LAND USE/LAND COVER CLASSIFICATION IN A HETEROGENEOUS AGRICULTURAL LANDSCAPE USING PLANETSCOPE DATA

I. T. Bueno [1,3 *], J. F. G. Antunes [1,2], A. P. S. G. D. D. Toro [3], J. P. S. Werner [3], A. C. Coutinho [2], G. K. D. A. Figueiredo [3], R. A. C. Lamparelli [1,3], J. C. D. M. Esquerdo [2,3], P. S. G. Magalhães [1]

[1] Interdisciplinary Center of Energy Planning, University of Campinas, Campinas 13083-896, SP, Brazil – (ibueno, lamparel, graziano)@unicamp.br

[2] Embrapa Digital Agriculture, Brazilian Agricultural Research Corporation, Campinas 13083-886, SP, Brazil – (joao.antunes, alex.coutinho, julio.esquerdo)@embrapa.br

[3] School of Agricultural Engineering, University of Campinas, Campinas 13083-875, SP, Brazil – (a265622, j164880)@dac.unicamp.br, gleyce@unicamp.br.

KEY WORDS: Clustering, Agricultural crops, OBIA, Random Forest, Spectro-temporal signature, Intra-class variability.

ABSTRACT:

Land use and land cover (LULC) classification has long been an essential topic in Earth Observation research and plays a key role in the sustainable development of agriculture. This study evaluated the accuracy of LULC classification based on an initial clustering step in a heterogeneous agricultural landscape using PlanetScope imagery while checking for variability among their Normalized Difference Vegetation Index (NDVI) temporal signatures. We adopt an object-based image analysis to generate image-objects and then extract statistical information of PlanetScope spectral bands and vegetation indices as input information for classification. The exploratory analysis focused on the double crop class and calculated the distance between NDVI temporal signatures of paired land parcels. We applied an unsupervised clustering technique along with Random Forest algorithm based on multiple tests to classify and analyse gains and losses in accuracies produced by these approaches. Our results showed that the initial clustering method outperformed the non-clustered classification of LULC in overall accuracy measures. The exploratory analysis demonstrated that double crops might present high intra-class variability and diverse crop calendars for neighbour land parcels. The accuracies achieved represent promising opportunities for the sufficiently accurate classification of such areas, and the knowledge of the intra-class variability allows the analyst to infer the temporal dynamics of crop fields. We reinforce that further work could assess other types of classifiers, especially in areas with a large number of crop types and distinct management practices.

1. INTRODUCTION

Land use and land cover (LULC) classification is an essential planning tool once it illustrates the spatial distribution of the Earth's surface attributes and plays a pivotal role in the sustainable development of agronomics, environment, and economics (Vali et al., 2020). Over the past few decades, accurate information on crop types and their spatial distribution have become a central issue for food security, policy making, and water and soil resource management (FAO, 2021).

Besides the accurate information, LULC maps for agricultural purposes must present a suitable spatial resolution and be updated frequently to deliver practical benefits. From this perspective, remote sensing data and classification techniques have the unique capability to automate the generation of such maps in a large-scale observation (Bellón et al., 2018) and the potential to discriminate a wide range of crop types (Brinkhoff et al., 2019).

A considerable amount of literature has been published on LULC classification of agricultural areas using remote sensing data (Orynbaikyzy et al., 2019). The increasing availability of public Earth Observation data and the advanced methods in digital image processing leveraged these studies from single imagery (Mathur and Foody, 2008) to time-series analysis (Cai et al., 2018), allowing it to capture the temporal dynamics of crops. In addition, the availability of very high-resolution images has provided fine-grained classifications and more effective target recognition under complex landscapes (Qin and Liu, 2022). For instance, the commercial constellation of Planet CubeSats provides a powerful combination of high spatial (3m pixel size) and temporal (daily) resolution imagery for fine-scale LULC classification and monitoring (Planet Team, 2022).

However, LULC classification of agricultural areas is not a trivial task since it depends on the size of the area and its complexity, for instance, the number of crop classes, parcel sizes, and environmental conditions. In addition, a single crop class can exhibit considerable variability in its within-parcel spatial distribution and spectro-temporal signatures due to different management practices and cropping intensity or even the presence of trees within the parcels. For instance, the growing demand for agricultural commodities has increased the incentive for intensification practices, and multiple-cropping has drawn interest in boosting production without expanding the planting area (Garrett et al., 2018). This system can demonstrate diverse crop calendars for neighbour land parcels, which may lead to similar spectro-temporal signatures to other LULC classes (Azar et al., 2016; Villa et al., 2015).

In order to deal with the high intra-class variability, crop mapping methods would largely benefit from spectral-temporal variability reduction steps prior to classification. Object-based approaches (so-called Object-Based Image Analysis – OBIA) are widely used to reduce spatial variability by dividing an image into groups of pixels that are spatially continuous and spectrally homogeneous (Blaschke, 2010). Moreover, unsupervised classification methods can reduce the spectral variability of crops. In this case, they rely on algorithms for grouping spectro-temporal signatures according to their intrinsic characteristics or similarity and not using a priori labelling. These methods appear as good alternatives for mapping agricultural patterns (Pascucci et al., 2018).

In this context, the objective of our study was to evaluate the accuracy of LULC classification based on an initial clustering step in a heterogeneous agricultural landscape while checking for variability among their Normalized Difference Vegetation Index (NDVI) temporal signatures. This paper first examines the temporal signatures of the double crop class and the intra-class variability, investigating the potential patterns of plant structure, growing behaviour, and management practices. Second, we evaluated the performance of an initial clustering-based classification, investigating the benefit of partitioning the spectro-temporal signatures dataset into initial sub-clusters. We hypothesized that grouping observations with similar spectro-temporal signatures will improve the performance of LULC classification over a non-clustered method because these clusters will reduce the intra-class variability, returning gains in accuracy.

This study addresses the following research questions: (1) How similar are NDVI temporal signatures derived from a dynamic agricultural land use class? (2) How does a LULC classification method based on an initial clustering scheme affect map accuracies in a heterogeneous agricultural landscape?

To answer these research questions, we extracted and classified PS spectro-temporal signatures in the southern region of Mato Grosso (MT) State, Brazil. We applied an unsupervised clustering technique along with RF algorithm based on multiple tests to classify and analyse gains and losses in accuracies produced by these approaches.

2. MATERIAL AND METHODS

2.1 Study area

The study area is the Gravataí farm located in the southern region of MT State, Brazil (Figure 1). Totalling 10,352 hectares, the area is occupied by LULC classes comprising agriculture parcels, pasture, native forest, roads, buildings, and bare lands.

The area presents sixteen land parcels that have been intensively managed as a double crop system based on the rotation of soybean in the summer season and corn during the winter season. Usually, the soybean-sowing period lasts from late October to mid-November according to the onset of the rainy season, which lasts from October to March; and, immediately after the soybean harvest, there is the corn-sowing period for harvest in May and June (called *safrinha* or second season in Brazil).

The crop expansion from the last five decades led MT to be Brazil's major source of agricultural production. Last year, MT produced 26.2% and 36.2% of the national production of soybeans and corn, respectively, making it the most productive state for each of these crops (IBGE, 2021).

The use of intensive practices, such as double cropping, has been widely adopted in MT, which is an additional challenge for accurate crop area mapping. Thus, the need for accurate crop maps and the agricultural attention upon MT reinforce the choice of our study area.

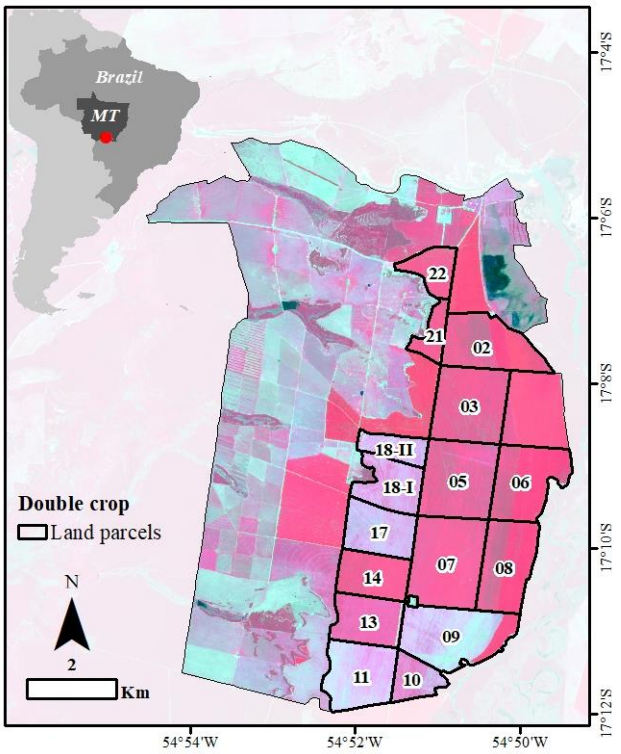

Figure 1. The Gravataí farm and its representation in PS false colour composition (Red = near-infrared, Green = red, Blue = green) from 16/01/2019.

2.2 PlanetScope data processing

We acquired PS images from 1st September 2018, to 31st August 2020, covering the study region. This time interval corresponds to the period of two agricultural years in the region: 2018-2019, and 2019-2020. We selected all surface reflectance images with less than 30% of cloud cover and masked them with the Unusable Data Mask product to generate cloud-free time series for each agricultural year. A total of 717 PS image files were downloaded.

A set of five PS-based spectral indices and four spectral bands was generated to classify the LULC in the study area (Table 1). These spectral variables can indicate vegetation condition, phenological dynamics, plant water, and chlorophyll content by providing information about crops' reflective characteristics at different wavelengths. In addition, these indices have been used in agricultural studies using PS data and have shown the potential to detect important vegetation features for classification and estimation purposes (Reis et al., 2020).

To overcome long image gaps in the time series due to cloud cover, we calculated a 15-day image composition by selecting the median value of each band/index in a 15-day interval. These image composites also provided a consistent time series with equal time intervals. In addition, we linearly interpolated potential remaining gaps in the time series using the nearest valid value before and after the time step.

Band/ Index	Formula	Ref
B	NA	(Planet Team, 2022)
G	NA	
R	NA	
NIR	NA	
EVI	$2.5((NIR - R)/(NIR + 6R - 7.5B + 1))$	(A. Huete et al., 2002)
GNDVI	$(NIR - G)/(NIR + G)$	(A. Huete et al., 2002)
MSAVI	$\frac{(2NIR + 1) - \sqrt{(2NIR + 1)^2 - 8(NIR - R)}}{2}$	(Qi et al., 1994)
NDVI	$(NIR - R)/(NIR + R)$	(Rouse et al., 1974)
SAVI	$1.5((NIR - R)/(NIR + R + 0.5))$	(A. R. Huete, 1988)

Table 1. Summary of the spectral indices and bands used in this study for LULC classification.

2.3 Image segmentation

Image segmentation is a crucial step in OBIA and divides an image into groups of pixels that are spatially continuous and spectrally homogeneous, also known as image-objects. The segmentation of a remote sensing image minimises the within-object variability compared to the between-object variability (Desclée et al., 2006).

We implemented image segmentation within the Google Earth Engine environment based on the Simple Non-Iterative Clustering (SNIC) algorithm (Achanta and Süsstrunk, 2017) performed on the NDVI images of each agricultural year, so-called multitemporal segmentation. This method created multitemporal image objects by segmenting multiple NDVI images of sequential periods. It incorporated spectral, spatial, and temporal information from NDVI images, which created objects based on the LULC dynamics in time (Bueno et al., 2019).

The multitemporal segmentation was based on setting a grid of seeds to establish a superpixel seed location spacing, influencing the image-objects size. Then, SNIC required setting some main parameters: the "compactness factor" was set to 0.5 and affects the object shape; the "connectivity" was set to 4 and defined the type of contiguity to merge adjacent objects, and a "neighborhoodSize" was set to 256 to avoid tile boundary artifacts. We set these parameters using an essay on trial-and-error to find the appropriate segmentation for the study area, then select the most suitable output based on a visual assessment (Duro et al., 2012).

We generated two sets of multitemporal segmentation, one for each agricultural year.

2.4 Ground Reference data

Field campaigns were conducted in June 2021 to collect terrestrial reference data points. The LULC information was collected based on the current land cover during the field campaigns and by interviewing local farmers to obtain the area's historical LULC. From the reference data, we set six LULC classes: cultivated pasture, double crop (soybean and corn), forest, integrated crop-livestock system (soybean and pasture), and others (roads, buildings, and bare lands).

2.5 Exploratory analysis

We conducted an exploratory analysis of the NDVI temporal signatures of the double crop class to obtain further in-depth information on their time dynamics and intra-class variability. From image-objects of this class, we extracted the mean NDVI and then generated their temporal signatures.

We calculated the mean NDVI temporal signature for each double crop land parcel by filtering their corresponding image-objects. In this case, we calculated the Euclidean distance of NDVI values (q_1 and p_1) for each date (q_2 and p_2), then obtained the Mean Euclidean Distance (MED) between NDVI temporal signatures of land parcels (Equation 1). Euclidean distance was chosen primarily due to its straightforward interpretation.

$$MED = \frac{\sum_{i=1}^{n} \sqrt{\left(p_{1_i} - q_{1_i}\right)^2 + \left(p_{2_i} - q_{2_i}\right)^2}}{n} \quad (1)$$

MED values between all pairs of land parcels generated a matrix of distances, which allowed us to infer their intra-class variability graphically. MED values were also scaled from 0 to 1 based on the absolute minimum and maximum values to relatively monitor their variability.

2.6 LULC classification

From image-objects, we extracted four descriptive attributes: mean, 95th and 5th percentile, and standard deviation based on each band/index value inside the objects, then generated their respective spectro-temporal signatures. The initial pool of variables was screened to limit the potential effects of multicollinearity by calculating correlations between pairs of variables using Pearson's R correlation coefficient. We removed those with R values greater than 0.90.

A clustering step partitioned the double crop image-objects into initial clusters based on their spectro-temporal signatures. We evaluated nine initial clusterings according to their number of clusters (NC), NC = {2, 3, 4, 5, 6, 7, 8, 9, 10}, and a non-clustered test (NC = 0). They were based on the k-means algorithm, a widely used clustering method that partitions the data into k groups such that the within-cluster sum of squares is minimized. For this study, k-means clustering was undertaken using the Hartigan-Wong algorithm in the 'stats' package (R Core Team, 2014).

We used the Random Forest algorithm (RF; Breiman (2001)) to classify the LULC in the study area. In this study, we tuned three RF parameters that control the structure of the algorithm: the number of trees to grow, or *Ntree*; the number of predictors sampled at each tree node, or *Mtry*; and the minimum size of terminal nodes. The control of the node size parameter defines the minimum number of observations in a terminal node. We used the following parameter values in the RF tuning: *Ntree* = {200, 600, 1000}; *Mtry* = $\{\sqrt{(p/4)}, \sqrt{(p/2)}, \sqrt{p}\}$ with p the total number of variables; and node size = {2, 6, 10}.

Ten RF classifications representing each initial clustering step were performed. We balanced the number of observations by sampling 100 image-objects per LULC class when available, then split the data into 70% for training, while 30% was used for the test set to assess the generalization error of the RF model. Observations of training and test sets were sampled from the first agricultural year (2018-2019), while all observations from the

last year (2019-2020) were set for prediction analysis and LULC maps illustration.

We created individual confusion matrices using the test dataset and computed overall accuracy, producer's accuracy of the double crop class (observations characterised as double crop at the reference dataset, but not assigned as such by the model), and user's accuracy (observations detected as a double crop by the model but not identified as such in the reference dataset). Since double crop clusters represent a unique LULC class, we expand the major diagonal of the confusion matrix of this class for clustered tests. This technique merged false negative and false positive observations into the true positive group allowing us to infer accuracies more properly (Congalton and Green, 2009).

Finally, we checked for linear associations between the overall accuracy and the number of clusters by fitting a linear regression model. The model performance was evaluated by the coefficient of determination (R^2).

RF analysis and tunning were undertaken using the *mlr* package (Bischl et al., 2016), while confusion matrices were derived using the *caret* package (Kuhn, 2008).

3. RESULTS

The first set of analyses examined the NDVI temporal signatures of the double crop class and the intra-class variability. This method was particularly useful since it allowed us to create a matrix of distance-based variability among all land parcels and infer the heterogeneity of the agricultural landscape (Figure 2b).

It is apparent from this matrix that some land parcels (e.g., 06, 08, 21, and 22) presented different NDVI temporal signatures compared to others (red tones). The most striking result to emerge from the matrix is the highest MED between parcels 08 and 11, consequently, the most heterogenous land parcel pair (Figure 2a). There was a visible time shift in soybean harvest along with different patterns in NDVI values. We also observed a time shift in corn-sowing and harvest between these land parcels, where parcel 08 started late than parcel 11.

The white tones indicate an intermediary variability between land parcels. Despite showing lower values of MED, those comparisons have a rather interesting outcome since they can present variability from a particular time period. For instance, parcels 14 and 22 returned a MED of 0.53 (Figure 2c). However, they might have adopted a different corn harvest practice or land cover before the soybean sowing while still presenting very similar summer season NDVI temporal signatures.

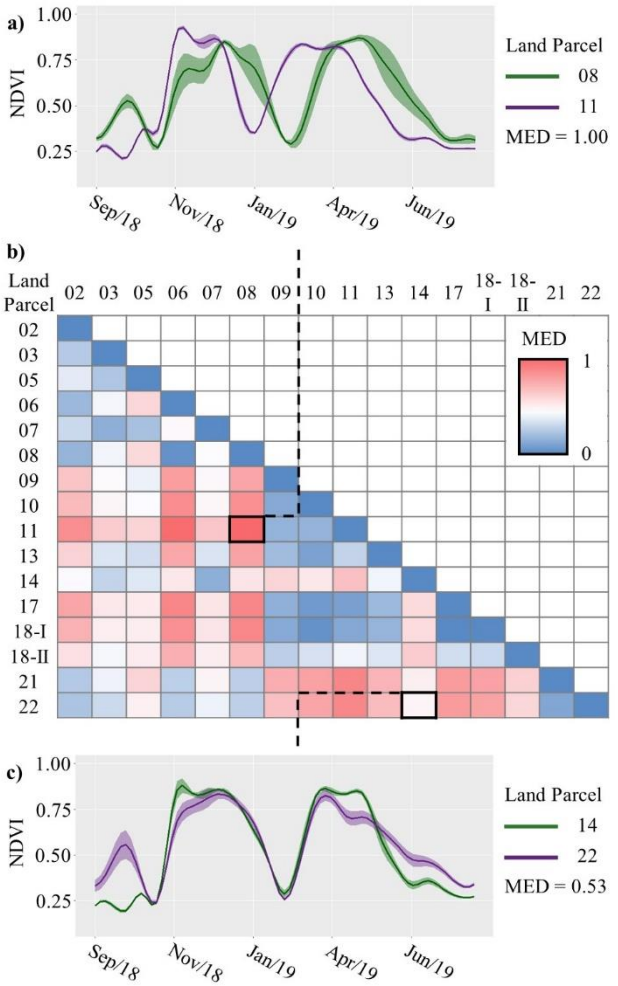

Figure 2. Exploratory analysis of the double crop land parcels' and their paired comparison of NDVI temporal signatures. b) Matrix of distance-based variability among land parcels, illustrating a) the lowest NDVI temporal signature variability between parcels 08 e 11, and (c) a moderate variability between parcels 14 and 22.

Accuracy measures of the test sets are displayed in Table 2. The initial clustering method outperformed the non-clustered classification of LULC in overall accuracy measures.

The most striking observation from this analysis was the classification output with eight initial double crop clusters that returned an overall accuracy of 98.8%, an increase of 3.2% compared to the non-clustered method. This test also returned the highest producer's accuracy of 99.7%, where, considering the 240 double crop observations sampled in the reference dataset, only one was omitted (0.3% of omission error of the double crop class). However, the non-clustered classification returned the highest user's accuracy with an inclusion error of 0.0%.

Although the clustered test with seven and ten clusters completed the top three highest accuracies, we found a weak linear association with R^2 of 0.47 between the overall accuracy and the number of clusters.

NC	OA	Double crop				Rank
		PA	OE	UA	CE	
0	95.6	86.7	13.3	100.0	0.0	10º
2	97.0	98.3	1.7	96.7	3.3	7º
3	96.4	96.7	3.3	96.7	3.3	9º
4	97.4	97.5	2.5	97.5	2.5	4º
5	97.3	98.0	2.0	98.0	2.0	5º
6	96.5	96.7	3.3	98.3	1.7	8º
7	97.8	98.6	1.4	99.0	1.0	2º
8	98.8	99.7	0.3	98.8	1.2	1º
9	97.1	98.9	1.1	97.8	2.2	6º
10	97.8	98.7	1.3	98.3	1.7	3º

Table 2. Accuracy results (in percentage) from cluster-based classification tests. (NC = number of clusters, OA = overall accuracy, PA = producer's accuracy, OE = omission error, UA = user's accuracy, and CE = commission error).

Figure 3 displays the map predictions for each test. Visual interpretation indicates that map predictions were similar and corroborated with the high overall accuracies of the test sets. However, we can observe a few cultivated pasture image-objects included in the double crop class, especially in the tests with 3, 5, and 9 clusters. These commission errors are a rather unexpected result and can not be compared with the user's accuracies of the test sets.

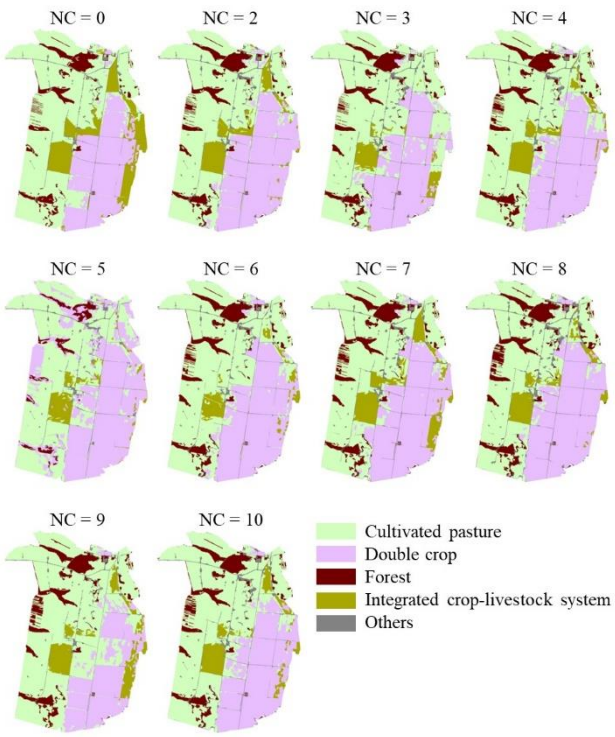

Figure 3. Map predictions of classification tests, where NC: number of clusters.

4. DISCUSSION

In this research, we evaluated the accuracy of LULC classification based on an initial clustering step in a heterogeneous agricultural landscape while checking for variability among their NDVI temporal signatures.

4.1 Effect of an initial clustering step on LULC classifications

Our results showed that an initial clustering method in a heterogenous agricultural class using PS data and RF algorithm produced more accurate results than a non-clustered LULC classification. These results also showed patterns related to plant structure, growing behaviour, and management practice, which allow for highly accurate classifications.

This finding broadly supports the work of other studies in this area, reducing the intra-class spectral-temporal variability of agricultural areas. Bellón et al. (2018) used multiple clustering approaches combining field-level segmentation and landscape stratification to reduce the spectral variability of crops. In that study, the authors demonstrated an improvement in classification accuracies for underrepresented and sparsely distributed cropping systems.

Although we achieved higher overall accuracies based on a higher number of clusters, we found a weak linear association between them. Nevertheless, the present results are significant in one major respect: the very satisfactory accuracy regardless of the clustering step, which indicates the potential of our method.

One unanticipated finding was that accuracy measures of predicted maps, such as the user's accuracy, did not fully mirror those we achieved from the test set. This result may be explained by the fact that we set a different agricultural year for prediction analysis and LULC maps illustration, which was entirely independent of the observations used to train and test the RF model in both time and space. Using yearly independent data sets leads to natural meteorological variations between agricultural years and different phenological responses of vegetation, which may directly affect the map predictions. These results reflect those of Han et al. (2022), who also observed this weather volatility, demonstrating its considerable impact on double crop systems yield variability.

4.2 Intra-class variability of NDVI temporal signatures

Soybeans are recognized as a typical summer crop in double-cropped rotations, particularly in combination with winter corn in the MT state. The exploratory analysis demonstrated that these agricultural land uses may present high intra-class variability and diverse crop calendars for neighbour land parcels. These results reflect those of Arias et al. (2020), who also found considerable variability in crop parcels related to management techniques and cultivars compositions using a crop classification approach based on temporal signatures from Sentinel-1 time series.

In our study, soybean-related NDVI temporal signatures varied little, although we observed a slight time shift between some land parcels during the harvest period. On the other hand, corn-related NDVI temporal signatures varied the most. Paired land parcels presented up to one month of time shift considering the sowing period. Harvest periods also varied depending on the management practice and season length of cultivars, e.g., plant residuals might be left to dry after the harvest or removed mechanically.

Another source of variability was the land cover prior to soybean sowing, where we observed land parcels with both low and high NDVI values. The first indicates the fallow period between rotations, and the latter may indicate a crop emergence of residual seeds left during harvest.

The variability among land parcels may partly be explained by the size of the crop area (approximately 2900 hectares) and the 15-day time interval of PS composites, which makes the time regularity of management practices impracticable.

We reinforce that such high variability should be better addressed by designing parcel-specific strategies and not based only on broad class-related spectro-temporal signatures.

Given the promising results, however, further investigation is needed to provide greater insight into the effects of this methodology on other crop dynamics. Moreover, more robust classifiers, such as deep learning architectures, should be considered to improve model accuracies and to better generalise the predictions to other areas.

5. CONCLUSION

The high intra-class variability of agricultural parcels has demanded reliable approaches for producing accurate LULC classification maps. The present study investigated the accuracy of LULC classification based on an initial clustering step in a heterogeneous agricultural landscape while checking for variability among their NDVI temporal signatures. In a heterogeneous agricultural area located in the state of MT, we examined the hypothesis that grouping observations with similar spectro-temporal signatures would improve the performance of LULC classification over a conventional method.

Results showed that some land parcels presented distinct NDVI temporal signatures.

The initial clustering scheme outperformed the non-clustered classification of LULC in overall accuracy measures regardless of the number of clusters, which addressed the complexity and variability of spectro-temporal signatures. These clusters identified parcel-specific dynamics of management practices and cropping intensity, reducing the uncertainty in classifying the area.

In order for countries to implement more effective and sustainable development of agriculture, there is a need to improve the accuracy of thematic maps produced through the process of LULC classification, especially in heterogeneous agricultural areas. The accuracies achieved with this approach represent promising opportunities toward the sufficiently accurate classification of such areas. In addition, the knowledge of the intra-class variability allows the analyst to classify agricultural areas and infer the temporal dynamics, whether caused by management practices or cropping intensity.

Further work could assess other types of classifiers, such as deep learning architectures, especially in areas with a large number of crop types and distinct management practices.

6. ACKNOWLEDGEMENTS

The authors are thankful for the financial support from FAPESP (Grants: 2021/15001-9 and 2017/50205-9). R. A. C. Lamparelli thanks the National Council for Scientific and Technological Development (CNPq) for the scholarship (Grant: 305271/2020-2)

7. REFERENCES

Achanta, R., Süsstrunk, S., 2017. Superpixels and polygons using simple non-iterative clustering. *30th IEEE Conference on Computer Vision and Pattern Recognition*, 4895–4904. https://doi.org/10.1109/CVPR.2017.520

Arias, M., Campo-Bescós, M. Á., Álvarez-Mozos, J., 2020. Crop Classification Based on Temporal Signatures of Sentinel-1 Observations over Navarre Province, Spain. *Remote Sensing*, *12*(2), 278. https://doi.org/10.3390/rs12020278.

Azar, R., Villa, P., Stroppiana, D., Crema, A., Boschetti, M., Brivio, P. A., 2016. Assessing in-season crop classification performance using satellite data: a test case in Northern Italy. *European Journal of Remote Sensing*, *49*(1), 361–380. https://doi.org/10.5721/EuJRS20164920.

Bellón, B., Bégué, A., Lo Seen, D., Lebourgeois, V., Evangelista, B. A., Simões, M., Demonte Ferraz, R. P., 2018. Improved regional-scale Brazilian cropping systems’ mapping based on a semi-automatic object-based clustering approach. *International Journal of Applied Earth Observation and Geoinformation*, *68*, 127–138. https://doi.org/10.1016/j.jag.2018.01.019.

Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., Casalicchio, G., Jones, Z. M., 2016: Mlr: Machine learning in R. *Journal of Machine Learning Research*, *17*, 1–5. https://github.com/mlr-org/mlr.

Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004.

Breiman, L., 2001. Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324.

Brinkhoff, J., Vardanega, J., Robson, A. J., 2019. Land Cover Classification of Nine Perennial Crops Using Sentinel-1 and -2 Data. *Remote Sensing*, *12*(1), 96. https://doi.org/10.3390/rs12010096.

Bueno, I. T., Acerbi Júnior, F. W., Silveira, E. M. O., Mello, J. M., Carvalho, L. M. T., Gomide, L. R., Withey, K., Scolforo, J. R. S., 2019. Object-Based Change Detection in the Cerrado Biome Using Landsat Time Series. *Remote Sensing*, *11*(5), 570. https://doi.org/10.3390/rs11050570.

Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C., Wardlow, B., Li, Z., 2018. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. *Remote Sensing of Environment*, *210*, 35–47. https://doi.org/10.1016/J.RSE.2018.02.045.

Congalton, R. G., Green, K., 2009. *Assessing the accuracy of remotely sensed data: principles and practices*. CRC Press/Taylor & Francis.

Desclée, B., Bogaert, P., Defourny, P., 2006. Forest change detection by statistical object-based method. *Remote Sensing of Environment*, *102*(1–2), 1–11. https://doi.org/10.1016/j.rse.2006.01.013.

Duro, D. C., Franklin, S. E., Dubé, M. G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment*, *118*, 259–272. https://doi.org/10.1016/j.rse.2011.11.020.

FAO, 2021. FAO Statistical Yearbook 2021: World Food and Agriculture. In *Food and Agriculture Organization of the United Nations (FAO)*. FAO. https://doi.org/10.4060/CB4477EN.

Garrett, R. D., Koh, I., Lambin, E. F., le Polain de Waroux, Y., Kastens, J. H., Brown, J. C., 2018. Intensification in agriculture-forest frontiers: Land use responses to development and conservation policies in Brazil. *Global Environmental Change*, *53*, 233–243. https://doi.org/10.1016/J.GLOENVCHA.2018.09.011.

Han, S.-S., Park, H.-J., Shin, T., Ko, J., Choi, W.-J., Lee, Y.-H., Bae, H.-S., Ahn, S.-H., Youn, J.-T., Kim, H.-Y., 2022. Effects of Tillage System, Sowing Date, and Weather Course on Yield of Double-Crop Soybeans Cultivated in Drained Paddy Fields. *Agronomy*, *12*(8), 1901. https://doi.org/10.3390/agronomy12081901.

Huete, A., Didan, K., Miura, T., Rodriguez, E. ., Gao, X., Ferreira, L., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1–2), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2.

Huete, A. R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*(3), 295–309. https://doi.org/10.1016/0034-4257(88)90106-X.

IBGE., 2021. *Produção Agrícola Municipal*. IBGE. https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9117-producao-agricola-municipal-culturas-temporarias-e-permanentes.html?=&t=destaques.

Kuhn, M., 2008. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, *28*(5), 1–26. https://doi.org/10.18637/jss.v028.i05.

Mathur, A., Foody, G. M., 2008. Crop classification by support vector machine with intelligently selected training data for an operational application. *International Journal of Remote Sensing*, *29*(8), 2227–2240. https://doi.org/10.1080/01431160701395203.

Orynbaikyzy, A., Gessner, U., Conrad, C., 2019. Crop type classification using a combination of optical and radar remote sensing data: a review. *International Journal of Remote Sensing*, *40*(17), 6553–6595. https://doi.org/10.1080/01431161.2019.1569791.

Pascucci, S., Carfora, M., Palombo, A., Pignatti, S., Casa, R., Pepe, M., Castaldi, F., 2018. A Comparison between Standard and Functional Clustering Methodologies: Application to Agricultural Fields for Yield Pattern Assessment. *Remote Sensing*, *10*(4), 585. https://doi.org/10.3390/rs10040585.

Planet Team., 2022. *Planet Imagery Product Specifications*. https://assets.planet.com/docs/Planet_Combined_Imagery_Product_Specs_letter_screen.pdf.

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., Sorooshian, S., 1994. A Modified Soil Adjusted Vegetation Index. *Remote Sensing of Environment*, *48*, 119–126.

Qin, R., Liu, T., 2022. A Review of Landcover Classification with Very-High Resolution Remotely Sensed Optical Images—Analysis Unit, Model Scalability and Transferability. *Remote Sensing*, *14*(3), 646. https://doi.org/10.3390/rs14030646.

R Core Team., 2014. *A language and environment for statistical computing*.

Reis, A. A., Werner, J. P. S., Silva, B. C., Figueiredo, G. K. D. A., Antunes, J. F. G., Esquerdo, J. C. D. M., Coutinho, A. C., Lamparelli, R. A. C., Rocha, J. V., Magalhães, P. S. G., (2020). Monitoring Pasture Aboveground Biomass and Canopy Height in an Integrated Crop–Livestock System Using Textural Information from PlanetScope Imagery. *Remote Sensing*, *12*(16), 2534. https://doi.org/10.3390/rs12162534.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., Haas, R. H., Schell, J. A., Deering, D. W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. *Third Earth Resources Technology Satellite-1 Symposium*, 309–317.

Vali, A., Comai, S., Matteucci, M., 2020. Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. *Remote Sensing*, *12*(15), 2495. https://doi.org/10.3390/rs12152495.

Villa, P., Stroppiana, D., Fontanelli, G., Azar, R., Brivio, P., 2015. In-Season Mapping of Crop Type with Optical and X-Band SAR Data: A Classification Tree Approach Using Synoptic Seasonal Features. *Remote Sensing*, *7*(10), 12859–12886. https://doi.org/10.3390/rs71012859.