CLASSIFICATION OF INTEGRATED CROP-LIVESTOCK SYSTEMS USING PLANETSCOPE TIME SERIES

João F. G. Antunes[1,2], Aliny A. Dos Reis[2,3], Henrique S. L. Almeida[2], João P. S. Werner[3], Gleyce K. D. A. Figueiredo[3], Júlio C. D. M. Esquerdo[1,3], Inacio T. Bueno[2,3], Ana P. S. G. D. D. Toro[3], Rubens A. C. Lamparelli[2], Alexandre C. Coutinho[1], Paulo S. G. Magalhães[2]

[1]Embrapa Digital Agriculture, Brazilian Agricultural Research Corporation – Embrapa, Campinas, SP, Brazil - joao.antunes@embrapa.br, julio.esquerdo@embrapa.br, alex.coutinho@embrapa.br
[2]Interdisciplinary Center of Energy Planning, University of Campinas – UNICAMP, Campinas, SP, Brazil - aareis@unicamp.br, h198850@dac.unicamp.br, ibueno@unicamp.br, lamparel@unicamp.br, graziano@unicamp.br
[3]School of Agricultural Engineering, University of Campinas – UNICAMP, Campinas, SP, Brazil - j164880@dac.unicamp.br, gleyce@unicamp.br, a265622@dac.unicamp.br

ABSTRACT

The new generation of orbital platforms has increased the opportunities for land cover classification using time series of satellite images in the last few years. In this study, we assessed the performance of high spatial and temporal resolution PlanetScope (PS) time series to map integrated crop-livestock systems (ICLS) and different land covers in the western region of São Paulo State, Brazil. To achieve this goal, 10-day and 15-day composite time series of the vegetation indices on both pixel and object-level were extracted from the PS images. The land cover classifications were performed using the Multi-Layer Perceptron (MLP) classifier, which achieved overall accuracies greater than 98.0%. The 10-day composite PS time series slightly outperformed the 15-day composite, returning overall accuracies of 99.1% and 98.6%, respectively. Although our method improved the discrimination of land parcels with ICLS, prediction maps returned misclassifications due to the hybrid unit of analysis, which will be improved in future works with the use of new deep learning algorithms that fully explore the temporal domain of the time series.

Key words — nano-satellites, NDVI, EVI, image composites, Multi-Layer Perceptron.

1. INTRODUCTION

Stimulated by the growing demand for more sustainable agricultural production, the integration of crops and livestock activities (ICLS) has been adopted in Brazil as a practice of intensification to achieve more sustainable agricultural systems [1]. ICLS are highly dynamic, as the land use changes over a short period, hampering the establishment of patterns in remotely-sensed data that can be standardized to represent and map such systems [2].

Using a time series of satellite images is one of the most efficient ways to map and monitor ICLS [2]. The new generation of orbital platforms, such as Planet CubeSat satellites, offers an unprecedented opportunity to monitor land cover dynamics with enhanced spatial detail and high temporal resolution [3], [4]. In this context, using a deep neural network classifier, we assessed the performance of high spatial and temporal resolution PlanetScope (PS) time series to map ICLS and different land covers in the western region of São Paulo State, Brazil. Specifically, we investigated the performance of 10 and 15-day composites of Normalized Difference Vegetation Index (NDVI) [5] and Enhanced Vegetation Index (EVI) [6] on model accuracies.

2. MATERIAL AND METHODS

The study area is located in the western region of the São Paulo State, Brazil (Figure 1), and covers approximately 7,300 hectares. The landscape comprises land parcels with remnants of native vegetation, crops, pasture, ICLS, and rural settlements of smallholder farmers.

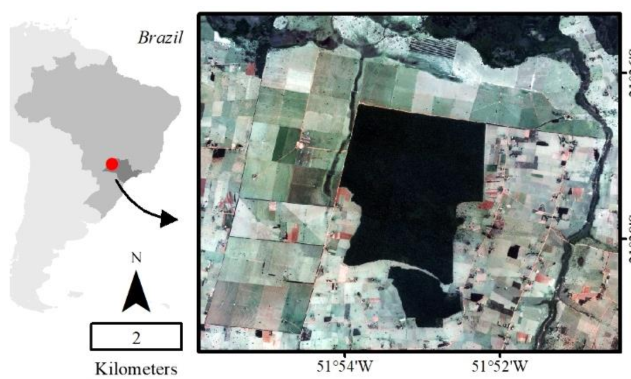

Figure 1. Location of the study area and RGB mosaic of PS images acquired in March 2019.

We considered the period from September 2018 to August 2019 (one agricultural year) to perform the land cover classification in the study area. During this period, 334 cloud-free PS multispectral images (surface reflectance product) covered the study area and were acquired for this study. PS is a constellation of nano-satellites with approximately 130 CubeSats 3U form factor (0.1 m by 0.1 m by 0.3 m), which have high spatial resolution of 3 m (nadir ground sampling distance), daily temporal resolution, and global coverage of the land surface. The PS sensor has four spectral bands: blue (455–515 nm), green (500–590 nm), red (590–670 nm), and near-infrared (780–860 nm).

For each PS image, we calculated the NDVI and the EVI to obtain the satellite image time series. It was considered that the time series of both indices would usefully increase the number of variables in the deep neural network classifier, improving the classification model performance. To generate time series with equal intervals and to overcome gaps due to frequent cloud cover in the study region, we generated the 10-day and 15-day composites by selecting the maximum value of the vegetation indices (NDVI and EVI) for each pixel in each interval of 10 and 15 days, respectively. In addition, the remaining gaps in the time series were linearly interpolated from the previous valid value to the next valid value.

Ground reference data from the current land cover and the historical land cover for this study was collected during two field in May 2019 and February 2020 by interviewing the local farmers. Eleven land cover classes were identified: ICLS, cultivated pasture, shrub-pasture, perennial crops, semi-perennial crops or annual crops, eucalyptus, primary forest, secondary forest, natural vegetation and wet areas, water, and others (buildings, roads, etc.).

We used a hybrid approach to perform the land cover classification in our study, as represented in Figure 2.

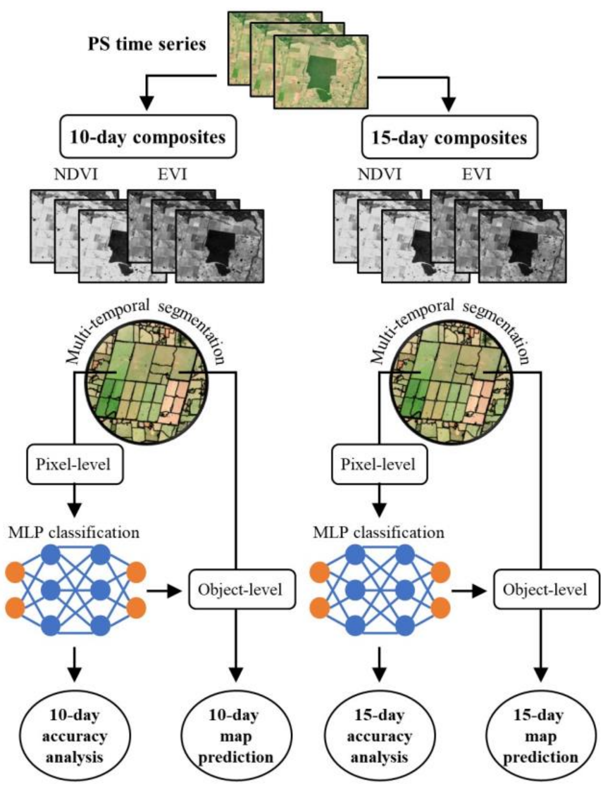

Figure 2. Schematic illustrating the classification method.

First, the deep neural network classifier was trained and tested using time series from the pixel level (pixel-based classification approach). Next, the land cover map was generated using average times series for each land parcel in an object-based classification approach. To delimit the land parcels, we employed a multi-temporal segmentation based on eight PS EVI mosaics in selected dates of the 2018-2019 period, and the region growing algorithm proposed by Baatz & Schäpe [7]. The resulting segments were manually edited, and the segments matching the ground reference data were selected as the reference parcels for this study. The pixels inside the reference parcels were used for training, validation, and testing the classification model. For each pixel, we identified its land cover class, and extracted the 10-day and 15-day composite EVI and NDVI time series.

The total number of samples (i.e., time series) in each land cover class (Table 1) was divided into 60% for training, 20% for validation, and 20% for testing.

Land cover class	Time series
Cultivated pasture	474,300
ICLS	396,717
Shrub-pasture	129,824
Eucalyptus	52,210
Natural vegetation and wet areas	45,829
Perennial crops	40,209
Primary forest	34,233
Secondary forest	33,782
Semi-perennial crops or annual crops	18,120
Others	7,928
Water	876
Total	**1,234,028**

Table 1. Number of time series extracted in each land cover class for training, validation, and testing the classification models.

To generate the land cover maps, we created prediction datasets by averaging the maximum values of NDVI and EVI in each image-composite (10 or 15 days) of the time series for each land parcel in our study area. As a result, each prediction dataset was composed of 4,983 time series.

Our study used the Multi-Layer Perceptron (MLP) classifier deep neural network as the classification algorithm. This model is a feedforward artificial neural network with more than one hidden layer, commonly referred to as the most basic deep neural network [8]. The neurons in a layer of a MLP are fully connected to all neurons in the other layers. The neural network architecture used in our study has two hidden layers and three dropout layers (10%, 20%, and 20%, respectively) between the layers mentioned above and the input layer to avoid overfitting, and all layers have 512 neurons. The ReLU activation function was used in the final fully-connected layer, followed by a Softmax layer with 11 neurons, one for each land cover class. The MLP models were trained using Adam optimizer, and the number of epochs and the batch size were set to 100 and 256, respectively. The parameters of Adam were fixed as: $\beta_1 = 0.9$, $\beta_2 = 0.999$, learning rate = 0.001, and $\varepsilon = 1\text{x}10^{-7}$.

The performance of the MLP models was evaluated both quantitatively and qualitatively. We used six metrics based on the testing dataset (pixel-based time series) to quantitatively assess the classification accuracy: overall accuracy (OA), macro-average (MA), weighted average (WA), precision, recall, and F1-score. Qualitative evaluation of the land cover maps was performed by visual inspection.

3. RESULTS

The 10-day composite PS time series resulted in higher accuracy of the MLP model (OA = 99.1%, MA = 98.6%, and WA = 99.9%) compared to the 15-day composite PS time series (OA = 98.6%, MA = 97.9%, and WA = 98.6%). The best model also resulted in higher accuracy metrics per land cover class (Table 2), with special attention on the ICLS class that reached 100.0% of accuracy in the precision, recall, and F1-score metrics based on the testing dataset.

The land cover maps generated by the MLP classifier using 10-day and 15-day compositions are illustrated in Figure 3a and Figure 3b, respectively. Despite the high OA, MA, and WA accuracy metrics for the two MLP models (greater than 98.0%), both models resulted in misclassifications between the land cover classes compared with the ground reference data in the qualitative evaluation of the land cover maps. Nevertheless, the MLP model using data from the 10-day composite EVI and NDVI time series resulted in fewer misclassifications than the 15-day composite MLP model.

4. DISCUSSION

This study explored the potential of using PS time series and a deep neural network classifier to map ICLS and different land covers in Brazil. The high accuracy of the MLP models emphasizes the importance of using high temporal resolution time series for the classification of dynamic agricultural systems such as ICLS. Manabe et al. [2] also highlighted the importance of suitable temporal resolution for seasonal agricultural mapping using satellite image time series.

Land cover class	Precision	Recall	F1-Score
Cultivated pasture	99.1%	99.1%	99.1%
Eucalyptus	99.0%	98.4%	98.7%
ICLS	100.0%	100.0%	100.0%
Natural vegetation and wet areas	100.0%	100.0%	100.0%
Others	99.2%	97.9%	98.5%
Perennial crops	99.0%	99.4%	99.2%
Primary forest	99.8%	97.4%	98.6%
Secondary forest	97.4%	99.9%	98.6%
Semi-perennial crops or annual crops	97.4%	92.2%	94.8%
Shrub-pasture	96.7%	97.7%	97.2%
Water	100.0%	100.0%	100.0%

Table 2. Class-based accuracy metrics of the best MLP model.

The high accuracy of the MLP models emphasizes the importance of using high temporal resolution time series for the classification of dynamic agricultural systems such as ICLS. Manabe et al. [2] also highlighted the importance of suitable temporal resolution for seasonal agricultural mapping using satellite image time series.

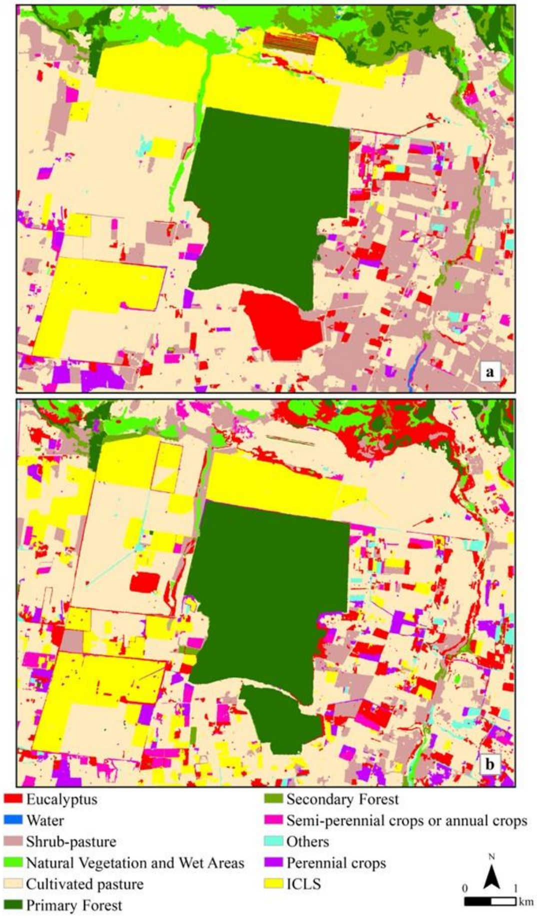

Figure 3. Land cover maps obtained by the MLP models using data from (a) 10-day composite EVI and NDVI time series, and (b) 15-day composite EVI and NDVI time series.

Although the class-based accuracy metrics show 100.0% accuracy for precision, recall, and F1-score for the ICLS class in the testing dataset (Table 2), some land parcels were misclassified as ICLS by the best MLP model (Figure 3a). These misclassifications may be related to the hybrid approach adopted in our study to perform the land cover classification, where the MLP model was trained using time series from pixels and the land cover map was generated from segments (spatial and spectral independent sets of data). Training the classification models using time series generated at the pixel level ensures a greater number of observations, which is an important factor to a more efficient learning of the MLP model. However, pixel-based classification approaches of high spatial resolution imagery usually result in mixed-class classifications of land parcels, so-called "salt-and-pepper" speckle [9].

The parcel-level classification approach improved some of the common pixel-based classification limitations, but it also included unwanted consequences of the segmentation process on the classification results since the size of image segments critically impacts the land cover mapping [10]. For example, the EVI and NDVI time series averages of large segments may present high internal (spectral) heterogeneity due to land cover mixture, which may have mischaracterized the land cover multi-temporal pattern and resulted in misclassifications of land parcels.

In addition, the MLP architecture may have not fully captured the temporal domain of the EVI and NDVI time series [11], [12]. Zhong et al. [11] reported that other deep neural network architectures such as the Convolutional Neural Networks (CNN) in one dimension and the Long Short-Term Memory (LSTM) can automatically learn the time characteristic and the multi-temporal patterns of the land covers, resulting in better classification results.

5. CONCLUSIONS

The performance of high spatial and temporal PS time series for mapping ICLS and different land covers in the western region of São Paulo State, Brazil, was assessed in this study. The 10-day composite EVI and NDVI time series resulted in enhanced land cover mapping in terms of accuracy metrics and the qualitative evaluation of classification outputs compared with the 15-day composite EVI and NDVI time series.

Although the MLP models have achieved high classification accuracy, some misclassifications between the land cover classes were observed due to the hybrid unit of analysis. We reinforce the use of new deep learning algorithms that fully explore the temporal domain of PS time series to improve land cover map predictions.

6. ACKNOWLEDGEMENTS

The authors are thankful for the financial support from FAPESP (Grants: 2021/15001-9, 2019/16870-0, 2018/24985-0, and 2017/50205-9) and would like to thank Jansle V. Rocha, Ricardo Torres, and Allan Pinto for their contribution to this study.

7. REFERENCES

[1] R. D. Garrett, M. T. Niles, J. D. B. Gil, A. Gaudin, R. Chaplin-Kramer, A. Assmann, T. S. Assmann, K. Brewer, P. C. de Faccio Carvalho, O. Cortner, R. Dynes, K. Garbach, E. Kebreab, N. Mueller, C. Peterson, J. C. Reis, and V. Snow, J. Valentim, Social and ecological analysis of commercial integrated crop livestock systems: Current knowledge and remaining uncertainty, *Agricultural Systems*, vol. 155, no. July, pp. 136–146, 2017, doi: 10.1016/j.agsy.2017.05.003.

[2] V. D. Manabe, M. R. S. Melo, and J. V. Rocha, Framework for Mapping Integrated Crop-Livestock Systems in Mato Grosso, Brazil, *Remote Sensing*, vol. 10, no. 9, pp. 1322, 2018, doi: 10.3390/rs10091322.

[3] Planet Labs PBC, *Planet Imagery Product Specifications*. https://assets.planet.com/docs/Planet_Combined_Imagery_Product_Specs_letter_screen.pdf. Accessed on: 10 May 2020.

[4] A. Sakuma and H. Yamano, Satellite constellation reveals crop growth patterns and improves mapping accuracy of cropping practices for subtropical small-scale fields in Japan, *Remote Sensing*, vol. 12, no. 15, pp. 2419, 2020, doi: 10.3390/RS12152419.

[5] J. W. Rouse Jr., R. H. Haas, J. A., Schell, and D. W. Deering, Monitoring vegetation systems in the Great Plains with ERTS. In: Fraden, S. C., Marcanti, E. P., Becker, M. A. (Eds.), *Third Earth Resources Technology Satellite-1 Symposium*, pp. 309–317, 1974.

[6] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, Overview of the radiometric and biophysical performance of the MODIS vegetation indices, *Remote Sensing of Environment*, vol. 83, no. 1–2, pp. 195–213, 2002, doi: 10.1016/S0034-4257(02)00096-2.

[7] M. Baatz and A. Schape, Multiresolution Segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation. In: Strobl, J., Blaschke, T. and Griesbner, G. (Eds.), *Proceedings Angewandte Geographische Informations-Verarbeitung*, XII, Wichmann Verlag, Karlsruhe, Germany, pp. 12-23, 2000.

[8] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT Press, 2016.

[9] M. Belgiu and O. Csillik, Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis, *Remote Sensing of Environment*, vol. 204, pp. 509–523, 2018, doi: 10.1016/J.RSE.2017.10.005.

[10] Y. Gao, J. F. Mas, N. Kerle, and J. A. Navarrete Pacheco, Optimal region growing segmentation and its effect on classification accuracy, *International Journal of Remote Sensing*, vol. 32, no. 13, pp. 3747–3763, doi: 10.1080/01431161003777189.

[11] L. Zhong, L. Hu, and H. Zhou, Deep learning based multi-temporal crop classification, *Remote Sensing of Environment*, vol. 221, pp. 430–443, 2019, doi: 10.1016/J.RSE.2018.11.032.

[12] C. Pelletier, G. Webb, and F. Petitjean, Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series, *Remote Sensing*, vol. 11, no. 5, pp. 523, 2019, doi: 10.3390/rs11050523.