NEAR REAL TIME YIELD ESTIMATION FOR SUGARCANE IN BRAZIL COMBINING REMOTE SENSING AND OFFICIAL STATISTICAL DATA

Marcio Pupin Mello[1,], Clement Atzberger[2], Antonio Roberto Formaggio[3]*

[1]The Boeing Company
Boeing Research & Technology - Brazil (BR&T-Brazil)
Estrada Dr Altino Bondesan 500 (Parque Tecnologico), Sao Jose dos Campos - SP, Brazil 12247-016
* marcio.p.mello@boeing.com

[2]University of Natural Resources and Life Sciences (BOKU)
Institute of Surveying, Remote Sensing and Land Information (IVFL)
Peter Jordan Straße 82, Vienna, Austria 1190

[3]National Institute for Space Research (INPE)
Remote Sensing Division (DSR)
Avenida dos Astronautas 1758, Sao Jose dos Campos - SP, Brazil 12227-010

ABSTRACT

Besides energy concerns, the cultivation of sugarcane in Brazil plays an important role in the development of its agriculture, economy, and environment. However, there is no research on a broad operational yield monitoring system for sugarcane in Brazil based on remote sensing technologies. This paper aims at proposing a method to fill this gap. Five municipalities located in the west corner of São Paulo State, Brazil, were used to run and test a new method, which estimates the yield based on a combination of remote sensing and official historical data without using crop masks. Using data from 2003 to 2012, results stated that sugarcane yield estimates from January to March using remote sensing data tends to match strongly with the official yield, with the benefit of being known before the harvest season. Although we only tested five municipalities, we do expect that the proposed method might be applicable to wider regions as well as to other crops.

Index Terms— IBGE, Canasat Project, agricultural statistics, yield forecast, image processing, energy crop, biofuel, sustainability.

1. INTRODUCTION

Although very much debated, agriculture has been pointed out as the main alternative to replace fossil-based fuel demand mitigating Green House Gas (GHG) emissions in the next decades through the use of bioenergy at a commercial-global scale [1]. In Brazil, sugarcane has been used as feedstock for ethanol production since the mid-1970's, when the first "oil crisis" lead the Brazilian government to develop and implement a broad nationwide research program named *PROÁLCOOL*, which stands for National Alcohol Program, in Portuguese [2]. As a consequence, Brazil is nowadays the world's largest sugarcane producer and, with the advent of flex-fuel cars in 2003, it has also experienced a rapid growth of sugarcane areas for ethanol production over the last decade [3].

Besides energy concerns, the cultivation of sugarcane in Brazil plays an important role in the development of its agriculture, economy, and environment [4], [5]. Given the characteristics of the sugarcane plantations in the South-central region of Brazil, remote sensing images have been used by the Canasat Project [6] since 2003 to elaborate accurate annual thematic maps of cultivated sugarcane areas in this region [3], [7]. Furthermore, the Canasat Project identifies, since 2006, harvested sugarcane areas with and without pre-harvest straw burning in São Paulo State at monthly intervals [8], [9]. Canasat also assessed, in 2012, changes in land use and cover due to the recent sugarcane expansion in the South-central region [10]. However, despite of these efforts, there is no research on a broad operational yield monitoring for sugarcane in Brazil based on remote sensing technologies.

This paper aims at proposing and testing a remote sensing based method to estimate sugarcane yield at a municipal scale through the combination of remote sensing and official historical data in São Paulo State, Brazil. The proposed method draws on earlier studies described by Kastens et al. [11] and Atzberger [12].

2. MATERIALS AND METHODS

The study area is located in the west corner of São Paulo State. The state is responsible for more than 50% of the national sugarcane acreage according to the Brazilian Institute of Geography and Statistics (IBGE [13]), which is the official agency for agricultural statistics in Brazil. Five municipalities were selected to run the analyses: *Caiuá*, *Marabá Paulista*, *Presidente Venceslau*, *Santo Anastácio*, and *Teodoro Sampaio*, as highlighted in Fig. 1.

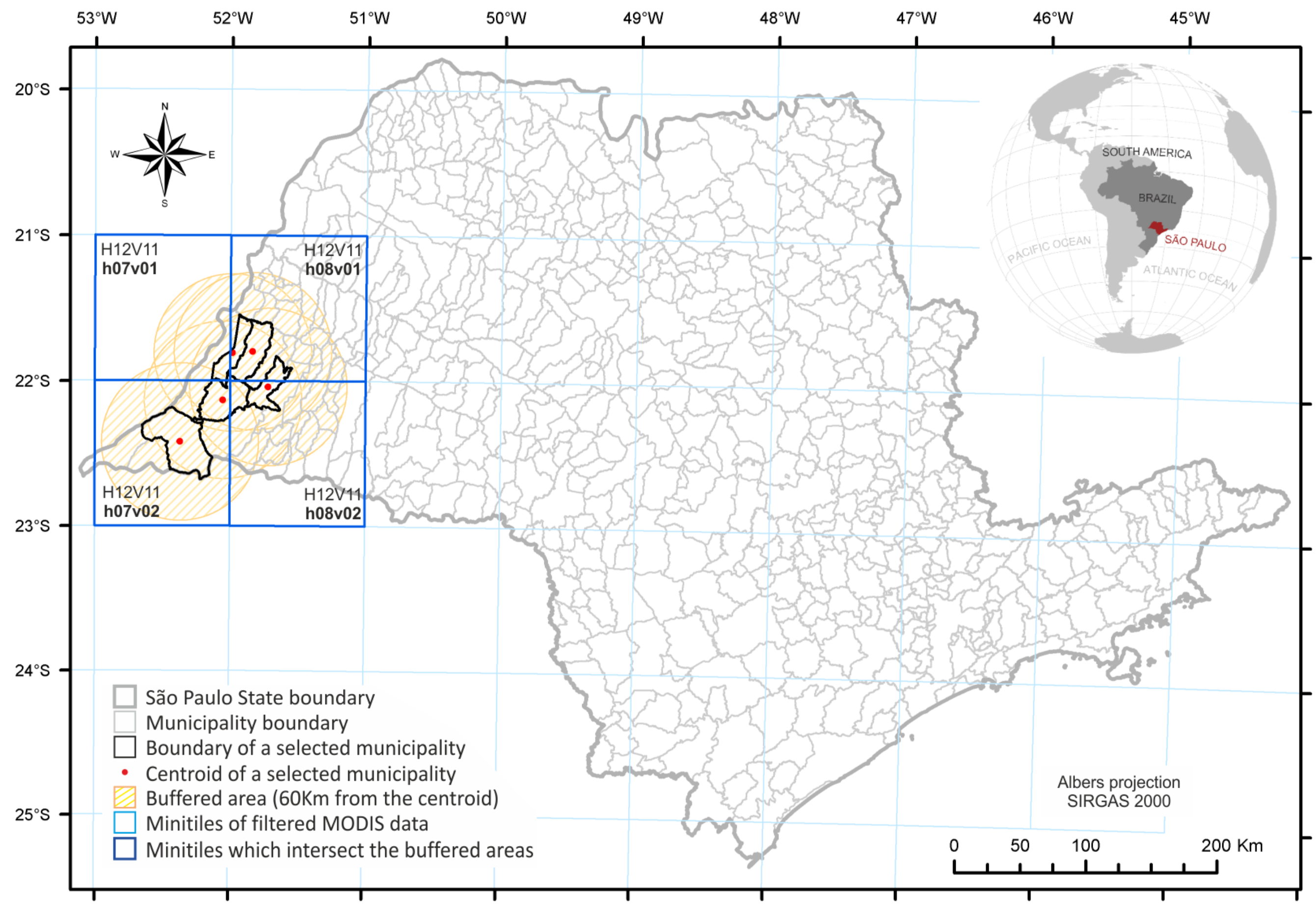

Fig. 1. Study area located in the west corner of the São Paulo State, Brazil, highlighting: a) (black borders) the five municipalities selected to run the analyses: *Caiuá*, *Marabá Paulista*, *Presidente Venceslau*, *Santo Anastácio*, and *Teodoro Sampaio*; b) (red dots) the centroid coordinate location of each selected municipality; c) (yellow circles) buffer of 60 Km based on the centroids; d) (blue squares) the four minitiles (1º x 1º) with MODIS filtered NDVI data that intersect all buffered area.

Remote sensing data used is based on 7-day smoothed real-time MODIS NDVI time series (MOD13) with regular weekly updates produced in gliding windows. The smoothed NDVI values refer to Mondays and are stored in grids of 1x1 degree called *minitiles*. The data is processed at the Institute of Surveying, Remote Sensing and Land Information (IVFL) of the University of Natural Resources and Life Sciences (BOKU), in Vienna [14]. We used data from the four *minitiles* that intersect the 60 km buffered area from the centroid of each of the five municipalities (see Fig. 1).

For each municipality, official yearly yield from 2003 to 2011, reported by IBGE [13], were used as reference to select, within the four *minitiles* (see Fig. 1), the 100 pixels from the NDVI time series that best match the official yield time series; these 100 pixels are called *proxies*. The quality of the match between official yield and remote sensing time series was based on the Root-Mean-Square Error (RMSE), after a Z-score normalization, as follows

$$Z_i = \frac{V_i - mean(\boldsymbol{V})}{sd(\boldsymbol{V})} \quad (1)$$

where: Z_i are the *i*-th *Z*-scored value; $\boldsymbol{V}$ is a vector with the observed values (either for official yield in a given municipality or for NDVI in a given pixel); *mean* is the average; and *sd* is the standard deviation values of $\boldsymbol{V}$. It is worth mentioning that proxies are not necessarily pixels in agricultural or sugarcane areas.

We ran weekly estimates of sugarcane yield using every Friday of 2012 as basis dates (i.e., 52 estimates), simulating a near-real-time processing. For each Friday, the proxies were selected; and the yield estimate was calculated using the median of Z-scored NDVI values for the proxies using an integration time of 10 weeks (starting on the last Monday before the Friday which was used as basis date). The official yield data for 2012 (not used to select proxies) was then used to assess the estimates regarding each Friday over the year 2012.

3. RESULTS AND DISCUSSION

Official sugarcane yield at municipality level in Brazil is usually released yearly by IBGE [13] with a delay of about one or two years after harvest. Regarding 2012, reported official yield for five municipalities were, in tons per hectare (ton.ha^{-1}): Caiuá = 60; Marabá Paulista = 66; Presidente Venceslau = 70; Santo Anastácio = 65; and Teodoro Sampaio = 60. The remote sensing approach proposed herein, however, allows weekly estimates of yield. Fig. 2 shows how sugarcane estimated yield throughout the year 2012 using the proposed approach differs (in %) from the

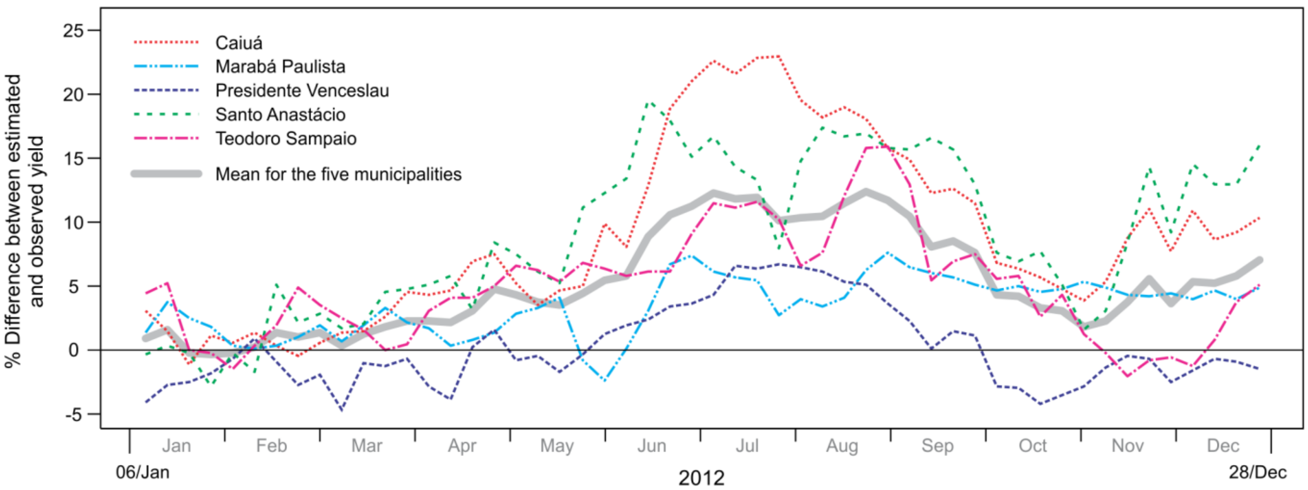

Fig. 2. Difference (given in %) between the sugarcane yield estimated based on the remote sensing approach proposed and the official yield published by IBGE [13].

official yield for the five municipalities evaluated; values above zero indicate overestimated yield. The thick gray line, representing the average of the difference for the five municipalities, revealed that the proposed remote sensing approach tended to overestimate the sugarcane yield along the evaluated year (2012).

Indeed, the mean difference for all 52 estimates along the year for each municipality was: Caiuá = 8.8%; Marabá Paulista = 3.4%; Presidente Venceslau = 0.2%; Santo Anastácio = 8.9%; and Teodoro Sampaio = 4.9%.

The period of full development for planted sugarcane goes until the end of March (which is also the end of the rainy season in São Paulo State), when sugar accumulation period usually ends. As sugarcane yield is strongly influenced by the sugar contents, which in its turn is strongly influenced by the water availability (rainy season), this period is critical to define yield [15]. In fact, Fig. 2 showed that the above-mentioned period presented the best estimates for remote sensing based yield (when the difference between remote sensing based and official yield were close to zero). From April to December, which defines the harvest season

Table. 1. Averaged results per municipality: **Obs.** is the observed yield yearly reported by IBGE [13], which is the official agency in Brazil; **Est.** is the yield estimated using the proposed new method; and **Diff.(%)** is the difference between **Est.** and **Obs.**, given in percentage. All values correspond to the average calculated considering two different periods: from January to March; and from April to December 2012.

Municipality	Obs.	Period averaged	Est.	Diff.(%)
Caiuá	60	January to March	60.8	1.3
		April to December	66.8	11.3
Marabá Paulista	66	January to March	67.1	1.6
		April to December	68.6	4.0
Presidente Venceslau	70	January to March	68.7	-1.9
		April to December	70.6	0.8
Santo Anastácio	65	January to March	65.9	1.4
		April to December	72.4	11.4
Teodoro Sampaio	60	January to March	61.1	1.8
		April to December	63.6	6.0

in São Paulo [3], remote sensing approach tended to overestimate sugarcane yield for all municipalities evaluated. Indeed, Table 1 summarizes results per municipality considering the two above-mentioned periods.

Apart from Presidente Venceslau, all municipalities presented better averaged-estimates for the period from January to March than for the period from April to December (see Table 1). Indeed, the fact that overestimated yield from June to September in that municipality compensated underestimated yield from October to December caused the Diff.(%) in Table 1 to be smaller than for the period from January to March (see Fig. 2). However Fig. 2 showed that, from January to March, the yield estimations, even for Presidente Venceslau, tended to stay close to the zero line, which means that this period may provide better estimations following the proposed method.

Although remote sensing based estimates in January seems to be sufficient to forecast sugarcane yield with good match with the official data, it is important to monitor yield throughout several months, especially until March, since heavy rains as well as frost in some growing areas may affect the sugar concentration and, consequently, yield [16]. It is worth mentioning that accurate and timely provided information to decision makers regarding energy crops (i.e., crop used as feedstock for biofuel production such as sugarcane in Brazil) possibly contributes to mitigate Green House Gas (GHG) emissions through the efficient use of sustainable bioenergy at a commercial-global scale, reducing fossil-based fuel demand and dependence.

4. CONCLUSION

The remote sensing approach proposed herein showed great potential to be used operationally for yield forecast. Results stated that sugarcane yield estimates from January to March using remote sensing data tends to match strongly with the official yield, with the benefit of being known before the harvest season. Although we only tested five municipalities, we do expect that the proposed method might be applicable to other regions and crops. The proposed method seems to represent a very accurate and robust method for yield forecast, providing decision makers relevant information to subsidize efficient energy crops management.

For future research we plan to test different normalization criteria, different integration periods for the time series, different number of *proxies* and also different crops and locations. In a leave-one-out approach, yield of all years will be estimated independently. It is also planned to run the method with yield data from different sources.

ACKNOWLEDGEMENTS

This research has been conducted under a cooperative project between Brazil's National Institute for Space Research (INPE) and the University of Natural Resources and Life Sciences (BOKU), in Vienna, Austria, in which the first author was a postdoctoral research fellow. Currently, the first author has joined Boeing Research & Technology – Brazil (BR&T-Brazil), and kindly acknowledges the opportunity to work on this subject within BR&T-Brazil.

5. REFERENCES

[1] P. Smith, D. Martino, Z. Cai, D. Gwary, H. Janzen, P. Kumar, B. McCarl, S. Ogle, F. O'Mara, C. Rice, B. Scholes, and O. Sirotenko, Agriculture. In: Metz, B.; Davidson, O.R.; Bosch P.R.; Dave R.; and Meyer L.A. (Eds.). *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press, 2007, pp. 497–540.

[2] J.R. Moreira and J. Goldemberg, "The alcohol program," *Energy Policy*, vol. 27, no. 4, pp. 229–245, 1999.

[3] B.F.T. Rudorff, D.A. Aguiar, W.F. Silva, L.M. Sugawara, M. Adami, and M.A. Moreira, "Studies on the rapid expansion of sugarcane for ethanol production in São Paulo State (Brazil) using Landsat data," *Remote Sensing*, vol. 2, no. 4, pp. 1057–1076, 2010.

[4] J. Goldemberg, "Ethanol for a sustainable energy future," *Science*, vol. 315, no. 5813, pp. 808–810, 2007.

[5] L.L. Nass, P.A.A. Pereira, and D. Ellis, "Biofuels in Brazil: An overview," *Crop Science*, vol. 47, no. 6, pp. 2228–2237, 2007.

[6] Canasat Project, *Sugarcane crop mapping in Brazil by Earth observing satellite images*. São José dos Campos, SP, Brazil. Available at: <www.dsr.inpe.br/laf/canasat/en/>. Accessed on: 13 Jan, 2014.

[7] M. Adami, M.P. Mello, D.A. Aguiar, B.F.T. Rudorff, and A.F. Souza, "A web platform development to perform thematic accuracy assessment of sugarcane mapping in South-central Brazil," *Remote Sensing*, vol. 4, no. 10, pp. 3201–3214, 2012.

[8] D.A. Aguiar, B.F.T. Rudorff, W.F. Silva, M. Adami, and M.P. Mello, "Remote sensing images in support of environmental protocol: Monitoring the sugarcane harvest in São Paulo State, Brazil," *Remote Sensing*, vol. 3, no. 12, pp. 2682–2703, 2011.

[9] M.P. Mello, C.A.O. Vieira, B.F.T. Rudorff, P. Aplin, R.D.C. Santos, D.A. Aguiar, "STARS: a new method for multitemporal remote sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 1897–1913, 2013.

[10] M. Adami, B.F.T. Rudorff, R.M. Freitas, D.A. Aguiar, L.M. Sugawara, and M.P. Mello, "Remote sensing time series to evaluate direct land use change of recent expanded sugarcane crop in Brazil," *Sustainability*, vol. 4, no. 4, pp. 574–585, 2012.

[11] J.H. Kastens, T.L. Kastens, D.L.A. Kastens, K.P. Price, E.A. Martinko, and R.Y. Lee, "Image masking for crop yield forecasting using AVHRR NDVI time series imagery," *Remote Sensing of Environment*, vol. 99, no. 3, pp. 341–356, 2005.

[12] C. Atzberger, "Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs," *Remote Sensing*, vol. 5, no. 2, pp. 949–981, 2013.

[13] IBGE, *Sistema IBGE de Recuperação Automática (SIDRA) - Produção Agrícola Municipal (PAM)*. Rio de Janeiro, RJ, Brazil. Available at: <http://www.sidra.ibge.gov.br/>. Accessed on: 07 Dec, 2013.

[14] F. Vuolo, M. Mattiuzzi, A. Klisch, and C. Atzberger, "Data service platform for MODIS Vegetation Indices time series processing at BOKU Vienna: current status and future perspectives," in *Proc. SPIE 8538, Earth Resources and Environmental Remote Sensing/GIS Applications III*, 2012, vol. 8538, pp. 85380A1–85380A10.

[15] B.F.T. Rudorff and G.T. Batista, "Yield estimation of sugarcane based on agrometeorological-spectral models," *Remote Sensing of Environment*, vol. 33, no. 3, pp. 183–192, 1990.

[16] L.A. Monteiro and P.C. Sentelhas, "Potential and Actual Sugarcane Yields in Southern Brazil as a Function of Climate Conditions and Crop Management," *Sugar Tech*, 2013 (in press).