



## ARTICLE

# Classification of sugarcane areas in Landsat images using machine learning algorithms

Ana Clara Arantes Villas Bôas de Barros,<sup>\*</sup><sup>1</sup> Marcelo Andrade da Silva,<sup>1</sup> Rodrigo Domiciano Marques,<sup>1</sup> Cristian Villegas,<sup>1</sup> and Ana Cláudia dos Santos Luciano<sup>2</sup>

<sup>1</sup>Department of Exact Sciences, “Luiz de Queiroz” College of Agriculture, University of São Paulo, Piracicaba, Brazil

<sup>2</sup>Department of Biosystems Engineering, “Luiz de Queiroz” College of Agriculture, University of São Paulo, Piracicaba, Brazil

\*Corresponding author. Email: anaarantes@usp.br

(Received: September 29, 2023; Revised: December 28, 2023; Accepted: January 22, 2024; Published: April 15, 2024)

### Abstract

Monitoring sugarcane areas through remote sensing is essential for planning and management of the national sugarcane industry. The use of machine learning algorithms has provided many benefits to remote sensing. This article aims to compare the prediction quality of three important machine learning methods in identifying sugarcane areas using Landsat images: Logistic Regression (LR), Decision Tree (DT) and Random Forest (RF). LR was applied in three versions: LR without penalization, LR with Ridge penalization (LR-R) and LR with Lasso penalization (LR-L). Data obtained in this study refer to a region of approximately 306,000 ha located in the state of São Paulo, Brazil, which was segmented into approximately 46,000 polygons (observations). Six spectral bands and vegetation indices observed along 17 months resulted in 102 covariates, which were reduced via Principal Component Analysis (PCA). In total, 19 Principal Components were chosen to account for 94.61% of the cumulative explained variance ratio and were used in machine learning methods to classify each polygon as sugarcane or other land covers. The method with the highest accuracy considering testing sample of 20% of data was RF (78.51%), followed by DT (72.30%), LR-L (69.64%), LR-R (69.64%), and LR (69.52%).

**Keywords:** Land cover prediction, Random Forest, Remote sensing, Sugarcane mapping.

## 1. Introduction

Sugarcane is of great importance in Brazilian agribusiness, as Brazil is the world’s largest sugar producer and one of the largest biofuel markets (Luciano *et al.*, 2019). This crop has production disseminated throughout the country, with greater concentration in the southeastern region, with the state of São Paulo accounting for 60% of national production (Silva *et al.*, 2011). According

to Evaluation Bulletins of the National Supply Company (in Portuguese, *CONAB - Companhia Nacional de Abastecimento*), the 2022/2023 harvested area in the state of São Paulo exceeded 4 million hectares, reaching productivity of 75.43 tons per hectare and production of around 312 million tons.

According to de Souza Dias *et al.*, 2015, sugarcane has been used as feedstock for the production of sugar and ethanol. The sugar produced in Brazil is used in the manufacture of foods and beverages and is also exported to several countries. Ethanol is a renewable fuel widely used in transport, helping to reduce the use of fossil fuels. Currently, Brazilian ethanol accounts for 30% of global production, being one of the most prominent biofuels produced on a large scale, second only to US ethanol production (Karp *et al.*, 2021).

Given the importance of sugarcane cultivation in Brazil, it is natural for the sector to increasingly invest in resources and technologies to improve production. The various studies on mapping and monitoring sugarcane plantation areas through remote sensing are good examples (Johnson *et al.*, 2014; Zhou *et al.*, 2015; Luciano *et al.*, 2018; Wang *et al.*, 2019; Luciano *et al.*, 2022). These studies have used satellite images to monitor large sugarcane plantation areas, vegetation indices and machine learning algorithms, achieving good results in mapping sugarcane areas. Moreover, the authors highlighted the importance of vegetation indices to identify sugarcane plantations and the good results from machine learning as random forest, artificial neural network, support vector machine and others.

This article focused on comparing the prediction quality of three important machine learning methods in identifying sugarcane areas using Landsat images: Logistic Regression, Decision Tree and Random Forest. Vegetation indices and spectral bands were obtained from Landsat satellite images in a region with large concentration of sugarcane plantations in the state of São Paulo.

This paper is organized as follows: Section 2 describes sugarcane data, machine learning methods used in reducing data dimensionality and classifying sugarcane areas, and the accuracy measures used to compare results. Section 3 provides the study results and discussion, including the interpretation of the new dimensions of data and comparison of the performance of methods. Some conclusions and remarks are shown in Section 4.

## 2. Materials and Methods

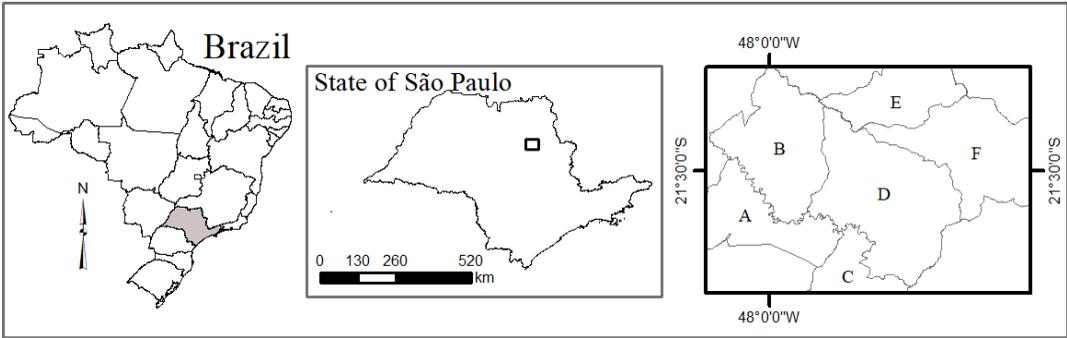
This section describes the dataset used and presents the machine learning methods used to perform data dimensionality reduction, identification of sugarcane areas and performance metrics of methods.

### 2.1 Sugarcane data

The test area selected for this study has approximately 306,000 ha and is located in the mid-northern region of the state of São Paulo, Brazil. Figure 1 shows the location of the study region.

The study region was segmented based on satellite images into approximately 46,000 polygons. Aiming to calibrate classification models, each polygon was classified as “sugarcane” or “other”, according to reference maps from the Canasat project (Rudorff *et al.*, 2010).

The dataset was obtained from spectral indices captured by the Landsat satellite, already processed by the method used in Luciano *et al.*, 2018. Spectral indices are measurements of the interaction of light with matter which are used to identify different types of materials, such as water, vegetation, and soil, and also to measure characteristics, such as vegetation health, soil moisture, and soil nutrient concentration. In remote sensing, spectral indices are used to identify different land cover types, such as vegetation, water and soil, and to identify plantations such as sugarcane, wheat, rice, soybeans and corn, among others (Luciano *et al.*, 2019; Lu *et al.*, 2022).



**Figure 1.** Location of the study region in the mid-northern portion of the state of São Paulo with overlapping municipal boundaries. A: Rincão; B: Guataporá; C: São Carlos; D: Luís Antônio; E: Cravinhos; F: São Simão.

To help discriminate between the two classes of the target variable (sugarcane or not), information on the annual growth cycle and on the previous crop was considered through spectral indices shown in Table 1 monthly collected between January 2015 and May 2016, totaling 17 months. The dataset has 102 predictor variables (17 months × 6 indices) and approximately 46,000 observations.

**Table 1.** Description of vegetation indices. *NIR*, *R*, *B*, and *G* are reflectances in the near-infrared, red, blue and green bands, respectively

Indices or bands	Identification	Equation or wavelength [ <i>min</i> , <i>max</i> ] ( $\mu\text{m}$ )
Enhanced vegetation index (Huete <i>et al.</i> , 1997)	EVI	$2.5 \cdot \frac{NIR - R}{NIR + 6 \cdot R - 7.5 \cdot B + 1}$
Normalized difference moisture index (Wilson & Sader, 2002)	NDMI	$\frac{NIR - SWIR1}{NIR + SWIR1}$
Normalized difference vegetation index (Rouse <i>et al.</i> , 1973)	NDVI	$\frac{NIR - R}{NIR + R}$
Normalized difference water index (McFEETERS, 1996)	NDWI	$\frac{G - NIR}{G + NIR}$
Short-wave infrared 1	SWIR1	[1.55, 1.75]
Short-wave infrared 2	SWIR2	[2.08, 2.35]

## 2.2 Machine learning methods

In this subsection, a brief description of the statistical and machine learning methods used in this paper will be presented.

### 2.2.1 Dimensionality reduction via PCA

Before running a machine learning algorithm, Principal Component Analysis (PCA, Johnson & Wichern, 2002) can be used to reduce data mass, with the least possible loss of information. This is helpful, for instance, to speed up the execution of the machine learning algorithm. PCA is a statistical technique of multivariate analysis that transforms a set of correlated variables into a new set of variables called principal components, which have important properties: (1) each principal component is a linear combination of all original variables; (2) they are independent of each other;

and (3) they are obtained with the purpose of retaining, in descending order, the maximum amount of information, in terms of the total variation contained in data.

In order to calculate the principal components, one first calculates the matrix of variances and covariance of the  $d$  considered variables. Commonly, the values of variables are standardized to ensure that they are weighted equally in the analysis. When values are standardized, the correlation matrix is used instead of the variance and covariance matrix. Then, the eigenvalues and eigenvectors of this matrix are calculated. Eigenvalues are the variances of principal components, while eigenvectors are vectors that point in the direction of principal components. Eigenvectors, also referred to as loading, correspond to the coefficients used in the linear combinations of the original variables to create each principal component (PC). The  $i$ th principal component is obtained by projecting data onto the  $i$ th eigenvector (sorted in descending order by the corresponding eigenvalues). For example, to calculate the first principal component, simply multiply the standardized dataset by the first eigenvalue.

### 2.2.2 Logistic regression

In the logistic regression model, the response variable  $Y$  follows a Bernoulli distribution, that is, the probability of success is given by  $P(Y = 1) = \pi$  and the probability of failure as  $P(Y = 0) = 1 - \pi$ . The probability mass function for the Bernoulli random variable  $Y$  with success probability  $\pi \in (0, 1)$  is given by:

$$P(Y = y) = \pi^y(1 - \pi)^{1-y}, \quad y \in \{0, 1\}.$$

This model is widely used in classification problems in different areas of knowledge. In practical problems associated with each response variable  $Y_i$  a vector of covariates  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{id})'$  is obtained, with  $i = 1, 2, \dots, n$ . Then, the Logistic Regression model can be written as, see McCullagh & Nelder, 1989:

$$P(Y_i = 1|\mathbf{x}_i) = \pi_i = \frac{\exp(\beta_0 + \sum_{j=1}^d \beta_j x_{ij})}{1 + \exp(\beta_0 + \sum_{j=1}^d \beta_j x_{ij})}, \quad i = 1, 2, \dots, n \quad (1)$$

where  $\beta_0$  is the intercept of the model and  $\beta_j$  is the parameter associated with the  $j$ th covariate, with  $j = 1, 2, \dots, d$ . To estimate the parameters of a Logistic Regression model, the maximum likelihood method can be used, which consists of obtaining parameters that maximize the likelihood function.

It is also possible to use penalties such as Lasso or Ridge regression to estimate the Logistic Regression parameters, which are techniques used to prevent overfitting and improve model interpretability. Both achieve this result by adding a penalty term to the loss function, but they differ in the type of penalty used and their effect on coefficients. While Lasso uses the sum of the absolute values of parameters, Ridge uses the sum of their squares.

Both regularization techniques have some specific characteristics. For example, Lasso can take coefficients close to zero and can even set some to exactly zero. This performs covariate selection by effectively removing irrelevant covariates from the model and can be especially useful for high-dimensional datasets with many correlated covariates. On the other hand, Ridge can take coefficients close to zero, but less aggressively than Lasso, as the coefficients tend to remain non-zero. Thus, Ridge does not perform explicit covariate selection. For more details see James *et al.*, 2013.

### 2.2.3 Decision tree

Decision tree was proposed by Breiman *et al.*, 1984 and can be applied for both classification (predicting categorical values) and regression problems (predicting continuous values). A decision tree is built by recursive partitioning on the covariate space. Each partitioning is named node and each end result is named leaf. Initially, the algorithm checks whether the condition at the first node

is satisfied. If so, go to the left. Otherwise, continue to the right. This continues until a single leaf is obtained.

A decision tree structure for classification problems can be created in two main steps: (i) the creation of a complete and complex tree and (ii) the pruning of this tree, in order to avoid overfitting. Formally, the first step consists of creating a partition of the covariate space generated by covariates  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{id})'$ , with  $i = 1, 2, \dots, n$ , into  $M$  regions distinct and disjoint denoted by  $R_1, R_2, \dots, R_M$ . The proportion of elements in the  $m$ th region belonging to the  $k$ th class,  $k=0,1$  for binary problems, is given by

$$\hat{p}_{m,k} = \frac{1}{N_m} \sum_{\mathbf{x}_i \in R_m} I(y_i = k),$$

where  $N_m$  is the number of elements in the region  $R_m$ . A suggested criterion for finding the best partition at each stage of the process is the Gini index, given by

$$\sum_R \sum_{k \in \{0,1\}} \hat{p}_{m,k}(1 - \hat{p}_{m,k}),$$

where  $R$  represents one of the regions induced by the tree. The objective is to minimize this index.

For the pruning stage, in general, the proportion of errors in the validation set is used as a risk estimate. More details about regression trees (predicting continuous values) and the pruning process can be found in Hastie *et al.*, 2001.

#### 2.2.4 Random forest

Although decision trees are methods of easy interpretation and simple understanding, they tend to have low predictive power when compared to other estimators. To overcome this limitation, another well-known method called Random Forest can be explored. This approach consists of obtaining  $B$  distinct trees and combining their results to improve the predictive power in relation to an individual tree. To create  $B$  distinct trees,  $B$  *bootstrap* samples from the original sample are used. Let  $\hat{C}_b(\mathbf{x})$  be the class prediction of the  $b$ th random forest tree, then the prediction function is given by

$$\hat{C}^B(\mathbf{x}) = \text{mode}\{\hat{C}_b(\mathbf{x}), \quad b = 1, 2, \dots, B\}.$$

More details about the random forest can be obtained in Hastie *et al.*, 2001.

In real life, the dataset is used to train a machine learning model/algorithm, which is used to predict outputs that are unknown on new datasets. To check the predictive capability of the model for new data, it is common to divide data into two parts, a larger one to train the method and a smaller one to validate it. This procedure is known as the Holdout method (Morettin & Singer, 2022; Izbicki & dos Santos, 2020; Lantz, 2019). This split ensures that the model would not simply memorize the training data, but truly understand the underlying rules, leading to robust and generalizable performance.

### 2.3 Confusion matrix and performance metrics

The confusion matrix is a table that presents the performance of a classification method (Morettin & Singer, 2022; Izbicki & dos Santos, 2020; Lantz, 2019). It is commonly used in supervised learning problems where the output data is known. For a classification problem with only two possible classes, the confusion matrix has four main elements: true positive (TP), where the method correctly classifies an instance as positive; false positive (FP), where the method incorrectly classifies an instance as positive; true negative (TN), where the method correctly classifies an instance as

**Table 2.** The structure of a confusion matrix for binary problems

		Prediction outcome	
		Positive	Negative
Actual value	Positive	True positive (TP)	False negative (FN)
	Negative	False positive (FP)	True negative (TN)

negative; and false negative (FN), where the method incorrectly classifies an instance as negative. An example of confusion matrix is shown in Table 2.

Based on the four elements of the confusion matrix, three performance metrics are used in this work: sensitivity, specificity and accuracy. These metrics are defined below.

- **Sensitivity:** measures how well the method can correctly identify instances of the positive class. In other words, it is the proportion of positive results, given that the actual value is positive. It is calculated as:

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

- **Specificity:** measures how well the method can correctly identify instances of the negative class. That is, it is the proportion of negative results, given that the actual value is negative. It is calculated as:

$$\text{Specificity} = \frac{TN}{TN + FP}$$

- **Accuracy:** measures the proportion of correctly classified instances in relation to the total number of instances. It is calculated as:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

### 3. Results and Discussion

Spectral bands and vegetation indices observed over 17 months resulted in 102 covariates (inputs) to predict the target variable (classification as sugarcane or not). After performing PCA (using the correlation matrix of data), principal components with eigenvalues greater than 1 were chosen to be applied in machine learning methods to classify each polygon of the study area as sugarcane (1 - positive) or other land cover (0 - negative). The 19 principal components accounted for 94.61% of the cumulative explained variance ratio, as shown in Table 3.

The loading of the four first principal components is shown in Figure 2. Component 1 can be defined as a contrast between vegetation indices (EVI and NDVI) and indices related to moisture and water detection (NDMI, NDWI, SWIR 1 and SWIR 2) through all months, from January/2015 to May/2016. A pattern in the loading of component 2 was observed, with lower (negative) values from November/2015 to January/2016, which correspond to rainy months in the state of São Paulo.

Seasonal pattern of loading related to vegetation and moisture/water indices may be suggested, as seen in the contrast along the months, especially in mid-fall/winter seasons (from May/2015 to September/2015, in component 3) and the first and second semesters of 2015, in component 4. Seasonal variation in the vegetation and moisture/water index should be carefully taken into account since the sugarcane areas under study presented diverse harvest management and cultural practices throughout the months.

The remaining components were used in the classification algorithms as previously described, but their loading values are not shown in Figure 2.

The 19 principal components were used to form the new database, which was split into training (80%) and testing (20%) subsets prior to applying the machine learning methods, including the

**Table 3.** Principal component, explained variance ratio and cumulative explained variance ratio

Principal component	Explained variance ratio	Cumulative explained variance ratio
1	29.82	29.82
2	8.35	38.17
3	8.02	46.19
4	7.32	53.51
5	5.15	58.66
6	4.77	63.43
7	4.21	67.64
8	3.87	71.51
9	3.05	74.56
10	2.91	77.47
11	2.51	79.98
12	2.41	82.39
13	2.19	84.58
14	2.11	86.69
15	2.00	88.69
16	1.85	90.54
17	1.56	92.10
18	1.41	93.51
19	1.10	94.61

calculation of the Lasso and Ridge penalty parameters of the logistic regression. The confusion matrices of prediction methods applied to the testing subset are presented side by side in Table 4.

**Table 4.** Confusion matrices for the classification algorithms. LR: logistic regression; LR R: logistic regression with Ridge; LR L: logistic regression with Lasso; DT: decision tree; RF: random forest

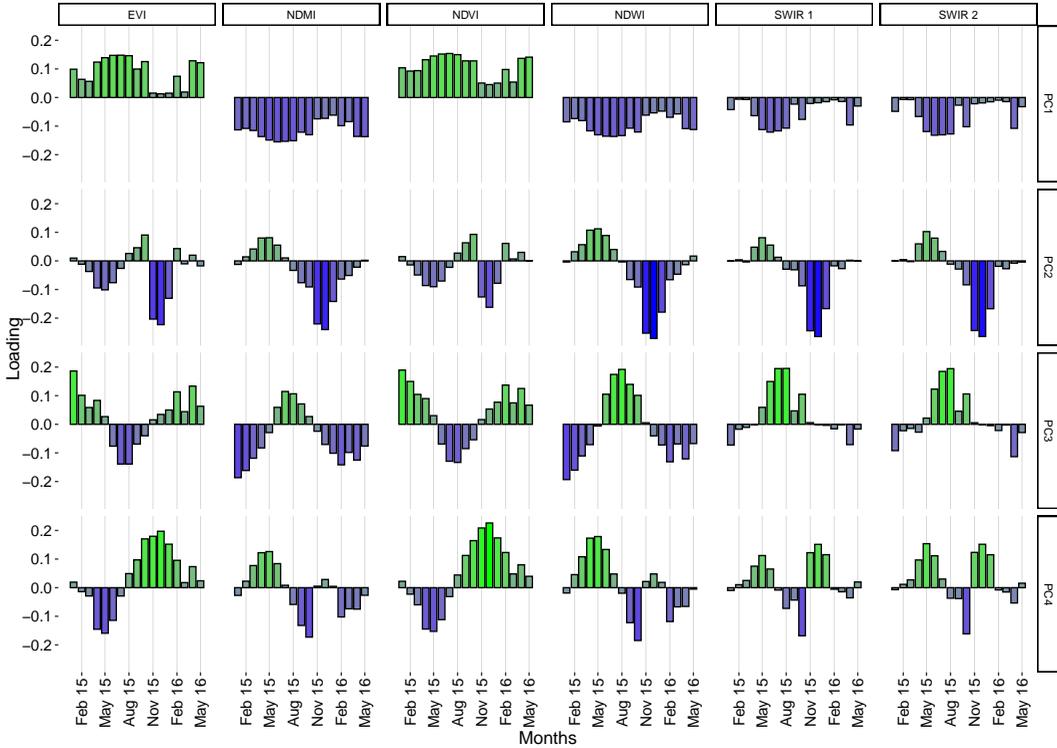
		Prediction outcome									
		LR		LR R		LR L		DT		RF	
		1	0	1	0	1	0	1	0	1	0
Actual value	1	4341	1476	4384	1433	4388	1429	4754	1063	5265	552
	0	1310	2013	1342	1981	1346	1977	1469	4754	1412	1911

The performance metrics, sensitivity, specificity and accuracy, with respective 95% confidence intervals, for the classification methods used are presented in Table 5.

The logistic regression, with or without Ridge and Lasso penalization, showed similar results for sensitivity (0.7462, 0.7537 and 0.7543, respectively), specificity (0.6058, 0.5961 and 0.5949, respectively) and accuracy (0.6952, 0.6964 and 0.6964, respectively) (Table 5). Consequently, penalization did not interfere with classification performance and thus can be considered unnecessary to classify sugarcane land cover under the conditions tested.

Although Decision tree accuracy, 0.7230 [0.7137, 0.7321], was greater than observed with logistic regression, the method's performance metrics were inferior to those obtained using Random Forest.

Random Forest presented greater sensitivity, 0.9051, compared to logistic regression and decision tree, and specificity, 0.5751, similar to values shown by other algorithms, which resulted in the greatest accuracy, 0.7851 [0.7766, 0.7935], observed among classification methods used (Table 5). Therefore, Random Forest can be considered the best option to classify sugarcane land cover using



**Figure 2.** Loading of the four first principal components.

**Table 5.** Performance metrics for classification algorithms. LR: logistic regression; LR R: logistic regression with Ridge; LR L: logistic regression with Lasso; DT: decision tree; RF: random forest

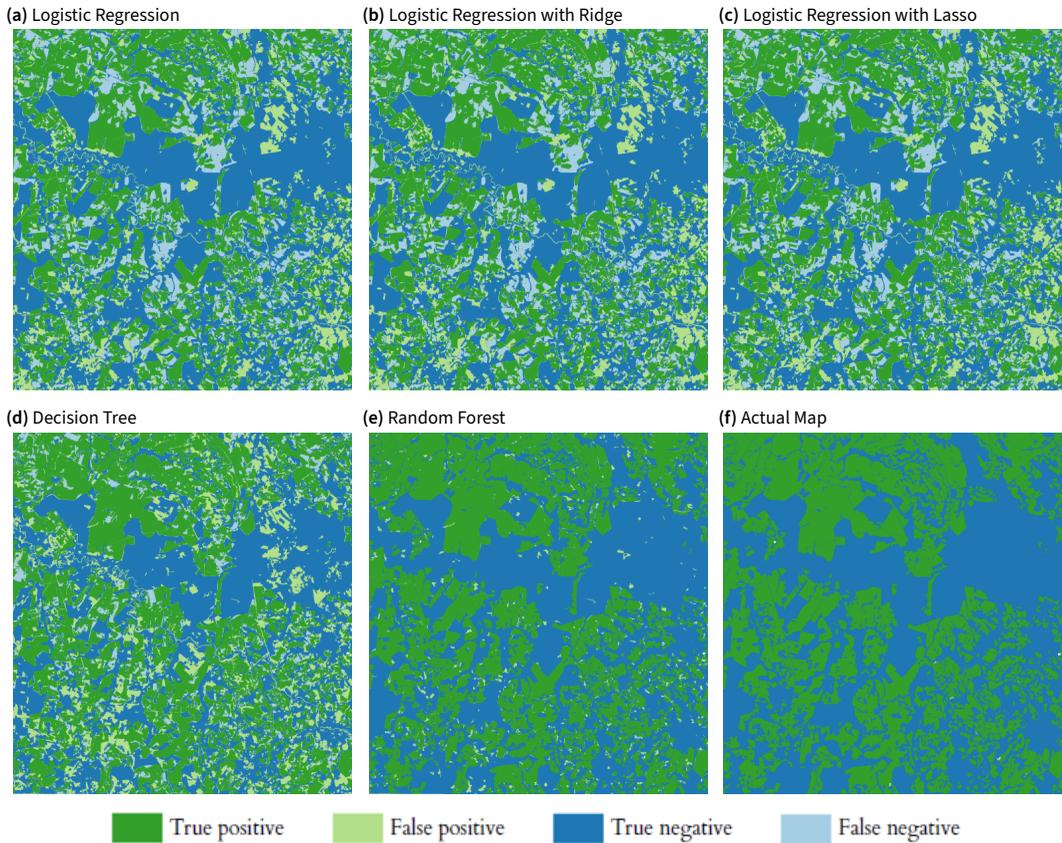
	LR	LR R	LR L	DT	RF
Sensitivity	0.7463	0.7537	0.7543	0.8173	0.9051
Specificity	0.6058	0.5961	0.5949	0.5579	0.5751
Accuracy	0.6952	0.6964	0.6964	0.7230	0.7851
	[0.6856, 0.7046]	[0.6868, 0.7058]	[0.6868, 0.7058]	[0.7137, 0.7321]	[0.7766, 0.7935]

PCA and vegetation indices under the conditions described in this study.

Figure 3 contains the maps of the study region, comparing the polygon classification predictions obtained by the five methods and the true classification. The maps were created using the trained methods and the complete dataset (training and testing) to fully form the maps. As the training set has greater accuracy, the images may appear to be more accurate than those described in Table 5.

The map of the study region shows areas highlighted in dark colors (dark green and dark blue), representing polygons that the methods correctly predicted as sugarcane (true positive) and non-sugarcane (true negative), respectively. Areas in green tones indicate positive predictions, that is, areas classified as sugarcane. The three maps referring to LR (maps a, b and c) are similar and have large number of polygons classified as false positive (light green) and false negative (light blue). In contrast, the map referring to DT (map d) has large number of polygons classified as false positive (light green). Finally, the map referring to the RF (map e) has few incorrectly classified polygons, resembling the actual map (map f).

The logistic regression model was used as a parametric approach for classification with and with-



**Figure 3.** Representation of polygons on the map of the study region with colors to present the prediction results of methods considered in the study, as well as the actual map.

out a penalty term (Lasso and Ridge). For a non-parametric approach, decision tree and random forest were used. The decision tree algorithm is simple and easy to interpret. Furthermore, by aggregating many decision trees, using methods such as random forest, the predictive performance of trees could be improved. As final consideration, it makes no sense to structurally compare parametric and non-parametric approaches because they make different assumptions about the methods.

## 4. Conclusions

In this paper, the prediction quality of logistic regression, decision tree and random forest was compared in identifying sugarcane areas using Landsat images. Methods were calibrated on training data covering 17 consecutive months, starting on January 2015, in a region of 306,000 hectares located in the state of São Paulo, Brazil, which was divided into 46 thousand polygons. The method with the highest accuracy considering test sample of 20% of data was RF, with 78.51%, followed by DT with 72.30%, LR-L with 69.64%, LR-R with 69.64% and LR with 69.52%. However, it is important to note that the performance of machine learning algorithms may vary depending on the specific dataset and the parameters used.

According to performance metrics used in the study, RF is the most recommended method among those investigated for sugarcane mapping, considering the spectral indices from January to May of the following year, totaling 17 months. RF provided satisfactory result for the test set (20%

of the total dataset). It is worth mentioning that the methodology presented here can be applied to map other crops using satellite images.

Future studies should be carried out in order to reduce the number of polygons incorrectly classified as sugarcane (i.e. false positives) by implementing a post-classification method that considers the classification of neighboring segments to confirm the prediction of each segment.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. *Classification and Regression Trees*. (Wadsworth, 1984).
2. De Souza Dias, M. O., Maciel Filho, R., Mantelatto, P. E., Cavalett, O., Rossell, C. E. V., Bonomi, A. & Leal, M. R. L. V. Sugarcane processing for ethanol and sugar in Brazil. *Environmental Development* **15**, 35–51. ISSN: 2211–4645. <https://www.sciencedirect.com/science/article/pii/S2211464515000147> (2015).
3. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning* (Springer New York Inc., 2001).
4. Huete, A., Liu, H. & van Leeuwen, W. *The use of vegetation indices in forested regions: issues of linearity and saturation in IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development* **4** (1997), 1966–1968 vol.4.
5. Izbicki, R. & dos Santos, T. M. *Aprendizado de máquina: uma abordagem estatística* 268 pp. (2020).
6. James, G., Witten, D., Hastie, T. & Tibshirani, R. *An Introduction to Statistical Learning: with Applications in R* (Springer, 2013).
7. Johnson, B. A., Scheyvens, H. & Shivakoti, B. R. An ensemble pansharpening approach for finer-scale mapping of sugarcane with Landsat 8 imagery. *International Journal of Applied Earth Observation and Geoinformation* **33**, 218–225. ISSN: 1569–8432. <https://www.sciencedirect.com/science/article/pii/S0303243414001366> (2014).
8. Johnson, R. & Wichern, D. *Applied multivariate statistical analysis* (Prentice hall Upper Saddle River, NJ, 2002).
9. Karp, S. G., Medina, J. D. C., Letti, L. A. J., Woiciechowski, A. L., de Carvalho, J. C., Schmitt, C. C., de Oliveira Penha, R., Kumlehn, G. S. & Soccol, C. R. Bioeconomy and biofuels: the case of sugarcane ethanol in Brazil. *Biofuels, Bioproducts and Biorefining* **15**, 899–912. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bbb.2195>. <https://onlinelibrary.wiley.com/doi/abs/10.1002/bbb.2195> (2021).
10. Lantz, B. *Machine learning with R: expert techniques for predictive modeling* (Packt publishing ltd, 2019).
11. Lu, M., Bi, Y., Xue, B., Hu, Q., Zhang, M., Wei, Y., Yang, P. & Wu, W. Genetic Programming for High-Level Feature Learning in Crop Classification. *Remote Sensing* **14**, 16 (2022).
12. Luciano, A., Picoli, M., Rocha, J., Franco, H., Sanches, G., Leal, M. & le Maire, G. Generalized space-time classifiers for monitoring sugarcane areas in Brazil. *Remote Sensing of Environment* **215** (July 2018).

13. Luciano, A. C. d. S., Campagnuci, B. C. G. & le Maire, G. Mapping 33 years of sugarcane evolution in São Paulo state, Brazil, using landsat imagery and generalized space-time classifiers. *Remote Sensing Applications: Society and Environment* **26**, 100749. ISSN: 2352-9385. <https://www.sciencedirect.com/science/article/pii/S235293852200057X> (2022).
14. Luciano, A. C. d. S., Picoli, M. C. A., Rocha, J. V., Duff, D. G., Lamparelli, R. A. C., Leal, M. R. L. V. & Le Maire, G. A generalized space-time OBIA classification scheme to map sugarcane areas at regional scale, using Landsat images time-series and the random forest algorithm. *International Journal of Applied Earth Observation and Geoinformation* **80**, 127–136. ISSN: 1569-8432. <https://www.sciencedirect.com/science/article/pii/S0303243418311917> (2019).
15. McCullagh, P. & Nelder, J. *Generalized Linear Models, Second Edition* (Chapman & Hall, 1989).
16. McFEETERS, S. K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing* **17**, 1425–1432. <https://doi.org/10.1080/01431169608948714> (1996).
17. Morettin, P. A. & Singer, J. d. M. *Estatística e ciência de dados* 454 pp. (2022).
18. Rouse, J., Hass, R., Schell, J. & Deering, D. Monitoring vegetation systems in the great plains with ERTS. *Third Earth Resour. Technol. Satell. Symp.* **1**, 309–317. <https://cir.nii.ac.jp/crid/1573105975083785728> (1973).
19. Rudorff, B. F. T., Aguiar, D. A., Silva, W. F., Sugawara, L. M., Adami, M. & Moreira, M. A. Studies on the Rapid Expansion of Sugarcane for Ethanol Production in São Paulo State (Brazil) Using Landsat Data. *Remote Sensing* **2**, 1057–1076. ISSN: 2072-4292. <https://www.mdpi.com/2072-4292/2/4/1057> (2010).
20. Silva, C. B., de Moraes, M. A. F. D. & Molin, J. P. Adoption and use of precision agriculture technologies in the sugarcane industry of São Paulo state, Brazil. *Precision Agriculture* **12**, 67–81 (2011).
21. Wang, M., Liu, Z., Ali Baig, M. H., Wang, Y., Li, Y. & Chen, Y. Mapping sugarcane in complex landscapes by integrating multi-temporal Sentinel-2 images and machine learning algorithms. *Land Use Policy* **88**, 104190. ISSN: 0264-8377. <https://www.sciencedirect.com/science/article/pii/S0264837719307185> (2019).
22. Wilson, E. H. & Sader, S. A. Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sensing of Environment* **80**, 385–396. ISSN: 0034-4257. <https://www.sciencedirect.com/science/article/pii/S0034425701003182> (2002).
23. Zhou, Z., Huang, J., Wang, J., Zhang, K., Kuang, Z., Zhong, S. & Song, X. Object-oriented classification of sugarcane using time-series middle-resolution remote sensing data based on AdaBoost. *PloS one* **10**, e0142069 (2015).