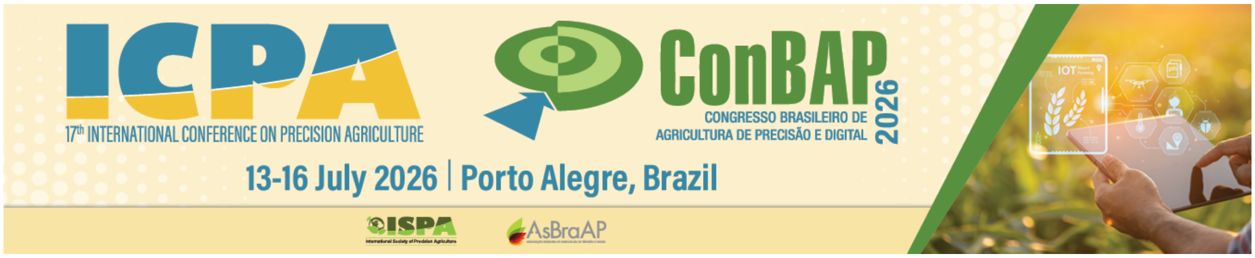


Co-registration of RGB UAV orthomosaics through a semi-automated affine method based on Ground Control Points and phase correlation validation

Bianca Batista Barreto[1]*; Bruno Ricardo Silva Costa[1]; José Paulo Molin[1]

[1] Laboratory of Precision Agriculture, "Luiz de Queiroz" College of Agriculture – University of São Paulo, SP, Brazil

A paper from the Proceedings of the
17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture
13-15 July 2026
Porto Alegre, Brazil

Abstract.

UAV images are crucial for Precision Agriculture (PA) purposes which require the monitoring of spatial variability regarding plant growth, especially to assess variation at plant level over time in perennial crops, such as banana plantations. However, spatial misalignments between orthomosaics from different dates requires post-processing image matching, i.e., co-registration, to ensure reliable spatiotemporal variability analysis. This study proposes a method for the co-registration of RGB UAV images using visually identifiable ground control points (GCPs), followed by quantitative evaluation of its performance through local phase correlation. The method was developed using a time series of four Red Green Blue (RGB) orthomosaics generated from images captured by a DJI Mavic 3M UAV over a 0.33 ha banana plantation, cv. "Nanica" (Musa acuminata, Cavendish group). Image surveys were conducted to monitor the growth of suckers after harvesting fruit-bearing stems until the next harvest in 20 randomly placed target plants within the plot. Two surveys were performed on the same day (before and after harvest), and subsequent surveys were conducted monthly thereafter. Checkerboard targets were positioned over fixed GCPs around the plot before each flight, ensuring their visibility across all orthomosaics and serving as fixed reference features for post-processing alignment. Orthomosaics were processed using Agisoft Metashape Professional (v. 2.2.2). The generated images were loaded into QGIS, and GCP centroid coordinates were manually identified in each orthomosaic using the QGIS Georeferencer tool. The orthomosaic from the first survey was assumed as the reference, and its image space and map space coordinates were exported as .points format files. These files were then used to perform affine transformations through a GDAL-based workflow implemented in the R programming language, guaranteeing that all aligned images shared an identical spatial grid, resolution, extent, and coordinate reference system (CRS). The alignment performance was then quantified using phase correlation applied to local windows centered on GCPs. Two patch sizes (40 and 60 pixels) were tested to estimate residual translational shifts between reference and target images in both pixel and meter units. The proposed co-registration method significantly improved the alignment between reference and target images. Using a 40-pixel window, mean displacement errors were reduced from 0.056–0.384 m (raw) to 0.010–0.284 m (aligned), representing decreases between 26% and 82%. This method proves to be a practical way for enhancing the geometric consistency and evaluating the alignment accuracy of time-series UAV RGB imagery, becoming particularly valuable for monitoring individual plants throughout the *growing season of perennial crops.*

Keywords.
Photogrammetry, Remote Sensing, Coordinate Reference System.

Introduction

PA employs technologies for monitoring and analyzing spatial and temporal variability within agricultural fields, enabling more efficient and site-specific crop management practices. Among these technologies, remote sensing has become an important source of information for characterizing crop conditions and evaluating changes over time (Weiss et al., 2020). In recent years, UAVs have emerged as a valuable platform for agricultural monitoring due to their ability to acquire imagery with ultra-high spatial resolution and flexible temporal frequency. These capabilities have facilitated multi-temporal analyses at the individual plant level, allowing detailed assessments of crop growth, canopy development, and vegetation dynamics throughout the production cycle (Velusamy et al., 2022).

The use of UAV imagery has facilitated crop monitoring at the individual plant level, allowing the extraction of information related to plant vigor, canopy architecture, growth dynamics, and productivity. This capability is particularly valuable in perennial crops, where plants remain in the field for multiple years and exhibit substantial temporal variability throughout their growth cycle. In banana crops, for example, the coexistence of mother plants, daughter plants and followers creates heterogeneous canopy structures and asynchronous development stages, making accurate temporal monitoring essential for crop management and research applications (Aeberli et al., 2021).

Multi-temporal UAV surveys provide an effective means of tracking crop development overtime However, reliable spatiotemporal analysis requires that corresponding objects occupy the same spatial position across all image's acquisitions. Although photogrammetric processing from independent UAV surveys often contain residual geometric discrepancies. These misalignments may arise from differences in flight conditions, camera positioning, image matching procedures, ground control point distribution, and limitations associated with onboard Global Navigation Satellite System (GNSS) positioning. Even small spatial offsets can compromise temporal analyses, especially when changes are evaluated at the individual plant scale (Ji et al., 2026).

Image co-registration is therefore a fundamental preprocessing step in multi-temporal remote sensing studies. The purpose of co-registration is to geometrically align images acquired at different dates so that homologous features spatially coincident throughout the image series. Several registration approaches have been proposed in the literature, including feature-based matching techniques, intensity-based methods, and transformations derived from ground control points GCPs. Among these approaches, GCP-based methods remain attractive due to their robustness, ease of implementation, and suitability for high-resolution UAV imagery (Grassi et al., 2025; Kazmi et al., 2026).

Beyond image alignment, the assessment of registration quality is equally important. Registration accuracy is commonly evaluated through root mean square error (RMSE) calculated from control points. However, this metric characterizes local residual displacements within the image. As a result, additional methods capable of evaluating local alignment accuracy are necessary to ensure the geometric consistency required for precise spatiotemporal analyses.

Phase correlation has emerged as an effective technique for estimating translational shifts between images through frequency-domains analysis. When applied to local image patches, phase correlation enables the quantification of residual displacements at specific locations, providing a detailed assessment of co-registration performance. This capability is particularly relevant for applications involving UAV time-series imagery, where accurate spatial correspondence is essential for monitoring subtle changes in plant growth and canopy development (Ju et al., 2026).

Therefore, this study proposes a practical methodology for the co-registration of RGB UAV orthomosaics using visually identifiable ground control point and affine transformation. The geometric consistency of the aligned images is quantitatively evaluated through local phase correlation analysis applied to image patches centers on control points. The proposed approach is tested using a time series of orthomosaics acquired over banana plantation, aiming to improve spatial alignment and provides a robust framework for assessing co-registration accuracy in multi temporal UAV datasets.

Methodology

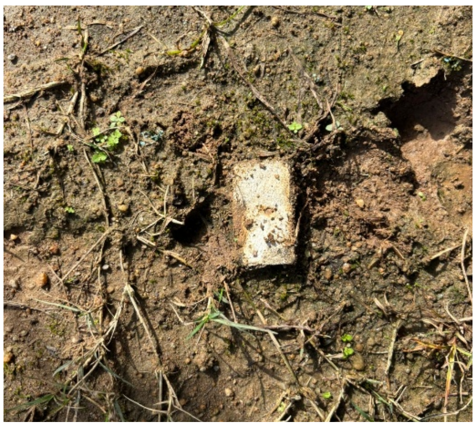
The study was conducted in an experimental banana field (*Musa spp., Prata subgroup*) located at the Agency for Agribusiness Technology (APTA) in Pariquera-Açu, São Paulo State, Brazil, (Figure 1). Multitemporal aerial surveys were performed using DJI Mavic 3 Multispectral UAV equipped with RGB sensors. Flights were conducted at an altitude of 40 m above ground level with 70% forward overlap and 80% side overlap, resulting in approximately 81 images per survey. The acquired imagery was processed into four orthomosaics and used for subsequent co-registration and temporal analysis. The first orthomosaic was used as the geometric reference, while the remaining three were co-registered to it.

Fig 1. Experimental area

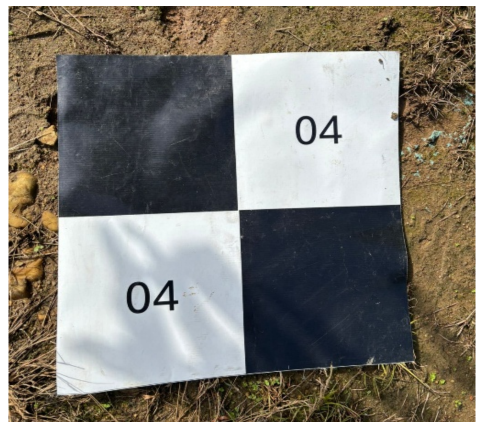
A co-registration procedure was development to improve the geometric consistency of a UAV-based RGB image time series acquired. Although all orthomosaics were individually georeferenced during the photogrammetric processing stage, residual geometric discrepancies between acquisition dates remained due to variations in image acquisition geometry, bundle adjustment solutions, and the distribution of ground control points. These discrepancies can compromise pixel-level temporal analyses and affect the monitoring of individual plants throughout the crop cycle. To establish a stable geometric reference framework, checkerboard

(Figure 2b) targets were positioned over permanently marked GCPs (Figure 2a) distributed around the experimental area before each UAV survey. The targets were designed to provide high visual contrast and facilitate their identification in the resulting orthomosaics. Because the targets remained fixed through the monitoring period, they served as invariant features for image alignment across dates.

All RGB orthomosaics were generated using Agisoft Metashape Professional (v. 2.2.2) following a standardized photogrammetric workflow. The orthomosaic corresponding to the first survey data was selected as the geometric reference. Subsequently, the centroids of all visible checkerboard targets were manually identified in each orthomosaic using the QGIS Georeferencer tool. For every target, image coordinates (pixel space) and geographic coordinates (map space) were recorded and exported.



(a)



(b)

Fig 2. In (a) Identification of ground control points on the field and (b) checkerboard

The co-registration workflow was implemented in the R environment using the terra package in combination with GDAL utilities. Initially, the reference coordinates extracted from the checkerboard targets were incorporated into a virtual raster structure through the insertion of GCPs. These GCPs established the correspondence between image coordinates in the target orthomosaic and their respective locations in the reference image. An affine transformation model was then applied to each target orthomosaic. The affine model was selected because it provides a robust solution for correcting systematic geometric distortions commonly observed in UAV-derived orthomosaics, including translation, rotation, scaling, and shear effects, while preserving linear geometric relationships. Transformation parameters were estimated from the GCP correspondences and applied using the GDAL warping procedure.

After geometric transformation, cubes convolution resampling was employed to interpolate pixel values and generate the final alignment orthomosaics. To ensure direct comparability among dates, all transformed images were forced to match the special properties of the reference orthomosaic, including coordinate reference system (EPSG:32723), spatial resolution, image extent, and pixel grid alignment. This procedure guaranteed pixel-to-pixel correspondence among all images in the time series, providing a consistent dataset for subsequent temporal analyses and individual plant monitoring.

The geometric accuracy of the co-registration process was evaluated using phase correlation analysis. This frequency-domain technique estimates translational displacement between two images by identifying the peak of the normalized cross-power spectrum and is widely used for measuring image registration accuracy.

For each checkerboard target, local image subsets centered on the target location were extracted both the reference and target orthomosaics. Two windows sizes (40x40 and 60x60 pixels) were evaluated to investigate the influence of the analysis area on displacement estimation. Phase correlation was applied independently to each image pair, yielding horizontal and vertical displacement components related to the reference orthomosaic. The estimates were converted from pixel units to metric using the orthomosaic ground sampling distance (GSD). Mean residual displacement values were then calculated before and after the co-registration procedure, allowing a quantitative assessment of alignment improvement. The percentage reduction in positional error was used as an indicator of co-registration performance across the image time series.

Results and Discussion

The proposed co-registration workflow successfully improved the geometric consistency of the UAV image time series. Visual inspection of the orthomosaics indicated a better correspondence between acquisition dates after the application of the affine transformation, particularly around the checkerboard targets used as reference features. Permanent structures and field boundaries also exhibited improved alignment, reducing the apparent positional shifts observed in the original orthomosaics. Figure 3 illustrates the effect of the co-registration procedure on image alignment. The reference orthomosaic (Figure 3a) was used as the geometric baseline for the alignment process. Before co-registration, noticeable positional discrepancies were observed between the reference and the target image (Figure 3b), resulting in misalignment of field features and reference targets. After applying the affine transformation, the aligned orthomosaic (Figure 3c) exhibited a substantially improved spatial correspondence with the reference image, reducing translational offsets and enhancing the overlap of common features. These results demonstrate the effectiveness of the proposed workflow in minimizing residual geometric errors among acquisition dates.

The observed improvements demonstrate the geometric discrepancies remaining after conventional photogrammetric processing can be substantially reduced through image-to-image co-registration. Although all orthomosaics were generated using the same photogrammetric workflow and ground control network, residual positional differences were still present between acquisition dates. These discrepancies are particularly relevant in studies involving individual plant monitoring, where small geometric offset may lead to inaccurate temporal comparisons.

The phase correlation results (Table 1) also demonstrated the suitability of using local image windows centered on stable reference targets to evaluate the images alignment quality. The 40x40 pixel window provided the most consistent assessment of co-registration performance in comparison to the 60x60 pixel window (results not presented). In this case, before co-registration, mean displacement errors ranged from 0.056 to 0.384 m in the evaluated image pairs. After image alignment, these values were reduced to between 0.010 and 0.284 m. Therefore, the magnitude of improvement varied according to the image pair, with positional error reductions ranging from 25.9% to 81.85%. The largest improvement was observed for Image 1, where the mean displacement error decreased from 0.056 m to 0.010 m, while Images 2 and 3 showed reductions from 0.309 m to 0.130 m and from 0.384 m to 0.284 m, respectively.

Table 1. Mean displacement errors before and after UAV orthomosaic co-registration estimated through phase-correlation analysis using a 40 × 40 pixel window.

Image pair	Raw displacement error (m)	Aligned displacement error (m)	Reduction (%)
Image 1	0.056	0.010	81.85
Image 2	0.309	0.130	57.86
Image 3	0.384	0.284	25.91

Overall, the proposed workflow produced a temporally consistent image dataset with improved pixel-to-pixel correspondence. This enhanced geometric consistency is expected to increase the reliability of subsequent analyses based on vegetation indices, canopy delineation, object-based image analysis, and plat-level monitoring throughout the crop cycle.

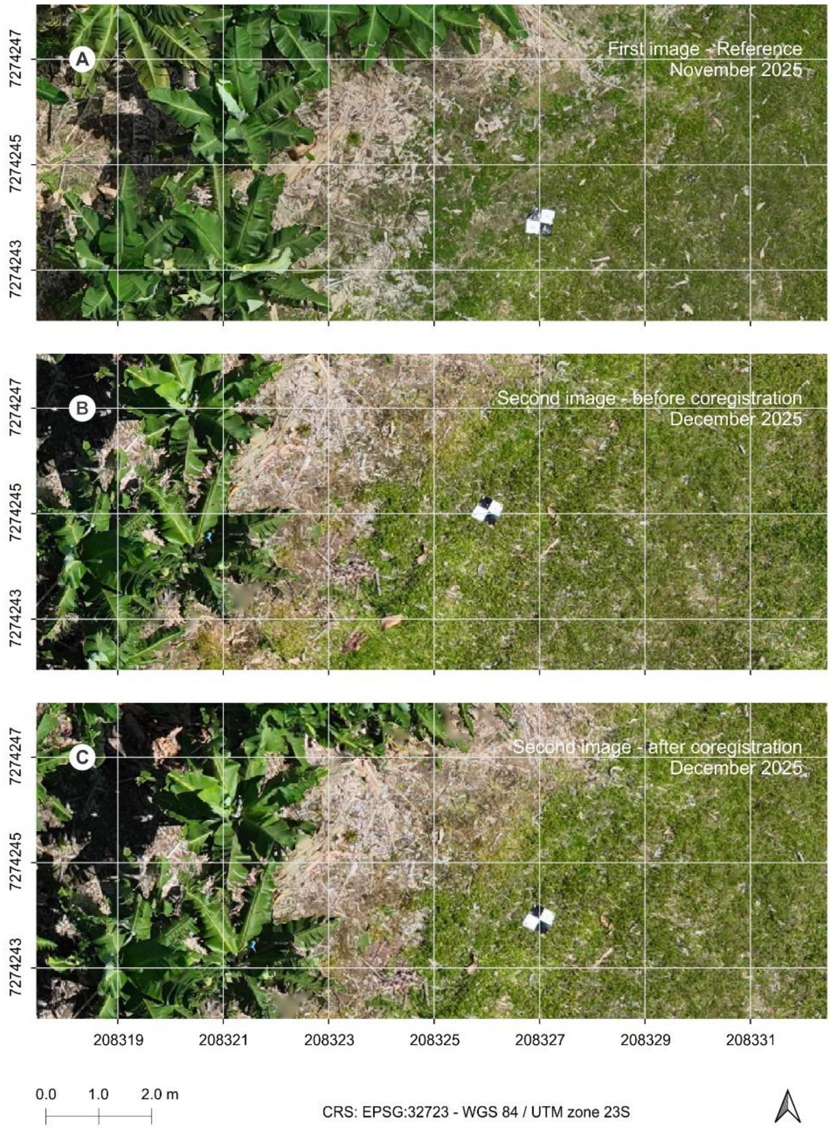


Fig 3. Local geometric alignment assessment at a checkerboard ground control point (GCP): (a) reference orthomosaic, (b) target orthomosaic before co-registration, and (c) target orthomosaic after co-registration.

Conclusion

The proposed co-registration workflow effectively improved the geometric consistency of the UAV-derived orthomosaic time series, reducing residual displacement errors by 26% to 82% compared to the original dataset. By combining checkerboard targets, affine transformation, and phase correlation analysis, the methods proved accurate image alignment and reliable pixel-to-pixel correspondence among acquisition dates. These results demonstrate the potential of the proposed approach to support temporal monitoring of individual plants and other applications that require high geometric consistency in UAV image time series.

Acknowledgments

The authors would like to thank Fapesp (Proc. 2022/09319-9; Proc. 2025/01992-4) for the funding. The authors thank to researcher area APTA – Pariquera-Açu for allowing image acquisition.

References

Aeberli, A.; Johansen, K.; Robson, A.; Lamb, D. W.; Phinn, S. Detection of banana plants using multi-temporal multispectral UAV imagery. Remote Sensing, v. 13, n. 11, p. 2123, 2021. DOI: https://doi.org/10.3390/rs13112123.

Grassi, M. Manzoni and S. Tebaldini, "Simultaneous Geometrical Autofocusing and Multi-Image Alignment for UAV-Based SAR Systems," 2025 IEEE Radar Conference (RadarConf25), Krakow, Poland, 2025, pp. 496-501, doi: 10.1109/RadarConf2559087.2025.11205027.

Ji, L.; Wang, X.; Hassan, H.; Deng, Z. A Multi-Sensor UAV Platform: Design, Testing, and Application for High-Throughput Plant Phenotyping. Drones 2026, 10, 372. https://doi.org/10.3390/drones10050372.

Kazmi, S.B.A., Basharat, M., Riaz, M.T., Hameed, N., Shafique, M., Ahmed, K.S. et al. (2026) Characterization and deformation monitoring of ancient Lohar Gali Landslide, Muzaffarabad, Pakistan. Earth Surface Processes and Landforms, 51(4), e70286. Available from: https://doi.org/10.1002/esp.70286

Lechun Zhang, Pengchao Chen, Ruitao Ma, Meiting He, Hang Zhu, Yubin Lan, A pixel-aligned co-registration and DSM-grid fusion framework for UAV multispectral and thermal imagery and point-cloud data: 3D Characterization of crop canopy water status, Computers and Electronics in Agriculture, Volume 241, 2026, 111290, ISSN 0168-1699, https://doi.org/10.1016/j.compag.2025.111290.

Velusamy, P.; Rajendran, S.; Mahendran, R.K.; Naseer, S.; Shafiq, M.; Choi, J.-G. Unmanned Aerial Vehicles (UAV) in Precision Agriculture: Applications and Challenges. Energies 2022, 15, 217. https://doi.org/10.3390/en15010217.

Weiss, M., Jacob, F., Duveiller, G. Remote sensing for agricultural applications: A meta-review, Remote Sensing of Environment, Volume 236, 2020, 111402, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2019.111402.