English Version

Mapping of the areas of soybean crop based on the spectral dynamics of the culture

Elói Lennon Dalla Nora[1]

Abstract

The objective of the present work was to evaluate, identify and map the area under soybean cultivation in the northern region of the State of Rio Grande do Sul. The study was developed based on multispectral data from the TM/Landsat-5 sensor and reference spectra of the various phases of phenological development of culture. The algorithm of supervised classification *Spectral Angle Mapper* (SAM) was applied successfully in one pre-processed TM/Landsat-5 sensor image. The procedure showed efficient capacity to identify in one period areas pertaining to one class, even under differentiated conditions of development. The classification process showed that approximately 42.66% of the area is under soybean cultivation and the SAM algorithm presents great potential to estimating the area under cultivation and the productivity of the crop.

Key-Words: Supervised classification; Remote sensing; Geoprocessing

Introduction

In the second half of the twentieth century, especially in the last three decades, soybean was a new agricultural product responsible for wide changes in Brazil's countryside. The growth of yield had been very impressive since the 1970s, which set Brazil among the world's largest producers and exporters of not only soybeans but also from some of derivatives in a shortly period of time (DALL'AGNOL et al., 2008).

The 2008/2009 crop area of soybean grown in Brazil was 21.73 million hectares, 2% above the acreage in the season before. This increase occurred primarily in major producing States such as Mato Grosso, Goiás and Paraná (CONAB, 2009). In the State of Rio Grande do Sul the cultivated area was about 3.73 million hectares with an expected total production of 7.67 million tons representing approximately 13.44% of national production (EMATER, 2009).

These figures, however, represent only a balance of cultivated areas and production values. The monitoring and prediction of soybean harvest in Brazil, as well as the other crops, is traditionally subsidized from surveys of empirical data conducted by entities linked to agricultural production. The municipal and state data collected are grouped into the entire country. Despite the great importance of these data for the economy, mostly because of the subjectivity in the evaluation by interviews determines a degree of uncertainty in the information generated (EPIPHANIO et al., 2002; IPPOLITI-RAMILO et al., 2003).

Based on these difficulties, on the last decades, many public and private institutions have directed their efforts to optimize the work of estimated crop areas using data from orbital sensors. The satellite images obtained by a constellation of remotely located sensors with different features are the basis of all the proposed methods with crop areas goals (HUETE et al., 2002; LOBELL et al., 2003; MOTTA et al., 2003).

Data from remote sensor represents a great alternative to the methods traditionally used because of low cost and speed which this information can be available to agribusiness. Though the major difficulties associated to this approach is the spatial data resolution used in some cases because they are kilometric and the changes in the agricultural calendar, besides the occasional presence of clouds (HUETE et al., 2002; KASTENS et al., 2005; XIAO et al., 2002)

This is a real dilemma since data showing an improved spatial resolution have a reduced temporal resolution discouraging the monitoring of all stages of plant development. As the agricultural calendar is different even in areas very close, many areas may be underestimated due to the low frequency in getting images. On the other hand high temporal

1 Center of Science of the Land System, National Institute of Spatial Researches. Av. dos Astronautas, 1758. Jardim da Granja, CEP 12227-010, São José dos Campos - SP.
Centro de Ciência do Sistema Terrestre, Instituto Nacional de Pesquisas Espaciais. eloidalla84@yahoo.com.br

resolution sensors currently available and with daily periods of revisits the presence of clouds can be outlined, but the maximum spatial resolution is reduced because of the altitude of the sensor. This condition has serious restrictions of application in some regions like the northern state of Rio Grande do Sul (DORAISWAMY et al., 2004).

This region, object of this study, represents the largest production area of soybean in the state. Though it is marked by small properties, many of them work as a family economy (MANTELLI, 2000). Because of the spatial heterogeneity of crop fields depending on the size of the properties, the technology used and the differences in the agricultural calendar, the classification methods used today have too much difficulty to differentiate cultures. This condition generates accuracy problems in the quantification of areas that can make these methods unfeasible or just a simple estimation.

However, the use of an automatic classifier of orbital images and improved spatial resolution capable to quantify soybean areas in different stages of development in a single period (scene) would be able to streamline this kind of study and increase the reliability of the crop areas. In this case, the high spatial resolution of images would identify crop areas under different conditions, even on small farms and low temporal resolution of the sensor would be offset by the ability of the algorithm in mapping areas at different phenological stages.

Based on this perspective, this study aimed to evaluate the performance of supervised classification *Spectral Angle Mapper* in the northern state of Rio Grande do Sul for the mapping of soybean crop. This algorithm uses target spectral data to discriminate a specific class and was subsidized with reference of all phenological development phases of soybean aiming to produce an estimate of the total cultivated area in a single pass of TM/Landsat-5 sensor.

Material and Methods

Location and description of the study area

This approach was developed using as basis the geographical area of eight counties located in the northern region of Rio Grande do Sul (Figure 1), representing a total area of approximately 410,598 hectares. This region is situated between 27° 34' and 28° 17' geographic coordinates of south latitude and 54° 03' and 52° 54' west longitude. This is an area characterized as the largest producer of soybean in the state and shows a similar pattern in the spatial distribution of physical characteristics.

In the study area the climate is classified as mesothermal and humid with weather changes at any time of the year. The annual average temperature is around 64.40 °F with amplitude about 51.80 °F. It shows an annual average of rain usually between 1.800 and 2.100 mm, distributed throughout the year (ELETROSUL, 1979).

The relief presents structural levels in basaltic rocks of the Serra Geral range formation. The basaltic rock is the largest geological unit, performing sequences of falls arranged in sub-horizontal with a thickness that ranges from 300 to 1,000 meters. In this area predominates deep oxisols, they are porous with appropriate conditions for a good deep root development (ELETROSUL, 1979).

Acquisition, processing and data analysis of TM\Landsat-5 sensor

The TM sensor image used in the process of identifying and classifying soybean crop field has been taken on Feb, 2nd 2008. During this period, on the study area the long cycle varieties that were seeded during November are at the stage of completely development covering all the soil. However, the late season varieties, seeded after the wheat or corn crop first harvest are in early stages of phenological development.

Before the methods of classification of the images of the TM sensor, the radiance data were converted to surface reflectance values, so corrected the effects of atmospheric scattering and absorption, with the help of application *Flash-Correcting Multispectral Data*, which is based on the model of radioactive transfer *MODTRAN*.

The supervised classification algorithm *Spectral Angle Mapper* is implemented in an image processing software called ENVI, version 4.5 (SULSOFT, 2009). This spectral classifier is based on the comparison of the image spectrum with a reference spectrum, from spectral libraries or final images members. The comparison is made by a

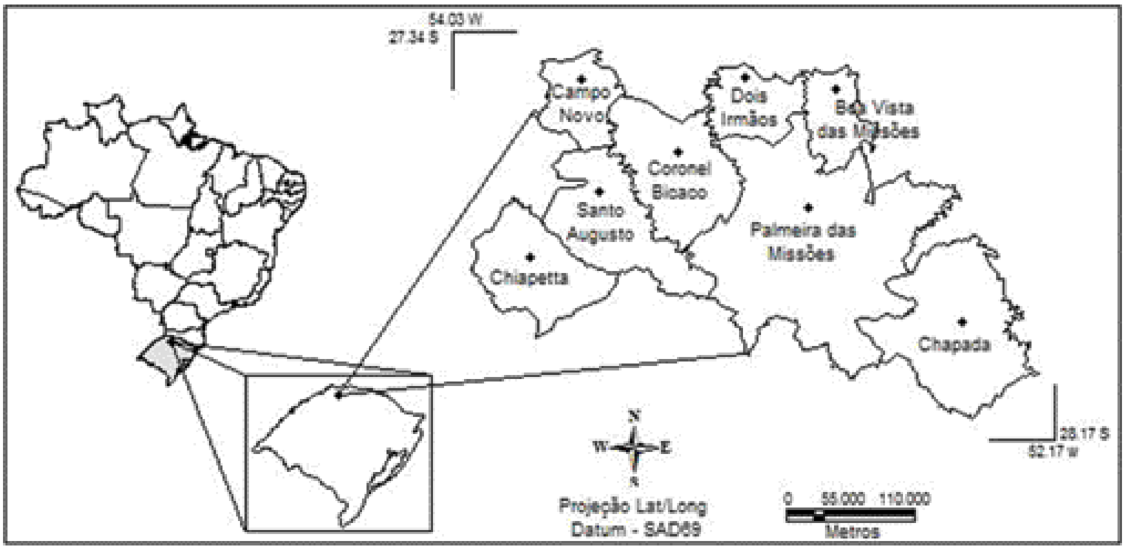

Figure 1. Study area location.

similarity criteria, thus the identification of the material is made according to the fit between the two spectral curves. As better the fit as greater the chance to exist the reference material in the image pixel. Basically they search the existence or not of the reference spectrum. The *SAM* mathematical formulation (KRUSE et al., 1993) comes from the Equation 1 and it is expressed as an angle instead of percentage.

Equation 1

$$SAM = \alpha = \cos^{-1} \frac{\sum_{i=1}^{i=nb} ER_i EI_i}{\sqrt{\sum_{i=1}^{i=nb} (EI_i)^2 \sum_{i=1}^{i=nb} (ER_i)^2}}$$

Thus we have the *SAM* value expressed in radians and as smaller the angle α is, the greater is the similarity between the waves. The angle α is determined by the arc cosine function has a range from 0^o to 90^o.

In this work we used the Boardman and Kruse (1994) methodology for the detection of reference members used in the classification process. The procedure can be divided in three stages: (a) reduction of the size spectrum (*Minimum Noise Fraction - MNF*) (CARVALHO JR et al., 2002; GREEN et al., 1988), (b) reduction of spatial size (*Purity Pixel Index - PPI*), and (c) manual identification (using a n-dimensional viewer). This approach aims to eliminate redundant factors, such as spatial and spectral, which do not change the set of feasible solutions to the reference members.

The reliability evaluation of the classification process was made by checking the mapped field works. A path of about 497 miles were made with a GPS (Global Positioning System) navigation support through which sample points were collected and tabulated on a confusion matrix to estimate the percentage of accuracy of mapping.

Results and Discussion

The heterogeneous aspect that characterizes the spatial distribution pattern of landscape elements is evidenced by the presence soybean crop areas in different stages of phenological development on TM sensor image. The spectral profile in Figure 2 illustrates this pattern where it is possible to observe the variations in spectral dynamics showed by varieties of soybean cultivation. These disparities are associated with conditions for planting and managing differentiated and more closely related to the different dates of cultivation.

You can see that the greatest disparities between the different stages of plant development are established in the region of the electromagnetic spectrum located between the red (660 nm) and near infrared (820-1165 nm). Areas in the final stages of development (maturity) have the lowest reflectance in the red area depending on the absorption of the wave length mainly by chlorophyll during photosynthesis. At this stage, the soil has completely covered the culture, reflecting the electromagnetic radiation lying completely in the near infrared region (820 nm), which the present vegetation is unable to absorb or transmit (PONZONI and SHIMABUKURO, 2007).

In the initial and intermediary phases of development one can observe gradual variations in reflectance. These variations are proportional interference soil in the spectral response of each pixel. The areas with a sparse canopy, depending on the stage of development, show increases in reflectance in the visible region and reductions in the infrared region. This feature is manifested in low availability due to chlorophyll (biomass) and spectral interference of the soil that typically has higher reflectance values than those presented by the vegetation.

The classification by spectral angle of the soybean crop areas got by the reference spectra and the *Spectral Angle Mapper* algorithm was performed with a maximum spectral angle equal to 0.08 radian. This value was selected after being tested different angles, major and minor, by being the best result in the delineation of soybean crop areas identified in the image. This means that 0.08 radian is the maximum acceptable angular value between the vector defined by the reference spectra (samples of class) and spectral vector defined by the value of the pixel being classified. Pixels with values greater than 0.08 radian were not classified and the final result of this process can be seen in Figure 3.

At the study results approximately 175,160.78 hectares are currently being used for planting soybeans. These values demonstrate the importance of culture in the regional economic development that is approximately 42.66% of the total area.

The estimated area consists in different stages of development. This characteristic enhances the capacity of the spectral angle classifier in the full monitoring of crop fields. The procedure adopted is efficient in relation to their ability in identify a single period; areas belonging to the same class even under different development conditions. Moreover, one can see the algorithm minimizes the effects of lighting since the intensity of shine does not interfere in the classification process.

The good performance of *SAM* classifier even in places with high level of spectral mixing shows as alternative to supervised and unsupervised classifiers have traditionally been used. The spectral

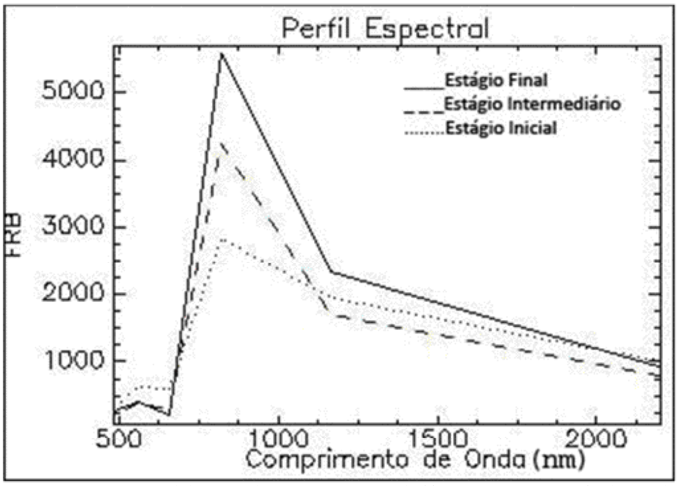

Figure 2. Spectral dynamic presented by soybean crop in different stages of phenological development in the image of TM/Landsat-5 sensor.

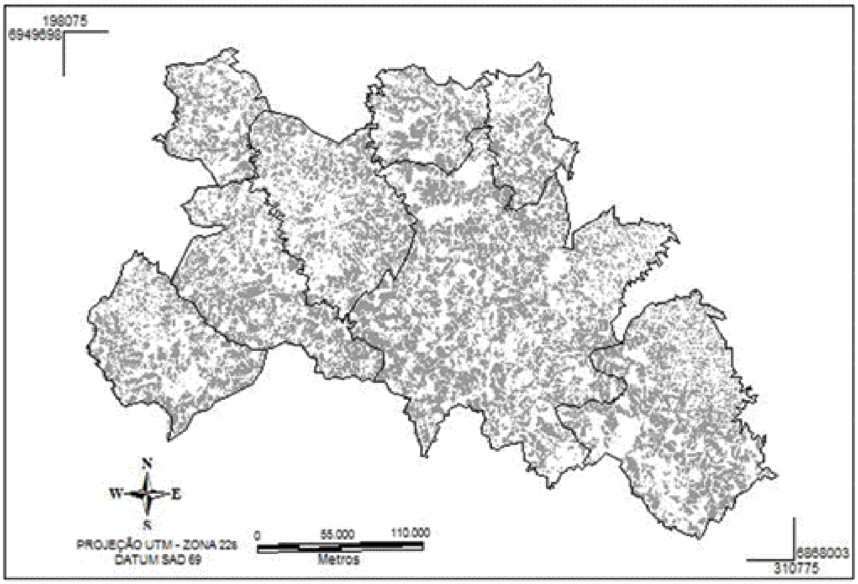

Figure 3. Spatial distribution of soybean crop areas by *Spectral Angle Mapper* algorithm.

mixing, caused by the development phase of the culture and other elements that constitute the average reflectance pixel, makes many of these classifiers able to successfully identify only cultures at specific stages of development. This condition becomes necessary many of classifications during the crop year.

According to Lamparelli et al. (2008), it was necessary 6 images of TM/ Landsat-5 sensor to estimate the soybean cultivation in the western region of Paraná State on the basis of classification algorithms "cobblestone" and "maximum likelihood" in the crop year evaluated. This condition of evaluation takes time and work for both ratings subsequent and editing of data series that causing delays in data availability and planning of agribusiness.

For this approach, the TM sensor spatial resolution allowed the definition of reference members as "purest" which minimizes the effects associated with the spectral mixture of the pixel. In this case is important to note the existence of a direct relationship between the consistency of results and quality of satellite data used besides the pre-processing adopted.

Many authors have reported difficulties in mapping of agricultural crops in southern region of Brazil with poor spatial resolution sensors due to the size of properties and topographic variations (YI et al., 2007). According to these authors, in these regions the use of moderate spatial resolution sensors is unable to provide precise estimates of crop areas according to local peculiarities. This condition was also reported by Lamparelli et al. (2008) showing lower values in the classification accuracy of soybeans using images of moderate spatial resolution and supervised classifiers.

However, the results were not produced an overall accuracy, since in some places were identified inconsistencies in the classification (Table 1.) The greater imprecision is still associated to the areas where the soil cover is very sparse and the soil spectral overlaps the characteristic spectrum of soybean in the period which the image was taken. This spectral mix produced appears as the most restrictive factor, however, in the image obtained by TM sensor, in the next step this problem could be minimized.

The validation of the proposed classification method for identifying soybean crop areas was assessed during field work in the study area during February 2009. Although field verification has been performed during crop year following the experiments, the studied area hardly showed big changes in the agricultural matrix, since historically is a major production center of soybean crop.

During the field analysis it was collected a

Table 1. Confusion matrix and global correlation obtained in the classification process

soybean crop areas		SUPERVISED CLASSIFICATION		
		Present	Absent	TOTAL
FIELD	Present	378	9	472
	Absent	12	51	28
	TOTAL	390	60	450

total of 450 sampling points distributed around the study area. The confusion matrix produced from this data and compared with values obtained in the laboratory shows an overall accuracy of 95.33% with a high correlation chart presented in Figure 3 related to their spatial distribution in the field.

Conclusion

1)The spatial/spectral resolution showed by TM/Landsat-5 sensor allows us to obtain spectral profiles able to characterize distinct stages of phenological development of soybean crop.

2) The use of the *Spectral Angle Mapper* algorithm based on reference spectra allows the identification and an effective mapping of the soybean crop areas even under different management and phenological development.

3)The spectral angle classifier, compared with the available data in the literature, is an alternative with potential gain for the techniques traditionally used in the estimation of crop areas.

4)The methodology presented is able to provide precise balance about the crop areas on time to the development of models of crop productivity and planning of agribusiness.

References

BOARDMAN, J.W.; KRUSE, F.A. Automated spectral analysis: A geologic example using AVIRIS data, north Grapevine Mountains,Nevada: **In:** ERIM - THEMATIC CONFERENCE ON GEOLOGIC REMOTE SENSING, 10, 1994, Ann Arbor. Environmental Research Institute of Michigan, Michigan: Proceedings, 1994. p.407-418.

CARVALHO JÚNIOR, O. A.; CARVALHO, A. P. F.; MENESES, P. R.; GUIMARÃES, R. F. Classificação e eliminação dos ruídos em imagens hiperespectrais pela análise seqüencial da transformação por fração de ruído mínima. **Revista Brasileira de Geofísica.** v.20, n.1, p.31-41, 2002.

COMPANHIA NACIONAL DE ABASTECIMENTO-CONAB. **Acompanhamento da safra brasileira 2008/2009.** Brasília: Ministério da Agricultura, Pecuária e Abastecimento, 2009. 42p. (Boletim Técnico, 9).

DALL'AGNOL, A.; HIRAKURI, M. H. **Realidade e perspectivas do Brasil na produção de alimentos e agroenergia, com ênfase na soja.** Londrina: Embrapa Soja, 2008.18p. (BoletimTécnico, 59).

DORAISWAMY, P.C.; HATFIELD, J.L.; JACKSON, T.J.; AKHMEDOV, B.; PRUEGER, J.; STERN, A. Crop condiction and yield simulations using Landsat and MODIS. **Remote Sensing of Environment**, v.92, n.4, p.548-59, 2004.

ELETROSUL-Centais Elétricas do Sul do Brasil-SA. **Bacia Hidrográfica do Rio Uruguai: Estudo de Inventário Hidroenergético**. CNEC. Apêndices I e II. 1979.

EMATER. Relatório de acompanhamento da safra 2008/2009 no Estado do Rio Grande do Sul. Porto Alegre: Secretaria da agricultura, pecuária, pesca e do agronegócio, 2009. 7p. (**Boletim Técnico**, 3).

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, n.2, p.187-97, 2002.

GREEN, A.A.; BERMAN, M.; SWITZER, P.; CRAIG, M.D.A transformation for ordering multispectral data in terms of images quality with implications for noise removal.IEEE **Transactions on Geoscience and Remote Sensing,** v.26, n.1, p.65-74, 1988.

HUETE, A.R.; DIDAN, K.; MIURA, T. Overview of the radiometric and biophysical performance of the MODIS vegetation index.**Remote Sensing of Environment**, v.83, n.2, p.195-213, 2002.

IPPOLITI-RAMILO, G.A.; EPIPHÂNIO, 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, n.7, p.1521-1534, 2003.

KASTENS, J.H.; KSTENS, T.L.; KSTENS, D.L.A.; PRICE, K.P.; MARTINKO, E.A.; LEE, R.Y. Image masking for crop forecasting using AVHRR NDVI time series imagery. **Remote Sensing of Enviroment,** v.99, n.3, p.341-356, 2005.

KRUSE, F.A., LEFKOFF, A.B., BOARDMAN, J.W., HEIEDBRECHT; K.B., SHAPIRO, A.T., BARLOON, P.J., GOETZ, A.F.H. The spectral image processing system (SIPS) – interactive visualization and analysis of imaging spectrometer data. **Remote Sensing of Environment**, vol. 283, n.2, p.145-163, 1993.

LAMPARELLI, R. A. C.; CARVALHO, W. M. O.; MERCANTE, E. Mapeamento de semeaduras de soja mediante dados Modis/Terra e TM/Landsat-5: Um comparativo. **Revista Engenharia Agrícola**. v.28, n.2, p.334-344, 2008.

LOBELL, D.B.; ASNER, J.I.; ORTIZ-MONASTERIO, J. I.; BENNING, T.L. Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. **Agriculture, Ecosystems and Environment**, v.94, n.2, p.205-20, 2003.

MANTELLI, J. **Evolução e tendências do setor agrário na região noroeste do Estado de Rio Grande do Sul: uma proposta de análise**. 2000. 202f. Tese (Doutorado em Geografia Humana) - Faculdade de Filosofia Letras e Ciências Humanas, Departamento de Geografia, Universidade de São Paulo, São Paulo, 2000.

MOTTA, J. L. G.; FONTANA, D. C; WEBER, E. Evolução temporal do NDVI/NOAA em áreas cobertas por pixels com proporções variáveis de soja. **Revista Brasileira de Agrometeorologia**. v.11, n.2, p.353-369, 2003.

PONZONI, F.J.; SHIMABUKURO, Y.E. **Sensoriamento Remoto no Estudo da Vegetação**. 1.ed. São José dos Campos: Parêntese, 2007. 128p.

SULSOFT PROCESSAMENTO DE DADOS. **Guia do ENVI em português**. Versão 4.5. PortoAlegre. Disponível em: <http://www.envi.com.br>. Acesso em: 15 jun. 2009.

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 imagery. **International Journal of Remote Sensing**, v.23, n.18, p.3579-3594, 2002.

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. **Revista Engenharia Agrícola**. v.27, n.3, p.753-763, 2007.