Integration of the SAR and optical sensors of the Sentinel constellation for land use classification in Lagoinha (SP)

Ana Carolina Vidal do Nascimento[12], Glauber José de Castro Gava[13], Isabella Clerici De Maria[14], Luciana Alvim Romani[15,] Jener Fernando Leite De Moraes[16]

Abstract

Information on land use and coverage is necessary to assist in the management process and assertive decision-making. Thus, the present study aimed to evaluate the fusion of Sentinel-1 (S1) and Sentinel-2 (S2) data in the mapping of land use and coverage of the municipality of Lagoinha (SP) using the Random Forest method. Three scenarios were tested for classification: data from (S1), (S2) and fusion of (S2+S1). To evaluate the accuracy of the classification, high-resolution images from Google Earth and S2 software were used. The overall accuracy of the classification from the combination of S2+S1 data was 94%, and the Kappa index was equal to 0.9. For the isolated images of S2 and S1, overall accuracies of 80% and 50% and Kappas index of 0.71 and 0.50 were obtained, respectively. The fusion of S1+S2 data showed high accuracy in mapping.

Key words: Remote Sensing; Sensor Fusion; Radar; Machine Learning; Land Use.

1. Introduction

The Sentinel-1 (S1) and Sentinel-2 (S2) satellites, launched by the European Space Agency (ESA), have improved the way in which land use and land cover monitoring and classifications are carried out (Gómez, 2017). Sentinel-1 has an active SAR (Synthetic Aperture Radar) sensor, which is useful for monitoring areas of dense vegetation in adverse weather conditions (Tavares et al., 2019). Sentinel-2 is equipped with a passive optical sensor with 13 multispectral bands, which performs planned analyses of vegetation types and their phenological and physiological conditions (Gava et al., 2024). Several studies have shown that the fusion of multispectral images from S2 with the

---

[12] Scholarship holder at Agronomic Institute. ORCID: 0009-0008-0244-8956. Email: anacavinagro@outlook.com.

[13] Agronomic Institute (IAC). ORCID: 0000-0002-3194-5432. Email: glauber.castro@sp.gov.br.

[14] Agronomic Institute (IAC). ORCID: 0000-0001-8093-1697. Email: isabella.maria@sp.gov.br.

[15] Embrapa Digital Agriculture. ORCID: 0000-0002-7386-3515. Email: luciana.romani@embrapa.br.

[16] Agronomic Institute (IAC). ORCID: 0000-0003-1112-4204. Email: jener.moraes@sp.gov.br

vertical-vertical (VV) and vertical-horizontal (VH) polarization bands, transformed into backscatter coefficients in decibels (dB) from S1, increased the accuracy of land use and land cover classifications (Steinhausen et al., 2018; Tavares et al., 2019). By combining the data from S1 and S2, it was possible to reduce ambiguity or error in classification, since the complementary information from S1 allowed greater differentiation of classes that presented similar spectral characteristics. For example, shadow areas in optical images were better identified in SAR images, allowing for more accurate classification (Tavares et al., 2019). The study of land use and land cover mapping in the Chennai Basin, India, carried out by (Steinhausen et al., 2018) found an overall accuracy of 92% when the images from S1 were combined with S2. Based on the above, the objectives of the present study were to classify and map land use and land cover, using the Random Forest (RF) algorithm in a manner: a) independent; and b) merging the data from the sensors onboard the S1 and S2 satellites, with subsequent evaluation of the accuracy of the products generated.

2. Methods

2.1. Study area

The study area is formed by the municipality of Lagoinha, State of São Paulo, whose central coordinates are latitude 23º05' S and longitude 45º12' W. This municipality has an area of 25,546 hectares (Figure 1).
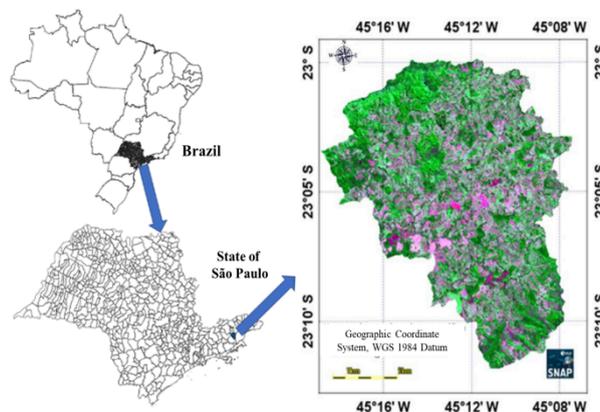

Figure 1. Location of the study area, municipality of Lagoinha (SP) and, RGB composition.

2.2. Data acquisition

The area of interest was constituted by the intersection of SAR (Sentinel-1) and optical (Sentinel-2) images. These images were acquired from ESA's Copernicus portal available at: (https://dataspace.copernicus.eu). The selection of Sentinel 1 and 2 images was based on cloud cover of less than 1% for the S2 images and the smallest date difference for the S1 images. Considering the lowest cloud cover rate, a Sentinel-2A image of level L2A was selected on May 5, 2024 with 0% clouds. The acquired S2 image had 13 spectral bands with spatial resolutions from 10 to 60 meters. The Sentinel-1A scene used was the Interferometric Wide Level 1 Single Look Complex, from May 1, 2024. This S1A image is C-band with two polarizations (VV and VH) and a spatial resolution of 5 x 20 meters. The high-resolution image mosaic, used to calculate the global accuracy and validate the classification, was accessed through the Google Earth software and the S2A image, updated on May 5, 2024.

2.3. Processing flowchart

The methodological procedures, briefly described in Figure 2, follow six steps, namely: a) data acquisition; b) pre-processing of the S1 and S2 images; c) integration of the SAR and optical sensors; d) classification of the training polygons; e) supervised classification using the Randon Forest algorithm; and f) assessment of the overall accuracy of the generated product. For the processing and integration of the S1 and S2 images, the Sentinel Application Platform (SNAP 11) software was used, employing the methodology described in (Tavares et al., 2019).
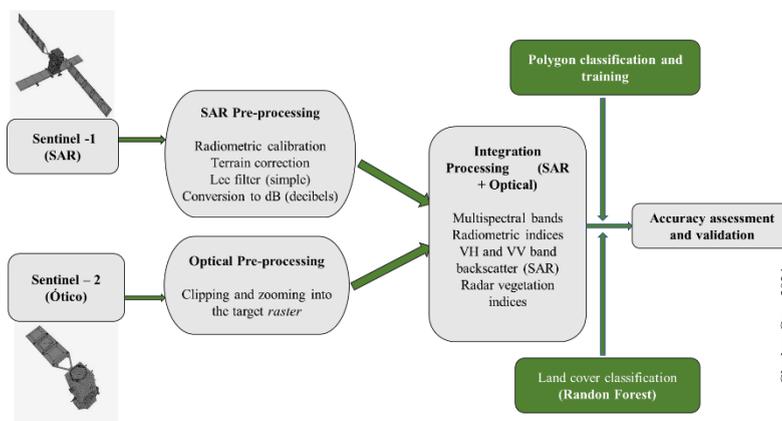

Figure 2. Flowchart of data processing steps.

In the integration (fusion) processing of the S1 and S2 images, the following sequence of operations was used for the S1 raster: a) Subset; b) Radiometric Calibration; c) Conversion to dB scale; and d) Range Doppler Terrain Correction (Figure 2). Finally, the SNAP 11 Collocation algorithm was applied. This last procedure created a stack between the input bands, in which the pixel values of a product (Slave) of S2 are resampled to the geographic raster of the second product (Master) of S1.

2.4. Definition of training classes

In this study, four land use and land cover classes were considered (Figure 3).

Figure 3. Land use classes and their RGB composition
of satellites S1 (SAR) and S2 (optical).

2.5. Classification, accuracy assessment and Kappa index

After the acquisition, pre-processing and integration of the S1 and S2 data (Figure 2), the images produced were exported to the QGIS Desktop software, where the categorization and training (machine learning) stage was performed for 50 polygons from each of the four classes, totaling 200 training areas. Subsequently, the following Python libraries were installed: Remotior Sensus; GDAL; OGR; Numpy; SciPy and Matplotlib on the computer for the correct execution of the supervised classification by Randon Forest with 700 trees, using the Semi Automatic Classification Plugin algorithm according to the protocol described by Congedo (2023). The Randon Forest classification was performed on the stacked images with 10 x 10 m spatial resolution, in three different combinations, namely: a) only the S1A data (with the bands: VV; VH and VV-VH); and b) data from S2A (with bands: B2; B3; B4; B5; B6; B8; B9; B11 and B12) and the fusion of S1A and S2A (with

bands: VV; VH; VV-VH; B2; B3; B4; B5; B6; B8; B9; B11 and B12). The calculation of the global accuracy and the confusion matrix of the classification maps produced were performed using the Accuracy Assessment of Thematic Maps (AcATaMa) algorithm, according to the protocol described in Llano (2024). To determine the models for assessing the global accuracy and the confusion matrix, 1,621 points were generated in stratified random sampling of each class, being 538 points for forest, 955 points for agriculture, 64 points for exposed soil and 64 points for water bodies, within the classified images. These points were validated using Google Earth software images and the RGB image from S2A updated on May 5, 2024 as reference. The Kappa index of each map was determined using the results of the confusion matrices from the three classification evaluations produced, according to the methodology in Cohen (1960).

3. Results and Discussion

The maps of land use and coverage classes (forest, agriculture, exposed soil and hydrography) in the municipality of Lagoinha classified with S1 (SAR), S2 (optical) and S1+S2 are presented in Figure 4.
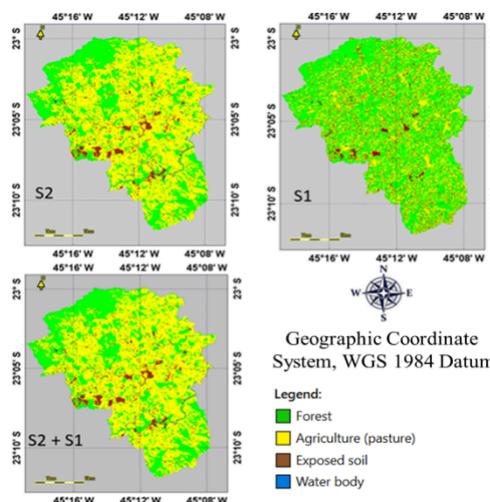

Figure 4. Map of classified image combinations S1 (SAR), S2 (optical) and S1+S2.

Visually, the classifications with the lowest incidence of noise were verified by the maps of S2 + S1 and S2, originating from the integrated information from Sentinel-2A and Sentinel-1A and from the independent data from Sentinel-2A, respectively. The S1 map with independent Sentinel-1A data presented a lot of noise throughout its image, not

separating the forest and agricultural classes well. This fact probably reduced the accuracy of this map when compared to the S2 + S1 and S2 maps. The classification which integrates the data from S2 + S1 presented a good separation and identification of the use and coverage classes, highlighting the contrast for the water bodies (Figure 4).

The results of the global accuracy and Kappa indexes, calculated for the maps produced (Figure 4), indicated that the classification that used only SAR data (map S1) obtained the lowest accuracy and Kappa index (Table 1). The classification that used only optical images (map S2) achieved significantly higher accuracy and Kappa when compared to the classification of map S1. The classification that used the integration of SAR and optical data (S1+S2) achieved the highest global accuracy and Kappa index (Table 1). Other studies obtained results similar to those of this research (Steinhausen et al., 2018; Tavares et al., 2019).

Table 1. Overall accuracy values of the Kappa index of the Randon Forest classification and Kappa quality ranking

Maps	Overall accuracy (%)	Index Kappa	Quality Kappa
S1 + S2	94	0.90	Excellent
S2	80	0.71	Very good
S1	50	0.25	Reasonable

4. Conclusion

The results demonstrated that the integration between SAR and optical data from the S1 and S2 satellites increased the accuracy of the characterization mapping of the Earth's surface, indicating great potential for this technique. The same technique will be applied to map the Jacupiranga and the Alto Alegre DATs. The land use and land cover map generated with greater accuracy for Lagoinha and, for the other DATs, will be the basis for monitoring agricultural crops.

Acknowledgements

References

COHEN, J. A. Coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, v. 20, n. 1, p. 37-46, Apr. 1960. DOI: https://doi.org/10.1177/001316446002000104.

CONGEDO, L. Semi-automatic classification plugin: a python tool for the download and processing of remote sensing images in QGIS. **Journal of Open Source Software**, v. 6, n. 64, 3172, 2021. DOI: https://doi.org/10.21105/joss.03172.

GAVA, G. J. de C.; ARLANCH, A. B.; OLIVEIRA, M. W. de; MARTINS, V. A.; MORAES, J. F. L. de. Monitoramento e estimativa de produtividade dos colmos da cana-de-açúcar por meio de imagens dos satélites Sentinel 2. In: SIMPÓSIO INTERNACIONAL SELPER, Belém, PA, 21., 2024. **Além do dossel - tecnologias e aplicações de sensoriamento remoto: anais.** Belém, PA: UFPA, 2024. 9 p.

GÓMEZ, M. G. C. **Joint use of Sentinel-1 and Sentinel-2 for land cover classification: a machine learning approach.** 2017. 60 p. Dissertação (Mestrado em Ciência da Geoinformação e Observação da Terra para Modelagem e Gestão Ambiental) – Lund University, Suécia.

LLANO, X. C. **SMByC-IDEAM. AcATaMa - QGIS plugin for Accuracy Assessment of Thematic Maps, version 24.10**. 2024. Available at: https://github.com/SMByC/AcATaMa. Accessed on: Apr. 24, 2025.

STEINHAUSEN, M. J.; WAGNER, P. D.; NARASIMHAN, B.; WASKE, B. Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions. **International Journal of Applied Earth Observation and Geoinformation**, v. 73, p. 595-604, 2018. DOI: https://doi.org/10.1016/j.jag.2018.08.011.

TAVARES, P. A.; BELTRÃO, N. E. S.; GUIMARÃES, U. S.; TEODORO, A. C. Integration of Sentinel-1 and Sentinel-2 for classification and LULC mapping in the urban area of Belém eastern Brazilian Amazon. **Sensors**, v. 19, n. 5, 1140, 2019. DOI: https://doi.org/10.3390/s19051140.