Proceedings of 8 ICSHMO, Foz do Iguaçu, Brazil, April 24-28, 2006, INPE, p. 861-868.

DETECTION OF TROPICAL SAVANNAH (CERRADO) PHYSIOGNOMIES IN THE LEGAL AMAZON BY THE APPLICATION OF THE VEGETATION AND MOISTURE INDICES WITH MODIS TIME SERIES DATA

Jorge Alberto Bustamante Becerra* and Regina Célia dos Santos Alvalá
Centro de Previsão de Tempo e Estudos Climáticos, Instituto Nacional de Pesquisas Espaciais (CPTEC/INPE)

1 INTRODUCTION

The Cerrado biome is the predominant natural vegetation of central Brazil and, before the recent extensive destruction for agricultural use, it covered approximately 2 X 10^6 Km^2, representing approximately 23% of the land surface of the country. This biome (Figure 1) is very important, since it represents the largest region of neo-tropical savanna in the world and has the second largest area of all biomes in South America (Oliveira & Marquis, 2002). The cerrado formations, which are extremely rich in biodiversity, consist of three major vertical domains: semideciduous woodland, shrubland, and grassland (Furley, 1996).

The climate of the cerrado biome, in general, is the transition of the Koppen's Savanna (Aw) and the Moonson (AM) subtypes of the Tropical Rain Climate (A). The average rainfall is 1500 mm (< 800 mm in NE to >2000 mm in NW extremes) with an average temperature range from 20 to 25ºC, and an annual evapotranspiration deficit (Thornthwaite's method) of 4-491 mm (Lopez and Cox, 1977). Approximately 90% of the rainfall is concentrated during the months of October to March, which results in two contrasting seasons: dry and wet.

Most of the cerrado region, in the last 50 years, has been converted especially for agriculture and pasture purposes (Nepstad et al., 1997; Mantovani and Pereira, 1998; Sano et al., 2001). This conversion has impacted the biological diversity, the hydrological cycle, the energy balance, the climate and the carbon dynamics at local and regional scales. Despite this high rate conversion, there have been few attempts in quantifying and monitoring the land-use changes of this threatened biome. Thus, the use of remote sensing is one possible method to monitor the vegetation conditions in the cerrado environment, as well as the identification of the current distribution of lad cover types and occurrences of land-use intensification and degradation.

On December 18, 1999, the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Terra platform, was launched as part of NASA's Earth Observing System (EOS). The MODIS vegetation index (VI) products, generated for improved vegetation monitoring , (Huete et al., 1994; Liu and Huete, 1995; justice et al., 1998) include the 16-day normalized difference vegetation index (NDVI) (Rousse et al., 1974), produced to extend the AVHRR-NDVI time series and a new index, the enhanced vegetation index (EVI) (Huete and Justice, 1999), which is resistant to atmospheric and canopy background effects,. MODIS also provides a set of short-wave infrared bands which allow the land surface water index (LSWI) to be computed as a moisture index.

Vegetation (NDVI and EVI) and moisture (LSWI) indices are spectral transformations, which were designed to enhance and quantify the "green" photosynthetic and vegetation water content (VWC) signal. These spectral indices are very useful for monitoring photosynthetic activity, VWC and seasonal and inter-annual vegetation behavior. These indices also have been used for land cover classification and change detection (Townshend et al., 1991). They are useful in deriving biophysical vegetation properties, such as fractional vegetation cover, leaf area index (LAI), fraction of absorbed photosynthetic active radiation (fPAR), biomass (Asrar et al., 1984; Sellers, 1985), and net primary production (NPP) (Tucker and Sellers, 1986; Running and Nemani, 1988).

An approach to characterize land cover types from multitemporal data set is the use of principal component analysis (PCA). PCA has been successfully employed in remote sensing for image data transformation, information compression, and change detection analysis. This multivariate statistic technique has proven to be particularly effective when applied to identify surface changes, both for short and long

* *Corresponding author address*: Jorge Alberto Bustamante Becerra, CPTEC/INPE, Rodovia Presidente Dutra, Km 40, SP-RJ 12630-000, Cachoeira Paulista, SP, Brazil.
E-mail: jabb@cptec.inpe.br

time image data series (Byrne et al., 1980; Eastman, 1992; Eastman and Fulk, 1993; Fung and LeDrew, 1987; Ingebritsen and Lyon, 1985; Singh, 1989).

In this study, the potential of multi-temporal MODIS VI products for mapping of various land use and land cover types, mainly the cerrado physiognomies, by the use of the PCA and isodata unsupervised classification techniques, is explored. The objective was twofold: (1) to document the current status (condition) of the cerrado physiognomies and the human activity areas, specially agriculture activities, in the northern region of Brazil (Tocantins State), using MODIS VI products acquired in 2004; and (2) to evaluate the potential of MODIS VI products to identify the cerrado physiognomies and other land use and land cover types. The cerrado vegetation of Tocantins State is similar to the one of central Brazil; therefore, this study would also help the efforts to map cerrado physiognomies in the biome as a whole by using MODIS VI products, as part of the objective of the cerrado physiognomies mapping project for the Legal Brazilian Amazon.

In addition, the map of the cerrado physiognomies, which includes the land use changes derived from the MODIS image analysis can be used to improve the land surface scheme (SSiB-Simplified Simple Biosphere Model-SSiB, Xue et al., 1991, 1996) coupled with the general (Kinter et al., 1988; Xue et al., 1991; Cavalcanti et al., 1995; Kinter et al.,1997) and regional (Messinger, 1984; Mesinger et al., 1988) circulation model used at the CPTEC/INPE.

2 STUDY AREA

The area selected for the study is the State of Tocantins, Brazil (Figure 1). The climate of the region is tropical, hot and humid, with a dry season from May to October. Most of the state territory is covered with open xeromorphic vegetation, tropical savanna named Brazilian cerrado. Small areas of tropical rain forest remain along the Araguaia River, while areas of semi-decidual forest remain in the North and Southeast of the State). The main economic activity has been extensive cattle farming together with traditional subsistence agriculture. However, commercial agriculture has been developed recently in the Central and the Southeastern part of the State.

3. MATERIALS AND METHODS

The main data used were two MODIS Vegetation Indices, that is, the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI). Both indices were determined for 16-days composition periods, 250 m resolution, and the near infrared (NIR) and mid-infrared (MIR) bands. Also, a moisture index, the land surface water index (LSWI), was computed using these bands.

The MODIS VI products were obtained for the period from January to November 2004. These data are originally projected onto an integerized sinusoidal (ISIN) mapping grid. The data have been re-projected into geographic projection (lat/long coordinates and WGS84 datum) by the use of the MODIS reprojected tool (available on the web: http:/edc.usgs.goc/ programs/sddm/modisdist/).

To cover the area of the study, a mosaic of two MODIS tiles (h13v09 and h13v10) was used. For the seasonal analysis, four mosaics were used: for the rainy season, from January 01 to 16, 2004, and from March 21 to April 05, 2004, and for the dry season, from July 11 to 26, 2004 and September 29 to October 14, 2004. For each period, 3 mosaics were elaborated, corresponding to NDVI, EVI and LSWI index images, totaling 12 mosaics.

The map of the cerrado formations of the Brasilia National Park-BNP (Ratana and Huete, 2004; Ferreira et al., 2004) was selected to identify the main cerrado physiognomies. Three polygons were drawn over each physiognomy type (cerrado grassland, wooded cerrado and cerrado woodland) and other three polygons were drawn over the gallery forest, agriculture-pasture and agriculture land cover classes. These polygons were plotted over the MODIS index images of the same area (BNP) to extract the pixel values for each index image and calculate the mean, minimum, maximum and standard deviation parameters.

The MODIS NDVI image of the study area (Tocantins State) was segmented into areas of connected pixels based on the minimum and maximum value of each BNP cerrado physiognomy. Three small sites of approximately 20 pixels were selected over each segmented area, thus, there were 3 sites for each physiognomy.

Six land cover types were identify from the MODIS time series data, including 3 major cerrado physiognomies, as follows: cerrado grassland (CG), wooded cerrado (WC), cerrado woodland (CW), forest gallery (FG) and cerrado converted into agriculture-pasture (AP) and agriculture (A).

Six land cover types were identify from the MODIS time series data, including 3 major cerrado physiognomies, as follows: cerrado grassland (CG), wooded cerrado (WC), cerrado woodland (CW), forest gallery (FG) and cerrado

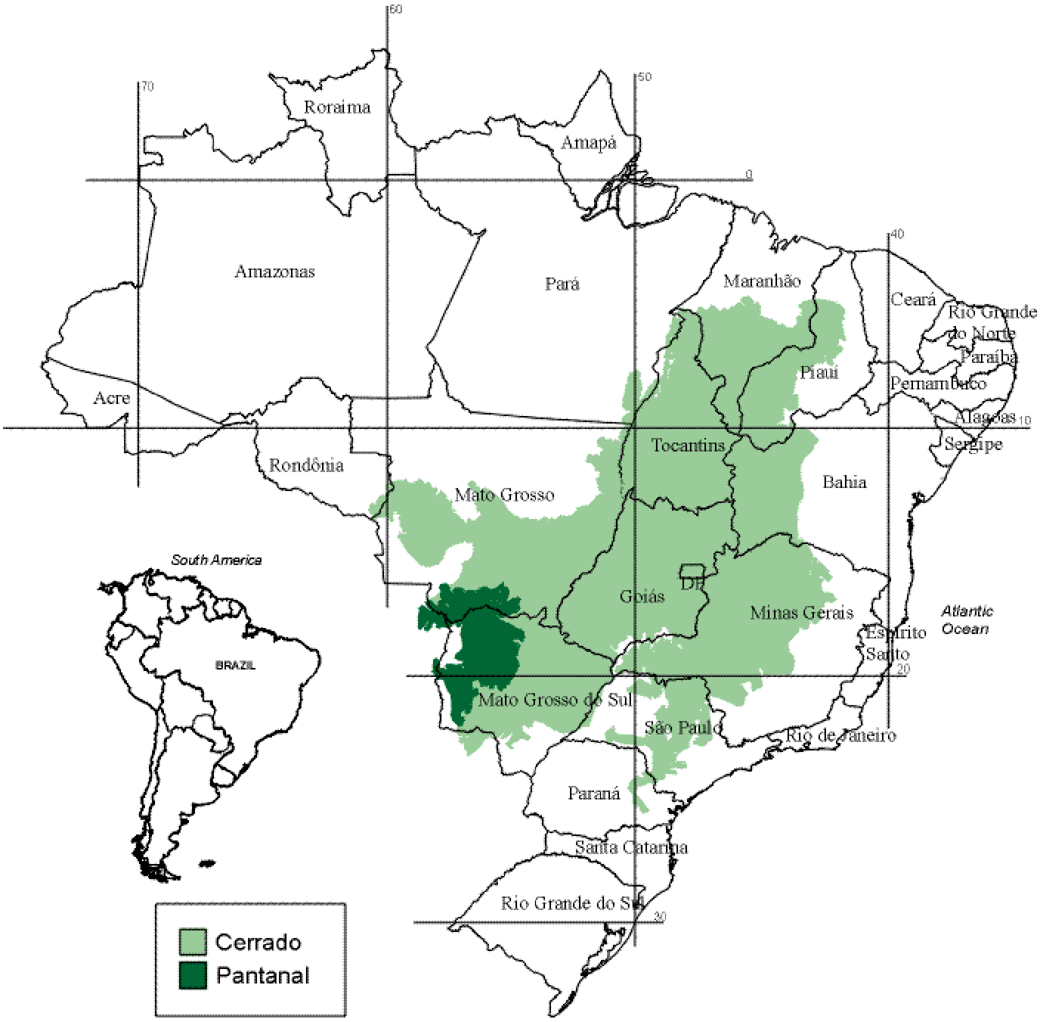

FIGURE 1. Study area (Tocatins State) within the cerrado biome. Extension of the cerrado biome over the Brazilian States, and location of Brazil in South America.

converted into agriculture-pasture (AP) and agriculture (A).

Principal component analysis (PCA) was applied on short time-series MODIS index data (NDVI, EVI e LSWI) to identify the accumulated greenness and the seasonal variability during the period of analysis.

PCA is a linear transformation of correlated variables into uncorrelated variables which does not change the number of variables (spectral or temporal bands). Because the transformed variables, which are named principal components (PCs), are uncorrelated, they can be considered to be independent (Johnson and Wichern, 1998). The transformed variables are ordered in terms of variance, so that the first PC represents the largest amount of variance within the data set. On a statistical basis, the multitemporal information present in the data is reflected by the variance. Therefore, at least a portion of the variance of the multitemporal image data set can be used to characterize land cover. If the PCA can extract factors that can be utilized to distinguish land cover, they will be objective and mutually independent.

The land cover classes were determined through an isodata unsupervised classification technique considering the first two principal components of the MODIS index images (NDVI, EVI and LSWI). Isodata (Iterative Self-Organizing Data Analysis Technique) unsupervised classification calculates class means evenly distributed in the data space and then iteratively clusters the remaining pixels using minimum distance techniques (Tou & Gonzales, 1974). The number of classes was set to 12, and either if a relative change in all of the cluster means did not exceed 1%, or after 20 iterations, the process was terminated.

To validate the results of the classification, the six land cover polygons (FG, CW, WC, GC, A-P and A) were plotted over the 12 spectral classes, obtained by isodata, to find some relationship between spectral and land cover classes.

Finally, for the determination of the seasonal patterns of each class evaluated, a seasonal profile was elaborated, from one transect that includes the main classes of vegetation cover.

4 RESULTS AND DISCUSSIONS

4.1 Principal component analysis

Figure 2(a) shows the loadings of the NDVI, EVI and LSWI eigenvectors plotted against the date (from January to October, 2004) for the first principal component (PC) extracted by the PCA, and applied to the entire study area. As different studies have demonstrated (e.g. Eastman, and Fulk, 1993, Hirosawa et al., 1996), the first PC tends to capture the accumulated greenness, which is the first major source of variation in the vegetation signal. The loadings of the first NDVI, EVI and LSWI eigenvectors were consistently positive and nearly identical.

The second principal component (Figure 2(b)) characterizes the seasonal contrast pattern of the vegetation (NDVI and EVI) and the moisture (LSWI) indices. The second NDVI, EVI and LSWI eigenvectors showed striking differences. While the loadings related to the vegetation and the moisture indices were, in absolute values, quite similar, the trends that they defined, with respect to time, were reversed.

For the vegetation indices (NDVI and EVI) second principal component, which correspond to approximately 12.4 % and 12.1% of the total information contained in the NDVI and the EVI data set, respectively, the negative loading values were predominantly related to the wet season months, while the positive ones were associated to the dry season months.

The LSWI second component, which contained approximately 8% of the total information of the LSWI dataset showed, in contrast, positive loading values associated to the wet season months, and negative ones associated to the dry season months.

These results corroborated the ability of the PCA approach to capture the seasonal contrast of the cerrado vegetation, and indicated that the vegetation (NDVI and EVI) and the moisture (LSWI) indices responded differently to the major seasonal variations of the vegetation. The vegetation indices (NDVI and EVI) seemed to be slightly more sensitive to the variation that occurs during the dry winter, while the moisture index (LSWI) tended to show slight higher response to the variation of the vegetation cover occurring during the rainy summer months.

4.2 Isodata unsupervised classification

The result of the isodata unsupervised classification of the first PC MODIS index data (NDVI, EVI and LSWI) included 12 spectral classes. Figure 3 shows the number of the spectral classes plotted against the six land cover classes.

These results showed a relationship of the high photosynthetic activity, the high vegetation cover (given by foliar phytomass) and the high humidity with the gradients of cerrado phytophysiognomies. This gradient varies from forested formations (forest gallery and cerrado woodland) to grassland formations (open cerrado grassland - campo limpo), according to the pattern showed by the Figure 2.

For the land use classes related to the anthropogenic activities (agriculture-pasture and agriculture), the vegetation cover of the agriculture-pasture class is greater than the one for the agriculture class. This can be explained by the fact that the agriculture class exhibits some periods of the year without vegetation cover (after the harvest period), differently from the agriculture-pasture class which almost always exhibits some degree of vegetation cover.

The proposed classification, when compared with the mean of the LULC classes, shows that the overlapping of classes exist between cerrado grassland and agriculture-pasture. Thus, no overlapping between the cerrado physiognomy classes occur. On the other hand, the comparison with the maximum and the minimum values of the LULC classes (wooded cerrado-WC and cerrado grassland-CG, and agriculture-pasture-AP and agriculture-A) shows a small overlapping of two subgroups of classes: WC with CG, and CG and AP with A.

These results show that there is some degree of overlapping between classes, like: WC with CG; and CG, AP with A. This overlapping may be explained by the degree of similarity of the spectral response of the LULC classes when it is analyzed the set of bands that compose the spectral indices utilized.

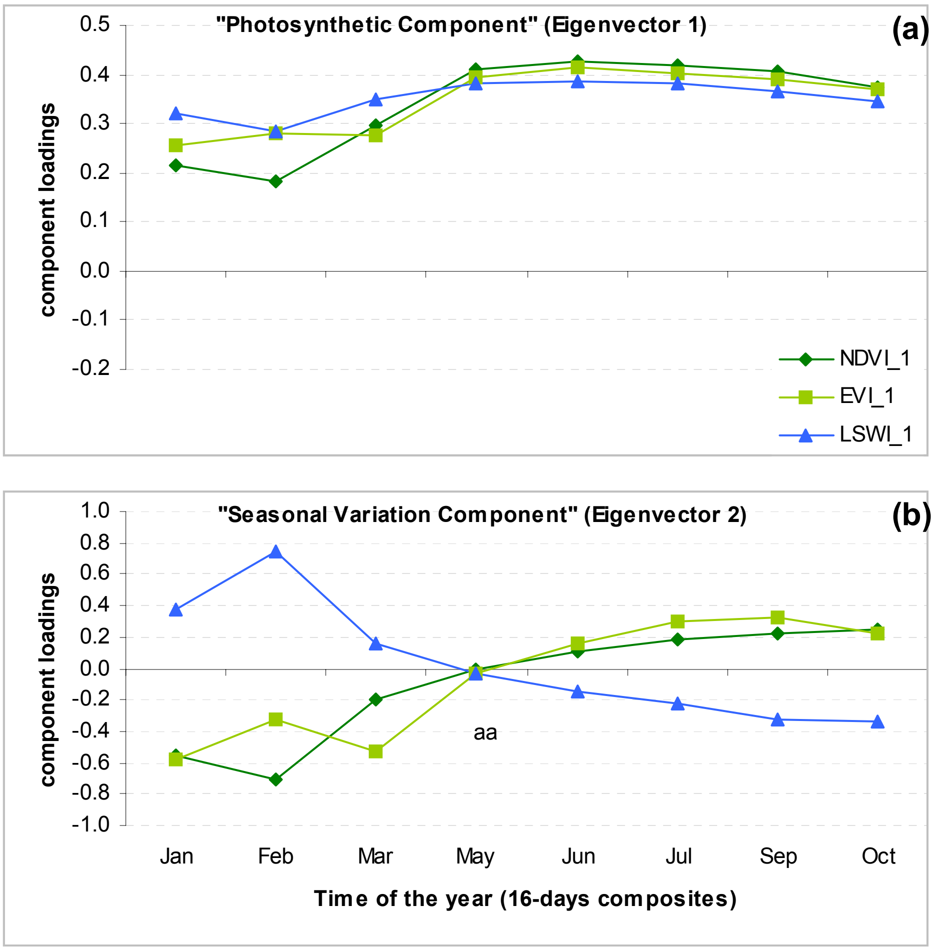

FIGURE 2. NDVI, EVI and LSWI eigenvector structures for the first (a) and second (b) principal components.

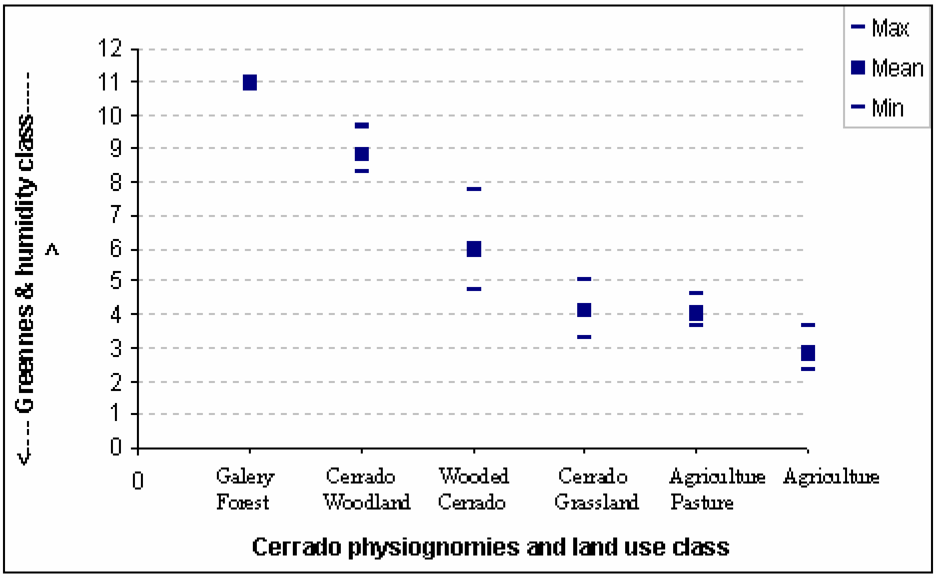

FIGURE 3. Relationship of the isodata unsupervised classification of the first principal component of the vegetation and the moisture indices against the LULC (cerrado physiognomies, forested formation and anthropogenic activities) classes.

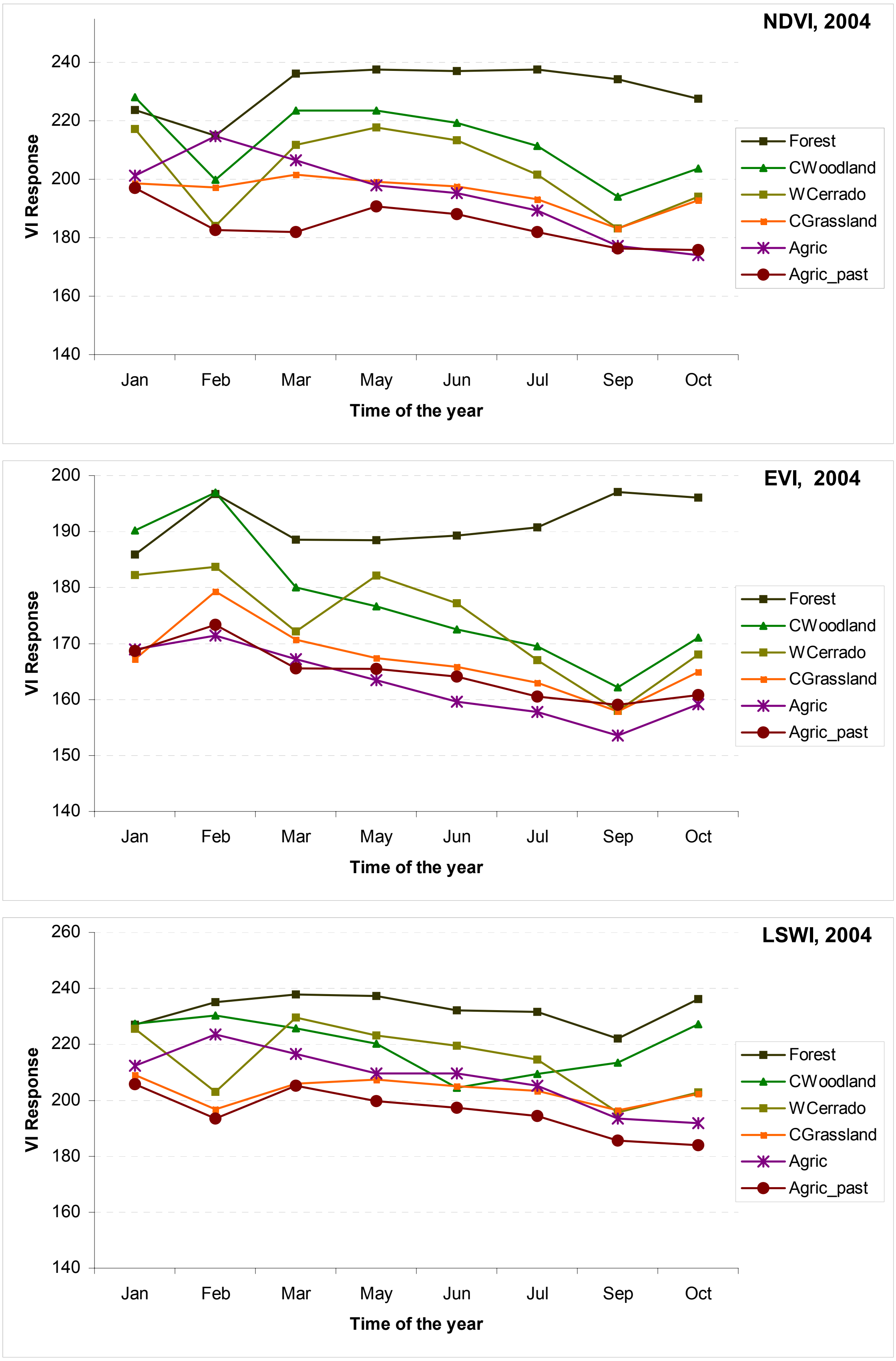

FIGURE 4. NDVI, EVI and LSWI seasonal profile for the major land cover types (i.e. cerrado physiognomies, forested formations, and anthropogenic activities)

4.3 Vegetation seasonal patterns

Figure 4 shows the spectral (NDVI, EVI and LSWI) indices response plotted against the time of the year (2004). This results, for the seasonal behavior of six the LULC classes (GF, CW, WC, CG, AP and A), showed the occurrence of difficulties to separate the forested formations (GF) from the agriculture areas, during the rainy season, and the agriculture from the savanna areas, during the dry season. This seasonal pattern is identified from the profiles of the 3 indices (NDI, EVI and LSWI) that shows similar tendency in the profile of the land cover classes in each index analyzed (Figure 4). Thus, firstly, a way to identify the Cerrado physiognomies of the Amazon region could be, through the separation of the Cerrado class from the other classes using the rainy season data and, thereafter, by the classification of the Cerrado physiognomies, using the dry season data.

5. CONCLUSIONS

The main objective of this study was to identify the land use land cover (LULC) of the tropical savanna of the Tocantins State, Brazil, using short time-series MODIS indices data (NDVI, EVI and LSWI). The method developed allows classifying the LULC based on the spectral temporal signature, derived from MODIS vegetation and moisture indices.

PCA and Isodata unsupervised classification techniques were applied to a short time-series MODIS data set to identify LULC classes, specially the cerrado physiognomies, for a tropical savanna region. When PCA was applied, two of the most important factors for characterizing the vegetation cover were clearly extracted from the first two PCs. The first one reflects the spatial variation of the indices (NDVI, EVI and LSWI), while the second one reflects the seasonal variation of these indices.

The analysis of unsupervised classification of the first PC of the MODIS vegetation (NDVI and EVI) and the moisture (LSWI) indices allow both grouping and identifying spectral classes. The association of these classes with the six LULC classes allow establishing a relationship between the high value of the spectral class and the high vegetation cover. This relationship follows a gradient varying from high to low spectral class, related with high to low vegetation cover.

The results show that the MODIS indices data are useful to monitoring the seasonal dynamics of the cerrado physiognomies and the anthropogenic activities (agriculture-pasture and agriculture). The seasonal profiles of the vegetation and the moisture indices of the savanna formations show an intermediate pattern between gallery forest and anthropogenic activities.

According to the seasonal behavior of the LULC classes, a way to identify the Cerrado physiognomies in the tropical savanna of the Amazon region could be, first, by the isolation of the Cerrado class from the other ones using data for the rainy season and, thereafter, by the classification of the Cerrado physiognomies, using data for the dry season.

Finally, these results show the relevance of the vegetation (NDVI and EVI) and the moisture (LSWI) indices, related to the biophysical parameters, to identify the LULC classes for the tropical savannas of the Amazon region. Nevertheless, to obtain a more complete mapping of the cerrado region and to avoid the overlapping of classes other factors, like soil, topography, heat spots and climate data (mainly precipitation) are required.

REFERENCES

Asrar, G., Fuchs, M., Kanemasu, E.T., and Hatfield, S.L., 1984. Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. *Agron. J.*, **76**, 300-306.

Byrne, G. F., Crapper, P. F., and Mayo, K. K. (1980), Monitoring land-cover change by principal components analysis of multiternporal Landsat data. *Remote Sens. Environ.* **10**:175-184.

Cavalcanti,I.F.A.; Nobre, P.; Abreu, M.L.; Quadro, M.; Pezzi, L.P., 1995. Vertical and horizontal resolution comparisons of CPTEC/COLA GCM. Proceedings of the twentieth annual climate diagnostics workshop, Seattle, Washington, Oct. 23-27, 1995, pp 73-76.

Eastman, R. J. (1992), Time series map analysis using standardized principal components, in *ASPRS/ACSM/RT 92 Technical Papers', Volunue 1: Global Change and Education, 3*-8 August, Washington, DC, pp. 195-204.

Eastman, R. J., and Fulk, M. (1993), Long sequence time series evaluation using standardized principal components. *Photogramm. Eng. Remote Sens.* **59**(8):1307-1312.

Fennessy, M. J., and J. Shukla, 1996: Impact of Initial Soil Wetness on Seasonal Atmospheric Prediction. COLA Tech. Rep. No. 34. Available from COLA, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705, USA, 31 pp.

Ferreira, L.G., Yoshioka, H., Huete, A., Sano, E.E., 2004. Optical characterization of the Brazilian Savanna physiognomies for improved land cover monitoring of the cerrado biome: preliminary assessments from an airborne campaign over an LBA core site. *Journal of Arid Environments*, **56**, 425–447.

Fung, T., and LeDrew, E. (1987), Application of principal components analysis to change detection. *Photogramm.Eng. Remote Sens.* **53**(12):1649-1658.

Furley, P.A., 1996, The nature and diversity of neotropical savanna vegetation with particular reference to the Brazilian cerrados. *Global Ecology and Biogeography*, **8**, 223-241.

Johnson, R.A. and Wichern, D.W., 1998. Applied multivariate statistical analysis. 4th. Ed. Prentice Hall, New Jersey, 816p.

Hirosawa, Y., Marsh, S.E., Kliman, D.H., 1996 Application of Standardized Principal Component Analysis to Land-Cover Characterization Using Multitemporal AVHRR Data. *Remote Sens. Env*. **58**:267-281.

Huete, A., Justice, C., 1999. MODIS Vegetation Index (MOD13) Algorithm Theoretical Basis Document, Greenbelt: NASA Goddard Space Flight Center, Technical Report, 129pp.

Huete, A.R., Justice, C.O., Liu, H., 1994. Development of vegetation and soil indices for MODIS-EOS. *Remote Sens. Envir.* **49**, 224–234.

Justice, C.O., Vermote, E., Townshend, R.D., Roy, D.P., Hall, D.K., Salomonson, V.V., Privette, J.L., Riggs, G., Strahler, A., Lucht, W., Myneni, R.B., Knyazikhin, Y., Running, S.W., Nemani, R.R., Wan, Z., Huete, A.R., van Leeuwen, W.J.D, Wolfe, R.E., Giglio, L., Muller, J., Lewis, P., Barnsley, M.J., 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sens.* **36** (4), 1–22.

Ingebritsen, S.E, and Lyon, R.J.P., 1985. Principal component analysis of multitemporal image pairs. *Int. J. Remote Sens*., **6**(5): 687-696.

Kinter,J.L.;J.Shukla; L.Marx; E.K. Schneider, 1988. A simulation of winter and summer circulations with the NMC global spectral model. *J.Atm.Sci.*, **45,** 2486-2522.

Kinter III, J., Dewitt, D., Dirmeyer, P., Fennessy, M., Kirtman, B., Marx, L., Schneider, E., Shukla, J. and Straus, D., 1997. The COLA atmosphere-biosphere genereal circulation model, volume 1, formulation. COLA. Technical COLA. Technical Report, 51, 45p.

Liu, H.Q., Huete, A.R., 1995. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sens.* **33**, 457–465.

Lopez, A.S. and Cox, F.R., 1977, A survey of the fertility status of surface soils under 'cerrado' vegetation in Brazil. *Soils Science Society of America Journal*, **41**, 742-747.

Mantovani, J.E., Pereira, A., 1998. Estimating the integrity of the Cerrado vegetation cover through the Landsat-TM data. In: Simposio Brasileiro de Sensoriamento Remoto, 9, Santos, SP (in Portuguese, CD-ROM version).

Mesinger, F., 1984: A blocking technique for representation of mountains in atmospheric models. *Rev. Meteor. Aeronaut.*, **44**, 195-202

Mesinger, F., Z. I. Janjic, S. Nickovic, D. Gavrilov, D. Deaven, 1988: The step-mountain coordinate: Model description and performance for cases of Alpine lee cyclogenesis and for a case of an Appalachian redevelopment. *Mon. Wea. Rev.*, **116**, 1493-1518.

Nepstad, D., Klink, C.A., Uhl, C., Vieira, I.C., Lefebebvre, P., Pedlowski, M., Matricardi, E., Negreiros, G., Brown, I.F., Amaral, E., Homma, A., Walker, R., 1997. Land use in Amazonia and the Cerrado of Brazil. Ciencia & Cultura (Sao Paulo) 49, 73–86.

Oliveira, P.S., & Marquis, R.J. (Eds.) (2002). *The Cerrados of Brazil-ecology and natural history of neotropical Savanna*. New York, NY, USA: Columbia University Press.

Ratana, P. and Huete, A., 2004. Seasonal dynamics of native and converted cerrado physiognomies with MODIS data. Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International, 4336-4339.

Rousse, J.W., Haas, R,H., Schell, J.A, Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS, Proceedings, Third Earth Resources Technology Satellite-1 Symposium, Greenbelt: NASA SP-351,3010-317.

Running, S.W. and Nemani, R.R., 1988. Relating seasonal patterns of the AVHRR vegetation index to simulated photosynthesis and transpiration of the forests in different climates, *Remote Sens. Env*., **24**, 347-367.

Sano, E.E., Barcellos, A.O., Bezerra, H.S., 2001. Assessing the spatial distribution of cultivated pastures in the Brazilian savanna. *Pasturas Tropicales* **22** (3), 2–15.

Sellers, P.J., 1985. Canopy reflectance, photosynthesis and transpiration, *Int. J. Remote Sens*., **6**, 1335-1372.

Singh, A. (1989), Digital change detection techniques using remotely sensed data: review article. *Int. J. Remote Sens.***10**:989-1003.

Townshend, J.R.G, Justice, C.,Li, W., Gurney C and McManus, J., 1991, Global landcover classification by remote sensing: present capabilities, and future possibilities, *Remote Sens. Environ*., **35**, pp. 243-256.

Tou, J. T. and R. C. Gonzalez, 1974. Pattern Recognition Principles, Addison-Wesley Publishing Company, Reading, Massachusetts.

Tucker, C.J. and Sellers, P.J., 1986. Satellite remote sensing of primary production. *Int. J. Remote Sens*, **7**, 1395-1416.

Xue,Y.; P.Sellers.J.L. Kinter; J.Shukla, 1991. A simple biosphere model for global climate studies. *J. Climate*, **4**, 345-364.