Analysis of defects in coffee beans compared to biplots for simultaneous tables[1]

Análise de defeitos nos grãos de café comparados por *biplots* para tabelas simultâneas

Carla Regina Guimarães Brighenti[2]* and Marcelo Angelo Cirillo[3]

ABSTRACT - The demand for high quality coffee has become a consolidated criterion to achieve the best prices. Currently, cooperatives evaluate the coffee beans mainly through the particle size and the number of defects in the sample. This evaluation type generates counting data that originates contingency tables from different periods or groups involving the same variables in the row and column and there may be interest in knowing if two tables are related and how much are related. These are the so-called combined tables. Statistical analysis techniques normally employed do not include categorical data in the combined tables. The aim of this study was to evaluate the incidence of different types of defects in samples of large flat coffee beans in two different harvests through the construction of biplots. The decomposition theory in single simultaneous values of double entry contingency tables was used. The results of defect counting in beans of 24 coffee samples from southern Minas Gerais, Brazil, were evaluated in the 2014 and 2015 harvests. Moreover, the association among defect types, considered within different total defect proportions in the sample, was verified based on the percentage in 17/18 sieves. It was also evaluated the relative sums of squares from the similarity and dissimilarity among the harvests. It is concluded that the simultaneous analysis technique allows better visualizing the common behavior and alterations among different harvests, distinguishing the defect types associated with each harvest and among different proportions of large flat beans.

Key words: *Coffea arabica* L.. Coffee quality. Particle size. SVD. Sieves.

RESUMO - A demanda por café de alta qualidade tornou-se um critério consolidado para atingir os melhores preços. Atualmente, as cooperativas avaliam os grãos de café através, principalmente, da granulometria e da quantidade de defeitos na amostra. Este tipo de avaliação gera dados de contagem que originam tabelas de contingência oriundas de épocas ou grupos distintos envolvendo as mesmas variáveis em linha e coluna e pode haver interesse em saber se duas tabelas estão relacionadas, e o quanto estão relacionadas. Estas são as denominadas tabelas combinadas. As técnicas de análise estatística, normalmente empregadas, não contemplam dados categóricos em tabelas combinadas. Assim, este trabalho tem por objetivo avaliar a incidência de diferentes tipos de defeitos em amostras de grãos de café chatos graúdos em duas safras diferentes, através da construção de *biplots*. Foi utilizada a teoria da decomposição em valores singulares simultânea de tabelas de contingência de dupla entrada. Avaliou-se os resultados da contagem de defeitos em grãos de 24 amostras de cafés provenientes do Sul de Minas Gerais, nas safras de 2014 e 2015. Além disso, verificou-se a associação entre os tipos de defeitos, considerados dentro de diferentes proporções totais de defeitos na amostra, tendo como base o percentual em peneiras 17/18. Avaliou-se ainda, as somas de quadrados relativas, a partir da similaridade e dissimilaridade entre as safras. Conclui-se que a técnica de análise simultânea permite visualizar melhor o comportamento comum e alterações entre diferentes safras, distinguindo os tipos de defeito associados a cada safra e entre diferentes proporções de grãos chatos graúdos.

Palavras-chave: *Coffea arabica* L.. Qualidade do café. Granulometria. DVS. Peneiras.

INTRODUCTION

The classification of coffee beans comprises three distinct phases: by types and defects, by bean quality (sieve) and by beverage quality (ABRAHÃO *et al.*, 2009; RAMOS *et al.*, 2016). In the classification by types and defects, the coloration and counting of defect beans (intrinsic defects) or impurities (extrinsic defects) contained in a sample of 300 g of benefited beans are evaluated, according to Brazilian Official Coffee Classification (COB) table (BRASIL, 2003).

The intrinsic defects of coffee beans are caused by incorrect use of agricultural and industrial processes and modifications of physiological or genetic origin, such as full or partly black, husks, greens, insect damage and immatures. Extrinsic defects are elements that are strange to the coffee beans (stones, clods, bark, broken and others), also known as impurities (CUSTÓDIO; GOMES; LIMA, 2007).

Cooperatives may have problems in assigning greater value to the coffee marketed by its members and if the number of defects found during the classification harms the sample quality. COB table establishes type 4 as the basis for classification, with a sum of defects equal to or less than 160 (CORADI; BOREM; OLIVEIRA, 2008).

Another important criterion in the commercialization of coffee, which is widely used in cooperatives, is the measurement of particle size through sieves (WACHHOLZ; POYER, 2014). COB table establishes the classification in flat beans or mocha (BRASIL, 2003). Batches with larger quantities of large flat beans are more valued (BÁRTHOLO; GUIMARÃES, 1997).

The high diversity of Brazilian harvests brings the need to compare them. Fernandes *et al.* (2003) quantified the chemical composition of coffee from different harvests and the effects on the quality of roasted coffee. In this respect, they used arabica coffee beans from the 88/89 and 2000 harvests and the conilon coffee 2000 harvest, and concluded that there were significant differences among the harvests results for some physical and chemical characteristics evaluated in the study.

Lima, Custódio and Gomes (2008) evaluated the effect of irrigation on coffee productivity and yield and did not notice significant differences among the harvests, but when comparing the total accumulated production of the harvests, there was a statistical difference between the control and irrigated treatments. Pedroso *et al.* (2009) evaluated the quality of coffee seeds produced under two planting densities and three irrigation regimes during three harvests (first, second and third production years) and observed different yields due to the harvests and a

higher incidence of mocha beans in non-irrigated plants in the year 2003. In both studies, the different harvests were considered as being a plot in the adopted design.

Custódio, Gomes and Lima (2007) studied the classification in relation to coffee type using the total number of defects in five harvests through different irrigation depths. However, it was not possible to discuss the similarity among harvests with the statistical technique used.

Defect counts can be summarized in contingency tables represented by random variables, and divided by a finite number of levels (BEH, 2008).

Tables from distinct harvests could be compared using a split-plot design, including harvests in plots. Naturally, it is understood that the variation among subplots smaller than that among plots is desirable. Thus, the factor that will be randomized in the subplots will supposedly require less amount of experimental material (HINKELMANN; KEMPTHORNE, 2005). However, it is not feasible to use techniques based on analysis of variance, not only due to the normality assumptions required for the experimental error, but also for not exploring the dependence interrelationships among categorical variables and their effects on size reduction. In other words, the inference through tests of multiple comparisons of means does not show results that allow concluding on the correlation among the response variables.

Another possibility would be the use of a more complex model that considers the non-continuous data and the covariance structure associated to the measures repeated in time. Diggle *et al.* (2002) suggest the use of marginal models, e.g., estimation of generalized equations, which in synthesis is a multivariate approach of the marginal models of quasi-likelihood. However, it is known that the use of these models does not contemplate the dimensional reduction.

Alternative option would be to use a principal component analysis assuming only one of the factors, such as defect type, without considering the harvests, since an analysis that considers all the factors, according to Konishi (2015), is impracticable because the scores would be estimated independently, disregarding the variation "among" and "within" the levels.

The isolated analysis of each harvest would provide only the individual profile, however, the singular value decomposition(SVD) adapted for a simultaneous analysis allows a similarity and/or dissimilarity being found also between the harvests through a graphical representation easy to interpret, using the biplots technique (GREENACRE, 2003), unlike what happens in the analysis of multiple factors (OSSANI *et al.,* 2017).

The aim of this study was to propose an alternative statistical method to analyze simultaneously the classification of defects in different percentages of large coffee beans, considering samples from different harvests.

MATERIAL AND METHODS

Data on coffee bean defects were obtained from a classification of 40 samples from producers of municipalities in the south of Minas Gerais in the 2014 and 2015 harvests from the Catuaí variety. Samples were selected in order to obtain a similar amount of each harvest, excluding those identified as outliers.

The classification of coffee bean size was done in a sample of 300 g from each producer, using sieves interspersed for those of flat beans from numbers 14 to 18 and mochas from numbers 9 to 13. The retention percentage of each sieve was evaluated individually and the sum of sieves 17 and above were attributed to flat (large flat) beans (BRASIL, 2003).

The counting of defects in coffee beans was performed using the COB classification (BRASIL, 2003) in order to determine the number of imperfect beans and impurities found in the sample.

The large flat beans from the retention in circular sieve P17/18 sieves, with higher commercial value, constituted the double entry contingency table for each harvest, considering the defect count and the percentage of large flat beans present in the sample, thus characterizing the two studied variables.

For this study, the defects were grouped into five categories: among the defects related to the cultivation were considered the insect damage, since they come from the pest attack (coffee berry borer) that can cause great damages to the coffee grower. Among the possible defects originated from the harvest were the greens, which come from premature harvest, and the full or partly black that are from delayed harvesting and prolonged fermentations, grouping these counts in a category called PVA, because some companies adopt the criteria for evaluating full or partly black and green beans (BANDEIRA *et al*., 2009), whose analysis may indicate the gain or discount in profitability of the analyzed batch.

Among the processing defects were considered broken/immatures, coming from inadequate drought and badly regulated huller, and husks, caused by genetic factors or by possible physiological causes. Furthermore, extrinsic defects or impurities were considered.

Regarding the percentage of bean in P17/18 sieves in the sample, the data were grouped into three categories: (i) lower than 20% (ii), from 20 to 30%, and (iii) higher than 30%.

Thus, there were two dimension contingency tables: five defects × three percentages were generated referring to the samples of the 2014 (S14) and 2015 (S15) harvests. Each cell was represented by a value corresponding to the total from a specific defect in the corresponding percentage of beans in P17/18 sieves.

In order to obtain the biplots for the defects and proportions of beans in the P17/18 sieve in the 2014 (S14) and 2015 (S15) harvests, centralization was initially performed in relation to the general average (LIU *et al*., 2014). Thus, each table (matrix) containing the transformed data Z_{5x3}, corresponding to the five categories of defects and the three categories of bean percentage in P17/18 sieves. Biplots were obtained using the SVD of Z, given by:

$$Z = U\Gamma V^T,$$

where: U and V are singular right and left vector matrices, each with *r* orthonormal columns, and Γ is the diagonal matrix of positive singular value γ in decreasing order of magnitude. When calculating the matrix $\hat{Z}_{nxp}$ using the first two singular values and corresponding singular vectors, it has

$$\hat{Z} = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix} \begin{pmatrix} v_1 & v_2 \end{pmatrix}^T$$

Then, $\hat{Z}$ is the approximate least squares matrix of Z. The biplot will be as accurate as the approximation of $\hat{Z}$ for Z (AGRESTI *et al*., 2008).

Following recommendations by Greenacre (2003), aiming of extracting the components that relate the joint influence of variables and differences between the S14 and S15 crops, simultaneously represented by S14 + S15, e S14 - S15, the SVD was performed considering the block matrix given by:

$$M = \begin{bmatrix} S14 & S15 \\ S15 & S14 \end{bmatrix} \tag{1}$$

The values of each cell were corrected by the average $\bar{c} = 0.5 * (\bar{\gamma}_{S14} + \bar{\gamma}_{S15})$ and a single SVD applied at M to simultaneously provide the components due to the union S14 + S15 and the difference S14 - S15, considering (2).

$$S14 + S15 = UD_\alpha V^T \text{ and } S14 - S15 = XD_\beta Y^T, \tag{2}$$

where U and X are left singular arrays of vectors and V and Y are matrices of right singular vectors, each with k orthonormal columns, where, D_α and D_β represent the diagonal matrices of positive singular values γ in descending order of magnitude.

It has $U^T U = V^T V = I$ and $X^T X = Y^T Y = I$, then the SVD of a block matrix of dimension 2i x 2j is given by:

$$\begin{bmatrix} S14 & S15 \\ S15 & S14 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} U & X \\ U & -X \end{bmatrix} \begin{bmatrix} D_\alpha & 0 \\ 0 & D_\beta \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} V & Y \\ V & -Y \end{bmatrix}^T \qquad (3)$$

It is important to emphasize that the singular vectors to the left and right of equation (3) are all orthogonal to each other due to the orthogonality of vectors in the SVDs (2) and the change in the signal of the singular vectors in X and Y corresponds to the difference matrix. The presence of the $1/\sqrt{2}$ multiplying the singular vectors on the left and right ensures correct normalization $\frac{1}{\sqrt{2}}\begin{bmatrix} U \\ U \end{bmatrix}^T \frac{1}{\sqrt{2}}\begin{bmatrix} U \\ U \end{bmatrix} = \frac{1}{2}U^TU + \frac{1}{2}U^TU = I$. Thus, the left and right singular vectors of equation (3) are orthonormal (GREENACRE, 2003).

SVDs of the sum S14 + S15 and the difference S14 - S15 are not separated, but interspersed according to the magnitude of the corresponding singular values, which are arranged in descending order. In the SVD of the block matrix, it is easy to distinguish the solution vectors corresponding to the sum and difference: singular vectors to the left and right corresponding to the sum have two identical copies of a vector, being these grouped in the same column, whereas in the corresponding singular vectors corresponding to difference, the vector grouped to the initial vector has opposite signals to it.

The sum of the squares of elements from the block M matrix, thus constructed, can be decomposed into two components: one component due to the sum S14 + S15, and one due to the difference S14 - S15.

$$2\sum_i \sum_j (s14)_{ij}^2 + 2\sum_i \sum_j (s15)_{ij}^2 = \sum_i \sum_j \left((s14)_{ij} + (s15)_{ij}\right)^2$$

After identified the vectors associated with the components, the coordinates for the construction of the biplots S14 + S15 and S14 - S15 were obtained. The contribution of each eigenvalue is calculated in order to not consider the total n generated eigenvalues, but only n/2 eigenvalues, since half of them refer to the union and the other half to the difference among harvests.

The results were obtained and the graphs were elaborated by a script given in the software R (R Core Team, 2016).

RESULTS AND DISCUSSION

The simultaneous contingency table representing the bean count in relation to the categorical levels described in the defect types and percentage of sieve beans with the respective counts to be used in the application of SVD is described in Table 1.

Preliminarily, the data analysis (Table 1) shows that the amount of impurities, i.e., defects of an extrinsic nature, was low, evidencing a minimum of quality among the producers participating in the evaluation.

For the data referring to the 2014 harvest, the total of defective beans observed in the 2014 harvest was higher than the quantity of beans referring to the 2015 harvest. However, there was an increase in the number of insect damage beans, which is harmful because it deals with a crop problem that requires more effective pest management. Possibly, this result derives from carelessness with post-harvest coffee or on the borer control. It is noted that the number of husks indicating genetic or physiological problems and the number of impurities, which is detrimental to the product's commercial value because it reduces its classification, was also increased (Table 1).

Table 1 - Defect counting as a function of the percentage of sieved beans (17/18), observed in a total of 24 batch samples from the southern region of Minas Gerais, Brazil, specific to the harvest samples of 2014 and 2015

| Harvest	Defect Types	Percentage (p) of sieved beans (17/18)			
		p<20%	20≤p≤30%	p>30%	Total
2014	Insect Damage	187	223	151	561
	Full or partly black/green(FPBG)	744	916	759	2419
	Broken	588	498	465	1551
	Husks	127	252	156	535
	Impurities	12	8	12	32
2015	Insect Damage	264	190	274	728
	Full or partly black/green(FPBG)	474	552	725	1751
	Broken	324	430	262	1016
	Husks	210	293	382	885
	Impurities	12	0	53	65

The most frequent defect in both the 2014 and 2015 harvests was full or partly black and green beans, followed by broken beans. This is corroborated by study performed by Custódio, Gomes and Lima (2007) that, among the classes of defects, found the green and full black beans in higher percentages for all studied harvests and slides. On the other hand, Pedroso *et al.* (2009) made the commercial classification, with type and number of defects for different coffee batches and found higher incidence of defects, such as insect damage, husks, green, broken, immature, being the broken in greater proportion in the different batches. The associations between the percentage of large flat beans and the incidence of defects in the samples were studied, initially and individually, for each harvest (Figures 1 and 2).

In the 2014 harvest (Figure 1), beans classified as broken and PVA are more associated with large flat beans, being the incidence of the first highest in beans and the other in large beans, corroborating the results obtained by Pimenta and Vilela (2002) when they affirmed that the harvested green coffees showed higher number of defects, hard drink, being rejected for the commercialization.

For the 2015 harvest, it is noted that the vector indicated in the direction of the category of percentage of large flat beans greater than 30% is more associated with defects from the PVA category.

The complete results of the simultaneous SVD of the 10x6 block matrix for all six dimensions are given in Table 2.

Figure 1 - Biplot for association of defect types of coffee beans and defect ratio in 17/18 sieves for the 2014 harvest
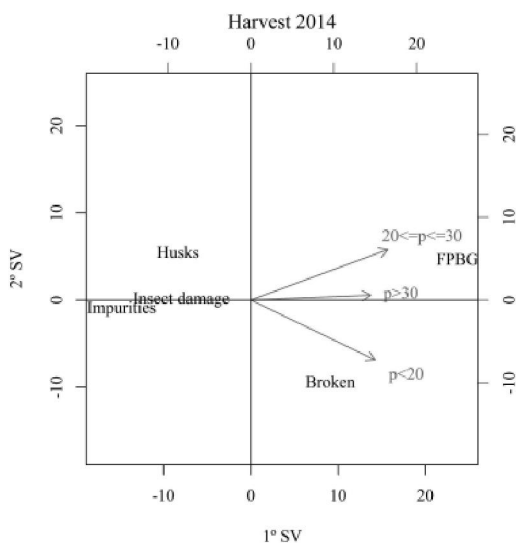
Figure 2 - Biplot for association of defect types of coffee beans and defect ratio in 17/18 sieves for the 2015 harvest
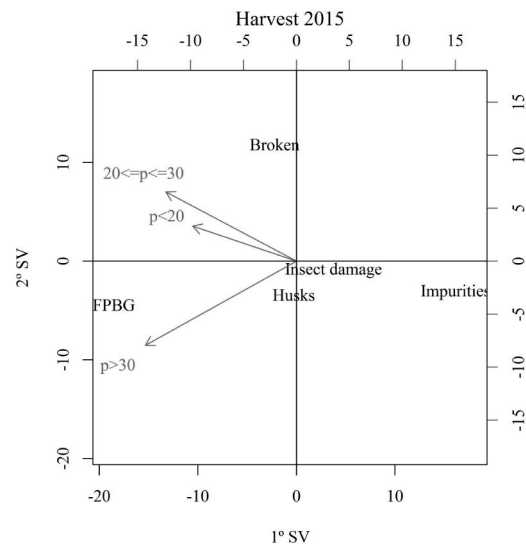
Note the block shape of the singular vectors, as given by equation (3). By checking the signal patterns only of singular vectors, left or right, and marking the signal from the upper block along as positive vectors (v), the six vectors are arranged in the following pattern of sub-vectors:

$$\begin{array}{c} \textit{Dimension} \\ \text{Harvest 2014} \\ \text{Harvest 2015} \end{array} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ \left[\begin{array}{cccccc} v & v & v & v & v & v \\ v & -v & -v & v & -v & v \end{array}\right] \end{array}$$

Thus, the dimensions of 1, 4 and 6 correspond to the matrix component sum (or union of harvests) and dimensions 2, 3 and 5 correspond to the difference matrix of harvests.

Once identified the eigenvectors for S14 + S15 and S14 - S15, it is possible to reproduce the respective covariance matrices.

The percentage of explained variability and the singular valuesobtained in each of the previously evaluated cases are presented in Table 3, in which it is possible to observe that the percentage of sample variation explained by the biplots showed values considered adequate.

Regarding the evaluation of the general profile of the number of coffee defects in relation to the harvests, a total sum of squares equal to 3,637,365 was obtained, calculated similarly to that used in the experimental approach; from this total, 3,231,715 refer to the sum of harvests S14 +S15. In the same way, the sum of squares of the difference among the harvests corresponds to 405,650.

Table 2 - Simultaneous SVD of the 10 x 6 block matrix of the data for 2014 and 2015 harvests

Dimension		1	2	3	4	5	6
Singular values		1779.2	552.8	302.5	243.9	92.5	81.7
Left singular vectors							
2014	Insect Damage	-0.144	0.111	-0.281	-0.135	-0.101	-0.377
	FPBG	0.521	-0.529	-0.440	0.260	-0.105	-0.203
	Broken	0.149	-0.396	0.301	-0.545	0.423	0.259
	Husks	-0.109	0.226	-0.359	0.330	0.535	0.475
	Impurities	-0.417	0.016	-0.086	0.090	0.117	-0.155
2015	Insect Damage	-0.144	-0.111	0.281	-0.135	0.101	-0.377
	FPBG	0.521	0.529	0.440	0.260	0.105	-0.203
	Broken	0.149	0.396	-0.301	-0.545	-0.423	0.259
	Husks	-0.109	-0.226	0.359	0.330	-0.535	0.475
	Impurities	-0.417	-0.016	0.086	0.090	-0.117	-0.155
Right singular vectors							
2014	p<20	0.370	-0.497	0.040	-0.527	0.502	-0.293
	20p≤p≤30	0.439	-0.407	-0.447	0.000	-0.367	0.554
	p>30	0.413	-0.296	0.547	0.472	-0.337	-0.327
2015	p<20	0.370	0.497	-0.040	-0.527	-0.502	-0.293
	20≤p≤30	0.439	0.407	0.447	0.000	0.367	0.554
	p>30	0.413	0.296	-0.547	0.472	0.337	-0.327

Table 3 - Singular values obtained in each of the cases evaluated separately

Data set	Singular values (sv)			Contribution to the first two axes (%)
	1°.	2°.	3°.	
Harvest 2014	1106.5	140.7	26.7	97.90
Harvest 2015	706.2	182.5	83.3	91.43
Union of harvests	1779.2	243.9	81.7	96.12
Difference of harvests	552.8	302.5	92.5	90.24

Given that the sum of eigenvalues is a measure of generalized variance, it is possible to verify that the sum of squares of the unified harvests is similar to the sum of squares of the singular values, i.e., $1779.2^2 + 243.9^2 + 81.7^2 = 3,231,715$. The same result is obtained when the sum of squares of the difference among harvests and sum of squares of the singular values obtained is verified: $[(552.8)^2 + (302.5)^2 + (92.5)^2] = 405,650$.

Depending on the obtained data, the biplots for the sum S14 + S15 and difference S14 - S15 of harvests were built.

Figures 3 and 4 correspond to the biplots using the structure of simultaneous contingency table, which allowed studying the associations considering the total variation of harvests and the variation of difference among the harvests.

Concerning the similarity of coordinates observed in the simultaneous biplot with the individual biplots, the dispersion of the categories of defects among the quadrants is shown, there is statistical evidence to affirm that the general profile was more pronounced by the incidence of defects observed in beans from the 2014 harvest. Nevertheless, the difference among harvests was more influenced by defects found in the 2015 harvest. However, the simultaneous biplot for the difference S14 - S15 showed an asymmetry supposedly caused in the category "Full or partly black/green- FPBG" observed in the individual biplots.

Rev. Ciênc. Agron., v. 49, n. 1, p. 62-69, jan-mar, 2018

67

Figure 3 - Biplot for association of defect types of coffee beans and defect ratio in 17/18 sieves for the unified harvests
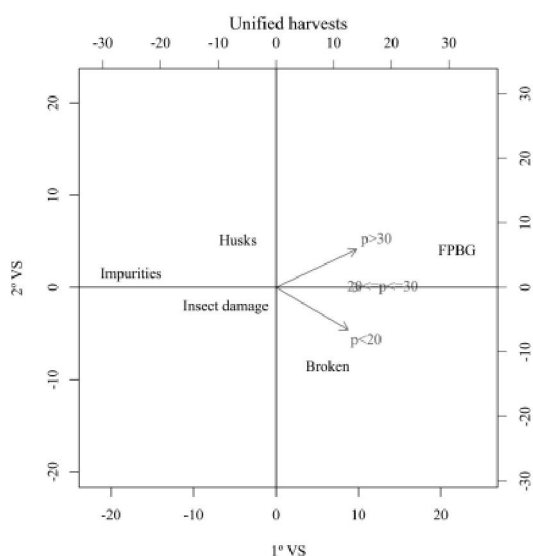
Figure 4 - Biplot for association of defect types of coffee beans and defect ratio in 17/18 sieves for the difference among harvests
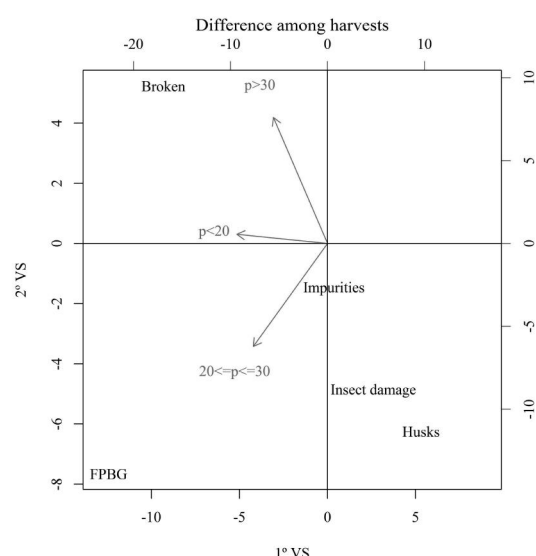
Several authors emphasize that coffee should be harvested at their optimum ripeness (cherry), because when harvested immature or dried on the plant can cause incidence of green, full or partly black beans, resulting in the worst defects for the coffee quality. The presence of 15% of green beans in the mixture varies the classification, from superior drink to the less acceptable drink; and from 60%, the drink is rated as very bad.

Another highlight refers to the length of vectors corresponding to the categories of large bean percentages. Being these with similar length, it can be stated that, in the two studied harvests, even in the simultaneous evaluation, there is evidence of similar degrees of influence among the used percentages. Therefore, regarding the association of vectors with the defect types, the simultaneous biplot did not show difference. The possible cause of intrinsic defects, according to Bártholo and Guimarães (1997), may be related to cultural practices and the culture physiology. Nunes *et al.* (2007) report that even in regions suitable for coffee cultivation, since it is considered a perennial crop, adverse weather conditions throughout the year, whether as rainfall, temperature variation and relative humidity, during the flowering, fruiting and ripening phases, can cause very uneven ripening. This causes a high percentage of green fruits in the harvest and undesirable fermentation in mature fruits, resulting in quality loss even before harvest.

It should be noted that in the 2015 harvest, corresponding to the 2014/15 crop year, there were droughts in some municipalities of the South of Minas Gerais, Brazil, damaging crops and their fruits, resulting in immature and small-sized beans. At that time, the market paid goodwill for the coffees with the highest ratio of large coffees (CONAB, 2015).

Given the little differences in the groups between the individual and the simultaneous biplots, it is suggested that some possible classification error of beans may have been occurred, so the biplot is an important instrument described to aid in the inspection of beans classified in the categories. Thus, Figure 3 allowed identifying the general behavior of defects, regardless of the differences among harvests. Impurities, husks and insect damage separated on the left side indicate a lower influence of these on the general profile of defects, regardless of the number of large flat beans.

Figure 4 illustrates the map of harvest differences. The greatest differences between the 2014 and 2015 harvests are for broken and PVA beans, since these defects are further away from the centroid. Note that the vectors are now in different positions in the biplot: all negative in the vertical axis, and contrasting the defect category broken mainly against the PVA defects, which come from the late or early harvest, i.e., the biplot identified a marked difference among harvests in the defects originating from harvest of the processing defect. It can also be noted that the "broken" defect is associated more with samples with higher percentage of large flat beans (p>30), indicating poor huller regulation or inadequate drying.

CONCLUSION

The statistical technique applied in this study to identify similarity and differences among harvests is efficient, evidencing the defect categories most incident in the two harvests. In sieve classification, the defect factor is more associated with the categories full or partly black, green and broken beans.

ACKNOWLEDGEMENTS

REFERENCES

ABRAHÃO, A. A. *et al.* Classificação física e composição química do café submetido a diferentes tratamentos fungicidas. **Coffee Science**, v. 4, n. 2, p. 100-109, 2009.

AGRESTI, A. *et al.* Simultaneous confidence intervals for comparing binomial parameters. **Biometrics**, v. 64, n. 4, p. 1270-1275, 2008.

BANDEIRA, R. D. C. C. *et al.* Composição volátil dos defeitos intrínsecos do café por CG/EM-headspace. **Química Nova**, v. 32, n. 2, p. 309-314, 2009.

BÁRTHOLO, G. F.; GUIMARÃES, P. T. G. Cuidados na colheita e preparo do café. **Informe Agropecuário**, v. 18, n. 187, p. 33-42, 1997.

BEH, E. J. Correspondence analysis of aggregate data: the 2x2 table. **Journal of Statistical Planning and Inference**, v. 138, n. 10, p. 2941-2952, 2008.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Instrução Normativa n. 8, de 11 de junho de 2003. **Diário Oficial da República Federativa do Brasil**, Brasília, 20 ago. 2003. Seção 1, p. 22-29.

COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira**: café. v. 1, n. 1, 2015. Available from: <http://www.conab.gov.br>. Acesso em: 12 jul. 2016.

CORADI, P. C.; BOREM, F. M.; OLIVEIRA, J. A. Qualidade do café natural e despolpado após diferentes tipos de secagem e armazenamento. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 12, n. 2, p. 181-188, 2008.

CUSTÓDIO, A. A. P.; GOMES, N. M.; LIMA, L. A. Efeito da irrigação sobre a classificação do café. **Engenharia Agrícola**, v. 27, n. 3, p. 391-701, 2007.

DIGGLE, P. *et al.* **Analysis of longitudinal data**. 2nd edition. Oxford: Oxford University Press, 2002. (Oxford Statistical Science Series, 25).

FERNANDES, S. M. *et al.* Constituintes químicos e teor de extrato aquoso de cafés arábica (*Coffea arabica* L.) e conilon (*Coffea canephora pierre*) torrados. **Ciência e Agrotecnologia**, v. 27, n. 5, p. 1076-1081, 2003.

GREENACRE, M. Singular value decomposition of matched matrices. **Journal of Applied Statistics**, v. 30, n. 10, 1101-1113, 2003.

HINKELMANN, K.; KEMPTHORNE, O. **Design and analysis of experiments**: advanced experiment design. New York: John Willey, 2005. v. 2.

KONISHI, T. Principal component analysis for designed experiments. **BMC Bioinformatics**, v. 16, n. 18, p. 1-9, 2015.

LIMA, L. A.; CUSTÓDIO, A. A. P.; GOMES, N. M. Produtividade e rendimento do cafeeiro nas cinco primeiras safras irrigado por pivô central em Lavras, MG. **Ciência e Agrotecnologia**, v. 32, n. 6, p. 1832-1842, 2008.

LIU, F. *et al.* Singular value decomposition based minutiae matching method for finger vein recognition. **Neurocomputing**, v. 145, n. 1, p. 75-89, 2014.

NUNES, E. L. *et al.* Zoneamento agroclimático da cultura do café para a bacia do rio doce. **Revista Brasileira de Meteorologia**, v. 22, n. 3, p. 297-302, 2007.

OSSANI, P. C. *et al.* Qualidade de cafés especiais: uma avaliação sensorial feita com consumidores utilizando a técnica MFACT. **Revista Ciência Agronômica**, v. 48, n. 1, p. 92-100, 2017.

PEDROSO, T. Q. *et al.* Qualidade de sementes de cafeeiro produzidas em diferentes densidades de plantio e regimes hídricos. **Coffee Science**, v. 4, n. 2, p. 155-164, 2009.

PIMENTA, C. J.; VILELA, E. R. Qualidade do café (*Coffea arabica* L.) colhido em sete épocas diferentes na região de Lavras-MG. **Ciência e Agrotecnologia**, v. 26, p. 1481-1491, 2002. Número especial.

R CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2016. Available from: <http://www.r-project.org>. Acesso em: 19 out. 2016.

RAMOS, M. F. et al. Discrimination of the sensory quality of the *Coffea arabica* L. (cv. Yellow Bourbon) produced in different altitudes using decision trees obtained by the CHAID method. **Journal of the Science of Food and Agriculture**, v. 96, n. 10, p. 3543-3551, 2016.

WACHHOLZ, L.; POYER, M. G. A importância das cooperativas cafeicultoras para os pequenos agricultores na exportação de café na região sul de Minas Gerais. **Revista Gestão & Sustentabilidade Ambiental**, v. 2, n. 2, p. 27-44, 2014.