Sugarcane Identification Through Time-Series of Landsat and ERS/ENVISAT Data

Lorenzo Iannini, Ramses Molijn, Arjan Tabak, Ali Mousivand, Ramon Hanssen

Department of Geoscience and Remote Sensing, Delft University of Technology
Stevinweg 1, 2628CN, Delft, The Netherlands

Abstract. The paper debates a novel approach for sugarcane identification and characterization based on multi-spectral and multi-temporal profile matching. The model accounts for asynchronous crop growth behavior and vigour diversity through the estimation of time offset parameter, and indicator scale and offset coefficients. The mathematical framework of the technique is first discussed in a generic application perspective. The approach is then applied to Landsat 5 TM and ERS/ENVISAT SAR time-series over the Orindiuva area attaining preliminary promising although perfectible results.

Keywords: sugarcane detection, optical and radar data assimilation, time-series processing .

1. Introduction

From the beginning of last decade sugarcane in Brazil experienced a dramatic expansion as a result of governmental politicies in favor of ethanol production as main alternative to fossil fuels. Sugarcane for ethanol occupies about 0.6% of the total area, about 1.7% of the total current agricultural area, in Brazil. From satellite images, the sugarcane expansion between 2000 and 2010 occurred mostly over pastures (70%), followed by annual crops (25%), citrus (1.3%), forest (0.6%) and sugarcane land under crop rotation (3.4%) (WALTER et al., 2014). Such phenomenon demanded the monitoring of direct and indirect Land Use Change (LUC) in order to identify and quantify LUC impacts such as greenhouse gas emissions, food security, biodiversity and stresses on soil and water conditions. Various works have been carried out on sugarcane classification utilizing spatio-temporal profiles based on optical data (mainly using SPOT and/or Landsat imagery) (ABDEL-RAHMAN; AHMED, 2008; RUDORFF et al., 2010; VIEIRA et al., 2012). One of the main works on mapping sugarcane fields in Brazil was carried out under the so-called Canasat project, (RUDORFF et al., 2010) a joint initiative by the Brazilian space agency (INPE) and several sugarcane industry and research related parties. It was found that this methodology was most accurate over large areas as compared to automated and semi-automated classifications, though no accuracy comparisons were given (AGUIAR et al., 2011). Generally from these works it was found that the discrimination between sugarcane and other semi- or perennial crops can be challenging, and this holds in particular for pastures, because of the similar temporal-spectral behavior. Furthermore, most of the approaches proposed for land use retrieval focus on a single remote sensing dataset or/and are not very applicable in a fast and automatized way to larger areas, either because the indicators cannot properly address the variability in the data or because human intervention represents a major component.

The aim of this work is that to move the first theoretical and practical steps towards an approach capable of encompassing multiple sensors with multiple temporal acquisitions and at the same time cope with data and crop variabilities. This latter represents indeed valuble information to be extracted through the proposed profile-based parametric model rather than nuisance. The choice was that of focusing on medium-high resolution sensors, in order to neglect, at least at this early stage, about mixed pixel effects while preserving good geometrical uderstanding of the land features. The Landsat 5 TM and the C-Band ERS/ENVISAT radar archive imagery were basing on an operative perspective. Whereas Landsat has been largely used for sugarcane caracterization (RUDORFF et al., 2010), only a limited amount of studies has been carried out on radar, mainly focused on how SAR signals are affected by sugarcane height
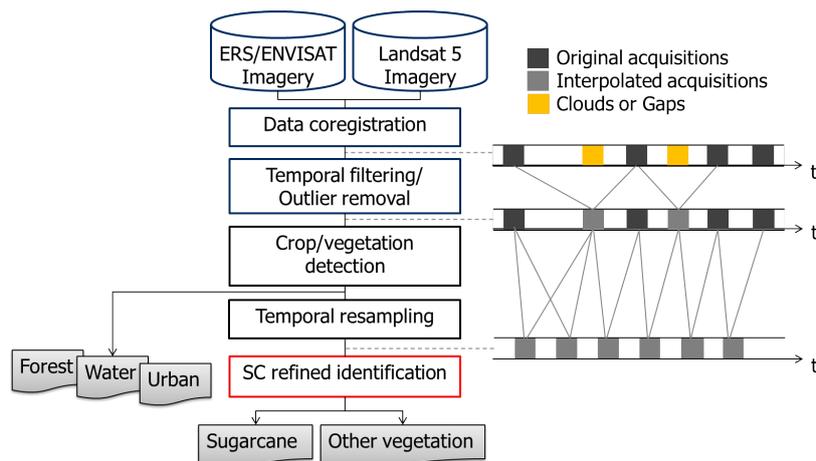
Figura 1: Flowchart of the operative framework proposed for sugarcane tracking

and harvest events (BAGHDADI et al., 2009; LIN et al., 2009), with contrasting sensitivity results. The role of SAR is deemed especially in those region where frequent cloud-coverage condition can hamper optical-base algorithms. The choice of a C-Band sensor has its rationale in view of the Satellite-1 sensor that will soon provide short revisit and wide coverage.

2. Methodology

The theoretical and operative aspects of the proposed methodology are here discussed. The former account for the mathematical formalization of the multi-temporal and multi-band model conceived for the detection and characterization of sugarcane fields. The operative flowchart, shown in Fig. 1, addresses the following preprocessing steps:

1. the initial stack coregistration

2. a temporal filtering procedure for noise and gap removal

3. a first classification step that assigns pixels to broad LU classes

4. a data filtering procedure that resamples the time-series on a uniform grid.

As further explained in the section, the implementation and rationale of steps 2-4 is that of pursuing computational efficiency within a matricial framework rather than increasing the algorithm classification performance. It is also important to remark that both the filtering and classification steps are performed on a per-pixel basis. The whole processing chain has been conceived and implemented within the MATLAB suite framework. The coregistration of the LS and SAR stacks has been carried out by visual identification of point correspondences. Such common ground control points (GCPs) have been used to estimate a first-order polynomial warping function, in both latitude and longitude, among the datasets by means of LSQ residual minimization technique. The six optical bands plus the derived NDVI and the SAR image set, so far arranged into stacks with size respectively $N_{lat} \times N_{lon} \times I_{LS}$ and $N_{lat} \times N_{lon} \times I_{SAR}$ in the lat,lon,image dimensions are then reshaped and temporarily stored in the 2D arrays with size $N_{pix} \times I_{LS}$ and $N_{pix} \times I_{SAR}$ with $N_{pix} = N_{lat} \cdot N_{lon}$. This allows to achieve greater computational efficiency since some of the stack filtering operations can be later performed through matricial algebra. Let then introduce the notation

$$y_p(i) \quad p = 1,..,N_{pix} \quad i = 1,...,I_{LS/SAR} \qquad (1)$$

to refer to the value stored by the generic indicator $y \in \{b_1....b_6, b_{NDVI}, b_{SAR}\}$ in the p-th pixel of the i-th stack acquisition. As simply sketched in Fig. 1, both the optical and SAR acquisitions are available on an irregular temporal grid. The sparsity depends in the optical case by extended cloud coverage, and in the case of SAR by the temporal gaps between the acquisitions. Some cloudy areas (in LS imagery) and void areas due to partial swath coverage (in SAR imagery) still need to be detected and properly treated. The devised approach consists in replacing such 'corrupted' samples with values interpolated from the closest valid acquisitions, time-wise. In such a way no pixel-dependent masking must be applied to the stack, but rather this latter can be processed uniformly. Since the sake for optimality in the interpolation kernel is out of the scope of this work, a simple linear interpolator has been adopted. The Fmask channel [ref] provided together with the LS reflectances has been used as cloud mask.

A first rough classification is then applied to the stack to divide the scene among 4 broad classes: urban, water, dense vegetation and shallow vegetation. The dense vegetation class mainly includes forest areas, whereas the shallow vegetation comprises agricultural areas, thus including sugarcane, pasture land and other shrublands/native vegetation areas. The classification is performed through supervised ML approach. The training set has been delineated by using the Canasat maps for the selection of the sugarcane fields and by visual inspection for the other classes. The following indicators extracted from the time-series are used by the classifier:

- SWIR mean: $\mu_{b_6} = \langle b_6 \rangle_i$

- NDVI mean: $\mu_{b_{NDVI}} = \langle b_{NDVI} \rangle_i$

- NDV standard deviation: $\sigma_{b_{NDVI}} = std_i(b_{NDVI})$

- SAR mean: $\mu_{b_{SAR}} = \langle b_{SAR} \rangle_i$

where the nomeclature $\langle . \rangle_i$ and std_i recall respectively the ensemble average and standard deviation operators along time. Since the 4 LU classes have been reasonably assumed stable within the investigated period, the operators have been applied over the full 4-year period. A quasi-ML classifier has been then applied at per-pixel assuming gaussian statistics with null covariance between the indicators. Figure 2 shows the estimated LU results around the Orindiuva area, here reported just on a qualitative level, with a close-up view on the area surrounding the city of Fronteira. It shall be remarked that the use of SAR brought significant benefits in the discrimination of agricuture/shallow vegetation from urban/anthropogenic areas, although room for further improvements can be easily foreseen with more elaborate indicators ad procedures. The reader should in fact consider that the aim of this first step is that of crop area detection. A more elaborate processing will be indeed applied only on such pixels, allowing to save computational resources, especially on densely forested or urbanized areas.

Before proceeding, it is necessary to register all the observations to the common uniformly sampled grid

$$T_k = t_0 + k \cdot \triangle t, \quad 0 \leq k \leq M - 1 \tag{2}$$

with T spanning the range $T_{span} = (M - 1)\triangle t = T_{year} + T_w$, where $T_{year} = 365$ days and T_w is the width of the time window explored in order to find the crop (in this case sugarcane) growth date, as later explained. The choice of the initial time t_0 and of the vector length M depends on the variability of the crop growth period, whereas the interval $\triangle t$ is arguably related to the density of the observations and to the smoothness of the indicator time-series. In this work the same temporal grid as the LS acquisitions, thus $\triangle t = 16$ days, has been adopted to
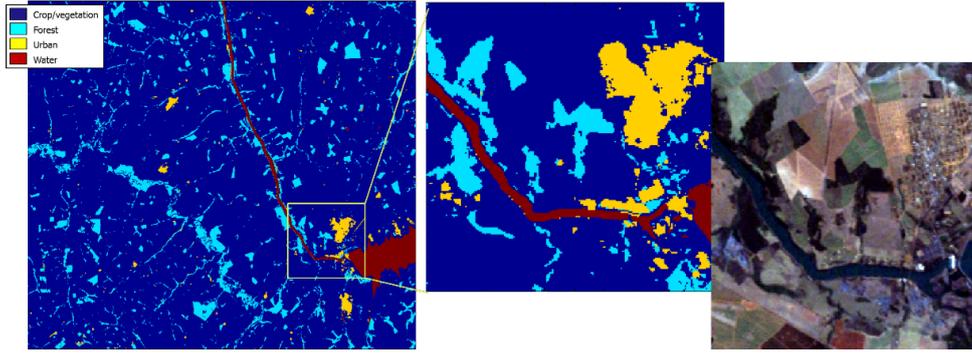
Figura 2: (Top) Classification map obtained from joint optical and SAR ML classifier. (Bottom) Close up of the LU map on the area surrounding the Fronteira city.

reduce interpolation artifacts of optical indicators. Such operation leads to the interpolated data sequence.

$$\mathbf{z}_{p,i} = \left[\begin{array}{ccc} z_{p,i}(T_1) & \cdots & z_{p,i}(T_M) \end{array} \right]^T. \tag{3}$$

for the generic i-th indicator.

2.1. The Crop Model

The mathematical foundation of the methodology consists on the assumption that, after proper filtering and interpolation, the sequences from crops of the same class share a similarity with respect to a temporal offset τ, an amplitude scaling coefficient A and an amplitude offset coefficient B in accordance with

$$z_{p,i}(t) = A_{p,i} \cdot r_i(t - \tau_p) + B_{p,i} + n_{p,i}(t - \tau_p) \tag{4}$$

where r_i is the normalized reference profile for crop type k defined over the time support $\Omega_r = [0, (M_r - 1)\triangle t]$ and comprised within the range:

$$0 = \min(z) \leq z \leq \max(z) = 1. \tag{5}$$

In a vector form, where more indicators are jointly accounted, (4) becomes

$$\mathbf{Z}_p(t) = \mathbf{A}_p \circ \mathbf{R}(t - \tau_p) + \mathbf{B}_p + \mathbf{N}_p(t - \tau_p) \tag{6}$$

with $\mathbf{Z}_p(t) = \left[\begin{array}{ccc} z_{p,1}(t) & \cdots & z_{p,N_I}(t) \end{array} \right]^T$, $\mathbf{A}_p = \left[\begin{array}{ccc} A_{p,1} & \cdots & A_{p,N_I} \end{array} \right]^T$, $\mathbf{R}(t) = \left[\begin{array}{ccc} r_1(t) & \cdots & z_{N_I}(t) \end{array} \right]^T$, $\mathbf{B}_p = \left[\begin{array}{ccc} B_{p,1} & \cdots & B_{p,N_I} \end{array} \right]^T$, $\mathbf{N}_p(t) = \left[\begin{array}{ccc} n_{p,1}(t) & \cdots & n_{p,N_I}(t) \end{array} \right]^T$, N_I standing for the number of indicators and $\circ$ recallingthe entry-wise (Hadamard) product. The noise on the profiles is assumed to behave as a zero-mean variable with covariance matrix

$$\mathbf{C}_n = E\left[\mathbf{N}_p(t)\,\mathbf{N}_p^T(t)\right] = \left[\begin{array}{ccc} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_{N_I}^2 \end{array} \right] \tag{7}$$

thus stationarity in time. Similarly, the variability in coefficients A, B and τ is defined through their covariances

$$\mathbf{C}_A = E\left[(\mathbf{A}_p - \mu_{\mathbf{A}})(\mathbf{A}_p - \mu_{\mathbf{A}})^T\right], \quad \mathbf{C}_B = E\left[(\mathbf{B}_p - \mu_{\mathbf{B}})(\mathbf{B}_p - \mu_{\mathbf{B}})^T\right] \tag{8}$$
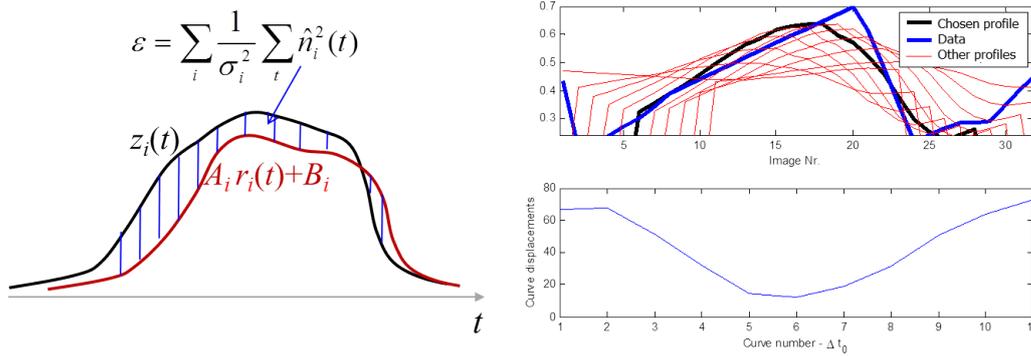
Figura 3: (**left**) Schematic representation of the error between fitted model and data time-series. (**top-right**) Results on a real data case. The fitted models for different τ parameters are reported. (**bottom-right**) Error (displacement) between data and model. The minimum is attained by the 6th curve, plotted in black in the graph above.

with $\mu_{\mathbf{A}} = E\left[\mathbf{A}\right], \mu_{\mathbf{B}} = E\left[\mathbf{B}\right]$. Notice that they are addressed as full matrices, i.e. the correlation between different indicators is accounted. The degree of similarity between the crop time-series z and the class reference r can then be defined through the quadratic distance

$$d_p = \hat{\mathbf{A}}_p^T(t;\hat{\tau})\mathbf{C}_A^{-1}\hat{\mathbf{A}}_p(t;\hat{\tau}) + \hat{\mathbf{B}}_p^T(t;\hat{\tau})\mathbf{C}_B^{-1}\hat{\mathbf{B}}_p(t;\hat{\tau}) \tag{9}$$

where $\hat{A}$, $\hat{B}$, $\hat{\tau}$ stand for the parameter estimates from the time-series. The latter is only indirectly involved in the metric calculation. Its estimation is however fundamental for the correct retrieval of the other coefficients. It is achieved through

$$\hat{\tau}_p = \underset{\tau}{\operatorname{argmin}}\left(\frac{1}{N_r}\sum\hat{\mathbf{n}}_p^T(t)\mathbf{C}_n^{-1}\hat{\mathbf{n}}_p(t)\right) \tag{10}$$

where $\hat{\mathbf{n}}_p(t) = \hat{\mathbf{Z}}_p(t) - \mathbf{Z}_p(t)$, with $\mathbf{Z}_p(t) = \begin{bmatrix} \hat{z}_{p,1}(t) & \cdots & \hat{z}_{p,N_I}(t) \end{bmatrix}^T$ with $\hat{z}_{p,i}(t)$ defined as

$$\hat{z}_{p,i}(t) = \hat{A}_{p,i}r_i(t - \hat{\tau}_p) + \hat{B}_{p,i} \tag{11}$$

i.e. the reconstructed data from the model estimates. The non-linear system in (10) is solved through exhaustive exploration of the τ domain by regular intervals of 16 days within a $T_w = 120 days$ window. For each τ, $\hat{A}$ and $\hat{B}$ are then retrieved by Weighted Least Squares (WLSQ) method from (6). As shown in Fig. 3 the best candidates are the ones providing the best fit with the data, i.e. minimizing (10).

The sugarcane identification will hence be based upon the set of indicators:

- N_I estimates $\hat{A}_p$

- N_I estimates $\hat{B}_p$

The time offset τ is not included in (8) since significant diversity is expected among the different sugarcane fields. The algorithm requires a calibration/training phase where the crop intrinsic parameters C_A, C_B and $r(t)$ are estimated from the data at disposal. The estimation procedure is carried out by means of an iterative optimization scheme. The procedure alternates the estimates of the $\{A, B, \tau\}$ coefficients with that of the signature $\{r, \mathbf{C}_A, \mathbf{C}_B\}$, all consistently attained in a WLSQ framework. The steps are iterated until a covergence on the parameters is reached.
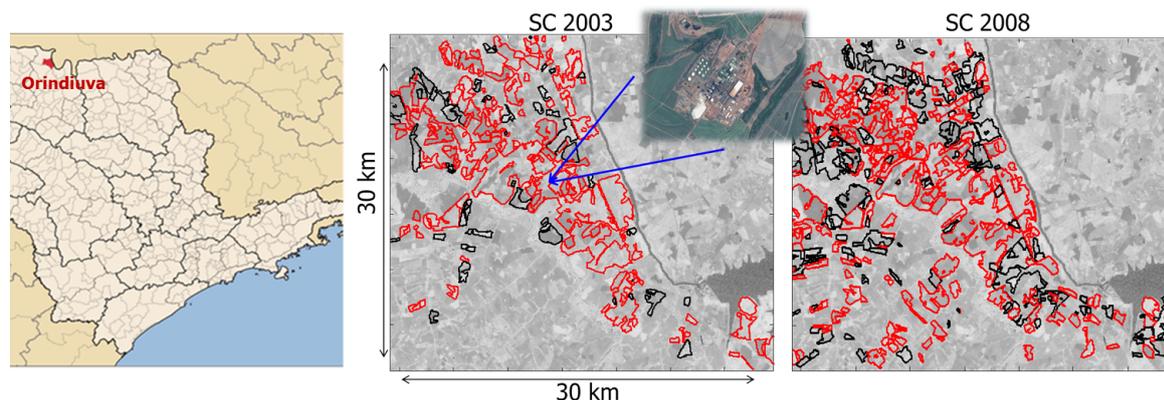
Figura 4: (**left**) Location of the Orindiuva municipality within the Sao Paolo state. (**Center-Left**) Frame of the test area showing the sugarcane fields (ratoon in red, renovated in black) for years 2003 and 2008 as extracted from the Canasat maps, overlayed to NDVI maps. The Moema mill location is also highlighted.

Tabela 1: Overview on the datasets selected for the work in the 2003-2006 period

Name	Description	Resolution	Availability
Landsat 5 TM	Multispectral optical sensor	30 m	1 every 16 days
ASAR/ERS SAR	C-Band radar sensor	4 x 20 m	40 images total
Canasat maps	Shapefiles of sugarcane fields	-	1 per year from 2003

3. Results and discussion

The area used for a preliminary assessment of the designed methodology covers the Orindiuva municipality and parts of neighboring municipalities for a frame extension of 30 x 30 km as shown in Fig. 4. The huge mill of Moema at the center of the municipality drove over the last decade a rapid sugarcane expansion especially in the surrounding municipalities. It should be in fact remarked that the acreage increase in the Orindiuva area is less pronounced, being this latter largely involved in 2003 in sugar production. As reported in Table 1, the remote sensing imagery and the land use information so far accounted comprise Landsat 5 TM calibrated time-series, ERS,ENVISAT SAR imagery and the Canasat maps. Since Canasat maps are only available in the timeframe from 2003 onwards whereas SAR data availability becomes critical after 2006, the time span 2003-2006 has been chosen for this preliminary assessment.

The pre-processing steps described in section 2 have been carried out on the two remote sensing datasets. The algorithm calibration procedure, aimed to extract the sugarcane signature, has been applied to a randomly selected set of samples from the 2004/2005 season basing on the corresponding Canasat map. The resulting reference profiles, reported in Fig. 5, register a clear pattern for most of the optical bands and for the derived NDVI, whereas in the case of SAR, although a trend is visible, its excursion is lower than the standard deviation. It must be recorded that such trend, with an increase in backscatter at the beginning of the growing season, is consistent over different years. This could be possibly due to the increase of the scene water content during the rain season but also to some particular sensitivity of radar to early sugarcane stages. Further investigation in this regard is conducted at the moment by the authors within the framework of ESA/CSA SOAR-EU-16850 project. The reference sugarcane signature, based on the 2004/2005 training process, is hence used to compute the metric (8) on
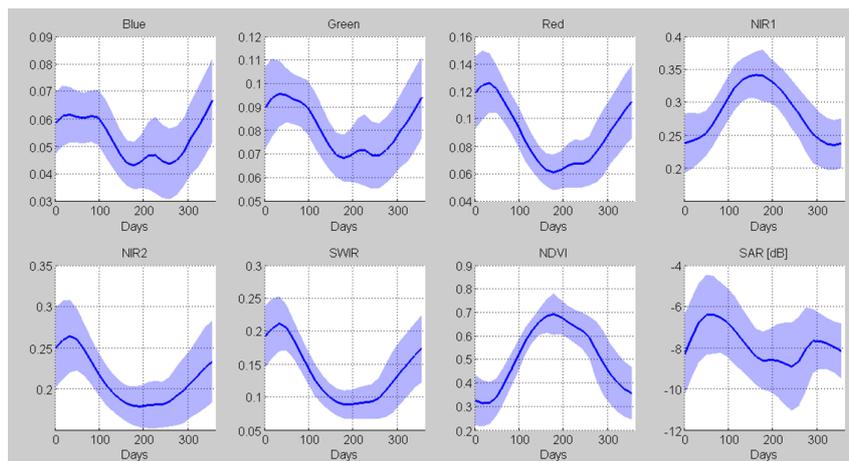
Figura 5: Estimated reference profiles for the 6 different LS bands, the NDVI indicator and the SAR backscatter. The profiles are based on 2004/2005 growth cycle. Mean value (blue line) and standard deviation (blue stripe) are represented.
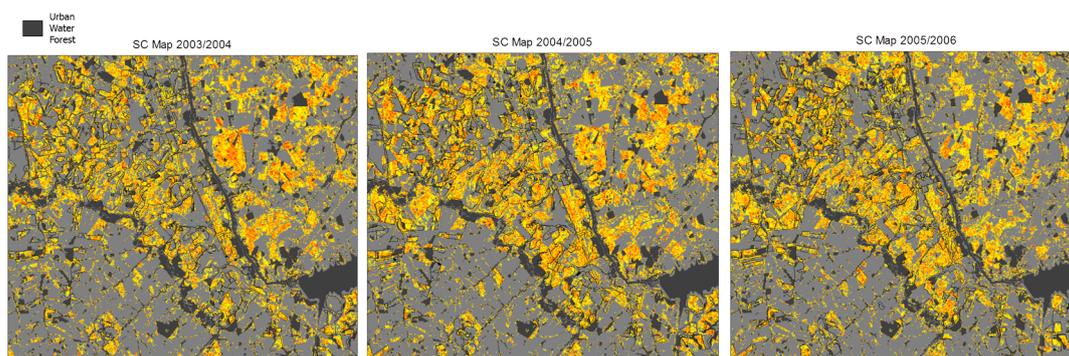
Figura 6: Sugarcane similarity maps produced by the application of the model, trained on the 2003/2004 season, to the whole dataset. The coloraxis is reversed with respect the error in 8. Red values correspond to maximum likelihood, yellow to good similarity, gray to low similarity. The shapes of the Canasat fields are also visible in black.

the crop pixel subset over the whole timeframe. The output of this last part of the approach, for a specific set of indicators, is therefore a sugarcane likelihood map at yearly basis. Different combinations of indicators have been tested with very contrasting results, which will be more insighfully analyzed at the symposium. The introduction of the SAR in the methodlgy so far had detrimental effects on the results, confirming that a deeper insight on the backscatter and consequently a better data exploitation (e.g detection and use of only suitable images), should be pursued. The indicators that bear the best performace are the SWIR band and the NDVI and their joint use. The results from this latter combination is qualitatively illustrated in Fig. 6. The values show a good correlation with the Canasat maps, although some significant discrepancy, both in terms of false positives and undetected fields can be found. The performance of such set is then represented quantitatively in Fig. 7 as a function of the selected threshold on d_p. The two probabilities, false alams and missing detection, cross their paths at 65
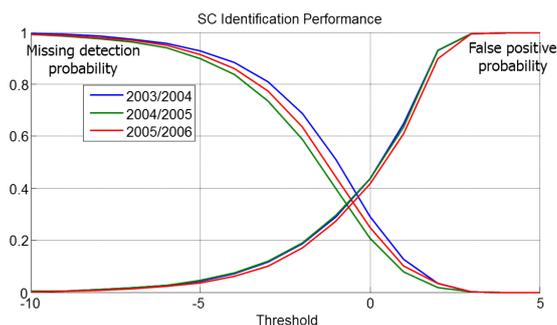
Figura 7: False alarm rate and missing detection probability by training the algorithm with the 2003/2004 imagery.

4. Acknowledgemts

The authors would like to acknowledge the European Space Agency (ESA) for providing the SAR data under the framework of project C1P.16849 and the Brazilian Space Agency for kindly supporting our research by providing the Canasat maps of the Sao Paolo region.

Referências

ABDEL-RAHMAN, E. M.; AHMED, F. B. The application of remote sensing techniques to sugarcane production: a review of the literature. *International Journal of Remote Sensing*, v. 29, n. 13, p. 3753–3767, jul. 2008.

AGUIAR, D. A. et al. Remote sensing images in support of environmental protocol: Monitoring the sugarcane harvest in sao paulo state, brazil. *Remote Sensing*, v. 3, n. 12, p. 2682–2703, 2011.

BAGHDADI, N. et al. Potential of sar sensors terrasar-x, asar/envisat and palsar/alos for monitoring sugarcane crops on reunion island. *Remote Sensing of Environment*, v. 113, n. 8, p. 1724–1738, ago. 2009.

LIN, H. et al. Monitoring sugarcane growth using envisat asar data. *Geoscience and Remote Sensing, IEEE Transactions on*, v. 47, n. 8, p. 2572–2580, Aug 2009.

RUDORFF, B. F. T. et al. Studies on the rapid expansion of sugarcane for ethanol production in sao paulo state (brazil) using landsat data. *Remote Sensing*, v. 2, n. 4, p. 1057–1076, 2010.

VIEIRA, M. A. et al. Object based image analysis and data mining applied to a remotely sensed landsat time-series to map sugarcane over large areas. *Remote Sensing of Environment*, Elsevier Inc., v. 123, p. 553–562, ago. 2012.

WALTER, A. et al. Brazilian sugarcane ethanol: developments so far and challenges for the future. *Wiley Interdisciplinary Reviews: Energy and Environment*, v. 3, n. 1, p. 70–92, jan. 2014.