Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Simultaneous selection index via REML/BLUP to identify superior clones in conilon coffee in agroecological system through multi-trait analysis

Antônio Carlos Silva Júnior
antonio.silva.c.junior@gmail.com
Empresa de Pesquisa Agropecuária de Minas Gerais – Unidade Regional da Zona da Mata – Viçosa (MG)
Waldênia Melo Moura
Empresa de Pesquisa Agropecuária de Minas Gerais – Unidade Regional da Zona da Mata – Viçosa (MG)
Luciana Gomes Soares
Empresa de Pesquisa Agropecuária de Minas Gerais – Unidade Regional da Zona da Mata – Viçosa (MG)
Sant´Anna Andrade
Empresa de Pesquisa Agropecuária de Minas Gerais – Unidade Regional da Zona da Mata – Viçosa (MG)
Romário Gava Ferrão
Pesquisa e Extensão de Vitória/Multivix, Vitória (ES)
Abraão Carlos Verdin Filho
Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural/Incaper, Marilândia (ES)
Amélia Gava Ferrão
Paulo Roberto Cecon
Universidade Federal de Viçosa – Departamento de Estatística – Viçosa (MG)
Leonardo Lopes Bhering
Universidade Federal de Viçosa – Departamento de Biologia Geral – Viçosa (MG)
Cosme Damião Cruz
Universidade Federal de Viçosa – Departamento de Biologia Geral – Viçosa (MG)

Research Article

Keywords: Coffea canephora, factor analysis, multicollinearity, genetic gain

Posted Date: July 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10068973/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Simultaneous selection index via REML/BLUP to identify superior clones in conilon coffee in agroecological system through multi-trait analysis

Antônio Carlos da Silva Júnior[1]* https://orcid.org/0000-0002-6316-8711, Waldênia de Melo Moura[1] https://orcid.org/0000-0002-5272-6837, Luciana Gomes Soares[1] https://orcid.org/0000-0002-5628-6250, Hugo Sebastião Sant´Anna Andrade[1] https://orcid.org/0009-0007-1318-4899, Abraão Carlos Verdin Filho[2] https://orcid.org/0000-0001-8800-2382, Romário Gava Ferrão[3] https://orcid.org/0000-0002-2865-7878, Maria Amélia Gava Ferrão[4] https://orcid.org/0000-0002-8863-6956, Paulo Roberto Cecon[5] https://orcid.org/0000-0001-8213-0199, Leonardo Lopes Bhering[6] https://orcid.org/0000-0002-6072-0996, Cosme Damião Cruz[6] https://orcid.org/0000-0003-3513-3391

1. Empresa de Pesquisa Agropecuária de Minas Gerais – Unidade Regional da Zona da Mata – Viçosa (MG), Brasil.
2. Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural/Incaper, Marilândia (ES), Brasil
3. Faculdades de Ensino, Pesquisa e Extensão de Vitória/Multivix, Vitória (ES), Brasil
4. Empresa Brasileira de Pesquisa Agropecuária/Embrapa Café, Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural/Incaper, Vitória (ES), Brasil.
5. Universidade Federal de Viçosa – Departamento de Estatística – Viçosa (MG), Brasil.
6. Universidade Federal de Viçosa – Departamento de Biologia Geral – Viçosa (MG), Brasil.

*Corresponding author antonio.silva.c.junior@gmail.com

Abstract

Selection efficiency could be greater if based on multiple traits but identifying clones that combine high performance and stability in many traits has been a difficult task. The aim of this study was to estimate and compare the applicability of multi-trait indices for selecting conilon coffee clones in different harvests. Thirty-six conilon coffee clones, which make up four clonal varieties, were evaluated over five consecutive harvests. The traits assessed were the severity of rust, cercosporiosis and miner bug attack, the vegetative vigour of the plants, the fruit ripening cycle, fruit ripening uniformity and yield. The experimental design used was a randomized block design with three replications. Statistical analyses consist of decomposing the singular value of the best linear unbiased prediction matrix (BLUP). Multi-trait stability indices based on factor analysis and the Smith-Hazel index were used to quantify genetic gain and selection efficiency. The broad-sense heritability was estimated based on the average of the clones. The genetic gain was calculated considering a selection proportion of 20%. The MTSI, FAI-BLUP, and SH indices were compared in terms of genetic gains and correlations among seven traits of conilon coffee. The percentage of predicted genetic gains among the indices ranged from 20.18 to 49.03, with a selection intensity of 20%. The multi-trait indices based on the factorial clone-ideotype distance are more suitable than the classic method for conilon coffee breeding programs. Conilon coffee clones 03, 04, 10, 23, 28, 30 and 33 have the greatest genetic potential to make up future conilon coffee cultivars.

Keywords: *Coffea canephora*, factor analysis, multicollinearity, genetic gain

Introduction

Coffee is a major global commodity with significant health and socioeconomic impacts (Paixão et al., 2022). Brazil leads in coffee production, contributing 70% of global exports (Conab, 2024). The primary species produced are *Coffea arabica* (71%) and *Coffea canephora* (29%) (Conab, 2024). This success is partly due to breeding programs that develop high-yield, stress-resistant cultivars with good beverage quality (Barbosa et al., 2019; Silva Junior et al., 2022). Breeders must understand the relationships between important traits to improve selection accuracy (Silva Junior et al., 2022).

Evaluating multiple traits maximizes selection efficiency by exploring genetic parameters. However, selecting high-performance genotypes for multiple traits is challenging (Rocha et al., 2018). A biometric methodology is needed to evaluate multi-trait data from various experiments accurately. Smith's (1936) and Hazel's (1943) classical index is commonly used for simultaneous selection but has limitations, such as weighting traits of economic importance and issues with multicollinearity (Rocha et al., 2018; Olivoto et al., 2019). Some parameters can be assigned economic weights, while others are assigned randomly (Bhering et al., 2012). Indices based on multi-trait genotype-ideotype distance and factor analysis avoid multicollinearity and do not require economic weights (Rocha et al., 2018; Olivoto et al., 2019).

To overcome these problems, newer selection indices based on genotype–ideotype distances and factor analysis have been proposed (Rocha et al., 2018; Olivoto et al., 2019). In these approaches, factor analysis is used to transform the original set of correlated traits into a smaller number of uncorrelated factors, simplifying the data structure. Each genotype is then evaluated by calculating its distance from an ideal genotype (the "ideotype") in this reduced factor space. Since factor analysis produces orthogonal axes, these indices inherently avoid multicollinearity, and no explicit economic weights are needed (Olivoto et al., 2019).

Applying such indices in conilon coffee breeding should improve the accuracy of selection by focusing on the most important trait combinations without arbitrary weighting (Carias et al., 2016; Piza et al., 2023). In particular, by capturing complex trait relationships and integrating data across multiple harvests, factor-analysis-based indices can help identify clones with consistently high yield, quality, and resistance under diverse conditions (Piza et al., 2023). Therefore, the objective of this study was to evaluate and compare the applicability of different multi-trait selection indices for selecting superior conilon coffee clones across multiple harvest seasons.

Material and methods

Field experiments

The experiment was set up in the Experimental Field of the Agricultural Research Company of the State of Minas Gerais (Empresa de Pesquisa Agropecuária do Estado de Minas Gerais, EPAMIG), in the municipality of Leopoldina - Minas Gerais- MG- Brazil (21º 31' 55" S, 42º 38' 35" W, 212m). The experimental design was in randomized blocks
with three replicates and 36 Conilon coffee (*Coffea canephora*) clones evaluated over five consecutive harvest seasons. This structure resulted in a total of 540 observations (3 replicates × 36 clones × 5 harvests), for the assessment of clones performance and stability across multiple production harvests. The experimental plot consisted of nine plants, spaced 1.0 x 2.5 m, plants and rows, respectively. Clones were obtained from the Conilon coffee genetic breeding program of the Capixaba Institute for Research, Technical Assistance and Rural Extension, Incaper (Ferrão et al., 2019). While the experiment was being conducted, cultivation was carried out by the recommendations for growing conilon coffee (Prezotti et al., 2007).

Based on the results of the soil analysis, there was no need to correct the acidity of the soil, whose soil pH ranged from 6.60 to 7.27. We prioritized the use of organic fertilization (poultry litter) as nitrogen, phosphate and potassium source, when necessary were supplemented it with chemical fertilizers. Foliar fertilization was carried out with zinc sulphate, copper sulphate, boric acid, potassium chloride and adhesive spread for fertilizers. All the fertilizers were applied in three doses during the rainy season at intervals of between 30 to 60 days. Weeds were managed by mowing between the rows, the residues of which were used as much.

The traits were evaluated in the harvests five consecutive, using scales and scores adapted from Ferrão et al. (2021): Leaf rust severity - RS (*Hemileia vastatrix*); Brown eye spot severity - CS (*Cercospora coffeicola*) and Leaf miner attack severity - LMS (*Leucoptera coffeella*), where scores range from 1 to 5 for the following traits: where 1: no symptoms, 2: few symptoms, 3: moderate symptoms, 4: intense symptoms; and 5: intense symptoms with defoliation. Moreover, we have vegetative vigour of the plants (VIG), using increasing scores from 1 (low) to 10 (high); fruit ripening cycle (FRC), using scores from 1 to 3, where 1: early, 2: intermediate and 3: late; ripening uniformity (FRU), with scores of 1: uniform and 2: non-uniform; and productivity, in 60 kg bags of processed coffee per hectare (sc. ha^{-1}) (PROD).

Statistical analyses

Statistical analyses were carried out by singular value decomposition of the matrix of best linear unbiased prediction (BLUP) for the effects of the clone (C) by harvest (environment) interaction (CE), generated by linear mixed-effects models (LMM). Each trait was analyzed using LMM, where the effects of the CE interaction were assumed to be random and the effects of the harvest (environment) and the block within the environment were assumed to be fixed effects, as described below (Olivoto and Nardino, 2021):

$$y = Xb + Zu + e,$$

Where **y** is an $\sum_{j=1}^{r}(cr)_{n\ x\ 1}$ vector $n\ x\ 1$ of response trait, i.e. the response of the i[th] clone in the j[th] block (i= 1, 2, ..., c; j= 1, 2, ..., r); $\boldsymbol{y} = [y_{11}, y_{12} \ldots, y_{cr}]'$; $\boldsymbol{b}$ is an 1 x r vector of unknown and unobservable fixed effects of block $\mathbf{b} = [\gamma_1, \gamma_2 \ldots, \gamma_r]'$; **u** is an $(1\ x\ c)_m$ vector $m\ x$ 1x unknown and unobservable random effects of clone $\mathbf{u} = [\alpha_1, \alpha_2 \ldots, \alpha_c]'$(where $m = c$, the number of clones); **X** is an $n\ x\ r$ design matrix of 0 and 1 relating **y** to **b**; **Z** is na $n\ x\ m$ design matrix of 0 and 1 relating **y** to **u**; and $\boldsymbol{e}$ is an $n\ x$ 1 vector of rando errors $\boldsymbol{e} = [y_{11}, y_{12} \ldots, \gamma_{cr}]'$.

Vectors $\boldsymbol{b}$ and $\boldsymbol{u}$ were estimated using the mixed model equation (Handerson, 1975).

$$\begin{bmatrix}\hat{b}\\ \hat{u}\end{bmatrix} = \begin{bmatrix}X'R^{-1}X & X'R^{-1}Z\\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1}\end{bmatrix}\begin{bmatrix}X'R^{-1}y\\ Z'R^{-1}y\end{bmatrix}$$

The estimates of the variance components in $\boldsymbol{G}$ and $\boldsymbol{R}$ were obtained by restricted maximum likelihood using the expectation-maximazation algorithm (Dempster et al., 1977).

Matrix $\widehat{\boldsymbol{u}}_{\boldsymbol{ce}}$, containing the effects of the CE interaction, was estimated as follows:

$$\hat{u}_{ce} = U_{gp}\Lambda_{pp}V_{\varepsilon p}^{T},$$

Where $\boldsymbol{\Lambda_{pp}}$ is the diagonal of the matrix with singular values of p, in descending order [$p \leq (\min c - 1;\ \varepsilon - 1)$]; $\boldsymbol{U_{cp}}$ and $\boldsymbol{V_{\varepsilon p}}$ are the orthogonal matrices with singular vectors of the clone $\hat{u}\hat{u}^{T}$ and singular vectors of the harvest (environment) $\hat{u}^{T}\hat{u}$, respectively.

Selection based on selection index

Smith-Hazel index

The Smith (1936) and Hazel (1943) index (SH), given economic weights and phenotypic and genotypic variance-covariance matrices, is computed as follows*:*

$$u = U^{-1}Gw$$

Where **U** and **G** are phenotypic and genetic covariance matrices, respectively, u and w are vectors of index coefficients and economic weights, respectively.

The genetic worth I of an individual clones based on traits x, y, ..., n, is calculated as:

$$I = u_x B_x + u_y B_y + \cdots + u_n B_n$$

where u is the index coefficient for the traits x, y, ..., n, respectively, and B is the individual clones BLUP's for the traits x, y, ..., n, respectively.

We first examined the eigenvalues of the correlation matrix to detect highly to address multicollinearity traits. This initial screening was followed by calculating the Variance Inflation Factor (VIF) as described by Montgomery and Peck (1981). Commonly, a VIF value greater than 10 is considered indicative of serious multicollinearity, as supported by James et al. (2013) and Vittinghoff et al. (2014). Traits exceeding our VIF threshold were deemed problematic and removed from subsequent analyses. The traits highlighted in bold (Table 2) were excluded from the Smith (1936) and Hazel (1943) selection index, represented by SH1, due to multicollinearity issues. Following their removal, the selection gain was recalculated, and the percentage gain attributable to this adjustment was estimated based on the revised SH1 index.

Multi-trait stability index based on factor analysis

The average performance and stability were considered a multi-trait approach based on the Euclidean distance between the clone and the ideotype, using the scores obtained in an exploratory factor analysis, as described below:

$$Y = \mu + Lf + \epsilon,$$

Where $\boldsymbol{Y}$ is the observation vector $q\ x\ 1$; μ is the mean vector of $q\ x\ 1$(representing the BLUPs); $\boldsymbol{L}$ is the factor loadings matrix $q\ x\ f$; $\boldsymbol{f}$ is the vector of common factors $q\ x\ 1$; and ϵ is the vector of residuals $q\ x\ 1$, where q and f is the number of traits and common factors, respectively.

The eigenvalues and eigenvectors were obtained from the correlation matrix. Initial loadings were estimated considering only factors with eigenvalues > 1. The varimax rotation criterion was used for analytic rotation and for estimating the final loadings. The scores for the clones were obtained as described below:

$$F = Z(A^{T}R^{-1})^{T},$$

Where $\boldsymbol{F}$ is a dimension matrix $c\ x\ f$ with the factor scores, where $\boldsymbol{c}$ and $\boldsymbol{f}$ correspond to the number of clones and factors, respectively; $\boldsymbol{Z}$ is the c x p matrix with standardized mean; **A** is a matrix $p\ x\ f$ of canonical loadings and $\boldsymbol{R}$ is the correlation matrix between the traits $p\ x\ p$.

MTSI index

The multi-trait stability index (MTSI) was estimated as described below:

$$MTSI_i = \left[\sum_{j=1}^{f}\left(F_{ij} - F_j\right)^2\right]^{1/2}$$

Where $\boldsymbol{MTSI_i}$ is the multi-trait stability index for the i[th] clone; $\boldsymbol{F_{ij}}$ is the j[th] point for the i[th] clone; and $\boldsymbol{F_j}$ is the j[th] ideotype point.

FAI-BLUP index

The multi-trait index based on factor analysis and clone-ideotype distance (FAI-BLUP) (Rocha et al. 2018) was used to identify superior clones. The formula is as follows:

$$L_{ij} = \frac{\frac{1}{I_{ij}}}{\sum_{i=1;j=1}^{i=n;j=m}\frac{1}{I_{ij}}}$$

Where $\boldsymbol{L_{ij}}$: likelihood of the i[th] clones (i = 1, 2, ..., n) to be similar to the j[th] ideotype (j = 1, 2, ..., m); $\boldsymbol{I_{ij}}$: clones-ideotype distance from i[th] clone to j[th] ideotype, based on standardized mean distance.

The clone-ideotype distances were obtained using the scores from an exploratory factor analysis. This analysis reduces the dimensionality of the data and calculates the Euclidean distance between each clone and the ideotype. The correspondence to the ideotype adopted for the MSTI (Multi-Trait Stability Index) was based on these distances, ensuring that the selection process accurately reflects the similarity between clones and the ideal type.

The variance components generated from the analysis were used to estimate the broad-sense heritability (H^2) based on the average of the clones:

$$H^{2(i)} = \frac{\sigma_g^{2(i)}}{\left(\sigma_g^{2(i)} + \frac{\sigma_\varepsilon^{2(i)}}{r}\right)}$$

Where, r represents the number of replications, $\sigma_g^{2(i)}$ genetic variance component of each trait and $\sigma_\varepsilon^{2(i)}$ the residual variance component of each trait.

Genetic Gain

The genetic gain GG (%) was estimated for all indices and estimated for each trait, considering only BLUP-predicted, and not phenotypic values, by the expression:

$$GG(\%) = \frac{(X_s - X_0)}{X_0}\ x\ 100$$

Where X_s: mean of the selected clones is typically a scalar, representing the average value of the trait for the selected clones, X_0: overall population mean, is also a scalar, representing the average value of the trait for the entire population.

To evaluate the overall efficiency of the selection indexes, the total genetic gain (GGT%) was calculated by summing the predicted genetic gains for all traits evaluated in each index. This helps to provide individual results in a single performance estimate, reflecting the ability of each index to promote simultaneous gains in multiple traits.

Comparing the indexes

We applied the coincidence index as outlined by Hamblin and Zimmermann (1986) and created a Venn diagram to illustrate the relationships among the indexes. The BLUPs for each genotype and trait were obtained as described by Olivoto and Nardino (2021). BLUP was used to estimate the indices MTSI, FAI-BLUP and SH, which were then compared in terms of genetic gains, and to estimate the Pearson's correlation between seven conilon coffee traits, as described below:

1- MTSI (Multi-Trait Stability Index): BLUPs were used to estimate the genetic values for each trait, which were then incorporated into the MTSI calculation to evaluate the stability and performance of the clones.

2- FAI-BLUP (Factor Analysis and Ideotype Distance): BLUPs were applied to obtain the genetic values for each trait, which were then used in the factor analysis to reduce data dimensionality and calculate the Euclidean distance between each clone and the ideotype.
3- SH (Selection Hybrid Index): BLUPs were used to estimate the genetic values for each trait, which were then combined into the SH to evaluate the overall genetic worth of the clones.
4- SH1 (Adjusted Smith-Hazel Index): BLUP-derived genetic values were used to compute SH1. Traits with low contribution or moderate correlations were excluded to reduce redundancy and improve stability and performance of the clones.

All data analysis was carried out using the R statistical package metan (Olivoto and Lucio, 2020).

Results

Overall performance, likelihood ratio tests and variance components

According to the likelihood ratio test (Table 1), all traits except LMS exhibited a significant clone effect (p-value < 0.05), indicating the presence of genetic variability among clones. The clone × harvest (CE) interaction effect was significant only for PROD, FRC, and FRU, suggesting that these traits are influenced by temporal variation across harvests. For most traits—excluding FRC, FRU, and LMS —the CE variance component (σ^2_{ge}) exceeded both the genotypic (σ^2_g) and residual variances (σ^2_r), highlighting CE variance as the predominant contributor to phenotypic variation. In contrast, LMS did not show a significant genotypic effect (p = 0.062), nor did it exhibit a significant CE interaction, similar to VIG, RS, and CS.

Pearson's correlation between traits using BLUP information

Pearson's correlation by BLUP between the traits assessed in the conilon coffee clones in five consecutive harvests is shown in Figure 1. The correlation was positive (p-value <0.001) between the traits PROD and VIG; FRC and VIG; RS and LMS. On the other hand, there was a negative correlation (p-value <0.001) between FRC, CS, FRU, RS; FRC and FRU; and CS with PROD and VIG.

Multi-trait stability index based on factor analysis

Table 2 shows that the traits belonging to the factors had communality estimates higher than 0.65, thus indicating that more than 65% of the variability of these traits was explained by the three factors. In this Table, each assessed trait was grouped into the three factors (FA) as follows: FA1: PROD, VIG and FRU; FA2: RS, CS and LMS; and FA3: FRC. The traits vegetative vigour (VIG), grain productivity (PROD) and fruit ripening uniformity (FRU) had a positive with the first factor. In this case, selecting individuals based on the factor is expected to result in plants with greater vigor, high yields and high ripening uniformity. The analyses revealed a significant positive correlation between PROD and VIG (r = 0.29), whereas no significant association was found between PROD and FRU (r = –0.04). In contrast, VIG and FRU exhibited a significant negative correlation (r = –0.17) (Figure 1). On the other hand, Factor 2 (FA2) was composed of RS (0.8130), CS (0.6160), and LMS (0.8280), representing the severity of foliar diseases. RS showed a positive and significant correlation with CS (r = 0.09) and LMS (r = 0.17). Despite this, the high factor loadings confirm that FA2 captures the shared variance between these traits, even when direct correlations are weak. Thus, factor analysis revealed two main axes: FA1, related to productive performance and vigor, and FA2, to the severity of foliar diseases, evidencing latent dimensions that are not fully manifested in simple correlations.

This result suggests that this factor can be included as a pseudo-phenotype in *C. canephora* breeding programs. In this context, pseudo-phenotype combines multiple correlated phenotypic traits into a single, interpretable trait. This composite trait is not directly observed but is inferred through methods like factor analysis, which identifies latent variables that explain shared variation among traits.

Table 2 presents the results of the Variance Inflation Factor (VIF) analysis. This assessment revealed that the traits VIG, CS, and LMS exhibited high collinearity, as indicated by VIF values exceeding the threshold of 10. Consequently, these traits were excluded from subsequent analyses. Following their removal, all remaining traits displayed VIF values below 10, suggesting the absence of multicollinearity among the retained traits.

The estimates of the selection differentials by the indices FAI-BLUP, SH and MSTI broad-sense heritability for seven conilon coffee traits are shown in Table 3. The selection differential percentage was not positive for all the traits. The negative SD% estimates observed for FRU (FAI-BLUP: -1.82%; MTSI: -10.76%), PROD (MTSI: -2.57%), and RS (SH: -2.16%; MTSI: -4.22%) indicate that the selected clones had lower means than the population for these traits. The SH and FAI-BLUP indices were not efficient in improving fruit ripening uniformity, as revealed by the negative selection differential percentages, SD% of

-1.82 and -10.76, respectively. On the other hand, the MSTI presented an SD% estimate of 6.72. Regarding PROD, the MTSI and SH indices were able to select clones with higher productivity, as reported by the positive selection differential percentage. The MTSI index was efficient in selecting clones with lower coffee berry borer severity, as indicated by the positive selection differential percentage (SD%: 9.09). The MTSI and FAI-BLUP indices were not effective in improving rust resistance, as shown by the negative selection differential percentage. These indices showed effectiveness in selecting clones with lower cercosporiosis severity, as demonstrated by the positive selection differential percentage.

Table 4 shows a comparison - via genetic gain percentages (GG%) - between the indices studied. The lowest percentage of total genetic gain was for the SH index, which may be related to the problem of multicollinearity between the traits. On the other hand, the highest total percentages of genetic gain were observed for the MTSI and FAI-BLUP indices, respectively. These indices presented the highest total genetic gains (49.03% and 29.51%, respectively), but they were not the best for all traits. For example, the SH1 index showed the highest gain for PROD (8.34%) and FRC (16.16%). The "NA" values indicate that certain traits were removed from the SH1 index due to multicollinearity problems. This implies that multicollinearity was affecting the accuracy of the estimates for these traits, and removing them was an attempt to improve the robustness of the index.

This reinforces the better performance of these indices for selecting clones with desired traits. There was a discrepancy in genetic gain between the traits for the different selection indices (Table 4). It should be noted that the trait FRC had the highest percentage gains for the FAI-BLUP, SH and SH1 indices. The MTSI index had the highest percentage of genetic gain for the FRU, VIG, FRC and PROD traits, respectively.

The VIF was not used with the direct objective of verifying genetic gain, but as an approach to identify multicollinearity among the characteristics included in the selection indices. The presence of multicollinearity can compromise the accuracy of the index coefficient estimates, which, consequently, affects the reliability of the estimated genetic gains. For example, in Table 4, it can be observed that the SH index showed the lowest total genetic gain, which may be related to multicollinearity among the variables. The SH1 index, which was adjusted based on the VIF results—removing traits with a high degree of collinearity—showed superior gains for PROD and FRC. The "NA" values indicate that these traits were excluded precisely because they presented high multicollinearity, as identified by the VIF. Therefore, VIF analysis was employed as a preliminary step to the construction of the indices, with the aim of improving the statistical robustness and interpretation of genetic gains. The practical influence of multicollinearity on genetic gain is evident in the discrepancy between the SH and SH1 indices, demonstrating that the removal of collinear traits can result in reliable estimates and, in some cases, greater genetic gains for specific traits.

Multi-trait indices based on clone-ideotype distance

Figure 2 shows the selection of conilon coffee clones using the multi-trait index based on clone-ideotype distance (FAI-BLUP). In this Figure, the black dotted line indicates conilon coffee clones assessed in different harvests and the red dots are the selected clones, 29, 28, 36, 33, 35, 32 and 31, by the FAI-BLUP index.

The clones selected by the multi-trait index were 30, 3, 33, 28, 4, 10, and 23 (Figure 3). The Figure shows the relationship of each factor in the multi-trait stability index (MTSI). The lower the proportion explained by a factor (closer to the outer edge), the closer the traits within that factor are to the ideotype. The dashed line shows the theoretical estimate if all factors contributed equally. Clones 30, 23, 3, and 28 show strength related to FA1 (Figure 3). This factor corresponds to traits VIG, FRU, and PROD (Table 4), which show positive genetic gain within FA1. On the other hand, FA2 grouped the traits related to disease in conilon coffee. In this factor, clones 4 and 10 had the best performance. Clones 4 and 30 had the lowest estimates within FA3, which represents the FRC trait (Table 4).

Comparison of multi-trait indices

Figure 4 shows the estimated coincident selection between the clones obtained by the three selection indices. Regarding the MTSI and FAI-BLUP indices, there was a coincidence between clones 23 and 30. Clone 29 showed similarity between the FAI-BLUP and SH indices, while clone 4 showed similarity between the MTSI and SH indices. On the other hand, clones 3, 32, 28 and 8 were individually selected for the MTSI, FAI-BLUP and SH indices, respectively. Clones 10, 28 and 33 showed an intersection in all used indices. These similarities can be attributed to the selection criteria used by the methods, by prioritizing similar traits, and by the genetic robustness of the clones.

Discussion

The selection efficiency could be greater if based on multiple traits but identifying clones that combine high performance and stability in many traits has been a difficult task. This study aimed to estimate

and compare the applicability of multi-trait indices for selecting conilon coffee clones in different harvests. As a cross-pollinated species, conilon coffee has a high level of genetic variability (Akpertey et al., 2022), a fact observed for most traits among the clones evaluated (Table 1). Our results revealed the presence of considerable variation among the clones for all the assessed traits, which is in line with reports in the literature regarding robusta coffee (Akpertey et al., 2020; Akpertey et al., 2022).

The existence and exploitation of this variability provide an opportunity to select clones with the potential to make up future conilon coffee cultivars. Important highlight that part of this variability was attributed to the higher estimates of CE variance (σ^2_{ge}) in relation to genotypic variance (σ^2_g) and residual variance (σ^2_r). This demonstrates the importance of variance components, where the "best" clones measured in different environments cannot be the same. The significant CE interaction effect emphasizes the need to evaluate clones in multiple environments. This ensures that selected clones perform well under different conditions, increasing the robustness of the selection process.

Conilon coffee breeding programs seek to select and recommend clones that present suitable traits, such as resistance to major pests and diseases, uniform maturity, high vegetative vigour and yield (Ferrão et al., 2019). Thus, correlation studies can reduce the number of traits to be evaluated through indirect selection.

In the present study, the correlation revealed the importance of correlations of the traits in conilon coffee in five harvests (Figure 1), as observed between the traits of PROD and VIG (0.29). The trait PROD showed a positive correlation with FRC of 0.10. Although this correlation is weak, it suggests that an increase in productivity is associated with a slight increase in FRC. This result may indicate that practices that increase productivity can also modestly improve FRC. Conversely, the negative correlation of -0.15 between PROD and CS indicates that more productive plants tend to be slightly less susceptible to diseases. This finding is relevant for breeding programs, as selecting high productivity plants may also contribute to disease resistance, despite the weak correlation.

The trait VIG showed a stronger positive correlation with FRC of 0.32, indicating that more vigorous plants tend to have significantly better fruiting (Figure 1). This result is important for management practices aimed at increasing plant vigor, as it can lead to higher fruit production. Additionally, VIG presented negative correlations with several undesirable traits. The negative correlation of -0.10 with RS suggests that more vigorous plants are slightly less susceptible to fungal infections. Although this correlation is weak, it can be useful for guiding management practices that increase plant vigor.

The negative correlation of -0.17 between VIG and FRU indicates that more vigorous plants tend to produce fewer defective fruits (Figure 1). This finding is relevant for harvest quality, suggesting that increasing plant vigor can improve fruit quality. The negative correlation of -0.34 between VIG and CS is the strongest among the observed negative correlations. Like this, more vigorous plants are significantly less susceptible to diseases, which is extremely relevant for breeding programs and management practices aimed at increasing plant vigor.

Alkimim et al. (2021) found no significant correlation between the traits of vegetative vigour submitted to the different strategies of selection methods, productivity and rust incidence. These authors found a negative correlation between productivity per plant and rust incidence. However, these traits can be grouped into common factors due to the correlation between them. Moura et al. (2018) observed positive genotype correlations between vegetative vigour and productivity, and negative correlations with leaf rust severity in 36 conilon coffee clones. The authors reported that leaf rust and cercosporiosis severity were positively correlated. They argue that an increase or decrease in the incidence of one of these diseases contributes in equal measure to the incidence of other diseases. Silva et al. (2018) obtained similar results when evaluating 71 conilon coffee genotypes.

Correlations between traits can be the result of pleiotropy or genetic linkage (Thorpe et al., 2024). Thus, if undesirable correlations are caused by genetic linkage, these correlations can be broken by recombination caused by crossing or selfing; consequently, these factors are not necessarily major impediments to breeding programs (Thorpe et al., 2024).

The results demonstrate that the selection indices evaluated (FAI-BLUP, SH, and MTSI) exhibited distinct behaviors regarding their efficiency in selecting superior clones for various agronomic traits of conilon coffee. Although the selection differential percentage (SD%) is commonly used as an indicator of genetic gain, its interpretation must consider the nature of each trait and the specific goals of the breeding program. Yadav et al. (2025) argued that the direction and magnitude of selection differentials should be interpreted based on the biological aspects of each trait, especially in complex traits such as disease resistance and productivity stability.

The presence of negative SD% estimates for FRU, PROD, and RS (Table 3) indicates that the selected clones had lower means than the overall population for these traits. In the case of fruit ripening uniformity (FRU), the SH and FAI-BLUP indices were not efficient, whereas MTSI showed a positive gain (SD%: 6.72), demonstrating greater selection capacity for this trait. For productivity (PROD), the MTSI

index presented a negative SD% (–2.57%), suggesting that it prioritized other traits over yield. On the other hand, the SH index was effective in selecting more productive clones.

It is important to emphasize that for traits such as disease severity, negative SD% estimates are desirable, as they indicate lower incidence in the selected clones. For rust severity (RS), the SH and MTSI indices showed negative SD% values, which can be interpreted as favorable genetic gain. This interpretation is supported by Falconer and Mackay (1996), who define the selection differential as the difference between the mean of selected individuals and the population mean, with the sign of this estimative being interpreted according to the biological context of the trait. Therefore, the efficiency of selection indices is not necessarily associated with obtaining positive SD% estimates, but rather with their ability to direct genetic gain in the desired direction for each trait.

The results of the factor analysis were not similar to those of the correlation between the evaluated traits (Figure 1 and Table 2). The traits grouped in the first factor—PROD, VIG, and FRU (Table 2)—did not show similarity based on the simple correlation analysis (Figure 1). A significant positive correlation was observed between PROD and VIG, while no statistically significant association was identified between PROD and FRU (Figure 1). Regarding the second factor, related to the severity of foliar diseases, weak correlations were found between RS and CS. Golino and Epskamp (2017) reported divergences between simple correlations and complex factor structures. Christensen and Golino (2021) argue that factor structures may diverge from simple correlations. In general, high factor loadings are expected to be associated with significant correlation coefficients. The results of this study indicated commonalities above 0.70, suggesting that the traits are well represented by the factorial model (Fabrigar et al., 1999). However, this similarity was not observed in the present study with conilon coffee. This divergence can be attributed to the high genetic variability of the species, which hinders the clear manifestation of latent structures through simple linear correlations. This indicates that they can be used to estimate genetic gain among conilon coffee traits. In this case, selecting individuals based on the factor is expected to result in plants with greater vigor, high yields and high ripening uniformity. This result suggests that this factor can be included as a pseudo-phenotype in *C. canephora* breeding programs.

The pseudo-phenotype is a composite trait, derived from several individual traits into a single measurement, used to simplify the selection process. Ferreira et al. (2005) pointed out that selecting genotypes based on a set of important traits is necessary to achieve simultaneous gains in several traits of agronomic interest. On the other hand, we should consider the second factor, "disease factor", since all the traits showed a high correlation in this factor. This result is corroborated by Ferreira et al. (2005) and Barbosa et al. (2019), who assessed the possibility of using selection indices in factorial complexes, in traits of genotypes *C. canephora* and *C. arabica*, respectively. These authors did not group the grain yield trait into any factor.

Breeding programs have presented a wide range of information on these traits. Consequently, there is a growing need to extract more knowledge about these traits. To this end, selection indices for multi-traits are commonly used to take advantage of correlated traits and improve the estimate of genetic gain. In addition, it can help guide the recommendation of a particular cultivar with broad adaptability and phenotypic stability.

The estimates found for heritability (Table 3) can be considered moderate to high (Resende, 1997), thus indicating the possibility of successful selection based on the assessed traits. The highest estimates of H^2 were for the FRC, VIG and FRU traits, respectively. Productivity had a low estimate of H^2, which indicates that this trait is controlled by several genes, widely influenced by the environment and, therefore, selection is more complex (Sousa et al. 2019; Paixão et al., 2022). Alkimin et al. (2021) obtained heritability estimates in the broad sense of low magnitude ($H^2 < 0.15$) for the trait's vegetative vigour, rust and cercosporiosis incidence, and fruit ripening time.

Senra et al. (2025) evaluated 175 clones and four clonal controls in two harvests. These authors used the restricted maximum likelihood method and the best linear unbiased prediction (REML/BLUP). Moderate heritability (0.21 - 0.38) was obtained for the trait's vegetative vigor, rust severity, brown eye spot severity, and productivity. This result was similar to that found in this study for these traits. Carias et al. (2016) evaluated eight half-sib families from the recurrent selection program and obtained estimates of 0.48 and 0.22 for productivity and vegetative vigor, respectively. These results were superior to those obtained in this study, except for the traits vegetative vigor. Carvalho et al. (2019) evaluated conilon coffee hybrids regarding the traits of rust severity, vegetative vigor and productivity, obtaining individual estimates of heritability in the strict sense of moderate magnitudes.

In the selection process, plant breeders usually manipulate multiple traits (Santchurn et al., 2014), and consequently, there are problems with multicollinearity, which is an obstacle faced with the classic Smith-Hazel index (Rocha et al., 2018). Multicollinearity between traits is a systemic issue in multivariate analysis and is likely to cause serious difficulties to the proper interpretation of results, with the risk of erroneous conclusions, misdirected research and inefficient conservation measures (Prunier et al., 2015).

The results show that when traits with multicollinearity by the SH were removed (Table 4), there was a selection gain. This indicates that the complexity of handling multiple traits simultaneously is one of the limitations that breeders face when applying this classic Smith-Hazel index. Volpato et al. (2019) reported that the Smith-Hazel index may be affected by multicollinearity among traits, which can distort the index weights and reduce selection efficiency. Olivoto et al. (2022) observed in a wheat crop that the SH index with high collinearity between traits provided undesirable gains for nine of the 14 traits. These authors obtained greater genetic gains using the FAI-BLUP index. On the other hand, FAI-BLUP indices in elephant grass showed greater genetic gain compared to SH (Rocha et al., 2018). Rocha et al. (2018) argue that the multi-trait index does not need economic weight. These authors obtained genetic gains using the FAI-BLUP index that were superior to the classic Smith-Hazel index in elephant grass cultivation.

The results (Table 4) indicate that the highest performances of these clones were based on multiple traits, without assigning weights, free from multicollinearity, and balanced genetic gains in a desirable direction for all traits. Previous studies have shown that selection based on multiple traits can lead to more robust and sustainable genetic gains (Darkwa et al., 2025). Additionally, the absence of multicollinearity improves the accuracy of estimates and the stability of coefficients, resulting in more efficient selection (Olivoto, 2020). The approach of balanced genetic gains is supported by studies that highlight the importance of considering multiple traits to achieve comprehensive improvements in crop quality and productivity (Olivoto, 2020).

Comparative analysis between selection indices is essential to validate the effectiveness of the adopted selection strategy. Although the SH1 index demonstrated superior gains compared to SH, this isolated comparison is not sufficient to support the removal of variables due to multicollinearity. It is necessary to evaluate the performance of SH1 against the SH index to verify if the gains obtained are consistent and competitive with the available alternatives. This comprehensive approach allows us to confirm whether the reduction of multicollinearity truly contributed to mitigating distortions and obtaining balanced genetic gain. By demonstrating that SH1 maintains or surpasses the results compared to SH, the robustness of the methodology and the scientific validity of its adoption as a preferred selection tool are reinforced. The results in Table 4 show that the MTSI index presented the highest total genetic gain (49.03%), standing out mainly for vegetative vigor (11.64%) and uniformity of maturation (21.7%). The FAI-BLUP and SH varieties showed more modest performance, with total gains of 29.51% and 20.18%, respectively, but still relevant for specific characteristics such as fruit set (15.1% and 15.2%). The SH1 variety, in turn, achieved a total gain of 29.02%, higher than the SH variety, and stood out for productivity (8.34%) and fruit set (16.16%), confirming that the removal of collinear variables contributed to genetic gains. From an agronomic point of view, these results have direct implications for breeding programs and field management. Increased vegetative vigor is associated with more robust plants with a greater capacity to withstand biotic and abiotic stresses, reducing costs with pesticides and increasing the longevity of coffee plantations. Improved uniformity of maturation contributes to more homogeneous harvests, reducing losses and ensuring higher sensory quality of the beverage. The gains in productivity and fruiting indicate clones with greater yield potential per hectare, which directly impacts the producer's profitability. Furthermore, the reduction in the severity of diseases such as rust and cercospora leaf spot means less need for chemical applications, favoring sustainable management practices and decreasing environmental impact.

In the context of *C. canephora* breeding, this result can be useful, especially if the aim is to identify vigorous clones with high grain yields and uniform ripening. However, ripening uniformity is crucial for the sensory quality of the beverage. Clones with uneven ripening can cause problems in coffee quality, as harvesting fruits at different stages of ripening can result in an inconsistent beverage quality. Therefore, it is important to consider ripening uniformity when selecting clones to ensure the sensory quality of the coffee.

In this study, different selection indices were used to estimate genetic gain and predict clone rankings over the course of the harvests. This is particularly relevant given the unpredictable weather patterns related to crop years, especially in temperate countries like Brazil. However, we evaluate the response to the selection of clones in conilon coffee, using linear mixed models with multi-trait information.

The significant clone effect validates the genetic basis for selecting superior clones. It supports the use of genetic indices such as MTSI, FAI-BLUP, and SHI to identify clones with desirable traits. The similarity in selection of genotypes observed in Figure 4, can be attributed to the following factors: First, the selection criteria used by the methods may value similar traits, indicating that these genotypes have highly desirable attributes (Olivoto et al., 2022). Second, the absence of multicollinearity allows a more accurate identification of genotypes with consistent traits (Rocha et al., 2018). Third, the genetic robustness of genotypes may make them adaptable and efficient in various conditions, leading to their selection by different methods (Bhering et al., 2012). However, principal component analysis and multivariate selection models, may help explain why these genotypes are superior in multiple traits. Furthermore, the genetic stability of these genotypes can be investigated to better understand their consistency in selection.

The coincidence of clones 10, 28, and 33 across the different selection indices (Figure 4) suggests genetic proximity between these clones, combined with the robustness of the biometric methods used. According to Volpato et al. (2019), the Smith-Hazel index tends to favor genotypes with high heritability and low multicollinearity between traits, which avoids distortions in the assigned weights and improves selection efficiency. When the methods share similar statistical foundations, it is common for genetically robust genotypes, such as the clones mentioned, to be selected repeatedly. This consistency between the methods reinforces the reliability of the results and the stability of the genotypes across different selection approaches.

The multi-trait indices based on factorial clone-ideotype distance are more suitable than the classic method for conilon coffee breeding programs, with the MTSI standing out for its greater efficiency in the selection of conilon coffee clones associated with greater genetic gain. According to the MTSI, conilon coffee clones 03, 04, 10, 23, 28, 30 and 33 have the greatest genetic potential to make up future conilon coffee cultivars.

References

Akpertey A, Anim-Kwapong E, Adu-Gyamfi PKK, Ofori A. 2020. Genetic variation among biparental Robusta coffee families and implications for variety development. Experimental Agriculture 56: 171-182. https://doi.org/10.1017/S0014479719000255.

Akpertey A, Anim-Kwapong E, Adu-Gyamfi PKK, Ofori A. 2z022. Genetic variability for vigor and yield of robusta coffee (*Coffea canephora*) clones in Ghana. Heliyon 8: e10192. https://doi.org/10.1016/j.heliyon.2022.e10192.

Alkimim ER, Caixeta ET, Sousa TV, Gois IB, Lopes SF, Sakiyama NS, Zambolim L, Alves RS, Resende MDV. 2021. Designing the best breeding strategy for *Coffea canephora*: Genetic evaluation of pure and hybrid individuals aiming to select for productivity and disease resistance traits. PLoS ONE 16: e0260997. https://doi.org/10.1371/journal.pone.0260997.

Barbosa IP, Costa WG, Nascimento M, Cruz CD, Oliveira AC B. 2019. Recommendation of Coffea arabica genotypes by factor analysis. Euphytica 215:1–10. https://doi.org/10.1007/s10681-019-2499-x.

Bhering LL, Laviola BG, Salgado CC, Sanchez CFB, Rosado TT, Alves AA. 2012. Genetic gains in physic nut using selection indexes. Pesquisa Agropecuária Brasileira, 47: 402–408. https://doi.org/10.1590/S0100-204X2012000300012.

Bocianowski J, Nowosad K, Szulc P, Tratwal A, Bakinowska E, Piesik D. 2019. Genetic parameters and selection of maize cultivars using Bayesian inference in a multi-trait linear model, Acta Agriculturae Scandinavica, Section B—Soil & Plant Science, 69: 465–478. https://doi.org/10.1080/09064710.2019.1601764.

Carias CMOM, Ferrão MAG, Fonseca AFA, Ferrão RG, Volpi PS. 2016. Prediction of genetic gains by mixed models in conilon coffee progenies. Coffee Science, 11:39-45. https://www.coffeescience.ufla.br/index.php/Coffeescience/article/view/961.

Carvalho HF, Silva FF, Resende MDV, Bhering LL. 2019. Selection and genetic parameters for interpopulation hybrids between kouilou and robusta coffee. Bragantia 78:52–59. https://doi.org/10.1590/1678-4499.2018124.

Christensen AP, Golino H. 2021. On the equivalency of factor and network loadings. Behavior Research Methods, 53: 1563–1580. https://doi.org/10.3758/s13428-020-01500-6.

CONAB—Companhia Nacional de Abastecimento. 2024. Acompanhamento da safra brasileira. https://www.conab.gov.br/info-agro/safras/cafe. Accessed 12 July 2024. l (*in Portuguese*).

Darkwa K, Adjei EA, Chamba EB, Sayibu A, Amegbor IK, Agyapong FA, et al. 2025. Genetic parameter estimates and selection gain for multiple traits in white Guinea yam (*Dioscorea rotundata*) in Ghana. Euphytica 221. https://doi.org/10.1007/s10681-025-03467-x.

Dempster AP, Laird NM, Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Methodological) 39: 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.

Fabrigar LR, Wegener DT, Maccallum RC, Strahan EJ. 1999. Exploratory factor analysis: A guide to best practice. Psychological Methods 4: 272–299. https://doi.org/10.1037/1082-989X.4.3.272.

Falconer DS, Mackay TFC. 1996. Introduction to quantitative genetics. Edinburgh, SC: Addison Wesley Longman.

Ferrão MAG et al. 2019. Origin, geographical disperson, taxonomy and genetic diversity of *Coffea canephora*. In: Ferrão, R. G. et al. Conilon coffee. 3rd ed. Vitória, ES: Incaper, 85-109.

Ferrão MAG, Mendonça RF, Fonseca AFA, Ferrão RG, Senra JFB, Volpi PC, Filho ACV and Comério M. 2021. Characterization and genetic diversity of *Coffea canephora* accessions in a germplasm bank in Espírito Santo, Brazil. Crop Breeding and Applied Biotechnology 21:1-9. https://doi.org/10.1590/1984-70332021v21n2a32.

Ferreira A, Cecon PR, Cruz CD, Ferrão RG, Silva MF, Fonseca AFA, Ferrão MAG. 2005. Simultaneous selection of *Coffea canephora* by combining factor analysis and selection indices. Pesquisa Agropecuária Brasileira 40:1189–1195. (*in Portuguese*). https://doi.org/10.1590/S0100-204X2005001200005.

Golino HF, Epskamp S. 2017. Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. PLoS ONE 12: e0174035. https://doi.org/10.1371/journal.pone.0174035.

Hamblin J, Zimmermann MJDO. 1986. Breeding common bean for yield in mixtures. In: Janick, J. (ed.), Plant Breeding Reviews. John Wiley & Sons 245–272. https://doi.org/10.1002/9781118061015.ch8.

Hazel LN. 1943. The genetic basis for constructing selection indexes. Genetics 28: 476–490. https://doi.org/10.1093/genetics/28.6.476.

Henderson CR. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics, 31: 423–447. https://doi.org/10.2307/2529430.

James G, Witten D, Hastie T, Tibshirani R. 2013. An Introduction to Statistical Learning: With Applications in R. Springer. https://doi.org/10.1007/978-1-0716-1418-1.

Liersch A, Bocianowski J, Woś H, Szała L, Sosnowska K, Cegielska-Taras T, Nowosad K, Bartkowiak-Broda I. 2016. Assessment of genetic relationships in breeding lines and cultivars of Brassica napus and their implications for breeding winter oilseed rape. Crop Science. 56:1540–1549. https://doi.org/10.2135/cropsci2015.08.0530.

Mohring J, Williams E, Piepho HP. 2015. Inter-block information: to recover or not to recover it. Theoretical and Applied Genetics 128:1541-54. https://doi.org/10.1007/s00122-015-2530-0.

Montgomery DC, Peck EA.1981. Introduction to Linear Regression Analysis. New York: Wiley.

Moura WM, Pedrosa AW, Oliveira RL, Cecon PR, Ferrão RG, Ferrão MAG, Verdin Filho AC. 2022. Selection of Conilon coffee clones for the Zona da Mata region of Minas Gerais, Brazil. Coffee Science l:e171995. https://doi.org/10.25186/.v17i.1995.

Olivoto T, ADC Lucio, JAG da Silva, VS, Marchioro VQ de Souza, EJ. 2019. Mean performance and stability in multi-environment trials I: Combining features of AMMI and BLUP techniques Agronomy Journal 111:2949-2960. https://doi.org/10.2134/agronj2019.03.0220.

Olivoto T, and Lúcio AD. 2020. metan: an R package for multi-environment trial analysis. Methods in Ecology and Evolution 11:783-789. https://doi:10.1111/2041-210X.13384.

Olivoto T, Nardino M 2021. MGIDI: toward an effective multivariate selection in biological experiments, Bioinformatics 37: 1383–1389, https://doi.org/10.1093/bioinformatics/btaa981.

Paixão PTM, Nascimento ACC, Nascimento M, Azevedo CF, Oliveira GF, Silva FL, Caixeta ET. 2022. Factor analysis applied in genomic selection studies in the breeding of Coffea canephora. Euphytica 218: 42. https://doi.org/10.1007/s10681-022-02998-x.

Patterson HD, Williams ER. 1976. A new class of resolvable incomplete block designs. Biometrika 63:83-92. https://doi.org/10.1093/biomet/63.1.83.

Piza JMV, Silva FF, Rocha RB, Ferrão MAG, Fonseca A. 2023. FAI-BLUP index for multi-trait selection in conilon coffee breeding. Pesquisa Agropecuária Brasileira, 58: e03095. https://doi.org/10.1590/0100-204x-eb03095.

Prezotti LC, Gomes JÁ, Dadalto GG, Oliveira JA. 2007 Liming and fertilization recommendation manual for the State of Espírito Santo - 5th approximation. (*in Portuguese*). Vitória. ES: SEEA/Incaper/Cedagro. 305p.

Prunier JG, Colyn M, Legendre X, Nimon KF, Flamand MC. 2015 Multicollinearity in spatial genetics: separating the wheat from the chaff using commonality analyses. Molecular Ecology 24: 263–283. https://doi.org/10.1111/mec.13029.

Resende MDV. 1997. Advances in forest biometric genetics. In: Meeting on Genetics and Breeding Topics, Esalq, Piracicaba- SP (*in Portuguese*).

Rocha JRASCR, Machado JC, Carneiro PCS. 2018. Multitrait index based on factor analysis and ideotype-design: proposal and application on elephant grass breeding for bioenergy. GCB Bioenergy 10:52-60. https://doi.org/10.1111/gcbb.12443.

Santchurn D, Ramdoyal K, Badaloo MGH, Labuschagne M. 2014. From sugar industry to cane industry: evaluation and simultaneous selection of different types of high biomass canes. Biomass & Bioenergy, 61:82–92. https://doi.org/10.1016/j.biombioe.2013.11.023.

Senra JF de B, Silva VAC, Esposti MDD, Ferreira A, Milheiros IS, Ramos IB, Oliveira RG de Benevenute, LM. 2025. Selection of differentiated maturity genotypes of Coffea canephora. Acta Scientiarum. Agronomy 47: e70812. https://doi.org/10.4025/actasciagron.v47i1.70812.

Silva Júnior AC, Sant'Anna IC, Silva Siqueira MJ, Cruz CD, Azevedo CF, Nascimento M, Soares PC. 2022. Multi-trait and multi-environment Bayesian analysis to predict the G x E interaction in flood-irrigated rice. PLoS ONE 17: e0259607. https://doi.org/10.1371/journal.pone.0259607.

Smith HF. 1936. A discriminant function for plant selection. Annals of Eugenics 7: 240–250. https://doi.org/10.1111/j.1469-1809.1936.tb02143.x.

Sousa TV, Caixeta ET, Alkimim ER, Alkimim ER, Oliviera ACB, Pereira AA, Sakiyama NS, Zambolim L, Resende MDV. 2019. Early selection enabled by the implementation of genomic selection in *Coffea arábica* breeding. Frontiers in Plant Science 9:1–12. https://doi.org/10.3389/fpls.2018.01934.

Thorpe HHA, Fontanillas P, Pham BK et al. 2024. Genome-wide association studies of coffee intake in UK/US participants of European ancestry uncover cohort-specific genetic associations. Neuropsychopharmacol. 49:1609–1618. https://doi.org/10.1038/s41386-024-01870-x.

Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. 2014. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. New York: Springer. https://doi.org/10.1007/978-1-4614-1353-0.

Volpato L, Resende MDV, Silva FFR, Dias KOG, Oliveira GHF. 2019. Inference of population effect and progeny selection via a multi-trait index in soybean breeding. Acta Scientiarum. Agronomy 41:e42750. https://doi.org/10.4025/actasciagron.v41i1.42750.

Yadav S, Dillon S, McNeil M, Dinglasan E, Mago R, Dodds P, Hickey L, Hayes BJ. 2025. Optimising parent selection in plant breeding: comparing metaheuristic algorithms for genotype building. Theoretical and Applied Genetics 138-242. https://doi.org/10.1007/s00122-025-05028-1.

Acknowledgment

We thank the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the financial support for this research and grants granted to the authors.

Author contributions

Conceptualization: Silva Junior, AC, Soares, LG, Moura, WM, Cruz, CD, Bhering, LL. **Methodology**: Silva Junior, AC, Soares, LG, Moura, WM, Cruz, CD, Bhering, LL. **Investigation**: Silva Junior, AC, Soares, LG, Moura, WM, Cruz, CD, Bhering, LL. **Data Curation**: Soares, LG, Andrade, HSS, Moura, WM. **Writing – Original Draft**: Silva Junior, AC, Soares, LG, Moura, WM, Cruz, CD, Bhering, LL. **Writing – Review & Editing**: Filho, ACV, Ferrão, RG, Ferrão, MAG, Cecon, PR Moura, WM. **Funding acquisition**: Moura, WM.

Conflict of interest

The authors declare that they have no conflict of interest.

Data availability

The data will be made available upon request by co-author Dra. Waldenia Melo Moura, who is responsible for the data.

Declaration of use of AI Technologies

The authors declare that they have not acquired AI technologies.

Table 1. Estimation of the p-value for the likelihood ratio test and variance components of morpho-agronomic traits in five harvests of conilon coffee in the Zona da Mata region in Minas Gerais.

	PROD	VIG	FRC	FRU	RS	CS	LMS
Clone	2.88E-04	8.75E-20	1.11E-19	2.63E-09	3.37E-09	1.09E-06	0.062
Clone: Harvests	1.37E-07	5.21E-01	1.70E-10	9.61E-03	5.54E-01	4.71E-01	1.000
σ_g^2	331	0.19	0.26	0.30	0.62	0.34	0.49
σ_{ge}^2	952	0.36	0.23	0.14	0.80	0.728	0.10
σ_r^2	188	0.01	0.13	0.096	0.30	0.205	0.14

RS: rust severity (*Hemileia vastatrix*); CS: cercosporiosis severity (*Cercospora coffeicola*); LMS: coffee leaf miner bug attack severity (*Leucoptera coffeella*); VIG: vegetative vigour of the plants; FRC: fruit ripening cycle; FRU: fruit ripening uniformity; PROD: productivity (sc.ha^{-1}); σ_{ge}^2: variance of GE: σ_g^2: genotypic variance; σ_r^2: residual variance.

Table 2. Variance inflation factor, eigenvalues, explained variance, factor loadings after varimax rotation and communalities obtained in factor analysis**.**

Trait	FA1	FA2	FA3	Communality	Uniquenesses	VIF
PROD	**0.6210**	-0.0243	0.1770	0.4180	0.5820	1.51
VIG	**0.7150**	0.4420	0.1610	0.7320	0.2680	**22.70**
FRC	-0.0392	0.0988	**0.7810**	0.6210	0.3790	1.43
FRU	**0.8240**	0.0564	0.2540	0.7470	0.2530	2.30
RS	-0.2570	**0.8130**	-0.2830	0.8070	0.1930	1.34
CS	-0.4220	**0.6160**	0.3650	0.6900	0.3100	**15.93**
LMS	-0.0825	**0.8280**	-0.2560	0.7580	0.2420	**14.01**
Eigenvalues	2.09	1.67	1.01	-	-	
Variation	29.90	23.90	14.40	-	-	
Cumulative %	29.90	53.80	68.20	-	-	

RS: rust severity (*Hemileia vastatrix*); CS: cercosporiosis severity (*Cercospora coffeicola*); LMS: coffee leaf miner bug severity (*Leucoptera coffeella*); VIG: vegetative vigour of the plants; FRC:

Fruit Ripening Cycle; FRU: Fruit Ripening Uniformity; PROD: productivity (sc. ha^{-1}); FA1: Factor 1; FA2: Factor 2; FA3: Factor 3; VIF: Variance Inflation Factor.

Table 3. Estimate of the percentage of selection differential and broad-sense heritability of indices for seven traits and five harvests of conilon coffee.

		MTSI		SH		FAI-BLUP	
Trait	H^2	Xo(Xs)	SD%	Xo(Xs)	SD%	Xo(Xs)	SD%
FRC	0.66	1.61(1.77)	17.0	1.61(1.98)	21.10	1.61(1.97)	14.16
FRU	0.46	1.83(1.90)	6.72	1.83(1.77)	-1.82	1.83(1.84)	-10.76
PROD	0.25	64.89(77.75)	21.20	64.89(86.23)	26.18	64.9(74.5)	-2.57
VIG	0.53	7.22(7.44)	5.05	7.22(7.66)	5.23	7.22(7.45)	5.87
RS	0.32	2.12(2.03)	-2.16	2.12(2.31)	6.78	2.12(2.16)	-4.22
CS	0.23	2.53(2.54)	18.40	2.53(2.44)	-1.43	2.53(2.59)	5.32
LMS	0.09	2.01(2.03)	9.09	2.01(2.01)	0.02	2.01(202)	-0.19

RS: rust severity (*Hemileia vastatrix*); CS: cercosporiosis severity (*Cercospora coffeicola*); LMS: coffee leaf miner bug severity (*Leucoptera coffeella*); VIG: vegetative vigour of the plants; FRC: Fruit Ripening Cycle; FRU: Fruit Ripening Uniformity; PROD: productivity, in 60 kg bags of processed coffee per hectare (sc. ha^{-1}); X_s: mean of the selected genotypes; X_0: overall population mean; H^2 : broad-sense heritability; SD%: selection differential percentage; MTSI: Multi-Trait Stability Index; SH: Smith (1936) and Hazel (1943) index; FAI-BLUP: Multi-Trait index based on factor analysis and clone-ideotype distance.

Table 4. Predicted genetic gains for the indices MTSI, FAI-BLUP, SH and SH1.

			GG (%)			
Factor	Trait	Goal	MTSI	FAI-BLUP	SH	SH1
FA1	**VIG**	**Increase**	11.64	1.64	3.11	**NA**
FA1	FRU	Increase	21.7	5.14	-1.47	1.70
FA1	PROD	Increase	5.02	3.75	2.34	8.34
FA2	**CS**	**Reduction**	1.31	0.53	-0.82	**NA**
FA2	RS	Reduction	2.06	2.66	2.82	2.82
FA2	**LMS**	**Reduction**	0.643	0.51	0.003	**NA**
FA3	FRC	Increase	6.87	15.1	15.2	16.16
GGT%			49.24	29.33	20.18	29.02

RS: rust severity (*Hemileia vastatrix*); CS: cercosporiosis severity (*Cercospora coffeicola*); LMS: coffee leaf miner bug severity (*Leucoptera coffeella*); VIG: vegetative vigour of the plants; FRC: Fruit Ripening Cycle; FRU: Fruit Ripening Uniformity; PROD: productivity (sc. ha^{-1}); **NA**: Traits in bold were removed from SH1 due to multicollinearity problems; GG (%): percentage of genetic gain; GGT%: percentage of total genetic gain; MTSI: Multi-Trait Stability Index; SH: Smith (1936) and Hazel (1943) index; FAI-BLUP: Multi-Trait index based on factor analysis and clone-ideotype distance.

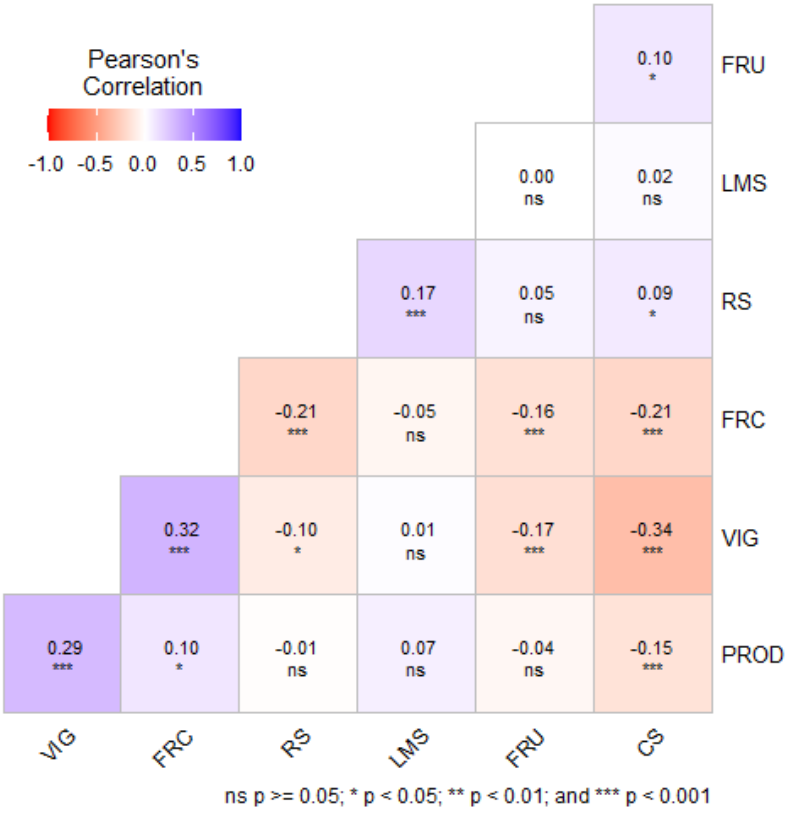


Figure 1. Pearson's correlation by BLUP between conilon coffee traits assessed in five consecutive harvests. RS: rust severity (*Hemileia vastatrix*); CS: cercosporiosis severity (*Cercospora coffeicola*); LMS: coffee leaf miner bug severity (*Leucoptera coffeella*); VIG: vegetative vigour of the plants; FRC: fruit ripening cycle; FRU: fruit ripening uniformity; PROD: productivity, in 60 kg bags of processed coffee per hectare (sc. ha^{-1}).

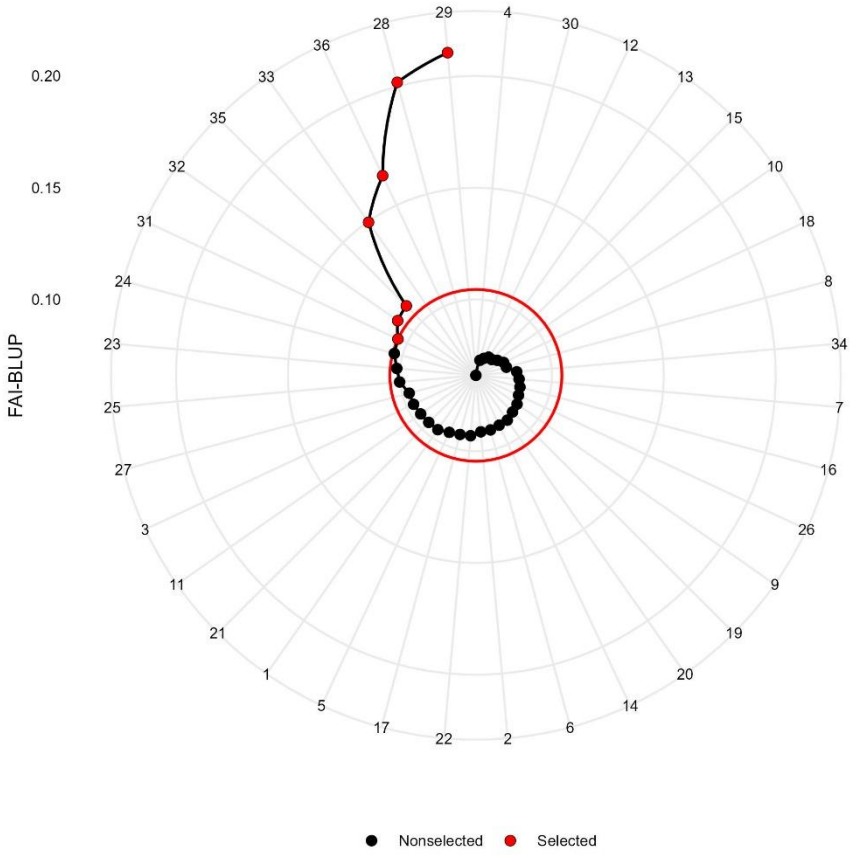


Figure 2. Selection of conilon coffee clones using the multi-trait index based on clone-ideotype distance (FAI-BLUP). Selection considering 20% selection intensity by the FAI-BLUP index, which is represented by the red line. The black dotted line indicates the conilon coffee clones assessed in different harvests. Selected clones correspond to the red dots.

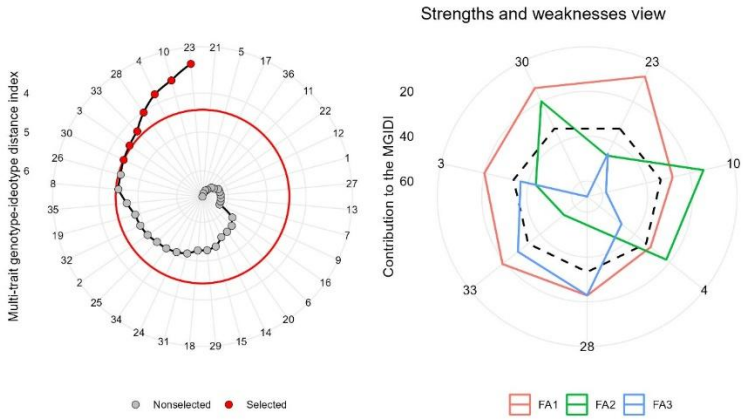
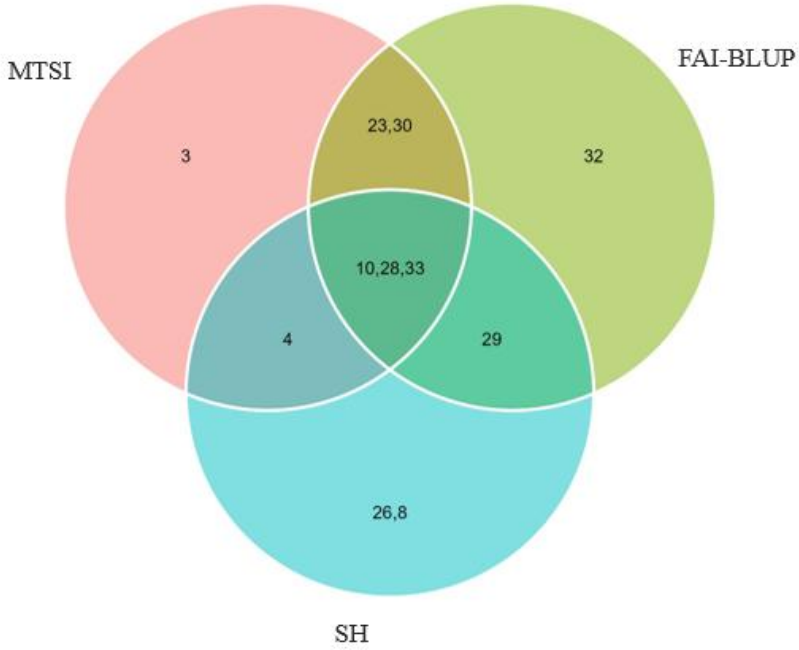


Figure 3. Genotypic classification of conilon coffee clones in ascending order for the MTSI index.



692

693 **Figure 4.** Index of selection coincidence between multi-trait indices based on factor analysis (MTSI, FAI-
694 BLUP) and clone-ideotype distance and Smith-Hazel (SH).