Waldner *et al.* (2015) e Maton *et al.*, (2015).

Para o Brasil, é possível encontrar diferentes *baselines*, calculadas em nível global (1-2) ou nacional (3-6), relacionadas ao mapeamento temático do País (Tabela 2.5). Estes dados são importantes fontes de informação secundária sobre o uso do pixel em determinado momento histórico.

O *GLCM* e *GFC* são os dados com resolução espacial mais fina entre os mapas temáticos produzidos em nível global de informação (HANSEN *et al.*, 2013; CHEN *et al.*, 2015). Estes dados já foram utilizados por Waldner *et al.*, (2015) e Maton *et al.*, (2015) como máscaras de plantio, para mascaramento e validação de áreas agrícolas na Argentina, França, Ucrânia, África do Sul e demais países.

Tabela 2.5 – Respectivas fontes de dados para construir as linhas para o cenário agrícola brasileiro. Estes dados sevem de informação secundária no momento da classificação visual do pixel na imagem *Landsat-like*.

Linhas de Base	Resolução espacial (m)	Link Web	Referência
1- GLCM	30	www.globallandcover.com	Chen <i>et al.</i> , (2015)
2- GFC	30	earthenginepartners.appspot.com	Hansen <i>et al.</i> , (2013)
3 -Canasat	30	http://www.dsr.inpe.br/laf/canasat/	Rudorff <i>et al.</i> , (2010)
4 - Cafesat	30	http://www.dsr.inpe.br/laf/cafesat/	Moreira <i>et al.</i> , (2010)
5 - Pivô central	30	mapas.cnpm.embrapa.br/	Guimarães e Landau, (2011)
6 - IVs	250	www.dsr.inpe.br/laf/series/	Adami <i>et al.</i> , (2012)

Onde: *Global Land Cover Map* (GLCM), *GFC (Global Forest Change)* e Série temporal dos índices de vegetação (IVs).

Em nível nacional, os dados vetoriais do Canasat, Citros e de pivô central, delineados manualmente para a safra 2012/2013, informam a distribuição espacial dos plantios de cana-de-açúcar no Centro-Sul do Brasil (LANDAU *et al.* 2013) e a distribuição espacial dos talhões de citros, para o estado de São Paulo. A distribuição espacial destes vetores está atrelada à classificação visual gerada sobre imagens *Landast-like* de 2012.

Por fim, as séries temporais dos produtos IVs do sensor MODIS podem ser usadas como subsídio na separação entre áreas com culturas perenes, semi-perenes, culturas anuais, culturas de inverno e pastagem (RUDORFF *et al.*, 2009). Neste caso, as curvas temporais dos IVs apresentam diferentes comportamentos/padrões no ano safra, isto é, de acordo com a fase das curvas e o período do ano, é possível constatar diferentes fases fenológicos das culturas supra referidas e mudanças do uso e ocupação do solo (RUDORFF *et al.* 2009; ADAMI *et al.*, 2012; ADAMI *et al.*, 2012).

2.4.8. Pré-processamento de dados OLI

Como falado anteriormente, as imagens do sensor OLI são compostas por 3,5 milhões de pixels, e estes pixels são coletados em nove bandas espectrais. Desta forma, todos os pixels possuem um valor, em nível de cinza, quantizado em cada banda deste sensor.

Ainda nesta linha de raciocínio, as imagens OLI compreendem 3,1 mil hectares, isto é, são necessárias 18 imagens *Landsat-like* para recobrir o estado de São Paulo, nas órbitas (total de 6) e pontos (total de 4).

A diferença entre as datas das imagens (pontos) e posições em relação ao equador faz com que os valores radiométricos de um mesmo dossel em diferentes imagens sejam variáveis, principalmente pela variação sazonal da vegetação e também pela geometria de iluminação (GALVÃO *et al.*, 2011; MOURA *et al.*, 2012).

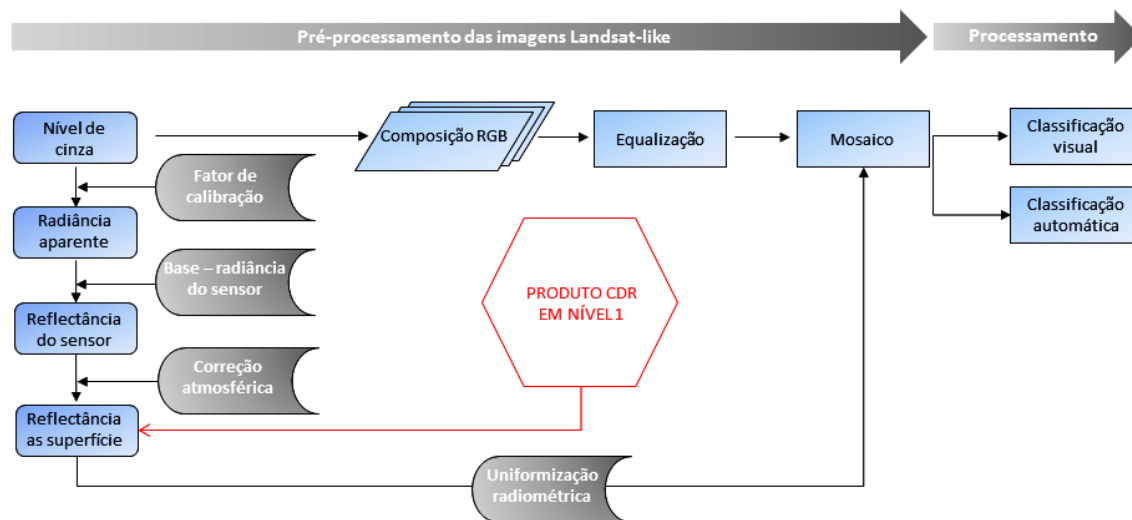
Estes problemas salientados no parágrafo supra referido, resultam na diferença radiométrica entre imagens adjacentes (YANG e LO, 2000; EIVAZI *et al.*, 2015), ou até mesmo nos valores radiométricos de um mesmo alvo localizado na sobreposição entre duas cenas (BAO *et al.*, 2012; KNORN *et al.*, 2009;).

Estas variações entre os valores radiométricos das imagens podem comprometer futuramente processos de classificação automática (problemas no reconhecimento das texturas das imagens) (CIHLAR, 2000; KNORN *et al.*, 2009) e na classificação visual de alvos em imagens multitemporais (diferentes cores para uma mesma cultura agrícola) (EBERHARDT, 2014).

Desta forma, alguns procedimentos operacionais são exigidos anteriormente ao processamento destes dados. Estes procedimentos são chamados de pré-processamento dos dados.

Basicamente, os passos descritos para o pré-processamento das imagens de sensoriamento, em grandes áreas (mais de uma imagem) é feito de duas principais formas (Figura 2.18).

Figura 2.18 – Passos utilizados para o pré-processamento e processamento de imagens Landsat-like.



Em destaque (vermelho) é ilustrado o produto dos dados CDR disponibilizados pela USGS em nível 1 de pré-processamento digital de imagens.

Fonte: adaptado de Florenzano (2011).

Percebe-se na Figura 2.18 que o produto CDR em nível 1 de pré-processamento é disponibilizado em reflectância da superfície, já corrigido geometricamente e radiometricamente (USGS, 2015). Demais procedimentos operacionais como uniformização radiométrica e equalização devem ser aplicados anteriormente ao processamento destes dados (SPRING, 2015).

Assim, dois conceitos sobre pré-processamento de imagens devem ser apresentados nesta presente pesquisa: equalização/realce e uniformização/normalização.

2.4.8.1. Equalização ou realce

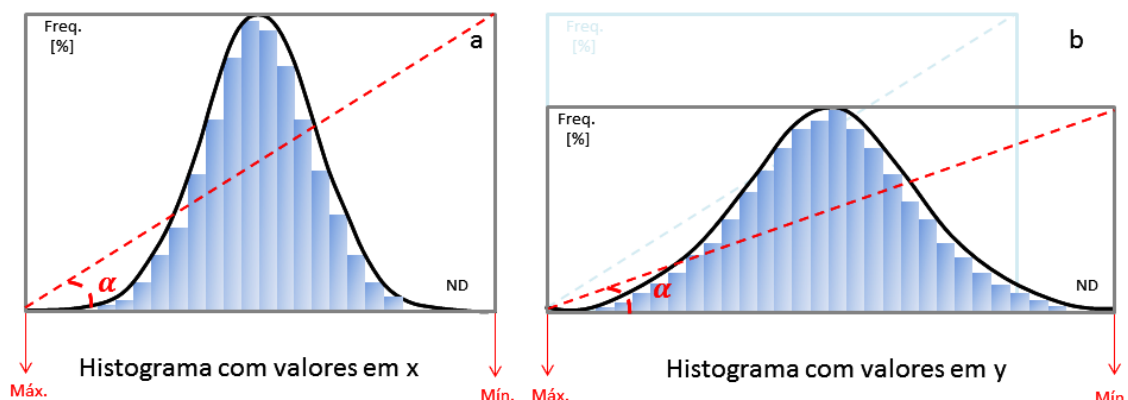
A equalização ou realce é uma técnica utilizada para aplicar às imagens adjacentes, contrastes semelhantes (GONZALES, 2007; TING *et al.*, 2015). A equalização tem por objetivo aproximar visualmente imagens com diferentes radiometrias e esse processo é realizado através da manipulação das curvas do histograma de cada imagem (TING *et al.*, 2015).

Normalmente, o contraste das imagens é ajustado pelos valores de máximo e mínimo dos canais RGB (EBERHARDT, 2014). Neste método, dependendo do grau de nitidez procurada pelo usuário nas imagens, os valores do contador digital, nos histogramas, são alargados ou achatados.

Isto ocorre pela conversão automática, pixel-a-pixel, dos contadores digitais por uma função linear ($Y = a \cdot x + b$), descrita pelo coeficiente de inclinação da reta e os valores de *máx.* e *mín.* (em vermelho) (Figura 2.19).

Na Figura 2.19, de a para b ocorreu um alargamento dos histogramas, alterando-se a amplitude entre os valores de máximo (*Máx.*) e mínimo (*Mín.*) e do ângulo de inclinação da reta.

Figura 2.19 – Histogramas hipotéticos da frequência (Freq.) dos níveis digitais (ND) de uma imagem de sensoriamento remoto que mostram os limites de máximo (*máx.*) e mínimo (*mín.*) dos valores de ND. Em (a) o histograma possui baixa amplitude entre os valores de *máx.* e *mín.*, e em (b) a amplitude é maior.



A reta de inclinação α sofre alteração em decorrência à variação da amplitude dos valores de *máx.* e *mín.*

Fonte: adaptado de Gonzales, (2007).

A função linear descrita logo acima é expressa por quatro parâmetros: y , x , a e b . Em que, y refere-se ao novo valor de contador digital, x refere-se ao antigo valor do contador digital, a é a tangente do ângulo α e b é fator de incremento (valor *mín.* – valor *máx.*) (SPRING, 1996). Este procedimento é iterativo e também subjetivo, pois depende do resultado inicial do contraste aplicado em

cada imagem pelo programa usado e também da acuidade visual de cada usuário do dado.

2.4.8.2. Normalização ou uniformização

A normalização radiométrica relativa é utilizada para relacionar linearmente duas imagens LL, adjacentes, de diferentes datas (CIHLAR, 2000; LONG et al. 2011). Para isto, uma regressão linear é construída por meio dos valores radiométricos de alvos pseudo-invariantes (APIs) entre as bandas das imagens adjacentes. Os APIs representam áreas temporárias e com reflectância constante (CANTY e NIELSEN, 2008; EIVAZI et al., 2015).

Após selecionar os APIs, o passo final da normalização radiométrica é realizado, aplicando-se a reta de regressão sobre umas das cenas utilizadas no processo. Este passo é realizado para todas as bandas que foram selecionadas no processamento dos dados (CIHLAR, 2000; LONG et al., 2011).

Um considerável número de trabalhos científicos foi desenvolvido, propondo seleção automática de APIs (CANTY e NIELSEN, 2008; EIVAZI et al. 2015). No entanto, para não se exaurir neste tópico, no capítulo 8, página 130, do livro de Achard e Hansen (2013) os autores consideram pixels de floresta nativa, como sendo os alvos menos variantes de uma cena de sensoriamento remoto, isto é, são alvos estáveis e que sofrem menos variação no tempo que demais alvos presentes.

Os dados do *Global Forest Cover* (GFC) (HANSEN et al., 2013) são a fonte mais recente para obtenção de pixels de floresta nativa em nível global. Porém, é importante relatar que os valores radiométricos dos alvos de floresta nativa podem variar no tempo, principalmente em função do tipo de vegetação arbórea (ACHARD e HANSEN, 2013) e do efeito direcional do ângulo solar e da topografia do local onde a floresta se encontra (VALERIANO, 2011). Assim, ao utilizar estes dados como API's, filtros devem ser aplicados anteriormente à sua seleção.

2.4.8.3. Aparência Equalizada (ApEq)

Um dos objetivos desta Tese foi desenvolver um método chamado Aparência Equalizada (ApEq), que fornece dados equalizados de imagens Landsat-like adjacentes, apenas utilizando um fator multiplicativo, calculado por meio dos valores de reflectância entre APIs de uma imagem base e APIs das imagens a serem normalizadas (imagens de conversão).

O fator multiplicativo é aplicado nos valores de máximo e de mínimo dos histogramas da imagem base. Os novos valores de máximo e mínimo obtidos para a imagem base são então utilizados para ajustar os histogramas das imagens de conversão.

Tal abordagem utiliza a técnica de uniformização radiométrica relativa e equalização de histogramas. A diferença é que aqui, nesta nova abordagem, a equalização é realizada de forma operacional e objetiva.

2.4.9. Processamento de dados OLI – classificação automática

2.4.9.1. Segmentadores de imagens de sensoriamento remoto

Na literatura, vários tipos de segmentadores de imagens de sensoriamento remoto e os mais importantes são baseados em propriedades similares ou descontínuas entre os pixels e seus vizinhos (BINS *et al.*, 1996; ESPÍNDOLA, 2006). A técnica de crescimento de regiões (*grow region - based*) utiliza a similaridade entre os pixels e a técnica detecção de bordas (*boundary - based*) utiliza a descontinuidade entre eles (MUÑOZ *et al.*, 2003).

Em cenas de sensoriamento remoto de paisagens naturais, os algoritmos baseados em crescimento de regiões produzem segmentos mais condizentes com a realidade de campo (MUÑOZ *et al.*, 2003). O algoritmo Crescimento de regiões de Bins (CRB) e Segmentador Mutli-resolução (SM-R), são exemplos de algoritmos de crescimento de regiões, e são os mais utilizados para segmentação de imagens no Brasil (PROJETO PRODES, 2015; SCHULTZ *et al.*, 2015a).

O CRB foi implementado em programas conhecidos como SPRING e GEODMA (BINS *et al.*, 1996; KORTING *et al.*, 2008) e SM-R de Battz e Schäpe (2000) foi implementado nos programas eCognition e GEODMA (KORTING *et al.*, 2008; TRIMBLE GEOSPATIAL, 2015).

No processo de segmentação CRB, o algoritmo utiliza os valores dos níveis de cinza de uma célula/pixel, chamado semente, e analisa a similaridade (S_m) entre ele e seus vizinhos (BINS *et al.*, 1996). Se o valor de similaridade entre eles se satisfazer, eles são agrupados em uma única célula. O crescimento da região é interrompido, devido a um valor de área (A_r) especificado anteriormente pelo usuário (ESPÍNDOLA, 2006).

A desvantagem deste método é que o segmentador analisa apenas a semelhança entre o pixel semente e seus vizinhos e não possui habilidade em analisar o contexto em que o pixel está inserido. A falta da análise do contexto faz com que informações espaciais e de textura não sejam incorporadas ao processo de segmentação pelo CRB (OLIVEIRA *et al.*, 2003).

Neste caso, o algoritmo SM-R, por possuir parâmetros como forma, compacidade, textura, etc (BATTZ e SCHÄPE, 2001), consegue simular padrões de comportamento dos objetos presentes na cena. Estes resultados são obtidos devido à preocupação do algoritmo estar relacionada na avaliação do contexto em que os objetos estão inseridos e não apenas na proximidade dos pixels que formam um único objeto (MYINT *et al.*, 2009).

Este algoritmo primeiramente agrupa os pixels pela similaridade entre eles e depois analisa as bordas dos segmentos por meio da descontinuidade dos pixels (BENZ *et al.*, 2004). Assim, o algoritmo irá agrupar regiões semelhantes, respeitando não apenas o valor radiométrico dos pixels, mas também o contexto (forma, compacidade, textura) em que eles estão inseridos (ESPÍNDOLA, 2006).

Para isto, é necessário que o usuário defina quatro parâmetros de segmentação: fator de escala (Fe), forma (Fm), compacidade (Cp) e peso das bandas(pb) (ESPÍNDOLA, 2006).

Conforme Baatz e Schäpe (2000) e Fernandes *et al.* (2012), Fe determina a heterogeneidade máxima permitida para os objetos da imagem; Fm consiste na união dos parâmetros Cp (razão da área do objeto e o raio da circunferência abrangente) e Sv (expressa o limite entre os objetos, sendo que, quanto mais fractal, maior o índice de forma); Cr é inversamente proporcional ao parâmetro Fm; por fim, Pb é atribuído pelo usuário em função da ponderação (entre zero e um) que deseja utilizar para cada banda utilizada.

Para o Fe, os valores variam de zero a um valor finito definido pelo usuário, enquanto que, para Fm, Cp, Sv e Cr, os valores variam de zero a um. É importante destacar que, na plataforma ED64 (TRIMBLE GEOESPATIAL, 2015), apenas é necessária a definição dos parâmetros Fe, Fm, Cp e Pb.

2.4.9.2. Algoritmos de classificação automática baseados em técnica de árvore de decisão

Como visto anteriormente, os modelos de conhecimento expressam padrões de comportamento nos dados (WITTEN e FRANK, 2005) depois de reconhecidos, estes padrões podem ser considerados para a classificação ou tomada de decisão em dados diferentes daqueles usados na tarefa de aprendizado.

O objetivo das técnicas aplicadas por Stefanski *et al.* (2013) e Schultz *et al.* (2015a), por exemplo, descrito na seção 2.4.4.1.4. (pg. 46 a 49), foi formalizar um conhecimento real sobre certo fenômeno e extrapola-lo para uma área de estudo.

Witten *et al.* (2011) relataram que, para formar modelos com padrões de comportamento replicáveis, estes métodos fundamentam-se em pressupostos estatísticos.

Os modelos de conhecimento mais encontrados na literatura que utilizam os pressupostos estatísticos são os aprendizados de máquina (*learn machine*), utilizadas no software *Waikato Environment Knowledge Analysis* (WEKA) (HALL et al., 2009).

Entre eles, os mais citados na literatura nacional e internacional são árvores de decisão (AD), em específico o algoritmo de mineração de dados J48 (C4.5) (Vieira et al., 2012), e o algoritmo *Random Forest* (RF) (LONG et al., 2013; IMMITZER et al., 2012; SCHULTZ et al., 2015a; dados não publicados). Estes dois algoritmos são supervisionados e baseados no método estatístico ensemble (BREIMAN, 2001; WITTEN e FRANK, 2005).

O método *ensemble* é utilizado para dividir uma amostra de dados de treinamento e, a partir dela, construir conjuntos de classificadores, combinando seus votos para classificar um novo exemplo (OSHIRO, 2013). O voto aqui corresponde à classe fornecida por cada um dos classificadores, para esse novo exemplo. Os métodos ensemble amplamente conhecidos são *bagging* (BREIMAN, 1996) e RF (BREIMAN, 2001).

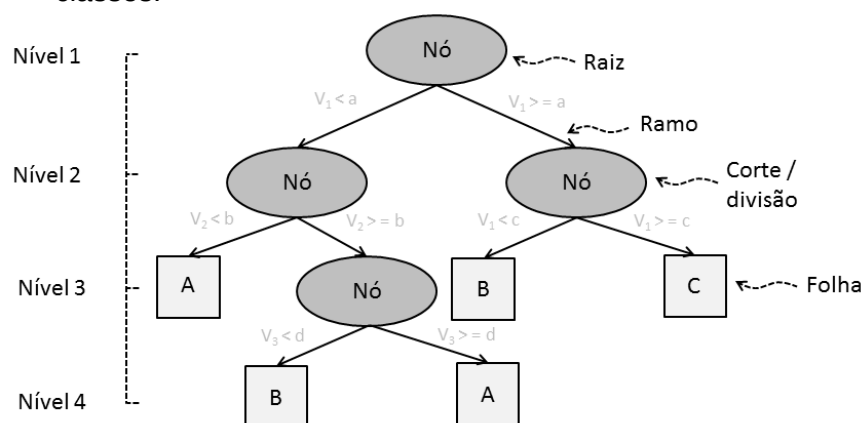
A partir de uma amostra inicial, a técnica *bagging* utiliza a amostragem aleatória com reposição (*bootstrap*) para formar diferentes subconjuntos de dados. Cada subconjunto é formado com o mesmo número de elementos do conjunto original. Os elementos que não foram selecionados para treinamento, são chamados de conjunto *out-of-bag* e utilizados para validação do modelo (BREIMAN, 2001).

O uso de dados de treinamento e validação, gerados pelas técnicas *bootstrap* e *bagging*, fornecem estimativas do erro de generalização dos classificadores (STEFANSKI et al., 2013). Termo conhecido também, como validação do modelo de conhecimento.

A árvore de decisão (AD) J48 utilizada técnica *bagging* sob um conjunto de dados de treinamento, e tem sido amplamente utilizada para mineração de dados *Landsat-like* e validada por diversos trabalhos objetivados a

classificação de diferentes usos e ocupação da terra (VIEIRA *et al.*, 2012). A AD J48 é construída em forma de árvore invertida, isto é, de cima para baixo encontram-se o nó raiz, os galho, demais nós e folhas da árvore (Figura 2.20).

Figura 2.20 – Uma árvore de decisão hipotética utilizada para classificação de A, B e C. Onde V são os diferentes atributos utilizados para separação das classes.



Fonte: = Adaptado de Chan e Paelinckx, (2008).

Cada um dos ramos de uma AD contém um teste sobre um ou mais atributos utilizados no treinamento do modelo (variáveis independentes). O teste é subdividido em três possibilidades (T_i), (1) todos os casos avaliados são identificados numa classe (c_j), (2) não contém casos na classe avaliada e (3) procura-se outra informação (atributo v_j), que no caso pertença a uma determinada classe (termo: divide para conquistar).

Desta forma, um nó de decisão identifica o teste para v , forma um o galho com um valor assumido para v e folhas que se referem ao número de instâncias/segmentos classificadas corretamente e erroneamente com o uso de v . Este procedimento ocorre sucessivamente, até que todos os segmentos sejam testados pelas três possibilidades (T_i), e assim pertençam a uma folha da árvore.

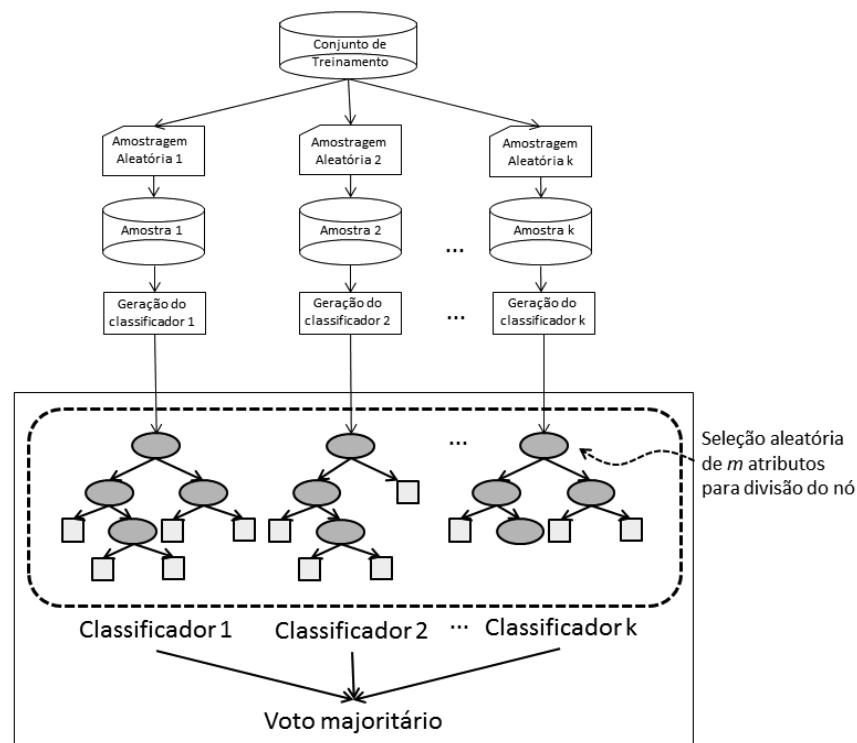
Segundo Breiman (1996) *bagging* é muito eficaz quando os classificadores utilizados não possuem um comportamento instável. Um classificador é conhecido como instável quando pequenas mudanças no conjunto de treinamento podem causar grandes mudanças na classificação gerada. Neste

caso, um único classificador instável (por exemplo, uma única AD) não é capaz de reproduzir o conhecimento gerado, uma vez que um classificador composto pode ter maior chance de acerto (WITTEN e FRANK, 2005).

Para tentar contornar este problema, outras técnicas de formação do modelo de conhecimento podem ser utilizadas, e uma delas é a técnica Random Forest (RF). O algoritmo RF é baseado em ADs, gera uma floresta com enésimas árvores e informa o grau de incerteza obtido em cada uma delas (BREIMAN, 2001).

Diferentemente da técnica *bagging*, na técnica RF, a cada nó formado na AD (Figura 2.21), um conjunto de m atributos são selecionados aleatoriamente e validados. O melhor atributo é, então, escolhido para dividir o nó (Figura 2.21). O valor de m é fixado para todos os nós subsequentes.

Figura 2.21 – Funcionamento do método Random Forest. Os classificadores são representados por hipotéticas árvores de decisão.



Fonte: Oshiro (2013)

O uso do *bagging* e seleção aleatória de atributos fazem com que as árvores sejam diferentes entre si e, portanto, diminui-se a correlação entre elas (WITTEN e FRANK, 2005). A baixa correlação entre as árvores da floresta tende a diminuir o erro de classificação (OSHIRO, 2013). Uma árvore com baixa taxa de erro é uma árvore forte (modelo de conhecimento replicável) e à medida que as árvores tornam-se fortes, reduz-se a taxa de erro da floresta (BREIMAN, 2001). Schultz *et al.* (2015a) compararam o resultado de 1025 modelos RF gerados de um único conjunto de amostras de treinamento. Estes autores obtiveram EG superior a 80 % na classificação de cana-de-açúcar, soja, amendoim, mandioca e outros na Mesorregião de Assis. Estes autores relataram também, que o modelo de maior valor EG foi utilizado para geração do mapa temático final com uma exatidão global de 78 %.

Desta forma, o modelo de conhecimento RF foi considerado o mais apto para ser utilizado na classificação automática de cana-de-açúcar, milho e soja no estado de São Paulo, nesta presente pesquisa.

3. MATERIAL E MÉTODOS

Para que se fossem alcançados os nove objetivos apresentados na presente pesquisa, foi necessário que mais de uma área de estudo fosse analisada. Assim, a pesquisa foi executada em nível estadual (área de estudo 1), nível municipal (área de estudo 2) e nível regional, isto é 15 municípios contínuos (área de estudo 3).

- Na área de estudo 1 foram respondidos os seguintes objetivos:

1. definir a melhor forma de estratificação e de alocação de pontos amostrais no estado de São Paulo, para reduzir o coeficiente de variação das culturas da cana-de-açúcar, milho e soja;
2. estabelecer o procedimento de obtenção de um mosaico bimensal de imagens sobre o estado de São Paulo, que permita a interpretação visual dos pontos amostrais.

Nestes dois primeiros objetivos as atividades de: (a) elaboração do painel amostral e obtenção das estatísticas agrícolas para as culturas da cana-de-açúcar, soja e milho, em quatro períodos do ano safra 2014/2015 (P1 - nov/dez, P2 - jan/fev, P3 - mar/abr e P4 - mai/jun), (b) aplicação da equalização/realce de todas as cenas multitemporais OLI e (c) classificação visual de pixels 5.000 pixels por períodos, em imagens multitemporais OLI, em ApEp; foram desenvolvidas.

- Na área de estudo 2 foram respondidos os seguintes objetivos:

3. desenvolver um protocolo de classificação, que englobe o ano safra 2014/2015, para a interpretação dos pixels amostrais nos mosaicos estaduais, aplicá-lo em quatro períodos bimensais (dezembro/janeiro, fevereiro/março, abril/maio, junho/julho) e gerar as estimativas correspondentes;
4. avaliar o desempenho do protocolo de classificação desenvolvido para as imagens de sensoriamento remoto tendo como referência dados validados em campo.

5. definir o método de segmentação mais adequado à situação agrícola da área de estudo utilizada para validação;
6. definir o método de classificação automática que tenha melhor desempenho para classificação dos pixels validados em campo;
7. avaliar os resultados da classificação automática com o uso dos dados validados em campo;

Nestes cinco objetivos subsequentes, as atividades de: (a) validação de campo de 200 pixels classificados em laboratório, (b) elaboração de um protocolo de classificação visual em laboratório de pixels com cana-de-açúcar, soja, milho e outros, nos quatro períodos da estimativa agrícola 2014/2015 e (c) classificação visual dos 200 pixels validados em campo por diferentes intérpretes utilizando para isto, o protocolo de classificação visual.

- Na área de estudo 3 foram respondidos os seguintes objetivos:

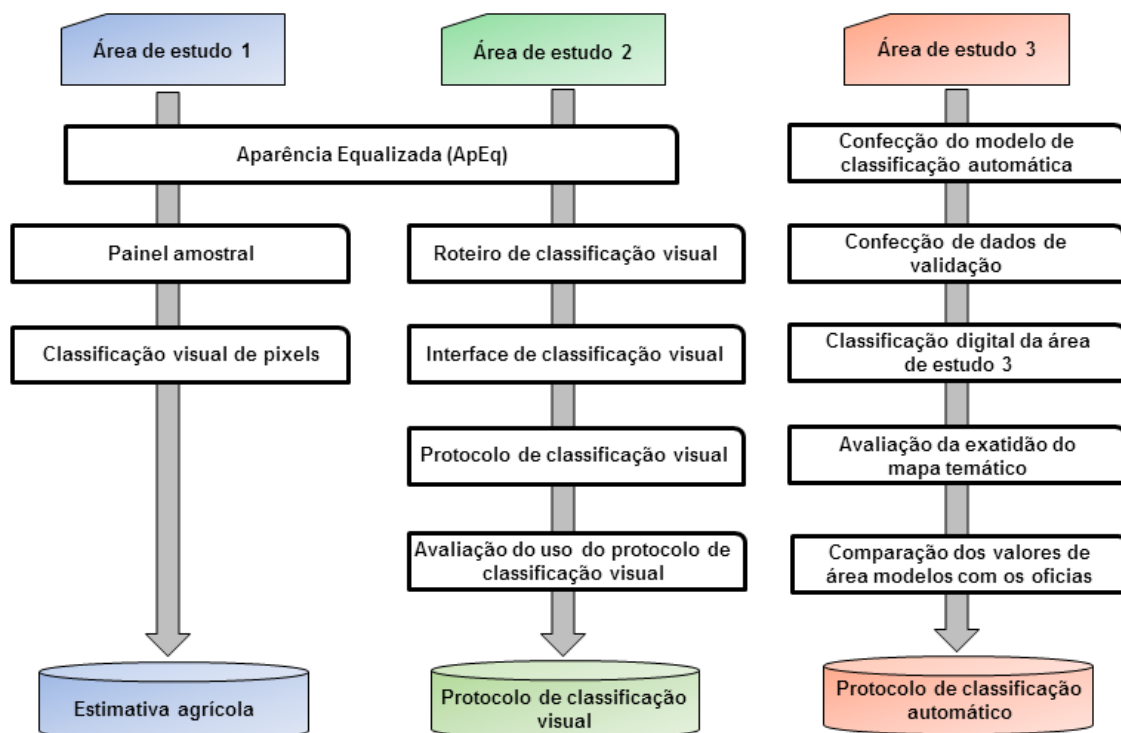
8. comparar os resultados da classificação visual com os dados da classificação automática;
9. comparar os dados de área ocupada com cana-de-açúcar e milho em 15 municípios localizados no extremo Sul do estado de São Paulo, estimados pela metodologia automática proposta, com os dados oficiais fornecidos pela Produção Agrícola Municipal (PAM) do IBGE.

E por fim, no que se refere à parte de classificação automática do trabalho, nos objetivos numerados de 8 a 9 foram desenvolvidas as atividades de: (a) classificação digital de cenas multitemporais OLI, localizada ao extremo Sul do estado de São Paulo (órbita/ponto 222/076), (b) treinamento do algoritmo *Random Forest* para classificar segmentos obtidos pela segmentação do SM-R da cena 222/076 com os 200 pixels validados em campo, (c) classificação visual de 200 pixels aleatórios, que foram utilizados posteriormente no processo de validação do mapa temático produzido pela classificação digital e

(d) comparação dos valores de área para as culturas da cana-de-açúcar e milho informados pelo IBGE e modelos pela classificação automática.

O fluxograma da Figura 3.1 mostra as atividades desenvolvidas nas três áreas de estudo da presente pesquisa.

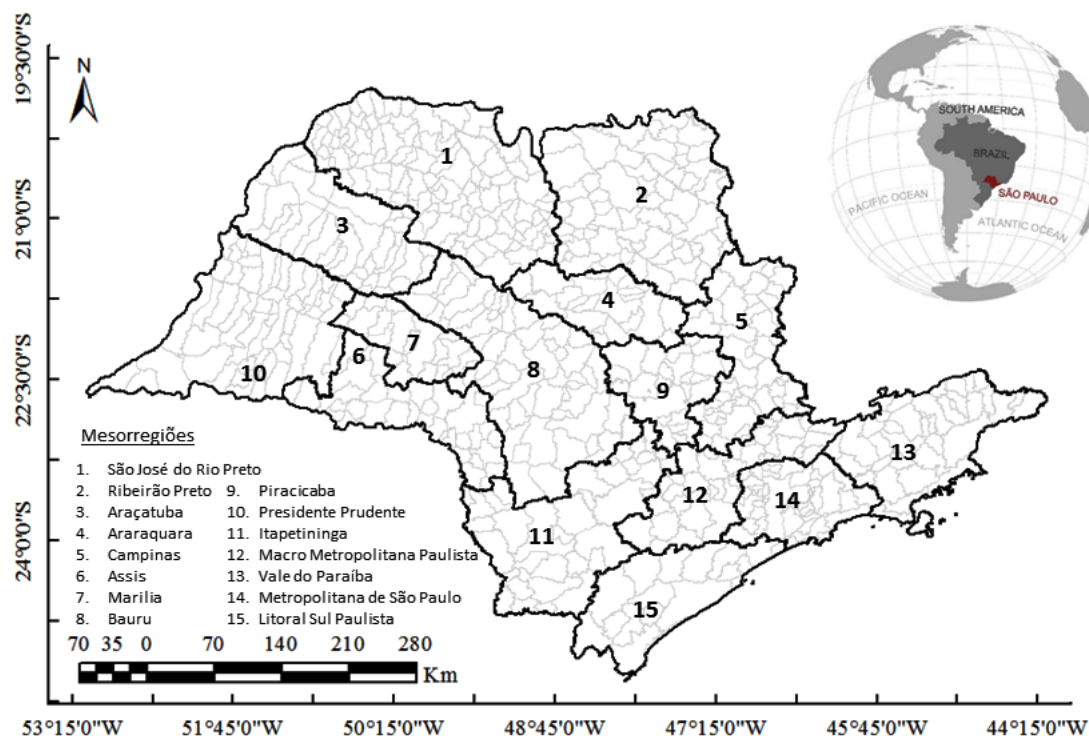
Figura 3.1 – Fluxograma das atividades desenvolvidas na presente pesquisa.



3.1. Área de estudo 1 – Estado de São Paulo

O estado de São Paulo possui 24,8 milhões de ha e está localizado entre os meridianos 44° 15' e 53° 15' e paralelos 19° 30' e 26° 30'. Além disso, o estado é subdividido em 15 diferentes mesorregiões (Figura 3.2).

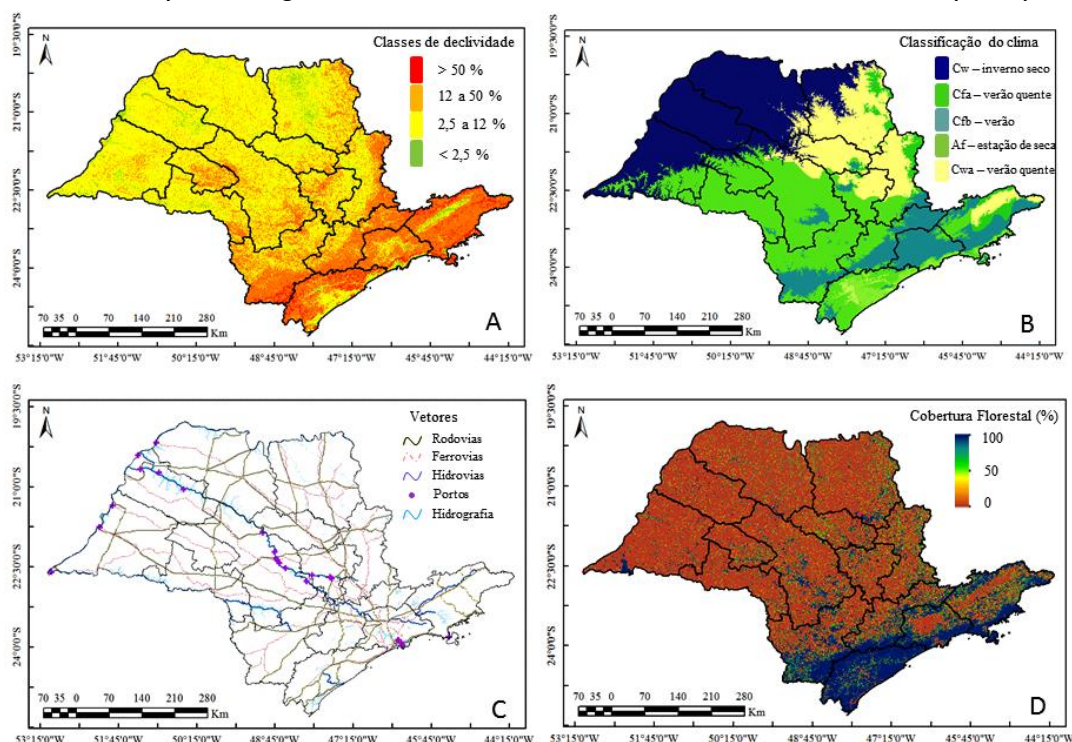
Figura 3.2 – Região de estudo, que compreende o estado de São Paulo e suas 15 mesorregiões.



O relevo do estado tem formações regulares e segundo Marques (1971) a maior proporção do estado possui declividade entre 2,5 a 12 % (Figura 3.3.a). Estes limites de declividade são toleráveis para o desenvolvimento da prática agrícola mecanizável no estado.

Segundo Alvarez *et al.*, (2013) o clima do estado é composto basicamente por cinco diferentes tipos de classificações (Figura 3.3.b). No oeste do estado ocorre predominância de períodos com inverno seco, mais ao sul é possível encontrar verões quentes e úmidos e o noroeste do estado é recoberto por verões quentes e mais secos. Os diferentes tipos de clima influenciam no zoneamento agrícola das culturas, principalmente na data de plantio da soja e do milho da safra e safrinha (MAPA, 2015).

Figura 3.3 – Algumas características relacionadas ao território do estado de São Paulo. (a) Classes de declividade conforme Marques (1971), (b) classes climáticas conforme Alvares *et al.* (2014), ramais logísticos e porcentagem de cobertura florestal conforme Hansen *et al.*, (2013).



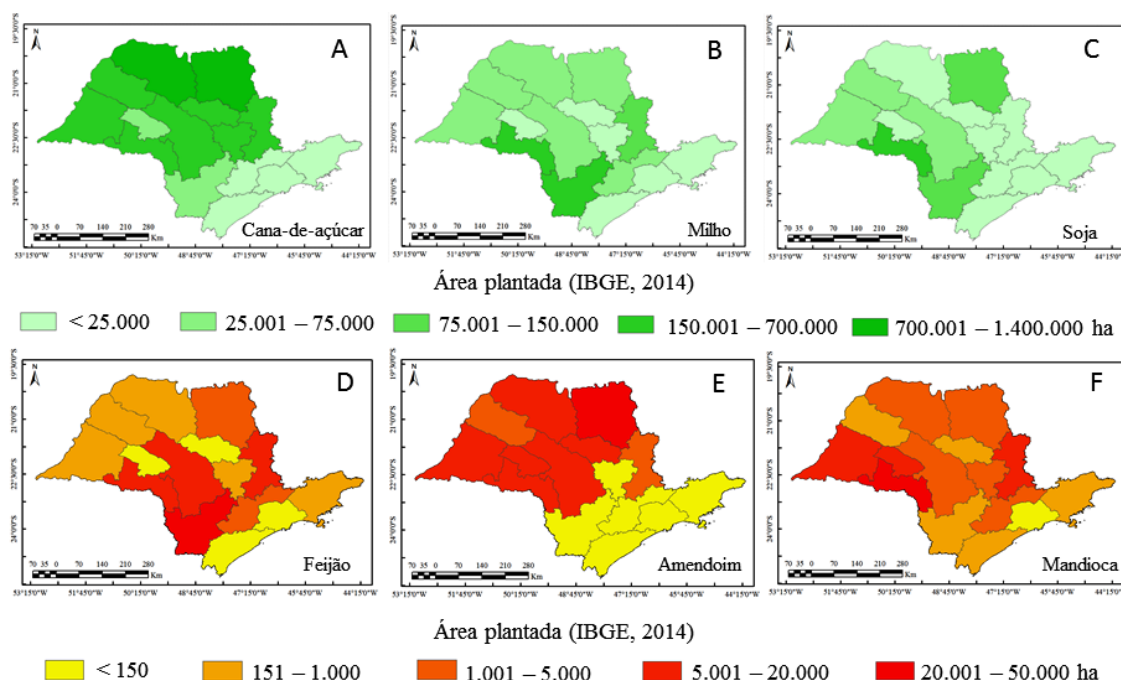
A malha viária do estado é composta por ferrovias, rodovias e hidrovias (Figura 3.3.c). São 4.098 km de ferrovias e 6.095 km de rodovias. A malha hidroviária do estado é composta de 60.130 km de hidrovias sendo que 56.229 km são considerados navegáveis com longo curso de viagem e apenas 2.736 km não são navegáveis. Além disso, o estado é abastecido de 28 portos, em que apenas seis são considerados terminal de uso privativo. Toda esta logística do estado permite o plantio e escoamento da produção de culturas agrícolas em todo território do estado de São Pulo.

A cobertura florestal do estado encontra-se concentrada na mesorregião Litoral Sul Paulista (HANSEN *et al.*, 2013) (Figura 3.3.d). As demais áreas são basicamente compostas por pastagens e áreas plantadas com culturas agrícolas, sejam elas temporárias, semi-perenes ou perenes (IBGE, 2015).

A aptidão agrícola do estado é baseada no plantio de cana-de-açúcar, milho, soja, feijão, amendoim e mandioca (Figura 3.4). A maior proporção de

mesorregiões do estado possui aptidão no plantio da cana-de-açúcar (Figura 3.4.a). As mesorregiões localizadas mais ao Sul e ao Norte, no plantio de milho e soja (Figura 3.4.b) e é importante relatar também a contribuição da mesorregião de Assis, no plantio de culturas de interesse secundário como o feijão, amendoim e mandioca.

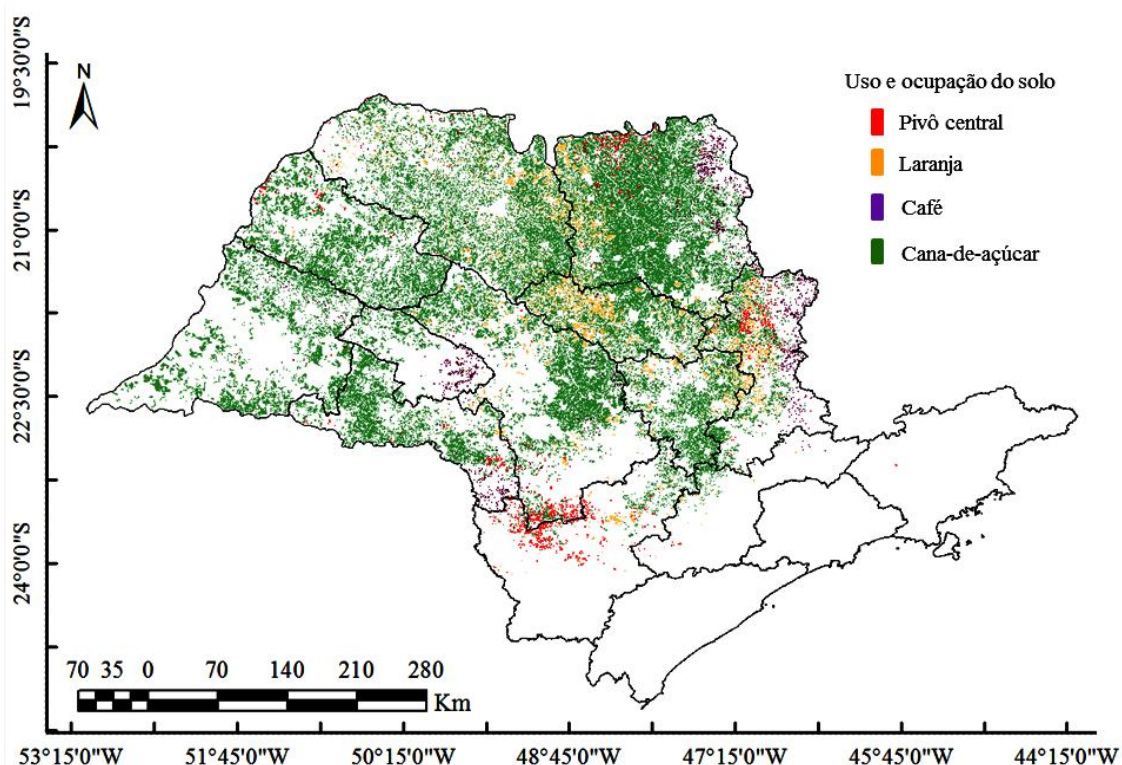
Figura 3.4 – Área plantada com (a) cana-de-açúcar, (b) milho, (c) soja, (d) feijão, (e) amendoim e (f) mandioca, no estado de São Paulo conforme dados da pesquisa agrícola municipal para o ano safra 2013/2014, agregados nas quinze mesorregiões do estado de São Paulo.



Ademais, alguns dados cadastrais confeccionados sobre imagens *Landsat-like*, como vetor de cana-de-açúcar do ano safra de 2012/2013 (CANASAT, 2014), de talhões de laranja de 2012 (FAVARO, 2013), de talhões de café de 2007 (CAFESAT, 2014) e áreas de pivô central de 2013 (LANDAU *et al.*, 2013) são apresentados na Figura 3.5.

Estes dados cadastrais, projetados em um mesmo *Datum* e projeção cartográfica, ilustram que ¼ do estado foi recoberto por um destes quatro importantes usos e ocupação do solo.

Figura 3.5 – Uso e ocupação do solo classificado em dados Landsat-like para o estado de São Paulo (CANASAT, 2014; CAFESAT, 2014; FAVARO, 2013; LANDAU *et al.*, 2013).



Depois de descritas algumas características agrícolas do estado de SP, o tópico logo abaixo descreve os materiais que foram usados para obter os resultados das estatísticas agrícolas, para o estado de SP, no ano safra 2014/2015.

3.1.1. Material

Para realização das pesquisas referentes aos "**objetivos 1 e 2**" foram necessários os seguintes insumos:

- Dados em formato matricial:

(1) Todas as imagens adquiridas no período de julho de 2014 a julho de 2015, pelo sensor OLI a bordo do satélite Landsat-8, nas 18 cenas necessárias para recobrir o estado de São Paulo;

(2) Todas as imagens de declividade do TOPODATA (<http://www.webmapit.com.br/inpe/topodata/>) que compõe o mosaico do estado de São Paulo;

(3) Quatro *tiles* de três produtos *Global Forest Change* (HANSEN et al., 2014) disponíveis no catálogo (<http://earthenginepartners.appspot.com/science-2013-global-forest>). Os produtos foram: cobertura percentual de árvore em 2000, perda de floresta constatada de 2000 a 2014 e incremento de floresta constatado de 200 a 2012.

- Dados em formato vetorial:

(1) Vetores de áreas disponíveis para colheita de cana-de-açúcar, do ano safra 2012/2013 (CANASAT, 2013);

(2) Vetores de áreas com Pivôs centrais em 2014 no estado de São Paulo. Disponível em <http://mapas.cnpm.embrapa.br/somabrasil/webgis.html> (Embrapa, 2014);

(3) Vetores de divisão político administrativo das mesorregiões do estado de São Paulo;

(4) Vetores das órbitas-pontos do sensor OLI, em sistema *Worldwide Reference System* (WRS-2).

(5) Dados cadastrais em formato de tabela, e vetor de área plantada com cana-de-açúcar, milho e soja para as 15 mesorregiões do estado de São Paulo, para o ano safra 2013/2014 (IBGE 2014).

- Demais insumos e capacidade humana:

- Série temporal do produto NDVI do MODIS disponível *on-line*, no *link*, <https://www.dsr.inpe.br/laf/series/>.

- Um computador Intel® Core™ i2-2400 CPU @ 3.10 GHz com 8 GB de memória RAM e 1 TB de HD, para processamento e armazenamento dos dados em ambiente SIG.
- Um HD externo de 2 TB para organização dos dados.
- *Softwares*: ArcGIS (ESRI CORPORATION, 2015), ENVI® (ITT VISUAL INFORMATION SOLUTION, 2015), R *software* (R DEVELOPMENT CORE TEAM, 2015).
- Um intérprete de classificação de culturas agrícolas em imagens OLI.

3.1.2. Métodos

Apenas para lembrar aos leitores, que a área de estudo 1 serviu para alcançar os dois primeiros objetivos específicos estabelecidos desta pesquisa. Para isto, os métodos foram divididos em três tópicos:

- (a) Formação do banco de dados em ambiente SIG (dados vetoriais e matriciais);
- (b) Estratificação, classificação e extrapolação da área ocupada com a cultura verde em pé de cana-de-açúcar, milho e soja, em São Paulo.
- (c) Pré-processamento das imagens OLI para obtenção de mosaicos com realce semelhante em ambas as imagens multitemporais;

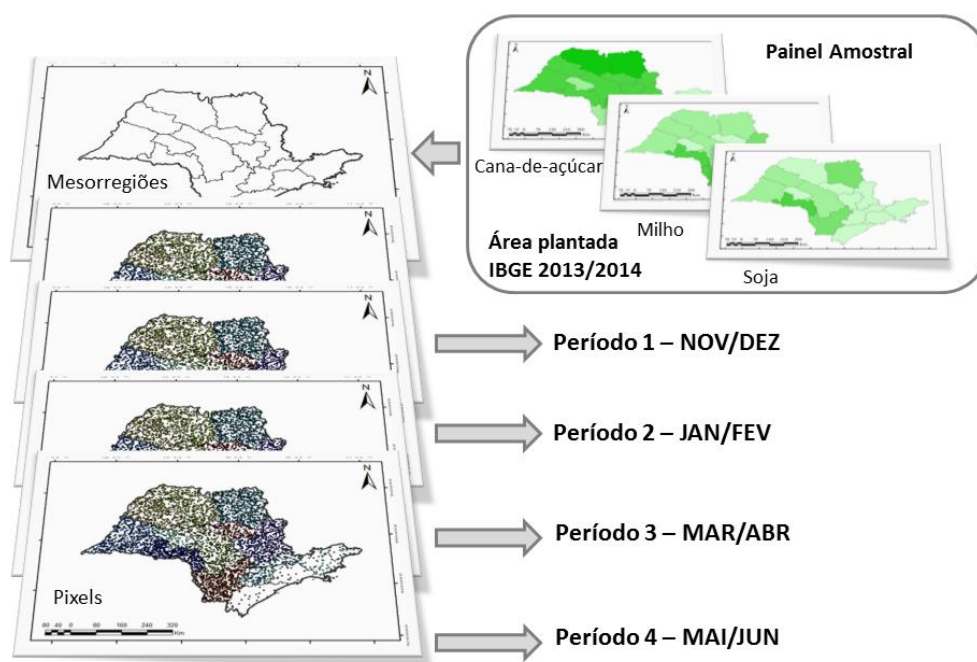
3.1.2.1. Formação do banco de dados: vetores

Os vetores da divisão político-administrativa das mesorregiões do estado de SP e das órbitas pontos OLI que recobriam o estado foram inseridos em ambiente SIG ArcGIS. Nele, todos os vetores foram projetados para a projeção cartográfica *Universal Transverse Mercator* (UTM) e Datum WGS84.

Devido à alta proporção de nuvens sobre as imagens OLI, foi definida uma regra para classificação dos pixels sorteados nas imagens multitemporais OLI. Ao iniciar o mês de novembro de 2014, o intérprete deveria classificar 5.000

pixels amostrais a cada dois meses (EBERHARDT, 2014; TRABAQUINI *et al.*, 2015).

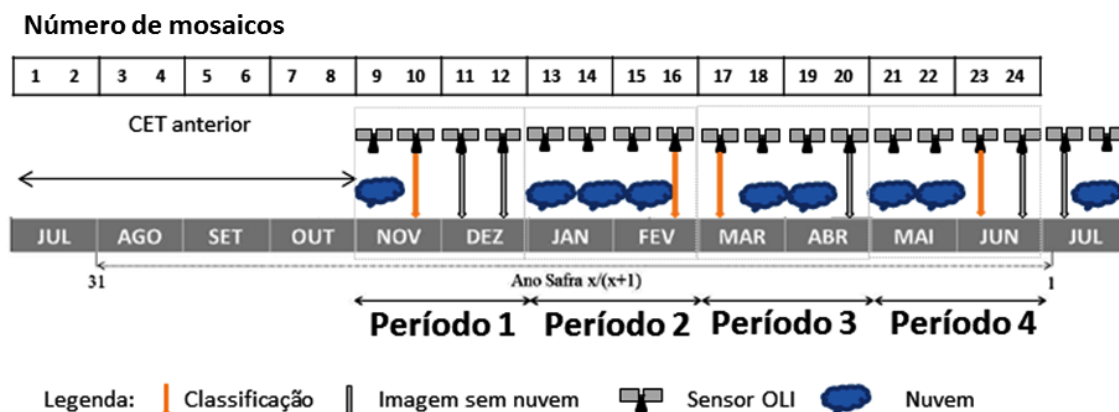
Figura 3.6 – Formação do banco de dados com dados vetoriais em ambiente SIG, a partir do painel amostral e da intensidade amostral de pixels definida por mesorregião do estado e para os quatro períodos da estimativa do ano safra 2014/2015.



O objetivo do uso desta regra de classificação foi de obter ao menos uma imagem OLI livre de nuvens de quatro disponíveis, em quatro períodos (P1 = nov/dez, P2 = jan/fev, P3 = mar/abr, P4 = mai/jun) para classificação, que compreenderam o calendário agrícola da cana-de-açúcar, milho e soja dentro de um ano safra 2014/2015, para o estado de São Paulo.

A Figura 3.7 ilustra os quatro períodos da estimativa 2014/2015, o número de mosaicos de imagens do sensor OLI sobre o estado de SP, os meses em que cada período está compreendido e como funciona a classificação visual com e sem a interferência de nuvens sobre os pixels a serem classificados em cada uma dos quatro períodos da estimativa agrícola.

Figura 3.7 – Figura ilustrativa que mostra a aquisição de imagens do sensor OLI para os quatro períodos da estimativa e regra de classificação visual (seta em laranja) utilizada para interpretação visual dos 5.000 pixels, por período do ano safra $x / (x+1)$.



Foram formados 24 mosaicos multitemporais sobre o estado. Os oito primeiros mosaicos foram utilizados como suporte na classificação visual, baseada no comportamento espectro-temporal das culturas da cana-de-açúcar, milho e soja.

Os quatro sorteios aleatórios formaram quatro diferentes arquivos vetoriais, que continham 5.000 pixels cada. Estes quatro arquivos vetoriais foram salvos em formato *shapefile* e cada um deles conteve uma lista, em formato *dbf*, com informações adicionais para cada ponto sorteado (Tabela 3.1).

Tabela 3.1 – Exemplo na organização da lista de pixels utilizada por mesorregião e por período para sorteio da amostra

ID	Latitude	Longitude	Mesorregião	Período	Classe
4	xx° xx' xx''	yy° yy' yy''	Assis	1	CN
(...)	(...)	(...)	(...)	1	CN
5.000	zz° zz' zz	ww° ww' ww''	Campinas	1	SJ

Na Tabela 3.1 percebe-se que no Período 1 (nov/dez) o identificador (ID) 4, localizado na mesorregião de Assis, foi classificado visualmente como a classe CN, que correspondeu à cultura verde em pé da cana-de-açúcar.

Em cada um destes arquivos vetoriais, o centro de cada pixel OLI foi considerado como o ponto possível de ser amostrado, conforme comentado em Luiz (2003). O total de pixels sorteados respeitou as condições impostas pelo painel amostral descrito por Luiz *et al.* (2015) na seção descrita abaixo

3.1.2.2. Painei amostral: definição da intensidade amostral

O número de pixels sorteados em cada uma das mesorregiões foi gerado pelo painei amostral descrito por Luiz *et al.* (2015) para o estado de São Paulo. De acordo com a proposta destes autores, o estado de São Paulo pode ser dividido em enésimas quadrículas imaginárias de 30 x 30 m (pixel OLI/Landsat). Cada uma destas quadrículas possui iguais chances de ser selecionada por um método amostral não estratificado.

No entanto, como visto anteriormente, as culturas agrícolas não são igualmente distribuídas na superfície do estado, isto é, existem mesorregiões que plantam mais de uma cultura e menos de outra, e desta forma, o método estratificado melhor correspondeu para o objetivo desta estimativa.

Desta forma, as mesorregiões foram definidas como estratos, e a Tabela 3.2 mostra o comportamento da área plantada como cana-de-açúcar, milho e soja, em cada um dos 15 estrados do estado de SP, para o ano base 2013, em São Paulo.

Luiz *et al.* (2015) descreveram de forma minuciosa como obter o número de amostras que devem ser sorteadas em cada um dos estratos, e ainda relacionando este número com os CVs% produzidos para cada cultura. Estes autores comentaram que ao utilizar os dados do IBGE de área plantada, ano base 2011, o processo de amostragem aleatório estratificado (AAE) pelo máximo corrigido (nc_h) foi aquele que obteve menor CVs% para estimativa de área com cana-de-açúcar, soja e milho em nível estadual.

Tabela 3.2 – Área total do território e áreas colhidas (ha) de cana-de-açúcar, milho e soja, e número teórico de pixels de 30 x 30 m, nas mesorregiões e no estado de São Paulo, em 2013 (fonte: IBGE, 2015).

Estrato/mesorregião	Território	Cana-de-açúcar	Milho	Soja	Número teórico de pixels
Araçatuba - SP	1.676.321	508.641	451.16	32.604	18.625.790
Araraquara - SP	945.529,9	381.709	18.200	4.790	10.505.888
Assis - SP	1.271.021	337.744	193.807	190.650	14.122.456
Bauru - SP	2.672.326	640.323	51.464	26.005	29.692.508
Campinas - SP	1.422.613	249.289	92.161	9.126	15.806.809
Itapetininga - SP	2.017.216	72.887	165.473	144.091	22.413.516
Litoral Sul Paulista - SP	1.320.089	27	400	0	14.667.659
Macro Metropolitana Paulista - SP	1.230.960	23.199	39.096	8.139	13.677.333
Marília - SP	717.017,1	70.580	13.106	4.441	7.966.857
Metropolitana de São Paulo - SP	929.819,2	10	532	0	10.331.324
Piracicaba - SP	904.563	313.375	24.658	4.369	10.050.700
Presidente Prudente - SP	2.403.430	525.294	39.311	31.506	26.704.778
Ribeirão Preto - SP	2.753.223	1.318.965	63.636	133.256	30.591.367
São José do Rio Preto - SP	2.939.470	970.380	67.655	20.825	32.660.778
Vale do Paraíba Paulista - SP	1.617.995	2.590	5.606	650	17.977.719
Total	24.821.593	5.415.013	820.221	610.452	275.795.480

Os parâmetros abaixo foram utilizados por estes autores, para se obter a intensidade amostral por mesorregião, pela nc_n .

- proporção (p_i) de cada cultura em cada mesorregião.
- o desvio padrão (s_{ih})
- os pesos de cada estrato (W_h)
- tamanho total da amostra (n_i)

A proporção (p_i) de cada cultura é calculada simplesmente dividindo-se a área plantada pela área total no estado, conforme demonstrado na Equação 3.1.

$$p_i = AC_i / A \quad \text{sendo } i = 1, 2 \text{ ou } 3 \quad (3.1)$$

onde A é a área total do estado, em hectares (ha); e AC_i é a área do estado ocupada com a cultura i (1 = cana-de-açúcar; 2 = milho; e 3 = soja).

Se ordenarmos as mesorregiões de São Paulo em ordem alfabética e chamarmos cada uma de estrato h e numerando-os de 1 até 15, os valores da proporção (p_{ih}), de cada cultura i em cada estrato h , podem ser obtidos de forma semelhante, como expresso na Equação 3.2.

$$p_{ih} = AC_{ih} / A_h \quad \text{sendo } h = 1, 2 \dots 15 \quad (3.2)$$

onde A_h é a área total do estrato h , em hectares (ha); e AC_{ih} é a área do estrato h ocupada com a cultura i .

No início, se considerarmos cada cultura separadamente, podemos supor que cada unidade na população (cada *pixel*, no presente caso) cai dentro de uma de duas classes C_i - significando que o *pixel* é ocupado pela cultura i , e C_i' - significando que o *pixel* não é ocupado pela cultura.

Conforme detalhado em Luiz (2009) e com base nas descrições dos cálculos envolvidos nas estimativas em amostragem de proporções, feitas por Cochran (1977), é possível afirmar que o desvio padrão (s_i), de cada cultura i no estado e o desvio padrão (s_{ih}), de cada cultura i em cada estrato h , podem ser obtidos pelas Equações 3.3 e 3.4, respectivamente.

$$s_i = \sqrt{[(p_i \cdot q_i)/n_i]} \quad \text{sendo } n_i \text{ o tamanho total da amostra e } q_i = (1-p_i) \quad (3.3)$$

$$s_{ih} = \sqrt{[(p_{ih} \cdot q_{ih})/n_{ih}]} \quad \text{sendo } n_{ih} \text{ o tamanho da amostra no estrato e } q_{ih} = (1-p_{ih}) \quad (3.4)$$

Ao dividirmos a população em estratos, é comum definir como peso do estrato nas estimativas populacionais a proporção entre ele e a população, tomando como base os tamanhos respectivos. Nesse caso, o tamanho é expresso pela área e para calcular os pesos de cada estrato (W_h), basta dividir a área de cada um pela área total do estado, conforme a Equação 3.5.

$$W_h = A_h / A \quad (3.5)$$

Para a definição da intensidade amostral por estrato de forma mais vantajosa para o estudo e curso, foi utilizado o conhecimento prévio do comportamento das variáveis a serem observadas, ou seja, as áreas de cana-de-açúcar, milho

e soja, no passado recente (Tabela 3.3). Com base nestes dados, os valores de p_{ih} , s_{ih} e W_h foram calculados e estão expostos na Tabela 3.3.

A partir dos dados da Tabela 3.3 foi possível fazer o cálculo do tamanho de amostra em cada estrato, segundo os critérios da proporcionalidade ao peso W_h e da otimização por cultura em função de p_{ih} e s_{ih} .

Tabela 3.3 – Proporção de área (p_{ih}) e respectivo desvio padrão (s_{ih}) para as culturas de cana, soja e milho, e peso (W_h) de cada estrato no Estado de São Paulo, 2013.

Mesorregiões	P_{ih}			S_{ih}^*			W_h
	Cana	Milho	Soja	Cana	Milho	Soja	
Araçatuba - SP	0,303	0,027	0,019	0,025	0,009	0,008	0,068
Araraquara - SP	0,404	0,019	0,005	0,027	0,008	0,004	0,038
Assis - SP	0,266	0,152	0,150	0,024	0,020	0,020	0,051
Bauru - SP	0,240	0,019	0,010	0,023	0,008	0,005	0,108
Campinas - SP	0,175	0,065	0,006	0,021	0,013	0,004	0,057
Itapetininga - SP	0,036	0,082	0,071	0,010	0,015	0,014	0,081
Litoral Sul Paulista - SP	0,000	0,000	0,000	0,000	0,001	0,000	0,053
Macro Metropolitana Paulista - SP	0,019	0,032	0,007	0,007	0,010	0,004	0,050
Marília - SP	0,098	0,018	0,006	0,016	0,007	0,004	0,029
Metropolitana de São Paulo - SP	0,000	0,001	0,000	0,000	0,001	0,000	0,037
Piracicaba - SP	0,346	0,027	0,005	0,026	0,009	0,004	0,036
Presidente Prudente - SP	0,219	0,016	0,013	0,023	0,007	0,006	0,097
Ribeirão Preto - SP	0,479	0,023	0,048	0,027	0,008	0,012	0,111
São José do Rio Preto - SP	0,330	0,023	0,007	0,026	0,008	0,005	0,118
Vale do Paraíba Paulista - SP	0,002	0,003	0,000	0,002	0,003	0,001	0,065
Total no Estado de São Paulo	0,194	0,034	0,023	0,017	0,008	0,006	--

* o cálculo supõe amostra aleatória simples, tamanho $n_{ih} = 333,33$ para cada estrato e $n = 5.000$ para o estado, ** os valores se referem à p_i e s_i ,

No caso da alocação proporcional, o tamanho da amostra nos estratos (nr_h) é igual para qualquer cultura, pois só depende do tamanho do estrato em relação ao estado, Para o seu cálculo, basta multiplicar o peso do estrato (W_h) pelo tamanho total da amostra (n_i), que nesse caso foi definido como igual para todas as culturas ($n_1 = n_2 = n_3 = 5.000$),

Para a alocação ótima, o tamanho da amostra em cada estrato (no_{ih}) é definido levando-se em conta também o desvio padrão da cultura i em cada caso, como expresso na Equação 3.6.

$$no_{ih} = n_i (W_h, S_{ih}) / \sum (W_h, S_{ih}) \quad (3.6)$$

Para obter-se o tamanho de amostra em um estrato pela *alocação máxima* (nx_h), basta calcular o máximo, respectivamente, entre os tamanhos de amostra ótimos das culturas naquele estrato (no_{ih}),

Entretanto, ao adotar-se o maior valor entre as três culturas (nx_h), ou o valor ótimo para o pior caso, foi encontrado um tamanho final da amostra é maior que 5,000, Para manter o tamanho total na dimensão preestabelecida, é proposta a alocação máxima corrigida (nc_h), que é obtida da aplicação da Equação 3.7.

$$nc_h = nx_h, n / \sum nx_h \quad (3.7)$$

Tabela 3.4 – Proporção de área (p_{ih}) e respectivo desvio padrão (s_{ih}) para as culturas de cana, soja e milho, e peso (W_h) de cada estrato no Estado de São Paulo, 2013.

Estrato (mesorregião)	Proporcional (nr_h)	Ótima (no_{ih})			Corrigida (nc_h)
		Cana	milho	Soja	
Araçatuba - SP	338	454	349	385	340
Araraquara - SP	190	273	167	112	204
Assis - SP	256	331	588	754	564
Bauru - SP	537	672	472	436	503
Campinas - SP	287	319	450	189	337
Itapetininga - SP	406	222	712	863	645
Litoral Sul Paulista - SP	266	4	30	0	22
Macro Metropolitana Paulista - SP	248	99	278	166	208
Marília - SP	144	126	124	93	94
Metropolitana de São Paulo - SP	187	2	29	0	22
Piracicaba - SP	182	253	189	104	189
Presidente Prudente - SP	484	585	392	454	438
Ribeirão Preto - SP	555	810	532	982	734
São José do Rio Preto - SP	592	814	567	410	609
Vale do Paraíba Paulista - SP	326	38	122	54	91
Total	4998	5002	5001	5002	5000

Como se pode ver na Tabela 3.4, onde aparecem todos esses valores calculados para um tamanho de amostra total proposto de $n=5.000$, a alocação individual ótima difere bastante de uma cultura para outra e mais ainda quando comparadas à alocação proporcional.

Tendo obtido o valor de amostras sorteadas por estrato, o próximo passo foi encontrar o valor de CV% obtido para cada cultura. Para cálculo do CVs% para cada uma das culturas avaliadas, cana-de-açúcar, milho e soja, o *número teórico* de pixels amostrados foi 5.000.

Este número definiu a intensidade amostral utilizada em cada um dos estratos, no entanto o número real de pixels amostrados dependeu do *número de pixels classificados* com cana-de-açúcar, milho e soja somados de uma classe chamada outros (não-cana., não-milho e não-soja), em cada um dos quatro períodos da estimativa agrícola 2014/2015.

Aqueles pixels classificados como nuvem, irão reduzir o número de pixels amostrados, e assim aumentar o CVs% de cada uma das culturas.

Apesar disso, Luiz *et al.* (2015) mostraram que mesmo com um valor de 3.000 pixels amostrados, o CVs% para cana-de-açúcar, milho e soja foi estatisticamente aceitável para o estado de São Paulo. Assim, percebe-se que ao lançar um valor de intensidade amostral de 5.000 pixels, o presente painel amostral terá 33,3% a mais de pixels sorteados. Este número servir para assegurar que o valor dos CVs% das estimativas, ainda seja inferior ao valor encontrado por Luiz *et al.* (2015).

3.1.2.3. Painel amostral: extrapolação

Após realizar todo o processo de classificação dos pixels estratificados, o próximo passo foi realizar a extrapolação dos resultados, em cada classe e período para valores de área ocupada com a cultura para todo o estado de São Paulo.

A Tabela 3.5 mostra a organização dos resultados do número de pixels classificados como cana, milho, soja, nuvem, outros e total, em cada mesorregião do estado (Est. = estrato) e em cada período (P.) da estimativa.

Ao organizar o número de pixels classificados em cada classe supra referida o próximo passo foi extrapolar estas classes por estrato, utilizando para isso o método da extrapolação por área descrito por Epiphany et al., (2002).

Tabela 3.5 – Proporção de área (p_{ih}) e respectivo desvio padrão (s_{ih}) para as culturas de cana, soja e milho, e peso (W_h) de cada estrato no Estado de São Paulo, 2013.

P.	Est.	Mesorregião	Número de pixels / classe visualmente interpretada					
			Cana	Milho	Soja	Nuvem	Outros	Total
1º	1	Araçatuba	1	1	1	3	123	340
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
1º	15	VPP	2	1	1	6	31	91
1º		Estado SP	767	15	93	230	3789	5000

O método da extrapolação (E_{ih}) é muito simples, ele depende apenas do valor da área de cada estrato (AC_{ih}), o número de pixels de cada classe (n_i) e tamanho da amostra dos estratos (nr_h). Com estes valores, faz-se a multiplicação da área de cada uma das mesorregiões/estratos (AC_{ih}) pelo número de pixels classificados com determinada classe (n_i), dividido pelo tamanho da amostra dos estratos (nr_h).

A Equação 3.8 foi utilizada no cálculo da extrapolação em área.

$$E_{ih} = (AC_{ih} \times n_i) / nr_h \quad (3.8)$$

A Tabela 3.6 mostra a organização dos dados, para obtenção dos valores em área de cada cultura, em cada período, por estrato e para o estado de São Paulo.

Tabela 3.6 – Proporção de área (p_{ih}) e respectivo desvio padrão (s_{ih}) para as culturas de cana, soja e milho, e peso (W_h) de cada estrato no Estado de São Paulo, 2013.

P.	Est.	Mesorregião	Área (ha)	Extrapolação - Área (ha)					
				Cana	Milho	Soja	Nuvem	Outros	Total
1º	1	Araçatuba	167.6321,1	4.930,4	4.930,4	4.930,4	14.791,1	606.433,8	29.582,1
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
1º	15	VPP	1.617.994,7	35.560,3	17.780,2	17.780,2	106.681,0	551.185,0	177.801,6
1º		Estado SP (milhões ha)	24,8	3,8	0,07	0,5	1,1	19,3	24,8

Após obter a estimativa de área ocupada com cada cultura, o painel amostral informou também o CVs% resultantes para as culturas da cana-de-açúcar, milho e soja, no estado de São Paulo.

No caso da alocação pelo máximo corrigido dos valores ótimos, conforme já descrito, o cálculo do CV_{cr} é o seguinte Equações 3.9.

$$CV_{cr} = [100.\sqrt{\Sigma(W_h^2.p_{ih}.q_{ih})/(nc_h-n_n)}]/p_i \quad (3.9)$$

Onde $(nc_h - n_n)$ refere-se ao número de pixels classificados.

3.1.2.4. Formação do banco de dados matricial

Como as imagens foram sistematicamente fornecidas pela USGS, um ou dois dias após a passagem do sensor sobre cada órbita do estado de São Paulo, e as estimativas agrícolas foram realizadas em tempo quase-real, o intérprete deveria conhecer a data das imagens para saber qual o tempo hábil que ele teria para execução de toda a parte de processamento dos dados e classificação visual (Tabela 3.7).

A Tabela 3.5 ilustra os 24 mosaicos do sensor OLI (linhas da Tabela 3.7) utilizadas neste trabalho, e a data final que cada mosaico foi finalmente construído. A data de construção do mosaico de todas as 18 imagens OLI de uma mesma visita foi referente à data das imagens oriundas, da última órbita formada no estado de São Paulo (órbita 219) (Tabela 3.7).

Estas imagens foram baixadas e armazenadas em formato *layer stacking* R-b5 G-b6 B-b4 com 16 bits de resolução radiométrica, tendo o nome que rege:

rgb_LC8(órbita)(ponto)(ano)(doy)_sr

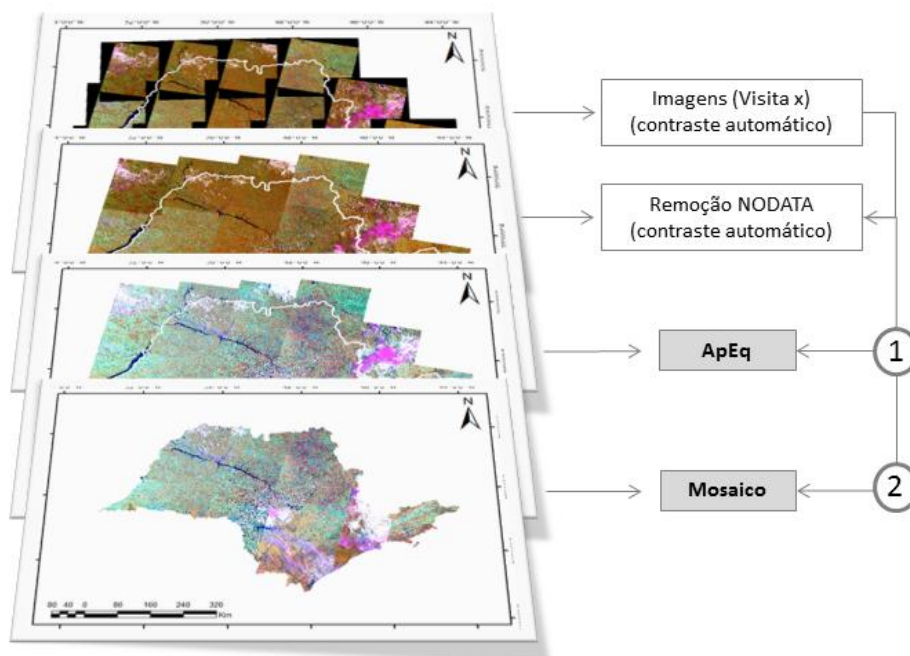
Tabela 3.7 – Datas das imagens OLI utilizadas nos quatros período da estimativa agrícola para cana-de-açúcar, milho e soja no ano safra 2014/2015, em São Paulo. Cada linha da Tabela refere-se a uma visita do sensor OLI sobre o estado de São Paulo.

Período/Data	Órbita/pontos						
	Nº	222/74/75/ 76/77	220/74/75/ 76/77	218/76	223/74/75 /76	221/74/75 /76	219/75/76
Banco de dados	1 ^a	07/jul/14	09/jul/14	11/jul/14	14/jul/14	16/jul/14	18/jul/14
	2 ^a	23/jul/14	25/jul/14	27/jul/14	30/jul/14	01/ago/14	03/ago/14
	3 ^a	08/ago/14	10/ago/14	12/ago/14	15/ago/14	17/ago/14	19/ago/14
	4 ^a	24/ago/14	26/ago/14	28/ago/14	31/ago/14	02/set/14	04/set/14
	5 ^a	09/set/14	11/set/14	13/set/14	16/set/14	18/set/14	20/set/14
	6 ^a	25/set/14	27/set/14	29/set/14	02/out/14	04/out/14	06/out/14
	7 ^a	11/out/14	13/out/14	15/out/14	18/out/14	20/out/14	22/out/14
	8 ^a	27/out/14	29/out/14	31/out/14	03/nov/14	05/nov/14	07/nov/14
Primeiro Período NOV/DEZ	9 ^a	12/nov/14	14/nov/14	16/nov/14	19/nov/14	21/nov/14	23/nov/14 ^a
	10 ^a	28/nov/14	30/nov/14	02/dez/14	05/dez/14	07/dez/14	09/dez/14
	11 ^a	14/dez/14	16/dez/14	18/dez/14	21/dez/14	23/dez/14	25/dez/14
	12 ^a	30/dez/14	01/jan/15	03/jan/15	06/jan/15	08/jan/15	10/jan/15 ^b
Segundo Período JAN/FEV	13 ^a	15/jan/15	17/jan/15	19/jan/15	22/jan/15	24/jan/15	26/jan/15 ^a
	14 ^a	31/jan/15	02/fev/15	04/fev/15	07/fev/15	09/fev/15	11/fev/15
	15 ^a	16/fev/15	18/fev/15	20/fev/15	23/fev/15	25/fev/15	27/fev/15
	16 ^a	04/mar/15	06/mar/15	08/mar/15	11/mar/15	13/mar/15	15/mar/15 ^b
Terceiro Período MAR/ABR	17 ^a	20/mar/15	22/mar/15	24/mar/15	27/mar/15	29/mar/15	31/mar/15 ^a
	18 ^a	05/abr/15	07/abr/15	09/abr/15	12/abr/15	14/abr/15	16/abr/15
	19 ^a	21/abr/15	23/abr/15	25/abr/15	28/abr/15	30/abr/15	02/mai/15
	20 ^a	07/mai/15	09/mai/15	11/mai/15	14/mai/15	16/mai/15	18/mai/15 ^b
Quarto Período MAI/JUN	21 ^a	23/mai/15	25/mai/15	27/mai/15	30/mai/15	01/jun/15	03/jun/15 ^a
	22 ^a	08/jun/15	10/jun/15	12/jun/15	15/jun/15	17/jun/15	19/jun/15
	23 ^a	24/jun/15	26/jun/15	28/jun/15	01/jul/15	03/jul/15	05/jul/15
	24 ^a	10/jul/15	12/jul/15	14/jul/15	17/jul/15	19/jul/15	21/jul/15 ^b

^a Início da classificação visual, por período; ^b Término das imagens disponíveis para classificação visual dos pixels, por período.

Após constarem as 18 imagens, de cada mosaico, no banco de dados, o pré-processamento nos dados era então aplicado, por mosaico. O fluxograma da Figura 3.8 ilustra como foi executado o armazenamento dos arquivos no banco de dados, pré-processamento e confecção do banco de dados para classificação visual, em ambiente SIG.

Figura 3.8 – Formação do banco de dados com dados matriciais em ambiente SIG, a partir das 18 imagens OLI em composição colorida R-b5 G-b6 B-b4 sobre o estado de São Paulo, remoção dos valores de *Nodata*, aplicação da ApEq e recorte do mosaico equalizado, para a abrangência da área do estado de São Paulo.



Dois procedimentos operacionais foram aplicados nas imagens: (1) equalização e (2) mosaico. A equalização foi feita pelo método da Aparência Equalizada (ApEq), que será descrito na seção a baixo e o mosaico para cada imagem foi construído pela interpolação do vizinho mais próximo.

3.1.2.4.1. Aparência Equalizada (ApEq)

O objetivo da ApEq foi equalizar e normalizar as imagens por meio de uma regra simples e interativa de processamento dos dados. Para isso, todas as imagens nomeadas *rgb_LC8(órbita)(ponto)(ano)(doy)_sr* de cada mosaico OLI, do ano safra 2014/2015, foram utilizadas.

Para cumprir tal objetivo proposto pela ApEq, dois passos foram executados com estas imagens:

O primeiro passo foi utilizar uma imagem base (IB) como referencial de imagem equalizada. As demais imagens (de conversão) (IC), inseridas no banco de dados, teriam que ter uma aparência equalizada semelhante à imagem base.

A IB foi definida como sendo a imagem que tivesse baixa porcentagem de nuvens e também, todas as classes almejadas neste presente estudo (cana-de-açúcar, milho e soja) deveriam estar presentes. Desta forma, a imagem 222/76 de 28 de Janeiro de 2014 foi utilizada como imagem base.

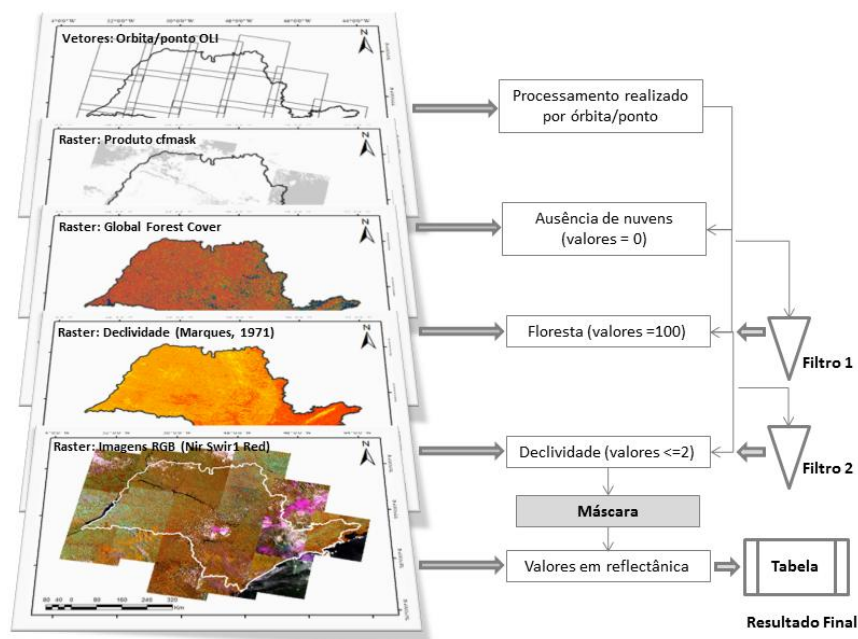
A equalização desta imagem foi feita manualmente, alterando-se os valores de máximo e mínimo das suas bandas em ambiente RGB. Posteriormente, a imagem com os novos histogramas foi salva em resolução radiométrica igual à inicial (16 *bits*), e foi chamada de ApEq_IB.

Para que as ICs fossem equalizadas, os valores de reflectância da superfície (σ) entre alvos semelhantes destas imagens com alvos semelhantes da IB foram extraídos. Os únicos alvos naturais que poderiam ser encontrados em todas as imagens e que se comportam nas imagens *Landsat-like* como menos invariantes possíveis foram os pixels de floresta natural (HANSEN *et al.*, 2013), Eivazi *et al.* (2015) chamaram estes alvos de alvos pseudo-invariantes (APIs).

Assim, os pixels de floresta das imagens ApEq_IC deveriam ter a cor semelhante (Aparência Equalizada) em todas as imagens, com os pixels de floresta da ApEq_IB.

A partir daí, o princípio da ApEq foi muito simples, de todos os pixels de floresta, não tiveram mudança na cobertura de floresta informada nos dados matriciais do GFC (ACHARD e HANSEN, 2013) em cada limite órbita/ponto das imagens *rgb_LC8(órbita)(ponto)(ano)(doy)_sr* do estado de SP, foram extraídos automaticamente os valores das bandas dos comprimento de onda de b4, b5 e b6. Para chegar a tais resultados, uma máscara de nuvens e de declividade foi construída (Figura 3.9),

Figura 3.9 – Passos descritos para aplicação da ApEq em imagens multitemporais OLI.



A ApEq utilizou as bandas b-4, b-5 e b-6 do sensor OLI em reflectância, o produto *cfmask* das imagens OLI em reflectância (USGS, 2015), o produto GFC em matriz 30 x 30 m (Hansen *et al*, 2013), as classes de declividade em matriz 30 x 30 m (Marques, 1971) e os vetores das órbitas/pontos do sensor OLI em sistema WRS-2.

Como os dados de floresta informados em cada pixel de 30 m pelo produto GFC foram computados para o ano de 2000, demais informações de perda e ganho de floresta até 2014, foram incorporados a este dado. Esta atualização foi realizada tendo como base os demais produtos presentes no catálogo *Global Forest Change*, obtidos pelo método publicado por Hansen *et al.*, (2013).

Para controlar o problema referente à presença de nuvens os valores de σ dos pixels de floresta, o produto *cf mask* de cada imagem OLI inicialmente salva no banco de dados, foi utilizado como filtro 1 (Figura 3.9). O filtro serviu para seleção de pixels com ausência (valor 0) e presença de nuvem (valor 1). Os valores de σ das três bandas supra referidas foram inicialmente selecionados apenas dos pixels que continham valor 0 no produto *cf mask*.

Para controlar o problema referente ao fator topográfico sobre σ de floresta, o produto declividade de Marques (1971) foi utilizado como filtro 2 ou secundário (Figura 3.9). Este filtro refinou ainda mais os valores radiométricos inicialmente selecionados no primeiro filtro, Apenas os pixels de floresta advindos do primeiro filtro e que estavam contidos na classe de declividade 1 (< 2,5%) e 2 (2,5 a 12%) foram selecionados,

O fluxograma da Figura 3.9 ilustra como a ApEq foi aplicada, 24 vezes, sobre todas 18 imagens de cada um dos 24 mosaicos utilizados na estimativa agrícola de 2014/2015. Todos estes passos descritos acima foram automaticamente processados pelo *script* 1, desenvolvido em linguagem *Python* (anexo 1).

Ao final do procedimento automático, a média dos σ em cada uma das três bandas (b4, b5 e b6) foi extraída em 16 bits de todos os pixels contidos que foram filtrados pelo procedimento descrito na Figura 3.9, contidos e cada cena OLI. Foi gerada uma lista com estas informações para cada mosaico do sensor OLI sobre o estado de SP. A Tabela 3.8 mostra a organização desta lista.

Tabela 3.8 – Exemplo na organização da lista das métricas obtidas dos valores de reflectância da superfície (σ) das bandas do b4, b5 e b6 para os pixels de floresta filtrados pelo passo 1 da ApEq, na imagens: Base (IB), 218/76 e 223/76.

Imagens	Mosaico (Nº)	Órbita	Ponto	Média		
				σ_{b4}	σ_{b5}	σ_{b6}
IB	--	222	76	366,56	1,123,56	989,56
	19	218	76	345,23	1,234,65	1,000,45
IC	(,,)	(,,)	(,,)	(,,)	(,,)	(,,)
	19	223	76	376,98	1,346	1,000,5

O **segundo passo** da ApEq foi corrigir os valores dos histogramas das ICs, pela imagem IB. Para isto, os valores de reflectância da superfície, das três bandas espectrais selecionadas ApEq_IB obtidos na lista do primeiro passo da ApEq, foram divididos pelos valores de reflectância de cada IC (Equação 3.10).

$$FM_k = \sigma_{IBk} / \sigma_{ICK} \quad [3.10]$$

Onde σ_{IBk} refere-se ao valor de reflectância da superfície da imagem base, em cada banda k e σ_{ICk} é o valor de reflectância de cada imagem de conversão, também em cada banda k .

O pressuposto desta divisão se baseia unicamente no valor do fator multiplicativo (Fm) gerado pela divisão de σ_{IBk} por σ_{ICk} . Este fator multiplicativo (Fm) por si só conseguiria diminuir os efeitos ruidosos causados por geometria de iluminação e sazonalidade sobre as ICs, para que os APIs destas imagens tivessem valores tonais e cores semelhantes aos APIs de IB.

A Tabela 3.9 ilustra a organização do procedimento operacional utilizado para obter os valores de FM a partir dos valores de σ de ApEq_IB e das ICs, obtidos no primeiro passo da ApEq.

Tabela 3.9 – Exemplo na organização da lista dos valores de Fm das bandas do b4, b5 e b6 para as imagens IB = 218/76 e IC = 223/76.

Imagem	Mosaico (Nº)	Órbita	Ponto	Fator multiplicativo (Fm)		
				Fm ob4	Fm ob5r	Fm ob6
IB	--	222	76	1,00	1,00	1,00
	19	218	76	1,06	0,91	0,99
IC	(,,)	(,,)	(,,)	(,,)	(,,)	(,,)
	19	223	76	0,98	0,83	0,99

Os Fm's resultantes em cada IC foram posteriormente utilizados para multiplicar os valores de máximo e mínimo dos histogramas das bandas da IB inserida em ambiente RBG. Os novos valores de máximo e mínimo, gerados a partir da multiplicação dos valores dos histogramas da IB foram então aplicados como novos valores na equalização das ICs inserida no SIG, também em ambiente RGB (Tabela 3.10).

Tabela 3.10 – Exemplo na organização da lista das métricas obtidas dos valores de Fm das bandas do b4, b5 e b6 para as imagens IB = 218/76 e IC= 223/76.

Mosaico (Nº)	Órbita	Ponto	ob4		ob5		ob6	
			Max	Min	Max	Min	Max	Min
	IB		1,234	-1,23	10,345	323,45	6,544,32	12,344
19	218	76	1,310,24	-1,31	9,414,19	294,35	6,473,08	12,21
(,,)	(,,)	(,,)	(,,)	(,,)	(,,)	(,,)	(,,)	(,,)
19	223	76	1,199,89	-1,20	8,635,51	270,00	6,473,08	12,21

Após a ApEq ser aplicada nas imagens, estas foram armazenadas em *layer stacking* R-b5 G-b6 B-b4 então salvas no formato *band sequential file* (BSQ), em 8 *bits* de resolução espectral, com o nome:

$$ApEq_LC8(órbita)(ponto)(ano)(doy)_sr$$

A hipótese do uso da ApEq na presente pesquisa é que a partir das imagens convertidas em ApEq, os pixels de floresta das ICs tendem a adquirir o mesmo histograma RGB dos pixels de floresta da IB. Demais alvos nas ICs também serão assim, automaticamente corrigidos para a equalização da IB.

Como o objetivo da ApEq era manter invariantes as cores dos pixels de floresta, consequentemente um talhão de soja, plantado no Norte do estado de São Paulo (órbita/ponto 220/74) terão a mesma cor de um talhão de soja de idades semelhantes plantado no Sul do Estado (órbita/ponto 221/76). Além da soja, a cana-de-açúcar e milho teriam este mesmo comportamento,

O método da ApEq foi validado conforme Eivazi *et al.*, (2015). Estes autores comentaram que a validação de mosaicos deve ser feita de duas formas:

- Comparativamente, pela fotointerpretação de alvos em imagens localizadas em diferentes órbitas/pontos;
- Visualmente, observando a linha formada entre as áreas de sobreposição de imagens adjacentes.

3.1.2.4.2. Construção dos mosaicos de imagens em ApEq

Os mosaicos foram construídos nos dias em que a última imagem de cada mosaico OLI foi inserida em ambiente SIG, pelo usuário. Neste mesmo dia, a ApEq foi aplicada em cada uma destas imagens e as com o nome *ApEq_LC8(órbita)(ponto)(ano)(doy)_sr* foram utilizadas no mosaico.

O mosaico das imagens foi construído no programa Envi 4,8, pelo uso das ferramentas *Map – mosaicking – georeferencead*.

A partir desta sequência de comandos, uma tela é mostrada com a possibilidade de inserir as imagens a serem utilizadas no mosaico. Quando todas as imagens foram inseridas nesta tela, foram eliminados os valores de *nodata* destas imagens. Além disso, quando as imagens foram inseridas automaticamente o programa altera os valores dos seus histogramas. Para que isto não ocorra, é necessário que o usuário defina, nos parâmetros do algoritmo, o termo “nenhum balanço” dos histogramas.

Na área de sobreposição formada entre as imagens adjacentes foi gerada uma regra, em que a área de sobreposição que tivesse menor interferência de nuvens (valor dado pelo produto *cf mask*) seria sobreposta à sobreposição com maior interferência de nuvens. O objetivo do uso desta regra foi reduzir o número de pixels com nuvens nos mosaicos confeccionados.

Posteriormente, cada um dos mosaicos foi recortado pelo polígono da unidade federativa do estado de São Paulo, e o arquivo foi salvo em formato TIF, em resolução espectral de 8 *bits*. Estes procedimentos foram realizados 24 vezes (dentro do ano safra 2014/2015).

Da visita 1 a 9 todos os mosaicos foram processados em lote e apenas a partir da décima visita do sensor OLI sobre SP que os dados foram processados por demanda, isto é, a cada dois meses. Cada mosaico formado era chamado de:

Visita_(n)_(mês de referência das imagens)_(ano)

Onde o n variou de 1 a 24 e os meses, de julho/2014 a julho/2015.

Demais dados matriciais como o Topodata e GFC passaram pelo mesmo processo descrito a cima.

3.2. Área de estudo 2 e 3 –Cândido Mota e municípios vizinhos

Mesmo a presente pesquisa englobando todo o estado de São Paulo, a área de estudo 1 foi utilizada pela necessidade de dados de campo mais concentrados (um único município) e que abrangiam as três culturas agrícolas estudadas (cana-de-açúcar, milho e soja). E na área de estudo 3, referente a 15 municípios do extremo sul do estado de São Paulo foi estudado o resultado da classificação automática de cana-de-açúcar e milho (culturas temporalmente presentes na imagem) para confecção de mapas contínuos com as culturas, em apenas uma imagem Landsat-like.

O município de Cândido Mota possui 59.629 ha e está localizado próximo ao município de Assis e à SP 270 – Rodovia Raposo Tavares ao Norte, e ao Sul faz divisa com o rio Paranapanema (Figura 3.10). Ainda na Figura 3.10, percebe-se que 15 municípios do estado de São Paulo listados nesta Figura, são recobertos pela imagem 222/076 do sensor OLI. Estes municípios possuem aptidão agrícola semelhante ao município de Cândido Mota (IBGE, 2015).

Figura 3.10 – Localização da área de estudo 2 (Município de Cândido Mota / número 15 da lista de municípios) e da área de estudo 3 (quinze municípios do estado de São Paulo recobertos pela imagem 222/076 do sensor OLI).

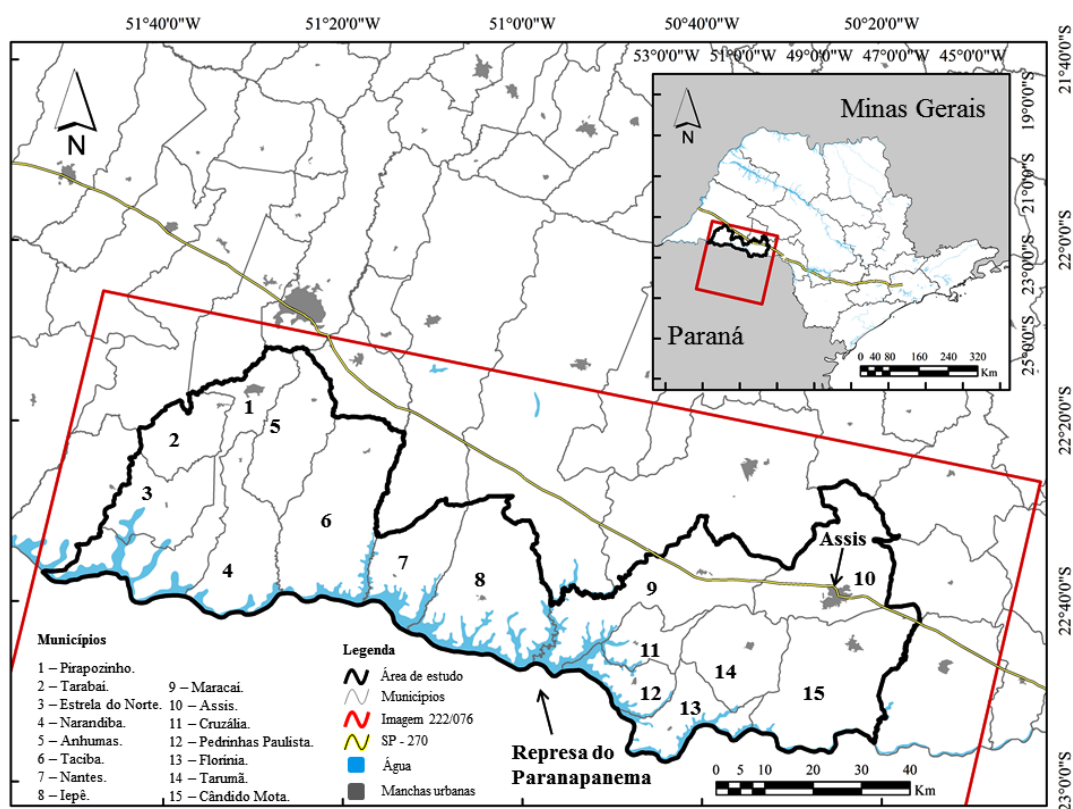


Figura 3.11 – (a) Cor do solo presente na região de estudo 2 e 3 e (b) erosões causadas em áreas agrícolas com declividade suave ondulada.



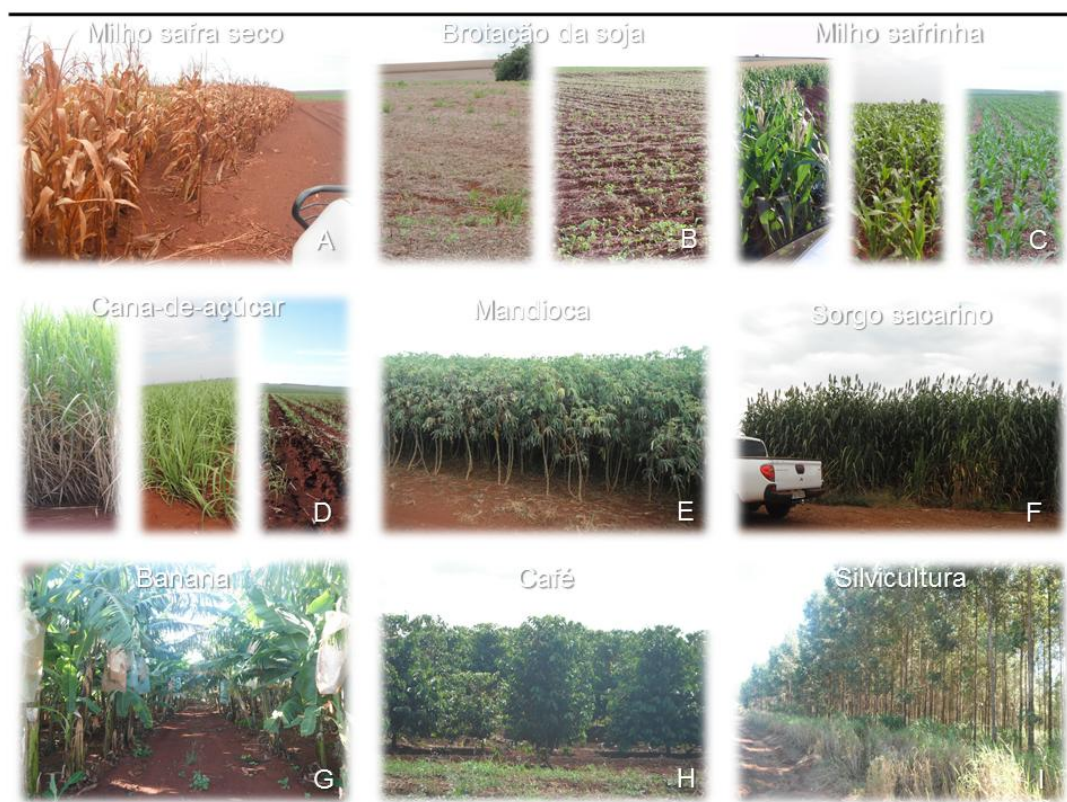
Fotos retiradas no trabalho de campo 2 (ver seção 3.2.2.2, pg. 101) realizado nos dias 13-17/abril/2015, no município de Cândido Mota.

O solo da região é classificado como Latossolo Vermelho-Eutroférrico (LVef1), constituído basicamente por basalto e óxidos ferrosos, o que confere a estes tipos de solo, coloração avermelhada (Figura 3.11).

O clima da região é o cfa (clima temperado húmido) com verões quentes (ALVARES *et al.* 2013) o que favorece o plantio de duas safras no mesmo ano (por exemplo, milho safrinha que sucede o plantio da soja).

As principais culturas plantadas na região agrícola do município são: milho (24.000 ha), soja (23.500 ha) cana-de-açúcar (21.000 ha), mandioca (1.045 ha), banana (250 ha) e café (200 ha) (IBGE, 2015) (Figura 3.12). É possível encontrar também, com pouca expressividade em área, talhões de sorgo sacarino nesta área (Figura 3.12.f)

Figura 3.12 – Principais culturas agrícolas/silvícolas fotografadas no trabalho de campo 2, datado no dia 13-17/abril/2015 no município de Cândido Mota.



3.2.1. Material

Para realização das pesquisas referentes aos "objetivos 3 a 9" foram necessários os seguintes insumos e capacidade de trabalho:

- Imagens de sensoriamento remoto (Tabela 3.11).

Tabela 3.11 – Imagens OLI e *RapidEye* usadas no estudo.

Identificação da imagem	OLI		<i>RapidEye</i> (órbital/ponto)			
	222/76	Mosaicos	285/15	286/15	287/15	288/15
			16/17/18	16/17/18	16/17/18	16/17/18
Reflectância da superfície (16 bits)	08/ago/14	--	--	--	--	--
	27/out/14	--	--	--	--	--
	15/jan/15	--	--	--	--	--
	21/abr/15	--	--	--	--	--
Mosaico R-b5 G-b6 B-b4) (8 bits)	08/ago/14	3 ^a	--	--	--	--
	09/set/15	5 ^a				
	11/out/14	7 ^a	--	--	--	--
	27/out/14	8 ^a				
	28/nov/14	10 ^a	--	--	--	--
	30/dez/14	12 ^a	--	--	--	--
	15/jan/15	13 ^a				
	04/mar/15	16 ^a	--	--	--	--
	21/abr/15	19 ^a				
	07/mai/15	20 ^a	--	--	--	--
Imagem R-b4 G-b5 B-b3 (12 bits)	--	--	03/jul/13	04/jul/13	05/jul/13	27/jul/13
Total do número de imagens	4	10	4	4	4	4

Onde: o termo “mosaicos” refere-se à combinação das 18 imagens que recobrem o estado de São Paulo, em ApEq na composição colorida falsa cor R-b5 G-b6 B-b4.

- O mosaico do TOPODATA para o estado de São Paulo (<http://www.webmapit.com.br/inpe/topodata/>).
- O mosaico dos produtos *Global Forest Cover* (Hansen *et al.*, 2013) disponíveis no catálogo (<http://earthenginepartners.appspot.com/science-2013-global-forest>).
- Vetores de área disponíveis para colheita de cana-de-açúcar, no ano base 2012 (CANASAT, 2013);

- Vetores de áreas com pivôs centrais no estado de São Paulo disponível em <http://mapas.cnpm.embrapa.br/somabrasil/webgis.html> (GUIMARÃES e LAU, 2011);
- Vetores de divisão político administrativo dos municípios do estado de São Paulo.
- Oito intérpretes para classificação de culturas agrícolas em imagens OLI (quatro experientes + quatro sem experiência).
- Um computador HD externo de 2 TB para organização dos dados.
- Um computador para processamento e armazenamento dos dados em ambiente SIG.
- *Softwares*: ArcGIS (ESRI CORPORATION, 2015), ENVI® (ITT VISUAL INFORMATION SOLUTION, 2015), R *software* (R DEVELOPMENT CORE TEAM, 2015).
- Dados validados em trabalhos de campo.
 - computador LENOVO YOGA i5, 6 GB RAM e 500 GB HD;
 - Software *Universal Maps Downloader* e *Global Mapper v.10*;
 - Imagens OLI com data mais próxima ao trabalho de campo;
 - imagens *RapidEye* datadas de 2013;
 - GPS *garmin*;
 - Carro modelo L200 *Triton*;
 - máquinas fotográficas modelo SONY e FUGI com GPS;
 - equipe composta por três membros, para direção, navegação e coleta de dados de campo;

3.2.2. Métodos

As áreas de estudo 2 e 3 serviram para cumprir com sete objetivos (3 a 9) específicos da presente pesquisa:

Para conquistar tais objetivos, duas campanhas de campo foram realizadas. Estas campanhas de campo serviram basicamente para entendimento da dinâmica agrícola do estado de São Paulo e para montagem de um bando de dados robusto para validação de dados agrícolas em imagens de sensoriamento remoto.

3.2.2.1. Trabalho de campo 1: microrregião de Assis

Este trabalho de campo foi realizado em março de 2014 e serviu mais como trabalho piloto para execução de outro trabalho de campo, executado um ano depois desta campanha (em abril de 2015).

Neste trabalho de campo 1, um total de 645 pixels aleatórios foram classificados sobre a mesorregião de Assis. Destes, apenas os pixels que estavam contidos no entorno da microrregião de Assis foram selecionados para validação (310 pixels).

No entanto, percebeu-se que este grande conjunto de pontos a se validar era oneroso em relação aos gastos de tempo e dinheiro. Desta forma, a validação foi direcionada para aqueles pixels contidos nos municípios que contiveram os maiores valores de área plantada com cana-de-açúcar, soja e milho na microrregião de Assis, em 2011. A partir daí, foram selecionados 91 pixels aleatórios sobre os municípios de Quatá, Paraguaçu Paulista e Campos Novos Paulista.

No processo de validação, foi constatada a similaridade espectral entre a soja, amendoim e mandioca vistas na imagem R-b5 G-b6 B-b4 de março/2014 com as culturas vistas em pleno vigor vegetativo, no campo. Demais detalhes estão em Schultz *et al.* (2015a), trabalho o qual foi escrito tendo por base estes dados de campo.

Esta informação foi valiosa no momento da montagem do protocolo de classificação feito para o ano safra 2014/2015. Além disso, todo o procedimento aprendido e executado para validação dos pontos em campo, foi novamente aplicado no Trabalho de campo 2, porém, agora realizado em nível municipal.

3.2.2.2. Trabalho de campo 2: município de Cândido Mota

Esta campanha de campo foi executada em um período de cinco dias, iniciando-se os trabalhos em 13 de abril/2015 e finalizando-os em 17 de abril/2015.

Anteriormente à visita de campo, os dados foram todos processados em ambiente de laboratório.

3.2.2.2.1. Atividades desenvolvidas em laboratório

Em ambiente SIG *Global Mapper*, inicialmente 200 pixels foram sorteados aleatoriamente sobre o município de Cândido Mota. Estes pixels foram salvos em formato *shapefile* que continha uma Tabela com informações sobre ID, posição geográfica e classe (em branco) de cada pixel sorteado.

Então estes pixels foram classificados sobre a imagem de 04 de março de 2015, por intérprete treinado. Apenas quatro classes poderiam ser indexadas em cada um destes pixels (cana-de-açúcar, milho, soja e outros). Imagens OLI, em composição colorida R-b5 G-b6 B-b4 também foram utilizadas no procedimento de classificação.

A imagem de 04 de março de 2015 foi inserida no banco de dados para servir como imagem base para navegação, em campo.

Imagens *RapidEye* foram baixadas do catálogo do INPE (<http://www.dgi.inpe.br/CDSR/>) e utilizadas principalmente para delimitação manual de vias alternativas às estradas gerais, para facilitar o processo de validação dos pontos, em campo. A resolução espacial de 6,5 m deste sensor

permitiu tal procedimento, anteriormente inviabilizado na resolução de 30 m das imagens OLI.

Ainda em ambiente SIG, a partir de todos estes dados, foi possível delimitar manualmente quatro quadrantes sobre o município de Cândido Mota-SP. Os quadrantes serviram para aperfeiçoar a logística da equipe de campo entre os pixels posicionados relativamente próximos e interligados, por fáceis vias secundárias de acesso. Estes quadrantes teoricamente englobavam uma quantidade aceitável de pixels a serem validados por dia.

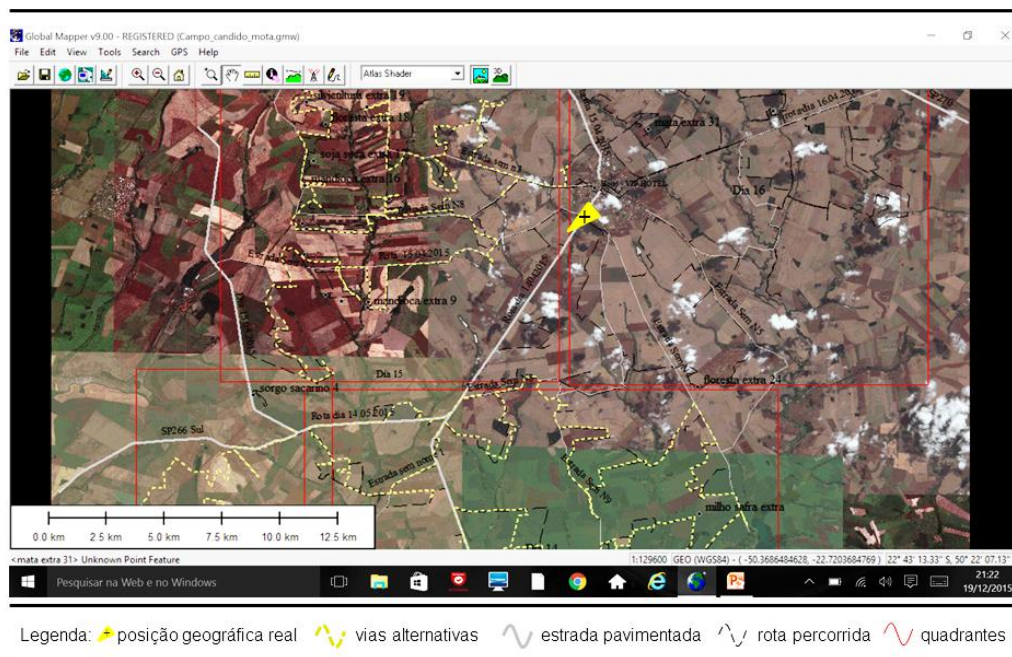
Todos estes dados em formato vetorial e matricial foram salvos em um único projeto em projeção UTM *datum* WGS84 no programa *Global Mapper*. Uma ficha de campo foi montada para facilitar a organização da coleta de dados no campo. Além disso, cartas imagem na escala 1:10.000 foram impressas, para servirem de dados auxiliares se algo impedisse o navegador, em utilizar o computador.

3.2.2.2. Atividades desenvolvidas em campo

O projeto salvo foi então utilizado em campo. O GPS *Garmin* foi acoplado ao computador por porta USB. No programa *Global Mapper*, pela aba *Track load* foi possível analisar, em tempo real, a posição geográfica/quadrante que a equipe estava em relação aos pontos a serem validados, em cada um dos dias de campo e ainda, sobre a imagem de navegação.

A Figura 3.13 ilustra como o navegador executava a tarefa de procura e validação dos pixels no campo, em tempo real.

Figura 3.13 – Ilustração do visor de navegação sobre imagens *Google Earth*, em ambiente *SIG Global Mapper* para validação de pixels no campo.



Em cada pixel validado, fotos foram retiradas e também, informação sobre hora da validação, cultura plantada, altura da planta, idade estimada, cultura plantada anteriormente (prazo de um a dois meses anteriores) e número das fotos foram coletadas na ficha de campo.

Estes dados então foram utilizados em laboratório, para execução de três diferentes atividades:

- Montagem do protocolo de classificação visual;
- Avaliação da classificação visual;
- Avaliação da classificação automática.

3.2.2.3. Montagem do protocolo de classificação visual de pixels

O protocolo de classificação visual desta pesquisa consistiu em um roteiro de classificação visual que deveria ser seguido pelo intérprete para classificação de pixels em imagens multitemporais OLI e também de uma interface de

classificação. A interface de classificação foi testada em dois diferentes ambientes: (1) laboratório em SIG e (2) laboratório *Web*.

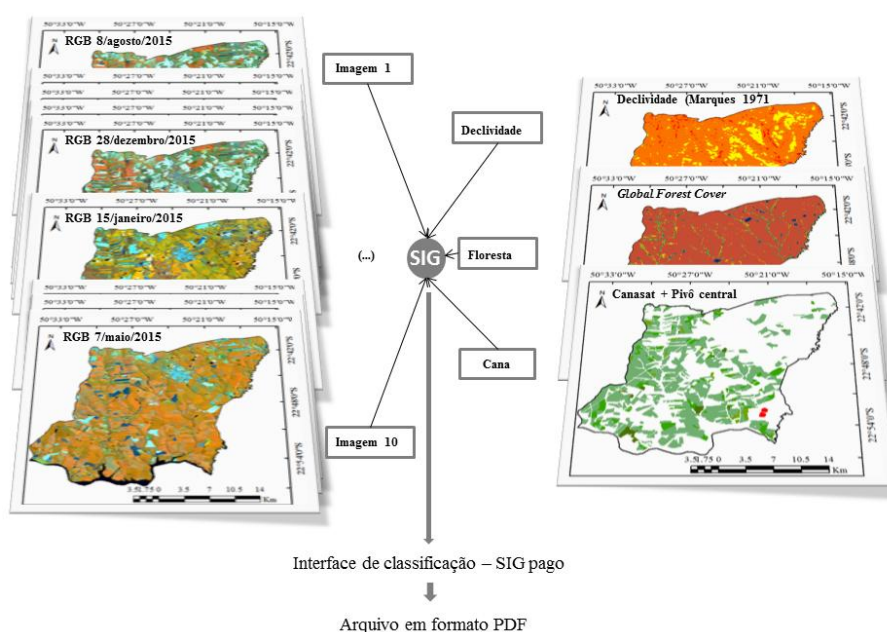
A principal diferença entre eles é que o 1 é licenciado mensalmente por meio de pagamento, e o 2 é grátis.

O protocolo de classificação foi montado, para que no futuro, o método utilizado pelo intérprete treinado, responsável pela classificação visual dos pixels no ano safra 2014/2015 possa ser replicado da mesma forma, por outro intérprete qualquer, que não seja *expert* em classificação de imagens de sensoriamento remoto.

3.2.2.3.1. Interface de classificação 1: Laboratório SIG

Para isso, cada um dos pixels validados no município de Cândido Mota no mês de Abril possuía uma história em relação ao tempo, que poderia ser descrita por diferentes tipos de dados. Os dados são produtos, de diferentes formatos computacionais, que já foram coletados ou gerados em determinado dia, destes pixels. Um total de seis tipos de dados foi utilizado para montagem do protocolo de classificação (Figura 3.14).

Figura 3.14 – Principais dados utilizados para montagem do protocolo de classificação visual 1.



No ambiente SIG ArcGIS, as visitas do sensor OLI da Tabela(Figura 3.7, página 87) serviram de dados espectro-temporais para se observar a variação do comportamento do matiz das culturas, no tempo.

Estas imagens, em formato R-b5 G-b6 B-b4, foram recortadas para a área do município de Cândido Mota. Demais imagens, das demais visitas OLI sobre o estado de São Paulo, que tiveram mais de 50 % da área recoberta com nuvem foram visualmente descartadas.

O CET das culturas foi avaliado sobre as imagens multitemporais, de acordo com os respectivos calendários agrícolas da cana-de-açúcar, milho e soja, descritos nas seções 2.4.3.4 e 2.4.3.5 (páginas 36-39), em cada um dos períodos da estimativa 2014/2015.

Outro dado gerado e fornecido pela USGS, juntamente com os dados radiométricos de cada banda do sensor OLI é o dado de cobertura de nuvens (*cf mask*). Este dado também foi indexado ao SIG, mas não foi útil para classificação dos pixels. Por este motivo, este tipo de dado foi descartado na interface de classificação.

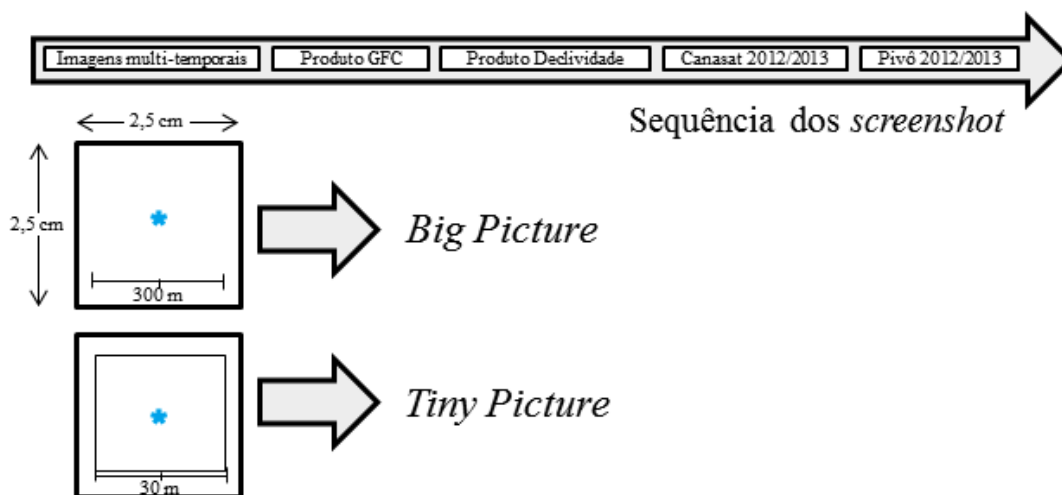
Dados em formato matricial de declividade (TOPODATA) e GFC, também foram inseridos a este conjunto de informações. O objetivo destes dados foi fornecer ao intérprete condições possíveis para separação de pixels com floresta natural de pixels agrícolas e também de visualizar se os pixels selecionados estavam contidos em faixa aceitável de declividade para mecanização (desenvolvimento da atividade agrícola) ou não. O produto GFC e a declividade do TOPODATA foram utilizados como dados excludentes.

Os dados vetoriais do CANASAT de 2013 e de Pivô central serviram também como dado auxiliar para constatação de usos e ocupação do solo, aplicados anteriormente aos pixels selecionados. Estes dados foram meramente informativos, e foram válidos apenas para mostrar com foi a aptidão agrícola do pixel, no momento da formação de tal produto.

Estes cinco dados (matriciais e vetoriais) foram utilizados como camadas independentes no programa ArcGIS (*Data frames*). Para cada um dos 200 pixels validados uma imagem de tela (*screenshot*) foi automaticamente capturada sobre estes quatros dados. Este método foi desenvolvido em programa *Python*, pelo *script 2* do anexo 2.

Cada *screenshot* foi delimitado no tamanho de 2,5 x 2,5 cm (em escala, 1:30.000 chamado de *Big Picture*, e 1:1.000 chamado de *Tiny Picture*) e foram inseridos uma ao lado do outro na seguinte ordem: imagens multitemporais em ordem cronológica, produto GFC, produto declividade, CANASAT e Pivô central. A imagem do *Tiny Picture* foi inserida abaixo da imagem do *Big Picture* (Figura 3.15).

Figura 3.15 – Quadrados de 2,5 x 2,5 cm que mostram as imagens OLI e demais produtos usados no protocolo de classificação visual 1 em duas diferentes escaladas de detalhamento, em big Picture e tiny Picture.



Para cada um dos 200 pixels, todos os *screenshots* foram exportados em um único arquivo em formato PDF. Desta forma, 200 diferentes arquivos PDFs foram exportados.

3.2.2.3.2. Interface de classificação 2: Laboratório Web

No presente trabalho desenvolvemos também um protótipo de interface de classificação de pixels em laboratório Web. O laboratório, nesta presente

pesquisa Web é subentendido como um modelo de SIG moderno, que possui um armazenamento fixo e local dos dados, é de livre acesso, vinculado a um nome de usuário e senha e que pode ser acessado, em qualquer computador via rede (*Web*).

Exemplos de SIG modernos são as interface elaborada pela equipe do projeto Canasat e Terraclass, para validação de áreas com cana-de-açúcar em imagens de sensoriamento remoto (ADAMI *et al.*, 2012) e para validação das classes de uso e ocupação do solo no Cerrado Brasileiro (MMA, 2015).

Ambas as interfaces, TERRACCLASS e CANASAT possuem dados geográficos georreferenciados em um sistema de coordenadas e Datum conhecidos. Todos os dados estão inseridos em um servidor fixo, é necessário o cadastramento para o seu uso; e assim pode ser acessado de qualquer computador via *links* <http://mapas.cnpm.embrapa.br/somabrasil/webgis.html> e <https://www.dsr.inpe.br/laf/validamapacanal> para o TERRACCLASS e CANASAT, respectivamente.

Ao observar a vantagem do uso deste SIG em ambiente de rede e o possível uso destes SIGs modernos no presente trabalho, todos os dados utilizados na interface de classificação formulada em SIG ArcGIS, foram utilizados neste protótipo.

Além dos dados já descritos (cinco no total), outros dados como a série temporal do NDVI, imagens *Google Earth* e imagens *RapidEye* foram também indexadas a este protótipo.

Estes novos dados referem-se: à série temporal do NDVI obtida do ano de 2007 até o ano de 2015, em 250 metros de resolução espacial e 16 dias de resolução temporal, para área de estudo; à imagens de altíssima resolução espacial de satélites como *GeoEye* e *IKonos* usadas em camadas no *Google Earth*; e imagens *RapidEye* em composição colorida salsa cor R-b4 B-b5 B-b3.

O entendimento do comportamento da série temporal é um fomento auxiliar para analisar a aptidão agrícola da cultura, no talhão agrícola em que o pixel

está inserido, no tempo (ADAMI *et al.*, 2012). O link <http://www.dsr.inpe.br/laf/> mostra o exemplo do SIG que utiliza a camada de imagens do *Google Earth* e o comportamento das curvas temporais do NDVI, de acordo com a posição geográfica selecionada pelo usuário, por meio de cliques.

Este protótipo foi desenhado posteriormente ao processo de classificação visual dos 200 pixels validados em campo, assim ele foi avaliado, posterior ao Protocolo 1.

3.2.2.4. Avaliação da classificação visual

A avaliação da classificação visual foi realizada tendo por base os dados obtidos no trabalho de campo 2. Os 200 pixels coletados em campo foram considerados a classificação de referência, e os dados classificados na imagem foram considerados a classificação do usuário.

As duas interfaces de classificação anteriormente descritas foram utilizadas para classificação visual dos pixels. Ademais, o roteiro de classificação descrito na seção de resultados (seção 4.3.3, página 154) foi aplicado para classificação dos pixels. As imagens utilizadas para a classificação dos pixels em laboratório SIG e *Web* foram datada em 21 de abril de 2015 e 07 de maio de 2015, respectivamente. Estas datas corresponderam ao 3º período da estimativa agrícola, do ano safra 2014/2015 (Figura 3.7, página 87). Todos os demais recursos disponíveis em cada uma das interfaces de classificação serviram para o roteiro de classificação dos pixels.

A diferença entre a data do trabalho de campo e a data da imagem utilizada para classificação dos pixels com a cultura verde e em pé foi de 8 e 24 dias para a interface em SIG e *Web*, respectivamente. Desta forma, a classe indexada ao pixel validado no campo ou classificado visualmente variava entre quatro possibilidades (cana-de-açúcar, milho, soja e outros).

Nenhuma regra foi aplicada para atender a diferença de dias entre as datas de validação em campo e classificação em laboratório SIG. No entanto, em laboratório *Web*, aqueles pixels que foram validados como milho no campo e

que tinham mais de 30 cm de altura, foram considerados como pixels de milho na imagem de 07/maio/2015.

No processo de classificação em laboratório, oito diferentes intérpretes foram testados quando à interpretação visual dos 200 pixels. Deste conjunto de intérpretes, quatro tinham conhecimento em classificação de culturas agrícolas em imagens multitemporais *Landsat-like* (Chamado de G1), e os outros quatro não (chamado de G2). A formação acadêmica de cada um dos intérpretes do G2 foi variável (Tabela 3.12).

Tabela 3.12 – Formação acadêmica dos intérpretes visuais do G2

Intérprete	Formação		
	Graduação	Escolaridade	Área de conhecimento
1	Engenheiro Agrícola	Doutorando	SR + Agricultura
2	Engenheiro Agrícola	Doutorando	SR + Agricultura
3	Geógrafo	Mestre	SR + Queimadas
4	Engenheiro Florestal	Doutorando	SR + Floresta

Uma orientação de apenas duas horas foi fornecida a cada um destes quatro intérpretes do G2, para que se fosse compreendido todo o procedimento de classificação visual.

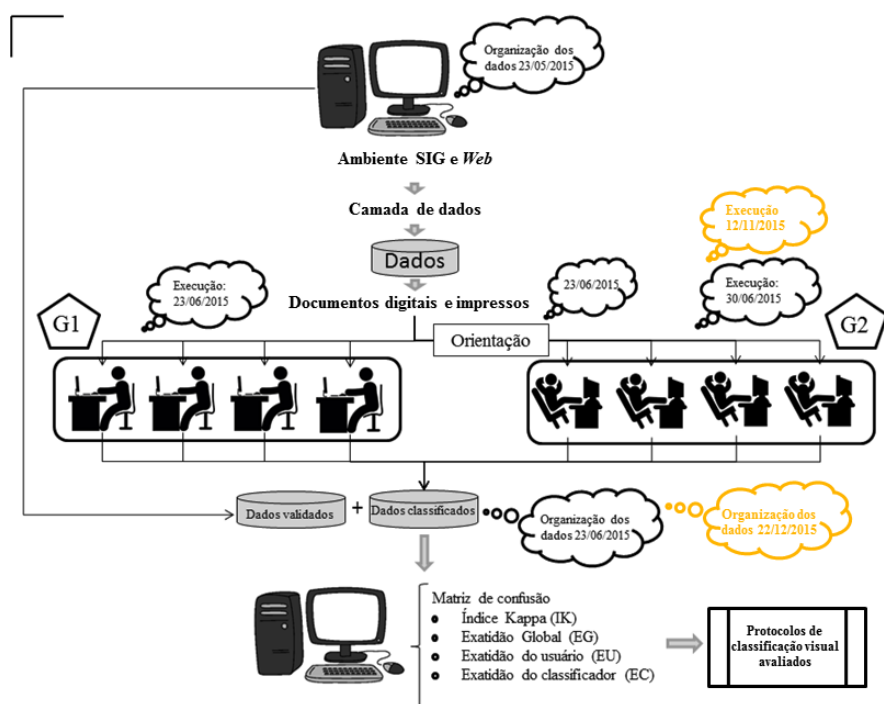
Cada um dos intérpretes de G1 e G2 recebeu o protocolo de classificação impresso em folha formato A4 (ver seção 4.3.3, página 1154), uma ficha para classificação de cada pixel e um documento em formato digital, em PDF, contendo os 200 pixels a serem classificados (interface de classificação em laboratório SIG). O mesmo G2 classificou os mesmos 200 pixels seguindo a interface de classificação 2, três meses depois de realizar a classificação na interface de classificação 1.

A Figura 3.16 mostra as datas de realização da classificação, dos procedimentos operacionais da pré-classificação e como ocorreu a classificação dos pixels, pelos dois grupos descritos.

Cada um dos intérpretes preencheu a ficha de classificação visual, e apenas uma classe foi interpretada para cada um dos 200 pixels. Os dados de

referência foram cruzados com os dados classificados pelo usuário, e então, oito diferentes matrizes de erro foram geradas. Cada uma destas matrizes de erro resultou em quatro diferentes parâmetros: Índice Kappa, Exatidão Global, Exatidão do Usuário e Exatidão do Classificador.

Figura 3.16 – Fluxograma das atividades desenvolvidas no processo de avaliação da classificação visual dos 200 pixels validados no campo e classificados pelos intérpretes do G1 e G2 utilizando o protocolo 1. A classificação visual decorrente do uso do protocolo 2 (destaque em laranja) foi avaliado apenas pelos intérpretes do G2



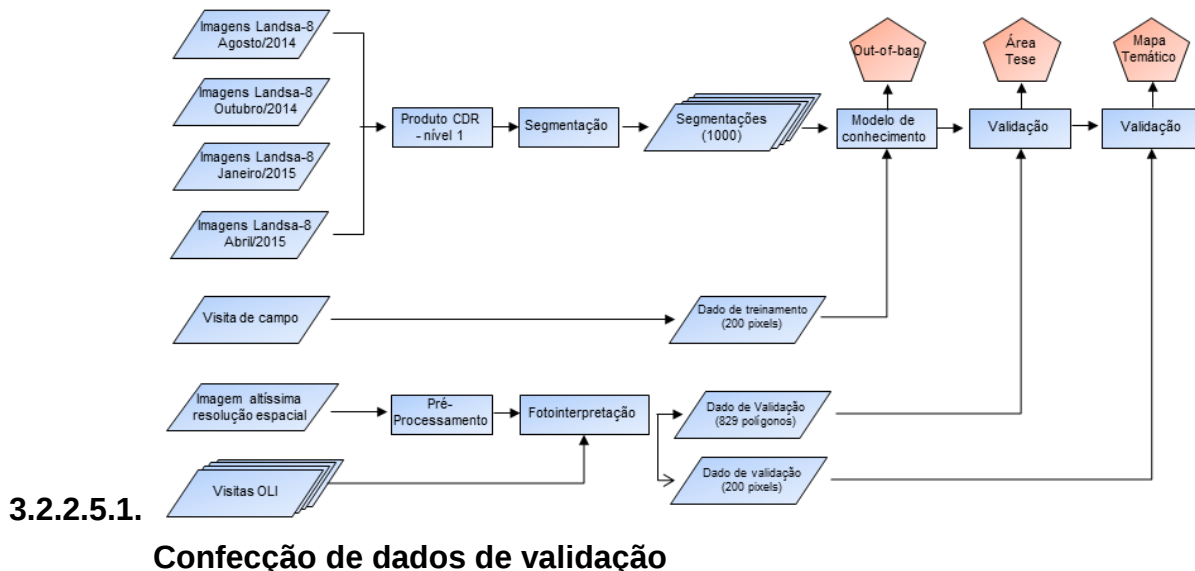
A matriz de erro teve por objetivo mostrar as dúvidas geradas pelos intérpretes na classificação visual dos pixels, em cada interface de classificação visual. Aqueles pixels com maior proporção de erros foram identificados e mostrados sobre recortes das imagens multitemporais OLI disponíveis nas interfaces, em formato *tiny* e *big picture*.

3.2.2.5. Classificação automática

Da mesma forma utilizada para a classificação visual, todos estes 200 pixels de referência também foram utilizados para gerar um modelo de conhecimento para classificação automática de pixels. O protocolo de classificação

automático foi definido pelo fluxograma da Figura 3.17. Cada um dos subitens ilustrados no fluxograma será descrito, subsequentemente.

Figura 3.17 – Fluxograma da metodologia aplicada na classificação automática dos pixels nas imagens multitemporais OLI.

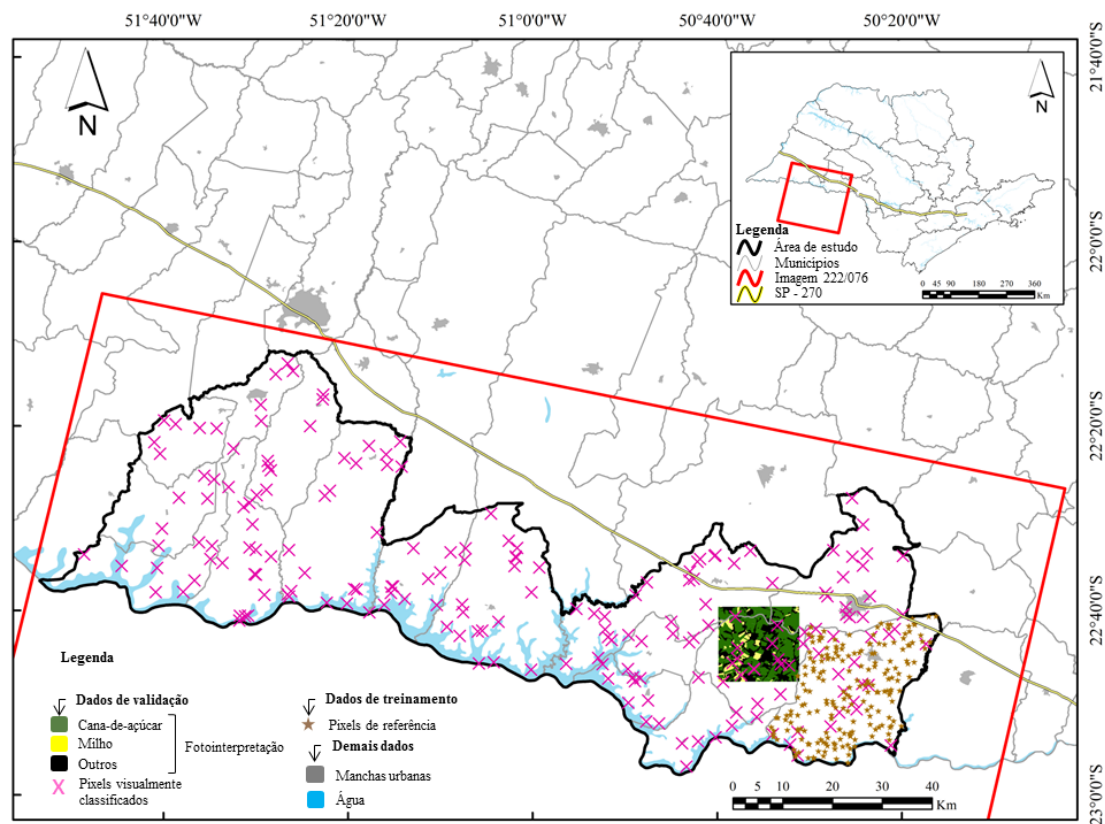


Para que fosse possível avaliar a qualidade da segmentação e criar amostras adicionais de dados de referência para validação independente, uma área teste de 15 km x 15 km localizada ao lado esquerdo do município de Cândido Mota e 200 pixels aleatórios, foram obtidos sobre a imagem 222/076 (Figura 3.18).

Na área teste, polígonos no formato dos talhões agrícolas foram manualmente digitalizados e visualmente indexados em uma das três classes possíveis (cana-de-açúcar, milho e outros) constatadas no trabalho de campo 2. A área teste ocupou uma área total de 225 km² (22.500 ha).

Os talhões agrícolas foram delineados sobre as imagens OLI e imagens *RapidEye* descritas na Tabela 3.11, pg. 98. O protocolo de classificação montado neste presente estudo, para classificação de cana-de-açúcar, milho e soja, foi utilizado no processo de fotointerpretação dos talhões e para classificação visual dos 200 pixels de validação, contidos na área de estudo 2. O resumo dos dados utilizados na presente pesquisa está expresso na Tabela 3.13.

Figura 3.18 – Informação da localização dos dados de referência (treinamento) e validação (área teste e pixels aleatórios, foto-interpretados).



A área teste foi delineada manualmente na tela do computador sobre imagens multi-temporais OLI e RapidEye; e os pixels de validação (cruzes em rosa) foram classificados visualmente obedecendo o roteiro de classificação da seção dos resultados.

Tabela 3.13 – Sumário descritivo das informações dos dados de referência e validação coletados na área de estudo 2 (município de Cândido Mota) e na área de estudo 3 (quinze municípios do estado de São Paulo recobertos pela imagem OLI 222/076).

Dados de referência	Tipo de dado	Interpretação	Cana-de-açúcar	Milho	Outros	Total
Dados de treinamento	Pontos	Visita de campo	53	77	70	200
Dados de validação (pixels)	Pontos	Visual	48	46	106	200
Dados de validação (polígonos)	Polígonos	Visual	401	100	328	829
Total			502	223	504	1229

Vale ressaltar que os dados de validação não foram utilizados como dados de calibração/treinamento no modelo de classificação, eles foram utilizados

apenas como dados independentes do treinamento, para avaliar o mapa temático final produzido.

3.2.2.5.2. Segmentação das imagens multitemporais e extração de atributos

No presente trabalho, nós utilizamos o Segmentador Multi-resolução (SM-R) implementado no programa eCognition® Developer 8.7 para segmentação das imagens. Este software é extensivamente utilizado em vários estudos em OBIA (VIERA *et al.*, 2012). O SM-R é uma sequência de métodos de agrupamento de regiões que se inicia a partir dos pixels. Estes segmentos são interativamente agrupados com seus vizinhos até o limiar de heterogeneidade definido inicialmente pelo usuário (SCHULTZ *et al.*, 2015a).

Foram utilizadas quatro imagens multitemporais provenientes da cena 222/076 no processo de segmentação: 08/ago/2014, 27/out/2014, 15/jan/2015, 21/abr/2015. Estas imagens tiveram menor proporção de cobertura de nuvens entre as demais imagens desta cena, no ano safra. A cobertura de nuvem foi analisada automaticamente por *script* desenvolvido em linguagem *Python* (Anexo 3). Apenas as bandas b4, b5 e b6 de cada uma destas imagens foram utilizadas no processo de segmentação. Estas três bandas possuem menor correlação dos dados da imagem (Schultz *et al.*, 2015a).

Em vez de ajustarmos manualmente os parâmetros de segmentação no SMR, um conjunto de segmentações foi obtido pela variação dos parâmetros. O fator de escala (fe) variou de 25 a 250 (variação de 5), a forma (fm) e compacidade (cp) ambos variaram os seus pesos de 0,1 a 0,9 (variação de 0,2). A combinação destes parâmetros gerou 1150 diferentes segmentações, que foram exportadas em formato *shapefile* e *GeoTIFFs*.

Mais tarde, para cada segmento e segmentação, 74 diferentes atributos foram calculados e exportados em formato de tabela:

- (1) O valor médio de cada banda para as quatro datas (6*4 – 24 variáveis);

- (2) Máxima diferença e brilho de todas as bandas (2 variáveis);
- (3) A homogeneidade do GLCM (Gray-Level Co-Occurrence Matrix), contraste do GLCM, entropia do GLCM, o 2nd momento angular do GLCM, a média do GLCM, o desvio padrão do GLCM, a correlação em todas as direções do GLCM (8 variáveis);
- (4) As frações sombra, solo e vegetação do MLME para cada data ($3 \times 4 = 12$ variáveis); e
- (5) Os índices de vegetação: IVverde, GNDVI, NDVI, SAVI, NDVI7, NDVI7, NDSVI para cada data ($7 \times 4 = 28$ variáveis).

Todos os 74 atributos foram extraídos para cada segmento e usados na subsequente análise da classificação.

3.2.2.5.3. Formação do modelo de conhecimento e classificação

O algoritmo *Random Forest* (RF) (BREIMAN, 2001) foi utilizado para gerar um modelo de classificação supervisionado. RF é um popular algoritmo que forma um conjunto de aprendizado composto por várias árvores de decisão (AD). Este algoritmo tem sido muito utilizado nestes últimos anos, na classificação de diferentes culturas agrícolas em imagens *Landsat-like* (LONG *et al.*, 2013; SCHULTZ *et al.*, 2015b).

A principal vantagem do uso do RF é que a construção de cada AD é baseada na amostragem aleatória com reposição do mesmo tamanho do conjunto dos dados de entrada. As árvores geradas são validadas pelas amostras não utilizadas para suas construções *out-of-bag* (OOB) e os votos da maioria das árvores são utilizados para validar o modelo.

Outros benefícios são oferecidos pelo algoritmo RF, como seleção aleatória de atributos, um produto final robusto em relação a variável de entrada e informações sobre a importância de cada atributo na construção do modelo de classificação. Vários estudos relataram que o uso de um número reduzido/otimizado de atributos melhorou a exatidão do mapeamento

(IMMITZER *et al.*, 2012). Assim, o RF permite a seleção de atributos para diminuir problemas como *overfitting* (SCHULTZ *et al.*, dados não publicados).

Os dois parâmetros que devem ser informados para que o modelo RF funcione são:

- *mtry* (número de variáveis de entrada), a configuração padrão foi mantida (raiz quadrada do número de variáveis de entrada).

- *ntree* (número de AD) foi fixada em 1.000.

Para cada um dos modelos RF gerado (total = 1150), foi aplicado um recurso para seleção de atributos, com base nos valores de importância de cada atributo “*Mean Decrease in Accuracy*”, descrito por Guyon e Elisseeff, (2003).

Este procedimento foi iniciado a partir do modelo com todos os atributos selecionados, calculando a importância de cada um deles e eliminando os menos importantes. A seleção de atributos foi replicada até que restou apenas um atributo. O modelo baseado no menor número de variáveis de entrada e que atingiu o 97,5th percentil de exatidão da classificação, entre todos os modelos testados, foi escolhido como modelo final (SCHULTZ *et al.*, 2015a). Nota-se que a seleção de atributos foi aplicada de forma independente em cada uma das 1025 segmentações.

Após o processo de seleção de atributos, cada uma das 1.000 árvores gerou dois resultados de classes para cada segmento. O primeiro resultado refere-se à classe mais comum indexada pelas 1.000 ADs, e a segunda, refere-se à segunda classe mais comum. Para a primeira e segunda classe, o número de menções foi contado.

O resultado da confiabilidade da classificação foi a diferença entre número contado pela classe 1 e pela classe 2. Valores desta diferença próximos de “1” (todas as ADs apontam para a mesma classe) implicam em alta confiabilidade; valores próximos de “0” indicam incerteza de classificação (as duas classes foram igualmente classificadas por muitas ADs).

Todo procedimento de construção do modelo de conhecimento e classificação foi realizado no software estatístico livre R 3.1.3, tendo como principais algoritmos auxiliares *randomForest* (LIAW e WIENER, 2002) *raster* (HIJMANS, 2014) e *caret* (KUHNS *et al.*, 2014).

3.2.2.5.4. Desempenho do modelo de classificação automática

Os modelos foram avaliados de duas formas:

- Baseado pelo método de amostragem (OOB), a partir dos pixels de referência (Figura 3.18, pg. 112).
- Baseado na fotointerpretação de talhões de referência e pixels aleatórios amostrais, como dados independentes de validação (Figura 3.18, pg. 112).
- Baseados nos resultados de área ocupada com a cultura verde na imagem.

Os erros gerados pela OOB foram usados na avaliação dos 1150 modelos de classificação. Para cada uma das 1150 segmentações uma matriz de confusão composta por 200 pixels foi calculada, pelos mesmos parâmetros utilizados para avaliar a classificação visual (IK, EG, EU e EP).

A exatidão do mapeamento temático foi realizada de duas formas: (a) pixel a pixel pela sobreposição entre o mapa temático final gerado para área teste com o mapa foto-interpretado e (2) amostral, pela comparação entre os 200 pixels visualmente classificados pelo modelo OBIA com estes mesmos pixels classificados visualmente.

4. RESULTADOS E DISCUSSÃO

Os resultados da presente pesquisa foram apresentados da seguinte forma:

1. Para responder os objetivos 1 e 2:

(a) Resultados para as estimativas agrícolas para cana-de-açúcar, milho e soja em quatro períodos do ano safra 2014/2015 para o estado de São Paulo, via painel amostral MoBARS.

(b) Resultados alcançados pelo uso da ApEq, em imagens OLI.

2. Para responder os objetivos 3 a 8:

(c) Protocolos desenvolvidos para classificação visual de pixels de cana-de-açúcar, soja, milho e outros, em imagens multitemporais OLI e dados auxiliares, no laboratório;

(d) Resultados alcançados pelo uso dos protocolos desenvolvidos nesta presente pesquisa, por intérpretes selecionados para classificação de 200 pixels validados no campo;

(e) Resultados alcançados pelo uso de OBIA/*Random Forest* na confecção de 1150 modelos de classificação automática com os 200 pixels selecionados em campo.

3. Para responder o objetivo 9:

(f) Obtenção de um mapa temático classificado via algoritmos de classificação e validado por dados independentes dos dados de treinamento do modelo.

4.1. Estimativa agrícola de cana-de-açúcar, milho e soja em São Paulo, no ano safra 2014/2015

Os resultados da estimativa agrícola de cana-de-açúcar, milho e soja no estado de São Paulo para o ano safra 2014/2015 foi apresentado em dois diferentes tópicos:

- (1) Serão apresentados os valores dos coeficientes de variação resultantes pela estimativa da área ocupada com as culturas de interesse, em São Paulo.
- (2) E também, o resultado da área ocupada com cana-de-açúcar, milho e soja em São Paulo e a comparação destes dados com os resultados encontrados pela Pesquisa Agrícola Municipal (PAM) informada pelo órgão oficial responsável pelas estimativas agrícolas nacionais, IBGE.

4.1.1. Coeficientes de variação para cana-de-açúcar, milho e soja

O resultado de área ocupada com cana-de-açúcar, milho e soja, em São Paulo foi obtido em cada um dos períodos da amostragem probabilística para o ano safra 2014/2015. Para chegar a tal resultado foi necessária a classificação visual de 5.000 pixels aleatórios em cada um dos quatro períodos da estimativa (Tabela 4.1).

Na Tabela 4.1 é importante ressaltar que não foi possível classificar os 5.000 pixels nos quatro períodos da estimativa, isto é, pixels com nuvem foram identificados em maior número na seguinte sequência: Período 1 (767) > Período 3 (78) > Período 2 (47) > Período 4 (8). Em média, 1,75 % dos pixels foram considerados pixels de nuvem (Tabela2).

Tabela 4.1 – Resumo do número de pixels classificados como CN (cana-de-açúcar), ML (milho), SJ (soja), OT (outros) e NV (nuvem) nas 15 mesorregiões do estado de São Paulo (estratos) e nos quatro períodos da estimativa agrícola do ano safra 2014/2015, dado em ordem decrescente do total de pixels sorteados por estrato.

Estrato	Área (Ha)	Total	Período 1					Período 2					Período 3					Período 4				
			CN	ML	SJ	OT	NV	CN	ML	SJ	OT	NV	CN	ML	SJ	OT	NV	CN	ML	SJ	OT	NV
Estrato 2	2.753.223	734	194	0	6	490	44	233	0	31	470	0	298	9	1	426	0	285	11	0	438	0
Estrato 11	2.017.216,4	645	13	5	43	577	7	13	9	70	548	5	28	20	5	571	21	20	43	0	582	0
Estrato 1	2.939.470	609	119	0	1	414	75	129	2	3	474	1	183	3	0	419	4	178	5	0	426	0
Estrato 6	1.271.021	564	98	0	26	417	23	103	5	107	322	27	123	31	2	390	18	131	83	0	349	1
Estrato 8	2.672.325,7	503	84	0	13	396	10	84	1	12	401	5	99	2	0	385	17	115	3	0	385	0
Estrato 10	2.403.430	438	56	1	1	372	8	61	0	7	370	0	79	3	0	356	0	87	7	0	344	0
Estrato 3	1.676.321,1	340	51	0	3	259	27	103	0	8	228	1	110	6	0	223	1	105	9	0	225	1
Estrato 5	1.422.612,8	337	53	3	3	273	5	36	3	9	287	2	48	4	0	285	0	55	9	0	273	0
Estrato 12	1.230.960	208	5	3	2	188	10	3	0	3	201	1	6	0	0	197	5	3	2	0	203	0
Estrato 4	945.529,9	204	49	0	1	148	6	78	0	2	124	0	80	0	0	124	0	78	0	0	125	1
Estrato 9	904.563	189	34	3	1	151	0	40	1	1	147	0	56	0	0	126	7	47	1	0	141	0
Estrato 7	717.017,1	94	11	0	0	76	7	8	0	2	84	0	17	0	0	77	0	11	0	0	83	0
Estrato 13	1.617.994,7	91	0	0	0	91	0	0	0	0	91	0	1	0	0	90	0	0	0	0	91	0
Estrato 14	929.819,2	22	0	0	0	20	2	0	0	0	21	2	0	0	0	22	2	0	0	0	22	2
Estrato 15	1.320.089,3	22	0	0	0	19	3	0	0	0	19	3	0	0	0	19	3	0	0	0	19	3
Estado SP	24821593,2	5000	767	15	100	3891	227	891	21	255	3787	47	1128	78	8	3710	78	1115	173	0	3706	8

Onde: nov/dez (Período 1), jan/fev (Período 2), mar/abr (Período 3) e mai/jun (Período 4) Estrato 1 - São José do Rio Preto, Estrato 2 - Ribeirão Preto, Estrato 3 - Araçatuba, Estrato 4 - Araraquara, Estrato 5 - Campinas, Estrato 6 - Assis, Estrato 7 - Marília, Estrato 8 - Bauru, Estrato 9 - Piracicaba, Estrato 10 - Presidente Prudente, Estrato 11 - Itapetininga, Estrato 12 - Macro Metropolitana Paulista, Estrato 13 - Vale do Paraíba, Estrato 14 - Metropolitana de São Paulo, Estrato 15 - Litoral Paulista; CN – cana-de-açúcar, ML – milho, SJ – soja, OT – Outros, NV – nuvem.

Demais pixels foram todos classificados entre as quatro classes possíveis, que seguem: cana-de-açúcar (CN), milho (ML), soja (SJ) e outros (OT) (Tabela1). Se forem somados todos os pixels classificados entre estas três culturas de interesse, nos quatro períodos, um maior número de pixels foi classificado como cana-de-açúcar (19,51 %), seguido de soja (1,82 %) e milho (1,44 %) (Tabela 4.2).

Percebe-se aqui, que a percentagem de pixels classificados com soja e milho foi muito semelhante ao número de pixels considerados nuvens (1,75 %). Isto mostra a dependência do uso do painel amostral em vez do mapa, para estimativa agrícola. No mapa é necessário que todos os pixels sejam classificados e a nuvem impede que toda área seja mapeada. Este problema não é constatado via painel amostral.

No painel amostral utilizado nesta presente pesquisa, o resultado do número de pixels classificados como nuvem teve influência direta apenas nos valores finais do coeficiente de variação da estimativa para cada um dos quatro períodos (Tabela 4.3). No entanto, mesmo com 1,75 % dos pixels classificados e considerados nuvem, o CV% da estimativa das culturas nos quatro períodos foi, em média, igual a 5,51 %. Este CV% foi inferior aos encontrados por Adami *et al.* (2005) e Adami *et al.* (2010). Estes referidos autores conseguiram CV% iguais a 6,6 % ao utilizar um painel amostral semelhante ao PREVS e CV% igual 6,98 % ao fazer uso de um painel amostral baseado no método Geosafras, para estimativa de soja.

Luiz *et al.* (2012) encontraram um CV% igual a 3,4 % para estimativa de soja no Rio Grande do Sul utilizando a abordagem Geosafras. Se compararmos o resultado encontrado por este autor com o resultado médio do CV% encontrado neste presente trabalho, percebemos que os dados de Luiz *et al.* (2012) possuíam menor variação estatística. No entanto, no presente trabalho, foram avaliadas três culturas consecutivamente, e se analisarmos o CV% da cana-de-açúcar (Tabela 4.3), cultura de maior expressividade em área no estado, vemos que em média, o CV% foi inferior (2,4 %) ao encontrado por Luiz *et al.* (2012).

Tabela 4.2 – Resumo do número de pixels classificados visualmente como cana-de-açúcar, milho, soja, outros e nuvem no estado de São Paulo, para os quatro períodos da estimativa agrícola do ano safra 2014/2015.

Pixel/Classe	Período 1	Período 2	Período 3	Período 4	Total	%
Cana-de-açúcar	767	891	1128	1115	3901	19,51
Milho	15	21	78	173	287	1,44
Soja	100	256	8	0	364	1,82
Outros	3891	3789	3713	3709	15102	75,51
Nuvem	227	43	73	3	346	1,73
Total	5000	5000	5000	5000	20000	100,00
Pixel sem nuvem	4773	4957	4927	4997	19654	98,27

Tabela 4.3 – Coeficientes de variação calculados pelo painel amostral MoBARS para as culturas agrícolas da cana-de-açúcar, milho e soja, nos quatro períodos da estimativa agrícola do ano safra 2014/2015

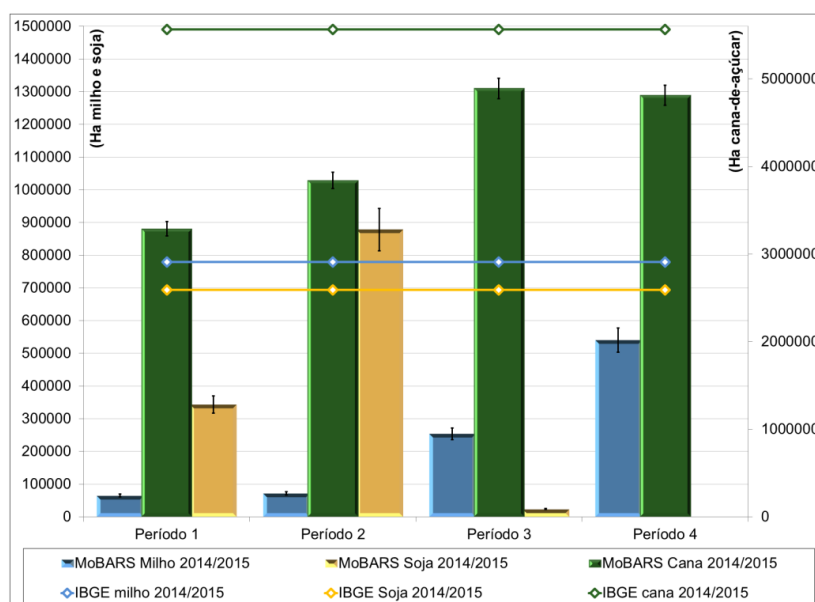
Cultura / Coeficiente de Variação (%)	Período 1	Período 2	Período 3	Período 4	Médio
Cana	2,44	2,39	2,40	2,38	2,40
Milho	6,97	6,85	6,87	6,82	6,88
Soja	7,49	7,36	7,38	7,34	7,39
Médio	5,64	5,53	5,55	5,51	5,56

O método proposto por Luiz *et al.* (2015) também chamado de painel amostral MoBARS mostrou superioridade em relação aos demais painéis utilizados em nível regional de estimativa agrícola no Brasil, como o PREVS e Geosafras. A superioridade foi evidenciada devido aos resultados dos CVs% para cada cana-de-açúcar, milho e soja nos quatro períodos avaliados forem menores ou próximos aos resultados já constatados pelos painéis amostrais supra referidos.

4.1.2. Resultados de área ocupada com cana-de-açúcar, milho e soja no estado de São Paulo, para o ano safra 2014/2015

A estimativa agrícola de cana-de-açúcar soja e milho para o ano safra 2014/2015 encontrada pelo painel amostral do método MoBARS foi apresentado na Figura 4.1. Nesta Figura, os dados de cada período foram ilustrados nas barras em cores, e as linhas representaram os dados obtidos pela PAM do IBGE, também para o ano safra 2014/2015.

Figura 4.1 – As barras ilustram a área ocupada com cultura verde de cana-de-açúcar, milho e soja pelo método MoBARS, por período: Período 1 (Nov/Dez), Período 2 (Jan/Fev), Período 3 (Mar/Abr) e Período 4 (Mai/Jun), no ano safra 2014/2015.



As linhas mostram o valor de área plantada com cana-e-açúcar, milho e soja no ano base de 2014, informado pelo IBGE.

Em números divulgados pelo IBGE, o valor de área plantada com cana-de-açúcar, milho e soja em São Paulo para o ano de 2014/2015 esteve próximo de 5,5 milhões, 0,8 e 0,7 milhão de hectares.

Dados estes números, inicialmente percebemos que ocorreu uma superestimativa dos valores de área ocupada com cultura verde de soja dado pelo método MoBARS no Período 2 da estimativa, se comparado ao IBGE (~ 200 mil de hectares). É importante ressaltar que mesmo com o CV(%) da soja para este período, os dados ainda ficam superestimados em (~ 100 mil de hectares).

Neste Período da estimativa, o alto vigor vegetativo da cultura da soja permite que esta cultura possam ser interpretadas nas imagens de 30 metros de resolução espacial (Rizzi e Rudorff, 2005; Sugawara et al., 2008; Schultz et al., 2016), e assim os pixels sejam facilmente identificadas, como soja pelo protocolo de classificação do MoBARS. No entanto, neste mesmo Período, a mandioca também está em alto vigor vegetativo e se assemelha espectralmente com soja (SCHULTZ *et al.*, 2015a). Desta forma, pode ter ocorrido comissão de pixels de mandioca (classe outros) na classe soja ao utilizar o protocolo de classificação.

Ao falar sobre os resultados de área ocupada com milho pelo método MoBARS, se forem somados os valores de área encontrados com esta cultura nos Períodos 1 e 2, os valores de área dados pelo IBGE e MoBARS tornam-se semelhante (~ 15 mil hectares superestimados pelo método MoBARS). No entanto, pelo CV% médio do milho (6,88 %), os dados obtidos pelo IBGE e MoBARS são considerados iguais. É importante ressaltar que o milho possui dois diferentes períodos de plantio (na safra e safrinha) bem identificados, fato que pode justificar a soma destes dois Períodos.

A área ocupada com cana-de-açúcar nos quatro períodos avaliados mostrou comportamento diferenciado para o milho e soja. Em todos os quatro períodos, a área resultante foi superior à metade do valor da área informada pelo IBGE, em 214/2015 (~ 5,5 milhões de hectares) e, além disso, a área ocupada com

cana-de-açúcar aumentou mais de 1 milhão de hectares no Período 3 e 4, se comparado aos Períodos 1 e 2. Este resultado mostra a maior presença, em número de pixels com VFA de cana-de-açúcar nos Períodos 3 e 4, que serão extrapolados em valor de área pelo painel amostral.

Esta afirmação corrobora com a literatura, pois como a cana-de-açúcar é colhida de abril a novembro e a intensificação da colheita ocorre em outubro (AGUIAR *et al.* 2008; RUDORFF *et al.* 2010; AGUIAR *et al.* 2011), tanto no Período 1 quanto no Período 2, os pixels de cana-de-açúcar ainda estarão adquirindo cor de VFA, fato este, que diminui o número de pixels classificados com cana-de-açúcar verde, em pé. Nos meses sequenciais, a cultura da cana-de-açúcar já adquiriu novamente cor de VFA, e assim os pixels serão computados com maior frequência, nos Períodos 3 e 4 da estimativa.

Ao comparar os resultados encontrados nos períodos 3 e 4, do valor em área ocupada com cana-de-açúcar (Figura 4.1), com os dados oficiais ocorreu subestimativa dos valores do MoBARS (~ 500 mil hectares). O menor valor em área ocupada com cana-de-açúcar nos Períodos 3 e 4 em comparação aos dados do IBGE pode estar atrelado à data da colheita da cana-de-açúcar. Como a colheita da cana-de-açúcar inicia-se em abril (RUDORFF *et al.*, 2010), talhões que até então não tinham sido colhidos nos Períodos anteriores, podem nos Períodos 3 e 4, já estarem colhidos. Assim, o intérprete não irá considerar este talhão como classe “cana”, e sim como classe “outros”.

Em resumo, os dados informados pelo painel amostral proposto pelo MoBARS é sobre monitoramento de área ocupada com cana-de-açúcar, milho e soja em pé, em Períodos bimestrais, e como visto, estes dados não se assemelham aos dados do IBGE. O mais plausível, e que poderá ser realizado no próximo ano, é comparar os dados bimestrais do MoBARS adquiridos no ano safra 2014/2015 aos dados encontrados no ano safra 2015/2016. Tomadas de decisão podem ser tiradas a partir de cada um dos quatro períodos da estimativa agrícola proposta (em tempo quase-real), e não mais anualmente como feita, após o lançamento dos dados oficiais.

- 4.2. Assim, a principal vantagem do painel amostral do MoBARS é fornecer dados estatísticos em tempo quase-real para tomada de decisões, sejam elas agrícolas, econômicas, sociais, etc. Estes dados estatísticos servem como suporte ao programa GEOGLAM, para gerenciamento global dos recursos agrícolas.**
- Aparência Equalizada em imagens OLI**

Nesta seção, foram avaliados os mosaicos produzidos ao usar o método da ApEq em cada uma das 24 visitas OLI sobre o estado de São Paulo, no ano safra 2014/2015. Neste caso, a ApEq foi aplicada para manter equalizadas as cenas OLI de diferentes órbitas/pontos, sendo baseadas estas imagens de conversão (IC), na equalização de uma imagem base (IB), e nos valores de reflectância da superfície de pixels de floresta da IB (σ_{ib}) e das IC (σ_{ic}). As imagens foram usadas em composição colorida falsa cor R-b5 G-b6 B-b4 do sensor OLI.

Para a apresentação dos resultados, apenas as informações provenientes do uso da ApEq do mosaico Nº 13 (Tabela 3.7, pg. 87) foram mostrados. É importante frisar que áreas com soja estiveram presentes nestas imagens datadas no mosaico Nº 13. Estas áreas foram utilizadas nos resultados visuais da ApEq e além disso, dados de campo puderam ser utilizados neste período, para comparação entre a cultura validada no campo, com a cultura presente na imagem, em composição colorida falsa cor R-b5 G-b6 B-b4.

Primeiramente, a IB foi apresentada em composição colorida R-b5 G-b6 B-b4 pelo contraste aplicado visualmente por intérprete treinado. Os valores de máximo e de mínimo de cada histograma da IB foram usados para os passos seguintes da ApEq.

Posteriormente, os valores de reflectância da superfície dos pixels de alvos pseudo-invariantes (APIs), das bandas b4, b5 e b6 foram analisados e discutidos. Como o nome diz, os alvos são pseudo-invariantes e a fonte da variação do valor destes pixels deve ser conhecida.

Em sequência, o mosaico de imagens Nº 13 foi aqui apresentado em forma de mosaico do estado de São Paulo. Como falado anteriormente, em cada um dos mosaicos do estado constaram 18 órbitas/pontos (cenas) OLI.

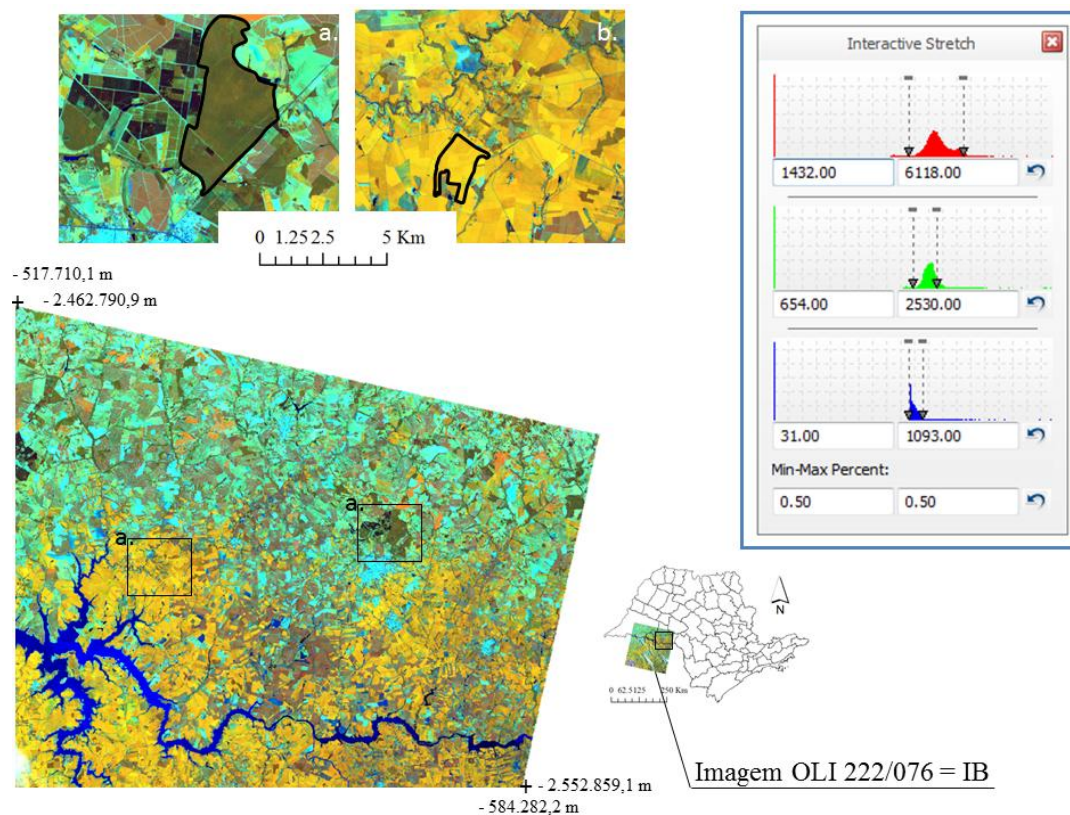
Por fim, recortes de imagens que mostram ilustrativamente áreas de soja (validadas em trabalhos de campo) e de outras culturas agrícolas localizadas em diferentes órbitas pontos foram selecionados do mosaico e comparados visualmente. Eivazi *et al.* (2015) comentaram que a comparação visual de imagens é umas das formas de avaliar o resultado final da equalização de diferentes cenas que compõem um mosaico de imagens Landsat-like.

4.2.1. Imagem Base usada no processo da ApEq

A imagem base localizada na órbita/ponto 222/076, data de 28/janeiro/2014 foi equalizada visualmente para que a floresta e a soja se apresentassem na cor desejada para fotointerpretação, na composição colorida R-b5 G-b6 B-b4. Os resultados estão ilustrados na Figura 4.2. Além disso, ainda nesta Figura, os valores de máximo (máx.) e mínimo (mín.) dos histogramas das bandas b-4, b-5 e b-6 foram salvos e usados para aplicação dos passos seguintes da ApEq.

Como os valores do histograma foram expressos em 2^{16} níveis de cinza (ND = 65.536), os resultados de cada um dos histogramas RGB presentes na Figura 4.2 referem-se aos maiores valores na b-5 (1432 – 6118), valores médios em b-6 (654 – 2530) e baixos valores em b-4 (31 – 1093). Estes valores destacam a maior presença de VFA informada pelos valores do histograma de b-5, áreas de vegetação seca provenientes dos valores de b-6 e áreas de solo exposto e água referente aos histogramas de b-4.

Figura 4.2 – Imagem base OLI localizada na órbita/ponto 222/076 e equalizada visualmente por intérprete treinado.



Em destaque, percebem-se áreas de floresta na cor marrom e áreas com soja na cor amarelo intenso. Estas cores foram resultantes devido ao ajuste dos valores de máximo e mínimo dos histogramas RGB na composição R-b5 G-b6 B-b4.

O contraste definido para IB mantém os alvos presentes nas imagens com cores semelhantes às encontradas por Sanches et al. (2005), Rizzi e Rudorff (2005), Rudorff e Sugawara (2008) e Rudorff et al. (2010). A cor dos alvos foi definida pelos valores de máx. e mín. dos histogramas b-4, b-5 e b-6. Estes valores foram considerados como padrão de uso para aplicação da ApEq.

O uso destes valores padronizados pode ser utilizado em qualquer trabalho de fotointerpretação que deseja manter alvos agrícolas com aceitável distinção em imagens coloridas, na composição colorida R-b5 G-b6 B-b4, para aplicação da ApEq.

4.2.2. Variação dos valores de reflectância da superfície dos pixels pseudo-invariantes

Conforme o método descrito anteriormente foram executados dois passos para que a ApEq fosse aplicada sobre as imagens: (1) Extração dos valores dos APIs e (2) aplicação do fator multiplicativo sobre os histogramas da imagem base (IB).

Os valores dos APIs foram extraídos das 18 imagens de conversão (IC) que compuseram o mosaico Nº 13 do estado de SP. Da mesma forma, os valores dos APIs de IB também foram extraídos. Os resultados dos valores dos APIs da IB e das ICs foram apresentados na Tabela 4.4.

Como os pixels são pseudo-invariantes, ocorreu variação dos valores de reflectância nas bandas utilizadas para ApEq, que neste caso são as bandas b4, b5 e b6 (Tabela 4.4). Esta variação se deu possivelmente pelas diferentes geometrias de visada das imagens (VALERIANO, 2011); ou devido à variação sazonal dos APIs, localizados em diferentes posições do estado de São Paulo (ACHARD e HANSEN, 2013).

Valeriano (2011) comentou que a distribuição da iluminação dada pelo relevo e geometria de iluminação afeta diretamente o brilho das imagens de sensoriamento remoto. A variação resultante pelo fator topográfico sobre os APIs, descrita por Valeriano (2011) foi considerada e minimizada no método da ApEq, no entanto, o ângulo azimutal (φ_i) e declinação do sol (δ) possivelmente causaram variações nos valores de reflectância da superfície dos APIs.

O FM foi utilizado como regra absoluta para minimização destes fatores sobre os alvos das imagens que compuseram o mosaico do estado de São Paulo. Isto é os 18 FMs calculados para estas imagens mostraram a discrepância dos valores de reflectância da superfície da IB para as ICs (Tabela 4.4).

Tabela 4.4 – Sumário descritivo das informações dos dados dos valores de reflectância da superfície (σ), fator multiplicativo (FM) e valores de máximo (máx.) e mínimo (mín.) das bandas b4, b5 e b6 do sensor OLI, para as imagens correspondentes ao mosaico Nº 13.

Imagens / bandas	Valores dos APIs			Fator multiplicativo			Histogramas ApEq					
	σ b5	σ b6	σ b4	FM b5	FM b6	FM b4	FM * b5		FM * b6		FM * b4	
	R	G	B	R	G	B	Máx. b5	Mín. b5	Máx. b6	Mín. b6	Máx. b4	Mín. b4
IB 222 076	3288.561	1912.365	417.356	--	--	--	1432*	6118*	654*	2530*	31*	1093*
IC 218 076	3281.042	1550.481	475.105	1.0022917	1.233401	0.8784508	1435.28	6132.02	806.64	3120.50	27.23	960.15
IC 219 075	3346.167	1780.917	422.000	0.9827846	1.0738094	0.9889953	1407.35	6012.68	702.27	2716.74	30.66	1080.97
IC 219 076**	--	--	--	--	--	--	--	--	--	--	--	--
IC 219 077	3306.417	1648.625	453.708	0.9945997	1.1599757	0.9198773	1424.27	6084.96	758.62	2934.74	28.52	1005.43
IC 220 074	3375.459	1771.189	438.595	0.9742558	1.0797068	0.9515758	1395.13	5960.50	706.13	2731.66	29.50	1040.07
IC 220 075	3337.471	1783.083	415.701	0.9853453	1.0725048	1.0039806	1411.01	6028.34	701.42	2713.44	31.12	1097.35
IC 220 076	3487.630	1774.835	426.348	0.9429214	1.0774892	0.9789094	1350.26	5768.79	704.68	2726.05	30.35	1069.95
IC 220 077	3870.723	1649.062	495.969	0.8495986	1.1596687	0.8414958	1216.63	5197.84	758.42	2933.96	26.09	919.75
IC 221 074	3848.571	1761.086	468.486	0.8544888	1.0859011	0.8908617	1223.63	5227.76	710.18	2747.33	27.62	973.71
IC 221 075	3686.545	1468.977	479.034	0.8920441	1.3018343	0.8712449	1277.41	5457.53	851.40	3293.64	27.01	952.27
IC 221 076	3645.351	1669.596	488.801	0.9021247	1.1454055	0.8538359	1291.84	5519.20	749.10	2897.88	26.47	933.24
IC 221 077	3490.143	1652.143	452.286	0.9422425	1.1575058	0.9227707	1349.29	5764.64	757.01	2928.49	28.61	1008.59
IC 222 074	3105.000	1671.667	441.000	1.0591179	1.143987	0.9463855	1516.66	6479.68	748.17	2894.29	29.34	1034.40
IC 222 075	3481.709	1708.443	400.119	0.9445249	1.1193616	1.0430801	1352.56	5778.60	732.06	2831.98	32.34	1140.09
IC 222 076	3279.602	1913.112	417.026	<u>1.0027317^a</u>	<u>0.9996094^a</u>	<u>1.0007925^a</u>	1435.91	6134.71	653.74	2529.01	31.02	1093.87
IC 223 074	3096.000	1525.000	512.333	1.0621967	1.2540098	0.8146181	1521.07	6498.52	820.12	3172.64	25.25	890.38
IC 223 075	3769.122	1614.220	542.732	0.8725006	1.1846995	0.7689914	1249.42	5337.96	774.79	2997.29	23.84	840.51
IC 223 076	3462.000	1517.000	506.000	0.9499021	1.2606229	0.8248142	1360.26	5811.50	824.45	3189.38	25.57	901.52

Onde: APIs (alvos pseudo-invariantes), IB (imagem base), IC (imagem de conversão), FM (fator multiplicativo) e ApEq (aparência equalizada). * valores encontrados pela equalização visual. ** imagem com 100 % de presença de nuvem.

Futuramente, se uma imagem for importada para ambiente SIG e nenhuma regra de atribuição for aplicada na equalização, o contraste definido foi aquele dado pelo programa utilizado. Apenas programas como SPRING necessitam que o usuário defina o contraste para as imagens em composição colorida (SPRING, 1996). O contraste automático causa ruídos sobre a cor resultante dos alvos, o que dificulta a modelagem de regras de classificação de pixels (EBERHARDT, 2014).

Desta forma, os diferentes FMs obtidos para as imagens tiveram como objetivo minimizar os impactos causados pela variação da sazonalidade e δ sobre o contraste automático sobre as imagens Landsat. O contraste em ApEq foi aplicado manualmente, utilizando-se os valores resultantes da multiplicação do FM de cada uma das bandas, sobre os valores de máx. e mín. dos seus histogramas.

4.2.3. Mosaico composto por dezoito imagens em ApEq

As Figuras 4.3, 4.4, 4.5, 4.6 e 4.7 ilustram o resultado do uso da ApEq, em diferentes escalas de detalhamento, para o mosaico do estado de São Paulo, composto por 18 imagens OLI.

Baseado na experiência em classificação visual de culturas agrícolas, em mosaicos composto por várias imagens de sensoriamento remoto, percebeu-se que a partir da ApEq é possível definir uma contraste universal para todas as 18 imagens que compuseram o estado de São Paulo (Figura 4.3). O uso deste método diminui o impacto causado pelas linhas de sobreposição entre imagens de diferentes órbitas (Figuras 4.4 e 4.5) e, além disso, as culturas como cana-de-açúcar e soja, ambas em pleno vigor vegetativo tiveram cores semelhantes em imagens localizadas em diferentes órbitas, neste caso 221/076 (Figura 4.6) e 220/075 (Figura 4.7).

Figura 4.3 – As 18 imagens OLI que compreendem o mosaico de Nº 13 usado nesta pesquisa, em aparência equalizada.

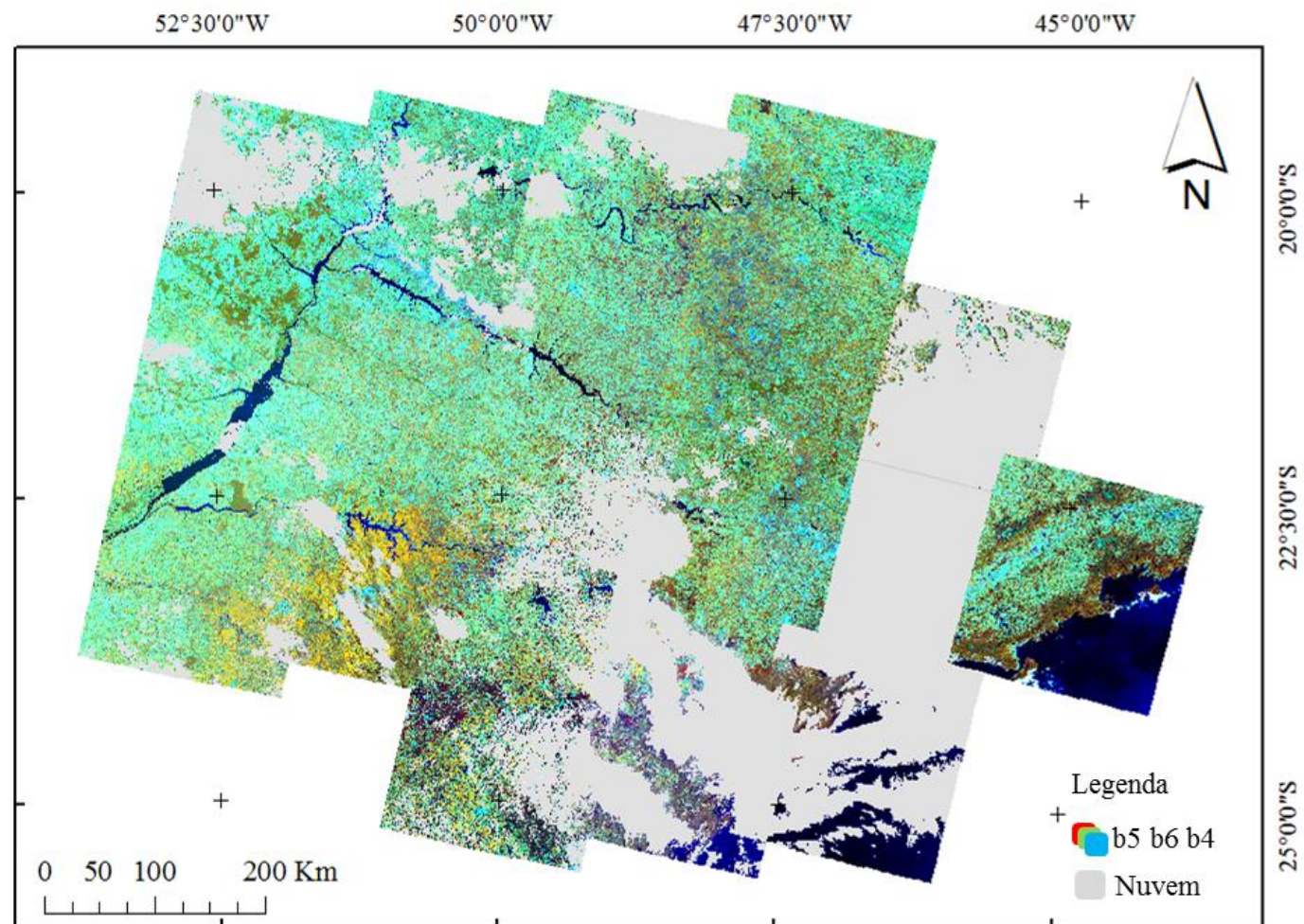


Figura 4.4 – Área de sobre posição das imagens OLI localizadas no extremo Sul do estado de São Paulo. As linhas em vermelho indicam as linhas de sobreposição entre imagens adjacentes.

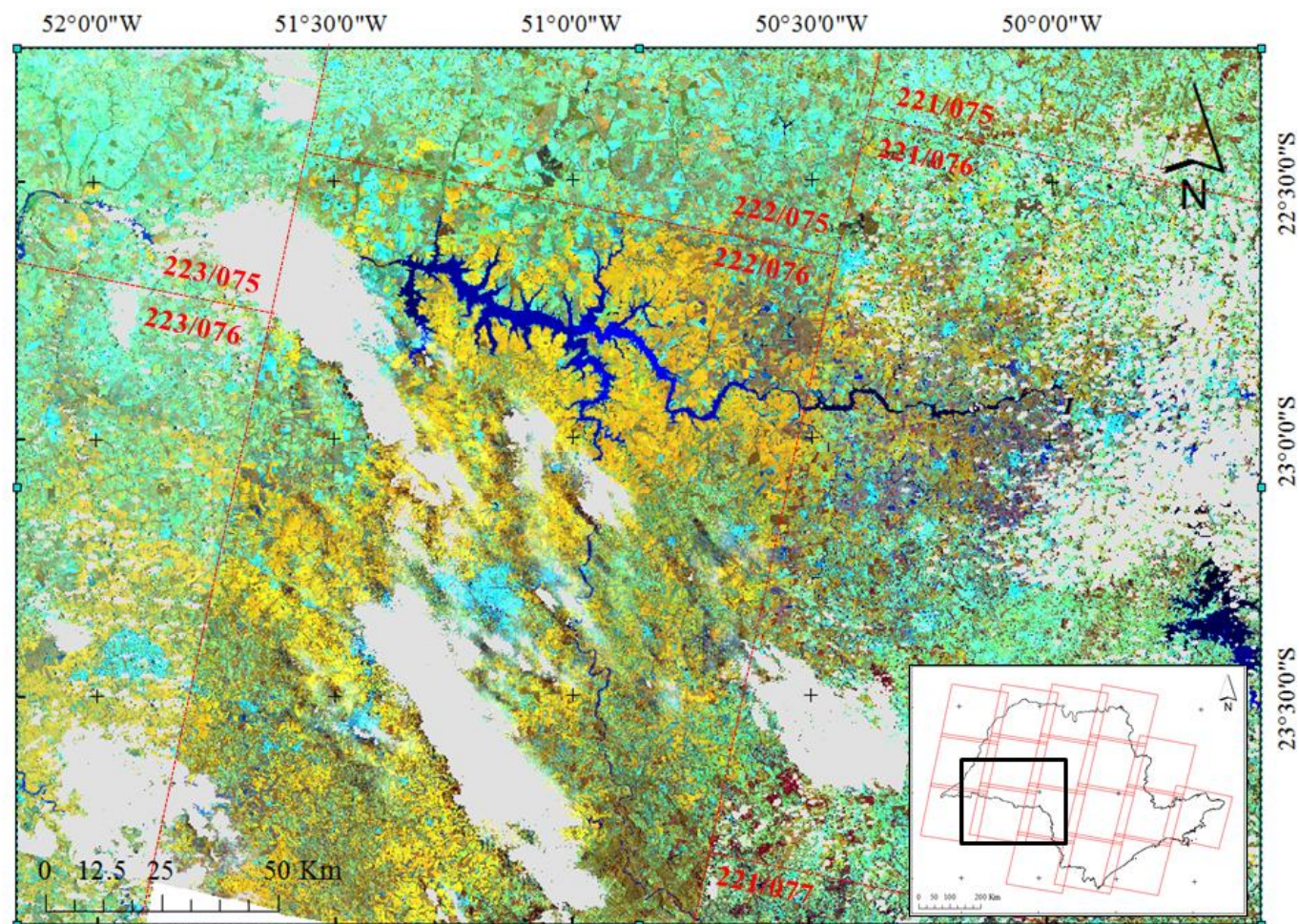


Figura 4.5 – Área de sobre posição das imagens OLI localizadas ao nordeste do estado de São Paulo. As linhas em vermelho indicam as linhas de sobreposição entre imagens adjacentes.

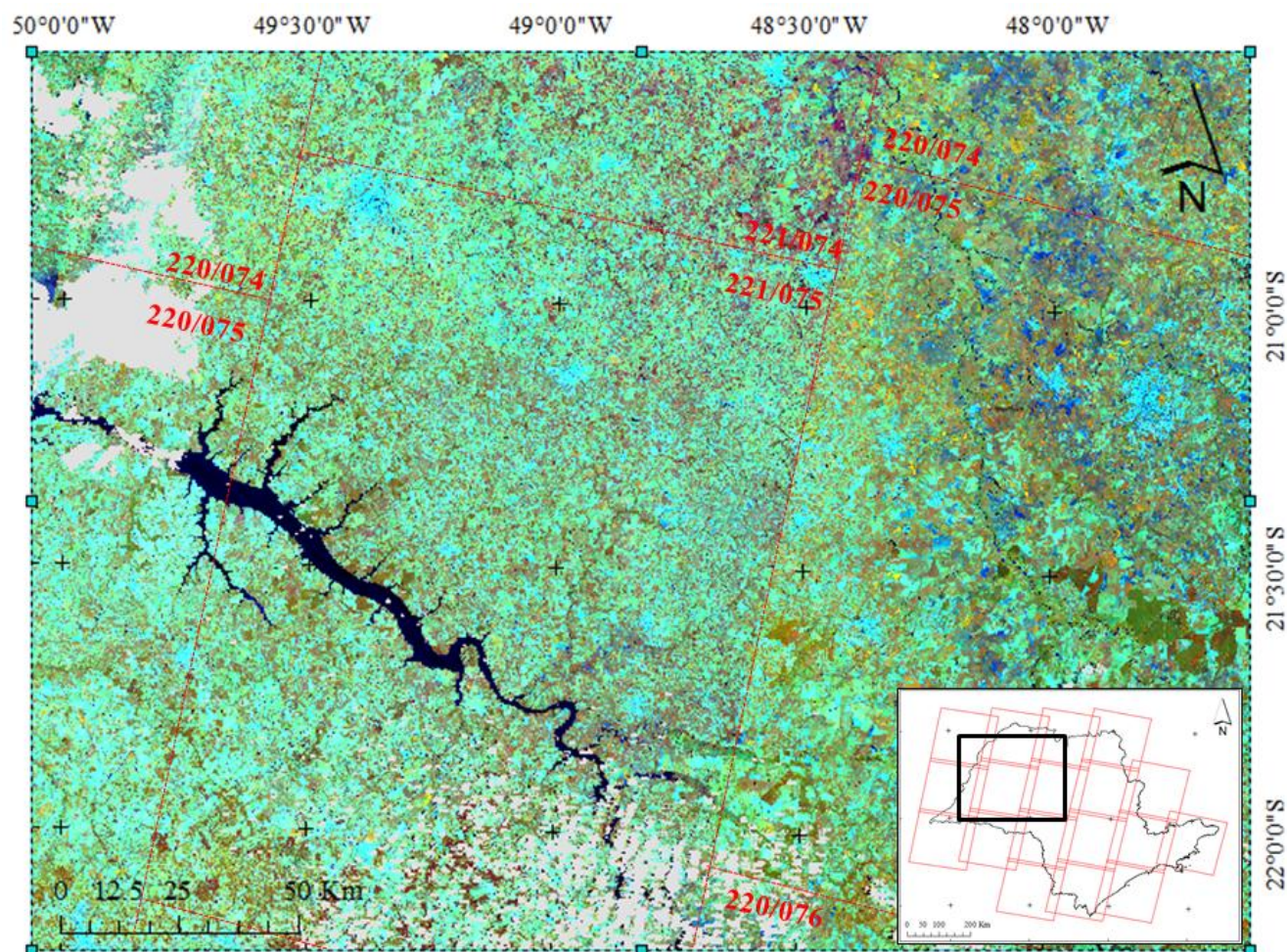


Figura 4.6 – Resultado das cores da imagem 221/076 do dia 24/janeiro/2015 em aparência equalizada. Em destaque as cores resultantes para a cana-de-açúcar (1), soja (2) e floresta nativa (3) validados em campo, no dia 13-17/abril/2015.

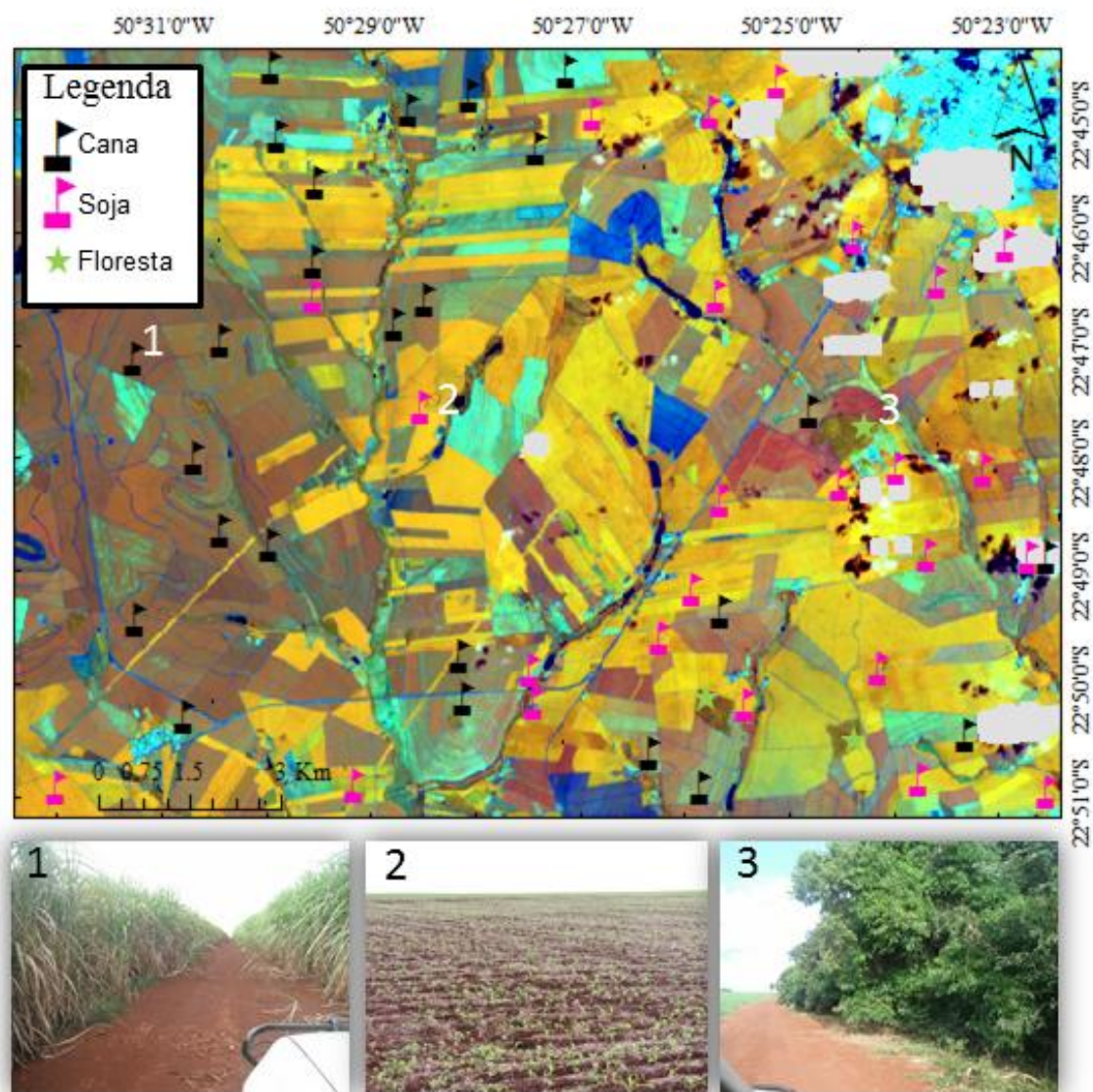
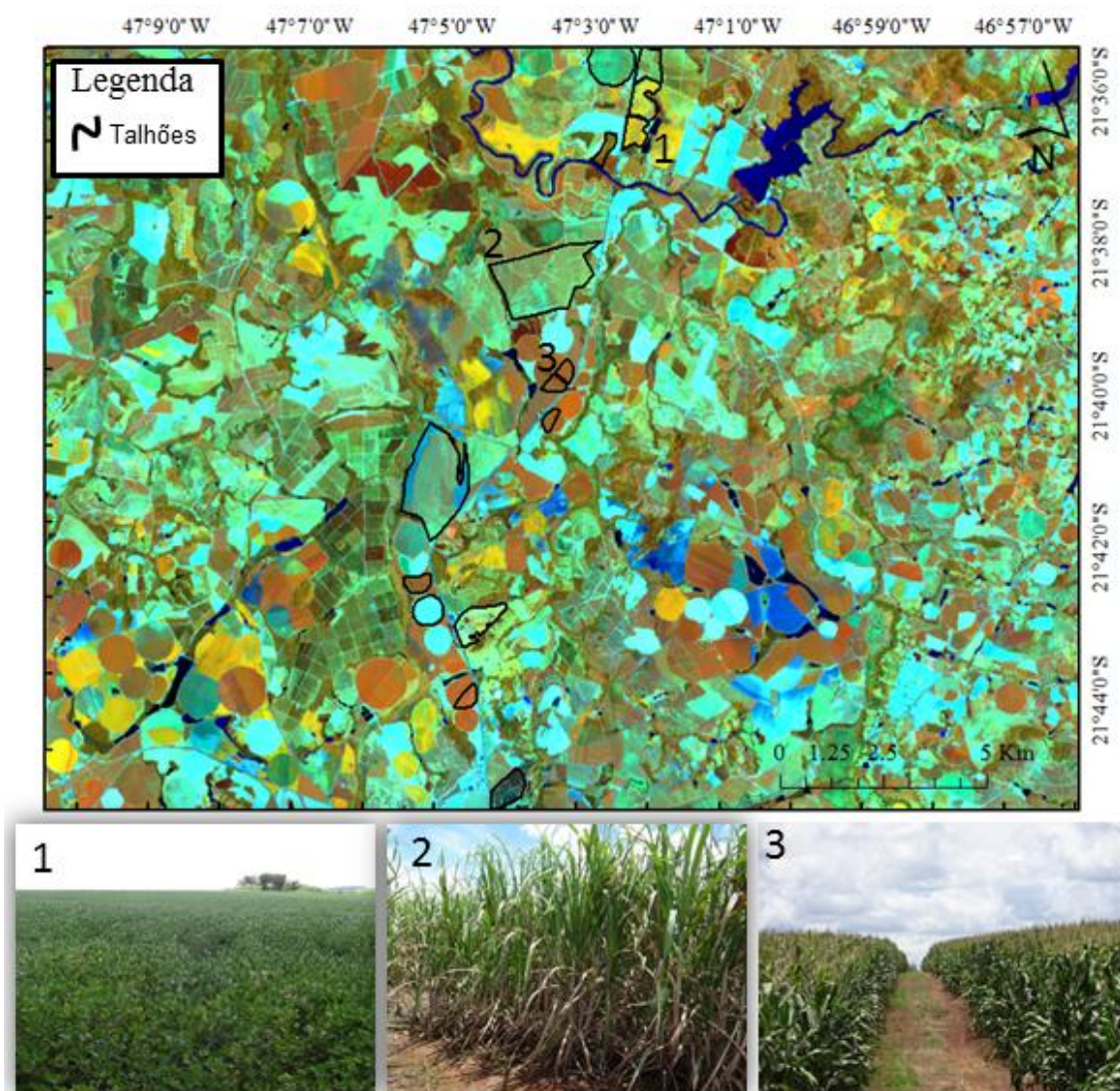


Figura 4.7 – Resultado das cores da imagem 220/075 do dia 17/janeiro/2015 em aparência equalizada. Em destaque as cores resultantes para a soja (1), cana-de-açúcar (2) e milho (3) validados em campo, no dia 13-17/abril/2015.



4.3. Definição do protocolo de classificação visual de imagens OLI

A campanha de campo realizada nos dias 17-21/abril/2015 serviu para validar 200 pixels OLI classificados visualmente, em laboratório. A imagem OLI utilizada para representar a reposta espectro-visual dos objetos validados em campo foi gerada 24 dias após o último dia do trabalho de campo. Demais imagens OLI anteriores a este período, serviram de base para montagem de uma série multitemporais de imagens OLI, também utilizadas nesta seção.

Para isto, os dados coletados em campo foram confrontados com as imagens OLI em composição colorida falsa cor R-b5 G-b6 B-b4, em ApEq. Os resultados foram apresentados da seguinte forma:

(1) Comportamentos espectro-visuais de:

- (a) Vegetação fotossinteticamente ativa (VFA) da cana-de-açúcar, milho e soja;
- (b) VFA de alvos semelhantes às culturas de interesse;
- (c) Alvos duvidosos como pixels de borda e pixels com culturas que apresentaram falhas de plantio;
- (d) Ausência de VFA.

(2) Correlação espectro-temporal-visual de:

- (a) Pixel de cana-de-açúcar colhida e reformada;
- (b) Milho de segunda safra (safrinha) e da soja;
- (c) Alvos semelhantes às culturas da cana-de-açúcar, milho e soja.

Todos estes dados foram relacionados ao calendário agrícola 2014/2015 da cana-de-açúcar, milho e soja em São Paulo e aos quatro períodos da estimativa agrícola do painel amostral definido pelo método MoBARS, descrito

por Luiz *et al.*, (2015). A partir deste ponto foi definido o protocolo de classificação visual das culturas de interesse.

- (3) Protocolo de classificação espectro-temporal de cana-de-açúcar, milho e soja.

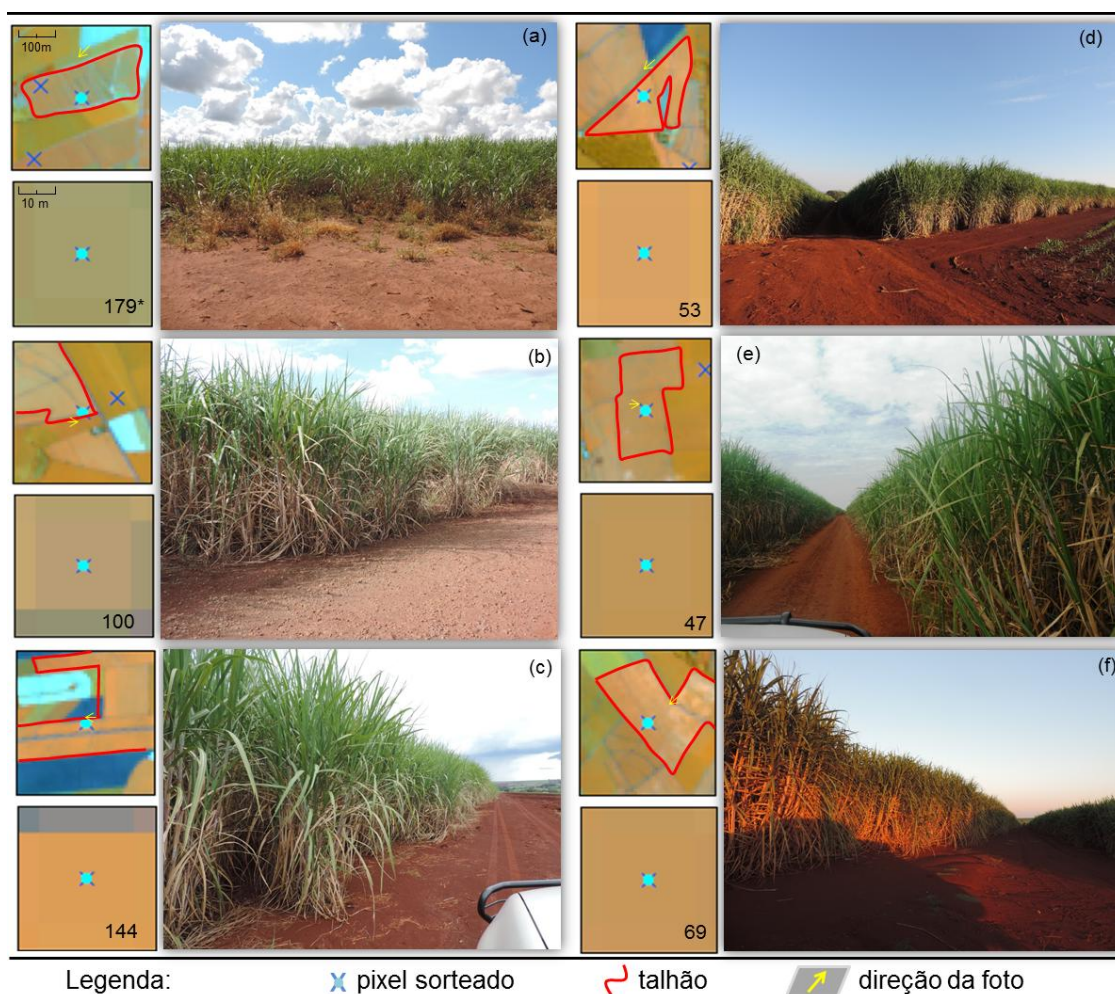
4.3.1. Comportamento espectro-visual

Ao observar os recortes das imagens de satélite (que correspondem à área do município de Cândido Mota/SP) na composição colorida falsa cor RGB (564), dos dias 07/maio/2015 (para cana-de-açúcar e milho) e 15/janeiro/2015 (para soja), em aparência equalizada, foi possível definir a cor que representa a VFA de cana-de-açúcar (Figura 4.8), de milho (Figura 4.9) e soja (Figura 4.10).

Na Figura 4.8 foi possível correlacionar fotos retiradas no campo, de talhões com cana-de-açúcar de diferentes alturas de dossel, com os pixels OLI em ApEq. Os talhões de cana-de-açúcar com menor altura de dossel (Figura 4.8.a e 4.8.b) apresentaram-se amarronzados, enquanto que os demais adquiriram gradação da cor magenta.

Os talhões de cor amarronzada apresentaram no campo, maior efeito de substrato que os talhões de cor magenta. Além disso, a gradação da cor magenta presente nos talhões de dossel mais alto (Figura 4.8.c a 4.8.f) esteve em função da palhada presente na base dos colmos. Sanches *et al.* (2005) comentaram que talhões de cana-de-açúcar de diferentes idades apresentam gradação da cor magenta, devido a presença de folhas secas e a irregularidade dos dosséis.

Figura 4.8 – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Talhões de cana-de-açúcar com: (a) 1,5 m; (b) 2,0 m; (c) 2,5 m; (d) 3,0 m; (e) 3,5 m e (f) 4,0 m.



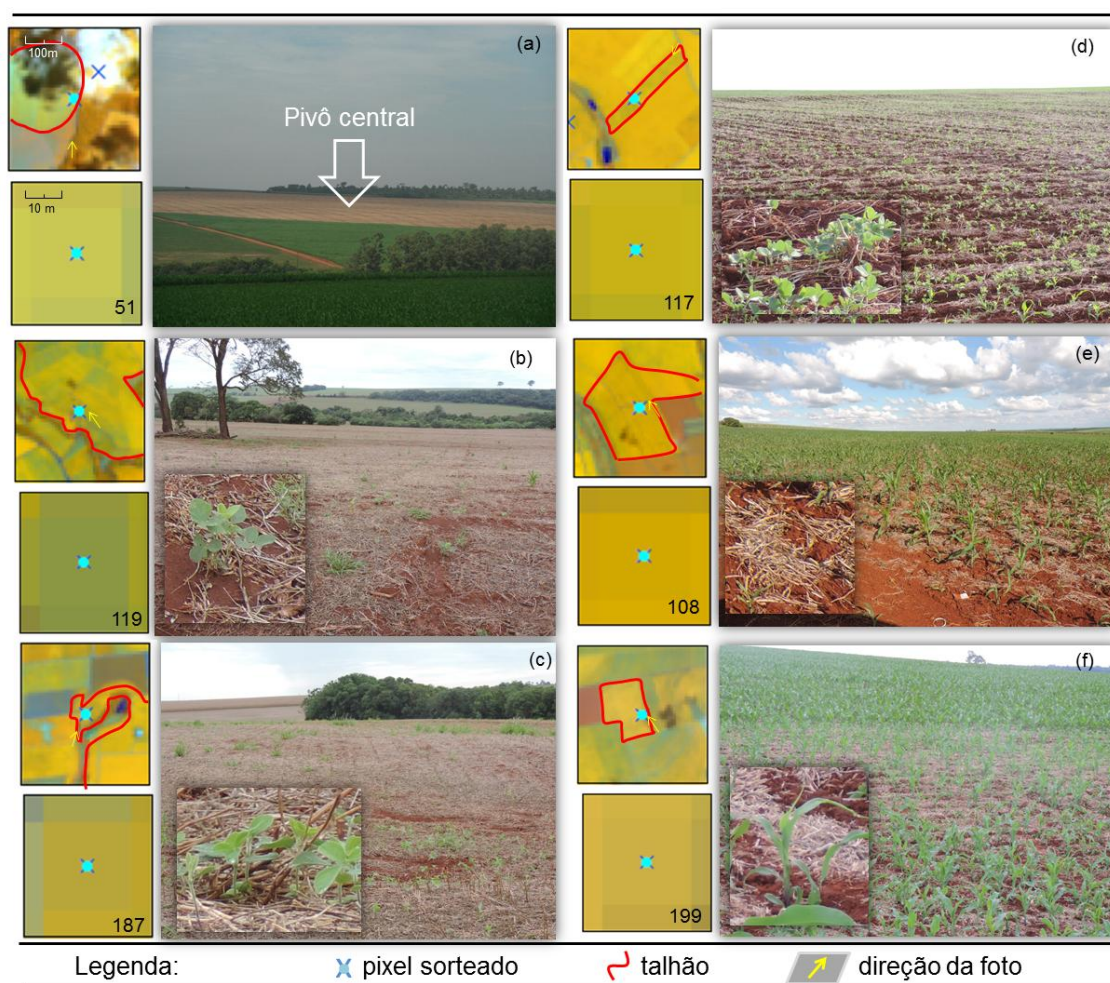
As imagens foram formadas em 07/maio/2015 e os dados de campo foram obtidos nos dias 17-21/abril/2015.

Figura 4.9 – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Talhões de milho com: (a) 0,3 m; (b) 1,2 m; (c) 1,5 m; (d) 2,0 m; (e) 2,5 m e (f) 2,5 m e pendoado.



As imagens foram formadas em 07/maio/2015 e os dados de campo foram obtidos nos dias 17-21/abril/2015.

Figura 4.10 – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. (a) Pivô central recoberto com palhada de soja. (b; c; d) Talhão com palhada e brotação da soja. (e; f) Talhão com palhada e milho de 30 cm.



As imagens foram formadas em 15/janeiro/2015 e os dados de campo foram obtidos nos dias 17-21/abril/2015. Talhões que haviam sido plantados com soja.

Na Figura 4.9 foi possível correlacionar talhões de milho de diferentes alturas de dossel, com os pixels OLI, em ApEq. Ao analisar os talhões de milho no campo, todos apresentaram visualmente o mesmo espaçamento entre plantas, no entanto, os dosséis diferiram apenas em relação às alturas (Figura 4.9.a a Figura 4.9.d).

Além disso, na Figura 4.9.f, as plantas de milho já se encontravam em estágio reprodutivo, em que o pendão estava parecendo. Nas imagens do sensor OLI para as Figuras 4.9.a e 4.9.b, os pixels apresentaram gradação da cor alaranjada, enquanto que na Figura 4.9.e o pixel se apresentou na cor marrom.

A gradação da cor alaranjada é resultado da variação da quantidade de substrato e VFA expostos para o sensor, em cada talhão. Quanto mais VFA recobre o pixel de 30 m, mais intenso é a cor resultante para a cultura avaliada (SANCHES *et al.*, 2005).

A influência do pendão na cor do pixel de VFA de milho (Figura 4.9.f) teve semelhança à influência da palhada presente na parte inferior do dossel de cana-de-açúcar. A estrutura de do pendão e da palha é diferente, porém a constituição é semelhante, isto é, depois de polonizado, o pendão assim como a palhada da cana-de-açúcar, são compostos por matéria seca (MAGALHÃES *et al.*, 2002).

Na Figura 4.10 foi possível correlacionar talhões de soja com as imagens OLI do dia 15/janeiro de 2015, em ApEq. Neste caso, os talhões de soja só foram validados, devido a presença de palhada e brotações de soja decorrentes da sua colheita, presentes no dia do trabalho de campo. A diferença temporal entre esta imagem e a campanha de campo foi próxima a três meses. A cor que a VFA da soja representou foi a gradação do amarelo.

Sanches *et al.* (2005), Rizzi e Rudorff (2005) e Sugawara *et al.* (2008) enfatizaram que, a soja, ao fechar completamente o dossel agrícola, apresenta matiz amarelo intenso. A gradação do matiz amarelo é dada pela influência do

substrato no talhão de plantas de soja em crescimento (SANCHES *et al.*, 2005).

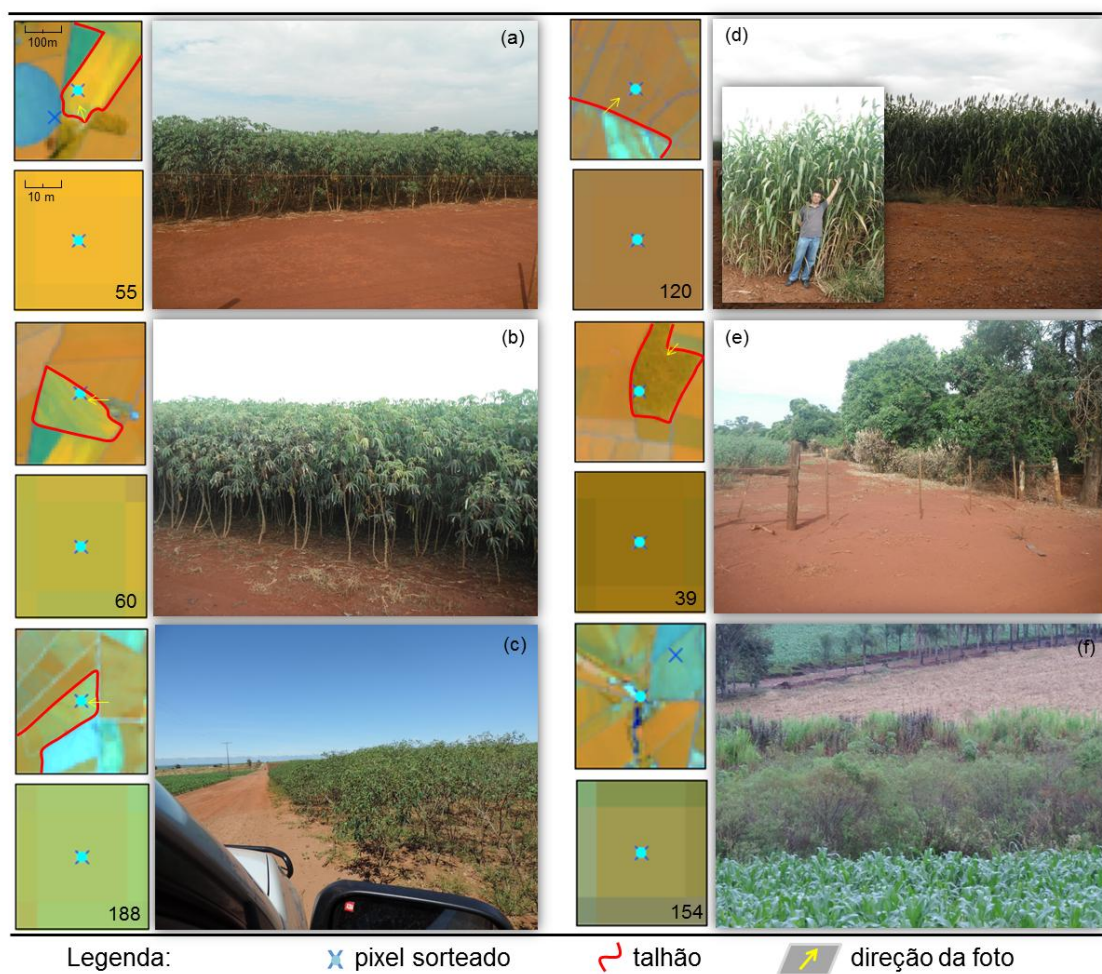
Áreas de floresta nativa e de culturas agrícolas como mandioca e sorgo sacarino (Figura 4.11.a, 4.11.b, 4.11.c e 4.11.d) possuem cor de VFA semelhante às culturas de interesse. Neste caso, a VFA de mandioca se assemelhou a soja; e a VFA de sorgo, se assemelhou a cana-de-açúcar e ao milho. A mandioca e a soja são plantas que apresentam estrutura de dossel planófilo, e o sorgo sacarino, o milho e a cana-de-açúcar são plantas erectófilas. A semelhança da estrutura do dossel destas culturas no campo aproxima os seus comportamentos espectral-visuais nas imagens.

Os pixels de áreas de floresta nativa (Figura 4.11.e), de capoeira suja (Figura 4.11.f), pixels localizados nas bordas entre talhões agrícolas e áreas de mata ciliar (Figura 4.12.a) ou fragmentos florestais (Figura 4.12.b), pixels representantes de áreas de pastos perenes (Figura 4.12.c) e pixels provenientes de áreas agrícolas com falhas de plantio (Figura 4.12.d) também se assemelham espectral-visualmente com as gradações das cores dos pixels de cana-de-açúcar e milho. Porém, aqueles pixels são considerados como classe “outros”.

Demais alvos como área de palhada resultante da colheita da soja (Figura 4.12.e) e da colheita do milho (Figura 4.12.f) apresentaram cores completamente diferentes das cores de VFA das culturas de interesse. Estes pixels são considerados como pixels de ausência da VFA (AVFA), que também pertencem a classe outros. A definição destes pixels de AVFA foi importante para o intérprete, principalmente para entender e definir o corte entre, pixels com VFA de milho, por exemplo, e pixels que ainda não atingiram a condição para serem classificados como milho.

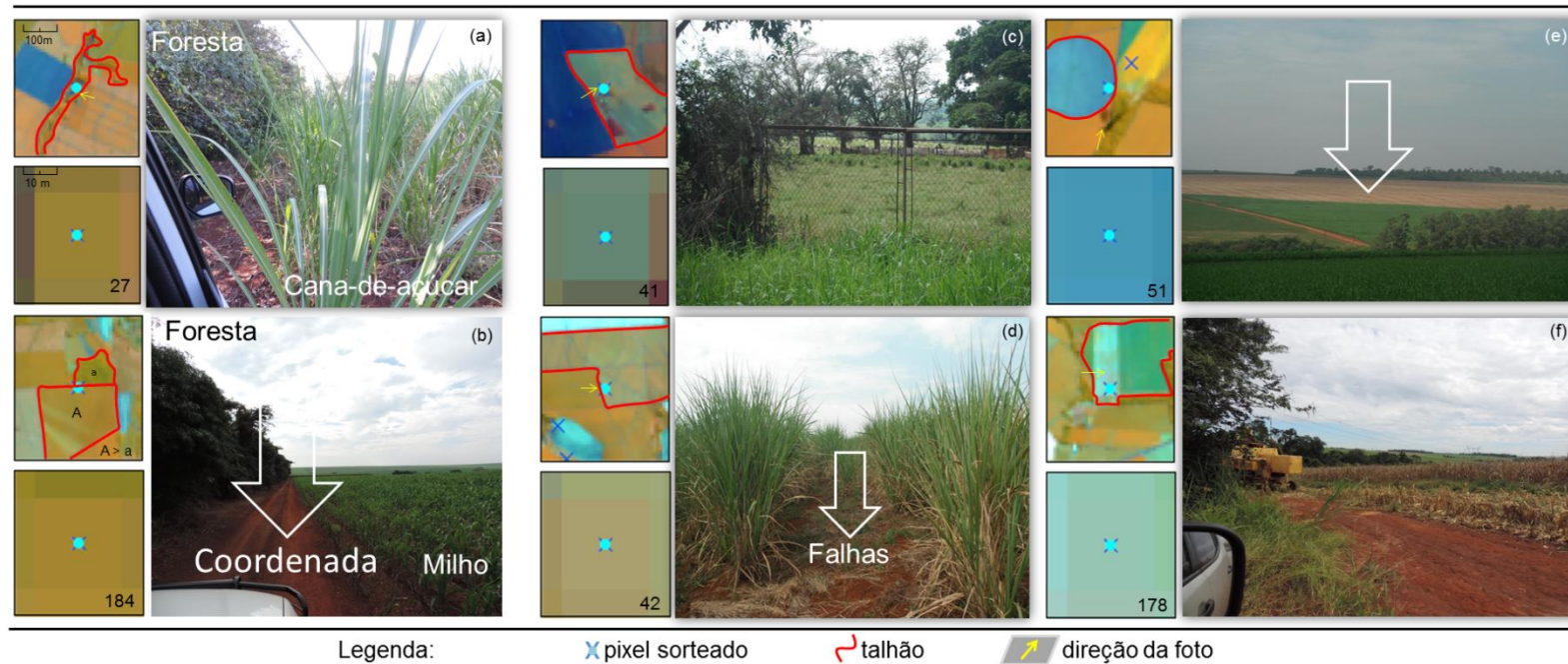
A Figura 4.13 mostra a presença da cultura verde em pé nos dias da campanha de campo e a resposta espectral-visual de pixels que não puderam ser considerados como pixels classificados como cana-de-açúcar (Figura 4.13.a a 4.13.c) ou milho (Figura 4.13.d a 4.13.f).

Figura 4.11 – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. (a; b; c) Talhões de mandioca com 2,0; 1,5 e 1,0 m, respectivamente. (d) Talhão de sorgo sacarino, considerado pixel “extra”. (e) Remanescente de floresta natural. (f) Área de capoeira.



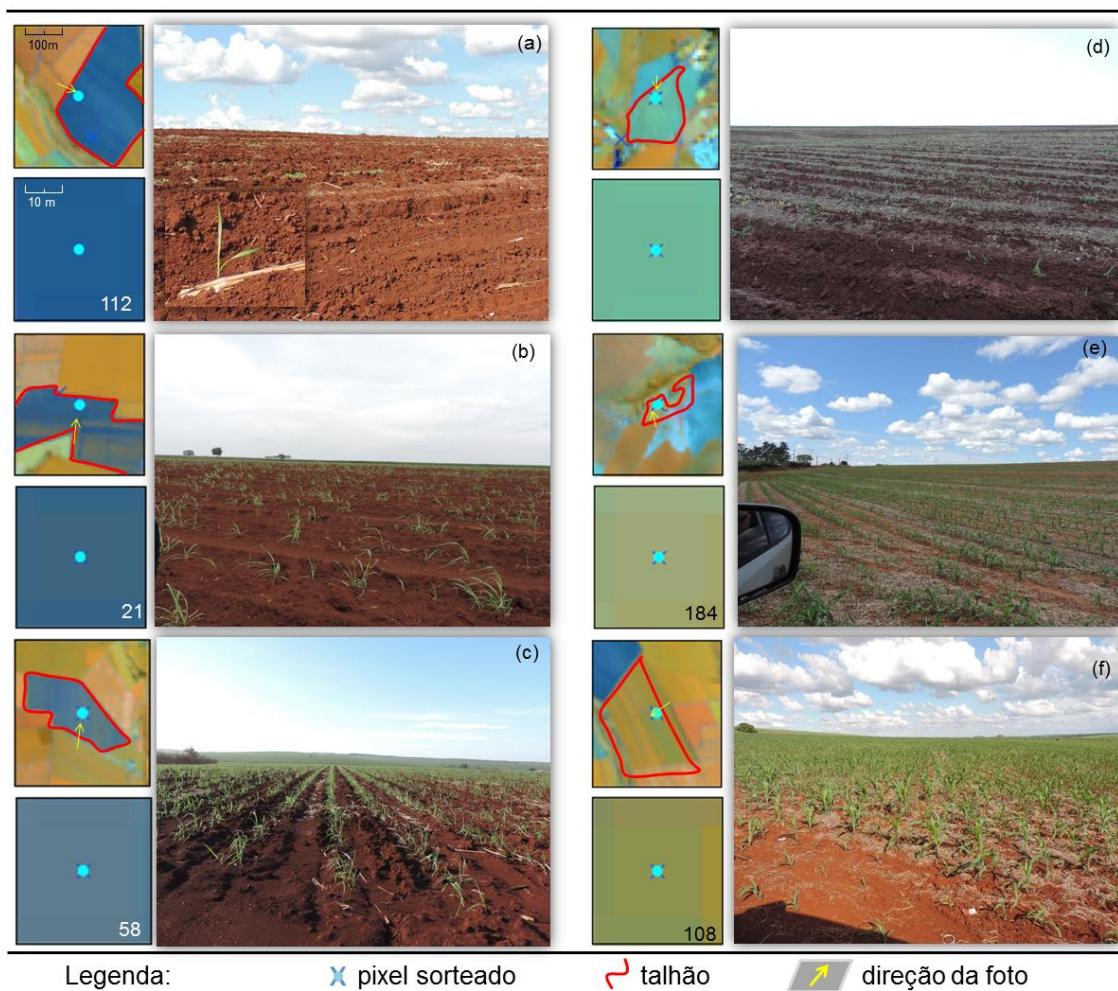
As imagens foram formadas em 07/maio/2015 e os dados de campo foram obtidos nos dias 17-21/abril/2015.

Figura 4.12 – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. (a) Mata ciliar. (b) Divisa entre remanescente florestal e talhão de milho. (c) Pastagem. (d) Talhão de cana-de-açúcar com falhas. (e) área de solo exposto; (f) talhão de milho seco em colheita.



As imagens foram formadas em 07/maio/2015 e os dados de campo foram obtidos nos dias 17-21/abril/2015.

Figura 4.13 – Recortes de imagens OLI/Landsat-8 nas escalas 1:30.000 e 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. (a) talhão de cana-de-açúcar recém-plantada. (b; c) talhões de cana-de-açúcar com plantas de 20 e 50 cm de altura, respectivamente. (d; e; f) Talhões de milho com plantas de 10, 20 e 50 cm de altura, respectivamente



As imagens foram formadas em 07/maio/2015 e os dados de campo foram obtidos nos dias 17-21/abril/2015..

Estes pixels que geraram confusão para o intérprete no momento da classificação visual só puderam ser classificados devido à diferença do comportamento espectro-temporal (CET), em mais datas de imagens na mesma composição colorida, de cada alvo analisado.

O CET da cana-de-açúcar, do milho e da soja foi definido pelo uso de 14 imagens multitemporais presentes em cada período (1, 2, 3 e 4) da estimativa agrícola do ano safra 2014/2015. Alguns pixels foram tomados como representantes das condições agrícolas destas culturas no tempo, e outros utilizados para comparação temporal de alvos duvidosos (que se assemelharam às culturas de interesse) e que puderam ser corretamente classificados pelo uso de imagens multitemporais.

4.3.2. Correlação espectro-temporal visual

A Figura 4.14 mostra a variação espectro-temporal-visual de pixels de cana-de-açúcar colhida com máquina. Na Figura 4.14 a variação das cores dos pixels no tempo, foi notória e seguiu um padrão. Foram definidos três principais rótulos destes pixels: (a) pixels ocupados com presença de VFA de cana-de-açúcar, (b) pixels em transição e (c) pixels com AVFA.

A cana-de-açúcar possui calendário semi-perene, isto é, pode permanecer 1 a 1,5 anos no campo, verde e em pé (SUGAWARA e RUDORFF, 2007; RUDORFF *et al.* 2010). A colheita da cana-de-açúcar, normalmente ocorre de abril a novembro (AGUIAR *et al.* 2011), e a intensificação da colheita ocorre em outubro (AGUIAR *et al.*, 2010).

A colheita da cana-de-açúcar realizada em diferentes meses (de abril a novembro), fez com que pixels de AVFA formassem uma espécie de “escada” na Figura 4.14. Neste caso, o padrão em escada foi chamado de “escada da colheita” da cana-de-açúcar. Os pixels que causaram dúvidas pela subjetividade da acuidade visual do olho humano, para decidir entre VFA e AVFA de “cana”, foram aqui chamados de pixels de transição (Figura 4.14); e os demais foram considerados como pixels de cana-de-açúcar ou com AFVA.

Após a colheita da cana-de-açúcar, a recuperação da VFA no tempo necessita de dois a três meses para adquirir novamente intensa atividade fotossintética (RUDORFF *et al.*, 2010). Em trabalhos de campo realizados no ano safra 2014/2015 este período mostrou-se, em alguns casos, ser mais longo. Talhões de cana-de-açúcar levaram de quatro a cinco meses para adquirirem novamente características de VFA.

Ao considerar o período tempo em que o talhão da cana-de-açúcar demora para atingir novamente a atividade fotossintética, foi elaborada uma regra de atribuição para classe cana-de-açúcar. Nesta regra, o intérprete deveria classificar o pixel de transição como cana-de-açúcar, apenas se ele tivesse sido considerado como AVF, decorrido três meses anteriores à data analisada (Figura 4.14). Para aplicar esta regra, faz-se necessário o uso de imagens multitemporais anteriores aos quatro períodos da estimativa (julho a outubro).

Para a Figura 4.15, o pixel de transição teve menor efeito sobre os problemas de classificação. A diferença entre a cor do solo exposto (escuro) e da VFA foi abrupta, isto é, passou do matiz preto para vermelho (Figura 4.15). Ainda na Figura 4.15, foi utilizado o pixel nº 6 para mostrar VFA da cana-de-açúcar estável durante todo o no ano safra.

Demais variações na gradação da cor dos pixels de cana-de-açúcar no tempo, estiveram em função de fatores topográficos, de iluminação do sol, alinhamento do plantio, espaçamento, tipos de cultivares, idades das soqueiras, etc (COLOMBO *et al.*, 2013). Estas variações causaram diferenças nos valores tonais dos pixels nas bandas utilizadas na composição colorida RGB, fato que dificilmente pode ser corrigido pela ApEq.

A Figura 4.16 mostra a variação do comportamento espectro-temporal-visual de pixels de soja e milho manejados sob o regime de safrinha do milho. Nesta Figura, a maior proporção de pixels seguiu um padrão semelhante, isto é, do mês de julho a novembro, estes pixels apresentaram-se com AVFA, de dezembro a janeiro permaneceram com VFA na gradação amarela e de abril a junho adquiriram gradações no matiz laranja.

Figura 4.14 – Recortes de imagens OLI/Landsat-8 na escala 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Os pixels de 30 x 30 m se localizam na parte central do recorte. A variação das cores dos pixels é resultante da quantidade de VFA da cana-de-açúcar presente em cada uma das 14 datas do ano-safra 2014/2015 e também da ausência de VFA correspondente a palhada gerada pela colheita mecanizada.

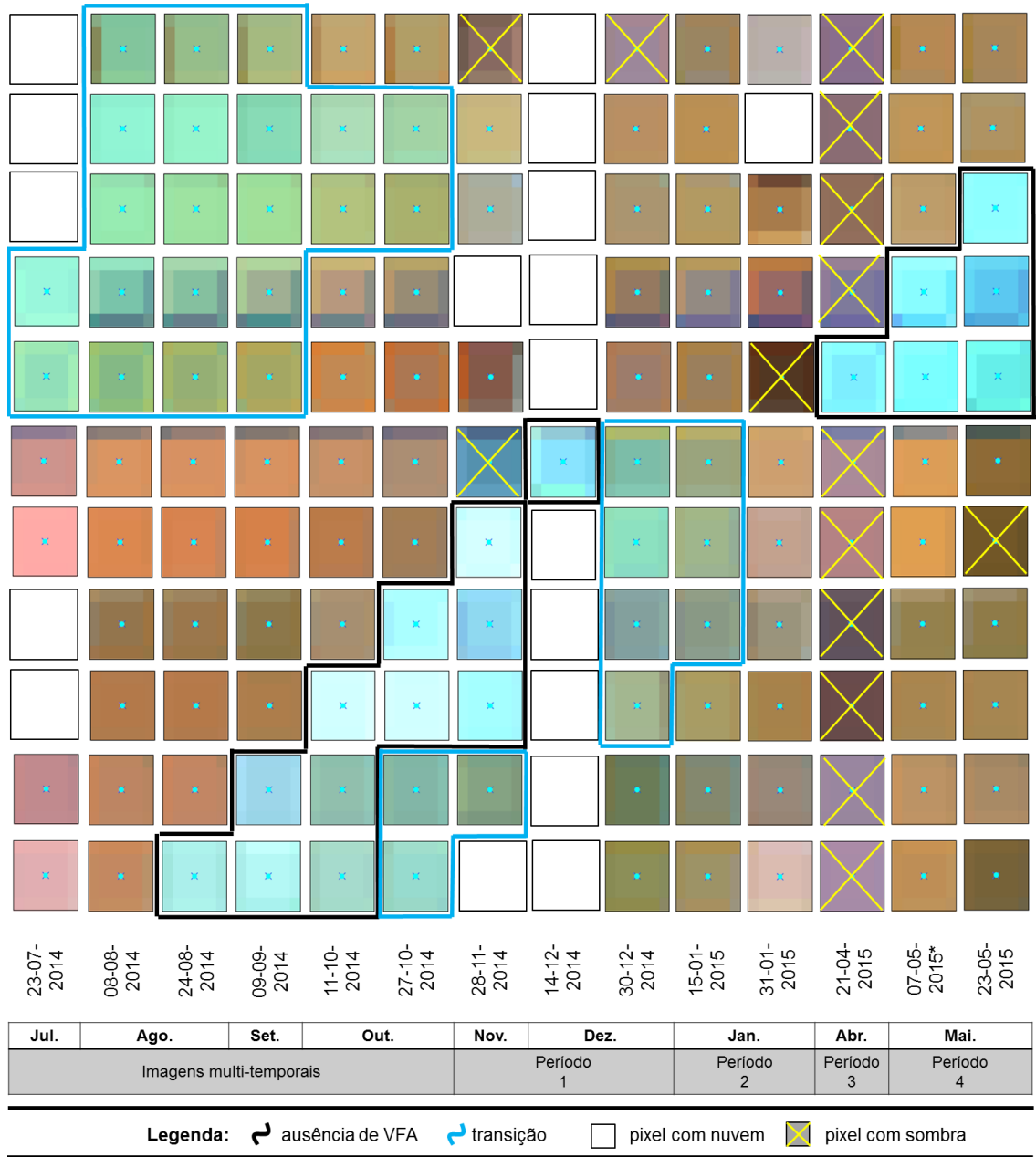


Figura 4.15 – Recortes de imagens OLI/Landsat-8 na escala 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Os pixels de 30 x 30 m se localizam na parte central do recorte. A variação das cores dos pixels é resultante da quantidade de VFA da cana-de-açúcar presente em cada uma das 14 datas do ano-safra 2014/2015 e também da ausência de VFA gerada pela exposição do solo.

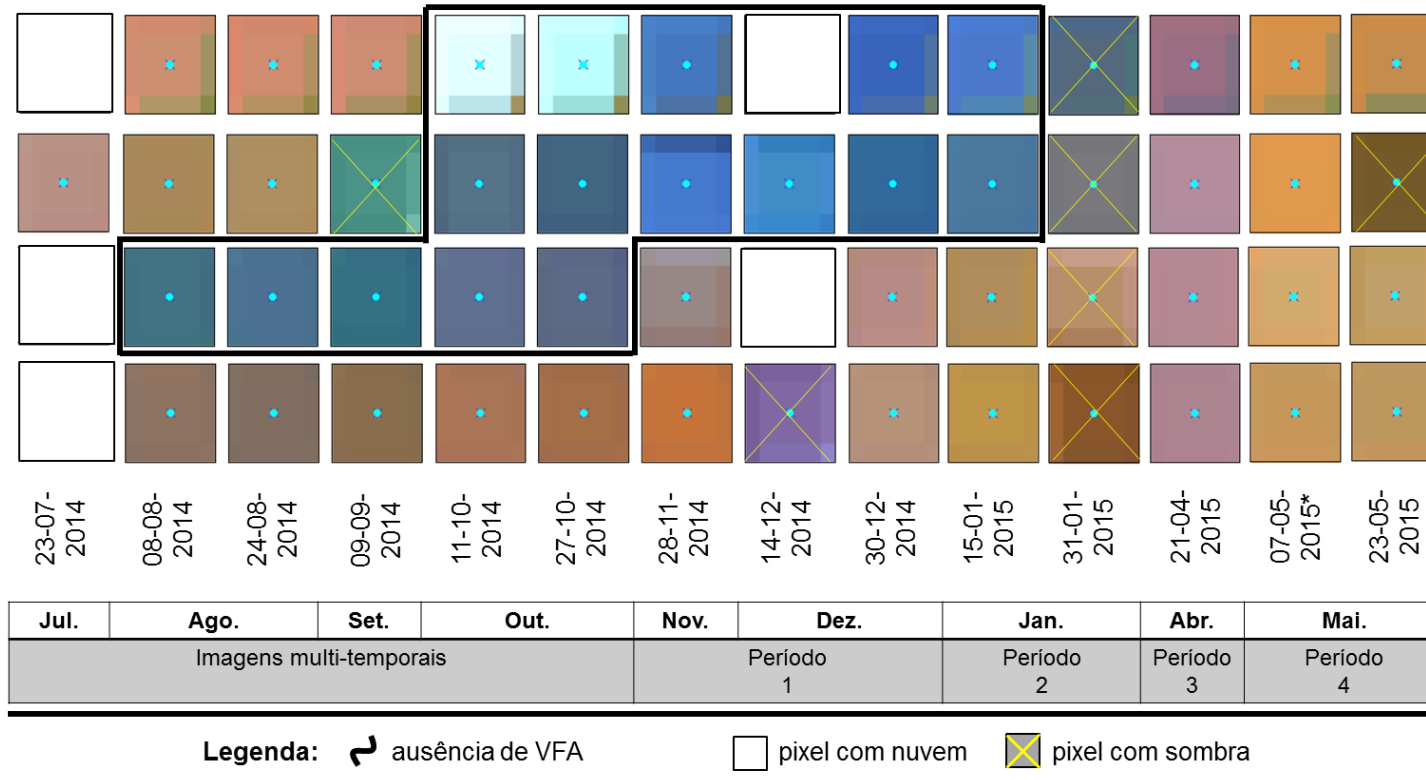


Figura 4.16 – Recortes de imagens OLI/Landsat-8 na escala 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Os pixels de 30 x 30 m se localizam na parte central do recorte. A variação das cores dos pixels é resultante da quantidade de VFA de soja e milho presente em cada uma das 14 datas do ano-safra 2014/2015 e da ausência da VFA devido ao milho seco, a palhada e ao solo exposto.

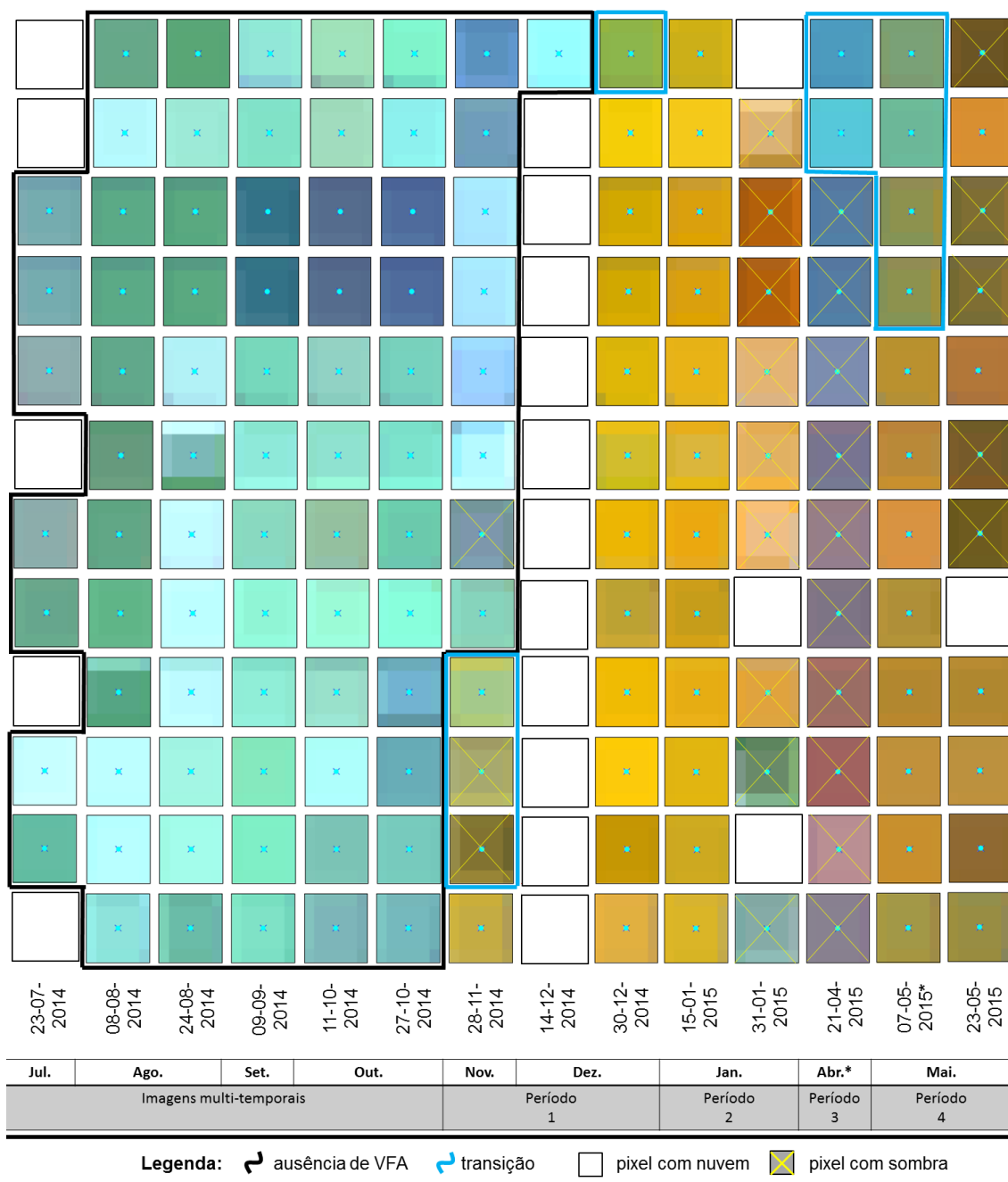
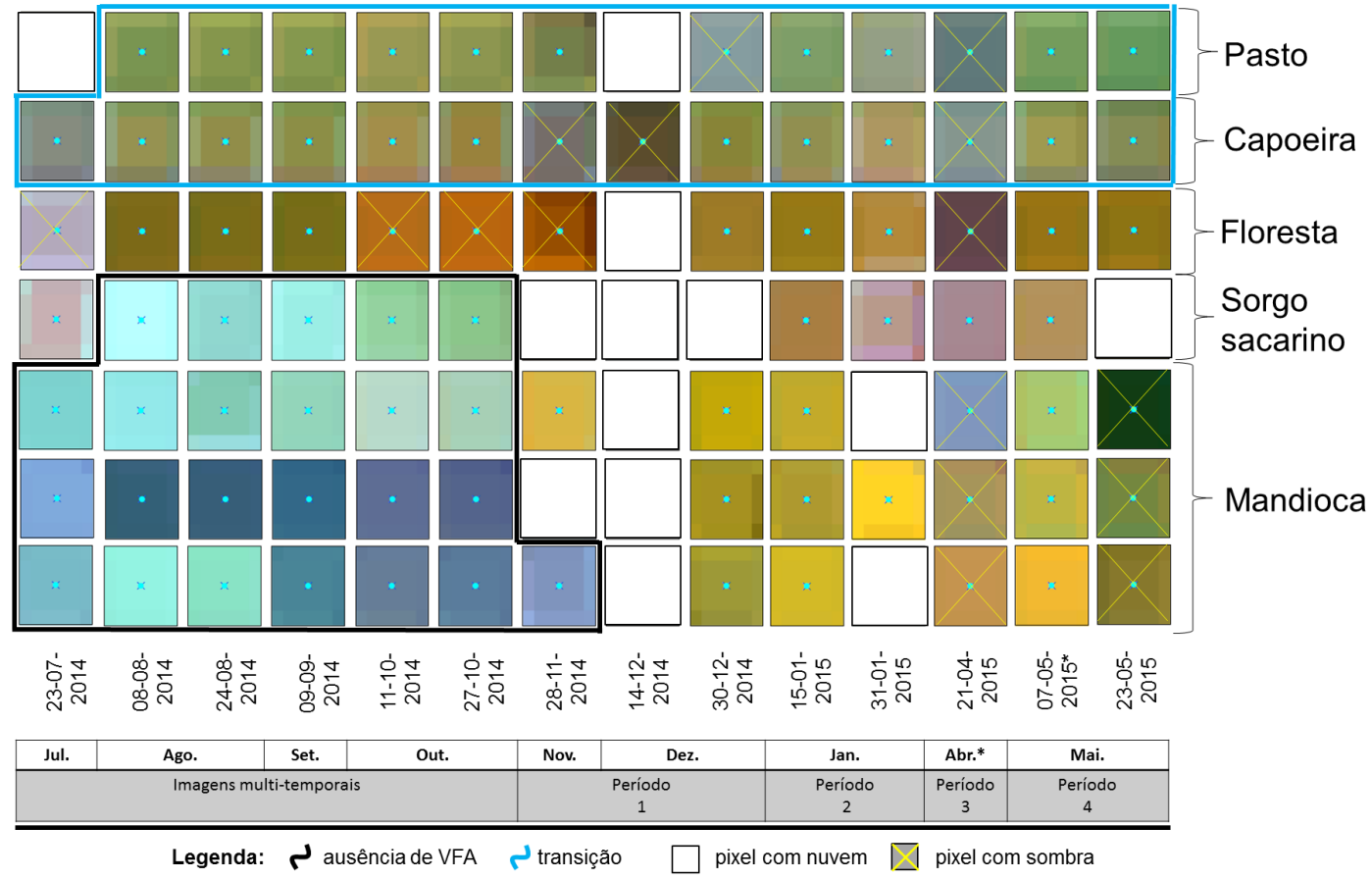


Figura 4.17 – Recortes de imagens OLI/Landsat-8 na escala 1:1.000 na composição colorida falsa cor RGB 564, em aparência equalizada. Os pixels de 30 x 30 m se localizam na parte central do recorte. A variação das cores dos pixels é resultante da quantidade de VFA de mandioca, sorgo sacarino, floresta, capoeira e pasto em cada uma das 14 datas do ano-safra 2014/2015 e da ausência de VFA devido a palhada e ao solo exposto.



O milho safrinha é produzido em mais de 73 % da área de milho plantada no estado de São Paulo (IEA, 2015). A soja, no sul do estado de São Paulo, é plantada de outubro a dezembro, e normalmente colhida em março. O plantio de milho safrinha normalmente ocorre depois da colheita da soja, nos meses de março a abril, e a colheita, até agosto. A tradição agrícola utilizada nesta área de estudo permite que seja definido um padrão de reconhecimento visual de pixels de soja e milho.

Na parte inferior da Figura 4.16 é perceptível que a gradação amarela aparece antecipadamente em alguns pixels, em relação aos demais. Como a soja possui uma ampla janela de plantio, alguns talhões agrícolas plantados com soja em outubro apresentaram característica de VFA, já no mês de novembro. Demais pixels de soja seguiram um padrão semelhante de comportamento, em que a VFA foi perceptível apenas em dezembro.

Os pixels que corresponderam à cor do milho safrinha apresentaram também gradação do seu matiz. A maior proporção dos pixels adquiriu a cor laranja nos meses de maio e junho, poucos pixels não tiveram características de VFA em maio. O efeito da nebulosidade sobre a imagem de abril (21/abril/2015) apresentou mudança na cor padrão dos pixels de milho. No entanto, nesta data, a cor magenta resultante em alguns pixels demonstrou alta atividade fotossintética das plantas milho.

A Figura 4.17 mostra a variação do comportamento espectro-temporal de demais alvos que representam a classe outros. A mandioca, por exemplo, se assemelha espectralmente com a soja, no entanto, a cultura da mandioca é semi-perene e permanece 1 a 1,5 anos, verde e em pé no campo. A soja por outro lado possui um calendário mais curto. A diferença do prolongamento da VFA da mandioca no tempo, em relação a soja é a principal regra de separação entre elas.

Ainda na Figura 4.17, percebeu-se a semelhança espectro-temporal-visual de pixels de sorgo sacarino e de cana-de-açúcar, comentados na Figura 4.14. Estas culturas possuem manejos culturais muito parecidos e são de difícil

distinção nas imagens multitemporais utilizadas. Além do pixel de sorgo sacarino, pixels de floresta nativa se assemelham a áreas plantadas com cana-de-ano-e-meio, e pixels de capoeira e pastagem são similares aos pixels de transição, comentados na Figura 4.17. A baixa intensidade de VFA presente nos pixels de pastagem e capoeira faz com que estes pixels sejam espectralmente semelhantes aos aqueles pixels em que a cultura está adquirindo vigor vegetativo.

4.3.3. Roteiro de classificação visual

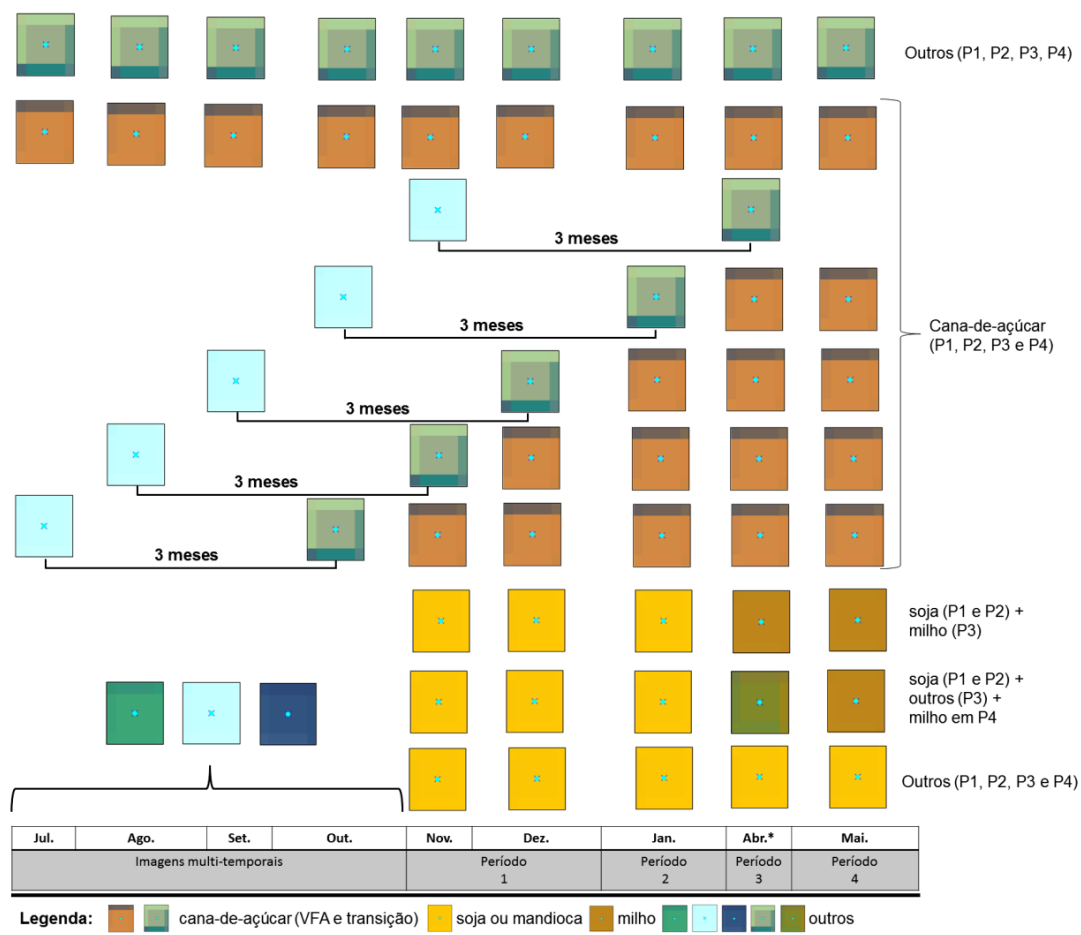
Ao analisar o CET da cana-de-açúcar dentro dos quatro períodos da estimativa agrícola, o Período 1 na Figura 4.14 foi o mais complexo para o intérprete, visto que neste período ocorreu a maior proporção de pixels que causaram dúvidas para uma mesma classe. Demais períodos (2,3 e 4) apresentaram apenas os dois rótulos de fácil separação, AVFA ou VFA.

O CET da soja se assemelhou ao da mandioca (outros) (Figura 4.16 e 4.17) e ao analisar estas duas culturas dentro dos quatro períodos da estimativa, apenas no período 4, estas culturas serão espectro-temporalmente distinguíveis. No Período 2, tanto a soja quanto a mandioca apresentaram intensa atividade fotossintética, o que torna estas duas classes de difícil separação visual. Existem semelhanças entre estas duas culturas nas demais imagens multitemporais anteriores ao Período 2. Demais informações, além da espectral devem ser utilizadas na separação destas duas culturas, no Período 2.

O CET do milho safrinha segue um padrão de fácil reconhecimento no tempo (Figura 4.16), isto é, o pixel de cor alaranjada em maio, posterior a um pixel amarelo em janeiro indica que o milho safrinha esteve ocupando o pixel, no Período 4 da estimativa. Este padrão pode ser também antecipado, isto é, a VFA de milho safrinha pode ser visualizada em pixels do Período 3 da estimativa. Porém, como a qualidade da imagem de Abril/2015 foi baixa, esta afirmação ainda depende de imagens com melhor qualidade que a utilizada no Período 3.

Em resumo, a Figura 4.18 mostra o roteiro de classificação visual para o conceito espectro-temporal visual de pixels de cana-de-açúcar, milho, soja e outros em imagens multitemporais OLI. Vale ressaltar que os pixels de transição foram considerados como pixels de cana-de-açúcar, apenas quando foi possível perceber a diferença de três meses entre a colheita da cana-de-açúcar (pixel com matiz ciano) e o pixel de transição (crescimento da cana-de-açúcar).

Figura 4.18 – Roteiro de classificação para os pixels representantes das classes de cana-de-açúcar, milho, soja e outros (transição e AVFA) na composição colorida R-b5 R-b6 B-b4.



4.3.4. Protocolo de classificação visual

Foram montados dois diferentes protocolos de classificação visual. O primeiro para classificação visual em *Laboratório de SIG pago* e a segunda para a classificação visual em interface produzido em *Laboratório Web*. A Figura 4.19

mostra a interface em Laboratório de SIG, isto é desenvolvido no ArcGIS; e as Figuras 4.20 e 4.21 mostram a interface em Laboratório *Web*, desenvolvido em ambiente *online*.

O protocolo de classificação em Laboratório em SIG foi definido aqui como resultado e segue a sequência 1 a 6.

Esta foi a sequência mais lógica para o intérprete classificar os pixels na tela do computador.

(1) Dois diferentes recortes das imagens OLI na composição colorida R-b5 G-b6 B-b4 para cada imagem multitemporal. Estes recortes mostram os pixels em escala 1:10.000 e 1:1.000. A maior escala mostra a cena em que o pixel está inserido e foi chamado de *Big Picture*. A menor escala mostra o pixel, e foi chamada de *Tiny Picture*.

(2) A sequência de imagens em *Tyni* e *BigPicture* mostra o comportamento do pixel no tempo, em ApEq e composição R-b5 G-b6 B-b4. O pixel deveria ser classificado como cana-de-açúcar, milho, soja ou outros na imagem do dia 21/abril/2015 respeitando o roteiro de classificação espectro-temporal das imagens OLI, descrito a cima.

(3) O produto GFC (Hansen *et al.*, 2014) foi utilizado para separação de pixels de borda, formados entre um talhão agrícola e um uso do solo qualquer presente na cena de sensoriamento remoto. Qualquer cor do pixel diferente de vermelho (Produto GFC) foi indicativo de que o pixel pertencia a classe outros.

(4) A declividade do TOPODATA (Valeriano, 2010) descrita pelas classes de declividade de Marques (1971) foi adequada para separação entre áreas aptas para agricultura (classe de declividade 0-2,5 % e 2,5-12 % de áreas não aptas para atividade agrícola, com declividade acima de 12,5 %. Aguiar *et al.* (2011) comentaram que acima de 12,5 % áreas de cana-de-açúcar colhidas com queima estão sendo utilizadas em outras atividade agrícolas, como plantio de culturas perenes.

Figura 4.19 – Interface de classificação desenvolvida em Laboratório SIG.

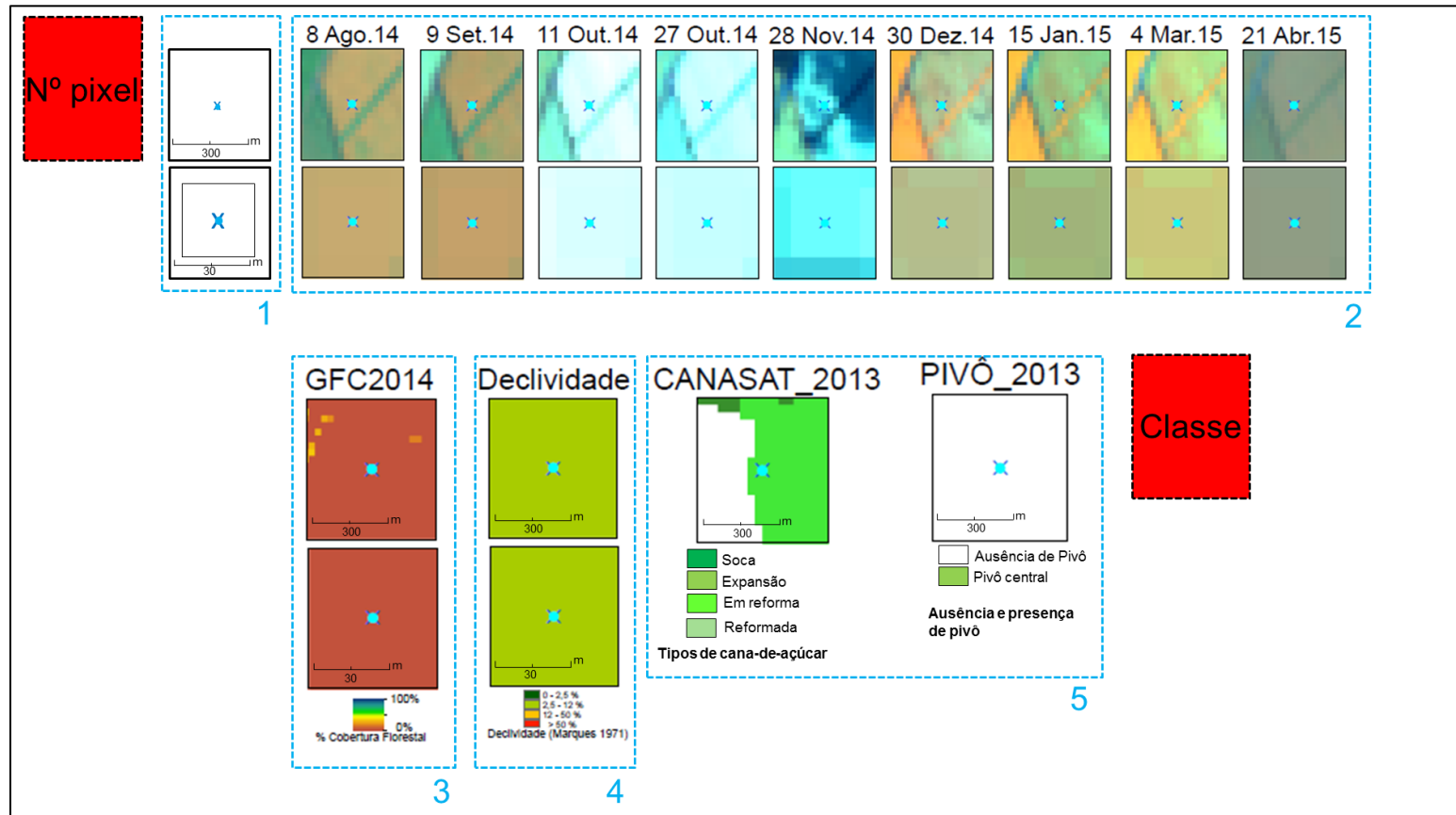


Figura 4.20 – Interface de classificação desenvolvida em Laboratório Web (... continua)

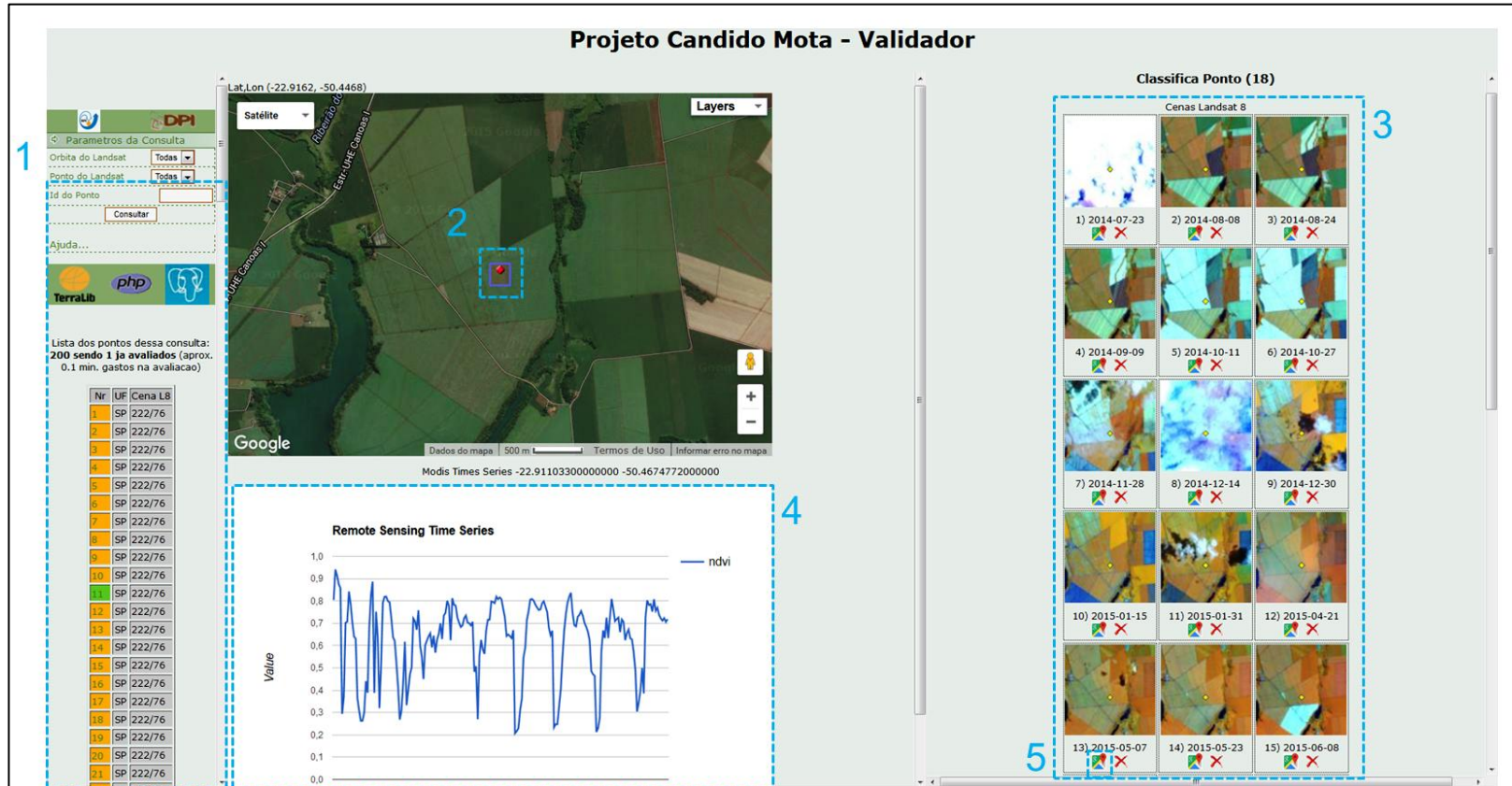
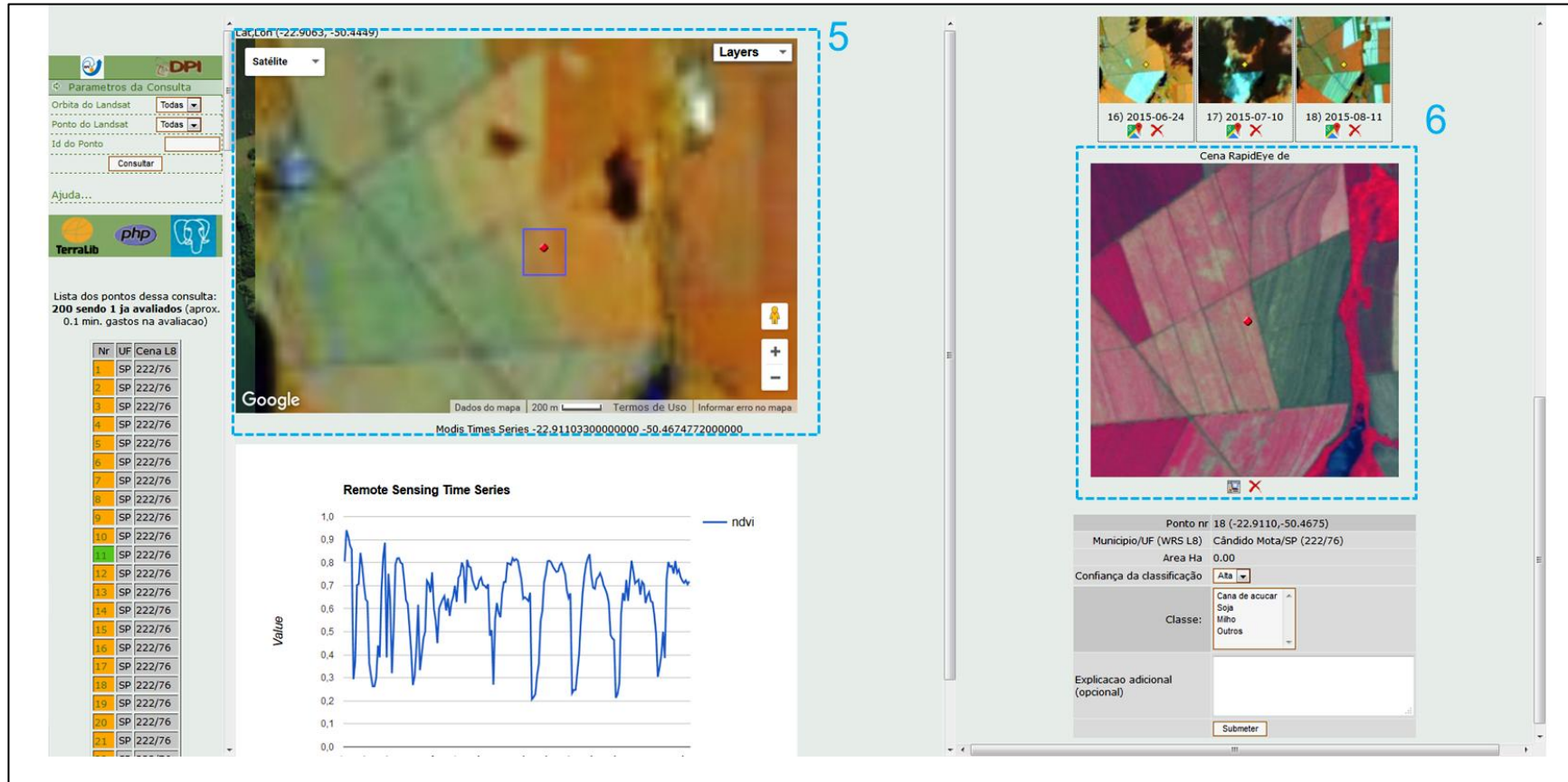


Figura 4.21 – Interface de classificação desenvolvida em Laboratório Web.



(5) Se o intérprete encontrar alguma dificuldade em classificar o pixels dados os tópicos acima, a história do pixel descrita nestes dois “quadrados” é de importância secundária, visto que serviram apenas para mostrar áreas que foram plantadas com cana-de-açúcar em 2012/2013 e foram manejadas sob regime de irrigação em pivô central em 2013. O calendário agrícola utilizado em áreas de pivô central segue padrão diferenciado de áreas não irrigadas, isto é, em áreas irrigadas é possível o plantio de mais de duas safras de diferentes culturas, em um mesmo ano safra.

O protocolo de classificação em Laboratório *Web* também foi definida como resultado e segue a sequência 1 à 6, ilustradas nas Figuras 4.20 e 4.21.

Esta sequência foi baseada no laboratório virtual projetado pelo Canasat para validação de áreas classificadas visualmente como cana-de-açúcar em imagens *Landsat-like* (ADAMI et al., 2011) e baseado também no classificador utilizado pelo Ministério do Meio Ambiente (MMA), para validar o projeto TerraClass Cerrado.

(1) No canto esquerdo superior da tela do classificador foi possível selecionar o pixel a ser classificado, pela aba “consultar”. O pixel a ser classificado foi então escolhido por meio de um clique sobre o número desejado. Ao clicar sobre o ponto, automaticamente foram mostradas as telas de 2 à 6.

(2) A imagem do *GoogleEarth* serviu de base para sobrepor o pixel OLI selecionado (indicador em vermelho) e o pixel MODIS (quadrado em azul). Pixels de borda entre áreas agrícolas das culturas de interesse e demais usos do solo puderam ser visualmente separados pelas imagens de altíssima resolução espacial do *GoogleEarth*.

(3) A série multitemporais de imagens OLI em ApEq na composição colorida R-b4 G-b5 B-b6 foi utilizada para iniciar a classificação dos pixels. O pixel deveria ser classificado como cana-de-açúcar, milho, soja ou outros na

imagem do dia 07/maio/2015 respeitando o roteiro de classificação espectro-temporal das imagens OLI, descrito na seção 4.3.3 (pg.153).

(4) Culturas que se assemelharam no aspecto visual e temporal dos dados OLI, tiveram que passar novamente por um novo processo de classificação. O uso da série temporal de 2007 a 2014 do NDVI foi utilizada neste procedimento. Os ciclos da curva temporal do NDVI das culturas temporárias é mais curto que o ciclo das culturas semi-perenes (RUDORFF et al., 2009; ADAMI et al., 2011). Desta forma, estes dois grandes grupos de culturas agrícolas podem ser diferenciados no tempo (ADAMI et al., 2011).

(5) Ao clicar sobre no número 5 da Figura 4.20, surgiu uma nova camada sobre a imagem *GoogleEarth* em 5, na Figura 4.21. Esta imagem é a imagem multitemporal base utilizada para classificação da cultura verde e em pé.

(6) A imagem *RapidEye* em composição colorida RGB 453 foi útil para visualização contextual dos talhões agrícolas, e separação de talhões de cana-de-açúcar das demais culturas. A imagem de 5 metros de resolução espacial delimita com mais precisão de detalhes os carregadores de dos talhões de cana-de-açúcar. O carregador é um indicativo do talhão, ser futuramente, utilizado para o plantio de cana-de-açúcar (SUGAWARA e RUDORFF, 2008).

4.3.5. Avaliação do uso dos protocolos de classificação visual

O resultado da classificação visual dos 200 pixels validados em campo foi apresentado da seguinte forma:

- (1) Foi obtida uma matriz de confusão para cada um dos intérpretes tanto para avaliação dos resultados obtidos pelo uso do protocolo 1, quanto para o protocolo 2. Lembrando que o protocolo 1 foi utilizado pelos Grupos 1 e 2, e o protocolo 2 foi utilizando apenas pelo Grupo 2. As matrizes de confusão foram quantificadas pela Exatidão Global (EG) e Índice Kappa (IK).

(2) Foram apresentados os recortes das imagens multitemporais e dos pixels que mais apresentaram discordâncias de classificação entre os intérpretes dos dois grupos, com a verdade de campo.

(3) Os Índices Kappa (IK) dos melhores resultados obtidos para o uso do protocolo 1 pelo grupo 1 e pelo Grupo 2; e para uso do protocolo 2 pelo grupo 2 foram comparados estatisticamente ao nível de 5 % de significância, pelo Teste de Z.

A Tabela 4.5 e 4.6 expõem as matrizes de confusão geradas pela classificação realizada pelo Grupo 1 e 2 respectivamente, ao usarem o protocolo de classificação 1. A EG dos quatro intérpretes do Grupo 1 variou de 96,65 % à 85,00 % e do Grupo 2, de 86,00 % à 65,83 %. Já era esperado que o resultado do Grupo 2 fosse inferior ao resultado do Grupo 1, no entanto o intérprete 4 da Tabela 4.5, obteve menor resultado de classificação que o intérprete 1 da Tabela 4.6.

Os intérpretes 1 e 2 do Grupo 2 possuíam o título de Engenheiros Agrícolas e o 3 e 4, o título de engenheiro florestal e geógrafo, respectivamente. A formação destes dois primeiros intérpretes teve influência no resultado de classificação, visto que demais fatores que poderiam influenciar a classificação dos intérpretes foram os mesmos. Assim, é perceptível que apenas duas horas de explicação sobre o protocolo de classificação 1, não foi suficiente para que os intérpretes 3 e 4 do Grupo 2, de diferentes formações profissionais conseguissem resultados similares aos intérpretes 1 e 2.

Os intérpretes do Grupo 1 erraram menos que os do Grupo 2, no entanto percebeu-se que existiu uma discrepância próxima a 20 pontos percentuais da classificação resultante do intérprete 1 do Grupo 1, para o 3 e 4 também do Grupo 1 (Tabela 4.5). Os erros cometidos por estes intérpretes (3 e 4) foram relacionados à classificação de pixels que pertenciam a classe outros, classificados como cana-de-açúcar e milho (Tabela 4.5). O mesmo ocorreu pelos intérpretes do Grupo 2, porém com maior frequência de acontecimentos (Tabela 4.6).

Tabela 4.5 – Matrizes de erro encontradas para quatro intérpretes do Grupo 1 que aplicaram o protocolo de classificação 1 para classificação visual de 200 pixels, nas classes cana-de-açúcar, milho, soja e outros, em imagens multitemporais OLI.

Grupo 1 – Protocolo 1		Cana	Milho	Soja	Outros	Total Linha	Usuário
Intérprete 1 - Exatidão Global = 96,65 %							
Referência	Cana	64	0	0	3	67	95,52
	Milho	0	40	0	1	41	97,56
	Soja	0	0	0	0	0	100,00
	Outros	2	1	0	89	92	96,74
	Total coluna	66	41	0	93	--	--
	Produtor	96,97	97,56	100,00	95,70	Kappa = 0,9448	
Intérprete 2 - Exatidão Global = 91,50 %							
Referência	Cana	59	3	0	5	67	88,06
	Milho	0	39	0	2	41	95,12
	Soja	0	0	0	0	0	100,00
	Outros	3	4	0	85	92	92,39
	Total coluna	62	46	0	92	--	--
	Produtor	95,16	84,78	100,00	92,39	Kappa = 0,8666	
Intérprete 3 - Exatidão Global = 87,00 %							
Referência	Cana	56	1	0	10	67	83,58
	Milho	5	31	0	5	41	75,61
	Soja	0	0	0	0	0	0,00
	Outros	2	3	0	87	92	94,57
	Total coluna	63	35	0	102	--	--
	Produtor	88,89	88,57	100,00	85,29	Kappa = 0,7917	
Intérprete 4- Exatidão Global = 85,00 %							
Referência	Cana	57	0	0	10	67	85,07
	Milho	4	33	0	4	41	80,49
	Soja	0	0	0	0	0	0,00
	Outros	5	6	0	80	92	86,96
	Total coluna	66	39	0	94	--	--
	Produtor	86,36	84,62	100,00	85,11	Kappa = 0,7631	
EG médio		90,04 %					

Tabela 4.6 – Matrizes de erro encontradas para quatro intérpretes do Grupo 2 que aplicaram o protocolo de classificação 1 para classificação visual de 200 pixels, nas classes cana-de-açúcar, milho, soja e outros, em imagens multitemporais OLI.

Grupo 2 – Protocolo 1		Cana	Milho	Soja	Outros	Total Linha	Usuário
Intérprete 1 - Exatidão Global = 86,00 %							
Referência	Cana	56	4	0	7	67	83,58
	Milho	0	29	0	12	41	70,73
	Soja	0	0	0	0	0	100,00
	Outros	2	3	0	87	92	94,57
	Total coluna	58	36	0	106	--	--
	Produtor	96,55	80,56	100,00	82,08	Kappa = 0,7750	
Intérprete 2 - Exatidão Global = 72,50 %							
Referência	Cana	56	5	0	6	67	83,58
	Milho	0	38	0	3	41	92,68
	Soja	0	0	0	0	0	100,00
	Outros	4	37	0	51	92	55,43
	Total coluna	60	80	0	60	--	--
	Produtor	93,33	47,50	100,00	85,00	Kappa = 0,5953	
Intérprete 3 - Exatidão Global = 67,00 %							
Referência	Cana	57	7	0	3	67	85,07
	Milho	7	31	0	3	41	75,61
	Soja	0	0	0	0	0	0,00
	Outros	27	19	0	46	92	50,00
	Total coluna	91	57	0	52	--	--
	Produtor	62,64	54,39	100,00	88,46	Kappa = 0,5071	
Intérprete 4 - Exatidão Global = 65,83 %							
Referência	Cana	57	5	0	4	66	86,36
	Milho	1	40	0	0	41	97,56
	Soja	0	0	0	0	0	0,00
	Outros	20	38	0	34	92	36,96
	Total coluna	78	83	0	38	--	--
	Produtor	73,08	48,19	100,00	89,47	Kappa = 0,5089	
EG médio		72,83 %					

A Figura 4.22 mostra as classificações resultantes de cada um dos intérpretes e a verdade de campo. Foi observado um padrão nos erros cometidos pelos intérpretes dos dois diferentes grupos. Os pixels número 122, 193, 194 e 198 na Figura 4.22 foram considerados os mesmos erros cometidos entre os dois Grupos. Os pixels 193 e 194 foram erroneamente classificados pelos quatro

intérpretes do Grupo 1 e os pixels 8, 15, 59, 77, 138 e 184 foram mal interpretados pelos quatro intérpretes do Grupo 2.

Figura 4.22 – Lista dos pixels classificados pelos oito intérpretes (Grupo 1 – 1, 2, 3 e 4; Grupo 2 – 1, 2, 3 e 4) que aplicaram o Protocolo de classificação visual 1 para interpretação de pixels das classes outros (verde), cana-de-açúcar (vermelho) e milho (amarelo). A coluna Rf. na Figura, refere-se aos dados coletados em campo.

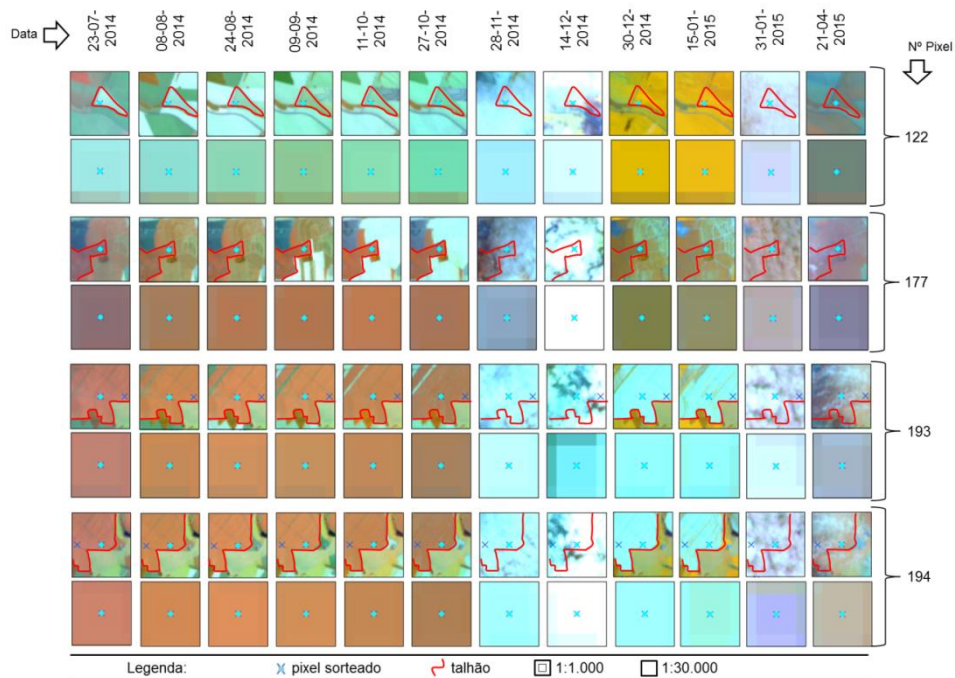
Protocolo de classificação 1												
Grupo 1				Grupo 2								
Outros				Cana-de-açúcar				Milho				
Px.	Rf.	1	2	3	4	Px.	Rf.	1	2	3	4	
1						7						
2						10						
3						11						
5						13						
6						16						
8						18						
9						20						
12						23						
15						24						
17						26						
19						28						
21						29						
25						32						
27						35						
30						42						
31						43						
34						47						
36						53						
37						59						
39						64						
40						67						
41						69						
46						76						
48						77						
51						78						
55						80						
56						81						
58						83						
60						85						
72						89						
74						91						
75						92						
79						100						
82						101						
86						103						
88						104						
90						110						
95						114						
97						120						
98						121						
99						124						
102						125						
105						127						
107						130						
108						134						
109						138						
111						144						
112						145						
113						149						
115						150						
117						155						
118						156						
119						159						
122						160						
123						162						
126						163						
129						167						
132						174						
133						176						
135						179						
136						180						
139						181						
140						184						
141						185						
142						191						
143						193						
146						194						
151												
152												
153												
154												
157												
164												
165												
166												
168												
169												
170												
172												
175												
177												
183												
186												
187												
188												
189												
195												
196												
197												
199												
200												

Legenda:

- cana-de-açúcar
- milho
- outros
- ~ mesma dúvida constatada entre grupos
- ~ mesma dúvida constatada dentro do grupo

As Figuras 4.23, 4.24 e 4.25 mostram os recortes de imagens multitemporais OLI em ApEq com duas escalas de observação, 1:1.000 e 1:30.000, associadas a cada um destes pixels citados a cima.

Figura 4.23 – Recortes de imagens multitemporais OLI em ApEq com duas escalas de observação, 1:1.000 (*tyni picture*) e 1:30.000 (*big picture*).



Estes recortes ilustram os pixels que foram erroneamente classificados pelos intérpretes devido à baixa qualidade da imagem de 21/04/2015.

Ao observar o tipo de dúvida que poderia ser relacionada a cada um destes erros cometidos pelos intérpretes responsáveis pela classificação visual dos pixels, eles foram divididos em três categorias:

- (1) Baixa qualidade visual da imagem na data pretendida para classificação (Figura 4.23);
- (2) Pixels de borda ou de transição (Figura 4.24);
- (3) Pixel presente em áreas agrícolas com falhas de plantio (Figura 4.25).

O pixel 122 pertencente à classe outros, foi classificado como milho pela grande maioria dos intérpretes do Grupo 1 e 2 (Figura 4.22). Já o pixel 177 considerado classe outros pela referência foi classificado como cana-de-açúcar por todos os intérpretes dos dois grupos; e os pixels 193 e 194 ambos considerados classe cana-de-açúcar, foram erroneamente classificados como outros pelo Grupo 2 (Figura 4.22).

Figura 4.24 – Recortes de imagens multitemporais OLI em ApEq com duas escalas de observação, 1:1.000 (*tyni picture*) e 1:30.000 (*big Picture*).

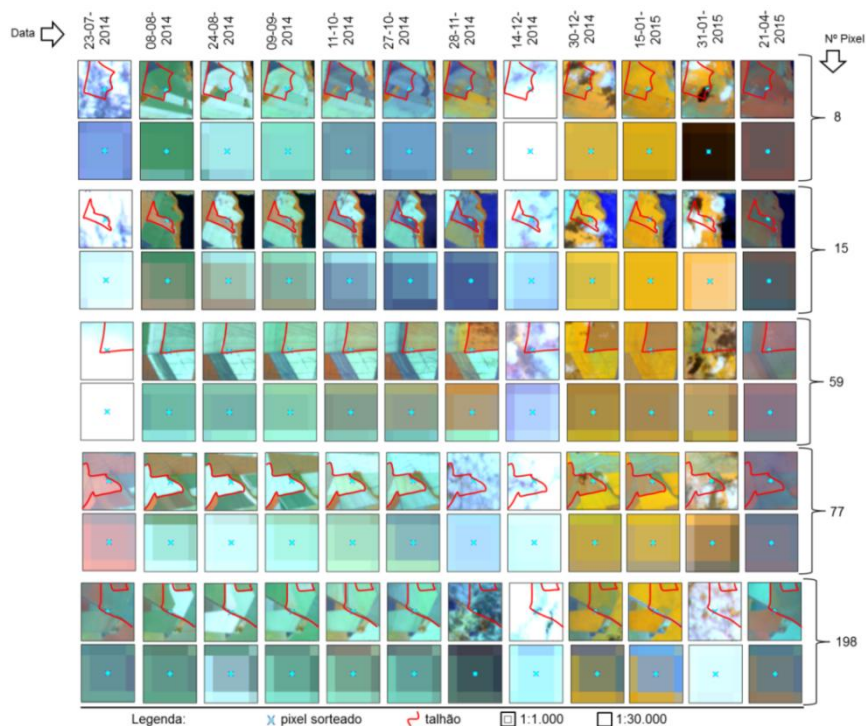
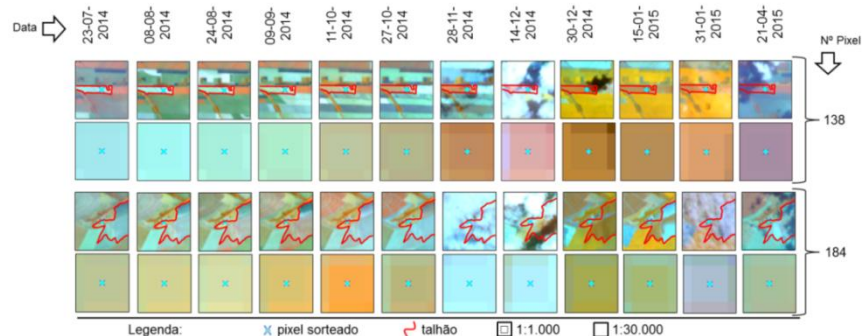


Figura 4.25 – Recortes de imagens multitemporais OLI em ApEq com duas escalas de observação, 1:1.000 (*tyni picture*) e 1:30.000 (*big picture*).



Estes recortes ilustram os pixels que foram erroneamente classificados pelos intérpretes devido à heterogeneidade espectral interna, do talhão agrícola.

Estes erros foram associados ao efeito causado pela qualidade da imagem utilizada para classificar a cultura verde e em pé pelos classificadores visuais. A existência da sombra no pixel 122 impossibilitou os intérpretes em separar o pixel ainda considerado com transição de VFA, da classe milho. Nos demais casos, a nuvem encobriu os pixels (Figura 4.24) e causou subjetividade para os intérpretes seguirem o protocolo de classificação, em Laboratório SIG.

Os pixels da Figura 4.23, além de terem séria influência de sombra sobre a imagem do dia 21/04/2015, o que gerou dúvidas no momento da interpretação, têm outros dois fatores de relevância que geraram subjetividade na classificação. A primeira delas refere-se ao pixel de transição já comentado no parágrafo acima, e o outro faz referência ao pixel de bordadura. Este mesmo problema foi relatado por Eberhardt (2014) para classificação visual de pixels de bordadura soja/floresta nativa o Rio Grande do Sul. Um exemplo disso foi evidenciado no pixel 77 (Figura 4.23), considerado borda entre um talhão de cana-de-açúcar/floresta nativa.

No caso de São Paulo, ao utilizar o protocolo de classificação 1, o maior problema na classificação visual foi encontrado na separação visual entre pixels de borda localizados entre dois diferentes talhões de milho, com diferentes idades (pixels 8, 15, 59, 77 e 198 da Figura 4.24). A dúvida dos intérpretes, principalmente do Grupo 2 foi se o pixel de borda seria considerado pixel do talhão de milho, com plantio mais recente (transição) ou considerado pixel com VFA (milho). Os intérpretes deveriam considerar os tamanhos dos talhões agrícolas (EBERHARDT, 2014). Nestes casos mostrados na Figura y, os pixels de borda foram considerados como pertencentes ao talhão de maior área em cada janela, em escala 1:30.000.

Os pixels da Figura 4.25 mostram a dificuldade do intérprete em associar uma classe a pixels que pertencem a talhões com falhas de plantio. Este problema ainda não foi solucionado para montagem de um novo protocolo de classificação, no entanto, demais dúvidas de classificação puderam ser novamente testadas quanto ao uso do protocolo de classificação 2, pelo Grupo 2.

Ao utilizar o Protocolo de classificação 2, o Grupo 2 obteve resultados de classificações superiores (Tabela 4.7) aos encontrados pelo Grupo 1 (Tabela 4.5) e a eles mesmos (Tabela 4.6), ao utilizar o protocolo 1. A exatidão Global variou de 88 a 95 %, o que comprova o avanço do uso do protocolo de classificação 2 para classificação dos pixels na imagem 07/junho/2015.

Tabela 4.7 – Matrizes de erro encontradas para quatro intérpretes do Grupo 2 que aplicaram o protocolo de classificação 2 para classificação visual de 200 pixels, nas classes cana-de-açúcar, milho, soja e outros, em imagens multitemporais OLI.

Grupo 2 – Protocolo 2		Cana	Milho	Soja	Outros	Total Linha	Usuário
Intérprete 1 - Exatidão Global = 96.50 %							
Referência	Cana	68	0	0	2	70	97.1
	Milho	1	74	0	2	77	96.1
	Soja	0	0	0	0	0	100.0
	Outros	2	0	0	51	53	96.2
	Total coluna	71	74	0	55	--	--
	Produtor	95.8	100.0	100.0	92.7	Kappa = 0,9470	
Intérprete 2 - Exatidão Global = 95.00 %							
Referência	Cana	67	2	0	1	70	95.7
	Milho	2	73	0	2	77	94.8
	Soja	0	0	0	0	0	100.0
	Outros	1	2	0	50	53	94.3
	Total coluna	70	77	0	53	--	--
	Produtor	95.7	94.8	100.0	94.3	Kappa = 0,9241	
Intérprete 4 - Exatidão Global = 94.50 %							
Referência	Cana	66	1	0	3	70	94.3
	Milho	2	73	0	2	77	94.8
	Soja	0	0	0	0	0	100.0
	Outros	0	3	0	50	53	94.3
	Total coluna	68	77	0	55	--	--
	Produtor	97.1	94.8	100.0	90.9	Kappa = 0,9167	
Intérprete 3 - Exatidão Global = 92.00 %							
Referência	Cana	65	2	0	3	70	92.9
	Milho	2	72	0	3	77	93.5
	Soja	0	0	0	0	0	100.0
	Outros	4	2	0	47	53	88.7
	Total coluna	71	76	0	53	--	--
	Produtor	91.5	94.7	100.0	88.7	Kappa = 0,8786	
EG médio		94,50 %					

A imagem de junho mostrou-se com uma qualidade superior à imagem utilizada para classificação dos pixels no Protocolo 1. Além disso, os ótimos resultados de classificação encontrados para todos os intérpretes testados quanto ao protocolo 2, mostra que o classificador utilizado em laboratório *online* é usual para esta aplicabilidade. Adami *et al.* (2013) e TERRACLASS (2015) comentaram que o uso do laboratório online foi de grande importância para validação de mapas temáticos produzidos visualmente.

Para o protocolo de classificação 2, não foi possível analisar um padrão nos erros cometidos pelos intérpretes, como ocorrido anteriormente para o uso do protocolo de classificação 1 (como mostrado na Figura 4.22). Esta falta de padrão fez com que não fosse possível analisar com maiores detalhes a interferência de fatores sobre a classificação executada pelo grupo 2.

Não foi possível obter 100 % de EG com o uso do protocolo de classificação 2, porém percebeu-se que os resultados obtidos por um grupo formado por pessoas com diferentes formações, numa segunda oportunidade de classificação, obtiveram resultados superiores a um grupo formado por pessoas que tinham conhecimento prático sobre classificação visual de imagens de sensoriamento remoto.

Desta forma, para satisfazer o objetivo desta tese, de classificar os pixels visualmente em laboratório, sem a necessidade de frequentes trabalhos de campo, os resultados são satisfatórios para que o método seja aplicado em escala estadual. No entanto, existe a necessidade do treinamento aplicado ao intérprete anteriormente a classificação visual.

Ao comparar estatisticamente os resultados dos IKs encontrados pelo Intérprete 1 do Grupo 1 com o resultado do IK encontrado para o Intérprete 1 do Grupo 2, também ao aplicar o Protocolo 1, ocorreu diferença significativa ao nível de 5 % de significância entre o uso deste mesmo protocolo (Tabela 4.8). Este resultado mostra que o protocolo de classificação 1, não deve ser utilizado por pessoas que não têm conhecimento sobre classificação de imagens multitemporais *Landsat-like*, para identificação de culturas agrícolas.

Tabela 4.8 – Testes para diferença estatística entre os resultados dos Kappas encontrados pelos intérpretes dos dois grupos com melhores resultados de exatidão global.

Comparação Estatística			Z	pvalor
Int. 1 Grupo 1 P1	(vs.)	Int. 1 Grupo 2 P1	3.85	0.0001 *
Int. 1 Grupo 1 P1	(vs.)	Int. 1 Grupo 2 P2	0.08	0.4685 ^{ns}
Int. 1 Grupo 2 P1	(vs.)	Int. 1 Grupo 2 P2	3.94	0.0001 *

Onde: Int. (intérprete), P1 e P2 (Protocolo 1 e 2, respectivamente), * significativo ao nível de 5 % de significância e ^{ns} não significativo ao nível de 5 % de significância.

No entanto, ao aplicar o protocolo 2 por pessoas que não possuíam conhecimento *a priori* sobre classificação visual de imagens *Landsat-like*, elas obtiveram os mesmos resultados encontrados por pessoas treinadas para classificação visual pelo protocolo de classificação 1. Isto comprova que o protocolo 2 pode ser usado por pessoas de um órgão agrícola qualquer para classificação de pixels em imagens *Landsat-like* objetivada à estimativa agrícola.

Além disso, os resultados de EG médios dos protocolos de classificação 1 e 2 aplicados pelos grupos de 72,83 % a 94,5 %, respectivamente. Eberhardt (2014) ao classificar visualmente soja e milho no Rio Grande do Sul obteve EG médio = 95,5 %. Resultado 1 % superior ao encontrado neste presente trabalho.

Eberhardt (2014) garantiu estes resultados em ambiente computacional de SIG pago, por pessoas que conheciam sobre a classificação visual de imagens de sensoriamento remoto e, além disso, o trabalho foi executado em uma área composta basicamente por soja. No presente trabalho, o melhor resultado foi obtido por pessoas que não tinham conhecimento *a priori* sobre classificação visual de imagens multitemporais *Landsat-like* e, além disso, os pixels foram classificados em área agrícola composta por duas culturas consideradas espectralmente similares (cana-de-açúcar e milho), fato este que torna a classificação visual mais difícil.

4.4. Classificação automática

A apresentação dos resultados da classificação automática se deu de quatro diferentes formas:

- (1) Resultados encontrados para os modelos de classificação automáticos gerados por OBIA/*Radom Forest* para classificação automática dos 200 pixels validados em campo;
- (2) Avaliação da robustez do melhor modelo de classificação RF gerado para classificar pixels independentes daqueles usados no treinamento;
- (3) Avaliação do resultado de área encontrado para cana-de-açúcar pelo melhor modelo de classificação, comparado com os dados de área plantada, disponibilizados pelo órgão oficial, responsável pelas estimativas agrícolas do Brasil.
- (4) Observação da classificação automática final de uma área teste comparada com o resultado dado pela fotointerpretação desta mesma área teste.

4.4.1. Exatidão dos modelos de classificação

As exatidões globais (EG) dos 1150 modelos de classificação validados pela técnica *out-of-bag* (OOB), após a seleção de atributos, apresentaram um desvio para o lado direito da curva de distribuição normal (Figura 4.26). A EG variou de 71 % a 90 % (mediana = 95 %). Se os parâmetros de segmentação não fossem apropriadamente escolhidos, seria possível a obtenção de menores resultados de EG da classificação. Por exemplo, quase metade das EGs produzidas entre 70 % e 99 % obteve menos de 13 % de acerto que os demais modelos gerados. A Tabela 4.9 mostra o resumo estatístico para o Kappa e a EG destes 1150 modelos de classificação.

Figura 4.26 – Distribuição da exatidão global OOB dos 1150 modelos de classificação RF baseados em diferentes segmentações.

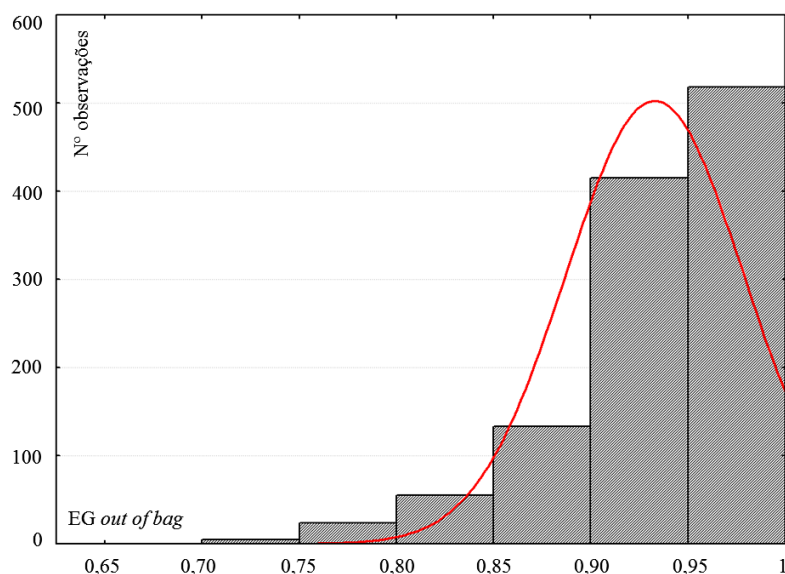


Tabela 4.9 – Resultados mínimos, medianos e máximos de exatidão global (EG) e coeficiente Kappa para 1150 segmentações e classificações (validação de 200 pixels pela técnica *out-of-bag*)

	Mínimo	Mediana	Máximo
EG	0.706667	0.945055	0.98995
Kappa	0.556452	0.915422	0.98462

Uma visão geral correspondente à variação dos 1150 diferentes modelos de classificação pode ser encontrada na Tabela 4.10. Os mais altos valores de exatidão (exatidão do usuário) foram obtidos para todas as classes. Em média, os pixels de milho foram os que geraram mais dúvidas para os modelos de classificado RF, seguido da classe outros e cana-de-açúcar.

Tabela4.10 – Estatística descritiva da exatidão global (EG) e coeficiente Kappa para 1150 segmentações e classificações (validação de 200 pixels pela técnica *out-of-bag*)

	Exatidão do Produtor			Exatidão do Usuário		
	Mínimo	Mediana	Máximo	Mínimo	Mediana	Máximo
Cana-de-açúcar	0.667	0.963	1.000	0.724	0.964	1.000
Milho	0.435	0.886	0.981	0.654	0.980	1.000
Outros	0.667	0.947	1.000	0.625	0.913	1.000

Entre as 1150 segmentações/classificações, o modelo que obteve o mais alto valor de EG foi identificado. O melhor resultado foi obtido quando os

parâmetros do algoritmo segmentador multirresolução (SM-R) foram: fator de escala =25, forma = 30 e compacidade = 10. Este modelo incluiu apenas três atributos (depois da seleção de atributos) dos 74 atributos utilizados como dados de entrada para formação do modelo de conhecimento RF. Os atributos da fração de vegetação da imagem OLI de janeiro e maio e NDSVI de outubro foram os três atributos selecionados para classificação dos 200 pixels pelo modelo de conhecimento.

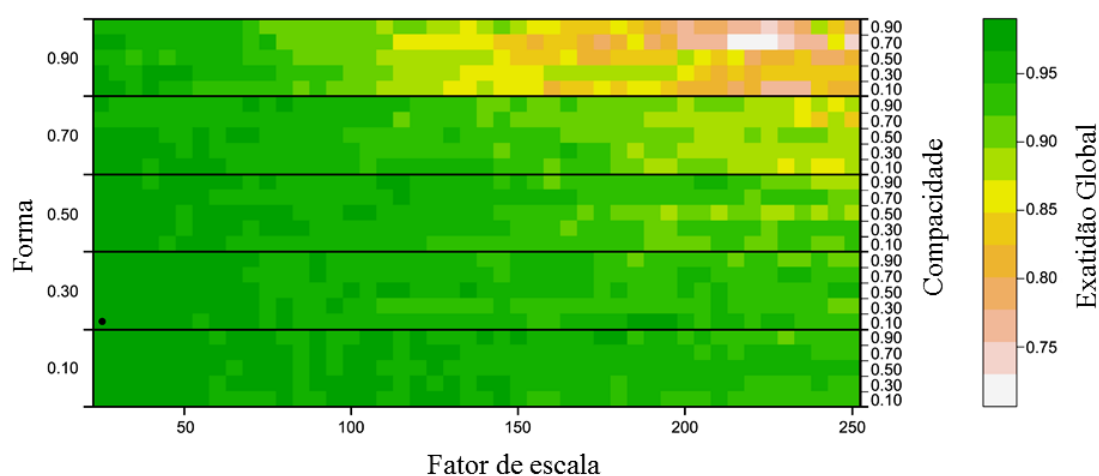
A fração vegetação esteve relacionada à vegetação fotossinteticamente ativa (VFA) das culturas agrícolas. Neste presente casado, com a VFA da soja em janeiro e do milho em maio. A presença de VFA decorrente ao plantio da soja em Janeiro foi um forte indicativo usado pelo classificador para classificar a VFA, em maio, como pixel pertencente à classe milho.

O NDSVI é um índice de vegetação relacionado com o estágio de senescência da cultura ou sobre a presença da palha ou da planta em estágio de seca (PEÑA *et al.*, 2014). Normalmente, a colheita de cana-de-açúcar é realizada até o mês de novembro (AGUIAR *et al.*, 2011). Nesta mesma época, áreas com culturas agrícolas temporárias estão em formação (VIERA *et al.*, 2012; SCHULTZ *et al.*, 2015a). Pixels referentes à colheita da cana-de-açúcar em outubro foram um indicativo para que o classificador selecionasse este pixel entre os demais e rotulasse como classe cana-de-açúcar a VFA presente em maio.

Schultz *et al.* (2015a) comentaram que apenas duas imagens localizadas em agosto e janeiro e 25 atributos não foram suficientes para classificar cana-de-açúcar, soja, mandioca, amendoim e outros, no Sul do estado de São Paulo. Viera *et al.* (2012) comentaram imagens de setembro, outubro, fevereiro e março e 9 atributos foram suficientes para classificar cana-de-açúcar no norte do estado de São Paulo. No presente trabalho, as três imagens multitemporais (outubro, janeiro e maio) foram eficientes na separação das três classes de interesse (cana-de-açúcar, milho e soja) usando apenas três atributos. Isto mostra a dependência de imagens multitemporais na classificação de culturas agrícolas em São Paulo usando OBIA e seleção de atributos.

A combinação de vários parâmetros de segmentação usados no algoritmo SM-R mostrou um distinto padrão com uso das exatidões produzidas pelo método OOB (Figura 4.27). A compacidade teve um impacto global muito menor sobre os dados se comparado aos parâmetros fator de escala e forma. A compacidade está relacionada aos alvos de bordas fractais, diferentemente daquelas bordas criadas pela ação antrópica (BAATZ e SCHÄPE, 2000).

Figura 4.27 – Exatidão Global (OOB) para todas as combinações dos parâmetros fator de escala, forma e compacidade, testados com os 200 pixels validados em campo.

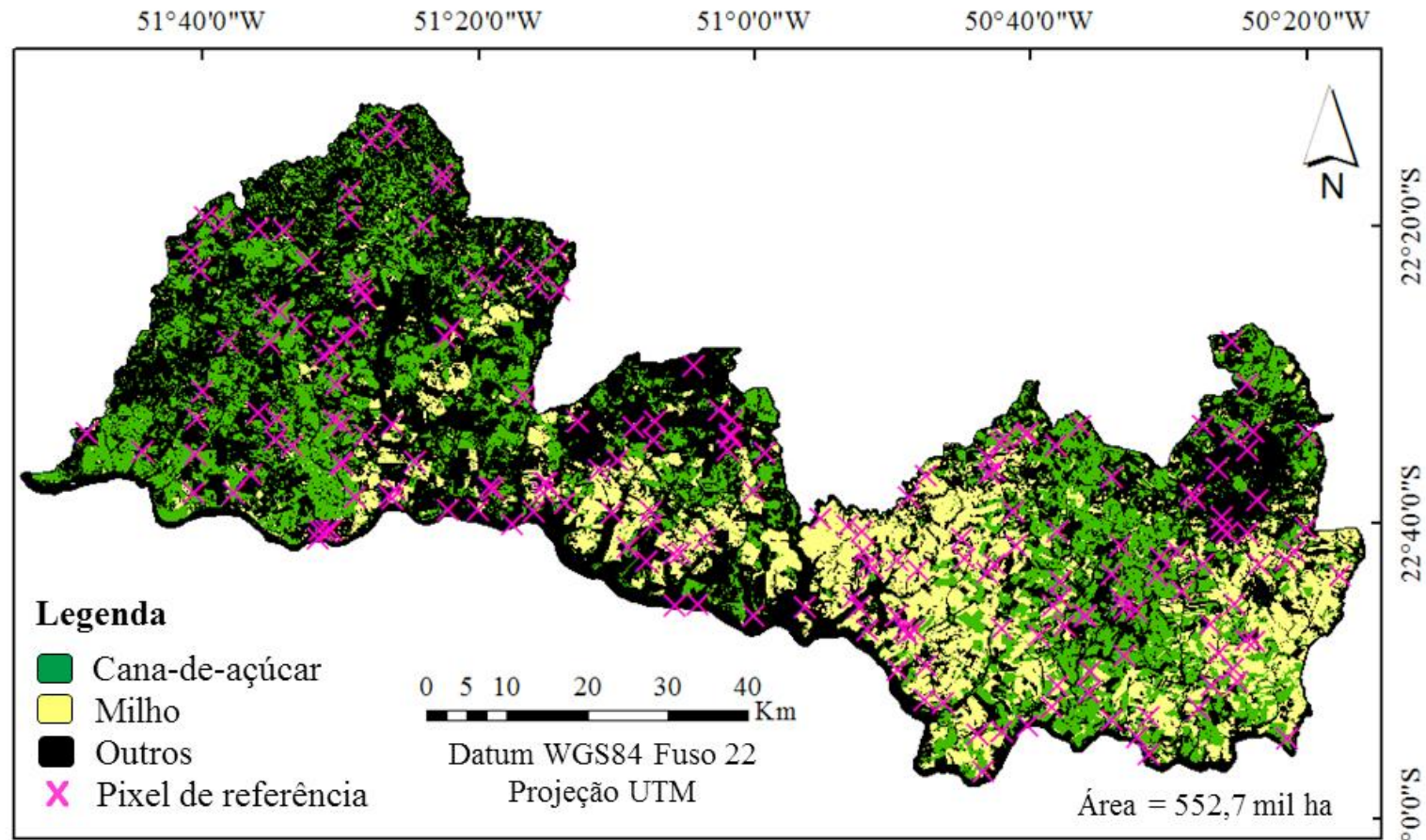


Para visualização dos resultados, cada quadrado representa um resultado individual formado pela combinação dos parâmetros e cores. O ponto em preto indica o melhor resultado de exatidão global dada pela combinação entre fator de escala = 25, forma = 30 e compacidade = 10.

4.4.2. Avaliação da robustez do modelo de classificação

O modelo que apresentou o melhor resultado de classificação pela validação OOB foi replicado para dados independentes daqueles usados para treinamento. Foram classificados visualmente 200 pixels aleatórios sobre 15 municípios do estado de São Paulo, localizados ao norte da imagem 222/076. O mapa temático (Figura 4.28) gerado para esta área, em resolução espacial de 30 x 30 m serviu de base para comparação entre a verdadeira classe dos 200 pixels independentes (pixels de referência), com os pixels classificados automaticamente pelo modelo replicado.

Figura 4.28 – Resultado da classificação obtida sobre a área de estudo usando a segmentação ótima para o modelo de classificação identificado pela validação OOB (e.g., fator de escala = 25, forma = 30 e compacidade = 10).



As cruzes destacadas em rosa indicam os pixels que foram visualmente interpretados utilizando a chave de classificação descrita na seção de resultados.

A matriz de confusão gerada entre o resultado dos pixels de referência e os pixels classificados automaticamente pelo melhor modelo de classificação encontrado é mostrada na Tabela 4.11.

Tabela 4.11 – Matriz de confusão formada com os dados de referência e os dados classificados automaticamente pelo melhor modelo de classificação RF selecionado.

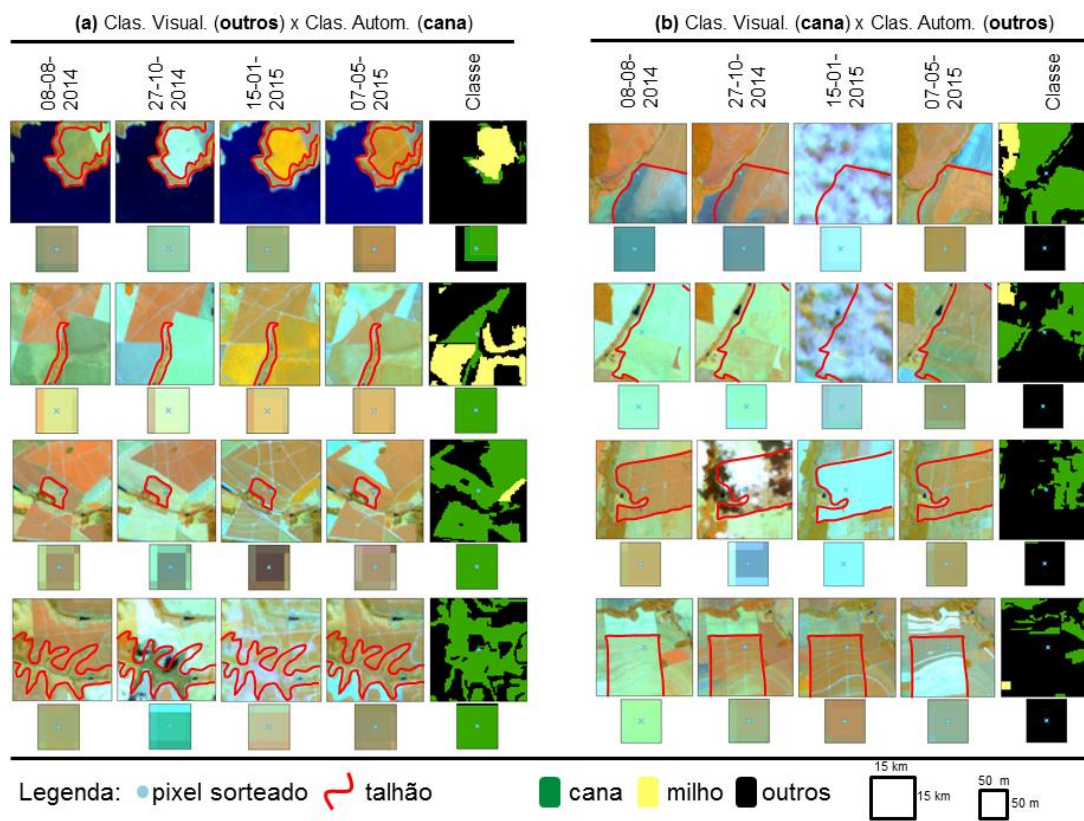
Melhor modelo		Cana	Milho	Soja	Outros	Total Linha	Usuário
Class. Automática - Exatidão Global = 86.50 %							
Referência	Cana	34	3	0	11*	48	70.83
	Milho	2	40	0	4	46	86.96
	Soja	0	0	0	0	0	0.00
	Outros	7**	0	0	99	106	93.40
	Total coluna	43	43	0	114	--	--
Produtor		79.07	93.02	0.00	86.84	Kappa = 0,77	

A exatidão global (EG) do mapeamento foi igual a 86,5 % (kappa = 0,77). Ao comparar este resultado de EG aos resultados encontrados pelos 1150 modelos de classificação validados aqui pela técnica OOB, ele está muito abaixo dos EGs encontrados por estes modelos. No entanto, este valor de EG se aproxima dos resultados encontrados por Schultz *et al.* (2015b), Vieira *et al.* (2012), Peña *et al.* (2014) e Long *et al.* (2013). Todos estes autores destacados utilizaram OBIA para classificação de suas áreas agrícolas e validaram os resultados utilizando pixels aleatórios. O resultado de EG encontrado por estes autores variou de 80% a 92 %.

Na Tabela 4.11, o maior número de erros de classificação automática cana-de-açúcar outros (* no total de 11 pixels) e também, ao classificar a classe cana-de-açúcar como outros (** no total de sete pixels). Demais erros foram inferiores a 4 pixels por classe (Tabela 4.11).

A Figura 4.29 ilustra recortes de imagens multitemporais OLI, em composição colorida R-b5 G-b6 B-b4 para as datas 08/08/2014, 27/10/2014, 15/01/2015 e 07/05/2015; e a classe (em cores) automaticamente selecionada para o pixel pelo melhor modelo de classificação RF.

Figura 4.29 – Erros de classificação gerados por OBIA na classificação dos 200 pixels de referência foto interpretados.



A coluna da esquerda (a) ilustra três erros cometidos pelo classificador automático ao classificar a classe outros como cana, e a da direita (b) segue o caminho inverso, classificou outros como cana.

Em todos os exemplos ilustrados na Figura 4.29.a, os pixels estavam localizados entre a borda de um talhão de cana-de-açúcar e de um alvo considerado como classe outros. Mesmo na classificação visual, estes pixels de borda são de difícil classificação. Para estes casos, regras de atribuição devem ser seguidas, para classificação visual destes pixels.

Em OBIA, regras de atribuição poderiam ser aplicadas ao usar dados de base (*baselines*), para guiar a segmentação e classificação apenas em áreas de interesse agrícola. A *baseline* serviria como filtro dos dados de entrada a serem classificados (WALDENER *et al.*, 2015; MATON *et al.*, 2015).

Na Figura 4.29.b, os pixels estavam localizados em áreas com cana-de-açúcar, no entanto foram classificados como outros. Nesta Figura as áreas seguiram diferentes padrões de comportamento da colheita. Então, é perceptível que as áreas de cana-de-açúcar apresentaram alta variabilidade intraclasse pelas imagens selecionadas. Mais pixels deveriam ser sorteados para que a classe tenha maior representatividade amostral e assim o modelo seja replicado em uma área com dados independentes dos dados de treinamento. Viera *et al.* (2012) comentaram que para uma áreas de 124 mil hectares foram necessários 396 segmentos aleatórios para treinamento do modelo de replicação e posterior replicação.

4.4.3. Avaliação do mapa temático

A imagem 222/076, localizada no extremo Sul do estado de São Paulo recobriu 15 municípios do estado, foram eles: Anhumas, Assis, Cândido Mota, Cruzália, Estrela do Norte, Florínia, Iepê, Maracaí, Nantes, Narandiba, Pedrinhas Paulistas, Piraporozinho, Taciba, Tarabai e Tarumã. A área total destes 15 municípios foi de 552,7 mil ha. A Figura 4.28 mostra o mapa temático obtido pelo melhor modelo RF ao classificar cana-de-açúcar e milho, na área de estudo.

A Tabela 4.12 mostra em resumo a área encontrada no mapa temático gerado automaticamente pelo modelo de classificação RF e os dados oficiais do IBGE, para o ano safra 2014/2015. O modelo de classificação RF estimou em 241.330,7 ha a área de milho e soja, enquanto que o número informado pelo IBGE, para estas duas culturas foi de 265.640 ha. A diferença constatada pela estimativa destes dois métodos foi de 9,15 % a mais para os dados do IBGE.

Tabela 4.12 – Área ocupada com a cultura da cana-de-açúcar e milho obtida por meio da classificação automática pelo melhor modelo RF em comparação à fornecida pelo IBGE, para a área de estudo.

Cultura	Mapa Temático	IBGE	Dif. (MT - IBGE)	Dif. (%)
Cana-de-açúcar	144934.89	164386.00	-19451.11	-11.8326
Milho	96395.76	101254.00	-4858.24	-4.79807
Total	241330.66	265640.00	-24309.34	-9.15124

O uso de imagens de Maio para informar área ocupada com cana-de-açúcar não foi suficiente para que todas as áreas destas culturas fossem mapeadas pelo modelo automático. Rizzi e Rudorff (2005) obtiveram diferença de -11,3 % para soja no Rio Grande do Sul, quando a área classificada automaticamente sobre imagens Landsat foi comparada com os dados oficiais do IBGE. Os autores comentaram que ao utilizar apenas imagens de Janeiro para classificação, áreas de soja de ciclo precoce foram omitidas nesta imagem. Desta forma, uma maior frequência de revisitas pode auxiliar o mapeamento automático de culturas de ciclo semi-perene (cana-de-açúcar).

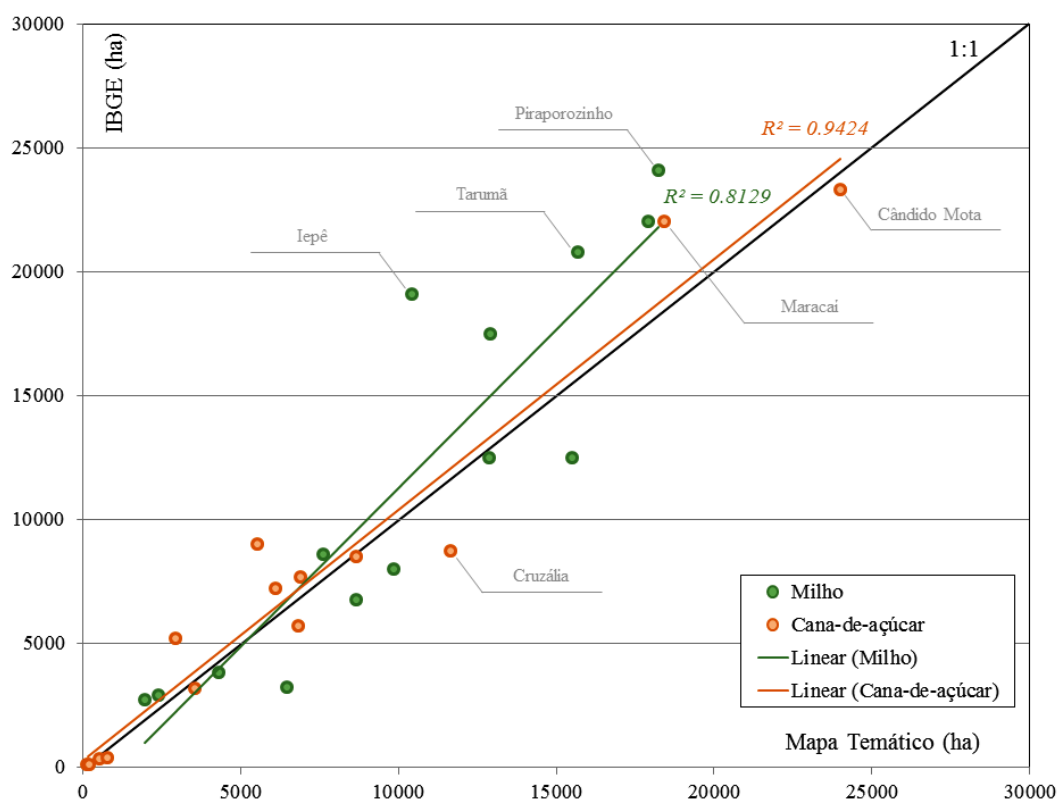
No caso do milho, os municípios localizados ao extremo Sul do estado de São Paulo tradicionalmente plantam o milho safrinha, depois da soja. O plantio de milho safrinha corresponde a 78 % do plantio de milho nestes municípios (CEPEA, 2015). Este fato possibilita argumentar que a imagem de Maio foi suficiente para informar áreas ocupadas com milho, na área de estudo.

Ainda na Tabela 4.12, a maior diferença constatada por classe, para as duas estimativas se deu para a classe cana-de-açúcar (11,8 %). Quando os dados de área informada pelo IBGE e pelo mapa temático foram plotados em gráfico de regressão linear, o R^2 (coeficiente de correlação linear) foi igual a 0,94 para o milho e 0,81 para a cana-de-açúcar (Figura 4.30).

Os municípios que mais se dispersaram em relação à curva 1:1 do gráfico foram Piraporozinho, Tarumã e Iepê, para a classe cana-de-açúcar; e Cândido Mota, Maracá e Cruzália, para a classe milho (Figura 4.30).

Howard *et al.* (2011), Jonhan *et al.* (2012) e Gusso *et al.* (2012) obtiveram R^2 superior à 84 % ao mapearem soja, milho e trigo na região de Iowa-EUA, soja e milho no Paraná e soja no Rio Grande do Sul, respectivamente. Os resultados destes autores foram semelhantes aos encontrados neste trabalho (cana-de-açúcar = 0,81 e milho = 0,94). No entanto, os autores referidos dispuseram de dados MODIS em seus métodos de mapeamento.

Figura 4.30 – Erros de classificação gerados por OBIA na classificação dos 200 pixels de referência classificados visualmente.

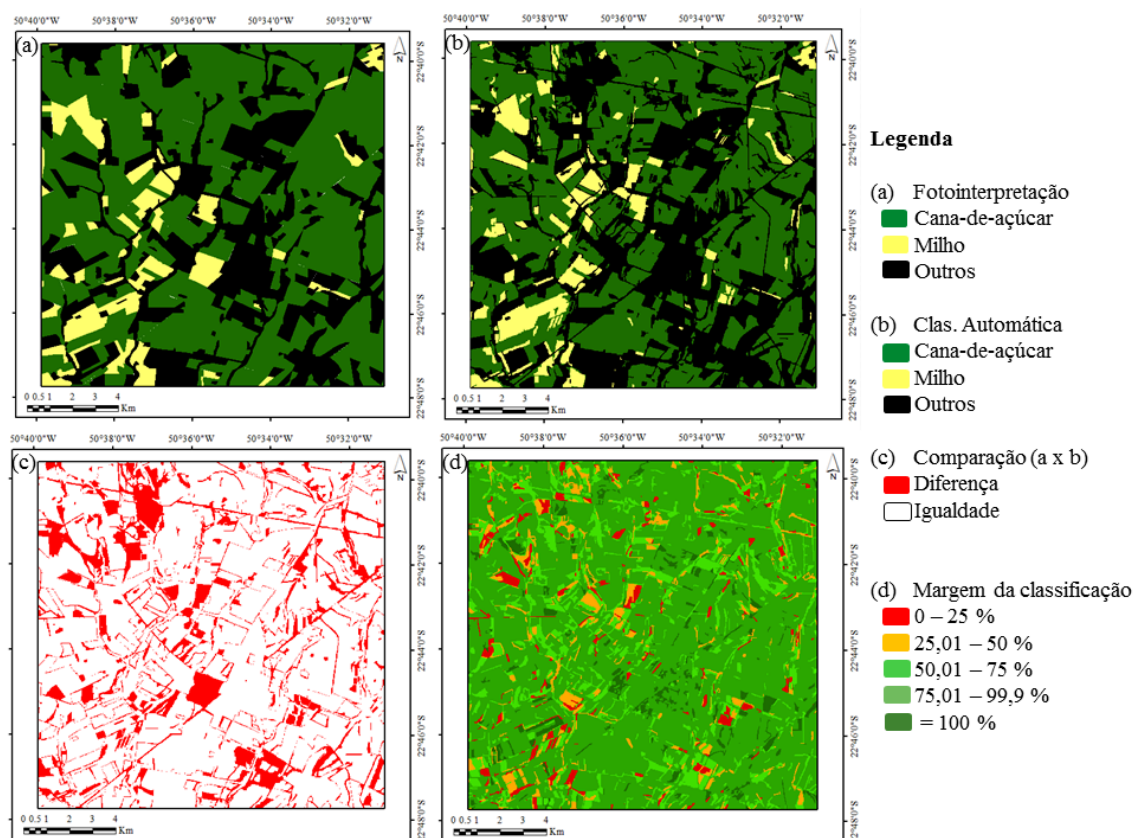


A coluna da esquerda (a) ilustra três erros cometidos pelo classificador automático ao classificar a classe outros como cana, e a da direita (b) segue o caminho inverso, classificou outros como cana.

Como área com cana-de-açúcar apresentou menor R^2 dado pela reta de regressão (Figura 4.30) e o município de Tarumã, vizinho do município de Candido Mota (local onde ocorreu coleta dos dados de treinamento) teve discrepância de – 24,6 % entre os dados do mapa temático e dados oficiais do IBGE, uma área teste de 15 x 15 km foi foto-interpretada sobre este município (Figura 4.31.a).

O mapa temático foi comparado pixel a pixel com a referência (Figura 4.31.b) e foram ilustradas as diferenças da classificação automática, para a referência (Figura 4.31.c). Além disso, foi calculada a sensibilidade do modelo pelo cálculo do acerto das 1.000 árvores de decisão utilizadas para classificar cada um dos segmentos da área teste (Figura 4.31.d).

Figura 4.31 – Área teste de 22.500 ha classificada sobre o município de Tarumã-SP. (a) Dados de referência foto-interpretados, (b) mapa temático, (c) divergências entre o mapa de referência e temático e (d) sensibilidade do modelo de classificação.



A área de discordância informada pela Figura 4.31.c foi de 21,38 %. Na Figura 4.31.c foram poucos os talhões que tiveram erros de discordância entre a referência foto-interpretada e o mapa classificado. Na Figura 4.31.c, as linhas em cor vermelha foram subintendidas com carregadores de cana-de-açúcar e estradas. Na Figura 4.31.a é perceptível que o foto-intérprete não delineou manualmente estas linhas.

Desta forma percebe-se que o modelo gerado teve habilidade de segmentador os talhões e classifica-los corretamente (Figura 4.31.d). Este fato possibilita argumentar e reafirmar que o uso de apenas uma imagem (Maio) para mapeamento de área ocupada com cana-de-açúcar não é suficiente para informar a área total da cultura plantada, isto porque a cultura da cana-de-açúcar é semi-perene e pode ser colhida em vários meses (abril/novembro) do ano safra.

5. CONCLUSÕES

Quanto ao o painel amostral definido por Luiz *et al.* (2015) para monitoramento de área ocupada com cana-de-açúcar, soja e milho no estado de São Paulo é possível obter dados de área para estas três culturas, a cada dois meses e com baixo valor de coeficiente de variação para a estimativa (~5,56 %). Além disso, o método permite o levantamento do valor de área para mais de uma cultura ao mesmo tempo, diferentemente do método Geosafras (apenas uma cultura). Por fim, os dados não podem ser comparados aos dados anuais do IBGE, visto que as variáveis avaliadas pelo IBGE e pelo método MoBARS são diferentes.

Quanto à montagem do banco de imagens para interpretação visual de pixels em laboratório, o uso da ApEq é um método de pré-processamento de uso obrigatório em imagens na composição colorida falsa cor R-b5 G-b6 B-b4. Nas imagens OLI em ApEq, as culturas agrícolas da cana-de-açúcar, milho e soja apresentaram-se nas gradações das cores magenta, alaranjada e amarela, respectivamente. Este método pode ser aplicado em qualquer trabalho que vise interpretação visual de imagens *Landsat-like*.

A partir de dados de campo coletados no mês de abril, foi possível a montagem de dois protocolos de classificação visual. Os protocolos de classificação foram definidos por meio de um roteiro de classificação e duas interfaces de classificação: SIG e *Web*. O protocolo de classificação em interface *Web* mostrou-se mais interativa ao usuário e com possibilidades de operacionalizar em nível estadual a classificação visual de pixels em imagens OLI, para estimativa agrícola de três culturas e redução dos custos das visitas de campo.

O protocolo de classificação em interface *Web* possibilitou o acerto, em média, de 94,5 % dos pixels validados em campo, por intérpretes com menos conhecimento sobre classificação visual de imagens *Landsat-like*. O resultado variou de acordo com a formação dos intérpretes que executaram o procedimento. Intérpretes formados em Engenharia Agrônômica obtiveram melhores resultados que Geógrafo e Engenheiro Florestal. Com isso, percebe-

se o roteiro de classificação deve ser apresentado e discutido com os intérpretes anteriormente à classificação visual.

Quanto ao uso de OBIA/RF para classificação de pixels agrícolas, as 1150 segmentações geradas a partir da variação dos parâmetros fator de escala (fe), forma (fm) e compacidade (cp) do segmentador S-MR classificadas por 1150 modelos de classificação RF possibilitou o acerto de 99 % dos pixels de validação (técnica *out-of-bag*). Este resultado foi obtido pelo uso dos parâmetros fe (25), fm (30) e cp (10) no algoritmo SM-R.

O acerto do mapa temático produzido foi de 86,5 % ($\kappa = 0,77$) (200 pixels independentes dos pixels usados no treinamento). A classificação visual dos pixels em laboratório mostrou-se superior (94,5 % de acerto) ao uso da técnica automatizada. Assim, com os resultados apurados pela classificação automática utilizada na presente pesquisa é possível constatar que este tipo de método automático, ainda não é viável para gerar estimativas agrícolas.

O mapa temático modelou resultado de área ocupada com milho, por município semelhante aos dados de área plantada com milho pela PAM, também por município. O coeficiente de correlação linear (R^2) entre estes dados foi superior a 0,9, enquanto que para a cana-de-açúcar o R^2 foi de apenas 0,81. Imagens provenientes de maio podem ser utilizadas para classificação automática de áreas com milho safrinha, por outro lado a cana-de-açúcar é uma cultura que deve ser mapeada em mais de uma imagem, no ano safra.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABICAB. **Cultivo do amendoim**. Mensagem recebida por: <<http://www.abicab.org.br/cultivo-de-amendoim/>> Acessado em 11 de mar. 2016.
- ACHARD, F. HANSEN, M.C. **Global forest monitoring from earth observation**. 1. ed. Boca Raton FL: CRC Press. 2013. 357 p.
- AGEITEC. **Plantio da cana-de-açúcar**. mensagem recebida por: <https://www.agencia.cnptia.embrapa.br/gestor/cana-de-acucar/arvore/CONTAG01_33_711200516717.html> Acesso em 11 de mar. 2015.
- ADAMI, M. **Estimativa de áreas agrícolas por meio de técnicas de sensoriamento remoto, geoprocessamento e amostragem**. 2004. p. 183. Dissertação (Mestrado na Área de Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2004. Disponível em: <<http://mtc-m12.sid.inpe.br/col/sid.inpe.br/jeferson/2003/08.05.10.38/doc/publicacao.pdf>>. Acesso em 11 de mai. 2015.
- _____. **Estimativa da data de plantio da soja por meio de séries temporais de imagens MODIS**. 2010. p. 163. Tese (Mestrado na Área de Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2004. Disponível em: <<http://mtc-m16d.sid.inpe.br/col/sid.inpe.br/mtc-m19/2010/09.15.21.47/doc/publicacao.pdf>>. Acesso em 11 de mai. 2015.
- ADAMI, M.; MELLO, M. P.; AGUIAR, D. A.; RUDORFF, B. F. T.; SOUZA, A. F. DE. A Web platform development to perform thematic accuracy assessment of sugarcane mapping in South-Central Brazil. **Remote Sensing**, v. 4, n. 12, p. 3201–3214, 19 out. 2012. Disponível em: <<http://www.mdpi.com/2072-4292/4/10/3201/>>. Acesso em: 13 mar. 2013.
- ADAMI, M.; MOREIRA, M. A.; FRIEDRICH, B.; RUDORFF, T.; FREITAS, C. Análise da eficiência dos estimadores de expansão direta e de regressão para áreas cultivadas do café, milho e soja no município de Cornélio Procopio, Estado do Paraná. **Agricultura de São Paulo**, v. 51, n. 2, p. 5–13, 2004.
- ADAMI, M.; MOREIRA, M. A.; RUDORFF, B. F. T.; FREITAS, CORINA DA C.; FARIA, R. T. DE; DEPPE, F. Painel amostral para estimativa de áreas agrícolas Marcos. **Pesquisa Agropecuária Brasileira**, v. 42, n. 1, p. 81–88, 2007.
- ADAMI, M.; RIZZI, R.; MOREIRA, M. A.; RUDORFF, B. F. T.; FERREIRA, C. C. Amostragem probabilística estratificada por pontos para estimar a área cultivada com soja. **Pesquisa Agropecuária Brasileira**, v. 45, n. 6, p. 585–592, 2010.

- ADAMI, M.; RUDORFF, B. F. T.; FREITAS, R. M.; AGUIAR, D. A.; SUGAWARA, L. M.; MELLO, M. P. Remote Sensing Time Series to Evaluate Direct Land Use Change of Recent Expanded Sugarcane Crop in Brazil. **Sustainability**, v. 4, n. 12, p. 574–585, 2 abr. 2012. Disponível em: <<http://www.mdpi.com/2071-1050/4/4/574/>>. Acesso em: 13 mar. 2013.
- AGUIAR, D. A.; RUDORFF, B. F. T.; SILVA, W. F.; ADAMI, M.; MELLO, M. P. Remote Sensing Images in Support of Environmental Protocol: Monitoring the Sugarcane Harvest in São Paulo State, Brazil. **Remote Sensing**, v. 3, n. 12, p. 2682–2703, 13 dez. 2011. Disponível em: <<http://www.mdpi.com/2072-4292/3/12/2682/>>. Acesso em: 20 mar. 2013.
- AGUIAR, D. A. DE; RUDORFF, B. F. T.; RIZZI, R.; SHIMABUKURO, Y. E. Monitoring of sugarcane harvested area through MODIS images. **Revista Brasileira de Cartografia**, v. 60, n. 04, p. 375–383, 2008.
- AGUIAR, E. B.; BICUDO, S. J.; CURCELLI, F.; FIGUEIREDO, P. G. Épocas de poda e produtividade da mandioca. **Pesq. agropec. Bras.** v. 46 n.11 Brasília Nov. , p. 1463–1470, 2011.
- ALVAREZ, C. A.; STEPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. DE M.; SPAROVEK, G. Koppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, p. 1–18, 2013.
- APARICIO , N.; VILLEGAS, D.; ROYO, C.; CASADEUS, J.; ARAUS, J. L. Effect of sensor view angle on the assessment of agronomic traits by ground level hyper-spectral reflectance measurements in durum wheat under contrasting Mediterranean conditions. **International Journal of Remote Sensing**, v. 25, n. 6, p. 1131-1152, mar. 2004.
- ARVOR, D.; JONATHAN, M.; MEIRELLES, M. S. P.; DUBREUIL, V.; DURIEUX, L. Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. **International Journal of Remote Sensing**, v. 32, n. 22, p. 7847–7871, nov. 2011.
- ATZBERGER, C. Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. **Remote Sensing**, v. 5, n. 2, p. 949–981, 22 fev. 2013. Disponível em: <<http://www.mdpi.com/2072-4292/5/2/949/>>. Acesso em: 13 mar. 2013.
- ATZBERGER, C.; FORMAGGIO, A. R.; SHIMABUKURO, Y. E.; UDELHOVEN, T.; MATTIUZZI, M.; SANCHEZ, G. A.; ARAI, E.; SENSING, R.; INFORMATION, L. Obtaining crop-specific time profiles of NDVI : the use of unmixing approaches for serving the continuity between SPOT-VGT and PROBA-V time series. **International Journal of Remote Sensing**, n. July, p. 37–41, 2014.
- BAO, N.; LECHNER, A. M.; FLETCHER, A.; MELLOR, A.; MULLIGAN, D.; BAI, Z. Comparison of relative radiometric normalization methods using pseudo-

invariant features for change detection studies in rural and urban landscapes. **Journal of Applied Remote Sensing**, v. 6, n. 1, p. 063578–1, 24 set. 2012. Disponível em: <<http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.JRS.6.0.63578>>. Acesso em: 23 dez. 2015.

BECKER-RESHEF, I.; JUSTICE, C.; SULLIVAN, M.; VERMOTE, E.; TUCKER, C.; ANYAMBA, A.; SMALL, J.; PAK, E.; MASUOKA, E.; SCHMALTZ, J.; HANSEN, M.; PITTMAN, K.; BIRKETT, C.; WILLIAMS, D.; REYNOLDS, C.; DOORN, B. Monitoring global croplands with coarse resolution earth observations: the Global Agriculture Monitoring (GLAM) project. **Remote Sensing**, v. 2, n. 6, p. 1589–1609, 18 jun. 2010. Disponível em: <<http://www.mdpi.com/2072-4292/2/6/1589/>>. Acesso em: 5 mar. 2013.

BENZ, U. C.; HOFMANN, P.; WILLHAUCK, G.; LINGENFELDER, I.; HEYENEN, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 58, n. 3-4, p. 239–258, jan. 2004. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0924271603000601>>. Acesso em: 10 jul. 2014.

BERNARDES, T.; ADAMI, M.; FORMAGGIO, A. R.; MOREIRA, M. A. Imagens mono e multitemporais Modis para estimativa da área com soja no Estado de Mato Grosso. **Pesquisa Agropecuária Brasileira**, n. 1, p. 1530–1537, 2011.

BINS, L.; FONSECA, L.; ERTHAL, G. Satellite imagery segmentation: a region growing approach. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 8., 1996, Salvador. **Anais...** São José dos Campos: INPE, 1996. p. 677-680. Disponível em: <p. 677-680. CD-ROM. ISBN 85-17-00014-5. (INPE-6231-PRE/2321)>. Disponível em: <<http://urlib.net/sid.inpe.br/deise/1999/02.05.09.30>>. Acesso em: 21 mar. 2016.

BLASCHKE, T. Object based image analysis for remote sensing. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 65, n. 1, p. 2–16, jan. 2010. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0924271609000884>>. Acesso em: 21 jan. 2014.

BREIMAN, L. Bagging predictors. **Machine Learning**, v. 24, n. 12, p. 123-140, 1996.

BREIMAN, L. Random forests. **Machine Learning**, v. 45, p. p. 5-32, 2001.

BREUNIG, F.M. **Influência da geometria de aquisição sobre índices de vegetação e estimativas de IAF com dados MODIS, HYPERION, e simulação PROSAIL para a soja**. 2011. p. 210. Tese (Doutorado na Área de Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2004. Disponível em: < <http://mtc->

m16d.sid.inpe.br/col/sid.inpe.br/mtc-
m19/2011/03.25.19.13/doc/publicacao.pdf>. Acesso em 10 de jan. 2015.

BREUNIG, F. M.; GALVÃO, L. S.; FORMAGGIO, A. R.; EPIPHANIO, J. C. N. Classification of soybean varieties using different techniques: case study with Hyperion and sensor spectral resolution simulations. **Journal of Applied Remote Sensing**, v. 5, p. 053533, 1 jan. 2011. Disponível em: <<http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.3604787>>. Acesso em: 3 abr. 2014.

BREUNIG, F. M.; GALVÃO, L. S.; FORMAGGIO, A. R.; EPIPHANIO, J. C. N. Variation of MODIS reflectance and vegetation indices with viewing geometry and soybean development. **Anais da Academia Brasileira de Ciências**, v. 84, n. 2, p. 263–74, jun. 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22473302>>.

BROWN, J. C.; KASTENS, J. H.; COUTINHO, A. C.; VICTORIA, D. D. C.; BISHOP, C. R. Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data. **Remote Sensing of Environment**, v. 130, p. 39–50, mar. 2013. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0034425712004336>>. Acesso em: 7 mar. 2014.

CAFESAT. **Mapa do café para o ano safra 2007**. Mensagem enviada por: <<http://www.dsr.inpe.br/laf/cafesat/mapa.html>> Acessado em 11 de abr. 2014.

CANASAT. **Área disponível para colheita de cana-de-açúcar ano safra 2012/2013**. Mensagem enviada por: <<http://www.dsr.inpe.br/laf/canasat/cultivo.html>> Acesso em 11 de jan. 2014.

CEPEA. **Dimensão do PIB do agronegócio do estado de São Paulo. Relatório de 2014**. Mensagem recebida por <http://www.cepea.esalq.usp.br/comunicacao/Cepea_RelatorioFinal_PIB%20SP.pdf> em 10 de out. de 2015.

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, mar. 2008. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0034425707003495>>. Acesso em: 17 ago. 2015.

CHAN, J. C.-W.; PAELINCKX, D. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. **Remote Sensing of Environment**, v. 112, n. 6, p. 2999–3011, jun. 2008. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0034425708000679>>. Acesso em: 18 jul. 2014.

CHEN, J.; CHEN, J.; LIAO, A.; CAO, X.; CHEN, L.; CHEN, X.; HE, C.; HAN, G.; PENG, S.; LU, M.; ZHANG, W.; TONG, X.; MILLS, J. Global land cover mapping at 30m resolution: A POK-based operational approach. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 103, p. 7–27, out. 2015. Disponível em:

<<http://linkinghub.elsevier.com/retrieve/pii/S0924271614002275>>. Acesso em: 27 jan. 2015.

CIHLAR, J. Land cover mapping of large areas from satellites: Status and research priorities. **International Journal of Remote Sensing**, v. 21, n. 6-7, p. 1093–1114, jan. 2000. Disponível em:

<<http://www.tandfonline.com/doi/abs/10.1080/014311600210092>>.

CLARK, R.N.; SWAYZE, G.A.; WISE, R.; LIVO, E.; HOEFEN, T.; KOKALY, R.; SUTLEY, S.J. **USGS digital spectral library splib06a**: U.S. Geological Survey, Digital Data Series 231. 2007. Disponível em:

<<http://speclab.cr.usgs.gov/spectral.lib06/ds231/index>>. Acesso em: 21 jun. 2014.

COLOMBO, R. Retrieval of leaf area index in different vegetation types using high resolution satellite data. **Remote Sensing of Environment**, v. 86, n. 1, p. 120–131, 30 jun. 2003. Disponível em:

<<http://linkinghub.elsevier.com/retrieve/pii/S0034425703000944>>. Acesso em: 11 set. 2015.

CONAB. **Acompanhamento da safra brasileira de grãos**: sexto levantamento - março/2014. Mensagem recebida por: <

http://www.conab.gov.br/OlalaCMS/uploads/arquivos/14_03_12_08_41_24_bol_etim_graos_marco_2014.pdf> Acesso em 11 de jan. 2015.

CONGALTON, R.; GU, J.; YADAV, K.; THENKABAIL, P.; OZDOGAN, M. Global land cover mapping: a review and uncertainty analysis. **Remote Sensing**, v. 6, n. 12, p. 12070–12093, 3 dez. 2014. Disponível em:

<<http://www.mdpi.com/2072-4292/6/12/12070/>>. Acesso em: 4 nov. 2015.

COCHRAN, W. G. **Sampling techniques**. 3. ed. New York: Wiley. 1977. 448 p.

CONGALTON RG, GREEN K. 2009. **Assessing the accuracy of remotely sensed data**: principles and practices. 2. ed. Boca Raton, FL: CRC Press. 2009. 210 p.

CRUSCO, D. **Sensoriamento remoto para análise multitemporal da dinâmica de áreas agrícolas**. 2006. p. 183. Dissertação (Mestrado na Área de Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2004. Disponível em: <http://mtc-m16b.sid.inpe.br/col/sid.inpe.br/MTC-m13@80/2006/08.25.14.03/doc/publicacao.pdf>>. Acesso em 11 de mai. 2007.

DURO, D. C.; FRANKLIN, S. E.; DUBÉ, M. G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. **Remote Sensing of Environment**, v. 118, p. 259–272, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.rse.2011.11.020>>.

EBERHARDT, I. D. R.; LUIZ, A. J. B.; FORMAGGIO, A. R.; IEDA DEL'ARCOR SANCHES. Detecção de áreas agrícolas em tempo quase real com imagens Modis no Rio Grande do Sul. **Pesquisa Agropecuária Brasileira**, v. 50, n. 7, p. 1–10, 2015.

EBERHARDT, I.D.R. **Estimativa em tempo quase real de área de milho e soja no Rio Grande do Sul, por sensoriamento remoto e amostragem**. 2015. p. 134. Dissertação (Mestrado na Área de Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2015. Disponível em: <<http://urlib.net/8JMKD3MGP3W34P/3HNHPQL>>. Acesso em 12 de mai. 2015.

EBERHARDT, I.D.R.; SCHULTZ, B.; IMMITZER, M.; RIZZI, R.; SANCHES, I.D.; FORMAGGIO, A.R.; ATZBERGER, C.; MELLO, M.P.; TRABAQUINI, K.; FOSCHIERA, W.; LUIZ, A.J.B. Cloud cover assessment for operational crop monitoring system in tropical areas. **Em análise pela revista Remote Sensing**. Acesso em 15 de fev. 2016.

EIVAZI, A.; KOLESNIKOV, A.; JUNTILA, V.; KAURANNE, T. Variance-preserving mosaicing of multiple satellite images for forest parameter estimation: Radiometric normalization. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 105, p. 120–127, jul. 2015. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0924271615000878>>. Acesso em: 31 ago. 2015.

EL-MAGD, I. A.; TANTON, T. W. Improvements in land use mapping for irrigated agriculture from satellite sensor data using a multi-stage maximum likelihood classification. **International Journal of Remote Sensing**, v. 24, n. 21, p. 4197–4206, 2003.

EMBRAPA. **Cultivo do milho**: plantio, espaçamento, densidade de sementes. Mensagem recebida por: <http://www.cnpms.embrapa.br/publicacoes/milho_1_ed/plantespaca.htm> Acessado em 11 de mar. 2015.

_____. **Cultivo da soja**: instalação da lavoura. Mensagem recebida por: <<http://www.cnpso.embrapa.br/producaosojapr/instalacao.htm>> Acessado em 11 de mar. 2015.

_____. **Cultivo da mandioca para a região Centro Sul do Brasil**. Mensagem recebida por

<https://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Mandioca/mandioca_centrosul/plantio.htm> Acessado em 11 de mar. 2015.

EPIPHANIO, J. C. N.; LUIZ, A. J. BARRETO; FORMAGGIO, A. R. Estimativa de áreas agrícolas municipais, utilizando sistema de amostragem simples sobre imagens de satélite. **Bragantia**, v. 3, n. 1, p. 187–197, 2000.

ESPÍNDOLA, G.M. **Ajuste dos parâmetros de algoritmos de segmentação de imagens por crescimento de regiões**. 2006. p. 65 Dissertação (Mestrado na Área de Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2004. Disponível em: < <http://mtc-m16b.sid.inpe.br/col/sid.inpe.br/MTC-m13@80/2006/12.21.11.14/doc/publicacao.pdf>>. Acesso em 11 de jan. 2015.

ESRI CORPORATION. **ArcGIS for desktop**. Mensagem enviada por: < <http://www.esri.com/software/arcgis/arcgis-for-desktop>> Acesso em 11 de jan. 2015.

FAGUNDES, L. K.; STRECK, N. A.; ROSA, H. T.; WALTER, L. C.; ZANON, A. J. Desenvolvimento , crescimento e produtividade de mandioca em diferentes datas de plantio em região subtropical. **Ciência Rural**, v. 40, n. 12, p. 2460–2466, 2010.

FAOSTAT. **Food balance**. Mensagem recebida por < <http://faostat.fao.org/>> em 12 de ago. de 2015.

FAVARO, F.N. **Dinâmica espacial da citricultura de São Paulo no período de 2007 a 2012**. 2013. p. 113 Dissertação (Mestrado na Área de Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2013. Disponível em: <http://mtc-m16d.sid.inpe.br/col/sid.inpe.br/mtc-m19/2013/09.03.18.29/doc/publicacao.pdf>>. Acesso em 07 de jan. 2014.

FERNANDES, R. R.; MANZON, G.; SANNA, T.; SILVA, F. Classificação orientada a objetos aplicada na caracterização da cobertura da terra no Araguaia. **Pesquisa Agropecuária Brasileira**, v. 47, n. 9, p. 1251–1260, 2012.

FIGUEIREDO, D.C. Projeto GeoSafras: aperfeiçoamento do sistema de previsão de safras da Conab. **Revista de Política Agrícola**, v.14, p.110-120, 2005.

FLORENZANO, T.G. **Iniciação em sensoriamento remoto**. 3. ed. São Paulo: Oficina do texto. 2011. 128 p.

FOOD BALANCE SHEET. **A handbook**. Mensagem recebida por. <http://www.fao.org/docrep/003/x9892e/x9892e00.htm>>. Acesso em 15 de dez. 2015.

FOODY, G. M. Thematic Map Comparison : Evaluating the Statistical Significance of Differences in Classification Accuracy. **Photogrammetric Engineering & Remote Sensing**, v. 70, n. 5, p. 627–633, 2004.

FORSTER, D.; KELLENBERGER, T.; BUEHLER, Y.; LENNARTZ, B. Mapping diversified peri-urban agriculture - potencial of object-based versus land cover/land use classification. **Geocarto International**, v. 25, n. 3, p. 171–186, 2010.

FROLKING, S.; XIAO, X.; ZHUANG, Y.; SALAS, W.; LI, C. Agriculturas land-use in China: a comparison of area estimates from ground-based census and satellite-borne remote sensing. **Global Ecology and Biogeography**, v. 8, p. 407–416, 1999.

GALFORD, G. L.; MUSTARD, J. F.; MELILLO, J.; GENDRIN, A.; CERRI, C. C.; CERRI, C. E. P. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. **Remote Sensing of Environment**, v. 112, n. 2, p. 576–587, fev. 2008. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0034425707002258>>. Acesso em: 2 mar. 2013.

GALLETTI, C.; MYINT, S. Land-use mapping in a mixed urban-agricultural arid landscape using object-based image analysis: a case study from Maricopa, Arizona. **Remote Sensing**, v. 6, n. 7, p. 6089–6110, 30 jun. 2014. Disponível em: <<http://www.mdpi.com/2072-4292/6/7/6089/>>. Acesso em: 24 abr. 2015.

GALVÃO, L. S.; DOS SANTOS, J. R.; ROBERTS, D. A.; BREUNIG, F. M.; TOOMEY, M.; DE MOURA, Y. M. On intra-annual EVI variability in the dry season of tropical forest: a case study with MODIS and hyperspectral data. **Remote Sensing of Environment**, v. 115, n. 9, p. 2350–2359, set. 2011. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0034425711001635>>. Acesso em: 23 dez. 2015.

GARAGORRY, F. L.; FILHO, H. C. Elementos de agrodinâmica. **SGE/EMBRAPA Brasília**, p. 1–72, 2008. Disponível em: <<http://www.embrapa.br/publicacoes/institucionais/titulos-avulsos/RELFIN1.pdf>>.

GAROFALO, D. F. T.; MESSIAS, C. A.; LIESENBERG, V.; BOLFE, É. L.; FERREIRA, M. C. Análise comparativa de classificadores digitais em imagens do Landsat - 8 aplicados ao mapeamento temático. **Pesquisa Agropecuária Brasileira**, v. 50, n. 7, p. 593–604, 2015.

GEBHARDT, S.; WEHRMANN, T.; RUIZ, M.; MAEDA, P.; BISHOP, J.; SCHRAMM, M.; KOPEINIG, R.; CARTUS, O.; KELLNDORFER, J.; RESSL, R.; SANTOS, L.; SCHMIDT, M. MAD-MEX: Automatic wall-to-wall land cover monitoring for the Mexican REDD-MRV program using all Landsat data.

Remote Sensing, v. 6, n. 5, p. 3923–3943, 30 abr. 2014. Disponível em: <<http://www.mdpi.com/2072-4292/6/5/3923/>>. Acesso em: 10 ago. 2015.

GITELSON, A. A.; STARK, R.; RUNDQUIST, D. Novel algorithms for remote estimation of vegetation fraction. **Remote Sensing of Environment**, n. 80, p. 76–87, 2002.

GONZALEZ, R.C.; WOODS, R.E. **Digital image processing**. 3. ed. Prentice Hall. 2007. 976 p.

GUIMARÃES, D. P.; LANDAU, E. C. Mapeamento de áreas irrigadas por pivôs centrais no Estado de Minas Gerais. **Boletim de Pesquisa**, v. 40, 2011.

GUSSO, A.; FORMAGGIO, A. R.; RIZZI, R.; ADAMI, M. Soybean crop area estimation by Modis/Evi data. **Pesquisa Agropecuária Brasileira**, v. 47, n. 3, p. 425–435, 2012.

GUYON, I.; ELISSEEFF, A. An introduction to variable and feature selection 1 introduction. **Journal of Machine Learning Research** 3, v. 3, p. 1157–1182, 2003.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I.H. The WEKA data mining software: an update. **SIGKDD Explorations**, v. 11, n. 1, p. 10–18, 2009.

HANSEN, M. C.; POTAPOV, P. V.; MOORE, R.; HANCHER, M.; TURUBANOVA, S. A.; TYUKAVINA, A. High-Resolution Global Maps of 21st-Century Forest Cover Change. **Science**, v. 134, n. November, p. 850–854, 2013.

HARALICK, R. M., SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 34, n. 3, p. 610-621, 1973.

HIJMANS, R.J. **Raster**: geographic data analysis and modeling; R package, 2014. Disponível em: < <https://cran.r-project.org/web/packages/raster/index.html>>. Acesso em: 10 ago. 2014.

HOWARD, D. M.; WYLIE, B. K.; TIESZEN, L. L. ICrop classification modelling using remote sensing and environmental data in the Greater Platte River Basin , USA. **International Journal of Remote Sensing**, v. 33, n. 19, p. 6094–6108, 2012.

HSIEH, P.; LEE, L. C.; CHEN, N. Effect of spatial resolution on classification errors of pure and mixed pixels in remote sensing. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 39, n. 12, p. 2657–2663, 2001.

INSTITUTO DE ECONOMIA AGRÍCOLA (IEA). **Estatística da produção paulista**. Mensagem recebida por: <
http://ciagri.iea.sp.gov.br/nia1/subjetiva.aspx?cod_sis=1&idioma=1> Acesso em 11 de jan. 2015.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Sistema IBGE de recuperação automática**. Mensagem recebida por: <
<http://www.sidra.ibge.gov.br/>> Acesso em 11 de jan. 2015.

IMMITZER, M.; ATZBERGER, C.; KOUKAL, T. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. **Remote Sensing**, v. 4, p. 2661–2693, 2012.

IPPOLITI-RAMILO, G. A.; EPIPHANIO, J. C. N.; SHIMABUKURO, Y. E. Landsat 5 Thematic Mapper data for pre planting crop area evaluation in tropical countries. **International Journal of Remote Sensing**, v. 24, n. 7, p. 1521–1534, jan. 2003. Disponível em:
<<http://www.tandfonline.com/doi/abs/10.1080/01431160010007105>>. Acesso em: 7 mar. 2014.

ITT VISUAL INFORMATION SOLUTION. **Increase agricultural efficiencies using satellite imagery (recorded)**. Mensagem enviada por: <
<http://www.exelisvis.com/Learn/EventsTraining/Webinars/TabId/397/ArtMID/1695/ArticleID/12215/ITT-Visual-Information-Solutions-signs-global-distributors-for-ENVI.aspx>> Acesso em 13 de abr. 2015.

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

JENSEN, J.R. **Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres**. Tradução, José Carlos Neves Epiphânio. 2. ed. São José dos Campos: Parênteses Editora. 2011. 598 p.

JOHANN, J. A.; ROCHA, J. V.; GARBELLINI, D.; AUGUSTO, R.; LAMPARELLI, C. Estimation of summer crop areas in the state of Paraná, Brazil, using multitemporal EVI/Modis images. **Pesquisa Agrícola Brasileira**, v. 47, n. 9, p. 1295–1306, 2012.

JUSTICE, C.; TOWNSHEND, J. R.; VERMOTE, E.; MASUOKA, E.; WOLFE, R.; SALEOUS, N.; ROY, D.; MORISETTE, J. An overview of MODIS Land data processing and product status. **Remote Sensing of Environment**, v. 83, n. 1-2, p. 3–15, nov. 2002. Disponível em:
<<http://linkinghub.elsevier.com/retrieve/pii/S0034425702000846>>.

KNORN, J.; RABE, A.; RADELOFF, V. C.; KUEMMERLE, T.; KOZAK, J.; HOSTERT, P. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. **Remote Sensing of Environment**, v.

113, n. 5, p. 957–964, 2009. Disponível em:
<<http://dx.doi.org/10.1016/j.rse.2009.01.010>>.

KORTING, T. S.; FONSECA, L. M.; ESCADA, M. I. S.; SILVA, F. C.; SILVA, M. P. S. GeoDMA: a novel system for spatial data mining. IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, Pisa, **Anais ...** 2008. p. 975 - 978. Disponível em: <
http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=4734031&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D4734031>
Acesso em 09 abr. 2013.

KUHN, M.; WING, J.; WESTON, S.; WILLIAMS, A.; KEEFER, C.; ENGELHARDT, A.; COOPER, T.; MAYER, Z. **Caret: Classification and Regression Training; R package**, 2014. Disponível em: < <https://cran.r-project.org/web/packages/caret/caret.pdf> >. Acesso em: 10 ago. 2014.

LAMPARELLI, R. A. C.; CARVALHO, W. M. O. DE; MERCANTE, E. Mapeamento de semeaduras de soja (*Glycinea max* (L.) Merr.) mediante dados MODIS/TERRA e TM/LANDSAT 5: Um comparativo. **Engenharia Agrícola**, v. 28, n. 2, p. 334–344, 2008.

LANDAU, E. C.; MOURA, L.; GUIMARÃES, D. P.; HIRSCH, A.; PIMENTA, F. M. Concentração geográfica de pivôs centrais no Brasil. **Boletim de Pesquisa**, No 69/CNPMS, 37p., 2013. Mensagem enviada por <
<http://www.infoteca.cnptia.embrapa.br/bitstream/doc/977535/1/bol69.pdf>>
Acesso em: 25/mai/2014.

LAPOLA, D. M.; MARTINELLI, L. A.; PERES, C. A.; OMETTO, J. P. H. B.; FERREIRA, M. E.; NOBRE, C. A.; AGUIAR, A. P. D.; BUSTAMANTE, M. M. C.; CARDOSO, M. F.; COSTA, M. H.; JOLY, C. A.; LEITE, C. C.; MOUTINHO, P.; SAMPAIO, G.; STRASSBURG, B. B. N.; VIEIRA, I. C. G. Pervasive transition of the Brazilian land-use system. **Nature Climate Change**, v. 4, n. 1, p. 27–35, 20 dez. 2013. Disponível em:
<<http://www.nature.com/doifinder/10.1038/nclimate2056>>. Acesso em: 22 jan. 2014.

LIAW, A.; WIENER, M. Classification and regression by randomForest. **R News**, v. 2, p. 18–22, 2002.

LILLESAND, T. M., KIEFER, R.W.; CHIPMAN, J. **Remote sensing and image interpretation**. 6. ed. New York: John Wiliv & Sons. 2007. 750 p.

LONG, J. A.; LAWRENCE, R. L.; GREENWOOD, M. C.; MILLER, P. R. Object-oriented crop classification using multitemporal ETM + SLC-off imagery and random forest. **GIScience and Remote Sensing**, n. August, p. 37–41, 2013.

LÖW, F.; DUVEILLER, G. Defining the spatial resolution requirements for crop identification using optical remote sensing. **Remote Sensing**, v. 6, n. 9, p.

9034–9063, 23 set. 2014. Disponível em: <<http://www.mdpi.com/2072-4292/6/9/9034/>>. Acesso em: 15 dez. 2015.

LUCON, O.; GOLDEMBERG, J. Sao Paulo--the “Other” Brazil: different pathways on climate change for state and federal governments. **The Journal of Environment & Development**, v. 19, n. 3, p. 335–357, 15 ago. 2010. Disponível em: <<http://jed.sagepub.com/cgi/doi/10.1177/1070496510378092>>. Acesso em: 22 dez. 2015.

LUIZ, A.J.B. **Estatísticas agrícolas por amostragem auxiliadas pelo sensoriamento remoto**. 2009. p. 112. Tese (Doutorado na Área de Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2003. Disponível em: <<http://mtc-m16c.sid.inpe.br/col/sid.inpe.br/jeferson/2004/02.04.16.08/doc/publicacao.pdf>>. Acesso em 10 de mai. 2014.

LUIZ, A.J.B.; EPIPHANIO, J.C.N. Amostragem por pontos em imagens de sensoriamento remoto para estimativa de área plantada por município. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 10., 2001, Foz do Iguaçu. **Anais ...** São José dos Campos: INPE, 2001, p. 111–118. Disponível em: <<http://marie.sid.inpe.br/col/dpi.inpe.br/lise/2001/09.13.10.55/doc/0111.118.116.pdf>> Acesso em 07 de jun. 2014.

LUIZ, A.J.B.; FORMAGGIO, A.R.; EPIPHANIO, J.C.N. Objective sampling estimation of crop area based on remote sensing: Advances and applications. In: DO PRADO, H.A.; LUIZ, A.J.B.; CHAIB FILHO, H.(eds.). **Computational methods for agricultural research: advances and applications**. Hershey: IGI Global, 2011. v. cap. 5, p. 73-95. ISBN 9781616928711. doi: <10.4018/978-1-61692-871-1.ch005>.

LUIZ, A. J. B et al. Estimativa amostral objetiva de área plantada regional, apoiada em imagens de sensoriamento remoto. **Pesquisa Agropecuária Agrícola** v. 47, n. 9, p. 1279–1287, 2012. Disponível em: <<http://www.scielo.br/pdf/pab/v47n9/13.pdf>>. Acesso em: 27 maio 2013.

LUIZ, AL. J. B.; SCHULTZ, B.; TRABAQUINI, K.; EBERHARDT, I. D. R.; FORMAGGIO, A. R. Método para estratificação em levantamentos agrícolas com mais de uma variável. **Documentos Embrapa Meio Ambiente**, v. 100, p. 1–140, 2015.

MAGALHÃES, P.S.; DURÃO, F.D.M.; CARNEIRO, N.P.; PAIVA, E. **Fisiologia do milho**. Sete Lagoas: EMBRAPA-CNPMS, 2002. 23 P. Embrapa (Circular Técnica 22).

MAPA. **Zoneamento agrícola**: portaria segmentada por UF. Mensagem recebida por <<http://www.agricultura.gov.br/politica-agricola/zoneamento-agricola/portarias-segmentadas-por-uf>> Acesso em 14 de abr. 2015.

MARQUES, J. Q. de A. **Manual brasileiro para levantamento da capacidade de uso da terra**: 3a aproximação. Rio de Janeiro: Escritório Técnico Brasil-EUA, 1971. 433 p.

MATTON, N.; CANTO, G.; WALDNER, F.; VALERO, S.; MORIN, D.; INGLADA, J.; ARIAS, M.; BONTEMPS, S.; KOETZ, B.; DEFOURNY, P. An Automated Method for Annual Cropland Mapping along the Season for Various Globally-Distributed Agrosystems Using High Spatial and Temporal Resolution Time Series. **Remote Sensing**, v. 7, n. 10, p. 13208–13232, 6 out. 2015. Disponível em: <<http://www.mdpi.com/2072-4292/7/10/13208/>>. Acesso em: 4 nov. 2015.

MCNAIRN, H.; PROTZ, R. Mapping corn residue cover on agricultural fields in Oxford County, Ontario, using thematic mapper. **Canadian Journal Remote Sensing**, v. 19, p. 152–159, 1993.

MCROBERTS, R. E. Satellite image-based maps: scientific inference or pretty pictures? **Remote Sensing of Environment**, v. 115, n. 2, p. 715–724, 15 fev. 2011. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0034425710003184>>. Acesso em: 23 jan. 2014.

MELLO, M. P.; RUDORFF, B. F. T.; VIEIRA, C. A. O.; AGUIAR, D. A. DE. Automatic classification of sugarcane harvest using spectral linear mixing model. **Revista Brasileira de Cartografia**, v. 2, n. 62, p. 181–188, 2010.

MENDONÇA, J. C.; FREITAS, R. M. DE; AGUIAR, D. A. DE; SOUSA, E. F. DE; MUNIZ, R. DE A.; ESTEVES, B. DOS S. Mapeamento das áreas de cana-de-açúcar na região norte fluminense-RJ por uso de técnicas de sensoriamento remoto. **Engenharia Agrícola**, v. 31, n. 3, p. 561–571, 2011.

MOREIRA, M. A.; BARROS, M. A.; RUDORFF, B. F. T. Geotechnologies in Coffee crop mapping at municipality scale. **Sociedade e Natureza**, v. 20, n. 1, p. 101–110, 2008.

MOREIRA, M.A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. 3. ed. Viçosa: UFV. 2007. 301 p.

MOREIRA, M. A.; RUDORFF, B. F. T.; BARROS, M. A.; VARIA, V. G. C. DE; ADAMI, M. Geotecnologias para mapear lavouras de café nos estados de Minas Gerais e São Paulo. **Engenharia Agrícola**, v. 30, n. 6, p. 1123–1135, 2010.

MOSLEH, M. K.; HASSAN, Q. K.; CHOWDHURY, E. H. Application of remote sensors in mapping rice area and forecasting its production: a review. **Sensors (Basel, Switzerland)**, v. 15, n. 1, p. 769–91, jan. 2015. Disponível em: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4327048&tool=pmc_entrez&rendertype=abstract>. Acesso em: 24 set. 2015.

- MOURA, Y. M.; GALVÃO, L. S.; DOS SANTOS, J. R.; ROBERTS, D. A.; BREUNIG, F. M. Use of MISR/Terra data to study intra- and inter-annual EVI variations in the dry season of tropical forest. **Remote Sensing of Environment**, v. 127, p. 260–270, dez. 2012. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0034425712003677>>. Acesso em: 23 dez. 2015.
- MULLER, C. C.; SILVA, G.; VILLALOBOS, A. G. Pesquisa agropecuária do Paraná - safra 1986/87. **Revista Brasileira de Estatística**, v. 49, n. 119, p. 55–84, 1988.
- MUÑOZ, X.; FREIXENET, J.; CUF, X.; MART, J. Strategies for image segmentation combining region and boundary information. **Journal Pattern Recognition Letters**, v. 24, p. 375–392, 2003.
- MYINT, S. W.; GOBER, P.; BRAZEL, A.; GROSSMAN-CLARKE, S.; WENG, Q. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. **Remote Sensing of Environment**, v. 115, n. 5, p. 1145–1161, 15 maio 2011. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0034425711000034>>. Acesso em: 25 jul. 2014.
- NAUGHTON D.; BRUNN, A.; CZAPLA-MYERS, J.; DOUGLASS, S.; THIELE, M.; WEICHELT, H.; OXFORT, M. Absolute radiometric calibration of the RapidEye Multispectral Imager using the reflectance-based vicarious calibration method. **SPIE Journal of Applied Remote Sensing**, v.5, p. 112-134, 2011.
- NAVULUR, K. **Multispectral image analysis using the object-oriented paradigm**. Boca Raton, FL: Taylor & Francis Group, 2007, 206p.
- NOVO, E.M.L. **Sensoriamento remoto: princípios e aplicações**. 4 ed. São Paulo: Blucher, 2009. 387 p.
- OLIVEIRA, J.; FORMAGGIO, A.; EPIPHANIO, J.; LUIZ, A. Index for the Evaluation of Segmentation (IAVAS): an application to agriculture. **Mapping Sciences & Remote Sensing**, v. 40, n. 3, p. 155–169, 1 set. 2003. Disponível em: <<http://bellwether.metapress.com/openurl.asp?genre=article&id=doi:10.2747/0749-3878.40.3.155>>. Acesso em: 29 jan. 2015.
- OLOFSSON, P.; FOODY, G. M.; STEHMAN, S. V; WOODCOCK, C. E. Remote Sensing of Environment Making better use of accuracy data in land change studies : Estimating accuracy and area and quantifying uncertainty using stratified estimation. **Remote Sensing of Environment**, v. 129, p. 122–131, 2013.
- ORTIZ, M.J., FORMAGGIO, A.R., EPIPHANIO, J.C.N. Classification of croplands through integration of remote sensing, GIS, and historical database. **International Journal of Remote Sensing**, v.18, n. 1, p. 95–105, 1997.

OSHIRO, T.M. **Uma abordagem para construção de uma única árvore a partir de uma Random Forest para classificação de bases e expressões gênicas**. 2009. p. 112. Dissertação (Mestrado em Bioinformática) – Universidade de São Paulo, Ribeirão Preto, 2011. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/95/95131/tde-15102013-183234/pt-br.php>>. Acesso em 10 de mai. 2015.

OTB CORE TIME. **Orfeo toolbox**. Mensagem recebida por: <<https://code.launchpad.net/~otb>> Acesso em 11 de mar. 2015.

PEÑA, J.; GUTIÉRREZ, P.; HERVÁS-MARTÍNEZ, C.; SIX, J.; PLANT, R.; LÓPEZ-GRANADOS, F. Object-based image classification of summer crops with machine learning methods. **Remote Sensing**, v. 6, n. 6, p. 5019–5041, 30 maio 2014. Disponível em: <<http://www.mdpi.com/2072-4292/6/6/5019/>>. Acesso em: 13 mar. 2015.

PEÑA-BARRAGÁN, J. M.; NGUGI, M. K.; PLANT, R. E.; SIX, J. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. **Remote Sensing of Environment**, v. 115, n. 6, p. 1301–1316, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.rse.2011.01.009>>.

PINO, F. A. Estatísticas agrícolas para o Século XXI. **Agricultura de São Paulo**, v. 46, n. 2, p. 71–105, 1999.

PONZONI, F. J.; SHIMABUKURO, Y. E. **Sensoriamento remoto aplicado ao estudo da vegetação**. São José dos Campos: Parêntese, 2007. 135 p.

POWELL, R. .; MATZKE, N.; DE SOUZA, C.; CLARK, M.; NUMATA, I.; HESS, L. .; ROBERTS, D. . Sources of error in accuracy assessment of thematic land-cover maps in the Brazilian Amazon. **Remote Sensing of Environment**, v. 90, n. 2, p. 221–234, mar. 2004. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S003442570300381X>>. Acesso em: 16 jul. 2015.

POWELL, S. L.; PFLUGMACHER, D.; KIRSCHBAUM, A. A.; KIM, Y.; COHEN, W. B. Moderate resolution remote sensing alternatives: a review of Landsat-like sensors and their applications. **Journal of Applied Remote Sensing**, v. 1, p. 012506, 1 nov. 2007. Disponível em: <<http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.2819342>>. Acesso em: 4 mar. 2013.

PROJETO PRODES. **Monitoramento da floresta amazônica brasileira por satélite**. Mensagem recebida por: <<http://www.obt.inpe.br/prodes/index.php>> Acesso em 11 de mar. 2015.

QI, J.; MARSELT, R.; HEILMAN, P.; BENDER, S. B.-; MONIN., S.; GOODRICH, D.; MARK WELTZ. RANGES Improves satellite-based information

and land cover assessments in Southwest United States. **EOS, Transactions, American Geophysical Union**, v. 83, n. 51, p. 605–606, 2002.

R CORE TEAM. R: **A Language and environment for statistical computing**; R Foundation for statistical computing. Mensagem recebida por: <<http://www.gbif.org/resource/81287>> Acesso em 11 de mar. 2015.

RIZZI, R. **Geotecnologia de um sistema de estimativa da produção de soja: estudo de caso do Rio Grande do Sul**. 2005 p. 214. Tese (Doutorado na Área de Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2005. Disponível em: <<http://www.obt.inpe.br/pgsere/Rizzi-R-2004/publicacao.pdf>>. Acesso em 13 de mai. 2015.

RIZZI, R.; RUDORFF, B. F. T. Estimativa da área de soja no Riop Grande do Sul por meio de imagens Landsat. **Revista Brasileira de Cartografia**, v. 3, n. 57, p. 226–234, 2005.

ROUSE, J. W., HAAS, R. H., SCHELL, J. A., DEERING, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. In: PROCEEDINGS OF THE THIRD EARTH RESOURCE TECHNOLOGY SATELLITE-1 SYMPOSIUM/NASA, Greenbelt, **Anais ...** 1974. p. 3017. Disponível em: <<http://catalog.hathitrust.org/Record/003603707>> Acesso em 12 de jan. 2013.

RUDORFF, B. F. T.; ADAMI, M.; AGUIAR, D. A.; GUSO, A.; SILVA, W. F.; FREITAS, R. M. Temporal Series of EVI/MODIS to Identify Land Converted to Sugarcane. IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, Cape Tow, **Anais ...** 2009. p. 225. Disponível em: <http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5417326&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D5417326> Acesso em 09 abr. 2013.

RUDORFF, B. F. T.; AGUIAR, D. A. DE; SILVA, W. F. DA; SUGAWARA, L. M.; ADAMI, M.; MOREIRA, MAURÍCIO A. Studies on the Rapid Expansion of Sugarcane for Ethanol Production in São Paulo State (Brazil) Using Landsat Data. **Remote Sensing**, v. 2, n. 4, p. 1057–1076, 9 abr. 2010. Disponível em: <<http://www.mdpi.com/2072-4292/2/4/1057/>>. Acesso em: 1 mar. 2013.

RUDORFF, C. DE M.; RIZZI, R.; RUDORFF, B. F. T.; SUGAWARA, L. M.; VIEIRA, C. A. O. Spectral-temporal response surface of MODIS sensor images for soybean area classification in Rio Grande do Sul State. **Ciência Rural**, v. 37, n. 1, p. 118–125, 2007.

RURDORFF, B. F. T.; SUGAWARA, L. M. Mapeamento da cana-de-açúcar na Região Centro-Sul via imagens de satélites. **Informe Agropecuário**, v. 28, n. 241, p. 79–86, 2007.

- SANCHES, I.D. **Sensoriamento remoto para levantamento espectro-temporal e estimativa de área de culturas agrícolas**. 2001. p. 144. Dissertação (Mestrado na Área de Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2003. Disponível em: <<http://www.obt.inpe.br/pgsere/Sanches-I-D-2004/publicacao.pdf>>. Acesso em 10 de mai. 2014.
- SANCHES, I. D. A.; GÜRTLER, S.; FORMAGGIO, A. R. Discriminação de variedades de citros em imagens CCD/CBERS-2 Discrimination. **Ciência Rural**, v. 38, n. 1, p. 103–108, 2008.
- SANCHES, I. D.; EPIPHANIO, J. C. N.; FORMAGGIO, A. R. Culturas Agrícolas em imagens multitemporais do satélite Landsat. **Agricultura de São Paulo**, v. 52, n. 1, p. 83–96, 2005.
- SANO, E. E.; FERREIRA, L. G.; ASNER, G. P.; STEINKE, E. T. Spatial and temporal probabilities of obtaining cloud-free Landsat images over the Brazilian tropical savanna. **International Journal of Remote Sensing**, v. 28, n. 12, p. 2739–2752, jun. 2007. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/01431160600981517>>. Acesso em: 7 ago. 2014.
- SBM. **Sociedade brasileira de mandioca**. Mensagem recebida por: <http://www.sbmandioca.org/pagina.php?id_menu_int=5> Acesso em 10 de jan. 2015.
- SCHULTZ, B.; FORMAGGIO, A.R.; ATZBERGER, C.; LUIZ, A.J.B.; GOLTZ, E.; MELLO, M.P. Dynamic of sugarcane harvested areas in São Paulo State, Brazil, over the last two decades. In: 2ND GLP OPEN SCIENCE MEETING, 2014, BERLIN. PROCEEDINGS OF GLOBAL LAND PROJECT 2ND OPEN SCIENCE MEETING, 2014, Berlin. **Anais ...** 2014. p. 512. Disponível em: <https://www.pik-potsdam.de/members/zemp/gpl_osm_book_of_abstracts>. Acesso em: 25 jan. 2015.
- SCHULTZ, B.; FORMAGGIO, A.R.; EBERHARDT, I.D.R.; SANCHES, I.D.; OLIVEIRA, J.C.; LUIZ, A.J.B. Classificação orientada a objetos em imagens multitemporais Landsat aplicada na identificação de cana-de-açúcar e soja. **Em análise pela revista**. Acesso em: 11 fev. 2016.
- SCHULTZ, B.; LUIZ, A.J.B.; SANCHES, I.D.; FORMAGGIO, A.R. Qualidade da classificação automática de imagens de sensoriamento remoto em trabalhos apresentados nas edições anteriores do SBSR. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 17., 2015, João Pessoa. **Anais ...** São José dos Campos: INPE, 2015. p. 2357– 2364. Disponível em: <<http://www.dsr.inpe.br/sbsr2015/files/p0477.pdf>>. Acesso em 03 de jan. 2016.
- SCHULTZ, B.; IMMITZER, M.; FORMAGGIO, A.; SANCHES, I.; LUIZ, A.; ATZBERGER, C. Self-guided segmentation and classification of multitemporal

landsat 8 images for crop type mapping in Southeastern Brazil. **Remote Sensing**, v. 7, n. 11, p. 14482–14508, 30 out. 2015. Disponível em: <<http://www.mdpi.com/2072-4292/7/11/14482/>>. Acesso em: 5 nov. 2015.

SPRING. **Processamento de imagens**. Mensagem recebida por: <http://www.dpi.inpe.br/spring/portugues/tutorial/introducao_pro.html> Acesso em 10 de fev. 2015.

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. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=103288>>.

SPAROVEK, G.; BERNDES, G.; KLUG, I. L. F.; BARRETTO, A. G. O. P. Brazilian agriculture and environmental legislation: status and future challenges. **Environmental science & technology**, v. 44, n. 16, p. 6046–53, 15 ago. 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20704198>>.

STEFANSKI, J.; KUEMMERLE, T.; CHASKOVSKYY, O.; GRIFFITHS, P.; HAVRYLUK, V.; KNORN, J.; KOROL, N.; SIEBER, A.; WASKE, B. Mapping Land Management Regimes in Western Ukraine Using Optical and SAR Data. **Remote Sensing**, v. 6, n. 6, p. 5279–5305, 6 jun. 2014. Disponível em: <<http://www.mdpi.com/2072-4292/6/6/5279/>>. Acesso em: 15 set. 2015.

STEFANSKI, J.; MACK, B.; WASKE, B. Optimization of object-based image analysis with random forests for land cover mapping. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 6, n. 6, p. 2492–2504, 2013.

SUGAWARA, L. M.; FRIEDRICH, B.; RUDORFF, T. Viabilidade de uso de imagens do Landsat em mapeamento de área cultivada com soja no Estado do Paraná. **Pesquisa Agropecuária Brasileira**, v. 43, n. 12, p. 1777–1783, 2008.

TAŞDEMİR, K.; MILENOV, P.; TAPSALL, B. A hybrid method combining SOM-based clustering and object-based analysis for identifying land in good agricultural condition. **Computers and Electronics in Agriculture**, v. 83, p. 92–101, abr. 2012. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0168169912000300>>. Acesso em: 15 fev. 2014.

TELES, T. S.; GALVÃO, L. S.; BREUNIG, F. M.; BALBINOT, R.; GAIDA, W. Relationships between MODIS phenological metrics, topographic shade, and anomalous temperature patterns in seasonal deciduous forests of south Brazil. **International Journal of Remote Sensing**, v. 36, n. 18, p. 4501–4518, 2015.

TING, C.-C.; WU, B.-F.; CHUNG, M.-L.; CHIU, C.-C.; WU, Y.-C. Visual contrast enhancement algorithm based on histogram equalization. **Sensors**, v. 15, n. 7,

p. 16981–16999, 13 jul. 2015. Disponível em: <<http://www.mdpi.com/1424-8220/15/7/16981/>>. Acesso em: 30 set. 2015.

TRABAQUINI, K.; LUIZ, A.J.B.; EBERHARDT, I.; SCHULTZ, B.; FORMAGGIO, A.R.; ATZBERGER, C. Metodologia para monitoramento agrícola com emprego de imagens orbitais e amostragem estatística. In. SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 17., 2015, João Pessoa. **Anais ...** São José dos Campos: INPE, 2015, p. 111–118. Disponível em: <<http://www.dsr.inpe.br/sbsr2015/files/p0877.pdf>> Acesso em 02 de dez. 2015.

TRIMBLE GEOSPATIAL. **Definiens eCognition developer** – version 8. Mensagem recebida por: <http://www.ecognition.com/sites/default/files/eCognition%20v8_Datasheet.pdf> Acesso em 02 de mar.2015.

VALERIANO, M.M. Anuais (I): espectralradiometria. **Espaço e Geografia**, v. 6, n. 1, p. 1–21, 2003.

VALERIANO, M.M. Cálculo do fator topográfico de iluminação solar para modelagem ecofisiológica a partir do processamento de Modelos Digitais de Elevação (MDE). In. SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 15., 2011, Curitiba. **Anais ...** São José dos Campos: INPE, 2011. p. 5933– 5940. Disponível em: < <http://www.dsr.inpe.br/sbsr2011/files/p0399.pdf>> Acesso em 24 de abr. 2015.

VIEIRA, M. A.; FORMAGGIO, A. R.; RENNÓ, C. D.; ATZBERGER, C.; AGUIAR, D. A.; MELLO, M. P. Object Based Image Analysis and Data Mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas. **Remote Sensing of Environment**, v. 123, p. 553–562, ago. 2012. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0034425712001800>>. Acesso em: 23 jan. 2014.

USGS. **Landsat surface reflectance high level data products**. Mensagem recebida por: < http://landsat.usgs.gov/CDR_LSR.php> Acesso em 11 de mai. 2004.

WALDNER, F.; CANTO, G. S.; DEFOURNY, P. Automated annual cropland mapping using knowledge-based temporal features. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 110, p. 1–13, dez. 2015. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0924271615002233>>. Acesso em: 4 nov. 2015.

WALTER, A.; DOLZAN, P.; QUILODRÁN, O.; DE OLIVEIRA, J. G.; DA SILVA, C.; PIACENTE, F.; SEGERSTEDT, A. Sustainability assessment of bio-ethanol production in Brazil considering land use change, GHG emissions and socio-economic aspects. **Energy Policy**, v. 39, n. 10, p. 5703–5716, out. 2011.

Disponível em:

<<http://linkinghub.elsevier.com/retrieve/pii/S0301421510005732>>. Acesso em: 25 jan. 2014.

WARDLOW, B. D.; EGBERT, S. L. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains.

Remote Sensing of Environment, v. 112, n. 3, p. 1096–1116, mar. 2008.

Disponível em:

<<http://linkinghub.elsevier.com/retrieve/pii/S0034425707003458>>. Acesso em: 2 mar. 2013.

WILKINSON, G. G. Results and implications of a study of fifteen years of satellite image classification experiments. **IEEE Transactions on Geoscience and Remote Sensing**, v. 43, n. 3, p. 433–440, mar. 2005. Disponível em:

<<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1396317>>.

WITTEN, I.H.; FRANK, E. **Data Mining Practical Machine Learning Tools and Techniques** (2nd Ed.). Elsevier: San Francisco, CA. 525 p. 2005.

YAN, L.; ROY, D. P. Automated crop field extraction from multitemporal Web Enabled Landsat Data. **Remote Sensing of Environment**, v. 144, p. 42–64, mar. 2014. Disponível em:

<<http://linkinghub.elsevier.com/retrieve/pii/S0034425714000194>>. Acesso em: 20 mar. 2014.

YANG, X.; LO, C. P. Relative radiometric normalization performance for change detection from multi-date satellite images. **Photogrammetric Engineering & Remote Sensing**, v. 66, n. 8, p. 967–980, 2000.

YI, J. L. R.; SHIMABUKURO, Y. E.; QUINTANILHA, J. A. Identificação e mapeamento de áreas de milho na região sul do Brasil utilizando imagens modis. **Engenharia Agrícola**, v. 27, n. 3, p. 753–763, 2007.

ZHANG, C.; XIE, Z. Combining object-based texture measures with a neural network for vegetation mapping in the Everglades from hyperspectral imagery.

Remote Sensing of Environment, v. 124, p. 310–320, set. 2012. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0034425712002180>>. Acesso em: 25 mar. 2014.

ANEXO A

Script 1 desenvolvido em linguagem *Python*.

ApEq – extração de valores APIs Desenvolvido por: Bruno Schultz e Gabriel Bertani

```
# Import system modules
import arcpy
from arcpy import env
from arcpy.sa import *
# Check out the ArcGIS Spatial Analyst extension license
arcpy.CheckOutExtension("Spatial")
arcpy.env.workspace = "Diretório"
text_file = open("Diretório/saída.txt", "w")
Hansen = "Hansen_GFC2014_treecover2000_20S_060W.tif"
rasters = arcpy.ListRasters("L*band*.tif")
rasters_nuvens = arcpy.ListRasters("L*cfmask.tif")
rasters_dec = "Declividade_MArques1971_WGS84.tif"
loss = "Hansen_GFC2014_loss_20S_060W.tif"
features = arcpy.ListFeatureClasses("OLI*")
for feature in features:
    arcpy.Clip_management(Hansen, "#", "Clip_Hansen_" + feature[:12] + ".tif", feature, "", "ClippingGeometry")
    arcpy.Clip_management(rasters_dec, "#", "Clip_Declividade" + feature[:12] + ".tif", feature, "", "ClippingGeometry")
    arcpy.Clip_management(loss, "#", "Clip_Loss" + feature[:12] + ".tif", feature, "", "ClippingGeometry")
    hansen_mask = Con("Clip_Hansen_" + feature[:12] + ".tif", 1, 0, "VALUE > 99") # gero máscara no HANSEN
    hansen_mask.save("Mask_hansen" + feature[:12] + ".tif")
    Decl_mask = Con("Clip_declividade" + feature[:12] + ".tif", 1, 0, "VALUE < 2") # gero máscara de DECLIVIDADE
    Decl_mask.save("Mask_Declividade" + feature[:12] + ".tif")
for raster_nuvem in rasters_nuvens:
    nuvem_mask = Con(raster_nuvem, 1, 0, "VALUE > 0") # gero máscara de NUVEM
    nuvem_mask.save("Mask_nuvem" + raster_nuvem[3:17] + ".tif")
    print(raster_nuvem)
for feature in features:
    for raster_nuvem in rasters_nuvens:
        if str(feature[4:7]) + feature[8:11] == str(raster_nuvem[3:9]):
            ND_mask = Con("Mask_nuvem" + raster_nuvem[3:17] + ".tif", "Mask_Declividade" + feature[:12] + ".tif", 0,
"VALUE = 0") # aplico máscara de nuvem sobre a máscara de Declividade
            ND_mask.save("ND_mask_" + raster_nuvem[3:17] + ".tif")
            NHD_mask = Con("ND_mask_" + raster_nuvem[3:17] + ".tif", "Mask_hansen" + feature[:12] + ".tif", 0, "VALUE =
1") # aplico máscara de floresta sobre a máscara de Declividade e nuvem
            NHD_mask.save("NHD_mask_" + raster_nuvem[3:17] + ".tif")
            NHDL_mask = Con("NHD_mask_" + raster_nuvem[3:17] + ".tif", "Clip_L_" + feature[:12] + ".tif", 0, "VALUE =
1") # aplico máscara de LOSS sobre a máscara de floresta, declividade e nuvem
            NHDL_mask.save("NHDL_mask_" + raster_nuvem[3:17] + ".tif")
for raster_nuvem in rasters_nuvens:
    for reflectancia in rasters:
        if str(reflectancia[3:17]) == str(raster_nuvem[3:17]):
            final = Con("NHDL_mask_" + raster_nuvem[3:17] + ".tif", reflectancia, "", "VALUE = 1") # aplico máscaras
geradas sobre os valores de reflectancia
            final.save("Final_" + raster_nuvem[3:17] + "Banda_" + reflectancia[29] + ".tif")
final = arcpy.ListRasters("Final*.tif")
for imagem in final:
    print(imagem)
    MEANResult = arcpy.GetRasterProperties_management(imagem, "MEAN")
    text_file.write("MEAN_Banda_" + imagem[26] + "_" + imagem[7:13] + ":" + str(MEANResult) + " ")
for raster_nuvem in rasters_nuvens:
    for reflectancia in rasters:
        if str(reflectancia[3:17]) == str(raster_nuvem[3:17]):
            print(reflectancia[3:17] + " ")
            print(raster_nuvem[3:17])
            print("Final_" + raster_nuvem[3:17] + "Banda_" + reflectancia[29])
            MEANResult = arcpy.GetRasterProperties_management("Final_" + raster_nuvem[3:17] + "Banda_" +
reflectancia[29] + ".tif", "MEAN")
            text_file.write("MEAN_Banda_" + reflectancia[29] + "_" + raster_nuvem[3:17] + ":" + str(MEANResult) + " ")
text_file.close()
```

#####

ANEXO B

Script 2 desenvolvido em linguagem *Python*.

```
##### Montagem dos screenshot Desenvolvido por: Bruno Schultz e Gabriel Bertani #####

import arcpy
from arcpy import env
from arcpy.sa import *
intervalo = range(1,200)
mxd = arcpy.mapping.MapDocument(r"diretório/arquivo_ArcGIS.mxd")
df = arcpy.mapping.ListDataFrames(mxd, "rgb*")
df_zoom = arcpy.mapping.ListDataFrames(mxd, "Zoom*")
for ponto in intervalo:
    for dataframe in df:
        lista = arcpy.mapping.ListLayers(mxd, "pontos*", dataframe)
        lista2 = arcpy.mapping.ListLayers(mxd, "rgb_imagens*", dataframe)
        print(lista)
        where = str ( "Id = " +str(ponto))
        print(where)
        arcpy.SelectLayerByAttribute_management(lista[0], "NEW_SELECTION", where )
        dataframe.scale = 30000
        dataframe.zoomToSelectedFeatures()
        lista[0].visible = True
        lista2[0].visible = True
        lixt = arcpy.mapping.ListLayoutElements(mxd, "TEXT_ELEMENT")
        for texto in lixt:
            if texto.text == "Id":
                texto.text = "Id " + str (ponto)
        arcpy.RefreshTOC()
        arcpy.RefreshActiveView()
    for dataframe in df_zoom:
        lista = arcpy.mapping.ListLayers(mxd, "pontos*", dataframe)
        lista2 = arcpy.mapping.ListLayers(mxd, "rgb_imagens*", dataframe)
        print(lista)
        where = str ( "Id = " +str(ponto))
        print(where)
        arcpy.SelectLayerByAttribute_management(lista[0], "NEW_SELECTION", where )
        dataframe.scale = 1000
        dataframe.zoomToSelectedFeatures()
        lista[0].visible = True
        lista2[0].visible = True
    arcpy.mapping.ExportToPDF(mxd, "Diretório\PDFs\\" + "Id_" + str(ponto) + "_" + ".pdf")
    for texto in lixt:
        if texto.text == "Id " + str (ponto):
            texto.text = "Id"
text_file.close()
```

#####

ANEXO C

Script 3 desenvolvido em linguagem *Python*.

Extração porcentagem nuvem Desenvolvido por: Bruno Schultz e Gabriel Bertani

```
import arcpy
from arcpy import env
from arcpy.sa import *
env.workspace = "Diretório"
lista = arcpy.ListFiles("*.txt")
print(lista)
filew = open("nome_saída.txt", "w")
filew.write (' PATH ROW DATE CLOUD_COVER \n')
for txt in lista:
    texto = txt.strip('u')
    filex = open('Diretório\\' + str(texto), 'r')
    for line in filex:
        if 'TARGET_WRS_PATH' in line:
            filew.write (line.split("=")[1].strip('\n'))
        if 'TARGET_WRS_ROW' in line:
            filew.write (line.split("=")[1].strip('\n'))
        if 'DATE_ACQUIRED ' in line:
            filew.write (line.split("=")[1].strip('\n'))
        if 'CLOUD_COVER ' in line:
            filew.write (line.split("=")[1])
filew.close()
```

#####