

PREVISÃO DE ÁREA PLANTADA COM SOJA, SAFRA 2017/18, NO SUL DE RONDÔNIA

Marcelo Vinicius Assis de Brito¹, Bárbara Laura Tavares¹, Valdir Moura¹, João Paulo Miranda de Castro², Ranieli dos Anjos de Souza¹

¹GREES - Grupo de Pesquisas Espaciais, Campus Colorado do Oeste, Instituto Federal de Rondônia - IFRO, BR 435, Km 66, Zona Rural, RO-399, s/n, Colorado do Oeste - RO, 76993-000, Brasil, marcelobrito03@gmail.com; barbaralauratavares@gmail.com; valdir.moura@ifro.edu.br; ranieli.muler@ifro.edu.br

²CIMO - Centro de Investigação de Montanha, Escola Superior Agrária, Instituto Politécnico de Bragança - IPB, Bragança, 5300-253, Portugal; jpmc@ipb.pt

RESUMO

O agronegócio exige que as organizações tenham dados precisos e no momento certo para nortear a tomada de decisões quanto à produção agrícola. Dessa forma, a estimativa de área a ser plantada com culturas anuais é uma das variáveis necessárias para a previsão de safras. Neste sentido, o presente trabalho objetivou estimar a área a ser plantada com soja (*Glycine max*) no município de Cerejeiras, Rondônia, safra 2017/18. Foram selecionadas imagens do sensor OLI/LANDSAT8 de agosto de 2017, período que coincide com a entressafra, e os solos encontram-se expostos devido preparo para plantio. As abordagens utilizadas foram: processamento com o modelo linear de mistura espectral (MLME) e aplicação da classificação supervisionada (MaxVer) sobre a imagem fração. Os resultados foram validados através da matriz de confusão que obteve fiabilidade global de 98,60% e Kappa 0,98, mapeando uma área de 32.664,37 ha de solo exposto apto para o plantio de soja.

Palavras-chave — estimativa, área plantada, solo exposto, modelo linear, maxver.

ABSTRACT

*The agribusiness requires organizations to have accurate and timely data to guide decision-making on agricultural production. Thus, the estimate of area to be planted with annual crops is one of the variables necessary for crop forecasting. In this sense, the present work aimed to estimate the area to be planted with soybean (*Glycine max*) in the municipality of Cerejeiras, Rondônia, crop 2017/18. OLI / LANDSAT8 images were selected from August 2017, a period that coincides with the off season, and the soils are exposed due to their preparation for planting. The approaches used were: processing with the linear model of spectral mixing (MLME) and application of the supervised classification (MaxVer) on the image fraction. The results were validated through a confusion matrix that obtained a global reliability of 98.60% and Kappa 0.98, mapping an area of 32,664.37 ha of exposed soil suitable for soybean planting.*

Key words — estimation, planted area, exposed soil, linear model, maxver.

1. INTRODUÇÃO

A disponibilidade de dados confiáveis em tempo hábil sobre a estimativa de área plantada com culturas anuais, tem grande valia para a elaboração de estratégias quanto ao armazenamento, comercialização e suporte às decisões governamentais [1].

De acordo com dados do 5º Levantamento da Safra de Grãos 2017/2018 da Companhia Nacional de Abastecimento (CONAB), somente o cultivo da soja (*Glycine max*) representou cerca de 58% das áreas de culturas de verão plantadas no país, demonstrando, assim, sua importância no cenário do agronegócio [2].

Atualmente, o Instituto Brasileiro de Geografia e Estatística (IBGE) é o órgão responsável pelos Censos Agropecuários no Brasil, realizado por meio de uma pesquisa que faz acompanhamento do desenvolvimento da safra agrícola, adotando-se diferentes procedimentos em seu processo de investigação, de acordo com as diversas características de cada cultura, tais como: levantamento de demanda de insumos agrícolas (e. g., sementes, fertilizantes, corretivos) junto aos produtores, associações de classe e cooperativas; levantamento e observações de campo; levantamento de informações sobre a comercialização, destino da produção colhida e outros dados correlacionados [3].

Desta forma, o sensoriamento remoto e as técnicas de geoprocessamento são uma alternativa para a estimativa de áreas plantadas, por serem capazes de fornecer informações em grandes extensões, possibilitarem o monitoramento das culturas com baixo custo e, permitirem análises em curto tempo. Os erros associados aos resultados subsidiados pelo sensoriamento remoto e geoprocessamento podem, portanto, serem menores que os erros associados às estimativas baseadas em diferentes fontes de dados, mas, é preciso avaliar cuidadosamente a eficiência dos modelos e os resultados alcançados com uso das geotecnologias. Por isso, a presente pesquisa teve por objeto estimar a área a ser plantada com soja (*Glycine max*) no município de Cerejeiras, Rondônia, safra 2017/18, compreendendo o período de entressafra, a fim

de utilizar o solo exposto, preparado para plantio, como área de análise para a previsão de área plantada.

2. MATERIAIS E MÉTODOS

A área estudo compreende o município de Cerejeiras (13°11'20"S, 60°48'44"W), estado de Rondônia, abrangendo uma área de 278.331 hectares (Figura 1). O município é caracterizado por ser uma área de transição do bioma Cerrado-Amazônia, apresentando clima tropical quente e úmido conforme o sistema de Koppen [4].

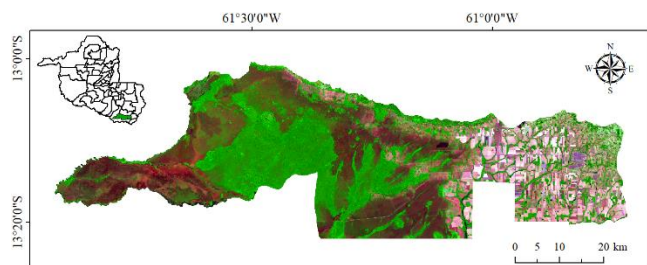


Figura 1. Área de Estudo, município de Cerejeiras.

Para o processamento dos dados foram utilizadas imagens do sensor *Operational Land Imager* (OLI) a bordo do satélite *LANDSAT/8*, órbita-ponto 230/069, de 08 de agosto de 2017, obtidas na plataforma da Divisão Geral de Imagem (DGI) do Instituto Nacional de Pesquisas Espaciais (INPE), projetadas no sistema UTM (WGS84), com uma resolução espacial de 30 m. Esta data foi escolhida por representar o período em que os solos de uso agrícola estão limpos, haja visto que o plantio da soja tem início em outubro, além disso, para obter imagens com baixa incidência de nuvens.

A metodologia adotada nesta pesquisa consistiu no pré-processamento das imagens com o delineamento amostral da área de estudo, foi utilizada uma máscara do Parque Estadual de Corumbiara, para que a área do parque contida no município fosse excluída da análise, uma vez que a região de interesse são as áreas consolidadas para uso agrícola.

Em seguida o Modelo Linear de Mistura Espectral (MLME) [5] foi aplicado para obtenção das imagens fração: solo, vegetação e sombra. Sobre a imagem fração foi aplicado o processo de classificação supervisionada de máxima verossimilhança (MaxVer), para com isso, obter a quantificação das áreas de preparo do solo para plantio de soja. Desta forma, foram delimitadas as classes: solo exposto (áreas de preparo do solo), floresta e uso consolidado (áreas com pastagens, outros usos agrícolas e áreas urbanas).

A precisão da classificação foi avaliada utilizando dois coeficientes: o Índice de Fiabilidade Global (Fg) [6], e o Índice de Concordância Kappa [6][7][8], que derivam de matrizes de confusão (*confusion matrix*).

O Kappa, além dos valores marginais, também contabiliza as relações entre as diferentes categorias, através de uma análise (categórica) multivariada das matrizes de confusão, verificando se uma classificação discriminou as

categorias de interesse com uma precisão significativamente maior do que a que se obteria aleatoriamente [9].

O esforço amostral (n) pode ser calculado segundo distribuições binomiais ou multinomiais [8], de acordo com a expressão: $n = B/4b^2$; em que, $n = \sum_{i=1}^k n_i$, B = percentil superior (α/k) da distribuição de χ^2 com 1 grau de liberdade, k = número de estratos, n_i = intensidade de amostragem no estrato i (distribuição binomial), b = erro admissível ($= 0,15$), nível de confiança $\alpha=0,05$.

A matriz de contingência registra os resultados da verificação no campo em colunas e os resultados da interpretação das imagens em linhas. Os valores da diagonal desta matriz quadrada (x_{ii}), representam a concordância entre os resultados da interpretação e da realidade observada. A fiabilidade global (Equação 1) da classificação expressa a relação entre o número de unidades de amostragem corretamente classificadas ($\sum x_{ii}$) (somatório dos pontos da diagonal) e o número total de pontos ($\sum x_{ij}$) [9].

$$Fg = \frac{\sum x_{ii}}{\sum x_{ij}} = \frac{N_{ii}}{N_{ij}} \quad (1)$$

Os valores acumulados por linha (x_{i+}) e por coluna (x_{+i}) e os valores da diagonal (x_{ii}), permitem avaliar os erros de omissão (erros de exclusão) e de comissão (erros de inclusão) para cada classe de ocupação do solo. Sendo geradas pelas Equações 2 e 3:

$$Eo_i = \frac{x_{+i} - x_{ii}}{x_{+i}} \quad (2)$$

$$Ec_i = \frac{x_{i+} - x_{ii}}{x_{i+}} \quad (3)$$

O Kappa, além dos valores marginais, também contabiliza as relações entre as diferentes categorias, verificando se uma classificação discriminou as categorias de interesse com uma precisão significativamente maior do que a que se obteria aleatoriamente [9]. Esse índice consiste numa medida de agregação do observador/classificador com os dados categóricos [10]. Suponhamos a classificação de uma imagem de satélite com N pixels em k categorias, assumindo um modelo de amostragem multinomial, a estimativa mais verossímil de Kappa é dada pelas Equações 4, 5 e 6:

$$Kappa = \frac{P_o - P_e}{1 - P_e} \quad (4)$$

$$P_o = Fg = \frac{N_{ii}}{N_{ij}} \quad (5)$$

$$P_e = \sum \left(\frac{x_{i+}}{N_{ij}} \cdot \frac{x_{+i}}{N_{ij}} \right) = \sum (p_{i+} \cdot p_{+i}); \quad (6)$$

Em que P_o a proporção observada de concordância, P_e a proporção esperada de concordância puramente aleatória, p_{ii} a percentagem de concordância observada na classe i , p_{i+} a probabilidade marginal na linha i , p_{+i} a probabilidade marginal na coluna i . Os valores admissíveis de Kappa [9], estão apresentados na Tabela 1. Podem ser estimados os valores marginais de Kappa tendo como base a interpretação,

perspectiva do utilizador ($Kappa_u$), ou a verificação de campo, perspectiva do produtor ($Kappa_p$), conforme as Equações 7 e 8, respetivamente.

$$Kappa_u = \frac{p_{ii} - p_{i+} \cdot p_{+i}}{p_{+i} - p_{i+} \cdot p_{+i}} \quad (7)$$

$$Kappa_p = \frac{p_{ii} - p_{i+} \cdot p_{+i}}{p_{i+} - p_{i+} \cdot p_{+i}} \quad (8)$$

Tabela 1. Qualidade da classificação em função de Kappa.

Kappa	Qualidade
≈ 0	Péssima
[0; 0,2]	Má
[0,2; 0,4]	Razoável
[0,4; 0,6]	Boa
[0,6; 0,8]	Muito Boa
[0,8; 1]	Excelente

Fonte: [6][7][8].

3. RESULTADOS

As assinaturas espectrais dos *pixels* puros (*endmembers*), utilizadas na aplicação do MLME, estão apresentadas na Figura 2. Durante a seleção dos *endmembers* notou-se que *pixels* com sombra apresentam resposta de fundo no espectro visível, quando sobreposto a áreas com vegetação.

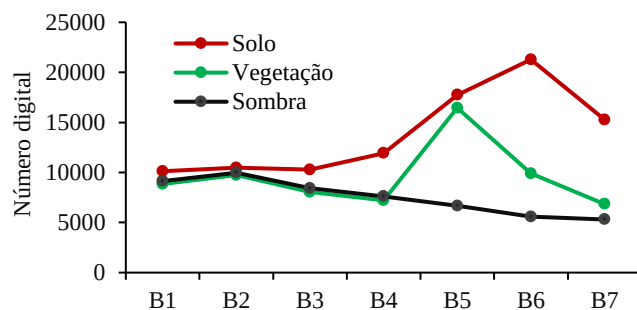


Figura 2. Resposta espectral dos endmembers, bandas do sensor OLI.

A seleção dos *endmembers* foi realizada diretamente na imagem, escolhendo *pixels* com a resposta espectral mais próxima da curva teórica esperada para alvos puros [11]. A vegetação apresenta maior resposta na faixa do infravermelho próximo, banda. O solo exposto já apresenta maior refletância no infravermelho médio, banda 6. Enquanto a fração sombra apresenta baixa resposta ao longo de todo espectro, devido à alta absorção da radiação incidente nos variados comprimentos de onda.

Considerando $k=3$ classes de ocupação do solo, nível de confiança $=0,05$, $b=0,15$ (erro máximo admissível), com 1 grau de liberdade, o número mínimo de amostras a efetuar para a verificação de campo seria de 64, tendo sido recolhidas 357 amostras (*pixels*), em média cerca de 119 por classe, conforme a Tabela 3. A amostragem foi realizada de forma

totalmente independente por interpretação visual sobre uma composição RGB654.

Após a classificação supervisionada com o classificador MaxVer, os resultados foram validados através da matriz de contingência que obteve fiabilidade global de 98,60%, e Kappa de 0,98. A classe de solo exposto obteve precisão de 96,58%, em que 3,42% destas áreas foram confundidas com área antropizada. O mapeamento da classe floresta foi exato, 100% das amostras foram classificadas corretamente. Já as áreas antropizadas apresentaram precisão de 99,17%, enquanto cerca de 0,83% das áreas foram confundidas com solo exposto. Na Tabela 2 podem ser observadas as distribuições dos acertos e erros da classificação em cada classe de uso.

Tabela 2. Resultados da classificação MaxVer.

Classes	Verdade de campo (%)			Total
	Área antropizada	Solo Exposto	Floresta	
Área antropizada	99,17	3,42	0,00	34,45
Solo exposto	0,83	96,58	0,00	31,93
Floresta	0,00	0,00	100,00	33,61
Total	100,00	100,00	100,00	100,00

Conforme Tabela 3, sob a perspectiva do utilizador obtiveram-se valores de Kappa parciais de 0,95, 0,99 e 1,00 para as 3 classes de área antropizada, solo exposto e floresta, respetivamente. Sob a perspectiva do produtor os valores correspondentes foram 0,99, 0,95 e 1,00.

Tabela 3. Indicadores de exatidão das classes de uso derivadas da classificação supervisionada (MaxVer).

Classes	Porcentagem		Pixels	
	Comissão	Omissão	Comissão	Omissão
Área antropizada	3.25	0.83	4/123	1/120
Solo exposto	0.88	3.42	1/114	4/117
Floresta	0.00	0.00	0/120	0/120

Após a classificação foi aplicada uma filtragem passa-alta para eliminação de *pixels* isolados, resultando no quantitativo de 32.664,37 ha de solo exposto apto para o plantio de soja (Figura 3).

4. DISCUSSÃO

Segundo o IBGE na safra 2017/18, o município de Cerejeiras apresentou área plantada de 28.000 ha [12], representando uma diferença de cerca de 17% entre os dados oficiais e os estimados por esta pesquisa. Os erros de comissão na classe solo exposto contribuem para o aumento na estimativa de área plantada com soja, contudo, os erros de comissão e omissão nesta classe totalizam 4,3%, valor abaixo da diferença observada (17%), indicando que, as áreas plantadas podem ser maiores do que as informadas pelos órgãos responsáveis pelo censo agrícola.

Atualmente, Cerejeiras é o terceiro município em área

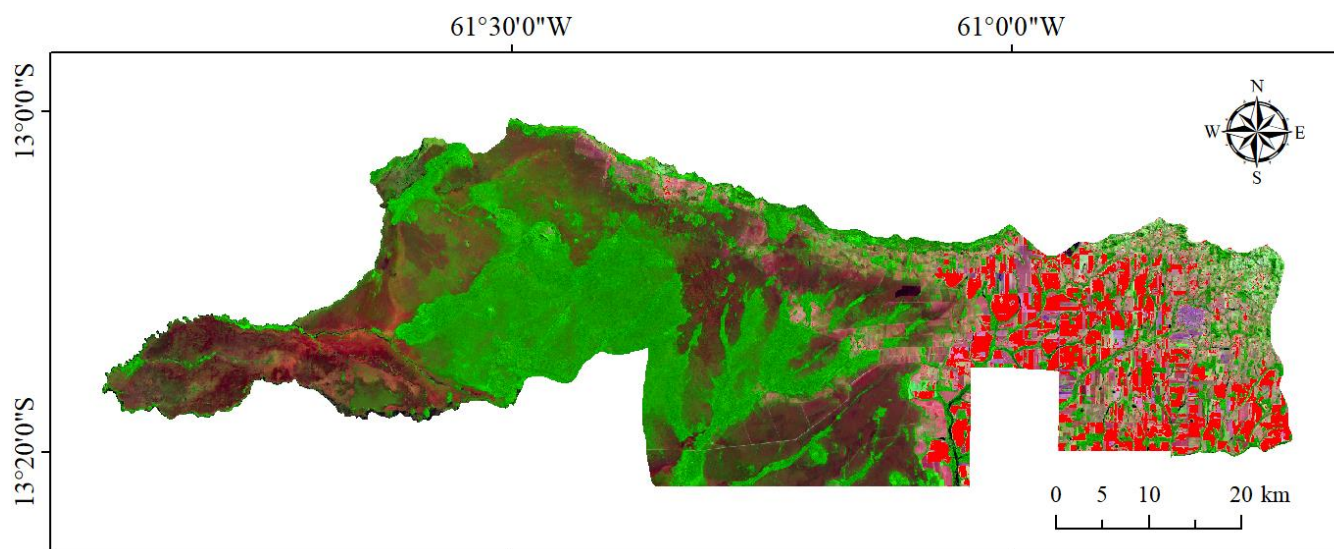


Figura 3. Área de solo exposto mapeada (em vermelho).

destinada ao plantio de soja no estado de Rondônia, ficando atrás apenas dos municípios de Vilhena e Corumbiara. A área levantada neste trabalho, como possível uso para o plantio de soja, representa 11,34% da área total do município.

Considerando que o Parque Estadual de Corumbiara ocupa 44% do território municipal, este percentual passa para 22% das áreas consolidadas, colocando a soja como uma cultura de forte predominância no agronegócio local.

5. CONCLUSÕES

Embora a estimativa levantada nesta pesquisa seja diferente da apresentada pelo órgão oficial, a classificação aplicada no período de preparo do solo mostrou-se uma metodologia eficaz para previsão de áreas plantadas com soja. Os erros observados neste estudo, podem ser minimizados explorando-se outros métodos e abordagens tais como, o uso de sensores com diferentes resoluções espaciais e espectrais.

O esforço amostral na etapa de treinamento, com coleta de amostras acima dos valores recomendados, garante confiabilidade aos resultados obtidos, indicando que o período entre safras pode ser utilizado para a estimativa de áreas plantadas com soja.

6. REFERÊNCIAS

- [1] Formaggio, A.R.; Moura, V.; Epiphany, J.C.N.; Fiorio, P.R.; Campos, R.C., “Dados tm/landsat na estimativa de áreas destinadas a culturas de verão, no estado de São Paulo”. Anais XI SBSR, p. 93–100, 2003.
- [2] CONAB - Companhia Nacional de Abastecimento. “5º Levantamento da Safra de Grãos 2017/2018”. v. 5, n. 9, p. 18, 2018.
- [3] IBGE, Instituto Brasileiro de Geografia e Estatística. “Pesquisas agropecuárias”. Rio de Janeiro: ed. 2, v.6, p. 92., 2002.
- [4] Bastos, T. X. e Diniz, T. D. “Avaliação do clima do estado de Rondônia para desenvolvimento agrícola”, *Belém*, n. 44, p. 5–27, outubro, 1982.
- [5] Shimabukuro, Y.E. 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.
- [6] Cohen, J. “A Coefficient of Agreement of Nominal Scales”. *Educational and Psychological Measurement*, v. 20, n. 1, p. 37-46, 1960.
- [7] Rosenfield, G.H.; Fitzpatrick-Lins, K.; Ling, H., “Sampling for Thematic Map Accuracy Testing”. *Photogrammetric Engineering & Remote Sensing*, v. 48, n. 1, p. 131-137, 1982.
- [8] Congalton, R.G. e Green, K. “Assessing the accuracy of remotely sensed data: principles and practice”. *Lewis publishers, New York*. p. 137, 1999.
- [9] Chuvieco, E., “Fundamentos de Teledetección Espacial”, Ediciones Rialp, S.A., Madrid, p. 375-395, 1990.
- [10] Rosenfield, G.H. e Fitzpatrick-Lins, K., “A Coefficient of Agreement as a Measure of Thematic Classification Accuracy”, *Photogrammetric Engineering & Remote Sensing*, v. 52, n. 2, p. 223-227, 1986.
- [11] Anderson, L. O.; Aragão, L. E. O. C.; Lima, A.; Shimabukuro, Y. E., “Detecção de cicatrizes de áreas queimadas baseada no modelo linear de mistura espectral e imagens índice de vegetação utilizando dados multitemporais do sensor MODIS/TERRA no estado do Mato Grosso”, *Acta Amazonica*, v. 35, p. 445–456, 2005.
- [12] IBGE, Instituto Brasileiro de Geografia e Estatística. “Produção Agrícola Municipal de 2017”. Rio de Janeiro, v.6, p. 92, 2018.