Acta Scientiarum
http://periodicos.uem.br/ojs/acta
ISSN on-line: 1807-8621
Doi: 10.4025/actasciagron.v41i1.39323

Statistical procedure for the composition of a sensory panel of blends of coffee with different qualities using the distribution of the extremes of the highest scores

Marcelo Ângelo Cirillo[1*], Mariana Figueira Ramos[1], Flávio Meira Borém[2], Felipe Mesquita de Miranda[2], Diego Egídio Ribeiro[2] and Fortunato Silva de Menezes[3]

[1]Departamento de Estatística, Universidade Federal de Lavras, Campos Universitário, Cx. Postal 3037, 37200-000, Lavras, Minas Gerais, Brazil. [2]Laboratório de Processamento de Produtos Agrícolas, Departamento de Engenharia, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil. [3]Departamento de Física, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil. *Author for correspondence. E-mail: macufla@des.ufla.br

ABSTRACT. The identification and interpretation of discrepant observations in sensory experiments are difficult to implement since the external effects are associated with the individual consumer. This fact becomes more relevant in experiments that involve blends, which scrutinize coffees with different qualities, varieties, origins, and forms of processing and preparation. This work proposes a statistical procedure that facilitates the identification of *outliers* while also evaluating the discriminatory powers of a sensory panel concerning the differentiation of pure blends and coffees. For this purpose, four experiments were performed that tested coffees with different qualities and varieties. The results suggest that the statistical procedure proposed in this work was effective for discriminating the blends relative to the pure coffees and that the effects of the concentrations and types of processing did not interfere with the statistical evaluations.

Keywords: canephora; arabica; mixture; models.

Received on September 5, 2017.
Accepted on December 14, 2017.

Introduction

The formation of numerous blends made from mixtures of coffees from, for example, the Arabica and Canephora species, with different concentrations is often carried out by roasters. Thus, diverse coffees can be generated and sold as new products. The differentiation of these products is undertaken by one of the programmes of the ABIC (Brazilian Coffee Industry Association); the organization distributes certifications for new blends to expand coffee consumption.

Given this perspective, the acceptance of the consumer of the produced coffee blends. Therefore, sensory experiments and acceptance tests are widely carried out to evaluate the quality of, the degree of acceptance of and the preference for a new particular blend. Alves et al. (2017) indicate that the coffee quality depends on many factors, which range from species choice, to crop variety and method of preparation of the coffee.

In carrying out these experiments, Lim (2011) described the potential errors that may arise from a sensory analysis, which are defined as follows: (a) errors of expectation, which occur when the taster has some previous knowledge of the coffee preparation and/or samples and (b) stimulus errors, which occur when samples are not presented in a consistent manner. In some situations, suggestion can also produce an effect, such that the taster is influenced by the facial expressions of other tasters. A higher frequency of these effects is found when the tasters have taken part in too few experiments or are not trained for tasting. Faced with these errors and undesirable effects, the results of a sensory experiment may be affected by observations classified as *outliers*. The inappropriate treatment of these observations results in imprecise and incoherent analyses. In most cases, the use of statistical methods and/or models, which are typically used to obtain results, is not possible.

Given this information, the researcher is faced with several strategies, such as the use of data transformations that can stabilize the data so that an adequate statistical analysis can be made. In this context, new statistical methodologies that are robust or sensitive to outliers can be used so that trained and untrained consumers can contribute to coffee quality studies. There are several studies that have considered this concept.

Liska et al. (2015) evaluated an experiment that tested four types of specialty coffees. The factors that were considered included sensory attributes such as aroma, body, sweetness and final score. By using a method called *boosting*, which is applied to the discriminant analysis, they concluded that the higher success rate observed for tasters after introductory training was approximately 80.63%, and there was a reduction in the rate of false negatives of approximately 19.37%.

Ossani, Cirillo, Borém, Ribeiro, and Cortez (2017) proposed the use of a multifactor technique applied to contingency tables (MFACT), which display the categorized data obtained from a sensory experiment conducted with different groups of consumers. This was done to identify similarities between four specialty coffees. They concluded that the use of this technique is feasible because it enables discrimination between the specialty coffees produced in different environments (altitudes) and processed using different methods and considers the heterogeneities of the consumers involved in the sensory analysis.

For the same experiments considering non-expert tasters, Ferreira et al. (2016) proposed a model using the distribution of extreme values, to identify the presence of outliers. The authors concluded that untrained consumers are not able to discriminate between specialty coffees, per the probabilities obtained.

In the case of the sensory analysis of specialty coffees with highly qualified tasters, Ramos, Ribeiro, Cirillo, and Borém (2016) used decision trees build via the CHAID approach. They concluded that the methodology used presented promising results regarding the accuracy and success rates when discriminating between samples of Arabica coffees. The sensory evaluations of the coffee had scores equal to or greater than 88 points, showing that the coffees produced at altitudes equal to or greater than 1.200 m could be discriminated in terms of high and low body intensities.

An example of an experiment that adds more complexity to the sensory analyses for both trained and untrained tasters is exemplified by a test of blends of coffees of different qualities and varieties in which the flavors are changed by the extent of roasting. This leads to a confounding variable and influences the tastings of the samples.

For example, in one study of blends, Ribeiro et al. (2014) state that the chemical compositions of the coffee beans are responsible for the formation of compounds linked to the flavours and aromas that the coffee presents during its tasting. The *C. arabica* species is known to present a balance between the desirable chemical compounds and is determined to have a pattern of higher quality than the *C. canephora* species.

It is true that there is still criticism of the formulation of blends of the Arabica and Canephora coffees in terms of the compositions of these coffees. In most situations, the formulations of these blends are made arbitrarily, resulting in products with different qualities, which risks compromising product acceptance.

Reporting the errors of the expectations, stimuli and the effects of suggestions mentioned above for a sensory analysis of the blends of coffees of different varieties using conventional statistical techniques based on an analysis of variance is inappropriate. This is because statistical experiments consider that all consumers are homogeneous and ignore their individual abilities. Moreover, the components used in the compositions of the blends are independent. In this context, the usual procedures to identify *outliers* are limited by these assumptions and, in most cases, do not provide an interpretation of the observations classified as *outliers.*

With the motivation of deriving new statistical methodologies to analyse the results of sensory experiments of blends, this work presents a new statistical procedure that can identify outliers and can serve as an instrument to evaluate the individuals that compose a sensory panel to differentiate and characterize the quality of blends in relation to pure coffees.

Material and methods

Description of the experiment used to study the acceptability of tasters in relation to the blends For the formulation of the blends, coffees with different quality standards were used. We considered the proportions of the following coffees: roasted and ground commercial Arabica (CAC), Canephora (CC) and specialty Arabica (SAC). Thus, four experiments were performed with the following differentiations (Table 1).

Table 1. Characterization of the compositions of the blends used in the experiments.

Experiment	Description of the specialty coffee used in the experiment
1	Yellow Bourbon processed naturally Concentration: 7% m/v (35 g/500 ml).
2	Yellow Bourbon processed naturally Concentration: 10% m/v (35 g/500 ml).
3	Yellow Bourbon processed via peeled cherry Concentration: 7% m/v (35 g/500 ml).
4	Yellow Bourbon processed via peeled cherry Concentration: 10% m/v (35 g/500 ml).

Given these concentrations, the preparations of the samples were conducted using drinking water at 93°C and no added sugar. The extraction time was 4 minutes, using a filtration preparation method. In this way, any risks related to allergic reactions or increases in glucose rates in the evaluating individuals were avoided. The preparations respected the hygiene standards imposed by the ethics committee under the protocol CAAE: 14959413.1.0000.5148.

To infer the effects of the concentrations of the beverages, 7 and 10% m/v (35 g 500 mL^{-1}) concentrations were used in the experiments. The experiments were evaluated using the compositions listed in Table 2. Thus, the identification of the blends in the joint analysis follows the encoding of the samples (k = 1,...,9) and (k = 10,...,18), respectively. The blends were analysed in experiments 1 and 2. This coding is continued for experiments 3 and 4.

Table 2. Composition of the blends formed by the following coffees: roasted and ground commercial Arabica (CAC), specialty Arabica (SAC) and Canephora (CC).

Experiment 1				Experiment 2			
Sample	SAC	CAC	CC	Sample	SAC	CAC	CC
1	1.000	0.000	0.000	10	0.340	0.330	0.330
2	0.670	0.330	0.000	11	0.000	0.000	1.000
3	0.340	0.330	0.330	12	0.340	0.000	0.660
4	0.500	0.500	0.000	13	0.000	1.000	0.000
5	0.500	0.000	0.500	14	0.670	0.330	0.000
6	0.340	0.660	0.000	15	0.340	0.660	0.000
7	0.340	0.000	0.660	16	1.000	0.000	0.000
8	0.000	1.000	0.000	17	0.500	0.000	0.500
9	0.000	0.000	1.000	18	0.500	0.500	0.000
Experiment 3				Experiment 4			
Sample	SAC	CAC	CC	Sample	SAC	CAC	CC
1	0.500	0.000	0.500	10	0.500	0.000	0.500
2	0.340	0.330	0.330	11	0.500	0.500	0.000
3	1.000	0.000	0.000	12	0.340	0.000	0.660
4	0.500	0.500	0.000	13	0.340	0.330	0.330
5	0.340	0.000	0.660	14	0.670	0.330	0.000
6	0.340	0.660	0.000	15	0.000	0.000	1.000
7	0.000	0.000	1.000	16	1.000	0.000	0.000
8	0.000	1.000	0.000	17	0.000	1.000	0.000
9	0.670	0.330	0.000	18	0.340	0.660	0.000

Each experiment was conducted in separate sessions at 24-hour intervals, due to the excessive number of evaluations. The group of evaluators in the test experiment comprised five qualified tasters. They were considered suitably qualified to differentiate the samples in the sensory experiments due to their experience in tasting experiments.

Each evaluator tasted approximately 20 mL of the prepared blend formulations at a temperature of approximately 65°C and served in disposable cups. After tasting each blend, the evaluator reported his/her evaluation on the appropriate data sheets.

The blends were evaluated, receiving scores ranging from 0 to 10 for the qualitative characteristics of the drink: body, acidity, flavour, bitterness and final score. The characteristic aftertaste represented the overall impression of the quality described by the evaluator.

Statistical procedure to identify *outliers* to be used in evaluating the sensory panels in relation to the discrimination between the blends in relation to pure coffee

Since the results of sensory evaluations are subject to the occurrences of outliers, we proceeded our analysis with a statistical analysis used by Veloso and Cirillo (2016), which consists of two successive re-scalings of multivariate data.

The performance of this procedure was initially represented by data in a matrix form, with the definition of the matrix $G^{(0)}$ such that each element, $g_{ij}^{(0)}$, corresponded to the average score of each sensory attribute obtained in 5 repetitions, which are characterized by the repetitions of the tasters. Thus, the multivariate observations were represented by indices (i=1,...,t=18), indicating the samples of the blends evaluated in relation to the j-th sensory attribute (j=1,..., p=5).

According to this notation, the vectors of the observations for each sample, the vectors of observations and the respective variables were $G_i^{(0)}=[G_{1j}^{(0)},G_{2j}^{(0)},\ldots,G_{ip}^{(0)}]$ and $G_j^{(0)}=[G_{1i}^{(0)},G_{2i}^{(0)},\ldots,G_{tj}^{(0)}]$. We note that the representations of these vectors suggest that the data matrix is structured in the following way: the "lines" of the average scores of the samples (blends) in the "columns" follow the evaluated attributes. The median absolute deviation (MAD) was applied to the vectors and the resulting vector is (2.1).

$$\mathrm{MAD}\left(g_{1j}^{(0)},\ldots,g_{tj}^{(0)}\right) = 1.4826 \times \mathrm{med}_j\left|g_j^{(0)} - \mathrm{med}_i(g_i^{(1)}\right| \qquad (2.1)$$

The quantile of the normal univariate standard distribution (Veloso & Cirillo, 2016) was found to correspond to 1.4826 and was obtained via the expression $1/(\Phi^{-1}(3/4))$. The median (med) of the observations of each variable assumed to be (i = 1,...,t) and (j = 1,...p). Subsequently, the first rescaling in the data was carried out by updating the elements $g_{ij}^{(0)}$ with $g^{ij(1)}$ in accordance with expression (2.2). Thus, the $G^{(1)}$ matrix is as follows.

$$g_{ij}^{(1)} = \frac{g_{ij}^{(0)} - \mathrm{med}\left(g_{1j}^{(0)},\ldots,g_{tj}^{(0)}\right)}{\mathrm{MAD}\left(g_{1j}^{(0)},\ldots,g_{tj}^{(0)}\right)}\ ; j = 1,\ldots,p \text{ and } i = 1,2,\ldots,t \qquad (2.2)$$

Assuming the $G^{(1)}$ matrix with elements given by $g_{ij}^{(1)}$, the covariance matrix is calculated. Then, the matrix of the eigenvectors was computed and applied to V. Next, the $G^{(2)} = G^{(1)} \times V$ matrix was obtained and the second rescaling was applied, yielding the $G^{(3)}$ matrix. Again, each element $g_{ij}^{(3)}$ was obtained by considering the median deviation MAD (2.1).

$$g_{ij}^{(3)} = \frac{g_{ij}^{(2)} - \mathrm{med}\left(g_{1j}^{(2)},\ldots,g_{tj}^{(2)}\right)}{\mathrm{MAD}\left(g_{1j}^{(2)},\ldots,g_{tj}^{(2)}\right)}\ ; j = 1,\ldots,p \text{ and } i = 1,2,\ldots,t \qquad (2.3)$$

With the rescaled data, the linear mixing model was adjusted with the inclusion of a process variable represented by the beverage concentrations. The beverages were prepared at 7 and 10% (m/v) concentrations. Thus, the adjusted model for each sensory attribute of the body, taste, acidity, bitterness and score was determined via $\eta(\gamma, x, z)$ (2.4), where the terms $\eta(\beta, x)$ corresponded to the linear model and considered the components x_r (r = 1, 2, and 3) indicating x_1 = SAC, x_2=CAC, x_3 = CC and $\eta(\alpha, z)$. The terms representing the adjusted model include the intercept, α, and the z process variable. Both models are estimated using the least squares method, where $e_r \sim N(0,\sigma^2_r)$. Further details of the model composition $\mu(\gamma,x,z)$ can be found in Dal Bello and Vieira (2011).

$$\mu(\gamma, x, z) = \mu(\alpha, z) + \mu(\beta, x)\ \text{, in which, (2.4)}$$

$$\mu(\beta, x) = \sum_{r=1}^{3} \hat{\beta}_r x_r + e_r$$

$$\mu(\alpha, z) = \hat{\alpha} + \hat{\alpha}_1 z$$

Thus, based on the predicted values obtained by the adjustments of the model (2.4) for each sensory attribute, the identifications of the blends, which were interpreted as outliers, were performed according to the procedure defined by Resende, Brighenti, and Cirillo (2017). The considered statistical measures were defined by the accumulated distributions of the highest order and independent scores, described in (2.5). In the context of this work, μ_i follows a normal distribution; thus, the parametric model can obtain a cumulative distribution, which is provided by the normal model.

$H_{max}(\hat{\mu}_i) = [H(\hat{\mu}_i)]^t, i = 1, \ldots, t = 18$ being $\hat{\mu}_{(1)}, \hat{\mu}_{(2)}, \ldots, \hat{\mu}_{(t)}$ (2.5)

The graphical representation of the observations were analysed and the outliers were determined. This was done using the Mahalanobs distance. The covariance matrix shown in (2.6) was used as a reference because $H^*(\hat{\mu}_i)$ is associated with $H_{max}(\hat{\mu}_i)$, as shown in the criteria adopted.

$$S = \begin{bmatrix} var(\hat{\mu}_i) & cov(\hat{\mu}_i, H^*(\hat{\mu}_i)) \\ cov(\hat{\mu}_i, H^*(\hat{\mu}_i)) & var(H^*(\hat{\mu}_i)) \end{bmatrix} \quad (2.6)$$

Thus, the Mahalanobs distance for each observation was determined via D_h (2.7) plotting, i.e., D_h vs. $H^*(\hat{\mu}_i)$, given that D_h~chi-square distribution. The specification limit was defined by the quantile of a chi-square distribution with p = 2 degrees of freedom and 1-α confidence level. Thus, it became possible to identify the points located above these quantiles as blends, which represents dissonant sensory results. The results are described and illustrated below.

$$D_h = \sqrt{\left(\tilde{H}^*(\hat{\mu}_i) - \tilde{\hat{\mu}}_i\right)^T S^{-1} \left(\tilde{H}^*(\hat{\mu}_i) - \tilde{\hat{\mu}}_i\right)} \text{ in which}$$

$$\tilde{H}^*(\hat{\mu}_i) = (H^*(\hat{\mu}_1), \ldots, H^*(\hat{\mu}_t))^T \text{ and } \widetilde{\hat{\mu}_i} = (\hat{\mu}_1, \ldots, \hat{\mu}_t)^T, \quad (2.7)$$

In appendix A, the data layout and experiment that identifies the variables are used in the script (appendix B)

Results and discussion

Due to its application in the roasting industry, to improve the qualities of the blends, the processing method to be used should be considered since the aroma and flavour of the coffee beverage is influenced by this process and the quality of the beverage is affected (Barbosa et al., 2019). Therefore, in accordance with the proposed method and considering the adjustment of the linear model with a processing variable (Conc), the results of the estimates that use the Kronecker model (Nepomucena, Silva, & Cirillo, 2013; Brighenti, Brighenti, Cirillo, & Santos, 2010) are described in Section 3.1 for specialty coffees processed naturally and in Section 3.2 for specialty coffee processed dry.

Specialty coffees processed naturally

Based on a linear model that was adjusted for each sensory attribute and with the process variable (concentrations of 7 and 10%), the results in Table 3 correspond to the estimates of the parameters for the proportions of the specialty Arabica coffee (X_{SAC}), the roasted and ground commercial Arabica coffee (X_{CAC}) and the Canephora (X_{CC}) when using a linear model and considering the interactions of the variables of interest.

Table 3. Estimates of the linear model for experiments 1-2.

Variables	Body	Flavour	Acidity	Bitterness	Score
X_{SAC}	-6.406	0.264	-0.838	2.834	-0.261
X_{CAC}	0.907	-3.052	1.422	-0.863	4.238
X_{CC}	2.354	-1.666	4.592	0.159	11.386
X_{SAC}×Conc	52.264	-4.884	11.144	-32.072	7.445
X_{CAC}×Conc	7.794	44.063	-19.731	10.307	-45.462
X_{CC}×Conc	-23.252	13.942	-50.658	-1.222	-133.387

For each variable, the outlier analysis, which enables the discrimination of the blends from pure coffees, considered the predicted rescaled values and followed an algebraic procedure (2.1 - 2.3) for each sensory attribute. The cumulative distribution of the highest scores (2.5 - 2.7) were determined according to the procedure described in the expressions (2.5) - (2.7), which is shown in Figures 1 to 5. The horizontal line corresponds to a specification limit, obtained as a quantile of the chi-square distribution with a confidence level of 95%.

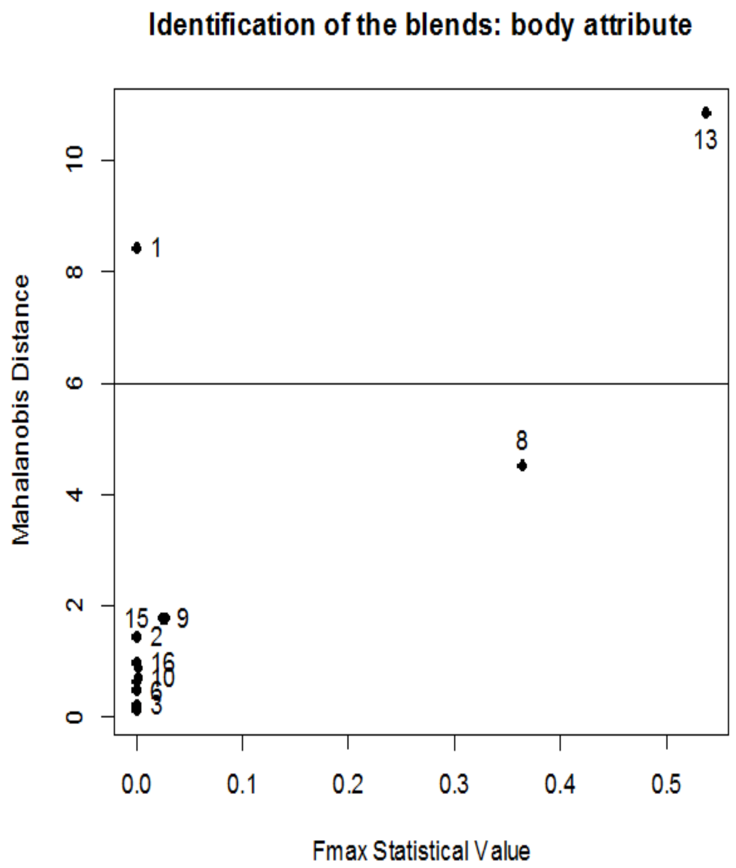

Figure 1. Identification of the dissonant blends of the sensory evaluations of the body variable for experiments 1-2.

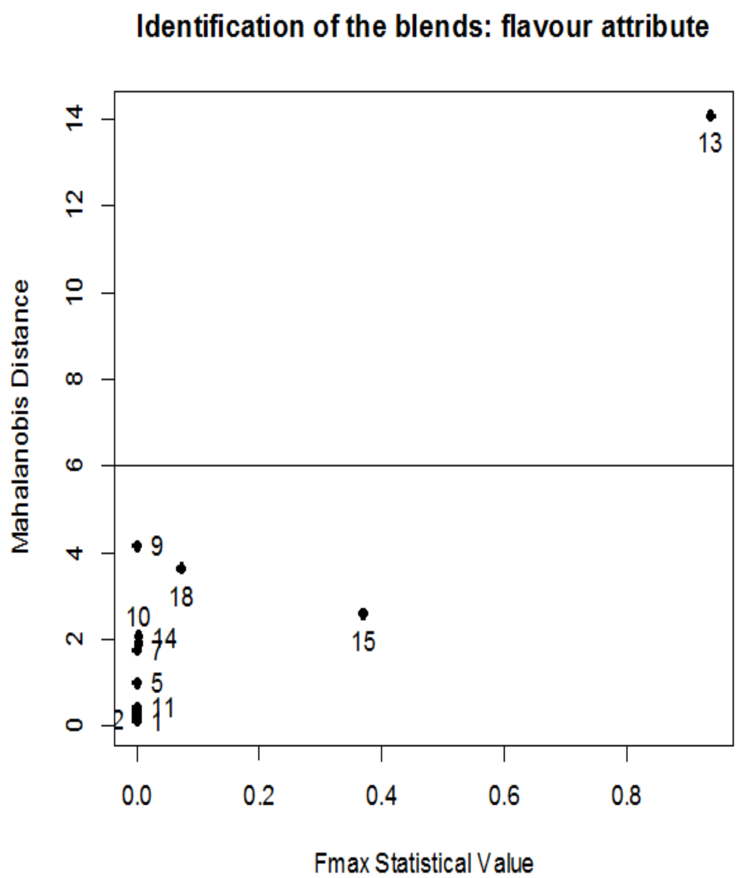

Figure 2. Identification of the dissonant blends of the sensory evaluations of the flavour variable for experiments 1-2.

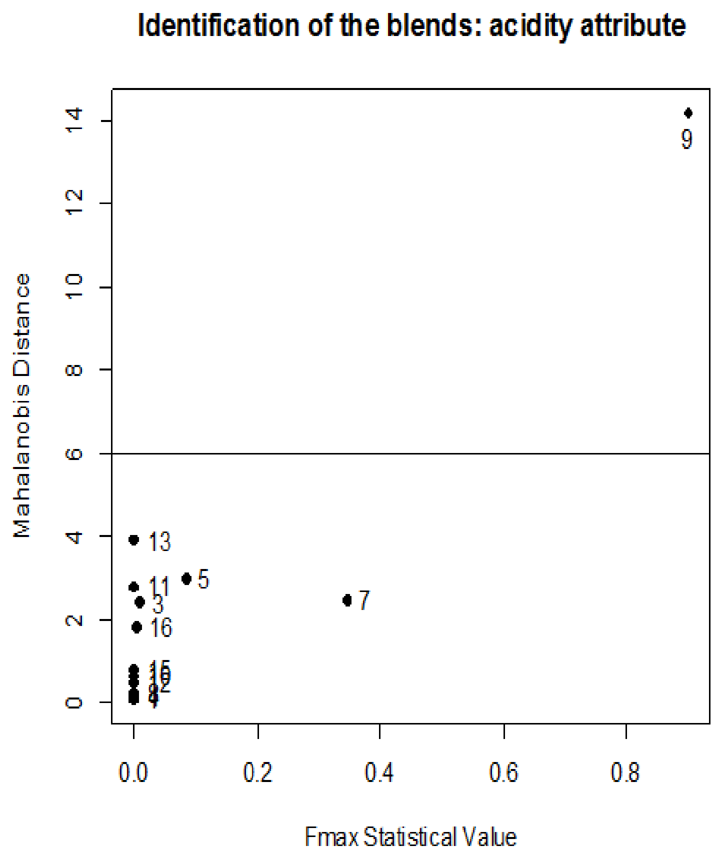

Figure 3. Identification of the dissonant blends of the sensory evaluations of variable acidity for experiments 1-2.

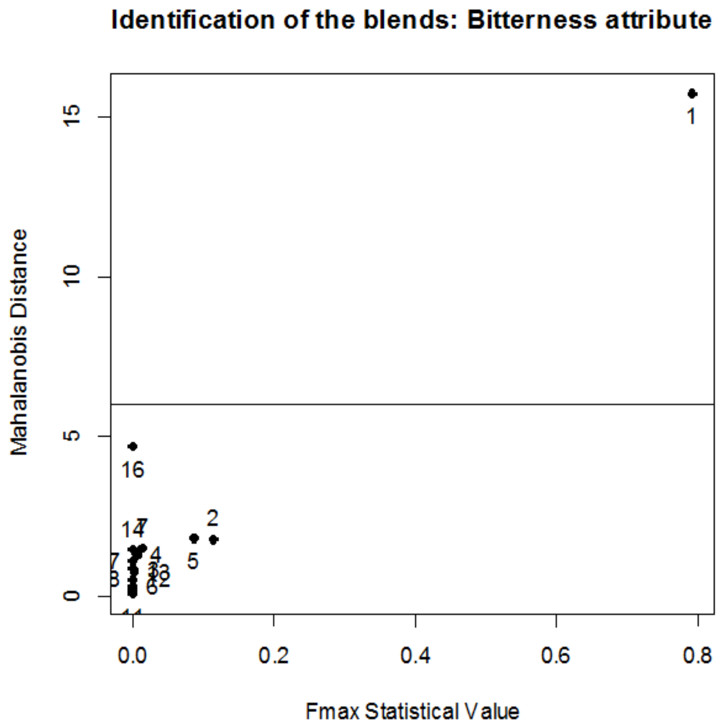

Figure 4. Identification of the dissonant blends of the sensory evaluations of the bitterness variables for experiments 1-2.

The general results shown in Figures 1 to 5 showed that the blends could be discriminated with respect to pure coffees, represented by samples No. 1, 9, and 13. In this context, and compared with the identifications given in Table 2, sample No. 9 (SAC = 0, CAC = 0, and CC = 1) showed equivalent results for Acidity and Final score.

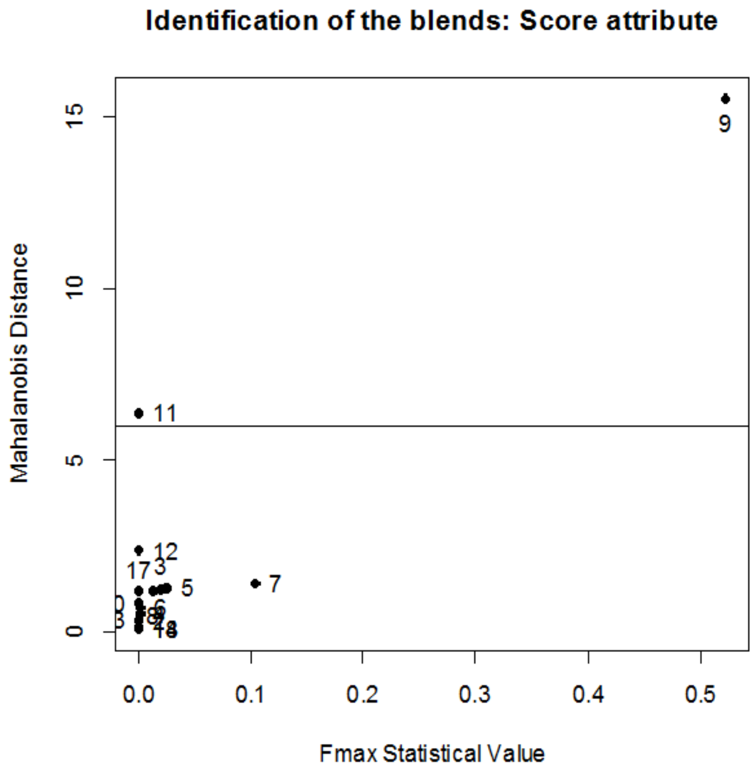

Figure 5. Identification of the dissonant blends of the sensory evaluations of the score variable for experiments 1-2.

In the case of the body and flavour attributes, samples No. 1 and 13, defined as (SAC = 1, CAC = 0, and CC = 0) and (SAC = 0, CAC = 1, and CC = 0), respectively, showed discordant results for Body and Flavour. The discrepancy for sample No. 1 was also observed for the bitterness attribute.

We note that the chemical composition of the blends were not evaluated in this study However, the compositions were noted in the literature to explain the differentiations of the blends characterized as "pure" in relation to the other samples. Alessandrini, Romani, Pinnavaia, and Rosa (2008) suggested that roasting causes "chemical, physical, structural and sensory changes in the beans that have a strong impacts on the qualities of the final product, which results in complex reactions that cause peculiar colours, aromas and flavours of the coffee".

Moura et al. (2007), evaluated the physical, chemical and sensory aspects of the blends of Arabica coffee with Robusta. They found that the caffeine content of the Robusta coffee was higher than that of Arabica, such that when Robusta is added to the blend; the caffeine content increased. Caffeine is odourless and has a very characteristic bitter taste, contributing to the significant bitter note of the flavour and the aroma of the coffee beverage (Monteiro & Trugo 2005).

Filho, Lucia, Saraiva, and Lima (2015) commented that some of the blends used by industries may be formulated using different proportions of a mixture of Arabica coffees and certain varieties of Canephora, such as Robusta. That Arabica coffee is more fruity and acidic, while Robusta coffee is more bitter and ensures a more full-bodied beverage.

Specialty coffee processed wet

Following the same specifications, the results described in Table 4 correspond to the estimates of the linear model parameters for the experiments. In this case, the specialty coffee was only processed when wet.

Table 4. Estimates of the linear model for experiments 3-4.

Variables	Body	Flavour	Acidity	Bitterness	Score
X_{SAC}	-4.586	-0.936	5.448	0.208	-6.279
X_{CAC}	-1.172	2.313	0.337	0.707	-1.429
X_{CC}	-0.329	-10.199	-2.238	2.074	-0.699
X_{SAC}×Conc	31.660	18.380	-64.534	3.210	79.441
X_{CAC}×Conc	35.485	-22.493	0.234	-14.430	13.250
X_{CC}×Conc	13.242	92.796	19.041	-26.197	4.698

Given the estimates of the model adjusted for each variable (Table 4), the breakdown of the blends of the wet-processed specialty coffee with the same concentrations of 7 and 10% m/v (35 g 500 mL^{-1}) is illustrated in Figures 6 to 10.

Note that, for all variables, the pure components of the blends, represented in samples No. 1, 3, and 7 (Experiment 3 - Table 2) and No. 16 and 17 (experiment 4 - Table 2) were identified against other blends. Thus, compared with the previous results (Figures 1 to 6), it is clear that the effect of the concentration and the type of processing did not affect the evaluations of the sensory abilities of the evaluators with respect to the discrimination between pure coffees and blends.

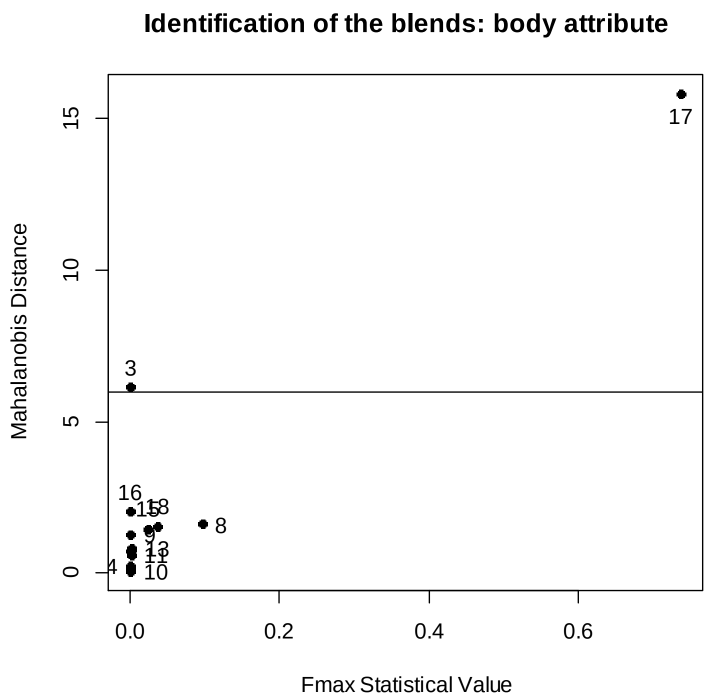

Figure 6. Identification of the dissonant blends of the sensory evaluations of the body variables of experiments 3-4.

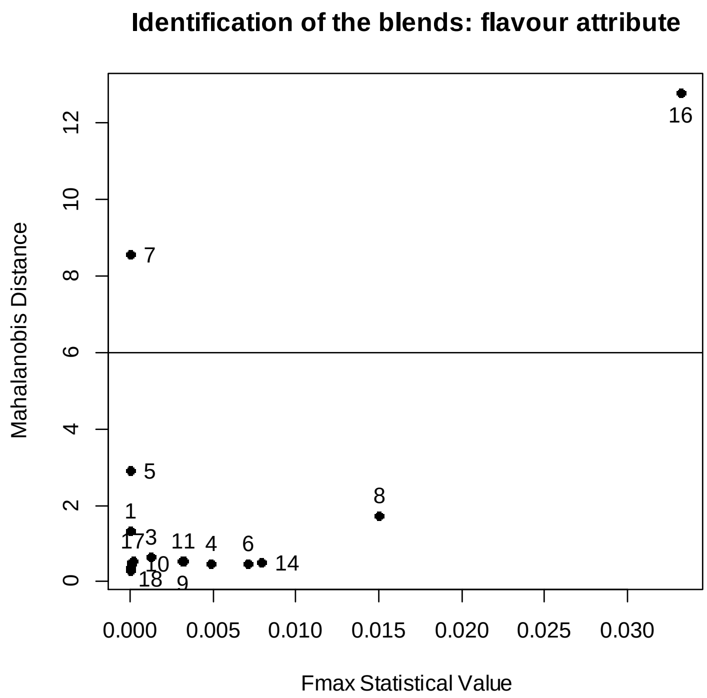

Figure 7. Identification of the dissonant blends of the sensory evaluations of the flavour variables of experiments 3-4.

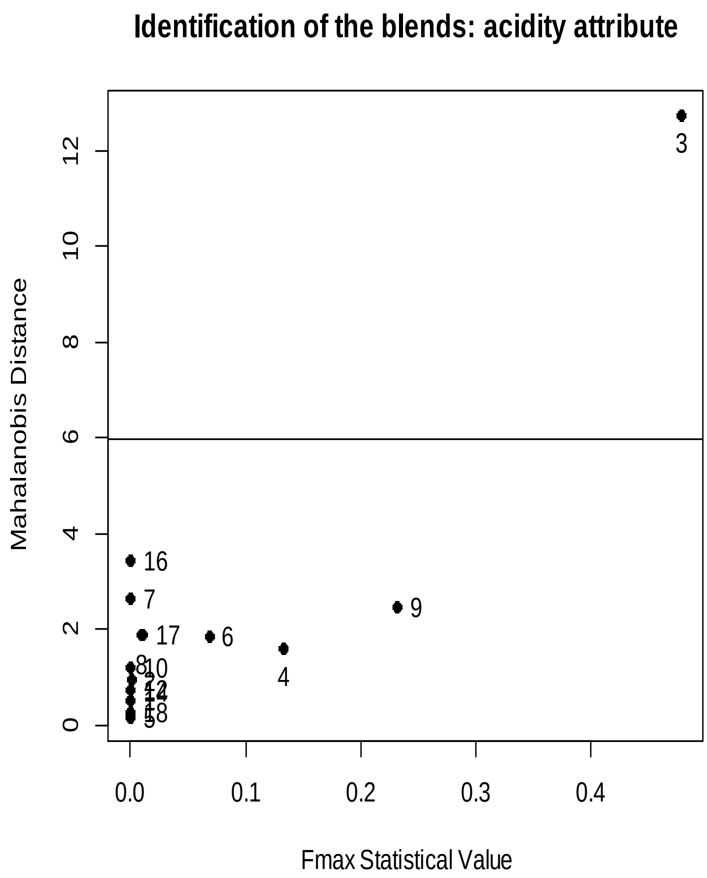

Figure 8. Identification of the dissonant blends of the sensory evaluations of the variable acidities of experiments 3-4.

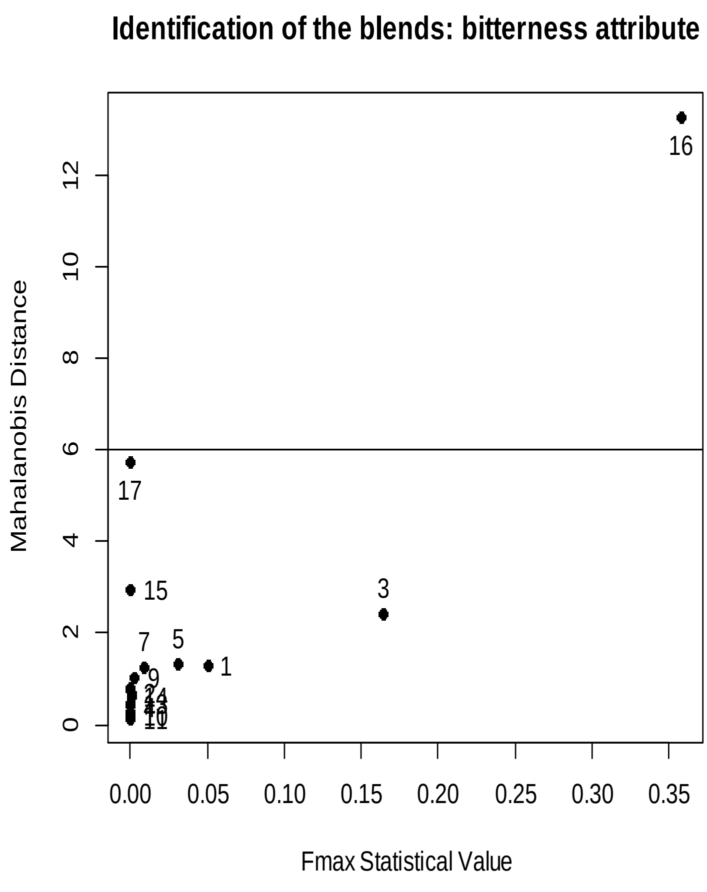

Figure 9. Identification of the dissonant blends of the sensory evaluations of the bitterness variables of experiments 3-4.

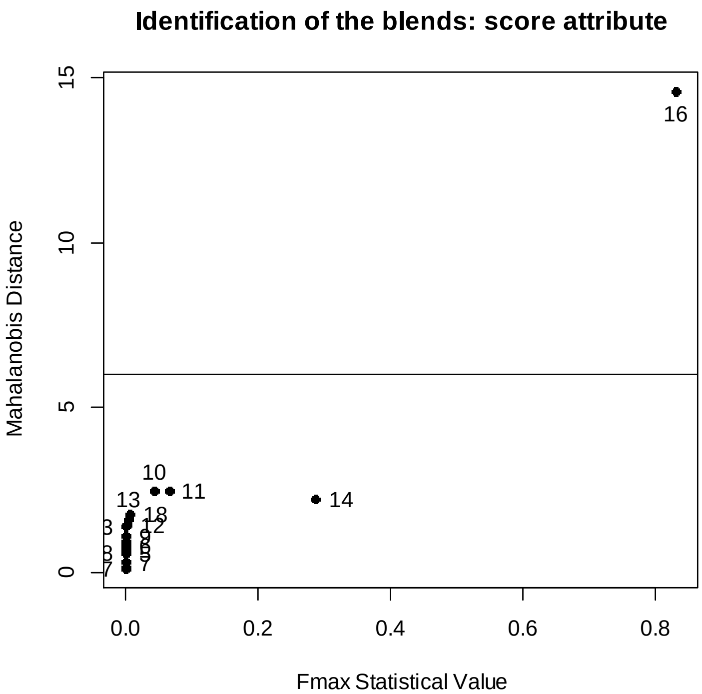

Figure 10. Identification of the dissonant blends of the sensory evaluations using the score variables for experiments 3-4.

In general, we note that the statistical methodology proposed in this work can be viewed as a tool for assessing the performances of tasters in a sensory panel. Additionally, that the discriminatory power of blends compared to pure coffees was efficient, which showed that there is no need to be retrained in order to exhibit repeatable results.

Conclusion

The procedure proposed in this work to identify outliers was adequate and feasible; it can be used in the analysis of sensory data by considering a normal distribution for sensory responses. According to these guidelines, the discrepant observations that were identified were interpreted as pure coffees.

The statistical procedure proposed in this study was efficient for the discrimination between blends and pure coffees. Based on this study, the effects of the concentrations and types of processing did not interfere with the statistical evaluations in any way; thus, the procedure was not invalidated.

Acknowledgements

We thank the CNPq for their aid via grant nº 304974/2015-3.

References

Alessandrini, L., Romani, S., Pinnavaia, G., & Rosa, M. D. (2008). Near infrared spectroscopy: an analytical to tool predicted coffee roasting degree. *Analytica Chemical Acta*, *625*(1), 95-102. DOI: 10.1016/j.aca.2008.07.013

Alves, G. E., Borém, F. M., Isquierdo, E. P., Siqueira, V. C., Cirillo, M. A., & Pinto, A. C. F. (2017). Physiological and sensorial quality of arabica coffee subject to different temperatures and drying airflows. *Acta Scientiarum. Agronomy*, *39*(2), 225-233. DOI: 10.4025/actasciagron.v39i2.31065

Barbosa, J. M., Borém, F. M., Alves, H. M. R., Cirillo, M. A., & Maguas, C. (2019). Isotopic signature of the relation between environment and the quality of spatial coffeea. *African Journal of Agricultural Research*, *14*(6), 354-360. DOI: 10.5897/AJAR2018.13633

Brighenti, D. M., Brighenti, C. R. G., Cirillo, M. A.,& Santos, C. M. B. (2010). Optimization of the components of an energetic diet for Africanized bees through the modeling of mixtures. *Journal of Apicultural Research*, *49*(4), 326-333. DOI: 10.3896/IBRA.1.49.4.05

Dal Bello, L. H. A., & Vieira, A. F. C. (2011). Optimization of a product performance using mixture experiments including process variables. *Journal of Applied Statistics*, *38*(8), 1701-1715. DOI: 10.1080/02664763.2010.518370

Ferreira, H. A., Liska, G. R., Cirillo, M. A., Borém, F. M., Ribeiro, D. E., Cortez, R. M., & Guiraldeli, C. H. C. (2016). Selecting a probabilistic model applied to the sensory analysis of specialty coffees performed with consumer. *IEEE Latin America Transactions*, *14*(3), 1507-1512. DOI: 10.1109/TLA.2016.7459642

Filho, L. T., Lucia, S. M. D., Saraiva, S. H., & Lima, R. M. (2015). Physical and chemical characteristics of espresso coffee beverages prepared from Arabica and Robusta coffee blends. *Revista Ceres*, *62*(4), 333-339. DOI: 10.1590/0034-737X201562040001

Lim, J. (2011). Hedonic scaling: a review of methods and theory. *Food Quality and Preference*, *22*(8), 733-747. DOI: 10.1016/J.FOODQUAL.2011.05.008

Liska, G. R., Menezes, F. S., Cirillo, M. A., Borém, F. M., Cortez, R. M., & Ribeiro, D. E. (2015). Evaluation of sensory panels of consumers of specialty coffee beverages using the boosting method in discriminant analysis. *Semina: Ciências Agrarias*, *36*(6), 3671-3680. DOI: 10.5433/1679-0359.2015v36n6p3671

Monteiro, M. C., & Trugo, L. C. (2005). Determination of bioactive compounds in commercial samples of roasted coffee. *Quimica Nova*, *28*(1), 637-641. DOI: 10.1590/S0100-40422005000400016

Moura, S. C. S. R., Germer, S. P. M., Anjos, V. D. A., Mori, E. E. M., Mattoso, L. H. C., Firmino, A., & Nascimento, C. J. F. (2007). Physical, chemical and sensorial evaluations of Arabica coffee blends with *Canephora* (Robusta) coffee blends. *Brazilian Journal of Food Technology*, *10*(4), 271-277.

Nepomucena, T. M., Silva, A. M., & Cirillo, M. A. (2013). Modelos ridge em planejamento de misturas: uma aplicação na extração da polpa de pequi. *Química Nova*, *36*(1), 159-164. DOI: 10.1590/S0100-40422013000100026

Ossani, P. C., Cirillo, M. A., Borém, F. M., Ribeiro, D. E., & Cortez, R. M. (2017). Quality of specialty coffees: a sensorial evaluation conducted with consumers using the MFACT technique. *Revista Ciência Agronômica*, *48*(1), 92-100. DOI: 10.5935/1806-6690.20170010

Ramos, M. F., Ribeiro, D. E., Cirillo, M. A., & Borém, F. M. (2016). 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*, *96*(10), 3543-3551. DOI: 10.1002/jsfa.7539

Resende, M., Guimaraes, C. R., & Cirillo, M. A. (2017). Procedure to identify outliers through cumulative distribution of extremes in a gamma response model. *Communications in Statistics - Simulation and Computation*, *46*(9), 6937-6946. DOI: 10.1080/03610918.2016.1217015

Ribeiro, B. B., Mendonça, L. M. V. L., Assis, G. A., Mendonça, J. M. A., Malta, M. R., & Montanari, F. F. (2014). Chemical and sensory evaluation of blends of *Coffea canephora* Pierre and *Coffea arabica* L. *Coffee Science*, *29*(2), 178-186.

Veloso, M. V. S., & Cirillo, M. A. (2016). Principal components in the discrimination of outliers: A study in simulation sample data corrected by Pearson's and Yates´s chi-square distance. *Acta Scientiarum. Technology*, *38*(2), 193-200. DOI: 10.4025/actascitechnol.v38i2.26046

Appendix A: Layout of the data in a matrix form to identify the variables to use to obtaining the results through the script used in the procedure of rescaling the data and plotting the graphs.

	X[,1]	X[,2]	X[,3]	X[,4]	Y1	Y2	Y3	Y4	Y5
Amostra	Conc	CE	CT	CC	Body	Flavour	Acidity	bitterness	Score
1	0.07	1.000	0.000	0.000	5.7	8.0	7.2	1.8	7.9
2	0.07	0.670	0.330	0.000	6.4	5.9	6.6	4.1	5.8
3	0.07	0.340	0.330	0.330	.	.	.	.	.
4	0.07	0.500	0.500	0.000	.	.	.	.	.
5	0.07	0.500	0.000	0.500	.	.	.	.	.
6	0.07	0.340	0.660	0.000	.	.	.	.	.
7	0.07	0.340	0.000	0.660	.	.	.	.	.
8	0.07	0.000	1.000	0.000	.	.	.	.	.
9	0.07	0.000	0.000	1.000	.	.	.	.	.
10	0.1	0.340	0.330	0.330	.	.	.	.	.
11	0.1	0.000	0.000	1.000	.	.	.	.	.
12	0.1	0.340	0.000	0.660	.	.	.	.	.
13	0.1	0.000	1.000	0.000	.	.	.	.	.
14	0.1	0.670	0.330	0.000	.	.	.	.	.
15	0.1	0.340	0.660	0.000	.	.	.	.	.
16	0.1	1.000	0.000	0.000	.	.	.	.	.
17	0.1	0.500	0.000	0.500	5.7	5.8	5.4	4.8	5.3
18	0.1	0.500	0.500	0.000	5.3	3.1	2.8	7	2.6

Appendix B: Script for obtaining the results with the functions to adjust models, rescale the data and construct graphs

```
# ### Function to adjust the Kronecker model ### #
 t=18  # Number of experimental points #
 ajuste_RSK <- function (X,resp)
 {
# X: experiment with mixture
# y: dependent variable
 aj <-lm(resp~ -1 + X[,2]+X[,3]+X[,4]+X[,2]:X[,1]+X[,3]:X[,1]+ X[,4]:X[,1])
    est=coef(aj) ; pred=fitted.values(aj)
    return (list(est=est,pred=pred))
 }
# ##  Rescaled data ##
res=function(X)
 {
   xmed   <- apply(X,2,median)
   xmad   <- apply(X,2,mad)
   Xij     <- matrix(0,nrow(X),ncol(X))
 for(j in 1:length(xmed))
  {
   for(i in 1:nrow(X))
   {
         Xij[i,j] <- ((X[i,j] - xmed[j])/xmad[j])
   }
  }
    return(Xij)
 }
# ###  Begin program   #### #
```

```
   X=as.matrix(dados[1:18,3:6])
  y=as.matrix(dados[1:18,7:11])
#
  chama_res=res(y)        #  1ª rescaled
  sigma=cov(chama_res)    #
  svdsigma <- svd(sigma);   V <- svdsigma$v;     Z <- chama_res%*%V
  yresc2=res(Z)           # 2ª rescaled
   sigma2=cov(yresc2)
  svdsigma2=svd(sigma2); V2=svdsigma2$v;    Z2= yresc2%*%V2
  yresc3=res(Z2)          # 3ª rescaled

# ########### Adjust the model ################### #

resp=yresc3[,5]  # variable with the rescaled data: Example: Body #
exemp=ajuste_RSK(X,resp)
ypred=exemp$pred   ;    mu=mean(ypred)  ;    desv=sd(ypred)

# ######## Identification of outliers by Fmax ########
Fac=pnorm(ypred,mu,desv,lower=T)
Fmax=(Fac)^t
D=cbind(Fmax,ypred)
S=cov(D)
DH=mahalanobis(D,colMeans(D),S)

# ##### Limit setting to discriminate outliers ######
Li=qchisq(0.95,2)

LY=range(DH)
plot(Fmax,DH,pch=16,main="Identification of the blends: score attribute",xlab = "Fmax Statistical Value",
ylab = "Mahalanobis Distance", ylim=LY)
abline(h =Li , lty = 1)
identify (Fmax,DH)
```