

Estimativa de áreas com culturas de verão no Paraná, por meio de imagens multitemporais EVI/Modis

Jerry Adriani Johann⁽¹⁾, Jansle Vieira Rocha⁽²⁾, Daniel Garbellini Duft⁽²⁾ e Rubens Augusto Camargo Lamparelli⁽³⁾

⁽¹⁾Universidade Estadual do Oeste do Paraná, Campus Cascavel, Rua Universitária, nº 2.069, Jardim Universitário, CEP 85819-110 Cascavel, PR. E-mail: jerry.johann@hotmail.com ⁽²⁾Universidade Estadual de Campinas (Unicamp), Faculdade de Engenharia Agrícola, Avenida Cândido Rondon, nº 501, CEP 13083-875 Campinas, SP. E-mail: jansle.rocha@feagri.unicamp.br, danielduft@gmail.com ⁽³⁾Unicamp, Núcleo Interdisciplinar de Planejamento Energético, Rua Cora Coralina, nº 330, CEP 13083-896 Campinas, SP. E-mail: rubens.lamparelli@gmail.com

Resumo – O objetivo deste trabalho foi estimar e mapear as áreas com as culturas de soja e milho, no Paraná, com uso de imagens multitemporais EVI/Modis. Foram avaliados os anos-safra de 2004/2005 a 2007/2008. Em razão da alta dinâmica temporal e da heterogeneidade de datas de semeadura das culturas no estado, foram utilizadas cenas que contemplavam as fases de pré-plantio e de desenvolvimento inicial das culturas, para gerar a imagem de mínimo EVI (IMIE), e cenas que consideravam o pico vegetativo das culturas, para gerar a imagem de máximo EVI (IMAE). Estas imagens foram utilizadas para gerar a composição colorida RGB (R, IMAE; GB, IMIE), o que permitiu a confecção de máscara das áreas com soja e milho. As estimativas das áreas de máscara por município foram comparadas com dados oficiais de produção agrícola municipal, tendo-se observado bons ajustes ($R^2 > 0,84$, $d > 0,95$, $c > 0,85$) entre os dados. Para a avaliação da exatidão espacial das máscaras, imagens Landsat-5/TM e AWiFS/IRS foram usadas como referência para construção da matriz de erros. Os resultados obtidos são indicativos de que a metodologia proposta é altamente eficiente e pode ser utilizada para mapeamento dessas culturas.

Termos para indexação: classificação de imagens, distribuição espacial de culturas, índice de vegetação, mapeamento, previsão de safras, sensoriamento remoto.

Estimation of summer crop areas in the state of Paraná, Brazil, using multitemporal EVI/Modis images

Abstract – The objective of this work was to estimate and map crop areas with soybean and corn in the state of Paraná, Brazil, using EVI/Modis images. The crop seasons from 2004/2005 to 2007/2008 were evaluated. Due to the high temporal dynamics and difference in sowing dates of the cultures within the state, scenes containing the pre-planting and initial crop development phases were used to obtain the minimum EVI image (IMIE), and scenes at the peak of the crop cycle were used to obtain the maximum EVI image (IMAE). These images were used to generate the RGB color composition (R, IMAE; GB, IMIE), which allowed for the creation of masks of the areas planted with soybean and corn. The estimation of masked areas by municipality was compared with the municipal agricultural production official data, and good fits ($R^2 > 0.84$, $d > 0.95$, $c > 0.85$) were observed between data. For spatial accuracy assessment, Landsat-5/TM and AWiFS/IRS images were used as references to build the error matrix. The obtained results indicate that the proposed methodology is highly efficient and may be used as a model for cropland mapping.

Index terms: image classification, crop spatial distribution, vegetation index, mapping, crop forecast, remote sensing.

Introdução

Os cultivos de verão são responsáveis por mais de 70% da produção de grãos no país (Companhia Nacional de Abastecimento, 2009b). Segundo dados do Instituto Brasileiro de Geografia e Estatística (2011), o Estado do Paraná produziu 10,63 milhões de toneladas de soja e 8,12 milhões de toneladas de

milho ao ano, no período compreendido entre 2005 e 2008. Essas culturas foram responsáveis por 92,5% da área e de 96,6% da produção dos cultivos de verão, no estado. Esses valores equivalem a 18,42 e 14,75% das áreas, e 19,16 e 23,88% das produções de soja e milho, respectivamente, no Brasil.

Atualmente, as estimativas de área e de produção agrícola do Brasil são realizadas pelo Instituto Brasileiro

de Geografia e Estatística (IBGE) e pela Companhia Nacional de Abastecimento (Conab), a partir de pesquisas com produtores rurais e cooperativas, de dados de financiamento agrícola e da utilização de dados históricos. A natureza desse tipo de informação torna as estimativas de safra sujeitas à subjetividade e à imprecisão. Além disso, os dados oficiais não trazem informações sobre a distribuição espacial da produção, imprescindíveis para a realização de estimativas de produtividade de culturas, dada a alta correlação desse parâmetro com as condições agrometeorológicas.

Apesar de a metodologia utilizada por órgãos oficiais ainda ser bastante empregada (Adami et al., 2010), a proposição e a incorporação de novos métodos é importante para complementar os dados da metodologia adotada em programas oficiais. O projeto Geosafra tem desenvolvido e testado metodologias objetivas, com uso de geotecnologias (Companhia Nacional de Abastecimento, 2009a). Várias metodologias também têm sido testadas com uso do sensoriamento remoto, com relativo êxito (Adami et al., 2004; Rizzi & Rudorff, 2005; Rojas et al., 2005; Lamparelli et al., 2008; Sugawara et al., 2008; Mercante et al., 2010; Esquerdo et al., 2011). Essas metodologias de mapeamento por imagens fornecem a localização e a quantificação das áreas plantadas com culturas de verão, o que pode direcionar e facilitar o trabalho de campo dos avaliadores do IBGE.

O desafio no cálculo das áreas cultivadas, para fins de previsão de safra, está no desenvolvimento de metodologia sistemática, em escala regional, que permita operacionalizar esta fase de acordo com o cronograma dos levantamentos oficiais. Para atender a esta finalidade, um dos sensores orbitais que mais tem se destacado nesses levantamentos é o “moderate resolution imaging spectroradiometer” (Modis) que, apesar de apresentar resolução espacial apenas moderada (250 m), é uma alternativa interessante em razão das características: cobertura global diária, elevada acuidade radiométrica, facilidade de aquisição, larga faixa de imageamento (imagens disponibilizadas gratuitamente, em cenas de 1.200x1.200 km) e disponibilidade de bandas espectrais próprias para o monitoramento da vegetação (Huete et al., 2002; Ferreira et al., 2008).

Diversos trabalhos utilizam o índice de vegetação “normalized difference vegetation index” (NDVI),

proposto por Rouse et al. (1973) para mapeamento de áreas vegetadas. Contudo, este índice apresenta, como problema, a possibilidade de saturação, ou seja, mantém-se praticamente o mesmo valor de NDVI a partir de índices de área foliar por volta de 4 ou 5, o que o torna insensível ao aumento da biomassa, mesmo quando há aumento da densidade do dossel da cultura (Gamon et al., 1995). Nesse sentido, o índice de vegetação “enhanced vegetation index” (EVI) tem se mostrado adequado para o monitoramento agrícola, pois substitui, com vantagens, o NDVI, por ser menos suscetível à saturação e mais sensível à variação da estrutura, arquitetura do dossel e fisionomia da planta (Huete et al., 2002).

Para o monitoramento da atividade agrícola é imprescindível a análise de múltiplas datas, para obter informações do comportamento espectro-temporal (perfil) ao longo de todos os estádios fenológicos da cultura (Holben, 1986; Jensen et al., 2002). Diversos trabalhos relatam a aplicação deste procedimento, para o sensor Modis (Yi et al., 2007; Ren et al., 2008; Wardlow & Egbert, 2008; Epiphanyo et al., 2010; Gusso et al., 2012).

O Estado do Paraná apresenta grande variabilidade de datas de semeadura de soja e milho (Paraná, 2011). Araújo et al. (2011) observaram que as datas de semeadura variam entre o primeiro decêndio de outubro (região Oeste do Paraná) e o terceiro decêndio de dezembro (região Centro-Sul do Paraná). Isso é importante para que se possa definir a ocorrência do pico vegetativo das culturas; informação utilizada no mapeamento.

O objetivo deste trabalho foi estimar e mapear as áreas com as culturas de soja e milho, no Paraná, com uso de imagens multitemporais EVI/Modis.

Material e Métodos

A área de estudo compreendeu o Estado do Paraná, na região Sul do Brasil (22°29' e 26°43'S; 48°2' e 54°38'W). Os municípios que apresentaram baixa produção de soja e milho (<6.800 Mg) e área colhida (<2.000 ha), segundo dados oficiais da Produção Agrícola Municipal do IBGE, não foram considerados (municípios em cinza) na espacialização dos dados de área colhida (Figura 1) e de produção (Figura 2). Dessa forma, foram considerados 322 municípios que representam 99% da produção e da área de soja e

milho do estado. O somatório de área colhida destes municípios foi denominado nível estadual.

Para o mapeamento e a estimativa de área colhida destas culturas de verão, foram obtidas imagens do sensor Modis, produto MOD13Q1 do “tile” h13v11, no formato “hierarchical data format” (HDF) (National Aeronautics and Space Administration, 2009). Este produto contém 12 tipos de imagens, entre elas, as de índice de vegetação EVI, “VI quality” (qualidade do índice de vegetação) e “pixel reliability” (confiabilidade do pixel), produzidos a partir da composição de 16 dias, com resolução espacial de 250 m. Estas imagens foram extraídas por meio da ferramenta “Modis reprojection tool” para o formato GeoTiff. As imagens “VI quality” e “pixel reliability” foram utilizadas para eliminar pixels com ruído, de acordo com proposta de Lu et al. (2007).

As 23 cenas anuais existentes (ou seja, composições de 16 dias) foram recortadas para o todo o Estado do Paraná (Figura 3). Para contemplar o ciclo de desenvolvimento destas culturas, utilizaram-se apenas as cenas 241 (29/8 do ano 1) a 113 (23/4 do ano 2);

portanto, foram utilizadas 15 imagens para cada um dos anos-safra (2004/2005 a 2007/2008), que foram agrupadas temporalmente (imagem composição), o que permitiu a geração do perfil espectral do EVI.

Em razão da grande variabilidade de datas de semeadura nas diferentes regiões do estado (Araújo et al., 2011), utilizaram-se diversas imagens de EVI para gerar a imagem de mínimo EVI e a imagem de máximo EVI. Este procedimento foi operacionalizado por um sistema de extração de dados de imagens em linguagem de programação “interactive data language” (IDL) (Esquerdo et al., 2011). Foram utilizadas as cenas de 257 (29/8/2006) a 321 (17/11/2007), para geração da imagem de mínimo EVI (menores valores de EVI das cenas), que contemplaram a fase de pré-plantio e desenvolvimento inicial das culturas; e de 321 (17/11/2006) a 81 (22/3/2007), para geração da imagem de máximo EVI (maiores valores de EVI dentre as cenas utilizadas), que consideraram o pico vegetativo (final de desenvolvimento, florescimento e formação de colheita) (Figura 3). Dessa forma, abrangeram-se os

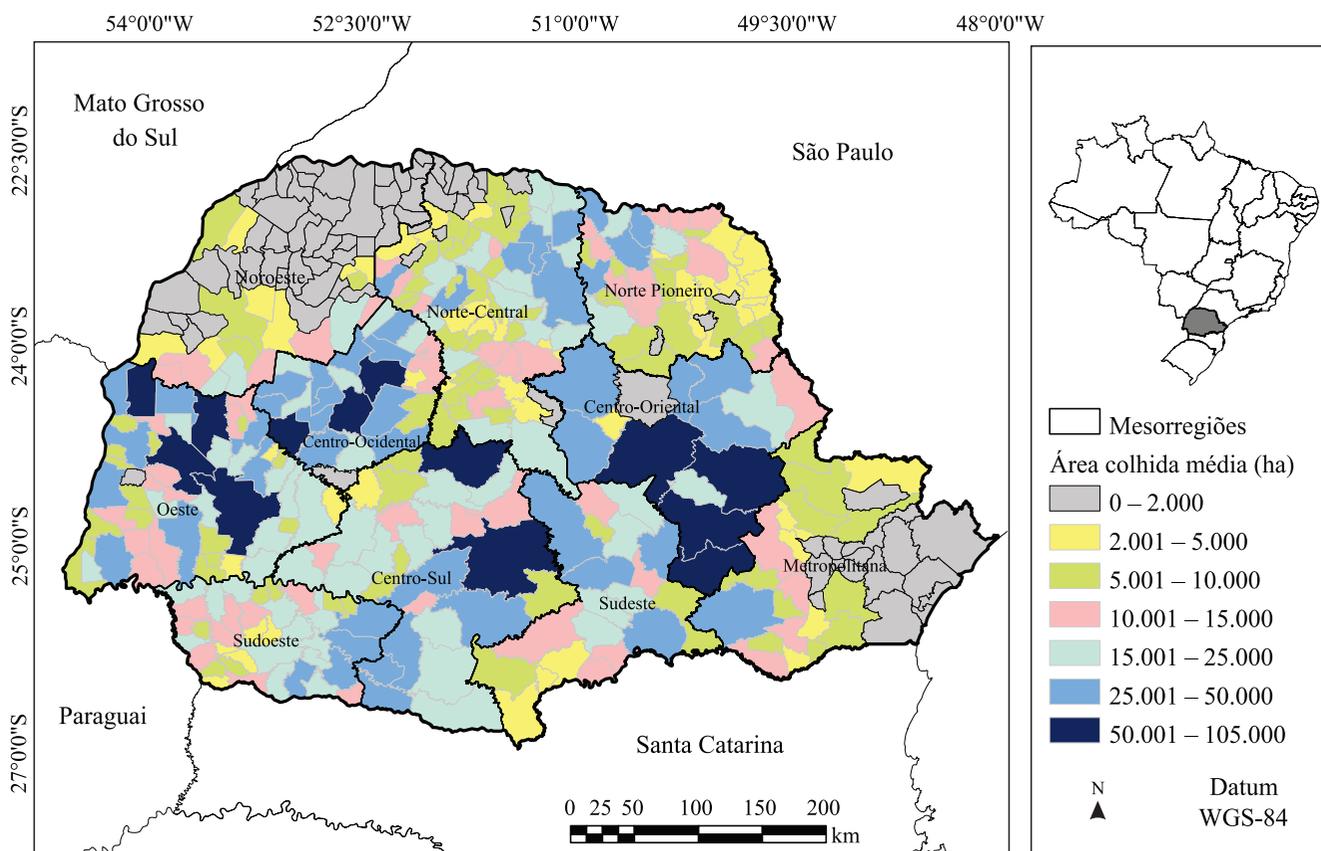


Figura 1. Mapa da área colhida média de 2005 a 2008, com as culturas de verão soja e milho, no Estado do Paraná. Fonte: Instituto Brasileiro de Geografia e Estatística (2011).

valores mínimos e máximos de EVI das culturas para cada uma das diferentes regiões do estado.

O programa ENVI (Exelis Visual Information Solutions, 2012) foi utilizado para gerar uma composição RGB, em que a imagem de máximo EVI foi colocada no filtro R (vermelho) e a imagem de mínimo EVI foi colocada nos filtros G (verde) e B (azul), tendo-se gerado uma composição colorida na qual as áreas em vermelho representaram as culturas de verão e as demais cores, os outros alvos (Figura 4). A composição colorida RGB, com valores originais de EVI, foi transformada para níveis de cinza (NC), que variaram de 0 a 255 no ENVI. Esta transformação teve como objetivo a normalização dos valores de acordo com a variação para mais ou para menos, em função das condições agrometeorológicas em cada ano-safra. Posteriormente, foram definidos os valores de corte. Este procedimento foi realizado para cada ano-safra, para obtenção do mapeamento das culturas de verão, denominado aqui de máscara.

A separação das áreas com culturas de verão (áreas vermelhas, Figura 4), expressa pela composição colorida RGB do EVI em NC, foi realizada com sistema de extração de dados, programada em linguagem IDL (Esquerdo et al., 2011). Basicamente, este sistema separou pixels da imagem a partir da definição de valores de corte, em NC, para o filtro R e para os filtros G e B, estabelecidos por simulação pelo usuário. Assim, para cada simulação, foram classificados os pixels da imagem que, simultaneamente, tiveram valores maiores que o corte definido no filtro R e menores que os cortes definidos nos filtros G e B, o que gerou as máscaras com as culturas de verão. O resultado de cada simulação foi confrontado, por meio de interpretação visual, com os mosaicos de imagens Landsat-5/TM e AWiFS/IRS (referência terrestre), que, nas composições coloridas de falsa cor RGB453, permitem identificar as áreas com as culturas de soja (cor amarela ou laranja) e milho (cor vermelha ou marrom).

Com as máscaras com as culturas de verão de cada ano-safra, procedeu-se à extração dos dados de

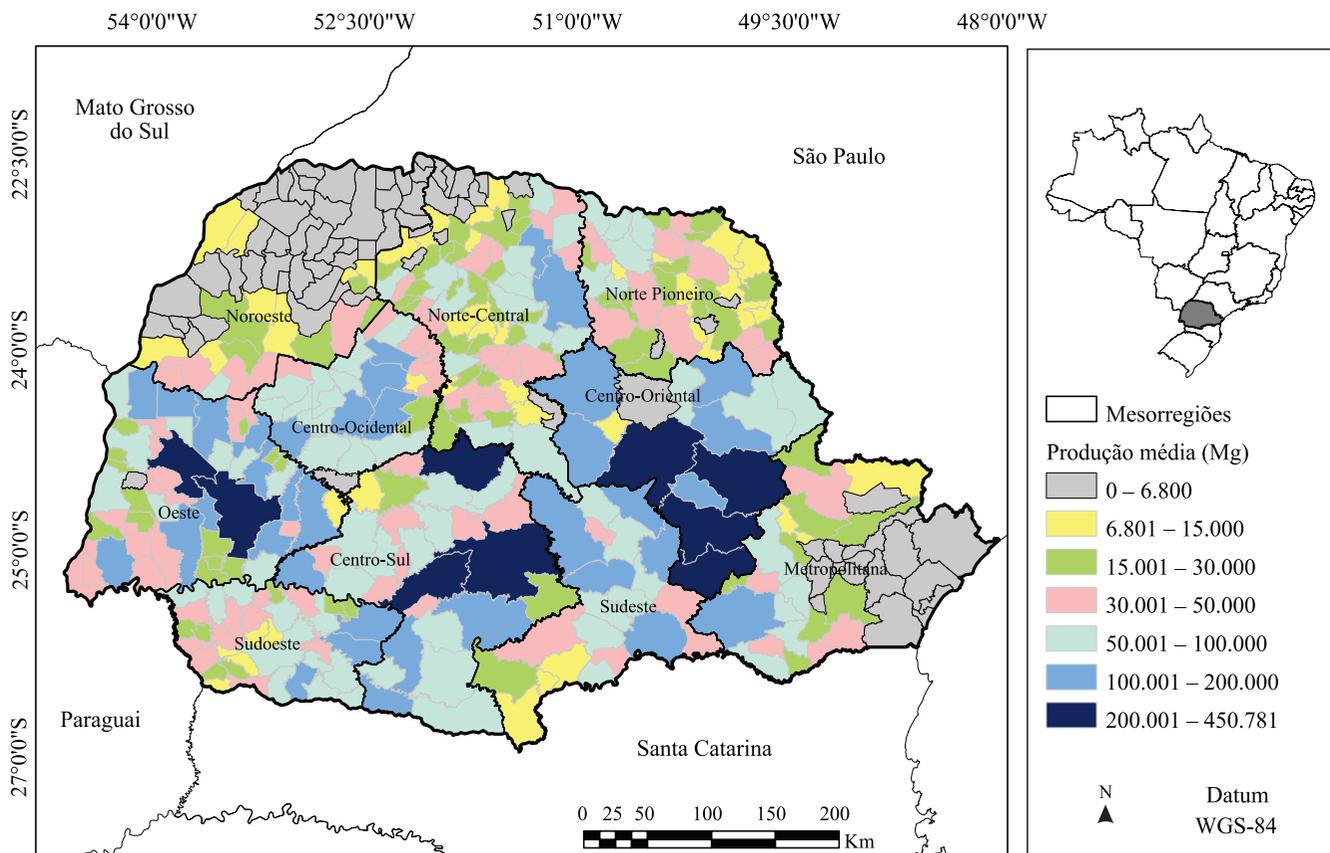


Figura 2. Mapa da produção média de 2005 a 2008, das culturas de verão soja e milho, no Estado do Paraná. Fonte: Instituto Brasileiro de Geografia e Estatística (2011).

área municipal de cada uma delas, tendo-se utilizado outro sistema de extração de dados desenvolvido em linguagem IDL. Portanto, para gerar a estimativa de área municipal, realizou-se a contagem da quantidade de pixels dentro de cada município, identificada na máscara como cultura de verão, multiplicada pela área de cada pixel (250x250 m = 62.250 m²).

As informações de área municipal, obtidas para cada ano-safra, foram comparadas com os dados municipais oficiais de área colhida – disponibilizados na produção agrícola municipal do IBGE (Instituto Brasileiro de Geografia e Estatística, 2011) –, por meio de modelo de regressão linear simples e de seu coeficiente de determinação (R²). Como indicadores estatísticos, foram utilizados o erro médio (ME), a raiz quadrada do erro médio quadrático (RMSE), o coeficiente de concordância (d) de Willmott et al. (1985) e o coeficiente de confiança (c) de Camargo & Sentelhas (1997), definidos, respectivamente, nas equações 1 a 4 mostradas adiante.

Para averiguação das exatidões espaciais (acurácia) das máscaras geradas, foram utilizados o índice de

exatidão global (EG), conforme equação 5; o índice de concordância Kappa (IK), definido na equação 6 (Congalton, 1991; Congalton & Green, 1999); e os erros de inclusão (EI) e omissão (EO), tendo-se tomado, como referência terrestre, imagens de média resolução espacial (Landsat-5/TM e AWiFS/IRS). Para obter erro amostral em torno de 5%, segundo Barbetta (2007), foram utilizadas cerca de 400 amostras (ou seja, 400 pontos), das quais metade foi sorteada aleatoriamente sobre o alvo de interesse (máscaras das culturas de verão), ou seja, houve amostragem estratificada aleatória. Assim, procurou-se evitar que os EO e os EI fossem mascarados na análise, por haver diferentes proporções entre os alvos avaliados.

Foram utilizadas as seguintes equações:

$$ME = \frac{1}{n} \times \sum_{i=1}^n (O - E) \tag{1}$$

$$RMSE = \left(\frac{1}{n} \times \sum_{i=1}^n (O - E)^2 \right)^{1/2} \tag{2}$$

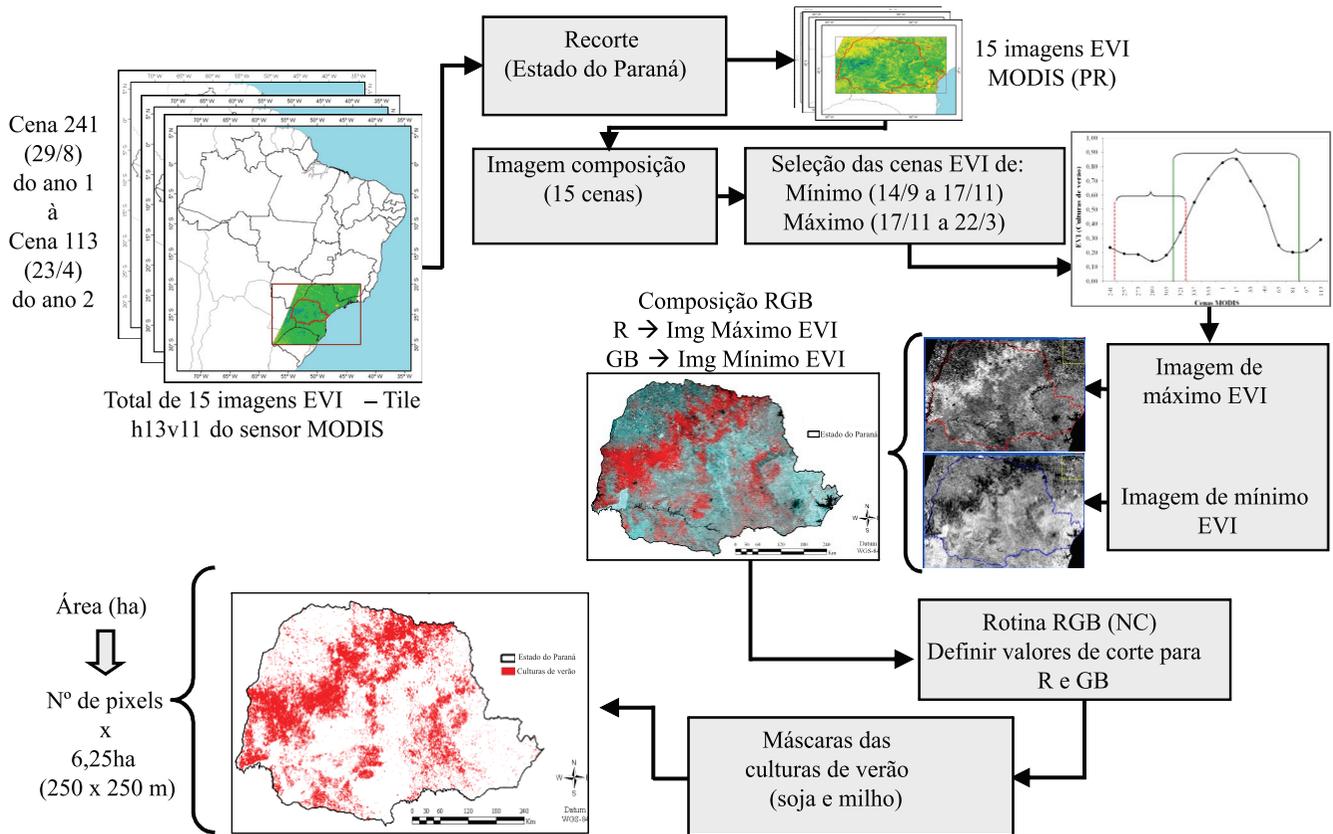


Figura 3. Fluxograma do procedimento de geração do mapeamento das culturas de verão soja e milho no Estado do Paraná.

$$d = 1 - \left[\frac{\sum_{i=1}^n (O - E)^2}{\sum_{i=1}^n (|E - \bar{O}| + |O - \bar{E}|)^2} \right] \quad (3)$$

$$c = r \times d \quad (4)$$

$$EG = (A/m)100 \quad (5)$$

$$IK = \left[n \times \sum_{i=1}^r x_{ij} - \sum_{i=1}^r (x_i \times x_j) \right] / \left[m^2 - \sum_{i=1}^r (x_i \times x_j) \right] \quad (6)$$

em que: n é o número de municípios; O é a área colhida (soja e milho) informada pelo IBGE; E é a área obtida pela máscara; \bar{O} é a área colhida média (soja e milho) dos municípios, informada pelo IBGE; A é o acerto geral (pontos amostrais com acertos); m é o número de pontos amostrais; r é o número de linhas da matriz de erro; x_{ij} é a observação na linha i e coluna j ; x_i é o total marginal da linha i ; e x_j é o total marginal da coluna j .

Resultados e Discussão

Na geração das máscaras para os anos-safra 2004/2005 a 2007/2008, entre as várias simulações realizadas de cortes de NC, para os canais da composição colorida RGB (Figura 4), o melhor resultado foi encontrado quando definiu-se 160 como valor de corte

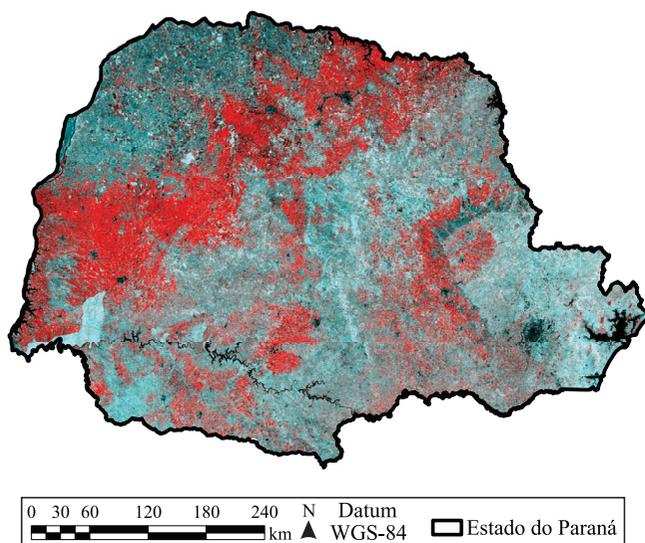


Figura 4. Composição RGB de imagens de máximo (R) e mínimo EVI/Modis (GB), para o Estado do Paraná, com base em dados da safra 2007/2008. As áreas vermelhas correspondem às culturas de verão.

para o filtro R (ou seja, valores de NC acima de 160), e 150 como valor de corte para os canais G e B (valores de NC abaixo de 150). Para as cenas que deram origem à imagem de máximo EVI (filtro R), os valores de NC deveriam estar acima de 160 e, para as cenas que deram origem à imagem de mínimo EVI (filtros G e B), os valores de NC deveriam estar abaixo de 150, simultaneamente.

Os resultados das distribuições espaciais pelo estado das áreas cultivadas com soja e milho evidenciaram a existência do chamado “cinturão da soja”, desde a região Oeste até a região Norte do Estado do Paraná. Além disso, observou-se concentração destas culturas nas regiões dos Campos Gerais (Centro-Oriental) e Centro-Sul do estado (Figura 5).

A maior diferença entre os dados do IBGE e os da máscara foi encontrada para o ano-safra 2004/2005 (610.875 ha), e a menor diferença foi observada para o ano-safra 2006/2007 (86.594 ha), o que representa, respectivamente, um montante de 11,48 e 1,65% da área total do estado com as culturas de soja e milho (Tabela 1). As áreas das máscaras municipais, em média, foram subestimadas em 1.897 ha (2004/2005), 1.659 ha (2005/2006), 269 ha (2006/2007) e 909 ha (2007/2008). Evidentemente, o ideal seria que estes erros percentuais fossem os menores possíveis. Porém, as grandes vantagens da metodologia proposta para estimativa das áreas das culturas incluem baixo custo (uso de imagens gratuitas), rapidez, objetividade e conhecimento da distribuição espacial pelo estado, o que não é possível quando se utiliza a metodologia dos órgãos oficiais.

Wardlow & Egbert (2008), ao comparar dados obtidos pelas máscaras de culturas de verão (geradas com imagens multitemporais de NDVI do Modis), no Estado de Kansas, EUA, com dados oficiais (USDA), verificaram subestimação de 4,4% para 2001. De acordo com estes autores, a resolução espacial do Modis (250 m) fez com que pixels de bordas de talhões não fossem selecionados pela máscara. Zhou et al. (2008), com o objetivo de separar áreas agrícolas das demais, na região de Yuli, China, utilizaram método de classificação supervisionada de imagens Landsat-5/TM, Landsat-7/ETM, CBERS 2 e Beijing 1, e encontraram diferenças de área, para os cinco anos avaliados, que variaram entre 3,2 e 11,7%, em comparação aos dados oficiais do país. Gusso et al. (2012) observaram diferenças máximas de 5,5% (anos-safra 2000/2001

a 2008/2009) entre as áreas estimadas de soja (223 municípios localizados na região norte do Rio Grande do Sul), com uso de imagens multitemporais de EVI e dados oficiais.

Os ajustes lineares realizados entre os dados do IBGE e os obtidos pelas máscaras (Figura 5 A, B, C e D) foram significativos a 1% de probabilidade. Os resultados apresentaram altos valores de R^2 – 0,841 (2004/2005) a 0,949 (2005/2006) –, o que indicou proximidade entre a área obtida pela máscara Modis e os dados oficiais (Figura 6). Gusso et al. (2012) encontraram valores de R^2 que variaram de 0,91 a 0,94, para ajustes lineares entre os dados oficiais e os obtidos pelas máscaras de soja (entre os anos-safra 2000/2001 e 2006/2007).

O coeficiente de correlação (r) indicou alta relação linear entre os dados ($r > 0,917$) (Tabela 1). Wardlow & Egbert (2008) encontraram $r = 0,76$, ao considerar todas as culturas de verão (milho, sorgo e soja). Entretanto, ao analisar cada uma das culturas separadamente, estes autores obtiveram $r = 0,92$ (milho), $r = 0,86$ (sorgo) e $r = 0,42$ (soja).

O índice de concordância (d), que mediu a exatidão dos valores estimados (máscaras das culturas de verão), em comparação aos dados oficiais (IBGE), mostrou que, para os anos-safra estudados, os valores apresentaram alta exatidão entre os dados (Tabela 1). Quanto ao coeficiente de confiança (c), os valores mostraram ótimo desempenho, conforme classificação proposta por Camargo & Sentelhas (1997).

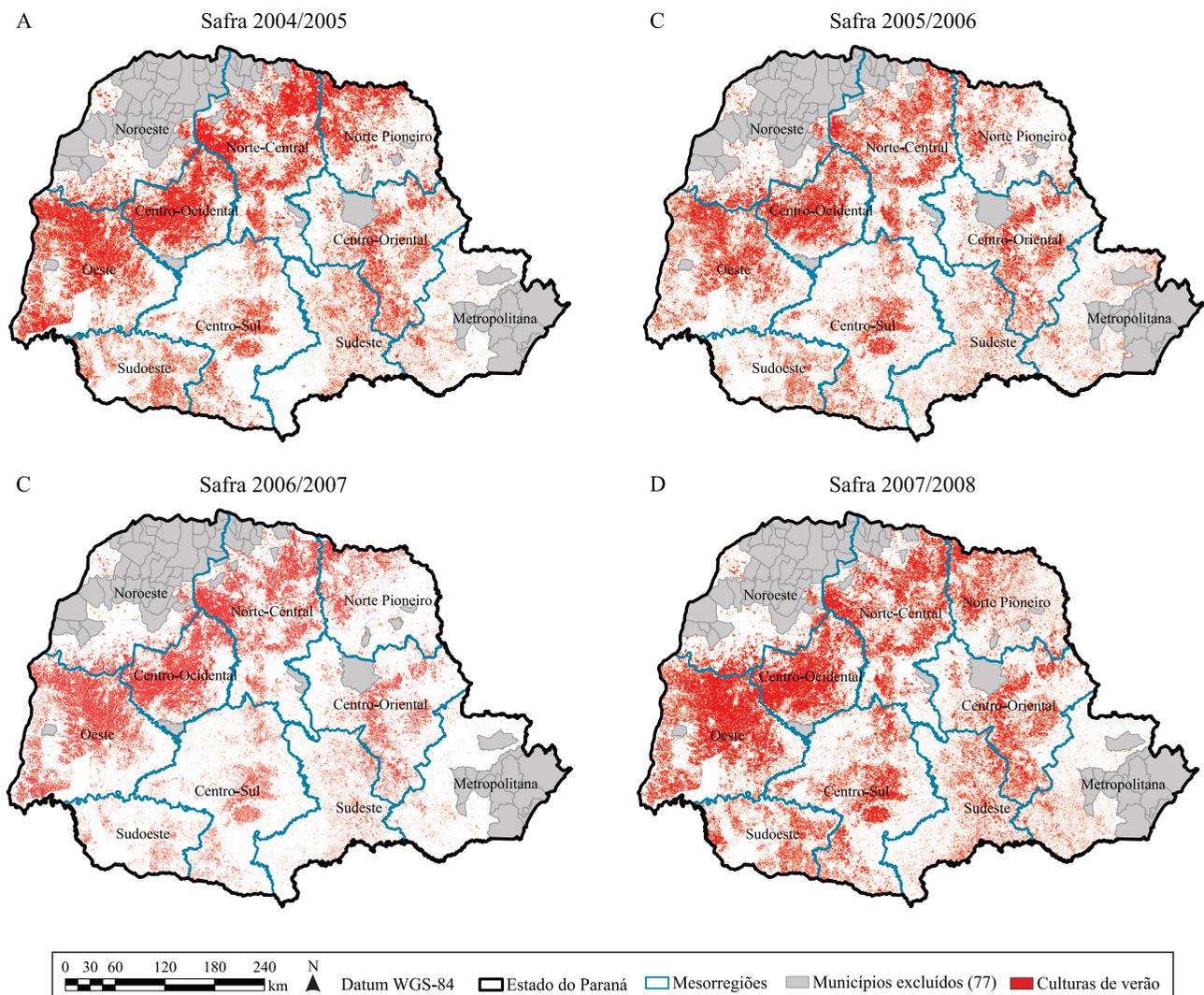


Figura 5. Máscara Modis para as culturas de verão soja e milho no Paraná: A, safra 2004/2005; B, safra 2005/2006; C, safra 2006/2007; e D, safra 2007/2008.

Pelo erro médio (ME), constatou-se que, em escala estadual, a área obtida pela máscara foi subestimada entre 269 ha (2006/2007) e 1.897 ha (2004/2005). O RMSE, que informa sobre o valor real do erro produzido pelo modelo, mostrou que, em média, houve variação entre 3.816 ha (2005/2006) e 6.573 ha (2004/2005) (Tabela 1). Gusso et al. (2012) encontraram valores de RMSE entre 3.328 ha (2000/2001) e 4.715 ha (2005/2006) nos modelos de regressão ajustados.

A análise do R^2 por mesorregião do Estado do Paraná mostrou que, em média, houve variação entre 0,717 (Sudoeste) e 0,954 (Centro-Occidental) para os quatro anos-safra (Tabela 2). Os piores ajustes (máscaras vs. IBGE) foram encontrados para a mesorregião Sudoeste, com R^2 mínimo = 0,579 (ano-safra 2006/2007) e R^2 máximo = 0,773 (ano-safra 2007/2008), o que pode ser explicado pela topografia acidentada da região e pelos pequenos tamanhos das propriedades na região, o que fez com que a resolução de 250 m do Modis não elegeisse todas as propriedades na região. Além disso, em 80% das mesorregiões, os maiores valores de R^2 máximo foram obtidos no ano-safra 2004/2005 e, no restante das mesorregiões (Centro-Sul e Sudoeste), no ano-safra 2007/2008 (Tabela 2). Já os valores de R^2 mínimo, foram encontrados em 50% das mesorregiões, para o ano-safra 2006/2007 (Centro-Sul, Curitiba, Noroeste, Norte Pioneiro e Sudoeste), em 30%, para 2004/2005 (Centro-Oriental, Norte Central e Oeste), e

em 20% para o ano-safra 2007/2008 (Centro-Oriental e Sudeste).

A soma da diagonal principal das matrizes de erros das 400 amostras utilizadas mostrou que 371 (2004/2005), 380 (2005/2006), 377 (2006/2007) e 379 (2007/2008) das amostras foram corretamente classificadas (Tabela 3), o que fez com que a exatidão global (EG) variasse entre 92,75% (2004/2005) e 95% (2005/2006). Zhou et al. (2008), ao utilizar imagens de média resolução espacial – Landsat-5/TM (30 m), Landsat-7/ETM (28,5 m), CBERS 2 (19,5 m) e Beijing 1 (32 m) –, encontraram EG que variou entre 88,90 e 95,20%. Gusso et al. (2012) compararam a máscara de soja obtida com EVI/Modis, do ano-safra 2000/2001, com o mapeamento realizado por Rizzi & Rudorff (2005) e encontraram EG de 82%. Isso indica que, no presente trabalho, as máscaras apresentaram ótima confiabilidade no mapeamento realizado pela metodologia proposta.

A acurácia também deve ser analisada, individualmente, por meio da análise dos erros de inclusão (EI) e dos erros de omissão (EO) presentes na classificação (Congalton & Green, 1999). Observou-se que as amostras sorteadas sobre a máscara das “culturas de verão” tiveram variação de 3,5% (ano-safra 2005/2006) a 7,5% (ano-safra 2004/2005) de EI (Tabela 3), ou seja, a máscara classificou erroneamente estes locais como sendo culturas de verão, mas, na realidade, eram outros alvos (mata, na grande maioria). Nas amostras sorteadas sobre as áreas “não culturas de verão”, houve variação entre 4,5% (ano-safra 2007/2008) e 7% (anos-safra 2004/2005 e 2006/2007) de EI, que representaram os valores de culturas de verão não selecionadas pela máscara. Verificou-se que a proporção de EO variou entre 3,6% (ano-safra 2005/2006) e 7,5% (ano-safra 2004/2005), ou seja, esta foi a proporção de amostras que foram excluídas da classe a que pertenciam (“culturas de verão” ou “não culturas de verão”). Portanto, o EI ocorre quando um objeto é incluído na classe a qual ele não pertence, e o EO quando um objeto é excluído da classe a qual pertence.

O índice Kappa (IK) apresenta vantagens sobre a EG, pois incorpora todos os elementos da matriz de erro (amostras classificadas de forma correta ou incorreta). Além disso, o IK avalia a acurácia temática, é mais sensível às variações de EI e EO, e avalia a coincidência espacial entre duas situações

Tabela 1. Área colhida oficial (Instituto Brasileiro de Geografia e Estatística, IBGE) e áreas obtidas pelas máscaras (ha) de soja e milho, e estatísticas obtidas na análise dos dados para os 322 municípios do Estado do Paraná, nos anos-safra de 2004/2005 a 2007/2008.

Variáveis ⁽¹⁾	Anos-safra			
	2004/2005	2005/2006	2006/2007	2007/2008
Área IBGE (ha)	5.320.555	5.298.318	5.247.792	5.291.236
Área máscara (ha)	4.709.680	4.764.205	5.161.198	4.998.479
Diferença (IBGE - Máscara)	610.875	534.113	86.594	292.757
Diferença (%)	11,48	10,08	1,65	5,53
Diferença média (ha)	1.897	1.659	269	909
r	0,9171	0,9742	0,9391	0,9407
ME	1.897	1.659	269	909
RMSE	6.573	3.816	6.423	6.013
d	0,9532	0,9843	0,9616	0,9659
c	0,8742	0,9589	0,9031	0,9086

⁽¹⁾r, coeficiente de correlação; ME, erro médio; RMSE, raiz quadrada do erro médio quadrático; d, coeficiente de concordância de Willmott et al. (1985); c, coeficiente de confiança de Camargo & Sentelhas (1997).

(Moreira, 2005). O IK, que avaliou a concordância ou discordância entre as duas situações avaliadas, variou entre 0,855 (2004/2005) e 0,900 (2005/2006) (Tabela 3), o que, segundo classificação proposta por Landis & Koch (1977), é de excelente qualidade ($IK > 0,81$). Assim, pode-se dizer que as máscaras geradas para as culturas de verão (soja e milho) aproximaram-se da realidade de campo.

Os valores foram ainda mais significativos quando comparados com trabalhos que utilizaram o mesmo sensor, porém com outra metodologia. Por exemplo, Lamparelli et al. (2008) encontraram valores de IK entre 0,60 e 0,80, ao mapear a cultura da soja no Paraná, com dados dos satélites Modis/Terra e Landsat-5/TM. Epiphany et al. (2010) estimaram a área plantada de soja no Estado do Mato Grosso, com imagens

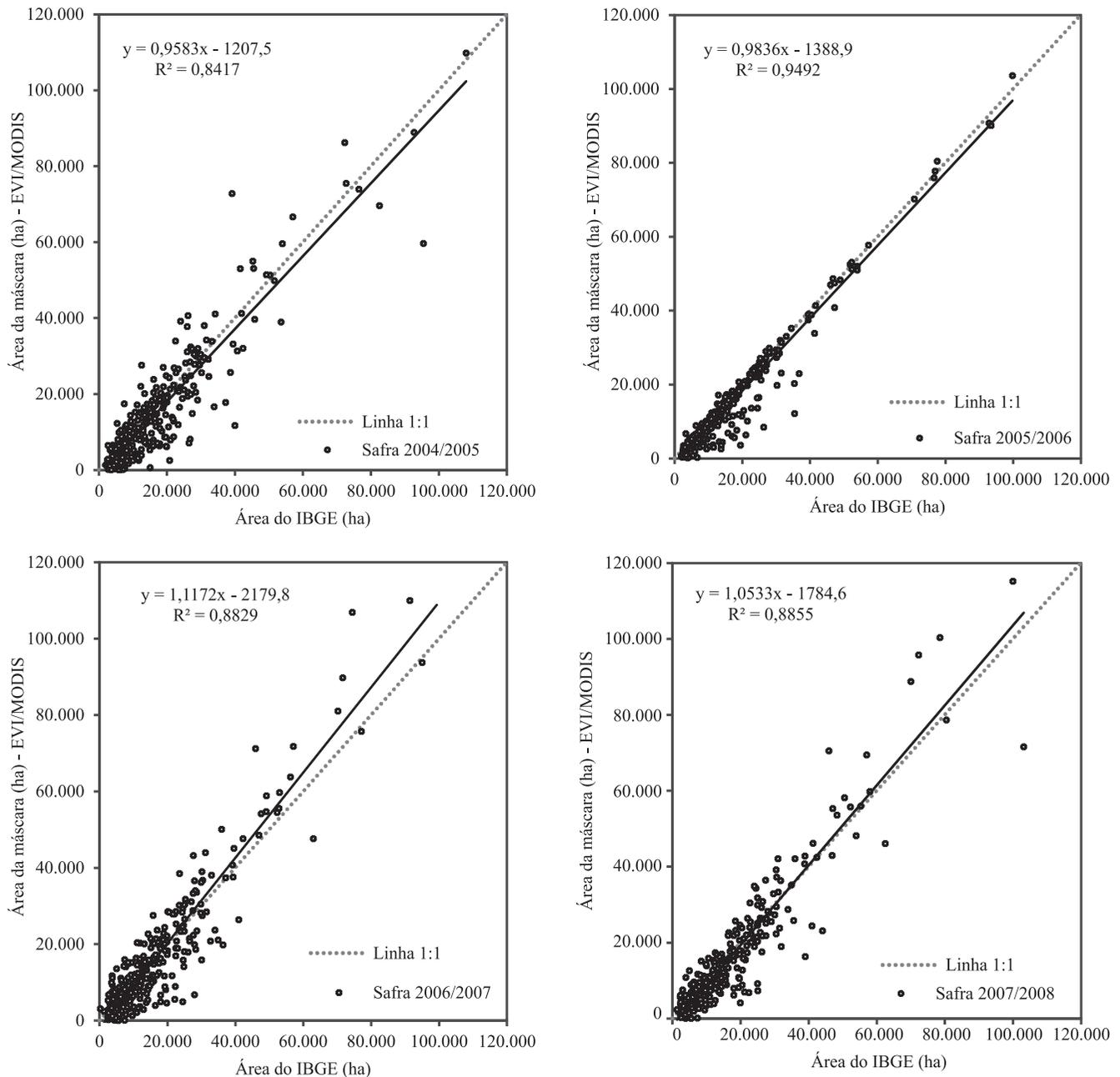


Figura 6. Relação entre a área colhida (ha) de soja e milho obtida com máscara Modis e os dados oficiais do IBGE nos 322 municípios avaliados nas safras: 2004/2005, 2005/2006, 2006/2007 e 2007/2008.

espectro-temporais do sensor Modis para a safra 2005/2006, e compararam os resultados com imagens Landsat-5/TM, tendo obtido EG de 80% e IK de 0,51. Wardlow & Egbert (2008) encontraram EG de 84% e IK de 0,76, para as culturas de verão, ao comparar suas

Tabela 2. Coeficiente de determinação (R^2) segundo mesorregiões do Estado do Paraná, nos anos-safra de 2004/2005 a 2007/2008.

Mesorregiões	Nº de municípios	Coeficiente de determinação (R^2)		
		Médio	Mínimo	Máximo
Centro-Occidental	24	0,954	0,924 ⁽⁴⁾	0,994 ⁽²⁾
Centro-Oriental	13	0,878	0,800 ⁽¹⁾	0,955 ⁽²⁾
Centro-Sul	29	0,929	0,899 ⁽³⁾	0,951 ⁽⁴⁾
Metropolitana	21	0,887	0,790 ⁽³⁾	0,947 ⁽²⁾
Noroeste	19	0,910	0,820 ⁽³⁾	0,985 ⁽²⁾
Norte Central	66	0,862	0,794 ⁽¹⁾	0,952 ⁽²⁾
Norte Pioneiro	43	0,908	0,853 ⁽³⁾	0,956 ⁽²⁾
Oeste	49	0,951	0,935 ⁽¹⁾	0,987 ⁽²⁾
Sudeste	21	0,811	0,644 ⁽⁴⁾	0,952 ⁽²⁾
Sudoeste	37	0,717	0,579 ⁽³⁾	0,773 ⁽⁴⁾

⁽¹⁾ano-safra 2004/2005; ⁽²⁾ano-safra 2005/2006; ⁽³⁾ano-safra 2006/2007; ⁽⁴⁾ano-safra 2007/2008.

Tabela 3. Matriz de confusão entre as máscaras de soja e milho obtidas com o Modis e a referência terrestre (imagens Landsat-5/TM e AWiFS/IRS), para os anos-safra de 2004/2005 a 2007/2008⁽¹⁾.

Máscara	Referência terrestre		Total geral	Erro de inclusão (%)
	Culturas de verão	Não culturas de verão		
Ano-safra 2004/2005				
Culturas de verão	185	15	200	7,5
Não culturas de verão	14	186	200	7,0
Total geral	199	201	400	
Erro de omissão (%)	7,0	7,5		93 ⁽¹⁾ ; 0,86 ⁽²⁾
Ano-safra 2005/2006				
Culturas de verão	193	7	200	3,5
Não culturas de verão	13	187	200	6,5
Total geral	206	194	400	
Erro de omissão (%)	6,3	3,6		95 ⁽¹⁾ ; 0,90 ⁽²⁾
Ano-safra 2006/2007				
Culturas de verão	191	9	200	4,5
Não culturas de verão	14	186	200	7,0
Total geral	205	195	400	
Erro de Omissão (%)	6,8	4,6		94 ⁽¹⁾ ; 0,88 ⁽²⁾
Ano-safra 2007/2008				
Culturas de verão	188	12	200	6,0
Não culturas de verão	9	191	200	4,5
Total geral	197	203	400	
Erro de Omissão (%)	4,6	5,9		95 ⁽¹⁾ ; 0,90 ⁽²⁾

⁽¹⁾Exatidão global (%). ⁽²⁾Índice Kappa.

máscaras com mapas de uso do solo gerados a partir de imagens Landsat.

Zhou et al. (2008) obtiveram IK que variou entre 0,762 (CBERS 2) e 0,896 (Landsat-7/ETM). Mello et al. (2010), ao utilizar recorte de cena Landsat-5/TM, com área de 225 km², na Região de Ribeirão Preto, SP, testaram método de classificação automática por meio da técnica de modelo linear de mistura espectral, para identificar áreas de cana-de-açúcar colhidas com e sem queima. Ao comparar o resultado obtido com o mapa gerado por interpretação visual (mapa referência), estes autores encontraram EG de 89,7% e IK de 0,845. Ou seja, ao se utilizar uma pequena área de estudo e um satélite com menor resolução espacial, estes resultados apresentaram menor acurácia que os encontrados no presente trabalho.

A metodologia aqui avaliada apresenta alto potencial de utilização para mapeamento de culturas de verão (soja e milho) e pode ser vista como uma importante ferramenta de complementação do levantamento realizado pelos órgãos oficiais responsáveis por estatísticas de produção agrícola, com vantagens de rapidez e objetividade, além da espacialização dessas áreas por município.

Conclusões

1. A utilização de dados espectrais multitemporais EVI/Modis permite o mapeamento e a estimativa das áreas cultivadas com as principais culturas agrícolas de verão (soja e milho), em escala municipal, no Estado do Paraná.

2. O uso de várias cenas para gerar as composições de mínimos e máximos EVI faz com que a grande variabilidade nas datas de semeadura de soja e milho não representem problemas para a estimativa de suas áreas de cultivo, e permite o mapeamento, de uma só vez, de todas as áreas com estas culturas no Estado do Paraná.

3. O procedimento de normalização dos dados de EVI permite o estabelecimento de valores de corte para os filtros de R (vermelho) e de G e B (verde e azul), para os quatro anos-safras, e possibilita a reprodução desta metodologia para outros anos safras.

4. Na comparação com os dados oficiais do Instituto Brasileiro de Geografia e Estatística, as máscaras de culturas de verão obtidas apresentam ótimos ajustes e excelente acurácia espacial, apesar de subestimarem a área cultivada com soja e milho.

Agradecimentos

À Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas pela oportunidade ao primeiro autor de realização do doutorado.

Referências

- ADAMI, M.; MOREIRA, M.A.; RUDORFF, B.F.T.; FREITAS, C. da C. Análise da eficiência dos estimadores de expansão direta e de regressão para áreas cultivadas com café, milho e soja no Município de Cornélio Procopio, Estado do Paraná. **Agricultura em São Paulo**, v.51, p.5-13, 2004.
- 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, p.585-592, 2010.
- ARAÚJO, G.K.D.; ROCHA, J.V.; LAMPARELLI, R.A.C.; ROCHA, A.M. Mapping of summer crops in the State of Paraná, Brazil, through the 10-day spot vegetation NDVI composites. **Engenharia Agrícola**, v.31, p.760-770, 2011.
- BARBETTA, P.A. **Estatística aplicada às ciências sociais**. Florianópolis: UFSC, 2007. 316p.
- CAMARGO, A.P. de; SENTELHAS, P.C. Avaliação de desempenho de diferentes métodos de estimativa da evapotranspiração potencial no Estado de São Paulo, Brasil. **Revista Brasileira de Agrometeorologia**, v.5, p.89-97, 1997.
- COMPANHIA NACIONAL DE ABASTECIMENTO. **GeoSafras**. 2009a. Disponível em: <<http://www.conab.gov.br/conabweb/index.php?PAG=81>>. Acesso em: 18 jun. 2009.
- COMPANHIA NACIONAL DE ABASTECIMENTO. **Levantamentos de safra**. 2009b. 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 of classifications of remotely sensed data. **Remote Sensing of Environment**, v.37, p.35-46, 1991.
- CONGALTON, R.G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices**. Boca Raton: CRC Press, 1999. 160p.
- EPIPHANIO, R.D.V.; FORMAGGIO, A.R.; RUDORFF, B.F.T.; MAEDA, E.E.; LUIZ, A.J.B. Estimating soybean crop areas using spectral-temporal surfaces derived from MODIS images in Mato Grosso, Brazil. **Pesquisa Agropecuária Brasileira**, v.45, p.72-80, 2010.
- ESQUERDO, J.C.D.M.; ZULLO JUNIOR, J.; ANTUNES, J.F.G. Use of NDVI/AVHRR time series profiles for soybean crop monitoring in Brazil. **International Journal of Remote Sensing**, v.32, p.3711-3727, 2011.
- EXELIS VISUAL INFORMATION SOLUTIONS. **ENVI**. Boulder: EXELIS, 2012. Available at: <<http://www.envi.com.br>>. Accessed on: 1 Feb. 2012.
- FERREIRA, L.G.; FERREIRA, N.C.; FERREIRA, M.E. Sensoriamento remoto da vegetação: evolução e estado-da-arte. **Acta Scientiarum. Biological Sciences**, v.30, p.379-390, 2008.
- GAMON, J.A.; FIELD, C.B.; GOULDEN, M.L.; GRIFFIN, K.L.; HARTLEY, A.E.; JOEL, G.; PEÑUELAS, J.; VALENTINI, R. Relationships between NDVI, canopy structure, and photosynthesis in three Californian vegetation types. **Ecological Applications**, v.5, p.28-41, 1995.
- GUSSO, A.; FORMAGGIO, A.R.; RIZZI, R.; ADAMI, M.; RUDORFF, B.F.T. Soybean crop area estimation by Modis/Evi data. **Pesquisa Agropecuária Brasileira**, v.47, p.425-435, 2012.
- HOLBEN, B.N. Characteristics of maximum value composite images from temporal AVHRR data. **International Journal of Remote Sensing**, v.7, p.1417-1435, 1986.
- HUETE, A.D.L.; DIDAN, K.; MIURA, T.; RODRIGUEZ, E.P.; GAO, X.; FERREIRA, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote Sensing of Environment**, v.83, p.195-213, 2002.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Banco de Dados Agregados: Sistema IBGE de Recuperação Automática – SIDRA**. 2011. Disponível em: <<http://www.sidra.ibge.gov.br>>. Acesso em: 05 abr. 2011.
- JENSEN, J.R.; BOTCHWAY, K.; BRENNAM-GALVIN, E.; JOHANNSEN, C.J.; JUMA, C.; MABOGUNJE, A.; MILLER, R.; PRICE, K.; REINING, P.; SKOLE, D.; STANCIOFF, A.; TAYLOR, D.R.F. **Down to Earth: geographic information for sustainable development in Africa**. Washington: National Academy Press, 2002. 155p.
- LAMPARELLI, R.A.C.; CARVALHO, W.M.O. de; MERCANTE, E. Mapeamento de sementeiras de soja (*Glycine max* (L.) Merr.) mediante dados MODIS/Terra E TM/Landsat 5: um comparativo. **Revista Engenharia Agrícola**, v.28, p.334-344, 2008.
- LANDIS, J.R.; KOCH, G.G. The measurement of observer agreement for categorical data. **Biometrics**, v.33, p.159-174, 1977.
- LU, X.; LIU, R.; LIU, J.; LIANG, S. Removal of noise by wavelet method to generate high quality temporal data of terrestrial MODIS products. **Photogrammetric Engineering and Remote Sensing**, v.73, p.1129-1139, 2007.
- MELLO, M.P.; RUDORFF, B.F.T.; VIEIRA, C.A.O.; AGUIAR, D.A. de. Classificação automática da colheita da cana-de-açúcar utilizando modelo linear de mistura espectral. **Revista Brasileira de Cartografia**, v.2, p.181-188, 2010.
- MERCANTE, E.; LAMPARELLI, R.A.C.; URIBE-OPAZO, M.A.; ROCHA, J.V. Modelos de regressão lineares para estimativa de produtividade da soja no Oeste do Paraná, utilizando dados espectrais. **Engenharia Agrícola**, v.30, p.504-517, 2010.
- MOREIRA, M.A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. Viçosa: UFV, 2005. 320p.
- NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. **Technical specifications: Moderate Resolution Imaging Spectroradiometer (MODIS)**. 2009. Available at: <<http://modis.gsfc.nasa.gov/about/design.php>>. Accessed on: 6 June 2009.
- PARANÁ. Secretaria de Estado da Agricultura e do Abastecimento. Departamento de Economia Rural. **Banco de dados da produção agropecuária no Paraná: situação mensal de plantio, colheita e comercialização de produtos agrícolas no Paraná**. 2011. Disponível em: <<http://www.agricultura.pr.gov.br/arquivos/File/deral/pss.xls>>. Acesso em: 6 dez. 2011.

- REN, J.; CHEN, Z.; ZHOU, Q.; TANG, H. Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. **International Journal of Applied Earth Observation and Geoinformation**, v.10, p.403-413, 2008.
- RIZZI, R.; RUDORFF, B.F.T. Estimativa da área de soja no Rio Grande do Sul por meio de imagens Landsat. **Revista Brasileira de Cartografia**, v.57, p.226-234, 2005.
- ROJAS, O.; REMBOLD, F.; ROYER, A.; NEGRE, T. Real-time agrometeorological crop yield monitoring in Eastern Africa. **Agronomy for Sustainable Development**, v.25, p.63-77, 2005.
- ROUSE, J.W.; HAAS, R.H.; SCHELL, J.A.; DEERING, D.W. Monitoring vegetation systems in the Great Plains with ERTS. In: EARTH RESOURCES TECHNOLOGY SATELLITE SYMPOSIUM, 1., 1973, Washington. **Proceedings**. Washington: NASA, 1973. v.1, p.309-317.
- SUGAWARA, L.M.; RUDORFF, B.F.T.; ADAMI, M. Viabilidade de uso de imagens do Landsat em mapeamento de área cultivada com soja no Estado do Paraná. **Pesquisa Agropecuária Brasileira**, v.43, p.1777-1783, 2008.
- 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, p.1096-1116, 2008.
- WILLMOTT, C.J.; ACKLESON, S.G.; DAVIS, R.E. Statistics for the evaluation and comparison of models. **Journal of Geophysical Research**, v.90, p.8995-9005, 1985.
- 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, p.753-763, 2007.
- ZHOU, O.; LI, B.; SUN, B. Modelling spatio-temporal pattern of landuse change using multitemporal remotely sensed imagery. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**. v.37, p.729-734, 2008.

Recebido em 27 de julho de 2011 e aprovado em 6 de agosto de 2012