

COMPARISON OF CLASSIFICATION ALGORITHMS OF IMAGES FOR THE MAPPING OF THE LAND COVERING IN TASSO FRAGOSO MUNICIPALITY, BRAZIL

P. R. M. Pereira¹; F. W. D. Costa²; E. L. Bolfe^{1,3*}; L. Macarringe¹; A. C. Botelho¹

¹University of Campinas – Unicamp / Graduate Programme in Geography, Campinas - SP, Brazil.
p.roberto18@hotmail.com; ebolfe@unicamp.br; lucrecio.macarringue@gmail.com; adielsonbotelho@hotmail.com

²Faculdade de Ciências e Tecnologia – UNESP PP, Presidente Prudente - SP, Brazil. geo.fwcosta@gmail.com

³Brazilian Agricultural Research Corporation – Embrapa Informática Agropecuária, Campinas - SP, Brazil.

Commission III, WG III/7

KEY WORDS: Digital Classification, Performance Indexes, Cerrado Biome, Maranhão State, Landsat 8.

ABSTRACT:

One of the main applications of satellite images is the characterization of terrestrial coverage, that from the use of classification techniques, allows the monitoring of spatial transformations of the terrestrial surface, this process being directly associated with the potential of classifiers to differentiate the most diverse data present in the images, a fundamental aspect for the use of remote sensing data. This article evaluates the performance of different classification algorithms in the mapping classes of land use and land cover in medium resolution images from the Landsat 8 program, the test area of this test corresponds to the Municipality of Tasso Fragoso (State Maranhão - Brazil), stands out for a typical vegetation cover of the Cerrado Biome, presents similar spectral patterns that induce high difficulty of class differentiation automatically. In this paper, were analyzed the machine learning algorithms C5.0 and Random Forest in comparison to traditional classification algorithms being the Minimum Distance and the Spectral Angle Mapper. The best results were generated by Random Forest with 90% accuracy and Kappa of 0.861, followed by the C5.0 algorithm. Traditional algorithms, on the other hand, presented a lower precision rate with global accuracy, not exceeding 75% of accuracy and Kappa varying between 0.507 and 0.627. The accuracy of the producer showed that all the algorithms, in major or minor tendency presented difficulties in to differentiate the areas, with rates of mistakes varying between 25 and 75%, being the main, the confusion with pastoral areas.

1. INTRODUCTION

The data associated to the covering and land use enables the detailed comprehension of the spatial organization, and are considered basilar information to many environmental and social economical applications, being it a thematic of relevant interest in the most diversified areas (Azzari, Lobell, 2017; Jin et al. 2017). The panoramic vision and the repeatability of the orbital images enables the obtaining of historical and current information and represent an important resource to the monitoring and natural resources, mainly in the regions of great extension, obtaining results with bigger quickness in relation to the field analysis with a cost relatively low (Li et al. 2016; Espinosa, Schröder, 2019).

The information extraction represents a challenge, many factors, such as the complexity of landscape, scale of information, image processing and approaches of classification may affect the success of a classification (Lu, Weng, 2007). And its analysis is directly associated to the development image classification techniques (Phriri et al. 2018; Noi, Kappas, 2018; Mishra et al., 2020).

The classification of an image consists into attribute meaning to a pixel set depending on its characteristics numerical, that is, giving the pixel a thematic class, solo, water, vegetation, from similar spectral properties, and represent them in map, table, or graphics (Mastella, Vieira, 2018). The automatic image classification is a complex process that can be affected by many factors. It is directly associated with the ability of the algorithms in the distinction of different patterns present in the image,

which represents a challenge, especially in environments with high spectral homogeneity, which directly interferes in the result.

With the increase of the availability of satellite images, with better temporal, radiometric and spatial resolution, caused the development of algorithms that overcame rudimentary classifiers, based only on the spectral characteristics of the images. During the last decade, the mapping of land covering using data resulting from orbital sensors was driven by a change of paradigm in the processes of classification, pointing out the appearance of machine learning algorithms (Maxwell et al. 2018), which present as alternatives to the traditional parametric algorithms (Crowson et al. 2020).

These groups of algorithms enable digital classification as of use of great volumes of data in a time relatively low and excellent performance in differentiation of classes with elevated similarity in remote sensing (Foody, 2002; Li et al. 2014). Methods such as decision and regression trees, vectors of changing, random forests are among the most used for the classification of remote sensing data and have been spreading mainly due to its capacity of processing time which allows automating the spatial analysis (Deng et al. 2019; Duro et al. 2012).

In this way, the general objective of this paper is to evaluate and compare the performance of classification algorithms: Minimal Distance (MD); Spectral Angle Mapper (SAM), Random Forest (RF), C5.0 Decision Tree for the mapping of classes of land covering to the cerrado biome as of Landsat 8 images.

*Corresponding author

2. MATERIAL AND METHODS

The municipality of Tasso Fragoso (Figure 1) owns a territorial extension of 4.382 Km², belonging to Micro region of Gerais de Balsas, is located in the south portion of the Maranhão state, having the coordinates 43°0'19.18"W; 42°40'7.96"W and 3°53'14.07"S, 03°18'22"S, and limited in the North with the municipality of Sambaíba; in South with municipality of Alto Parnaíba and in the West with the municipality of Balsas (Maranhão State) in the East with Piauí state (IBGE, 2017).

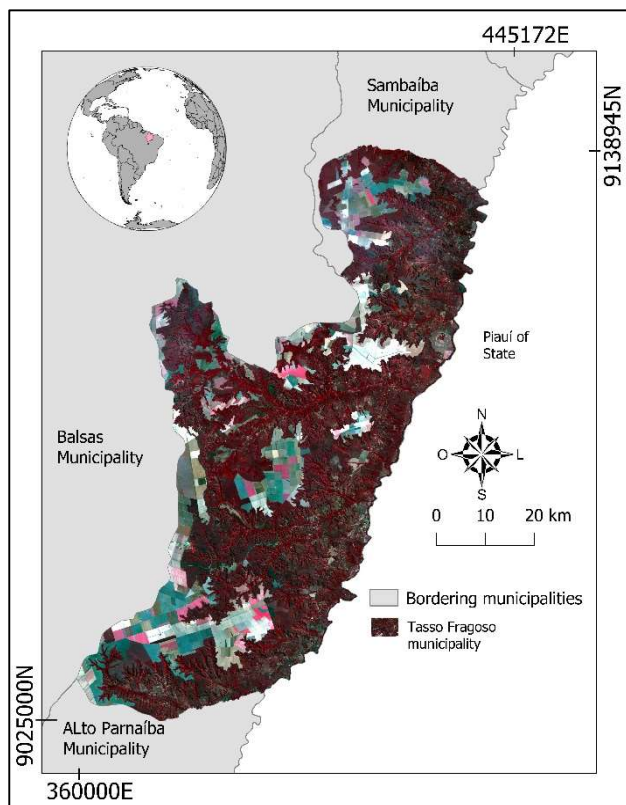


Figure 1. Localization of Tasso Fragoso municipality – MA (Brazil).

With a predominantly agricultural economy, the municipality owns a gross domestic product per capita of R\$ 116.445,6, highlighting in state scenario with the second major grains producer, predominating the growing of cotton and soybean. In the of 2019 the soybean growing represented 70% of all the area destined to tillage of the municipality, which represented 19% of all produced in Maranhão state, with more than 10 million tons (IBGE, 2019).

For this study, it was used two scenes from the Landsat - 8 satellite, OLI sensor point/orbit 221/65 and 221/66, bands 2, 3 and 4 (visible) are used; 5 (near infrared); 6 and 7 (medium infrared); from August 22nd, 2020, percentage of clouds below 5% per scene and with up to 1% in the area of interest, freely available in the American Geological Survey-USGS.

2.1 Methodological Procedures

The methodology was based in techniques of digital application treatment of images as described by Florenzano (2011) which corresponds to: i) atmospheric correction; ii) merge; iii) contrast enhancement, iv) combined images; v) extraction of attributes, vi) collection of samples, vii) classification and validation.

2.1.1 Atmospheric correction

It was used the Dark Object Subtraction - DOS (Chaves Jr et al. 1988), to reduce the effect of atmospheric scattering, using only parameters related to the digital number of the image. At DOS it is assumed the existence of pixels with null values, such as shadows caused by topography or clouds, which present values higher than expected, because of the effect of atmospheric scattering, and that reference for the correction of atmospheric scattering from the subtraction of these values by the image, and the results obtained in the atmospheric attenuation process in different targets, mainly in the blue and green bands, as defined by Maia (2017).

2.1.2 Contrast enhancement

This process aims to reduce the effects of pixel grouping in delaminated regions of the histogram, which cause low variance and generate low contrast problems between the present features and has as objective to improve the visual information present in the images but changes the gray scale and changes the pixel grouping limits within the histogram (Meneses and Almeida, 2012). It was applied linear enhancement filter was applied in which it offers a distribution of the gray level values according to a linear function of 1st following equation 1.

$$Y = (x - \text{minor}) * \left(\frac{2^n - 1}{\text{major} - \text{minor}} \right) \quad (1)$$

Where, The Y value represents the pixel value in the new histogram.

x corresponds to the gray level value of the original image.

n represents a radiometric resolution of the sensor.

larger and smaller are the original limits of the values in the histogram.

2.1.3 Segmentation

The segmentation of images constitutes a process of subdivision into discrete regions, from the grouping of pixels that have similar spectral and spatial characteristics. It was used the technique for growing regions (Baat, Shape, 2000) that starts from a "seed" pixel and groups pixels with similar characteristics, defined using similarity criteria that correspond to the Euclidean distance of the pixel values that will compose each segment, and area that consists of the minimum size that each segment presents. For this test, 0.30 similarity and an area of 10 pixels were established, which corresponds to a minimum 90 hectares for the composition of a segment.

2.1.4 Attribute extraction

Extracting attributes from Remote Sensing images is an important step in the classification process, and its function is to identify aspects about the structure arrangement of surfaces and their relations with their neighbors, because they show the differences and similarities between the segments or objects, which go beyond the values of digital numbers. It was extracted of 26 attributes for each image band, being 13 spatial attributes and 5 spectral and 8 textural, which generated 156 attributes. Aiming to minimize the effects of high existent correlation among generated attributes, it was used the technique of Main Analysis Components, in order to reduce the total level of information which generated a total of 15 main non correlated components.

2.1.5 Classification

It was evaluated four supervised algorithms of classification, being two of the type of decision tree: Random Forest (Breiman, 2001) C5.0 (Quinlan, 1992); and two of traditional approaching: Minimal Distance and Spectral Angle Mapper. The choose of algorithms was associated to the availability and usability, bearing in mind that these ones already are implemented in packages of algorithms of a diversity of geographical information. The C5. is a nonparametric binary classifier of decision tree, considered simple and intuitive (Maxwell et al. 2018), which has as principle of classification a sequential classification of attributes of samples using data of gaining to determinate the best attribute to define classes, forming only a decision tree, that will be the base to the classification (Quinlan, 1992).

Random Forest –RF is part from a set of algorithms of classifier per conjunct (Li et al. 2014) uses as intelligent strategies of classification considering different samples randomly selected in order to train various distinct decision trees which combined, supply elevated accuracy (Zhang, Yang, 2020; Melville et al. 2018). The MD corresponds to a supervised parametric classifier which uses the Euclidian distance to associate a segment to a class, considering only the medium of the attributes of the used samples, which may present elevated error rate in environments of elevated homogeneity. Whereas the SAM is a supervised algorithm which seeks to analyze spectral similarities of image that must be classified to a one spectral library of reference created from the images, and in areas with similar spectrally aims may cause significant reduction of matching (Lenzi, Nunes, 2016).

2.1.6 Collection of samples

To train the classification algorithms used in this test, 801 samples were used, coming from a set of 1000 points created randomly using the radom points tool present in TerraView 5.5.1 and distributed at the municipal limit. All points belonging to the same segment were excluded, thus seeking to avoid repetition of the data. The identification of the class to which each sample belongs was subsidized by the image Landsat false composition color RGB-453 (Infrared, medium infrared and red), as shown in Sano et al. (2009). The randomness of the sample data caused a difference in the total of samples per class, with 163 samples being used for forest formations; 179 samples for country formations; 399 samples for temporary cultures; and 60 samples for hydraulic bodies, this difference in the sample quantity by class was caused by the spatial distribution that each class presented in the area.

The choice of the class that each sample would represent was guided by an interpretation key adapted from Sano et al. (2009), according to criteria of spatial and phytophysiognomics variability, these being: i) Forest formations presented a reddish color, varying from dark red to medium red, a heterogeneous pattern with a rough texture and without defined shape. It represents dense to medium sized vegetation areas corresponding to riparian forest in the area's river channels; ii) Country formations showed coloration ranging from green to bluish green and brown, heterogeneous pattern and rough texture without a defined shape, and may also present reddish spot pigmentation, composed of low-density grasses and shrubs, mixing small vegetation; iii) Temporary crops have homogeneous patterns, well-defined regular shape, and smooth texture, presented coloration ranging from blue to brown, and may also have reddish, magenta, or white tones, depending on the vegetative stage of the culture; iv) Water bodies it presented a color ranging from bluish, greenish to black, with a sinuous shape (channels) or not, smooth texture and very uniform pattern.

2.1.7 Evaluation of performance

The accuracy of the classifications and the level of performance of the classifiers were analyzed from the elaboration of the confusion matrix (Congalton, Green, 1991), which corresponds to a classic and binary model of performance evaluation in validating the classification of satellite images (Andrade et al. 2014). 891 points were used chosen from a universe of 1000 points created at random and spatially distributed to cover more than 85% of the area, 212 points for forest formations, 258 points for rural formations, 324 points for temporary crops and 38 points for the water bodies. As a methodological criterion, all points belonging to the same segment were excluded, thus avoiding repetition of the data, and influencing the result.

The identification of the classes in which the validation points represent was performed manually based on a set of CBERS 4A images (INPE, 2020) that has a spatial resolution of 8 meters, these were subjected to a fusion process to present a spatial resolution of 2 meters using the pan sharpening image fusion algorithm. The validation samples were cross tabulated with the results of the classifications and supported the creation of the confusion matrix, the correctly classified data were classified positioned on the main diagonal, and the incorrectly classified data were inserted at the top and bottom of the matrix and made it possible to evaluate the kappa performance indices (2), Global Accuracy (3), producer precision (4) and user (5).

$$k = \frac{n \sum_{i=1}^r x_{ii} - \sum_{i=1}^r x_{i+} x_{+i}}{n^2 - \sum_{i=1}^r (x_{i+} * x_{+i})} \quad (2)$$

$$E_g = \left(\frac{\sum_{i=1}^r x_{ii}}{n} \right) \quad (3)$$

$$K_p = \left(\frac{n * n_{ii} - n_{i+} * n_{+i}}{n * n_{i+} - n_{+i} * n_{+i}} \right) \quad (4)$$

$$K_u = \left(\frac{n * n_{ii} - n_{i+} * n_{+i}}{n * n_{i+} - n_{+i} * n_{+i}} \right) \quad (5)$$

Where n is the total number of used samples
 k is the value of kappa index
 E_g is the global index of classification
 K_p represents the accuracy of producer by class
 K_u is the accuracy of user per class
 $\sum_{i=1}^r x_{ii}$ is the sum of main diagonal
 $\sum_{i=1}^r (x_i x_j)$ represents the product of the sum of the line by the column of each representative class
 n_{ii} is the total number of samples correctly classified of the class
 n_{i+} is the total of classified samples of the class
 n_{+i} is the total number of collected samples of the class k

The global accuracy indicates the percentage of successes relating the total of samples correctly classified and with total number of used samples (Mao et al. 2020; Zhang, Yang, 2020). Foody (2002) highlights that a classification must reach elevated success rate, with indexes superior to 85% to be considered acceptable, indicating that smaller rates indicate levels of confusion by producer and user inconsistent. Whereas the Kappa coefficient shows the statistic differences between the classifications and the map of reference. According to Silva Júnior et al. (2014), the kappa index presents vantages about the global accuracy due to incorporate all the elements of the confusion matrix, being sensible to the variations of errors of consumer and producer. The results were compared to the values of performance established by Landis and Koch (1997), in which attributes qualitative features to the levels of classification

performance, that indicate the quality of the thematic map to the obtained kappa value, being: Very bad (<0.000); Bad (0.00 – 0.200); Regular (0.201-0.400); Good (0.401-0.600); Very Good (0.601-0.800); and, Excellent (>0.801).

It was analyzed if the performance of the classification generated in relation to the reference data are significant, using the Z test (6) with a confidence rate of 95%, as shown by Congalton, Green (1991).

$$Z = \frac{k}{\sqrt{\sigma_{(k)}^2}} \quad (6)$$

Where k is correspondent to the value of the kappa index generated for the classification
 $\sigma_{(k)}^2$ is the sampling of the generated kappa coefficient

The $\sigma_{(k)}^2$ is calculated following the equation 7 as described by Amaral et al. (2009).

$$\sigma^2 = \frac{1}{n} \left[\frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2-\theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4-4\theta_2^2)}{(1-\theta_2)^4} \right] \quad (7)$$

Where:

$$\theta_1 = \frac{\sum_{i=1}^r x_{ii}}{X^2}$$

$$\theta_3 = \frac{\sum_{i=1}^r x_{i+1} x_{i+1}}{X^2}$$

$$\theta_3 = \frac{\sum_{i=1}^r x_{ii} (x_{i+} + x_{+i})}{X^2}$$

$$\theta_4 = \frac{\sum_{i=0}^r \sum_{i=1}^r x_{ii} (x_{i+} + x_{+i})^2}{X^3}$$

According to Landis and Kock (1977) IF $Z \geq Z_{\alpha/2}$ the classification is significantly better than a random distribution where $\alpha/2$ is the level of confidence of 1,96 in the two sides of the curve. In the Z test is the number of liberty grades is assumed to be infinite.

In order to evaluate if there are statistical differences of obtained the kappa indexes, was used the Z test (Congalton, Green 1991), that evaluate the level of static significance between the performance of the classifications (8).

$$Z = \frac{K_2 - K_1}{\sqrt{\sigma_{K_2}^2 + \sigma_{K_1}^2}} \quad (8)$$

Where K_1 is the value of kappa of the classification 1
 K_2 represents the value of the index kappa
 σ^2 is the variance of the kappa index referent to the clarifications 1 and 2

3. RESULTS AND DISCUSSION

3.1 Visual Analysis and Patterns of Land Covering

The classifications generated by the used algorithms (figure 2), showed differences visually perceptible in the spatial pattern of distribution of classes. The resulting classifications of the algorithm's RF and C5.0 present representations visually near to the covering of the real land present in the area. Whereas the resulting classifications of the algorithms SAM and MD showed visual confusions, with elevated overestimation of the classes forest covering and hydrous bodies, which is also evidenced when patterns of distribution are observed by class, as shown in table 1.

	Area in Km ²			
	RF	C5.0	MD	SAM
FF	1240.06	1067.21	2111.73	1163.35
FC	1857.9	1868.47	1386.11	1781.27
CT	1242.43	1362.46	829.81	845.78
CH	46.52	82.4	55.31	590.31

Table 1. Spatial distribution of use and land covering according to the tested algorithms.

The country formations presented patterns of spatial distribution relatively near in the classifiers C5.0, RF and SAM, covering approximately 42%, 42% and 40% of the municipality area. This class presented significant reduction when is observed the resulting classification of the MD algorithm where it covered only 31% of the area. Similar pattern is noted with the forest formations, SAM, C5.0 and RF that cover respectively 26%, 24% and 28%, whilst the classification by MD represents 18% of the entire area. The temporary crops represent 28% and 31% in the classifications by RF and C5.0. Whereas the classifications SAM and MD represented only 18% and 19% of the total covering.

Whereas the hideous bodies presented constant area to the classifiers RF, C5.0 and MD were corresponded to little more than 1% of the area. Only the classifier SAM presented overestimation with covering rates superior to 13% of the area, indicating elevated rates of mistakes by the producer and consumer mainly for the algorithms SAM and MD.

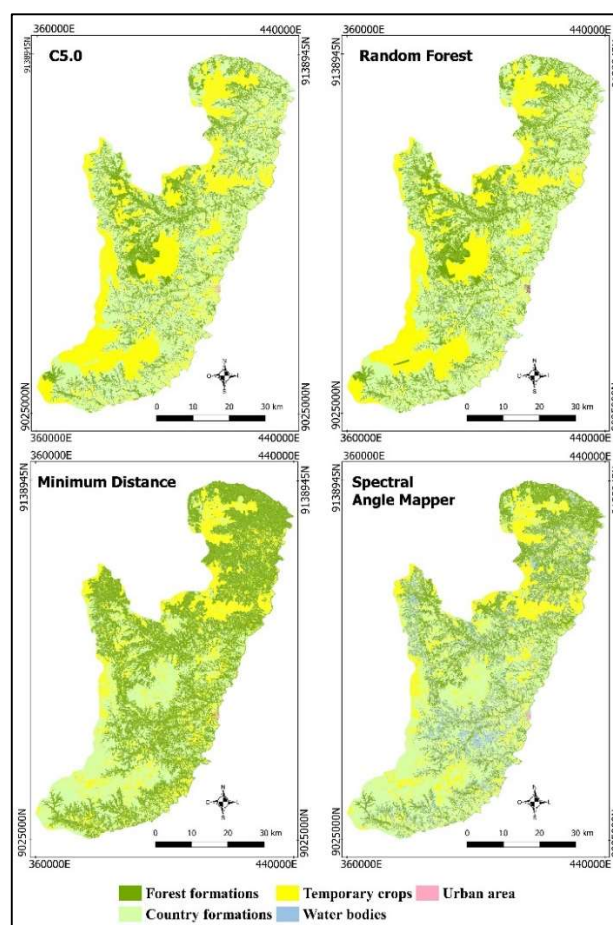


Figure 2. Use and land covering generated from the algorithms of classification algorithms C5.0, RF, SAM and MD.

3.2 Analysis of Performance

All the classifications presented acceptable values of kappa varying between 0.50 and 0.870, and the global accuracy presented fluctuations from 66 to 90% (table 2). The tests of hypothesis in function of the results of the kappa indexes and kappa variance (table 2), showed that all the classifiers are significantly bigger than zero, with a level of confidence of 95%. The algorithm RF is significantly better than the other ones (59.56), followed by C5.0 (49.3780), this result corroborates with the values of kappa index and smaller variance, which also showed the superiority of these algorithms before the others.

	Kappa	Global Accuracy	Var. Kappa	Quality
RF	0.8617	90.68	0.000209272	Excellent
C5.0	0.7832	85.52	0.000304507	Very good
SAM	0.6276	74.64	0.000455307	Very good
MD	0.5057	66.33	0.000470549	Good

Table 2. Performance of the algorithms of classification.
Source: data from research, 2020.

The Z test also showed that all classifications have statistical differences, with the biggest differences between the RF and MD classifiers, which showed the highest level of significance. Among RF over C5.0, there is a slight superiority (table 3).

Classifier	Z	p-value
RFxC5.0	3.4629	0.0002672*
RFxDM	13.6537	0.0000000*
RFxSAM	9.0774	0.0000000*
C5.0xDM	9.9679	0.0000000*
C5.0xSAM	5.6419	0.0000000*
DMxSAM	4.0091	0.0000305*

Table 3. Test de significance of the kappa coefficient, Significant to 0.05 of probability, in other words, if H0: Kappa₁-Kappa₂ = 0; H1: Kappa₁ - Kappa₂ < 0. Source: data from research, 2020.

The RF classifier presented the best rates for both global accuracy (90.68%); and kappa index (0.861) and lower statistical variance, evidencing also when analyzing the levels of statistical significance presented by Z Test. Similar results were found by, Li et al. (2014), when supervised and unsupervised algorithms were compared with Landsat 5 and 8 data, where Random Forest presented the best performance parameters, with emphasis on the segmentation addressing, with an accuracy greater than 0.800. Ge et al. (2020) when comparing only algorithms for machine learning and deep learning also showed a high consistency of classification from the use of RF with a rate of 96.20% accuracy with a rate lower than neural networks.

According to Andrade et al. (2014) the superiority of the classifiers by machine learning, as shown by Kappa and E.G, is directly associated with the model of class identification and separation. Whilst the MD and SAM, define fixed rules for class differentiation, requiring sample data that have relatively homogeneous information with normal distribution of attributes (Zanotta et al. 2019); C5.0 and RF do not depend on normal distribution, focusing on the structure of the training data it is assumed pre-established hypotheses. Even though with acceptable kappa index and global accuracy, classifications showed high rates of confusion when were analyzed conditional kappa values, producer, and user accuracy, as shown in the confusion matrices in tables 4, 5, 6, 7.

		Reference				Σ
		FF	FC	CT	CH	
Classification	FF	217	5			222
	FC	3	191	3		197
	CT		72	359		431
	CH		0	0	41	41
Σ		220	268	362	41	

Table 4. Matrix of confusion of the algorithm Random Forest.

		Reference				Σ
		FF	FC	CT	CH	
Classification	FF	210	4			214
	FC	9	149			158
	CT		115	362		477
	CH	1			41	42
Σ		220	268	362	41	

Table 5. Matrix of confusion of the C5.0 algorithm.

		Reference				Σ
		FF	FC	CT	CH	
Classification	FF	203	118	1	2	324
	FC	16	42	54		112
	CT		108	307		415
	CH	1		39		40
Σ		220	268	362	41	

Table 6. Matrix of confusion of the MD algorithm.

		Reference				Σ
		FF	FC	CT	CH	
Classification	FF	198	29			227
	FC	17	119	55		191
	CT		106	307		413
	CH	5	14	0	41	60
Σ		220	268	362	41	

Table 7. Matrix of confusion of the SAM algorithm.

When is analyzed the errors of commission and omission (table 8 and 9), it is noted that, to a greater or lesser extent, all the algorithms showed relatively high performance for three of the four classes analyzed with rates above 85% of accuracy. As with the Kappa indices and global accuracy, the RF and C5.0 classifications achieved better results with lower error rates for both producer and user (consumer).

Class	Classifiers			
	RF	C5.0	MD	SAM
FF	0.99	0.95	0.92	0.90
FC	0.71	0.56	0.16	0.44
CT	0.99	1.00	0.85	0.85
CH	1.00	1.00	0.95	1.00

Table 8. Level of accuracy of the producer.

Class	Classifiers			
	RF	C5.0	MD	SAM
FF	0.98	0.98	0.63	0.87
FC	0.97	0.94	0.38	0.62
CT	0.83	0.76	0.74	0.74
CH	1.00	0.98	0.98	0.68

Table 9. Level of accuracy of the producer.

The dissipated hydrous bodies presented better performance in the RF, C5.0 and MD classifications. Only one classification by MD classifier had a production error of 5% from the producer and 2% from the user. This rate is not observed for a classification by SAM with a user (consumer) error greater than 30%, which indicates a high overestimation of this class.

The accuracy of producer in Forest formations was greater than 90% in all classifications, however this parameter is not observed when is analyzed the user (consumer) accuracy, for the classification by MD with user error of more than 10%, which indicates even overestimation 5%.

The Temporary crops obtained a high producer accuracy in the classifications by C5.0 with 100% accuracy, however the user error was 24%, indicating that of the classified areas for this class, only 76% correspond to reality, which indicates overestimation of this class for this classifier. In the classifications by MD and SAM, the temporary cultures presented a low performance for both producer and user, which indicate a high underestimation.

When it is analyzed the inconsistencies per class, the lowest hit rate was for country formations, with an omission error of 85%, 56% and 44% in the classifications by MD; SAM and C5.0 respectively, while the user (consumer) error showed an error of 63%, 38 and 6%. Only the RF classification showed acceptable accuracy rates for this class with a producer and user above 95%.

When the levels of confusion were approached among the classes, it is noted that the biggest inconsistencies were for the rural formations, which presented a high error rate with the temporary cultures, varying between 26 and 42%. With highlight to the classifier's SAM, MD and C5.0.

This factor is associated with the high spectral proximity that these classes present. The rural formations of the Cerrado Biome are characterized by a vegetation cover space usually composed of grasses and shrubs, which do not fully cover the soil, providing high reflectance, patterns close to areas destined for cultivation.

4. CONCLUSION

Bearing in mind that the objective of this study was to analyze the performance between the classification algorithms, for the mapping of land cover and use, it can be affirmed that the kappa measures had quite similar behavior, which is not evident when observing global accuracy measures.

The RF algorithm showed better performance when was compared to the algorithms C5.0, MD and SAM, with a value considered excellent. C5.0 and MD were classified as very good, and SAM qualified as good. However, there was a significant inconsistency in the classification, which indicates an improvement in the classification parameters. However, high visual confusion was observed in three of the four classifiers analyzed, given that this was not evidenced by the global accuracy and kappa indexes, indicating a need to improve the classification parameters.

The Z test showed that all classifiers presented strong significance in relation to the reference samples, with significant superiority of the RF when compared with the classifications resulting from the other tested algorithms. The superiority of the RF algorithm also evidenced by the kappa values and global accuracy.

The MD and SAM classifications presented the greatest inconsistency both visually and by the indices tested, with a significant overestimation of the classes rural formations, forest formations and water bodies and underestimation of the temporary cultures classes, with a rate higher than 30%, this data can be explained by the high spectral proximity between classes, in addition to this associated with similarity parameters used to differentiate classes at the time of classification.

Despite showing the high potential of the classifiers for machine learning such as Random Forest and C5.0, which are generally available in open access geographic information systems such as the Orfeo toolbox and Geodma, it was evidenced in this study, the need for improvement of training and validation sampling parameters when used in medium resolution images, such as Landsat data, aiming that, despite presenting an accuracy rate greater than 85%, significant confusion was observed between the classes. This data shows the need for an improvement in the classification adjustments of the tested algorithms, also considering the representative attributes of the classes.

ACKNOWLEDGEMENTS

This study was partially financed by the Coordination for the Improvement of Higher Education Personnel (CAPES - Finance Code 001) and by the National Council for Scientific and Technological Development (CNPq - Research Productivity Scholarship).

REFERENCES

- Andrade, A. C., Francisco, C. N., Almeida, C. M. 2014. Desempenho de classificadores paramétrico e não paramétrico na classificação da fisionomia vegetal. *Revista Brasileira de Cartografia*, 66(2), 349-363.
- Azzari, G., Lobell, D. B., 2017. Landsat-based classification in the cloud: na opportunity for a paradigm shift in land cover monitoring. *Remote Sensing of Environment*, 202(1), 64-74. doi.org/10.1016/j.rse.2017.05.025
- Baatz, M.; Schäpe, A. 2000. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. *Journal of Photogrammetry and Remote Sensing*, 58(1), 12-23. doi:10.12691/jgg-6-3-2
- Breiman, L. 2001. Random Forests. Machine Learning. Springer. 45(1), 5-32.
- Congalton, R., Green. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 49(12), 1671-1678. doi.org/10.1016/0034-4257(91)90048-B
- Chavez JR., P.S. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24, 459-479, 1988. doi.org/10.1016/0034-4257(88)90019-3
- Crowson, M., Hagensierker, R.; Waske, B. 2019. Mapping land cover change in northern Brazil with limited training data. *Inter. Journal of Applied Earth Observation and Geoinformation*. 78, 202-214. doi.org/10.1016/j.jag.2018.10.004

- Deng, Z., Zhu, X., He, Q., Tang, L. 2019. Land use/land cover classification using time series Landsat 8 images in a heavily urbanized area. *ScienceDirect Advances in Space Research*. 63, 2144–2154. doi.org/10.1016/j.asr.2018.12.005
- 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*, Ottawa. 118(1), 259-272. doi:10.1016/j.rse.2011.11.020
- Foody, G. M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1), 185-201. doi.org/10.1016/S0034-4257(01)00295-4
- Florenzano, T. G. 2007. Iniciação em Sensoriamento Remoto. São Paulo: Oficina de Textos.
- IBGE. Brazilian Institute of Geography and Statistics. 2017. <https://cidades.ibge.gov.br/brasil/ma/tasso-fragoso/panorama> access in: 23/01/2021
- IBGE. Brazilian Institute of Geography and Statistics. 2019. Municipal agricultural production. <https://sidra.ibge.gov.br/pesquisa/pam/tabelas>. Access in: 24/01/2021
- Jin, S., Yang, L., Xhu, Z. Homer, C. A., 2017, A land cover change detection classification protocol for upadating Alaska NLCD 2001 to 2011. *Remote Sensing of Environment*, 195(1), 44-55. doi.org/10.1016/j.rse.2017.04.021
- Landis, J.R. e Kock, G.G., 1977 The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. doi.org/10.2307/2529310
- Lenzi, I. L. C., Nunes, G. M. 2016. Comparação entre os classificadores Support Vector Machine e Spectral Angle Mapper aplicado à diferenciação das fitofisionomias do parque estadual do Araguaia (MT) *Revista Eletrônica Geoaraguaia. Barra do Garças-MT*. 6(1), 99-114. doi:10.13140/RG.2.1.1915.8003
- Li, C.; Wang, J.; Wang, L.; Hu, L.; Gong, P., 2014. Comparison of classification algorithms and training sample sizes in urban land classification with landsat thematic mapper imagery. *Remote Sensing*. 6(2), 964-983. doi.org/10.3390/rs6020964
- Li, X., Chen, W., Cheng, X., Wang, L., 2016, A Comparison of Machine Learning Algorithms for Mapping of Complex Surface-Mined and Agricultural Landscapes Using ZiYuan-3 Stereo Satellite Imagery. *Remote Sensing*. 16(8), 514. doi:10.3390/rs8060514
- Lu, D.; Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*. 28(5), 823-870. doi: 10.1080/01431160600746456
- Maia, M. A., Rodrigues, N. B., Richter, M. Rubim, I. B. 2017. Modelos de correção atmosférica aplicados em imagens do sensor OLI/ Landsat 8 a partir do uso de programas gratuitos: uma análise comparativa. SBSR, XVIII Simpósio Brasileiro de Sensoriamento Remoto. Santos.
- Mastella, A. F. M., Vieira. 2018. Acurácia temática para classificação de imagens utilizando abordagens por pixel e por objetos. *Rev. Bras. de Cartografia*, 70 (1), 1618-1643. doi: 10.14393/rbcv70n5-44559
- Maxwell, A. E.; Warner, T. A.; Fang, F. 2019. Implementation of machine-learning classification in remote sensing: an applied review. *International Journal of Remote Sensing*, 39(9), 2784-2817. doi.org/10.1080/01431161.2018.1433343
- Melville, B., Lucieer, A., Aryal, J. 2018. Object-based random forest classification of Landsat ETM+ and WorldView-2 satellite imagery for mapping lowland native grassland communities in Tasmania, Australia. *Int J Appl Earth Obs Geoinformation*. 66(1), 46-55. doi.org/10.1016/j.jag.2017.11.006
- Mishra, P. K., Rai, A., Rai, S.C., 2019. Land use and land cover change detection using geospatial techniques in the Sikkim Himalaya, India. *The Egyptian Journal of Remote Sensing and Space Sciences*. 23(2), 133-143. doi.org/10.1016/j.ejrs.2019.02.001
- Noi, P. T.; Kappas, M., 2018. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors*, 18(1), 18. doi:10.3390/s18010018
- Phriri, D., Morgenroth, J., Xu, C., Hermosilla, T., 2018. Effects of pre-processing methods on Landsat OLI-8 land cover classification using OBIA and random forests classifier. *Int J Appl Earth Obs Geoinformation*. 73(1), 170-178. doi.org/10.1016/j.jag.2018.06.014
- Quinlan, J. R. 1992. C4.5: programs for machine learning. London: *Morgan Kaufmann Publishers*. v. 2021
- Sánchez-Espinosa, A., Schröder, C., 2019. Land use and land cover mapping in wetlands one step closer to the ground: Sentinel-2 versus Landsat 8. *Journal of Environmental Management*. 247(1), 484–498. doi.org/10.1016/j.jenvman.2019.06.084
- Sano, E. E.; Rosa, R.; Brito, J. L. S.; Ferreira, L. G., Bezerra, H. S. 2009. Mapeamento da cobertura vegetal natural e antrópica do bioma Cerrado por meio de imagens Landsat ETM+-. *SBSR, Simpósio Brasileiro de Sensoriamento Remoto*, Natal: Instituto Nacional de Pesquisas Espaciais, 1199-1206
- Silva Junior, C. A., Frank, T., Rodrigues, T. C. S. Discriminação de áreas de soja por meio de imagens EVI/MODIS e análise baseada em geo-objeto. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 18(1), 44-5. doi.org/10.1590/S1415-43662014000100007
- Zhang, F. Yang, X. 2020. Improving land cover classification in an urbanized coastal area by Random forests: The role of variable selection. *Remote Sensing of Environment* 251(1), 112105. doi.org/10.1016/j.rse.2020.112105