

MAPPING OF SUMMER CROPS IN THE STATE OF PARANÁ, BRAZIL, THROUGH THE 10-DAY SPOT VEGETATION NDVI COMPOSITES

GLEYCE K. D. ARAÚJO¹, JANSLE V. ROCHA², RUBENS A. C. LAMPARELLI³,
AGMON M. ROCHA⁴

ABSTRACT: The search for low subjectivity area estimates has increased the use of remote sensing for agricultural monitoring and crop yield prediction, leading to more flexibility in data acquisition and lower costs comparing to traditional methods such as census and surveys. Low spatial resolution satellite images with higher frequency in image acquisition have shown to be adequate for cropland mapping and monitoring in large areas. The main goal of this study was to map the Summer crops in the State of Paraná, Brazil, using 10-day composition of NDVI SPOT Vegetation data for 2005/2006, 2006/2007 and 2007/2008 cropping seasons. For this, a supervised digital classification method with Parallelepiped algorithm in multitemporal RGB image composites was used, in order to generate masks of Summer cultures for each 10-day composition. Accuracy assessment was performed using Kappa index, overall accuracy and Willmott's concordance index, resulting in good levels of accuracy. This methodology allowed the accomplishment, with free and low resolution data, of the mapping of Summer cultures at State level.

KEYWORDS: remote sensing, time series, crop yield estimates.

MAPEAMENTO DE CULTURAS DE VERÃO NO ESTADO DO PARANÁ POR MEIO DE COMPOSIÇÕES DECENIAIS DE NDVI DO SENSOR SPOT VEGETATION

RESUMO: A busca por menor subjetividade em estimativas de área tem aumentado a utilização do sensoriamento remoto para monitoramento agrícola e previsão de safras, pois proporciona maior agilidade na aquisição de dados e menor custo em relação a métodos tradicionais de censos e pesquisas. Imagens de satélite de baixa resolução espacial e alta periodicidade são adequadas para o mapeamento, acompanhamento e desenvolvimento de culturas em áreas extensas. O objetivo deste trabalho foi mapear culturas de verão no Estado do Paraná por meio de composições deceniais de imagens NDVI do satélite SPOT *Vegetation* para safras de 2005/2006, 2006/2007 e 2007/2008. Para isso, foi utilizado um método de classificação digital supervisionada com algoritmo Paralelepípedo em composições RGB multitemporais das imagens, de forma a gerar máscaras das culturas de verão para cada composição decenal. A verificação da acurácia das máscaras foi realizada utilizando índice Kappa, exatidão global e índice de concordância de Willmott, resultando em bons índices de acerto. Essa metodologia possibilitou realizar, com dados gratuitos e de baixa resolução espacial, o mapeamento de culturas de verão em nível estadual.

PALAVRAS-CHAVE: sensoriamento remoto, séries temporais, previsão de safras.

INTRODUCTION

According to the Food and Agriculture Organization of the United Nations (FAO, 2009), Brazil is one of the largest producers of soybean and corn. According to the survey carried out by the National Company of Supplying (CONAB - Companhia Nacional de Abastecimento, 2010), the estimated planted area in crop 2009/2010 for the soybean crop is 23.06 million hectares, equivalent

¹ Mestre em Engenharia Agrícola, Doutoranda, FEAGRI/UNICAMP.

² Eng^o Agrícola, Professor, FEAGRI/UNICAMP.

³ Eng^o Agrícola, Pesquisador, CEPAGRI/UNICAMP.

⁴ Profissional especializado em Geoprocessamento - FEAGRI/UNICAMP.

Recebido pelo Conselho Editorial em: 9-4-2010

Aprovado pelo Conselho Editorial em: 18-4-2011

to an increase of 6.7% in the previous crop. For the corn crop is estimated to be 8,159.7 million hectares of planted area, with a decrease of 12% area regarding the last crop.

Given the importance of both grains in Brazilian and international scenario, it is essential to have a crop prediction system able to state accurate figures on production aimed especially to avoid losses as well as the search for new technologies aimed at reducing risk.

The agricultural monitoring using remote sensing holds the advantage over traditional methods of crop prediction, such as censuses or surveys at field. According to SANO et al. (1998) this kind of product provides updated information and a relatively low cost.

Satellite images provide important data for crop prediction models, because through a series of images, you can track the development of a culture in wide areas. According to ESQUERDO (2007), the existence of spectral information generated from sensors with high temporal resolution allows evaluating the spectral behavior of the cultures over all their phenological stages.

According to BRITO & SAGE (2005) the knowledge of land use by mapping a specific region is important due to the understanding and organization of space and the changes that have occurred there.

Vegetation indexes are also used for agricultural cultures identification and mapping (ROCHA, 2006; YI et al., 2007; LAMPARELLI et al., 2008; FERNANDES, 2009). The Normalized Difference Vegetation Index (NDVI), proposed by ROUSE et al. (1973), is based on the ratio of the reflectance at a wavelength equivalent to red, and near infrared, such as eq.(1). According to PONZONI & SHIMABUKURO (2009) the NDVI is greatly used to build seasonal and temporal profiles of vegetation activities, which admits inter-annual comparisons between them, and also allows detecting phenological activities, length of growing season, vegetative peak, physiological changes of leaf, and period senescence.

$$NDVI = \frac{(NIR - R)}{(NIR + R)} \quad (1)$$

where,

NIR - reflectance in the near infrared band, and
R - reflectance in red band.

The choice of sensor is also crucial, especially when it comes to agricultural tracking of short cycle cultures, due to its fast vegetative growth. Sensors on orbiting platforms such as *Système Pour l'Observation de la Terre (SPOT) Vegetation*, which has low spatial resolution and high temporal resolution, facilitate the tracking of vegetation cover, allowing the generation of vegetation indexes regularly, encouraging analysis and agricultural tracking. Table 1 shows the main characteristics of the sensor.

TABLE 1. Main characteristics of vegetation's sensor.

Band	Spectral Resolution (μm)	Spatial Resolution	Temporal Resolution	Imaged Range	Orbit
Blue	0.43 – 0.47				822 km,
Red	0.61 – 0.68				Synchronous with the sun,
Near infrared	0.78 – 0.89	1 km	1 day	2,250 km	98.7 ° Inclination, Equatorial Crossing 10:30
Short-wave infrared	1.58 – 1.75				

Source: Adaptated from JENSEN (2009).

The accuracy of a thematic map shows the proximity of a particular measure to its real value, therefore, the reliability of a map is linked to its accuracy. In this context, it is necessary to perform a statistical procedure on the product of a classification of digital images to determine the accuracy or exactness of this classification (BERNARDES, 2006).

This paper aims to map the Summer culture in the state of Paraná (Brazil's State), using ten-day period compositions of NDVI of the SPOT Vegetation sensor for 2005/2006, 2006/2007, and 2007/2008 seasons.

MATERIALS E METHODS

The area under study was the state of Paraná, located in southern Brazil, one of the major producers of soybeans in the country. The Summer crops seasons 2005/2006, 2006/2007, and 2007/2008 were studied.

Compositions of ten-day period of NDVI of the SPOT Vegetation sensor was used, (VGT product - S10) which has a spatial resolution of 1 km and temporal resolution of ten days. The images of this sensor already have atmospheric corrections; however, the radiometric resolution of the Vegetation sensor is 8 bits in a range of gray level (NC). The conversion to the real value of NDVI was made by eq.(2), available at VITO (2009).

$$\text{NDVI} = 0.004 \text{ NC} + (-0.08) \quad (2)$$

Where,

NC - 8-bit digital value.

To determine the period of interest, the average profile of NDVI of the Summer culture in the state, corresponding to a period of nine years (1999 to 2007, during which there is availability of images of the sensor used) was drawn. Thus, it was possible to identify the average date of sowing of the Summer culture over the years. Referring to Figure 1, it is possible to note that the Summer culture from 1999 to 2002 was initiated in September. Since 2003, there was an advance in sowing period for the month of October. So, it was possible to choose the best season for the acquisition of images, including the entire vegetative cycle for the 2005/2006, 2006/2007, and 2007/2008 crops.

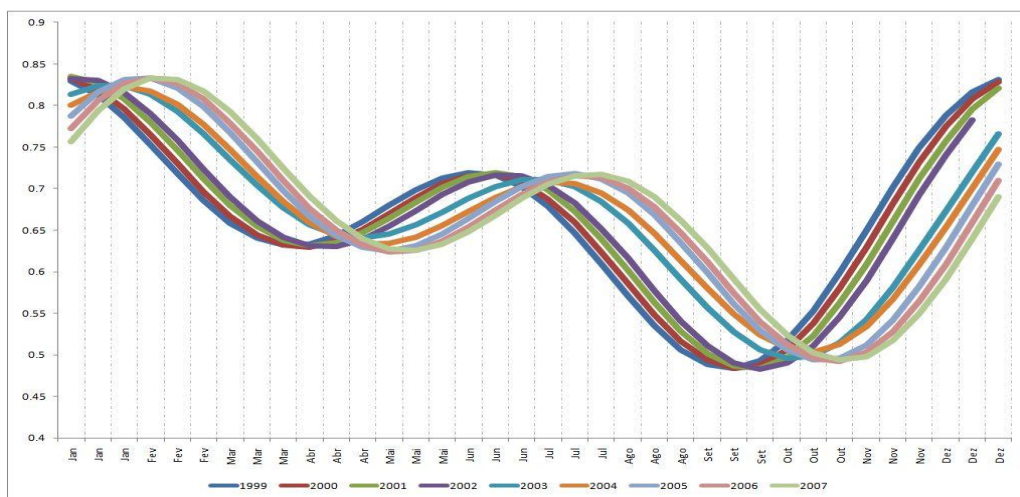
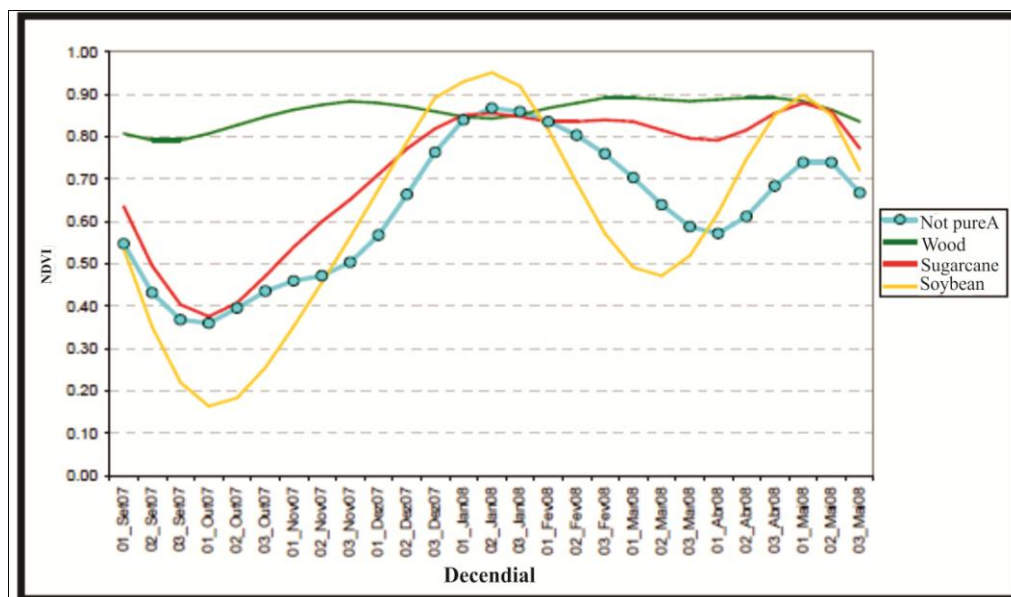


FIGURE 1. Average profile of NDVI for the state of Paraná.

Before starting the mapping, the analysis of the behavior of NDVI throughout its vegetative cycle was carried out. This analysis was necessary to obtain the behavior of the NDVI of the cultures in question, in view of the low spatial resolution of sensor used, since there may be other targets along with the Summer culture in a single pixel, influencing its spectral behavior over profile.

GENOVESE et al. (2001) integrated data of NDVI from NOAA-AVHRR satellite (spatial resolution of $4.4 \text{ km}^2 \times 4.4 \text{ km}^2$) with a land use map generated from the CORINE ("Co-ordination of Information on the Environment" program mapping land use in Europe, coordinated by the European Commission), with spatial resolution of $100 \text{ m} \times 100 \text{ m}$, to reduce problems with non-pure pixels. Thus, the CNDVI (CORINE-NDVI) was created, which is based on the proportion of the targets of the map. According to JOHANN et al. (2009), the study of the behavior of different targets is relevant to remote sensing, because from this knowledge it is possible to build masks from these targets based on NDVI values.

The authors also stated that the mixture of different targets (non-pure pixels) and in different proportions can influence the final behavior of the profile of the pixel. This can often lead to difficulties in generating masks of different agricultural cultures, since the mixture of targets can mischaracterize, in some cases, the expected behavior of a particular pixel. Thus, this pixel is not selected in the generation of masks. Figure 2 illustrates a pixel with the influence of other targets along its spectral profile (blue curve 'Não_puroA') and profiles of pixels without influence of targets.



Source: Adaptated from JOHANN et al., 2009.

FIGURE 2. Influence of other targets in a pixel.

Due to the non-uniform distribution of rainfall and different cycles, the Summer culture is sown at different times within the state. Thus, it is more effort to map the culture in order to include the entire season, opting to build masks for 10-day period for the state of Paraná.

Based on this, multi-temporal compositions were created in RGB color, composed of three images every 10-day period, since the beginning of the sowing until the period covered by the peak of the vegetative cycle of the Summer culture. In the G channel, the images were placed where the NDVI profile pointed to the beginning of the sowing; in the B channel, the images were placed after the beginning of the sowing.

From the image of these two channels, the third image was placed in the R channel, where it should be the image with the highest NDVI during the season, that is, the vegetative peak of the culture to that profile. Thus, only Summer culture gained prominence in the composition. Figure 3 shows the schemes for composition of images in RGB. A total of eight compositions were created for each crop.

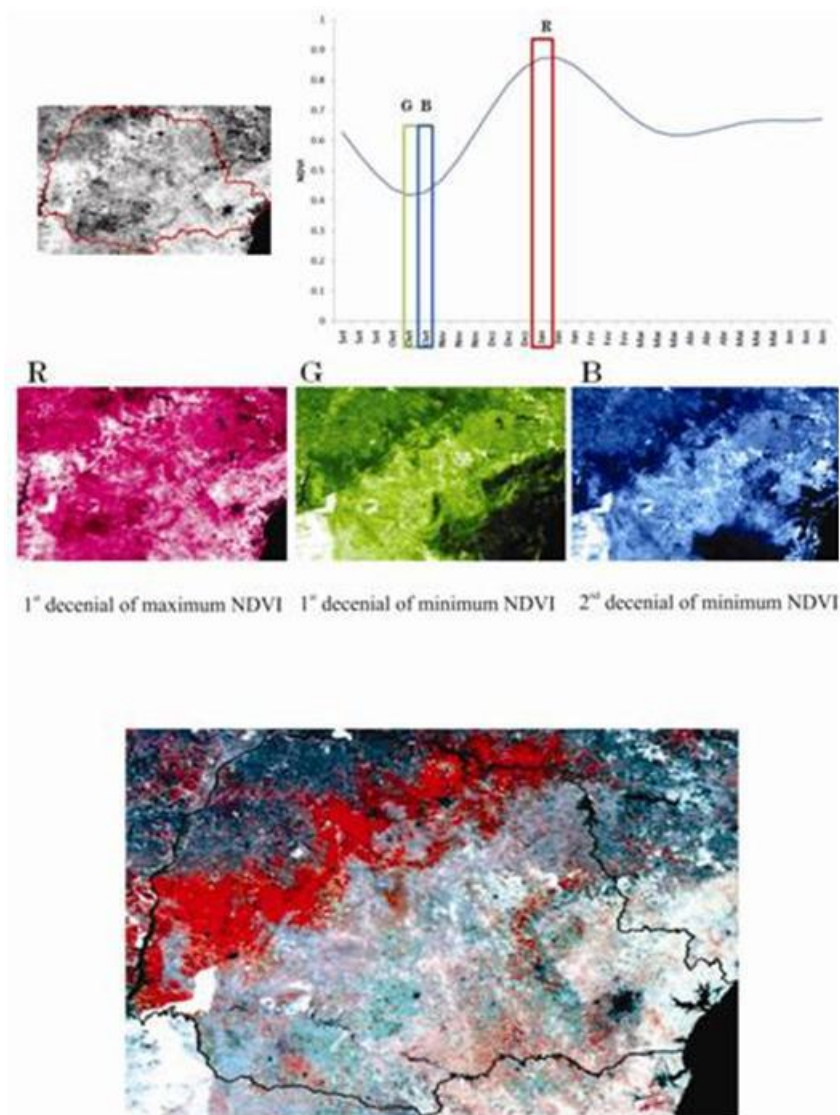


FIGURE 3. Scheme for composition of images in RGB.

For mapping the Summer culture in the compositions in RGB, we used the supervised digital classification method called Parallelepiped. This classifier is deterministic (MATHER, 1999), which identifies the minimum and maximum values of the histograms of the samples of the classes in each band, which are used to define the vertices of a parallelepiped in the space of attributes. Hence, the pixel whose gray level value is within this parallelepiped will be classified as belonging to a particular class in each spectral band selected for classification.

In this procedure, it was necessary to select areas of training, in this case the highest vegetative value of the Summer culture within each composition. The algorithm made use of these regions to create an image of supervised classification with two classes: 'Summer Culture' and 'Other Cultures'.

Then, we selected only the class 'Summer culture', which was converted to a vector file, making it possible to create a mask of culture for each composition. All the processes of classification and creation of the mask were made through the software *Envi 4.5*. In order to obtain the final mask of each season, an overlap of 10-day period masks was made.

This operation allowed the production of a new plan. In this case, we performed a cumulative sum of the images using the *Overlay* option of the software *Idrisi Kilimanjaro*. Figure 4 presents the main steps of culture mapping.

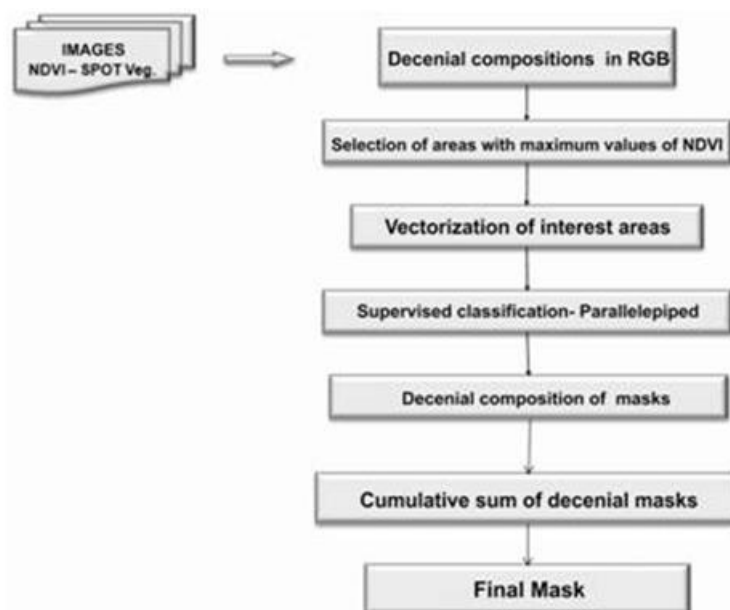


FIGURE 4. Stages of construction of the mask.

Before starting the process of evaluating the accuracy of the masks, making a mosaic with images of the Landsat 5/TM sensor in the period in question for the state of Paraná was necessary, for comparison purposes between the masks generated by the classifier. In the evaluation of accuracy, 400 sampling points were distributed according to the stratified sampling technique, within the State. From these points, we constructed a contingency matrix between the reference (mosaic of images Landsat 5/TM for the State), and supervised classification (mask of culture).

The verification of the accuracy of the masks generated by digital classification was made by obtaining matrices of errors, calculating, thereby, the rate of overall accuracy (Eq. 3) and Kappa (Eq. 4) of the results. According to CONGALTON (1991), a matrix of errors may be used as a starting point for applying a series of statistical techniques, both descriptive and analytical. Table 2 shows the evaluation of the quality of the proposed classification by LANDIS & KOCH (1977).

For checking the concordance of the mask, the Willmott's concordance index (d) was also used, given by a mathematical approximation that evaluates the accuracy and the separation of the estimated values in relation to those observed (WILLMOTT, 1981), where the variation is from zero (no concordance) to one (perfect concordance). For extraction of the planted area of the masks, we used an automatic process of extracting data from images through a system of programming routines made in a language called IDL (Interactive Data Language) in *Envi* based on the study of ESQUERDO et al. (2006).

The Willmott's concordance index is calculated by Eq. 5, where the data of planted area of the mask were compared with official data from the IBGE (Brazilian Institute of Geography and Statistics).

TABLE 2. Quality of classification associated with values of kappa.

Kappa index	Quality
0	Very bad
0.01 a 0.20	Bad
0.01 a 0.40	Reasonable
0.41 a 0.60	Good
0.61 a 0.80	Very good
0.81 a 1.00	Excellent

Source: LANDIS & KOCH (1977).

$$EG = \frac{A}{n} 100 \tag{3}$$

Where,

A - general hit (sample points with hits), and
 n - number of sampling points.

$$K = \frac{\left[n * \sum_{i=1}^r x_{ij} - \sum_{i=1}^r (x_{i.} * x_{.j}) \right]}{\left[n^2 - \sum_{i=1}^r (x_{i.} * x_{.j}) \right]} \tag{4}$$

where,

n - number of observations;
 r - number of rows of the matrix of error;
 x_{ij} - observation in row i and column i;
 x_{i.} - marginal total of the row I, and
 x_{.j} - marginal total of the column j;

$$d = 1 - \frac{\sum_{i=1}^n (p_i - o_i)}{\sum_{i=1}^n (|p_i - \bar{o}| + |o_i - \bar{o}|)^2} \tag{5}$$

where,

p_i - estimated or predicted value;
 o_i - observed or measured value;
 n - number of observations;
 \bar{o} - arithmetic mean of observed values, and
 \bar{p} - arithmetic mean of the predicted values.

RESULTS AND DISCUSSION

The result of the applied methodology for each 10-day period is shown in Figure 5, where it is possible to observe the evolution of planting throughout the state over the 10-day periods.

After the performed processing on the images, and from the supervised classification and overlap of the several dates of the images in each of the monitored crops, masks of Summer culture were built, as shown in Figures 6, 7, and 8 (seasons 2005/2006, 2006/2007, and 2007/2008, respectively).

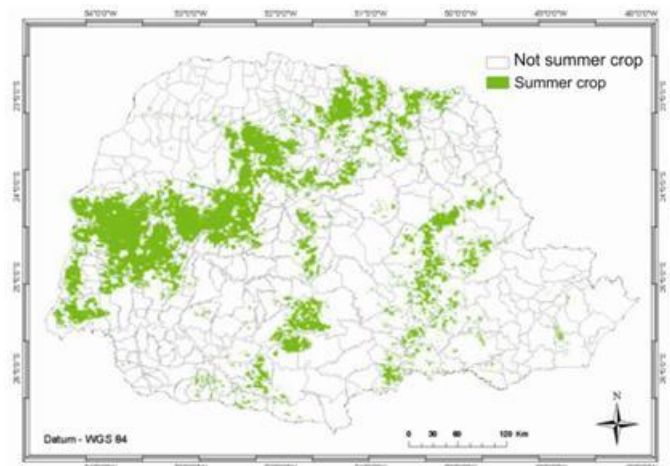
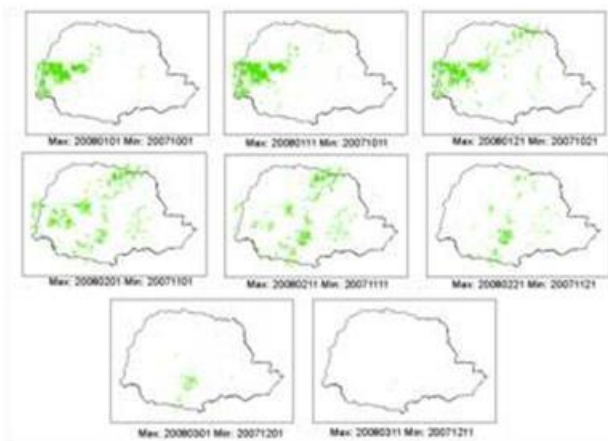


FIGURE 5. Masks decadal of the summer crop in the state of Paraná.

FIGURE 6. Final mask for the 2005/2006 season.

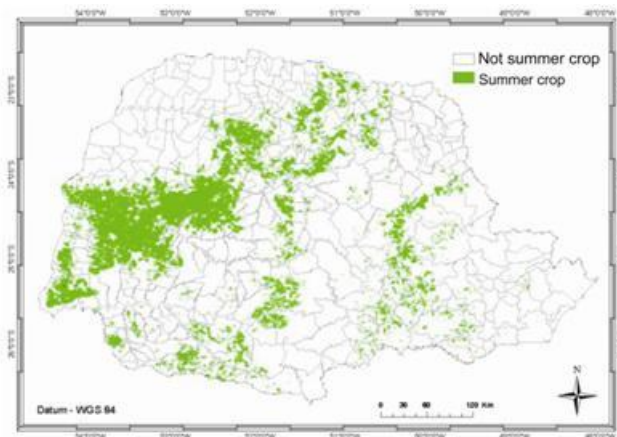


FIGURE 7. Final mask for the 2006/2007 season.

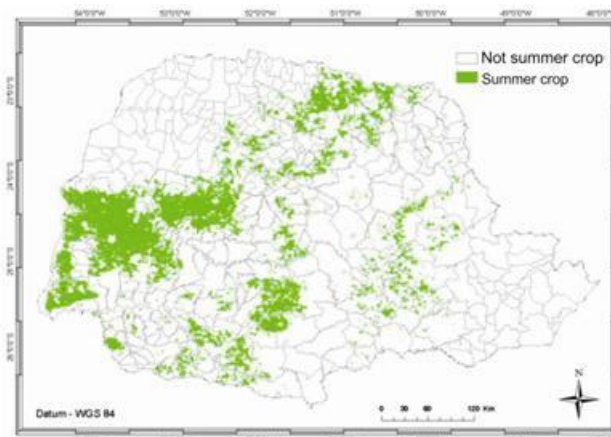


FIGURE 8. Final mask for the 2007/2008 season.

Out of the total of 400 samples used for the construction of the matrix of errors, analyzing the sum of main diagonal of the matrix of errors, it was observed that crops in 2005/2006, 2006/2007, and 2007/2008, respectively, 363, 364 and 372 samples were correctly classified. Based on these figures, we obtained overall accuracy indexes (EG) of 0.91, 0.91 and 0.93. That is, in these three crops, 91%, 91% and 93% of the samples were correctly identified by the classifier when compared with the Landsat 5/TM images, showing good reliability of the mapping compared to the reference of surface.

The index Kappa (K) was determined to assess the accuracy of the data, allowing viewing of the results of a thematic classification in relation to Landsat 5/TM images. The result indicated a Kappa coefficient of 0.70, 0.75 and 0.77 for the 3 studied crops. The Kappa index refers to a method for assessing concordance or discordance between two situations, in this way, there was a concordance of 70%, 75% and 77% of the samples in the area of interest, therefore, we may say that the classification in the studied area approached of the reference of surface.

The Willmott's concordance index (d) measured the dispersion of data regarding the line 1:1, that is, we measured the accuracy of estimated values (mask of Summer culture) in relation to the observed (official data from IBGE). Thereby, we observe values of 'd' 0.93 for the 2005/2006 crop, 0.93 for the 2006/2007 crop, and 0.99 for the 2007/2008 crop.

According to the parameters required by LANDIS and KOCH (1977) the result of the accuracy of the masks was considered "very good". The values of concordance index 'd' were close to 1, that is, there was high concordance between the values estimated by the mask in relation to official data from IBGE. On this basis, these masks may be used to support the identification of Summer cultures in the state of Paraná. Tables 3, 4, and 5 show the matrix of errors for each crop.

TABLE 3. Confusion matrix for the 2005/2006 season.

Reference Mask	Reference Image		Grand Total	Error of Inclusion
	No Summer Culture	Summer Culture		
No Summer Culture	241	7	248	2.80%
Summer Culture	30	122	152	19.7%
Grand total	271	129	400	-
Error of omission	11.1%	5.4%	-	-

TABLE 4. Confusion matrix for the 2006/2007 season.

Reference Mask	Reference Image		Grand Total	Error of Inclusion
	No Summer Culture	Summer Culture		
No Summer Culture	243	11	254	4.30%
Summer Culture	25	121	146	17.10%
Grand total	268	132	400	-
Error of omission	9.30%	8.30%	-	-

TABLE 5. Confusion matrix for the 2007/08 season.

Reference Mask	Reference Image		Grand Total	Error of Inclusion
	No Summer Culture	Summer Culture		
No Summer Culture	159	25	184	13.60%
Summer Culture	3	213	216	1.40%
Grand total	162	238	400	-
Error of omission	1.90%	10.50%	-	-

It is worth to accentuate that the SPOT Vegetation sensor was chosen for mapping in this paper due to its high temporal resolution, which is extremely important in agriculture, especially when it comes to short phenological cycle crops.

This methodology may be applied at operational level and low cost in several regions, in view of the temporal resolution of the Vegetation sensor, and especially because it is available for free via the Internet.

CONCLUSIONS

The use of 10-day period compositions in RGB shown to be effective for mapping of summer cultures. The assessment of accuracy for the three masks proved to be accurate, good levels of accuracy in the interpretation of mapping in relation to the data found in the reference image were discovered. In the concordance between the observed data in relation to the IBGE data estimated by the masks, the results were very close to 1, that is, the masks are close to the reality of the surface.

The method used allowed to identify important areas of occupation of the Summer culture in the state. Also, note that there are possible differences in season planting of the crop in some regions of the state.

REFERENCES

- BERNARDES, T. *Caracterização do ambiente agrícola do Complexo Serra Negra por meio de sensoriamento remoto e sistemas de informação geográfica*. 2006, 119 f. Dissertação (Mestrado) - Universidade Federal de Lavras, 2006. Disponível em: <<http://www.epamig.br/geosolos/MaterialSite/Teses/Tiago/Tiagodissertacao.pdf>>. Acesso em: 8 abr. 2010.
- BRITO, J.L.S.; PRUDENTE, T.D. Análise temporal de uso do solo cobertura vegetal do município de Uberlândia- MG, utilizando imagens ETM+ Landsat 7. *Sociedade e Natureza*, Uberlândia, v.17, p.37-46, 2005.
- CONAB. *Companhia Nacional de Abastecimento*. 2009 <http://www.conab.gov.br/conabweb/download/safra/4graos_07.01.10.pdf>. Acesso em: 27 mar. 2010.
- CONGALTON, R.G. A review of assessing the accuracy classifications of remotely sensed data. *Remote Sensing of Environment*, New York, v.37, p.35-46, 1991
- ESQUERDO, J.C.D.M.; ANTUNES, J.F.G.; BALDWIN, D.G.; EMERY, W.J.; ZULLO JÚNIOR, J. An automatic system for AVHRR land surface product generation. *International Journal of Remote Sensing*, Basingstoke, v.27, p.3.925-3.942, 2006.

- ESQUERDO, J.C.D.M. *Utilização de perfis multi-temporais do NDVI/AVHRR no acompanhamento da safra de soja no oeste do Paraná*. 2007. 186 f. Tese (Doutorado em Engenharia Agrícola) - Universidade Estadual de Campinas, Campinas, 2007.
- FAO. FOOD AND AGRICULTURE ORGANIZATION. *Faostat*. Dados estatísticos sobre a produção mundial da soja. Disponível em: <<http://faostat.fao.org/site/567/default.aspx#ancor>>. Acesso em: 21 jan. 2009.
- FERNANDES, J.L. *Monitoramento da cultura de cana-de-açúcar no Estado de São Paulo por meio de imagens SPOT Vegetation e dados meteorológicos*. 2009. 114 f. Dissertação (Mestrado em Engenharia Agrícola) - Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas, 2009.
- GENOVESE, G.; VIGNOLLES, C.; NÉGRE, T.; PASSERA, G. A methodology for a combined use of normalised difference vegetation index and CORINE land cover data for crop yield monitoring and forecasting: a case study on Spain. *Agronomie*, Paris, v.21, p.91-111, 2001.
- IBGE. INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *SIDRA, banco de dados área plantada municipal*. Disponível em: <<http://www.sidra.ibge.gov.br/bda/tabela/listabl.asp?c=1612&z=t&o=11>>. Acesso: 10 jan. 2009.
- JENSEN, J.R. *Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres*. Tradução de J.C.N. Ephiphany. São José dos Campos: Parêntese, 2009. 604 p.
- JOHANN, J.A.; ARAÚJO, G.K.D.; ROCHA, J.V. Avaliação do perfil temporal de NDVI decendial do sensor SPOT Vegetation em pixels “puros” e “não puros” derivados de mapa de uso da terra realizado a partir do sensor Landsat-5/TM. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 14., 2009, Natal. *Anais...* São José dos Campos: INPE, 2009. p.223-225.
- LAMPARELLI, R.A.C.; CARVALHO, W.M.O. de; MERCANTE, E. Mapeamento de semeaduras de soja (*Glycine max* (L.)Merr.) mediante dados MODIS/Terra E TM/Landsat 5: um comparativo. *Engenharia Agrícola*, Jaboticabal, v.28, n.2, p.334-344, 2008.
- LANDIS, J.R.; KOCH, G.G. The measurement of observer agreement for categorical data. *Biometrics*, Washington, v.33, n.1, p.159-174, 1977.
- MATHER, P.M. *Computer processing of remotely-sensed images: An introduction*. 2nd ed. Chichester: John Wiley & Sons, 1999. 292 p.
- PONZONI F.J.; SHIMABUKURO, Y.E. *Sensoriamento Remoto no estudo da vegetação*. São José dos Campos: Parêntese, 2009. 136 p.
- ROCHA, J.V. *Crop monitoring in South America (MERCOSUR + Bolivia)*. Ispra: JRC, 2006. (Bulletin, 7).
- ROUSE, J.W.; HAAS, R.H.; SCHELL, J.A.; DEERING, D.W. Monitoring vegetation systems in the Great Plains with ERTS. In: Earth RESOURCES TECHNOLOGY SATELLITE-1 SYMPOSIUM, 3., 1973, Washington. *Proceedings...* Washington: NASA, 1974. v.1, p.309-317.
- SANO, E.E; ASSAD, E.D.; ORIOLI, A.L. Monitoramento da Ocupação Agrícola. In: *Sistema de informações geográficas: aplicações na agricultura*. Brasília: EMBRAPA - SPI/EMBRAPA Cerrados, 1998. p. 179-190.
- VITO – PRODUTO VGT-S10. Disponível em: <<http://www.vgt.vito.be/faq/faq.html>>. Acesso em: 29 jan. 2009.
- WILLMOTT, C.J. On the validation of models. *Physical Geography*, Norwich, v.2, p.184-194, 1981.

YI, J.L.R.; SHIMABUKURO, Y.E.; QUINTANILHA, J. A. Identificação e mapeamento de áreas de milho na região sul do Brasil utilizando imagens MODIS. *Engenharia Agrícola*, Jaboticabal, v.27, n.3, p.753-763, 2007.