Assessing the potential of Google satellite embeddings for mapping sugarcane in Brazilian production areas

Grazieli Rodigheri[1], João Paulo da Silva[1], Jayme Garcia Arnal Barbedo[2], Luciana Alvim Santos Romani[2].

[1] Postdoctoral research at Embrapa Digital Agriculture, Av. Dr. André Tosello, 209 - Cidade Universitária, Campinas - SP, 13083-886, grazieli.rodigheri@colaborador.embrapa.br, joao-paulo.silva@colaborador.embrapa.br

[2] Researcher at Embrapa Digital Agriculture, Av. Dr. André Tosello, 209 - Cidade Universitária, Campinas - SP, 13083-886, luciana.romani@embrapa.br, jayme.barbedo@embrapa.br

A paper from the Proceedings of the

17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture

13-15 July 2026

Porto Alegre, Brazil

Abstract.

The increasing availability of remote sensing (RS) data with higher spatial resolution, combined with advances in artificial intelligence (AI), has been an essential tool in driving the development of the agricultural sector, such as precision agriculture (PA). Among the most relevant information derived from these approaches, crop mapping plays an essential role in crop monitoring, management strategies and yield forecasting. However, the large amount of data required for training classification models, as well as the complexity of agricultural systems and the demand for computing power, are still factors that can limit the performance of models and their application in large and varied areas. On this point, advances in foundation models (FMs) have opened new possibilities in data processing and analysis. FMs are generalist AI models, pre-trained on large volumes of unlabeled data, which can be adapted to a wide range of downstream tasks. Recently, Google DeepMind released the AlphaEarth Foundations, an AI model that provides ready-for-analysis geospatial embeddings integrating multi-source spatial and temporal information, the Google Satellite Embedding dataset. Although considered highly promising for earth observation applications, few studies have investigated the ability and efficiency of using this dataset in cropland areas classification. Thus, the objective of this study was to evaluate the performance of Google satellite embeddings dataset in agricultural areas classification, focusing on mapping sugarcane in the municipality of Alto Alegre - SP, Brazil. This municipality was chosen due to its role as one of the study regions of the Center for Science in Digital Agriculture Development (Semear Digital), funded by Fapesp. The first step consisted of collecting and analyzing land use and land cover samples for the study area, related to the 2023/2024 crop season. Three different approaches were evaluated for mapping crops: (a) a conventional approach, using spectral bands and vegetation and texture indices compositions; (b) the use of Google satellite embeddings dataset; and (c) the combination of both strategies. For the conventional approach, the Sentinel-2 satellite image collection was filtered to cover the study period (July 2023 to September 2024) *and cloud-masked. Subsequently, vegetation indices were calculated to extract the spatial features of satellite imagery for crop mapping and, based on them, a segmentation process was performed using the Simple Non-Iterative Clustering (SNIC) algorithm. Temporal aggregations were generated for the entire feature space and texture indices were calculated for start and peak of the crop season. Since it is a ready-to-analyze dataset, the Google image embeddings were averaged for the years 2023 and 2024 and used directly in the classifier. The accuracy of the classifications was assessed based on confusion matrix analysis, overall accuracy (OA), and comparison with the MapBiomas classification. The use of satellite embeddings provides smoother classifications, with higher accuracy (OA = 0,9820), better field edge delineation and optimized computational performance. These findings provide an important contribution to the crop mapping area, highlighting the potential of new approaches that combine remote sensing and artificial intelligence to provide more accurate and easily retrievable estimates.*

Keywords.
Machine learning, Vegetation indices, Foundation Models, Sentinel-2, Agriculture.

Introduction

Accurate crop classification maps are key information for advancing precision agriculture. This data can assist crop development monitoring, severity weather impact detection (Chaves et al., 2025), management strategies, enhance crop yield and support public policy formulation in the agricultural sector. Several studies have developed increasingly accurate mapping techniques (Yuan et al., 2023; Qu et al., 2024). This progress was mainly due to advances in remote sensing technology, the availability of sensors with better resolutions, combined with the development of user-friendly cloud-based platforms, which led to greater access of Analysis Ready Data (ARD) and easy-to-use machine learning algorithms (Goelick et al., 2017; Tassi and Vizzari, 2020). However, the large amount of data required to train classification models, as well as performance in real-world scenarios and high computing power for large-scale applications, are still challenges that compromise the accuracy of classification maps (Meshram et al., 2021). Therefore, there are still several gaps in crop mapping that need to be investigated in order to generate more reliable and easily obtainable estimates.

Overall, most land use and land cover classifications use a combination of multiple spectral bands, vegetation indices and climate information to construct the feature space used by machine learning algorithms (Prudente et al., 2022; Vizzari, Lesti, Acharki, 2024). Although effective, this approach requires preprocessing steps to generate cloud-free mosaics and reduce noise in microwave measurements (Zeng et al., 2020; Filipponi, 2019). In most cases, this strategy results in a lack of information about phenological dynamics (Zeng et al., 2020) and transient events that downstream models require.

Recent advances in Artificial Intelligence (AI), particularly the emergence of Foundation Models (FMs), have opened up new possibilities to agriculture studies (Li et al., 2024, Yin et al., 2025). FMs are pre-trained on large volumes of unlabeled data and learn general-purpose representations that can be adapted to a wide range of downstream tasks (Jakubik et al., 2023; Chang et al., 2025), such as classification and change detection. Based on these models, Google DeepMind introduced AlphaEarth Foundations, an AI model that provides ready-for-analysis geospatial embeddings (Brown et al., 2025). These embeddings integrate spatial and temporal information from multiple sources across each pixel's annual trajectory, offering a richer and more contextualized representation than traditional isolated spectral indices. Despite its strong potential for environmental and agricultural applications, its practical use and comparative performance remain relatively unexplored in literature.

Thus, the objective of this study was to evaluate the performance of Google satellite embeddings dataset in classifying agricultural areas, particularly in mapping sugarcane, in the municipality of Alto Alegre - SP, Brazil.

Materials and methods

Study area

The study area covered the municipality of Alto Alegre, located in the northwest of the state of São Paulo, Brazil (Figure 1). The region integrates one of the Agrotechnological Districts (DATs) studied by the Science Center for Development in Digital Agriculture (CCD-AD/Semear Digital), an initiative created by the Brazilian Agricultural Research Corporation (Embrapa) and the São Paulo Research Foundation (FAPESP). This is a multi-institutional initiative that seeks to provide solutions to improve access to digital technologies and services for small farmers in Brazil, focusing on solving real problems in farming (EMBRAPA, 2026).

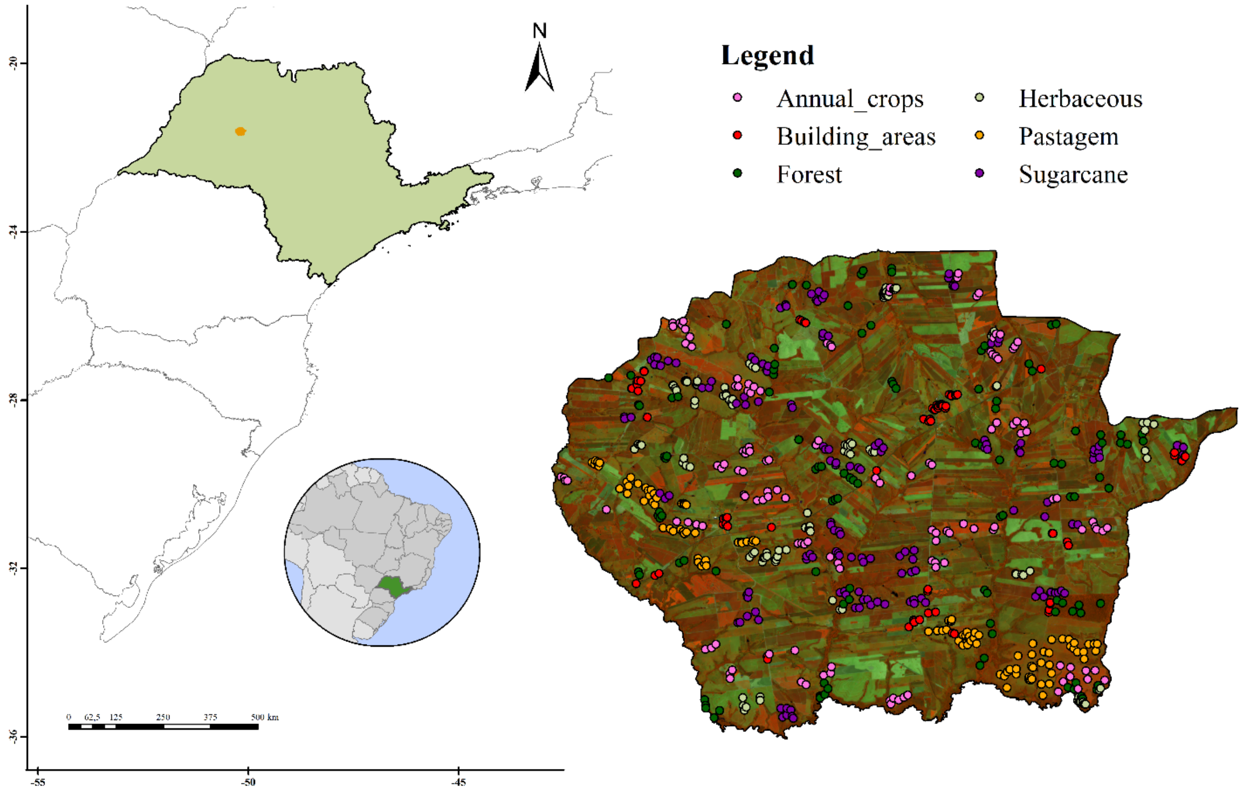

Fig 1. Location of the municipality of Alto Alegre-SP, Brazil, and overview of the selected samples and their corresponding classes for classification purposes.

The study period covered the 2023/2024 crop season to compare the final results with the MapBiomas classification.

Satellite Data

To conduct this study, two different satellite databases were selected: 1) images from the Sentinel-2 satellite; and 2) images from the Google embeddings dataset. Sentinel-2 (S2) satellite images were used both for selecting training and validation samples and for generating the classification. Level-2A surface reflectance images were selected to cover the period from July 2023 to September 2024, with temporal resolutions of 10, 20, and 60 m and a temporal resolution of five days. To ensure cloud-free and higher-quality data, we used the Cloud Score+ dataset (Pasquarella et al., 2023) to identify and remove pixels with clouds and cloud shadows from the image collection.

The Google Embeddings images are generated by the Alpha Earth Foundation Model, a deep neural network pre-trained on large-scale Earth observation data through self-supervised learning. This model compresses each pixel into a compact numerical vector, called an embedding, and stores the results as multi-band GeoTIFF files, where each band corresponds to one dimension of the embedding space (Wu, 2026). This representation enables various spatial analysis, including segmentation, per-pixel classification and change detection.

In this study, Google Embeddings were obtained from Google Earth Engine and used exclusively for the classification process. Since these images are produced annually, the 2023 and 2024 acquisitions were selected and averaged to produce a single composite image, which was then incorporated into the classification model.

Training and Validation Sample Data

The training and validation samples were selected through systematic visual inspection of false color compositions (NIR-SWIR-RED images of S2 images) on different dates and NDVI time-series. This approach took into account the agricultural calendar of different crops and the spectral and spatial patterns corresponding to each class. Reddish, greenish, and bluish spectral tones were considered as indicators of (i) vigorous vegetation, (ii) soil influence, and (iii) the presence of water or bare soil, respectively, according to Chaves et al. (2025).

After sample selection, the sits package (Simoes et al., 2021) was used to extract time series for all sampled points, which were then plotted to evaluate their consistency with the assigned land-use profiles. Analyzing the time series of collected samples, we noticed inconsistencies among built-up areas (urban areas or exposed soil), which were filtered to eliminate potential misclassifications (Figure 2). Despite the elimination of noisy data, all samples remained within a balanced range between 10 and 20% (Table 1).

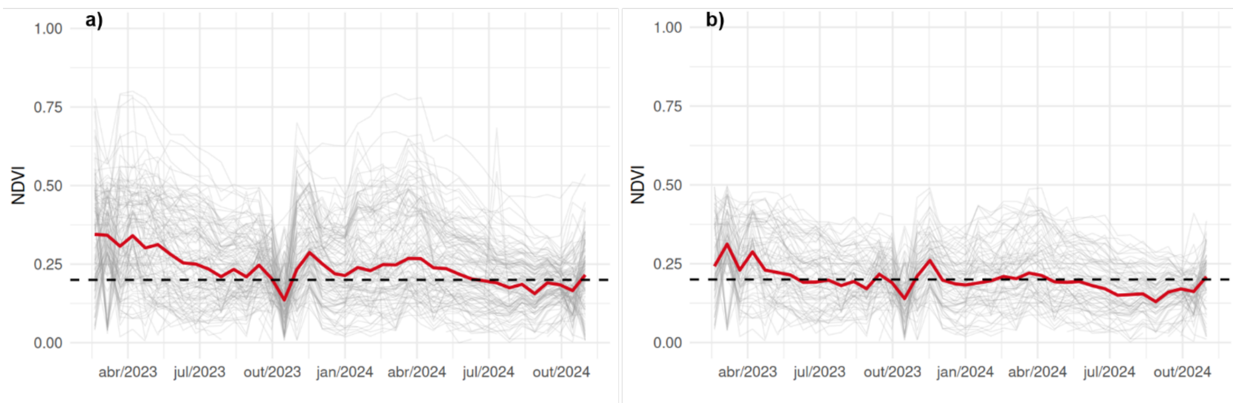

Fig 2. Built-up areas samples before (a) and after (b) noisy elimination.

Table 1. Number and proportion of collected samples

Class	N° Before filtering	N° After filtering	Proportion (%)
Sugarcane	144	144	20.7
Annual Crops	136	136	19.6
Pasture	105	105	15.1
Forests	109	109	15.7
Built-up areas	120	80	11.5
Herbaceous vegetation	121	121	17.4

Methodology

In this study, a comparative analysis was performed on two different datasets to compose a feature space for the classification process: S2 satellite images and Google embeddings dataset. Thus, three different classifications were generated, the first using only data derived from S2 satellite images, named conventional approach, the second using only Google Embeddings images and the third using a combination of both datasets. The methodological workflow presents the main steps carried out in this study, specifically: a) dataset composition and segmentation; (b) classification; and (c) accuracy assessment (Figure 3). All steps were developed using the Google Earth Engine (GEE) cloud processing platform.

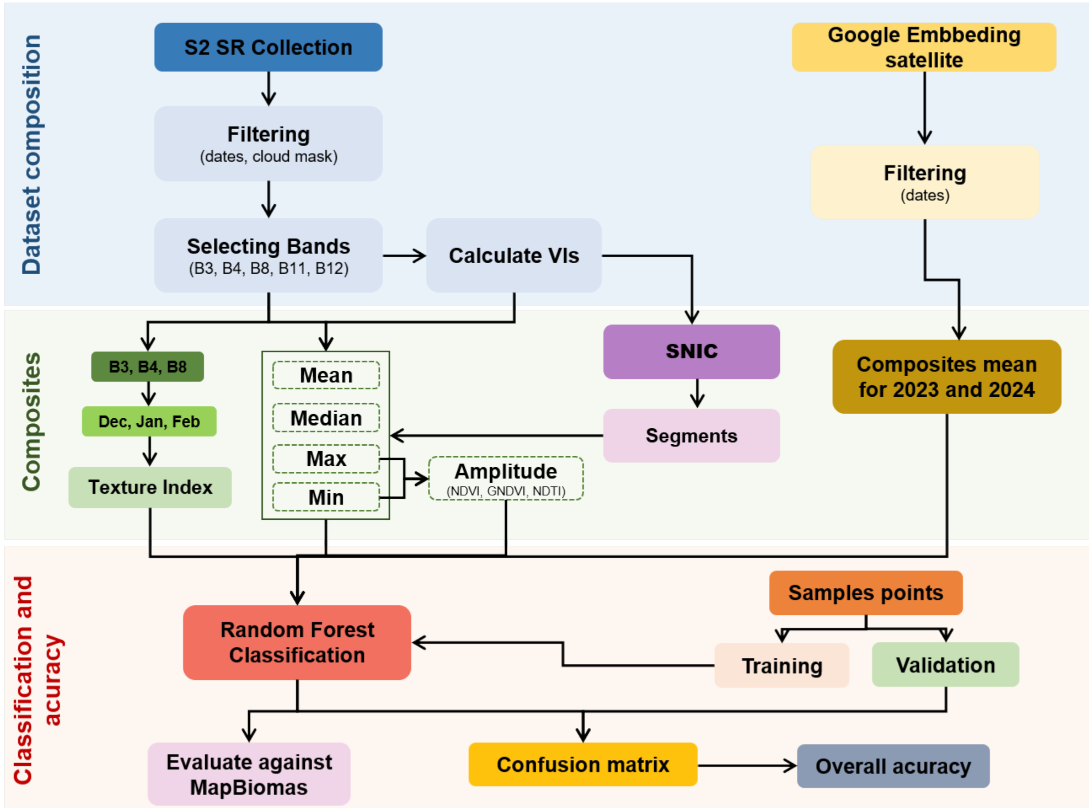

Fig 3. Flowchart showing the methodological steps performed in this study.

Dataset composition

To compose S2 image feature space, the green (B3), red (B4), near-infrared (NIR) (B8), and shortwave infrared (B11 and B12) bands were initially selected, with spatial resolutions ranging from 10m to 20m. From the selected bands, vegetation indices were calculated (Table 2) for all the available images and a segmentation using the Simple Non-Iterative Clustering (SNIC) algorithm was performed. Temporal aggregations were performed for each band, index and average index of each patch generated by SNIC. They included maximum, minimum, mean, and median values calculated quarterly and for the entire time series. In addition, amplitudes were calculated across the whole study period only for vegetation indices.

Table 2. Vegetation indices calculated for S2 images

Index	Equation	Source
Normalized Difference Vegetation Index (NDVI)	(NIR - Red) / (NIR + Red)	Rouse et al., 1974
(NIR - Red)/(NIR + Red)	(NIR - Green) / (NIR + Green)	Gitelson et al., 1996
Normalized Difference Tillage Index (NDTI)	(SWIR1 - SWIR2) / (SWIR1 + SWIR2)	Van Deventer, 1997

Subsequently, texture indices were calculated using the Gray Level Co-occurrence Matrix (GLCM) approach by linearly combining the NIR, Red, and Green bands (Tassi e Vizzari, 2020; Tassi et al., 2021).

$$(0.3 \times NIR) + (0.59 \times RED) + (0.11 \times GREEN) \quad \text{Eq. (1)}$$

The GLCM was applied to three single-date images, one representing the early development and growth stage of the crops and two representing the peak development of the crops in general

(December 27, January 27 and February 26). For GLCM application, the S2 images were rescaled from 0 to 255 (8-bit resolution) and percentiles of 2 and 98 were used to perform histogram equalization and improve textural analysis (Vizzari, Lesti, Acharki, 2024). The window size for textural analysis was set at 3 pixels and five GLCM indices were selected. All components of the feature space have been summarized in Table 3.

Table 3. Vegetation indices calculated for S2 images

Bands and Indices	Metrics	Aggregation	Number of bands
B3, B4, B8, B11, B12	Mean, median, maximum and minimum	Quarterly	100
NDVI, GNDVI, NDTI	Mean, median, maximum, minimum and amplitude	Annual and quarterly	75
NDVI_mean*, GNDVI_mean*, NDTI_mean*	Mean, median, maximum and minimum	Annual and quarterly	72
Texture indices	Contrast, Correlation, Variance, Inverse Difference Moment and Entropy	Three single image	15
Total:			262

*Cluster segments mean

Classification procedure

The classification stage was performed using the Random Forest (RF) classification algorithm. First, a subset of the sample points was selected to train the classifier and another one to validate the classification, considering a 7:3 split ratio, respectively. Then, we use the GEE's “explain” function to construct a graph of feature importance, known as the “mean decrease in Gini impurity”, which is generated as part of the tree-growing routine (Deschamps et al. 2012). Once the importance of the features was analyzed, the classification was performed using a configuration of 100 trees in the RF algorithm.

Validation assessment

To investigate the reliability of classification, the accuracy of the classification was evaluated using the confusion matrix, overall accuracy and F1 score index, as followings equations:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad \text{(Eq. 2)}$$

$$\text{F1 score} = (2 \times \text{precision} \times \text{recall}) / (\text{precision} + \text{recall}) \quad \text{(Eq. 3)}$$

$$\text{precision} = TP / (TP + FP) \quad \text{(Eq. 4)}$$

$$\text{recall} = TP / (TP + FN) \quad \text{(Eq. 5)}$$

where TP, TF, FP, and FN refer to True positive and negative and False positive and negative values from the confusion matrix. Finally, the areas mapped as sugarcane were compared to those mapped by the MapBiomas project. The MapBiomas initiative produces annual land use maps covering the full calendar year (January through December). In addition, its classification considers the purpose of each given area, meaning that areas designated for sugarcane cultivation are mapped as sugarcane even during renovation years (a year typically cultivated with temporary crops). Given this, the years 2023 and 2024 were selected for comparison. As both years yielded nearly identical results, only the 2024 comparisons are presented in the results.

Results and Discussion

Training and Validation samples analysis

Overall, most of the selected samples presented a consistent time series pattern for each class (Figure 4). However, a sharp decline in forest and herbaceous vegetation was observed between October and November 2023. This may be associated with the region's dry season, which typically occurs during the winter (Figure A8, see the Appendix section). Since only two fire hotspots were recorded in the municipality for the year 2023, we discarded the possibility of burning at those locations (INPE, 2026).

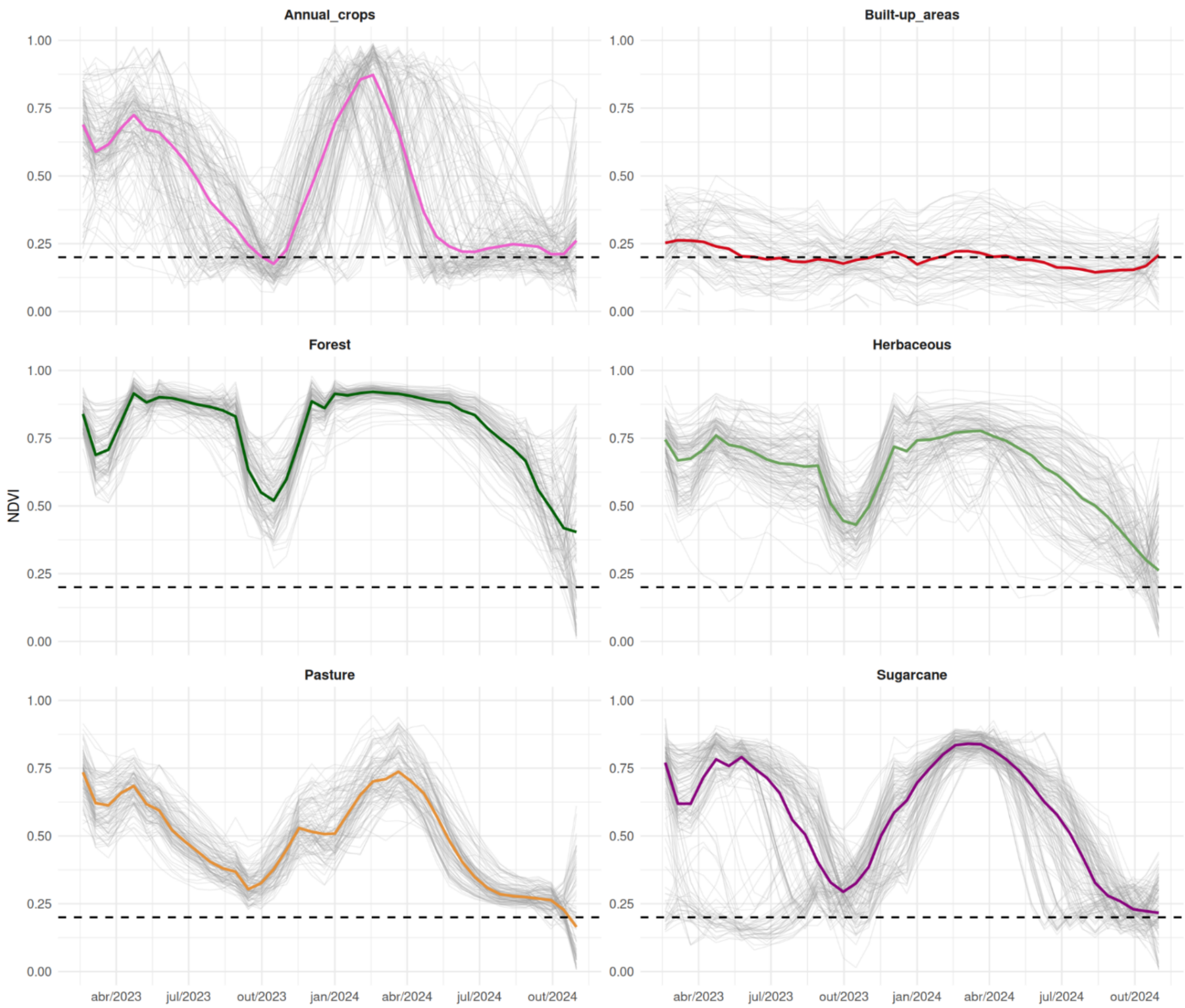

Fig 4. Time series of annual crops, forest, pasture, built-up areas, herbaceous vegetation and sugarcane samples collected for classification.

Classification and accuracy analysis

The importance analysis revealed that most characteristics were significant, except for textural characteristics, as shown in Figure 5. Overall, the SWIR bands (B11 and B12) were the most important characteristics, with most values higher than 1. Notably, only a few bands of the embeddings showed the highest importance among all characteristics (outliers in embeddings boxplot), highlighting the importance of such characteristics in crop classification. Vegetation indices also showed good significance overall, with NDVI having some more important characteristics than other indices (outliers in NDVI boxplot). Considering these results, the texture characteristics were eliminated from the feature space.

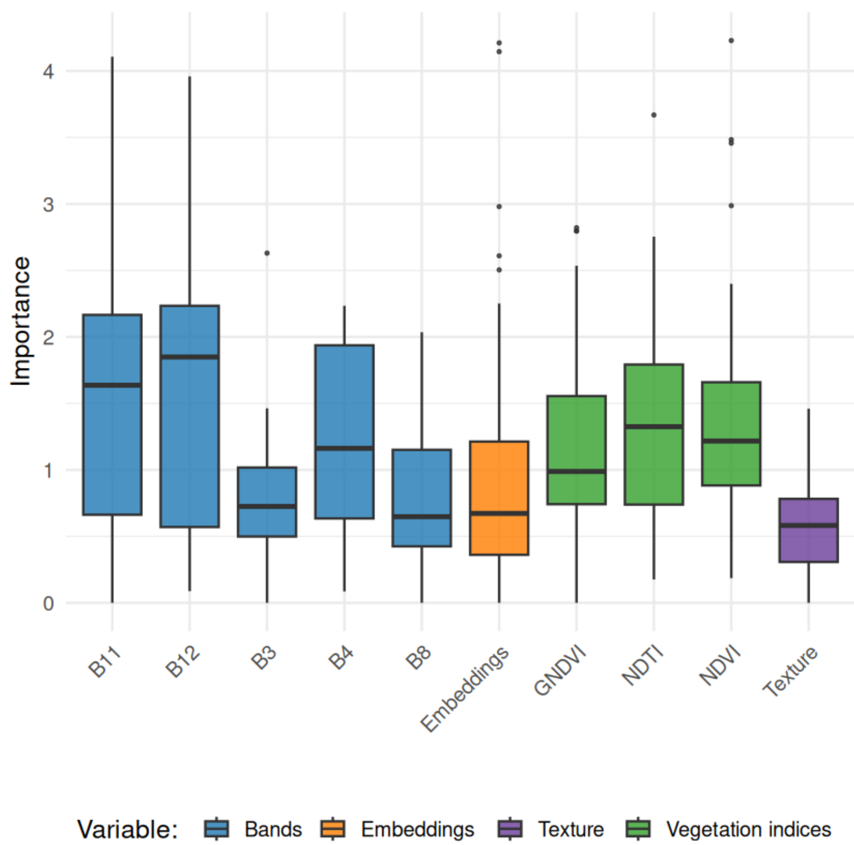

Fig 5. Feature importance for the RF crop classifications.

Although the classifications showed similar levels of accuracy (Table 4), the map using Google Embeddings exhibited maps with well-defined edges and spatial consistency, eliminating the common noise found in other methods (e.g., the salt-and-pepper noise) (Figure 6).

When comparing the mapped areas using different classification approaches against sugarcane areas mapped by MapBiomas, the classification using only embeddings exhibited the highest error (Table 4; Figure 7). This difference is mainly since MapBiomas classifies areas undergoing renovation as sugarcane, even if they are being temporarily used to grow other crops (MapBiomas, 2025).

Table 4. Área e RSME

Approach	Area (ha)	Accuracy	F1 score	MapBiomas corres.* (%)
Bands + IVs	10308,06	0,9819	0,9821	55,77
Google Emb.	8974,19	0,9820	0,9821	49,87
Combination	10479,43	0,9820	0,9821	57,42

*Correspondence with MapBiomas areas indicating areas mapped as sugarcane in both classifications.

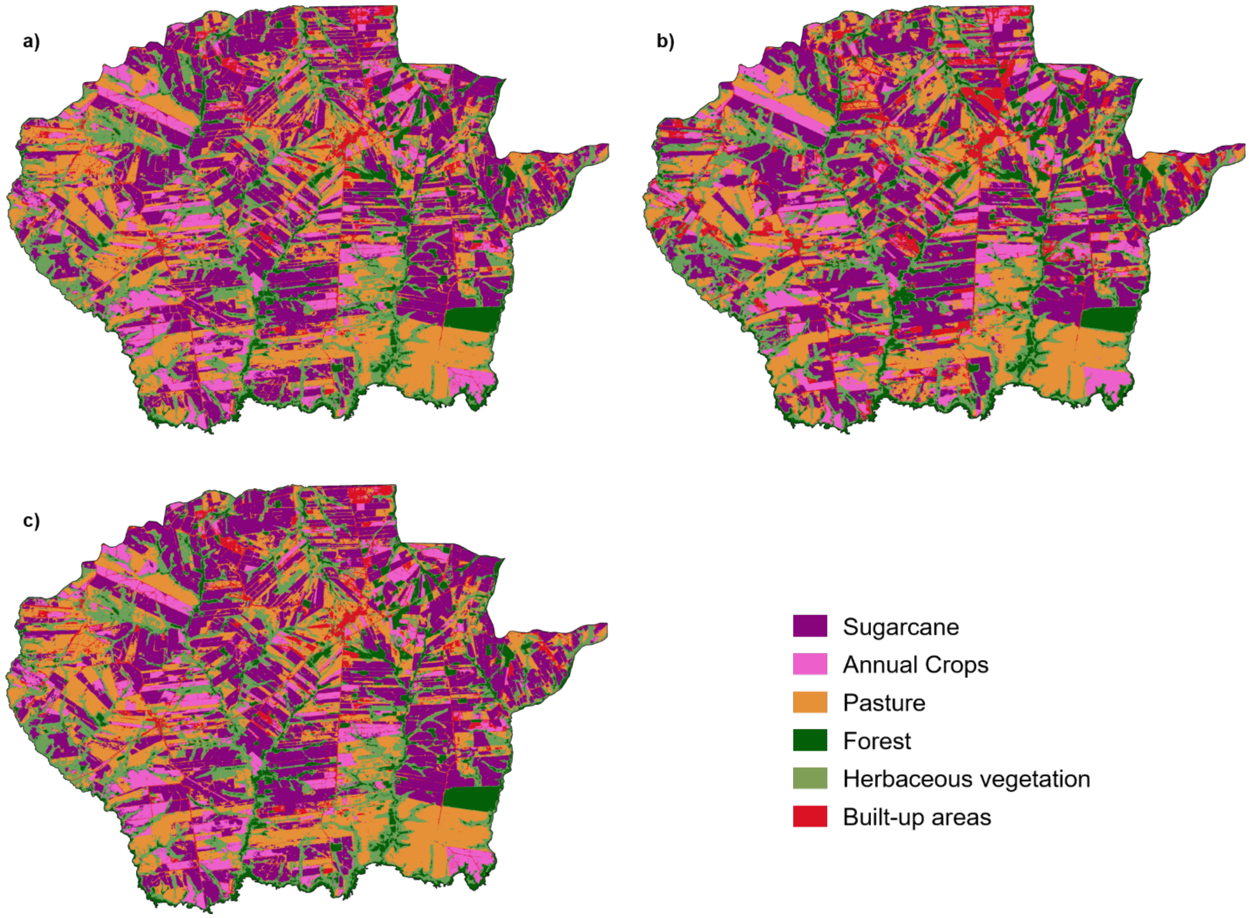

Fig 6. Land use and land cover maps using metrics from bands + IVs (A), only Google Embeddings (B) and metrics + Google Embeddings (C)

Despite the limitations identified, the use of Google Embeddings yielded more consistent and less noisy results, demonstrating its potential for land-use mapping. These vectors encode key surface characteristics, allowing the identification of regions with similar patterns. The main advantage is that, once generated, these features serve as universal, reusable elements that enable simple operations, such as classification, clustering and change detection, without the need to retrain the underlying model. Additionally, further advantages of using such data include automatic data harmonization and ease of acquisition and processing, as it does not require cloud masking or special aggregation (Brown et al., 2025; Ma et al., 2026).

Nevertheless, some limitations remain. These include interpreting the physical meaning of the bands generated by the foundation models and the time lag in the generated embeddings, since they take an entire year to be computed.

As new datasets of satellite image embeddings have become available, future work may evaluate other pre-computed embeddings that also cover the entire globe, such as the Clay Foundation Model (Clay Foundation, 2024) and TESSERA (Feng et al., 2025).

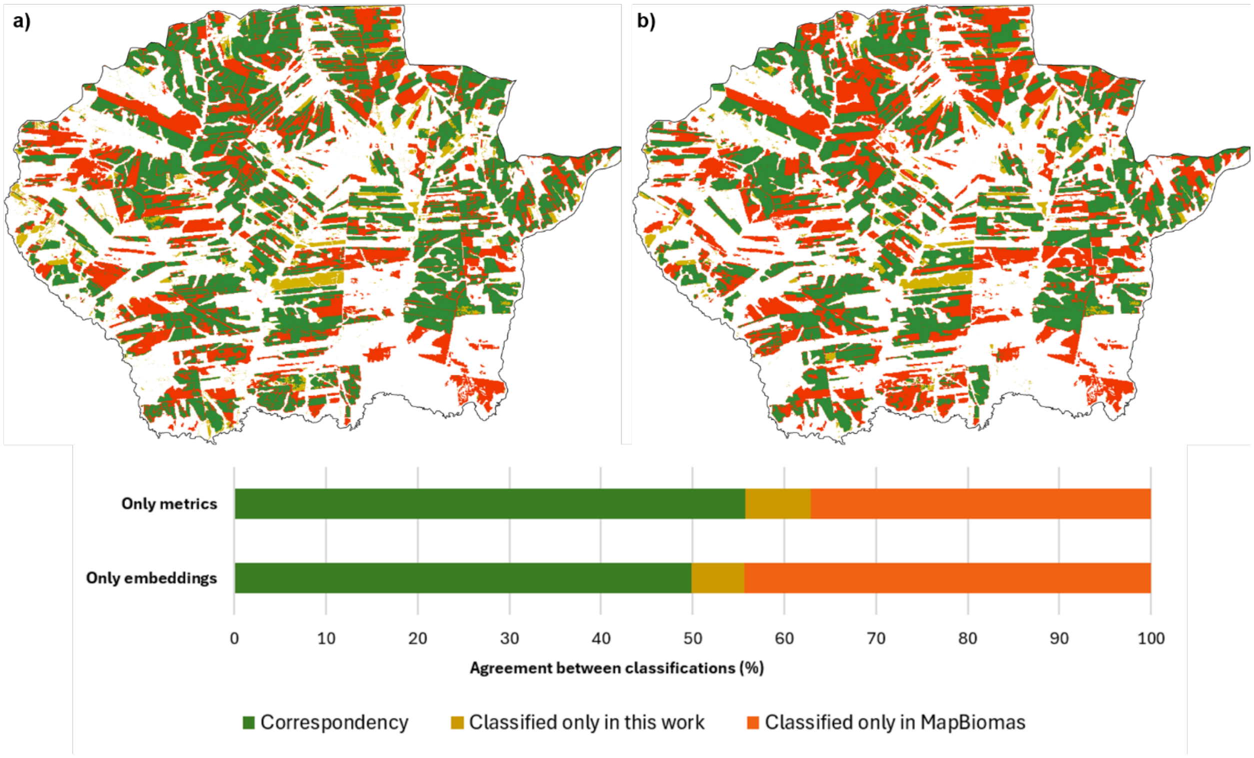

Fig 7. Comparison of sugarcane areas mapped for MapBiomas vs our classification using metrics (A) and only embeddings (B)

Conclusion or Summary

The use of the Google Embeddings Dataset demonstrated better edge delineation and reduced classification noise compared to other approaches, while remaining computationally efficient and scalable for mapping large areas. These representations offer a practical and powerful data source for mapping crops.

However, further research is needed to better interpret the semantic meaning of the vectors produced by the base models. Additionally, the evaluation of alternative pre-computed embeddings that may offer complementary advantages could also be further explored.

Overall, this study's findings contribute to the advancement of agricultural crop mapping methodologies, with direct implications for the advancement of precision agriculture and decision-making support.

Acknowledgments

The authors wish to thank the São Paulo Research Foundation (FAPESP), grants numbers 2022/09319-9 (Semear Digital Project), 25/21967-4 (Postdoctoral scholarship).

Code availability

The code was developed in Google Earth Engine and R language. Thus, we made de scripts available on GitHub (https://github.com/grazirodigheri/sugarcane-mapping-google-embeddings).

References

Brown, C. F., Kazmierski, M. R., Pasquarella, V. J., Rucklidge, W. J., Samsikova, M., Zhang, C., Shelhamer, E., Lahera, E., Wiles, O., Ilyushchenko, S., Gorelick, N., Zhang, L. L., Alj, S., Schechter, E., Askay, S., Guinan, O., Rebecca Moore, R., Boukouvalas, A. & Kohli, P. (2025).

Alphaearth foundations: An embedding field model for accurate and efficient global mapping from sparse label data. https://arxiv.org/pdf/2507.22291

Chang, Y. C., Stewart, A. J., Bastani, F., Wolters, P., Kannan, S., Huber, G. R., Wang, J., & Banerjee, A. (2025, August). On the generalizability of foundation models for crop type mapping. In IGARSS 2025-2025 IEEE International Geoscience and Remote Sensing Symposium (pp. 948-953). IEEE. https://doi.org/10.1109/IGARSS55030.2025.11242260

Chaves, M. E., Adami, M., Oldoni, L., Rodigheri, G., Prudente, V., Santana, C., Garcia, A., Covre, R. M. & Sanches, I. D. (2025). GEEadas: GEE-based Automatic Detection of Adverse-frost Stress. Remote Sensing Applications: Society and Environment, 101799. https://doi.org/10.1016/j.rsase.2025.101799

Clay Foundation (2024). Clay v1.5: An Open Foundation Model for Earth. Available at: clay-foundation.github.io/model.

Deschamps, B., McNairn, H., Shang, J., & Jiao, X. (2012). Towards operational radar-only crop type classification: comparison of a traditional decision tree with a random forest classifier. Canadian Journal of Remote Sensing, 38(1), 60-68. https://doi.org/10.5589/m12-012

EMBRAPA. Semear Digital Project. Available online: https://www.semear-digital.cnptia.embrapa.br/ (accessed on 20 January 2026).

Feng, Z., Atzberger, C., Jaffer, S., Knezevic, J., Sormunen, S., Young, R., Lisaius, M.C., Immitzer, M., Jackson, T., Ball, J., Coomes, D.A., Madhavapeddy, A., Blake, A., & Keshav, S. (2026). Tessera: Temporal embeddings of surface spectra for earth representation and analysis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 34818-34831).

Feng, Z., Atzberger, C., Jaffer, S., Knezevic, J., Sormunen, S., Young, R., Lisaius, M.C., Immitzer, M., Jackson, T., Ball, J. and Coomes, D.A., 2026.

Filipponi, F. Sentinel-1 GRD preprocessing workflow. Multidiscip. Digit. Publ. Inst. Proc. 2019, 18, 11. https://doi.org/10.3390/ECRS-3-06201

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific data, 2(1), 150066. https://doi.org/10.1038/sdata.2015.66

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote sensing of Environment, 58(3), 289-298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote sensing of Environment, 202, 18-27. https://doi.org/10.1016/j.rse.2017.06.031

INPE. Banco de Dados de queimadas. Disponível em: http://www.inpe.br/queimadas/bdqueimadas. Acesso em: 22 de jan. 2026

Jakubik, J., Roy, S., Phillips, C. E., Fraccaro, P., Godwin, D., Zadrozny, B., ... & Ramachandran, R. (2023). Foundation models for generalist geospatial artificial intelligence. arXiv preprint arXiv:2310.18660. https://arxiv.org/pdf/2310.18660

Li, J., Xu, M., Xiang, L., Chen, D., Zhuang, W., Yin, X., & Li, Z. (2024). Foundation models in smart agriculture: Basics, opportunities, and challenges. Computers and Electronics in Agriculture, 222, 109032. https://doi.org/10.1016/j.compag.2024.109032

MapBiomas, 2025, "MapBiomas General “Handbook” - Algorithm Theoretical Basis Document (ATBD) - Collection 10 - Agriculture - Appendix", https://doi.org/10.58053/MapBiomas/DXS8Z9

Meshram, V., Patil, K., Meshram, V., Hanchate, D., & Ramkteke, S. D. (2021). Machine learning in agriculture domain: A state-of-art survey. Artificial Intgence in the Life Sciences. https://doi.org/10.1016/j.ailsci.2021.100010

Prudente, V. H. R., Skakun, S., Oldoni, L. V., Xaud, H. A., Xaud, M. R., Adami, M., & Sanches, I. D. A. (2022). Multisensor approach to land use and land cover mapping in Brazilian Amazon. ISPRS Journal of Photogrammetry and Remote Sensing, 189, 95-109. https://doi.org/10.1016/j.isprsjprs.2022.04.025

Qu, T., Wang, H., Li, X., Luo, D., Yang, Y., Liu, J., & Zhang, Y. (2024). A fine crop classification model based on multitemporal Sentinel-2 images. International Journal of Applied Earth Observation and Geoinformation, 134, 104172. https://doi.org/10.1016/j.jag.2024.104172

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. 1974. In Proceeding Of Erts-1 Symposium.

Simoes, R., Camara, G., Queiroz, G., Souza, F., Andrade, P. R., Santos, L., Carvalho, A., & Ferreira, K. (2021). Satellite image time series analysis for big earth observation data. Remote Sensing, 13(13), 2428. https://doi.org/10.3390/rs13132428

Tassi, A., & Vizzari, M. (2020). Object-oriented LULC classification in Google earth engine combining SNIC, GLCM, and machine learning algorithms. Remote Sens. 12, 3776. https://doi.org/10.3390/rs12223776

Van Deventer, A. P., Ward, A. D., Gowda, P. H., & Lyon, J. G. (1997). Using thematic mapper data to identify contrasting soil plains and tillage practices. Photogrammetric engineering and remote sensing, 63, 87-93.

Vizzari, M., Lesti, G., & Acharki, S. (2025). Crop classification in Google Earth Engine: leveraging Sentinel-1, Sentinel-2, European CAP data, and object-based machine-learning approaches. Geo-Spatial Information Science, 28(3), 815-830. https://doi.org/10.1080/10095020.2024.2341748

Wu, Q. (2026). GeoAI with Python: A Practical Guide to Open-Source Geospatial AI. Independently published. PDF edition ISBN 979-8993859729; Print edition ISBN 979-8253507414. Available at book.opengeoai.org.

Yin, S., Xi, Y., Zhang, X., Sun, C., & Mao, Q. (2025). Foundation models in agriculture: A comprehensive review. Agriculture, 15(8), 847. https://doi.org/10.3390/agriculture15080847

Yuan, Y., Lin, L., Zhou, Z. G., Jiang, H., & Liu, Q. (2023). Bridging optical and SAR satellite image time series via contrastive feature extraction for crop classification. ISPRS Journal of Photogrammetry and Remote Sensing, 195, 222-232. https://doi.org/10.1016/j.isprsjprs.2022.11.020

Zeng, L., Wardlow, B. D., Xiang, D., Hu, S., & Li, D. (2020). A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. Remote Sensing of Environment, 237, 111511.

Appendix

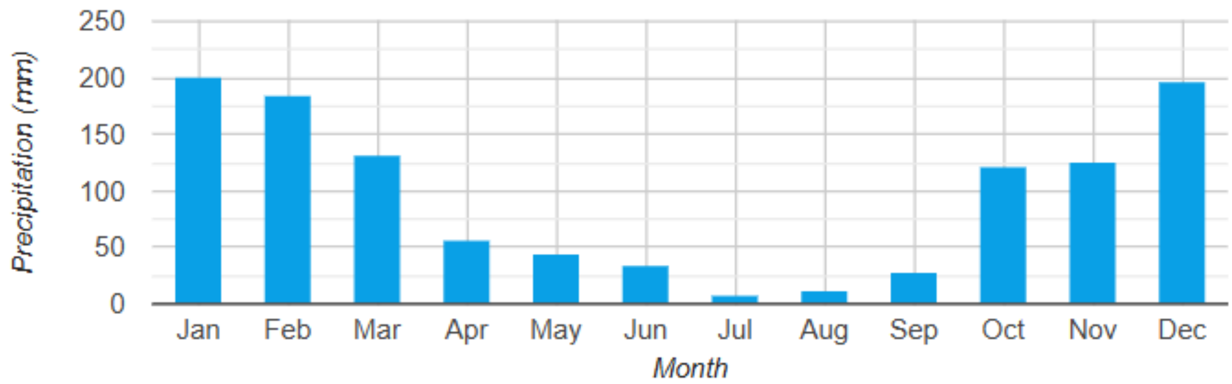

Fig A8. Monthly precipitation for 2023 for Alto Alegre using CHIRPS data obtained from Google Earth Engine dataset. Source: Funk et al., 2026.