

Metodologia para geração de máscaras de culturas de verão para o ano-safra 2007/2008 no Estado do Paraná por meio de índices de vegetação do sensor MODIS

Daniel Garbellini Duft¹;
Jerry Adriani Johann^{1,2};
Jansle Vieira Rocha¹;
Rubens Augusto Camargo Lamparelli³

¹ Universidade Estadual de Campinas- FEAGRI/UNICAMP,
C.P. 6011, CEP:13081-970 - Campinas, SP, Brasil
danielduft@gmail.com; jansle.rocha@feagri.unicamp.br;

² Universidade Estadual do Oeste do Paraná – UNIOESTE,
C.P. 0701, CEP: 85819-110 – Cascavel, PR, Brasil
jerry.johann@hotmail.com;

³ Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura – CEPAGRI,
Campinas, SP, Brasil
rubens@cpa.unicamp.br

Abstract: This paper proposes a methodology to map planted areas with corn and soybean in Paraná state, Brazil, using MODIS images. EVI images of 2007/2008 cropping season were gathered. The cropland mask was obtained using the difference of spectral response between top of crop vegetative cycle and planting date. The procedure was to use images of minimum and maximum EVI in a RGB composition. This composition was digitally classified to generate a cropland mask for corn and soybean, and the classification accuracy assessment carried out using LANDSAT 5/TM and AWiFS images as reference to build error matrix and calculate Kappa Index of Agreement (KIA) and Overall Accuracy (OA). Checking the area calculated from the cropland mask against IBGE official statistics showed that the mask area underestimated in 4,64% the official data, with an $R^2=0,895$. Comparing the cropland area of five municipalities located in main producing areas over the state, has shown that IBGE figures have some inconsistency with reality mapped from Landsat images. The mask calculated areas were closer to reality than IBGE figures. The $OA=94,72\%$ and $KIA=0,8945$ revealed that the cropland mask has good spatial accuracy, showing that the methodology proposed is useful for agricultural cropland mapping.

Palavras-chave: remote sensing, EVI, cropland mask, sensoriamento remoto, máscara de culturas.

1. Introdução

As culturas de verão são responsáveis por mais de 70% da produção de grãos no país. Isso deve-se principalmente pela aptidão climática do período que traz uma maior produtividade às culturas (CONAB, 2009).

Segundo IBGE (2010), o estado do Paraná tornou-se o maior produtor de grãos do país na safra 2010 sendo responsável por 21% da produção do país. As culturas de verão foram responsáveis por 27 milhões de toneladas dos 30 milhões produzidos na safra, evidenciando assim a importância das mesmas. Ainda segundo o IBGE (2009) a agricultura já é responsável por 19% do PIB do estado que por sua vez tem uma economia que é 6% do total da economia do país. Analisar e monitorar essas culturas neste estado são ações estratégicas para uma previsão de colheita e produção. Os métodos atuais de previsão de safras baseiam-se em pesquisas com produtores rurais e utilização de dados históricos, sendo sujeitos a subjetividade e imprecisões.

O cenário político e econômico da agricultura atual exige que haja agilidade e qualidade na aquisição de dados que possam servir de instrumento para tomadas de decisões. As estimativas de área cultivada e de produção são extremamente importantes para que possa ser feito um planejamento mais preciso, seja do setor público, ou do setor privado, para o suporte da safra.

O sensoriamento remoto proporciona a utilização de imagens e dados de sensores instalados em satélites, necessários para um monitoramento eficaz em áreas agrícolas. Para este monitoramento, um dos sensores mais utilizados é o MODIS (*Moderate Resolution Imaging Spectroradiometer*), instalado no satélite TERRA da NASA. Com este sensor é possível gerar produtos capazes de captar vegetação, ou seja, é possível estudar as áreas cultivadas e analisá-las para que seja possível reconhecer culturas. Isso pode ser feito através do *Enhanced Vegetation Index* (EVI) que é um índice de vegetação, gerado a partir de imagens de satélites, utilizado em estudos de áreas agrícolas e que fornece informações sobre a biomassa da cultura, levando em conta interações com o solo e corrigindo ruídos atmosféricos. É calculado em uma relação matemática que leva em conta a reflectância na banda do espectro do vermelho, azul e do infravermelho, além de fatores que fazem as correções já citadas.

Apesar das suas virtudes em relação a tempo e precisão, o sensoriamento remoto, na maioria das vezes, ainda é caro e de difícil manejo. Para que seja possível uma previsão de safra útil aos órgãos e empresas competentes é necessário que o processo para tal seja algo fácil e baixo ou nenhum custo. O grande desafio então é aliar qualidade aos atributos necessários no levantamento de dados oficiais, ao tornar isso possível, a previsão de safras passará a colaborar para uma maior eficiência da agricultura brasileira.

O objetivo deste trabalho é mostrar uma maneira prática para conseguir-se mapear áreas de culturas de verão no Estado do Paraná, através de imagens disponibilizadas gratuitamente, para o ano-safra 2007/2008.

2. Metodologia de Trabalho

O estudo foi realizado para os 399 municípios do Estado do Paraná. Para analisar os dados, os municípios foram subdivididos em 4 grupos, de acordo com o quartis, levando-se em conta área plantada de milho e soja (ha) informada pelo IBGE (2008). Assim, a subdivisão ficou da seguinte forma: do mínimo (Min) até o quartil 1 (Q1) estão 25% dos municípios com menor área plantada; do Q1 a mediana (Md = 50%) estão os municípios com esta proporção de área plantada; da Md ao quartil 3 (Q3) os municípios com área plantada de soja e milho entre 50% e 75%; e, finalmente entre o Q3 e o máximo(Max), os municípios com maior área plantada do Paraná. A Figura 1 mostra a espacialização das subdivisões em quartis.

Para o mapeamento e estimativa de área plantada das culturas de soja e milho, no período de verão, no Estado do Paraná, foram utilizadas as imagens do sensor MODIS, produto MOD13Q1 do "Tile" h13v11 (NASA, 2009). Este produto contém composições de imagens de 16 dias do índice de vegetação EVI, com uma resolução espacial de 250m.

Após a extração do produto EVI para as 23 imagens do ano-safra 2007/08 (cenas 214/2007 a 225/2008), estas foram recortadas, compiladas e analisadas. Entretanto, para este estudo foram utilizadas somente as cenas 241 (29/08/06) a 113 (23/04/07) gerando-se o perfil espectral multi-temporal de EVI como exemplifica a Figura 2.

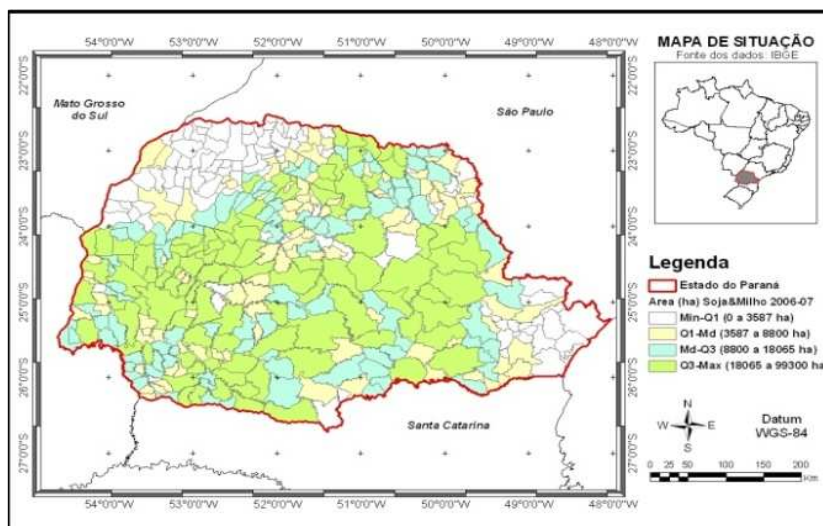


Figura 1: Mapa de localização dos 399 municípios do Estado do Paraná dividido por quartis

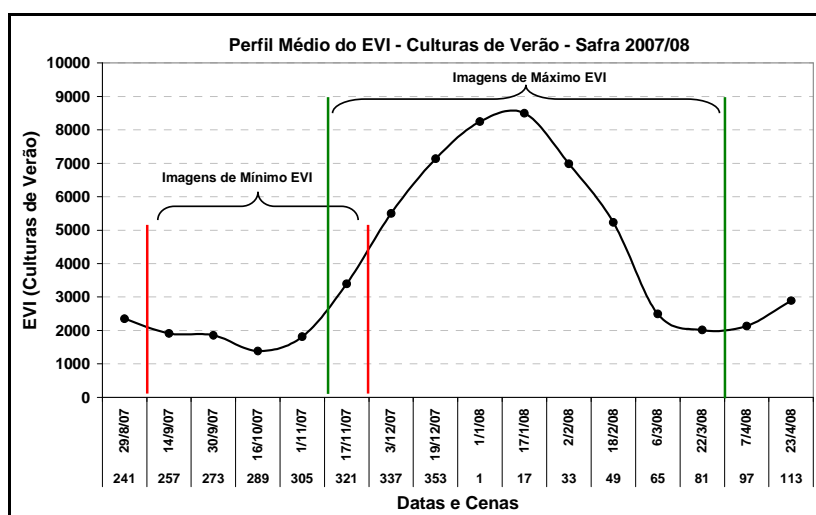


Figura 2: Perfil temporal de EVI de talhões de Soja e Milho no município de Cascavel/PR.

Os perfis de EVI da média dos pixels de municípios tipicamente produtores de soja e milho são muito parecidos com o EVI de pixels puros da cultura da soja e do milho. Um problema é que os perfis dessas culturas apresentam uma curva semelhante, o que dificulta a separação das mesmas. Além disto, diferentes épocas de plantio, bem como cultivares com diferentes ciclos de maturação influenciam no comportamento do EVI e conseqüentemente no perfil espectral em todo o Estado. Neste sentido, faz-se necessário uma amostragem por regiões do Estado para poder-se conhecer o comportamento espectral multi-temporal do EVI, trabalho este demorado e que muitas vezes não traz resultados satisfatórios para a análise global das culturas desejadas.

Para solucionar tal problema, propôs-se a utilização de múltiplas datas para mapear-se as áreas de plantio (máscaras) das culturas. O procedimento consistiu em utilizar várias imagens para gerar a chamada “Imagem de Mínimo EVI” e “Imagem de Máximo EVI” ao longo do Estado, fazendo assim uma composição destas imagens.

Para composição destas imagens, como ilustra a Figura 2, foram utilizadas das cenas 257 (29/08/07) a 321 (17/11/07) para geração da “Imagem de Mínimo EVI” e das cenas 321 (17/11/07) a 81 (22/03/08) para geração da “Imagem de Máximo EVI”. Desta forma garantiu-

se abranger os mínimos e máximos das culturas para cada uma das diferentes regiões do Estado, uma vez que a data de plantio e tipos de cultivares não são uniformes. Através do programa ENVI, a “Imagem de Máximo EVI” foi colocada no canal R, enquanto que a “Imagem de Mínimo EVI” foi colocada nos canais G e B da composição colorida. Através de um sistema de extração de dados de imagens desenvolvido por Esquerdo (2007) em linguagem IDL, foi extraída essa composição RGB em nível de cinza (NC). Foram então gerados cortes em valores de NC (0-255) a fim de garantir valores que representassem apenas as culturas de verão, ou seja, valores de NC que restrinjam valores de EVI máximo e mínimo capaz de fazer esta separação.

A verificação e avaliação dos resultados foram baseadas em três parâmetros: a comparação com dados municipais de área plantada disponibilizados pelo IBGE, o Índice de Exatidão Global (EG) e o Índice Kappa (IK) (Congalton, 1991) sendo estes dois últimos gerados pelo confronto da máscara com imagens do sensor Landsat5-TM e AWFIS-IRS. Os dados do IBGE foram obtidos através da base de dados SIDRA, de cada um dos 399 municípios, para as datas estudadas. Para compará-los aos dados da máscara foi feita uma extração de área da máscara para cada município do Estado, esses dados foram então cruzados e avaliados pelo coeficiente de determinação (R^2).

Além disto, com o intuito de analisar mais detalhadamente os resultados encontrados pela máscara gerada, foram selecionados 5 municípios divididos pelas áreas de maior produção do Estado, segundo o IBGE, para que sua área plantada com soja e milho fossem vetorizadas manualmente em um mosaico de imagens do LANDSAT5/TM a fim de obter-se a área real.

3. Resultados e Discussão

Para confeccionar a máscara foram separadas as áreas de cultura (milho+soja) e não-cultura (outros alvos). A Figura 3 mostra o resultado da composição colorida RGB (Canal R → Imagem de Máximo EVI; Canais GB → Imagem de Mínimo EVI). Observa-se que a imagem resultante tem coloração vermelha, que representa as áreas de cultura e as demais colorações (tons de verde e azul) representam as áreas com outras culturas ou alvos (não-cultura).

Assim, para conhecer a máscara das culturas fez-se necessário a extração dos pixels com coloração vermelha. Para isto, considerou-se o nível de cinza (NC) de cada uma das imagens (de máximo e mínimo EVI) e através de simulações. Neste sentido, pesquisou-se os limites de corte (mínimo e máximo) para o canal R e outro para o canal GB, uma vez que esses estão com a mesma imagem. Como cada pixel da imagem (Figura 3) possui NC variando de 0 a 255, o procedimento faz com que sejam escolhidos valores limítrofes: no canal R valores de NC acima do escolhido e nos canais G e B valores abaixo do escolhido. Desta forma, serão então extraídos os pixels que possuem tais restrições, o que dará origem à máscara para as culturas de verão de interesse.

Dentre inúmeras simulações realizadas, o melhor resultado de corte (mínimo e máximo) para geração da máscara das culturas de verão (soja + milho) no Estado do Paraná foi obtido com R160G150B150. O resultado da máscara destas culturas para o ano-safra 2007/08 é mostrado na Figura 4. Observa-se claramente numa comparação entre as Figuras 3 e 4, que a áreas de coloração vermelha na Figura 3 coincidem com a máscara na Figura 4.

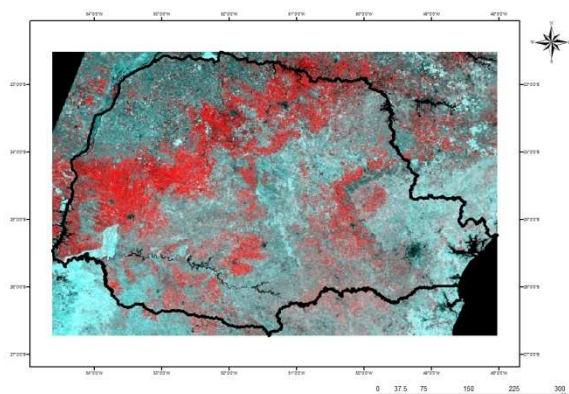


Figura 3: Composição RGB de compilações das Imagens de Máximo EVI e de Mínimo EVI.

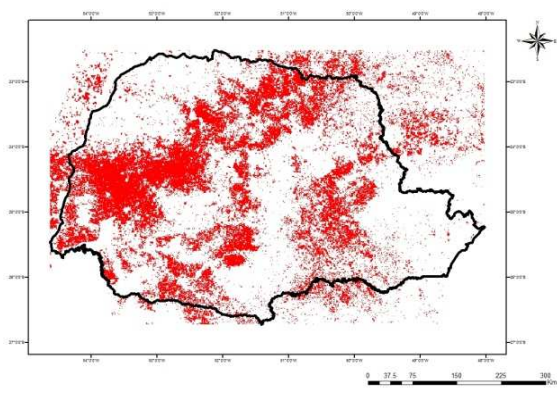


Figura 4: Máscara para as culturas de soja e milho para a safra 2007/2008 no Paraná.

A máscara gerada a partir desta combinação (R160G150B150) teve sua área plantada (ha) extraída para cada município a qual foi confrontada com a área plantada (ha) oficial informada pelo IBGE.

A Figura 5 mostra a relação entre as áreas obtidas pela máscara e a informada pelo IBGE. O valor de R^2 encontrado para esta relação foi de 0,895, mostrando um bom resultado para estudos de mapeamento de culturas. Segundo o IBGE, para o ano-safra 2007/08 foram plantadas 5.351.671 ha de soja e milho (período de verão). A máscara encontrou uma área de 5.103.273 ha, ou seja, ela subestimou apenas em 4,64% a área total do Estado destas culturas.

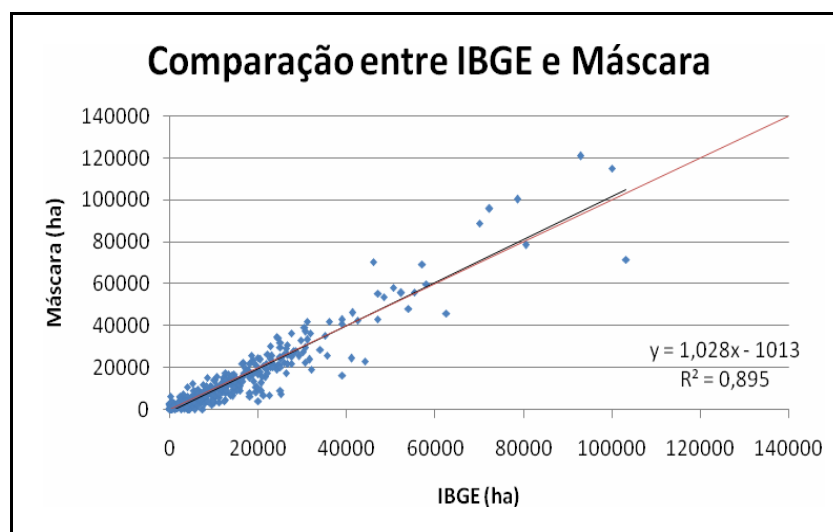


Figura 5: Relação entre a área plantada (ha) por município de soja+milho da máscara e os dados oficiais do IBGE.

Já para a averiguação da exatidão espacial da máscara gerada determinou-se o índice de Exatidão Geral (EG) e o Índice de Concordância Kappa (IK) tomando-se como referência imagens de alta resolução espacial (LANDSAT 5/TM e AWiFS/IRS). Para que o erro amostral desta averiguação fosse de $\pm 5\%$ determinou-se que seriam necessárias um total de 400 amostras (ou seja, 400 pontos). Desta forma, foram sorteadas aleatoriamente em dois estratos de 200 amostras, dentro e fora da máscara. Desta forma, evitou-se que erros de omissão e inclusão fossem mascarados na análise por haver diferentes áreas envolvidas.

Os resultados deste procedimento são apresentados na Tabela 1. Obteve-se uma $EG=94,72\%$ e um $IK=0,8945$ que segundo Landis e Koch (1977) são considerados excelentes. Os valores são ainda mais significativos se comparados com outros trabalhos que utilizaram o mesmo sensor para o mesmo fim, porém com outra metodologia. Por exemplo, Lamparelli et al. (2008) também trabalhando com mapeamento de soja no Paraná, com dados

dos satélites *MODIS/Terra e TM/Landsat 5* encontraram valores de kappa que variam de 0,6 a 0,8 que foram bem menores que o encontrado por esta metodologia. Já no trabalho de Lindsey e Mueller (2008) o valor de kappa máximo encontrado foi de 0,7832. Portanto é possível concluir que a metodologia apresentada possui alto potencial de utilização para mapeamento de culturas de verão.

Quanto a matriz de erros apresentada na Tabela 1, pode-se observar que os erros de inclusão e omissão são bem parecidos, isso quer dizer que a máscara está equilibrada, sem erros demais seja por omissão ou por inclusão. Valores de no máximo 6% são responsáveis por gerar este valor de IK e EG.

Tabela 1: Matriz de erros para 400 pontos aleatórios em dois extratos

ESPERADO (Máscara)	REAL (Landsat+AWIFS) Padrão		Total Geral	Erro de Inclusão
	Culturas de verão	Outras culturas		
Culturas de verão	191	12	203	5,9%
Outras culturas	9	188	197	4,6%
Total geral	200	200	400	-
Erro de Omissão	4,5%	6,1%	-	-

As áreas encontradas pela vetorização das culturas de verão (soja e milho) para os municípios de Campo Mourão, Cascavel, Castro, Guarapuava e Londrina (apresentados na Figura 6), foram cruzados com os valores da máscara e do IBGE. Através da Tabela 2 é possível observar que os valores de área oriundos da vetorização de áreas de culturas de verão sobre as imagens LANDSAT5/TM (reais) são mais próximos aos valores de área da máscara do que dos valores de área lançados pelo IBGE. Isso sugere que o método de levantamento de área do IBGE pode estar trazendo erros devido a subjetividade embutida na coleta de dados. Já a diferença entre a máscara e os valores vetorizados pode ser explicada pelo tamanho do pixel do MODIS (250mx250m) que é de 6,25ha. Desta forma, talhões menores que este valor não são selecionados pela máscara, assim como outros alvos em meio a grandes talhões são ignorados. Porém, pelo resultado pode-se afirmar que em um contexto geral a máscara possui um bom nível de acerto.

Comparando-se os valores de área do IBGE com a máscara e da vetorização com a máscara, é possível observar que, para o município de Londrina a diferença de área entre o IBGE e a máscara é de 53%, enquanto a diferença de área entre a vetorização e a máscara é de apenas 5%. Já a maior diferença entre máscara e vetorização está no município de Castro (18%). Isso acontece devido ao relevo irregular, que é utilizado para o plantio das culturas de verão.

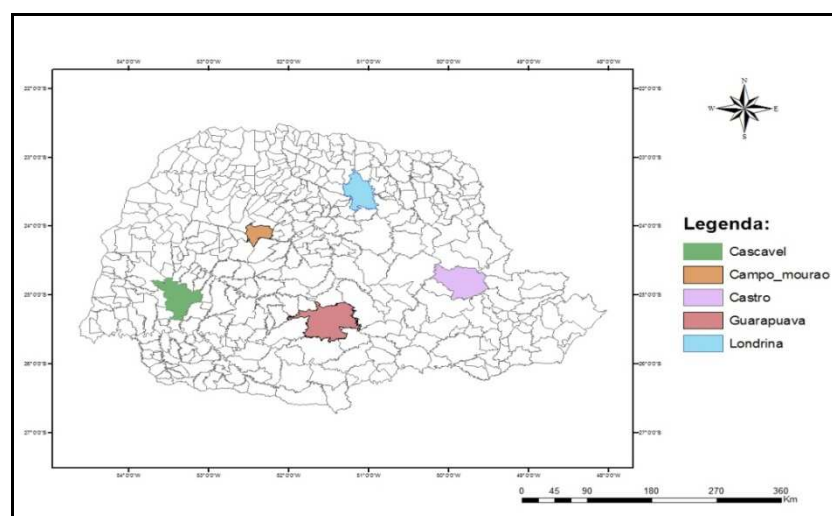


Figura 6: Municípios selecionados para análise de área

Tabela 2: Análise entre dados do IBGE, da máscara e vetorizados

Municípios	Área Colhida (ha) - IBGE			Área (ha)	
	Milho	Soja	Milho+Soja	Máscara	Vetorizada
Campo Mourão	5500	48500	54000	48113	46214
Cascavel	10000	82850	92850	120888	115766
Castro	30861	72300	103161	71531	84677
Guarapuava	29800	48750	78550	100288	95455
Londrina	6000	40000	46000	70475	66899

A vetorização dos talhões é considerada como valor real, pois utiliza imagens de alta resolução espacial, a Figura 7 mostra um exemplo de como foi feita a vetorização para os talhões de soja e milho nos municípios citados na Tabela 2. Já a Figura 8 mostra como a máscara comportou-se em relação a vetorização.

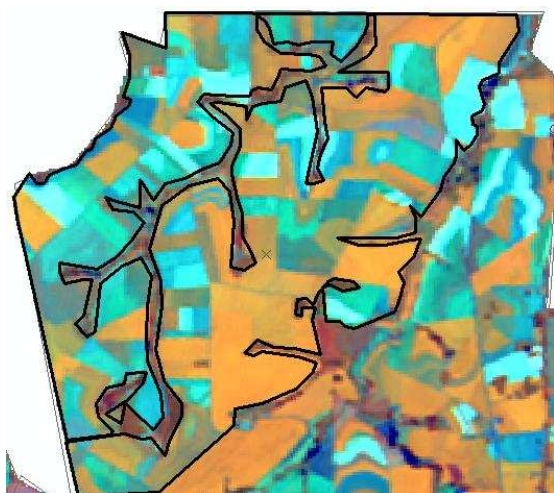


Figura 7: Figura esquemática de vetorização de talhões de culturas de soja e milho para região do município de Cascavel.

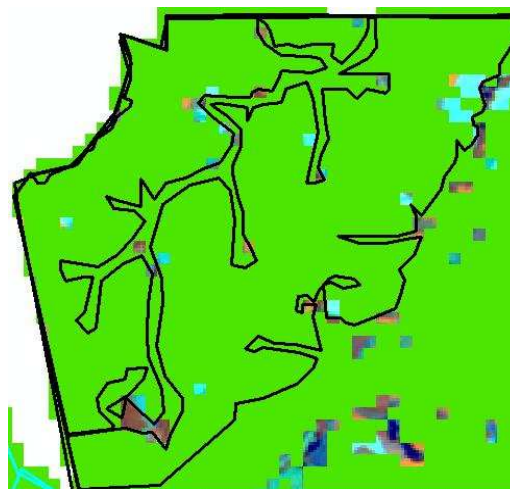


Figura 8: Máscara para as culturas de soja e milho para a safra 2007/2008 sobreposta à vetorização dos talhões da figura 7.

4. Conclusões

A metodologia proposta permitiu mapear as áreas das culturas da soja e milho para o ano-safra 2007/2008 no Estado do Paraná, com bom índice de exatidão, colocando-se como uma alternativa aos órgãos oficiais responsáveis por estatísticas de produção agrícola, pois apresenta as vantagens da rapidez e objetividade, além da espacialização dessas áreas por município.

Agradecimentos

Ao GEO- Grupo de Estudos em Geoprocessamento pela oportunidade de realizar o trabalho e ao CNPq pela bolsa de Iniciação Científica através do PIBIC/UNICAMP.

Referências Bibliográficas

CONAB – COMPANHIA NACIONAL DE ABASTECIMENTO. **Levantamentos de Safra**. 2009. Disponível em: <<http://www.conab.gov.br/conteudos.php?a=1253&t=2>>. Acesso em 12 out. 2010.

CONGALTON, R.G. A review of assessing the accuracy classificatons of remotely sensed data. **Remote Sensing Environment**, v. 37, p.35-46, 1991.

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

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Banco de Dados Agregados - Sistema IBGE de Recuperação Automática – SIDRA**. 2009. Disponível em: <<http://www.sidra.ibge.gov.br>>. Acesso em 09 out. 2009.

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Levantamento Sistemático da Produção Agrícola – Indicadores econômicos**. 2010. Disponível em: <<http://www.ibge.gov.br/home/estatistica/indicadores/agropecuaria/lspa>>. Acesso em 12 out. 2010.

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

LINDSEY, M; MUELLER, R. M. **Agricultural Land-Use Classification for California Using AWiFS and MODIS Data**. University of Maryland. Maryland, 2009. Disponível em: <http://www.nass.usda.gov/research/sarsblog/Lindsey_AAG08_CaCDL.pdf>. Acesso em 20 out. 2010.

NASA - NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. **Technical specifications: Moderate Resolution Imaging Spectroradiometer (MODIS)**, 2009. Disponível em: <<http://modis.gsfc.nasa.gov/about/design.php>>. Acesso em 06 jun. 2009.

LAMPARELLI, R.A.C.; CARVALHO, W. M. O. de; MERCANTE, E.. Mapeamento de semeaduras de soja (*Glycine max* (L.)Merr.) mediante dados *MODIS/Terra E TM/Landsat 5*: um comparativo. **Engenharia Agrícola**, Jaboticabal, v.28, n.2, p.334-344, abr./jun. 2008.