
# Applying Precision Agriculture principles to assess spatial variability of soil fertilization profiles in coffee farms at a regional scale

Helena Fantin Gebler[1] , Bruno Ricardo Silva Costa[1], José Paulo Molin[1].

Luiz de Queiroz College of Agriculture - University of São Paulo, Department of Biosystems Engineering, Precision Agriculture Laboratory - Piracicaba, 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.***

Precision and digital Agriculture have made significant progress in the last years, resulting in the extensive collection of georeferenced data on climate, soil, and plants that often are underused. These data provide an opportunity to identify agronomic variability across farms at a regional scale, supporting the delineation of management priorities and the adoption of site-specific practices, which are fundamental principles of Precision Agriculture (PA). In this context, using a regional-scale soil fertility database, this study aimed to assess whether soil fertility differs among coffee farms in the municipality of Caconde, São Paulo State, Brazil. Data from 156 farms in 2024 were analyzed, aggregating information at the farm level based on averages of pH, organic matter (OM), phosphorus, potassium, calcium, and magnesium. The farms were grouped using cluster analysis (k-means), considering their geographic coordinates, with k = 3 defined by the silhouette method. The non-parametric Kruskal-Wallis test was applied, followed by Dunn's test ($\alpha = 0.05$), in addition to the estimation of the effect size ($\varepsilon^2$). PERMANOVA was also performed for multivariate analysis. The results indicated significant differences between regions and the identification of spatial variability patterns on a regional scale proved to be capable of translating soil data into useful information and presents promising expectations regarding the analysis of databases comprising several farms in relevant agricultural regions, aligned with PA principles.

***Keywords:***
*Clustering analysis, regional analytics, Coffee*

## Introduction

In the last two decades, the development and increasing efficiency of Precision and Digital Agriculture technologies has been observed in a broad range of farming systems.(Council of the European Union, 2025). One of the consequences of this was the popularization of digital data collection, resulting in the current scenario where agricultural companies usually have large georeferenced databases concerning climate, soil and plant traits that often are underused (Cherubin et Al. 2022). These datasets are also produced by agricultural cooperatives such as the Cooxupé (Regional Cooperative of Coffee Growers in Guaxupé), which has accumulated thousands of georeferenced soil analyses over the years, representing individual small plots data and covering, for example, cooperative members. These datasets provide the opportunity to apply a fundamental Precision Agriculture (PA) principle, i.e., the identification of spatial variability patterns, (Brasil, 2016) at the level of agricultural regions which could not only enhance the value of databases, but also generate a decision-making and information management tool at the regional level. Thus, this research aims to evaluate whether there is evidence, at regional scale, of differences in soil fertility of coffee farms within the municipality of Caconde – SP, through soil analysis data.

## Material and methods

### Study site and dataset description

The farms from this study are located in the municipality of Caconde, São Paulo state, Brazil. Soil Fertility data from 156 coffee farms evaluated in 2024 were analyzed from Cooxupe's database (Fig.1). Plot-level data were aggregated at the farm level using the mean values of pH; organic matter (OM); Phosphorus (P); Potassium (K); Calcium (Ca) and Magnesium (Mg).

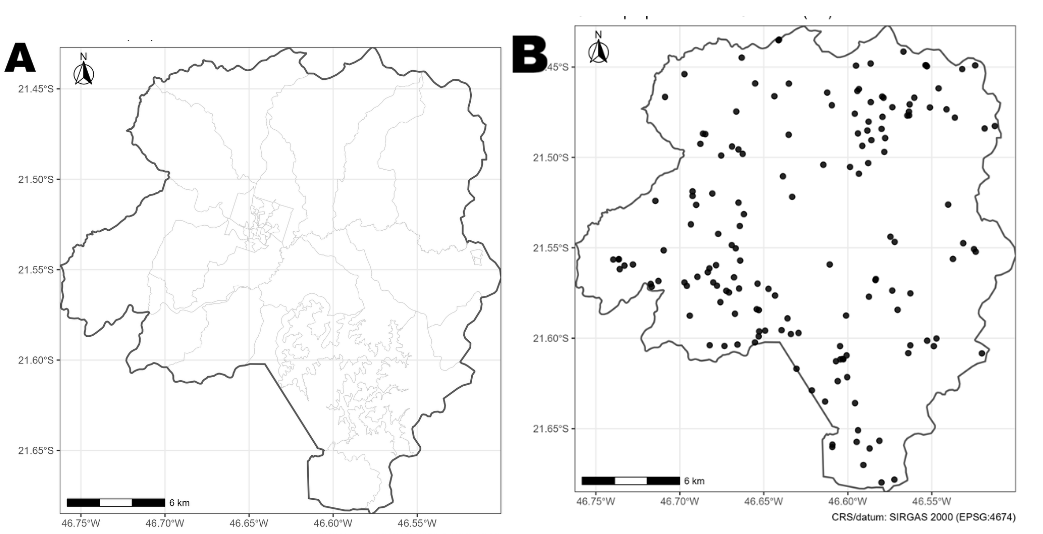


**Fig 1. A=Municipality of Caconde, São Paulo, Brazil. B= Coffee farms within Caconde from which the data were extracted.**

### Statistical tests

Subsequently, the farms were grouped assuming their central coordinates (longitude, latitude) though k-means clustering to define local regions within the municipality.The optimal number of clusters (i.e., regions) was determined using the silhouette method, selecting the solution with the highest average silhouette width. After clustering analysis, the non-parametric Kruskal-Wallis's test was performed to compare whether there are differences between the regions concerning each individual soil attribute, followed by the post-hoc Dunn test if differences were significant (assuming $\alpha = 0.05$ for both tests). The effect size for the possible differences was estimated via epsilon-square ($\varepsilon^2$) and categorized as "small" ($\varepsilon < 0.08$), "medium" ($0.08 \leq \varepsilon < 0.26$) or "large" ($\varepsilon \geq 0.26$). The non-parametric PERMANOVA test was also applied to look for differences between regions assuming the multivariate space of soil attributes. All the statistical analyses were calculated using the RStudio software at its version 4.3.1.

.

## Results and discussion

### Descriptive statistics and zoning analysis

According to the silhouette method (Fig 2.), the best number of clusters was defined as k = 3, leading to the creation of three zones throughout the territory. Figure 3 provides a comparative overview of soil fertility attributes across the three identified regions by clustering analysis. Through tables 1 and 2, the results indicated significant differences between the regions, with "small" to "medium" effect size, for Ca ($p < 0.001$; $\varepsilon^2 = 0.088$), P ($p = 0.001$; $\varepsilon^2 = 0.076$), Mg ($p = 0.003$; $\varepsilon^2 = 0.062$), K ($p = 0.005$; $\varepsilon^2 = 0.055$) e pH ($p = 0.007$; $\varepsilon^2 = 0.052$), except for OM ($p = 0.628$). The results from Table 3 show that regions 1 and 2 were similar to each other and different from region 3, which have lower mean ranks for all individual soil attributes. PERMANOVA confirmed the difference between the regions ($F = 5.48$; $R^2 = 0.067$; $p = 0.001$), explaining 6,7% of total variation ($R^2 = 0.067$). Thus, regions 1 and 2 present a similar soil fertility multivariate profile ($p = 0.726$), while region 3 was different from region 1 ($R^2 = 0.093$; $p = 0.003$) and 2 ($R^2 = 0.069$; $p = 0.003$). The differences between the three groups are visible in Figure 4.

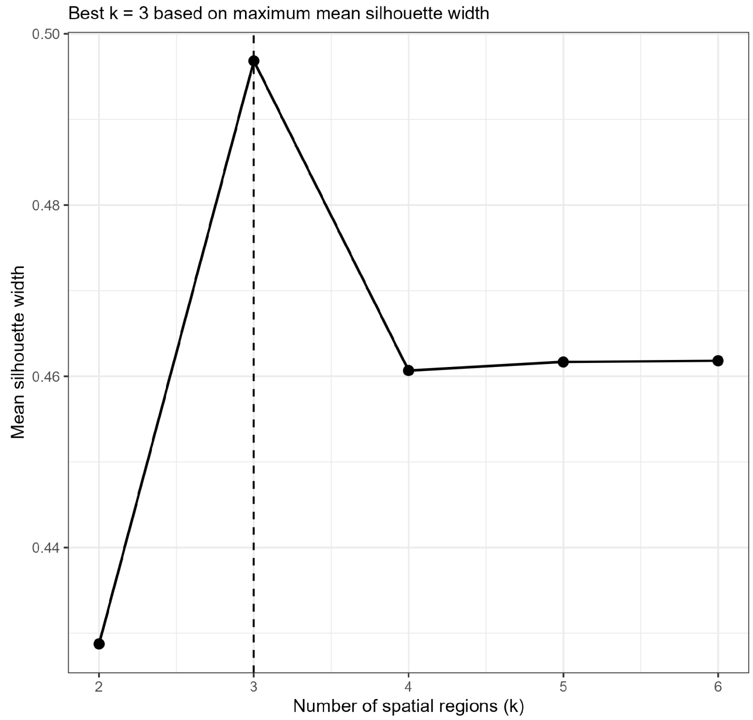


**Fig 2. Silhouette analysis to select the optimal number of groups (i.e., regions) for k-means clustering of coffee farms within the municipality of Caconde.**

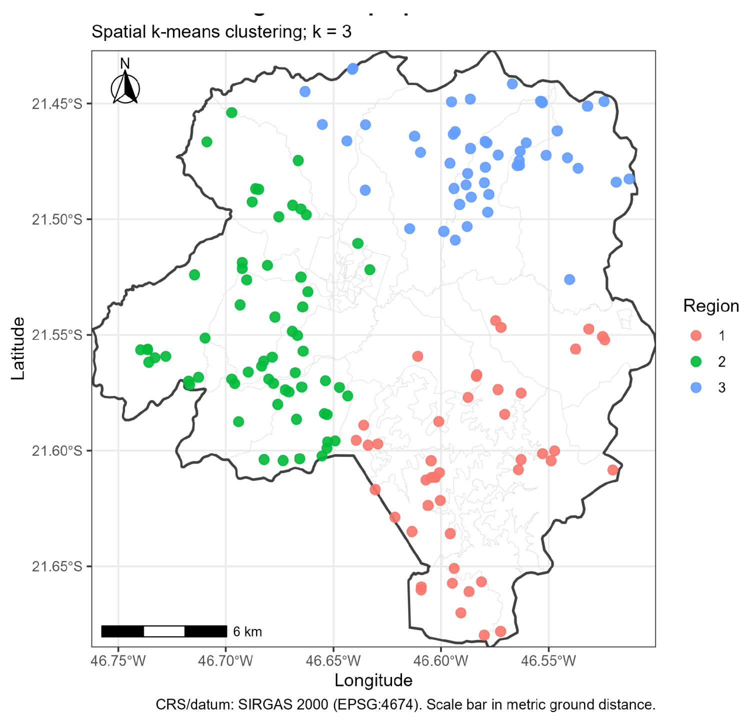


**Fig 3. Regional k-means clustering of coffee farms within the municipality of Caconde.**

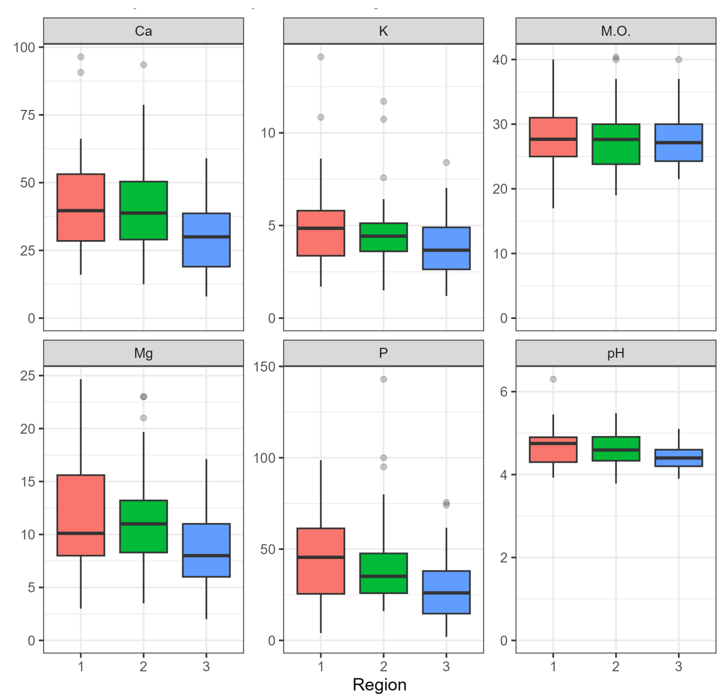


**Fig 4. Boxplots of soil fertility attributes across the three spatial regions identified by k-means clustering of coffee farms within the municipality of Caconde.**

**Table 1. Results of the Kruskal-Wallis's test concerning the comparison between regional clusters of coffee farms within the municipality of Caconde regarding soil fertility attributes.**

| Variable | H | df | ρ value | ε² | Magnitude |
|---|---|---|---|---|---|
| pH | 10.013886210149 | 2 | 0.00669132675698806 * | 0.0523783412428039 | Small |
| OM. | 0.931492543762606 | 2 | 0.627666520490518 | 0 | Trivial |
| P | 13.5531953515941 | 2 | 0.00114014745802363 * | 0.0755110807293735 | Small |
| K | 10.4649019752138 | 2 | 0.00534041994558455 * | 0.0553261567007435 | Small |
| Ca | 15.4744409390352 | 2 | 0.000436282551977642 * | 0.0880682414316027 | Medium |
| Mg | 11.4187513096441 | 2 | 0.00331474144181508 * | 0.0615604660761054 | Small |

H=Kruskal-Wallis statistic. Df= degrees of freedom. ρ value=statistical significance. ε²= Effect size.

**Table 2. Results of the Dunn's post-hoc test concerning pairwise comparisons between the regional clusters of coffee farms within the municipality of Caconde regarding soil fertility attributes.**

| Soil attribute | Group 1 | Group 2 | Z statistic | Effect size (r) | p-value | Adjusted p-value | Significance |
|---|---|---|---|---|---|---|---|
| pH | 1 | 2 | -0.4930 | 0.0395 | 0.6220 | 0.6220 | ns |
| | 1 | 3 | -2.8800 | 0.2310 | 0.0040 | 0.0118 | * |
| | 2 | 3 | -2.6100 | 0.2090 | 0.0091 | 0.0182 | * |
| P | 1 | 2 | -0.8270 | 0.0663 | 0.4080 | 0.4080 | ns |
| | 1 | 3 | -3.4500 | 0.2760 | 0.0006 | 0.0017 | ** |
| | 2 | 3 | -2.8800 | 0.2310 | 0.0040 | 0.0079 | ** |
| K | 1 | 2 | -0.5480 | 0.0438 | 0.5840 | 0.5840 | ns |
| | 1 | 3 | -2.9600 | 0.2370 | 0.0030 | 0.0091 | ** |
| | 2 | 3 | -2.6400 | 0.2110 | 0.0083 | 0.0165 | * |
| Ca | 1 | 2 | -0.1940 | 0.0155 | 0.8460 | 0.8460 | ns |
| | 1 | 3 | -3.3900 | 0.2710 | 0.0007 | 0.0016 | ** |
| | 2 | 3 | -3.4600 | 0.2770 | 0.0005 | 0.0016 | ** |
| Mg | 1 | 2 | 0.1510 | 0.0121 | 0.8800 | 0.8800 | ns |
| | 1 | 3 | -2.7400 | 0.2190 | 0.0062 | 0.0124 | * |
| | 2 | 3 | -3.1100 | 0.2490 | 0.0019 | 0.0056 | ** |

Z: Dunn test statistic; r: effect size estimated as r=Z/N; Adjusted p-value: p-value corrected for multiple comparisons using the Holm procedure; ns: not significant; * p < 0.05; ** p < 0.01.

**Table 3. Results of the pairwise PERMANOVA for multivariate soil fertility differences between the regional clusters of coffee farms within the municipality of Caconde**

| Group 1 | Group 2 | F | R2 | ρ value | ρ adjusted |
|---|---|---|---|---|---|
| 1 | 2 | 0.504732091537651 | 0.00478397586090961 | 0.72 | 0.72 |
| 1 | 3 | 9.41669685158884 | 0.0928515436207615 | 1 | 3 |
| 2 | 3 | 8.10657640325202 | 0.069223921083111 | 1 | 3 |

Pseudo-F: PERMANOVA test statistic; R²: proportion of multivariate variance explained by the regional grouping (effect size); Adjusted p-value: p-value corrected for multiple comparisons using the Holm procedure; ns: not significant; * p < 0.05; ** p < 0.01.

Based on the fertility indices of the farms in the database, it was possible to statistically prove the variability between farms and establish three regions of minimum internal variability within the municipal territory. Through the establishment of these variability regions, the application of PA technologies is justified, for example, the adaptation of management zones for use on a regional scale. Furthermore, the results demonstrate the importance of data collection and filtering, as they create a regional soil profile to be used not only by landowners and the cooperative, but also by other social strata to understand and define the historical and economic use of land in the municipality.

## Conclusion

Statistical analysis of soil fertility data revealed the variability in fertility within the municipal territory, allowing the creation of three regions and confirming the research hypothesis. Despite the limitations of the data, which did not present management information and required the approximation of geographic coordinates, the results of this research demonstrate the possibility of using data at the regional level for the application of PA technologies and reinforces the importance of collecting quality data. In order to further explore this line of research, it is suggested that the area of data exploration be increased in order to understand the spatial limits of this methodology.


### Acknowledgments

This study was financed, in part, by the São Paulo Research Foundation (FAPESP), Brazil. Process Numbers 2022/09319-9 and 2025/24469-5. The authors thank the Coffee Producers Cooperative in Guaxupé LTDA – Cooxupé for data availability.

## References

BRASIL (2016). Agricultura de Precisão. Ministério da Agricultura, Pecuária e Abastecimento. *Secretaria do Desenvolvimento Agropecuário e Cooperativismo.* Brasília, Mapa/ACS. Available at: https://www.gov.br/agricultura/pt-br/acesso-a-informacao/acoes-e-programas/cartas-de-servico/desenvolvimento-agropecuario-cooperativismo-e-associativismo-rural/agricultura-de-precisao. Accessed 01 June 2026.

Cherubin M. R., Damian J. M., Tavares, T. R., Trevisan, R. G., Colaço, A. F., Eitelwein, M. T., et al. (2022). Precision Agriculture in Brazil: The Trajectory of 25 Years of Scientific Research. *Agriculture*, 12(11), 1882, doi: https://doi.org/10.3390/agriculture12111882.

Council of the European Union (2025). From screens to fields: how digitalisation is transforming agriculture. *Brussels: Council of the European Union*. Available at: https://www.consilium.europa.eu/media/31jkkbxi/2024_971-art-agriculture-02.pdf Accessed 01 June 2026.