próximo
artigo

Crop patterns extraction derived by classic Fourier analysis of EVI-MODIS time-series data to support crop discrimination

John Mauricio Arenas-Toledo [1]
José Carlos Neves Epiphanio [1]

[1] Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12245-970 - São José dos Campos - SP, Brasil
{john, epiphanio}@dsr.inpe.br

Abstract. The current economic panorama with market crises, food crises, bio-fuels expansions, commodities crush down, etc. makes absolutely relevant for any country to setup agriculture information in a quick and operational way. In this complicated scenario we proposed an approach to perform crop discrimination based on crop patterns of major annual crops in Mato Grosso State, known as one of the largest world agriculture frontier. This region is a large agriculture producer, especially of soybean, cotton and maize. These annual crops have a short cycle, which makes crop monitoring hard to achieve only by using medium spatial resolution imagery because there is a coincidence with a period of high cloud cover, particularly during the summer season. Lower-order harmonic terms derived from Time-series of EVI MODIS were related to crop patterns. We found that cotton areas were modeled by first-order term and succession soybean and second maize crop known as *safrinha* were modeled by second-order term. Per-pixel classifications of harmonic terms reached accuracies of 90% for harmonic terms.

Keywords: annual crops, Fourier series, harmonic terms, per-pixel classification
Palavras-chave: culturas anuais, series de Fourier, termos harmônicos, classificação por pixel

1. Introduction

Modeling of seasonal profiles using time-series data obtained by satellite imagery, especially AVHRR-NDVI, have being carried on by several researchers using classic Fourier analysis or harmonic analysis (Azzali and Menenti, 2000; Moody and Johnson, 2001; Jakubauskas et al., 2001). With this high temporal resolution data, including MODIS, others authors proposed wavelet analysis (Sakamoto et al., 2005; Galford et al., 2008). Hermance (2007) proposed a method to stabilized superior-order terms for non-classical Fourier analysis. Bradley et al. (2007), who used the Hermance (2007) approach, performed a NDVI annual average of 12-years to produce a curve fitting to define the onset of greenness. Galford et al. (2008) stated several considerations about limitations in the method follow by Bradley et al. (2007) because it presumes periodicity over the years and also the phenology as a function of previous years, and in the study area it is not accurate for annual crops.

Harmonic processing decomposes temporal curves in amplitude, variance and phase metric terms. These harmonic terms are significant for incorporating information and knowledge of seasonal profiles of crop cycles (Moody and Johnson, 2000; Jakubauskas et al., 2001). As MODIS products and bands are radiometrically and geometrically corrected (Justice et al., 1997), they are more suitable than AVHRR data as input for time-series modeling in recent years.

In *Mato Grosso State*, which is one of the largest agriculture frontier in the world, we tested the Fourier analysis approach with some restrictions according to the annual crop targets tested to extract crop patterns as support of crop discrimination and posterior mapping jointly with medium spatial resolution data.

2. Study area

The study area contains four municipalities in the southern region of *Mato Grosso* State, which is located in Central Brazil (Figure 1). They are located between latitude 14°32’ S and 16°25’ S and longitude 53°40’ W and 55°27’ W. The whole area is 1,884,900 ha. The municipalities are large producers of soybean, cotton and maize. The other vegetal units besides crop areas correspond to *cerrado* (savannas), grasslands and ciliar forest. The climate is warn semi-humid, average temperature around 21° C, rainfall (mean) is 1500 mm with a bimodal rainy season, dry period from May to September. Altitude varies from 140 m to 900 m.

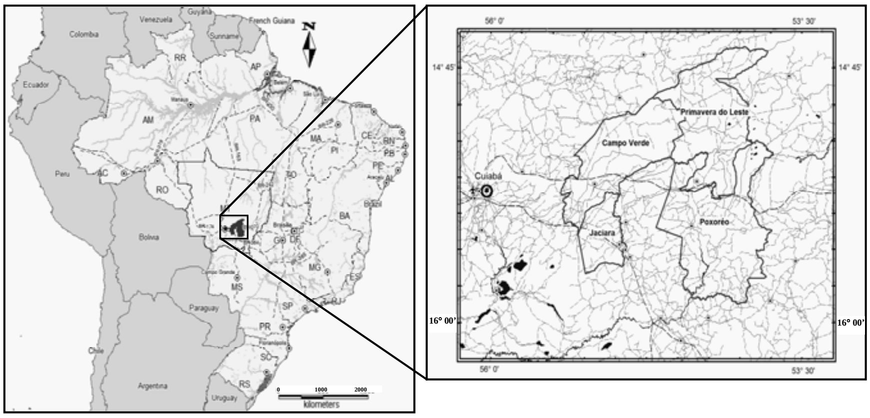


Figure 1. Study area, southern part of *Mato Grosso* State, Brazil.

3. Data and methods

3.1 Crop Calendar

The crop calendar for the study area exhibits three major annual crops: soybean, cotton and maize (first and second periods). For soybean, the sowing for non-irrigated crops began in middle September with an intense period between October and November. Soybean harvest period embraces January to April. Cotton sowing began from the middle of November till middle of February; intense sowing peak occurs in the final of December and beginning of January. Harvesting period occurs from middle April till middle August. Maize in *Mato Grosso State* has two different crop periods, in last years the maize of second period has larger production: First period or summer maize has a sowing period of three months, from October to December; November is the most intense sowing month. Harvesting period is from February to June. The second period of maize (*safrinha*) represents the major part of maize in *Mato Grosso* (USDA-WAP, 2007). Its sowing period is from January until middle March, with intense peak on February; harvesting period occurs from May to middle of September. This second maize is sowed in soybean-harvested areas, and is the most frequent succession practice: maize after soybean. Central pivot crops have two or three crops in one year: beans, soybean, peanuts are usually irrigated crops.

SOWING AND HARVEST CALENDAR

CROP	PHASE	June: 21 to Sept: 22						Sept: 23 to Dec: 21						Dec: 22 to Mar: 20						Mar: 21 to June: 20					
		Winter						Spring						Summer						Autumn					
		JULY		AUG		SEPT		OCT		NOV		DEC		JAN		FEB		MAR		APR		MAY		JUNE	
		15	30	15	30	15	30	15	30	15	30	15	30	15	30	15	30	15	30	15	30	15	30	15	30
SOYBEAN 105 to 135 days	SOW.								S	S															
	HARV.																H	H							
COTTON 150 to 180 days	SOW												S	S											
	HARV	H																							H
MAIZE 1st 120 to 180 days	SOW.									S	S														
	HARV.																		H	H					
MAIZE 2nd 120 to 180 days	SOW.															S	S								
	HARV.	H																							H

Legend: sowing; S S intense sowing; harvest; H H intense harvest

Source: CONAB/SUINF/GEASA

Figure 2. Crop calendar for soybean, cotton and maize for the study area

3.2 Field work campaign

The major ground-truth data collected in the fieldwork was information of land cover units with special attention to crop type units. In addition, qualitative data like crop early/late stages, crop condition, etc., was gathered in the fieldwork. We used information from two field work campaigns. The first one captured information about summer crops especially soybean (15 to 21 of January 2007), they were collected 150 points. The second field campaign was done in the transition crop period (one crop is harvested and other crop is sowed on the same plots). The field work occurred from 27/Feb to 06/Mar of 2007. In this campaign there were collected 1500 points. At that time, several soybean fields were already harvested and some maize fields were also sowed in formerly soybean fields. Some fields had late soybean still on ground. Cotton fields were in the middle of their phenologic stage.

These large ground-truth data set were used jointly with medium resolution imagery (Landsat TM and CBERS-CCD) to generate a reference map obtained by on-screen digitalization of vector polygons using images of several dates.

3.3 Method

Multitemporal data were collected of MODIS EVI (Enhanced Vegetation Index) 16-days composite data of 250 m which is embedded in the MODIS product (MOD13Q1), for the tile h12v10. This MODIS product for EVI is corrected for sun-target-sensor variations, angle reduction, and a BDRF schema option (Justice et al., 1998; Huete et al., 2002).

The EVI formerly called as SARVI2, was develop to optimize the vegetation signal and minimize atmospheric influence and background brightness effects (Huete et al., 1994). The equation for this vegetation index is as follow:

$$EVI = G\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C1\times\rho_{red} - C2\times\rho_{blue} + L} \tag{1}$$

Where ρ is surface reflectance in sensor bands: NIR - Near Infrared; red and blue in the optical region; L is the canopy background adjustment factor, and C1 and C2 are the coefficients of the aerosol resistance term; L=1, C1 = 6, C2 = 7.5, and G (gain factor) = 2.5 (Huete et al., 1994).

EVI-MODIS data which is originally available in HDF format (**H**ierarchy **D**ata **F**ormat) and sinusoidal grid projection was reprojected and subsetted to contain the study area frame. To perform this processing we used the **M**ODIS **R**eprojection **T**ool (MRT) software, and the selected output format was Geotif.

The dataset evaluated correspond, according to the crop calendar, to the period of September/2006 to June/2007. In this period 18 composite images were available as input of harmonic analysis model via Fourier Discrete Transformation (DFT). In this case, differences in processing data (collections V004 and V005) did not affect our results because it was modeled by coarse seasonal behavior curves but not decimal numeric values, which differences in processing for these two collections, eventually could derive. Further details of collection changes are commented in Didan and Huete (2006).

For the observed discrete data, registered at regular t times, the interval from origin time is t_i= 1 to N. For this sampled discrete data, N is the period of this discrete Fourier series, y(t)={1,2,..,N}

$$y(t) = c_0 + \sum_{k=1}^{M} A_k \cos\left(\frac{2\pi kt}{N}\right) + B_k \sin\left(\frac{2\pi kt}{N}\right) \tag{2}$$

where k is restricted by Nyquist frequency. Thus the calculus of t has two possibilities: when t is even $M_{max}=k=N/2$, for odd values $M_{max}=k=N\text{-}1/2$. C_0 is called the additive term and represents the mean of the whole time-series set; A_k and B_k are the Fourier coefficients and can be calculated as follows:

for any non-negative integer *k*:

$$A_k = \frac{2}{N} + \sum_{t=1}^{N} y_{(t)} \cos\left(\frac{2\pi t}{N}\right) \tag{3}$$

$$B_k = \frac{2}{N} + \sum_{t=1}^{N} y_{(t)}\, sen\left(\frac{2\pi t}{N}\right) \tag{4}$$

The amplitude for the *k-nt* term is:

$$c_k = \sqrt{A^2{}_k + B_k{}^2} \tag{5}$$

The percentage variance for each harmonic term is the simple ratio of amplitude and the double-value of population variance.

$$\%\ Variance\ = \frac{c^2{}_k}{2s^2} \tag{6}$$

The phase angle can be calculated according to the range angle:

$$\phi_k = \begin{cases} \tan^{-1}\dfrac{B_k}{A_k} & A_k > 0 \\ \tan^{-1}\dfrac{B_k}{A_k} \pm \pi\ (180°) & A_k < 0 \\ \dfrac{\pi}{2}(90°) & A_k = 0 \end{cases} \tag{7}$$

Once the harmonic term images were generated, points of typical fields (1780 in total) crops were selected trying to avoid spectral mixture between vegetal units and assuring their spectral behavior being analyzed over time. Some of these typical points (around 660) were used to collect harmonic term values to generate comparison between crop patterns single and double, and also between EVI curve and Harmonic term modeling. Individual and average curves for cotton and soybean/maize succession were compared. Finally, there were performed per-pixel classifications: ISODATA (Iterative Self-Organizing Data Analysis Technique) and Gaussian MLC (Maximum Likelihood Classifier) over harmonic terms images (1st-term, 2nd-tem amplitude, and 1st-term phase). Accuracy assessment, compared to reference map, was performed using Congalton and Green (1999) approach with 530 random points.

4. Results and Discussion

Major annual crops for study area are soybean, cotton and maize; most of cotton fields have a previous management practice which is sowed millet for being incorporated as mulch. As observed in crop calendar, the timing is different to each of these crops (Figure 3); EVI images follow that crop dynamics as shown ahead. Soybean is sowed in the last trimester of the year (early-medium and late soybean cultivars). The green-peak of most soybean areas is in late December and early January (2007-017, Figure 3). Soybean harvesting period is between January (early soybean) and April (late soybean). Cotton was sowed between the end of the year 2006 and beginning of the 2007; in March, cotton areas exhibit the green-peak of its cycle (2007-065, Figure 3). In general, after soybean is harvested, farmers sowed maize. In image 2007-113 (Figure 3), cotton and maize fields are displayed with higher EVI values than other vegetated units. In image 2007-177, annual crops have been harvested; the dry season is in its onset, and some few areas under central-pivot irrigation are still active.

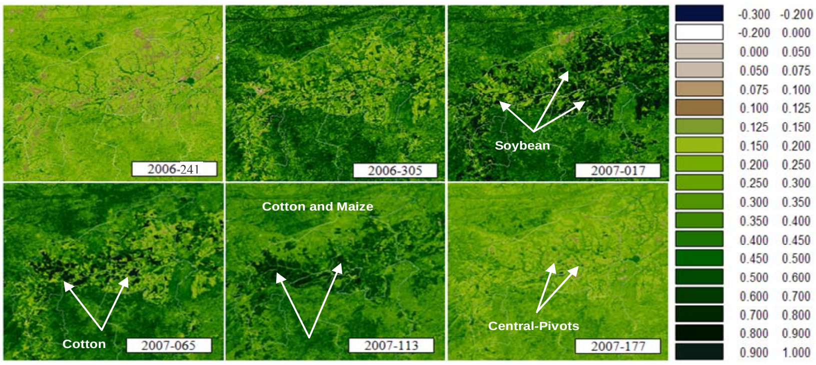


Figure 3. Dynamics of EVI composition images at different dates during the crop (period 2006-2007), days are expressed in day of year format (DOY).

In this region, a particular point over an agriculture field-plot is not stable over time, indeed, cotton and soybean fields area redistributed along a period of 3 to 4 years, varying every year in the whole context. This means that single and double patterns or even fallow in smaller percent areas do not follow a strict tendency for a specific resolution point. Because of this, time-series for several years are not as symmetric and periodical as would be recommended for techniques as DFT, and it is more consistent to analyze a single year context isolated as we proposed.

4.1 Harmonic terms

Harmonic processing decomposes temporal curves in amplitude, variance and phase metric terms. According to the typical points collected (1720), 328 points for cotton and 328 points for soybean/maize were used to generate mean curves of EVI for cotton and soybean/maize. The remaining points were over other vegetal units including semi-perennial sugarcane crop fields. Cotton profile for the 328 points exhibits a great variability in the period from October to December (Figure 4); this is the period when for most fields it is planted a gramineous plant, normally millet. The “green-up” phase shows the highest variability in the whole cotton cycle, which may indicate some delays in sowing timing. In harmonic modeling, both cases (Figure b,d) show that much of variance is captured by the first-order terms, with low influence values for the second-order term (Figure c,e). The sum of first four amplitude terms plus additive terms (Figure b,d) accounts for almost 70% of variance (Figure c) and more than 90% in the millet/cotton case (Figure e).

Soybean and maize succession planting exhibits a bimodal behavior in average seasonal profile (Figure 5a). After regular soybean initiates to green-up, composition of December-3 composition shows a high EVI variability due to the influence of late and even early soybean cultivars, observed also in the senescence stage on February-2 composition. Soybean green-peak covers the period from middle December to middle January, as maize does in April. March data expressed a high variability due to variance in sowing timing for maize crop areas. In harmonic modeling the first three harmonic terms contains 96.4% of the total variance (Figure 5c); the second-order term exhibit the highest variance percent value as a single term (Figure 5c). Second-order amplitude image presents a two-cycle wave pattern (figure 5e), highly coincident with actual soybean/maize succession pattern.

Other vegetal units such as forest, cerrado and natural grassland, and the semi-perennial sugarcane crop exhibited a heterogeneous and more complex profile in EVI curves. Those classes could be not modeled by just a single harmonic term, and variance is distributed in several order terms; many of them even in the higher terms which are susceptible to noise and spurious data (Hermance, 2007).

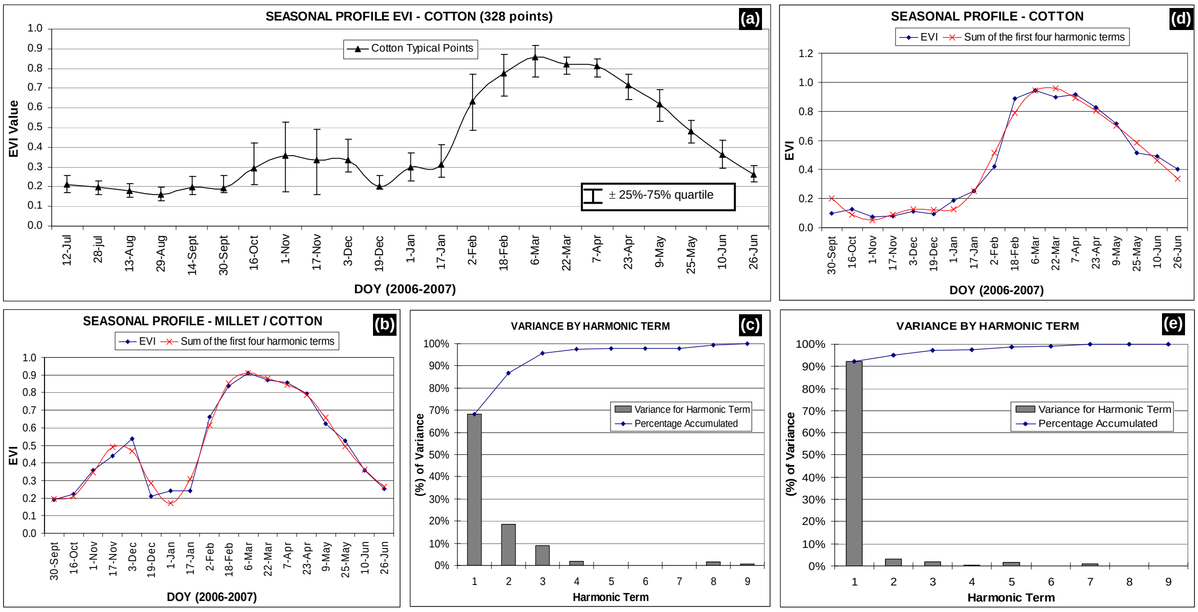


Figure 4. Comparison example of cotton fields with different management practices. (a) Seasonal profile of cotton crop, (b) variance for harmonic terms of cotton crop, (c) seasonal profile of millet and cotton crop, (d) variance for harmonic terms of millet and cotton crop.

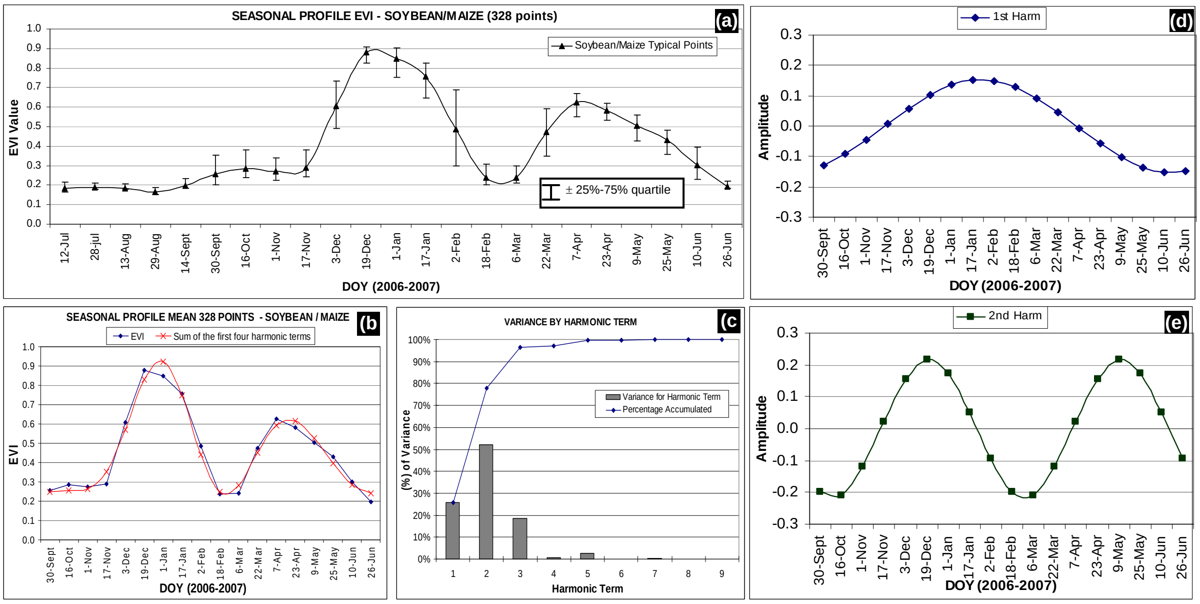


Figure 5. (a) Seasonal profile of mean typical EVI data for crop succession soybean/maize, (b) profile of EVI and sum of the first four harmonic terms, (c) variance for the nine harmonic terms, (d) 1st-order amplitude term, (d) 2nd-order amplitude term.

Phase term represents the angle in which the green-peak occurs in a particular vegetation unit. The angle refers to the time in which this peak occurs, one for each harmonic term. We used first-order phase as complementary axe to evaluate the separability for the two patterns set besides amplitude terms. As phase is a circular data, this data must be interpreted knowing the inherent conditions of its nature. In multiyear data or range selection criterion (e.g. standard year: startup in January), all phase data for annual crop target are circular. In this particular case, first-order phase data, according to the range time-series selection, were linear and phase data could have a valid value in the entire scale.

Once harmonic term images were produced, the reference map polygons were overlaid to harmonic images. Crop patterns matched with amplitude term values in image (Figure 6): cotton fields show high values for 1st-order term and low values for 2nd-order;

soybean/maize succession is the opposite case, low values for 1st-order term and high for 2nd-order term. This apparent affinity was tested by several per-pixel classifications (ISODATA and MLC). The accuracy assessment was accomplished using a multinomial formula (Congalton and Green, 1999) with 95% as a level of confidence, with 5% as desired precision for three classes (cotton, soybean/maize, others), obtaining 530 points.

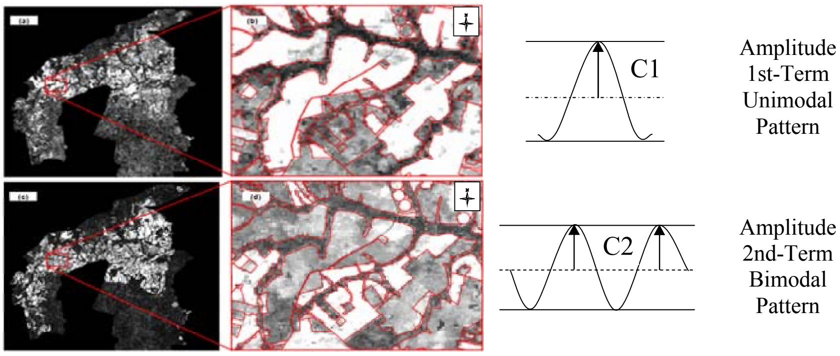
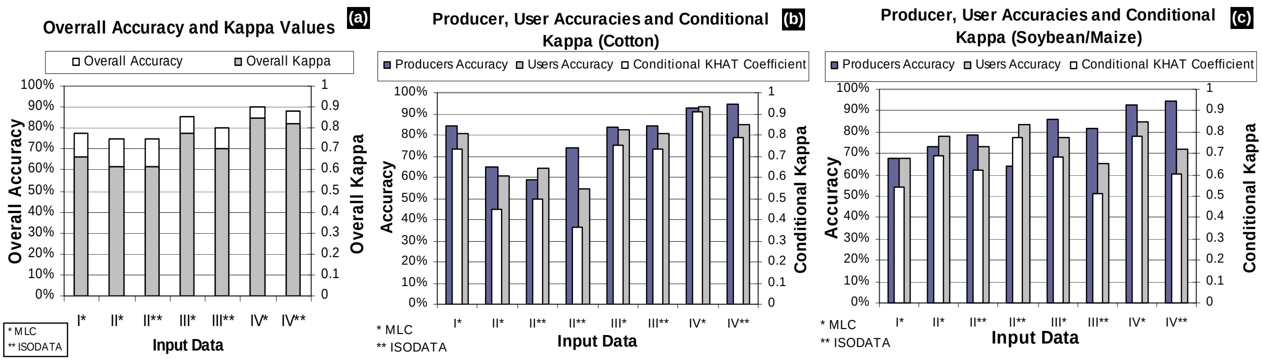


Figure 6. (a) First-order amplitude term image, (b) detail of cotton fields in bright. (c) Second-order amplitude term image, (d) detail of most of the soybean/maize fields in bright.

Accuracy assessment for the three classes reported as the highest overall accuracy of 90% with overall KHAT coefficient of 0.8479 was obtained for MLC of 3-band image composed by first and second order amplitude images plus first-order phase image (schema IV). Using only one single harmonic term, first-order amplitude showed the highest overall accuracy (77.74%) with an overall KHAT coefficient of 0.6604 (Figure 6a). In analysis by classes, highest accuracies for the MLC–IV in cotton conditional Kappa was superior to 0.9; for soybean/maize succession there was a tendency of higher commission errors, and the highest conditional Kappa was 0.78 for the MLC–IV.



X Axe [I: 1st Amplitude Term; II : 2nd Amplitude Term; III: 1st+2nd Amp. Term; IV: 1st+2nd Amp-1st Phase]
Figure 6. (a) Overall accuracy and kappa values; producer, users, conditional kappa for (b) cotton and (c) soybean/maize. II** are two different ISODATA classifications recoded to three classes (6 to 3 and 4 to 3).

5. Conclusions

Crop patterns (single and double), which involve management practices and crop cycle characteristics, are compelling to discriminate annual crops by harmonic analysis in study area. EVI-MODIS time-series of classic Fourier analysis captures differences in seasonal profiles between annual crops and other class units.

Just few lower harmonic terms are suitable for modeling annual crops; management practices such as millet before cotton had impact over individual harmonic term.

Individual amplitude images were successful to minimize differences in crop calendar for same crop type. Early, medium and late soybean cultivars were collected together in just a single term image.

Supervised (ML) and ISODATA standard classifiers applied over harmonic terms data showed high accuracies. It was not necessary to apply filters to reduce isolated class pixels in harmonic terms classifications as is usual in per-pixel approaches. MLC was slightly superior to ISODATA.

Best classification results were obtained when it was possible to include amplitude and phase harmonic terms, because they isolated specific crop patterns that can not be reached exclusively with spectral bands in punctual dates.

Gross estimation using harmonic terms derived by MODIS data is very suitable to be applied in *Mato Grosso State* and regions with similar dynamics and conditions. Knowledge of the agriculture context in the region is essential. It is proposed to explore this alternative in Central region of Brazil.

Acknowledgements

Special thanks to Dr. Silvia Pardi and Dr. Manoel Sousa for the Fourier algorithms that they implemented using the IDL software, and provided the compiled file to be used in this research.

References

Azzali, S.; Menenti, M., Mapping vegetation-soil-climate complexes in Southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data. **International Journal of Remote Sensing**, v.21, p.973–996, 2000.

Congalton, R.G.; Green, K. **Assessing the accuracy of remotely sensed data:** principles and practices. Boca Raton, Florida: Lewis Publishers, 1999. 137p. ISBN 0873719867.

Bradley,B.A.; Jacob,R.W.; Hermance, J.F.; Mustard, J.F. A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. **Remote Sensing of Environment**, v.106, p. 137-145, 2007.

Congalton, R.G.; Green, K. **Assessing the accuracy of remotely sensed data**: principles and practices. Boca Raton, Florida: Lewis Publishers, 1999. 137p. ISBN 0873719867.

Didan, K.; Huete, A. **MODIS vegetation index product series, collection 5 change summary**. TBRS Lab., University of Arizona, 2006. Available in <http://landweb.nascom.nasa.gov/QA_WWW/forPage/MOD13_VI_C5_Changes_Document_06_28_06.pdf>. Accessed in 15 Jan 2008.

Galford, G.L.; Mustard, J.F.; Melillo, J.; Gendrin, A.; Cerri, C.C.; Cerri, C.E.P. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. **Remote sensing of environment**, v.112, p.576-587, 2008.

Hermance, J. F. Stabilizing high-order, non-classical harmonic analysis of NDVI data for average annual models by damping model roughness**. International Journal of Remote Sensing**, v. 28, n.12, p. 2801-2819, 2007.

Huete A.; Justice C.; Liu H. Development of vegetation and soil indices for MODIS-EOS. **Remote Sensing of Environment**, v.49, p.224-234, 1994.

Huete, A.; Didan, K.; Miura, T.; Rodriguez, E. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. (Special Issue) **Remote Sensing of Environment**, v.83, p.195-213, 2002.

Jakubauskas, M.; Legates, D.; Kastens, J. Harmonic analysis of time-series AVHRR NDVI data. **Photogrammetric Engineering and Remote Sensing**, v.67, p.461–470, 2001.

Justice, D. H.; Salomonson, V.; Privette, J.; Riggs, G.; Strahler, A.; Lucht, R.; Myneni, R.; Knjazihhin, Y.; Running, S.; Nemani, R.; Vermote, E.; Townshend, J.; Defries, R.; Roy, D.; Wan, Z.; Huete, A.; Van Leeuwen, R.; Wolfe, R.; Giglio, L.; Muller, J.P.; Lewis, P.; Barnsley, M. The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research**. IEEE Transactions on Geoscience and Remote Sensing**, v.36, p.1228–1249, 1998.

Moody, A.; Johnson, D. Land-surface phenologies from AVHRR using the discrete Fourier transform. **Remote Sensing of Environment**, v.75, p. 305-323, 2001.

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

United States Department Of Agriculture. World Agricultural Production (USDA-WAP). **Circular series - report Mar, 2007.** Available in: <http://www.fas.usda.gov/wap/circular/2007/07-01/production.pdf>.Accessed in Dec 2007.