IMPACT OF CLOUDS ON FOREST DEGRADATION MEASUREMENT ERRORS USING THE LINEAR SPECTRAL MIXING MODEL IN SENTINEL-2A IMAGERY

Débora Joana Dutra[1], Juan Carlos Guerra[2], Egídio Arai[3], Andeise Cerqueira Dutra[1], Poliana Domingos Ferro[1], Yosio Shimabukuro[1,3], Thales Sehn Körting[1,3], Luiz Eduardo Oliveira e Cruz de Aragão[1,3], Liana Oighenstein Anderson[3,4]

[1]Remote Sensing Postgraduate Program (PGSER), Coordination for Education, Research and Outreach (COEPE), Brazil's National Institute for Space Research (INPE) debora.joana, andeise.dutra, poliana.ferro {@inpe.br}, [2]Division of Impacts, Adaptation and Vulnerabilities (DIIAV), INPE, São José dos Campos, Brazil - juan.carlos {@inpe.br}, [3]Remote Sensing Postgraduate Program (PGSER), Coordination for Education, Research and Outreach (COEPE), Brazil's INPE; Earth Observation and Geoinformatics Division (DIOTG), Earth Sciences General Coordination (CGCT), Brazil's National Institute for Space Research (INPE) - egidio.arai, yosio.shimabukuro, thales.korting, luiz.aragao {@inpe.br}, [4]National Center for Monitoring and Early Warning of Natural Disasters, São José dos Campos - liana.anderson{@cemaden.gov.br}

ABSTRACT

Degradation affects ecological integrity and carbon storage, impacting forest's climate regulation capacity. This study evaluated forest degradation from wildfires and logging in the Gurupi Biological Reserve, eastern Amazonia, using Sentinel-2A imagery and a linear spectral mixing model (LSMM). The model decomposed pixels into vegetation, soil, and shade fractions and was applied to detect degraded areas. However, cloud cover reduces accuracy of detection, especially in areas with high probability of cloud occurrence. The error fraction showed overestimations where image fraction ranged between 50% and 70%. As cloud cover exceeded 50%, both amplitude error and standard deviation increased, highlighting the impact of clouds on LSMM results. These findings stress the need for improved data or alternative methods to enhance degradation monitoring in cloud-prone regions.

Keywords *— Degradation, linear spectral mixing model, cloud cover.*

1. INTRODUCTION

Forest degradation is a phenomenon that, although often less noticeable than deforestation, has significant implications for the ecological integrity and functioning of forest ecosystems. In the Amazon, forest degradation is driven by multiple factors, including selective logging, wildfires, extreme drought, and activities related to land conversion for agriculture and cattle ranching [1]. By altering forest structure, these processes contribute to biodiversity loss, increasing the vulnerability of remaining areas to wildfires, and reducing carbon storage capacity, a critical factor to mitigate climate change [2].

As the world's largest tropical forest, the Amazon rainforest is important for maintaining global climate balance. Its degradation has significant regional and global implications, leading to the loss of essential ecosystem services, including carbon storage [3].

In this context, identifying and monitoring forest degradation patterns are essential for formulating conservation and sustainable management policies.

Satellite imagery analysis has proven to be a valuable tool for detecting changes in forest cover, allowing the remote and large-scale identification of affected areas. Among the available methodologies, the linear spectral mixing model stands out for its ability to decompose the spectral response of a pixel into different components, such as exposed soil, vegetation, and shade [4]. This approach allows for a detailed analysis of changes in forest structure and facilitates the distinction between different levels of degradation [5].

One of the main challenges in using optical satellite data for forest degradation studies is the frequent cloud cover in tropical regions such as the Amazon. Clouds can obscure imagery, causing errors that affect land cover classification and forest degradation detection [6]. They can also distort pixel spectral signatures, complicating the Linear Spectral Mixing Model (LSMM) decomposition process and potentially misclassifying forest components [7].

In this study, we utilized the LSMM to address two objectives: (a) to delineate and quantify the soil, vegetation, and shade fractions in the Gurupi Reserve region associated with forest degradation, and (b) to investigate the influence of varying cloud cover percentages on the accuracy of these measurements. The outlined objectives seek to evaluate the model's capability in identifying land cover components and to assess the impact of cloud cover on the precision of the results

2. MATERIAL AND METHODS

2.1. Study Area

The Gurupi Biological Reserve is a conservation area in Maranhão state, Brazil (Figure 1), and is considered one of the last significant portions with significant preservation of the eastern Amazon Rainforest.

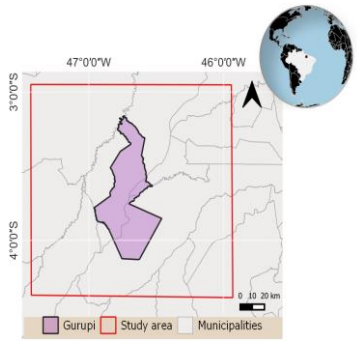


Figure 1 - Study Area

2.2.Methods

2.2.1 Selection of Sentinel-2A Level 2 Data for the Dry Season of the Study Area

Sentinel-2 satellite images were selected during the dry season, from August to October 2020, when conditions are more prone to fires (Figure 2). The images were filtered for cloud cover below 10%, and a median filter was applied to generate a temporal composite, smoothing the series and reducing noise [8].

2.2.2 Selection of Endmembers for the Linear Spectral Mixing Model

Endmembers represent spectral signatures of different components in the image. Endmembers were selected from Sentinel-2A images based on prior knowledge of the study area, in combination with spectral reflectance analysis and extraction algorithms [9].

2.2.3 Application of the Linear Spectral Mixing Model (LSMM)

The LSMM was applied to Sentinel-2 images, using visible and infrared bands within the Google Earth Engine (GEE) platform to decompose each pixel into a linear combination of the selected endmembers [8]. This method quantifies the fraction of each component, enabling the identification of potentially burned areas, water bodies, and cloud shadows (Eq. 1).

$$r_i = \sum_{j=1}^{m} x_j a_{i,j} + e_i \quad \text{(Eq. 1)}$$

where, r_i = spectral reflectance of the pixel for the i-th spectral band; $a_{i,j}$ =spectral reflectance of the j-th component in the pixel for the i-th spectral band; x_j =value of the j-th element in the pixel; e_i = error for the i-th spectral band; j = 1,2, ..., m (m = number of components considered in the pixel);

2.2.4 Comparison with DETER data

The Sentinel-2A cloud probability product was employed to evaluate the influence of cloud probability on the fraction images derived from the LSMM. Cloud probability thresholds were used to categorize the data, assessing how different levels of cloud cover affect the accuracy of fraction estimates. Additionally, DETER data served as ground truth to identify regions affected by fires, logging degradation, and deforestation. The analysis compared cloud probability and fraction values within these identified areas to determine the extent of the impact on the model's accuracy.

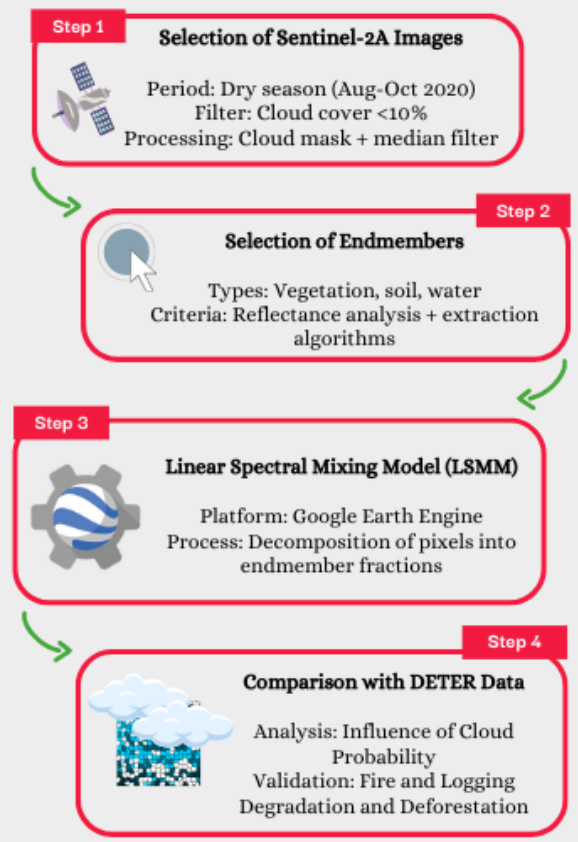


Figure 2 - Flowchart of the applied methodology

3. RESULTS

Figure 3 shows that the greatest differences between the chosen endmembers and the theoretical curve occur in bands 11 and 12 of Sentinel 2A, which correspond to the shortwave infrared regions of the soil endmember's spectral curve. However, when compared with the curve of the OLI sensor proposed by Shimabukuro and Ponzoni (2021) [4], the behavior of the targets in the corresponding spectral bands of both sensors shows a similar pattern.

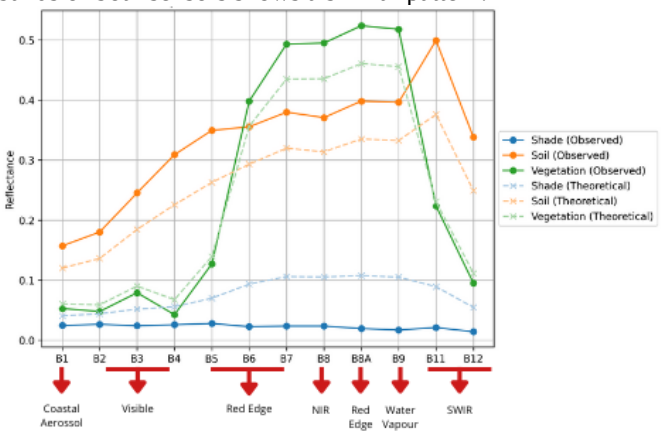


Figure 3 - Endmember curve

The error fraction (Figure 5) presented overestimations in areas where image fraction is between 50% and 70%. These results were primarily observed in regions characterized by

mixings of shade/soil and vegetation/soil components. Conversely, the vegetation/shade mixing presented underestimates (with negative values). The main factor contributing to these results in the study region is the presence of clouds, with probabilities of occurrence ranging from 67% to 83%.

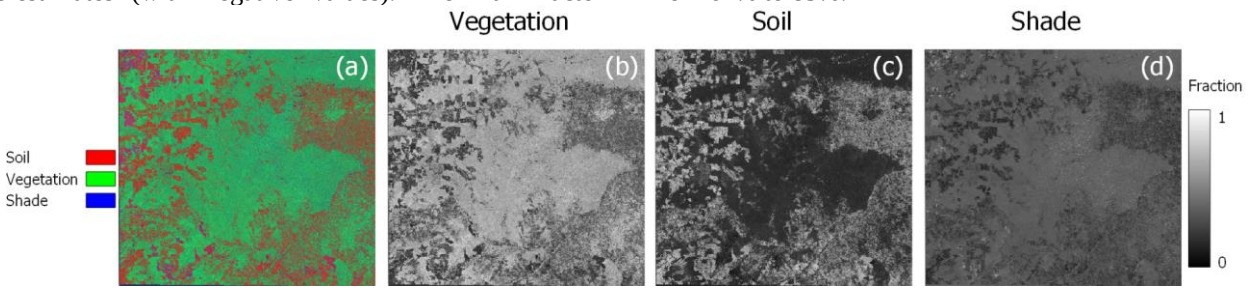


Figure 4 - (a) Linear Spectral mixing Model, (b) Vegetation; (c) Soil; (d) Shade fraction

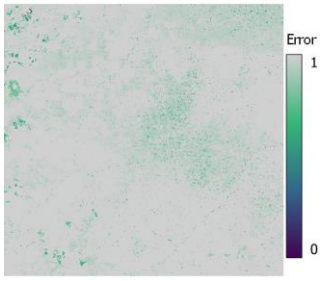


Figure 5 - Linear Spectral Mixing Model Error

The plots illustrate how cloud cover affects fractional cover estimates for vegetation, soil, and shade across five disturbance classes: Burned Scar, Selective Logging, Deforestation, Degradation, and Disorderly Selective Logging (Figure 6). As cloud cover exceeds 50%, the amplitude error and standard deviation increase significantly, indicating that cloud presence introduces greater variation in the LSMM results. This effect is particularly notable in Deforestation and Selective Logging, where higher variability is observed due to the complexity of their spectral characteristics. The estimates show lower variability due to clouds when their cover is below 50%, with relatively stable results across all classes

Figure 7 shows (a) a cloud probability map, while Panels (b), (c), and (d) depict the changes in fractional cover for vegetation, soil, and shade, respectively, as cloud cover increases. The graphs reveal a consistent decrease in fractional values, particularly at higher cloud percentages. Panel (e) of Figure 7 summarizes the impact of cloud cover on each disturbance class, showing significant reductions in vegetation and soil fractions between 60-80% cloud cover, with shade dropping to near zero at higher levels of cloud cover. This highlights the increased difficulty in interpreting fractional estimates as cloud presence grows.

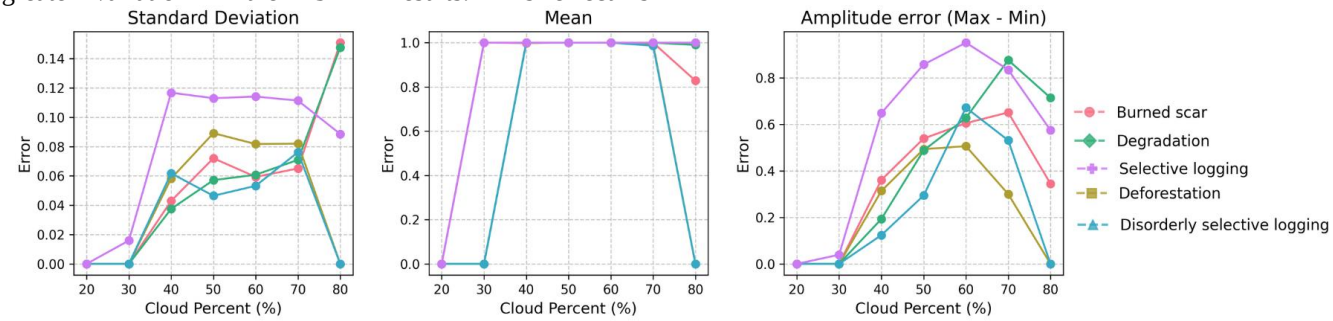


Figure 6 -Standard deviation, mean, and amplitude of error estimates for different DETER classes as a function of cloud cover percentage

4. DISCUSSION

Overall, the selected endmembers respected the expected theoretical curve suggested by Shimabukuro et al. 2020 [10]. Regarding the fraction images (Figure 4a), the selected endmembers could portray the vegetation, soil, and shade regions, similar to Ferro et al. 2024 [11]. Figure 4b shows the regions of Amazonian vegetation in the study region, covering the Gurupi reserve region. Figure 4b highlights the soil component, enabling the identification of croplands, pasturelands, and roads.

These results demonstrate the substantial impact of increasing clouds beyond a 10% cover threshold on vegetation, soil, and shade fraction detection. The vegetation fraction is the most affected, showing the greatest percentage decrease as cloud cover increases. In contrast, the soil fraction exhibits a moderate decline, and the shade fraction shows more complex variability [9]. These findings highlight the necessity of considering cloud cover when interpreting satellite images to improve the accuracy of environmental analyses by Shimabukuro et al. (2023)[10].

5. CONCLUSIONS

This study applied the linear spectral mixing model to Sentinel-2A imagery in the Gurupi Biological Reserve, identifying key pixel component fractions such as vegetation, soil, and shade. However, the quality of the

input data significantly compromised the model's accuracy, particularly due to cloud cover, which led to overestimations and underestimations in certain areas. These limitations stemmed from the data rather than the method and affected the model's ability to capture spectral signatures related to fires, logging, and deforestation. The results highlight the need for improved data acquisition or pre-processing techniques to manage cloud cover interference better and accurately represent the complexity of land cover changes, thus enhancing the effectiveness of forest degradation monitoring and conservation efforts in critical regions like the Amazon.

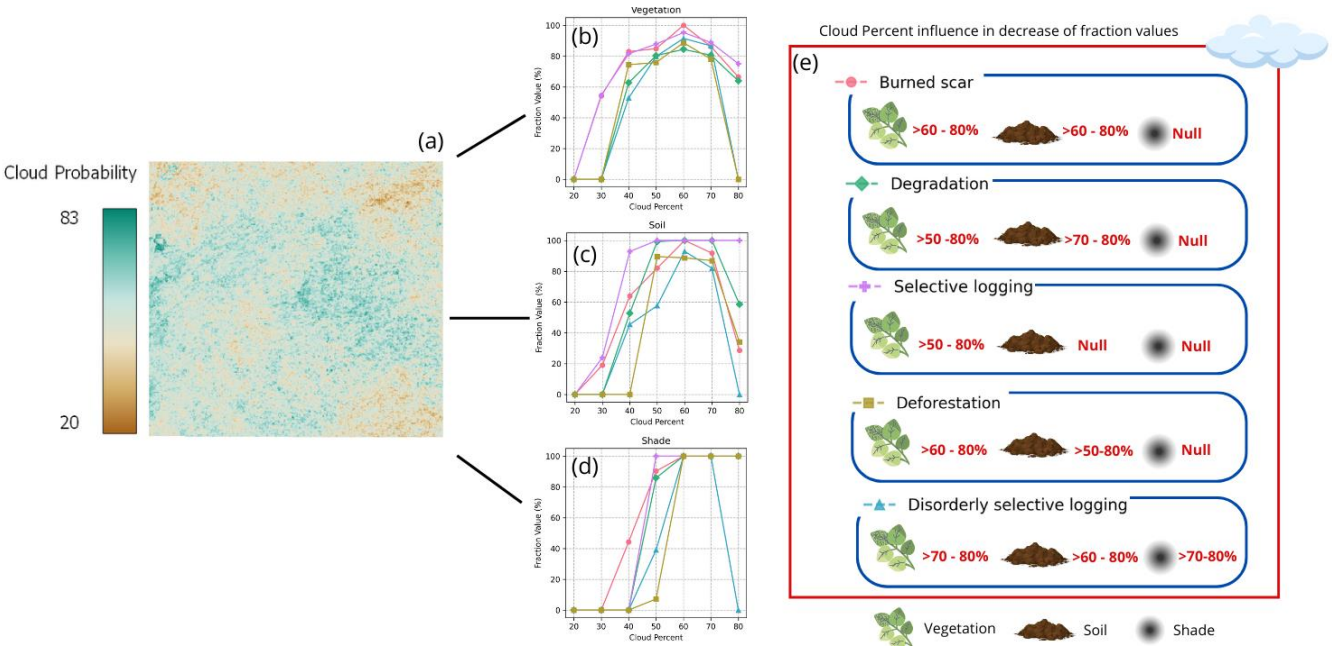


Figure 7 -Influence of cloud cover on fractional cover estimates across various disturbance classes. (a) Cloud probability maps range from 20% to 83%, with green tones indicating lower probabilities and brown tones representing higher probabilities. (b) vegetation, (c) soil, and (d) shade fractional cover values decrease as cloud cover increases, with distinct classes represented by different colored lines: burned scar (pink), degradation (green), selective logging (purple), deforestation (brown), and disorderly selective logging (blue), (e) Summary of the effect of cloud cover on fractional values, indicating significant declines in vegetation and soil fractions at 60-80% cloud cover, and near-zero shade fractions under high cloud cover conditions.

6. ACKNOWLEDGEMENTS

DJD thanks Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and Process 88887.948259/2024-00

7. REFERENCES

[1] D. M. Lapola *et al.*, "The drivers and impacts of Amazon forest degradation", p. eabp8622, Jan. 2023.

[2] L. E. O. C. Aragão *et al.*, "21st Century drought-related fires counteract the decline of Amazon deforestation carbon emissions", *Nat Commun*, vol. 9, nº 1, p. 536, fev. 2018

[3] L. E. O. C. Aragão, C. H. L. S. Junior, e L. O. Anderson, "Brazil's challenge to restrain deforestation and fires in the Amazon during COVID-19 pandemic in 2020: Environmental, social implications and their governance", 2020

[4] Y. E. Shimabukuro e F. J. Ponzoni, *Modelo linear de mistura espectral*. Oficina de Textos, 2021.

[5] Y. E. Shimabukuro, A. C. Dutra, e E. Arai, "Modelo Linear de Mistura Espectral: Conceitos Teóricos, Algoritmos e Aplicações em Estudos na Amazônia Legal", *Rev. Bras. Cartogr.*, vol. 72, p. 1140–1169, dez. 2020

[6] G. P. Asner, "Cloud cover in Landsat observations of the Brazilian Amazon", *International Journal of Remote Sensing*, vol. 22, nº 18, p. 3855–3862, Jan. 2001

[7] M. Xu, M. Pickering, A. J. Plaza, e X. Jia, "Thin Cloud Removal Based on Signal Transmission Principles and Spectral Mixture Analysis", *IEEE Trans. Geosci. Remote Sensing*, vol. 54, nº 3, p. 1659–1669, Mar. 2016

[8] A. Shah *et al.*, "Comparative analysis of median filter and its variants for removal of impulse noise from grayscale images", *Journal of King Saud University - Computer and Information Sciences*, vol. 34, nº 3, p. 505–519, Mar. 2022

[9] Y. E. Shimabukuro e J. A. Smith, "The least-squares mixing models to generate fraction images derived from remote sensing multispectral data", *IEEE Trans. Geosci. Remote Sensing*, vol. 29, nº 1, p. 16–20, Jan. 1991

[10] Y. E. Shimabukuro *et al.*, "Assessment of Burned Areas during the Pantanal Fire Crisis in 2020 Using Sentinel-2 Images", *Fire*, vol. 6, nº 7, p. 277, Jul. 2023.

[11] P. D. Ferro *et al.*, "Regional-Scale Assessment of Burn Scar Mapping in Southwestern Amazonia Using Burned Area Products and CBERS/WFI Data Cubes", *Fire*, vol. 7, nº 3, p. 67, fev. 2024.