

Estimativa de área de soja por classificação de imagens normalizada pela matriz de erros

João Francisco Gonçalves Antunes⁽¹⁾, Erivelto Mercante⁽²⁾, Júlio César Dalla Mora Esquerdo⁽¹⁾, Rubens Augusto de Camargo Lamparelli⁽³⁾ e Jansle Vieira Rocha⁽⁴⁾

⁽¹⁾Embrapa Informática Agropecuária, Caixa Postal 6041, CEP 13083-886 Campinas, SP. E-mail: joao.antunes@embrapa.br, julio.esquerdo@embrapa.br ⁽²⁾Universidade Estadual do Oeste do Paraná, Caixa Postal 701, CEP 85819-110 Cascavel, PR. E-mail: erivelto.mercante@unioeste.br ⁽³⁾Universidade Estadual de Campinas (Unicamp), Núcleo Interdisciplinar de Planejamento Energético, Campus Unicamp, Caixa Postal 6166, CEP 13083-860 Campinas, SP. E-mail: rubens@nipeunicamp.org.br ⁽⁴⁾Unicamp, Faculdade de Engenharia Agrícola, Avenida Cândido Rondon, nº 501, Cidade Universitária Zeferino Vaz, CEP 13083-875 Campinas, SP. E-mail: jansle.rocha@feagri.unicamp.br

Resumo – O objetivo deste trabalho foi estimar a área plantada com soja por meio da normalização da matriz de erros gerada a partir da classificação supervisionada de imagens TM/Landsat-5. Foram avaliados oito municípios no Estado do Paraná, com dados referentes à safra de 2003/2004. As classificações foram realizadas por meio dos métodos paralelepípedo e máxima verossimilhança, dando origem à “máscara de soja”. Os valores do índice Kappa dos oito municípios ficaram acima de 0,6. As estimativas de área de soja, corrigidas por matriz de erros, apresentaram alta correlação com as estimativas oficiais do estado e com as estimativas geradas a partir de um método alternativo denominado “expansão direta”. A estimativa de área de soja por meio da normalização da matriz de erros apresenta menor custo e pode subsidiar métodos convencionais na estimativa menos subjetiva de safras.

Termos para indexação: *Glycine max*, cultura da soja, geotecnologia, índice Kappa, previsão de safras, TM/Landsat-5.

Soybean crop area estimation through image classification normalized by the error matrix

Abstract – The objective of this work was to estimate soybean crop area by the normalization of the error matrix generated from the supervised classification of TM/Landsat-5 images. Eight municipalities of the state of Paraná, Brazil, were evaluated using data from the 2003/2004 crop season. Classifications were carried out using the parallelepiped and maximum likelihood methods, resulting in a “soybean mask”. Kappa index values for the eight municipalities were above 0.6. Estimated soybean areas, corrected by the error matrix, were highly correlated with official estimates of the state and with estimates generated from an alternative method called “direct expansion”. Soybean crop area estimation by the normalization of the error matrix is less costly and can aid conventional methods in estimating harvests in a less subjective manner.

Index terms: *Glycine max*, soybean crop, geotechnology, Kappa index, crop forecasting, TM/Landsat-5.

Introdução

A soja é uma das principais culturas do agronegócio brasileiro, com grande importância econômica nas exportações de grãos e seus derivados. O Paraná é o segundo maior produtor do país (Companhia Nacional de Abastecimento, 2012), e a região oeste do estado responde por mais de 23% de sua produção (Paraná, 2012).

A estimativa da área plantada é uma das principais informações empregadas na previsão de safras agrícolas. Em nível federal, ela é fornecida por órgãos oficiais do Governo, como a Companhia Nacional de

Abastecimento (Conab), vinculada ao Ministério da Agricultura, Pecuária e Abastecimento, e o Instituto Brasileiro de Geografia e Estatística (IBGE). Na esfera estadual, as estimativas são obtidas e publicadas por órgãos ligados às Secretarias de Estado, como é o caso do Departamento de Economia Rural (Deral) da Secretaria de Estado da Agricultura e do Abastecimento do Paraná (Seab).

Nos últimos anos, a previsão de safras no Brasil tem incorporado geotecnologias para diminuir a subjetividade, melhorar a precisão das estimativas e reduzir custos (Companhia Nacional de Abastecimento, 2011). O sensoriamento remoto e os métodos

tradicionais de classificação digital de imagens e avaliação de sua acurácia já são utilizados em alguns países (Gallego, 2004; Gallego et al., 2008; Wu et al., 2012).

A matriz de erros, bem como as métricas de exatidão global e o índice Kappa, vem sendo utilizada para determinar a acurácia de classificações digitais mediante a utilização de imagens de satélites (Jupp, 1989; Congalton & Green, 1999; De Wit & Clevers, 2004; Serra & Pons, 2008; Peña-Barragán et al., 2011). Essa matriz permite avaliar o desempenho da classificação realizada para uma classe individual, particularmente quando um pequeno número de classes de uso do solo é de interesse, como, por exemplo, na estimativa de área de uma cultura agrícola (Ceballos-Silva & López-Blanco, 2003). Chen & Goodchild (2007) também desenvolveram um método de calibração de estimativa de área de culturas agrícolas ao utilizar proporções da matriz de erros. De acordo com estes autores, as classes mais homogêneas podem se beneficiar desse ajuste e proporcionar melhores estimativas de áreas do que classes heterogêneas.

A estimativa de área por meio de levantamentos de campo é baseada em amostras, em decorrência da limitação de tempo e do custo envolvido. Nas classificações digitais, as áreas são estimadas por contagem de pixels de toda a região em estudo, sem a necessidade de amostragem. No entanto, erros na classificação da imagem podem implicar em imprecisão nas estimativas. Para corrigir a estimativa de área, Conese & Maselli (1992) transformaram matrizes de erros em matrizes de proporções, com o intuito de corrigir as probabilidades de designação das classes simultaneamente.

Ippoliti-Ramilo et al. (2003) fizeram análise das estimativas de área de culturas de verão com auxílio de dados de sensoriamento remoto, em levantamentos regionais do Estado de São Paulo. Estes autores avaliaram a eficiência relativa do método, em comparação a dados oficiais, e observaram que bons níveis de acurácia são alcançados, desde que as estimativas sejam corrigidas por matrizes de erros. Shao et al. (2003) também analisaram a influência da utilização de dados marginais da matriz de erros no produto final de classificação de imagens orbitais e concluíram que a estimativa de área poderia tornar-se mais confiável se fossem utilizadas informações da

matriz de erros localizadas fora da diagonal principal, para corrigir seu valor.

O objetivo deste trabalho foi estimar a área plantada com soja [*Glycine max* (L.) Merr.] por meio da normalização da matriz de erros gerada a partir da classificação supervisionada de imagens TM/Landsat-5.

Material e Métodos

A área de estudo foi compreendida por oito municípios da região oeste do Paraná (Assis Chateaubriand, Cascavel, Mamborê, Maripá, Palotina, Rancho Alegre D'Oeste, Toledo e Ubitatã), responsáveis por mais de 11% da área plantada de soja no estado (Figura 1). Esses municípios foram escolhidos por terem sido monitorados pelo projeto GeoSafras, coordenado pela Conab, que tem como um dos seus objetivos a geração de metodologias para subsidiar estimativas de safras com caráter menos subjetivo e, portanto, reduzir incertezas (Companhia Nacional de Abastecimento, 2011).

Para o estudo, foram utilizadas seis imagens TM/Landsat-5, órbita/ponto 223/77, adquiridas nas datas: 21/11/2003, 7/12/2003, 23/12/2003, 24/1/2004, 9/2/2004 e 25/2/2004. As imagens foram selecionadas de forma a abranger todo o período do ciclo de desenvolvimento da cultura da soja na região. A escolha do ano-safra 2003/2004 foi favorecida pelo número satisfatório de imagens com baixa cobertura de nuvens, o que não é muito comum nas safras de verão. As datas utilizadas englobaram todo o ciclo de desenvolvimento da soja, considerando variações das datas de semeadura e colheita que ocorrem na região, o que permitiu a classificação da maior parte das áreas da cultura presentes nas cenas.

O processamento das imagens foi feito com o programa “the environment for visualizing images” (Envi) e incluiu correção geométrica e transformação radiométrica, em que os valores digitais da imagem foram convertidos para valores físicos de fator de refletância aparente no topo da atmosfera, conforme Chander & Markham (2003). Não foi feita correção atmosférica, pois as classificações foram realizadas para cada imagem separadamente, com amostras e parâmetros da classificação únicos para cada data.

Em seguida, foram geradas composições coloridas RGB-453, das quais foram coletadas as amostras de

treinamento em áreas de cultivo de soja previamente definidas e com conhecimento de campo, no Município de Cascavel. A partir dessas amostras, foram executadas as classificações supervisionadas com os métodos paralelepípedo e máxima verossimilhança (MaxVer), escolhidos pela facilidade de parametrização e execução. O método paralelepípedo funciona bem quando as classes têm intervalos de valores digitais mais definidos. O método MaxVer trabalha com probabilidades e permite atuar com certo intervalo em torno dos agrupamentos das classes. Foram utilizadas a banda 3 (0,63–0,69 μm), a banda 4 (0,76–0,90 μm) e a banda 5 (1,55–1,75 μm), que apresentaram menor correlação entre si no caso dos alvos agrícolas, além da banda 7 (2,08–2,35 μm), que foi menos influenciada pelos efeitos de absorção e pelo espalhamento dos aerossóis e do vapor de água presentes na atmosfera. Ao final, foram gerados 12 mapas temáticos referentes a duas classificações em cada uma das seis cenas TM/Landsat-5. Estes mapas foram então sobrepostos em um único mapa temático, denominado de “máscara de soja”. Nas classificações pelo método MaxVer, utilizou-se limiar de aceitação de 90%, o que significa que os pixels que não alcançaram probabilidade de 90% de pertencer à classe “soja”, pré-determinada pela

amostra de treinamento, foram associados à classe de pixels classificados como “não soja”.

Após a geração da “máscara de soja”, procedeu-se à verificação da acurácia das classificações e estimativa de área plantada. Como referência terrestre, foram utilizados 100 pontos amostrais por município, a partir dos quais foi possível realizar a avaliação da acurácia das classificações por meio das métricas de exatidão global, índice Kappa e valores dos erros de inclusão e omissão (Richards & Jia, 2006).

Para o cálculo da área plantada, empregou-se o método denominado direto, proposto por Jupp (1989), que normaliza a matriz de erros pelas linhas. A probabilidade de um pixel, previamente atribuído a uma dada classe, pertencer a todas as classes examinadas pode ser estimada pela padronização das linhas e do total de pixels das classes de uso do solo determinadas pela imagem de satélite. Dessa forma, o método direto é considerado mais natural para corrigir as probabilidades de designação das classes, por meio das matrizes de proporções.

A Tabela 1 representa a matriz de erros e, entre parênteses, as proporções para as classes soja e não soja, geradas pela divisão dos elementos das linhas por N_i , conforme a equação $p(r_j | c_i) = n_{ij}/N_i$, construída para a correção das estimativas de áreas, em que: $p(r_j | c_i)$ é a proporção da classe de referência j representada na classe espectral i , na amostra; n_{ij} é o número de observações da classe de referência j que ocorrem na classe espectral i , na amostra; N_i é o número total de observações que ocorrem na classe espectral i , na amostra.

Supondo que o número total de pixels da classe espectral i na máscara de soja seja T_i , por município, então, a área estimada corrigida da classe de referência j em pixels é calculada pela equação

$$A_j = \sum_{i=1}^c p(r_j | c_i) T_i.$$

Assim, ao se considerar T_1 o total de pixels da classe soja na máscara, e T_2 o total de pixels da classe não soja na máscara, ambos por município, as estimativas de áreas corrigidas das classes soja e não soja foram calculadas pelas equações $A_1 = p(r_1|c_1)T_1 + p(r_1|c_2)T_2$ e $A_2 = p(r_2|c_1)T_1 + p(r_2|c_2)T_2$, respectivamente.

Para uma simples checagem, neste caso específico, a soma da área das classes soja e não soja tem de ser igual à área do município.

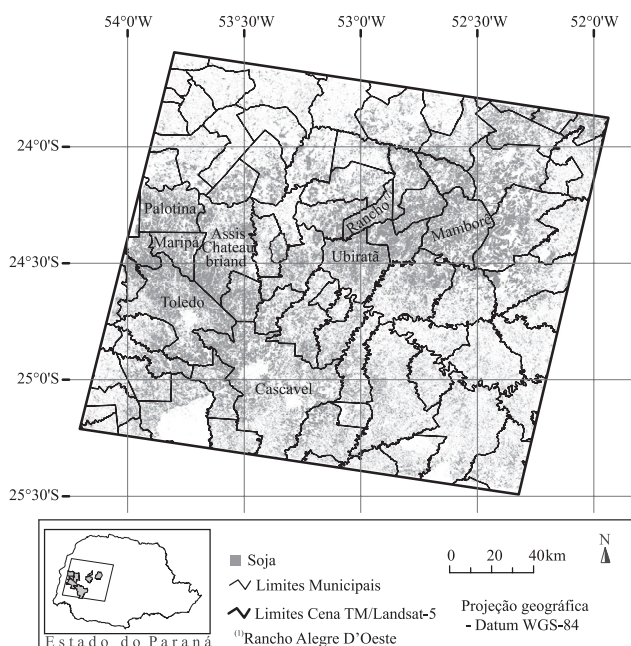


Figura 1. Máscara de soja gerada a partir da classificação de imagens TM/Landsat-5, para oito municípios do Estado do Paraná, com dados referentes à safra de 2003/2004.

Para fins de comparação, também foi estimada a área de soja de cada um dos oito municípios pelo método da expansão direta que, a partir de dados amostrais, reproduz, para toda a população, as proporções relativas encontradas nas amostras (Carfagna & Gallego 2005). Este método de estimativa de área combina os princípios estatísticos da teoria da amostragem com as características inerentes aos dados de sensoriamento remoto e foi utilizado, experimentalmente, para a safra de 2003/2004, pela Conab. Foram utilizados os 100 pontos do painel amostral reconhecidos como soja nas imagens TM/Landsat-5 e em levantamento de campo, com base na área total do município obtida do IBGE. A estratégia foi encontrar um paralelo entre os dois métodos que pudesse fornecer um indicativo sobre a acurácia da estimativa de área corrigida por matriz de erros, uma vez que o método de expansão direta já foi utilizado com bons resultados (Rizzi et al., 2006) e depende menos da disponibilidade de imagens.

De acordo com Epiphanyo et al. (2002), a estimativa por expansão direta, também com uso da máscara de soja, consiste na multiplicação da proporção de pontos identificados como soja a partir da referência terrestre pela área municipal, calculada como $A_{Soja} = (p_{Soja}/p_m)A_m$, em que: A_{Soja} é a área estimada da cultura de soja no município (ha); p_{Soja} é o número de pontos amostrais identificados como soja no município; p_m é o número total de pontos amostrais no município, igual a 100 pontos no presente trabalho; e A_m é a área total do município (ha).

Tabela 1. Matriz de erros e proporções para as classes soja e não soja.

| Máscara (i) | Pontos de referência (j) | | Total linha |
|--------------|--------------------------------------|--------------------------------------|----------------|
| | Soja (1) | Não soja (2) | |
| Soja (1) | $n_{11} (p(r_1 c_1) = n_{11}/N_1)$ | $n_{12} (p(r_2 c_1) = n_{12}/N_1)$ | N_1 |
| Não soja (2) | $n_{21} (p(r_1 c_2) = n_{21}/N_2)$ | $n_{22} (p(r_2 c_2) = n_{22}/N_2)$ | N_2 |
| Total | | | 100 |

Tabela 2. Métricas de acurácia da classificação e área de soja em cada município.

| Município | Erro de inclusão | | Erro de omissão | | Exatidão global | Índice Kappa | Área de soja (ha) |
|-----------------------|------------------|----------|-----------------|----------|-----------------|--------------|-------------------|
| | Soja | Não soja | Soja | Não soja | | | |
| Assis Chateaubriand | 0,060 | 0,520 | 0,356 | 0,111 | 0,710 | 0,708 | 48.724,11 |
| Cascavel | 0,270 | 0,143 | 0,250 | 0,156 | 0,810 | 0,808 | 82.434,42 |
| Mamborê | 0,050 | 0,425 | 0,230 | 0,115 | 0,800 | 0,799 | 46.961,46 |
| Maripá | 0,417 | 0,269 | 0,333 | 0,345 | 0,660 | 0,650 | 16.324,11 |
| Palotina | 0,327 | 0,386 | 0,315 | 0,400 | 0,646 | 0,635 | 34.002,09 |
| Rancho Alegre D'Oeste | 0,032 | 0,676 | 0,296 | 0,143 | 0,768 | 0,767 | 14.565,33 |
| Toledo | 0,176 | 0,327 | 0,276 | 0,214 | 0,750 | 0,746 | 66.565,89 |
| Ubiratã | 0,034 | 0,463 | 0,250 | 0,083 | 0,790 | 0,789 | 36.069,84 |

Para a avaliação da qualidade da classificação que gerou a máscara de soja, foram escolhidos os valores de referência do índice Kappa propostos por Landis & Koch (1977).

Para medir a precisão entre as duas estimativas de área, foi utilizado o erro relativo (ER), calculado pela subtração da segunda estimativa (matriz de erros) pela primeira (expansão direta), cujo resultado foi dividido pelo valor da primeira estimativa, multiplicado por 100, para ser expresso em porcentagem. Como a análise foi realizada de forma pareada entre duas estimativas de área, o ER é equivalente ao coeficiente de variação. Para fins de comparação, foi adotada a escala de Pimentel-Gomes (2000), em que o erro relativo é considerado baixo quando for menor que 10%, médio de 10 a 20%, alto de 20 a 30% e muito alto quando for maior que 30%.

Resultados e Discussão

O erro quadrático médio (RMS), encontrado na etapa de correção geométrica, foi de aproximadamente 10 m. Ao se considerar um pixel de 30 m do TM/Landsat-5, o RMS foi satisfatório, pois representou deslocamento máximo de um terço do pixel. Essa acurácia é essencial quando se trabalha com imagens multitemporais, para que a coleta dos dados seja sempre realizada na mesma posição geográfica.

Ao se analisar os dados dos erros de inclusão e omissão, ressalta-se que o menor valor para o erro de inclusão, para a classe soja, foi obtido no Município de Rancho Alegre D'Oeste (Tabela 2). O erro de omissão ocorrido neste município foi de 0,296. Isso significa que, dos 70,4% de áreas classificadas como soja, 96,8% corresponderam, de fato, a essa classe. Nos demais municípios, os valores dos erros para as classes soja e não soja variaram, respectivamente, entre 0,034 e

0,676, no erro de inclusão, e entre 0,083 e 0,40, no erro de omissão.

A exatidão global corresponde à concordância entre os pontos classificados como soja ou não soja na máscara e os da referência terrestre, presentes na diagonal principal da matriz de erros, por cada município. Os valores ficaram em torno de 0,742. O índice Kappa envolve a matriz de erros completa, e incorpora os elementos das linhas e das colunas marginais referentes aos erros de classificação, além dos que se situam na diagonal principal. Os valores para todos os municípios ficaram acima de 0,635.

Com base em Landis & Koch (1977), pode-se dizer que a classificação que deu origem à máscara de soja foi de qualidade “muito boa”. Entretanto, segundo Foody (2002), ao revisar métodos de análise de acurácia de classificação de imagens, os níveis recomendados para as métricas de exatidão global e o índice Kappa devem estar em torno de 0,85, para que se obtenha classificação com dados precisos. No presente trabalho, os classificadores apresentaram limitações ao efetuar o reconhecimento da cultura, apesar de o padrão da soja ter sido bem delimitado nas amostras de treinamento.

Ao se considerar o limitante de acurácia inerente aos classificadores digitais, a área da máscara de soja dos oito municípios foi corrigida por meio de matrizes de erros, transformadas em matrizes de proporções (Tabela 3), conforme Jupp (1989). Assim, todas as probabilidades de designação das classes foram corrigidas simultaneamente, com conseqüente correção das estimativas de área. Ao tomar como exemplo o Município de Cascavel, a área com soja foi de 78.388,29 ha e a sem soja de 131.680,26 ha. Para a análise dos resultados, utilizou-se, também, o método da expansão direta para a estimativa da área plantada de soja. Por esse método, no Município de Cascavel, ao se

Tabela 3. Matriz de erros e proporções, em Cascavel, PR⁽¹⁾.

| Máscara (i) | Pontos de referência (j) | | | | Total linha |
|--------------|--------------------------|-----------|--------------|-----------|-------------|
| | Soja (1) | | Não soja (2) | | |
| | Nº de pixels | Proporção | Nº de pixels | Proporção | |
| Soja (1) | 27 | 0,730 | 10 | 0,270 | 37 |
| Não soja (2) | 9 | 0,143 | 54 | 0,857 | 63 |
| Total | | | | | 100 |

⁽¹⁾As estimativas de áreas corrigidas das classes soja e não soja foram calculadas pelas equações $A_1 = p(r_1|c_1)T_1 + p(r_1|c_2)T_2$ e $A_2 = p(r_2|c_1)T_1 + p(r_2|c_2)T_2$, respectivamente, em que: $p(r_j | c_i)$ é a proporção da classe de referência j representada na classe espectral i, na amostra.

considerar a área total do município de 210.000,00 ha, a área com soja foi de 77.700,00 ha.

Ao final, as estimativas de área de soja calculadas a partir da máscara de soja, baseada na matriz de erros, foram comparadas àquelas obtidas por expansão direta e com a estimativa oficial realizada pela Seab/Deral (Paraná, 2005) (Tabela 4). Pôde-se observar que os valores de área encontrados na máscara de 2003/2004, após correção, aproximaram-se dos valores oficiais, o que confirma a eficácia da correção.

A correlação entre as estimativas de área de soja foram extremamente altas, superiores a 0,98 (Tabela 5). Como os valores de p foram menores que 0,05 para todas as comparações, a associação entre as estimativas foi considerada significativa, a 95% de probabilidade. Isso indica que as estimativas são fortemente correlacionadas, apresentam mesma tendência e movem-se, em perfeita proporção, na mesma direção.

Tabela 4. Estimativas de áreas de soja na safra de 2003/2004.

| Município | Área de soja (ha) | | |
|-----------------------|-------------------|-----------------|-----------------|
| | Seab/Deral | Matriz de erros | Expansão direta |
| Assis Chateaubriand | 72.000,00 | 70.918,97 | 70.810,00 |
| Cascavel | 82.000,00 | 78.388,29 | 77.700,00 |
| Mamborê | 54.000,00 | 57.744,34 | 57.646,00 |
| Maripá | 21.500,00 | 12.797,04 | 11.928,00 |
| Palotina | 43.300,00 | 34.987,46 | 35.154,00 |
| Rancho Alegre D'Oeste | 15.400,00 | 20.560,44 | 20.726,00 |
| Toledo | 68.600,00 | 72.194,60 | 69.426,00 |
| Ubiratã | 51.200,00 | 48.409,06 | 49.628,00 |

Tabela 5. Erro relativo entre as estimativas de áreas de soja na safra de 2003/2004.

| Município | Erro relativo (%) | | |
|--------------------------------|------------------------|------------------------|------------------------|
| | SD x ME ⁽¹⁾ | SD x ED ⁽²⁾ | ME x ED ⁽³⁾ |
| Assis Chateaubriand | -1,5 | -1,7 | -0,2 |
| Cascavel | -4,4 | -5,2 | -0,9 |
| Mamborê | 6,9 | 6,8 | -0,8 |
| Maripá | -40,5 | -44,5 | -6,8 |
| Palotina | -19,2 | -18,8 | 0,5 |
| Rancho Alegre D'Oeste | 33,5 | 34,6 | 0,8 |
| Toledo | 5,2 | 1,2 | -3,8 |
| Ubiratã | -5,5 | -3,1 | 2,5 |
| Coefficiente de correlação (r) | 0,98 | 0,98 | 0,99 |
| Valor de probabilidade (p) | 0,000 | 0,000 | 0,000 |

⁽¹⁾SD x ME, Seab/Deral x Matriz de erros. ⁽²⁾SD x ED, Seab/Deral x Expansão direta. ⁽³⁾ME x ED, Matriz de erros x Expansão direta.

A estimativa de área de soja, por meio de matriz de erros obtida para os municípios de Maripá e Palotina, foi inferior à estimativa da Seab/Deral em 40,5 e 19,2%, respectivamente. Já para o Município de Rancho Alegre D'Oeste, a área foi superestimada em 33,5%. Para os demais municípios, o ER ficou dentro da variação aceitável, e o menor valor foi observado para o Município de Assis Chateaubriand (Tabela 5).

Na comparação entre as estimativas de área de soja da Seab/Deral e a obtida por expansão direta, os valores do ER seguiram a mesma tendência de estimativa de área por meio de matriz de erros (Tabela 5). As estimativas para os municípios de Maripá e Palotina também foram menores que as estimativas oficiais, enquanto a estimativa do Município de Rancho Alegre D'Oeste superestimou em 34,6% a área de soja estimada pela Seab/Deral. O Município de Toledo apresentou o menor ER.

Ao se analisar o ER entre as estimativas de área de soja por meio de matriz de erros e por expansão direta, observou-se valores muito baixos para todos os municípios. O maior ER foi apresentado para o Município de Maripá, onde a estimativa por expansão direta subestimou em 6,8% a área de soja estimada por matriz de erros. Apesar das diferenças encontradas em relação à estimativa da Seab/Deral para os municípios de Maripá, Palotina e Rancho Alegre D'Oeste, observou-se grande similaridade entre as estimativas para todos os municípios, com maior correlação na análise pareada, com $r=0,99$ (Tabela 5).

Verificou-se que os erros apresentados pelos classificadores podem ser reduzidos por meio da metodologia proposta por Jupp (1989), que pondera os valores de área pelos erros encontrados na matriz. Cabe lembrar que essas estimativas basearam-se em métodos objetivos com uso de imagens de satélite, com resolução espacial de 30 m, ao contrário das estimativas da Seab/Deral, que são obtidas de forma subjetiva por meio de pesquisas de intenção de plantio, realizadas por amostragens. Mesmo assim, na maior parte dos municípios analisados, as estimativas de área se aproximaram dos dados oficiais (Tabela 4). Entretanto, os métodos objetivos podem trazer incorreções e incertezas que só podem ser identificadas a partir da comparação com um dado de referência confiável e representativo da realidade. Xiao et al. (2002) observaram, em escala municipal, diferenças entre os valores de área estimada por satélites e os valores das estatísticas oficiais, que ora subestimavam (7%), ora

superestimavam (92%). Ippoliti-Ramilo et al. (2003) e Rizzi et al. (2006) também obtiveram resultados conflitantes com os dados oficiais.

Apesar de os dados oficiais da Seab/Deral terem sido usados como referência para a comparação dos resultados apresentados, a alta correlação entre as estimativas geradas pelos métodos objetivos e o seu conflito com os dados oficiais em alguns municípios, em especial no caso de Maripá, indicam que a estimativa oficial pode não estar bem representada pelas amostragens, ao subestimar ou superestimar as previsões oficiais. Neste sentido, a adoção de uma referência confiável, determinada por mapeamento de imagens de satélite de maior detalhamento espacial, parece ser uma boa solução para determinar o nível de incertezas dos métodos objetivos e subjetivos de previsão de safras agrícolas.

Cabe destacar que, ao final do ciclo de desenvolvimento da soja na região em estudo, foi possível obter a estimativa de área plantada com a cultura antes da época da colheita, diferentemente do levantamento oficial que se estende além do final da safra.

Conclusões

1. A classificação temporal de imagens TM/Landsat-5 possibilita a geração de máscaras de cultivo e a obtenção de estimativas de área a partir da normalização da matriz de erros.

2. A estimativa de área de soja por meio da matriz de erros é obtida antes da colheita, e apresenta alta correlação com os dados oficiais.

3. As estimativas de área de soja, por meio da matriz de erros, aproximam-se mais das estimativas de área por expansão direta, o que indica melhor precisão.

4. A estimativa de área, a partir da normalização da matriz de erros, é viável e pode subsidiar os métodos convencionais na estimativa menos subjetiva de safras.

Agradecimentos

À Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas, pela disponibilização da infraestrutura do Laboratório de Geoprocessamento.

Referências

CARFAGNA, E.; GALLEGU, F.J. Using remote sensing for agricultural statistics. *International Statistical Review*, v.73, p.389-404, 2005.

- CEBALLOS-SILVA, A.; LÓPEZ-BLANCO, J. Delineation of suitable areas for crops using a multi-criteria evaluation approach and land use/cover mapping: a case study in Central Mexico. **Agricultural Systems**, v.77, p.117-136, 2003.
- CHANDER, G.; MARKHAM, B. Revised Landsat 5 TM radiometric calibration procedures and postcalibration dynamic ranges. **IEEE Transactions on Geoscience and Remote Sensing**, v.41, p.2764-2677, 2003.
- CHEN, D.M.; GOODCHILD, M. Calibrating area estimate bias on categorical maps using the contingency table. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v.34, p.1-4, 2007.
- COMPANHIA NACIONAL DE ABASTECIMENTO. **Boletim de acompanhamento de safras: grãos: safra 2011-2012: décimo levantamento**. Brasília: Conab, 2012. 29p. Disponível em: <<http://www.conab.gov.br/conteudos.php?a=1253&t=>>>. Acesso em: 6 ago. 2012.
- COMPANHIA NACIONAL DE ABASTECIMENTO. **Projeto GeoSafras**. Disponível em: <<http://www.conab.gov.br/conteudos.php?a=1079>>>. Acesso em: 19 jul. 2011.
- CONESE, C.; MASELLI, F. Use of error matrices to improve area estimates with maximum likelihood classification procedures. **Remote Sensing of Environment**, v.40, p.113-124, 1992.
- CONGALTON, R.G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices**. New York: Lewis Publishers, 137p. 1999.
- DE WIT, A.J.W.; CLEVERS, J.G.P.W. Efficiency and accuracy of per-field classification for operational crop mapping. **International Journal of Remote Sensing**, v.25, p.4091-4112, 2004.
- EPIPHANIO, J.C.N.; LUIZ, A.J.B.; FORMAGGIO, A.R. Estimativa de áreas agrícolas municipais utilizando sistema de amostragem simples sobre imagens de satélite. **Bragantia**, v.61, p.187-197, 2002.
- FOODY, G.M. Status of land cover classification accuracy assessment. **Remote Sensing of Environment**, v.80, p.185-201, 2002.
- GALLEGO, F.J. Remote sensing and land cover area estimation. **International Journal of Remote Sensing**, v.25, p.3019-3047, 2004.
- GALLEGO, J.; CRAIG, M.; MICHAELSEN, J.; BOSSYNS, B.; FRITZ, S. (Ed.). **Best practices for crop area estimation with remote sensing**. Ispra: Joint Research Center, 2008. Available at: <http://www.earthobservations.org/documents/cop/ag_gams/GEOSS%20best%20practices%20area%20estimation%20final.pdf>. Accessed on: 3 Aug. 2012.
- IPPOLITI-RAMILO, G.; 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, p.1521-1534, 2003.
- JUPP, D.L.B. The stability of global estimates from confusion matrices. **International Journal of Remote Sensing**, v.10, p.1563-1569, 1989.
- LANDIS, J.R.; KOCH, G.G. The measurement of observer agreement for categorical data. **Biometrics**, v.33, p.159-174, 1977.
- PARANÁ. Secretaria da Agricultura e do Abastecimento. Departamento de Economia Rural. **Estimativas de safras**. Disponível em: <<http://www.seab.pr.gov.br>>. Acesso em: 27 abr. 2005.
- PARANÁ. Secretaria da Agricultura e do Abastecimento. Departamento de Economia Rural. **Soja (2009-2010)**. Disponível em: <<http://www.seab.pr.gov.br>>. Acesso em: 6 ago. 2012.
- 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, p.1301-1316, 2011.
- PIMENTEL-GOMES, F. **Curso de estatística experimental**. 14.ed. Piracicaba: ESALQ, 2000. 477p.
- RICHARDS, J.A.; JIA, X. **Remote sensing digital image analysis: an introduction**. 4th ed. New York: Springer-Verlag, 2006. 439p.
- RIZZI, R.; RUDORFF, B.F.T.; ADAMI, M. Estimativa da área de soja no Estado do Rio Grande do Sul por um método de amostragem. **Ciência Rural**, v.36, p.30-35, 2006.
- SERRA, P.; PONS, X. Monitoring farmer's decisions on Mediterranean irrigated crops using satellite image time series. **International Journal of Remote Sensing**, v.29, p.2293-2316, 2008.
- SHAO, G.F.; WE, W.C.; WU, G.; ZHOU, X.H.; WU, J.G. An explicit index for assessing the accuracy of cover-class areas. **Photogrammetric Engineering and Remote Sensing**, v.69, p.907-913, 2003.
- WU, S.; WU, B.; FENG, Q.; XUE, H.; DU, Y.; LI, Y. Determining crop area proportion based on the specific sampling method applied. **Journal of Agricultural Science**, v.4, p.123-130, 2012.
- XIAO, X.; BOLES, S.; FROLKING, S.; SALAS, W.; MOORE, B.; LI, C.; HE, L.; ZHAO, R. Landscape-scale characterization of cropland in China using Vegetation sensor data and Landsat TM images. **International Journal of Remote Sensing**, v.23, p.3579-3594, 2002.

Recebido em 28 de julho de 2011 e aprovado em 3 de agosto de 2012