Revista Brasileira de Cartografia
ISSN 1808-0936 | https://doi.org/10.14393/revbrascartogr
Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto

Mapping Pinus spp. Forestry and Land Cover Classes Using High-resolution PlanetScope Satellite Data: Experimenting Images from Different Seasons and Machine Learning Methods

Mapeamento de Florestas de Pinus spp. e Classes de Cobertura do Solo Usando Dados de Satélite PlanetScope de Alta Resolução: Experimentando Imagens de Diferentes Estações e Métodos de Aprendizado de Máquina

Andressa Kossmann Ferla [1], Fábio Marcelo Breunig [2], Rafaelo Balbinot [3], Ricardo Dal'Agnol da Silva [4]

1 Universidade de Santa Maria, Frederico Westphalen, Brasil. andressaferla94@gmail.com.
ORCID: https://orcid.org/0000-0003-1801-3355
2 Universidade Federal do Paraná, Curitiba, Brasil. fabiobreunig@ufpr.br.
ORCID: https://orcid.org/0000-0002-0405-9603
3 Universidade Federal de Santa Maria, Frederico Westphalen, Brasil. rafaelo.balbinot@gmail.com.
ORCID: https://orcid.org/0000-0001-6209-8232
4 University of California Los Angeles, Los Angeles, Estados Unidos. ricds@hotmail.com.
ORCID: https://orcid.org/0000-0002-7151-8697

Received: 12.2022 | Accepted: 11.2023

Abstract: The Remote Sensing and machine learning techniques are cost-effective ways of mapping land use and cover, especially forestry areas. This is essential for the management and planning of such resources. The purpose of this study was to identify which classifier (Random Forest or Support Vector Machine) reach the best accuracy in land use and cover classification and determine which is the best season of year for *Pinus spp*. forest mapping. PlanetScope multispectral image was used with 3.7 m of spatial resolution, collected over the coastal region of Rio Grande do Sul state. The input variables for the classifiers were the four spectral bands: RGB and NIR, and the NDVI vegetation index. In both classifiers, high accuracy values were obtained, as well as for all seasons of the year. The Random Forest classifier obtained better results in the spring and summer seasons, while in the autumn and winter seasons there was no significant difference between the classifiers for the classification of *Pinus spp*. forests. The results reached an adequate precision to be used for the management and monitoring of the land use and cover in the municipality of São José do Norte, RS.
Keywords: Mapping. Pine. Machine Learning. GIS. Remote Sensing.

Resumo: O uso do Sensoriamento Remoto e técnicas de aprendizado de máquina são amplamente utilizadas para o mapeamento do uso e ocupação do solo, e em especifico, de áreas de silvicultura, de forma ágil e acessível. Esse resultado é fundamental para a gestão e planejamento desses recursos. O objetivo do trabalho foi identificar qual classificador (*Random Forest* ou *Support Vector Machine*) atinge melhor acurácia na classificação do uso do solo e determinar qual é a melhor época do ano para o mapeamento de florestas de *Pinus spp*. Foram utilizadas imagens multiespectrais PlanetScope, contendo 3,7 m de resolução espacial, coletadas sobre a região costeira do estado do Rio Grande do Sul. As variáveis de entrada utilizadas nos classificadores foram as quatro bandas espectrais: RGB e NIR e o índice de vegetação NDVI. Em ambos os classificadores se obteve altos valores de acurácia, bem como em todas as épocas do ano. O classificador *Random Forest* obteve melhores resultados nas estações primavera e verão, já nas estações outono e inverno não houve diferença significativa entre os classificadores para a classificação de florestas de *Pinus sp.*. Os resultados atingiram precisão adequada para serem utilizados de forma confiável no gerenciamento e monitoramento da cobertura da terra na região costeira do Rio Grande do Sul.
Palavras-chave: Mapeamento. Pinus. Aprendizado de Máquina. SIG. Sensoriamento Remoto.

1 INTRODUCTION

Land use maps contribute to the forwarding of land use policies, planning, conservation, agricultural monitoring, agricultural management, and forestry management. However, accurate mapping of heterogeneous landscapes is a challenge (CHEN et al., 2021; SCHULZ et al., 2021). To improve the accuracy of a land cover classification using satellite images, it is important to select suitable predictor variables from calibrated remote sensing data, in order to encompass the representative statistical parameters of each class of land use and land cover. As high spatial resolution data usually have restricted spectral information (few bands), the addition of spatial features becomes critical (LI et al., 2014; DHINGRA; KUMAR, 2019). Individual bands, vegetation indices, or even transformed images (XIE et al., 2019) represent spectral variables. For Xie et al. (2019), several studies explore the classification of tree species using high spatial resolution image and with several variables such as spectral signatures, vegetation indices, textural metrics, among other auxiliary data; however, not all input variables are useful and even some combinations of predictive variables can lead to a decrease in classification accuracy (BREUNIG et al., 2011). Luo et al. (2022) reports that excessive input resources can decrease classification accuracy and increase computation time. Therefore, developing feasible, assertive, and low-complexity classification protocols is essential for the process of technology transfer to the productive sector.

Remote sensing has been used for a long time to map land use and land cover, allowing it to cover large territorial extensions and with a good cost-benefit ratio (XIE et al., 2019). Sheykhmousa et al. (2020) exposes that the classification of an image becomes a costly task due to the large volume of data, the landscape heterogeneity, as well as limited and unbalanced training data between classes. The competence and computational cost of image classification are influenced by different classification algorithms, sensor types, training samples, input resources and by pre- and post-processing techniques, auxiliary data, target classes, and the final product accuracy (LI et al., 2014; SHEYKHMOUSA et al., 2020). High spatial resolution images can improve the quality of forest distribution mapping (XU et al., 2021). On the other hand, due to some implications, such as difficulty in automatic classification, number of available bands, high cost of purchasing images, and computational cost for processing large volumes of data; these images are little used in forestry applications. However, on local and regional scales, high spatial resolution data achieve better results (XIE et al., 2019). In this context, recent satellite constellations have generated a large volume of data in different spectral ranges (PERKO et al., 2018; SOZZI et al., 2019; WULDER et al., 2019).

Promoting the miniaturization of satellites and associating technological advances, satellite constellations have emerged in different areas of application. The PlanetScope satellite constellation operates with more than 150 satellites (3U model) in different generations (PLANET TEAM, 2019). Given the large number of satellites, images are basically taken every day with a spatial resolution of 3.7 meters. Although the recent generation has increased the number of bands, they are generally limited to visible and near infrared (NIR). This fact may limit the different tree species classification accuracy, due to the spectral similarity between species (XIE et al., 2019; HILL et al., 2010). Therefore, in a complementary manner, phenological information can increase the separability between tree species, given that the phenological cycle of forest species causes a variation in reflectance at different dates. It is noteworthy that for some species, phenological differences are costly to obtain, and sometimes not noticeable with orbital sensors (HADDAD et al., 2022; GRABSKA et al., 2019).

Remote sensing techniques, especially optical images, have long been analyzed and applied for forest assessment (HILL et al., 2010). In this context, several classifiers have been created and used for land cover mapping; however, their individual performance may vary according to input data, reference class definitions, and environmental conditions of the region of interest. A classifier may work well in a given geographic area and may not work well in another (SCHULZ et al., 2021). Xie et al. (2019) mentions that it is currently not conclusive which classification algorithm provides the best performance and which different seasonal characteristics and different data sources lead to a better land cover classification or forest classification. The image acquisition time may be more important than the number of images in a multitemporal analysis (GRABSKA et al., 2019). In this context, the Random Forest (RF) (BELGIU; DRĂGU, 2016; LARY et al., 2016; OK; AKAR; GUNGOR, 2012) and Support Vector Machine (SVM) (PAL; FOODY, 2010; MONNET;

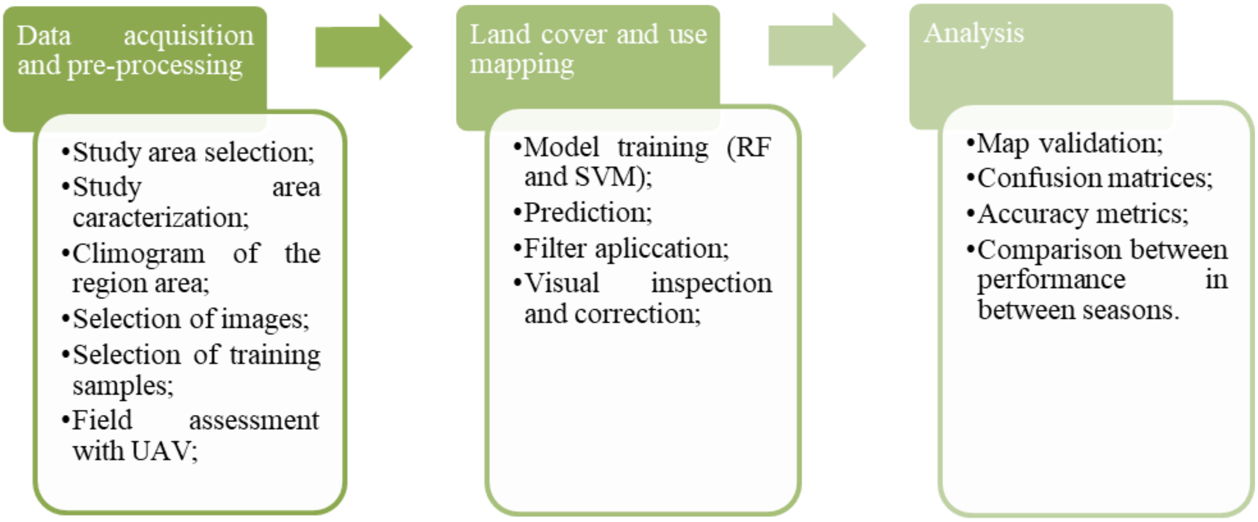
CHANUSSOT; BERGER, 2011; MOUNTRAKIS; IM; OGOLE, 2011) machine learning algorithms have been showing promising results, being widely used by the scientific community.

In this perspective, this study aims to evaluate the performance of two machine-learning methods (Random Forest and Support Vector Machines) and the use of high-resolution PlanetScope satellite imagery acquired in different seasons on mapping land cover and use classes with distinction of Pinus spp. forestry species from other land use and covers in southern Brazil at the Pampa biome. We aim to evaluate the impact of seasonality on the accuracy of classifications, as well as to identify the best date for the separation of Pine Forestry, using high-resolution PlanetScope spatial images. This is important to achieve accurate and cheap land cover classification for landscape planning in the region.

2 MATERIALS AND METHODS

The developed workflow basically involved three steps (Figure 1). The first phase of data acquisition, followed by processing and, finally, with post-classification.

Figure 1 – Flowchart of the methodology.



Source: The authors (2022).

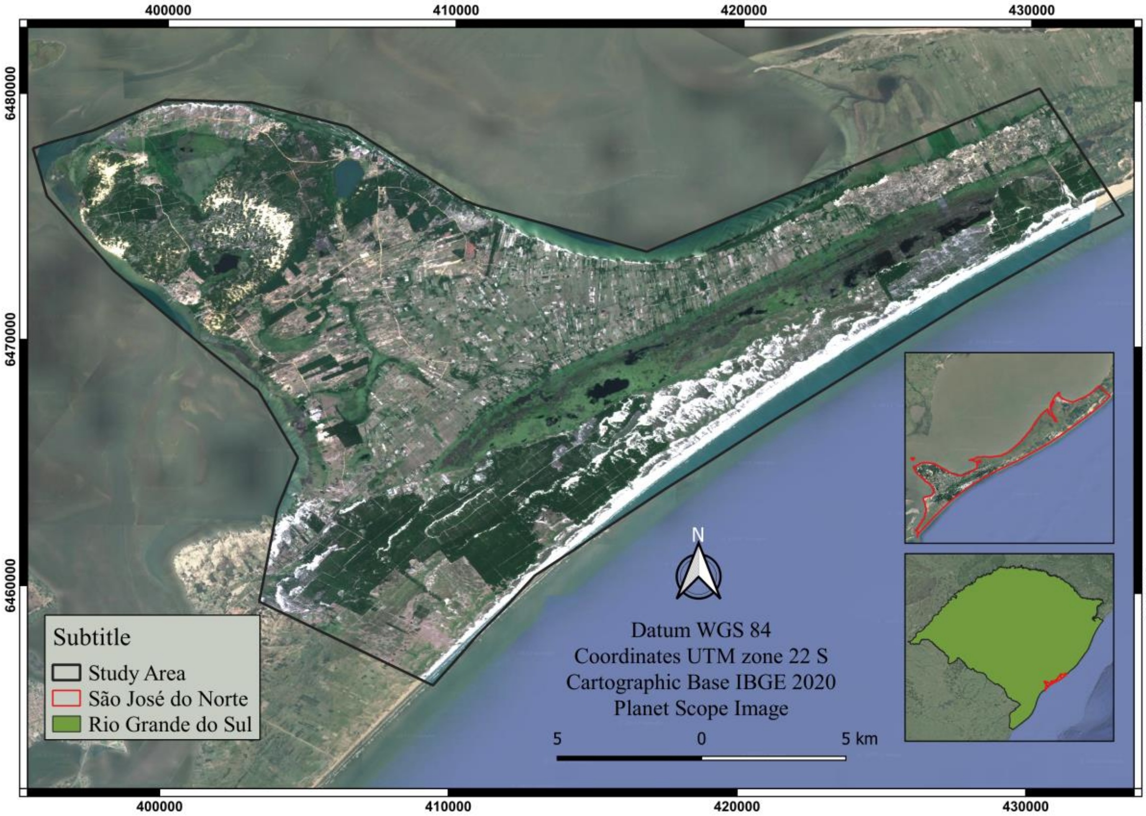
2.1 Study area

The study area is located in the municipality of São José do Norte, on the coastal plain of Rio Grande do Sul, in the Pampa Biome (Figure 2). The region belongs to the Marine Coastal System on a peninsula located between Lagoa dos Patos and the Atlantic Ocean. According to the IBGE census (2021), São José do Norte has a population of 27,866 inhabitants, with a 1,071.824 km² area. The municipality's economic base in the primary sector is fishing, crop production, especially onion and rice production, livestock, mining, and forestry (cultivation of *Pinus sp.*).

According to Pinheiro e Silva (2018) São José do Norte is the largest representative in the production of logs and gum in the region. It has more than 70% of its area dedicated to agricultural activities and forestry, with these activities being the basis of the economy; the cultivation of exotic forests covers an area of approximately 11,500 hectares. São José do Norte has large extensions of wetlands, which form part of the natural physiognomy of the municipality, normally bordering the lakes and performing the transition of the lakes with other formations (TAGLIANI; VICENS, 2003). The wetlands have a diversity of macrophyte plant communities that vary according to the water regime, morphometry, among other physical characteristics common to each system (GIANUCA; TAGLIANI, 2012). In this study, an area within the municipality of São José do Norte that has a large amount of planted forest was delimited for carrying out the research.

The study area is located at an altitude between 3 and 5 meters above mean sea level. The soil is predominantly sandy, of recent formation. According to Alvares *et al.* (2013), São José do Norte has hot summers; the classification according to Köppen and Geiger is Cfa. According to Inmet (2022), the average annual temperature is 16.48 ºC, and the average annual precipitation is 1,466 mm. Thus, data from the weather station closest to the study area from the municipality of Pelotas - RS were collected, and a climogram was created to identify the temperature and precipitation in the region between 10 and 30 years.

Figure 2 – Location map of the study area, in the city of São José do Norte – RS. True-color composite of the PlanetScope image from 1/8/2021 is presented.



Source: The authors (2022).

2.2 Data acquisition and pre-processing

The study used high spatial resolution remote sensing images of the PlanetScope constellation, operated by Planet Labs (PLANET LABS, 2019). The level 3B product was used – surface reflectance, with orthorectified and scaled images, with a 3.7 meters spatial resolution, radiometric resolution of 12 *bits*, covering four different bands: *Red* (590 – 670 nm), *Green* (500 – 590 nm), *Blue* (455 – 515 nm), and NIR (780 – 860 nm). The images were taken without the presence of clouds in the four seasons of the year, on the dates: autumn: 5/12/2021, winter: 7/29/2021, spring: 10/28/2021, and summer: 1/12/2022. The seasonal approach aims to enhance the discrimination between forestry areas and other land cover and uses. Therefore, knowing the phenological cycle of the plants under study is considered essential, so that the most appropriate dates can be selected for collecting information (JENSEN, 2011).

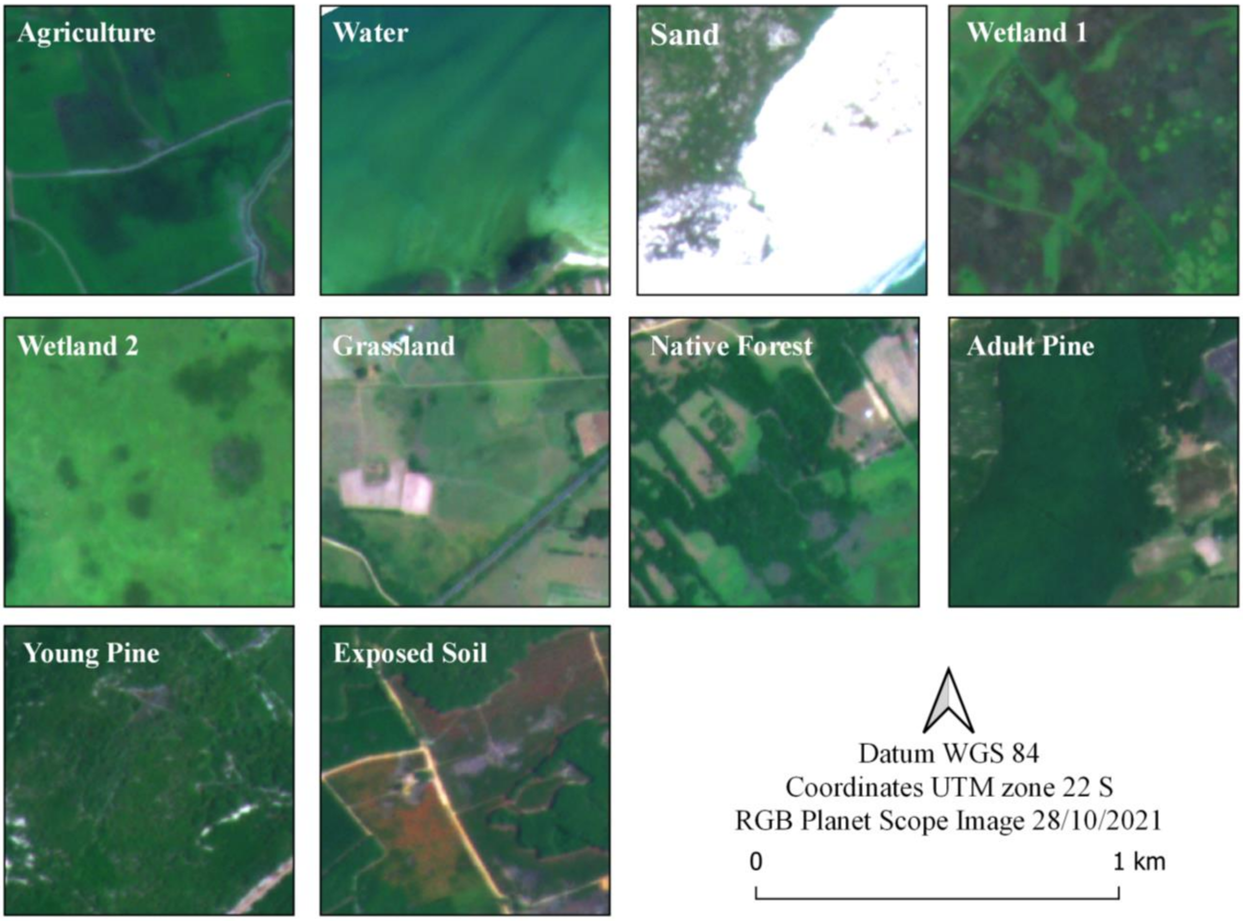
With the purpose of distinguishing the native forest from the forestry a hierarchical classification we defined two hierarchical levels for the land cover and use mapping, the first level having eight classes, and a second level with more detailed classes for vegetation (three classes). Pixel samples were collected through the photo-interpretation of the true-color composites of PlanetScope imagery.

For the first level, 2,979 pixel samples were collected, of which, 393 water samples were collected from the sea, lagoon, and lakes, 513 samples of sand, 247 samples of Exposed Soil, these in areas prepared for planting or recently cut and burned areas, 324 samples of Grassland, 54 samples of Agriculture, 1,155 Forest samples, 57 samples of Wetland 1 and these having brown color, 236 samples of Wetland 2, these being greenish, the wetlands were separated into two classes, so that there was no confusion between other classes, as they are spectrally different. For the second level, the Forest samples were separated into: 556 samples of

 DOI: http://dx.doi.org/10.14393/rbcv75n0a-67769

Native Forest, 723 samples of Adult Pine, and 230 samples of Young Pine, the pine forestry classes were separated according to the canopy closure, in which Adult Pine has a closed canopy and Young Pine has a less closed canopy. They were randomly collected over the entire image, 70% of them were used for training and 30% were used for validation. With the help of Google Earth software and images taken with a Phantom 4 UAV, during a field survey, a careful distinction and visual interpretation of the classes was carried out. According to Novo (2010), although digital processing tools have great agility in the process of extracting information from large amounts of data, this extraction often only becomes complete with the visual analysis of the images. All processing was conducted in the QGIS 3.22.4 software (SHERMAN, 2002).

Figure 3 – Example of the land cover classes that were photo-interpreted with the PlanetScope images.



Source: The authors (2022).

2.3 Land cover and use mapping

For each of the four images (from each seasons) data processing was carried out involving the computation of vegetation indices and hierarchical classification. For this, the R programming language was used, through the RStudio software (RStudio 4.1.2 (2021-11-01)). The Normalized Difference Vegetation Index (NDVI; ROUSE et al., 1973) is based on the normalized difference between the reflectance of the red and near-infrared bands, allowing to differentiate pixels that comprise healthy and unhealthy vegetation, being closely associated with pigments (LEITE; SANTOS; SANTOS, 2017). Furthermore, the Enhanced Vegetation Index (EVI; HUETE et al., 2002) was calculated, which tends to improve sensitivity for high biomasses (JANSEN; VANDERWEL, 2011). Due to results not statistically superior to the NDVI, the study continued only with the NDVI, in order to simplify the process.

Considering the classification processes, for both classifiers (RF and SVM) and both classifications, the input spectral variables were the *Red*, *Green*, *Blue*, and NIR and NDVI bands. It should be noted that tests were carried out with the inclusion of textural metrics: *Mean*, *Homogeneity*, *Second Moment*, and *Correlation*, in a 3x3 window using the Red and NIR bands; however, the results of both classifications with eight and three classes did not significantly differ from the classification without textural metrics, according to the Hudson & Ramm test (1987). Analyzing the small percentage difference and the large computational demand, we decided not to use the textural metrics in the classification in order to reduce the classification processing time. Through the “caret” package (KUHN, 2015), in R, the samples were randomly divided into 70% training samples and 30% validation samples, based on each class data. The “caret” package has the function of developing parameter adjustments, selecting the values that maximize the precision according to the selected validation; it also allows to define the number of values used for each adjustable parameter, even if the specific values cannot to be selected (FERNÁNDEZ-DELGADO et al., 2014). In addition to the CARET package, the “raster”, “rgeos”, “randomForest”, and “e1071” packages were used (MEYER et al., 2021).

The RF classifier uses a tree structure in which each internal node is symbolized by a question related to the value of an attribute, and each external node is related to a class (*FACELI* et al., 2021). The RF uses a subset by randomly selecting datasets from samples and training variables producing multiple decision trees. It has the ability to successfully deal with high dimensionality of data and multicollinearity, being fast and insensitive to overfitting (*FACELI* et al., *2021*). Its performance can be affected by two parameters, which are the number of variables randomly selected at each node (*mtry*) and the number of decision trees (*ntree*) (BEHERA et al., *2021*). In the study, *ntree* was set as equal to 500, and *mtry* equal to 2.

The SVM is a linear binary classifier, which from a given test sample assigns a class of one of two probable labels. In remote sensing, the sample to be labeled is usually the individual pixel coming from the multispectral or hyperspectral image (MOUNTRAKIS et al., 2011). It is a training algorithm that automatically configures the ability of the classification function, leveraging the margin between the training examples and the class limit, after removing some atypical or nonsensical examples from the training data if necessary (CORTES; VAPNIK, 1995). In the study for each autumn, winter, spring, and summer image, the C-classification and linear kernel parameters were used. The resulting classification function depends on said support patterns, which are training examples that come closest to the decision limit and are commonly a subset of small training data (CORTES; VAPNIK, 1995).

After the models were trained, they were applied to predict the land cover and use classes for the whole region for each of season. As post-processing, a median filter was applied to homogenize the classes by using a moving window of 3x3 pixels. After that, a thematic vector map was created to facilitate the polygon area calculations.

2.4 Analysis

For the hierarchical classification for both classifiers, the validation of the results was firstly performed through visual inspection, followed by the evaluation of the confusion matrix and the Kappa index, in the classifications of first and second level, 890 and 451 pixels were used for validation, respectively; these were not used in the model training process. To identify whether the results of the classifications between seasons and classifiers have statistically significant results, the statistical test by Hudson & Ramm (1987) was performed.

With the purpose of increasing the accuracy of the classification, a test was performed using the summer image, adding textural metrics, comparing the values through the Hudson & Ramm (1987) test, with the Kappa index values of the summer image, without textural metrics. Finally, to quantify the classified areas, the image that obtained the best classification was vectorized to perform the area calculation using the field calculator.

3 RESULTS AND DISCUSSION

In both classifications, the best classification accuracy was observed in the RF classifier, in the eight-class classification mean Kappa = 0.95 and in the three-class classification mean Kappa = 0.69 compared to the SVM classifier, classification eight classes mean Kappa = 0.94 and classification three-class mean Kappa = 0.59. In Table 1 we can analyze the results of the Kappa index and Accuracy of classifications.

Summer image, adding textural metrics, comparing the values through the Hudson & Ramm (1987) test, with the Kappa index values of the summer image, without textural metrics, both values indicated in Table 1. There is no statistically significant difference between them.

Table 1 - Kappa index and global accuracy resulting from the classification using the Random Forest (RF) and Support Vector Machine (SVM) classifiers of each image referring to each season of the year and hierarchical levels.

Image	**Classifier**	**Kappa**	**Accuracy**
Autumn	*Random Forest 1st hierarchical level*	0.94	0.96
	Random Forest 2nd hierarchical level	0.65	0.79
	Support Vector Machine 1st hierarchical level	0.91	0.93
	Support Vector Machine 2nd hierarchical level	0.43	0.68
Winter	*Random Forest 1st hierarchical level*	0.93	0. 95
	Random Forest 2nd hierarchical level	0.68	0.81
	Support Vector Machine 1st hierarchical level	0.92	0.94
	Support Vector Machine 2nd hierarchical level	0.58	0.76
Spring	***Random Forest 1st hierarchical level***	**0.96**	**0.97**
	Random Forest 2nd hierarchical level	**0.76**	**0.86**
	Support Vector Machine 1st hierarchical level	0.95	0.96
	Support Vector Machine 2nd hierarchical level	0.72	0.84
Summer	***Random Forest 1st hierarchical level***	**0.98**	**0.98**
	Random Forest 2nd hierarchical level	**0.67**	**0.80**
	Support Vector Machine 1st hierarchical level	0.96	0.97
	Support Vector Machine 2nd hierarchical level	0.63	0.79
Summer with textural metrics	*Random Forest 1st* hierarchical level	0.98	0.98
	Random Forest 2nd hierarchical level	0.72	0.84

Elaboration: The authors (2022).

Through the Hudson & Ramm (1987) statistical test, it can be identified whether there is a statistically significant difference, for the classification of eight classes the summer and autumn images, the RF classifier presented a statistically significantly higher Kappa index than the SVM classifier, for the spring and winter images there was no significant difference between the Kappa of the classifiers. For the classification containing three classes, in the images of the autumn and winter seasons, the RF classifier presented a significantly higher kappa than the SVM, and for the spring and summer season images, there is no statistically significant difference between the Kappa of the classifiers.

The seasons that obtained the best accuracy and Kappa index results were spring (Kappa = 0,96 from RF model) and summer (Kappa = 0,98 from RF model), of which summer had the best accuracy in the first level classification and spring had the best accuracy in the second level classification. In the classification of first level, the Hudson & Ramm (1987) test demonstrated that the Kappa of the summer season image is significantly higher than that of spring. In the classification of second level, the RF classification in spring presented a significantly higher Kappa index than in summer. Such results can be related to the effect of the large presence of non-photosynthetically active vegetation in spring and, due to the greening rates in summer.

According to the Confusion Matrix (Table 2) all classes obtained excellent classification results The Forestry class presented the lowest misclassification due to its high greenness index throughout the seasons. Both native Forest and pine forestry keep their leaves throughout the year, while the grassland, for example, suffers with the low temperatures in the winter, compared to other seasons of the year (as we can analyze in the climogram generated for the study region, Figure 5) – presenting low NDVI values. The classes that did not represent vegetation were easily detected during all seasons. On the other hand, wetland, grassland, and pine were confused with each other, due to the fact that they are spectrally similar. In this region, wetlands usually have low depth and allow the development of vegetation, leading to a confusion with acicular forestry and even with grassland areas.

Table 2 - Confusion matrix of the summer image RF classification without textural metrics, with eight sample classes.

Class	Agriculture	Water	Sand	Wetland 1	Wetland 2	Grassland	Forest	Exposed Soil	Kappa
Agriculture	15	0	0	0	0	0	1	0	
Water	0	116	0	0	0	0	0	1	
Sand	0	0	153	0	0	1	0	0	
Wetland 1	0	0	0	66	0	1	1	0	
Wetland 2	0	0	0	1	17	0	0	0	
Grassland	0	0	0	3	0	95	0	6	
Forest	1	1	0	0	0	1	345	0	
Exposed Soil	0	0	0	0	0	3	1	72	
Kappa									0.9783

Elaboration: The authors (2022).

Then, in the second classification with three classes, Table 3, where the Forest class was reclassified into Native, Adult Pine, and Young Pine, there was little confusion between Native and Adult Pine and vice versa, while the Young Pinus class had much more confusion between the other classes, mostly among the Adult Pine class. This is justified by the fact that they are the same forest species and are practically the same spectrally speaking.

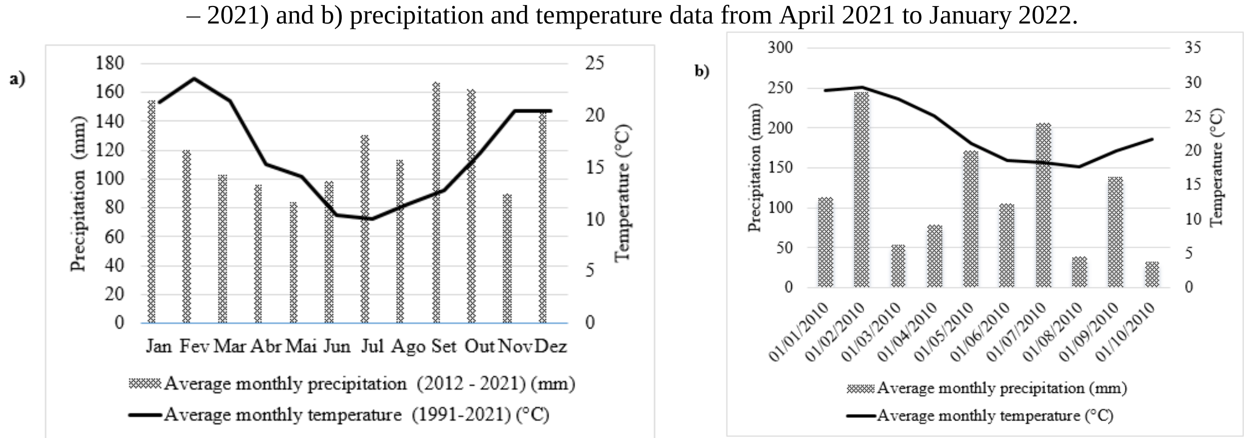
Table 3 - Confusion matrix of the summer image RF classification without textural metrics, with three sample classes.

Class	Native Forest	Adult Pine	Young Pine	Kappa
Native Forest	151	14	13	
Adult Pine	8	186	30	
Young Pine	7	16	26	
Kappa				0.6749

Elaboration: The authors (2022).

Considering the seasonal variations of precipitation and temperature for the region for the period with images (Figure 5 b), we verified that the highest monthly accumulated precipitation occurs in the months of August and September. As the temperature slowly begins to increase in August, vegetation begins to sprout causing vegetation greening, in which we expected that there would be greater confusion between vegetation classes; however, this was not the case, the spring season, which begins in September, was the one that obtained the best results in the classification. This is because the forestry and the native forest do not suffer as much from the cold as from the fields and wetlands.

Figure 5 – Climograms for the study area, showing a) precipitation and temperature data for a period of 10 years (2012 – 2021) and b) precipitation and temperature data from April 2021 to January 2022.



Source: Inmet (2022).

Observing the confusion matrices of the SVM classifier using three classes, Table 5, it classified only four training samples as Young Pine, generating a lower precision in the classification. However, accuracies were similar, as reported by the models presented in Breunig et al. (2020). When visually analyzing the maps presented in Figure 6, referring to the classifications in the different seasons of the year, we observed that they are visually similar on a large scale. When observing the details of each class, it appears that the SVM classifier

 DOI: http://dx.doi.org/10.14393/rbcv75n0a-67769

confused the Young Pine with Adult Pine classes (arrow indicated in Figure 6), and both classifiers in the spring and summer seasons confused Young Pine with Native Forest. Still, through the visual interpretation it can be identified through the RGB image that these are the Young Pine and Adult Pine classes, which can be proven by the confusion matrix, Table 4.

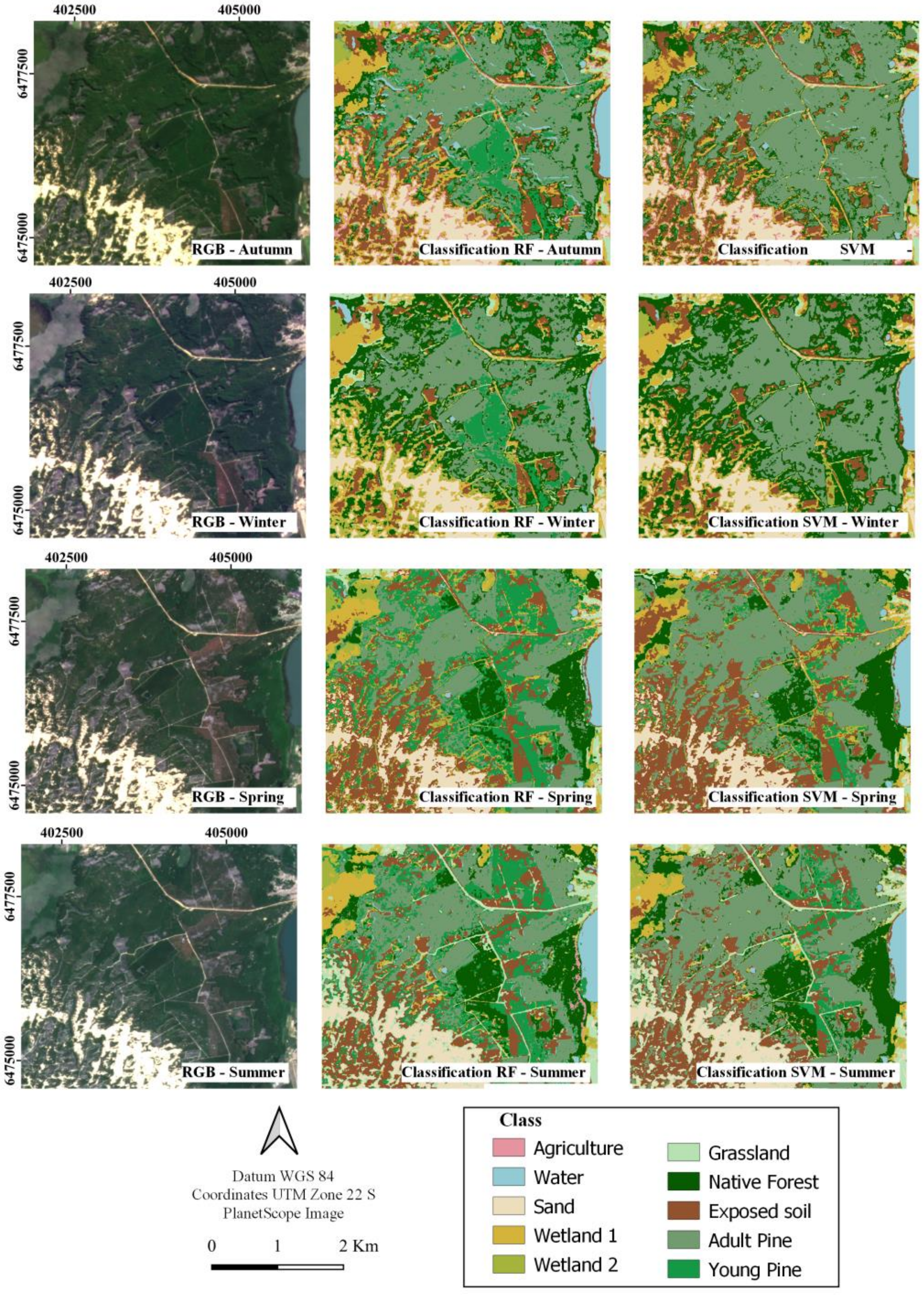
Considering the results associated with commercial forestry areas, we observed that both classifiers were able to map with high accuracy the large forest masses existing in the study area.

Table 4 - Confusion matrix of the summer image SVM classification without textural metrics, with three sample classes.

Class	Native Forest	Adult Pine	Young Pine	Kappa
Native Forest	152	15	14	
Adult Pine	14	201	51	
Young Pine	0	0	4	
Kappa				0.6333

Elaboration: The authors (2022).

Figure 6 – Results of classification with the Random Forest (RF) and Support Vector Machine (SVM) model for the Autumn, Winter, Spring, and Summer seasons, using PlanetScope multispectral images.



Source: The authors (2022).

Studies such as that by Fagan et al. (2018), Souza et al. (2016), Xie et al. (2019) and Banskota, Wynne and Kayastha (2011) used other variables, such as texture, geometric variables, lidar data, time series, and spectral variables from hyperspectral images, using narrower bands to map different species through supervised classifiers. The adjustment of different mapping methods varies with the environmental situation, with the input data attributes and with the specific application requirements of map use (SCHULZ et al., 2021). The use of high-resolution PlanetScope images and the RF algorithm provided high accuracy for mapping land use and land cover in the region of São José do Norte, RS. The use of images from summer and spring are advised to obtain the best results using images from a single date. The use of multiple dates as predictors for the classification process could be explored in future studies.

The SVM and RF classifier models operate from Machine Learning and are generally the most popular and with relatively high performance, being similar in this regard (SHEYKHMOUSA et al., 2020). In the study by Adam et al. (2014), in which they used high spatial resolution images, both RF and SVM were unable to fully deal with the high spectral variation, due to landscape heterogeneity. In this context, the PlanetScope images, with a spatial resolution of 3.7 m and considering the smaller spectral variance in the study area, these algorithms allowed to achieve high accuracy and Kappa index.

Pine plantations are very important for the region's economy, as seen by its representativeness in land use and cover, covering around 20% of the area (Table 5). Considering the spatial distribution, the main areas of native forest are found around wetlands 1 and 2, around the lagoon and in consortium with the areas and native grassland. The region identified as exposed soil is largely associated with areas prepared for short cycle plantings (e.g., Onion) or to regions that went through the forest harvesting process. Thus, PlanetScope images, associated with the use of machine learning algorithms, allowed the quantification and qualification of land use and land cover in the region of São José do Norte, RS. The developed methodology can be used in conjunction with multi-temporal PlanetScope images for forest management practices on tracking the development of new plantations, their growth over time changing from young to adult stages, and until the trees are felled.

Table 5 - Relationship of area and coverage percentage of the different land use and land cover classes for 10/28/2021.

Land use and land cover class	Area (Km^2)	Area (ha)	Area (%)
Agriculture	2.27	226.97	0.55
Water	41.51	4,150.98	10.14
Sand	37.94	3,794.19	9.27
Wetland	62.62	6,262.31	15.30
Green Wetland	29.22	2,922.31	7.14
Grassland	85.29	8,528.73	20.83
Native Forest	23.06	2,306.54	5.63
Exposed Soil	43.31	4,331.09	10.58
Adult Pine	42.02	4,201.06	10.26
Young Pine	42.13	4,212.80	10.29
Total	**409.38**	**40,936.98**	**100.00**

Elaboration: The authors (2022).

4 CONCLUSION

The results of the classification using the RF and SVM models demonstrate that Planet Scope's high-resolution multispectral images allow obtaining high accuracy rates, even when mapping a heterogeneous and complex landscape. High accuracy values were achieved in both classifiers and in all seasons. The Random Forest algorithm obtained results statistically greater for the Spring and Summer seasons than Support Vector Machine, while the Autumn and Winter seasons did not have a statistically significant difference. The spring and summer seasons showed better discrimination between classes.

The highly accurate classification data obtained in this study contribute with reliable information about the extension of the main land classes located in the municipality of São José do Norte and the developed methodology can be used for further environmental monitoring, forestry management, and landscape planning.

Acknowledgements

We thank Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) (FAPERGS project – UFSM/FW) for the master's scholarship. We also thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) (process 305084/2020-8) for the financial support. We thank Planet Labs for providing the PlanetScope images (Education and Research Standard – 503533). We are grateful to the FLOPAL Model Plantation project registered on the supporting technology foundation – FATEC under the number 9.33.0002 for providing the facilities for the conduction of this experiments and data analysis.

Authors Contribution

This article was prepared based on the contributions of all authors. A. K. F.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization. F. M. B.: Conceptualization, Acquisition of financing, Project administration, Resources, Supervision. R. B.: Acquisition of funding, Resources, Supervision, Writing - review and editing. R. D. S.: Formal analysis, Software, Supervision, Writing - review and editing.

Conflict of Interest

The authors have no conflicts of interest to declare. The funders did not interfere in the development of the study, in the analysis or interpretation of the data, in the writing of the manuscript, or in the decision to publish the results. This article is part of the study of the dissertation project of the first author, Andressa Kossmann Ferla, developed under Fábio Marcelo Breunig and presented in 2023 at the Postgraduate Program in Environmental Science and Technology (UFSM).

References

ADAM, E.; MUTANGA, O.; ODINDI, J.; ABDEL-RAHMAN, E. M. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers. **International Journal of Remote Sensing**, v. 35, n. 10, p. 3440–3458, 2014. DOI. 10.1080/01431161.2014.903435.

ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; DE MORAES GONÇALVES, J. L.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2013. DOI. 10.1127/0941-2948/2013/0507.

BANSKOTA, A.; WYNNE, R. H.; KAYASTHA, N. Improving within-genus tree species discrimination using the discrete wavelet transform applied to airborne hyperspectral data. **International Journal of Remote Sensing**, v. 32, n. 13, p. 3551–3563, 2011. DOI. 10.1080/01431161003698302.

BEHERA, M. D.; BARNWAL, S.; PARAMANIK, S.; et al. Species-Level Classification and Mapping of a Mangrove Forest Using Random Forest—Utilisation of AVIRIS-NG and Sentinel Data. **Remote Sensing**, v. 13, n. 11, p. 2027, 2021. DOI. 10.3390/rs13112027.

BELGIU, M.; DRĂGUŢ, L. Random forest in remote sensing: A review of applications and future directions. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 114, p. 24–31, 2016. DOI. 10.1016/j.isprsjprs.2016.01.011.

BREUNIG, F. M. Classification of soybean varieties using different techniques: case study with Hyperion and sensor spectral resolution simulations. **Journal of Applied Remote Sensing**, v. 5, n. 1, p. 053533, 2011. DOI. 10.1117/1.3604787.

BREUNIG, F. M.; GALVÃO, L. S.; DALAGNOL, R.; DAUVE, C. E.; et al. Delineation of management zones in agricultural fields using cover–crop biomass estimates from PlanetScope data. **International Journal of Applied Earth Observation and Geoinformation**, v. 85, p. 102004, 2020. DOI. 10.1016/j.jag.2019.102004.

BREUNIG, F. M.; GALVÃO, L. S.; DALAGNOL, R.; SANTI, A. L.; et al. Assessing the effect of spatial resolution on the delineation of management zones for smallholder farming in southern Brazil. **Remote Sensing Applications: Society and Environment**, v. 19, p. 100325, 2020. DOI. 10.1016/j.rsase.2020.100325.

CHEN, L.; LI, S.; BAI, Q.; et al. Review of Image Classification Algorithms Based on Convolutional Neural Networks. **Remote Sensing**, v. 13, n. 22, p. 4712, 2021. DOI. 10.3390/rs13224712.

CHUVIECO, E. Fundamentos de teledection espacial. 2.a ed. Madrid: RIALF SA, 1990

CONGALTON, R. G. Accuracy assessment and validation of remotely sensed and other spatial information. **International Journal of Wildland Fire**, v. 10, n. 4, p. 321, 2001. DOI. 10.1071/WF01031.

CORTES, C.; VAPNIK, V. Support-vector networks. **Machine Learning**, v. 20, n. 3, p. 273–297, 1995. Disponível em: <http://link.springer.com/10.1007/BF00994018>.

DHINGRA, S.; KUMAR, D. A review of remotely sensed satellite image classification. **International Journal of Electrical and Computer Engineering (IJECE)**, v. 9, n. 3, p. 1720, 2019. DOI. 10.11591/ijece.v9i3.pp1720-1731.

FACELI, K; LORENA, A. C; GAMA, J.; ALMEIDA T.A.; et al. **Inteligência artificial**: uma abordagem de aprendizado de máquina. 2.a ed. Rio de Janeiro, 2021.

FAGAN, M. E.; MORTON, D. C.; COOK, B. D.; et al. Mapping pine plantations in the southeastern U.S. using structural, spectral, and temporal remote sensing data. Remote Sensing of Environment, v. 216, p. 415–426, 2018. DOI. 10.1016/j.rse.2018.07.007.

FERNÁNDEZ-DELGADO, M.; CERNADAS, E.; BARRO, S.; AMORIM D. Do we need hundreds of classifiers to solve real world classification problems? **The journal of machine learning research**, v. 15, n. 1, p. 3133-3181, 2014.

FOODY, G. M. Status of land cover classification accuracy assessment. **Remote Sensing of Environment**, v. 80, n. 1, p. 185–201, 2002. DOI. 10.1016/S0034-4257(01)00295-4.

GIANUCA, K. S.; TAGLIANI, C. R. A. Análise em um Sistema de Informação Geográfica (SIG) das alterações na paisagem em ambientes adjacentes a plantios de pinus no Distrito do Estreito, município de São José do Norte, Brasil. **Revista da Gestão Costeira Integrada**, v. 12, n. 1, p. 43–55, 2012. Disponível em: <http://scielo.pt/scielo.php?script=sci_arttext&pid=S1646-88722012000100005&lng=pt&nrm=iso>. ISSN 1646-8872.

GRABSKA, E.; HOSTERT, P.; PFLUGMACHER, D.; OSTAPOWICZ, K. Forest Stand Species Mapping Using the Sentinel-2 Time Series. **Remote Sensing**, v. 11, n. 10, p. 1197, 2019. DOI. 10.3390/rs11101197.

HADDAD, I.; GALVÃO, L. S.; BREUNIG, F. M.; et al. On the combined use of phenological metrics derived from different PlanetScope vegetation indices for classifying savannas in Brazil. **Remote Sensing Applications: Society and Environment**, v. 26, p. 100764, 2022. DOI. 10.1016/j.rsase.2022.100764.

HILL, R. A.; WILSON, A. K.; GEORGE, M.; HINSLEY, S. A. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. **Applied Vegetation Science**, v. 13, n. 1, p. 86–99, 2010. DOI. 10.1111/j.1654-109X.2009.01053.x.

HUETE, A.; DIDAN, K.; MIURA, T.; et al. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote Sensing of Environment**, v. 83, n. 1–2, p. 195–213, 2002. DOI. 10.1016/S0034-4257(02)00096-2.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Cidades de Estados (São José do Norte)**. Disponível em: https://www.ibge.gov.br/cidades-e-estados/rs/sao-jose-do-norte.html. Acesso em: 23 junho 2022.

INSTITUTO NACIONAL DE METEREOLOGIA (INMET). **Banco de Dados Meteorológicos**. Disponível em: https://bdmep.inmet.gov.br. Acesso em: 23 junho 2022.

JANSSEN, L. LF; VANDERWEL, F. JM. Accuracy assessment of satellite derived land-cover data: a

 DOI: http://dx.doi.org/10.14393/rbcv75n0a-67769

review. **Photogrammetric Engineering and Remote Sensing.** v. 60, n. 4, 1994.

JENSEN, J. R. **Sensoriamento Remoto do Ambiente**: uma perspectiva em recursos terrestres. 2. ed. São José dos Campos: Parêntese, 2011.

KUHN, M. Caret: classification and regression training. **Astrophysics Source Code Library**, p. ascl: 1505.003, 2015.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, p. 159-174, 1977.

LARY, D. J.; ALAVI, A. H.; GANDOMI, A. H.; WALKER, A. L. Machine learning in geosciences and remote sensing. Geoscience Frontiers, v. 7, n. 1, p. 3–10, 2016. DOI. 10.1016/j.gsf.2015.07.003.

LEITE, A. P.; SANTOS, G. R.; SANTOS, J. É. O. Análise temporal dos índices de vegetação NDVI e SAVI na Estação Experimental de Itatinga utilizando imagens Landsat 8. **Revista Brasileira de Energias Renováveis**, v. 6, n. 4, 2017. DOI. 10.5380/rber.v6i4.45830.

LI, M.; ZANG, S.; ZHANG, B.; LI, S.; WU, C. A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information. European Journal of Remote Sensing, v. 47, n. 1, p. 389–411, 2014. DOI. 10.5721/EuJRS20144723.

LUO, H.; LI, M.; DAI, S.; et al. Combinations of Feature Selection and Machine Learning Algorithms for Object-Oriented Betel Palms and Mango Plantations Classification Based on Gaofen-2 Imagery. **Remote Sensing**, v. 14, n. 7, p. 1757, 2022. DOI. 10.3390/rs14071757.

MEYER, D. et al. Package ‘e1071’. **The R Journal**, 2019.

MONNET, J.-M.; CHANUSSOT, J.; BERGER, F. Support Vector Regression for the Estimation of Forest Stand Parameters Using Airborne Laser Scanning. **IEEE Geoscience and Remote Sensing Letters**, v. 8, n. 3, p. 580–584, 2011. DOI. 10.1109/LGRS.2010.2094179.

MOUNTRAKIS, G.; IM, J.; OGOLE, C. Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, v. 66, n. 3, p. 247–259, 2011. DOI. 10.1016/j.isprsjprs.2010.11.001.

NOVO, E. M. L. de Moraes. **Sensoriamento Remoto**: princípios e aplicações. 4. ed. São Paulo: Blucher. p. 387, 2010.

OK, A. O.; AKAR, O.; GUNGOR, O. Evaluation of random forest method for agricultural crop classification. **European Journal of Remote Sensing**, v. 45, n. 1, p. 421–432, 2012. DOI. 10.5721/EuJRS20124535.

PAL, M.; FOODY, G. M. Feature Selection for Classification of Hyperspectral Data by SVM. **IEEE Transactions on Geoscience and Remote Sensing**, v. 48, n. 5, p. 2297–2307, 2010. DOI. 10.1109/TGRS.2009.2039484.

PERKO, R.; RAGGAM, H.; SCHARDT, M.; MICHAEL ROTH, P. Very High Resolution Mapping with the Pléiades Satellite Constellation. **American Journal of Remote Sensing**, v. 6, n. 2, p. 89, 2018. DOI. 10.11648/j.ajrs.20180602.14.

PINHEIRO, R. M.; DA SILVA, M. D. Paisagens Ameaçadas Da Restinga Da Lagoa Dos Patos (Rs): Ecologia Da Paisagem Como Contribuição Para O Zoneamento Ecológico Econômico Do Litoral Médio / Threatened Landscapes Of The Restinga Da Lagoa Dos Patos (Rs): Landscape Ecology As Contribution To Th. **Geographia Meridionalis**, v. 4, n. 2, p. 269, 2019. DOI. 10.15210/gm.v4i2.14483.

PLANET LABS. 2019. “Planet Surface Reflectance Product.” San Francisco: Planet Labs. Disponível em: https://assets.planet.com/marketing/PDF/Planet_Surface_Reflectance_Technical_White_Paper.pdf. Acesso em: 23 junho 2022.

ROUSE JR, J. W. et al. Paper a 20. In: **Third Earth Resources Technology Satellite-1 Symposium: The Proceedings of a Symposium Held by Goddard Space Flight Center at Washington, DC on**. 1973. p. 309.

R CORE TEAM (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

SCHULZ, D.; YIN, H.; TISCHBEIN, B.; et al. Land use mapping using Sentinel-1 and Sentinel-2 time series in a heterogeneous landscape in Niger, Sahel. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 178, p. 97–111, 2021. DOI. 10.1016/j.isprsjprs.2021.06.005.

SHERMAN, Gary. QGIS **- A Free and Open Source Geographic Information System**. Viena: GNU - Free Software Foundation, Inc. Disponível em: <http://www.qgis.org/en/site/>. , 2002.

SHEYKHMOUSA, M.; MAHDIANPARI, M.; GHANBARI, H.; et al. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 13, p. 6308–6325, 2020. DOI. 10.1109/JSTARS.2020.3026724.

SOUZA, C. G.; CARVALHO, L.; AGUIAR, P.; ARANTES, T. B. Algoritmos de Aprendizagem de Máquina e Variáveis de Sensoriamento Remoto para o Mapeamento da Cafeicultura. **Boletim de Ciências Geodésicas**, v. 22, n. 4, p. 751–773, 2016. DOI. 10.1590/S1982-21702016000400043.

SOZZI, M.; KAYAD, A.; GIORA, D.; SARTORI, L.; MARINELLO, F. Cost-effectiveness and performance of optical satellites constellation for Precision Agriculture. Precision agriculture '19. **Anais...** . p.501–507, 2019. The Netherlands: Wageningen Academic Publishers. DOI. 10.3920/978-90-8686-888-9_62.

TAGLIANI, C. R. A.; VICENZ, R. S. Mapeamento da vegetação e uso do solo nos entornos do estuário da Laguna dos Patos, RS, utilizando técnicas de processamento digital de imagem do SIG SPRING. **Anais eletrônicos, XI Simpósio Brasileiro de Sensoriamento Remoto**, Belo Horizonte - MG, p. 1461 – 1468, 2003.

WULDER, M. A.; LOVELAND, T. R.; ROY, D. P.; et al. Current status of Landsat program, science, and applications. **Remote Sensing of Environment**, v. 225, p. 127–147, 2019. DOI. 10.1016/j.rse.2019.02.015.

XIE, Z.; CHEN, Y.; LU, D.; LI, G.; CHEN, E. Classification of Land Cover, Forest, and Tree Species Classes with ZiYuan-3 Multispectral and Stereo Data. **Remote Sensing**, v. 11, n. 2, p. 164, 2019. DOI. 10.3390/rs11020164.

XU, K.; ZHANG, Z.; YU, W.; et al. How Spatial Resolution Affects Forest Phenology and Tree-Species Classification Based on Satellite and Up-Scaled Time-Series Images. **Remote Sensing**, v. 13, n. 14, p. 2716, 2021. DOI. 10.3390/rs13142716.

Biography of the main author

Andressa Kossmann Ferla, born in 1994. She has a degree in Environmental and Health Engineering from the Federal University of Santa Maria campus Frederico Westphalen, and a master's degree in Environmental Science and Technology from the same University. She has experience in Image Classification, Land Use and Cover Map, Fire Risk Map, Environmental Monitoring, Remote Sensing Image Interpretation. She is currently a Fellow of the Institutional Training Program-PCI – INPE in Santa Maria-RS.