MAPPING LAND USE CLASSES BY ANALYZING MODIS LST TIME-SERIES

MAPEAMENTO DE CLASSES DE USO DO SOLO POR MEIO DE ANÁLISE DE SÉRIES TEMPORAIS DE DADOS MODIS-LST

Denis Araujo Mariano[*1], William Foschiera[2], Maurício Alves Moreira[3]

Remote Sensing Division (DSR), National Institute for Space Research (INPE), Av. dos Astronautas, 1758, São José dos Campos, CEP: 12227-010, SP, Brazil. Emails: [*1]denis.mariano@usp.br; [2]wfoschiera@gmail.com; [3]mauricio@dsr.inpe.br

ABSTRACT

The current paper presents a method to discriminate land use classes (LCCs) by analysing Land Surface Temperature (LST) time-series derived from the Moderate Resolution Imaging Spectro radiometer (MODIS). We used Terra and Aqua LST daytime and night time data (M D11A2) with 8-day temporal and 1km spatial resolution. The physical basis behind the method is the heat transfer between soil, plant and atmosphere over time. There are two approaches, inter-daily and intra-daily LST variation. We tested daytime and day-night difference time-series, being the latter more efficient on discriminating classes. Regarding the satellites, Aqua proves on being more efficient due the passage hour for daytime. In sense, the couple Aqua/Difference yielded better results. However, the performance is strongly dependent upon the targets' acreage due to the high thermal mixing effect. Despite the limitations, this approach shows potential on being coupled to traditional vegetation indices (VI) based methods for furthering the biophysical meaning and relationships between vegetation and the electromagnetic spectrum. It also brings new findings about vegetation thermal behaviour throughout the time.

Keywords: Land surface temperature, MODIS, time-series, Python, agriculture, thermal.

INTRODUCTION

Mapping land cover classes is a recurrent task and is always evolving due to new methods and sensors. Several works rely on direct image interpretation, which is based on the interpreter knowledge about the crops and datasets. Another common technique is based upon VI, so that many works use one or combinations of various(WARDLOW and EGBERT, 2008). Most of the VIs are based on the relationship between the red and near infra-red channels showing the contrast between the high radiation absorption by the chlorophyll in the red channel and high reflectivity by the leaf structure and its components on near infra-red channel (TUCKER, 1979). Albeit the aforementioned techniques are well consolidated and have proven their effectiveness, in some situations they can fail or be highly time-consuming depending on the method and experience of the interpreter. In sense, these techniques rely only upon optical remote sensing channels. However, the thermal analysis approach for mapping purposes was not in depth investigated yet. A comprehensive work of land cover classes discrimination and change analysis was performed by Lambin and Ehrlich (1996). They coupled VI and LST images from the Advanced Very High Resolution Radiometer (AVHRR) and analyzed 10 years of data for the African continent showing that the use of LST improves the capacity to discriminate land cover classes as compared to the VI applied solely. Nemani and Running (1997) compared the correlation between VI and LST derived data from AVHRR for different LCCs yielding good agreement. Nonetheless, the mentioned studies were developed always combining traditional VIs to LST data. Indeed, the biophysical meaning of LST was not in depth explored being often overshadowed by VIs high efficiency.

The physical basis behind the vegetation thermal behaviour explores the difference between the emissivity of the soil and the plant canopy. According to Tang and Li (2014), land surface emissivity is the effectiveness of a surface on emitting thermal radiation, and spectral emissivity is the ratio of energy radiated by a particular material to energy radiated by a black body at the same temperature. Another property of the material is the

thermal inertia (TI) which is the resistance to temperature variations; it is defined in function of the material's specific heat capacity, thermal conductivity and bulk density. Water has high TI which makes its temperature fluctuation slow when compared to other common surface materials. For this reason, water bodies can be heater than the neighbour land surface during the night. Crop canopies have higher TI than dry soil showing lower diurnal temperature variation due to the content of water in its leaf structure. Moreover, as observed by Murray and Verhoef (2007), the vegetation canopy significantly influences the soil heat flux by reducing the irradiance reaching the soil surface during the day and acting as an insulator barrier between soil and atmosphere during the nighttime.

The Land Surface Temperature products provided by sensors onboard satellites are mainly dependent on the object's albedo, emissivity, thermal inertia and exogenous factors such as wind and relief (KUENZER; DECH, 2013). However, LST is highly inverted correlated to the target moisture coupled with its intrinsic properties, therefore, its products have been used in applications for drought detection on agriculture and forests. Wan *et al*. (2004) used LST and a traditional VI combined for monitoring drought in the USA. Many other works have applied LST solely or combined with other VIs and even precipitation data for detecting and monitoring drought on a multitemporal approach (KOGAN, 1995; EZZINE *et al*., 2014).

The present work aims at mapping LCCs relying on its heterogeneous thermal properties throughout the season.For doing so, we are relying on MODIS-LST time-series analysis using a developing a conceptually and computationally straightforward method to perform the mapping task. The specific objectives are: *i)* characterize the LST temporal behavior of the LCCs; *ii)* validation through assessment of spatial agreement between LST derived classification and reference map.

MATERIALS AND METHODS

Study area and base datasets

The study focused on the Northwest, Central-North, Central-West and West mesoregions of Paraná state, which is located in the south of Brazil, however, the analysis were performed by municipality and then aggregated into microregions. The three latter regions comprise the traditional and most important grain production area of the state which occurs mostly on clay texture soils. During the last two decades the Northwest region has shown an expansion of sugarcane on areas of medium texture soils. Nonetheless, a few sugarcane expansion areas can be found on the Central-North area. Forest areas are scarce, varying in size and type, usually protected by the Federal government. Albeit being traditional grain areas, the region contains a good range of acreages for each of the LCCs tending to difficult the mapping, so that being a good area to test the methodology proposed in this work.

The reference map used is the product MCD12Q1 for the year 2011 (FRIEDL *et al*., 2002). For deepening the analysis about the relationship between soil textures and LST, we used a soil map which was elaborated by the Brazilian Agricultural Research Corporation (EMBRAPA) and is available at *www.itcg.pr.gov.br*. In the studied region there is a predominance of Latosols followed by Argisols. Soybean and maize are produced almost totally on clay soils (Argisols, Latosols, Neosols and Nitosols), sugarcane on medium texture soils (Latosols and Argisols) and Forests do not follow any pattern of occurrence regarding soil textures.

MODIS LST datasets

For the current work, a time series of LST data from September (185) to March (105) for the season of 2011/2012 were analyzed for both Terra (morning) and Aqua (afternoon) were used. The MOD11A2 (Terra) and MYD11A2 (Aqua) are 8-day compositions of LST from the daytime and nighttime, and emissivity of bands 31 (10780 – 11280 nm) and 32 (11770 – 12270 nm), all of them at 1 km nominal spatial resolution. In this study we used LST day and nighttime data from both satellites.

All the images were firstly reprojected to geographic/WGS84 using the MODIS Reprojection Tool. Then, the

files were organized for posterior processing by a Python script that was designed to compute the difference of LST between Day and Night for each 8-day image (hereafter LST-Difference). The rationale behind that is that greater temperature differences indicate dryer surfaces or lower LAI or both (SANDHOLT *et al*., 2002). As we are concerned on developing a simple methodology for mapping crops using LST, the Day datasets (hereafter LST-Day) were also analyzed solely by focusing only on the variation between days rather than intra-day. The Python script is available in our repository at *www.bitbucket.org/geopaitos/python-remotesensing*.

Temporal profiles of LST-Day and LST-Difference for LCCs

In this step, seasonal profiles of LST-Day and LST-Difference for each LCC were analyzed, there are four cases: LST-Day and LST-Difference for Terra and Aqua. In order to minimize the irradiance variation, the samples were acquired at about the same latitude (-23°40' to -25°10').

The soil classes were summarized into clay and medium texture due to the intrinsic capacity of water retention which is a key factor for heat transferring between surfaces. Within the study region, not all the LCCs occurs on the two considered soil types, so that, only sugarcane, forest and annual agriculture were analyzed for both soil classes and pasture only for the medium texture soils. Moreover, the annual agriculture on medium texture soils is seldom verified, due to this the samples may carry great uncertainty derived from the intrinsic heterogeneity and small acreage. Additionally, a special care was taken on sampling Forests, based on the assumption of its leaf area index (LAI) steadiness over a season, an ancillary images of VIs were used for selecting Forests whose VI were similar among samples despite of the soil class.

Using season statistics images for mapping LCCs

The chosen period is expected to comprehend the maximum phenology variation of the most dynamic classes (summer annual agriculture, sugarcane and pasture). The sowing and harvest of annual crops are within this period, sugarcane can present considerable variation as well, pasture and forest tend to be steadier than annual and semi-perennial agriculture. The sefactors will allow us to use time-series descriptive statistics for separating classes.

The metrics peer (LST-Day or LST-Difference) with satellite (Terra or Aqua) that yielded better results were used for mapping and hence validation. The measures adopted for assessing the LST variation for both intra and inter-daily throughout the time-series are the variance and median. In this sense, a more steady land cover class (e.g. Forest) is likely to show lower phenology variance than agriculture during a given season, due to this, variance can be a good measure to assess the phenology steadiness. Moreover, the median would separate the steady classes due to their lower and higher LST over time. For the studied season, the statistics were calculated in a pixel basis resulting in one variance (σ^2) and one median image for the season. Then, we combined these two datasets by testing thresholds for better separating the classes. The validation scheme was the simply comparison between the yielded and base maps. The outputs for the analysis are a difference map and contingency tables.

RESULTS AND DISCUSSION

LST-Day and LST-Difference temporal profiles for LCCs

The variances and medians of all time-series are presented in the Table 1 and the temporal profiles in the Figure 1. Even though not all LCCs are representative over the considered soil types, Forest-clay and Forest-medium could be used efficiently to assess the soil texture influence on the LST behaviour throughout the season. As observed in the Figure 1, for all combinations the separation between Forest and the other targets were clear and effective, moreover, the Forest-clay class usually presented lower LST-Day and LST-Difference than Forest-medium. This let us assume that clay textures lead to slower heat transferring between surfaces due to its higher water retention capacity than of the medium texture soils (MURRAY; VERHOEF,

2007). As we are not concerned on assessing the LST data quality which has been already assured (WAN, 2014), it is unnecessary to rely on meteorological stations temperature data, hence, the Forest behaviour throughout the season serves as a comparative pattern for the other classes. Considering this assumption and analyzing the plots in the Figure 1, we could observe a reasonable correlation among the other LCCs and Forests.

As observed in the Table 1, the Annual-clay class shows the higher variance for all the four cases whereas Forest classes show the lowest variances as expected. The odd output is the class Annual-medium whose variance differs considerably from the Annual-clay. However, the Annual-medium class as previously mentioned is unlikely to exist in the study area; therefore, the samples are too small due to the limited acreage. Additionally, as annual agriculture is not traditional in the medium texture area, the sowing and harvest dates do not follow a strict pattern since they are higher dependent on the rain occurrence during the expected sowing season. For these reasons, the Annual-medium class analysis is inconclusive. Other important observation is in how the variance between classes differs in each of the four cases. The peer Aqua-Difference yielded the highest amplitudes between classes for both variance and median, being so chosen for further analysis and mapping.

Some factors explain such behaviour differences between the four combinations. The use of nighttime data presents an information gain about the land surface. As reported by Wang *et al.* (2006), difference between day and night temperatures can be closely related to the surface moisture. In our case, the surface moisture comprises both soil and plant. Since health vegetation together with its LAI is also correlated with moisture, the LST-Difference can also be related to the canopy and soil moisture together. For these reasons, LST-Difference yielded better results than LST-Day.

Regarding the difference in the performance between Terra and Aqua datasets, there are some reasons for that. The Terra satellite acquire images from the land about 10:30 and 22:30 whereas Aqua at 13:30 and 1:30. This means that at the moment of the Aqua passage during the day the land received 3 more hours of irradiance than when Terra passes. Moreover, these three hours are indeed the most intensive which means closer to solar nadir. This can be observed when comparing the LST-Day plots of Terra and Aqua (Figure 1) where the latter usually presents higher values.

Variance amplitudes could clearly separate Forests (lowest variances) and Sugarcane and pasture (medium variances) from Annual agriculture (higher variances), as seen in the Table 1. Then, to effectively separate Forests from the other classes, the medians were used. As sugarcane does not represent large areas in the studied region, we simplified the final results by clumping sugarcane and pastures. Even so, four classes (Annual, Sugarcane, Pasture and Forests) could be separated by a thresholds combination of median and variance. The threshold selection was performed by trial and error yielding the following results: *1°)* $\sigma^2 > 21$ separates Annual agriculture and $\sigma^2 < 7$ for Forests and Pasture; *2°)* for the remaining LCCs we used as median thresholds the following rules: *median* < 13.5 to isolate Forests from Pastures; *3°)* and finally, *median* > 17.1 isolates Sugarcane. After these findings, we proceeded with the mapping task.

Table 13. LST time-series statistics for LCCs.
Tabela 1. Estatísticas das séries temporais de LST para as LCCs.

	Terra – Day		Terra – Difference		Aqua – Day		Aqua – Difference	
	Median	Variance	Median	Variance	Median	Variance	Median	Variance
AM	31.62	14.60	13.89	8.22	35.73	12.65	19.00	10.15
AC	30.98	26.41	13.70	19.72	36.05	32.23	19.26	33.23
SC	30.79	15.95	12.32	10.32	35.17	12.63	17.75	14.40
SM	30.84	14.87	12.37	7.65	34.63	12.32	18.28	10.86
FC	25.18	10.74	9.45	4.62	30.13	8.04	13.37	6.51
FM	27.14	11.19	7.85	5.15	26.80	8.62	11.56	3.77
PM	31.58	17.58	14.04	7.59	35.05	12.26	17.93	5.38

A = Annual, S = Sugarcane, F = Forest and P = Pasture. M = medium and C = clay.

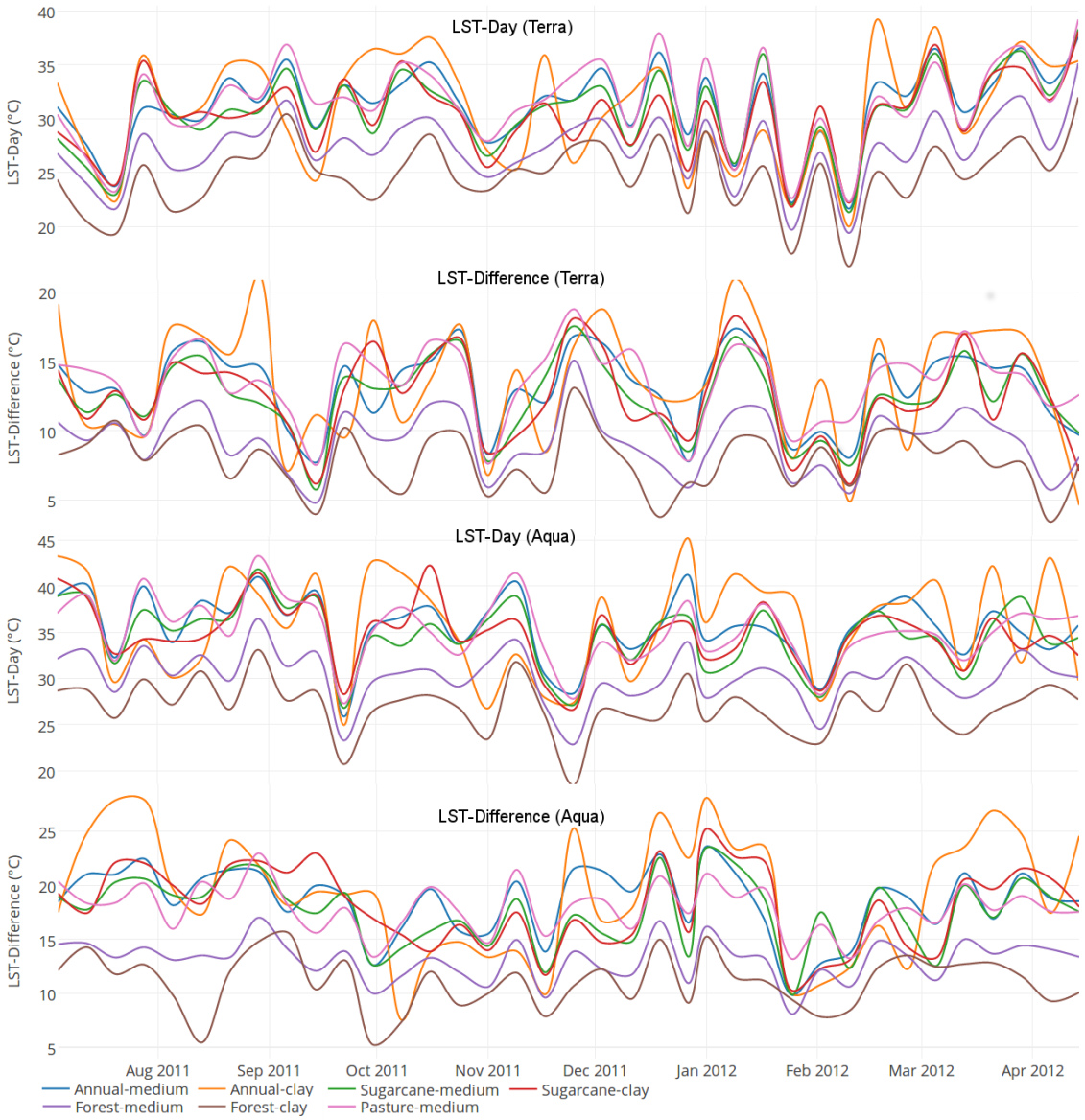

Figure 12.LST time-time series for LCCs.
Figura 1. Séries temporais de LST para as LCCs.

Mapping and performance analysis

Accuracy assessment of maps derived from coarse resolution images, as observed by Xiao *et al.* (2005) is a daunting task due to the LCCs fragmentation and spectral-mixing in the pixel. To validate the proposed method, we compared the MCD12Q1 map to the one yielded by our method. However, several factors tend to ruin such analysis. The percentage of agreement between base and produced map is not straightforward comparable. There are some municipalities where some classes are almost inexistent, then, a simple commission or omission errors lead to great percentage error. Due to this reason, we presented the yielded map (Figure 2) and a table resuming the results and errors (Table 2) aggregated by microregions.

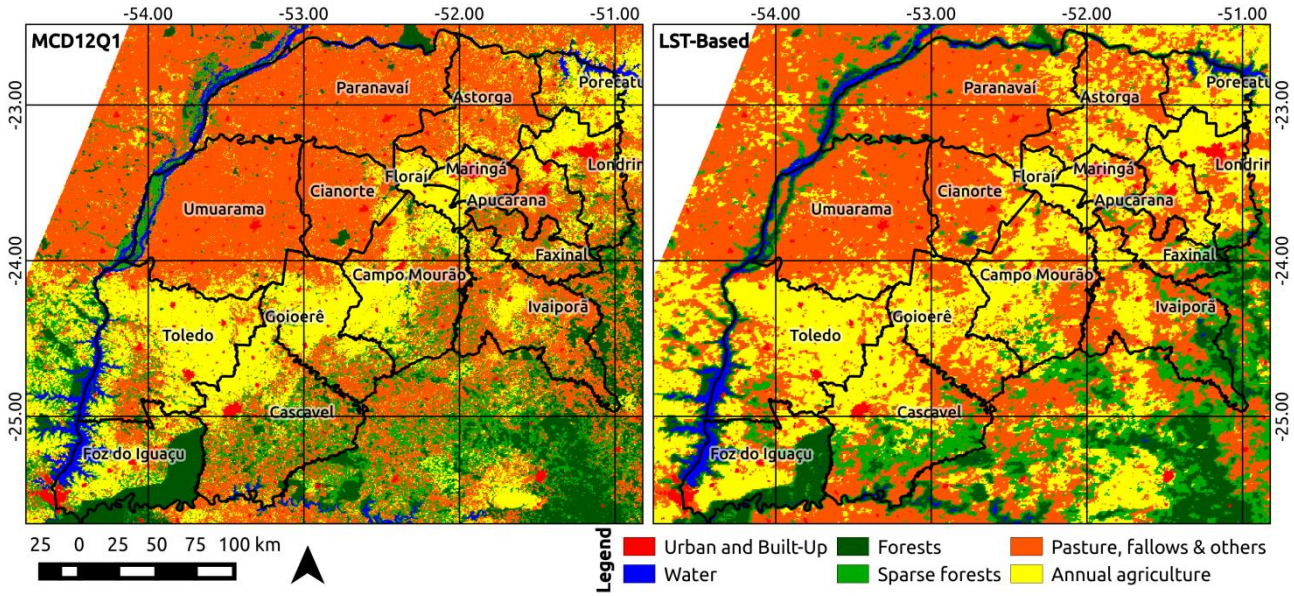

Figure 13.MCD12Q1 and LST-Based maps.
Figura 2. Mapa do MCD12Q1 e mapa gerado usando a metodologia baseada em LST.

With regards to agreement, the map derived from LST tends to fail where different classes comprise more complex mosaics (noncontiguous) or where a specific class is rare. For both cases, the reason that leads to error is the same, this is caused due to the MODIS-LST product inaptitude on detecting small areas; moreover, the surrounding targets exerts heavy influence on the central pixel by contaminating it, even more than in optical products (DENG and WU, 2013), it is a property of the heat which is always changing towards the equilibrium. In short, abrupt transitions between adjacent targets are unlikely to occur in thermal products.The Table 2 resumes the results seen in the Figure 1.

Table 14. Agreement percentage between methods for each class aggregated by microregions.
Tabela 2. Concordância percentual entre os métodos para cada classe agrupadas em microrregiões.

	Forest			Sugarcane / Pasture			Agriculture		
	MCD (ha)	LST (ha)	%	MCD (ha)	LST (ha)	%	MCD (ha)	LST (ha)	%
Paranavaí	22840.0	40810.3	178.7	909832.3	804050.1	88.4	51447.5	116206.4	225.9
Umuarama	47303.1	40711.6	86.1	871223.3	795956.0	91.4	66277.4	132669.4	200.2
Cianorte	19871.6	7950.0	40.0	344877.6	319414.7	92.6	40503.0	74686.9	184.4
Goioerê	57232.7	17074.7	29.8	165524.8	194016.3	117.2	261911.6	234688.8	89.6
C. Mourão	107380.8	41876.3	39.0	255463.9	258924.6	101.4	337962.2	348733.8	103.2
Astorga	12613.5	9311.2	73.8	393686.5	278243.8	70.7	98698.3	214565.3	217.4
Porecatu	15639.3	8720.2	55.8	60916.8	65209.9	107.0	136081.1	120575.5	88.6
Floraí	2088.9	353.4	16.9	13700.1	26425.6	192.9	113076.1	63058.9	55.8
Maringá	5336.2	775.7	14.5	56504.5	57190.4	101.2	79685.6	91071.5	114.3
Apucarana	21535.1	13884.3	64.5	127710.8	90291.9	70.7	69341.3	116257.0	167.7
Londrina	47651.7	20012.1	42.0	126620.9	124465.7	98.3	157030.9	173364.2	110.4
Faxinal	40840.4	35521.4	87.0	137929.8	95996.2	69.6	47000.7	74137.8	157.7
Ivaiporã	124476.0	102551.8	82.4	358846.5	316367.0	88.2	130136.5	153473.2	117.9
Toledo	81780.9	40390.1	49.4	177842.2	243357.5	136.8	555513.4	494837.5	89.1
Cascavel	264962.1	219036.4	82.7	328894.6	291915.6	88.8	235051.9	296879.2	126.3
Foz do Iguaçu	237976.0	55326.4	23.2	88487.0	103256.6	116.7	165158.1	182794.9	110.7

Note: the percentages are calculated considering MCD acreage as reference.

In the Table 2, percentages above and below100 % represents overestimation and underestimation by the

LST method, respectively. From the Table 2 analysis, it is clear that the microregions whose texture soil is medium, such as Paranavaí, Umuarama, Cianorte and Astorga, presented lower agreement for agriculture. This is a case II situation; so, simple commission or omission errors lead to low agreement. For regions which show high Forest fragmentation such as Maringá, Floraí and Goioerê the LST method fails by underestimating such small isolated areas. From the Figure 2 and Table 2 coupled analysis, we can conclude that the LST derived method performance is highly dependent on the adjacency and acreage of the LCCs. In short, the more aggregated the areas, the less susceptible to contamination, the more efficient is the LST derived method.

CONCLUSIONS

We proposed a method for mapping LCCs using only time-series analysis of MODIS-LST data. Then, to separateclasses we relied on the calculation of the time-series variance and median which are effective in representing the time-series unsteadiness and central trend, respectively. This simple statistic applied to the thermal domain were not yet in depth explored, thus, the present work brought new findings about temperature and vegetation relationships over time. The proposed method does not proves itself on being more efficient on mapping agriculture than traditional VIs methods, however, on achieving good results, it demonstrates that thermal data can effectively translate biophysical attributes of vegetation using a different physical basis. The analysis showed that Forests has lower seasonal temperature variation than annual agriculture, pasture and sugarcane stay in the middle of this variance scale. So, for future works, Forests can be used as a proxy for assessing the climate variation over an area since it shows lower variance than other targets, this helps to dissociate the thermal variability into climatic and phenology components.

The analysis let us conclude that the difference between daytime and nighttime temperatures is more suitable than daytime solely on dissociating land cover classes. Further, MODIS-Aqua data were more responsive to intra-daily temperature variation than MODIS-Terra due to the time of satellites passage over the area. Day temperature is usually higher at 13:30 than 10:30 due to the accumulated irradiance, for this reason Aqua showed better results and were then used for mapping the area.

As main drawback, the method performance is strongly influenced by neighbour contamination on a central pixel, so, the boundaries between classes are not so clear when compared to optical remote sensing. This is due to the low spatial resolution of the sensor and the fact that heat is physically always flowing towards equilibrium. Indeed, despite the shortcomings, the proposed method uses non-traditional datasets for mapping vegetation classes achieving good results depending on the target. Moreover, this approach can be combined to traditional methods yielding better results and furthering the biophysical relations between vegetation and the electromagnetic spectra.

REFERENCES

DENG, C.; WU, C. Examining the impacts of urban biophysical compositions on surface urban heat island: A spectral unmixing and thermal mixing approach. *Remote Sensing of Environment*, v. 131, p. 262-274, 2013.

EZZINE, H. *et al.* Seasonal comparisons of meteorological and agricultural drought indices in Morocco using open short time-series data. *International Journal of Applied Earth Observation and Geoinformation*, v. 26, p. 36-48, 2014.

FRIEDL, M. *et al.* Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, v. 83, n. 1-2, p. 287-302, 2002.

KOGAN, F. Application of vegetation index and brightness temperature for drought detection. *Advanced Space Research*, v. 15, n. 11, p. 91-100, 1995.

KUENZER, C.; DECH, S. (Org.). Thermal Infrared Remote Sensing. Dordrecht: Springer Netherlands, 2013. v. 17. 537 p.

LAMBIN, E.F.; EHRLICH, D. The surface temperature-vegetation index space for land cover and land-

cover change analysis. *International Journal of Remote Sensing*, v. 17, n. 3, p. 463-487, 1996.

MURRAY, T.; VERHOEF, A. Moving towards a more mechanistic approach in the determination of soil heat flux from remote measurements. Agricultural *and Forest Meteorology*, v. 147, n. 1-2, p. 80-87, 2007.

NEMANI, R.; RUNNING, S. Land cover characterization using multitemporal red, near-IR, and thermal-IR data from NOAA/AVHRR. *Landscape Parametrization*, v. 7, n. February, p. 79-90, 1997.

SANDHOLT, I. *et al.* A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sensing of Environment*, v. 79, n. 2-3, p. 213-224, 2002.

TANG, H.; LI, Z.-L. Quantitative Remote Sensing in Thermal Infrared. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. 281 p.

TUCKER, C.J. Red and photographic infrared linear combinations for monitoring vegetation.*Remote sensing of Environment*, v. 8, n. 2, p. 127-150, 1979.

WAN, Z. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sensing of Environment*, v. 140, p. 36-45, 2014.

WAN, Z. *et al.* Using MODIS Land Surface Temperature and Normalized Difference Vegetation Index products for monitoring drought in the southern Great Plains, USA. *International Journal of Remote Sensing*, v. 25, n. 1, p. 61-72, 2004.

WANG, K.; LI, Z.; CRIBB, M. Estimation of evaporative fraction from a combination of day and night land surface temperatures and NDVI: A new method to determine the Priestley–Taylor parameter. *Remote Sensing of Environment*, v. 102, n. 3-4, p. 293-305, 2006.

WARDLOW, B.D.; EGBERT, S.L. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment*, v. 112, n. 3, p. 1096-1116, 2008.

XIAO, X. *et al.* Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sensing of Environment*, v. 95, n. 4, p. 480-492, 2005.