COFFEE CROP DETECTION AND MAPPING USING SENTINEL-2 DATA AND SPATIAL COHERENCE

Bruno de Oliveira Schneider
Federal University of Lavras, Department of Applied Computing, Cx. Postal 3037, 307203-202, Lavras, MG, Brazil
Corresponding author: bruno@ufla.br

Marcelo de Carvalho Alves
Federal University of Lavras, Agricultural Engineering Department, Cx. Postal 3037, 307203-202, Lavras, MG, Brazil

Vítor Augusto Ferreira
Federal University of Lavras, Agricultural Engineering Department, Cx. Postal 3037, 307203-202, Lavras, MG, Brazil

Abstract: In this work, we propose a methodology for mapping coffee crops using Sentinel-2 data and Random Forest (RF) classification, focusing on improving accuracy in regions outside the training area. We evaluate the ability of the model to generalize to the same biome, testing its performance in spatially distinct areas near the municipality of Lavras, MG, Brazil. During the experiments, we develop a new technique called spatial coherence, which incorporates information from neighboring pixels into the classification process to reduce salt-and-pepper (isolated) errors. This approach improves traditional RF classification by combining spectral data with spatial context, resulting in more accurate and contiguous coffee crop maps. Our methodology, implemented in R using Sentinel-2 imagery from June 2022, demonstrates the potential to bridge the gap between academic studies and practical large-scale mapping of coffee plantations.

Keywords: Crop mapping, coffee, land use, random forest, supervised classification, ranger R package.

Received: October 18, 2024 - Accepted: May 26, 2025

INTRODUCTION

The coffee sector is a significant part in Brazil's economy. Keeping track of coffee plantations and how it changes over time is useful for establishing public policies and understanding how the sector responds to them, similar to what is been done with deforestation (FINER et al., 2018). Satellite remote sensing data have been used to map coffee plantations and could play a major role in mapping coffee over vast areas and it has been done with reported high accuracy, despite its many challenges. However, accurate, up to date maps are not easy to find, as there seems to be a gap between academic classifications and practical mapping.

A practical problem in crop mapping using supervised machine learning is that training is often done on sampled data from the classification data. Such methodology is impractical because it requires that manual classification from experts and sampling to be tightly coupled, which is hard to do in large areas. We used all data (no sampling) from a manual classification to generate a Random Forest classification model that is used for classification on a different area (data set).

Random Forest classification (BREIMAN, 2001) is very robust for large number of samples and require little to no parameter configuration (SHI; YANG, 2016). It is often used for land use/land cover classification and it has been used before to map coffee (CHEMURA; MUTANGA, 2017; KELLEY; PITCHER; BACON, 2018; BOURGOIN et al., 2020) under different conditions, usually classifying very distinct classes such as water bodies, urban areas, bare soil, and a few distinct crops. Using distinct classes from a spectral point of view, tends to produce good accuracy, but we wanted to test it in coffee mapping. Instead of differentiating coffee from very distinct other land cover, we wanted it to find coffee crops. However, since there are different coffee growing systems with distinct spectral signatures, coffee is often more than just one class, so we also wanted a system that is able to identify more than one coffee class.

Random Forest is a pixel-based classification algorithm and as such, it is prone to produce a noisy classification with *salt-and-pepper* error, that is, false positives and false negatives, that appear isolated in the classification area. Since crops usually extend in relatively large areas, isolated pixels have a high probability of being errors. Data used in this work had a spatial resolution of 10m and a coffee patch of 100m^2 is considered unlikely. Since Random Forest is a machine learning algorithm, we "taught it" to take in consideration the classification of neighbor pixels, therefore reducing the probability of producing small patches in the final classification, and we call this novel feature *spatial coherence*.

This work goes beyond just classifying pixels, it requires less expert effort in generating the classification model while also producing a map that is more object-based than usual output from Random Forest classification.

MATERIALS AND METHODS

We used freely available data from Sentinel-2 together with classification polygons for training and Random Forest classification. Classification polygons that demarcate coffee plantations kindly provided to us from previous mappings from `EMATER/EPAMIG` projects in form of a `shapefile` (.shp, .prj, .shx ... files). This data used to be available from *Geoportal do Café*[1]. Since the classification polygons were 3 years old, polygons in the area of interest were verified and modified by experts were inconsistencies were found.

Sentinel-2 Data

The experiments reported here used data from Sentinel-2 MSI instrument, tile 23KMS, recorded on June 25, 2022. The tile is almost free of clouds and only areas free of clouds within it were used. The data is near the city of Lavras, MG, Brazil, and contains many coffee plantations. Coffee mapping was made by classification using random forests algorithm.

[1] `https://portaldocafedeminas.emater.mg.gov.br/`

Tools used

Classification polygons were verified manually against CBERS-4A satellite images (pan-sharpened 2m resolution from June 2022) using `QGIS`. Images from Google Maps were eventually used in the verification process of less identifiable crops because they had higher resolution than the pan-sharpened images, even though the exact date and resolution of images in Google Maps were undisclosed. Google Maps was accessed during the months of July to December, in 2022.

Processing (training and classification) was done using `R` language version 4.2.2, `Terra` library version 1.7.3 (for geographic raster and vector processing) and `ranger` library version 0.14.1 (random forests implementation).

Study area and Data

Four manually classified regions were produced for the experiments, with extents in WGS 84/UTM zone 23S being:

- training region 1 (TR1): 486700 to 492624 x, 7650110 to 7652390 y (1430.7 ha);
- training region 2 (TR2): 498534 to 500530 x, 7646218 to 7647919 y (363.8 ha);
- training region 3 (TR3): 500298 to 502017 x, 7657617 to 7659499 y (333.2 ha);
- classification region (CR): 504138 to 509636 x, 7648014 to 7650800 y (1587.8 ha).

These regions are around the urban area of Lavras (Figure 1). Some coffee plantations were difficult to verify over satellite images because the plants were too small or had few leaves. We expected these plantations to share little characteristics with other areas, so they were marked as a different class of coffee (Figure 2). 11 months later, these classes were confirmed to be coffee plantations using new satellite images. When mapping coffee, is usually useful to have classification of more than one class, because properties change a lot over the year and different kinds of coffee systems are likely to appear on classification (ESCOBAR-LÓPEZ et al., 2022).

https://doi.org/10.31422/taae.v9i3.64

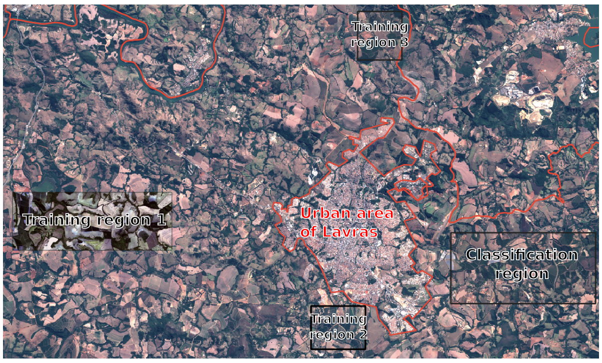

Figure 1: Study areas in the city of Lavras shown over the Sentinel-2 RGB data used. Three training regions (TR1, TR2 & TR3) were used as input data for the creating of Random Forest classification models, while the classification region (CR) was used to validate the models results.

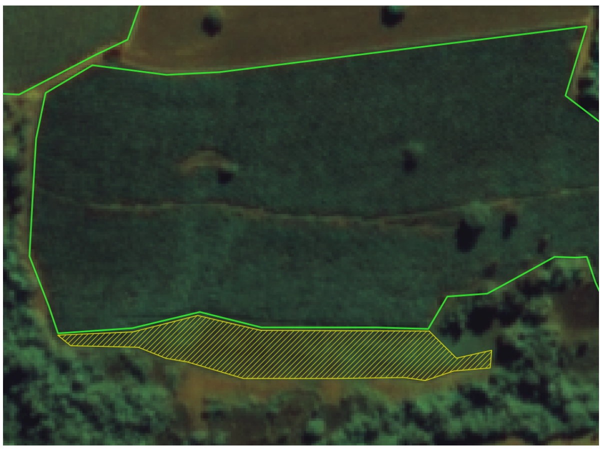

Figure 2: CBERS-4A pan sharpened image (June 21, 2022) showing coffee plantation polygons of both coffee classes: class 1 (hashed yellow) where coffee canopies were not observable in manual classification; class 2 (green) where canopies are visible.

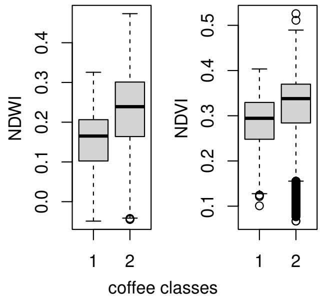

Figure 3: Vegetation indices for the two coffee classes

Coffee polygons contain noise such as big trees inside the plantation, their shadows and farm tracks. They were left as is, for a more practical experiment on coffee mapping than using only pure pixels samples.

We tested *Normalized Difference Vegetation Index* (NDVI) and *Normalized Difference Water Index* (NDWI) vegetation indices on both coffee classes to confirm they would be distinct in the data (Figure 3).

Manual classification

Experts validated and updated coffee classification results from other's previous work by drawing coffee their crop polygons over most recent cloudless CBERS-4A pan sharpened images of the area. Then, the polygons were edited in order to remove previous areas which no longer were coffee crops, as well as add new ones or carefully align crop border with the satellite image where border precision was deemed faulty. In a few cases where the CBERS-4A spatial resolution was deemed insufficient for manual classification, Google Maps (higher resolution imagery) was used to look at the area for decision making. QGIS was used to edit crop polygons.

Road tracks, big trees and other noise found inside coffee crops were left inside the coffee polygons, unless they were on the border and could be easily discarded.

Coffee polygons were labeled according to coffee class (numbered 1 and 2). This vector classification data was later converted to raster data using `Terra`'s `rasterize` function.

```r
# Produce raster classification data from polygons, using resolution and
# extent from band 2, where pixel values come from "Class" field and
# pixels outside polygons get a value of zero.
rasterClasses <- rasterize(polygons, band2, field="Class", background=0)

# Group raster classification data with other raster input data
trainData <- c(rasterClasses, bands, indices)

# Create classification model
model <- ranger(formula, data=trainData, importance="impurity",
                classification=TRUE)
```

Figure 4: Creation of a classification model using `ranger` library, where `formula` is a control string defining input and output labels, `trainData` is the multidimensional labeled raster data from satellite bands, vegetation indices, classification and neighborhood (if available).

Classifier model creation

Using the `ranger` library, labeled data is used to create a classification model (Figure 4). This process could also be described as training the classifier. After a model has been build, it is used for classification (Figure 5).

We created Random Forest classification models using different sets of bands, including the creation of vegetation indices and our new neighborhood data. Classification was done in the classification region except for confirmation of the *out of bag error* (OOB error) which is an error measurement for the Random Forest algorithm. Accuracy measurements in this work are always *producer accuracy*, as opposed to *user accuracy*.

Due to the cost of creating a manual classification, which is needed for training and for computing the accuracy of results, small regions of the multi spectral data from Sentinel-2 were used. Using a localized subset of data of training is considered a valid, effective strategy for classification of geospatial data using machine learning (RAMEZAN; WARNER; MAXWELL, 2019). The training regions 2 and 3 were introduced later in our experiments as a way to provide new, diverse samples and also change proportion of samples of coffee/non coffee classes used for training.

Traditional Random Forest classification

In an initial experiment, the *training region 1* was used to train a model for classification.

```r
classification <- predict(model, data=classifData,
                          type="response")
```

Figure 5: Classification of data using `ranger` library, where `model` is the classification model created previously and `classifData` is the multidimensional labeled raster data from satellite bands, vegetation indices and neighborhood (if available).

All pixels inside the polygons were used as coffee plantation samples (2 classes of coffee), all other pixels were used as “not coffee” class. All 13 Sentinel-2 MSI spectral bands were used in the training. The `ranger` library reported an *OOB error* of 0.0311. To verify if this error measurement would be consistent with overall accuracy, a second experiment where nine tenths of the pixels in the *training region 1* were used for training, while the remaining one tenth were used for classification, resulting in an overall accuracy of 96.9%, compatible with the *OOB error* computed.

Since Random Forest was proving effective in classifying coffee crops of distinct classes (as well at "not coffee" class with everything else in the region), we tested how well the classifier model would perform classifying a region that was distinct from the training region. So using all samples from TR1, we classified all pixels in the *classification region* (CR). This produced an overall accuracy of 87.34%, which is significantly worse, specially considering both regions were in the same satellite image (relatively close to one another).

Further experimentation was done in order to see what would improve classification accuracy in an area distinct from training, because we consider this an important practical aspect of crop mapping, which is usually not tackled in the literature.

Removing less important bands

Some of the input data is not very relevant for coffee classification. We wanted to test if removing some bands available from Sentinel-2 data would improve classification.

Previous work may employ some discriminant analysis such as Fisher’s Discriminant Ratio (FDR) (BOELL et al.,

https://doi.org/10.31422/taae.v9i3.64

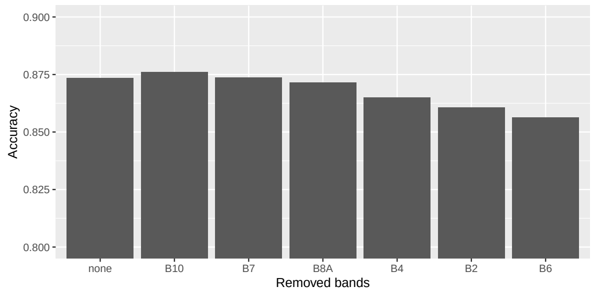

Figure 6: Changes in accuracy when removing less important bands. The y axis does not start at zero to enhance the differences. The x axis is cumulative, meaning that in each column, all bands removed previously are still absent from data.

2018) or Principal Component Analysis (PCA) and then use the high ranking data, while the rest is discarded. As the independent variables are ranked by importance in `ranger`'s implementation of the Random Forest algorithm, we used the importance value computed by it to choose bands for removal from training, the idea is that using less data would leave the algorithm with more memory to work the most important bands and would also reduce the noise that the training algorithm has to deal with. Band 10 (SWIR/Cirrus 1.375μm) was ranked the less important band in classification. Removing it resulted in a marginal improvement in accuracy. Further removing next less important bands did not improve accuracy, as shown in Figure 6.

Adding samples of coffee classes to the training

We used all pixels in training region for training, instead of using a small sample as usually done because we wanted to let the classification manage the noise that naturally occurs on crop delimitation for satellite images. Noise includes crop borders, farm tracks inside the crops, shadows from big trees and differences in vigor of the actual coffee trees. Random Forest is known for its robustness related to overfitting, so a classification would benefit from a large number of examples. Also, the non-coffee class is very diverse, because we did not create distinct classes for urban areas, rivers, lakes and other vegetation covers.

Because the number of samples for non coffee class was far greater then both coffee classes, we experimented proportionally increasing the number of coffee samples. To do so, all coffee pixels from training regions 2 and 3 were added to the training data. Increasing the number of coffee samples had a significant improvement in accuracy.

Adding vegetation indices to the training

Vegetation indices are commonly used for land cover classification, but since they are not independent data, and are instead computed from the bands already available for training, there is always a discussion about its benefits. We wanted to experiment if vegetation indices would improve accuracy. We chose NDVI and NDWI indices because the first is a well established index and the latter is related to water. Adding both indices to the training data resulted in a marginal improvement in accuracy.

Spatial Coherence

Using more samples of coffee crops, adding vegetation indices and removing the less important bands provided a good increase in accuracy. However, looking at a classification result, it is clear that accuracy would benefit form neighborhood information.

Neighborhood information is used in textural classification, but we wanted to use in a different way because textural classification usually requires choosing good parameters and is significantly more computation intensive.

A pixel is clearly less likely to be a sample of coffee plantation if there are none or few coffee pixels nearby. So we devised a way to make this information avaliable to the Random Forest classifier.

The polygons for coffee plantations were rasterized, creating a raster image where there are ones in coffee areas and zeroes elsewhere. Because we had two classes for coffee, two separate images were created, one for each class. These images were then convoluted (a focal operation was applied), so that each pixel gets the sum of its neighbors, creating information about how many nearby pixels are also from the same coffee class. We used a

https://doi.org/10.31422/taae.v9i3.64

5x5 pixel neighborhood, so that each pixel has 24 neighboors and its neighboorhood data is a number from 0 to 24 that indicates how many of its neighboors are coffee samples from the same class. Each coffee class produces one layer of neighborhood information.

The neighborhood information is clearly beneficial to the classifier, but to use it for classification one would need to know the classification output before the classification. To get around this problem, the classification was done in two steps. In the first step, classification is done without neighborhood information. The resulting classification is deemed an intermediary classification and used to create neighborhood information. In the second step training is done again (a second classification model is created), this time using neighborhood information. Then the classification is done again, but using neighborhood information from the intermediary classification (Figure 9). We called the second classification, a classification with *spatial coherence* because of the neighborhood knowledge it contains. The classification done on the second step does not contain the salt-and-pepper effect of single, misclassified pixels, borders are more accurate and holes inside plantations are smaller. Spatial coherence creates an eroding effect on false positives while having a dilating effect on false negatives, as seen in Figure 7.

Neighborhood information may be used in an iterative manner because subsequent classifications tend to be better than the previous. In our test region, accuracy increased from 92.4% to 94.3% with one iteration (adding neighborhood information), and then to 94.4% with two iterations (neighborhood updated – see Figure 8). Such a marginal enhancement, however, suggests that using it as an iterative method may not be worth the extra computation required. Figure 9 depicts a methodology overview of the classification steps used in this work.

Spatial coherence creates more continuous areas, feasible for creation of polygons, essentially producing results that look like object-based classification.

RESULTS

Performing classification on the same area produced the highest accuracy. The loss of accuracy noticed when classifying on a different area could be mitigated using several techniques, including our novel spatial coherence modification to the training data. Random Forest classification has shown to be very robust, using large amounts of training samples, which simplify the sampling process by including as many different samples as possible. However, balancing the percentage of coffee and non-coffee samples resulted in greater accuracy improvements than any other tuning strategy.

Removing bands form the training data is very dependent on the training data, removing a few of the less important bands, according to the internal Random Forest importance, we found that removing band 10 from training data improved accuracy by 0.3%. After this initial improvement, removing further bands in order of importance, produced a slight decrease in accuracy (Figure 6). In similar experimentation, with different scenarios than reported here, removing up to the 3rd less important band, resulted in marginal classification improvement.

Removing further bands decreased accuracy. Adding NDVI and NDWI bands accounted for another 0.3% increase in accuracy. Sentinel-2 Band 1 (Coastal aerosol 0.44μm) was surprisingly ranked first in importance for coffee mapping in every run.

Proportion of samples of different classes is important in the accuracy. Increasing the number of training samples from 16,605 (11.6% of all samples) to 35,510 (21.9%) resulted in 4.74% increase in overall accuracy and was the most useful improvement for classification.

Our novel neighborhood data increased accuracy in 1.9% for one iteration and 1.95% for two iterations, creating mapped regions more likely to be represented as coffee plantation polygons.

The loss of accuracy observed for classifying a region different from the training region could be recovered by using the techniques described. Table 1 is a summary of accuracy evolution as the classification was enhanced.

https://doi.org/10.31422/taae.v9i3.64

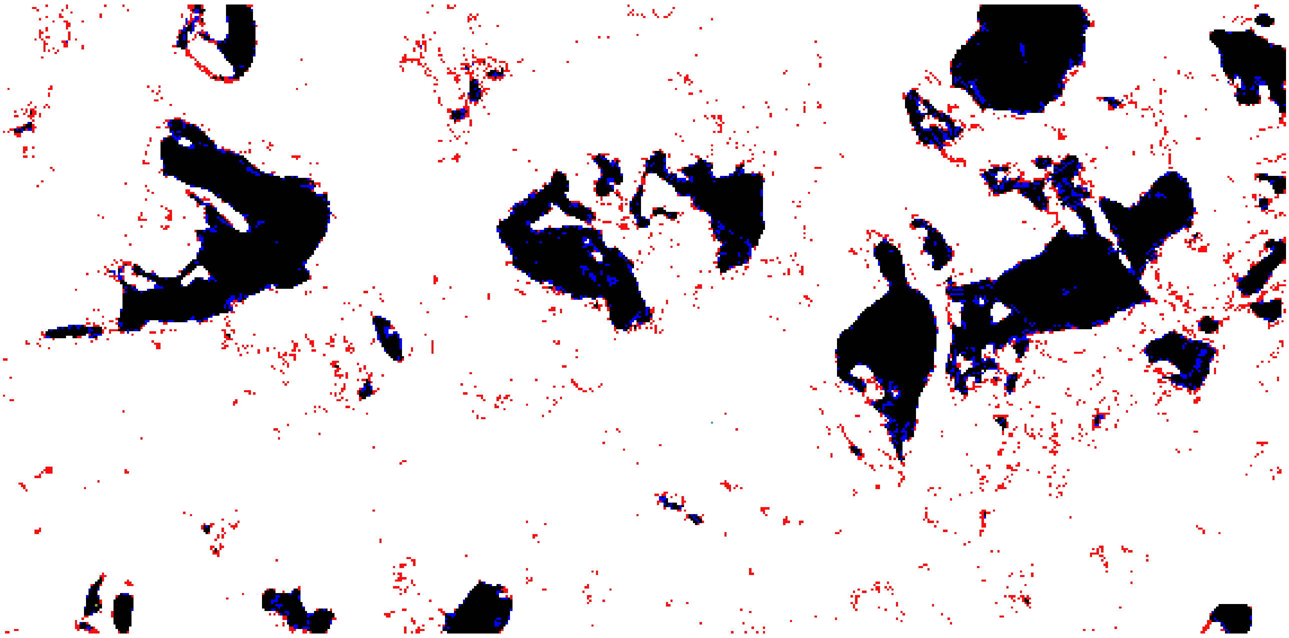

Figure 7: Effect of applying spatial coherence - red pixels were removed from coffee classes and blue pixels were inserted.

Figure 8: Effect of second iteration of spatial coherence - red pixels were removed from coffee classes and blue pixels were inserted.

Description	Accuracy	Difference
Training and classifying on Training Region 1, 13 bands	0.9694	–
Training on TR1, classifying on Classification Region, 13 bands	0.8734	-0.0960
Training on TR1, classifying on Classification Region, 12 bands	0.8760	+0.0026
Training on TR1 and TR2, classifying on Classification Region, 12 bands	0.8931	+0.0171
Training on TR1, TR2 and TR3, classifying on Classification Region	0.9234	+0.0303
Added NDVI and NDWI to data	0.9242	+0.0008
Added Neighborhood data	0.9429	+0.0187
Second iteration of neighborhood data	0.9437	+0.0008

Table 1: Timeline of classification accuracy.

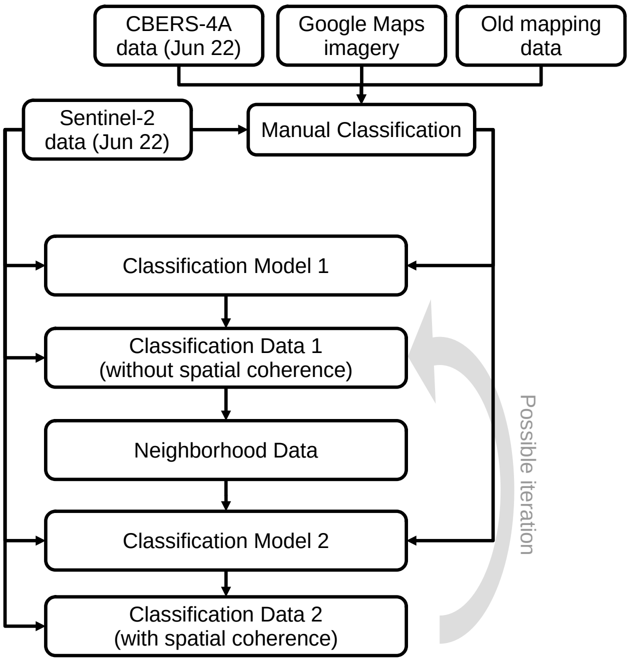

Figure 9: Methodology Overview

Figure 10: Coffee plantation (upper region, delimited by a polygon) and forest (lower region) create distinct shadow patterns on Sentinel-2 10m resolution images.

DISCUSSION

The decline in accuracy when classifying distinct regions (from 96.9% to 87.3%) reveals a critical limitation of machine learning models in remote sensing: the contextual dependence of training. Even in nearby areas within the same biome, subtle variations in topography, agricultural management, or coffee growth stages can impact classification. Spatial coherence mitigated this issue by incorporating local geographic information, reducing errors by 1.9%. This suggests that hybrid methods (spectral + spatial) are essential for large-scale applications. The accuracy of supervised classification models can vary significantly when applied to regions different from those used in training, even when the areas are geographically close and belong to the same satellite image. This behavior has already been observed by (MAXWELL; WARNER; FANG, 2018), who highlights the challenges of spatial extrapolation in machine learning-based classifications.

The integration of spatial and spectral information represents a promising path in enhancing classification accuracy. Forested regions emerged as the primary source of false positives in coffee classification as also noted by others (SOUZA et al., 2016); however, a distinct shadow pattern differentiates coffee plantations from forested areas. This distinction arises from the structured arrangement of coffee trees, as illustrated in Figure 10. Future research should further investigate this phenomenon to refine classification methodologies and improve the robustness of remote sensing applications in agricultural mapping.

Spatial data is also used in textural classification. It has been done for surface texture, i.e. canopy height variation, using SAR data (SILVA et al., 2009), and also done for image texture, i.e. variation of intensity in pixels (LELONG; THONG-CHANE, 2003; TSAI; CHEN, 2017). SAR data has also been used to identify forests (DOSTÁLOVÁ et al., 2016). We think texture based methods still need further work, as there are many non trivial ways for representing texture in useful ways for classification. It is also dependent on image resolution, size of analyzed windows, sun/shadow arrangements and land slope (BAETA et al., 2017). Spatial information is also behind object-based classification methods such as done by Wang et al. (WANG et al., 2018).

Using Sentinel-2 data, the band 10 was ranked less important for coffee classification. Removing less important bands from training should be done iteratively, as the importance of the remaining bands may change on the new model. We noticed that removing less important bands quickly exhausts its benefits. Removing one band was best for this work. We

https://doi.org/10.31422/taae.v9i3.64

performed other classification tests with more coffee classes that benefit from removing at most three bands from the training data. Since improvement was small it seems that the general rule is that Random Forest classification benefits from having more spectral bands. We have no explanation for the consistent first place in importance rank for band 1 and believe this merits further investigation.

Band 1 (0.44 μm), typically used for atmospheric correction, was the most relevant in the classification. One hypothesis is that it captures variations in moisture or atmospheric particles associated with young coffee plants, but further studies are needed to validate this correlation.

Unlike studies that use pure samples (ideal, carefully selected, data samples that represent each category), this work incorporated real-world noise such as tree shadows or trails, into the training process. This made the model more robust for practical applications, where pure data is impractical to obtain for large scale classification.

The use of all available samples for training, including areas with noise such as trails and shadows, proves to be a practical alternative in large-scale mapping contexts. This approach is supported by the robustness of the Random Forest algorithm, which stands out for its ability to handle intraclass variability and noisy data (BELGIU; DRĂGUŢ, 2016).

The use of Sentinel-2 data, with its adequate spatial resolution and multiple spectral bands, has also proven to be an effective basis for agricultural mapping, as previously demonstrated by (IMMITZER; VUOLO; ATZBERGER, 2016) in studies on tree species identification. The importance of balancing class proportions in training data was once again verified in this work, aligning with the methodological recommendations of (RAMEZAN; WARNER; MAXWELL, 2019), who emphasizes that sample balance and diversity are key factors for the performance of classifiers in studies using remote sensing data.

The final accuracy (94.4%) surpasses similar studies (CHEMURA; MUTANGA, 2017), (ESCOBAR-LÓPEZ et al., 2022), which achieved 89% and 91%, but direct comparison is limited due to data heterogeneity. The addition of vegetation indices (NDVI/NDWI) had little impact, suggesting that the original Sentinel-2 bands already capture sufficient information for coffee mapping.

CONCLUSION

Coffee mapping using Sentinel-2 data and Random Forests classification algorithm achieves a good accuracy, comparable with what is been reported in the literature. Although accuracy may vary significantly depending on training data, classification may be fine tuned using simple methodologies. The disadvantages of pixel based classification may be mitigated using spatial coherence in the training.

ACKNOWLEDGMENT

The authors would like to thank EPAMIG for the coffee crop polygons over which our updated data was based. No funding was received. The authors declare no conflict of interest.

REFERENCES

BAETA, Rafael et al. Learning Deep Features on Multiple Scales for Coffee Crop Recognition. In: 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). [S.l.: s.n.], 2017. P. 262–268. DOI: 10.1109/SIBGRAPI.2017.41.

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

BOELL, M. et al. Exploiting Feature Extraction Techniques for Remote Sensing Image Classification. **IEEE Latin America Transactions**, v. 16, n. 10, p. 2657–2664, 2018. DOI: 10.1109/TLA.2018.8795147.

BOURGOIN, Clément et al. Assessing the ecological vulnerability of forest landscape to agricultural frontier expansion in the Central Highlands of Vietnam. **International Journal of Applied Earth Observation and Geoinformation**, v. 84, p. 101958, 2020. ISSN 1569-8432. DOI: https://doi.org/10.1016/j.jag.2019.101958. Available from: <https://www.sciencedirect.com/science/article/pii/S0303243419307202>.

BREIMAN, Leo. Random forests. **Machine learning**, Springer, v. 45, p. 5–32, 2001.

CHEMURA, Abel; MUTANGA, Onisimo. Developing detailed age-specific thematic maps for coffee (Coffea arabica L.) in heterogeneous agricultural landscapes using random forests applied on Landsat 8 multispectral sensor. **Geocarto International**, Taylor & Francis, v. 32, n. 7, p. 759–776, 2017. DOI: 10.1080/10106049.2016.1178812.

DOSTÁLOVÁ, A. et al. FOREST AREA DERIVATION FROM SENTINEL-1 DATA. **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. III-7, p. 227–233, 2016. DOI: 10.5194/isprs-annals-III-7-227-2016. Available from: <https://isprs-annals.copernicus.org/articles/III-7/227/2016/>.

ESCOBAR-LÓPEZ, Agustín et al. Identifying Coffee Agroforestry System Types Using Multitemporal Sentinel-2 Data and Auxiliary Information. **Remote Sensing**, v. 14, n. 16, 2022. ISSN 2072-4292. DOI: 10.3390/rs14163847. Available from: <https://www.mdpi.com/2072-4292/14/16/3847>.

FINER, Matt et al. Combating deforestation: From satellite to intervention. **Science**, v. 360, n. 6395, p. 1303–1305, 2018. DOI: 10.1126/science.aat1203. eprint: https://www.science.org/doi/pdf/10.1126/science.aat1203. Available from: <https://www.science.org/doi/abs/10.1126/science.aat1203>.

IMMITZER, Markus; VUOLO, Francesco; ATZBERGER, Clement. Tree species classification with Random Forest using Sentinel-2 data. **Remote Sensing**, MDPI, v. 8, n. 2, p. 166, 2016. DOI: 10.3390/rs4092661.

KELLEY, Lisa C.; PITCHER, Lincoln; BACON, Chris. Using Google Earth Engine to Map Complex Shade-Grown Coffee Landscapes in Northern Nicaragua. **Remote Sensing**, v. 10, n. 6, 2018. ISSN 2072-4292. DOI: 10.3390/rs10060952. Available from: <https://www.mdpi.com/2072-4292/10/6/952>.

LELONG, Camille C. D.; THONG-CHANE, Audrey. Application of textural analysis on very high resolution panchromatic images to map coffee orchards in Uganda. In: 2003 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2003, Toulouse, France, July 21-15, 2003. [S.l.]: IEEE, 2003. P. 1007–1009. DOI: 10.1109/IGARSS.2003.1293994. Available from: <https://doi.org/10.1109/IGARSS.2003.1293994>.

MAXWELL, Aaron E; WARNER, Timothy A; FANG, Feng. Implementation of random forest classification in remote sensing for land use change. **Remote Sensing**, MDPI, v. 10, n. 11, p. 1733, 2018. DOI: 10.3390/rs10111733.

RAMEZAN, Christopher A; WARNER, Timothy A; MAXWELL, Aaron E. Evaluation of Sampling and Cross-Validation Tuning Strategies for Regional-Scale Machine Learning Classification. **Remote Sensing**, MDPI, v. 11, n. 2, p. 185, 2019. DOI: 10.3390/rs11020185.

SHI, Di; YANG, Xiaojun. An Assessment of Algorithmic Parameters Affecting Image Classification Accuracy by Random Forests. **Photogrammetric Engineering and Remote Sensing**, 2016. DOI: 10.1016/S0099-1112(16)82032-7.

SILVA, Wagner F. et al. Discrimination of agricultural crops in a tropical semi-arid region of Brazil based on L-band polarimetric airborne SAR data. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 64, n. 5, p. 458–463, 2009. Theme Issue: Mapping with SAR: Techniques and Applications. ISSN 0924-2716. DOI: https://doi.org/10.1016/j.isprsjprs.2008.07.005. Available from: <https://www.sciencedirect.com/science/article/pii/S0924271608000932>.

SOUZA, Carolina Gusmão et al. Algoritmos de Aprendizagem De Máquina e Variáveis de Sensoriamento Remoto para o Mapeamento da Cafeicultura. **Boletim de Ciências Geodésicas**, Universidade Federal do Paraná, v. 22, n. 4, p. 751–773, Oct. 2016. ISSN 1982-2170. DOI: 10.1590/S1982-21702016000400043. Available from: <https://doi.org/10.1590/S1982-21702016000400043>.

TSAI, Du-Ming; CHEN, Wan-Ling. Coffee plantation area recognition in satellite images using Fourier transform. **Computers and Electronics in Agriculture**, v. 135, p. 115–127, 2017. ISSN 0168-1699. DOI: 10.1016/j.compag.2016.12.020.

WANG, Wenjie et al. Improving Object-Based Land Use/Cover Classification from Medium Resolution Imagery by Markov Chain Geostatistical Post-Classification. **Land**, v. 7, n. 1, 2018. ISSN 2073-445X. DOI: 10.3390/land7010031. Available from: <https://www.mdpi.com/2073-445X/7/1/31>.