SOYBEAN CROP AREA ESTIMATION AND MAPPING IN MATO GROSSO STATE, BRAZIL

A. Gusso [a, *], J. R. Ducati [a,b]

[a] CEPSRM-UFRGS, Center for Remote Sensing and Meteorological Research,
Federal University of Rio Grande do Sul, 15044, Porto Alegre, Brazil – anibal.gusso@ufrgs.br
[b]Institute of Physics, Federal University of Rio Grande do Sul, Porto Alegre, Brazil, - ducati@if.ufrgs.br

Commission WG VII/5

KEY WORDS: Crop Area, Algorithm, LUCC mapping, Remote Sensing, MODIS

ABSTRACT:

Evaluation of the MODIS Crop Detection Algorithm (MCDA) procedure for estimating historical planted soybean crop areas was done on fields in Mato Grosso State, Brazil. MCDA is based on temporal profiles of EVI (Enhanced Vegetation Index) derived from satellite data of the MODIS (Moderate Resolution Imaging Spectroradiometer) imager, and was previously developed for soybean area estimation in Rio Grande do Sul State, Brazil. According to the MCDA approach, in Mato Grosso soybean area estimates can be provided in December (1st forecast), using images from the sowing period, and in February (2nd forecast), using images from sowing and maximum crop development period. The results obtained by the MCDA were compared with Brazilian Institute of Geography and Statistics (IBGE) official estimates of soybean area at municipal level. Coefficients of determination were between 0.93 and 0.98, indicating a good agreement, and also the suitability of MCDA to estimations performed in Mato Grosso State. On average, the MCDA results explained 96% of the variation of the data estimated by the IBGE. In this way, MCDA calibration was able to provide annual thematic soybean maps, forecasting the planted area in the State, with results which are comparable to the official agricultural statistics.

---

* Corresponding author.

1. INTRODUCTION

1.1 Current Methodologies for Crop Area Estimation

The agricultural production of Brazil has an important role in the world commodities market. Until now, a fully operational method for planted area forecast is not known. In this way, it is relevant to develop objective methods that can provide precise and timely estimates of crop production. The Brazilian agencies CONAB (Companhia Nacional de Abastecimento - National Company of Food Supply) and IBGE (Instituto Brasileiro de Geografia e Estatística - Brazilian Institute of Geography and Statistics are both responsible for the official agricultural statistics in Brazil. Currently, institutional methodology for crop forecasting in Brazil suffers from a large subjective component of the technical staff involved. It is well known that there is a confidence difficulty associated with subjective methods, because they do not present an associated error measurement. In other hand, the completely statistical-based methods suffer from not having a conceptual physics anchorage when its results are submitted to spatial-temporal analysis, which commonly, throws them back to less objective approaches.

Usually, the major limitation to optical remote sensing is high cloud cover incidence during key periods in order to identify annual crops (Sano et al., 2007). The best solution found to overcome the cloud cover issue is to increase the temporal resolution of orbital sensors, but due to technical constraints this will reduce spatial resolution. The Moderate Resolution Imaging Spectroradiometer (MODIS) sensor, aboard the Terra satellite provides an adequate imaging configuration, with an almost-daily revisit with a moderate spatial resolution of 250m, which its favourable characteristics to map crops in large scale. Even when MODIS is not a viable option for detailed cropland mapping due to its limitation to resolve smaller field sizes, it still can provide cropland presence over large areas (Gusso et al., 2012; Rudorff et al., 2011; ABIOVE, 2010; Pittman et al., 2010). The geometric quality of the images allows time series composition that performs good geolocation of pixels (Justice et al., 2002).

Many studies have used MODIS imagery for agricultural crop surveys and monitoring. Lobell and Asner (2004) evaluated the impact of land use and soil cover for agricultural classification and observed an inverse relationship of the Mean Square Error (MSE) and field size, tending to stabilize for fields larger than 500ha. Those authors also concluded that MODIS images have considerable advantages over Landsat images in the characterization of extensive agricultural crops, mainly due to their higher temporal resolution. Doraiswamy et al. (2005) evaluated the quality of MODIS data to provide information on both crop yield and area. In the USA, Wardlow et al. (2007) investigated the applicability of MODIS/EVI time series data to map agricultural lands, and concluded that 16-day composites of MODIS images gave sufficient spatial, spectral and temporal information to adequately separate crop fields from other land uses and also expressed the phenology and climate characteristics of the region.

1.2 Soybean Crop Area Mapping in Brazil

In Brazil, several works from government and Non-governmental Organizations efforts have tried to solve the problem of crop area mapping and forecast. Epiphanio et al. (2010) used MODIS in a temporal-spectral surface response approach for mapping soybean in Mato Grosso State and obtained an overall accuracy of 80%. Even so, most of these studies are modelled to few crop years and/or for a limited region, indicating good potential of MODIS data for crop

forecast, but not actually proving its usefulness within a routine and systematic crop forecast system (Gusso et al., 2012; Gusso et al., 2009), not able to provide timely soybean area estimates in an operational mode. Gusso et al. (2012) developed the MODIS Crop Detection Algorithm, which is a systematic forecast system, with R^2 greater than 0.95 and overall accuracy of 82% for several crop years in Rio Grande do Sul State, but also found mapping accuracy dependence from the mean field size (Lobell and Asner 2004), by using 250 meters spatial-resolution. On the modelling of the physically-driven components of soybean crop areas those authors observed that the difficult is to understand the main dynamic processes involved, from one harvest to another, which leads to changes in vegetation index. In this sense, the objective in this study is to evaluate the performance of the MCDA procedure adaptation to Mato Grosso State, in order to estimate crop area and mapping of soybean in local and regional scales.

2. MATERIAL AND METHODS

2.1 Study Area

Mato Grosso State accounts for almost 27% of the total area for Brazilian soybean production, being the major soybean producer in the country (IBGE, 2011). It has 146 municipalities, as shown in Figure 1.
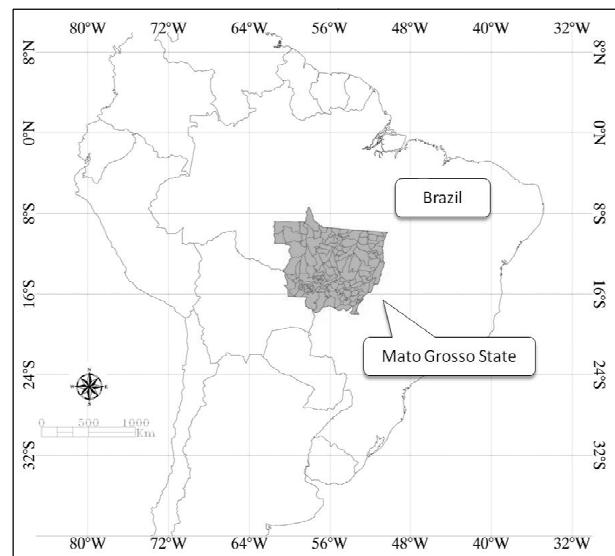

Figure 1.Mato Grosso State, Brazil, and its 146 municipalities

Typically, the sowing period for soybean goes from early September to late October, based on agricultural zoning for different soils, regions, and cultivars (MAPA, 2009). Depending on sowing date, which is function of rain occurrence, maximum plant growth is observed from early January to early February. The natural vegetation cover in the region is the one typical of the Cerrado biome, which is also a cause of confusion with soybean development during the rainy season (Epiphanio et al., 2010). The soybean cycle development profiles are found to be analogous to those given by Wardlow et al. (2007). Soybean fields can be distinguished from other vegetation classes, including natural grazing fields and crops, from the application of an algorithm which uses data acquired in different dates. This algorithm is based on the soybean EVI temporal profile, especially, on the low EVI values during sowing period and high EVI values at maximum crop development.

2.2 Materials

As through fully described in Gusso et al., (2012) a set of several types of data was needed to accurately represent the physical conditions and management practices found in Mato Grosso: i) EVI data from 2000 to 2011 were extracted from MODIS sensor on board of Terra satellite, product MOD13Q1-collection 5 for 6 image Tiles (H11V09, H11V10, H12V09, H12V10, H13V09, H13V10) covering all Mato Grosso State; ii) Shuttle Radar Topography Mission (SRTM) data were used to generate a slope map with 90 meters spatial resolution, according to Rabus et al. (2003), in order to exclude areas improper for mechanization (slope >12%); iii) annual soybean agricultural statistics, in state and municipality level, (IBGE, 1989) for the entire study area were used to compare and evaluate the results obtained from the present soybean area estimation procedure; iv) yearly rainfall data (2000 to 2009) in 10 days accumulated precipitation for 11 meteorological stations, during the period from September to October were used to refine the period of initial sowing from INMET (Instituto Nacional de Meteorologia - Brazilian Institute of Meteorology); v) as a second source of information data, annual soybean agricultural statistics, at state level, (CONAB, 2011) for the entire study area were used also to compare and evaluate the results obtained from the present soybean area estimation procedure and IBGE; and vi) and TM images obtained from INPE (www.dgi.inpe.br) to perform crosschecking with crop area maps.

The EVI data were chosen due to its very know potential to mitigate cloud cover effect, atmospheric and soil background effects (Huete et al. 2002; Justice et al. 2002). The EVI data is obtained from the MOD13Q1-V005 product, which is a 16-days composition with the best radiometric and geometric pixels selected. MODIS images and products were pre-processed by the National Aeronautics and Space Administration (NASA) and are available at no charge at https://wist.echo.nasa.gov/api/.

2.3 Methodology

The SRTM data were used to generate the land slope variable in order to exclude from the analysis all areas with slope greater than 12%, once soybean is a highly mechanized crop and requires relative smooth land to allow the traffic of farm implements. The soybean classification procedure used in this study was the previously named MODIS Crop Detection Algorithm (MCDA). The present work aims to adapt and test the MCDA, in a way to evaluate how MCDA adheres to both conditions A and B described in Gusso et al. (2012). According to MCDA approach, the variables are the following: Amp (amplitude, which is the difference between maximum and minimum EVI values); Lmin (is the lower minimum EVI value in a minimum image) and Umin (the upper minimum which is the maximum EVI value in a minimum image). Details of this procedure are available in Gusso et al. (2012). In order to adjust these two conditions into the algorithm for Mato Grosso, some parameters have to be defined based on the EVI temporal profile of the soybean. Adjustments of MCDA parameters, from now on, will be referred here as MCDA calibration.

For Mato Grosso, the sowing period for each crop year must be defined based on rainfall, in agreement with the soybean zoning provided by MAPA (2009). In the study area, the sowing period normally starts during September. In order to cover the maximum vegetation development period, it is used four consecutive EVI images from DOY (Day of Year) 353 to 033. The challenge is to obtain the best Amp value that includes not only pure soybean pixels with high values in the maximum EVI

image and low values in the minimum EVI image, but also mixed pixels located at the border of soybean fields.

The procedure used to verify the map accuracy was overlaying Landsat TM5 images and the classification maps of the MCDA. The main orbits used were 224069 for Feb/7/2001, Jul/31/2006, Jan/23/2007 and Jan/12/2009; 227069 for Feb/12/2001, Jan/30/2002, Feb/7/2005, Jan/1/2009 and Jan/20/2010.

It should be emphasized that MCDA calibration procedure is only completed when Mato Grosso parameters, which were chosen to represent the physical-driver components defined in the MCDA, can be used as the same input for all the analysed crop years. Therefore, once identified the parameters from the physically-driven components, no post-adjustment is allowed in a way to constrain dynamical adjustment process of the algorithm (Gusso et al., 2009). If some further adjustment is needed, in order to plot a better fitting of crop areas with Landsat TM5 images and statistics from IBGE, for one harvest or more, then this new parameter must run for all tested crops. The selected values are 3550, 500 and 3900 for Amp, Lmin and Umin respectively, on the MCDA calibration for Mato Grosso Sate.

For forecasts, soybean area estimation and mapping can be provided right after the maximum mean EVI images became available, which normally occurs in early February. Even with a delay of about 20 days which is expected to acquire the MOD13Q1 product, the soybean estimation should be released no later than early March (2nd forecast of the MCDA). Alternatively to the 2nd forecast, the soybean area at a first estimate can be provided in mid-December of each crop year based on the minimum mean EVI image of the current crop year, as described in Gusso et al., (2012). In doing so, in order to completely cover the sowing period it was used DOY 225 to 289 from MODIS/EVI. Typically, maize is sown earlier then soybean or right after soybean (MAPA, 2009), so, discrimination can be done based on EVI temporal profiles. However, when maize is sown during the same period as soybean, then it might be erroneously tagged, particularly for very well developed crop fields. Soybean will reach its maximum EVI values from DOY 0353 to 049. Soybean has significant higher values than dry maize cultivation and can be well discriminated, particularly during the period of DOY 353 to 049.

3. RESULTS AND DISCUSSION

3.1 State Area Estimates

The soybean area estimates of the MCDA (2nd forecasts) compared to CONAB and IBGE estimates, for the entire State of Mato Grosso, from 2001/02 to 2008/09, are presented in Figure 2.
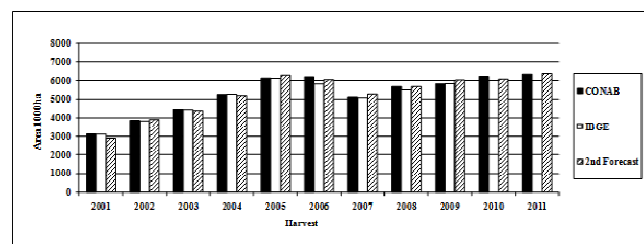
Figure 2. Comparison between MCDA (2nd forecast), IBGE and CONAB soybean area estimates for Mato Grosso State

The MCDA estimates for 2nd forecast in Mato Grosso fit quite well with the state estimates of IBGE and CONAB, with an observed maximum difference of -7.4% from IBGE, for 2001 harvest estimate. A maximum overestimate of 4.6% from CONAB was observed for 2006 harvest. Due to the spatial distribution characteristics of the classified soybean area maps, municipal statistics from IBGE are able to be compared to MCDA soybean area estimates. Soybean area was estimated by municipality from 2000/01 to 2008/09 crop years and compared to official estimates provided by IBGE (http://www.sidra.igbe.gov.br) using regression analysis. A difficult to a quick update of current harvest is that IBGE municipality data are only published around a year after the end of the soybean season.

3.2 Municipalities Area Estimates

No early estimate could be provided for the 2000/01 because this was the first crop year for which MODIS data became available. MCDA calibration for Mato Grosso demonstrated that in comparing to current official methods for soybean area estimation in Brazil, the MCDA procedure represents not only a considerable gain in time but also in spatial information which can afford agrometeorological and LUCC studies. No regression analysis was made after 2009 because no later municipal information for crop years 2009/10 and 2010/11 were released by IBGE. In an overall comparison of MCDA with municipal data from IBGE, MCDA results are slightly underestimated. The coefficients of determination (R^2) were between 0.93 in 2002 and 0.98 in 2005, indicating good agreement between the estimates. Figure 3 presents a comparing of MCDA classification and Landsat TM5 over soybean crop areas for 2010/2011 harvest.
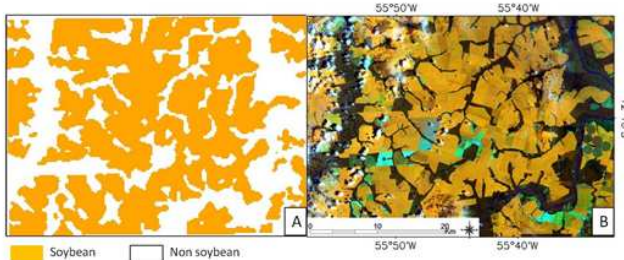
Figure 3. Sector enlargement in Mato Grosso State, for comparison between (A) soybean map classification from MCDA (2nd forecast) in 2010/2011 harvest and (B) Landsat TM5, orbit 227068, image (RGB/453) in January, 7, 2011.

For an average of the nine harvests studied, the MCDA explained 96% of the variation of the data estimated by the IBGE. The test of b0 = 0 indicated that b0 was significantly different from zero (α = 0.05) for all tested crop years. This parameter with a positive value around 4,000ha indicates that in general, municipalities with lower values of soybean area (typically < 50,000ha) are slightly overestimated.
The test of b1 = 1 indicated that b1 was significantly different from 1 (α = 0.05) for all tested crop years with most values around 0.88. This result indicates that, for municipalities with large soybean areas (typically > 200,000ha) the MCDA underestimated soybean area in relation to IBGE. The Root Mean Square Error (RMSE), for both forecasts of MCDA, in all crop years, were around 4,000ha, indicating that the MCDA estimates is consistent. In fact, fields with smaller areas are more subject to errors than those with larger areas, in agreement with Lobell and Asner (2004). Despite the moderate spatial

resolution of MODIS and the soybean field size heterogeneity, the averaged behaviour of R^2(Adjusted) = 0.95 and b1 indicates that MCDA is robust and that its physically-driven components representation is consistent.

Chang et al. (2007) compared soybean area estimates generated by MODIS and by the USDA/NASS, and observed R^2 values ranging from 0.44 to 0.94 and RMSE varying from 41,465 to 120,955ha for the entire USA country. Regression analysis for the municipal soybean estimates from the MCDA procedure using only the 2nd Forecast, and from IBGE for crop years 2000/01 to 2008/09 is presented in Figure 4.
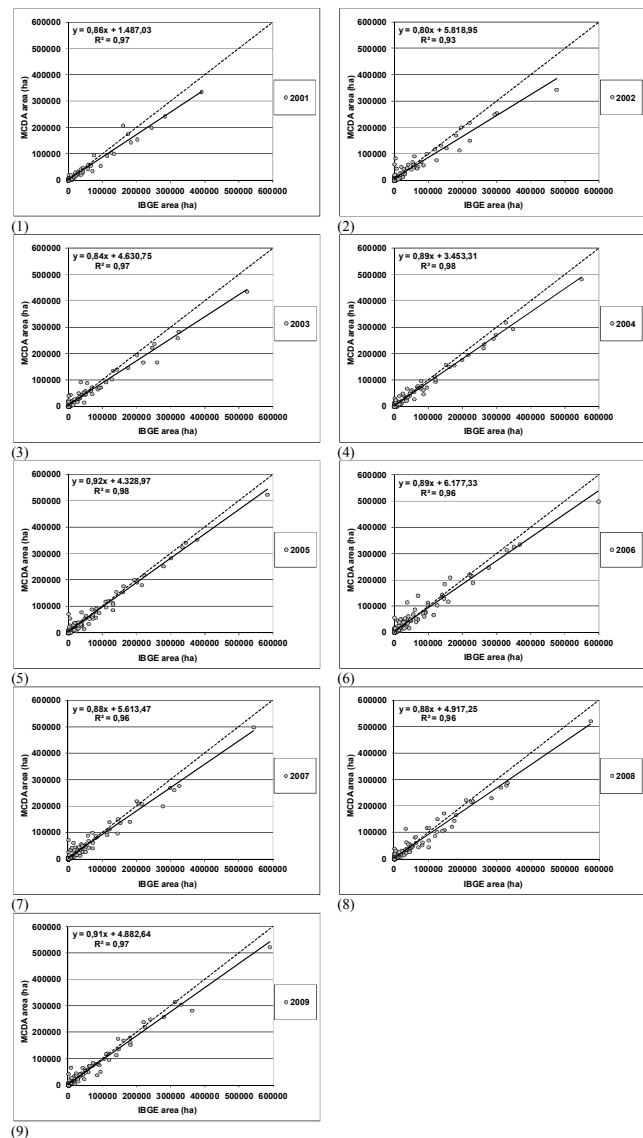
Figure 4. Regression analysis between soybean area estimates by MCDA (2nd forecast) and IBGE, for Mato Grosso state for the crop years 2000/01 to 2008/09.

It is important to emphasize that the parameters defined in the MCDA for detection crop areas are constant, as a fixed criteria, despite of soybean crop area increases and independently of the dynamics of the crop years in Mato Grosso, during the period we studied the municipal data (from crop year 2000/01 to 2008/09). Further adaptations of MCDA to other crops like maize, rice crops, cotton and sugar-cane are necessary.

4. CONCLUSION

The MCDA calibration to Mato Grosso is based on a consistent and objective approach for estimating historical soybean crop areas using MODIS/EVI data.

The MCDA methodology is able to provide annual thematic soybean maps, forecasting the crop area and so, assisting official soybean crop areas estimates, providing reliable spatial information.

5. REFERENCES

Brazilian Vegetable Oils Industry Association (ABIOVE) - Soy Moratorium Report: Mapping & Monitoring of Soy Plantings in the Amazon Biome in the Third Year. In: http://www.abiove.com.br/english/sustent/relatorio09/moratoria09_relatorio_jul10_us.pdf (accessed 13 Dez. 2011).

Chang, J., Hansen, M. C., Pittman, K., Carroll, M., & DiMiceli C., 2007. Maize and Soybean Mapping in the United States Using MODIS Time-Series Data Sets. *Agronomy Journal*, 99, p. 1654-1664.

Doraiswamy, P. C., Sinclair, T. R., Hollinger, S., Akhmedov, B., Stern, A., & Prueger, J., 2005. Application of MODIS derived parameters for regional crop yield assessment. *Remote Sensing of Environment*, 97, p. 192-202.

Epiphanio, R. D. V., Formaggio, A. R., Rudorff, B. T. F., Maeda, E. E., Luiz, A. J. B., 2010. Estimating soybean crop areas using spectral-temporal surfaces derived from MODIS images in MatoGrosso, Brazil. *Pesquisa Agropecuária Brasileira* (PAB), 45(1), p.72-80.

Gusso, A., Adami, M., Formaggio, A. R., Rizzi, R., Rudorff, B. T. F. Soybean area estimation and mapping by means MODIS/EVI data. *Pesquisa Agropecuária Brasileira*, 47, p.425-435.

Gusso, A., Adami, M., Formaggio, A. R., Rudorff, B. F. T., Rizzi, R., 2009. Análise de séries temporais EVI/MODIS para a estimativa da área cultivada com soja no Rio Grande do Sul. In: *Simpósio Brasileiro de Sensoriamento Remoto (SBSR),* Natal, RN, Brazil, INPE pp. 207-213. http://marte.dpi.inpe.br/col/dpi.inpe.br/sbsr@80/2008/11.18.01.43.47/doc/207-213.pdf (accessed 10 Nov. 2010).

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83, p. 195-213.

Instituto Brasileiro de Geografia e Estatística - IBGE (1986) Folha SH. 22, Porto Alegre, SH 21 Uruguaiana and SI 22, Lagoa Mirim. RADAMBRASIL. Rio de Janeiro: IBGE.

Instituto Brasileiro de Geografia e Estatística - IBGE (2009). SIDRA - Automatic Data Recovery System. In: http://www.sidra.ibge.gov.br/ (accessed 1 Oct. 2010).

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., Roy, D. P., & Morisette, J. T., 2002. An overview of MODIS Land data processing and product status. *Remote Sensing of Environment*, 83, p. 3-15.

Lobell, D. B., & Asner, G. P., 2004. Cropland distributions from temporal unmixing of MODIS data. *Remote Sensing of Environment*, 93, p. 412-422.

Ministério da Agricultura, Pecuária e do Abastecimento - MAPA (2009). Zoneamento Agrícola para a cultura de soja no Estado do Mato Grosso, ano-safra 2009/2010. In (p. 1): D.O.U., July, Diário Oficial da União 21/07/2009.

Moreira, M. A., Barros, M. A., Rosa, V. G. C., & Adami, M., 2007. Tecnologia de informação: imagens de satélite para o mapeamento de áreas de café de Minas Gerais. *Informe Agropecuário*, 28, 27-37.

National Aeronautics and Space Administration - NASA (2009). Warehouse Inventory Search Tool. In, https://wist.echo.nasa.gov/api/ (accessed 22 Apr. 2009).

Pittman, K.; Hansen, M. C.; Becker-Reshef, I.; Potapov, P. V. & Justice, C. O. 2010. Estimating Global Cropland Extent with Multi-year MODIS Data. *Remote Sensing*, 2, 1844-1863.

Raabus, B. M., Eineder, A. R. R., 2003. The Shuttle Radar Topography Mission - a new class of digital elevation models acquired by spaceborne radar. *Photogrammetric Engineering & Remote Sensing*, 57, p. 241-262.

Rudorff, B. F. T.; Adami, M.; Aguiar, D. A.; Moreira, M. A.; Mello, M. P.; Fabiani, L.; Amaral, D. F. & Pires, B. M., 2011. The soy moratorium in the Amazon biome monitored by remote sensing images. *Remote Sensing*, 3, 185-202.

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., Ohno, H., 2005. A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment*, 96, p. 366-374.

Sano, E. E., Ferreira, L.G., Asner, G.P., & Steinke, E.T., 2007. Spatial and temporal probabilities of obtaining cloud-free Landsat images over the Brazilian tropical savanna. *International Journal of Remote Sensing*, 28, p. 2739-2752.

Sugawara, L. M., Rudorff, B. F. T., & Adami, M., 2008. Viabilidade de uso de imagens do Landsat em mapeamento de área cultivada com soja no Estado do Paraná. *Pesquisa Agropecuária Brasileira*, 43, p. 1763-1768.

Vieira, C. A. O., 2000. Accuracy of remotely sensing classification of agricultural crops: a comparative study. p. 323. Dr. Thesis In: Physical Geography. Nottingham, England: University of Nottingham.

Wardlow, B. D., Egbert, S. L., & Kastens, J. H., 2007. Analysis of time-series MODIS 250m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sensing of Environment*, 108, p. 290-310.

Wolfe, R. E., Nishihama, M., Fleig, A. J., Kuyper, J. A., Roy, D. P., Storey, J. C., & Patt, F. S., 2002. Achieving sub-pixel geolocation accuracy in support of MODIS land science. *Remote Sensing of Environment*, 83, p. 31-49.