University of São Paulo
"Luiz de Queiroz" College of Agriculture

Orbital, aerial, and proximal sensing applied to monitoring the spatial variability of coffee plantations

Maurício Martello

Thesis presented to obtain the degree of Doctor in Science. Area: Agricultural Systems Engineering

Piracicaba
2022

Maurício Martello
Environmental Engineer

Orbital, aerial, and proximal sensing applied to monitoring the spatial variability of coffee plantations

Advisor:
Prof. Dr. **JOSÉ PAULO MOLIN**

Thesis presented to obtain the degree of Doctor in Science. Area: Agricultural Systems Engineering

Piracicaba
2022

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Martello, Maurício

Orbital, aerial, and proximal sensing applied to monitoring the spatial variability of coffee plantations / Maurício Martello. - - Piracicaba, 2022.

86 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”.

1. Agricultura de precisão 2. Sensoriamento Remoto 3. Geoprocessamento 4. Colheita mecanizada I. Título

ACKNOWLEDGEMENTS

My sincerest thanksfulness,

University of São Paulo - "Luiz de Queiroz" College of Agriculture (USP-ESALQ) and Agricultural Systems Engineering Graduate Program PPGESA) for providing the infrastructure and the environment necessary to get my Doctor in Science degree.

My advisor Prof. Dr. José Paulo Molin for for accepting me to be part of his research team. Thank you for encouraging and supporting my ideas throughout my doctorate. You are an inspiration for me as a scientist, professor and person.

The Coordination for the Improvement of Higher Education Personnel (CAPES), for granting the scholarship.

To Terrena Agro, Yara, Jacto, Planet Scope and Guima Coffee companies, for providing the equipment, imagens, and experimental fields over the three years of research. Especially to Mauro Mara, who on behalf of Terrena supported the project from the beginning. To Evaldo, Lucimar, Ricardo, Mariana, Salvador, Renato, and other people from the Guima Café farm who supported the field activities, always helping in the exchange of experiences and sparing no effort to help with whatever was needed. To Dr. Gustavo Portz for all the setup help and tips for using the sensor provided by Yara.

To all friends from Precision Agriculture Laboratory (LAP), including the guys from the mechanization and precision agriculture group (gMAP), for the pleasant working environment of the last years. In particular, Helizani C. Bazame, Tiago R. Tavarez, Leonardo F. Maldaner, Marcelo C.F. Wei and Matheus Acorsi for their support in data collection and processing.

To all employees of the Department of Biosystems Engineering, for cordiality and support during the daily activities, especially to Agnaldo F. Degaspari, Áureo S. de Oliveira, Davilmar A. D. Colevatti, Francisco de Oliveira, José G. Gomes, and Juarez R. do Amaral.

To my friends who are part of my family in Piracicaba, thank you very much for your support and companionship.

To my parents Mauri and Jaqueline, for the example of dedication, honesty and perseverance, models to be followed and, to my brothers Filipe, João and Antônio.

And especially to my love Graciele Angnes for her support and unconditional love. Thank you for encouraging me to walk this path so far, thank you also for your patience and support for the writing of this thesis, and also for all the emotional support during this journey. This thesis is dedicated to you! I love you!

Thank You All!

SUMÁRIO

RESUMO

Sensoriamento orbital, aéreo e proximal aplicado ao monitoramento da variabilidade espacial de lavouras de café

O Brasil é o maior produtor e exportador de café do mundo. Nos últimos anos cresceu a necessidade de modernização da produção agrícola, baseando-se na variabilidade espacial dos atributos do solo e das plantas, visando aumentar eficiência da lavoura e rentabilidade do produtor rural. Neste sentido, diferentes técnicas de sensoriamento (orbital, aéreo e proximal) têm sido testadas para mapear a variabilidade espacial de características de plantas e solo em diferentes sistemas de produção, permitindo obter alta frequência de dados com rapidez e baixo custo. No entanto, para a cultura do café os estudos que abordam essas técnicas em conjunto para gerar informações que subsidiem melhorias na gestão espacializada da lavoura de café são escassos. Por ser uma cultura perene, a produção de café é o resultado integrado dos diversos fatores envolvidos no manejo da cultura, solo, clima e da própria planta, como por exemplo o comportamento de bienalidade produtiva, com anos de alta e baixa produção. A criação de um sistema que permita obter essa informação de forma rápida e não invasiva é fundamental para uma gestão eficiente da lavoura. Nesse sentindo, este estudo buscou avaliar o potencial de uso de dados obtidos por diferentes sensores, buscando identificar atributos e feições da variabilidade presente nas lavouras de café que expressem relação direta com produtividade e consequentemente identifiquem a bienalidade produtiva (variabilidade temporal). Para tanto, o estudo foi dividido em cinco capítulos que abordam essas lacunas a partir de diferentes perspectivas. Inicialmente apresenta-se uma revisão breve sobre os principais trabalhos publicados na área, em seguida, no segundo capítulo buscou-se avaliar a qualidade dos dados obtidos por meio de um monitor de produtividade embarcado em uma colhedora de café. Os dados obtidos pelo monitor, de volume de café colhido, apresentaram alta correlação com dados obtidos com as células de carga, validando o método e permitindo mapear três safras consecutivas, que possibilitaram avançar na compreensão da variabilidade espacial e temporal da produtividade cafeeira em áreas comerciais. Visando explorar a potencialidade de mapear a variabilidade produtiva antes da colheita, os capítulos três, quatro e cinco, utilizaram técnicas de sensoriamento remoto para predizer e mapear a variabilidade espacial de uma lavoura comercial de café, com diferentes sensores e em três níveis de aquisição dos dados. No capítulo três foram explorados os sensores ópticos ativos (AOS) embarcados em um trator, em nível proximal. No quarto capítulo foi abordado o uso de imagens aéreas obtidas por aeronaves remotamente pilotadas (RPA) e no quinto capítulo foram utilizadas imagens orbitais de alta resolução espacial. De forma geral, os dados de sensores ópticos ativos apresentaram alta correlação com a produtividade, bem como permitiram monitorar temporalmente a variabilidade espacial da área de estudo, identificando regiões que apresentam inversão produtiva (bienalidade). Os resultados apresentados com o uso das imagens aéreas obtidas por RPA demonstraram o potencial das imagens na extração de parâmetros biofísicos dos cafeeiros. Também permitiram a individualização das plantas visando à extração de dados de altura e volume, permitindo observar a relação espaço-temporal das variáveis estudadas com dados de produtividade durante três safras consecutivas. As imagens orbitais de alta resolução espacial apresentaram a potencialidade em predizer a produtividade do café com alta assertividade um ano antes da colheita. De maneira geral os resultados encontrados nesse trabalho reforçam a importância em conhecer a variabilidade espacial produtiva nas áreas de café, pois esse tipo de informação auxilia na busca de possíveis causas dessa variabilidade para que a cultura possa ser manejada considerando essas diferenças espaciais e temporais. Este estudo apresentou as variações espaço-temporais dos dados de sensores orbitais, aéreos e terrestres em uma área comercial de café, bem como sua relação com mapas de produtividade gerados com alta densidade de dados, permitindo estimar a produtividade e identificar as variações produtivas ocasionadas pela bienalidade.

Palavras-chave: Agricultura de precisão; Sensoriamento Remoto; Geoprocessamento ;Colheita mecanizada

ABSTRACT

Orbital, aerial, and proximal sensing applied to monitoring the spatial variability of coffee plantations

Brazil is the largest coffee producer and exporter in the world. In recent years, the need to modernize agricultural production has grown, based on the spatial variability of soil and plant attributes, in order to increase crop efficiency and rural producer profitability. In this sense, different sensing techniques (orbital, aerial and proximal) have been tested to map the spatial variability of plant and soil characteristics in different production systems, allowing to obtain high frequency data quickly and at low cost. However, for the coffee culture, studies that approach these techniques together to generate information that subsidize improvements in the spatialized management of the coffee plantation are scarce. As it is a perennial crop, coffee production is the integrated result of the various factors involved in the management of the crop, soil, climate and the plant itself, such as the behavior of productive biennials, with years of high and low yields. The creation of a system that allows obtaining this information quickly and non-invasively is essential for an efficient management of the crop. In this sense, this study sought to evaluate the potential use of data obtained by different sensors, seeking to identify attributes and features of the variability present in coffee plantations that express a direct relationship with yield and, consequently, identify the biennial yield (temporal variability). To this end, the study was divided into five chapters that address these gaps from different perspectives. Initially, a brief review of the main works published in the area is presented, then, in the second chapter, we sought to evaluate the quality of the data obtained through a yield monitor embedded in a coffee harvester. The data obtained by the monitor, of the volume of coffee harvested, showed a high correlation with data obtained with the load cells, validating the method and allowing the mapping of three consecutive harvests, which made it possible to advance in the understanding of the spatial and temporal variability of coffee yield in commercial areas. Aiming to explore the potential of mapping production variability before harvest, chapters three, four and five used remote sensing techniques to predict and map the spatial variability of a commercial coffee crop, with different sensors and at three levels of data acquisition. In chapter three, active optical sensors (AOS) embedded in a tractor, at a proximal level, were explored. In the fourth chapter, the use of aerial images obtained by remotely piloted aircraft (RPA) was addressed and in the fifth chapter, high spatial resolution orbital images were used. In general, data from active optical sensors showed a high correlation with yield, as well as allowing the temporal monitoring of the spatial variability of the study area, identifying regions that present inversion of yield (biennial). The results presented with the use of aerial images obtained by RPA demonstrated the potential of images in the extraction of biophysical parameters from coffee trees. They also allowed the individualization of plants aiming at extracting height and volume data, allowing the observation of the space-time relationship of the variables studied with yield data during three consecutive seasons. High spatial resolution orbital images showed the potential to predict coffee yield with high assertiveness one year before harvest. In general, the results found in this work reinforce the importance of knowing the productive spatial variability in coffee areas, as this type of information helps in the search for possible causes of this variability so that the culture can be managed considering these spatial and temporal differences. This study presented the spatio-temporal variations of data from orbital, aerial and terrestrial sensors in a commercial coffee area, as well as their relationship with yield maps generated with high data density, allowing to estimate productivity and identify production variations caused by biennial.

Keywords: Precision Agriculture; Remote Sensing; Geoprocessing;Mechanized Harvesting

1. GENERAL INTRODUCTION

Coffee has high relevance in the Brazilian agricultural scenario. For many years it was the main local agricultural crop and, therefore, has an extremely relevant cultural and social role for the people involved in the production process. Currently, Brazil is the largest producer and exporter in the world and the second-largest consumer of beverages. The estimate for production is 53.4 million bags of coffee in the 2022 harvest (CONAB, 2022).

As it is a perennial crop, coffee production is the integrated result of the various factors involved in the management of the crop, soil, climate, and the plant itself (SANTINATO, 2016). Morphologically and physiologically, the coffee tree has some particularities that influence the variability of its productivity, for example the flowering of the coffee tree (CAMARGO; CAMARGO et al., 2001), which originates the coffee tree fruits, generally does not present uniformity. This fact interferes with the yield variation, as well as fruit maturation at harvest time (SANTINATO; FERNANDES, 2012). In addition, the plant bears fruit and grows simultaneously, and the fruiting drain is sovereign over the growth drain, which means that in years of high yield, there is little vegetative growth and, consequently, lower yield in the following crop, this variation is called biennial effect (PEREIRA, et al., 2011; VALADARES et al., 2013). In this sense, knowing the high and low yield areas can bring benefits from the adoption of management strategies, guiding the appropriate regulations for the harvesting process, increasing operational efficiency and reducing costs.

One approach to mapping spatial variability is through techniques from Precision Agriculture (PA) (MOLIN; AMARAL; COLAÇO, 2015). PA is one of the ways to face the demand in the increase of food production, since it helps the management of the area starting from the diagnosis of the causes of the variability until proposing solutions for the coexistence or correction of this variability, allowing the optimization of the use of resources available, increasing profitability and production sustainability (GEBBERS; ADAMCHUK, 2010). Researchers have sought to understand the spatial variability present in agricultural areas for more than a century (MERCER; HALL, 1911; WAYNICK; SHARP, 1919), but recently the in-depth study of techniques to map this variability with greater precision has been advancing. Among the main targets studied in agricultural areas, soil and plants stand out (KHOSLA et al., 2010).

Many research works have shown that in some cases soil attributes can be mapped through the use of sampling and geostatistics, allowing the identification of regions of high and low fertility.. Many of these studies involving spatial variability, especially in annual crops, are reported from the 1980s onwards, but in coffee crops there are reports from the 2000s onwards. Molin et al. (2002) observed significant yield variability in relatively small areas. These results indicated that the use of variable rate fertilizers in the studied areas was an important management option for the climate-plant-soil system. Silva et al. (2008) corroborated the results observed by Molin et al. (2002), emphasizing the need for differentiated and localized application of fertilizer. Soil sampling results indicated low range values of soil chemical attributes. The evaluation of the application of phosphorus and potassium at a variable rate in coffee was carried out in a field by Molin et al. (2010) and the results indicate savings of 23% in phosphorus and a 13% increase in potassium consumption when compared to the application of fertilizers at a fixed rate.

Regarding plant response, Silva, Lima and Queiroz (2011) and Silva, Lima and Bottega (2013) evaluated the spatial variability of the nutritional status of arabica coffee varieties using the content of foliar macro and micronutrients and relating them to the decrease in yield, bark percentage and crop yield. Silva, Lima and Bottega

(2013) observed nutritional imbalance related to the deficiency or excess of some nutrients (e.g., Ca, B and Zn) in plant tissue. The authors also point out that spatial analysis made it possible to evaluate the nutritional status of coffee plants considering the spatial variability existing in the field, providing information that allowed understanding the performance of the crop in the study area and generating information for future actions, aiming to avoid excess and deficiency of some nutrients that contributed to reduce the coffee yield.

It is observed that the works that study the behavior of spatial variability in microscale in coffee plantations evaluate attributes from punctual samples and with low sampling density, often limiting the generation of reliable data to map the variability. The costs of obtaining high-density data and the time required to collect these samples often prevent their practical application (KHOSLA; ALLEY, 1999; KHOSLA et al., 2002), making an alternative approach to obtaining data on a large scale necessary.

In the agricultural context, the application of remote sensing is based on the interaction of electromagnetic radiation (REM) with soil, vegetation and water (MULLA, 2013). The sensors can be embedded in different platforms, so their distance from the target can be classified according to the altitude levels, being orbital, aerial and proximal (FLORENZANO, 2002). In viticulture, researchers are using data from remote sensing at the orbital, aerial and proximal levels (ANASTASIOU et al., 2018; ANASTASIOU et al., 2019; BROOK et al., 2020) and the results demonstrate the achievement of these data with good precision, agility and affordable cost, even with a focus on nobler goals, such as product quality.

In coffee farming, the use of remote sensing has already been studied. However, at the terrestrial level, limited number of studies were found using embedded proximal sensors for high-density data collection. Putra et al. (2018), in Indonesia, individually evaluated 40 coffee plants with an active optic sensor (AOS) CropSpec (Topcon, Japan) and correlated them with leaf nitrogen and chlorophyll contents, finding promising results. In Brazil, Silva et al. (2015a) evaluated the use of the Normalized difference vegetation index (NDVI) using a AOS GreenSeeker (Trimble, USA) to observe the development of 30 coffee plants in southern Minas Gerais state. The authors point out that measurements using vegetation indices and AOS can be used to infer the physiological state of coffee trees. In another study, Silva et al. (2015b) related a time series of the Leaf Area Index (LAI) to the NDVI and found that the NDVI can be used to estimate the LAI, highlighting that measurements on the sides of the plants are more associated with LAI. Recently Oliveira et al. (2021) using AOS and artificial neural networks, managed to reasonably estimate the height and diameter of coffee trees using different indices in different parts of the plant.

At the aerial level, pioneering works point to the potential of aerial photographs (AMARAL, 1964; AMARAL; CHIARINI, 1968; AMARAL; VERDADE, 1966), highlighting the first work using this type of information for coffee in Brazil. Amaral (1964) used a set of aerial photographs taken in June 1962 at an approximate scale of 1:25000. In the work, the author sought to map the distribution and characteristics of the coffee culture in the municipality of Campinas, SP using photointerpretation techniques. In 1966, Amaral and Verdade (1966) sought to map the coffee growing situation in 13 municipalities in the northeast region of the state of São Paulo using aerial photographs. The work showed information regarding geographic distribution, number of plants, occupied area, age estimate, use of conservation practices and relationship with relief and soils.

Over time, technologies for collecting images have evolved, opening space for the development of research more focused on local mapping of small areas with a high level of detail. Aiming to map the weed infestation in the coffee crop, Sartori, Galo and Imai (2009) used multispectral aerial images, obtained from a digital video camera onboard a conventional aircraft.

In recent years, the number of studies using Remotely Piloted Aircraft (RPA) for various applications in agriculture has grown (GÓMEZ-CANDÓN; DE CASTRO; LÓPEZ-GRANADOS, 2014; SANKARAN et al., 2015). Studies highlight several uses for sensors embedded in small aircraft, such as identifying the appropriate time to harvest grapes (JOHNSON; HERWITZ, 2003), estimating the canopy height of olive trees (DÍAZ-VARELA et al., 2015), monitor rice and barley crop growth (BARETH et al., 2015; BENDIG et al., 2013), variability of water stress in vineyards (SANTESTEBAN et al., 2017), and estimate of biomass production in maize crops (LI et al., 2016). For Bendig et al. (2015) data collection with RPA fills a gap in the observational scale in remote sensing, providing high spatial and temporal resolution data, which are necessary in monitoring crop growth.

In the coffee culture, the work of Herwitz et al. (2004), using aerial images collected by an RPA with a wingspan of more than 35 m, mapping coffee areas south of Kauai, Hawaii, USA, found a positive correlation between a vegetation index (obtained with the RGB sensor) and the coffee maturation stage. Also noteworthy is the approach to crop management, indicating the use of images for locating invasive plants and monitoring irrigation. Another approach to aerial images is highlighted by Carrijo et al. (2017), who used Machine Learning techniques to classify the areas that had fruits on the tree. The flight was performed between the rows, obtaining images of the side of the plants. Recently, Cunha et al. (2019) used aerial images to model the crop canopy and extract plant volume information. The authors point out that the method is fast and allows obtaining information from large areas compared to the manual method. Barbosa et al. (2021) evaluated the potential of the practical application of RPA and RGB vegetation indices (VIs) in the monitoring of a coffee crop. The authors observed that the use of low-cost RPA and RGB cameras has potential for monitoring the coffee production cycle, providing producers with information in a more accurate, quick and simple way.

On the other hand, remote sensing at the orbital level stands out for its high coverage rate, allowing the monitoring of large areas allied to a temporal database. Studies using satellite images with the aim of estimating the areas cultivated with coffee portray the difficulty in discriminating the areas of coffee plantations from the other targets. Studies using images from the Landsat-5-sensor TM (Thematic Mapper) satellite point out the difficulty in classifying areas with perennial crops (coffee and citrus), due to the low spatial resolution of these images (TARDIN; ASSUNÇÃO; SOARES, 1992). Studies using spectral and temporal analysis of TM sensor images highlight that band 4 images (Near Infrared) showed better differentiation with the other targets. The authors identified high variability between crops, which can be attributed to the difference in age, spacing and cultivar (MOREIRA; ADAMI; RUDORFF, 2004; FERRAZ et al., 2012).

However, in recent years, some authors have obtained good results using orbital images. Tsai and Chen (2017) analyzed satellite images of coffee growing regions at different times with different lighting and cultivation stages in Brazil, Africa, Vietnam and Hawaii, using a machine vision scheme to recognize the coffee plantation area. Tridawati et al., (2020) mapped the distribution of coffee plantations from Multi-Resolution, Multi-Temporam and Multi- Sensor data using a random forest algorithm. These authors used pan-sharpened orbital images from GeoEye-1, multi-temporal Sentinel 2, and DEMNAS.

In addition, the estimation of coffee yield from orbital images can be made using vegetation indices. The study by Nogueira et al. (2018) reports the use of vegetation indices obtained with images from the Landsat-8 satellite - OLI sensor (Operational Land Imager). The authors point out that the best phases to determine coffee yield were the phenological phases of dormancy and flowering.

Hunt et al. (2020), conducting a literature review, highlighted the variety of employed approaches and sensors that specific researchers have previously to accurately map coffee. However, they do not identify any routine methods that consistently produced accurate results, thus making a challenging method recommendation. They also found significant gaps in the literature, including mapping coffee using other sensor types, such as synthetic aperture radar (SAR), studies differentiating between multiple production systems. Silva et al. (2021) evaluated the correlation of different phenological stages spectral responses of coffee in center pivot, under distinct apparent brightness conditions, with the yield sampled in field. They observed that coffee yield could be predicted from the first month of its production phase. Crusiol et al. (2021) investigated the potential of using high spatial resolution daily NDVI-time-series from Planet CubeSat images for crop monitoring.

However, it is still a challenge to integrate the existing knowledge of the use of each of these sensing techniques to jointly understand the spatial variability present in the field and to correlate it with coffee yield. As surveys are generally conducted with punctual sampling and with low sampling density, an alternative for collecting yield data with high density and for large areas is to use yield monitors built into the harvesters. Sartori et al. (2002) proposed a yield monitor, which became commercially available a few years ago, allowing the evaluation of coffee yield through a sensor system that measures the volume of harvested fruits.

In view of the above, it is highlighted that the motivations for carrying out this research were: i) the need for practical methods, using data from the harvester and remote sensors at their different levels of collection to enable the collection of high density data on variability spatial yield in coffee area; ii) the lack of research that validates the performance of datasets obtained by remote sensing at differente levels of collection and correlate with high-density yield data in commercial coffee plantations.

The thesis consists of five capters addressing these gaps from different perspectives. The (i) first chapter presents a brief introduction on the research topic studied. The (ii) second chapter of this thesis, entitled "Obtaining and validating high-density coffee yield data", aims to evaluate the quality of the data obtained through a yield monitor built into a coffee harvester and generate yield maps to express and advance the understanding of the spatial and temporal variability of coffee yield in commercial areas. This chapter was published in the scientific journal Horticulturae (Martello, M.; Molin, J.P.; Bazame, H.C. 2022. doi.org/10.3390/horticulturae8050421). The (iii) third chapter, entitled "Use of active optical sensor in coffee for monitoring crop yield" aimed to evaluate the potential of AOS to map the spatial and temporal variability of coffee crop yields, as well as to establish guidelines for the acquisition of AOS data for sensing the side of coffee plant, allowing the evaluation of large commercial fields. The (iv) fourth chapter, entitled "Assessing temporal and spatial variability of coffee plants using RPA-based RGB imaging", aimed to explore the use of RGB aerial images to obtain 3D information of coffee crop, deriving plant height and volume information together with yield data. The (v) fifth and final chapter of this thesis is entitled "Coffee yield estimation using high resolution orbital imaging and machine learning". The objective of this study was to adjust yield prediction models based on a time series of satellite image data (spectral bands and vegetation indices - VIs) and high density data collected using random forest and multiple linear regression, indicating the appropriate phenological stage of coffee crop to obtain satellite image data. After these chapters, Final Considerations are made, presenting a summary of the results found and perceptive for future work.

REFERENCES

AMARAL, Á. Z. do. Distribuição e características da cultura cafeeira no município de campinas levantadas pela foto interpretação. **Bragantia**, v. 23, n. unico, p. 271–279, 1964.

AMARAL, A. Z. do; VERDADE, F. da C. Situação da cafeicultura em alguns municípios da região nordeste do Estado de São Paulo, levantada com auxílio da fotografia aérea. **Bragantia**, v. 25, n. 1, p. 95–106, 1966.

AMARAL, Á. Z. do; CHIARINI, J. V. Distribuição e características da cultura cafeeira no Estado de São Paulo, levantadas através de fotografias aéreas. **Bragantia**, v. 27, n. 2, p. 267–277, 1968.

ANASTASIOU, E.; BALAFOUTIS, A.; DARRA, N.; PSIROUKIS, V.; BINIARI, A.; XANTHOPOULOS, G.; FOUNTAS, S. Satellite and proximal sensing to estimate the yield and quality of table grapes. **Agriculture**, v.8(7):94, 2018.

ANASTASIOU, E.; CASTRIGNANÒ, A.; ARVANITIS, K.; FOUNTAS, S. A multi-source data fusion approach to assess spatial-temporal variability and delineate homogeneous zones: A use case in a table grape vineyard in Greece**. Sci. Total. Environ.**, v. 684, 155–163, 2019.

BARBOSA, B.D.S.; FERRAZ, G.A.S.; MENDES DOS SANTOS, L.; SANTANA, L.S.; BEDIN MARIN, D.; ROSSI, G.; CONTI, L. Application of RGB Images Obtained by UAV in Coffee Farming. **Remote Sensing**, 13, 2397, 2021.

BARETH, G.; AASEN, H.; BENDIG, J.; GNYP, M. L.; BOLTEN, A.; JUNG, A.; MICHELS, R.; SOUKKAMÄKI, J. Low-weight and UAV-based Hyperspectral Full-frame Cameras for Monitoring Crops: Spectral Comparison with Portable Spectroradiometer Measurements. Photogrammetrie. **Fernerkundung - Geoinformation**, v. 2015, n. 1, p. 69–79, 2015.

BENDIG, J.; WILLKOMM, M.; TILLY, N.; GNYP, M. L.; BENNERTZ, S.; QIANG, C.; MIAO, Y.; BARETH, G. Very high resolution crop surface models (CSMs) from UAV-based stereo images for rice growth monitoring In Northeast China. **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. XL, p. 4–6, 2013.

BENDIG, J.; YU, K.; AASEN, H.; BOLTEN, A.; BENNERTZ, S.; BROSCHEIT, J.; GNYP, M. L.; BARETH, G. combining UAV-based plant height from crop surface models visible. **International Journal of Applied Earth Observation and Geoinformation,** v. 39, p.79-87, 2015.

BROOK, A.; DE MICCO, V.; BATTIPAGLIA, G.; ERBAGGIO, A.; LUDENO, G.; CATAPANO, I.; BONFANTE, A. A Smart Multiple Spatial and Temporal Resolution System to Support Precision Agriculture from Satellite Images: Proof of Concept on Aglianico Vineyard. **Remote Sens. Environ**.v. 240, 111679, 2020.

CAMARGO, Â. P.; CAMARGO, M. B. P. DE. Definição e esquematização das fases fenológicas do cafeeiro arábica nas condições tropicais do Brasil. **Bragantia**. v. 60, n. 1, p. 65-68, 2001.

CARRIJO, G. L. A.; OLIVEIRA, D. E.; DE ASSIS, G. A.; CARNEIRO, M. G.; GUIZILINI, V. C.; SOUZA, J. R. Automatic detection of fruits in coffee crops from aerial images. In: LARS 14TH LATIN AMERICAN ROBOTICS SYMPOSIUM AND 2017 5TH SBR BRAZILIAN SYMPOSIUM ON ROBOTICS, LARS-SBR 2017, Curitiba. **Anais**…Piscataway, NJ:IEEE, 2017.p.196-206

COMPANHIA NACIONAL DE ABASTECIMENTO—(CONAB) Acompanhamento da Safra Brasileira: Café. Prim. Levant. 2022 Brasília, 1, 60, 2022.

CUNHA, J. P. A. R. da; SIRQUEIRA NETO, M. A.; HURTADO, S. M. C.; CUNHA, J. P. A. R. da; SIRQUEIRA NETO, M. A.; HURTADO, S. M. C. Estimating vegetation volume of coffee crops using images from unmanned aerial vehicles. **Engenharia Agrícola**, v. 39, n. spe, p. 41–47, set. 2019.

CRUSIOL, G. T.; SUN, L.; CHEN, R.; SUN, Z.; ZHANG, D.; CHEN, Z.; WUYUN, D.; NANNI, R. M.; LIMA NEPOMUCENO, A.; BOUÇAS FARIAS, J.R. Assessing the potential of using high spatial resolution daily NDVI-time-series from Planet CubeSat images for crop monitoring. **Int. J. Remote Sens.** v. 42, p.7114–7142, 2021.

DÍAZ-VARELA, R.; DE LA ROSA, R.; LEÓN, L.; ZARCO-TEJADA, P. High-Resolution Airborne UAV Imagery to Assess Olive Tree Crown Parameters Using 3D Photo Reconstruction: Application in Breeding Trials. **Remote Sensing**, v. 7, n. 4, p. 4213–4232, 2015.

FERRAZ, G.A.S.; SILVA, F.M.; ALVES, M.C.; BUENO, R.L; COSTA, P.A.N. Geostatistical analysis of fruit yield and detachment force in coffee. **Precision Agriculture**, v. 13, n. 1, p. 76–89, 2012.

FLORENZANO, T. G. Imagens de satélites para estudos ambientais. São Paulo: **Oficina de Textos**, 2002.

GEBBERS, R.; ADAMCHUK, V. I. Precision agriculture and food security. **Science** (New York, N.Y.), v. 327, n. 5967, p. 828–31, 2010.

GÓMEZ-CANDÓN, D.; DE CASTRO, A. I.; LÓPEZ-GRANADOS, F. Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat. **Precision Agriculture**, v. 15, n. 1, p. 44–56, 2014.

HERWITZ, S. .; JOHNSON, L. .; DUNAGAN, S. .; HIGGINS, R. .; SULLIVAN, D. .; ZHENG, J.; LOBITZ, B. .; LEUNG, J. .; GALLMEYER, B. .; AOYAGI, M.; SLYE, R. .; BRASS, J. . Imaging from an unmanned aerial vehicle: agricultural surveillance and decision supportghts (coffee seeds encased in endocarp) yielded by each field. Mechanical sorting by the mill allowed the grower to track and assign ripeness level to parchment Weight. **Computers and Electronics in Agriculture**, v. 44, n. 1, p. 49–61, 2004.

HUNT, D.A., TABOR, K., HEWSON, J.H., WOOD, M.A., REYMONDIN, L., KOENIG, K., SCHMITTHARSH, M., FOLLETT, F. Review of remote sensing methods to map coffee production systems. **Remote Sensing**. v.12, 2041, 2020.

JOHNSON, L.; HERWITZ, S. Collection of ultra high spatial and spectral resolution image data over California vineyards with a small UAV. **Proc. 30th International Symposium on Remote Sensing of Environment** (ISRSE), n. Ccd, p. 3–5, 2003.

KHOSLA, R., & ALLEY, M. M. Soil-specific nitrogen management on Mid-Atlantic Coastal Plain soils. **Better Crops**, V. 83, P.6–7, 1999.

KHOSLA R., FLEMING, K. L., DELGADO, J. A., SHAVER, T. M., & WESTFALL, D. G. Use of site-specific management zones to improve nitrogen management for precision agriculture**. Journal of Soil and Water Conservation**, v.57, p.513–518, 2002.

KHOSLA, R., WESTFALL, D.G., REICH, R.M., MAHAL, J.S., GANGLOFF, W.J. Spatial Variation and Site-Specific Management Zones. In: Oliver, M. (eds) **Geostatistical Applications for Precision Agriculture**. Springer, Dordrecht, p.195-209, 2010.

LI, W.; NIU, Z.; CHEN, H.; LI, D.; WU, M.; ZHAO, W. Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. **Ecological Indicators**, v. 67, p. 637–648, ago. 2016.

MERCER, W.B. AND HALL, A.D. (1911). The Experimental Error of Field Trials. **Journal of Agricultural Science**, v.4, p.107-132, 1911.

MOLIN, J. P.; RIBEIRO FILHO, A. C.; TORRES, F. P.; SHIRAISI, L. E.; SARTORI, S.; SARRIÉS, G. A. Mapeamento da produtividade de café e sua correlação com componentes da fertilidade do solo em duas áreas piloto. In: Simposio de Agricultura de Precisão, 3, 2002, Piracicaba. Avanços na Agricultura de Precisão no Brasil no Período de 1999-2001.**Anais...** Piracicaba: USP/Esalq, p. 58-65. 2002.

MOLIN, J. P.; MOTOMIYA, A. V. de A.; FRASSON, F. R.; FAULIN, G. D. C.; TOSTA, W. Test procedure for variable rate fertilizer on coffee. Acta Scientiarum. **Agronomy**, v. 32, n. 4, p. 569–575, 2010.

MOLIN, J. P.; AMARAL, L. R.; COLAÇO, A. F. Agricultura de precisão. **Oficina de Textos**, São Paulo, 2015.

MOREIRA, M. A.; ADAMI, M.; RUDORFF, B. F. T. Análise espectral e temporal da cultura do café em imagens Landsat. **Pesquisa Agropecuária Brasileira**, v. 39, n. 3, p. 223–231, 2004.

MULLA, D. J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. **Biosystems Engineering**, v. 114, n. 4, p. 358–371, 2013.

NOGUEIRA, S.M.C., MOREIRA, M.A., VOLPATO, M.M.L. Relationship between coffee crop productivity and vegetation indexes derived from OLI/Landsat-8 Sensor Data with and without topographic correction. **Eng. Agríc.** 38, 387–394, 2018.

OLIVEIRA, M.F., SANTOS, A.F., KAZAMA, E.H., ROLIM, G. S., SILVA, R.P. Determination of application volume for coffee plantations using artificial neural networks and remote sensing. **Comput. Electron. Agric**. 184, 106096, 2021.

PEREIRA, S.O.; BARTHOLO, G.F.; BALIZA, D.P., SOGREIRA, F.M.; GUIMARÃES, R. J. Productivity and coffee biannuality depending on the crop spacing. **Pesqui Agropecuária Bras**. V.46(2), p.152-160, 2011.

PUTRA, B. T. W.; SONI, P.; MORIMOTO, E.; PUJIYANTO, P. Estimating biophysical properties of coffee (Coffea canephora) plants with above-canopy field measurements, using CropSpec. **International Agrophysics**, v. 32, n. 2, p. 183–191, 2018.

SANKARAN, S.; KHOT, L. R.; ESPINOZA, C. Z.; JAROLMASJED, S.; SATHUVALLI, V. R.; VANDEMARK, G. J.; MIKLAS, P. N.; CARTER, A. H.; PUMPHREY, M. O.; KNOWLES, R. R. N.; PAVEK, M. J. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. **European Journal of Agronomy**, v. 70, n. March, p. 112–123, 2015.

SANTESTEBAN, L. G.; DI GENNARO, S. F.; HERRERO-LANGREO, A.; MIRANDA, C.; ROYO, J. B.; MATESE, A. High-resolution UAV-based thermal imaging to estimate the instantaneous and seasonal variability of plant water status within a vineyard. **Agricultural Water Management**, v. 183, p. 49–59, 2017.

SANTINATO, F. **Inovações tecnológicas para cafeicultura de precisão**. Tese (Doutorado em Agronomia). Faculdade de Ciências Agrárias e Veterinárias – UNESP, Jaboticabal,SP.125p. 2016.

SANTINATO, R., FERNANDES, A.L.T. Cultivo do cafeeiro irrigado por gotejamento. 2. Ed. Uberaba: **Autores**, 388p., 2012.

SARTORI, L. R.; GALO, M. de L. B. T.; IMAI, N. N. Mapeamento de plantas daninhas em cultura de café a partir de imagens multiespectrais de escalas grandes usando redes neurais artificiais. **Revista Brasileira de Cartografia**, v. 61, n. 2, 2009.

SARTORI, S.; FAVA, J. F. M.; DOMINGUES, E. L.; RIBEIRO FILHO, A. C.; SHIRAISI, L. E. Mapping the Spatial Variability of Coffee Yield with Mechanical Harvester. In: Proceedings of the World Congress of Computers in Agriculture and Natural Resources. **Anais...** ASAE, Foz do Iguaçu - PR: 2002.

SILVA, F. M.; SOUZA, Z. M.; FIGUEIREDO, C. A. P.; VIEIRA, L. H. D. S.; OLIVEIRA, E. Variabilidade espacial de atributos químicos e produtividade da cultura do café em duas safras agricolas. **Ciência e Agrotecnologia**, v. 32, n. 1, p. 231–241, 2008.

SILVA, S. A.; LIMA, J. S. de S.; QUEIROZ, D. M. Spatial variability in nutritional status of arabic coffee based on dris index. **Rev. Ceres**, v. 58, n. 2, p. 256–261, 2011.

SILVA, S.; LIMA, J. S.; BOTTEGA, E. . Yield mapping of arabic coffee and their relationship with plant nutritional status. **Journal of soil science and plant nutrition**, v. 13, 2013.

SILVA, L. C.; FILHO, M. P. da S.; AGUIAR, F. A.; NAVES, V. L.; LIRA, J. M. S.; RAMBAL, S.; MOUILLOT, F.; BARBOSA, J. P. R. A. D. Uso do índice de vegetação por diferença normalizada (ndvi) na avaliação do desenvolvimento de cafeeiros do sul de minas gerais. In: IX SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 2015, Curitiba. **Anais**... Consórcio pesquisa café, 2015a.

SILVA, L. C.; FILHO, M. P. da S.; AGUIAR, F. A.; NAVES, V. L.; LIRA, J. M. S.; RAMBAL, S.; PENNACCHI, J. P.; MOUILLOT, F.; BARBOSA, J. P. R. A. D. Conhecendo o iaf do cafeeiro arábica através do índice de vegetação por diferença normalizada. In: IX Simpósio de Pesquisa dos Cafés do Brasil, Curitiba. Anais...Concórcio pesquisa café, 2015b.

SILVA, P.A. A.; ALVES, M.C.; SILVA,F.M.; FIGUEIREDO, V.C. Coffee yield estimation by Landsat-8 imagery considering shading effects of planting row's orientation in center pivot. **Remote Sensing Applications: Society and Environment,** V. 24, 2021.

TARDIN, A. T.; ASSUNÇÃO, G. V.; SOARES, J. V. Análise preliminar de imagens TM visando à discriminação de café, citrus e cana-de-açúcar na região de Furnas, MG. **Pesquisa Agropecuária Brasileira**, v. 27, 9, p. 1355–1361, 1 set. 1992.

TRIDAWATI, A.; WIKANTIKA, K.; SUSANTORO, T.M.; HARTO, A.B.; DARMAWAN, S.; YAYUSMAN, L.F.; GHAZALI, M.F. Mapping the Distribution of Coffee Plantations from Multi-Resolution, Multi-Temporal, and Multi-Sensor Data Using a Random Forest Algorithm. **Remote Sensing**, v. 12, 2020.

TSAI, D.M., CHEN, W.L. Coffee plantation area recognition in satellite images using Fourier transform. **Comput. Electron. Agric**. V.135, 115–127, 2017

VALADARES S.V, NEVES J.C.L, ROSA G.N.G.P, MARTINEZ H.E.P, VENEGAS, V.H.A, LIMA P. C.Productivity and biennial production of dense coffee plantations under different doses of N and K. **Pesqui Agropecuária Bras**. V. 13, 48(3)p.96-303, 2013.

WAYNICK, D.D. AND SHARP, L.T. Variability in Soils and its Significance to Past and Future Soil Investigations. II. **Variations in Nitrogen and Carbon in Field Soils and Their Relation to the Accuracy of Field Trials**. University of California Publications in Agricultural Sciences, v. 4: p.121–139, 1919.

2. OBTAINING AND VALIDATING HIGH-DENSITY COFFEE YIELD DATA

ABSTRACT

Coffee producers are ever more interested in understanding the dynamics of coffee's spatial and temporal variability. However, it is necessary to obtain high-density yield data for deci-sion-making. The objective of this study is to evaluate the quality of yield data obtained through a yield monitor onboard a coffee harvester, as well as to evaluate the potential of the data col-lected over three harvests. The yield monitor validation data showed a high correlation (above R^2 0.968) when compared with the data obtained by a wagon instrumented with load cells. It was also possible to obtain yield maps for three consecutive seasons, allowing the identification of their internal variability, as well as classifying regions that show alternating yield patterns be-tween years as the expression of the biennial yield behavior manifested inside and along the field, in addition to the spatial variability. This result indicates that, in addition to knowing the spatial yield variability, the biennial variance information must also be considered in the strat-egies for site-specific management. Regions that presented high yield variance should be alter-nated according to the productive year (high and low yield) and not only in consideration of their yield variability as on the regions with more stable yield behavior over time. The use of yield data can help the producer make more assertive decisions for crop and farm management.

Keywords: Precision agriculture; Coffee yield monitor; Yield map validation; Mechanical harvesting; Coffee biennial cycle

2.1. INTRODUCTION

Brazil is among the world's leading producers of Coffea arabica L., second-largest consumer, and leader in world exports (CONAB, 2021). The expected area cultivated with coffee in Brazil for the year 2022 is 1.82 million hectares, with 78% of the Arabica variety and 22% of the Conilon (or Robusta) variety (CONAB, 2022).

As a perennial crop, coffee yield is the integrated result of various factors involved in the crop management, soil, climate, and the plant itself (SANTINATO, 2016). The coffee crop has particulari-ties that contribute to high spatial variability of yield, and one of them is its biennial na-ture (DE CARMARGO; DE CARMARGO, 2001). Due to the physiological characteristics of the plant, the crop alternates between years of high and low yield and this biennial period has a marked effect on the coffee yield variability (RENA; MAESTRI, 1985). The previous knowledge of high and low yield areas has its benefits, for example, the adoption of management strategies that consider the potential for yield and profitability. In addition, a broad knowledge of high-quality coffee production techniques is indispensable for modern coffee farming and, for that, modern management tools, such as precision agriculture (PA), must be used (MOLIN et al., 2010).

PA techniques offer solutions for the differentiated management of areas, in which inputs are applied according to demands where and when they are needed. The main ob-jective of PA is to provide economic and environmental benefits (LOWENBERG; ERICKSON 2019). Thus, agricultural practices with greater precision can maximize the potential of each portion of the field, making the crop more profitable, favoring cost reduction and minimizing environmental impacts (MURUNGAN; GARG; SINGH 2017).

For some crops, such as grains, techniques for monitoring the spatial variability of yield are already relatively consolidated and available on the market, unlike perennial crops such as coffee (MOLIN; FAULIN, 2013; COLAÇO et al., 2020). The yield map is considered the most complete and true in-formation to visualize the crop variability, as it shows the actual response of the crop to the conditions presented during the season and, thereby, affected by the different produc-tion factors (MOLIN; AMARAL; COLAÇO 2015). In mechanized harvesting, this information can be obtained throughout the harvest with the georeferenced measurement of the product flow in the harvester.

In Brazil, with the increasing reduction of labor availability for agriculture, the coffee plantations, traditionally cultivated in sloped areas, have migrated to flatter areas, signif-icantly expanding mechanized harvesting (ROSA et al., 2010). This technological advance was made possible by the development of the first coffee harvester in 1979 (SILVA et al., 2008).

According to Silva et al. (2008), this technological development in mechanization has been favorable for the coffee sector, but when it comes to yield mapping, most of the scien-tific approaches that aim to identify spatial variability and yield map of coffee plants, use manual collected point sampling techniques with low density, in the order of 2.2 to 4.6 points ha^{-1} (FERRAZ et al., 2012; CARVALHO et al., 2017; FERRAZ et al., 2017). Pioneering investigations for coffee yield mapping with mechanized harvesting were carried out by Balastreire et al. (2002), measuring the weight of stripped coffee, and by Sartori et al. (2002), Molin et al. (2010), and Angnes et al. (2021), measuring the harvested volume and which became commercial for some time. Both obtained consistent results quantifying the spatial variability of coffee yield. However, in addition to the raw data obtained during the harvesting, Sartori et al. (2002) and Molin et al. (2010) used a Conversion Factor (CF) to transform field data obtained by the harvester into processed coffee yield data. Traditionally, producers already sample known volumes of coffee that are then dried, peeled, and weighed to obtain the CF (FAULIN; MOLIN; STANISLAVSKI 2014). In addition to the CF, the samples can also be classified according to the degree of maturation to quantify the classes of harvested fruits, such as unripe, ripe, and overripe fruits (BAZAME et al., 2021).

Recently, Santana et al. (2021) compiled academic research on PA in Brazilian coffee production in the period from 2000 to 2021 and observed that remote sensing and the investigation of soil and climate spatial variability, along with yield data sampled with aid of geostatistics, predominated with few studies focusing on yield maps from harvesters. In addition, Molin et al. (2010) and Angnes et al. (2021) showed that the observation of areas with different productive potential within the stands, and the possibility of site-specific treatment indicates the potential of using PA concepts for the management of coffee areas.

In this sense, the objectives of this study were to evaluate the quality of coffee volume data obtained using a yield monitor embedded in a coffee harvester. The other objective was to generate yield maps along production cycles to advance the understanding of the spatial and temporal variability of coffee yields in commercial areas. The spatialized production information in the form of yield maps is fundamental for understanding the behavior of the variability of a crop.

2.2. MATERIALS AND METHODS

The study was conducted in commercial coffee areas in the municipality of Patos de Minas, Minas Gerais state, Brazil (reference coordinates: Latitude 18°32'28.55" S, Longitude 46°3'51.17" W, coordinates reference system WGS 84; altitude greater than 1000 m). The climate of the study areas is classified as AW, tropical with dry winter and rainy summer, according to the Köppen climate classification (ALVARES et al., 2013). The monthly normals temperature (1991–2020) range from 18.9 °C in the coldest month (June) to 23.4 °C in the warmest month (October) with average annual temperature around 21.6 °C (INMET 2022).

The yield data were collected for the crop harvests of 2019, 2020, and 2021, with a single pass of the harvester. The experimental area of 10.24 ha is characterized by a flat to a slightly corrugated surface and equipped with drip irrigation. The area was cultivated with the species *Coffea arabica L.*, IAC Catuaí 144 variety, planted in 2006 and had its first harvest in 2009.

Data were obtained using a harvester K3 Millennium (Jacto, Pompeia, Brazil), equipped with a yield monitor (SARTORI et al., 2002). In the coffee harvester, two systems of vibrating rods detach the fruits from the branches on both sides of the coffee plants. The fruits are collected below the plants, transported by horizontal belts, and driven by two elevators, one on each side of the plants. When the fruits get to the top of the machine, they are blown to remove light impurities and then fall into a small reservoir that feeds the paddle conveyor. The paddle conveyor then takes the fruits to a container that moves parallel to the harvester (Figure 1—green flow). On top of the temporary reservoir, there is an ultrasonic sensor which monitors the presence or absence of fruits and governs the hydraulic motor (Figure 1c) trigger of the paddle conveyor, ensuring its 2.791 L volume compartments are full. The datalogger collected data (number of full cells in a given period) at a frequency of 0.05 Hz, with the geographic coordinates obtained by the GNSS receiver. The work speed is obtained by the displacement sensor installed on a non-drive wheel of the harvester.

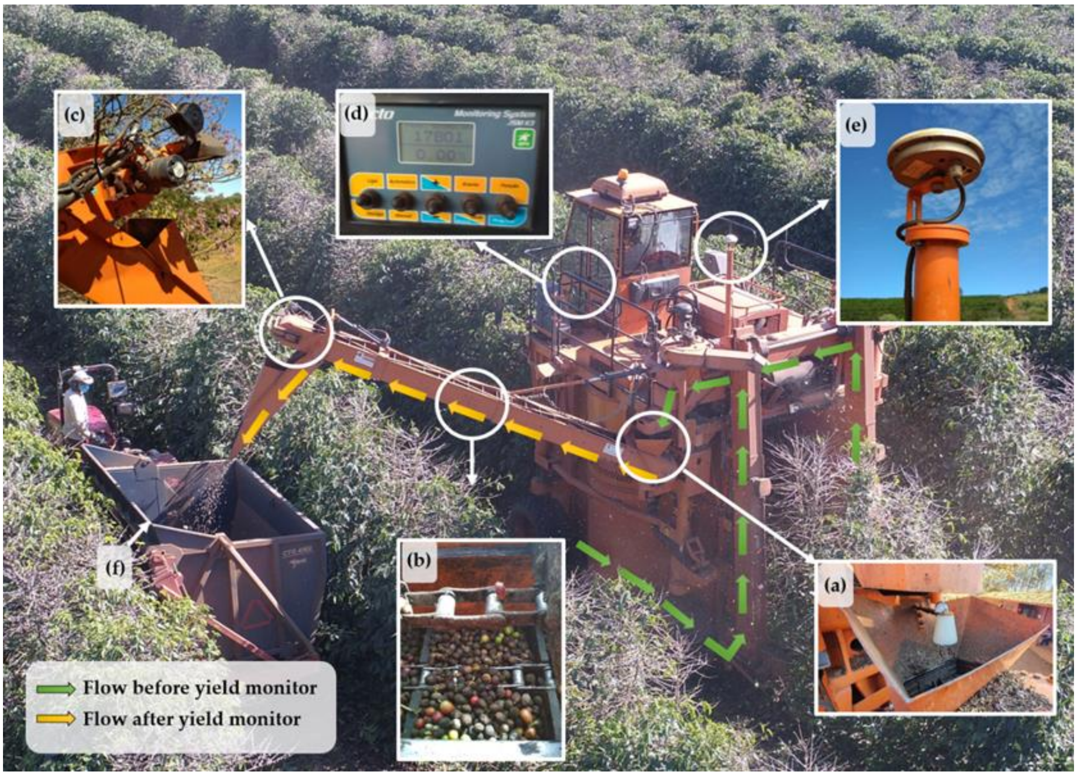

Figure 1. Flow of coffee fruits in the harvester and details of the components of the yield monitor, temporary reservoir, and ultrasonic sensor (a); paddle conveyor (b); hydraulic motor (c); data logger (d); GNSS receiver (e); coffee wagon (f).

At the same time, an experiment was carried out to compare the data obtained with the yield monitor with the simultaneous weighing using a coffee wagon instrumented with load cells (Figure 2). The measurement of harvester yield data was carried out using a big-bag fixed to a metal structure supported on four load cells with individual reading capacity of up to 1000 kg and with an accuracy of 10 g, on a coffee wagon, similar to the work performed by (BALASTREIRE et al., 2002). The load cells were calibrated and connected to a data logger responsible for sending the data at a frequency of 5 Hz to a notebook that performed the integration with the positioning data obtained by a GNSS SMART6-L™ (NovAtel Inc., Calgary, AB, Canada) with TerraStar-C correction that results in an accuracy of ±0.09 m (MALDANER et al., 2021a) (Figure 2e).

The evaluation of the quality of the data generated by the monitor was carried out in two crop rows, one with low and the other with high yield of coffee fruits of the Catuaí 144 variety, with a spacing between lines of 4.0 m and 0.5 m between plants, with an approximate population of 5000 plants ha^{-1}. In the low yield line, therefore, with low grain flow, data were collected in an extension of approximately 250 m and, in the high flow area, in an extension of approximately 87 m.

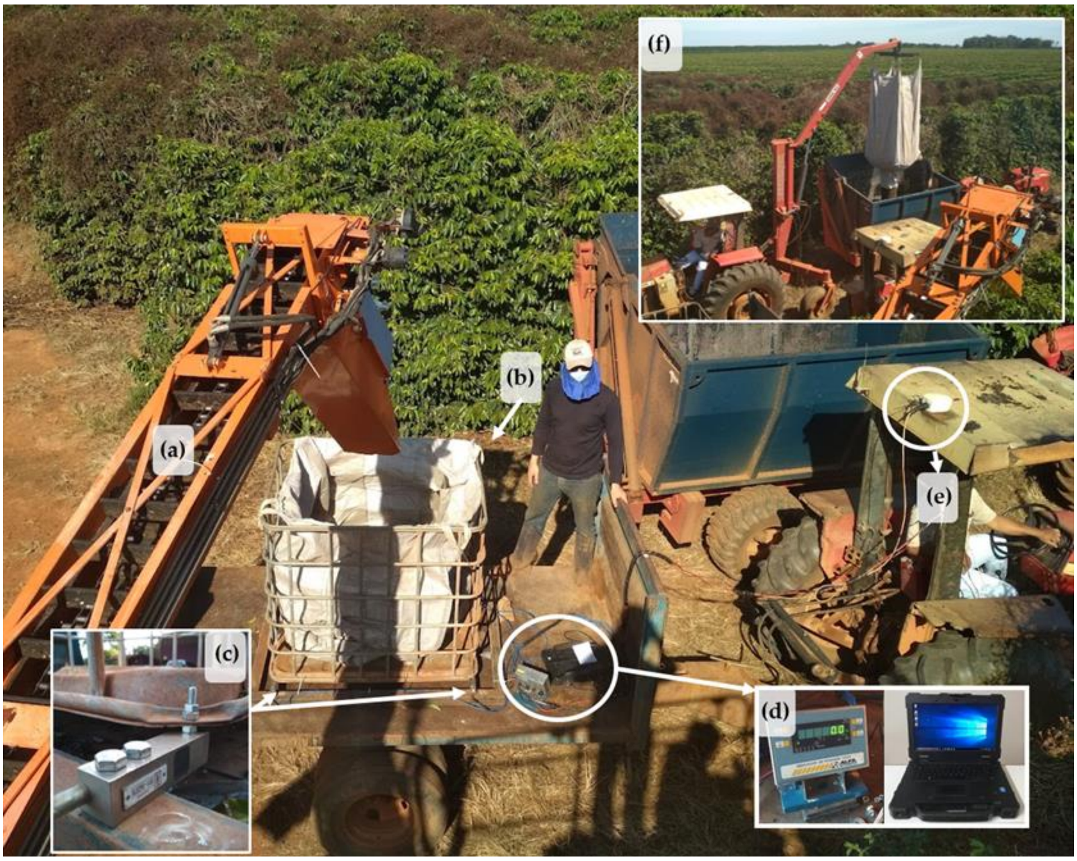

Figure 2. Details of the coffee wagon instrumentation used to measure the harvester's yield data, loading (**a**); big-bag (**b**); load cells (**c**); data logger (**d**); GNSS receiver (**e**); unloading (**f**).

In order to compare the data obtained by the harvester and the load cells, it was necessary to standardize the frequency between the data obtained using the distance traveled as an adjustment factor. The data were analyzed with the addition of a standard offset according to the delay distance between the harvester and the instrumented wagon and the data without the offset. To compare the yield estimates performed by the yield monitor versus the load cells, a linear model was adjusted, and its coefficient of determination (R^2) calculated.

To obtain the weight values of processed coffee (peeled and dried grains), it was necessary to obtain the CF. For this, 23, 44, and 31 samples were collected during the 2019, 2020, and 2021 harvests, respectively, at each coffee wagon, so random samples were collected throughout the area (Figure 3).

Each sample consists of one liter of raw coffee that was weighed and sent to dry until it reached approximately 12% moisture and then peeled and weighed again to obtain the CF. The final weight of the processed coffee was determined using Equation (1):

$$X=\left(\left(10000 * \frac{V_i}{D_i * W}\right) * CF\right)/1000 \quad (1)$$

where X is the yield of processed coffee (Mg ha^{-1}); V_i= volume accumulated over a time interval (L); $D_i$is the distance traveled in the same time interval (m), *W* is the width of the coffee line (4 m), and *CF* is (kg L^{-1}).

Figure 3. Collection flow and processes for obtaining the CF.

Yield data was filtered, eliminating headland maneuver data. Data close to roads that divide and cross the area and the discrepant data were filtered using the MapFilter 2.0 software using global filtering with a threshold of 100%, based on the methodology proposed by Maldaner and Molin (2020).

Descriptive statistical and geostatistical analyses followed by visual assessment of yield maps were performed to identify patterns in spatial variability (COLAÇO et al., 2019). The variability expressed by the coefficient of variation (CV) was classified as low (CV < 12%), median (12% < CV < 62%), and high (CV > 62%), according to Warrick and Nielsen (1980). Spatial dependence index (SDI) was calculated from the semivariograms as C0/(C0 + C1), where C0 is the nugget variance (non-spatial variance), and C0 + C1 is the sill variance (spatially dependent variance). SDI was interpreted as strong (<0.25), moderate (between 0.25 and 0.75), or weak (>0.75), according to Cambardella et al. (1994). The model was determined based on the lowest root mean squared error (RMSE) value.

Yield data were interpolated using the software Vesper 1.6 (MINASNY et al., 2005) using the ordinary kriging method with a spatial resolution of 3.0 × 3.0 m. For the temporal analysis of the data, the interpolated maps were normalized with values 0 to 1 from Equation (2):

$$xn_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (2)$$

where xn_i is the normalized yield at point *i* of the grid, and x_i is the yield at point *i* of the grid.

The identification of the temporal patterns of yield in the study area was first performed through the punctual correlation between the yield of the high-yield years. For this, the adjusted Pearson correlation coefficient was calculated for a linear model. At the same time, the normalized difference in yield between the years was calculated from Equation (3):

$$Dif_{t,n} = (xn_{i,t1} - xn_{i,t2})^2 \quad (3)$$

where $Dif_{t,n}$ is the normalized difference between the evaluated years, $xn_{i,t1}$ is the normalized yield at point *i* of the grid in year $t1$, and $xn_{i,t2}$ is the normalized yield in grid point *i* in year $t2$.

To evaluate the temporal behavior and stability of yield between the years, the average temporal variance (Equation (4)) was calculated, following the methodology proposed by Blackmore et al. (2003).

$$\sigma_i^2 = \frac{\sum_{21}^{19}(Y_{t.i}-\bar{Y}_t)^2}{3} \tag{4}$$

where: σ_i^2 is the temporal variance at grid point i, t is the time in years between 2019 and 2021, Y is the yield in years t at point i, and $\bar{Y}_t$ is the mean of the yield for the whole field in years t.

2.3. RESULTS AND DISCUSSION

2.3.1. Quality of Yield Data

In order to compare the data obtained by the yield monitor and the data obtained by the load cells, they were synchronized. This synchronization was necessary due to the path that the coffee fruits travel internally in the harvester. The harvest speed was approximately constant, and the displacement of the paddle conveyor was controlled by the ultrasound sensor, which depends on the amount of fruit being harvested. In this sense, it was necessary to adjust the distance between the beginning of the harvest and the begin-ning of the discharge to the coffee wagon.

The coffee wagon moved approximately 10 m until the first data was recorded by the harvester's yield monitor. The behavior of the coffee fruit flow obtained in the high yield row (Figure 4a) shows that 519 L and 359 kg of coffee were harvested at a distance of ap-proximately 87 m, representing a fruit flow of 5.96 L m^{-1} and 4.12 kg m^{-1}, respectively. For the low-yield row (Figure 4b), the data indicated a harvest of 435 L and 331 kg of coffee in 219 m, corresponding to a flow of 1.98 L m^{-1} and 1.51 kg m^{-1}, respectively, approximately three times lower than in the high-yield row. In addition, the distance lag between the beginning of the harvester's data recording and the arrival of material on the instrumented wagon was approximately 36 m, a distance 3.6 times greater than in the high flow condition. In this case, higher variability in fruit flow was also observed. The first peak of variation was observed for both the load cells and by the yield monitor shortly after the arrival of the first fruits to the coffee wagon when the harvester was already in operation and the fruit flow suddenly increased.

The high frequency of data collection obtained with the load cells on the wagon made it possible to clearly observe the variations in the flow of coffee fruits and, in some places, the offset between the data from the monitor and the load cells (Figure 4b). The fact that the measurement was made at different locations along with the coffee fruit flow most likely explains these differences. The ultrasonic sensor effectively controls the grain flow to en-sure that the paddle conveyor compartments remain full until unloading. The load cells measure the mass flow being unloaded at the end of the paddle conveyor. The flicker effect is caused by the lack of harvested coffee fruit to fill the cells, causing the paddle conveyor to not move. Consequently, small variations in the comparison of signals are expected, as the systems are not measuring the flow of fruits harvested at the same location and with the mechanical interference of the flow by the yield monitor.

The results are also presented considering the addition of the offset in the distance between the harvester and instrumented wagon data. The offsets adopted were 10 m and 36 m for the high and low flow lines, respectively (Figure 4c, d). After correcting the load cells' spatial offset, the data became more similar to those of the yield monitor, especially for the low flow line. Regardless, the yield monitor performed well in capturing the yield variability in the field.

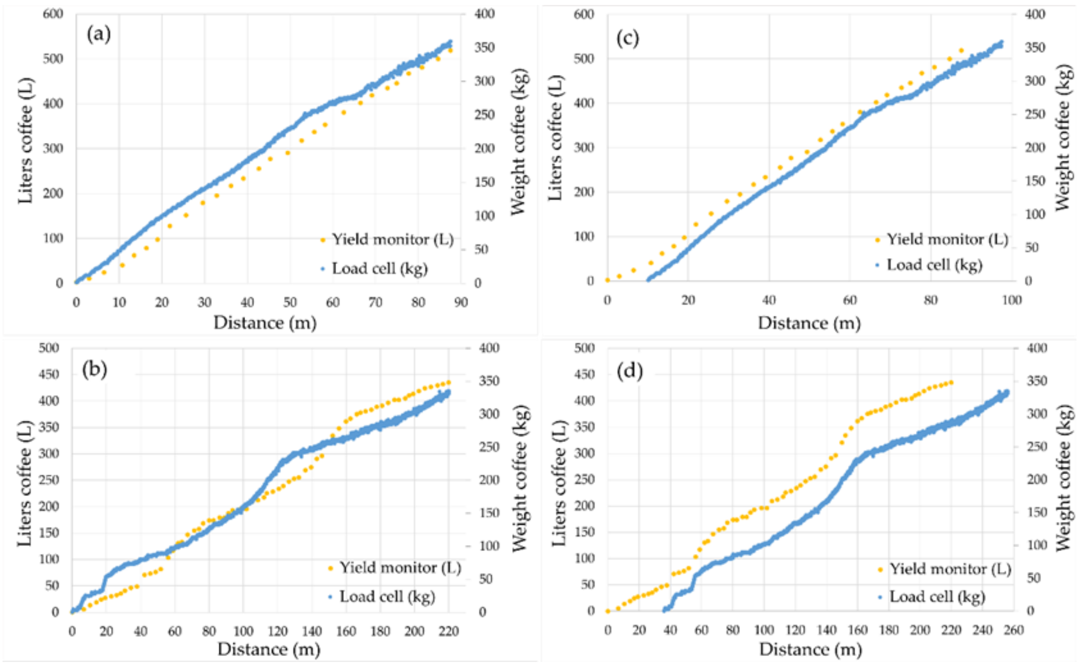

Figure 4. Distribution of coffee fruit harvest data obtained by the yield monitor and by load cells in the (**a**) high- and (**b**) low-yield rows, and with the addition of offset in the distance to the load cells wagon in the (**c**) high- and (**d**) low-yield rows.

The coffee fruit volume data obtained by the yield monitor and the fruit weight data obtained by the load cells on the wagon were initially correlated separately for each area without the addition of the offset. The fruit flow was regular for the high-yield row (Figure 5a), with no great variations in yield along the way, resulting in R^2 of 0.995. In the low-yield row (Figure 5b), the R^2 was 0.968. The joint analysis of the data referring to the two collections (Figure 5c) resulted in an R^2 of 0.975, indicating a high correlation between the data obtained at high frequency by weighing and those on fruit volume obtained by the yield monitor at relatively low frequency.

As for the data considering the offset, the differences are small for the high-yield row (Figure 5d) when compared to those without the offset (R^2 0.997), which can be explained by the constant and linear coffee fruit flow. For the low-yield row (Figure 5e), it is possible to notice a difference between the data that were not adjusted (Figure 5b). As the area presented greater variation in the fruit flow, the adjustment between the fruits leaving the harvester and the fruits entering the instrumented wagon proved to be efficient to perform the comparison, obtaining an R^2 of 0.997. When the data from the two collections are grouped (Figure 5f), it is possible to observe a good fit, with R^2 of 0.994.

The figures evidence the fact that the coffee mass data obtained from the load cells allowed greater detailing of the small flow changes during harvest. This behavior can be explained due to the low frequency in the collection of data from the yield monitor, as already observed by Maldaner et al. (2021b) when using an instrumented wagon to measure yield data in sugarcane harvesters.

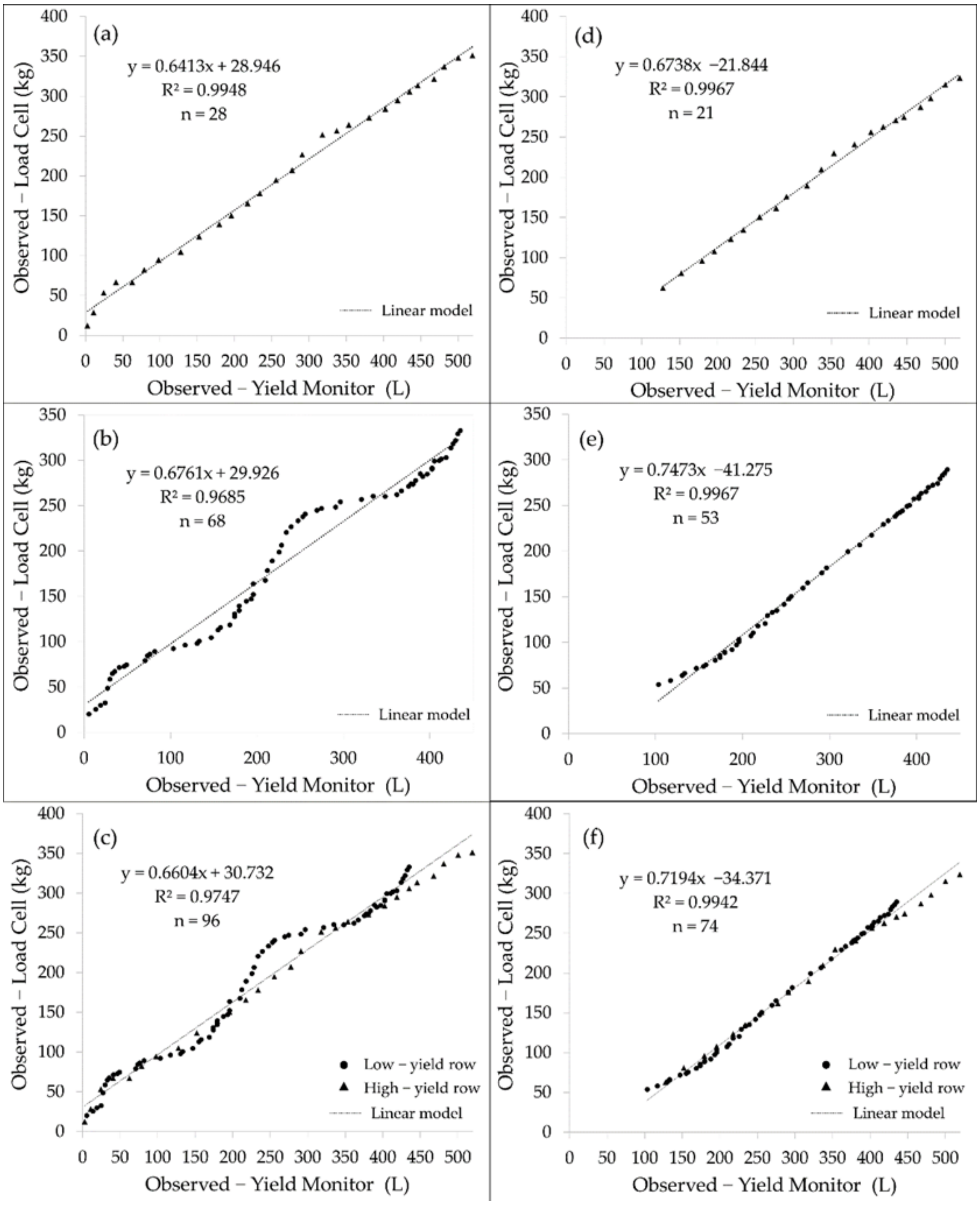

Figure 5. Linear regression of coffee harvest data obtained by the yield monitor and by load cells for the (**a**,**d**) high-yield row, (**b**,**e**) low-yield row, (**c**,**f**) and all data together (**a**–**c**) before and (**d**–**f**) after correcting the offset.

2.3.2. Temporal and Spatial Yield Variability

The data referring to coffee fruit samples obtained in the field over the three harvests and the calculated CF are presented in Table 1. It was possible to observe that in the years 2019 and 2021, the area was harvested when more than 80% of the fruits were overripe or dry. On the other hand, the 2020 harvest happened when 49.1% and 22.2% of the coffee fruits were classified as ripe and unripe, respectively. Analyzing the results, the 1.0 L of coffee samples convert to a higher mass of dried and peeled grains when the harvest pre-sented a higher percentage of overripe and dry coffee fruits, therefore, resulting in higher CF.

Table 1. Coffee fruit weight, maturity stage, and CF from samples collected over three consecutive harvests (2019 to 2021).

	Coffee	SampleDry	CoffeeProcessed	CoffeeUnripe	CoffeeRipe	CoffeeOverripe	Coffee
	Weight (g)	Humidity (%)	Weight—CF (g L^{-1})	Fruits (%)	Fruits (%)	Fruits (%)	
2019 (n = 23)	470.0	11.9	159.5	6.1	7.5	86.4	
SD	39.2	0.26	14.8	4.4	3.9	8.3	
CV (%)	8.3	2.2	9.3	72.2	51.7	9.4	
2020 (n = 44)	677.5	12.0	127.7	22.2	49.1	28.7	
SD	22.9	0.1	9.4	8.6	11.2	7.6	
CV (%)	3.4	1.2	7.4	39.0	22.9	26.5	
2021 (n = 31)	492.5	12.0	143.5	7.4	8.5	84.1	
SD	33.3	0.2	12.5	3.8	4.2	8.4	
CV (%)	6.8	1.5	8.7	51.8	48.7	10.1	

n: number of samples. SD: standard deviation. CV: coefficient of variation.

As highlighted by Pezzopane et al. (2005), as coffee fruits ripen, natural water loss is expected. Therefore, in the years 2019 and 2021, the amount of overripe coffee was greater and with the driest fruits. According to Silva et al. (2021), unripe coffee has humidity between 60–70%, ripe coffee 45–55%, and overripe coffee 30–40%. In the harvest years of 2019 and 2021, the higher percentage of overripe fruits led to a coffee fruit density ranging between 470.0 and 492.5 g L^{-1}. On the other hand, for the 2020 harvest, most of the fruits were in the ripe and unripe stages, resulting in an average coffee fruit density of 677.5 g L^{-1}. This variation in the maturation stage also influences the size of the fruit. The overripe coffee fruits are usually smaller than the unripe and ripe coffee fruits, therefore, due to the smaller volume, more fruits are sampled per liter, resulting in higher CF values after drying and processing.

The need to obtain a CF from field samples to convert volume to yield data may be one of the factors limiting the adoption of yield monitors embedded in harvesters. As an alternative to this challenge, Bazame et al. (2021) proposed the use of images associated with artificial intelligence to classify the grains and determine the stage of maturation in high spatial resolution.

During the 2019 harvest, data from some regions were not recorded by the yield monitor due to operational problems, resulting in a smaller amount of data when compared to the other years. However, even with this loss, the density of the filtered yield data ranged from 535 to 760 points ha^{-1}. The density of data collection was still high when compared to scientific approaches that also aimed to identify spatial variability and map coffee yield. Ferraz et al. (2012), Carvalho et al. (2017), and Ferraz et al. (2017) used manually collected point sampling techniques, and data collection ranged from 2.2 to 4.6 points ha^{-1}.

Descriptive statistics of raw and filtered yield data are presented in Table 2. Data showed variations between years, with 2019 being the year with the lowest average yield (1.29 Mg ha^{-1}). In the following year, the yield of processed coffee reached 2.03 Mg ha^{-1}, and in 2021, it again decreased to 1.44 Mg ha^{-1}. According to the classification of Warrick and Nielsen (1980), the values of coefficient of variation (CV) ranging from 43.43% to 49.04% indicate that there is median variability in yield in the area. Carvalho et al. (2017) highlighted that this may be a problem when only the average yield is considered for management decisions carried out for the entire area.

Table 2. Descriptive statistics of yield data after filtering for the three evaluated crops.

Year	Dataset	n	n ha⁻¹	Mean	Min	Max	SD	CV (%)
				Mg ha⁻¹				
2019	Original	6398	624.8	1.90	0	303.88	4.96	260.88
	Filtered	5479	535.0	1.29	0.12	2.68	0.57	44.35
2020	Original	8242	804.8	2.20	0.00	164.63	3.90	176.94
	Filtered	7789	760.6	2.03	0.09	4.06	0.88	43.43
2021	Original	7748	756.6	1.54	0.00	7.18	0.89	57.90
	Filtered	6913	675.1	1.44	0.13	2.90	0.71	49.04

n: number of samples. SD: standard deviation. CV: coefficient of variation.

Observing the average annual yield data, it is possible to identify the behavior of the crop, going from a year of low yield in 2019, to a year of high yield in 2020, and again showing low yield values in the following harvest (2021). In general, the coffee yield has a biennial behavior, with an alternation between a year of high yield, followed by a year of low yield. This variation is explained by the morphophysiological behavior of the coffee plant (PEREIRA et al., 2011; VALADARES et al., 2013).

In addition to the biennial variability, the yield data for the three years also show a large yield spatial variability in the area (Figure 6). Several factors can influence this variability, including the biennial characteristic (CAMARGO; CAMARGO 2001), the occurrence of diseases, pests and weeds (FIALHO et al., 2011; CARVALHO et al., 2012; LOPES et al., 2012), and soil fertility and foliar nutrition (WADT; DIAS 2012). The biennial participation is expressive in the variation of the coffee yield. Due to the physiological characteristics of the plant, it must have high vegetative activity during one year, so that the following year it can produce well (RENA; MAESTRI 1985). These high-density data (between 552.8 and 792.1 points ha^{-1}) show this behavior of the crop with a high level of detail.

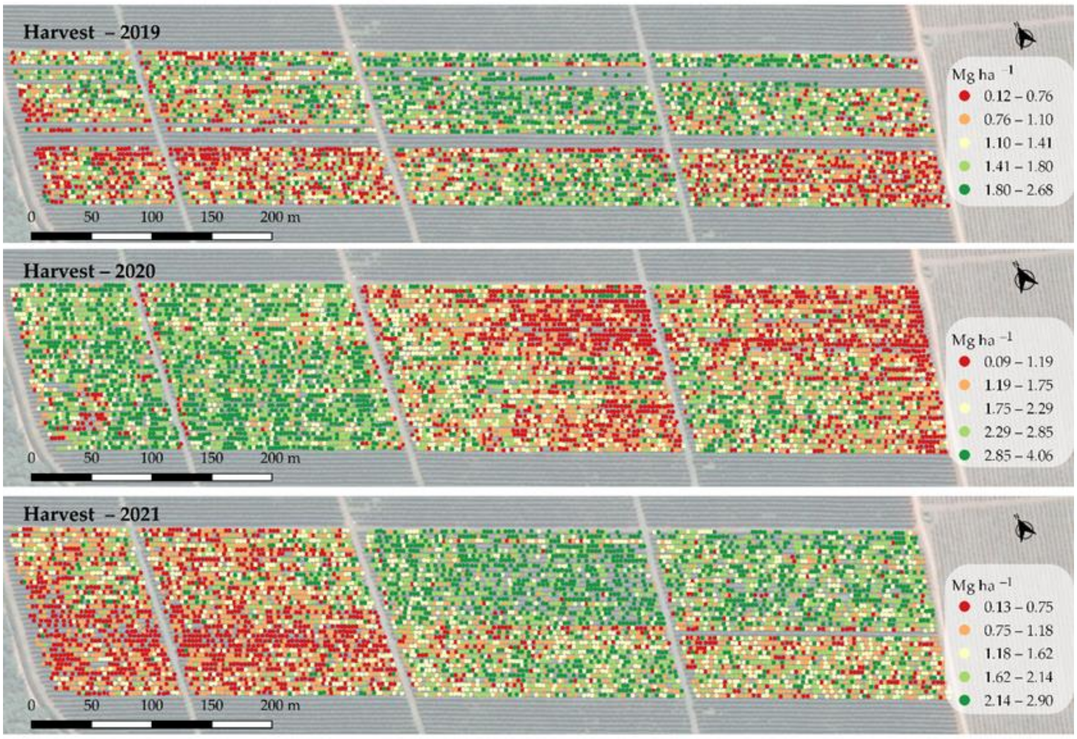

Figure 6. Filtered processed coffee yield map data obtained by the yield monitor for the 2019 harvest; 2020 and 2021.

The semivariograms of yield data (Table 3) fitted exponential models with RSME of 0.022 Mg ha^{-1} for the 2019 harvest, with RSME of 0.011 Mg ha^{-1} for 2020 and 0.016 Mg ha^{-1} for 2021 using the Gaussian model. The spatial dependence index (SDI) was moderate, similar to the results found by Molin et al. (2010), in which the authors used yield data obtained by a yield monitor and found a moderate SDI value fitted to the spherical model. In contrast, Carvalho et al. (2017) collected point yield data samples, with a density of 4.5 points ha^{-1} at regular spacings and reported strong spatial dependence for yield in two seasons, with semivariograms fitted to a spherical model. Silva et al. (2008), who sampled fruits of four plants around the crossing points of a sampling grid of 25 m (11 points ha^{-1}) to calculate the average yield per plant, also found a strong spatial dependence for the yield of two coffee plantations by adjusting spherical models.

Table 3. Semivariogram model and variables used to interpolate coffee yield data.

Year	Model	Range	Sill [1]	Nugget [2]	RSME (Mg ha^{-1})	SDI
2019	Exponential	59.9	0.343	0.214	0.022	moderate
2020	Gaussian	316.4	1.031	0.563	0.011	moderate
2021	Gaussian	202.2	0.596	0.345	0.016	moderate

1 Sill = maximum observed variability in data; 2 nugget = sources of error or variation at distances smaller than the sampling interval.

After adjusting the semivariograms, it was possible to estimate the yield in unsampled locations by interpolation and to standardize the size of each pixel of the maps for further spatial and temporal analysis.

The results of interpolation and normalization of yield data are shown in Figure 7. Regions with values close to 1 represent the maximum yield, while regions close to 0 represent the minimum yield for the year. A visual comparison shows some similarities between the years. The regions that had high yields in 2019 and 2021 were the same regions that had the lowest yields in 2020. Carvalho et al. (2017) also observed similar results. The explanation for this alternation may be related to fruiting. Plants that produced a lot in 2019 (regions with green coloring in Figure 7) used their reserves for fruiting, negatively influencing the growth of branches and, consequently, reducing yield in 2020 (regions with red coloring in Figure 7). According to Matiello et al. (2008), the coffee plant goes through a phase of low metabolism after high yields, and this behavior can cause this alternation of yield. It is important to point out that the biennial yield behavior is expressed inside and along the field, in addition to the spatial variability.

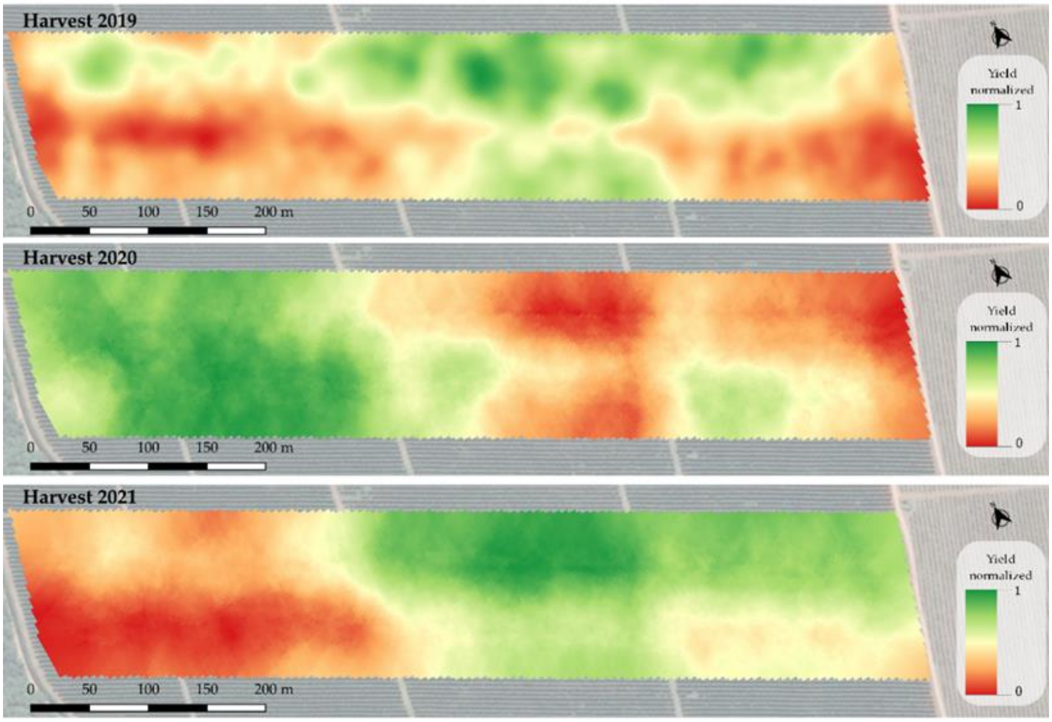

Figure 7. Normalized yield map for the harvest in 2019 (a), 2020 (b), and 2021(c).

When the normalized yield data are individually compared among the years, it is possible to observe this yield alternation in the area. In consecutive years (Figure 8a,c) it is possible to notice the yield inversion pattern, presenting negative correlation values r of −0.53 when comparing between the harvests of 2019 and 2020, and r of −0.87 in the comparison between the harvests of 2020 and 2021. These results corroborate the results found by Carvalho et al. (2017), which evaluated coffee yield over two years and also found a negative correlation between sequential years, with an r of −0.69. When the data from the odd years (2019 and 2021) are compared (Figure 8b), the behavior is opposite, with a positive correlation of r of 0.75. This indicates that there is an alternation of yield within the same area during the years and that the alternation regions are repeated.

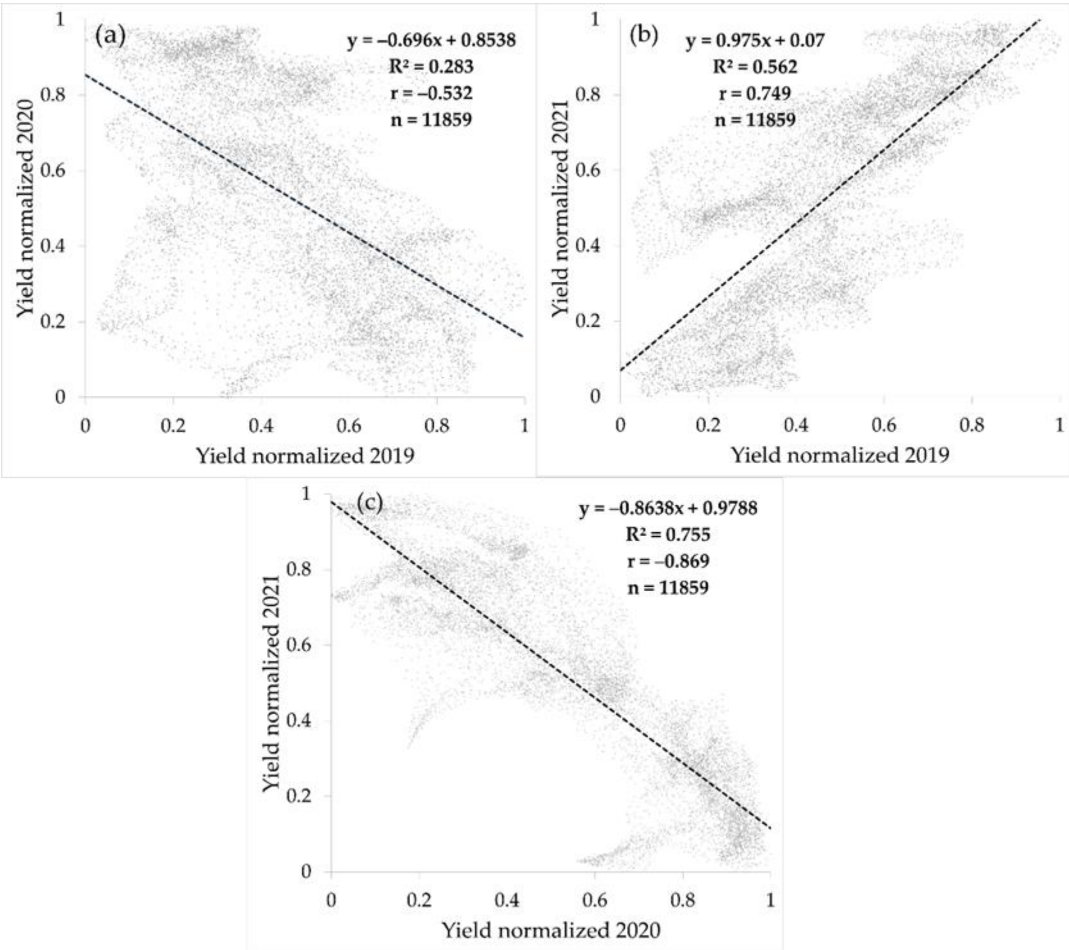

Figure 8. Correlation of normalized yield data between crops of 2019 and 2020 (a), 2019 and 2021 (b), and 2020 and 2021 (c).

The maps of the normalized difference in yield are shown in Figure 9. Values close to 1 indicate that the region has a greater difference in yield between the years evaluated, whereas values close to 0 show a low difference in yield between the years. Likewise, in the correlation data, it was possible to notice the alternating behavior of yield between the sequential years 2019 and 2020 (Figure 9a) and 2020 and 2021 (Figure 9c). The map of the difference between the odd-numbered years (2019 and 2021) showed a result of low difference, indicating the similarity between the productive regions (Figure 9b). This yield difference between the years presents an additional degree of difficulty in managing the spatial variability of coffee plantations, since temporal variability influences and requires a longer sequence of data for a better understanding of the variability and for making decisions.

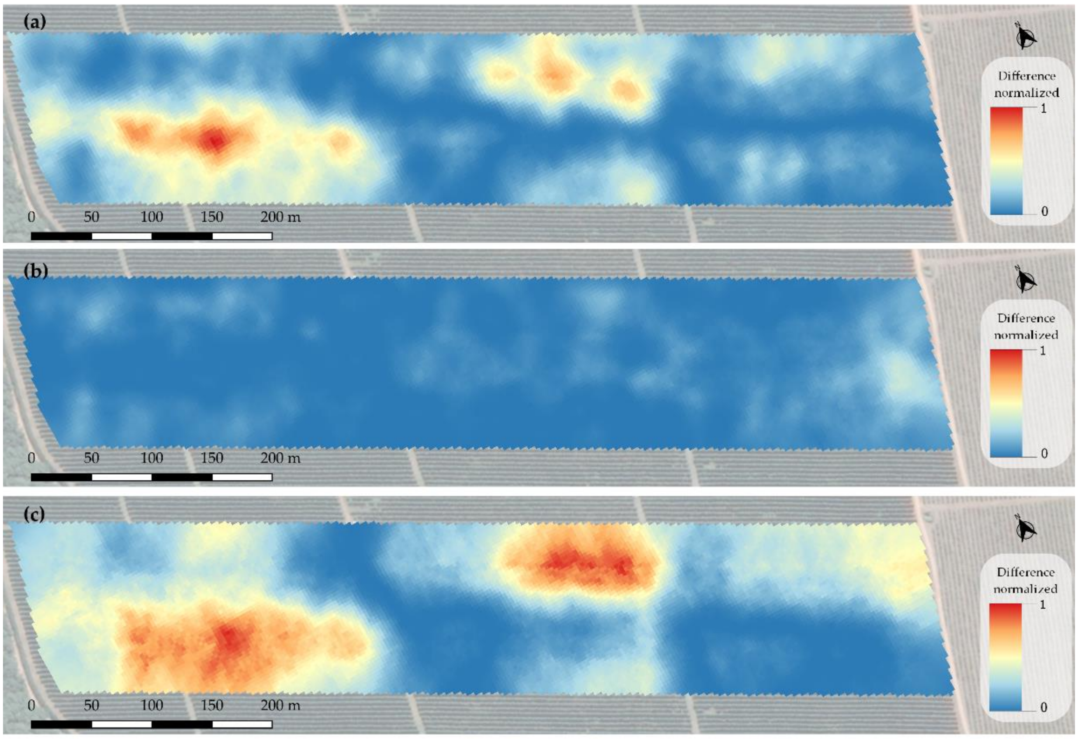

Figure 9. Normalized difference map between yield data between crops 2019 and 2020 (a), 2019 and 2021 (b), and 2020 and 2021 (c).

Following the methodology proposed by Blackmore et al., (2003), Figure 10 shows the map with the temporal variance yield for the three years evaluated. The regions that have the greatest variance in yield data are highlighted in the red areas, and the regions that have the most stability in yield data over the three years are highlighted in the blue areas. The regions with the highest variance were also the regions that showed the greatest differences between sequential years (Figure 9a,c). This result may indicate that, in addition to knowing the production variability, the variance information must also be considered in the strategies for site-specific management. The regions that presented high yield variance should possibly be conducted differently, receiving localized treatments alternated according to the productive year (high and low yield) and not only in consideration of their yield variability. As for the regions with low variance in the yield data, it is assumed that they present a more stable yield behavior over time, so they could be conducted taking into account only the yield variability.

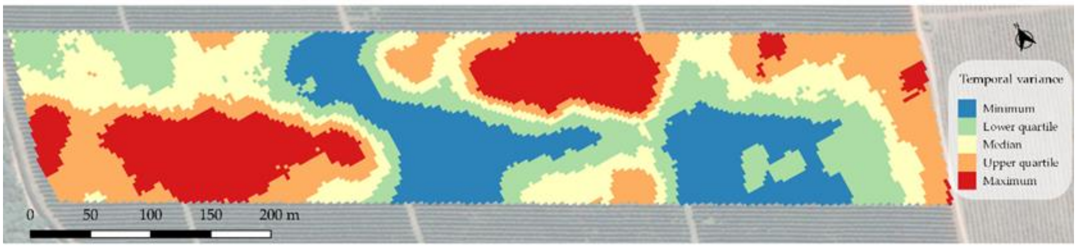

Figure 10. Yield map temporal variance along the three harvests.

2.4. CONCLUSIONS

The data obtained by the volumetric yield monitor embedded in a harvester showed a high correlation with those obtained with the load cells. This fact qualifies the yield data, allowing the identification of the fruit flow variation during harvest and the spatialization of these data to obtain reliable yield maps, with high data density.

It was possible to obtain yield data in a commercial area during three consecutive seasons, allowing the identification of its internal variability, as well as classifying regions that present alternating yields between the years evaluated. This result indicates that, in addition to knowing the spatial yield variability, the biennial variance information must also be considered in the strategies for site-specific management. Regions that presented high yield variance should be conducted differently, being alternated according to the productive year (high and low yield) and not only in consideration of their yield variability. This type of information helps in the search for possible causes of this variability so that the crop can be managed considering these spatial and temporal differences.

Future research aimed at estimating coffee yields should consider methods to identify fruits that were not measured on the monitor due to fruit drop before and/or during harvest, as well as methods to unify yield data for areas that perform more than one pass of the harvester during the same harvest. Moreover, it is necessary to develop faster and more assertive techniques to identify the degree of ripeness of the fruits, thus, in addition to the yield, it would be possible to obtain a new layer regarding the quality of the fruits, a subject that was not addressed in detail in this work and that is of great interest to coffee producers.

REFERENCES

ALVARES, A. C.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M.; SPAROVEK, G. Koppen's climate classification map for Brazil. **Meteorologische zeitschrift,** v. 22, n.6, p.711–728, 2013.

ANGNES, G.; MARTELLO, M.; FAULIN, G. D. C.; MOLIN, J. P.; ROMANELLI, T. L. Energy efficiency of variable rate fertilizer application in coffee production in Brazil. **AgriEngineering**, v.3, p815–826, 2021.

BALASTREIRE, L. A.; SCHUELLER, J. K.; AMARAL, J. R.; LEAL, J. C. G.; BAIO, F. H. R. Coffee Yield Mapping. ASAE Annu. Meet. 2002.

BAZAME, H. C.; MOLIN, J. P.; ALTHOFF, D.; MARTELLO, M. Detection, classification, and mapping of coffee fruits during harvest with computer vision. **Comput. Electron. Agric.** v.183, p.106066, 2021.

BLACKMORE, S.; GODWIN, R. J.; FOUNTAS, S. The analysis of spatial and temporal trends in yield map data over six years. **Biosystem. Engineering**, v. 84, p.455–466, 2003.

CAMARGO, A. P.; CAMARGO, M. B. P. Definition and outline for the phenological phases of arabic coffee under Brazilian tropical conditions. **Bragantia**, v.60, p.65–68, 2001.

CAMBARDELLA, C. A.; MOORMAN, T. B.; NOWAK, J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA, A. E. Field-scale variability of soil properties in central Iowa soils. **Soil Science Society of America Journal**, v.58, p.1501–1511, 1994.

CARVALHO, A. M.; MENDES, A. N.; BOTELHO, C. E.; OLIVEIRA, A. C.; REZENDE, J. C.; REZENDE, R. M. Agronomic performance of coffee cultivars resistant to coffee rust in Minas Gerais state, Brazil. **Bragantia**, v.71, p.481–487, 2012.

CARVALHO, L. C.; SILVA, F. M.; FERRAZ, G. A. S.; STRACIERI, J.; FERRAZ, P. F.; AMBROSANO, L. Geostatistical analysis of Arabic coffee yield in two crop seasons. **Rev. Bras. De Eng. Agrícola E Ambient**. v. 21, p.410–414, 2017.

COLAÇO, A. F.; MOLIN, J. P.; ROSELL-POLO, J. R.; ESCOLÀ, A. Spatial variability in commercial orange groves. Part 1: Canopy volume and height. **Precision Agriculture**, v.20, p.788–804, 2019.

COLAÇO, A. F.; PAGLIUCA, L. G.; ROMANELLI, T. L.; MOLIN, J. P. Economic viability, energy and nutrient balances of site-specific fertilization for citrus. **Biosystem Eng.**, v.200, p.138–156, 2020.

COMPANHIA NACIONAL DE ABASTECIMENTO—(CONAB) Acompanhamento da Safra Brasileira: Café. Bol. Safra 2021 **Café Brasília**, v.8, p.59, 2021.

COMPANHIA NACIONAL DE ABASTECIMENTO—(CONAB) Acompanhamento da Safra Brasileira: Café. Prim. Levant. 2022 Brasília, 1, 60, 2022.

FAULIN, G. C.; MOLIN, J. P.; STANISLAVSKI, W. M. Sample density and method for obtaining of the correction factor used in the coffee (Coffea Arabica L.) yield map. In **Anais** do Congresso Brasileiro de Agricultura de Precisão—ConBAP; São Pedro, São Paulo, Brazil, 2014.

FERRAZ, G. A. S.; DA SILVA, F. M.; CARVALHO, L. C. C.; ALVES, M. D. C.; FRANCO, B. C. Spatial and temporal variability of phosphorus, potassium and of the yield of a coffee field. **Eng. Agric**. v.32, p.140–150, 2012.

FERRAZ, G. A. S.; DA SILVA, F. M.; DE OLIVEIRA, M. S.; CUSTÓDIO, A. A. P.; FERRAZ, P. Spatial variability of plant attributes in a coffee plantation. **Rev. Ciênc. Agron.** v.48, p.81–91, 2017.

FIALHO, C. M. T.; SILVA, G. R.; FREITAS, M. A. M.; FRANÇA, A. C.; MELLO, C. A. D.; SILVA, A. A. Competition of weeds with coffee plants, in two times of infestation. **Planta Daninha**, v.28, p.969–978, 2011.

GASPARI-PEZZOPANE, C. D.; MEDINA FILHO, H. P.; BORDIGNON, R.; SIQUEIRA, W. J.; AMBRÓSIO, L. A.; MAZZAFERA, P. Environmental influences on the intrinsic outturn of coffee. **Bragantia,** v. 64, p.39–50, 2005.

INMET. Brazil Climate Normals 1991-2020. 2022 Available online: https://portal.inmet.gov.br/normais (accessed on 30 April 2022).

LOPES, P. R.; ARAÚJO, K. C. S.; FERRAZ, J. M. G.; LOPES, I. M.; FERNANDES, L. G. Producing agroecological coffee in Southern Minas Gerais: Alternative systems forintensive production of agrochemicals. **Rev. Bras. De Agroecol.** v.7, p. 25–38, 2012.

LOWENBERG-DEBOER, J.; ERICKSON, B. Setting the record straight on precision agriculture adoption. **Agron. J**. v.111, p.1552–1569, 2019.

MALDANER, L. F.; MOLIN, J. P. Data processing within rows for sugarcane yield mapping. **Scientia Agricola**, v.77, n.5, 2020.

MALDANER, L. F.; CANATA, T. F.; DIAS, C. T. S.; MOLIN, J. P. A statistical approach to static and dynamic tests for Global Navigation Satellite Systems receivers used in agricultural operations. **Scientia Agricicola**, v.78, n.5, 2021a.

MALDANER, L. F.; CANATA, T. F.; MOLIN, J. P. An approach to sugarcane yield estimation using sensors in the harvester and zigbee technology. **Sugar Tech**, v.24, p. 813–821, 2021b.

MATIELLO, J. B.; GARCIA, A. W. R.; ALMEIDA, S. R. **Adubação Racional na Lavoura Cafeeira**, 1st ed.; Bom Pastor: Varginha, Brazil, p. 114, 2008.

MINASNY, B.; MCBRATNEY, A. B.; WHELAN, B. M. VESPER Version 1.62; Australian Centre for Precision Agriculture, McMillan Building A05; The University of Sydney: Sydney, Australia, 2005.

MOLIN, J. P.; MOTOMIYA, A. V. D. A.; FRASSON, F. R.; FAULIN, G. D. C.; TOSTA, W. Test procedure for variable rate fertilizer on coffee. Acta Scientiarum. **Agronomy,** v.32, p.569–575, 2010.

MOLIN, J. P.; FAULIN, G. D. C. Spatial and temporal variability of soil electrical conductivity related to soil moisture. **Sci. Agric.**, v.70, p.01–05, 2013.

MOLIN, J. P.; AMARAL, L. R.; COLAÇO, A. Agricultura de Precisão, 1st ed.; **Oficina de Textos**: São Paulo, Brazil, p. 224, 2015.

MURUGAN, D.; GARG, A.; SINGH, D. Development of an Adaptive Approach for Precision Agriculture Monitoring with Drone and Satellite Data. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**,v.10,n.12, p.5322–5328, 2017.

PEREIRA, S. O.; BARTHOLO, G. F.; BALIZA, D. P.; SOGREIRA, F. M.; GUIMARÃES, R. J. Growth, productivity and bienniality of coffee plants according to cultivation spacing. **Pesquisa Agropecuária Brasileira**, v.46, p.152–160, 2011.

RENA, A.B.; MAESTRI, M. Fisiologia do cafeeiro. **Inf. Agropecuário**, v.11, p.26–40, 1985.

ROSA, V. G. C.; MOREIRA, M. A.; RUDORFF, B. F. T.; ADAMI, M. Coffee crop yield estimate using an agrometeorological-spectral model. **Pesqui. Agropecuária Bras.,** v.45, p.1478–1488, 2010.

SANTANA, L. S.; FERRAZ, G. A. S.; TEODORO, A. J. D. S.; SANTANA, M. S.; ROSSI, G.; PALCHETTI, E. Advances in Precision Coffee Growing Research: A Bibliometric Review. **Agronomy**, v.11, p.1557, 2021.

SANTINATO, F. **Inovações Tecnológicas Para Cafeicultura de Precisão**. Thesis. Faculdade de Ciências Agrárias e Veterinárias—UNESP, Jaboticabal, São Paulo, Brazil,125p, 2016.

SARTORI, S.; FAVA, J. F. M.; DOMINGUES, E. L.; RIBEIRO FILHO, A. C.; SHIRAISI, L. E. Mapping the spatial variability of coffee yield with mechanical harvester. In World Congress on Computers in Agriculture and Natural Resources; Foz do Iguaçu, Paraná, Brazil, 2002; pp. 196–205.

SILVA, F. M.; SOUZA, Z. M.; FIGUEIREDO, C. A. P.; VIEIRA, L. H. S.; OLIVEIRA, E. Spatial variability of chemical attributes and coffee productivity in two harvests. **Ciência E Agrotecnologia**, v.32, p.231–241, 2008.

SILVA, J. S.; MORELI, A. P.; DONZELES, S. M. L.; SOARES, S. F.; VITOR, D. G. Harvesting, Drying and Storage of Coffee. In Food Engineering Series; **Springer**: Berlin/Heidelberg, Germany,p. 1–64, 2021.

VALADARES, S. V.; NEVES, J. C. L.; ROSA, G. N. G. P.; MARTINEZ, H. E. P.; VENEGAS, V. H. A.; LIMA, P. C. Productivity and biennial production of dense coffee plantations under diferente levels of N and K. **Pesquisa Agropecuária Brasileira**, v. 98, p.296–303, 2013.

WADT, P. G. S.; DIAS, J. R. M. Regional and inter-regional DRIS norms for nutritional evaluation of Conilon coffee. **Pesquisa Agropecuária Brasileira**. V.47, p.822–830, 2012.

WARRICK, A. W.; NIELSEN, D. R. Spatial variability of soil physical properties in the field. In **Applications of Soil Physics**; Hillel, D., Ed.; Academic Press, New York, NY, USA, pp. 319–344, 1980.

3. USE OF ACTIVE OPTICAL SENSOR IN COFFEE FOR MONITORING CROP YIELD

ABSTRACT

Monitoring the spatial variability of agricultural variables is one of the main steps for implementing PA techniques. The active optical sensors (AOS) can be compatible with its instrumentation directly on agricultural machines, which makes it possible to obtain data with high frequency. Studies have already been developed using AOS for annual crops, however there are few studies on perennial crops such as coffee. This study aimed to evaluate the potential of AOS to map the spatial and temporal variability of coffee crop yields, as well as to establish guidelines for the acquisition of AOS data for sensing the side of coffee plant, allowing the evaluation of large commercial fields. The study was conducted in a commercial coffee area of 10.24 ha, cultivated with the Catuaí 144 variety. Data collection was performed with six Crop Circle ACS 430 sensors (Holland Scientific, Lincoln, NE, USA) and two N-Sensor NG sensors (Yara International, Dülmen, Germany). Seven field expeditions were made for data collection with the optical sensors over the years 2019 to 2021, collecting data during the flowering, fruit filling, fruit maturation phases (pre-harvest) and post-harvest. The results showed that the different faces of the same plant present different correlations with its yield. The face with the highest exposure to solar radiation presented a slightly higher correlation with yield when compared with the face with less exposure. In adition, it was observed that the vegetation indices acquired in the post-harvest and flowering stages have positive correlations with the yield of that next year, while vegetation indices acquired during fruit filling and pre-harvest stages have negative correlations with coffee crop yields of that year.

Keywords: Precision agriculture; Sensors; Spatial variability; Temporal variability

3.1. INTRODUCTION

Precision Agriculture (PA) approaches seek to understand the spatial and temporal variability present in agricultural variables within a production field. This knowledge allows for the localized management of inputs, aiming to increase production efficiency and, therefore, the profitability and sustainability of agricultural production (GEBBERS, ADAMCHUK, 2010). Monitoring the spatial variability of agricultural variables is one of the main steps for implementing PA techniques. This step routinely demands a high spatial density of data for the construction of reliable maps that characterize details of the crop. In addition, monitoring large-scale crops (e.g., hundreds of hectares) requires practical, inexpensive technologies that are compatible with the automation of the monitoring process.

Remote sensing approaches are classic alternatives for monitoring agricultural variables on large extensions of land, allowing spatial resolution from a few meters (e.g., 2 - 30 m for satellite images) up to sub-metric values (e.g., remotely piloted aircraft images). On the other hand, proximal sensing using active optical sensors (AOS) can be compatible with its instrumentation directly on agricultural machines, which makes it possible to obtain data with high frequency (e.g., 5 Hz), practically more than one piece of information per plant. In addition, as they are active, they do not depend on external radiation sources. AOS is less affected by variability in lighting and shading, which is common in field data acquisition, e.g. the presence of clouds or variation in solar position, factors to consider when using orbital and aerial remote sensing approaches.

In agriculture, AOS was initially developed to identify changes in the nitrogen (N) nutritional status and estimate its demand. This is possible because these sensors are closely related to photosynthetic activity and vegetative vigor of plants, both factors related to N absorption in most crops (AULA et al., 2020). An advantage of these sensors is that they do not depend on post-processing steps, as imaging sensors do. This allows the synchronization of collection, diagnosis, prescription, and application in one single step, making real-time interventions possible (PALLOTTINO et al., 2019). Studies have already been developed using AOS for crops such as maize (KITCHEN et al., 2010; SCHMIDT et al., 2011; SOLARI et al., 2008; TEAL et al., 2006), wheat (BERNTSEN et al., 2006; GIRMA et al., 2006; GROHS et al., 2009; RAUN et al., 2001), rice (XUE et al., 2014), cotton (TREVISAN et al., 2018) and sugarcane (AMARAL et al., 2015; AMARAL et al., 2017; PORTZ et al., 2012). However, these approaches have not been explored in depth for coffee cultivation, with only a few scientific studies evaluating the performance of AOS to estimate biometric parameters of individual plants (PUTRA et al., 2018; OLIVEIRA et al., 2021), that is, which is not applicable for large scales such as in commercial areas.

Unlike annual crops, AOS sensing for coffee should be performed with equipment turned to the side view of the plants since positioning the sensor at the top of the plant is unfeasible due to the height of the canopy. Although the lateral positioning of active sensors has already been adopted in some studies that monitored perennial crops (ANASTASIOU et al., 2018; DARRA et al., 2021; YU et al., 2021), there is still a need to establish an optimized protocol for the acquisition and processing of these data for coffee crop in order to evaluate its spatial variability over its biennial cycle. For example, in crops with planting lines oriented in an east-west direction, there is no consensus on using data from only one of the side faces (e.g., using only the face that does not receive direct sunlight) or if the best strategy is to use a methodology to calculate the average of both faces. This fact must be considered, as there is a consensus that the face that receives more solar radiation has a different behavior, resulting in different vegetative vigor if the same plant is evaluated from opposite faces/sides (SANTINI et al., 2019; CHAVES et al., 2012; DAMATTA, 2004; CANNELL, 1976; MONTOYA et al., 1961; CASTILLO; LOPEZ, 1966). Which of these faces best represents the productive potential of the crop? In addition, there is also no consensus on the height of the plant that best represents the yield variability; that is, which part of the coffee plants should be scanned? The lower, intermediate, or upper third of coffee plants? To evaluate spatio-temporal variability using AOS, these questions still need to be answered. Finally, it is important to mention that, to the best of our knowledge, there is no long-term research evaluating the response of AOS in the coffee crop over the biennial cycle typical of this crop.

In this context, this study aims to evaluate the potential of AOS to map the spatial and temporal variability of coffee crop yields, as well as to establish guidelines for the acquisition of AOS data for sensing the side of the coffee plant, allowing the evaluation of large commercial fields. More specifically, this study aims to (i) evaluate and propose a strategy for acquiring AOS data aimed at the side of coffee plants; (ii) to assess the spatial relationship between data from different commercial AOS and coffee yield throughout its biennial cycle over three seasons.

3.2. MATERIALS E METHODS

3.2.1. Study area

The study was conducted in a commercial coffee area of 10.24 ha, cultivated with the species *Coffea arabica L.*, IAC Catuaí 144 variety. The coffee trees were spaced by 0.5 m between plants and 4.0 m between rows, resulting in a density of approximately 5,000 plants ha^{-1}, with a drip irrigation system. The area (Figure 11) is located in the municipality of Patos de Minas, state of Minas Gerais, Brazil. The local climate is classified as Aw, according to the Köppen classification (ALVARES et al., 2013), characterized by a tropical environment with dry winters and rainy summers. The coffee lines were planted in east-west direction so that one side of the plants receives direct solar radiation most of the year and the other side receive indirect radiation (Figure 11a).

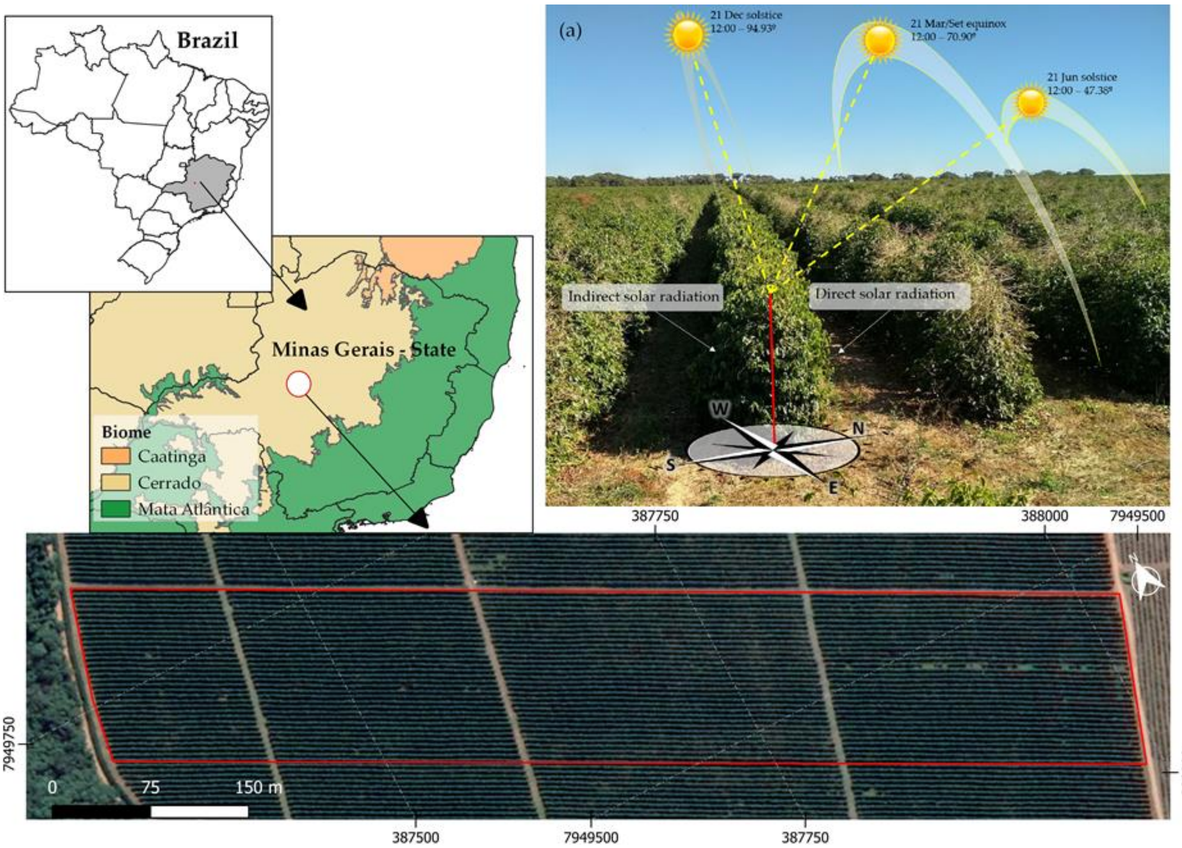

Figure 11. Map of the study area location; (a) detail for the face with an incidence of direct and indirect solar radiation and the predominance of annual solar incidence in the area.

3.2.2. Tractor instrumentation with AOS

Data collection was performed with six Crop Circle ACS 430 sensors (Holland Scientific, Lincoln, NE, USA) and two N-Sensor NG sensors (Yara International, Dülmen, Germany). The sensors were embedded in an agricultural tractor. Sensor specifications are shown in Table 4. The arrangement of sensors on the machine was previously defined in a 3D environment to avoid overlapping the lighting emitted by the active sensors (Figure 12a). Details of sensor positioning are shown in Figure 12.

The Crop Circle sensors were positioned at three different heights to simultaneously scan the lower (0 to 0.8 m, designated as TI), middle (0.8 to 1.6 m, designated as TM) and upper (1.6 to 2.4 m, designated as TS) thirds of the coffee plants. The N-Sensor sensors were positioned to scan only the middle third of the coffee plants. Half of the sensors collected data from the right row with the tractor moving between the coffee rows and the other half from the left row. All equipment worked with a data acquisition frequency of 5 Hz, obtaining approximately 1 point on each side of the plant every 0.3 m (considering an average traveling speed of 1.5 m s^{-1}).

The Crop Circle sensors were integrated with the GEOSCOUT GLS-420 mapping system (Holland Scientific, Lincoln, NE, USA) for data acquisition, and the NDVI (ROUSE et al., 1974) and NDRE (BARNES et al., 2000) vegetation indices were calculated, being designated in this study as CC-NDVI and CC-NDRE, respectively. The N-Sensor data acquisition was made in Yara N-Sensor software running on Microsoft Windows operating system. The sensor provides a vegetation index (designated in this work as NS), which the company does not open its specifications. All data were georeferenced using a Global Navigation Satellite System (GNSS) receiver (SMART6-L™, NovAtel Inc., Calgary, AB, Canada) with TerraStar-C (NovAtel Inc., Calgary, AB, Canada) correction that allows the accuracy of ±0.09 m (MALDANER et al., 2021a).

Table 4. Technical specifications of sensors

Specification	Crop Circle ACS 430	N-Sensor NG
Light source	Polychromatic modulated LED	LED (light-emitting diodes)
	670 nm (RED)	
Spectral Bands	730 nm (RedEdge)	730 nm (RedEdge)
	780 nm (NIR)	770 nm (NIR)

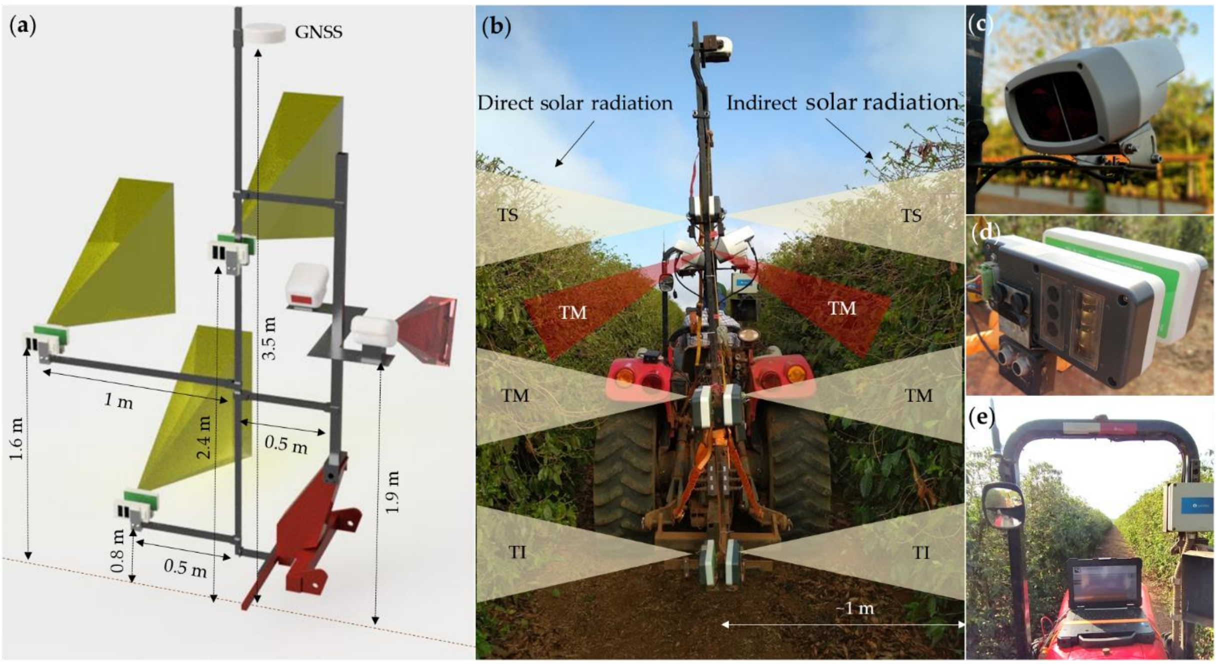

Figure 12. Schematic layout of sensors coupled on the tractor; (a) position of each sensor together with the GNSS receiver, showing the field of view of the Crop Circle sensors in yellow and the field of view of the N-Sensor NG sensors in red; (b) position of the plant scanned by each piece of equipment, showing the collections carried out in the lower third (TI), middle third (TM) and upper third (TS) of the coffee plant; (c) N-Sensor sensor; (d) Crop Circle sensor; (e) Equipment data logger.

3.2.3. Data acquisition with AOS and yield measurement

Seven field expeditions (designated from C1 to C7) were made for data collection with the optical sensors over the years 2019 to 2021 (Figure 13), collecting data during the flowering, fruit filling, fruit maturation phases (pre-harvest) and post-harvest. Harvests were carried out in 2019, 2020, and 2021 using a K3 Millennium harvester (Jacto, Pompeia, Brazil), equipped with a yield monitor that measures the volume of harvested grains converted to weight using a conversion factor. More details of the yield monitor and the conversion from volumetric to gravimetric units are described by Martello et al. (2022). Yield data collected from maneuvers and roads were manually removed. Subsequently, the discrepant data were filtered using the software MapFilter 2.0 (MALDANER et al., 2022), using global filtering with a threshold of 100% (MALDANER; MOLIN, 2020).

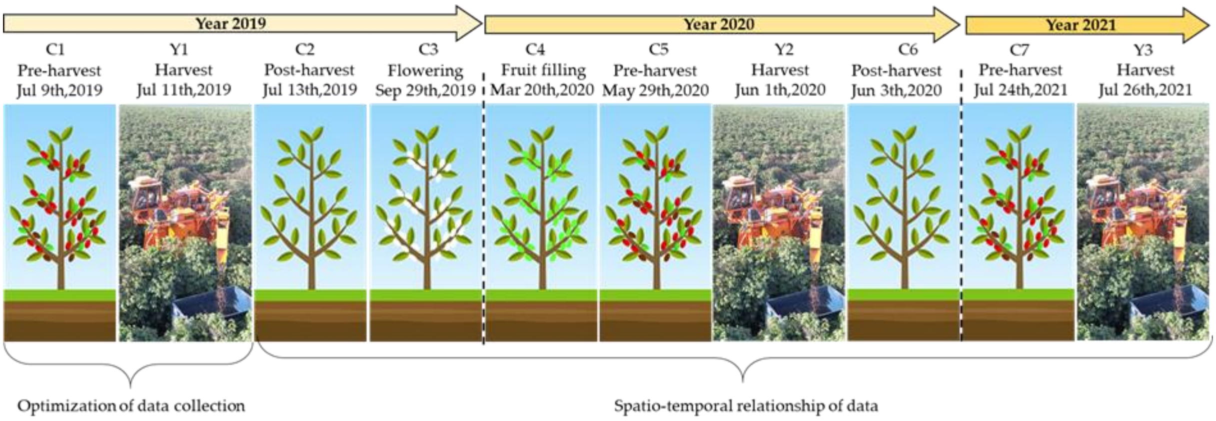

Figure 13. Flowchart of dates and coffee's phenological stages at data collection with optical sensors and yield measurement.

3.2.4. Data analysis

3.2.4.1. Optimizing AOS positioning on the side of coffee plant for inferring its yield

Three data acquisition scenarios were evaluated: (i) using only data obtained on the face exposed to direct sunlight (A1); (ii) using only the data obtained on the face with indirect sun exposure (A2); and (iii) using the average of the two faces for the same plant (A3). To categorize the sensor data as a function of the face exposed to direct and indirect sunlight, an algorithm was developed that recognizes the direction of the tractor's displacement as a function of the azimuth of its coordinates. The points were classified according to the direction of the tractor displacement (Figure 14a) and, then, it was applied an offset of 2 m to reproject the position to the center of the row at 90º and 270º from the alignment (Figure 14b). This was necessary because the equipment positioned to the right and left of the tractor scan different faces as the direction of displacement is reversed.

The optimization of sensor height was performed using the A3 scenario. These optimizations were performed using the sensors data obtained in the C1 stage that were compared with the yield data obtained in the Y1 harvest. To ensure that the yield data obtained by the yield monitor were compared with the data obtained by the sensors, a polygon (measuring 3 m wide by 4 m long) was generated for each yield point (Figure 14e). Then, the average of the sensor data within this polygon was calculated, as shown in Figure 14e. The correlation between the crop yield and the sensors' data of the different scenarios described above was evaluated in 5382 points. The optimizations regarding the sensor height and data acquisition scenarios were performed using the sensors' raw data without spatial filters to avoid possible interference. The scenario that presented the best correlation with the yield data was used in the subsequent analyzes (described below).

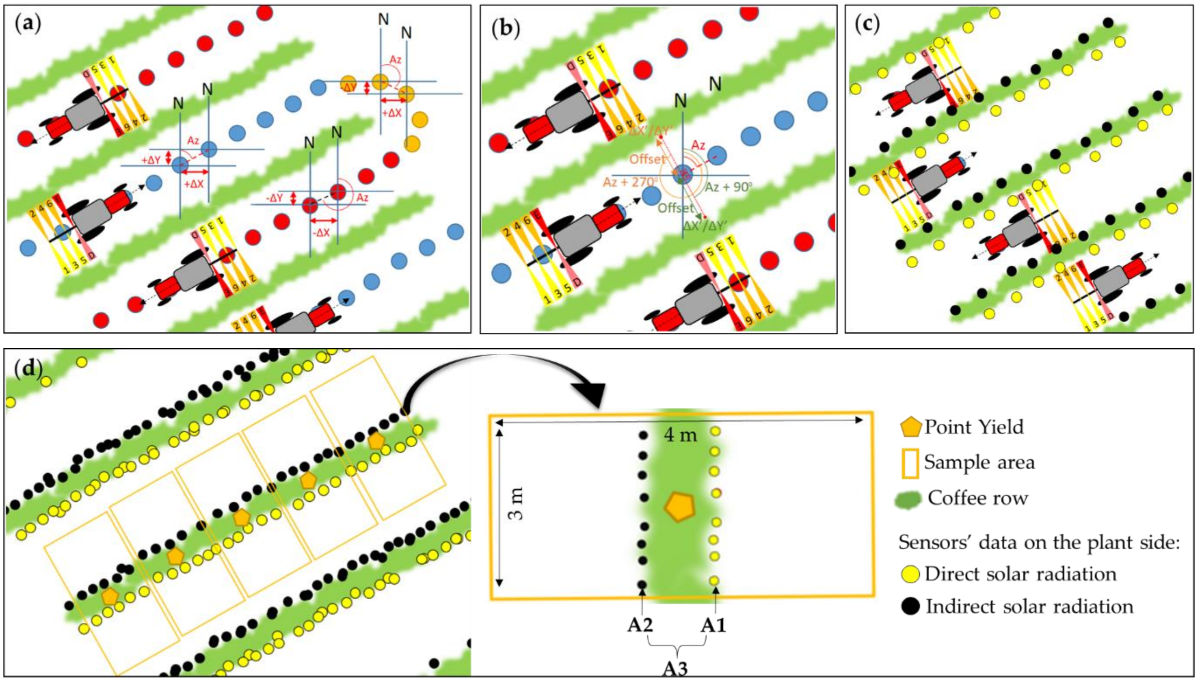

Figure 14. Schematic of data processing: (a) calculation of the direction and azimuth of the sensors' data points, blue and red points indicate the position where the data were obtained from the sensors before post-processing for the categorization of the plant side. In the red dots the sensors to the right of the tractor obtained information from the side with direct solar radiation, and to the left obtained information from the side with indirect solar radiation. In the blue points, the sensors on the right obtained information from the side with indirect solar radiation and the sensors on the left from the side with direct solar radiation; (b) applying the offset and direction to each point; (c) categorization of points according to plant side; (d) polygons created over the yield data point to sample the sensors data, obtaining the data of scenarios A1 (mean using only yellow points), A2 (mean using only black points), and A3 (mean using yellow and black points).

3.2.4.2. Spatio-temporal relationship of AOS data with coffee crop yield

Using the optimal scenarios of sensor height and data acquisition, the AOS data obtained from C1 to C7 were filtered using the MapFilter software (MALDANER et al., 2022) and then interpolated using ordinary kriging and a 3x3 m grid (similar grid to the one used for the yield data, described in Section 3.2.4.1). The AOS data were filtered, using global filtering with a threshold of 100% and local filtering with a threshold of 10% and a radius of 10 m. The spatio-temporal relationship of AOS data with coffee crop yield was evaluated by comparing the AOS maps of C1, C2, C3, C4, and C5 with yield map of Y2, as well as comparing the yield map of Y3 with the AOS maps of C5, C6, and C7. These data were compared using Pearson's correlation coefficient and the ability of the AOS data to predict yield was verified using simple linear regressions (n = 11,859, for all datasets). These comparisons permit to evaluate the temporal relationship of the sensor data with the coffee yield, showing its temporal oscillation as a function of its biennial seasonality.

The calibrations of the predictive models were validated with full cross-validation using as quality indicators the coefficient of determination (R^2), the residual prediction deviation (RPD), and the root means square error (RMSE). The RPD was calculated as the ratio between the standard deviation of the measured yield and the RMSE obtained in the prediction. Four RPD classes adapted from Chang et al. (2001) were used to evaluate the quality of models: poor models (RPD < 1.40), reasonable models (1.40 ≤ RPD < 2.00), good models (2.00 ≤ RPD < 3.00), and excellent models (RPD ≥ 3.00).

3.3. Results and Discussion

3.3.1. Optimizing AOS positioning on the side of the coffee plant for yield inference

Table 5 presents the correlation between the coffee yield data (in Y1) and the different scenarios of data acquisition (A1, A2 and A3) and sensor height (TI, TM, and TS). We observed that, in general, both commercial sensors showed better results in the A3 scenario (-0.34 ≤ r ≤ -0.21) in relation to the A1 (-0.34 ≤ r ≤ -0.17) and A2 (-0.27 ≤ r ≤ -0.15) scenarios. The A3 scenario represents the average value of the vegetation index obtained on each face of the plant, which probably promotes a better characterization of the average vigor of the plant and, consequently, a better relationship with its yield. These results suggest that, when AOS sensors are used to infer about coffee crop yield, it is necessary to post-process its data in order to organize them so that a plant is represented by the average of the vegetation indices collected on each of its sides. This post-processing is necessary because when traveling in the inter-row, the sensors mounted on the right and left of the tractor collect data from different coffee rows. These results even allow suggesting to AOS manufacturers that this data post-processing algorithm may be included in the intelligence of sensors used on the side of perennial crops. This tool would benefit conducting further studies comparing the relationship of the different plant faces with the vigor and yield of these crops.

When comparing only the A1 and A2 scenarios, the results showed that the different faces of the same plant present different correlations with its yield. The face with the highest exposure to solar radiation (A1) presented a slightly higher correlation with yield when compared with the face with less exposure (A2). This result may be linked to the fact that the face that receives the greatest amount of direct radiation tends to present a higher yield compared to the shaded face, with r oscillating from -0.34 to -0.17 for the A1 scenario and from -0.27 to -0.15 to A2 (Table 2). The higher yields on faces with greater sun exposure are explained by the fact that this face presents higher photosynthetically active radiation (CHAVES et al., 2012) and higher carbohydrate assimilation (DAMATTA, 2004), when compared with the face that receives less radiation. In addition, shaded faces form and fewer nodes per branch and fewer flower buds at existing nodes (MONTOYA et al., 1961; CASTILLO; LOPEZ, 1966).

Regarding the vegetation index, the CC-NDVI was the one that presented the smallest correlations with the yield data. This behavior may be linked to the saturation of this index that has already been observed in other crops with intense vegetative growth, such as soybean (MORLIN et al., 2019), rice (LU et al., 2017) and vine (JUNGES et al., 2017). An alternative to avoid saturation is to use vegetation indices that use the red-edge region instead of the near-infrared, such as the NDRE index (TASKOS et al., 2014). Our results corroborate this, showing higher correlations of yield with NDRE than with NDVI in all data acquisition scenarios (Table 2).

Regarding the optimization of sensor positioning, we observed that the data obtained in the TM of the plant showed a higher correlation with yield (r of -0.21 and -0.34 for NDVI and NDRE, respectively) in relation to those obtained in TI (r of -0.07 and -0.11 for NDVI and NDRE, respectively) and TS (r of -0.07 and -0.26 for NDVI and NDRE, respectively) (Table 2). The higher correlation between yield and TM is probably related to the fact that TM has a higher presence of vegetative structures and chlorophyll accumulation compared to TI and TS (SANTINI et al., 2019). Additionally, TI has a predominance of older branches and TS of younger branches, which are naturally less productive.

Table 5. Pearson's correlation between coffee yield data and AOS data obtained in different acquisition scenarios and positioning in relation to plant height.

Sensor	VI	Optimization of the acquisition scenario			Optimization of sensor positioning		
		A1	A2	A3	TI	TM	TS
Crop Circle	CC-NDVI	-0.17	-0.15	-0.21	-0.07	-0.21	-0.07
	CC-NDRE	-0.33	-0.27	-0.34	-0.11	-0.34	-0.26
N-Sensor	NS	-0.34	-0.27	-0.34			

3.3.2. Spatio-temporal relationship of AOS data with coffee yield

Table 6 shows the spatio-temporal relationship of the AOS data with the yield obtained in the 2020 (Y2) and 2021 (Y3) harvests. When analyzing the correlation of the sensor's output with the coffee yield data, we observed that although the correlation remains strong (with r absolute greater than 0.70) throughout the cycle, its sign is inverted in the middle of the coffee development (between C3 and C4). In other words, direct relationships with positive correlation ($0.73 \leq r \leq 0.91$) are observed with C1 (pre-harvest regarding Y1), C2 (post-harvest regarding Y1), and C3 (flowering) with the yield-Y2, and indirect relationships with negative correlation ($-0.93 \leq r \leq -0.77$) are seen with C4 (fruit filling) and C5 (pre-harvest regarding Y2) with the yield-Y2 (Table 3). This behavior was confirmed when assessing the data of the subsequent season, where the OAS data of C5 (pre-harvest regarding Y2) and C6 (post-harvest regarding Y2) presented positive correlations ($0.72 \leq r \leq 0.90$) with the yield-Y3 and the data acquired in C7 (pre-harvest regarding Y3) had negative correlations ($r = -0.71$) with yield-Y3 (Table 6). The inversion of the relationship between OAS output and coffee yield occurred after the flowering, i.e., after the beginning of the rains in the Cerrado region, which occurs at the end of September. This moment of the coffee cycle is characterized by intense biomass production, both vetative and reproductive structures (CANNELL et al., 1976). This generates an increase of biomass in the plant, increasing the value of the vegetation index and, consequently inverting the relation of the index with the yield of that year. The increase in the vegetation index value in C4 (fruit filling) and C5 (pre-harvest regarding Y2) can be seen in Figure 15, which shows that this behavior both for the regions that will have higher yields and those that will have lower yields.

Another interesting behavior observed in this study is the relationship of the AOS data of a specific collection (e.g., C1 and C5) with the yield of different seasons. The AOS data collected in C5, during the pre-harvest of 2020, showed a negative correlation ($r = -0.81$) with the yield of that same year (Y2), whereas it showed a positive correlation ($r = 0.72$) with the yield of the following season (Y3). The same behavior can be observed with C1 data, when compared with Y1 (Table 2) and Y2 (Table 3). This oscillation is an effect of the biennality of coffee production, in which a plant that showed high yield in one year, will show reduced yield in the next season (PEREIRA, et al., 2011; VALADARES et al., 2013).

Table 6. Temporal relationship of the AOS sensors data with the yield obtained in the 2020 (Y2) and 2021 (Y3) crop. The results show the performance of yield prediction (obtained in the full cross-validation) using simple linear regression models.

	In comparinson with Y2					**In comparinson with Y3**		
	Pre-harvest Y1 C1	**Post-harvest Y1 C2**	**Flowering C3**	**Fruit filling C4**	**Pre-harvest Y2 C5**	**Pre-harvest Y2 C5**	**Post-harvest Y2 C6**	**Pre-harvest Y3 C7**
	Crop Circle (CC-NDRE)							
r	0.91	0.82	0.73	-0.77	-0.81	0.72	0.88	*
R²	0.84	0.68	0.54	0.59	0.65	0.52	0.77	*
RMSE	0.20	0.28	0.34	0.32	0.29	0.30	0.21	*
RMSE%	9.96	13.96	16.78	15.74	14.50	20.38	13.99	*
RPD	2.47	1.76	1.47	1.56	1.70	1.45	2.11	*
	N-Sensor							
r	0.88	0.87	0.89	-0.77	-0.93	0.90	0.85	-0.71
R²	0.77	0.75	0.79	0.59	0.86	0.81	0.72	0.50
RMSE	0.24	0.25	0.23	0.32	0.19	0.19	0.23	0.31
RMSE%	11.68	12.20	11.38	15.78	9.19	12.93	15.45	20.84
RPD	2.11	2.02	2.16	1.56	2.68	2.28	1.91	1.41

* Data not acquired due to equipment problems during the field expedition.

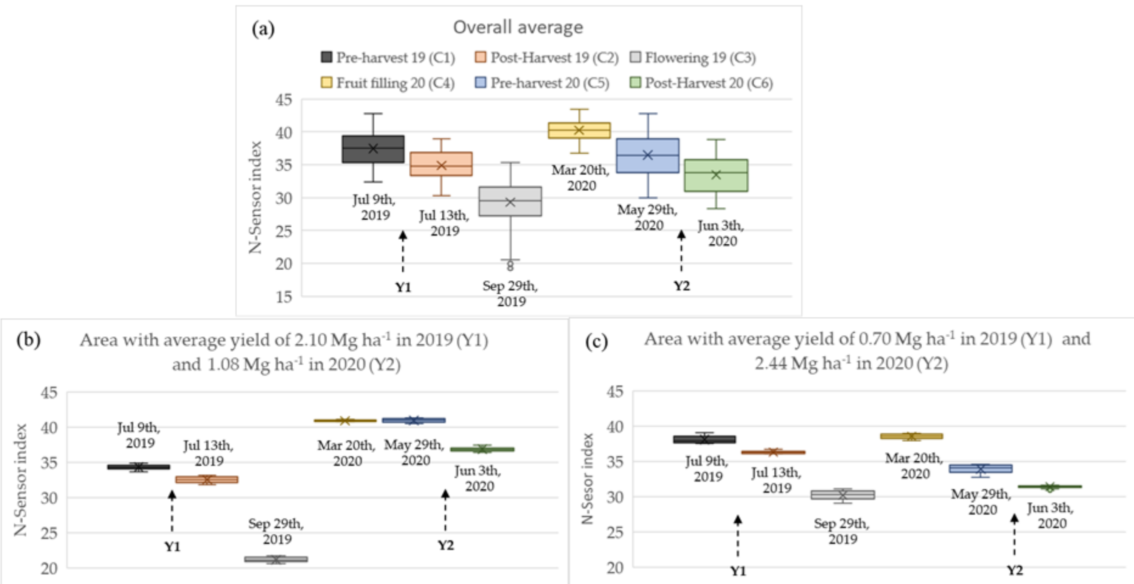

Figure 15. Temporal behavior of the N-Sensor vegetation index (from C1 to C6) showing the boxplots of all field data (a), the boxplots of an area with high yield in 2019 and low yield in 2020 (b), and the boxplots of an area with low yield in 2019 and high yield in 2020 (c). For the graphs generated in (b) and (c), 16 polygons (the ones created over the yield data point specified in Figure 14d) with similar yields were evaluated.

In summary, our results showed that during C2 (post-harvest) and C3 (flowering) the vegetation indices evaluated correlate positively with the yield of that year, so that sites with low and high biomass will respectively show low and high yields at the harvest of that respective year. On the other hand, during C4 (fruit filling) and C5 (pre-harvest), the plant with higher vigor will have lower yields than the plant with lower vigor. This is because the relationship between vegetation index (i.e., plant vigor) and coffee yield is inverse in phenological stages after flowering. It was also observed that the relationship between the yields of different seasons will invert the relationship with the vegetation indices observed at a specific time, i.e., if the index collected in the phenological stage "X" of 2020 shows a positive correlation with the 2020 yield, its correlation with the 2021 yield will be negative.

The AOS performance for coffee yield prediction ranged from reasonable ($1.40 \leq RPD < 2.00$) to good ($2.00 \leq RPD < 3.00$) for both sensors over the entire coffee cycle. Despite the inverse correlation between vegetation indices and coffee yields in C4 and C5, the predictive potential of these stages does not change compared to the other stages, keeping RPD values with reasonable and good performances ($1.56 \leq RPD \leq 2.68$). In general, the performances obtained with the N-Sensor were superior to those obtained with the Crop Circle, with RPD varying between 1.41 and 2.68 for the N-Sensor and between 1.45 and 2.47 for the Crop Circle. This difference between both sensors can be explained due to differences regarding internal calibrations and filtering, besides the difference in the spectral region of the indices.

Regarding the AOS maps showed in Figure 16, it is possible to identify the regions that present low (red regions) and high vegetative vigor (regions in green). It is observed that throughout the coffee cycle (from C1 to C6) these regions are inverted, i.e., places in the field that showed low vigor, become more vigorous and vice-versa. This inversion is visually evident when we observe the maps obtained in C3 and C4 (Figure 16c, d, h, i). It is important to emphasize that the temporal behavior presented in Figure 15, is still occurring, i.e., we observe in the maps that the index values in C4 (e.g., CC-NDRE values oscillating between 0.39 and 0.44) are clearly higher than in C3 (e.g., CC-NDRE values oscillating between 0.26 and 0.37). What occurs spatially, for the reversal of vigor in different regions, is that zones of higher vigor in C3 (e.g., that had CC-NDRE values around 0.36) have a subtle increase in vegetation index value in C4 (e.g., showing CC-NDRE values around 0.38), while regions of lower vigor in C3 (e.g., that had CC-NDRE values around 0.31) have a greater increase in vegetation index value in C4 (e.g., showing CC-NDRE values around 0.41).

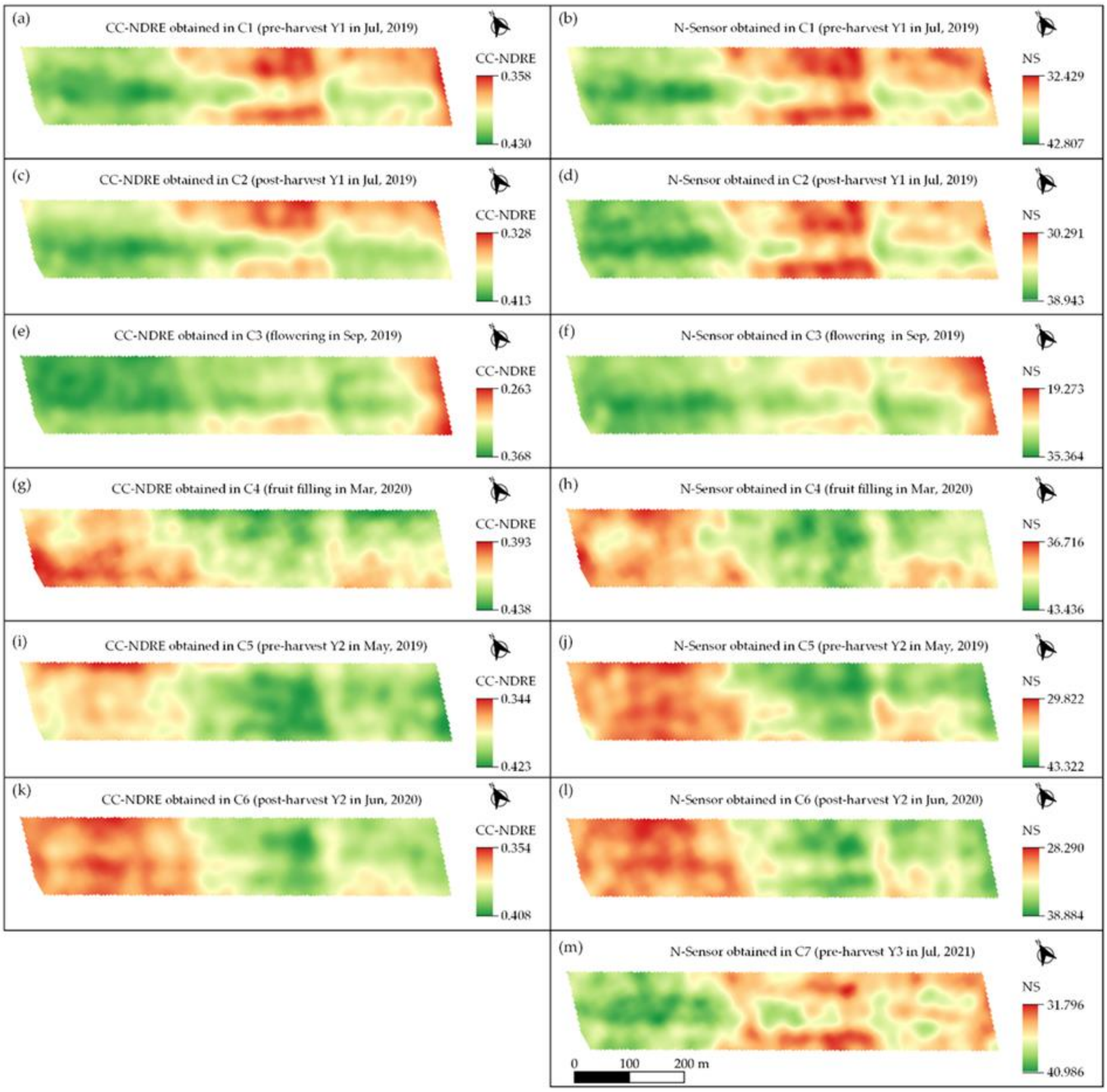

Figure 16. AOS data interpolated. (a) CC-NDRE obtained in C1; (b) NS obtained in C1; (c) CC-NDRE obtained in C2; (d) NS obtained in the C2; (e) CC-NDRE obtained in C3; (d) CC-NDRE obtained in C4; (e) CC-NDRE obtained in C3; (f) N-Sensor data obtained in C3; (g) CC-NDRE obtained in C4; (h) N-Sensor data obtained in C4; (i) CC-NDRE obtained in C5; (j) N-Sensor data obtained in C5; (k) CC-NDRE obtained in C6; (l) N-Sensor data obtained in C6; (m) N-Sensor data obtained in C7.

The spatio-temporal evaluation of coffee yield also showed that, in a single field, there can exist coffee plants with inverted bienniality (Figure 17 and 18). In other words, there are plants that have high yield potential in 2019 that, consequently, will have low potential in 2020, but in this same field there are also plants with low yield potential in 2019 that will have high potential in 2020.

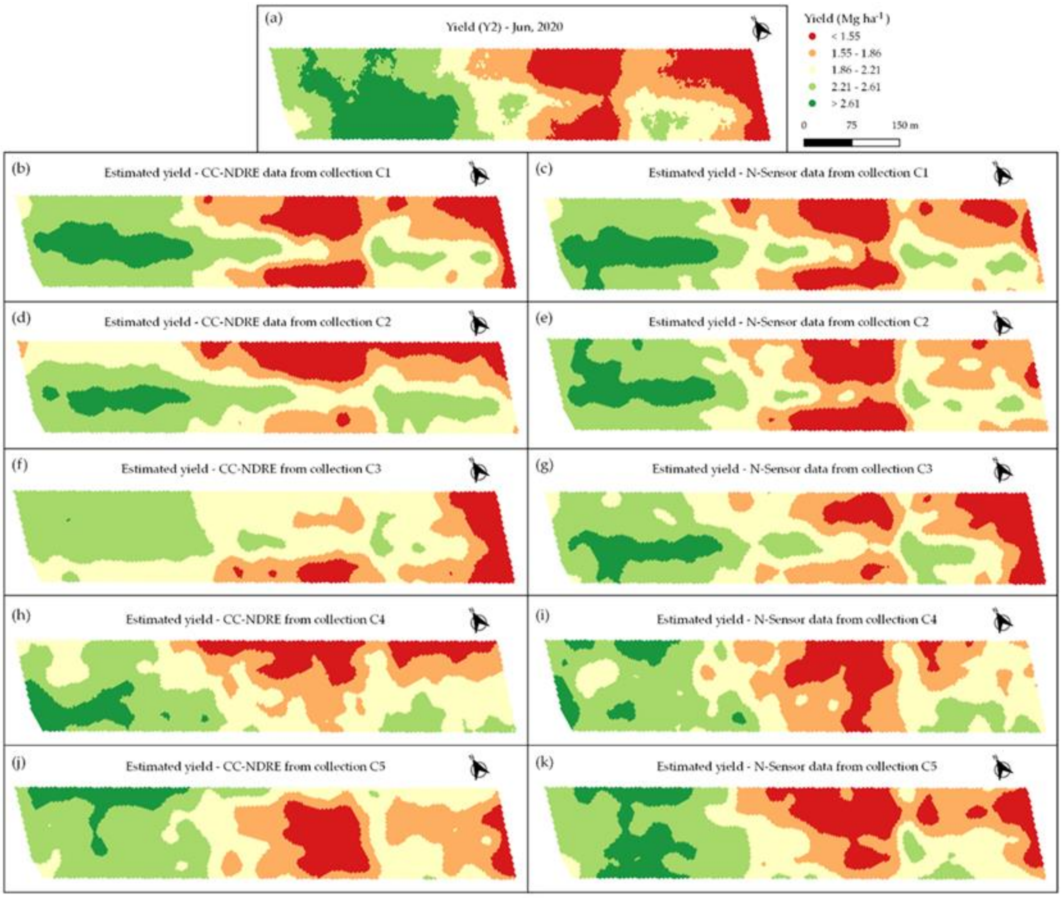

Figure 17. (a) Yield data obtained with yield monitor in Y2; (b) estimated yield with CC-NDRE in C1; (c) estimated yield with N-Sensor in C1; (d) estimated yield with CC-NDRE in C2; (e) estimated yield with N-Sensor in C2; (f) estimated yield with CC-NDRE in C3; (g) estimated yield with N-Sensor in C3; (h) estimated yield with CC-NDRE in C4; (i) estimated yield with N-Sensor in C4; (j) estimated yield with CC-NDRE in C5; (k) estimated yield with N-Sensor in C5;

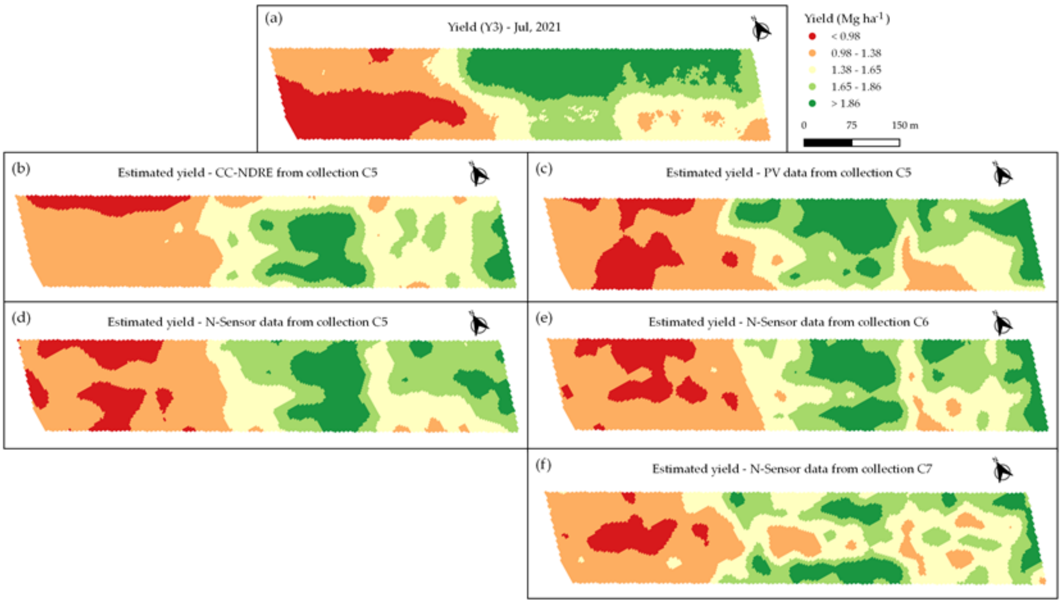

Figure 18. Figura 18. (a) Yield data obtained with yield monitor in Y3; (b) estimated yield with CC-NDRE in C5; (c) estimated yield with N-Sensor in C5; (d) estimated yield with CC-NDRE in C6; (e) estimated yield with N-Sensor in C6; (f) estimated yield with N-Sensor in C7;

This behavior can be observed both in the maps of Figurese 17-18 and in Figure 15b, c. In the field expeditions, this behavior was also visualized along different coffee lines (Figure 19), where plants with different behavior compared to their neighbors were found. This finding shows that one of the causes of variation over short distances in coffee is the existence of different biennial patterns, which can cause variability from plant to plant. In this work, this type of variability was ignored since spatial filters and interpolation were performed to make the sensor information compatible with that of yield data, which had higher spatial resolution. However, this plant-to-plant spatial variability within the crop row can be identified and managed by using proximal sensor systems and should be explored in future studies.

Figure 19. Photo indicating the presence of variability at short distances within a single row of coffee.

The high resolution of the sensor data allowed the identification of spatial variations in the productive potential of coffee plants, as well as the identification that in the same plot there are plants with inverted biennial behavior. It is evident that AOS sensors have a great potential to be explored in coffee farming, and that the present study sought to provide an initial approach regarding the potential of using these sensors for PA management. Future studies may address the use in other scenarios (e.g., crops with different row orientation, mountain coffee, shaded coffee), as well as exploiting this information to create management strategies that optimize the management of inputs, as has been done for annual crops (ANASTASIOU et al., 2018; TREVISAN et al., 2018; PALLOTTINO et al., 2019; YU et al., 2021).

3.4. CONCLUSIONS

A procedure for acquiring and processing active optical sensors (AOS) data that are positioned on the side of coffee plants has been proposed and used for the spatio-temporal evaluation of coffee yield over its biennial cycle. The results of this study suggest that sensors should be positioned to collect data from the middle third of the coffee plants (scenario TM), as well as it should be averaged the AOS data obtained from both sides of the coffee plant (scenario A3) for a better correlation of vegetation indices with coffee yield.

Regarding the temporal relationship of AOS data with yield, it was observed that the vegetation indices acquired in the post-harvest (C2) and flowering stages (C3) have positive correlations with the yield of that year, while vegetation indices acquired during fruit filling (C4) and pre-harvest stages (C5) have negative correlations with coffee crop yields.

Assessing the spatial variability of the coffee field with the tools used in this study (AOS equipment and yield monitors), it was observed that one of the causes of coffee variability present in the field is the existence of plants with different productive potential behaviors due to their uneven bienniality. This variation can be observed and, therefore, managed by using proximal sensors. This study presents the spatio-temporal variations of OAS data in large-scale coffee plantations, as well as their relationship with the coffee crop yield. More studies are needed to go deeper into practical applications of these data in order to support optimized input management through precision agriculture practices in coffee crops.

REFERENCES

ALVARES, A. C.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M.; SPAROVEK, G. Koppen's climate classification map for Brazil. **Meteorologische zeitschrift,** v. 22, n.6, p.711–728, 2013.

AMARAL, L. R.; MOLIN, J. P.; SCHEPERS, J. S. Algorithm for Variable-Rate Nitrogen Application in Sugarcane Based On Active Crop Canopy Sensor. **Agronomy Journal**, v. 107, p. 1-1, 2015.

AMARAL, L. R.; TREVISAN, R. G.; MOLIN, J. P. Canopy sensor placement for variable-rate nitrogen application in sugarcane fields. **Precision Agriculture**, v. 19, p. 1-14, 2017.

ANASTASIOU, E.; BALAFOUTIS, A.; DARRA, N.; PSIROUKIS, V.; BINIARI, A.; XANTHOPOULOS, G.; FOUNTAS, S. Satellite and proximal sensing to estimate the yield and quality of table grapes. **Agriculture**, v.8, n.7, p.94, 2018.

AULA, L.; OMARA, P.; NAMBI, E.; OYEBIYI, F.B.; RAUN, W.R., Review of active optical sensors for improving winter wheat nitrogen use efficiency**. Agronomy**, 10(8), p.1157, 2020.

BARNES, E. M.; CLARKE, T. R.; RICHARDS, S. E.; COLAIZZI, P.D.; HABERLAND, J.; KOSTRZEWSKI, M.; WALLER, P.; CHOI, C.; RILEY, E.; THOMPSON, T.; et al. Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. In Proceedings of the 5th International Conference on Precision Agriculture, Bloomington, MN, USA, 16–19 July 2000.

BERNTSEN, J. et al. Algorithms for sensor-based redistribution of nitrogen fertilizer in winter wheat. **Precision Agriculture**, v. 7, n. 2, p. 65-83, 2006.

CANNELL, M. G. R. Crop physiological aspects of coffee bean yield—a review. **Kenya Coffee**, v. 41, p.245–253, 1976.

CASTILLO, Z. J.; LOPEZ, A. R. Nota sobre el efecto de la intensidad de la luz en la floracio´n del café. **Cenicafe**, v.17, p.51–60, 1966.

CHANG, C. W.; LAIRD, D. A.; MAUSBACH, M. J.; HURBURGH, C. R. Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. **Soil Science Society of America Journal**, 65(2), 480-490, 2001.

CHAVES, A. R.; MARTINS, S. C.; BATISTA, K. D.; CELIN, E. F.; DAMATTA, F. M. Varying leaf-to-fruit ratios affect branch growth and dieback, with little to no effect on photosynthesis, carbohydrate or mineral pools, in different canopy positions of field-grown coffee trees. **Environmental and Experimental Botany**, Viçosa, v. 77, p.207-218, abr. 2012

DAMATTA, F. M. Ecophysiological constraints on the production of shaded and unshaded coffee: a review. **Field Crop Res.**, 86, p. 99-114, 2004.

DARRA, N.; PSOMIADIS, E.; KASIMATI, A.; ANASTASIOU, A.; ANASTASIOU, E.; FOUNTAS, S. Remote and proximal sensing-derived spectral indices and biophysical variables for spatial variation determination in vineyards. **Agronomy**, v.11, n.4, p.741, 2021.

GEBBERS, R.; ADAMCHUK, V. I. Precision agriculture and food security. **Science**, v. 327, n. 5967, p. 828–31, 2010.

GIRMA, KEFYALEW et al. Mid-season prediction of wheat-grain yield potential using plant, soil, and sensor measurements. **Journal of Plant Nutrition**, v. 29, n. 5, p. 873-897, 2006.

GROHS, DANIEL S. et al. Model for yield potential estimation in wheat and barley using the GreenSeeker sensor. **Engenharia Agrícola**, v. 29, p. 101-112, 2009.

JUNGES, A.H.; FONTANA, D.C.; ANZANELLO, R.; BREMM, C. Normalized difference vegetation index obtained by ground-based remote sensing to characterize vine cycle in Rio Grande do Sul, Brazil. **Ciênc. E Agrotecnol.**, v.41, p.543–553, 2017.

KITCHEN, N. R.; SUDDUTH, K. A.; DRUMMOND S. T.; SCHARF, P. C.; PALM, H. L.; ROBERTS, D. F.; VORIES, E. D. Ground-based canopy reflectance sensing for variable-rate nitrogen corn fertilization. **Agronomy Journal**, v.102, n.1, p. 71-84, 2010.

LU, J.; MIAO, Y.; SHI, W.; LI, J.; YUAN, F. Evaluating different approaches to non-destructive nitrogen status diagnosis of rice using portable RapidSCAN active canopy sensor. **Sci. Rep.,** v.7, p.14073, 2017.

MALDANER, L. F.; CANATA, T. F.; DIAS, C. T. S.; MOLIN, J. P. A statistical approach to static and dynamic tests for Global Navigation Satellite Systems receivers used in agricultural operations. **Scientia Agricicola**, v.78, n.5, 2021.

MALDANER, L. F.; CANATA, T. F.; MOLIN, J. P. An approach to sugarcane yield estimation using sensors in the harvester and zigbee technology. **Sugar Tech**, v.24, p. 813–821, 2022.

MALDANER, L. F.; MOLIN, J. P. Data processing within rows for sugarcane yield mapping. **Scientia Agricola**, v.77, n.5, 2020.

MARTELLO, M.; MOLIN, J. P.; BAZAME, H. C. Obtaining and Validating High-Density Coffee Yield Data. **Horticulturae**, v. 8, n. 5, p. 421, 2022.

MONTOYA, L. A., SYLVAIN, P. G., UMANÃ, R. Effect of light intensity and nitrogen fertilization upon growth differentiation balance in Coffea arabica L. **Coffee (Turrialba)**, v. 3, p. 97–104, 1961.

MORLIN C., F. et al. Comparison between vegetation indices for detecting spatial and temporal variabilities in soybean crop using canopy sensors. **Precision Agriculture,** v. 21, n. 5, p. 979-1007, 2020.

OLIVEIRA, M.F., SANTOS, A.F., KAZAMA, E.H., ROLIM, G. S., SILVA, R.P. Determination of application volume for coffee plantations using artificial neural networks and remote sensing. **Comput. Electron. Agric**. 184, 106096, 2021.

PALLOTTINO, F.; ANTONUCCI, F.; COSTA, C.; BISAGLIA, C.; FIGORILLI, S.; MENESATTI, P. Optoelectronic Proximal Sensing Vehicle-MountedTechnologies in Precision Agriculture: A Review. **Computers and Electronics in Agriculture**, v. 162, p.859–873, 2019.

PEREIRA, S. O.; BARTHOLO, G. F.; BALIZA, D. P., SOGREIRA, F. M.; GUIMARÃES, R. J. Productivity and coffee biannuality depending on the crop spacing. **Pesquisa Agropecuária Brasileira**, V.46, n.2, p.152-160, 2011.

PORTZ, G.; MOLIN, J. P.; JASPER, J. Active crop sensor to detect variability of nitrogen supply and biomass on sugarcane fields. **Precision Agriculture (Print)**, v. 13, p. 33-44, 2012.

PUTRA, B. T. W.; SONI, P.; MORIMOTO, E.; PUJIYANTO, P. Estimating biophysical properties of coffee (Coffea canephora) plants withabove-canopy field measurements, using CropSpec. **Interntional Agrophysics**, v.32, p.183–191, 2018.

RAUN, W. R. et al. In-season prediction of potential grain yield in winter wheat using canopy reflectance. **Agronomy Journal**, v. 93, n. 1, p. 131-138, 2001.

ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the great plains with ERTS. In Proceedings of the Earth Resources Technology Satellite—1 **Symposium**, Washington, DC, USA, 10–14 December, pp. 309–317, 1974.

SANTINI, P. T.; ALMEIDA, L. G.; SOUZA, K. R. D. de; BARBOSA, J. P. R. A. D.; ALVES, J. D. Spatio-temporal variability of carbohydrate and chlorophyll content in the coffee canopy. **Coffee Science**, v. 14, n. 3, p. 366–372, 2019.

SCHMIDT, J.; BEEGLE, D.; ZHU, Q.; SRIPADA, R. (2011). Improving in-season nitrogen recommendations for maize using an active sensor. **Field Crops Research**, v.120, n.1, p.94-101, 2011.

SOLARI, F.; SHANAHAN, J. F.; FERGUSON, R. B.; ADAMCHUK, V. I. An active sensor algorithm for corn nitrogen recommendations based on a chlorophyll meter algorithm. **Agronomy Journal**, v.102, n.4, p.1090-1098, 2010.

TASKOS, D. G.; KOUNDOURAS, S.; STAMATIADIS, S.; ZIOZIOU, E.; NIKOLAOU, N.; KARAKIOULAKIS, K.; THEODOROU, N. Using active canopy sensors and chlorophyll meters to estimate grapevine nitrogen status and productivity. **Precision Agriculture**. 2014, 16, 77–98, 2015

TEAL, R. K.; TUBANA, B.; GIRMA, K.; FREEMAN, K. W.; ARNALL, B. D.; WALSH, O.; RAUN, W. R. In-Season prediction of corn grain yield potential using Normalized Difference Vegetation Index. **Agronomy Journal**, v.98, p.1488-1494, 2006.

TREVISAN, R. G.; VILANOVA JÚNIOR, N. S.; EITELWEIN, M. T.; MOLIN, J. P. Management of plant growth regulators in cotton using active crop canopy sensors. **Agriculture**, v.8, n.7, p.101, 2018.

VALADARES S.V.; NEVES J. C. L.; ROSA G. N. G. P.; MARTINEZ H. E. P; VENEGAS, V. H. A; LIMA P. C. Productivity and biennial production of dense coffee plantations under different doses of N and K. **Pesquisa Agropecuária Brasileira**, v. 13, n.48(3), p.96-303, 2013.

XUE, L.; LI, G.; QIN, X.; YANG, L.; ZHANG, H. Topdressing nitrogen recommendation for early rice with an active sensor in south China. **Precision Agriculture**, 15(1), 95-110, 2014.

YU, R., BRILLANTE, L., TORRES, N., & KURTURAL, S. K. Proximal sensing of vineyard soil and canopy vegetation for determining vineyard spatial variability in plant physiology and berry chemistry. **OENO One**, v.55, n.2, p.315–333, 2021.

4. ASSESSING TEMPORAL AND SPATIAL VARIABILITY OF COFFEE PLANTS USING RPA-BASED RGB IMAGING

ABSTRACT

The biophysical parameters of the coffee plants can provide important information to guide crop management. An alternative to traditional methods to obtain this type of information can be the 3D modeling of the coffee canopy using aerial images from RGB cameras attached to remotely piloted aircraft (RPA). Studies have shown the potential of this technique for grains like maize, rice, barley, and perennial crops such as grapes and olive trees, but in coffee crop most studies have not addressed the use of this approach for temporal analysis and its relationship with crop yield data. This study aimed to explore the use of RGB aerial images to obtain 3D information of coffee crop, deriving plant height and volume information together with yield data during three growing seasons. The study was developed in a commercial production area of 10.24 ha, located in the cerrado biome, state of Minas Gerais. Seven data acquisition campaigns were conducted during the years 2019, 2020 and 2021. The flights were made at 70 m above ground level with lateral and longitudinal overlap of 75% and 80%, respectively. The images were processed obtaining canopy surface models (CSMs) derived into plant height and volume data for each campaign. To validate the plant height results, 120 sample points were used for reference ground measurements. The results showed that it is possible to extract plant height of coffee plants with R^2 0.86 and RMSE 0.4 m. It was possible to monitor the temporal variability of coffee plant height and volume based on aerial images and correlate this information with yield data. The results of the modeling analysis demonstrated the possibility of using these variables to help understand the spatial variability of coffee yield within the field.

Keywords: Structure from motion; UAV; Crop surface model; Precision agriculture

4.1. INTRODUCTION

Remote sensing applications in agriculture have been trending over the last decades, providing rapid and efficient insights for crop scouting based on the study of spectral response of soil, water and vegetation (MULLA, 2013). Sensors can be attached on different platforms, being classified as proximal, aerial and satellite-based according to their distance in relation to the target object (FLORENZANO, 2002). Reflecting recent technology advances, the use of remotely piloted aircraft system (RPA) gained significant attention in agriculture, being established as an effective tool for small-scale remote sensing (GÓMEZ-CANDÓN, DE CASTRO E LÓPEZ-GRANADOS, 2014; SANKARAN et al., 2015, COLOMINA E MOLINA, 2014).

The use of RPA has been considered promising tool with easy data acquisition and higher operational flexibility in relation to classical platforms, providing reliable data (COLOMINA E MOLINA, 2014; SCHIRRMANN et al., 2016). The application of remote sensing analysis coupled with aerial platforms data provide an efficient, rapid and non-destructive assessment through the spectral response of plants and its relationship with various environmental factors (TILLY et al., 2014; POSSOCH et al., 2016), mainly because the spectral patterns of a crop vary due to several factors, such as development stage, vegetative vigor, and management practices (BERNI et al., 2009; BALLESTEROS et al., 2018).

Several studies have explored the use of RPA with different on-board sensors for crop monitoring. The applications vary from identifying the appropriate time to harvest grapes (JOHNSON; HERWITZ, 2003), estimating the canopy height of olive trees (DÍAZ-VARELA et al, 2015), monitoring the growth of rice and barley crops

(BARETH et al., 2015; BENDIG et al., 2013), spatial variability of water stress assessments in vineyards (SANTESTEBAN et al., 2017), and estimation of biomass production of maize (LI et al., 2016).

Some of these studies explore the use of consumer-grade red-green-blue (RGB) cameras for aerial imaging, in which centimeter spatial resolution images are obtained and processed to generate digital elevation models (VERHOEVEN, 2011; BENDIG, et al., 2014). Based on structure from motion (SfM) algorithms, RGB aerial images are processed exploring photogrammetric methods specifically developed for RPA imagery, resulting in 3D point clouds that can represent crop canopy height (BENDIG, et al., 2015). According to Hoffmeister et al. (2010), crop height information is generated by subtracting a digital terrain model (DTM) from a digital surface model (DSM). As crop height information has been successfully used as an estimate of plant biomass in many crops, several studies have explored the use of CSM data derived from RGB imagery as a non-destructive method to assess crop biomass, with successful results for onion, barley, wheat and maize (SCHIRRMANN et al., 2016; BENDIG, et al., 2014; BENDIG, et al., 2015; LI et al., 2016; BALLESTEROS et al., 2018).

In coffee production, the use of aerial imagery to explore spatial and temporal variability is highlighted in studies developed by Cunha et al. (2019), Santos et al. (2020) and Bento et al., (2022), in which biophysical variables of coffee plants were estimated using RGB imagery data. Cunha et al. (2019) sought to estimate the biophysical variables of coffee plants in order to improve pesticide application efficiency. Cunha et al. (2019) and Santos et al. (2020) highlighted that the height and volume of the canopy of coffee plants could be determined in a practical and precise way through digital processing of images captured by RPA, and that the determination of these variables is fundamental, since they can be related to coffee plant responses, such as biomass reduction, water stress, occurrence of pests and diseases, nutritional deficiencies, changes in chlorophyll content, and ultimately crop yield. Bento et al. (2022) characterized recently planted coffee cultivars using RPA-based vegetation indices. They also calculated height and diameter of the canopy of recently planted coffee plants in three developmental stages based on RPA imagery, achieving good correlations between plant height and volume from CSMs and field data.

Although some studies have explored the use of RGB imagery to estimate plant height and volume for coffee crop, they were limited to point-wise sampling comparison within a single growing season. Therefore, the objective of this study was to analyze the potential of RGB imagery data acquired over three growing seasons to monitor the spatial and temporal variability of biophysical parameters of coffee crops and their relation with high-density yield data.

4.2. MATERIALS AND METHODS

4.2.1. Study area

The area is located in the municipality of Patos de Minas, state of Minas Gerais, Brazil (Figura 20a). The local climate is classified as Aw, according to the Köppen classification (ALVARES et al., 2013), characterized by a tropical environment with dry winters and rainy summers. The study was conducted in a commercial coffee area of 10.24 ha, cultivated with the species *Coffea arabica L.*, IAC Catuaí 144 variety. The coffee trees were spaced by 0.5 m between plants and 4.0 m between rows, resulting in a density of approximately 5,000 plants ha^{-1}, with a drip irrigation system. In total, 35 rows of coffee with approximately 760 m in length each were evaluated.

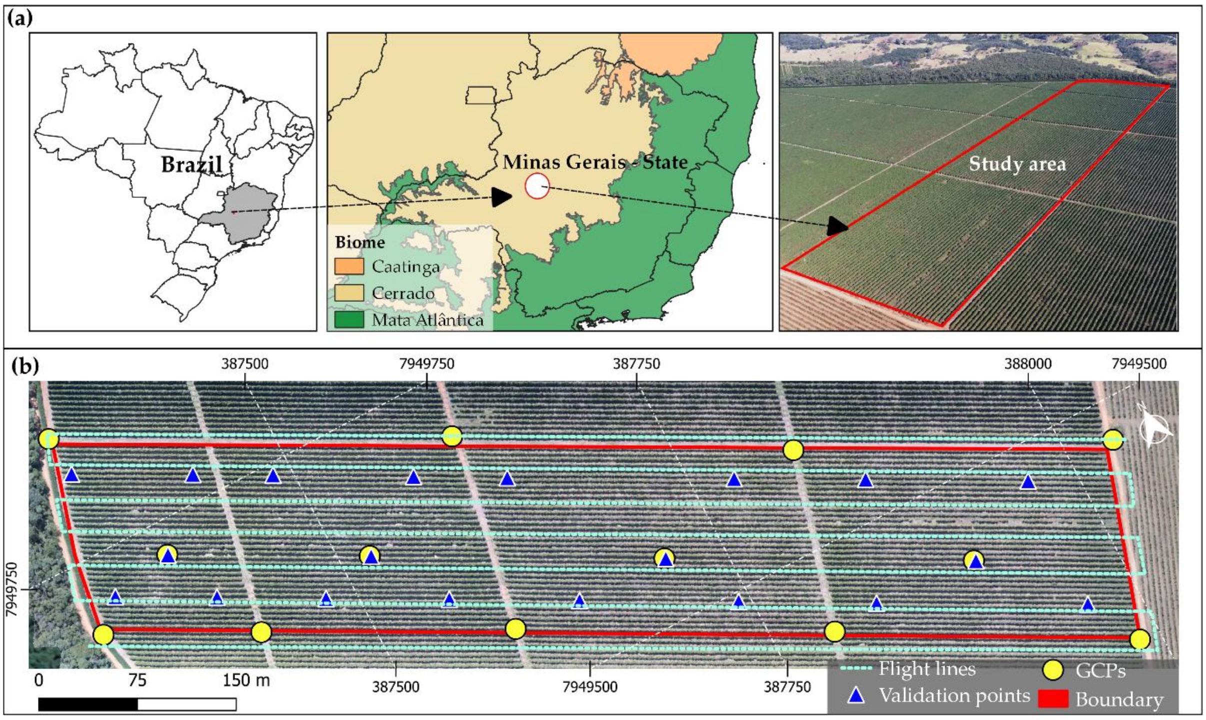

Figure 20. (a) general map of the study area; (b) map with details regarding the flight plan, ground control points and height validation points.

4.2.2. RPA Flight and Image Acquisition

Seven campaigns were carried out between 2019 and 2021 (Table 7) to capture aerial images, providing datasets at different growth stages throughout multiple growing seasons. All flights were conducted close to solar noon with sunny and calm wind weather conditions.

A DJI Phantom 4 Advanced quadcopter (DJI, Shenzhen, Guangdong, China) was used during the campaigns to capture RGB images. The built-in camera has a spectral resolution within the visible spectrum (350~700 nm - RGB), featuring a 12.8 x 9.6 mm CMOS sensor (Complementary Metal Oxide Semiconductor) controlled by a mechanical shutter, which produces images with up to 20 megapixels (4864 × 3648 pixels). The flight campaigns were planned using DroneDeploy mobile software (DroneDeploy Inc., San Francisco, CA, USA). All the flights were conducted at 70 m above ground level, with side and forward overlap of 75 and 80%, respectively. The aforementioned settings provided a ground sample distance (GSD) of around 0.02 m with over 400 images captured (Table 1) To ensure positional accuracy during the georeferencing procedures, 13 ground control points were previously located in the field with their coordinates measured by a pair of RTK GNSS receiver (Topcon GR-3, 0.01 m accuracy, Topcon Corporation, Tokyo, Japan).

Table 7. Details of each collection referring to date, phenological stages and amount of images obtained.

Collection	Stage	Date	Local Time	Images
C1	Pre-harvest	Jul 9th,2019	11:57	407
C2	Post-harvest	Jul 13th,2019	11:10	406
C3	Flowering	Sep 29th,2019	11:02	401
C4	Fruit filling	Mar 20th,2020	13:47	401
C5	Pre-harvest	May 29th,2020	13:15	402
C6	Post-harvest	Jun 3th,2020	12:15	404
C7	Pre-harvest	Jun 24th,2021	12:49	405

4.2.3. Canopy Surface Model Generation

To produce the digital elevation models (DEM), the RGB images obtained were processed using Agisoft PhotoScan software (Agisoft LLC, São Peterburgo, Russia), which features structure from motion algorithm (SfM) for 3D digital reconstruction of scenes (MALAMBO et al., 2018). The workflow consists of importing the images into the software to perform image alignment, followed by the GCPs insertion to optimize the previous obtained image alignment. Furthermore, the dense cloud can be generated and derived into a mesh surface that represents the above surface structures (DSM). In order to obtain the digital terrain model (DTM), the dense cloud previously generated needs to be filtrated, maintaining only ground level points (BENDIG et al., 2013). To define appropriate parameters for the filtering process, an empirical approach was implemented, testing different combinations of maximum angle, maximum distance and cell size, observing the correspondence to the features overlayed on the orthomosaic. Thus, the following parameters were defined: maximum angle of 1.5°; maximum distance of 0.25 m; and cell size of 3 m. The selected points were then interpolated and derived into a smoothed mesh surface obtained with a 5-points threshold that represents the DTM. Both DSM and DTM were exported as TIFF file with a 0.03 m pixel^{-1} fixed spatial resolution. Figure 21 presents a flowchart of image processing.

Using map algebra (Equation (5)) between the DSM and the DTM, the so-called crop surface model (CSM) was obtained, which represents the plant height information (BENDIG et al., 2013).

$$CSM=DSM-DTM \tag{5}$$

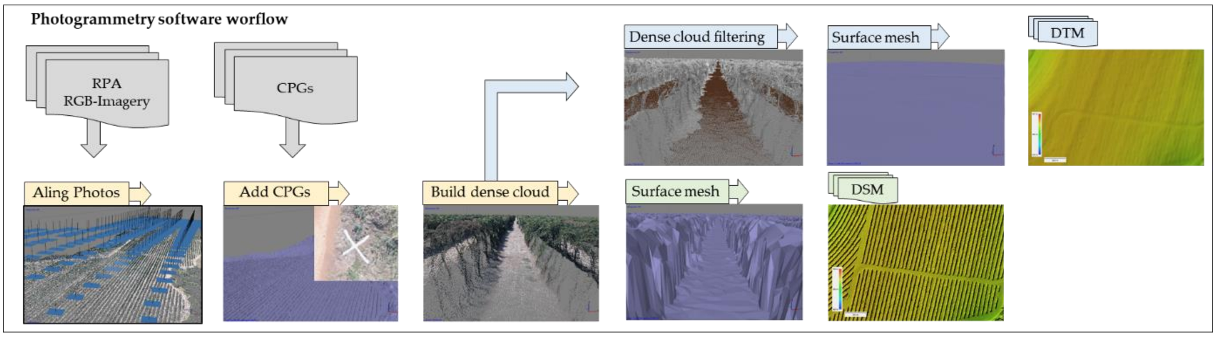

Figure 21. Flowchart of aerial image processing to obtain the DTM and DSM.

4.2.4. Extraction of plant height and plant volume information

Before extracting plant height information from CSMs we first developed a process aiming to individualize each coffee tree using QGIS (QGIS Development Team, 2022) (Figure 22). The first step consisted of clipping portions of the CSM with values ≥ 0.2 m, generating a mask feature layer separating each coffee row. Furthermore, a transversal clip at every 0.5 m (assuming regular spacing between plants) was applied to the previous obtained mask layer in order to separate each coffee plant within the row. This way, each plant was individualized based on a polygon feature maintaining the variable width metrics of every tree. Based on the obtained mask feature, zonal statistics were obtaining, calculating the maximum plant height, average plant height and area within each polygon representing individual trees. By multiplying the average plant height by the area, the plant volume of each tree was then obtained.

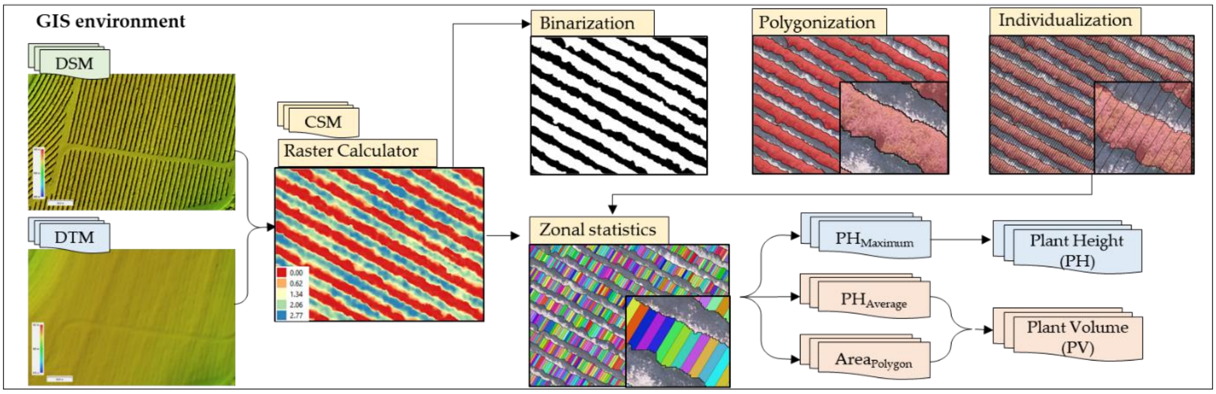

Figure 22. Flowchart of plant individualization processing and extraction of PH and PV metrics.

4.2.5. Validation of plant height derived from CSMs

The validation of plant height information, 20 sampling points were distributed across the field during C1 (Figure 20b). At each sampling location, six plants were selected and had their maximum height measured using a topographic ruler, with the value being compared to the CSM-based plant height. Based on the results of the 120 samples, a linear regression was obtained between ground based plant height and CSM-based plant height, with the accuracy evaluated based on root mean squared error (RMSE).

4.2.6. Spatio-temporal relationship of Image data with coffee crop yield

Yield data was obtained based on the yield monitor system installed on the K3 Millenium coffee harvester (Jacto, Pompeia, Brazil). The yield datasets were first filtered and then interpolated using ordinary kriging with a spatial resolution of 3 m according to the methodology described by Martello et al. (2022). The same methodology was applied to the PH and PV data derived from CSMs. Datasets obtained from C1 to C7 were filtered using the MapFilter software (MALDANER et al., 2022), employing a global filtering with a threshold of 100% and a local filtering with a threshold of 10% and a search radius of 10 m. The filtered datasets were then interpolated using ordinary kriging with a 3 m spatial resolution using the Vesper 1.6 software (MINASNY et al., 2005).

The spatio-temporal relationship of PH an PV data with coffee crop yield was evaluated comparing the PH an PV maps of C1 with yield map of Y1, the PH and PV maps of collections C1, C2, C3, C4 and C5 were also compared with the yield map of Y2, as well as comparing the yield map of Y3 with the PH an PV maps of C5, C6, and C7. These data were compared using Pearson's correlation coefficient and the ability of the PH an PV data to predict yield was verified using simple linear regressions (n = 11,859, for all datasets). These comparisons permit to evaluate the temporal relationship of the sensor data with the coffee yield.

The calibrations of the predictive models were validated with full cross validation using as quality indicators the coefficient of determination (R^2), the residual prediction deviation (RPD), and the root-mean-square error (RMSE). The RPD was calculated as the ratio between the standard deviation of the measured yield and the RMSE obtained in the prediction. Four RPD classes adapted from Chang et al. (2001) were used to evaluate the quality of models: poor models (RPD < 1.40), reasonable models (1.40 ≤ RPD < 2.00), good models (2.00 ≤ RPD < 3.00), and excellent models (RPD ≥ 3.00).

4.3. RESULTS AND DISCUSSION

4.3.1. Plant Height Validation

The regression model between CSM-based plant height and reference measurements yielded an R^2 of 0.86, with RMSE = 0.41 m (Figure 23). Other studies that explored the use of SfM algorithms to derive plant height information from RPA-based imagery obtained similar results in terms of precision for maize (R^2 = 0.88) (LI et al., 2016), black oat (R^2 = 0.68 – 0.92) (ACORSI et al., 2019), wheat (R^2 = 0.87 – 0.96) (SCHIRRMANN et al., 2016), and grassland (R^2 = 0.56-0.70) (GRÜNER et al., 2019).

Some other studies that also employed RGB aerial images to assess plant height in coffee production achieved R^2 values between 0.52 and 0.87, and RMSE between 0.07 and 0.09 m (BARBOSA et al., 2021; BENTO et al., 2022; SANTOS et al., 2020). Although the higher residuals obtained in our study indicates a lower accuracy in relation to the aforementioned studies, the high precision achieved (R^2 = 0.86) demonstrates that the plant height measurements are mainly affected by a systematic error that can be easily corrected using an empirical line calibration model to achieve equivalent accuracy.

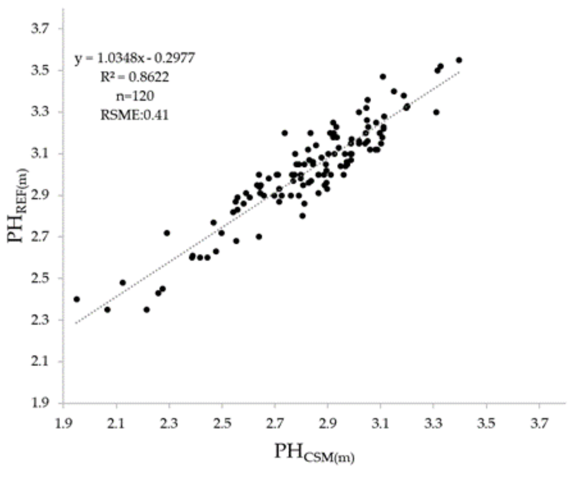

Figure 23. Linear regression between plant height deviated from CSM (PH_{CSM}) and manual height measurements (PH_{REF})

4.3.2. Exploratoy Analysis

The descriptive statistics for PH and PV from all datasets are presented in Table 8. The original PV data presented a CV from 16.89 to 22.23%, while the CV for the PH data presented a variation of 8.52 to 13.47%. With the filtered datasets the CV was smaller, presenting values from 8.5 to 15.79% for PV and between 6.17 and 10.53% for PH. The reduction in the variation of values is a result of the filtering process, which aims to exclude discrepant values in order to improve the understanding of the real variability of the data. It is also noticeable that the filtering process removed more readings from PV than PH. Since the estimation of PV uses two variables (average height and plant area) the variations found in the original values were greater and consequently the filtering step eliminated more discrepant points.

The average plant height increased over time, starting from an average of 2.67 m (C2 - year 2019) all the way to 2.99 m (C7 - year 2021). The average PV also showed variations during the evaluation period, but in this case, the volume of the plant suffered a negative impact after harvest. The average values of the volume before harvest, C1 and C5 (2.31 m³ and 2.79 m³), are greater them the average values of the datasets after harvest, C2 and C6 (2.20 m³ and 2.75 m³), in which a decrease in the average volume of 0.11 and 0.04 m³ was observed. This negative variation can be explained by the impact that the coffee trees suffer during the mechanized harvest of the fruits. As observed by Santinato et al. (2014), during the harvest the plant is exposed to constant vibrations to detach the fruits. These vibrations combined with the traffic of machines in the area can cause damages to branches and loss of leaves, which can justify the decrease in the average volume of the plants.

From the end of harvest until the first rains the plant remains in a vegetative resting stage and the change in plant volume is small (C2 to C3), but as soon as the rainy season begins (C3) the plant moves to an active vegetative stage (C4), and this behavior can be noticed when comparing the average volume of the C3 collection, that was 2.25 m³, and changes to 2.80 in the C4 collection, an average increase of 0.55 m³.

Table 8. Descriptive statistics of raw and filtered PH and PV data.

Variable	Unit	n	Mean	Min	Max	SD	CV (%)	n	Mean	Min	Max	SD	CV (%)
		Raw Data						Filtered Data					
PV-C1	m³	51196	2.30	0.04	4.64	0.48	21.10	27329	2.31	1.44	3.42	0.31	13.50
PV-C2	m³	51196	2.16	0.02	4.50	0.49	22.59	26205	2.20	1.29	3.23	0.32	14.59
PV-C3	m³	51196	2.21	0.02	4.91	0.49	22.23	29441	2.25	1.19	3.29	0.36	15.79
PV-C4	m³	51196	2.86	0.10	5.71	0.51	17.96	29099	2.80	1.75	3.86	0.30	10.86
PV-C5	m³	51196	2.84	0.18	5.58	0.49	17.42	28143	2.79	1.81	3.67	0.26	9.31
PV-C6	m³	51196	2.79	0.04	5.84	0.47	16.89	29002	2.75	1.91	3.63	0.23	8.50
PV-C7	m³	51196	2.83	0.02	5.77	0.58	20.58	25495	2.81	1.69	3.97	0.36	12.92
PH-C1	m	51196	2.66	0.44	3.66	0.32	12.13	42364	2.71	1.66	3.41	0.26	9.55
PH-C2	m	51196	2.60	0.36	4.06	0.35	13.47	40196	2.67	1.52	3.37	0.28	10.53
PH-C3	m	51196	2.66	0.32	3.70	0.33	12.52	43160	2.71	1.72	3.43	0.27	9.98
PH-C4	m	51196	2.93	0.65	3.86	0.25	8.52	46971	2.95	2.00	3.59	0.20	6.78
PH-C5	m	51196	2.82	0.48	3.75	0.24	8.54	45982	2.85	2.00	3.41	0.18	6.47
PH-C6	m	51196	2.89	0.61	3.97	0.25	8.59	45703	2.93	2.09	3.54	0.18	6.17
PH-C7	m	51196	2.94	0.33	3.90	0.33	11.29	42838	2.99	1.9	3.7	0.27	9.11

Figure 24 presents the correlation matrix between PH (Figure 24a) and PV (Figure 24b) over the seven campaigns. A positive correlation between the variables for all datasets can be observed, with r values ranging from 0.53 to 0.94 for PH and 0.46 to 0.91 for PV. The lowest correlation values were found when comparing the data from C2 with those from C5 and C6, with r values of PH 0.53 and PV 0.46 respectively. This decrease in correlation can be explained by the physiological aspects of coffee plants, which needs two years to produce the fruit, being the first year a vegetative stage and the second year the reproductive stage. Therefore, when comparing the data with an interval of two years (C2 - C7) the results show a higher correlation between datasets, with r values of 0.88 PH and 0.78 PV.

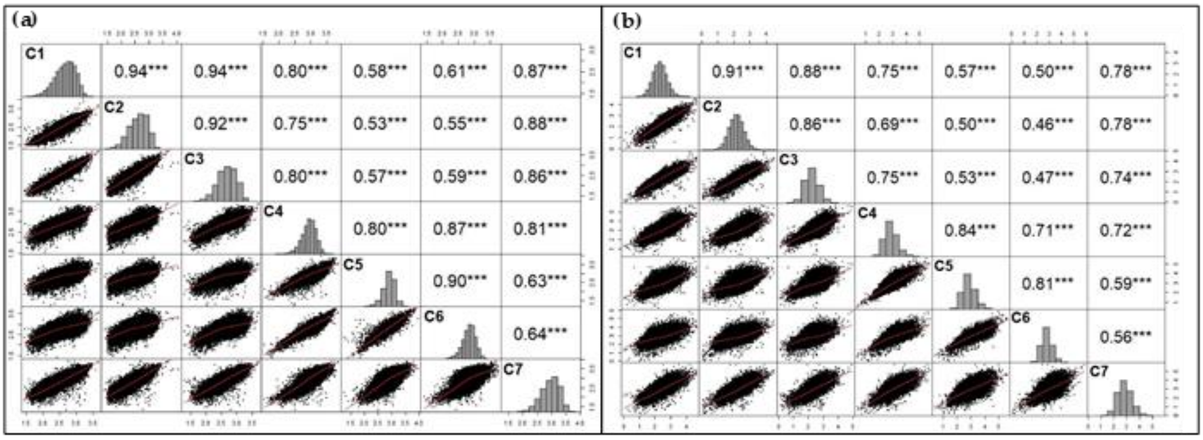

Figure 24. Correlation matrix, *** means p-value <0.001. (a) among the PH data; (b) among the PH data.

4.3.3. Spatio-temporal relationship of image data with coffee crop yield

The interpolated maps of PH and PV from all campaigns (C1 to C7) are presented in Figure 25, in which the spatial behavior from all campaigns can be observed. The PV maps from C1, C2 and C3 present a similar spatial behavior, with the same regions of high (regions in green) and low plant volume (regions in red). After collection C3 (Flowering) due to the beginning of the rainy season in the Cerrado region, plants present an increase in vegetative vigor (CANNELL, 1976), which may influence the behavior of plant height, specially during C4, where the spatial pattern of PH begins to display differences in relation to previous collections (C1, C2 and C3). These differences increase in C5 and C6, because the area was showing less variations in height, with a predominance of greener regions (Figure 25i,k) and with low CV 6.47 and 6.17% respectively (Table 8). During C7, once more the area presents a similar behavior to collections C1, C2 and C3, with regions of high and low plant height well defined.

Comparing the PV and PH maps from the same campaign it is possible to observe that they have some differences in spatial behavior. The PV data present a greater definition of areas with high (regions in green) and low volume of plants (regions in red), consequently presenting higher CV values (Table 8) in comparison with the PH datasets. The behavior observed in the PH maps is repeated in the PV maps, starting with C4 it is possible to identify that the area presents a distinct behavior in comparison with collections C1, C2, C3 and C7, also presenting some isolated regions that previously presented low volume (regions in red) in collections C1, C2, C3 and C7. During C4, C5 and C6, the same regions presented inverse behavior with high volume (regions in green). This inversion may be linked to the biennial behavior of the plant, which presents opposite behavior in years of high and low yield (PEREIRA et al., 2011; VALADARES et al., 2013).

Another point that is important to highlight is the possibility that this PH and PV information can be used as a guide for the application of phytosanitary products that use plant volume and height as a factor for determining the application dose (GIL et al., 2007, SIEGFRIED et al., 2007). The results presented in Figure 25 generally present well-defined regions with PH and PV variation, indicating that it could be used to direct spraying at a variable rate.

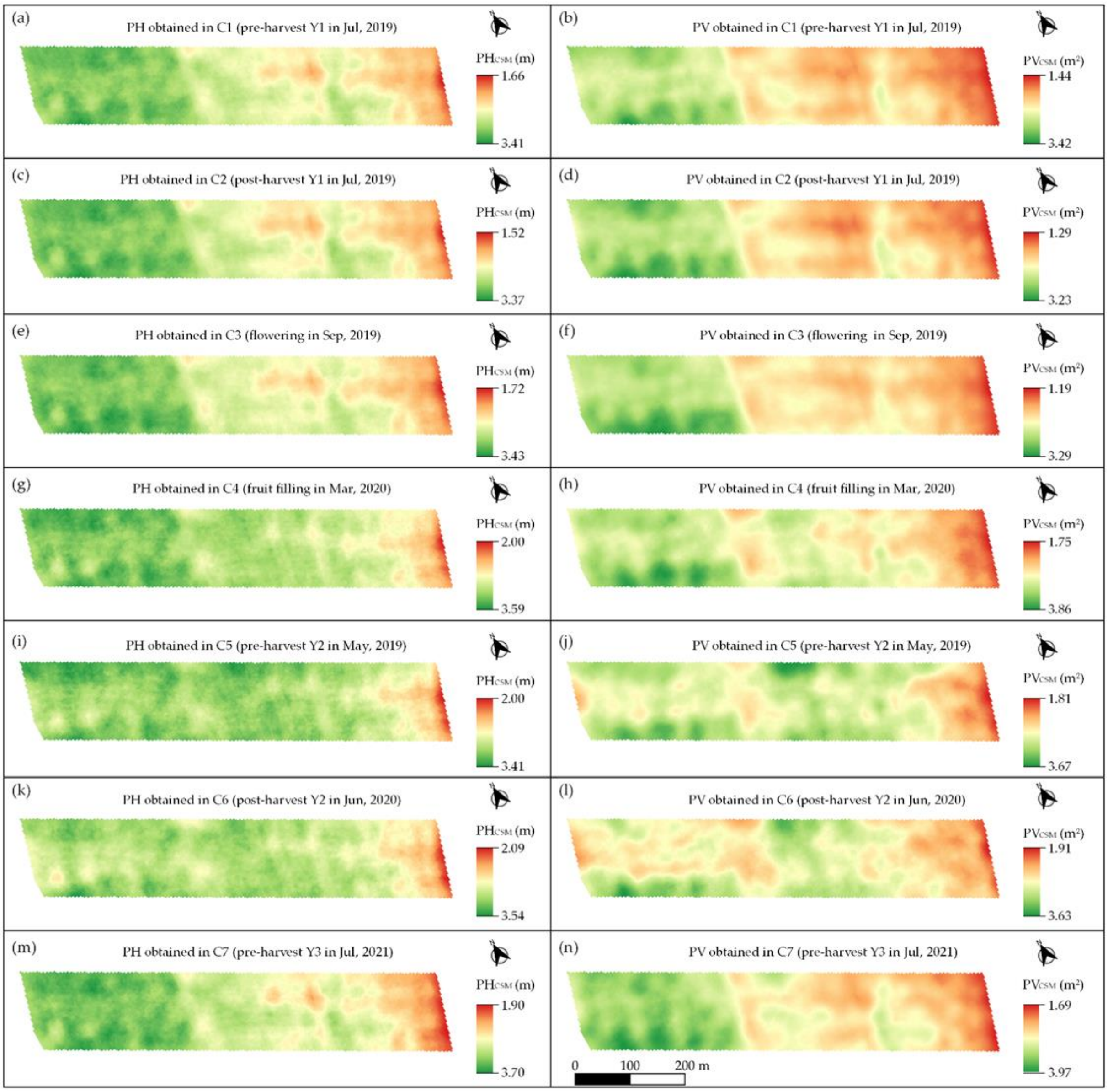

Figure 25. PH and PV interpolated data. (a) PH obtained in C1; (b) PV obtained in C1; (c) PH obtained in C2; (d) PV obtained in the C2; (e) PH obtained in C3; (f) PV obtained in the C3; (g) PH obtained in C4; (h) PV obtained in C4; (i) PH obtained in C5; (j) PV obtained in the C5; (k) PH obtained in C6; (l) PV obtained in the C6; (m) PH obtained in C7; (n) PV obtained in the C7.

Table 9 presents the spatio-temporal relationship of the PH and PV together with yield data obtained from harvests of 2019 (Y1) 2020 (Y2) and 2021 (Y3). When analyzing the correlation of PH and PV data with yield, we observe that the correlation shows negative values for yield data Y1 and Y3, and direct correlation (positive) with yield data from Y2. This result may indicate that the volume and plant height have a more static behavior not following the reversal of yield as shown in Figures 26, 27 and 28, and therefore, show direct correlation with a year and indirect correlation with the following year, when yield zones invert, as observed by Martello et al. (2022).

Another interesting result observed in this study is the relationship of PH and PV data from a specific campaign (e.g. C1 and C5) with Y2 dataset. The PH and PV from C5, during the pre-harvest of 2020, showed low correlation ($r = 0.43$ and 0.28) with the yield of that same year (Y2), while it showed higher values when compared to the data from the C2 collection (one year earlier) with a direct correlation of $r = 0.90$ for PV and $r = 0.92$ for PH. The C5 and C6 datasets showed low values of correlation, both for Y2 and Y3 yield. This might indicate that for this area the best time to conduct flight campaigns aiming to predict crop yield, would be in the odd years, because when we observe the data of Y3 yield the highest values of correlation were found in collection C7 (pre-harvest 2021). Despite the inverse correlation, the predictive potential using the volume data from C7 showed reasonable performance (RPD = 1.56) in predicting Y3 yield.

The performance of the models using PH and PV data for yield prediction showed distinct behavior, with best models found for C1 and C2 datasets when estimating the yield of Y2, with RPD greater than 2.00. According to the classification proposed by Chang et al. (2001), values of RPD between 2 and 3 indicates good models for prediction. The C3 dataset versus the yield of Y2 and the C7 dataset versus the yield of Y3 showed reasonable results ($1.40 \leq RPD < 2.00$). In general, the performance of the PV data presented the lowest RMSE and RMSE% values for yield estimation in comparison with the PH data.

Table 9. Temporal relationship of the PH and PV data with the yield obtained in the 2019 (Y1), 2020 (Y2), and 2021 (Y3) crop. The results show the performance of yield prediction (obtained in the full cross-validation) using simple linear regression models.

In comparinson with Y1		In comparinson with Y2					In comparinson with Y3		
	C1	C1	C2	C3	C4	C5	C5	C6	C7
	PH_{CSM}								
r	-0.25	0.88	0.90	0.86	0.67	0.43	-0.15	-0.15	-0.65
R^2	0.06	0.77	0.81	0.75	0.45	0.18	0.02	0.02	0.42
RMSE	0.28	0.24	0.22	0.25	0.37	0.45	0.43	0.43	0.33
RMSE%	21.38	11.88	10.72	12.37	18.18	22.27	29.10	29.13	22.38
RPD	1.03	2.07	2.30	1.99	1.35	1.11	1.01	1.01	1.32
	PV_{CSM}								
r	-0.32	0.90	0.92	0.86	0.69	0.28	-0.10	-0.07	-0.77
R^2	0.10	0.80	0.84	0.73	0.48	0.08	0.01	0.00	0.59
RMSE	0.27	0.22	0.20	0.26	0.36	0.48	0.43	0.43	0.28
RMSE%	20.92	10.93	9.77	12.75	17.77	23.62	29.30	29.38	18.84
RPD	1.05	2.25	2.52	1.93	1.39	1.04	1.01	1.00	1.56

Figure 26 shows the Y1 yield results, as well as the predicted yield maps obtained from the linear regression using PH and PV independent variables. The maps were generated using the same color scale, promoting a more reliable comparison. Visually the estimated maps did not present spatial similarity with the measured yield data, which corroborates with the results found in Table 9.

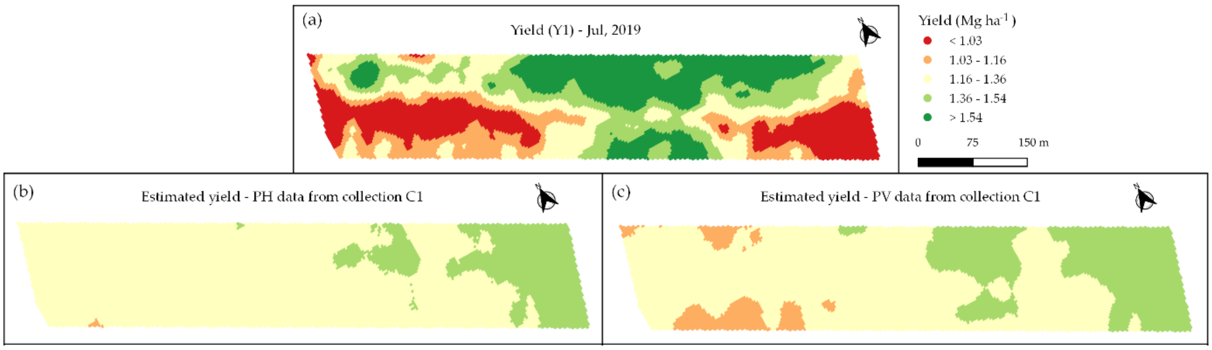

Figure 26. (a) Harvest yield data Y1 (Jul,2019); (b) yield estimated with PH - C1; (c) yield estimated with PV - C1.

The predicted maps for yield Y2 are presented in Figure 27. These maps were obtained using datasets from C1 to C5, in which is possible to observe changes in the spatial behavior of the estimated data, with C1 and C2 presenting great spatial similarity with the observed yield data, and after the collection C3 the maps began to present greater differences in comparison with Y2 due to lower accuracies for yield prediction. This behavior is in agreement with the analytical data presented in Table 9. In addition, the visual analysis indicates that the PV data presented smoother transitions between classes, unlike the PH results.

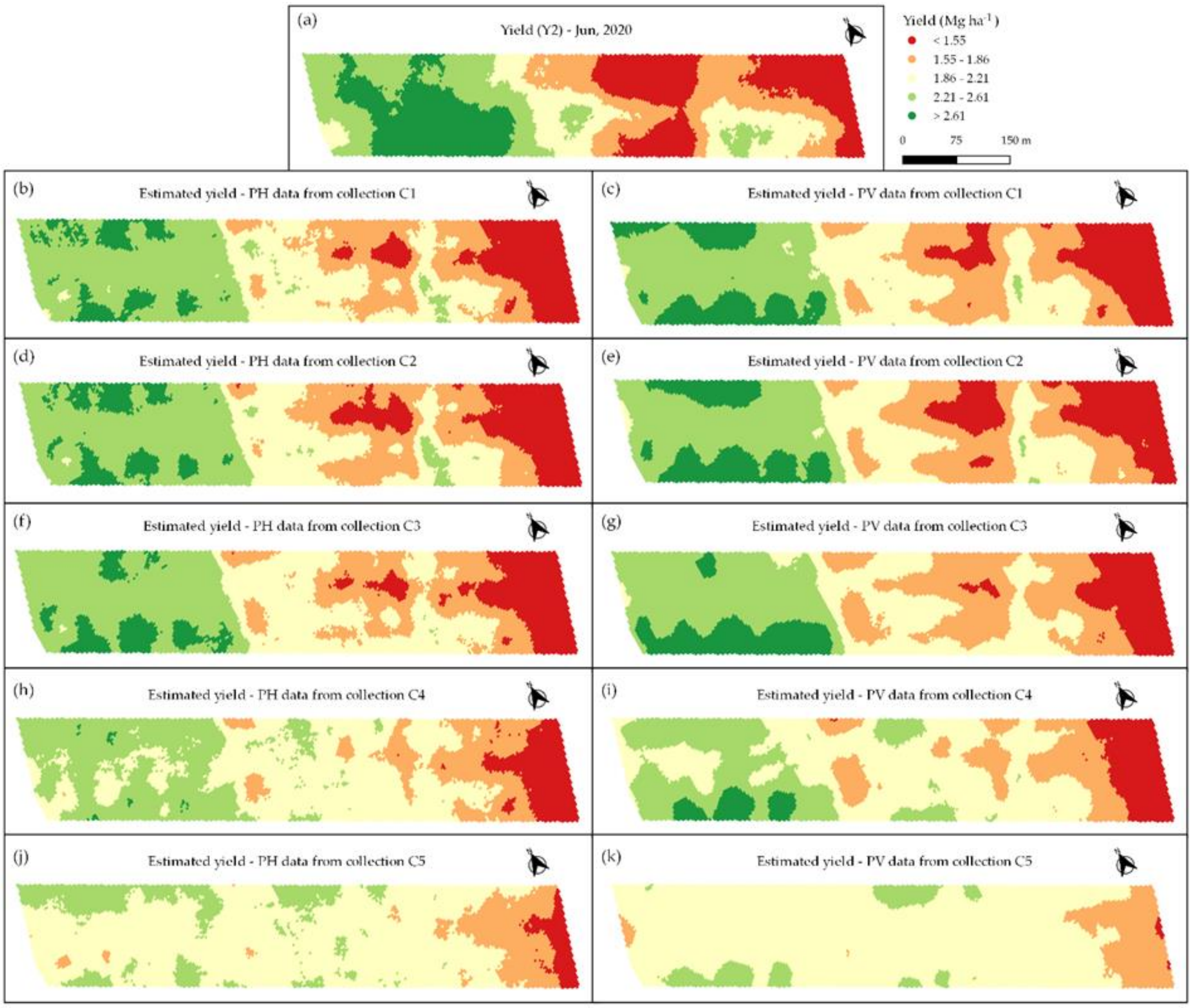

Figure 27. (a) Harvest yield data Y2 (Jun,2020); (b) yield estimated with PH - C1; (c) yield estimated with PV - C1; (d) yield estimated with PH – C2; (e) yield estimated with PV – C2; (f) yield estimated with PH – C3; (g) yield estimated with PV – C3; (h) yield estimated with PH – C4; (i) yield estimated with PV – C4; (j) yield estimated with PH – C5; (k) yield estimated with PV – C5.

The yield data from Y3, along with the predicted yield results using datasets from C5, C6 and C7 are presented in Figure 28. Similar to the results discussed for the maps of Figure 26bc and Figure 27jk, the results from C5 and C6 presented low spatial similarity with the observed yield data, whereas the results from C7 presented a greater similarity with observed yield in comparison to other campaigns. The PV results found in C7 present greater similarity with Y3 in comparison with the PH data. Nonetheless, even with greater similarity the PV results still have limitations for yield prediction.

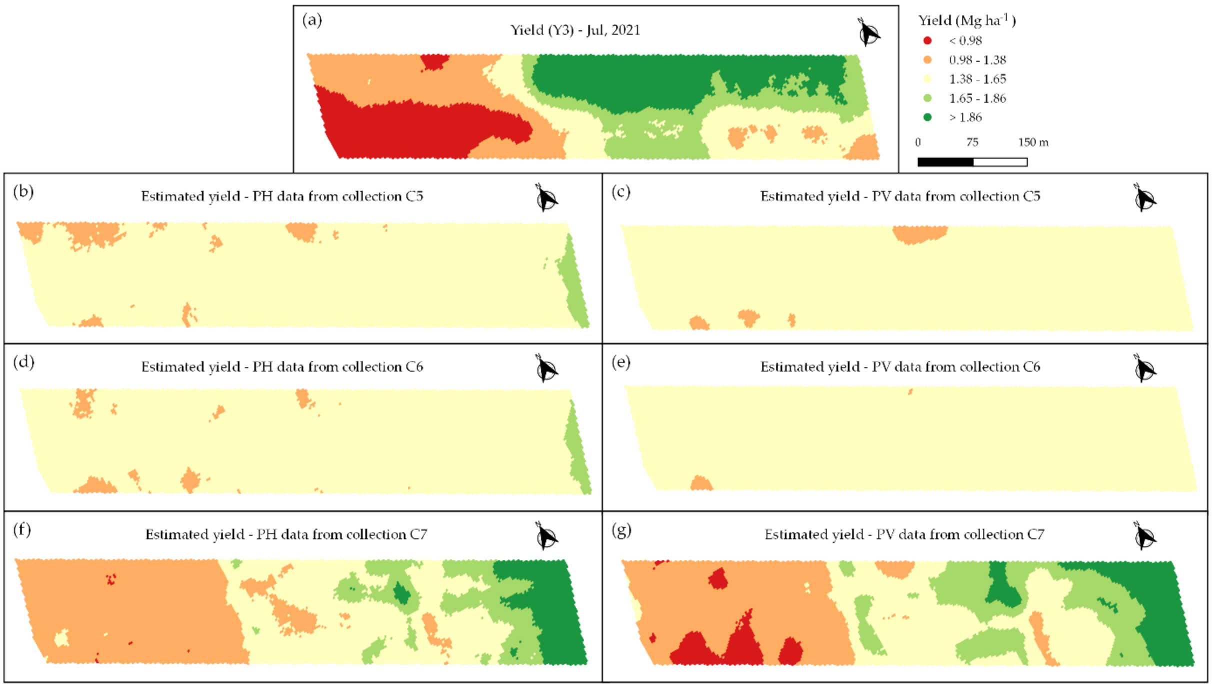

Figure 28. (a) Harvest yield data Y3 (Jul,2021); (b) yield estimated with PH – C5; (c) yield estimated with PV – C5; (d) yield estimated with PH – C6; (e) yield estimated with PV – C6; (f) yield estimated with PH – C7; (g) yield estimated with PV – C7.

Furthermore, the biophysical parameters extracted by the images in this work can also be used to help understand several limitations in coffee crop, as highlighted previously by Santos et al. (2020). As an example, we can mention water deficit (PELOSO et al., 2017), and the solar radiation interception dynamics (PILAU et al., 2016). Moreover, it can also serve as a guideline for site-specific management practices, especially for spraying activities using variable rate systems based on plant height and volume information (GIL et al., 2007, SIEGFRIED et al., 2007). Therefore, we emphasize the importance of further studies that allow validating these types of results as an information source for PA practices.

4.4. CONCLUSIONS

The results presented in this work demonstrated the potential of using aerial images to extract biophysical parameters of coffee plants, as well as allowing the observation of the spatio-temporal relationship of the studied variables with yield data during three consecutive growing seasons.

The data analysis suggests that additional variables are necessary to better understand the spatial variability, because the higher correlations obtained with yield data were observed only during odd years, indicating that plant height and volume alone are not sufficient to characterize the spatial variability from year to year.

Further studies are still needed to evaluate the feasibility of this tool in other production systems, as well as the use of this information to implement site-specific management strategies towards the optimization of inputs.

REFERENCES

ACORSI, M.G.; MIRANDA, F.D.A.; MARTELLO, M.; SMANIOTTO, D.A.; SARTOR, L.R. Estimating biomass of black oat using UAV-based RGB imaging. **Agronomy**, v.9, p.344, 2019.

ALVARES, A.C.; STAPE, J.L.; SENTELHAS, P.C.; GONÇALVES, J.L.M.; SPAROVEK, G. Koppen's climate classification map forBrazil. **Meteorologische zeitschrift**, v. 22, n.6, p.711–728, 2013.

BALLESTEROS, R.; ORTEGA, J.F.; HERNANDEZ, D.; MORENO, M.A. Onion biomass monitoring using UAV-based RGB imaging. **Precision Agriculture.** v.19, p.840–857, 2018.

BARBOSA, B.; FERRAZ, G.A.S.; DOS SANTOS, L.M.; SANTANA, L.; MARIN, D.B.; ROSSI, G.; CONTI, L. Application of rgb images obtained by uav in coffee farming. **Remote Sensing**, v. 13, n. 12, p. 2397, 2021.

BARETH, G.; AASEN, H.; BENDIG, J.; GNYP, M. L.; BOLTEN, A.; JUNG, A.; MICHELS, R.; SOUKKAMÄKI, J. Low-weight and UAV-based Hyperspectral Full-frame Cameras for Monitoring Crops: Spectral Comparison with Portable Spectroradiometer Measurements. Photogrammetrie - Fernerkundung - **Geoinformation**, v. 2015, n. 1, p. 69–79, 2015.

BENDIG, J., YU, K., AASEN, H., BOLTEN, A., BENNERTZ, S., BROSCHEIT, J., GNYP, M.L., BARETH, G. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. **Int. J. Appl. Earth Obs.** v.39, p.79–87, 2015.

BENDIG, J.; BOLTEN, A.; BENNERTZ, S.; BROSCHEIT, J.; EICHFUSS, S.; BARETH, G. Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging. **Remote Sensing**, v.6, p.10395–10412, 2014.

BENDIG, J.; WILLKOMM, M.; TILLY, N.; GNYP, M. L.; BENNERTZ, S.; QIANG, C.; MIAO, Y.; BARETH, G. Very high resolution crop surface models (CSMs) from UAV-based stereo images for rice growth monitoring In Northeast China. **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. XL, p. 4–6, 2013.

BENTO, N.L.; FERRAZ, G.A.S.; BARATA, R.A.P.; SOARES, D.V.; SANTOS, L.M.D.; SANTANA, L.S.; FERRAZ, P.F.P.; CONTI, L.; PALCHETTI, E. Characterization of Recently Planted Coffee Cultivars from Vegetation Indices Obtained by a Remotely Piloted Aircraft System. **Sustainability**, v. 14, n. 3, p. 1446, 2022.

BENTO, N.L.; FERRAZ, G.A.S.; BARATA, R.A.P.; SOARES, D.V.; SANTANA, L.S.; BARBOSA, B.D.S. Estimate and Temporal Monitoring of Height and Diameter of the Canopy of Recently Transplanted Coffee by a Remotely Piloted Aircraft System. **AgriEngineering**, v. 4, n. 1, p. 207-215, 2022a.

CANNELL, M.G.R. Crop physiological aspects of coffee bean yield—a review. **Kenya Coffee**, v. 41, p.245–253, 1976.

CHANG, C. W.; LAIRD, D. A.; MAUSBACH, M. J.; HURBURGH, C. R. Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. **Soil Science Society of America Journal**, 65(2), 480-490, 2001.

COLOMINA, I.; MOLINA, P. Unmanned aerial systems for photogrammetry and remote sensing: a review ISPRS J. Photogramm. **Remote Sensing**, v. 92, n.6, p.79-97, 2014.

CUNHA, J. P. A. R. DA; SIRQUEIRA, M. A; HURTADO, S. M. C. Estimating vegetation volume of coffee crops using images from unmanned aerial vehicles. **Engenharia Agrícola**, v. 39, n. spe, pp. 41-47, 2019.

DÍAZ-VARELA, R.; DE LA ROSA, R.; LEÓN, L.; ZARCO-TEJADA, P. High-Resolution Airborne UAV Imagery to Assess Olive Tree Crown Parameters Using 3D Photo Reconstruction: Application in Breeding Trials. **Remote Sensing**, v. 7, n. 4, p. 4213–4232, 2015.

FLORENZANO, T. G. Imagens de satélites para estudos ambientais. São Paulo: **Oficina de Textos**, 2002

GIL, E.; ESCOLÀ, A.; ROSELL, J. R.; PLANAS, S.; VAL, L. Variable rate application of plant protection products in vineyard using ultrasonic sensors. **Crop Protection**, v. 26, n. 8, p. 1287-1297, 2007.

GÓMEZ-CANDÓN, D.; DE CASTRO, A. I.; LÓPEZ-GRANADOS, F. Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat. **Precision Agriculture**, v. 15, n. 1, p. 44–56, 2014.

GRÜNER, E.; ASTOR, T.; WACHENDORF, M. Biomass prediction of heterogeneous temperate grasslands using an SfM approach based on UAV imaging. **Agronomy**, v. 9, n. 2, p. 54, 2019.

HOFFMEISTER, D.; BOLTEN, A.; CURDT, C.; WALDHOFF, G.; BARETH, G. High-resolution Crop Surface Models (CSM) and Crop Volume Models (CVM) on field level by terrestrial laser scanning. In Sixth International Symposium on Digital Earth: **Models, Algorithms, and Virtual Reality** Vol. 7840, pp. 90-95, 2010.

JOHNSON, L.; HERWITZ, S. Collection of ultra high spatial and spectral resolution image data over California vineyards with a small UAV. **Proc. 30th International Symposium on Remote Sensing of** Environment (ISRSE), n. Ccd, p. 3–5, 2003.

LI, W., NIU, Z., CHEN, H., LI, D., WU, M., ZHAO, W. Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. **Ecological Indicators**, V.67, p.637–648, 2016.

MALAMBO L., POPESCU S.C., MURRAY S.C., PUTMAN E., PUGHNA, HORNE D.W., BISHOP M. Multitemporal field-based plant height estimation using 3D point cloudsgenerated from small unmanned aerial systems high-resolution imagery. **International Journal of Applied EarthObservation and Geoinformation**, v.64, p.31-42, 2018

MALDANER, L. F.; CANATA, T. F.; MOLIN, J. P. An approach to sugarcane yield estimation using sensors in the harvester and zigbee technology. **Sugar Tech**, v.24, p. 813–821, 2022.

MARTELLO, M.; MOLIN, J. P.; BAZAME, H. C. Obtaining and Validating High-Density Coffee Yield Data. **Horticulturae**, v. 8, n. 5, p. 421, 2022.

MINASNY, B.; MCBRATNEY, A.B.; WHELAN, B.M. VESPER Version 1.62; Australian Centre for Precision Agriculture, McMillan Building A05; The University of Sydney: Sydney, Australia, 2005.

MULLA, D. J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. **Biosystems Engineering**, v. 114, n. 4, p. 358–371, 2013.

PELOSO, A. F.; TATAGIBA, S. D.; AMARAL, J. F. T. Limitations of vegetative growth in arabica coffee plants promoted by water deficit. **Revista Engenharia na Agricultura**, v.25 n.2, p.139–147, 2017.

PEREIRA, S.O.; BARTHOLO, G.F.; BALIZA, D.P., SOGREIRA, F.M.; GUIMARÃES, R. J. Productivity and coffee biannuality depending on the crop spacing. **Pesquisa Agropecuária Brasileira**, V.46, n.2, p.152-160, 2011.

PILAU, F. G., & ANGELOCCI, L. R. Patterns of interception of solar radiation by coffee trees as a function of leaf area. **Coffee Science**, v.11, n.1, p.127–136, 2016.

SANKARAN, S.; KHOT, L. R.; ESPINOZA, C. Z.; JAROLMASJED, S.; SATHUVALLI, V. R.; VANDEMARK, G. J.; MIKLAS, P. N.; CARTER, A. H.; PUMPHREY, M. O.; KNOWLES, R. R. N.; PAVEK, M. J. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. **European Journal of Agronomy**, v. 70, n. 3, p. 112–123, 2015.

SANTESTEBAN, L. G.; DI GENNARO, S. F.; HERRERO-LANGREO, A.; MIRANDA, C.; ROYO, J. B.; MATESE, A. High-resolution UAV-based thermal imaging to estimate the instantaneous and seasonal variability of plant water status within a vineyard. **Agricultural Water Management**, v. 183, p. 49–59, 2017.

SANTINATO, F. Inovações tecnológicas para cafeicultura de precisão. **Tese** (Doutorado em Agronomia). Faculdade de Ciências Agrárias e Veterinárias – UNESP, Jaboticabal,SP, 125p. 2016.

SANTOS, L. M., BARBOSA, B. D. D. S., DIOTTO, A. V., MACIEL, D. T., XAVIER, L. A. G. Biophysical parameters of coffee crop estimated by UAV RGB images. **Precision Agriculture**, v. 21, n. 6, p. 1227-1241, 2020.

SCHIRRMANN, M.; GIEBEL, A.; GLEINIGER, F.; PFLANZ, M.; LENTSCHKE, J.; DAMMER, K.-H. Monitoring agronomic parameters of winter wheat crops with low-cost UAV imagery. **Remote Sensing**, v. 8, n. 9, p. 706, 2016.

SIEGFRIED, W.; VIRET, O.; HUBER, B.; WOHLHAUSER, R. Dosage of plant protection products adapted to leaf area index in viticulture. **Crop Protection**, v. 26, n. 2, p. 73-82, 2007.

TILLY, N.; AASEN, H.; BARETH, G. Fusion of plant height and vegetation indices for the estimation of barley biomass. **Remote Sens**. 7, 11449–11480, 2015.

VALADARES SV, NEVES JCL, ROSA GNGP, MARTINEZ HEP, VENEGAS, V.H.A, LIMA P. C. Productivity and biennial production of dense coffee plantations under different doses of N and K. **Pesquisa Agropecuária Brasileira**, v. 13, n.48(3), p.96-303, 2013.

VERHOEVEN, G. Taking computer vision aloft Archaeological three-dimensional reconstructions from aerialphotographs with photoscan. **Archaeol. Prospect**. 18, 67–73, 2011.

5. COFFEE YIELD ESTIMATION USING HIGH RESOLUTION ORBITAL IMAGING AND MACHINE LEARNING

ABSTRACT

The objective of this study was to adjust yield prediction models based on a time series of satellite image data (spectral bands and vegetation indices - VIs) and high density data collected using random forest and multiple linear regression, and indicate the appropriate phenological stage of coffee crop to obtain satellite image data. The study was developed in a commercial production area of 10.24 ha, located in the cerrado biome, state of Minas Gerais, Brazil. Data were obtained using a harvester equipped with a yield monitor that measures the volume of coffee harvested with a resolution of approximately 3.0 m. The data were collected during the 2019, 2020 and 2021 harvest and then converted to weight of processed coffee using a conversion factor. Satellite images from the PlanetScope (PS) platform were used. The data set was composed of high-sampling coffee yield, spectral bands (red, green, blue and NIR), normalized difference vegetation index NDVI and green normalized difference vegetation index GNDVI. RF regression and MLR were fitted to the different data sets composed of the temporal series (spectral bands and VIs) and the selected periods based on the indication of the most suitable period (VIs). The results showed that regardless the regression model used to fit predictive models, they all present similar data, allowing to reproduce yield maps, in general, with similar spatial patterns which can be used not only as method to estimate yield, but also to indicate the spatial variability with high and low yielding zones. In aditions, the more appropriate time to obtain satellite image data was dry season months, July and August.

Keywords: Remote sensing; Precision agriculture; NDVI ; GNDVI

5.1. INTRODUCTION

Coffee crop is the interaction among several factors in the involved in the management of the culture, soil, climate and the plant itself, presenting particularities that contribute to having high spatial variability of yield (SANTANA et al., 2021; SANTINATO, 2016). In this sense, knowing the areas of high and low yield can bring benefits for farms from the adoption of management strategies (CHEMURA et al., 2018; PHAM et al., 2019; MARIN et al., 2021). However, knowledge about coffee yield before harvest is still a challenge for the coffee sector (BAZAME et al., 2021).

Remote sensing (RS) is a potential source of data for monitoring agricultural areas. The RS (MULLA, 2013) at the orbital level stands out for its high coverage rate, allowing the monitoring of large areas combined with a temporal database. In particular, orbital imaging is commonly used in agriculture to identify spectral variations resulting from large-scale soil and crop characteristics (KAYAD et al., 2019; DAMIAN et al., 2020; SANTANA et al., 2021), supporting diagnostics for agronomic parameters of crops and helping farmers make better management decisions (EL-GHANY et al., 2020; FABBRI et al., 2020; KARTHIKEYAN et al., 2020; SARAIVA et al., 2020). However, some of the main limitations related to orbital images are the lack of field data helping as a source for the calibration of predictive models (LOBELL et al., 2015). Furthermore, empirical models for predicting agronomic parameters may have spatial and temporal constraints for application in different fields and seasons (COLAÇO; BRAMLEY, 2018; LUO et al., 2018; LAWES et al., 2019; PADARIAN et al., 2019).

For the coffee crop, the assessment of spatial variability is challenging due to the limited solutions adapted, mainly for yield mapping. Crop yield can be considered the most important data layer to start the investigation of spatial variability within the field (VEGA et al., 2019), to delimit management zones and enhance site-specific management strategies (BRAMLEY et al., 2019; JEFFREY et al., 2020; MOMIN et al., 2019). Currently, the coffee crop yield, by hand or machine harvesting, may be estimated from in-field samples (SILVA et al., 2008; FERRAZ et al., 2012; CARVALHO et al., 2017) or exclusively on mechanical harvesting, using a yield monitor (SARTORI et al., 2002; MARTELLO et al., 2022). Sampling fruits infield is a destructive analysis process that is onerous and expensive, thus impractical for large areas (RAHNEMOONFAR; SHEPPARD, 2017). The yield monitor available on the market measures the volume of harvested fruits using a conveyor belt with cells of known volume (SARTORI et al., 2002; MOLIN et al., 2010; MARTELLO et al., 2022). This method is associated with a high investment cost and its use is exclusive to a single brand of coffee harvester, in addition to requiring calibration, accessibility of data processing for farmers or users and lack of knowledge to manage data from more than one harvester.

Studies have reported the strength of the linear correlation between vegetation indices (VIs) using orbital images and coffee yield prediction. Bernardes et al. (2012) evaluated possible correlations between coffee yield and MODIS-derived vegetation indices in Brazil. Nogueira et al. (2018) reported the use of vegetation indices obtained with images from the Landsat-8 satellite - OLI sensor (Operational Land Imager) to estimate yield. Recently, Silva et al. (2021) evaluated the correlation of different phenological stages spectral responses of coffee in center pivot, under distinct apparent brightness conditions, and observed that coffee yield could be predicted. However, they used low-density yield samples as ground truth data.

An approach to be tested is similar to the solution proposed for sugarcane crops (CANATA et al., 2021), of using a method to estimate sugarcane yield based on satellite images and yield monitor data at high density. They also used machine learning (ML) techniques that can provide greater prediction accuracy compared to pattern statistics, as well as identify datasets (HUNT et al., 2019; JEONG et al., 2016; HOCHACHKA et al., 2007).

Therefore, the objective of this work was: (a) to adjust yield prediction models based on a time series of satellite image data (spectral bands and vegetation indices – Vis) and high density data collected using random forest (RF) and multiple linear regression (MLR) algorithms and, (b) indicate the appropriate phenological stage of coffee crop to obtain satellite image data.

5.2. MATERIALS AND METHODS

The study was conducted in a commercial area located in the municipality of Patos de Minas, Minas Gerais state, Brazil, with central geographic coordinates at 18°32'28.55"S latitude and 46°3'51.17"W longitude, coordinates reference system WGS 84 and altitude of 1025 m (Figure 29). The climate of the study areas is classified as AW, tropical with dry winter and rainy summer, according to the Köppen climate classification (ALVARES et al., 2013). The monthly normal temperatures (1991–2020) range from 18.9 °C in the coldest month (June) to 23.4 °C in the warmest month (October) with average annual temperature around 21.6 °C (INMET, 2022). The area of interest had 10.24 ha cultivated with the species *Coffea arabica L.* (IAC Catuaí 144 variety) planted in 2006 and had its first harvest in 2009, with drip irrigation system.

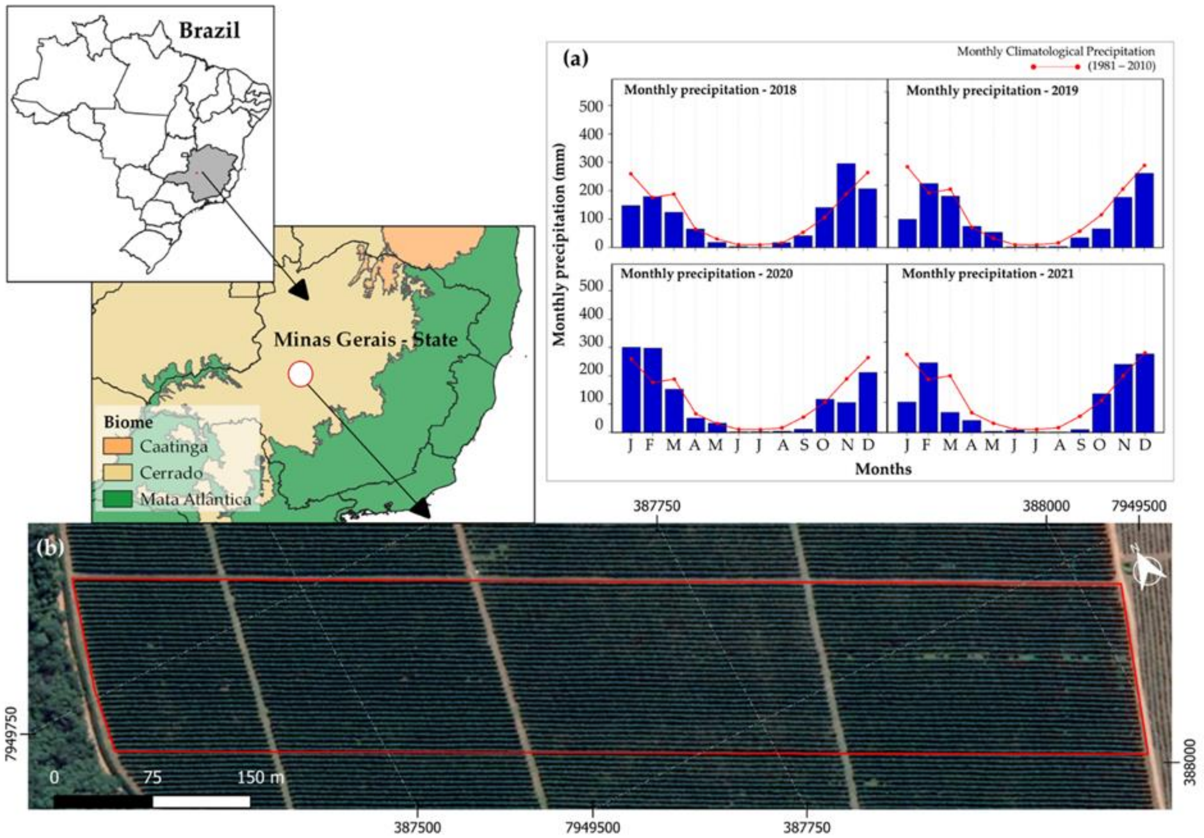

Figure 29. Location of coffee field study site in Brazil. (a) Historical data of monthly precipitation for the years 2018, 2019, 2020 and 2021 and the line with the monthly climatological precipitation of 30 years (1981-2010) for the study area, (Source: CPTEC/INPE, INMET). (b) Study area, the red line represents the area boundary.

5.2.1. Field data

Data were obtained using a harvester K3 Millennium (Jacto, Pompeia, Brazil), equipped with a yield monitor that measures the volume of coffee harvested with a resolution of approximately 3.0 m. The data were collected during the 2019, 2020 and 2021 harvest and then converted to weight of processed coffee using a conversion factor (MARTELLO et al. 2022). Headland maneuver and data close to roads that divide and cross the area were firstly removed. Discrepant data inside the field were filtered using the MapFilter 2.0 software using global filtering with a threshold of 100%, based on the methodology proposed by Maldaner and Molin, 2020. Yield data were interpolated using the Vesper software 1.6 (MINASNY et al., 2006) using the ordinary kriging method with a spatial resolution of 3.0 m × 3.0 m, and with the grid obtained from the central coordinates of the satellite image pixels.

5.2.2. Orbital data

In this study, satellite images from the PlanetScope (PS) (PLANET TEAM, 2017). All PS images had 3.0 m spatial resolution, and each image included four spectral bands: blue (455–515 nm), green (500–590 nm), red (590–670 nm), and near-infrared (780–860 nm). The images obtained were from the sensor Dove Classic – PS2, product Orto Scene - Analytic - Level 3B; these images underwent a series of processes by the vendor, including sensor and radiometric correction, atmospheric correction and conversion to top of atmosphere reflectance (TOA) and geometric correction. This constellation can present daily temporal resolution, but only PS images that did not have clouds were selected, resulting in 33 selected images from 2018 to 2021 seasons; dates are shown in Table 10.

Table 10. Table 1. Dates of the Planet Scope orbital images obtained.

Harvest 1 (2018-2019)	**Harvest 2 (2019-2020)**	**Harvest 3 (2020-2021)**
07/30/18	07/28/19	07/27/20
08/31/18	08/31/19	08/31/20
09/29/18	09/22/19	09/28/20
10/27/18	10/14/19	10/26/20
12/21/18	12/31/19	12/18/20
01/30/19	01/14/20	01/30/21
02/25/19	02/21/20	02/01/21
03/15/19	03/19/20	03/21/21
04/27/19	04/25/20	04/28/21
05/28/19	05/28/20	05/25/21
06/30/19	06/28/20	06/28/21

* No images found without cloud cover in the month of November for the three years.

5.2.3. Statistical analysis

The data set was composed of high-sampling coffee yield, spectral bands (red, green, blue and NIR), normalized difference vegetation index NDVI (ROUSE et al., 1974) and green normalized difference vegetation index GNDVI (GITELSON et al., 1996).

Before fitting yield models, Pearson correlation coefficient (r) was calculated among variables (spectral bands x yield, NDVI x yield and GNDVI x yield) aiming to find those that present the highest linear correlation values with yield and indicate the most suitable periods to obtain satellite imagery data aware of the cloud-covered and temporal resolution limitations. To find the most suitable period it was chosen to use VIs since they can be considered as a dimensionality reduction method which might facilitate the interpretation of the data correlation.

RF regression and MLR were fitted to the different data sets composed of the tempora series (spectral bands and VIs) and the selected periods based on the indication of the most suitable period (VIs). According to Breiman (2001), RF is a classification algorithm composed of several decision tree classifiers, where each tree depends on the values of a random vector sampled independently of the input vector with the identical distribution for all trees within the forest. In this study, RF regression was implemented in Rstudio (R Core Team, 2018) using the "randomForest" package (LIAW; WIENER, 2002). The coffee yield predicted value is the mean fitted response from all the individual trees that resulted from each bootstrapped sample.

Yield predictive models were compared considering the coefficient of determination (R2), root mean squared error (RMSE – equation (6)) and mean absolute error (MAE – equation (7)). It was calculated these parameters for training (2/3), test (1/3) and entire dataset (3/3). Yield maps were generated using the geographic information system (GIS) Quantum GIS – QGIS (QGIS Development Team, 2018)

$$RMSE = \{n^{-1}[\sum(y_i - \hat{y})^2 + \cdots + (y_n - \hat{y})^2]\}^{0.5} \quad (6)$$

where RMSE = root mean squared error, n = number of samples, y = observed variable response and $\hat{y}$ = predicted variable response.

$$MAE = \{n^{-1}[\sum(|y_i - \hat{y}|) + \cdots + (|y_n - \hat{y}|)]\} \quad (7)$$

where MAE = mean absolute error, n = number of samples, y = observed variable response and $\hat{y}$ = predicted variable response.

A flowchart corresponding to the coffee yield prediction and mapping procedure is shown in Figure 30. It presents the process through the stages of data collection using satellite imagery (including pre-processing and data selection), georeferenced coffee yield sampling, data merging (satellite imagery and coffee yield sampling data), data splitting (train and test data) and RF e MLR regression application.

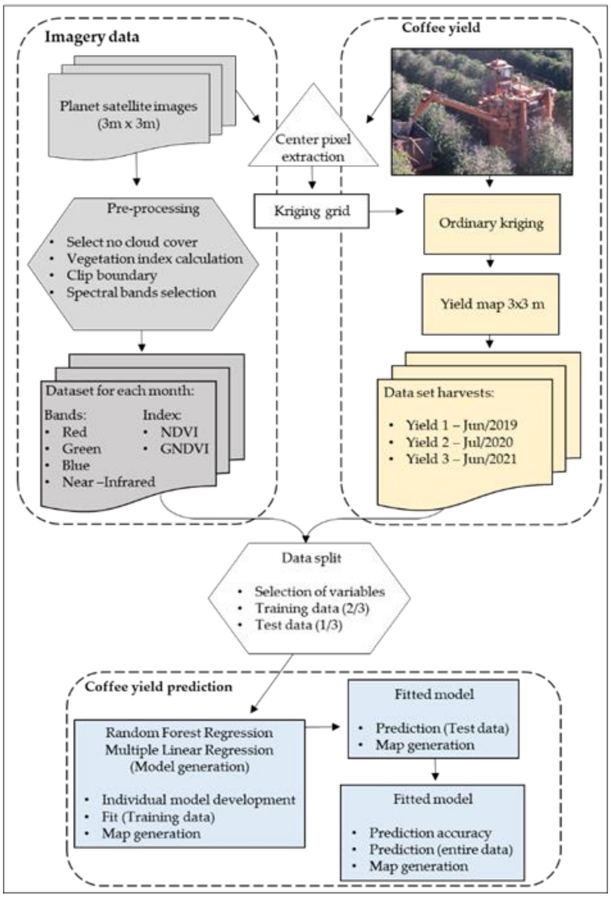

Figure 30. Coffee yield prediction and mapping flowchart

5.3. RESULTS E DISCUSSION

Figure 31 shows the Pearson correlation values among spectral bands, VIs and yield from harvests 2019, 2020 and 2021. It can be seen that highest correlation values with yield are found one year before the harvest (usually in june/july) regardless the season. July and August were the months that presented the highest correlation values with yield, fact related to the plant phenology and already reported by some authors since it can be inferred from this stage the vegetative vigor of the plant (BERNARDES et al., 2012; NOGUEIRA et al., 2018; SILVA et al., 2021).

For harvest 2019, done in (Jun/2019), the highest values of r were 0.64 and 0.65 for NDVI and GNDVI on July and 0.67 for both NDVI and GNDVI for August. In harvest 2020 done in (Jul/2020), the highest r values were 0.89 and 0.85 for NDVI and GNDVI, respectively, on July and 0.85 for NDVI and 0.83 for GNDVI on August. Note that a similar behavior was found for harvest 2021 (Jun/2021), presenting r values of 0.8 and 0.78 for NDVI and GNDVI on July and 0.8 for both NDVI and GNDVI on August.

July and August were also the best months indicated by Silva et al. (2021) to be used to predict coffee yield. These months are related to the phenological stage (dormancy bud phase) of the plant on which the potential production of the next year is already established due to flowering induction occurred in the previous months (OLIVEIRA et al., 2014). The coffee plant enters a mandatory dormancy stage during the period of water deficit (months of dry winter in Brazil) until the rainfall to break flower bud dormancy (CAMARGO; CAMARGO, 2001; LIMA et al., 2021).

They present intense vegetative growth during one year, allowing to produce grains more intensively (reproductive stage) in the next year (RENA & MAESTRI, 1985). During the dormancy phase, plagiotropic branches (productive branches) go into senescence. Due to the vegetative growth stage until the next reproductive stage, ortotropic branches (vegetative stage) receive more nutrients than the plagiotropic branches so it can form new branches and leaves (CAMARGO; CAMARGO, 2001; PEREIRA et al., 2011; GASPARI-PEZZOPANE et al., 2005). The alternation of r values (negative and positive) in the temporal sequence can be explained based on the phenological process of the culture where it can be found positive values right after the prior harvest season and negative values during the development of the crop. These results indicate that the dormancy phase (July and August) based on VIs can be an indicative of potential yield in qualitative terms to the next harvest.

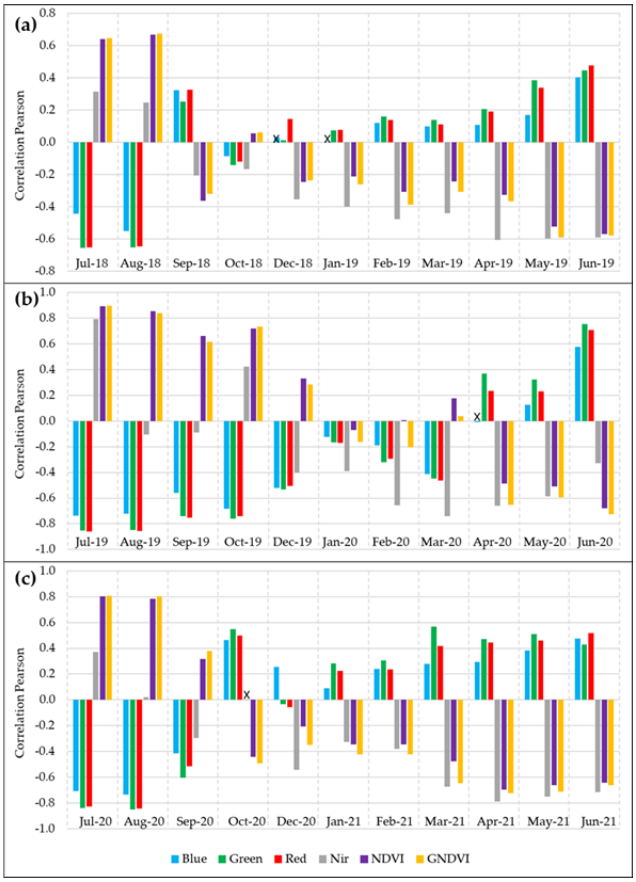

Figure 31. Pearson correlation between spectral bands and vegetative indices with coffee yield at a significance level of 5%. Cross-marks (X) mean that the variables were not significant. Nir: Near infrared, NDVI: Normalized Difference Vegetation Index, GNDVI: Green Normalized Difference Vegetation Index. (a) from the 2019 harvest; (b) from the 2020 harvest; (c) from the 2021 harvest.

Based on the results from NDVI and/or GNDVI, it can be inferred that yield modelling could be conducted using only two months (July and August) for linear models, in case of limited data from satellite imagery which brings benefits since these months comprehend the dry season, not limiting the availability of images due to cloud coverage on the desired scene.

From the results of the linear correlation, yield predictive models were fit according to the different types of dataset: a) temporal series of spectral bands, NDVI, GNDVI and b) data from July and August for NDVI and GNDVI. It was selected the use of July and August because they presented the highest Pearson correlation with yield.

Table 11 shows RMSE, R^2 and MAE results for training, test and entire data set applying RF and MLR to predict coffee yield based on different types of variables and months within seasons. Note that the highest R^2 are found in the models based on the RF regression regardless the data set, similarly found in Wei et al. (2020) and Canata et al. (2021). These results highlight that even if there are some linear correlation among yield and satellite imagery data (spectral bands or vegetation index), the RF in this context is likely to be used instead of MLR regression.

Comparing the results relying on the temporal series (11 months) and only on the two best months (July and August), it can be noted that the fewer variables used, the lower the accuracy regardless the data set and regression model.

Table 11. Root Mean Squared Error (RMSE), Coefficient of Determination (R2), Mean Absolute Error (MAE) results for training, test and entire data set applying Random Forest and Multiple Linear Regression to predict coffee yield based on different types of variables and months within harvests.

Variable	Model	Hs	Mt	Training Data Set (2/3)			Test Data Set (1/3)			Entire Data Set (3/3)		
				RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE
		1	11[a]	0.04	0.99	0.03	0.09	0.91	0.07	0.06	0.96	0.04
Spectral Bands		2	11[a]	0.05	0.99	0.04	0.13	0.93	0.10	0.09	0.97	0.06
		3	11[a]	0.05	0.99	0.03	0.12	0.93	0.09	0.08	0.97	0.05
		1	11[a]	0.04	0.98	0.03	0.10	0.87	0.08	0.07	0.94	0.05
			2[b]	0.10	0.89	0.08	0.20	0.51	0.16	0.14	0.76	0.11
NDVI		2	11[a]	0.06	0.99	0.05	0.15	0.91	0.11	0.10	0.96	0.07
	RF		2[b]	0.10	0.96	0.08	0.21	0.81	0.16	0.15	0.91	0.11
		3	11[a]	0.07	0.98	0.05	0.16	0.86	0.12	0.11	0.94	0.08
			2[b]	0.12	0.93	0.09	0.25	0.68	0.19	0.17	0.84	0.13
		1	11[a]	0.05	0.97	0.04	0.12	0.83	0.09	0.08	0.93	0.05
			2[b]	0.10	0.90	0.08	0.19	0.53	0.16	0.14	0.77	0.11
GNDVI		2	11[a]	0.06	0.98	0.05	0.15	0.90	0.12	0.10	0.96	0.07
			2[b]	0.10	0.96	0.08	0.21	0.82	0.16	0.15	0.91	0.11
		3	11[a]	0.07	0.97	0.06	0.17	0.84	0.14	0.12	0.93	0.08
			2[b]	0.12	0.93	0.09	0.24	0.69	0.19	0.17	0.85	0.12
		1	11[a]	0.12	0.81	0.10	0.12	0.81	0.10	0.12	0.81	0.10
Spectral Bands		2	11[a]	0.17	0.88	0.13	0.17	0.88	0.14	0.17	0.88	0.14
		3	11[a]	0.16	0.86	0.13	0.16	0.86	0.13	0.16	0.86	0.13
		1	11[a]	0.14	0.77	0.11	0.14	0.77	0.11	0.14	0.77	0.11
			2[b]	0.20	0.49	0.17	0.20	0.50	0.16	0.20	0.50	0.17
NDVI		2	11[a]	0.19	0.86	0.15	0.19	0.86	0.15	0.19	0.86	0.15
			2[b]	0.21	0.83	0.16	0.21	0.82	0.17	0.21	0.82	0.16
	MLR	3	11[a]	0.21	0.77	0.17	0.21	0.76	0.16	0.21	0.76	0.17
			2[b]	0.24	0.70	0.19	0.23	0.70	0.19	0.24	0.70	0.19
		1	11[a]	0.14	0.74	0.11	0.14	0.74	0.12	0.14	0.74	0.11
			2[b]	0.20	0.51	0.16	0.20	0.52	0.16	0.20	0.51	0.16
GNDVI		2	11[a]	0.18	0.87	0.14	0.19	0.86	0.15	0.18	0.86	0.15
			2[b]	0.21	0.82	0.17	0.21	0.82	0.17	0.21	0.82	0.17
		3	11[a]	0.22	0.75	0.17	0.22	0.75	0.17	0.22	0.75	0.17
			2[b]	0.24	0.70	0.19	0.23	0.70	0.19	0.24	0.70	0.19

Reg = regression; Hs = Harvest; Mt = Months; RF = Random forest regression; MLR = Multiple linear regression
a it was considered satellite imagery data from January, February, March, April, May, June, July, August, September, October and December.
b it was considered satellite imagery data from July and August.

Figures 32, 33 and 34 represent the yield maps from yield monitor and different yield predictive models. From Figures 32A, 33A and 34A it can be visualized the biennial bearing effect of coffee yield, a phenomena expected and largely reported in the literature (BERNARDES et al., 2012; CARVALHO et al., 2020) on which harvest 1 and 3 could be considered as low yielding season and harvest 2 as high yielding season. However, usually the effect is reported in low spatial resolution, different from this work that shows results based on high-density data. Thus, from this results it can be inferred not only the variability between years, but also the variability within the field which can provide data for decisions makers to improve crop management (MOLIN et al. 2010; ANGNES et.al 2021; MARTELLO et al. 2022).

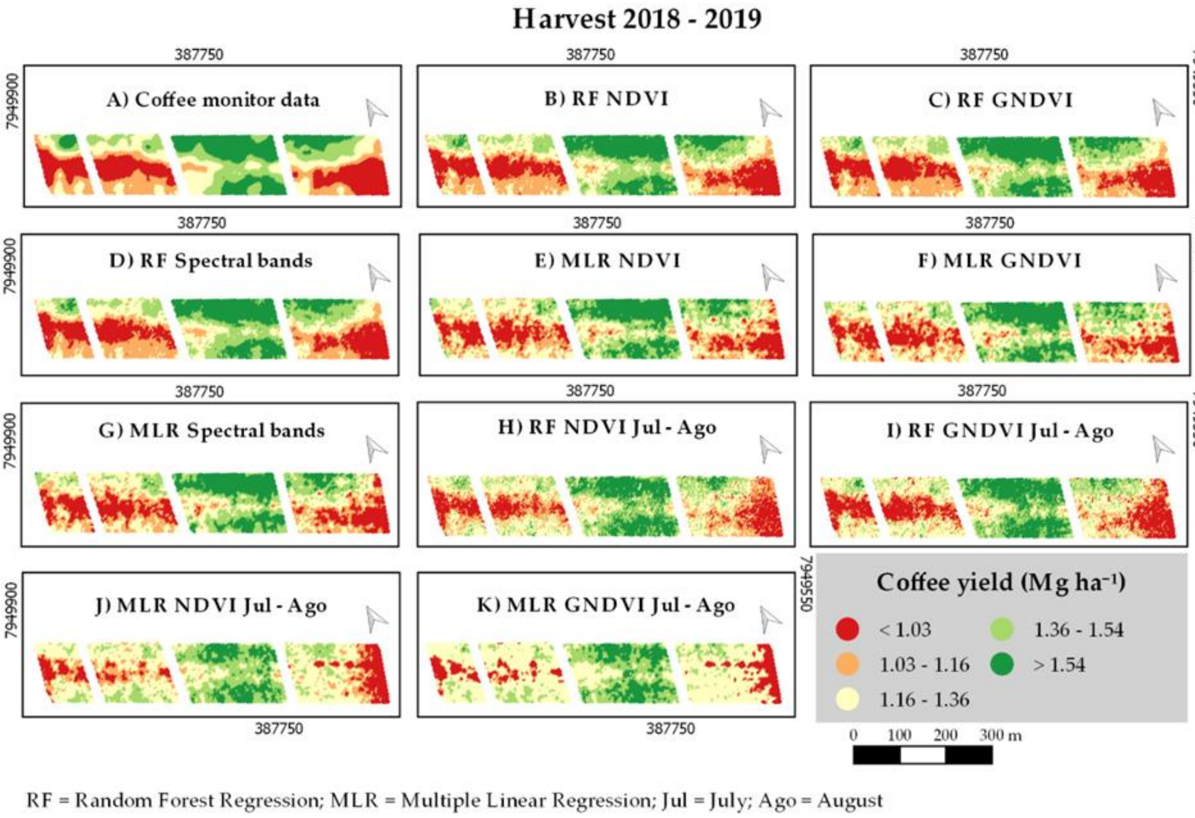

Figure 32. Figure 5. Coffee yield (Mg ha-1) maps generate for harvest 1 (2018-2019) from (A) Coffe monitor data(B) RF regression model based on spectral bands, (C) RF regression based on Normalized Difference Vegetation Index (NDVI), (D) RF regression based on Green Normalized Vegetation, (E) MLR based on spectral bands, (F) MLR based on NDVI, (G) MLR based on GNDVI, (H) RF regression based on NDVI obtained from July(jul) and August(ago), (I) RF regression based on GNDVI obtained from July and August, (J) MLR based on NDVI obtained from July and August, Index (GNDVI) and (K) MLR based on GNDVI obtained from July and August.

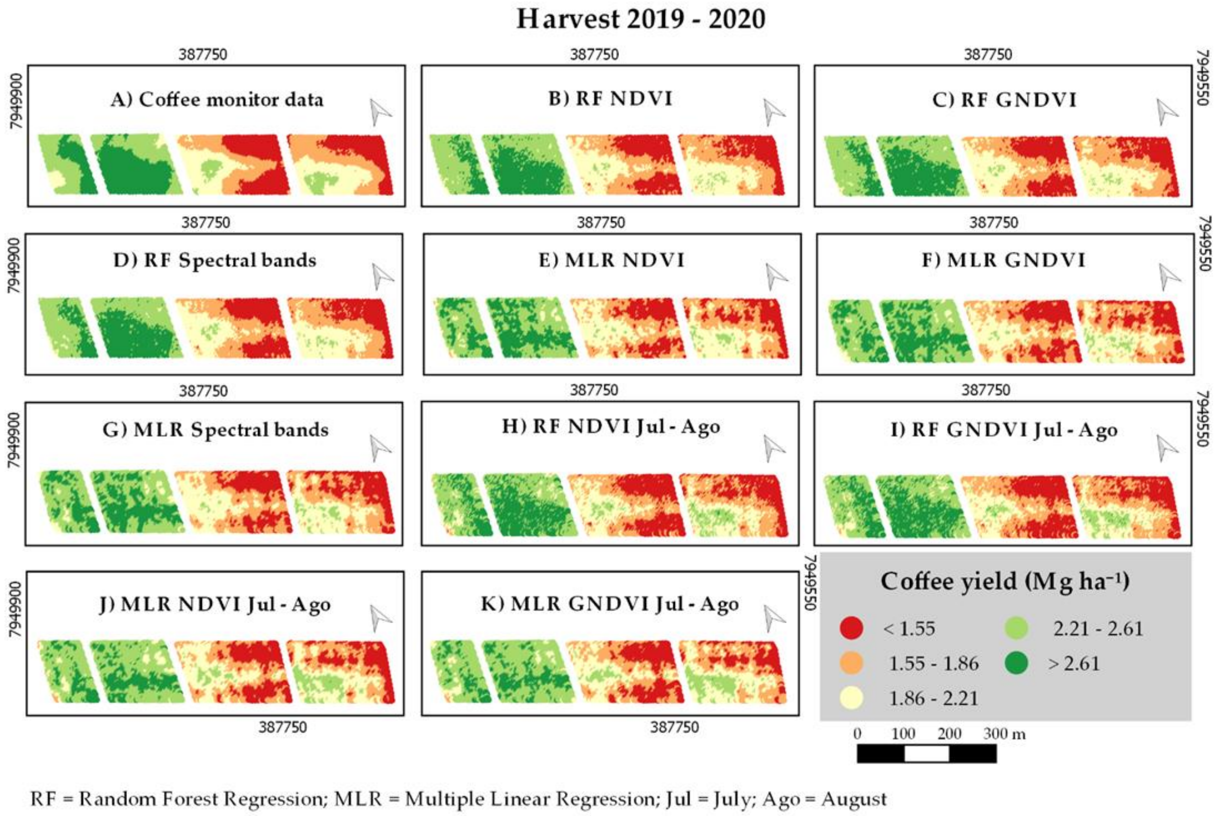

Figure 33. Figure 6. Coffee yield (Mg ha-1) maps generate for harvest 2 (2019-2020) from (A) Coffe monitor data(B) RF regression model based on spectral bands, (C) RF regression based on Normalized Difference Vegetation Index (NDVI), (D) RF regression based on Green Normalized Vegetation, (E) MLR based on spectral bands, (F) MLR based on NDVI, (G) MLR based on GNDVI, (H) RF regression based on NDVI obtained from July(jul) and August(ago), (I) RF regression based on GNDVI obtained from July and August, (J) MLR based on NDVI obtained from July and August, Index (GNDVI) and (K) MLR based on GNDVI obtained from July and August.

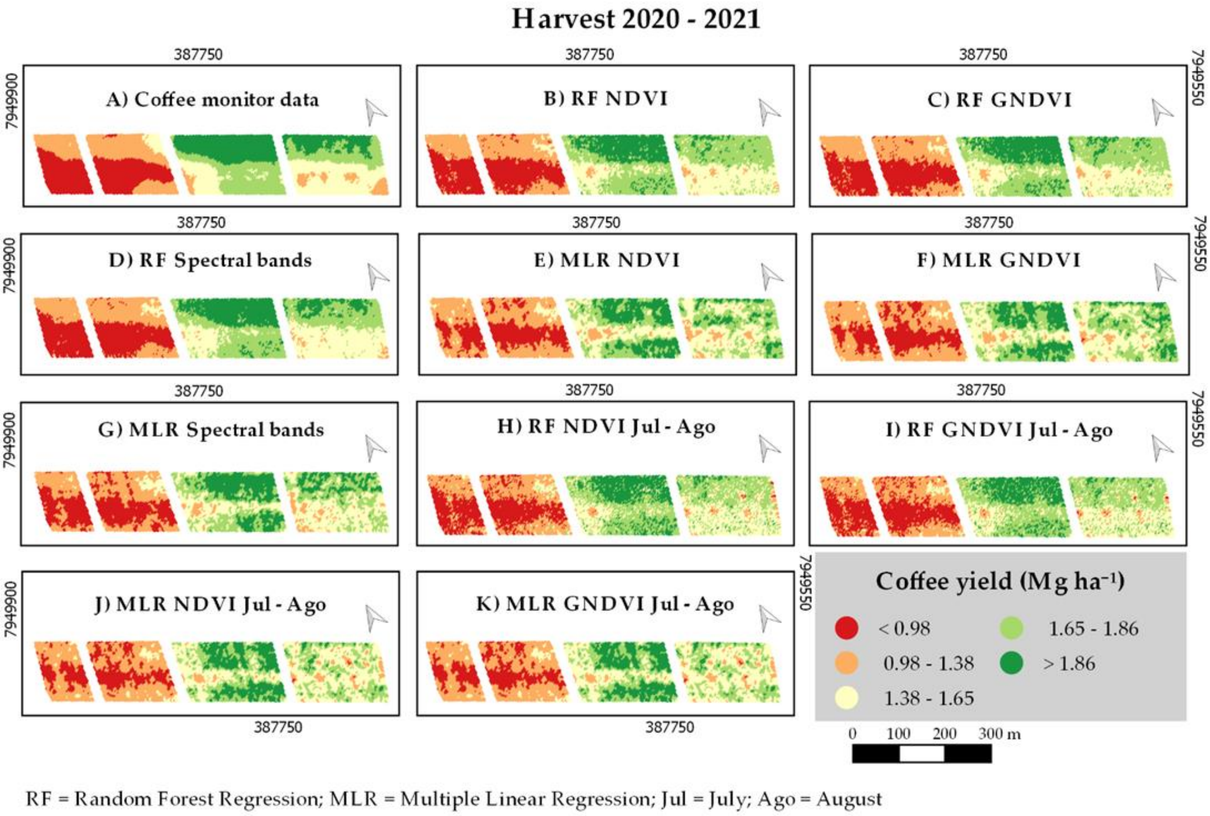

Figure 34. Figure 7. Coffee yield (Mg ha-1) maps generate for harvest 3 (2020-2021) from (A) Coffee monitor data(B) RF regression model based on spectral bands, (C) RF regression based on Normalized Difference Vegetation Index (NDVI), (D) RF regression based on Green Normalized Vegetation, (E) MLR based on spectral bands, (F) MLR based on NDVI, (G) MLR based on GNDVI, (H) RF regression based on NDVI obtained from July(jul) and August(ago), (I) RF regression based on GNDVI obtained from July and August, (J) MLR based on NDVI obtained from July and August, Index (GNDVI) and (K) MLR based on GNDVI obtained from July and August.

Looking at the maps from Figures 32, 33 and 34, it can be noted that RF regression models presented smoother results when compared to MLR, regardless the database on which the larger the number of available variables (spectral bands), the smoothers the results. The results from VIs and MLR are also interesting to be used, but aware of the computing power to process data, the use of spectral bands to fit coffee yield prediction models should be first accounted, as shown in other studies but for different crops: carrot (WEI et al., 2020) and sugarcane (CANATA et al., 2021).

In addition, it is highlighted that regardless the regression model used to fit predictive models, they all present similar data, allowing to reproduce yield maps, in general, with similar spatial patterns which can be used not only as method to estimate yield, but also to indicate the spatial variability with high and low yielding zones.

The use of VIs in this work allowed to identify qualitatively potential management zones as they present strong linear correlation with yield (Figure 31) due to established importance for coffee physiology if the image is acquired in the ideal crop phenological phase. Thus, NDVI/GNDVI data collection after harvesting (dry season – July and August – dormancy phase) allow to improve management decisions regarding the crop to ensure it to express their yield potential as much as possible. It is also possible to infer qualitatively about high and low yielding zones one-year prior the harvest season, a valuable data to be used by the decision maker to enhance crop management considering the crop variability.

Note that the strategy described in this study, using NDVI/GNDVI data obtained during the dormancy phase of coffee plants that allow to identify the qualitative yield potential zones before the flowering occurs, can be critical to decision-making in coffee production systems. Irrigation planning can be developed based on this information to match the crucial moment that enhance coffee quality (RODRIGUES et al., 2017, MIRANDA et al., 2020). Also, the optimal soil correctives and fertilizers application can be planned using the zones prediction, following the principles of PA (ISPA, 2022). The high spatial density in which the qualitative zones can be predicted using this strategy, before the blossoming and coffee fruits development, can guide the prescription of coffee management along the season.

The dormancy phase is critical to the coffee plant as it is waiting for the first rains to bloom (DAMATTA et al., 2007). Aware that blossoming occurs uneven within the field and it is related to yield, efforts studying the application of site-specific irrigation based on the yield potential zones, which can be developed based on NDVI/GNDVI data obtained on the dry season) should be done.

Several studies have been conducted evaluating the irrigation and coffee attributes qualitatively and quantitatively. For example, Rodrigues et al. (2017) found that better coffee quality is found in irrigated areas, but the watering time is crucial. In addition, several morphological structures are also enhanced which can also be expressed in more yield. Damatta et al. (2007) also found that irrigated areas may yield better coffee quality, but they highlighted that the timing to provide water can affect negatively the quality. Therefore, as already stated by the previous authors, that irrigation can improve the yield expression in terms of quality and quantity, as long as irrigation occurs at the correct time, an additional data layer (yield potential) could also be used to guarantee that yield expression following the principles of PA.

Another management that can be guided by the yield map is the fertilization (MOLIN et al., 2010). Therefore, fertilizer application should be distributed unevenly according to their yield potential based on the NDVI/GNDVI map.

In terms of quantitative data, it was demonstrated in this work that the use of field data (harvester yield data) with satellite images applying regression models are suitable for estimating coffee yield. Future approaches can explore the use of these models to extrapolate the yield estimate to nearby areas, as well as evaluate the potential of the technique for applying a model to the next year.

5.4. CONCLUSIONS

It was possible to observe with the set of Planet Scope orbital images and the yield data during three harvests, that there is a direct correlation between the NDVI and GNDVI indices with yield one year before harvest, especially in the months of July and August (post-harvest and the dormancy phase of the plant).

Using regression models (RF and MLR) it was possible to estimate the coffee yield, and the RF regression models showed the highest R^2 values compared to the MLR. Comparing the results based on the time series (11 months) and only on the two best months (July and August), it is noted that the fewer variables used, the lower the accuracy independent of the dataset and regression model. However, when it is not possible to obtain a set of annual images, the results showed that it is possible to estimate productivity with images during the dormancy phase of the plant one year before harvest.

In addition, it is noteworthy that regardless of the regression model used to adjust the predictive models, they all present similar data, allowing to reproduce yield maps, in general, with similar spatial patterns that can be used not only as a method to estimate yield, but also to indicate the spatial variability with zones of high and low productivity, observing the behavior of productive alternation in the area and indicating the presence of biennial yield.

REFERENCES

ALVARES, A. C.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M.; SPAROVEK, G. Koppen's climate classification map for Brazil. **Meteorologische zeitschrift,** v. 22, n.6, p.711–728, 2013.

ANGNES, G.; MARTELLO, M.; FAULIN, G.D.C.; MOLIN, J.P.; ROMANELLI, T.L. Energy efficiency of variable rate fertilizer application in coffee production in Brazil. **AgriEngineering**, v.3, p.815–826, 2021.

BAZAME, H. C.; MOLIN, J. P.; ALTHOFF, D.; MARTELLO, M. Detection, classification, and mapping of coffee fruits during harvest with computer vision. **Comput. Electron. Agric.** v.183, p.106066, 2021.

BERNARDES, T.; MOREIRA, M. A.; ADAMI, M.; GIAROLLA, A.; RUDORFF, B. F. T. Monitoring Biennial Bearing Effect on Coffee Yield Using MODIS Remote Sensing Imagery. **Remote Sensing**, v.4, p.2492 – 2509, 2012.

BRAMLEY, R.G.V.; OUZMAN, J.; GOBBETT, D.L. Regional scale application of the precision agriculture thought process to promote improved fertilizer management in the Australian sugar industry. **Precision Agriculture**, v.20, p.362–378, 2019.

BREIMAN, L. Random forests. Machine learning, 45(1), 5-32, 2001.

CAMARGO, A. P.; CAMARGO, M. B. P. Definition and outline for the phenological phases of arabic coffee under Brazilian tropical conditions. **Bragantia**, v.60, p.65–68, 2001.

CANATA, T. F.; WEI, M. C. F.; MALDANER, L. F.; MOLIN, J. P. Sugarcane Yield Mapping Using High-Resolution Imagery Data and Machine Learning Technique. **Remote Sensing**, v.13, p.232,2021.

CARVALHO, L.C.; SILVA, F.M.; FERRAZ, G.A.; STRACIERI, J.; FERRAZ, P.F.; AMBROSANO, L. Geostatistical analysis of Arabic coffee yield in two crop seasons. **Rev. Bras. Eng. Agrícola E Ambient**, v.21, p.410–414, 2017.

CARVALHO, H. F.; GALLI, G.; FERRAO, L. F. V.; NONATO, J. V. A.; PADILHA, L.; MALUF, M. P.; RESENDE JR, M. F. R.; FILHO, O. G.; FRITSCHE-NETO, R. The effect of bienniality on genomic prediction of yield in arabica coffee. **Euphytica**, 216, 2020.

CHEMURA, A.; MUTANGA, O.; ODINDI, J.; KUTYWAYO, D. Mapping spatial variability of foliar nitrogen in coffee (Coffea arabica L.) plantations with multispectral Sentinel-2 MSI data. **ISPRS Journal of Photogrammetry and Remote Sensing**, 138, 1-11, 2018.

COLAÇO, A. F.; BRAMLEY, R. G. Do crop sensors promote improved nitrogen management in grain crops?. **Field Crops Research**, 218, 126-140, 2018.

DAMATTA, F. M.; RONCHI, C. P.; MAESTRI, M.; BARROS, R. S. Ecophysiology of coffee growth and production. **Brazilian Journal of Plant Physiology**, v.19, p.485 – 510, 2007.

DAMIAN, J.M.; PIAS, O.H.D.C.; CHERUBIN, M.R.; DA FONSECA, A.Z.; FORNARI, E.Z.; SANTI, A.L. Applying the NDVI from satellite images in delimiting management zones for annual crops. **Sci. Agricoia** 77, 2020.

EL-GHANY, A.; NESREEN, M.; EL-AZIZ, A.; SHADIA, E.; MAREI, S. S. A review: Application of remote sensing as a promising strategy for insect pests and diseases management. **Environmental Science and Pollution Research**, 27, 33503-33515, 2020.

FABBRI, C.; MANCINI, M.; DALLA MARTA, A.; ORLANDINI, S.; NAPOLI, M. Integrating satellite data with a Nitrogen Nutrition Curve for precision top-dress fertilization of durum wheat. **European Journal of Agronomy**, 120, 126148, 2020.

FERRAZ, G. A. S.; DA SILVA, F. M.; CARVALHO, L. C. C.; ALVES, M. D. C.; FRANCO, B. C. Spatial and temporal variability of phosphorus, potassium and of the yield of a coffee field. Eng. Agric. v.32, p.140–150, 2012.

GASPARI-PEZZOPANE, C.D.; MEDINA FILHO, H.P.; BORDIGNON, R.; SIQUEIRA, W.J.; AMBRÓSIO, L.A.; MAZZAFERA, P. Environmental influences on the intrinsic outturn of coffee. **Bragantia**, v.64, p.39–50, 2005.

GITELSON, A.A.; KAUFMAN, Y.J.; MERZLYAK, M.N. Use of a green channel in remote sensing of global vegetation from EOS- MODIS. **Remote Sens. Environ.** v.58, p.289–298, 1996.

HOCHACHKA, W.M.; CARUANA, R.; FINK, D.; MUNSON, A.; RIEDEWALD, D.; SOROKINA, D.; KELLING, S. Data-mining discovery of pattern and process in ecological systems. J. **Wildl. Manag**. 71, 2427–2437, 2007.

HUNT, D.A.; TABOR, K.; HEWSON, J.H.; WOOD, M.A.; REYMONDIN, L.; KOENIG, K.; SCHMITT-HARSH, M.; FOLLETT, F. Review of Remote Sensing Methods to Map Coffee Production Systems. **Remote Sensing**, v.12, p.2041, 2020.

INMET. Brazil Climate Normals 1991–2020. 2022. Available online: https://portal.inmet.gov.br/normais (accessed on 30 April 2022).

ISPA International Society of Precision Agriculture. **Precision Agriculture Definition**. https://www.ispag.org/about/definition.

JEFFREY, G. R.; GRIFFIN, T. S.; FLEISHER, D. H.; NAUMOVA, E. N.; KOCH, M.; WARDLOW, B. D. Mapping sub-field maize yields in Nebraska, USA by combining remote sensing imagery, crop simulation models, and machine learning. **Precision Agriculture**. v. 21,p. 678–694, 2020.

JEONG, J.H.; RESOP, J.P.; MUELLER, N.D.; FLEISHER, D.H.; YUN, K.; BUTLER, E.E.; TIMLIN, D.J.; SHIM, K.M.; GERBER, J.S.; REDDY, V.R.; et al. Random forests for global and regional crop yield predictions. **PLoS ONE** 11, e0156571, 2016.

KARTHIKEYAN, L.; CHAWLA, I.; MISHRA, A. K. A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. **Journal of Hydrology**, 586, 124905, 2020.

KAYAD, A.; SOZZI, M.; GATTO, S.; MARINELLO, F.; PIROTTI, F. Monitoring within-field variability of corn yield using Sentinel-2 and machine learning techniques. **Remote Sens.** 11, 2873, 2019.

LAWES, R. A.; OLIVER, Y. M.; HUTH, N. I. Optimal nitrogen rate can be predicted using average yield and estimates of soil water and leaf nitrogen with infield experimentation. **Agronomy Journal**, 111(3), 1155-1164, 2019.

LIAW, A.; WIENER, M. Classification and regression by randomForest. **R news**, 2(3), 18-22, 2002.

LIMA, A. A.; SANTOS, I. S.; TORRES, M. E. L.; CARDON, C. H.; CALDEIRA, C. F.; LIMA, R. R.; CHALFUN-JUNIOR; A. Drought and re-watering modify ethylene production and sensitivity, and are associated with coffee anthesis. **Environmental and Experimental Botany**, 181, 104289, 2021.

LOBELL, D.B.; THAU, D.; SEIFERT, C.; ENGLE, E.; LITTLE, B. A scalable satellite-based crop yield mapper. **Remote Sens. Environ.** 164, 324–333, 2015.

LUO, X.; YE, Z.; XU, H.; ZHANG, D.; BAI, S.; YING, Y. Robustness improvement of NIR-based determination of soluble solids in apple fruit by local calibration. **Postharvest Biology and Technology**, 139, 82-90, 2018.

MALDANER, L.F.; MOLIN, J.P. Data processing within rows for sugarcane yield mapping. **Scientia Agricola**, v. 77, n.5, 2020.

MARIN, D. B.; FERRAZ, G. A. S.; GUIMARAES, P. H. S.; SCHWERZ, F.; SANTANA, L. S.; BARBOSA, B. D. S.; ROSSI, G. Remotely Piloted Aircraft and Random Forest in the Evaluation of the Spatial Variability of Foliar Nitrogen in Coffee Crop. **Remote Sensing**, 13(8), 1471, 2021.

MARTELLO, M.; MOLIN, J. P.; BAZAME, H. C. Obtaining and Validating High-Density Coffee Yield Data. **Horticulturae**. v.8, p.421 – 427, 2022.

MINASNY, B.; MCBRATNEY, A. B.; WHELAN, B. M. VESPER Version 1.62; Australian Centre for Precision Agriculture, McMillan Building A05; The University of Sydney: Sydney, Australia, 2005.

MIRANDA, F. R.; DRUMOND, L. C. D.; RONCHI, C. P. Synchronizing coffee blossoming and fruit ripening in irrigated crops of the Brazilian Cerrado Mineiro Region. **Australian Journal of Crop Science**, 14(4), 605-613, 2020.

MOLIN, J.P.; MOTOMIYA, A.V.D.A.; FRASSON, F.R.; FAULIN, G.D.C.; TOSTA, W. Test procedure for variable rate fertilizer on coffee. Acta Scientiarum. **Agronomy**, v.32, p.569–575, 2010.

MOMIN, M.A.; GRIFT, T.E.; VALENTE, D.S.; HANSEN, A.C. Sugarcane yield mapping based on vehicle tracking. **Precision Agriculture**. 20, 896–910, 2019.

MULLA, D. J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. **Biosystems Engineering**, v. 114, n. 4, p. 358–371, 2013.

NOGUEIRA, S.M.C.; MOREIRA, M.A.; VOLPATO, M.M.L. Relationship between coffee crop productivity and vegetation indexes derived from OLI/Landsat-8 Sensor Data with and without topographic correction. **Eng. Agríc.** 38, 387–394, 2018.

OLIVEIRA, R. R.; CESARINO, I.; MAZZAFERA, P.; DORNELAS, M. C. Flower development in Coffea arabica L.: new insights into MADS-box genes. **Plant reproduction**, 27(2), 79-94, 2014.

PADARIAN, J.; MINASNY, B.; MCBRATNEY, A. B. Transfer learning to localise a continental soil vis-NIR calibration model. **Geoderma**, 340, 279-288, 2019.

PEREIRA, S. O.; BARTHOLO, G. F.; BALIZA, D. P.; SOGREIRA, F. M.; GUIMARÃES, R. J. Growth, productivity and bienniality of coffee plants according to cultivation spacing. **Pesquisa Agropecuária Brasileira**, v.46, p.152–160, 2011.

PHAM, Y.; REARDON-SMITH, K.; MUSHTAQ, S.; COCKFIELD, G. The impact of climate change and variability on coffee production: a systematic review. **Climatic Change**, 156(4), 609-630, 2019.

PLANET TEAM. Planet Application Program Interface. In Space for Life on Earth; Planet Team: San Francisco, CA, USA, 2017; Available online: https://api.planet.com

QGIS DEVELOPMENT TEAM. QGIS Geographic Information System. Open Source Geospatial Foundation Project. 2018. Available online: http://qgis.osgeo.org

R CORE TEAM. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2018.

RAHNEMOONFAR, M.; SHEPPARD, C. Deep count: fruit counting based on deep simulated learning. **Sensors**, 17(4), 905, 2017.

RENA, A.B.; MAESTRI, M. Fisiologia do cafeeiro. **Inf. Agropecuário**, v.11, p.26–40, 1985.

RODRIGUES, W. N.; BRINATE, S. B. V.; MARTINS, L.D.; COLODETTI, T.V., TOMAZ, M. A. Genetic variability and expression of agro-morphological traits among genotypes of Coffea arabica being promoted by supplementary irrigation. **Genetics and Molecular Research**. 16, 1–12, 2017.

ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the great plains with ERTS. In Proceedings of the Earth Resources Technology Satellite—1 S**ymposium**, Washington, DC, USA, 10–14 December, pp. 309–317, 1974.

SANTANA, L. S.; FERRAZ, G. A. S.; TEODORO, A. J. D. S.; SANTANA, M. S.; ROSSI, G.; PALCHETTI, E. Advances in Precision Coffee Growing Research: A Bibliometric Review. **Agronomy**, v.11, p.1557, 2021.

SANTINATO, F. **Inovações tecnológicas para cafeicultura de precisão**. Tese (Doutorado em Agronomia). Faculdade de Ciências Agrárias e Veterinárias – UNESP, Jaboticabal,SP.125p. 2016.

SARAIVA, M.; PROTAS, É.; SALGADO, M.; & SOUZA, C. Automatic mapping of center pivot irrigation systems from satellite images using deep learning. **Remote Sensing**, 12(3), 558, 2020.

SARTORI, S.; FAVA, J. F. M.; DOMINGUES, E. L.; RIBEIRO FILHO, A. C.; SHIRAISI, L. E. Mapping the Spatial Variability of Coffee Yield with Mechanical Harvester. In: Proceedings of the World Congress of Computers in Agriculture and Natural Resources. **Anais...** ASAE, Foz do Iguaçu - PR: 2002.

SILVA, F. M.; SOUZA, Z. M.; FIGUEIREDO, C. A. P.; VIEIRA, L. H. D. S.; OLIVEIRA, E. Variabilidade espacial de atributos químicos e produtividade da cultura do café em duas safras agricolas. **Ciência e Agrotecnologia**, v. 32, n. 1, p. 231–241, 2008.

SILVA, P.A. A.; ALVES, M.C.; SILVA, F.M.; FIGUEIREDO, V.C. Coffee yield estimation by Landsat-8 imagery considering shading effects of planting row's orientation in center pivot. **Remote Sensing Applications: Society and Environment,** V. 24, 2021.

VEGA, A.; CÓRDOBA, M.; CASTRO-FRANCO, M.; BALZARINI, M. Protocol for automating error removal from yield maps. **Precision Agriculture**. 2019, 19, p. 1 – 15, 2019.

WEI, M. C. F.; MALDANER, L. F.; OTTONI, P. M. N.; MOLIN, J. P. Carrot Yield Mapping: A Precision Agriculture Approach Based on Machine Learning. **Ai**. 1, 229 – 241, 2020.

6. FINAL CONSIDERATIONS

Mapping the spatial variability of an agricultural area is not always a simple task, in this sense research is developed seeking to improve ways to identify these variations in the field, through the development of new techniques, sensors, and tools. Among the most diverse layers that can be mapped, the spatial variability of yield can be considered extremely important information, as it presents the results that the producer obtained in his last harvest, and serves as information for decision-making in the management of following seasons. However, obtaining yield data in sufficient density to map these variations remains a challenge for modern agriculture, especially in perennial crops such as coffee, that presents the extra challenge of bienual behaviour.

Strategies using yield monitors embedded in harvesters can be a viable alternative in large-scale data collection with high data density. In this sense the results of the second chapter of this work showed that the data obtained by the volumetric yield monitor built into a coffee harvester have a high correlation with those obtained using load cells. This fact qualifies the yield data, allowing the identification of the fruit flow variation during harvest and the spatialization of this data to obtain reliable yield maps. In addition to this validation, yield maps were also obtained in a commercial area in the Cerrado Mineiro region during three consecutive seasons. With the yield maps, it was possible to identify productive variability with well-defined regions of high and low productive potential in an area of approximately 10.2 h that always received the same management by the producer. In addition to this finding, by analyzing the yield maps in the temporal scale, it was possible to classify regions that present alternating yields between the years evaluated. This result indicates that, in addition to knowing the spatial variability of yield, the biennial variance information must also be considered in site-specific management strategies. The regions that presented high yield variance should be conducted in a differentiated way, alternating according to the productive year (high and low yield) and not only considering their general yield variability.

These results showed the potential to know the productive variability, but these maps can only be obtained at the time of harvest, and if there is a failure in the monitoring system, this information is lost. In addition, several activities related to the management of the area could be previously planned in case it was possible to estimate yield with some assertiveness before the actual harvest. In this sense, the results of chapters three, four, and five showed the potential of remote sensing to predict and map the spatial variability of a commercial coffee crop, with different sensors and at three levels of data acquisition (proximal, aerial, and orbital).

The results of chapter 3 showed the potential to use data from active optical sensors (AOS) positioned on the side of the coffee plants, in this specific area in which the thesis was developed, the results suggest that the sensors should be positioned to collect data from the third average of the coffee plants, as well as the average of the AOS data obtained from both sides of the coffee plant, must be calculated for a better correlation of vegetation indices with coffee yield. It was also possible to observe that one year before harvest, the AOS data showed a positive correlation with yield, indicating the potential to obtain coffee yield predictions. In addition, the results also indicated that the AOS data were able to identify the alternation of productive behavior temporally.

Another approach to monitoring and identifying yield variability in coffee areas is using aerial images. The results presented in chapter 4 demonstrated the potential of images in the extraction of biophysical parameters of coffee trees, as well as the individualization of plants aiming at extracting height and volume data, allowing the observation of the space-time relationship of the variables studied with yield data during three consecutive seasons, but the results point to the need for additional variables to better understand the spatial variability, since the highest correlations obtained with yield data were observed only in odd-numbered years, indicating that plant height and volume alone do not are sufficient to temporally describe the spatial variability. In addition, obtaining aerial images with RPA is a laborious task and requires high computational power for image processing. In this sense, an alternative for monitoring coffee areas is to use high-resolution orbital images. The results of chapter 5 showed the potential of using orbital images with a high spatial resolution (3 m) to predict coffee yield with high assertiveness even one year before harvest, and also indicated that it is possible to estimate the yield of the next crop with images after the current harvest (July and August).

In general, the results found in this work reinforce the importance of knowing the productive spatial variability in coffee areas, as this type of information helps in the search for possible causes of this variability so that the culture can be managed considering these spatial and temporal differences. This study presents the space-time variations of data from orbital, aerial, and terrestrial sensors in a commercial coffee area, as well as their relationship with yield maps generated with high data density. However, more studies are needed to deepen the practical applications of these data, to support the optimized management of inputs through precision agriculture practices in coffee plantations, as well as to evaluate the feasibility of these tools in other production systems (coffee mountain, rainfed coffee, non-mechanized areas, etc). Future research aimed at estimating coffee yield should also consider the fruits that fall to the ground and those that remain on the plant after harvest. In addition, it is necessary to develop faster and more assertive techniques to identify the degree of maturation of the fruits, issues that were not addressed in detail in this work and that are of great interest to coffee producers. Finally, future studies that seek to use techniques similar to those presented in this thesis must continually improve and optimize data acquisition and processing systems for their full applicability to farmers.