ARTIFICIAL INTELLIGENCE AND ORBITAL IMAGES APPLICATION FOR ANALYSIS OF SPATIAL LAND USE AND COVERAGE PATTERNS

Roberta Aparecida Fantinel[1*], Rudiney Soares Pereira[2], Ana Caroline Paim Benedetti[3], Fernando Coelho Eugenio[4], Juliana Marchesan[5], Mateus Sabadi Schuh[1]

[1*]Universidade Federal de Santa Maria, Programa de Pós-Graduação em Engenharia Florestal, Santa Maria, Rio Grande do Sul, Brasil - *fantinel.ar@gmail.com
[2]Universidade Federal de Santa Maria, Departamento de Engenharia Rural, Santa Maria, Rio Grande do Sul, Brasil - rudiney.spereira@gmail.com
[3]Universidade Federal de Santa Maria, Departamento de Ensino, Santa Maria, Rio Grande do Sul, Brasil - nacaroline@politecnico.ufsm.br
[4]Universidade Federal de Santa Maria, Departamento de Engenharia Agrícola, Cachoeira do Sul, Rio Grande do Sul, Brasil - fernando.eugenio@ufsm.br
[5]Universidade Federal de Santa Maria, Doutora em Engenharia Florestal, Santa Maria, Rio Grande do Sul, Brasil - marchesan.ju@gmail.com
[1]Universidade Federal de Santa Maria, Programa de Pós-Graduação em Engenharia Florestal, Santa Maria, Rio Grande do Sul, Brasil - mateuschuh@gmail.com

Received for publication: 10/02/2021 – Accepted for publication: 07/02/2022

Resumo

Uso da inteligência artificial e imagens orbitais para análise de padrões espaciais de uso e cobertura da terra. O objetivo do estudo foi analisar o desempenho de diferentes algoritmos de aprendizado de máquina (AM) na predição dos padrões do uso e cobertura da terra a partir de dados espectrais de séries temporais dos sensores *Thematic Mapper* (TM) e *Operational Land Imager* (OLI). Foi utilizado o software QGIS, onde iniciou-se a importação das imagens TM/Landsat 5 nos anos de 2004 e 2009 e OLI/Landsat 8 para os anos de 2015 e 2019, a fim de obter informações para caracterizar e diferenciar os padrões de uso e cobertura da terra. Posteriormente realizou-se o treinamento e teste dos algoritmos, *Random Forest* (RF), *Support Vector Machine* (SVM), *K-Nearest Neighbors* (K-NN) e *Naive Bayes* (NB), nas proporções de 80%-20%, 70%-30%, 60%-40% no software KNIME. O desempenho foi analisado com base na acurácia global e no índice de Kappa. O RF e o SVM para os anos de 2004 e 2009 apresentaram o melhor desempenho (acurácia global), enquanto que para os anos de 2015 e 2019, foram o K-NN e a RF. Os valores do índice Kappa indicaram que as classificações dos algoritmos variaram de 0.80 – 1.00. A proporção de 60% (treinamento) e 40% (teste) foi a que proporcionou os melhores resultados para todas as datas analisadas. Os dados provenientes dos pixels amostrado dos padrões do uso e cobertura da terra das imagens dos sensores TM e OLI demonstraram ser eficientes para o processo de ML no software KNIME.
Palavras-chave: Aprendizado de máquina, *Software* livre, Landsat, Sensoriamento Remoto.

Abstract

The study aimed to analyze the performance of different machine learning (ML) algorithms in predicting land use and land cover patterns from time series spectral data from Thematic Mapper (TM) and Operational Land Imager (OLI) sensors. The QGIS software was used, where the import of TM / Landsat 5 images began in 2004 and 2009 and OLI/Landsat 8 for 2015 and 2019, to obtain information to characterize and differentiate usage patterns and land cover. Subsequently, training and testing of the algorithms, Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (K-NN), and Naive Bayes (NB), were carried out in the proportions of 80%-20%, 70%-30%, 60%-40% in the KNIME software. The performance was analyzed based on global accuracy and the Kappa index. The RF and SVM for the years 2004 and 2009 showed the best performance (global accuracy), while for the years 2015 and 2019, they were the K-NN and the RF. The Kappa index values indicated that the classifications of the algorithms varied from 0.80 – 1.00. The proportion of 60% (training) and 40% (test) was the one that provided the best results for all the dates analyzed. The data from the pixels sampled from the land use and land cover patterns of the TM and OLI sensor images proved to be efficient for the ML process in the KNIME software.
Keywords: Machine learning, Landsat, Free software, Remote Sensing.

INTRODUCTION

Changes in the environment, such as agricultural expansion and deforestation, directly impact ecosystem processes, thus, it becomes crucial to map changes in land use and land cover to monitor and manage these changes (ABDI, 2020; BELMAKER *et al*., 2015). Thus, temporal and spatial changes in land use and land cover have been the focus of attention in several studies over the years, thus, decision support systems, in a rapid process of evolution, have allowed the use of technologies for active form in the management and planning of the territory (BUĞDAY and ERKAN BUĞDAY, 2019). Among these technologies, there is the Remote Sensing (RS) where significant improvements in the spatial, spectral and temporal resolution of the data, has allowed more accurate results. In addition, the development of artificial intelligence (AI) techniques and the use of machine learning algorithms (ML) in the analysis of RS data have been gaining more prominence in the scientific community. This evolution has increased the availability of data, as well as a huge source of information that allows us to have more comprehensive images of patterns of land use and land cover. In recent years, supervised and unsupervised ML algorithms have drawn attention to several applications in RS, including land use and land cover mapping (ABDI, 2020).

Andrade *et al*. (2014) emphasizes that in the supervised method, the user defines the classes and the respective samples that will make up the training. The unsupervised, for Ge et al. (2020) the algorithm uses statistics and data grouping techniques, not having the need for prior knowledge of the classes to be created thus, algorithms using supervised techniques have been used to improve the classification of land use and land cover patterns, such as Support Vector Machines (SVM), Random Forest (RF), Naive Bayes (NB), and the K-Nearest Neighbor (K-NN). As reported in the studies by Li *et al*. (2014) and Zhu *et al*. (2012), these algorithms have demonstrated excellent performance in the analysis of RS data.The advancement of software, open source languages and modern algorithms, according to Bunting *et al*. (2014) make it possible to study and understand the patterns of land cover in the most diverse scales of time and space.

To use these techniques, it is necessary to have a data processing environment that operates simply and intuitively, citing as an example, the Konstanz Information Miner (KNIME) platform, which allows quick and interactive changes in the analysis and interpretation of results. KNIME is a data mining tool developed in 2004, by the University of Konstanz, Germany, being launched in 2006. It is a platform for predictive analysis of free software, open-source, having the General Public License (GLP) (KNIME, 2019). The workflow creation and execution flow in KNIME are based on the idea of adding nodes (nodes), where they are linked through their inputs and outputs. According to Feltrin (2015), each node performs processing based on a data algorithm, which is capable of interacting with other nodes (when connected), thus allowing the generation and recording of complex data processing workflows. Thus, it can be said that research focused on the use of KNIME linked to the Remote Sensing area, are still little explored and, especially, their studies are little disseminated, from the theoretical, methodological, and practical point of view, in high impact on the scientific environment in the area of agrarian sciences.

The justification for the use of KNIME was given by the ample support material for users, by the capacity of integration in different programming languages in the same workflow environment, and mainly by the scarcity, due to the lack of studies focused on the forest area until the present time, using data from remote sensor time series. In this context, the work aimed to analyze the ability of machine learning algorithms RF, SVM, NB and K-NN to predict land use and land cover patterns using spectral time series data from Thematic Mapper (TM) sensors from Landsat 5 and Operational Land Imager (OLI) of Landsat 8, through open source software, in order to validate its use in a Geographic Information Systems environment

MATERIAL AND METHODS

Characterization of the study area

The area chosen for the study is the municipality of Dona Francisca, which belongs to the region of the Quarta Colônia and is located in the Central Depression of the State of Rio Grande do Sul, which is 210 km away from the capital Porto Alegre. It has 29º 36’41” south latitude and 53º 21’ 03” west longitude as reference coordinates (Figure 1). The territorial area of the municipality according to the Brazilian Institute of Geography and Statistics (IBGE, 2010) is 114.35 km².

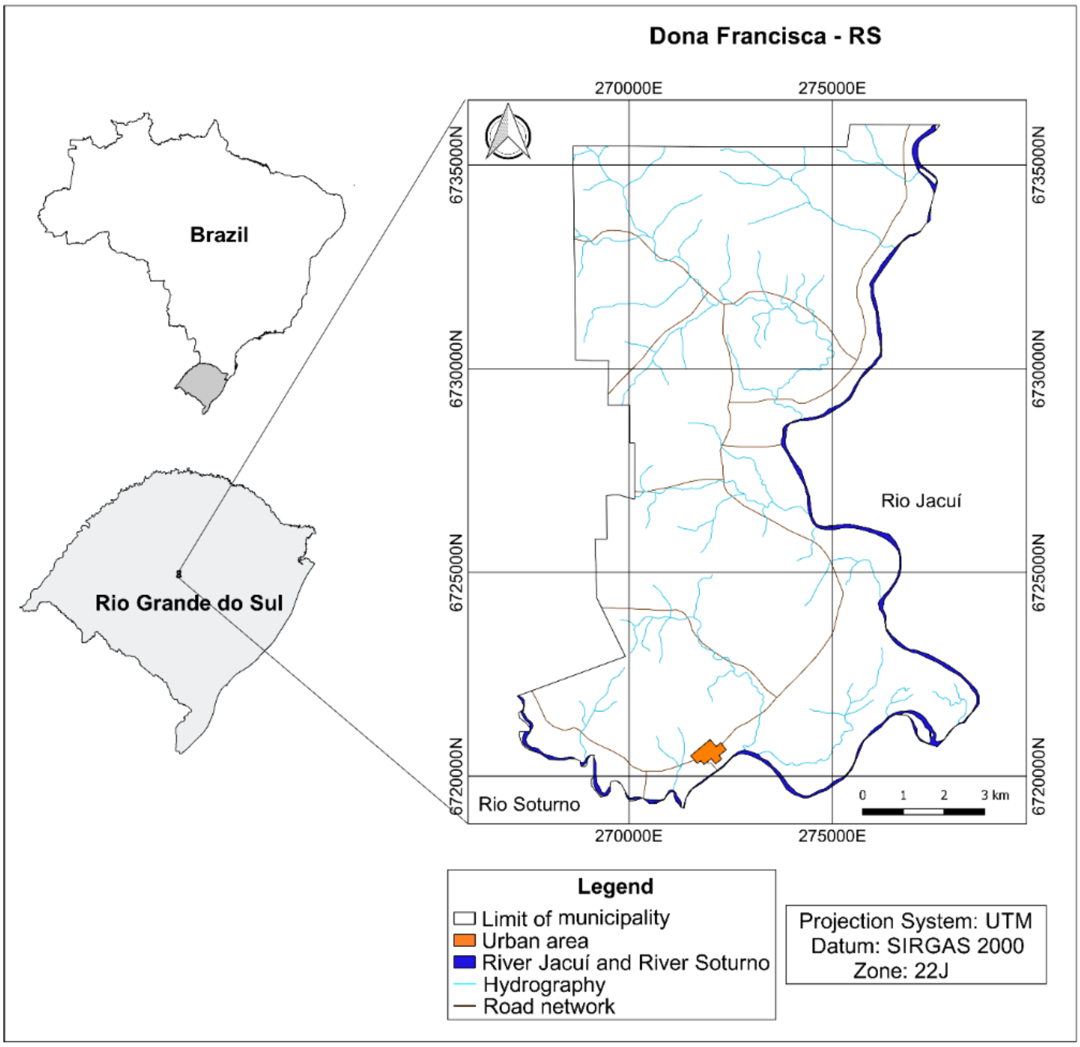


Figure 1. Location of the study area, municipality of Dona Francisca – RS.
Figura 1. Localização da área de estudo, município de Dona Francisca – RS.

According to the Köppen classification, the municipality of Dona Francisca has a humid subtropical climate - Cfa, where the hottest months are from December to February with an average temperature of 22ºC, while the months June, July, and August are the colder, with an average temperature of -3 ° C to 18 ° C. The average annual precipitation is approximately 1.722 mm (ALVARES *et al.*, 2013).

The relief of the municipality is characterized by plain alluvial-colluvial, Depression of the Jacuí River, Serra Geral, and the Plateau of Campos Gerais (IBGE, 2019). The rugged area is part of the Serra Geral formed by successive basalt spills, with Morro Santo Antônio and Morro Formoso present (LAGO and FARENZENA, 2008). The municipality of Dona Francisca belongs to the Guaíba hydrographic region and the Alto Jacuí hydrographic basin, with the Jacuí and Soturno rivers as main watercourses, which are used for water supply, energy supply, fishing, and irrigation.

Data processing

Figure 2 shows this work proposed methodology flowchart. In the QGIS software, spectral data of TM/Landsat 5 and OLI/Landsat 8 images are taken, considering different patterns of land-use and land-cover. This tabulated information served as input for the training and validation of four machine learning algorithms, RF, NB, K-NN, and SVM in the KNIME software environment. Following, each of the methodological procedures is detailed, to facilitate their understanding and reproduction.

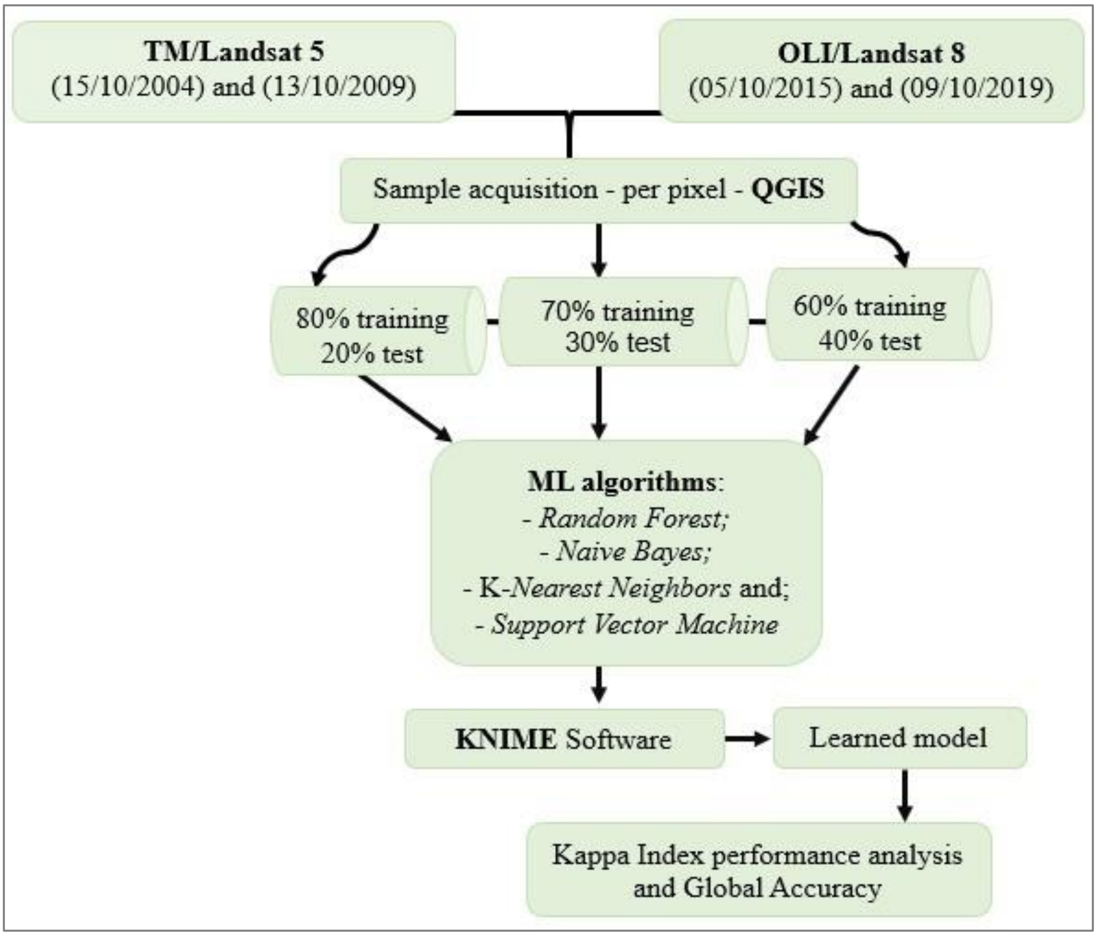


Figure 2. Methodological procedures flowchart.
Figura 2. Fluxograma dos procedimentos metodológicos.

Identification of usage patterns in the QGIS software

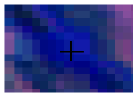
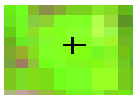
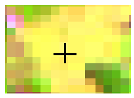
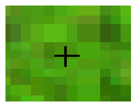
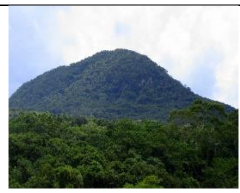
The geographic database was built using QGIS software version 2.14 Essen, and vector data from the State of Rio Grande do Sul with their respective municipal physical limits were imported. Thus, the area of interest was delimited, with cartographic projection UTM (Universal Transverse Mercator), Datum SIRGAS 2000, and spindle 22 J. The images (Table 1) of the TM/Landsat 5 and OLI/Landsat 8 sensors were acquired free of charge on the website of the National Institute for Space Research (INPE) and on the website of the Remote Sensing Center of the United States Geological Survey (USGS). For the images acquisition, the choice of images that had a maximum of 5% cloud cover was taken into account, as well as the season, with images from the same period, october, being selected.

Table 1. Comprehensive images of Landsat 5 and Landsat 8 and their dates.
Tabela 1. Imagens abrangentes do Landsat 5 e Landsat 8 e as respectivas datas de aquisição.

Scene	Órbit	Point	Date of acquisition	Satellite	Sensor
1	222	081	15/10/2004	Landsat 5	TM
			13/10/2009		
			05/10/2015	Landsat 8	OLI
			09/10/2019		

Subsequently, the pixels of the synthetic images of the TM/Landsat 5 RGB 543 and OLI/Landsat 8 RGB 654 sensors were sampled, taking into account only the spectral information of each pixel (numerical value), to obtain information to characterize and differentiate land use and land cover patterns (Table 2). 250 pixels were acquired at random for each time series in the years 2004, 2009, 2015, and 2019.

Table 2. Land use and land cover patterns are seen in the RGB 543 images from the TM/Landsat 5 sensor and RGB 654 from the OLI/Landsat 8 sensor.
Tabela 2. Padrões de uso e cobertura da terra visualizados nas imagens RGB 543 do sensor TM/Landsat 5 e RGB 654 do sensor OLI/Landsat 8.

Pattern	Characteristics	Sampled pixel	Sample real
Water	Natural and/or artificial reservoirs and, in some cases, areas used for rice cultivation.		
Agriculture	Agricultural areas are covered by vegetation.		
Field	Undergrowth is composed of grasses and pastures.		
Forest	Arboreal vegetation formations, including native and planted forests.		
Exposed soil	Agricultural areas in the preparation or fallow, gullies, or erosive processes.		

Then, the data generated using the point sampling tool were saved in an dataset. In this dataset, the geographic coordinates (X and Y), land use and coverage patterns (water, agriculture, field, forest, and exposed soil), and the six spectral bands (TM/Landsat 5 - Bands 1, 2, 3, 4, 5 and 7 and OLI/Landsat 8 - Bands 2, 3, 4, 5, 6 and 7, with their respective numerical values (spectral information for each pixel).

Machine Learning in KNIME platform

Through the samples from the reference pixels of the satellite images for the different dates, it was possible to prepare the data for reading by the miner, and thus, start the machine learning process on the KNIME platform. The reference pixel samples for training and testing were set at 80% to 20%, 70% to 30%, and 60% to 40%, respectively. The nodes used and their operation in KNIME is described below:

1) File Reader: responsible for reading data from a repository, in this case, the file entry was in *.csv format;

2) Partitioning: the input table is divided into two partitions, where this node is used to produce a set of training data (where the model is generated) and the set of test data (evaluation of the trained model).

3) Naive Bayes, Random Forest, Support Vector Machine, and K-Nearest Neighbors Learner: these nodes create a model from the data provided to the four machine learning algorithms;

4) Naive Bayes, Random Forest, Support Vector Machine and K-Nearest Neighbors Predictor: predicts the class per line based on the model learned and;

5) Scorer: reports a series of precision statistics, such as shows the global accuracy and Cohen's Kappa index.

It is important to note that at each node there is a semaphore below it, which indicates the state of readiness of that node. If it is in the green tone, it means that the node has already rotated and completed its work, in yellow the node is properly configured and ready to rotate, however, it has not yet completed its work, and if it is red, the node has not yet been configured and cannot be executed until it has been configured correctly.

To assess the accuracy of the classification quality, the Kappa index proposed by Landis and Koch (1977) was considered: Rubbish (< 0.00); Bad (0.0 – 0.20); Reasonable (0.20 – 0.40); Good (0.40 – 0.60); Very good (0.60 – 0.80); Excellent (0.80 – 1.00).

RESULTS

The four RF, NB, SVM, and K-NN algorithms were used to classify each data set from the pixel values of the synthetic composition of the RGB 543 TM/Landsat 5 sensor bands (Figure 3 (a) (b)) and RGB 654 OLI/Landsat 8 sensor (Figure 4 (c) and (d)) 250 pixels were collected for each time series (2004, 2009, 2015, and 2019), grouped into five classes of use and coverage patterns: water, agriculture, field, forest, and exposed soil.

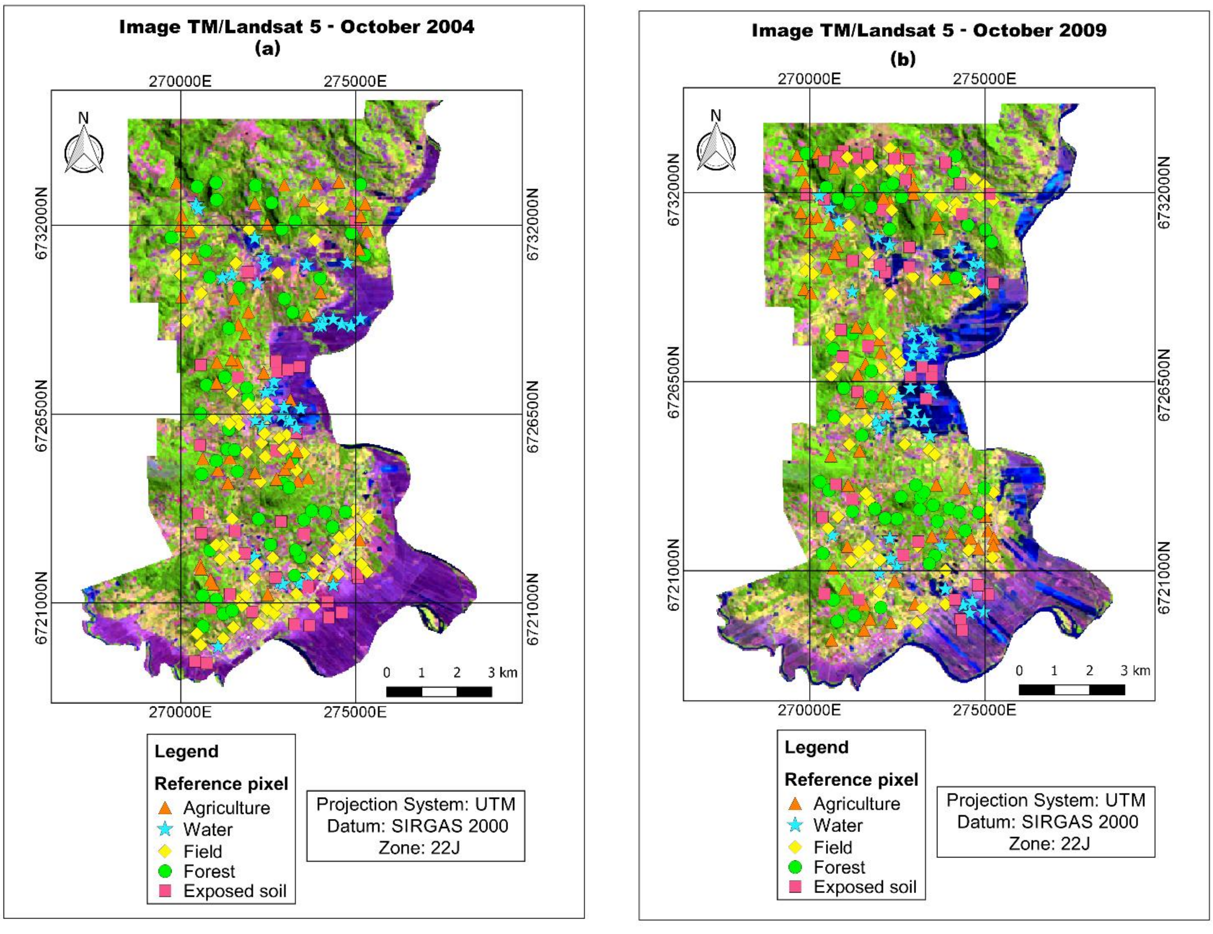


Figure 3. Identification of the sampled pixels as a reference in the RGB 543 synthetic image of the TM/Landsat 5 sensor, in the different patterns of land use and coverage for the years 2004 (a) and 2009 (b), municipality of Dona Francisca - RS.

Figura 3. Identificação dos pixels amostrados como referência na imagem sintética RGB 543 do sensor TM/Landsat 5, nos diferentes padrões de uso e cobertura da terra para o ano de 2004 (a) e 2009 (b), município de Dona Francisca – RS.

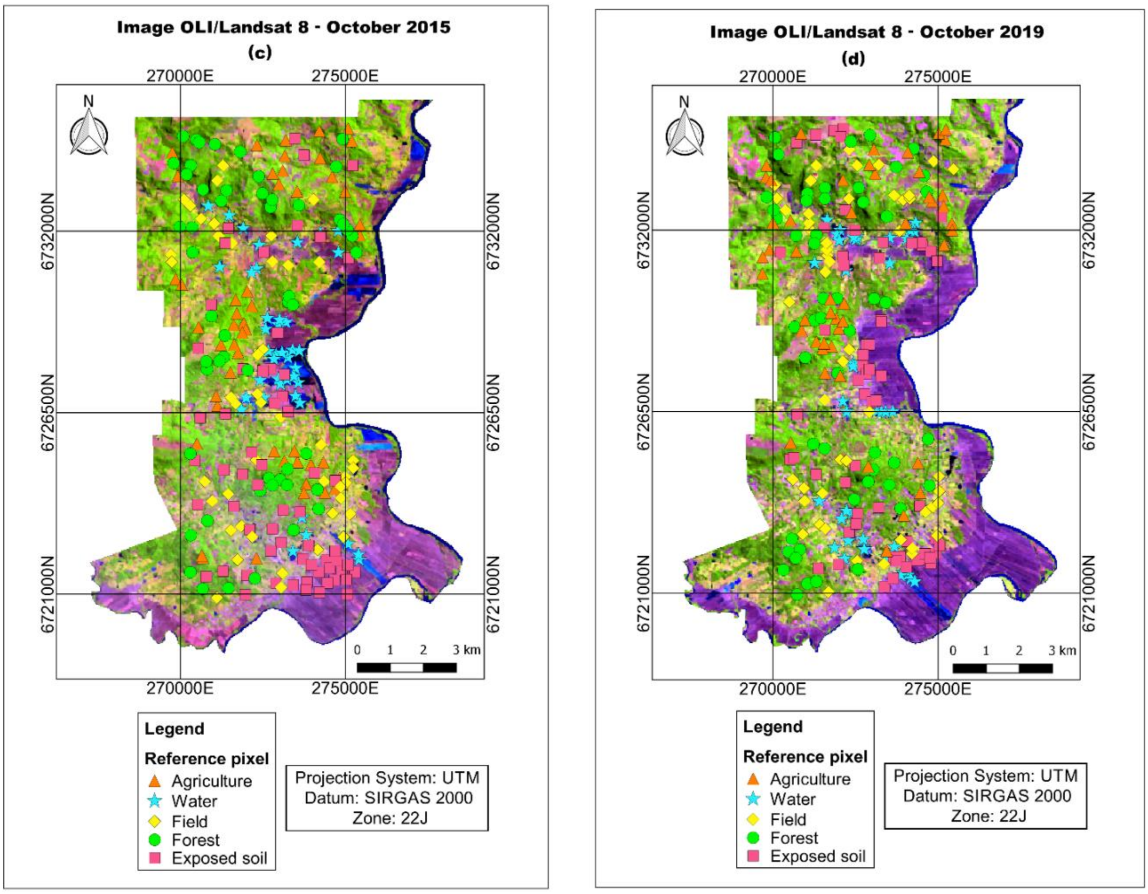


Figure 4. Identification of the sampled pixels as a reference in the RGB 654 synthetic image of the OLI/Landsat 8 sensor, in the different patterns of land use and coverage for the year 2015 (c) and 2019 (d), municipality of Dona Francisca - RS.

Figura 4. Identificação dos pixels amostrados como referência na imagem sintética RGB 654 do sensor OLI/Landsat 8, nos diferentes padrões de uso e cobertura da terra para o ano de 2015 (c) e 2019 (d), município de Dona Francisca – RS.

The results of AM predictions of land use and land cover patterns for the TM/Landsat 5 sensor (2004 and 2009) showed an accuracy of 0.85 and 0.86, respectively, classified as excellent, according to Landis and Koch (1977). The RF obtained the best results, showing good accuracy on the two dates analyzed, followed by the SVM (Table 3). The NB and K-NN algorithms showed bottom quality to the accuracy performance when compared to the others and, consequently, greater inconsistencies of the pixel values in mainly predicting the agriculture, field, and forest patterns.

Table 3. Global accuracy and Kappa index in the predictions of land use and land cover patterns, using the KNIME software – TM/Landsat 5 sensor, October 2004 and 2009.

Tabela 3. Acurácia global e índice de Kappa nas predições dos padrões de uso e cobertura da terra, utilizando o software KNIME - sensor TM/Landsat 5, outubro de 2004 e 2009.

October 15, 2004							
RF		**SVM**		**NB**		**K-NN**	
Test 20%							
AG	**Kappa**	**AG**	**Kappa**	**AG**	**Kappa**	**AG**	**Kappa**
89.5%	0.86	87%	0.83	84.5%	0.80	88.5%	0.85
Test 30%							
91.4%	0.89	88%	0.85	86.2%	0.82	86.2%	0.82
Test 40%							
92%	0.89	90%	0,87	88%	0.84	89.3%	0.86

October 13, 2009							
Test 20%							
90.5%	0.88	91.5%	0.89	82.5%	0.78	88%	0.84
Test 30%							
92.5%	0.90	90.8%	0.88	84.5%	0.80	89.1%	0.86
Test 40%							
93.3%	0.91	93.3%	0.91	86%	0.82	90.5%	0.88

Note AG: Global accuracy.

The classifications generated by the AM algorithms in the prediction of the land use and land cover patterns of the OLI/Landsat 8 sensor for the years 2015 and 2019, demonstrated an excellent accuracy according to Landis and Koch (1977), less for the NB, in which was considered to be very good on both dates. Table 4 shows that the best classification performances regarding the number of pixels, belonging to each pattern, were obtained by the K-NN and RF algorithms, respectively.

Table 4. Global accuracy and Kappa index in predictions of land use and land cover patterns, using the KNIME software – OLI/Landsat 8 sensor, October 5, 2015, and October 9, 2019.

Tabela 4. Acurácia global e índice de Kappa nas predições dos padrões de uso e cobertura da terra, utilizando o software KNIME - sensor OLI/Landsat 8, outubro de 2015 e 2019.

October 5, 2015							
RF		**SVM**		**NB**		**K-NN**	
Test 20%							
AG	**Kappa**	**AG**	**Kappa**	**AG**	**Kappa**	**AG**	**Kappa**
91%	0.88	83.5%	0.79	78%	0.72	91.5%	0.89
Test 30%							
93.7%	0.92	88%	0.85	79%	0.74	92%	0.90
Test 40%							
90.6%	0.88	90%	0.87	84.6%	0.80	94%	0.92
October 9, 2019							
Test 20%							
88%	0.84	84%	0.79	77%	0.70	90.5%	0.88
Test 30%							
90.8%	0.88	81%	0.75	80%	0.74	91%	0.89
Test 40%							
91.3%	0.89	86%	0.82	81.3%	0.76	92%	0.90

Note AG: Global accuracy.

DISCUSSION

The patterns of use and coverage of the TM Landsat 5 and OLI Landsat 8 sensor images showed subtle variations between all the algorithms used. In 2004, the RF and SVM algorithms had similar results to those of Pal (2005), when using images from the Landsat Enhanced Thematic Mapper Plus (ETM +) for the classification of different pixel-based land cover, where he found that the RF classifier obtained 88.4% accuracy and that it can achieve a classification accuracy very similar to SVM (87.9%).

Moreira *et al.* (2014), when using six dates from Landsat 5 to map the use and land cover of the Rio Vieira basin - MG, they obtained SVM accuracy ranging from 83.51% to 93.23%, similar data were found in this study using images from Landsat 5, in which they ranged from 87% to 93.3%. This similarity in performance may be linked to the spatial resolution of TM/Landsat sensor images. Another fact that the RF presents better results according to Zha *et al.* (2020), is related to the greater capacity for generalization compared to other ML methods.

In an evaluation of the Landsat 5 images in the classification of 14 different land cover classes in southern Spain, Galiano *et al.* (2012) found that the RF algorithm produces accurate classifications, with 92% global accuracy. According to Duro *et al.* (2012), the RF can achieve the best accuracy results in the classifications, while the NB has the worst results, a fact found in this study.

The NB algorithm on all studied dates presented confusion between the pixel values in the forest, agriculture, water, and exposed soil patterns. These last two may be linked to the fact that many properties in the municipality, in October, plant pre-germinated rice, being carried out by flood irrigation, thus, the low performance of the NB may be related to the similarity spectral of these humid areas, causing confusion between exposed soil and water patterns. Thus, to improve the classification performance of the NB, more homogeneous and large samples can lead to a more accurate parameter estimate, since this algorithm is sensitive to the size and uniformity of the training samples (QIAN *et al*., 2015).

Cracknell and Reading (2014), found that the NB algorithm did not obtain the best accuracy results, about RF, K-NN, SVM, and artificial neural networks (ANN), in which they varied from 0.54% to 0.64%, using Landsat 7 ETM + images to map geological areas. Müller *et al.* (2015), analyzing the RF algorithm in Landsat 8 time series to classify crops, pasture, and natural vegetation obtained global accuracy of up to 93%. Evaluating several classifiers for mapping invasive woody species, in the Azores in Portugal, Gil *et al*. (2014), found the best accuracy results for K-NN ranging from 89% to 94%, when compared with SVM and Maximum Likelihood (MLC). Sarmiento *et al*. (2014), showed that the K-NN algorithm presented good indexes (90% of correctness) for the mapping of coffee areas in Campos Gerais, in the State of Minas Gerais. Heydari and Mountrakis (2018) compared the NB, KNN, SVM, Tree Ensemble and ANN algorithms to classify land cover and identified SVM and KNN as the best performing methods for Landsat classifications.

For the classification to be considered adequate, according to the authors Hentz *et al*. (2015); Gao (2009), it is sought that their values are high, that is, the higher the value of global accuracy, the greater the number of pixels classified correctly. Analyzing in general, it was possible to observe that in the four dates analyzed (2004, 2009, 2015, and 2019) there was an increase in precision by increasing the proportion of test data from the ML algorithms. Thus, our results indicate that the proportion of 60% (training) and 40% (test) is more suitable for predicting land use and land cover patterns using Landsat images in areas with similar characteristics to the study.

Another point to be highlighted is the precision of the algorithms concerning the analyzed time series, and it is possible to infer that, over the years, there was a decrease in accuracy, except K-NN. This fact may be linked to the different radiometric resolutions of the TM (8-bit) and OLI (16-bit) sensors, in which the K-NN algorithm allowed to cover the largest variations in the digital levels linked to the images with higher radiometric resolution (OLI), while that the other algorithms had difficulties in predicting the patterns of land use and land cover in these images (OLI). Thus, to increase the accuracy in the identification of land use and land cover patterns in OLI/Landsat 8 images, we recommend increasing the sampling of pixels, which may assist in improving the performance of the algorithms, since, to contemplate the greatest variability in terms of gray levels in the image.

It is important to note that the KNIME software, with its intuitive approach, facilitated the application of ML tools, used to predict land use and land cover patterns from time series. Besides, the software allowed to expand the capacity to extract information, which can assist in strategic planning for territorial management.

CONCLUSION

- The use of machine learning algorithms to predict land use and land cover patterns for the spectral data of the Thematic Mapper (TM) sensor of Landsat 5, proves to be efficient, highlighting the RF and SVM algorithms, which showed higher values for the parameters of global accuracy and Kappa index. Regarding the spectral data from the Landsat 8 Operational Land Imager (OLI) sensor, the K-NN and RF algorithms showed satisfactory results.
- When analyzing the responses of the NB algorithm, it was observed that there was greater confusion in predicting the patterns, when compared to the others.

REFERENCE

ABDI, A.M. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. **GIScience & Remote Sensing**, v. 57, n. 1, p. 1-20, August, 2020.

ALVARES, C.A.; STAPE, J.L.; SENTELHAS, P.C.; GONÇALVES, J.L.M.; SPAROVEK, G. Köppen's Climate Classification Map for Brazil. **Meteorologische Zeitschrift**, Sttutgart, v. 22, n. 6, p. 711-728, 2014.

ANDRADE, A.C.; FRANCISCO, C.N.; ALMEIDA, C.M. Desempenho de classificadores paramétricos e não paramétricos na classificação da fisionomia vegetal. **Revista Brasileira de Cartografia**, Rio de Janeiro, v. 66, n. 2, p. 349-363, Julho, 2014.

BELMAKER, J.; ZARNETSKE, P.; TUANMU, M.N.; ZONNEVELD, S.; RECORD, S.; STRECKER, A.; BEAUDROT, L. Empirical evidence for the scale dependence of biotic interactions. **Global Ecology and Biogeography**, v. 24, n. 7, p. 750-761, April, 2015.

BUĞDAY, E.; ERKAN BUĞDAY, S. Modeling and simulating land use/cover change using artificial neural network from remotely sensing data. **CERNE**, Lavras, v. 25, n. 2, p.246-254, Apr-Jun, 2019.

BUNTING, P.; CLEWLEY, D.; LUCAS, R.M.E GILLINGHAM, S. The Remote Sensing and GIS Software Library (RSGISLib). **Computers and Geosciences**, New York, v. 62, p. 216-226, January, 2014.

CRACKNELL, M. J.; READING, A. M. Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. **Computers and Geosciences**, New York, v. 63, p. 22-33, 2014.

GALIANO, R.V.F.; GHIMIRE, B.; ROGAN, J.; CHICA-OLMO, M.; RIGOL. J.P.S. An assessment of the effectiveness of a random forest classifier for land-cover classification. **Journal Photogrammetry & Remote Sensing**, 67: 93–104, 2012.

GAO, J. **Digital analysis of remotely sensed imagery**. New York, USA: McGraw-Hill Professional Publishing, 2009. 439 p.

GE, G.; SHI, Z.; ZHU, Y.; YANG, X.; HAO, Y. Land Use/Cover Classification in an Arid Desert-Oasis Mosaic Landscape of China Using Remote Sensed Imagery: Performance Assessment of Four Machine Learning Algorithms. **Global Ecology and Conservation**, v. 22: p. e00971, 2020.

GIL, A.; YU, Q.; ABADI, M.; CALADO, H. Using aster multispectral imagery for mapping woodyinvasive species in Pico da Vara natural reserve (Azores Islands, Portugal). **Revista Árvore**, Viçosa, v. 38, n. 3, p. 391-401, 2014.

HEYDARI, S.S; MOUNTRAKIS, G. Effect of classifier selection, reference sample size, reference class distribution and scene heterogeneity in per-pixel classification accuracy using 26 Landsat sites. **Remote Sensing Environment**, v. 204 p. 648–658, 2018.

HENTZ, Â. M. K.; DALLA CORTE, A. P.; BLUM, C. T.; SANQUETTA, C. R. Técnicas orientadas ao objeto para levantamento da fragmentação florestal na sub-bacia Alto Iguaçu, Paraná. **Geociências**, São Paulo, v. 34, n. 4, p.883-896, 2015.

IBGE. **Instituto Brasileiro de Geografia e Estatística**. 2019. Disponível em: <www.cidades.ibge.gov.br>. Acesso em: 08 de jan. de 2020.

KNIME. **Konstanz Information Miner**. 2019. Disponível em: <http://www.knime.org/. >. Acesso em: 15 de jan. de 2020.

LAGO, R.B.; FARENZENA, D. 2008. Contrastes socioeconômicos do conjunto habitacional Gentil Tessele e área central da cidade de Dona Francisca, RS. **Disciplinarum Scientia**, Santa Maria, v. 9, n. 1, p. 17-34.

LANDIS, J.; KOCH, G. G. The measurements of agreement for categorical data. **Biometrics**, Washington, v. 33, n. 3, p. 159-179, March, 1977.

LI, C.; WANG, J.; WANG, L.; HU, L.; GONG, P. Comparison of classification algorithms and training sample sizes in urban land classification with Landsat Thematic Mapper imagery. **Remote Sensing**, Suíça, v. 6, n. 2, p. 964-983, January, 2014.

MOREIRA, A. A.; NERY, C. V. M.; FERNANDES, F. H. S. Aplicação do algoritmo Support Vector Machine na análise espaço-temporal do uso e ocupação do solo na bacia do Rio Vieira. **Caminhos de Geografia**, Uberlândia, v. 15, n. 50, Junho, 2014.

MÜLLER, H.; RUFIN, P.; GRIFFITHS, P.; BARROS, A.J.S.; HOSTERT, P. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. **Remote Sensing of Environment**, v. 156, p. 490-499, January, 2015.

PAL, M. Random forest classifier for remote sensing classification. **International Journal of Remote Sensing**, v. 26, n. 1, p. 217-222, May, 2005.

QIAN, Y.; ZHOU, W.; YAN, J.; LI, W.; HAN, L. Comparing machine learning classifi ers for object-based land cover classifi cation using very high resolution imagery. **Remote Sensing**, Suíça v. 7, n. 1, p. 153-168, 2015.

SARMIENTO, C.M.; RAMIREZ, G.M.; COLTRI, P.P.; SILVA, L.F.L.; NASSUR, O.A.C.; SOARES, J.F. Comparação de classificadores supervisionados na discriminação de áreas cafeeiras em Campos Gerais – Minas Gerais. **Coffee Science**, Lavras, v. 9, n. 4, p. 546- 557, Outubro-Dezembro, 2014.

ZHA, H.; MIAO, Y.; WANG, T.; LI, Y.; ZHANG, J., SUN, W.; KUSNIEREK, K. Improving unmanned aerial vehicle remote sensing-based rice nitrogen nutrition index prediction with machine learning. **Remote Sensing**, Suíça v. 12, n. 2, p. 1-22, January, 2020.

ZHU, Z.; WOODCOCK, C.E.; ROGAN, J. KELLNDORFER, J. Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data. **Remote Sensing of Environment**, n. 117, p. 72-82, February, 2012.