Alves Campos, Katia; Ramalho de Morais, Augusto; Arantes Paixão, Crysttian
Alternative for the evaluation of coffee seedlings using Fisher's discriminant analysis
Revista Ciência Agronômica, vol. 47, núm. 2, abril-julio, 2016, pp. 299-306
Universidade Federal do Ceará
Ceará, Brasil

Alternative for the evaluation of coffee seedlings using Fisher's discriminant analysis[1]

Alternativa para avaliação de mudas de cafeeiro por meio de análise discriminante de Fisher

Katia Alves Campos[2]*, Augusto Ramalho de Morais[3] and Crysttian Arantes Paixão[4]

ABSTRACT - One of the applications of Fisher's linear discriminant function (FDF) is its use in transforming multivariate data into a new univariate variable. This then makes possible a new option for the variance analysis of multivariate data, in addition to the multivariate analysis of variance (MANOVA). The aim of this work was to select groups of seven characteristics of quality in coffee seedlings using six criteria for selection, to use the FDF to transform such groupings of characteristics into a new variable, and then to compare interpretation of the results obtained from the univariate and multivariate analyses of variance of the characteristics and this new variable, with a view to its use in evaluating coffee seedlings. A randomised block design was used to assess the effect of organic fertiliser on the formation of seedlings in coffee cv. Catuaí Vermelho IAC-44, evaluating the following characteristics: seedling height, diameter, root length, dry weight of shoots and roots, leaf area, number of leaves and total dry weight. According to the selection criteria used, different subsets of the selected characteristics are possible. The use of the FDF is shown to be viable in discriminating between treatments. Univariate analysis of the new variable obtained with the FDF and multivariate analysis (MANOVA) was able to detect differences between the treatments, however, it is simpler to apply FDF methodology.

Key words: Variable selection. Multivariate analysis. Data transformation. Analysis of variance. *Coffea arabica.*

RESUMO - Uma das aplicações da função discriminante linear de Fisher (FDF) é a sua utilização, para a transformação de dados multivariados, em nova variável univariada, de maneira a possibilitar uma nova opção de análise de variância dos dados multivariados além da análise de variância multivariada (MANAVA). O objetivo deste trabalho é selecionar agrupamentos de sete características de qualidade de mudas de cafeeiro, por meio de seis critérios de seleção, utilizar a FDF para transformar tais agrupamentos das características em uma nova variável e posteriormente comparar a interpretação dos resultados obtidos com as análises de variância univariada e multivariada das características mensuradas e essa nova variável, com vista em seu uso na avaliação de mudas de cafeeiro. O delineamento utilizado foi em blocos casualizados para avaliar o efeito de adubações orgânicas na formação de mudas de cafeeiro cv. Catuaí Vermelho IAC-44, avaliando as características altura de mudas, diâmetro, comprimento radicular, peso seco da parte aérea e da raiz, área foliar, número de folhas e peso seco total. De acordo com o critério de seleção usado, podem-se ter diferentes subconjuntos de características selecionados. A utilização da FDF se mostra viável na discriminação de tratamentos. A análise univariada da nova variável obtida pela FDF e a análise multivariada (MANAVA) foram capazes de detectar diferenças entre tratamentos, contudo, a aplicação da metodologia da FDF é mais simples.

Palavras-chave: Seleção de variável. Análise multivariada. Transformação de dados. Análise de variância. *Coffea arabica.*

INTRODUCTION

In experiments with coffee seedlings, usually more than one response variable are measured in order to improve characterisation of the plants and treatments being used. These experiments have as options in the evaluation, univariate analysis of variance (ANOVA), and multivariate analysis of variance (MANAVA), which is rarely used. Both of these options, univariate and multivariate analysis, have disadvantages relating to conclusions from the results they provide. In univariate analysis, each response variable is analysed individually and a joint conclusion is arrived at. It must be pointed out that the level of significance of this joint conclusion will not be known, as it derives from a combination of results. Whereas in the case of multivariate analysis, and despite the level of significance being determined, since the analysis is carried out using all the characteristics, it is more complex than univariate analysis, demanding a more careful interpretation of the results.

One option is to use the transformation of multivariate data into a new variable which would accumulate information on the characteristics, and facilitate interpretation of the results. Among the transformations possible, the Fisher linear discriminant function (FDF) was addressed by Pimentel-Gomes (2009), who demonstrated the technique in an experiment in which they evaluated levels of nitrogen and phosphorus in order to compare three different treatments.

The initial proposal by Fisher (1936) was to use weighting, employing the canonical variable, to obtain a variable capable of classifying a new individual in one of two pre-defined populations. This technique was later extended to g populations, where $G \geq 2$.

By applying FDF to multivariate observations, this G-dimensional space can be reduced to a one-dimensional space whose new variable displays a maximum value for the F-test in univariate analysis of variance, within certain restrictions such as the homogeneity of covariance matrices, and data with multivariate normal distribution (MANLY, 2008; PIMENTEL-GOMES, 2009).

When comparing commercial broiler hybrids with experimental hybrids, Fonseca *et al*. (2002b) using FDF, affirmed the superiority of commercial hybrids by means of six characteristics, four of which were significant in individual analyses. Torres Filho *et al*. (2005) used FDF to study genetic divergence among strains of pig.

Using multivariate techniques, including FDF, Carneiro *et al*. (2006) evaluated genetic divergence in sheep populations by means of seven growth characteristics. In an experiment for intercropping lettuce and carrots, Bezerra Neto *et al*. (2007) pointed out that the use of the multivariate method proved to be informative and had a greater discriminating capacity. Binotto, Lucio and Lopes (2010) evaluated relationships in eucalyptus seedlings between characteristics and a quality index, highlighting dry phytomass as being the most correlated.

In the work of Santana *et al*. (2011), FDF was used as an additional technique to the univariate analysis of variance, to evaluate the use of foliar fertilisation in coffee seedlings during their time in the nursery, and by Silva *et al*. (2012) when comparing the composition of substrates and managements in three cultivars of *Coffea arabica* L.

In the various studies employing FDF as a statistical technique, it was used either as a way to get yet one more response variable for the univariate analysis or as a multivariate technique to aid in analysis. The present work demonstrates the feasibility of using FDF in an experiment with seven characteristics which is not found in other reference work, and particularly not with coffee. It is worth noting that using simulations, Campos (2012) demonstrated the feasibility of adopting FDF in the analysis of experiments, when he demonstrated that in the analysis of variance of transformed data, differences were found which were compatible to the differences highlighted when using MANAVA.

The aim of this work therefore was to select clusters of quality characteristics in coffee seedlings, use FDF to transform these characteristics into a new variable, and compare interpretation of the results obtained with univariate (ANOVA) and multivariate (MANAVA) analysis of variance of the characteristics being measured and with this new variable.

MATERIAL AND METHODS

In order to apply the proposed methodology, data was used from a trial where the aim of the original study had been to test three sources of organic fertiliser: (A) rotted manure from laying hens (70 L m^{-3}), (B) rotted manure from dairy cattle (300 L m^{-3}), (C) worm humus (200 L m^{-3}) and (D) a control treatment, which only received a fertiliser common to all the other treatments. The experiment was carried out from September 2008 to March 2009 in a randomised block design with five replications, at the seedling nursery of the Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais, *Campus* Machado, located at 21º40'29" S and 45º55'11" W, at an altitude of 820 m. Seeds from the coffee cultivar Catuaí Vermelho IAC-44 were used, with each lot comprising four polyethylene bags.

For the formation and development of the seedlings, the following cropping treatments were carried

out: sowing, preparation of the substrate, cover for the beds, and daily irrigation during the periods of seed germination and seedling emergence. Seeds were sown directly, with two seeds per polyethylene bag (20 cm high by 10 cm wide), followed by thinning when the seedlings reached the "Jaguar Ear" stage (unfolded cotyledons). The substrate was made from a distropheric Red Latosol (700 L m^{-3}) taken from a gully, and fertilisation was by the addition of natural phosphate (5 kg m^{-3}) and potassium sulphate (0.5 kg m^{-3}) (MATIELLO *et al.*, 2010).

At 180 days after planting, the following characteristics were taken: seedling height (HGT) measured with a ruler in centimetres, considering the root collar to the apical bud; collar diameter (DIAM), measured with a digital calliper, in millimetres; root length (ROOT) measured with a ruler from the cap of the primary root to the collar, in centimetres; shoot dry weight (SDW) and root dry weight (RDW), both determined on a digital balance after drying in an oven to constant weight, in grams; leaf area (AREA) determined as per Silva, Leite and Ferreira (2008); and number of true leaves (NLEAF). The variable, total seedling dry weight (TDW), was also obtained in grams by summing RDW and SDW.

Analysis of variance was carried out for each of the eight characteristics as per Pimentel-Gomes (2009). Also evaluated were the assumptions for error normality by the Shapiro-Wilk test (1%), and homogeneity of variance using Bartlett's test (1%). The mean values for those treatments where the F-test was significant were compared by Tukey's test (5%).

These eight characteristics were grouped so as to cover all possible subsets, using combinations of from two to seven characteristics, as shown in Table 1. Subsets that had linearly dependent characteristics were excluded, so that combinations in which TDW appeared together with SDW or RDW were not considered. In total, 151 combinations were formed, and the FDF with its respective analysis of variance was obtained for each one.

To estimate the FDF for each group, it was necessary to find the eigenvector *t* which maximizes the ratio (t'Ht)/(t'Rt), where H and R represent respectively the sum of squares and the sum of products matrix due to the effects of the treatments and the residuals. Such a vector can be found by means of the Lagrangian method, which employs maximisation of the numerator, with the state of the denominator taken to be a constant. Determination of the roots of the characteristic polynomial det(R^{-1} - $\lambda I = 0$), is equal to the resolution of the thus-formed system, where I is the identity matrix and λ the largest eigenvalue for $R^{-1}H$ (PADOVANI; ARAGON, 2005).

The eigenvector found in this way expresses the weighting coefficients for each characteristic, and can be interpreted as the coefficient of a multiple regression model that serves to identify the variables which contribute most to distinguish between treatments (SIMEON; PADOVANI, 2008); or it can be used to transform the multivariate data into univariate data (CARNEIRO *et al.,* 2006; FONSECA *et al.,* 2002b; PIMENTEL-GOMES, 2009; SANTANA *et al.,* 2011; SILVA *et al.,* 2012; TORRES FILHO *et al.*, 2005).

The transformed data obtained with the Fisher linear discriminant function for the 151 combinations of characteristics, underwent different tests for univariate and multivariate analysis. The tests carried out in both analyses were: analysis of variance (ANOVA and MANAVA);

Table 1 - Number of characteristics in each subset, the number of possible subsets, and subsets formed from groups ranging from 1 to 7 characteristics of seedling quality, to build the Fisher linear discriminant function and the multivariate analysis of variance

No. of Characteristics	No. of Subsets	Subset
1	8	HGT, DIAM, ROOT, SDW, RDW, AREA, NLEAF, TDW
2	26	HGT+DIAM, HGT+ROOT, ..., NLEAF+TDW
3	45	HGT+DIAM+ROOT, HGT+DIAM+SDW, ..., AREA+NLEAF+TDW
4	45	HGT+DIAM+ROOT+SDW, HGT+DIAM+ROOT+RDW, ..., SDW+RDW+AREA+NLEAF
5	26	HGT+DIAM+ROOT+SDW+RDW, HGT+DIAM+ROOT+SDW+AREA, ..., ROOT+SDW+RDW+AREA+NLEAF
6	8	HGT+DIAM+ROOT+SDW+RDW+AREA, ..., DIAM+ROOT+SDW+RDW+AREA+NLEAF
7	1	HGT+DIAM+ROOT+SDW+RDW+AREA+NLEAF

residual normality (Shapiro-Wilk test); homogeneity of variance (Bartlett's test); with multiple comparison (Tukey's test) being used to complement the analysis of variance where necessary. In the case of multivariate analysis, the homogeneity of the covariance matrices was also tested (SIMEÃO; PADOVANI, 2008). Significance levels employed in the tests were 1% for normality and homogeneity, and 5% for multiple comparison.

In order to select which characteristics of seedling quality should be combined to be used in the multivariate studies and make the FDF more informative, six selection criteria were applied: Akaike information criterion (AIC), Mallows Cp criterion (CHARNET *et al.,* 2008) and the four dimensionless selection criteria τ^2, ρ^2, ζ^2 and ξ^2 associated with the MANAVA F-tests, whose estimators were taken from Huberty (2002) and are described below in equations 1 to 4:

$$\hat{\rho}_1{}^2 = \theta_1, \tag{1}$$

$$\hat{\tau}^2 = 1 - \Lambda^{1/r}, \tag{2}$$

$$\hat{\zeta}^2 = U/(U+r), \tag{3}$$

$$\hat{\xi}^2 = V/r, \tag{4}$$

where: θ_1 is the estimate of the Roy maximum eigenvalue θ test; Λ is the estimate with Wilk's Λ test; U is the estimate with the Hotelling-Lawley test; V is the estimate with Pillai's test, r being the rank of the sum of the squares and products matrix.

After selection, in order to compare the results obtained by each method, the p-values obtained in the four approximate F-tests were examined (Roy's maximum eigenvalue θ, Wilk's Λ, Hotelling-Lawley *(U)* and Pillai *(V)*, for which statistics can be found in Manly (2008) among others) for the multivariate analysis of variance, with the p-values obtained for the univariate analysis, and with data processed by the respective FDF.

In order to show which FDF gave the most information, a check was carried out of the assumptions of the univariate analysis of variance, and an estimation made of the percentage explanation given by the transformation which was estimated for the ratio $\lambda_1/\Sigma\lambda_i$, suggested by Padovani and Aragon (2005), where λ_i are the non-zero eigenvalues of $R^{-1}H$, and λ_1 the greatest eigenvalue.

RESULTS AND DISCUSSION

To facilitate presentation of the data, each characteristic of plant quality was numbered: HGT - 1, DIAM - 2, ROOT - 3 SDW - 4, RDW - 5 AREA - 6, NLEAF - 7 and TDW - 8. The characteristics combined for evaluation by multivariate analysis (MANAVA) or Fisher's linear discriminant function (FDF) were denoted by CC. Thus, for example, CC.34 indicates formation of the subset obtained through study of the combined characteristics of root length (ROOT) and shoot dry weight (SDW). In the same way, data obtained with the Fisher linear discriminant function, after applying the transformations to a subset of characteristics, were denoted by FDF. Adopting the above subset as an example, after data transformation this was referred to as FDF.34.

For the assumptions of error normality, the p-values obtained using the Shapiro-Wilk test were in the range [0.1072, 0.9999], which ensured normality of the errors, even though a critical level of significance of 10% had been adopted for all characteristics and also for the transformed data.

With the test for homogeneity of variances, it was found that except for the data transformed by FDF.1238, whose p-value was 0.0060, the other p-values were contained in the interval [0.0153, 0.8023], a range which did not include the level of 1% considered in this work to be the cut-off level. This indicated that variances were homogeneous at the minimum level of 1.53% for the combination CC.1234, and at a level greater than 80% for the combination CC.2345.

By means of the F-test, significant differences were detected between the mean values for the treatments for all characteristics, except for the number of leaves (p = 0.4646), with the range of p-values, excluding any not significant, being between [0 0000, 0.0051].

Among the univariate characteristics, NLEAF (7) had the smallest dimensionless criterion ($\tau^2 = 0.1855$), as well as not being homoscedastic; this was followed by ROOT (3), HGT (1), DIAM (2) and RDW (5), which appeared with the worst classifications in all subsets. The combined characteristics classified in last place in the selections were CC.37 ($\rho^2 = 0.6969$ and $\zeta^2 = 0.5463$), CC.2357 (by the criteria, $\tau^2 = 0.4070$ and $\xi^2 = 0.2468$), and the worst classified, CC.13, by the Mallows Cp criterion (55016.13). It was found that root length (3) is a characteristic which is frequently seen in groups with the worst classifications; the use of this characteristic was also questioned by Peixoto and Peixoto (2009) along with other the root measurements, volume and diameter, which, although informative, should be disregarded, in view of the insurmountable errors accrued due to difficulties in their measurement.

Generally, the selection criteria do not coincide. In this example, based on the estimated values for the dimensionless indices, the subsets CC.1234567 ($\rho^2 = 0.9802$), CC.67 ($\tau^2 = 0.8429$ and $\zeta^2 = 0.9478$) and CC. 45 ($\xi^2 = 0.7359$) were selected as being the most informative.

The Akaike criterion (AIC = -74.8417) made it possible to classify the combination of characteristics CC.245 as the most informative. The subset CC.1234567, containing all the characteristics that are linearly independent, was also considered the most significant using Mallows Cp. Table 2 shows the statistics from each test for the top four sets of groups of characteristics.

None of the selected sets adopted TDW as being among the most important characteristics, a result that should be better investigated, given that in almost all the work with coffee seedlings, this is one of the most common variables. This fact is probably due to the presence of the characteristics SDW and RDW, whose sum is the TDW, suggesting that when analysing the individual parts it is not necessary to include total dry matter. This reinforces comments by Gomes *et al.* (2002) and Bernardino *et al.* (2005), who, although they consider dry matter to be an important characteristic, claim that it should be considered separately, in order to avoid distortions, especially in species with a greater number of leaves.

Another point that was worthy of note was studying non-destructive quality characteristics in seedlings with a view to meeting the need of experiments that require continuity. Miglioranza *et al.* (2010), studying which features most determined the quality of coffee seedlings, claimed that height, seedling diameter, leaf area and dry weight were correlated, but that weight, being destructive, should be avoided. However, the combination of dry weights (CC.45), and the combination of collar diameter and dry weights (CC.245) were both selected; of the non-destructive characteristics in this study, only the combination of leaf area and leaf number (CC.67) was selected statistically, using the criteria τ^2 and ζ^2.

With the multivariate analysis (MANAVA) of the selected sets, differences were found between treatment effects by means of the four approximate F-tests (p<0.05), with Pillai's test being the most critical. Table 3 shows the p-values for the four approximate F-tests: Roy's maximum eigenvalue θ, Wilk's Λ, Hotelling-Lawley (U) and Pillai (V), and also the p-values obtained with the F-test for the univariate analysis of variance of the transformed data (p<0.01).

The p-values relative to the analyses of variance of the new univariate variable, and obtained by transforming the data by the FDF for some of the selected subsets (Table 3), show that the assumptions of error normality and homogeneity of variance were accepted, as all p-values were greater than 0.05 (p>0.05); specifically, the lowest value being 0.14 with the transformation carried out on the data for dry weight (FDF.45).

It was also found that with the F-test in the univariate analyses of the new variable (FDF), as well as the approximate F-tests in the multivariate analyses (Table 3), the hypothesis of treatment-effect equality was rejected, with a value for p of <0.018. Therefore, with similar results to the F-test, the five tests displayed the same behaviour, with the univariate analysis of the new variable giving results consistent with multivariate analysis. It is worth noting that the first discriminant function for each of the chosen subsets returned a percentage explanation of the transformed data greater than 80% (Table 3).

The estimated term for calculation of the new variable transformed by the Fisher linear discriminant function and chosen by the ρ^2 and Mallows Cp criteria, was FDF.1234567 = - 0.0143 HGT + 0.1032 DIAM + 0.0102 ROOT + 0.5559 SDW - 0.8220 RDW + 0.0266 AREA - 0.0590 NLEAF. It can be seen that the greatest weightings were assigned by the variables SDW and RDW, and the smallest weightings by height (HGT) and root length (ROOT).

The allocation of these weightings in the Fisher linear discriminant function can be explained through the applied work of several authors, for example, the studies developed by Gomes *et al.* (2002) with eucalyptus

Table 2 - Estimated values for the statistics ρ^2, τ^2, ζ^2 and ξ^2, Akaike information criterion (AIC) and Mallow's Cp with their respective sort orders (*), for the four best grouped sets of characteristics in the multivariate tests

Combined Characteristics	ρ^2		τ^2		ζ^2		ξ^2		AIC		Mallows Cp Difference	
45	0.89	123°	0.79	8°	0.83	101°	0.74	1°	-67.77	2°	12.73	52°
67	0.97	40°	0.84	1°	0.95	1°	0.53	19°	154.66	28°	3724.04	144°
245	0.90	119°	0.66	71°	0.78	116°	0.51	27°	-74.87	1°	13.09	55°
1234567	0.98	1°	0.56	136°	0.88	66°	0.29	149°	480.38	151°	0.00	1°

1 - seedling height, 2 - collar diameter, 3 - root length, 4 - shoot dry weight, 5 - root dry weight, 6 - leaf area, 7 - number of leaves and 8 - total dry weight. (*) Classification is made by individual decision for each criterion: the dimensionless criteria ρ^2, τ^2, ζ^2 and ξ^2 should be closer to one; the AIC information criterion chosen will be that with the lowest relative value; for Mallows Cp the best statistic will be closest to zero

Table 3 - p-values for the tests of error normality (Shapiro-Wilk), homogeneity of variance (Bartlett), F-test, and percentage explanation for the new variable (FDF) formed from selected subsets of the characteristics and, for the Roy, Wilks, Hotelling-Lawley and Pillai criteria, from the multivariate analysis of variance (MANAVA)

Transformation	Tests for the new variable (FDF)				Combined Characteristics	Tests for MANAVA			
	Shapiro-Wilk	Bartlett	F-Test	$\lambda_l/\Sigma\lambda_i$ (%)		Roy θ	Wilks Λ	Hotelling-Lawley (U)	Pillai (V)
FDF.45	0.1455	0.5526	0.000	84.1	CC.45	0.000	0.0000	0.0000	0.0000
FDF.67	0.4489	0.3116	0.000	96.8	CC.67	0.000	0.0000	0.0000	0.0035
FDF.245	0.4422	0.7526	0.000	80.3	CC.245	0.000	0.0000	0.0000	0.0010
FDF.1234567	0.3254	0.7446	0.000	84.9	CC.1234567	0.000	0.0003	0.0000	0.0183

seedlings ready for planting, demonstrated the importance of studying dry matter, SDW and RDW, as these represent the hardiness of the plant, and are correlated directly with the survival and the initial performance of the seedlings after planting in the field. While the lower value assigned to the weighting for seedling height may be confirmed by a study from Fonseca *et al*. (2002a) with plants of the woody species, *Trema micrantha*; the authors suggest that this is a feature that should be of restricted use, since "taller but weaker plants may be selected, and the smaller, more vigorous plants discarded". Other authors who comment on the importance of height as a characteristic, are Birchler *et al*. (1998), who state that despite being a measurement which is easily made and non-destructive, and which can provide an approximation of photosynthesis and air transpiration, it has the disadvantage of not taking the stem architecture into account. Furthermore, for root length, Peixoto and Peixoto (2009) focused on the importance of using root measurements, but due to the lack of standardisation in their collection, also advised against their use.

However, when only non-destructive characteristics are considered, the Fisher linear discriminant function selected by the criteria ζ^2 and τ^2 is FDF.67 = 0.6100 AREA - 0.7924 NLEAF, where the weightings are similar, albeit of opposite sign. These characteristics of seedling quality are mentioned, as the leaf is an important organ for plants, being the main organ involved in photosynthesis and evapotranspiration responsible for the exchange of gasses between the plant and the environment (PEREIRA; VILLA NOVA; SEDIYAMA, 1997).

With the mean values shown in Table 4, it was found that from the equality of treatment effects on the number of leaves, the individual characteristics can be seen as three more ways to structure classification by treatment effect. For example, for the height variable, it was not possible to distinguish between treatments B and C (cattle manure and worm humus respectively). For the values, transformed by the Fisher linear discriminant function, FDF.1234567 and FDF.67, it is possible to detect the differences between the effects of these treatments as expected, given that they are different sources of organic fertiliser with different compositions, in a substrate with no addition of supplementary organic fertiliser.

The FDF procedure constitutes a very interesting alternative, primarily because of its practical aspects, as it was possible to detect significant differences between

Table 4 - Mean values for the characteristics height (HGT), stem diameter (DIAM), root length (ROOT), shoot dry matter (SDW), root dry matter (RDW), leaf area (AREA), number of leaves (NLEAF), total dry matter (TDW) and for two new variables obtained with the Fisher discriminant function, with their respective comparisons by Tukey's test (5%) for the treatments

Treament	HGT	DIAM	ROOT	SDW	RDW	AREA	NLEAF	TDW	FDF.1234567	FDF.67
A	25.9 a	3.0 a	23.8 a	1.1 a	0.3 a	29.2 a	8.0 a	1.3 a	0.85 a	11.21 a
B	23.6 ab	2.6 b	20.9 b	0.8 a	0.2 b	26.3 a	7.5 a	1.0 a	0.73 a	9.96 a
C	18.0 bc	2.4 b	20.2 b	0.4 b	0.1 c	14.3 b	7.5 a	0.6 b	0.28 b	3.13 b
D	16.6 c	2.3 b	20.0 b	0.4 b	0.1 c	13.0 b	6.8 a	0.5 b	0.23 b	2.80 b

A - rotted bird manure, B - rotted cattle manure, C - worm humus, and D - control. Mean values followed by the same letters in a column do not differ by Tukey's test at 5%

treatments, which had not been found with the individual analyses. It should be noted, that in research the main purpose is to choose the treatment or treatments that stand out from the rest, and not just focus on the mean values returned, as these are simply estimates. In this respect, the new FDF variable was more informative than the individual characteristics, as it made it possible to easily separate the treatments.

CONCLUSIONS

1. The Fisher linear discriminant function (FDF) is a useful statistical technique to be used in the analysis of experiments with coffee seedlings. The FDF makes it possible to transform multivariate data into univariate data with high practical efficiency, facilitating the decision-making process. In this respect, analysis of variance detects differences compatible with multivariate analysis of variance, highlighting the ease of decision making;

2. The proposed method was shown to be effective in discriminating between the different treatments.

ACKNOWLEDGEMENTS

REFERENCES

BERNARDINO, D. C. S. *et al*. Crescimento e qualidade de mudas de *Anadenanthera macrocarpa* (Benth.) Brenan em resposta à saturação por bases do substrato. **Revista Árvore**, v. 29, n. 6, p. 863-870, 2005.

BEZERRA NETO, F. *et al*. Desempenho de sistemas consorciados de cenoura e alface avaliados através de métodos uni e multivariados. **Horticultura Brasileira**, v. 25, p. 514-520, 2007.

BINOTTO, A. F.; LÚCIO, A. D.; LOPES, S. J. Correlations between growth variables and the Dickson quality index in forest seedlings. **Cerne**, v. 16, n. 4, p. 457-464, 2010.

BIRCHLER, T. *et al*. La planta ideal: revision del concepto, parametros definitorios e implementaction practica. **Investigacion Agraria, Sistemas y Recursos Forestales**, v. 7, n. 1/2, p. 109-121, 1998.

CAMPOS, K. A. **Função discriminante de Fisher como alternativa à análise de variância multivariada**. 2012.

106 f. Tese (Doutorado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2012.

CARNEIRO, P. L. S. *et al*. Estudo de populações de ovinos Santa Inês utilizando técnicas de análise multivariada. **Revista Científica de Produção Animal**, v. 8, n. 1, p. 40-50, 2006.

CHARNET, R. *et al*. **Análise de modelos de regressão linear -** com aplicações. 2. ed. Campinas: Ed. UNICAMP, 2008. 356 p.

FISHER, R. A. The use of multiple measurements in taxonomic problems. **Annals of Eugenics**, v. 7, n. 2, p. 179-188, 1936.

FONSECA, E. P. *et al*. Padrão de qualidade de mudas de Trema micrantha (L.) Blume, produzidas sob diferentes períodos de sombreamento. **Revista Árvore**, v. 26, n. 4, p. 515-523, 2002a.

FONSECA, R. *et al*. Avaliação de frangos de corte utilizando técnicas de análise multivariada: I Características de carcaça. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 54, n. 5, p. 525-529, 2002b.

GOMES, J. M. *et al*. Parâmetros morfológicos na avaliação de qualidade de mudas de *Eucalypitus grandis*. **Revista Árvore**, v. 26, n. 6, p. 655-664, nov. 2002.

HUBERTY, C. J. A history of effect size indices. **Educational and Psychological Measurement**, v. 62, p. 227-240, 2002.

MANLY, B. J. F. **Métodos estatísticos multivariados**: uma introdução. 3. ed. Porto Alegre: Bookman, 2008. 229 p.

MATIELLO, J. B. *et al*. **Cultura de café no Brasil**: manual de recomendações. Edição 2010. Rio de Janeiro: MAPA: Procafé; Varginha: Fundação Procafé, 2010. 542 p.

MIGLIORANZA, E. *et al*. Estudos de parâmetros para determinação da qualidade de mudas de *Coffea arabica* L. cv Mundo Novo sob efeito do sombreamento. **Semina: Ciências Agrárias**, v. 14, n. 11, p. 5-8, 2010.

PADOVANI, C. R. P.; ARAGON, F. F. Programa computacional para método de discriminante de Fisher. **Revista Energia na Agricultura,** v. 20, n. 1, p. 1-10. 2005.

PEIXOTO, C. P.; PEIXOTO, M. F. S. P. Dinâmica do crescimento vegetal: princípios básicos. *In*: CARVALHO, C. A. L. *et al*. (Org.) **Tópicos em ciências agrárias**. Cruz das Almas: Nova Civilização, 2009, v. 1, p. 37-53.

PEREIRA, A. R.; VILLA NOVA, N. A.; SEDIYAMA, R. **Evapotranspiração**. Piracicaba: FEALQ: ESALQ: USP, 1997. 70 p.

PIMENTEL-GOMES, F. **Curso de estatística experimental**. 15. ed. Piracicaba: FEALQ, 2009. 451 p.

SANTANA, S. L. A. *et al*. Fertilização foliar em mudas de cafeeiro com organominerais líquidos. **Tecnologia & Ciência Agropecuária**, v. 5, n. 3, p. 9-13, set. 2011.

SILVA, A. R.; LEITE, M. T.; FERREIRA, M. C. Estimativa da área foliar e capacidade de retenção de calda fitossanitária em cafeeiro. **Bioscience Journal**, v. 24, n. 3, p. 66-73, 2008.

SILVA, E. A. *et al*. Desenvolvimento de mudas de cafeeiro *Coffea arabica* L sob diferentes composições de substratos. **Enciclopédia Biosfera**, v. 8, n. 14, p. 337-346, 2012.

SIMEÃO, S. F. A. P.; PADOVANI, C. R. Utilização da função discriminante quadrática em ciências experimentais. **Revista Energia na Agricultura**, v. 23, n. 1, p. 116-134, 2008.

TORRES FILHO, R. A. *et al*. Estudo da divergência genética entre linhas de suínos utilizando técnicas de análise multivariada. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 57, n. 3, p. 390-395, 2005.