SOJAMAPS: PROJECT OF MONITORING OF SOYBEAN AREAS IN BRAZIL USING BIG DATA IN THE CLOUD COMPUTING

Carlos Antonio da Silva Junior [1], Mendelson Lima [1], Jerry A. Johann [2], Fernando Saragosa Rossi [1], José Francisco de Oliveira-Júnior [3], Antonio Hérbete Sousa Leonel Junior [1] Alex Paludo [2], Willyan Ronaldo Becker [2], Kleberson Rodrigo do Nascimento [2], Rodrigo Bruno Zanin [1], Paulo Eduardo Teodoro [4], Luciano Shozo Shiratsuchi [5], Muhammad Shakir [6], Givanildo de Gois [7] and Marcos Rafael Nanni [8]

[1]State University of Mato Grosso, Cáceres-MT, carlosjr@unemat.br; [2]State University of Western Paraná, Cascavel-PR, Jerry.Johann@unioeste.br; [3]Federal University of Alagoas, Maceió-AL, jose.junior@icat.ufal.br; [4]Federal University of Mato Grosso do Sul, Chapadão do Sul-MS, paulo.teodoro@ufms.br; [5]Louisiana State University, Baton Rouge-LA, LShiratsuchi@agcenter.lsu.edu; [6]University of Peshawar-Institute of Space Technology, Islamabad-PAK, mshakirgeo@gmail.com; [7]Federal University of Fluminense, Volta Redonda-RJ, givanildogois@gmail.com; [8]State University of Maringá, Maringá-PR, mrnanni@uem.br

ABSTRACT

The main objective of this research was to use time-series images with medium spatial resolution and values with maximum and minimum vegetation index relating to the dynamic aspects of the soybean crop in Brazil via Earth Engine and making the data query available on-line high performance. Three sensors (MODIS, OLI and MSI) and cloud-based JavaScript processing were used for time-series composition. The largest soybean areas were detected in the Center-West and Southern regions of Brazil. Automated monitoring via Earth Engine was satisfactory, showing the ability to identify in near real-time what areas were grown with soybeans after harvesting.

Key words **—** Time-serie, agriculture, Earth Engine, automated mapping.

RESUMO

O objetivo principal dessa investigação foi utilizar série-temporal de imagens com média resolução espacial e valores com máximos e mínimos de índice de vegetação relacionando com os aspectos dinâmicos da cultura da soja no Brasil via EE e disponibilizando a consulta dos dados em plataforma on-line de alta performance. Foram utilizados três sensores (MODIS, OLI e MSI) e processamento via JavaScript em nuvem para composição da série-temporal. As maiores áreas de soja foram detectadas nas regiões Centro-Oeste e Sul do Brasil. O monitoramento automatizado via Earth Engine foi satisfatório, mostrando a capacidade de identificar em tempo quase real quais áreas foram cultivadas com soja após sua colheita.

Palavras-chave **—** *Série temporal, agricultura, Earth Engine, mapeamento automatizado.*

1. INTRODUCTION

Currently, one of the most important agricultural crops in Brazil is the soybean crop (*Glycine max* L. Merr.) - [1], since it is fundamental as a product of animal feed, vegetable oil and human consumption. Its economic importance moves the agroindustrial market of the country and is a product of export to international market, with this brings the movement of capital benefit to the producing regions [2].

The structure and solidity of the international market for soy products has led to the establishment of monoculture in Brazil, as well as the emergence of new technologies that have stimulated the cultivation of this oilseed and the creation of regions of agricultural consortia, i.e., Central Brazil and MATOPIBA (junction of the states of Maranhão, Tocantins, Piauí and Bahia), besides the association of the management and the effectiveness of the producers in the activity [3]. According to Food and Agriculture Organization (FAO), Brazil has prospects for increasing grain production, especially soybeans, due to the amount of arable land. Thus, the agricultural culture that has grown the most in Brazil in the last three decades, reaching 49% of the area planted with grain in the country [4].

In Brazil, the monitoring of agricultural areas by means of orbital images is very useful, given its territorial extension and enormous diversity of cultures. Studies on soybean cultivation can be confirmed around the world, addressing topics such as discrimination and quantification of areas, identification of phenological stage, estimation of productivity by chlorophyll content, among others [5].

Thus, the use of remote sensing data for the monitoring of agricultural dynamics requires fundamentally the continuous imaging of vegetation [6]. In addition, the soybean crop has some characteristics that favor its recognition by remote sensors [7]. For Gorelick et al. [8] supercomputers and high-performance computing systems are becoming abundant and large-scale cloud computing is universally available as a commodity. Google Earth Engine is a cloud-based platform that facilitates access to high-performance computing resources for the processing of very large geospatial datasets using remote sensing data without having to deal with the IT difficulties surrounding it currently facilitating the monitoring of agricultural areas in near real-time.

Therefore, the main objective of this research was to use time-series of images with medium spatial resolution and values with maximum and minimum vegetation index related to the dynamic aspects of the soybean crop in Brazil via Earth Engine and making the data query in an online platform high performance.

2. MATERIAL AND METHODS

By remote sensing data, mostly derived from orbital images, the monitoring of areas with soybean crop cultivation was sought. Initially, the project started with the state of Mato Grosso (crop yr^{-1} 2016/17), using data from a set of satellites and its time-series, allied with more than 800 randomly sampled field points. More tests of the algorithm were performed in the state of Paraná, through statistical validation [7]. For this, the model was applied using the spectral bands of Red (620-670 nm) and Near Infrared (840-880 nm) for the construction of the Perpendicular Vegetation Index (PVI). After that, the temporality that coincides with the beginning of the soil preparation until the harvest of the soybean crop was used, seeking to know the phenological phases by means of the spectrum-temporal behavior (Figure 1), then called PCEI - Perpendicular Crop Enhancement Index (Equation 1).

$$\text{PCEI} = \text{g}.\frac{\left(\text{Max}\,\dfrac{\rho_{\text{IVP}} - a\rho_{\text{V}} - b}{\sqrt{1+a^2}} + \text{S}\right) - \left(\text{Min}\,\dfrac{\rho_{\text{IVP}} - a\rho_{\text{V}} - b}{\sqrt{1+a^2}} + \text{S}\right)}{\left(\text{Max}\,\dfrac{\rho_{\text{IVP}} - a\rho_{\text{V}} - b}{\sqrt{1+a^2}} + \text{S}\right) + \left(\text{Min}\,\dfrac{\rho_{\text{IVP}} - a\rho_{\text{V}} - b}{\sqrt{1+a^2}} + \text{S}\right)} \quad (1)$$

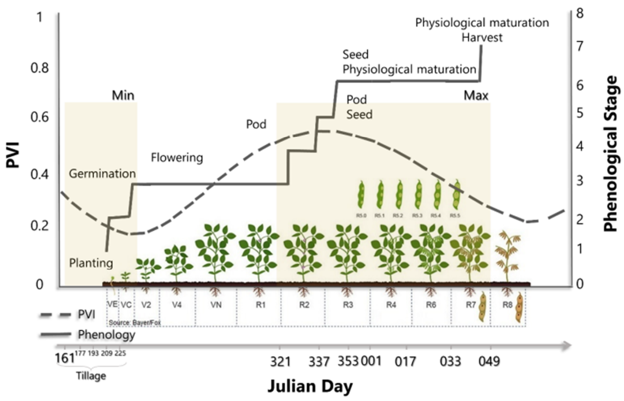

Figure 1. Spectrum-temporal characteristic of soybean as a function of vegetation index and phenological phases.

In order to identify the areas with soybean cultivation by means of a time-series allied to its phenological phase, it was sought to know the crop season of each Brazilian state through the sowing and harvest dates obtained at CONAB-Companhia National of Supply, thus constituting images of different periods for each state (Figure 2).

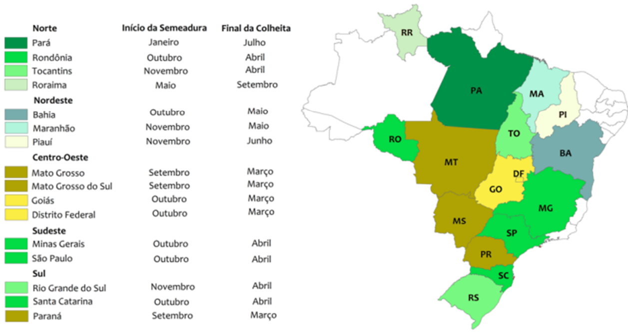

Figure 2. Crop seasons of soybean crop for the different Brazilian states separated by regions.

However, for the states of Pará and Roraima, the areas of soy that will always be monitored in relation to the previous year. For the construction of the mapping of the soybean areas of Brazil, data from the MODIS sensor (Products MOD13Q1 and MYD13Q1) were used by means of 13 tiles, with the intention of using as a basis of interpretation. Therefore, OLI/Landsat-8 data were used, using the median technique of composition of best scenes between the crops months of each state ("<=20%/cloud"). The spectral bands used were B4 (640-670 nm) and B5 (850-880 nm), through a total of 255 scenes for composition of the producing states. In addition, images of the MSI/Sentinel-2 were used in some cases where states with relatively small areas were identified and vectored when it was not possible via OLI in the clouds and *vice-versa*. At the end, the vectors obtained through the MODIS, OLI and MSI sensors were combined for the total monitoring composition and the polygons less than 6.25 ha.

In the last three crop years, several areas were traveled in the states of Mato Grosso, Mato Grosso do Sul and Paraná (Figure 3), in order to obtain control points for the monitoring of land use and occupation with soybean . Then, using field data, mapping validations were applied via Kappa parameter.

The results obtained through the vegetation index and its time-series were vectorized and transformed into the .kml format, where they were made available on API Google Maps platform for spatial consultation by municipality at http://pesquisa.unemat.br/gaaf/sojamaps, with direct link in the server of the State University of Mato Grosso.

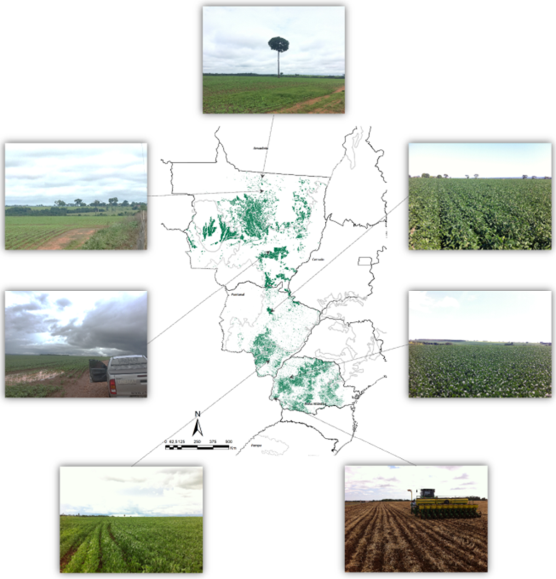

Figure 3. Detail of the areas covered by the states of Mato Grosso, Mato Grosso do Sul and Paraná.

3. RESULTS AND DISCUSSION

The monitoring of areas with soybean cultivation is being carried out via JavaScript [8] and the result of part of the Brazilian areas is presented in Figure 4. Similar mapping is carried out by the Global Food Security-Support Analysis Data project [9] through detection of croplands in general. However, the rough resolution coarse resolution nature of the map products with significant and few details are presented in areas, especially in Brazil.

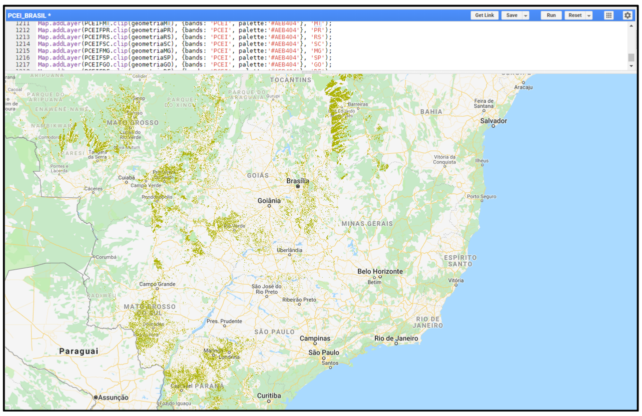

Figure 4. Detail of the monitoring of soybean areas in cloud computing (Earth Engine).

Figure 5 shows the municipalities of the state of Mato Grosso that present the largest extensions of areas cultivated with soy. A total of 120 municipalities were identified with soy in the state. The municipality of Sorriso, located in the northern mesoregion of the state, has the largest cultivated area, with a total of 650,995 ha, representing 6.22% of the total soybean area of Mato Grosso.

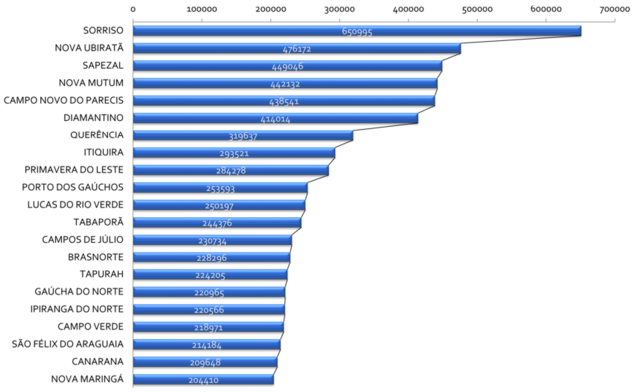

Figure 5. Relation of the 21 municipalities of Mato Grosso with the largest areas of soybean cultivation monitored in the 2017/18 crop year.

Figure 6 shows the municipalities of the state of Mato Grosso do Sul that present the largest extensions of soybean cultivated areas. The municipality of Maracaju, located in the southwest mesoregion of the state, has the largest cultivated area, with a total of 270,162 ha, representing 10.62% of the total soybean area of Mato Grosso do Sul.

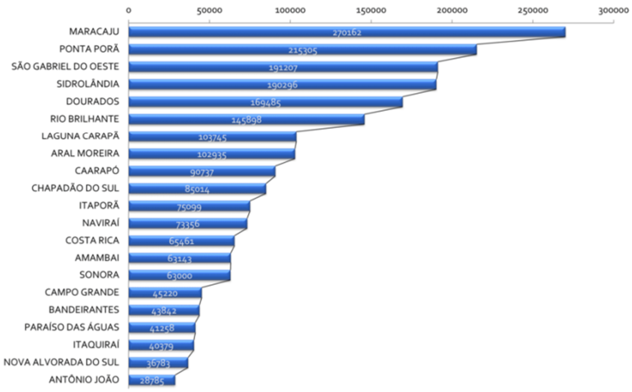

Figure 6. Relation of the 21 municipalities of Mato Grosso do Sul with the largest areas of soybean cultivation monitored in the 2017/18 crop year.

Figure 7 shows the municipalities of the state of Paraná that present the largest extensions of areas cultivated with soybean. The municipality of Cascavel, located in the western mesoregion of the state, presents the largest cultivated area, with a total of 95,309 ha, representing 1.97% of the total soybean area of Paraná. This smaller percentage is due to the enormous amount of municipalities in the state, adding 399, causing in smaller areas and still as a consequence the relief and small territories of areas available for cultivation when compared with the state of Mato Grosso.

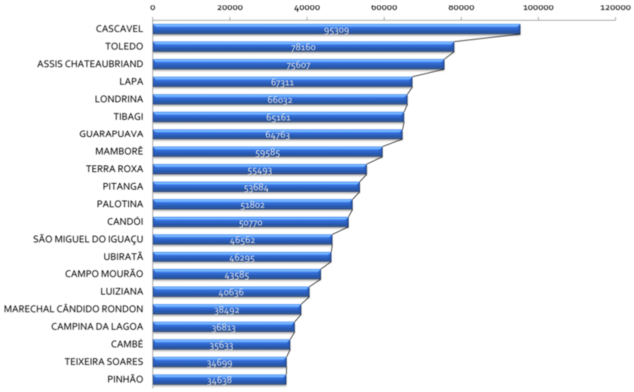

Figure 7. Relation of the 21 municipalities of Paraná with the largest areas of soybean cultivation monitored in the 2017/18 crop year.

Figure 8 shows the Brazilian states that grow soybean areas. A total of 16 states plus the Distrito Federal cultivate agricultural culture. The largest areas are concentrated in the state of Mato Grosso and a new region called MATOPIBA is promising in production.

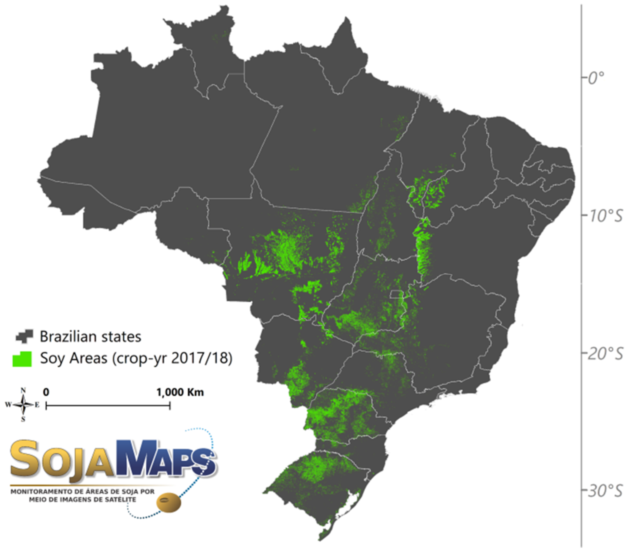

Figure 8. Areas of Brazilian soybean monitored by the SojaMAPS project.

It can be noticed that the Central-West and South regions of Brazil are the largest areas cultivated with soybean (Figure 9). The variations in the areas occur more frequently in the South of Brazil and with discrepancy in the state of Mato Grosso. The Northeast, North and Southeast regions present the smallest areas of soybeans, however, each crop year presents sowing in new areas. A total of 33.347.199 ha of soybean was mapped in the SojaMAPS project in the Brazilian territory, where 55.14% are in the Cerrado. The monitoring of the areas via remote sensing in the cloud was able to detect Kappa parameter accuracy higher than 0.98.

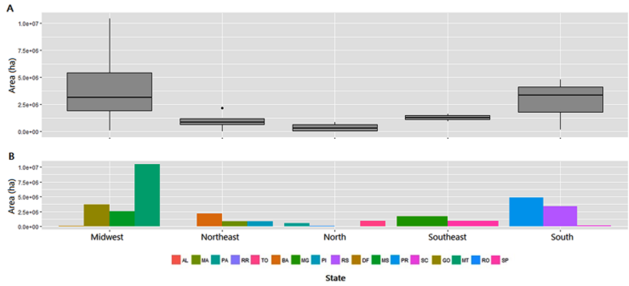

Figure 9. List of regions and their areas that grow soybeans.

4. CONCLUSIONS

Automated monitoring using time-series with Earth Engine in orbital images linking three sensors was satisfactory, showing the ability to identify in near real-time which areas were grown with soybean after harvesting. Using the Google Maps API tool is imperative in spreading spatial data with high-quality viewing of Big Data.

6. REFERENCES

[1] Silva Junior, C.A.; Nanni, M.R.; Oliveira-Junior, J.F.; Cezar, E.; Teodoro, P.E.; Delgado, R.C.; Shiratsuchi, L.S.; Shakir, M.; and Chicati, M.L., "Object-based image analysis supported by data mining to discriminate large areas of soybean", *International Journal of Digital Earth*, v.1, n.1, p.1-23, 2018.

[2] Pontes, H.L.J.; Carmo, B.B.T.; and Porto, A.J.V., "Problemas logísticos na exportação brasileira da soja em grão", *Revista Eletrônica Sistemas & Gestão*, v.4, n.2, p.155-181, 2009.

[3] Gazola, K.S.; Melz, L.J.; Torres, A.L.; and Anunciato, K.M., "Evolução do nível de concentração e produção da soja no Brasil e Mato Grosso do ano 2000 a 2010", *Revista UNEMAT de Contabilidade*, v.1, n.2, p.139-157, 2012.

[4] FAO: Food and Agriculture Organizations of the United Nations. Perspectivas agrícolas 2015-2024. Available in: http://www.agri-outlook.org/.

[5] Silva Junior, C.A.; Moreira, E.P.; Frank, T.; Moreira, M.A.; and Barcellos, D., "Comparação de áreas de soja (*Glycine max* (L) Merr.) obtidas por meio da interpretação de imagens TM/Landsat e MODIS/Terra no município de Maracaju (MS)", *Bioscience Journal*, v.30, n.5, p.707-716, 2014.

[6] Brown, L.A.; Dash, J.; Ogutu, B.O.; and Richardson, A.D, "On the relationship between continuous measures of canopy greenness derived using near-surface remote sensing and satellite-derived vegetation products", *Agricultural and Forest Meteorology*, v.247, n.1, p.280-292, 2017.

[7] Silva Junior, C.A.; Nanni, M.R.; Teodoro, P.E.; and Silva, G.F.C., "Vegetation Indices for Discrimination of Soybean Areas: A New Approach", *Agronomy Journal*, v.109, n.4, p.1331-1343, 2017.

[8] Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R., "Google Earth Engine: Planetary-scale geospatial analysis for everyone", *Remote Sensing of Environment*, v.202, n.1, p.18-27, 2017.

[9] Zhong, Y.; Giri, C.; Thenkabail, P.S.; Teluguntla, P.; Congalton, R. G.; Yadav, K.; Oliphant, A. J.; Xiong, J.; Poehnelt, J.; and Smith, C., "*NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security-support Analysis Data (GFSAD) @ 30-m for South America: Cropland Extent Product (GFSAD30SACE)*", NASA EOSDIS Land Processes DAAC, 2017.