

Sample size estimation of fruit maturation for Arabica's coffee

Gustavo Pucci Botega¹ , Juliana Costa de Rezende Abrahão^{2*} , Thiago Tavares Botelho¹ , Cesar Elias Botelho² , Guilherme Soares Salvador¹ , Flávia Maria Avelar Gonçalves¹ 

1. Universidade Federal de Lavras  – Departamento de Biologia – Lavras (MG), Brazil.

2. Empresa de Pesquisa Agropecuária de Minas Gerais  – Lavras (MG), Brazil.

Received: July 12, 2024 | **Accepted:** Dec. 10, 2024

Section Editor: Gabriel Constantino Blain 

***Corresponding author:** julianacosta@epamig.br

How to cite: Botega, G. P., Abrahão, J. C. R., Botelho, T. T., Botelho, C. E., Salvador, G. S. and Gonçalves, F. M. A. (2025). Sample size estimation of fruit maturation for Arabica's coffee. *Bragantia*, 84, e20240230. <https://doi.org/10.1590/1678-4499.20240230>

ABSTRACT: This study aimed at establishing the ideal sample size for evaluating the maturation cycle in *Coffea arabica*, and investigating the errors associated with different sample sizes, in addition to verifying the possibility of using the clustering method to separate genotypes according to the maturation stage. Two experiments were analyzed: one with F2:3 progenies using visual maturation assessment through fruit counting, and another with cultivars using image processing for maturation assessment. To determine the ideal sample size for this trait, we used the estimation of the errors associated with maturation, using the bootstrap technique. Subsequently, the *K-means* algorithm was tested as an alternative for clustering genotypes into maturation classes. The application of the bootstrap technique in order to estimate the error associated with maturation revealed that the adoption of a 450-mL sample size resulted in an associated error of approximately 5%, indicating that it is an adequate size for character assessment. The implementation of *K-means* as a clustering tool offers a promising perspective for Arabica coffee plant breeding programs. A more comprehensive analysis, which not only assesses the proportion of ripe fruits, but also considers the distribution of different maturation stages, provides a more accurate understanding of the maturation process. This allows a more precise identification of genotypes with the most suitable performance for different growing conditions, as well as enabling adjustments in harvest management and post-harvest processing, optimizing coffee quality.

Key words: *Coffea arabica*, *k-means*, plant breeding, bootstrap, maturation stage, ideal sample.

INTRODUCTION

In Arabica coffee plant breeding, it is essential to evaluate a large number of traits. Some of the most relevant include yield, resistance/tolerance to biotic and abiotic stresses, grain size, fruit maturation cycle, and uniformity. The maturation cycle is particularly important, as it is fundamental to conduct selections based on specific goals. The release of cultivars with different maturation cycles is important for producers, as it allows them to scale production and maximize efficiency and profitability (Waha et al. 2020).

In the reproductive phase, the coffee tree undergoes multiple flowering events, with one main flowering followed by several secondary ones. The number and timing of these flowerings are influenced by phenological factors, including climatic conditions and the genetic variability of each cultivar. This phenological pattern results in the simultaneous presence of fruits at various stages of maturation during the harvest period, including green, light green, ripe (which can be red or yellow), raisin, and dry stages (Pezzopane et al. 2003).

Knowledge of the behavior of genotypes concerning the phenological cycle, including aspects such as maturation uniformity and cycle length (classified as early, medium, or late), along with agronomic attributes, is essential to support breeding research. The maturation cycle of coffee is determined by the interval between flowering and fruit maturation and

can vary based on the time of year when flowering occurs, as well as from one year to another. Comparatively, cultivars are classified into early, medium, and late cycles: early-maturing cultivars, such as 'Bourbon Amarelo IAC J10,' 'Catucaí 785-15,' and 'Siriema AS1'; cultivars that mature between early and medium, such as 'Acaíá IAC 474-19,' 'Catucaíam 24137,' and 'MGS Paraíso 2'; medium-maturing cultivars, such as 'Catucaí Amarelo 2SL,' 'IAC 125 RN,' and 'Mundo Novo IAC 376-4'; cultivars maturing between medium and late, like 'Catuaí Amarelo IAC 62,' 'Topázio MG 1190,' and 'Catiguá MG2'; and late-maturing cultivars, including 'Acauã,' 'Arara,' and 'TPR 100' (Carvalho et al. 2022).

A common approach to maturation phenotyping involves either subjective evaluations (Adunola et al. 2024) or fruit sampling, which quantifies the number of fruits at different maturation stages (Nogueira et al. 2005, Petek et al. 2006). However, the optimal sample size for such studies remains unknown and is often determined empirically. Sampling is crucial for studying any trait, as analyzing an entire population is typically impractical. To ensure that findings are applicable to the broader population, samples must be representative, enabling more accurate estimation of population parameters. While larger sample sizes enhance precision and the statistical power of tests, they also tend to increase data collection costs. Therefore, it is essential to strike a balance between sample size, the desired level of precision, and the statistical power required (Sun et al. 2019).

It is necessary to use precise approaches in order to evaluate maturation, thus allowing the selection of genotypes with the desired maturation cycle. The use of clustering can be an alternative, as it allows the genotypes to be grouped for different maturation cycles. This can be accomplished by statistical clustering techniques, such as *K-means*, a clustering algorithm that aims to separate a set of data into *k* clusters (*k* is a previously specified number), so that the clusters are similar to each other (Abbas 2008). Thus, genotypes can be selected according to the researcher's purpose.

With this in mind, the study aimed at establishing the ideal sample size for evaluating the maturation cycle in *Coffea arabica*, and investigating the errors associated with different sample sizes, in addition to verifying the possibility of using the clustering method to separate genotypes according to the maturation stage.

MATERIAL AND METHODS

Experiment description

Two experiments were analyzed in the present study. In the first experiment (E1), 10 progenies of the $F_{2.3}$ generation were evaluated from the first cycle of recurrent selection from the Coffee Plant Breeding Program (Universidade Federal de Lavras/Empresa de Pesquisa Agropecuária de Minas Gerais) (Rezende 2025a). The second experiment (E2) consisted of 21 cultivars developed by the *Coffea arabica* Plant Breeding Programs from Instituto Agronômico de Campinas, Empresa de Pesquisa Agropecuária de Minas Gerais/Universidade Federal de Viçosa/Universidade Federal de Lavras, Fundação Procafé (formerly IBC) and Instituto de Desenvolvimento Rural do Paraná (formerly Iapar) (Rezende 2025b).

E1 was set in a randomized complete block design, with three replications, and plots of 12 plants at 3.5×0.7 m spacing. In the beginning of June, a random sample of 1.8 L of fruit was collected from each plot. The harvest time was set in June to align with the typical maturation period of coffee fruits in the region where the study was conducted. Each sample was subdivided into six 300-mL subsamples to allow for a manageable and precise analysis. The counting of fruits in each sample was carried out by three different evaluators, who counted the number of unripe fruits (including green and light green), ripe and dry fruits (including raisin and dry fruits) for each sample (Fig. 1). To estimate the maturation percentage, it was considered the mean of the three evaluators.

E2 was carried out in a randomized complete block design, with three replications, and plots of eight plants at 3.6×0.7 m spacing. A random sample of 1 L of fruit was collected from each plot at the time of harvest, and this sample was separated into two subsamples of 500 mL each. The count was conducted using image processing, in which it was possible to calculate the percentage of unripe fruits (including green and light-green fruits), ripe and dry fruits (including raisin and dried fruits). To accurately count the fruits at each stage of maturation, an application specifically developed for this task was used. The application allows counting through sequential mouse clicks and has additional functions, such as the option to reset the count and enter large numbers. After counting how many fruits are in each image, the information is stored, and, at the end of the process, a spreadsheet is generated with all the information acquired.



Figure 1. Maturation stages of coffee cherries evaluated in the samples. (a) Yellow-colored cherries, encompassing the maturation categories: green, light green, yellow-ripe, and dry fruits (including raisin and dry fruits). (b) Red-colored cherries, representing the same maturation categories: green, light green, red-ripe, and dry fruits (including raisin and dry fruits).

The adoption of different evaluation methods (visual and image-based) reflects a strategy of complementarity in the breeding program. While visual evaluation is crucial for capturing details in the early stages, the use of digital images—an increasingly common tool in breeding programs—is a modern approach that enhances precision, reproducibility, and efficiency in the evaluation process, particularly in more advanced stages, as with commercial cultivars.

Image acquisition in E2

To determine the maturation stage, photos were taken using a phenotyping platform built as a wooden ‘box’ measuring 80 cm wide, 80 cm long, and 60 cm high, illuminated by four 18 W fluorescent lamps with a 6,500 K color temperature. The lamps were arranged in a square at the top of the ‘box’ (Fig. 2). An opening was made in the upper center to insert a camera. The fruits were uniformly arranged for the photos to be captured. The images were taken with a Canon EOS 60D camera using the following settings: 1/13 shutter speed, no flash, F10 aperture, 3:2 aspect ratio, 200 ISO, with a resolution of $5,184 \times 3,456$ pixels, and were saved in JPG format. The camera was connected to and controlled by a portable computer.

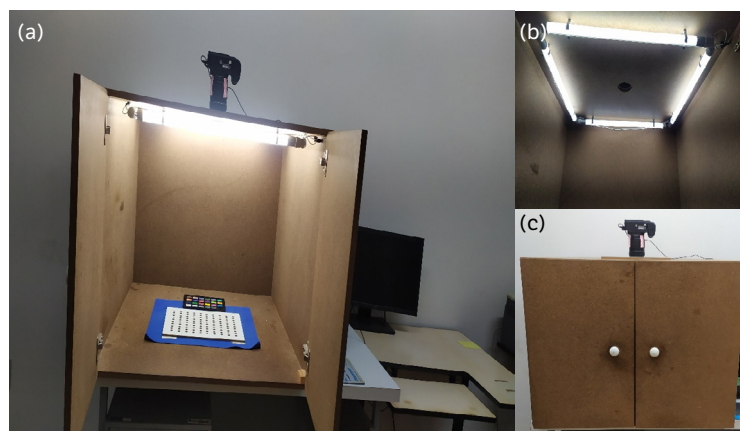


Figure 2. Phenotyping platform used for capturing images of coffee fruits at different maturation stages. (a) General view of the open box with fluorescent lamps arranged at the top. The background consists of light blue ethylene-vinyl acetate (EVA), with white EVA to highlight the fruits. (b) Detail of the 18 W, 6,500 K fluorescent lamps. (c) Closed box with the camera positioned at the central top opening for image capture.

The platform is an adaptation of Mendoza and Aguilera (2004). To highlight the fruits, white ethylene-vinyl acetate (EVA) was used in a rectangular shape, while light blue EVA was used as the background for the platform (Fig. 2a). Additional information such as treatment, repeat, sample number, sample size, and date of analysis were also included. Printed numbers were used to obtain this information, and a 10-cent coin was placed next to the samples to serve as a real reference.

Simulation via bootstrap

To determine the ideal sample size, the Bootstrap resampling method was employed. Bootstrap is a widely used statistical technique that approximates the distribution of a function of observations by resampling with replacement from the original dataset (non-parametric bootstrap). This method is particularly advantageous in cases in which the sample size is small or the underlying distribution is unknown, offering robust estimates of standard errors and confidence intervals for population parameters (Efron 1992). In this study, we opted for the non-parametric bootstrap, as the goal was to maintain the original sample's empirical distribution without making assumptions about the data's underlying distribution.

While the Bootstrap method is flexible and intuitive, it is important to recognize potential limitations, such as its reliance on the representativeness of the original sample. If the initial sample is not sufficiently large or diverse, the resampling process may introduce bias or overestimate the accuracy of the results. In this case, care was taken to ensure that the sample was representative of the population under study, minimizing potential issues related to bias. Overall, Bootstrap provides a reliable approach for this analysis, though it is important to interpret results in light of these considerations.

In the present study, 1,000 resamples were carried out, with replacement for each sample size (20 to 900 fruits). Subsequently, the error associated with the percentage of maturation was estimated with a 95% confidence interval. Each experimental plot was treated as an independent sample. Therefore, the distribution of fruits at different maturation stages was obtained for each of these resamplings.

Statistical analysis

Data was analyzed using the mixed model approach considering each maturation stage as an independent variable. Variance components were estimated from restricted maximum likelihood (REML) using the expectation-maximization (EM) algorithm, and prediction of breeding values via best linear unbiased predictor (E-BLUP). The model used was Eq. 1:

$$y = X\beta + Z\gamma + e \quad (1)$$

where: $y_{(n \times 1)}$: the vector of phenotypic observations; $\beta_{(b \times 1)}$: the vector for the fixed effect of repetitions; $\gamma_{(g \times 1)}$: the vector for the random effect of the genotypes; $e_{(g \times 1)}$: the error vector; $X_{(n \times b)}$: the block incidence matrix; $Z_{(n \times g)}$: the incidence of genotype effects' matrix.

After obtaining the E-BLUPs for each genotype and for each variable, the *K-means* algorithm was used to classify the genotypes according to their maturation status. For the present study, *k* was specified as 3, aiming to group the genotypes into early, medium, and late, based on information on the percentage of green, ripe, and dry. The algorithm works by randomly initializing *k* centroids, in which each centroid represents the center of a group. Then, the data points are assigned to the group whose centroid is the closest, based on a statistical metric, and in the present study Euclidean distance was used. The centroids are then recalculated with the mean of all points in the group, and the process is repeated until the groups do not change or the maximum number of iterations is reached (Abbas 2008).

Python software and the sklearn library were used to execute the K-means algorithm (Pedregosa et al. 2011). Additionally, matplotlib.pyplot was used to create the figures (Hunter 2007). Statistical analyses were carried out in the R software, using the sommer library (Covarrubias-Pazaran 2018).

RESULTS

The results obtained in both experiments demonstrated a remarkable convergence (Fig. 3). The associated error decreased as the sample size increased. A substantial reduction in error occurred until the evaluation of 300 fruits. By adopting a sample size of 300 fruits, it is estimated that the margin of error in estimating ripeness is approximately 5% (Fig. 3). For example, if we consider a plot with 50% of fruits in the ripe stage, we can infer that the real mean would be, on average, between 45 and 55%.

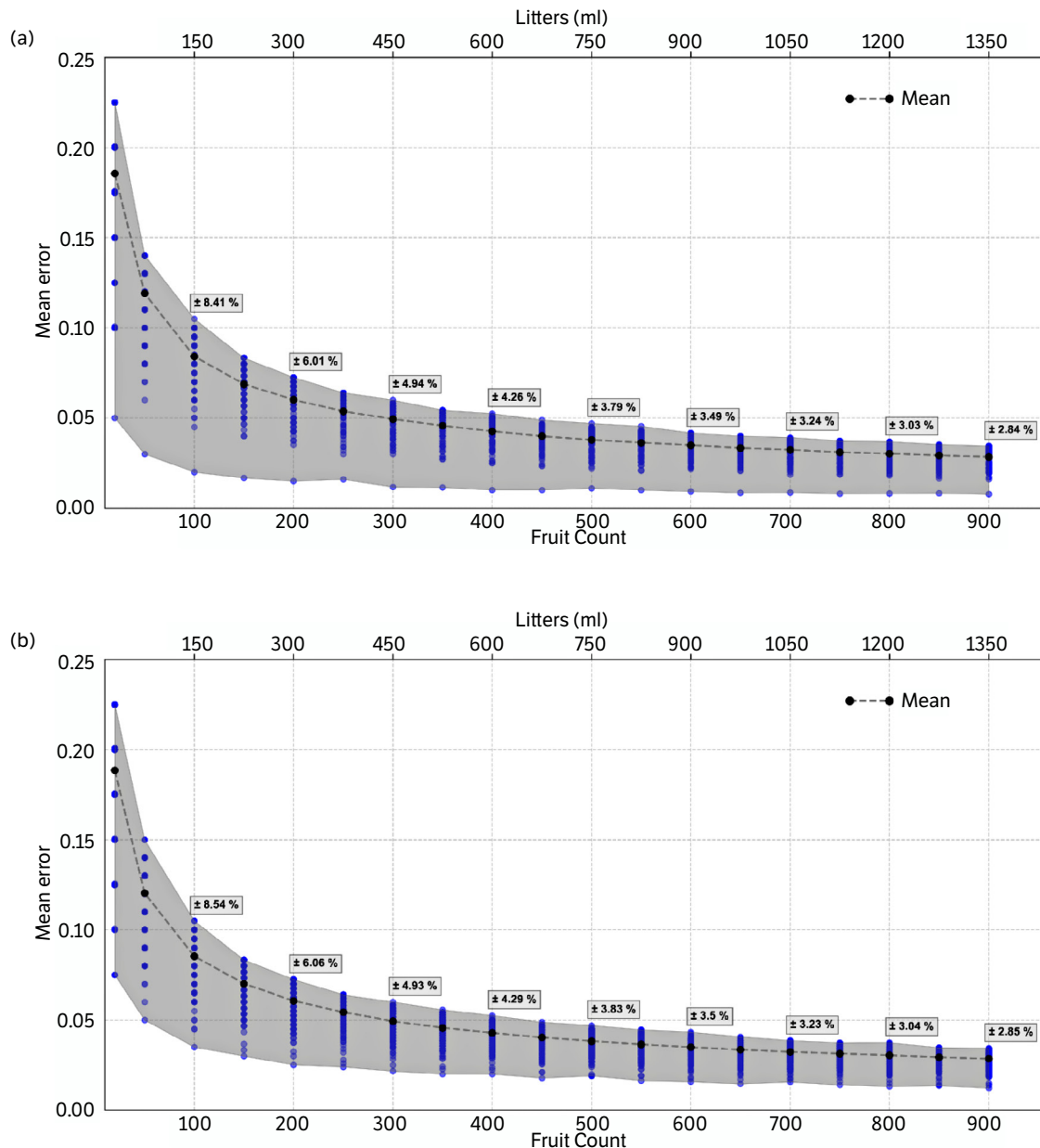


Figure 3. Error estimate associated with estimating the mean at 95% probability for the maturation stages: (a) experiment 1, and (b) experiment 2. Each point represents a sample. The values in the boxes represent the mean error associated with the corresponding sample size.

In E1, the mean for the maturation stage ranged from 7.5 (progeny 40) to 37.3% (progeny 10) for green ones, 28.6 (progeny 41) to 41.5% (progeny 8) for ripe ones, and from 25.3 (progeny 10) to 62.8% (progeny 40) for dry ones (Fig. 4). When only considering the mean maturation stage of the progenies, without considering the clustering, progenies 10, 4, and 18 presented the majority of fruits in the mature stage (37.4, 38.5, and 40.1%, respectively). This would lead us to classify them as progenies of medium maturity. The other progenies (9, 24, 8, 42, 41, 23, and 40) presented a higher percentage of fruits in the dry stage, and, in this case, they could be classified as early ripening. However, when considering the *K-means* clustering, changes in this classification are observed, as the clustering considered the three maturation stages to classify the genotypes. Therefore, progenies were classified into late, medium, and early clusters. Progeny 10 was in the late maturation cluster, progenies 4, 8, 9, 18, and 24 were in the medium maturation cluster, and progenies 23, 40, 41, and 42 were in the early cluster (Fig. 4).

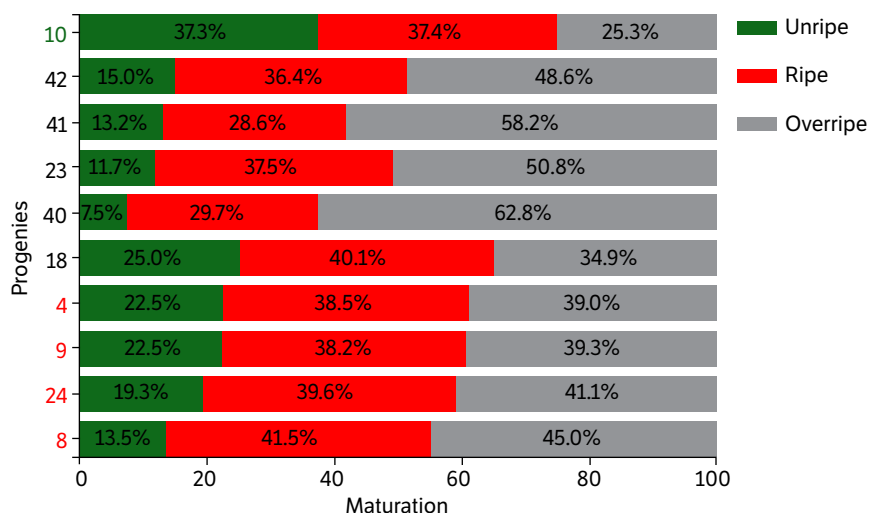


Figure 4. Assessment of progenies regarding maturation stages (Experiment 1). The colors of the progenies on the Y-axis represent the clusters performed by the K-means algorithm.

In E2, the mean for the maturation stage ranged from 17 (Guará) to 40.2% (IPR 103) for green ones, 35.6 (Oeiras) to 63% (Catiguá MG-2) for ripe ones, and from 17.8 (IPR 103) to 42.1% (Rubi MG 1192 and Oeiras) for dry ones (Fig. 5). When only considering the mean maturation stage of the progenies, without considering the clustering, we observed that cultivars Oeiras MG 6851 and Rubi MG 1192 were classified as early, as they present most fruits in the dry stage of maturation (40% of fruits). Cultivars IPR 100 and IPR 103 were considered late, as most fruits were in the green maturation stage (42 and 38.6% of fruits). All other cultivars had a higher percentage of fruit in the mature stage, considered medium ripening (Fig. 5). On the other hand, when considering the K means clustering, 10 cultivars were classified as late, six as medium, and four as early cultivars (Fig. 5). Therefore, considering the interpretation by the K means clustering, cultivars Acauã Novo, Pau Brasil MG, Catuaí Amarelo IAC 62, IPR102, MGS Travessia, Araponga MG1, Topázio MG 1190, and Catiguá MG1 are now classified as late. Cultivars Paraíso MG 419-1 and Catuaí Vermelho IAC 99 are now considered early (Fig. 4). For the other cultivars, the classification was the same in both methods.

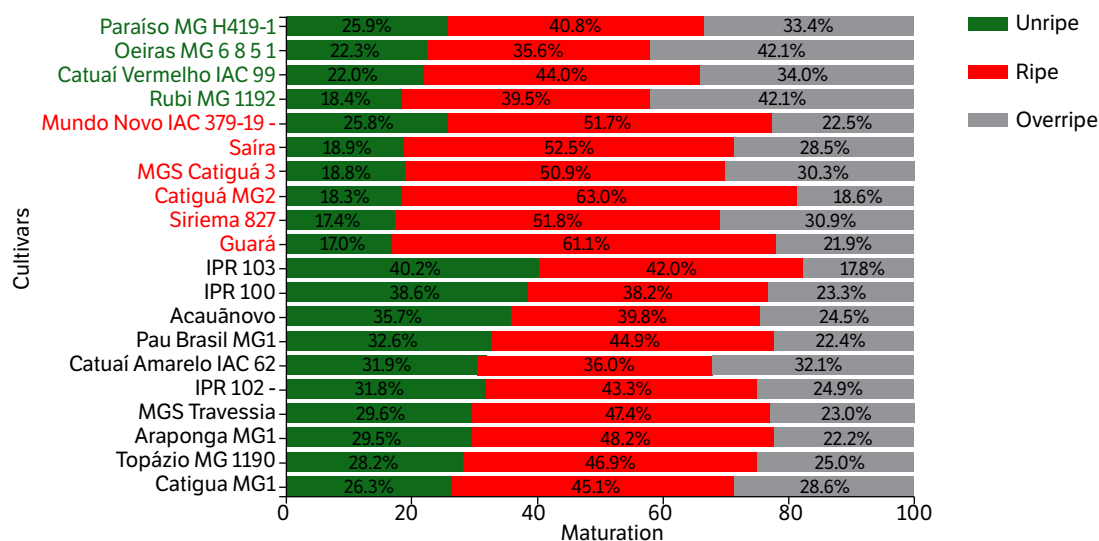


Figure 5. Assessment of cultivars regarding maturation stages (Experiment 2). The colors of the cultivars names on the Y-axis represent the clusters performed by the K-means algorithm.

DISCUSSION

Subjective determination of maturation degrees is commonly used in phenotyping for the maturation season, in which scores from 1 to 5 are assigned, according to the number of green, cherry, and dry fruits (Petek et al. 2006, Carvalho et al. 2022). In this type of assessment, in addition to the scores varying depending on the evaluation period of this characteristic, small differences between plants may not be detected, making the assessment subjective. Therefore, the best way to evaluate the character is to quantify the fruits in terms of their degree of maturation, using a representative sample. Studies that researched fruit maturation adopted different sample sizes, ranging from 300 (Costa et al. 2013) to 1,000 mL (Nogueira et al. 2005).

In the present study, we used the estimation of the error associated with maturation using the bootstrap technique to determine the ideal sample size for this trait. By applying this approach, we obtained a clear understanding of the error associated with each sample size, as all other estimated parameters depend on the phenotypic value of the plot. In both E1 and E2, the errors associated with the maturation estimates were quite similar across the different sample sizes. This result is particularly notable because, in E1, the plants in the progeny plots in the $F_{2,3}$ generation segregate, leading to possible variations in maturation between individual plants within the same plot. However, in this study, data was collected at the plot level, not at the level of individual plants. Therefore, despite the potential segregation within the plot, no significant differences in error were observed when compared with the cultivar experiment (E2), in which there is no segregation within the plots. This suggests that, at the plot level, segregation does not significantly affect the overall error estimates for maturation.

Using a 450-mL sample resulted in an approximate 5% error in the maturation estimate, indicating that this volume is suitable for assessing this trait. However, when opting for visual evaluation, which involves manually counting fruits in the green, cherry, and dry maturation stages, sample size may become a limiting factor due to the increased cost of phenotyping with larger samples. Moreover, increasing sample size can lead to a higher occurrence of random errors associated with evaluators (Raudys and Jain 1991). To reduce inter-evaluator variability and improve the reliability of estimates, it is essential to implement well-defined protocols and provide prior training (Beijbom et al. 2015).

Image analysis, on the other hand, offers significant advantages in terms of accuracy and efficiency. Automated capture of phenotypic information from fruits eliminates the subjectivity inherent in manual evaluation (Huang et al. 2020). Furthermore, the ability to store and revisit images enables more detailed analyses at later stages, allowing for the assessment of not only maturation but also other relevant traits, such as fruit size, shape, and color, which are important for genetic improvement and market demands. Automation reduces human error, enhances analytical precision, and allows for an increase in sample size without raising evaluation costs (Cooper et al. 2019), thereby providing greater confidence in genotype maturation estimates.

The adoption of both visual and image-based evaluation methods reflects a strategy of complementarity within the breeding program. Visual evaluation was prioritized in E1, particularly in the early stages of the breeding process, such as the evaluation of segregating populations like the $F_{2:3}$ progenies. At this stage, progenies exhibit a high degree of variability, and subtle characteristics—such as variations in fruit size, shape, or skin texture—can be more effectively captured through direct observation by trained evaluators. These details are crucial for selecting superior individuals to advance in the breeding cycle.

In contrast, image-based analysis was employed in E2, which involved more genetically uniform commercial cultivars. The use of digital images at this stage enhances precision, reproducibility, and efficiency, particularly when analyzing a large number of samples. While image analysis can capture high-resolution details and offers greater standardization, it may not always adequately identify nuanced traits that are critical in early breeding stages, such as slight morphological differences or irregularities. As the breeding process advances, and the material becomes more homogeneous, image-based methods prove superior due to their scalability and objectivity.

In addition to considering the balance between sample size and precision of the visual assessment, it is important to highlight that, when choosing the sample size, the intrinsic variability of the population of fruits under study must be also considered. More heterogeneous populations may require larger sample sizes to ensure a better representation of the variability present (Singh and Masuku 2014). By carefully selecting representative plots and stratifying the sampling process in our study, we ensured that the intrinsic variability within the population was adequately captured. Thus, the sample size,

though not large, was sufficient to provide reliable estimates of maturation characteristics, as it reflects the population's diversity in a precise manner (Singh and Masuku 2014).

The *K-means* algorithm was tested as an alternative for clustering genotypes into maturation classes and provide researchers with an efficient methodology for identifying coffee genotypes that align with the assumptions established in their research. By applying the *K-means* algorithm to form clusters, researchers can carefully select early, medium, or late genotypes, according to the specific purposes of their scientific investigations.

Figures 4 and 5 show the evaluation of different progenies and cultivars in relation to their maturation stages. There are two main ways of interpreting this data. In the first way, the individual means of each maturation stage can be analyzed separately. The second approach involves the use of *K-means* clustering, which classifies genotypes considering all three maturation stages together. Therefore, the interpretation of the genotypes' maturation stage may vary depending on which of these two methods is used, since *K-means* clustering provides an aggregated perspective that incorporates information from all maturation stages.

In the present study, when only the means of the cultivars' maturation stages are considered, we observed that 20% of the cultivars presented a classification different from those reported in the literature (Sera et al. 2017, Carvalho et al. 2022). This variation may be mainly due to the nature of the coffee, which has multiple flowerings, meaning that climatic conditions directly affect the classification. However, when considering *K-means* clustering, eight of the 21 cultivars did not fit the classifications reported in the literature. The change in classification occurred mainly among cultivars considered to be of medium maturity, which were grouped as late via *K-means*. It is important to highlight that the presented classification of the maturation cycle (early, medium, and late) reflects exclusively the specific conditions under which the experiments were conducted. This classification is not intended to suggest a new categorization for the tested cultivars, but rather to reflect the environmental influences observed in this particular study.

Three clusters were used to differentiate genotypes into maturation stages: early, medium, and late. However, the number of clusters to discriminate genotypes may be greater (Petek et al. 2006, Carvalho et al. 2022). In the research by Carvalho et al. (2022), the following classification was proposed: early (including 'Bourbon Amarelo IAC J10', 'Catucaí 785-15', 'Sirriema ASI'), medium early ('Acaia IAC 474-19', 'Catucaí am 24137', 'MGS Paraíso 2'), medium ('Catucaí Amarelo 2SL', 'IAC 125 RN', 'Mundo Novo IAC 376-4'), medium late ('Catucaí Amarelo IAC 62', 'Topázio MG 1190', 'Catiguá MG2'), and late ('Acauã', 'Arara', 'IPR 100').

The diversity of the genetic materials studied is highlighted in the supplementary tables (Rezende 2025a and b). Color characteristics were thoroughly evaluated through images on a developed phenotyping platform (Botega 2023)¹, based on the analysis of 36,876 individual images of coffee fruits at different stages of maturation. This method resulted in a more accurate classification model. Therefore, the use of a computational image approach combined with the use of *K-means* clustering can be a more accurate and standardized alternative for determining the maturation stages of coffee cultivars. Botega (2023)¹ compared the use of the phenotyping platform with the visual classification carried out by three different evaluators and observed differences in judgments for each stage of maturation, revealing a lack of consistency and agreement in visual evaluations carried out by different evaluators, which compromises accuracy assessment of maturation stages for different genotypes.

Although the percentage of ripe fruit is often used as a primary indicator to determine ripening time, this approach may have limitations when dealing with the complex maturation stages in plant breeding program populations (Melese and Kolech 2021). In these programs, it is common to deal with a wide diversity of genotypes, ranging from early to late cultivars, which can coexist in the same population and be harvested together. Therefore, it is important to employ more comprehensive and precise approaches to evaluate maturation, enabling the careful selection of genotypes that meet the desired requirements regarding maturation time. In the present study, the *K-means* algorithm considered not only the percentage of ripe fruits, but also the distribution of the different stages of maturation and provided a more accurate understanding of

¹ Botega, G. P. (2023). Visão computacional aplicada a análise de frutos de *C. arabica*. PhD Thesis (Doctorate in Genetics and Plant Breeding) – Universidade Federal de Lavras, Lavras.

maturation in the genotypes and cultivars studied. Furthermore, it opens up possibilities for efficient and targeted selection by considering particularities of each cultivar and growing conditions.

CONCLUSION

Samples greater than 450 mL of fruit presented an associated error of approximately 5% or less, a value considered acceptable for assessing the maturation character. Using the *K-means* technique to cluster data into different maturation cycles can be an excellent alternative for researchers, allowing for accurate and efficient analysis and decision-making. The different stages of maturation must be considered for a more accurate assessment of genotypes and more efficient selection.

CONFLICT OF INTEREST

The authors declare no competing interests.

AUTHORS' CONTRIBUTION

Conceptualization: Botega, G. P. and Gonçalves, F. M. A.; **Methodology:** Botelho, T. T. and Salvador, G. S.; **Formal analysis:** Botega, G. P.; **Writing – original draft:** Botega, G. P. and Abrahão, J. C. R.; **Writing – review & editing:** Gonçalves, F. M. A. and Botega, G. P.; **Supervision:** Gonçalves, F. M. A. and Botelho, C.E.; **Final approval:** Abrahão J.C. R.

DATA AVAILABILITY STATEMENT

The data are available in <https://doi.org/10.6084/m9.figshare.28152959> and <https://doi.org/10.6084/m9.figshare.28152962>.

FUNDING

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior 
Finance code 001

Conselho Nacional de Desenvolvimento Científico e Tecnológico 
Grant No.: 317001/2021-3

Consórcio de Pesquisa Café
Grant No.: Sincov 888689.2019

Instituto Nacional de Ciência e Tecnologia do Café
Grant No.: 465580/2014-9 (CNPq)

Fundação de Amparo à Pesquisa do Estado de Minas Gerais 
Grant No.: APQ 03605-17



ACKNOWLEDGMENTS

Not applicable.

REFERENCES

- Abbas, O. A. (2008). Comparisons between data clustering algorithms. *International Arab Journal of Information Technology*, 5, 320-325.
- Adunola, P., Tavares Flores, E., Riva-Souza, E. M., Ferrão, M. A. G., Senra, J. F. B., Comério, M., Espindula, M. C., Verdin Filho, A. C., Volpi, P. S., Fonseca, A. F. A., Ferrão, R. G., Munoz, P. R. and Ferrão, L. F. V. (2024). A comparison of genomic and phenomic selection methods for yield prediction in *Coffea canephora*. *Plant Phenome Journal*, 7, e20109. <https://doi.org/10.1002/ppj2.20109>
- Beijbom, O., Edmunds, P. J., Roelfsema, C., Smith, J., Kline, D. I., Neal, B. P. and Kriegman, D. (2015). Towards automated annotation of benthic survey images: Variability of human experts and operational modes of automation. *PloS One*, 10, e0130312. <https://doi.org/10.1371/journal.pone.0130312>
- Carvalho, C. H. S., Bartelega, L., Sera, G., Matiello, J., de Almeida, S. R., Santinato, F. and Hotz, A. (2022). Catálogo de cultivares de café arábica. Brasília: Embrapa Café. Documentos, 16.
- Cooper, L. A., Holderness Jr., D. K., Sorensen, T. L. and Wood, D. A. (2019). Robotic process automation in public accounting. *Accounting Horizons*, 33, 15-35. <https://doi.org/10.2308/acch-52466>
- Costa, J. C., Carvalho, C. H. S., Matiello, J. B., Almeida, S. R., Carvalho, S. P. and Baliza, D. P. (2013). Comportamento agrônômico de progênies e cultivares de cafeeiro com resistência específica à ferrugem. *Coffee Science*, 8, 183-191.
- Covarrubias-Pazaran, G. (2018). Software update: moving the R package sommer to multivariate mixed models for genome-assisted prediction. *BioRxiv*, 354639. <https://doi.org/10.1101/354639>
- Efron, B. (1992). Bootstrap methods: another look at the jackknife. In S. Kotz and N. L. Johnson (Eds.). *Breakthroughs in statistics* (p. 569-593). New York: Springer. https://doi.org/10.1007/978-1-4612-4380-9_41
- Huang, Y., Ren, Z., Li, D. and Liu, X. (2020). Phenotypic techniques and applications in fruit trees: a review. *Plant Methods*, 16, 107. <https://doi.org/10.1186/s13007-020-00649-7>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9, 90-95. <https://doi.org/10.1109/MCSE.2007.55>
- Melese, Y. Y. and Kolech, S. A. (2021). Coffee (*Coffea arabica* L.): Methods, objectives, and future strategies of breeding in Ethiopia. *Sustainability*, 13, 10814. <https://doi.org/10.3390/su131910814>
- Mendoza, F. and Aguilera, J. M. (2004). Application of image analysis for classification of ripening bananas. *Journal of Food Science*, 69, E471-E477. <https://doi.org/10.1111/j.1365-2621.2004.tb09932.x>
- Nogueira, Â. M., Carvalho, S. P., de Bartholo, G. F. and Mendes, A. N. G. (2005). Avaliação da maturação dos frutos de linhagens das cultivares Catuaí Amarelo e Catuaí Vermelho (*Coffea arabica* L.) plantadas individualmente e em combinações. *Ciência e Agrotecnologia*, 29, 18-26. <https://doi.org/10.1590/S1413-70542005000100002>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R. and Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Petek, M. R., Sera, T., Sera, G. H., Fonseca, I. C. de B. and Ito, D. S. (2006). Seleção de progênies de *Coffea arabica* com resistência simultânea à mancha aureolada e à ferrugem alaranjada. *Bragantia*, 65, 65-73. <https://doi.org/10.1590/S0006-87052006000100009>

- Pezzopane, J. R. M., Pedro Júnior, M. J., Thomaziello, R. A. and Camargo, M. B. P. (2003). Escala para avaliação de estádios fenológicos do cafeeiro arábica. *Bragantia*, 62, 499-505. <https://doi.org/10.1590/S0006-87052003000300015>
- Raudys, S. J. and Jain, A. K. (1991). Small sample size effects in statistical pattern recognition: Recommendations for practitioners. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13, 252-264. <https://doi.org/10.1109/34.75512>
- Rezende, J. (2025a). List of *Coffea arabica* progenies used in Experiment 1 with their respective parents. figshare. <https://doi.org/10.6084/m9.figshare.28152959.v1>
- Rezende, J. (2025b). List of *Coffea arabica* cultivars used in Experiment 2 with the respective provider institution and parents. figshare. <https://doi.org/10.6084/m9.figshare.28152962.v1>
- Sera, T., Sera, G. H., Fazuoli, L. C., Machado, A. C. Z., Ito, D. S., Shigueoka, L. H. and Silva, S. A. (2017). IPR 100 - Rustic dwarf Arabica coffee cultivar with resistance to nematodes *Meloidogyne paranaensis* and *M. incognita*. *Crop Breeding and Applied Biotechnology*, 17, 175-179. <https://doi.org/10.1590/1984-70332017v17n2c26>
- Singh, A. S. and Masuku, M. B. (2014). Sampling techniques & determination of sample size in applied statistics research: An overview. *International Journal of Economics, Commerce and Management*, 2, 1-22.
- Sun, P., Congalton, R. G. and Pan, Y. (2019). Using a simulation analysis to evaluate the impact of crop mapping error on crop area estimation from stratified sampling. *International Journal of Digital Earth*, 12, 1046-1066. <https://doi.org/10.1080/17538947.2018.1499827>
- Waha, K., Dietrich, J. P., Portmann, F. T., Siebert, S., Thornton, P. K., Bondeau, A. and Herrero, M. (2020). Multiple cropping systems of the world and the potential for increasing cropping intensity. *Global Environmental Change*, 64, 102131. <https://doi.org/10.1016/j.gloenvcha.2020.102131>