



Identifying Panel Inconsistency in Sensory Profiles using Multivariate Analysis of Variance (MANOVA) and follow-up Canonical Variate Analysis (CVA)

D.R. Fernando^{1*}, S. Samita¹ and T.U.S. Peiris²

¹Postgraduate Institute of Agriculture, University of Peradeniya, Sri Lanka

²Bio-Statistics Unit, Faculty of Livestock Fisheries & Nutrition, Wayamba University of Sri Lanka

ARTICLE INFO

Article history:

Received: 22 June 2021

Revised version received: 21 September 2021

Accepted: 21 October 2021

Available online: 01 January 2022

Keywords:

Canonical Variate Analysis
Consistent assessors
Multivariate statistical methods,
Panel disagreement
Sensory evaluation

Citation:

Fernando, D.R., Samita, S., and Peiris, T.U.S. (2022). Identifying Panel Inconsistency in Sensory Profiles using Multivariate Analysis of Variance (MANOVA) and follow-up Canonical Variate Analysis (CVA). *Tropical Agricultural Research*, 33(1): 80-91.

DOI: <http://doi.org/10.4038/tar.v33i1.8537>

Fernando, D.R. 
<https://orcid.org/0000-0002-4034-2799>



ABSTRACT

Inconsistency in sensory evaluation is a serious problem and it often leads to loss of large revenue. Screening tasters before the sensory evaluation is the only remedy to overcome this inconsistency. This paper aims at showing how the above requirement can be achieved by using Multivariate Analysis of Variance (MANOVA) and follow-up Canonical Variate Analysis (CVA). The approach was illustrated using sensory profiles of Sri Lankan tea. The principle behind the approach is that any discrepancy between assessors on several attributes is detected simultaneously using MANOVA, and the discrepancy interacts with the factors such as products or region is detected using CVA. Data used for the study consisted of sensory scores given by 8 tea tasters for 13 tea growing regions on 6 attributes; colour, brightness, strength, flavour, aroma, quality. Samples from four factories represented a region, and data were collected for a one-year period on a monthly basis. Data from each month were analyzed separately. The Wilk's Lambda statistics of MANOVA revealed assessor effect as well as assessor \times region interaction effect ($P < 0.05$) in every month indicating the inconsistency among assessors. The CVA for each region, specifically the 95% confidence regions of CV bi-plots, clearly identified clusters of assessors. Based on the location of these clusters in bi-plots, assessors who are suitable for different attributes were also identified. MANOVA followed by CVA, can effectively be used to identify discrepancies between assessors, discrepancy interacts with factors such as geographical region, and selecting consistent assessors depending on product or region and season.

*Corresponding author: dshafernando@yahoo.com

INTRODUCTION

Food product quality is the leading factor that governs the food market. Since product quality and human responses to products are assessed in sensory evaluations, organoleptic properties play an important role in food quality concerns (Silva *et al.*, 2014). Hence, a sensory evaluation involves the application of well-established experimental design together with statistical analysis of sensory data (Tham *et al.*, 2016). The use of organoleptic evaluation to identify food quality is rapid (Duan *et al.*, 2018). Further, as the sensory quality is a result of the interaction between the food and the human sensation, instrumental assessments of organoleptic properties are dubious and no viable alternatives to expert panels have been developed yet (Peiris *et al.*, 2018). Even though it is expected that a sensory panel should provide results that are accurate, discriminating and precise, problems in human assessors such as bias, influence of external factors, change in perceptiveness, and the time required to develop such professionals can significantly distort industrial production and mislead scientific studies (Di Donfrancesco *et al.*, 2014). Therefore, assessors' consistency becomes of paramount interest in sensory profiling (Wilderjans & Cariou, 2016). There are several studies in the literature showing assessors' inconsistency (Kaskela *et al.*, 2019; López-López *et al.*, 2018; Wilderjans & Cariou, 2016). Further, it has found some significant differences among assessors in scoring even the same sample in blind studies on sensory evaluation of tea (Peiris *et al.*, 2018). This inconsistency in sensory evaluation then becomes a serious problem and it may often lead to loss of large revenue.

Many statistical approaches have been suggested to overcome this problem, such as methods to improve the panel performances and screening methods to identify the inconsistency. On improving the panel performance, Schlich (1996) proposed the Control of Assessor Performance (CAP) method to validate assessor compromises and Peiris *et al.*, (2018) suggested that Mandel's bundle of lines model to correct individual bias and thereby to improve the performances (Peiris *et al.*, 2018; Schlich, 1996). Even though it is vital to apply these methods to improve the assessor performance, practically it is not taking place in the industry. However, before improving the assessor performance, it is imperative to identify assessors who have deviated. Thus, screening assessors before the sensory evaluation processes becomes the first remedy to overcome this inconsistency. A number of the methods have been suggested to identify the inconsistency. Consonance index (Dijksterhuis, 1995), egg-shell plot (Lea *et al.*,

1995), and linear mixed models (Latreille *et al.*, 2006) are some of them. Panel concordance analysis (PANCA), which is an extension of PCA is also a tool that has been used to identify disagreement between the panelists (Derks, 2010). In another study (Romano *et al.*, 2011), a parallel factor (PARAFAC) model was used to investigate individual differences between assessors. Data obtained from sensory tests are naturally fuzzy and uncertain. Therefore, rational methods like fuzzy logic, neural network, data aggregation, classification, and clustering methods have been proposed to screen the inconsistency and also to solve the vagueness related to the sensory field (Tham *et al.*, 2016). Due to its adaptability, the analysis of variance (ANOVA) has been employed frequently in the sensory analysis (Peiris *et al.*, 2018). Assessor main effect in the classical assessor model of ANOVA describes the disagreement between assessors, and the assessor \times product interaction describes how the disagreement varies depending on the product (Kermit & Lengard, 2005).

The major disadvantage of most of the methods suggested is that each attribute has to be considered separately and thereby correlation between attributes are ignored. In addition, several factors such as region, product and assessor cannot be considered simultaneously with most of those methods. In the case of PARAFAC method the disadvantage is that, although several attributes as well as several factors can be considered simultaneously, interpretation of factor loadings is difficult as factor rotation is not possible with PARAFAC method (Wilderjans & Cariou, 2016). However, Wilderjans & Cariou (2016) presented a clustering around the latent variable method to detect panel disagreement in three-way sensory data (CLV3W method), and this method provides solutions to most of the limitations with PARAFAC method (Wilderjans & Cariou, 2016). However, CLV3W method detects the discrepancy only for a particular cluster of attributes and also the method is applicable only for three-way data. Therefore, it would be much more interesting to identify discrepancies among assessors in a general case without limiting to three-way data. The approach of the method suggested here is the use of MANOVA and canonical variate analysis (CVA). The principle behind the approach is that any discrepancy between assessors concerning all attributes is detected simultaneously using MANOVA, and the discrepancy interacts with other factors such as products or region is detected using CVA. In addition to the discrepancy, clusters of consistent assessors for each region and month can also be identified using CVA. This study aimed at identifying panel inconsistency in sensory profiles

using Multivariate Analysis of Variance (MANOVA) and follow-up Canonical Variate Analysis (CVA).

METHODOLOGY

Multivariate Assessor Model

In conventional sensory data, i^{th} assessors assign scores for j^{th} products on K^{th} attributes in M replicates. Then univariate ANOVA model for the k^{th} attribute is given by;

$$y_{ijm}^k = \mu^k + \alpha_i^k + \beta_j^k + (\alpha\beta)_{ij}^k + \varepsilon_{ijm}^k$$

where y_{ijm}^k is the k^{th} attribute score for j^{th} product and m^{th} replicate from i^{th} assessor, α_i^k is the effect of i^{th} assessor on k^{th} attribute, β_j^k is the effect of j^{th} product on k^{th} attribute, $(\alpha\beta)_{ij}^k$ is the interaction effect of i^{th} assessor and j^{th} product on k^{th} attribute, and ε_{ijm}^k is the random error associated with y_{ijm}^k assumed to be $\varepsilon_{ijm}^k \sim IIDN(0, \sigma^2)$. In this study, regions were considered equivalent to products. Further, fixed effects were assumed for factors, assessor and region. Hence, the partitioning of total sum of squares is of the form; $SS_{Total} = SS_{Assessor} + SS_{Region} + SS_{Assessor \times Region} + SS_{Error}$,

which can be given as;

$$\begin{aligned} \Sigma Y_{ijm}^2 - \frac{y_{...}^2}{IJM} &= \left(\Sigma \frac{y_{i..}^2}{JM} - \frac{y_{...}^2}{IJM} \right) + \left(\Sigma \frac{y_{.j.}^2}{IM} - \frac{y_{...}^2}{IJM} \right) \\ &+ \left\{ \left(\Sigma \frac{y_{ij.}^2}{M} - \Sigma \frac{y_{i..}^2}{JM} \right) \right. \\ &- \left. \left(\Sigma \frac{y_{.j.}^2}{IM} - \frac{y_{...}^2}{IJM} \right) \right\} \\ &+ \left(\Sigma Y_{ijm}^2 - \Sigma \frac{y_{ij.}^2}{M} \right), \end{aligned}$$

where $\left\{ \left(\Sigma \frac{y_{ij.}^2}{M} - \Sigma \frac{y_{i..}^2}{JM} \right) - \left(\Sigma \frac{y_{.j.}^2}{IM} - \frac{y_{...}^2}{IJM} \right) \right\}$ is the SS for the Assessor \times Region interaction. The $SS_{Assessor \times Region}$ interaction compares the panel deviations.

Data were analyzed with MANOVA to test the significance of the assessor effect and Assessor \times Region interaction for all attributes. Thereby, discrepancy between assessors and the discrepancy interacting with the regions were detected. The H_0 for MANOVA were, $\alpha_i = \mathbf{0}$ for all i , $\beta_j = \mathbf{0}$ for all j , and $(\alpha\beta)_{ij} = \mathbf{0}$ for all i and j , where α_i is the vector of effect of i^{th} assessor, β_j is the vector of effect of j^{th} region

and $(\alpha\beta)_{ij}$ is the vector of interaction effect of i^{th} assessor and j^{th} region. The MANOVA needs to hold the assumption of multivariate normal distribution. The attribute scores were tested for multivariate normality using the method Mardia's test and it was found the data follow the multivariate normal distribution. Moreover, to validate the method suggested in this paper, the consistency of the outcome was studied by analyzing the data for each month separately.

As a follow up of MANOVA in the presence of interaction between assessors and region, Canonical Variate Analysis (CVA) was performed separately for each region, and thereby consistent assessors for each region and attributes were identified. The MANOVA and follow up CVA were performed using SAS University edition software package. Canonical option was included in the MANOVA statement of PROC MANOVA to estimate significant canonical variates. Using original scores of observations for attributes, CVA scores for observations were computed and then mean CVA score for each assessor were computed. Confidence regions were derived using the method as described by Seahl (Seal, 1964) for a 95% confidence (Johnson et al., 2007). After that, the assessor mean scores and 95% confidence regions for true assessor means were plotted on the graph of first two CVs. Non-overlapping 95% confidence circles would provide a guide to depict the discrepancy of assessors in two canonical space accounting for the greatest group separation. Moreover, attribute coefficients of each of the two CVA were also plotted on the CVA bi-plot and this plot was used as the guide to group the uniform assessors for each attribute.

Conventional cluster analysis will also be performed in order to evaluate the capacity of identifying panel inconsistency as MANOVA and CVA. Hierarchical agglomerative clustering (HAC) was the selected clustering method to cluster assessors in attribute-wise. Chi-square was the distance measure used for clustering. Ward's clustering method was selected as it is the only one among the agglomerative clustering methods that is based on a classical sum-of-squares criterion (Murtagh & Legendre, 2014). Cluster analysis was performed using R software with the function "hclust". The function, "fviz_nbclust" in R package of factoextra plots within sums of squares against the number of clusters. Elbow method was used to select optimal number of clusters. Result of the dendrogram was confirmed with "Duda" index in the NbClust function. Smallest number of clusters which had higher index values compared to critical values were determined as optimal number of clusters.

Data used for this study

Data on Sri Lankan tea tasters (Peiris et al., 2018) were used for the study. Data set consisted of scores on a 5-point rating scale given by eight well experienced tea tasters for samples from 13 tea growing regions of Sri Lanka on six sensory attributes. For each region there were samples from four factories for a one-year period on monthly basis. This scale data has been advocated by several authors within the scope of data exploration, for instance Cariou & Wilderjans (2018). Sensory attributes considered were colour, brightness, strength, flavour, aroma and quality (Cariou & Wilderjans, 2018). Data of assessor 1 and 3 were not available for the month of September. The 13 regions considered in the study were, Nuwaraeliya (NE), Udapussellawa (UP),

Lindula/Thalawakelle (LT), Agarapathana (AGP), Bogawanthalawa (BGW), Maskeliya (MSK), Malwatta Valley (MWV), Bandarawela (BW), Uva Medium (UM), Mid Country (MC), Sabaragamuwa (SG), Matara/ Akuressa (MA), Galle/ Elpitiya (GA).

RESULTS AND DISCUSSION

Multivariate normality of the data was tested using Mardia's test. Table 1 describes the Mardia's test statistics for each month. According to Table 1, $P > 0.05$ for both statistics in all months except for months April, July and August. Since the data for 9 months satisfy multivariate normal distribution, it was assumed that the data were suitable for the analysis suggested in the methodology.

Table 1: Mardia's test statistics and associated probabilities

Month	Test Statistic		Inference
	Skewness	Kurtosis	
January	64.08 (0.21) ^a	-0.09 (0.92) ^a	YES
February	51.75 (0.64) ^a	-0.63 (0.53) ^a	YES
March	58.23 (0.39) ^a	-1.15 (0.25) ^a	YES
April	47.68 (0.01) ^a	5.80 (0.01) ^a	NO
May	57.42 (0.42) ^a	0.01 (0.99) ^a	YES
June	45.24 (0.85) ^a	-1.00 (0.32) ^a	YES
July	88.80 (0.003) ^a	2.24 (0.03) ^a	NO
August	88.58 (0.004) ^a	0.70 (0.48) ^a	NO
September	44.66 (0.86) ^a	0.76 (0.44) ^a	YES
October	41.18 (0.93) ^a	-0.83 (0.41) ^a	YES
November	39.31 (0.96) ^a	-0.18 (0.85) ^a	YES
December	51.16 (0.66) ^a	0.52 (0.60) ^a	YES

^a Values in parenthesis are probability associated with the test statistics

Table 2: Mean and variance of scores for each sensory attributes for each tea growing Region

Region		colour	brightness	strength	flavour	aroma	quality
Nuwaraeliya	Mean	1.71	4.18	2.80	3.60	3.72	3.77
	Variance	0.52	0.54	0.82	1.23	1.05	1.13
Udapussellawa	Mean	3.09	3.36	3.12	2.36	2.54	2.61
	Variance	0.50	0.47	0.42	1.31	1.19	1.17
Lindula/ Thalawakelle	Mean	3.50	3.48	3.45	2.49	2.64	2.81
	Variance	0.55	0.33	0.38	1.24	1.06	0.99
Agarapathana	Mean	3.41	3.29	3.36	2.43	2.56	2.72
	Variance	0.61	0.53	0.47	1.33	1.23	1.15
Bogawanthalawa	Mean	3.44	3.53	3.47	2.45	2.59	2.74
	Variance	0.43	0.33	0.39	1.36	1.19	1.17
Maskeliya	Mean	3.49	3.56	3.57	2.58	2.75	2.87
	Variance	0.53	0.47	0.47	1.44	1.34	1.16
Malwatta Valley	Mean	2.95	3.53	3.24	2.52	2.66	2.72
	Variance	0.78	0.60	0.63	1.46	1.33	1.35
Bandarawela	Mean	3.45	3.15	3.17	2.24	2.42	2.39
	Variance	0.59	0.71	0.61	1.22	1.21	1.11
Uva Medium	Mean	2.20	3.77	3.06	2.85	2.94	3.07
	Variance	0.53	0.42	0.62	1.41	1.27	1.24
Mid Country	Mean	2.63	3.37	3.02	2.17	2.24	2.38
	Variance	0.53	0.46	0.48	1.19	0.98	1.12
Sabaragamuwa	Mean	2.95	3.08	3.09	1.95	2.07	2.14
	Variance	0.66	0.60	0.42	0.91	0.85	0.88
Matara/Akuressa	Mean	2.90	3.31	3.12	2.15	2.19	2.38
	Variance	0.60	0.50	0.51	1.11	1.04	1.07
Galle/Alpitiya	Mean	3.07	3.08	3.09	1.94	2.02	2.16
	Variance	0.64	0.44	0.43	0.99	0.84	0.98

Table 2 presents the mean and variance of scores given by assessors for each tea growing region of all months.

The Wilk's Lambda statistic (Λ) for the three possible effects, with associated P values, separately for each month is given in Table 3. According to Table 3, small Λ with $P < 0.0001$ for all three effects indicates significant assessor, region and assessor \times region effects in every month. This implies not only a discrepancy between assessors but also the discrepancy between assessors varies across regions in all months. Upon detecting the discrepancy between assessors across regions, follow-up CVA analysis was performed for each region to detect assessors who are consistent. The resulted first and second mean CV scores for each region in each month are summarized in Appendix 1 (see the online

version). Based on CVA bi-plots for each region and each month, consistent assessors were grouped. As an illustration, the bi-plot for the region *Bandarawela* in the month of January is given in Figure 1. According to Figure 1, all circles are overlapping indicating all assessors are consistent. Similarly, Figure 2 displays a situation where one assessor is clearly deviated from the others, and others are more or less consistent. Figure 3 displays another possibility where two groups of assessors are clearly identified. Likewise, the outcome from all those plots is summarized in Tables 4 – 6. Accordingly, the regions where all assessors were in agreement, are listed in Table 4. It can be notified that assessors were consistent in evaluating samples from *Malwatta Valley* and *Bandarawela* almost throughout the year. This hints that experience of assessors in evaluating these two regions is at optimum level.

Table 3: Wilk's Lambda statistics (Λ) for three effects of MANOVA

Month	Source	Wilk's Λ	P
January	Assessor	0.09	<.01
	Region	0.15	<.01
	Assessor \times Region	0.07	<.01
February	Assessor	0.08	<.01
	Region	0.20	<.01
	Assessor \times Region	0.11	<.01
March	Assessor	0.08	<.01
	Region	0.10	<.01
	Assessor \times Region	0.06	<.01
April	Assessor	0.06	<.01
	Region	0.11	<.01
	Assessor \times Region	0.04	<.01
May	Assessor	0.06	<.01
	Region	0.20	<.01
	Assessor \times Region	0.03	<.01
June	Assessor	0.07	<.01
	Region	0.26	<.01
	Assessor \times Region	0.12	<.01
July	Assessor	0.06	<.01
	Region	0.16	<.01
	Assessor \times Region	0.08	<.01
August	Assessor	0.03	<.01
	Region	0.14	<.01
	Assessor \times Region	0.06	<.01
September	Assessor	0.12	<.01
	Region	0.19	<.01
	Assessor \times Region	0.10	<.01
October	Assessor	0.03	<.01
	Region	0.20	<.01
	Assessor \times Region	0.04	<.01
November	Assessor	0.05	<.01
	Region	0.12	<.01
	Assessor \times Region	0.01	<.01
December	Assessor	0.03	<.01
	Region	0.13	<.01
	Assessor \times Region	0.03	<.01

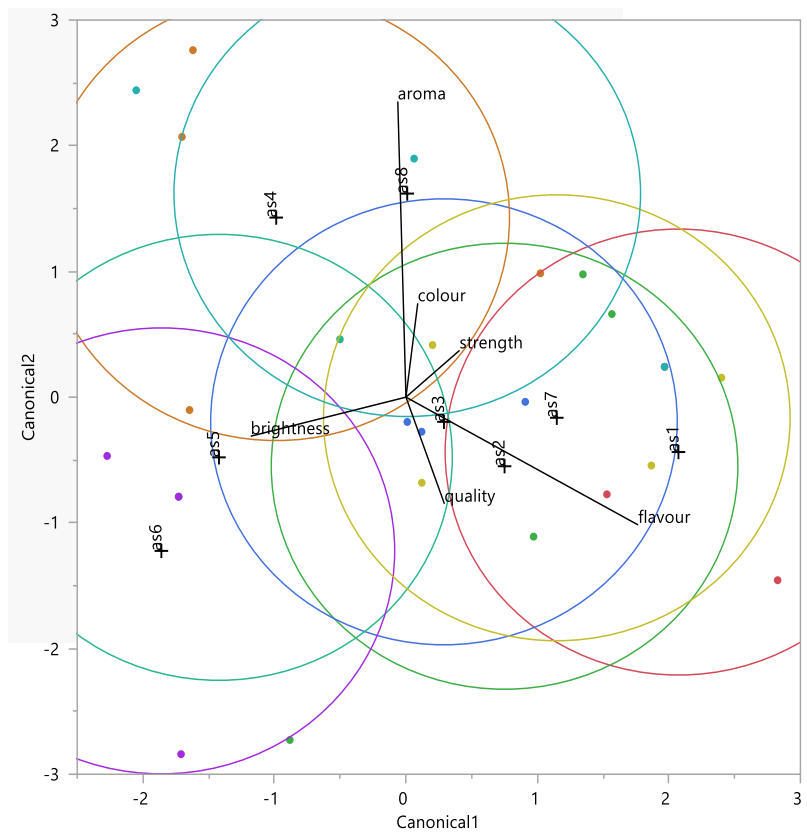


Figure 1: Canonical bi-plot for the region *Bandarawela* in January where all assessors were in an agreement.

Table 4: Regions for those assessors were in an agreement.

Months	Tea growing regions												
	NE	UP	LT	AGP	BGW	MSK	MWV	BW	UM	MC	SG	MA	GA
Jan								√	√				
Feb				√			√	√	√				√
Mar			√			√	√						
Apr	√								√			√	√
May		√											
Jun							√	√		√		√	
Jul							√	√		√			
Aug							√	√					
Sep							√						
Oct							√						
Nov							√	√					
Dec	√						√						

Situations similar to Figure 2 where a single assessor has deviated for many regions are summarized in Table 5. From the scatter points of attribute coefficients in the CVA bi-plot of Figure 2,

it can be interpreted that the expertness of assessor 3 is more towards the attribute of strength and assessors 1, 7, 5 and 8 are towards flavour.

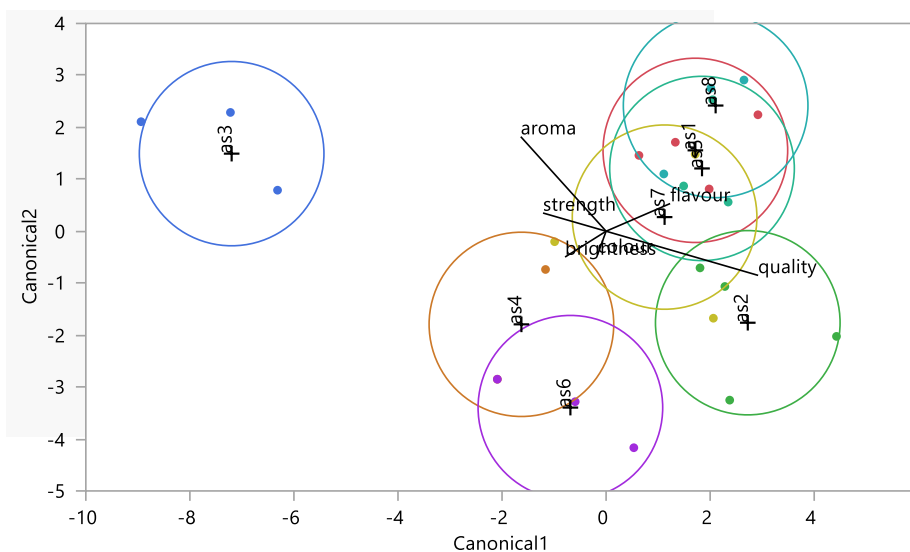


Figure 2: Canonical bi-plot for the region *Agarapathana* in April with single assessor is deviated

Table 5: Regions in each month where only one assessor is significantly different

Months	Tea growing regions												
	NE	UP	LT	AGP	BGW	MSK	MWV	BW	UM	MC	SG	MA	GA
Jan						4	8						3
Feb			3								8	5	
Mar				6						5			
Apr		3		3	3			8		8	5		
May	2					3		8		1	8		3
Jun	1			3		6							
Jul	2	3											
Aug		2											
Sep													
Oct													
Nov	2											2	
Dec										3			

In the two-group case of Figure 3, the detailed interpretation can be made as assessors 3, 4, and 6 are consistent and deviated from assessor 1, 2, 5, 7, and 8 who are separately consistent. Such situations of two groups are summarized in Table

6. According to the Figure 3, it may further be interpreted that the group consisting of assessors 3, 4 and 6 are not suitable to assess any attribute considered in the study and the group consisting of assessors 1, 2, 5, 7, and 8 are suitable to assess all attributes except strength.

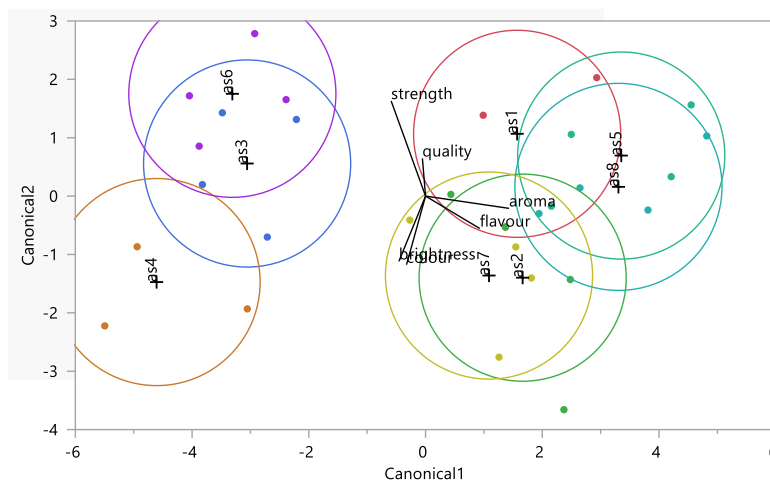


Figure 3: Canonical bi-plot of two groups of assessors for *Bogawanthalawa* in October

Table 6: Regions in each month with two different groups of assessors

Month	Assessors in group 1	Assessors in group 2	Tea growing region
Jan	2,5,6	1,3,4,7,8	SG, UP, GA
	4,6	1,2,3,5,7,8	AGP
Feb	3,4,6	1,2,5,7,8	UP
Mar	2,7	1,3,4,5,6,8	BW
	2,5,7,8	1,3,4,5,6	UP
Apr	3,4,6	1,2,5,7,8	MWV
	3,4,6	1,2,5,7,8	BGW
Jun	4,6	1,2,3,5,7,8	GA
Jul	2,3,4,6	1,5,7,8	MSK
Aug	2,3,4,6	1,5,7,8	LT
	3,4,6	1,2,5,7,8	MC
Sep	4,6	2,5,7,8	GA
	4,6,7	2,5,7,8	UP
	4,6	2,5,7,8	LT
	4,6	2,5,7,8	SG
Oct	4,6	1,2,3,5,7,8	AGP
	1,3,4,6	2,5,7,8	BW
	3,4,6	1,2,5,7,8	BGW
Nov	3,4,6	1,2,5,7,8	UP
	3,4,6	1,2,5,7,8	MSK
Dec	3,4,6	1,2,5,7,8	MA
	3,4,6	1,2,5,7,8	MSK
	1,3,	2,4,5,7,8	UM

In the conventional cluster analysis, groups of assessors were created for each attribute separately. Hence, the resulting partitions of assessors will depend only on the chosen attribute. This finding will be a special interest of the panel leaders when it is required to select uniform set of panelists with similar capacity of assessment of a particular attribute. However, discrepancy

interacts with other factors such as products or region is not detected by the cluster analysis. Groups of assessors formed for each attribute are presented in Table 7. Therefore, MANOVA and CVA would be opted over the cluster analysis in identifying panel inconsistency and also the changes of inconsistency with tea growing region for attributes simultaneously.

Table 7: Number of clusters created for each attribute in all months.

Month/Attribute	Number of Clusters					
	Color	Brightness	Strength	Flavour	Aroma	Quality
January	2	2	3	3	3	3
February	2	2	2	3	3	3
March	2	2	2	2	3	2
April	2	2	2	2	3	2
May	2	2	2	3	2	2
June	3	2	2	3	3	3
July	2	2	2	2	4	2
August	2	2	2	2	3	3
September	2	2	2	2	2	2
October	2	2	2	2	3	5
November	3	2	2	3	3	2
December	2	2	2	4	3	3

Canonical Variate Analysis (CVA) has been introduced as a natural alternative to Principal Component Analysis (PCA) for sensory profiling data (Peltier *et al.*, 2015). This paper also highlighted that the multivariate approaches have to be used in sensory data analysis as several attributes have to be considered simultaneously, and especially they have identified the discrimination of product using CVA. Multivariate methods to evaluate assessors in sensory data are based on multivariate assessor model (Romano *et al.*, 2015), which is an extension of the univariate multiplicative assessor model with several attributes. Although, Wilderjans & Cariou, (2016) a follow up from (Romano *et al.*, 2015), assessed the panel inconsistency among assessors using the approach clustering around latent variable. The limitation of that method is that, it is appropriate only for three-way data, which means that analysis will be allowed only with two factors and the response variables. The usefulness of the method suggesting in present paper overcomes that limitation as this method does not have limitations of number of factors that can be considered. The

data used in the present study consisted of scores assigned for six attributes on the same scale. However, it is possible in experiments that scores for different attributes are assigned on different scales. In such a case scale effect can be removed by the method reported by Wilderjans & Cariou, (2016) and adjusted data can be used in the method suggested in this paper (Wilderjans & Cariou, 2016). Very few literature was found on MANOVA and CVA in identifying panel discrepancy in sensory analysis. Nevertheless, present study demonstrated that the canonical bi-plots clearly express the discrepancies among assessors on the first two canonical variates.

Among the discrepancies found for different regions, highest level discrepancy was found for the region *Galle/Elpitiya* in the month of November and *Uva Medium* in October, where five groups were identified among the eight assessors. This shows the extent to which discrepancies among assessors exist and how important it is to have methods to detect such discrepancies. Tea auctioneers will be highly benefitted by such

methods as consistent outcome from assessors is crucial to have a high price for the product. Another crucial aspect of sensory panel evaluation is identifying suitable assessors to assess certain specific attributes which are of immense valuable in tea exporting. Such expertness to assess certain attributes has to be achieved by proper training and then the tea exporters may also need to evaluate whether their tea tasters have achieved that competence after the training. Methods for these kinds of purposes are very useful, and the statistical method suggested in this paper can provide solutions to those requirements.

CONCLUSIONS

This study clearly demonstrated that disagreements exist among assessors and assessments could vary depending on factors such as geographical region and the month of the year. Effective methods are required to identify discrepancies between assessors, and selecting consistent assessors depending on product such as region and season. The method suggested in this paper, which is the use of MANOVA followed by CVA, can effectively be used to serve the above purpose. Moreover, assessors who have specific expertise to identify specific attributes can also be selected from this method. MANOVA paves the way to detect discrepancies, and CVA bi-plots provides a strategy to identify consistent assessors suitable for different products such as for different regions. The repeated analysis performed by means of separate analysis for each month has shown that the method can be used under different circumstance such as two groups of assessors or three groups of assessors, with some groups suitable and some groups not suitable, or one group suitable for certain attributes and another group suitable for some other attributes. As the number of observations in this study were 2496 for a month, fulfilling distributional assumptions for the data may not be a problem. However, multivariate normality was preserved in this data, and in case if number of observations is not that large, testing data for suitability can be done before proceeding with the analysis. Hence, in situations where general, or attribute specific agreement among assessors is required, such assessors can be identified using MANOVA followed by CVA, and this approach could be recommended as a statistical tool to explore the disagreement among assessors and identifying suitable assessors for sensory evaluations.

ACKNOWLEDGEMENT

Sharing the data collected by Dr. T. U. S. Peiris has been very much appreciated.

REFERENCES

- Cariou, V., & Wilderjans, T. F. (2018). Consumer segmentation in multi-attribute product evaluation by means of non-negatively constrained CLV3W. *Food Quality and Preference*, 67, 18-26.
- Derks, E. P. P. A. (2010). PANCA: Panel concordance analysis. *Food Quality and Preference*, 21(3), 324-329.
- Di Donfrancesco, B., Gutierrez Guzman, N., & Chambers Iv, E. (2014). Comparison of Results from Cupping and Descriptive Sensory Analysis of Colombian Brewed Coffee. *Journal of Sensory Studies*, 29(4), 301-311.
- Dijksterhuis, G. (1995). Assessing panel consonance. *Food Quality and Preference*, 6(1), 7-14.
- Duan, Y., Song, X., Yang, Y., Li, L., Zhao, Y., Wang, Y., Fang, W., & Zhu, X. Q. (2018). Correlation between Sensory Evaluation Scores and Flavor Attributes of Black Teas from Around the World. *Journal of Food Processing and Technology*, 9, 1-9.
- Johnson, H. E., Lloyd, A. J., Mur, L. A., Smith, A. R., & Causton, D. R. (2007). The application of MANOVA to analyse Arabidopsis thaliana metabolomic data from factorially designed experiments. *Metabolomics*, 3(4), 517-530.
- Kaskela, J., Vainio, A., Ollila, S., & Lundén, J. (2019). Food business operators' opinions on disclosed food safety inspections and occurrence of disagreements with inspector grading. *Food Control*, 105, 248-255.
- Kermit, M., & Lengard, V. (2005). Assessing the performance of a sensory panel-panellist monitoring and tracking. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 19(3), 154-161.
- Latreille, J., Mauger, E., Ambroisine, L., Tenenhaus, M., Vincent, M., Navarro, S., & Guinot, C. (2006). Measurement of the reliability of sensory panel performances. *Food Quality and Preference*, 17(5), 369-375.
- Lea, P., Rødbotten, M., & Næs, T. (1995). Measuring validity in sensory analysis. *Food Quality and Preference*, 6(4), 321-326.
- López-López, A., Sánchez Gómez, A., Montaña, A., Cortés-Delgado, A., & Garrido Fernández, A.

- (2018). Sensory profile of green Spanish-style table olives according to cultivar and origin. *Food Research International*, 108.
- Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *Journal of classification*, 31(3), 274-295.
- Peiris, T. U. S., Walgampaya, C. K., Banks, D. L., Thattil, R. O., & Abeysinghe, I. S. B. (2018). Use of Mandel's bundle of lines model to improve agreement of a panel of tea tasters. *Journal of the National Science Foundation of Sri Lanka*, 46(1).
- Peltier, C., Visalli, M., & Schlich, P. (2015). Canonical Variate Analysis of Sensory Profiling Data. *Journal of Sensory Studies*, 30.
- Romano, R., Næs, T., & Brockhoff, P. B. (2015). Combining analysis of variance and three-way factor analysis methods for studying additive and multiplicative effects in sensory panel data. *Journal of Chemometrics*, 29(1), 29-37.
- Romano, R., Vestergaard, J., Kompany-Zareh, M., & Bredie, W. (2011). Monitoring Panel Performance Within and Between Sensory Experiments by Multi-Way Analysis. In (pp. 335-342).
- Schlich, P. (1996). Defining and Validating Assessor Compromises About Product Distances and Attribute Correlations. *Data Handling in Science and Technology*, 16, 259-306.
- Seal, H. (1964). Multivariate statistical analysis for biologists. *Multivariate statistical analysis for biologists*.
- Silva, A. Y. S. L., Binduhewa, A. M. C. U., & Subodinee, A. A. M. (2014). A study to recruit and train the product oriented sensory panel. *International Journal of Multidisciplinary Studies*, 1(2).
- Tham, H., Tang, S., Teo, K., & Loh, S. (2016). A review on intelligent sensory modelling. IOP Conference Series: Earth and Environmental Science,
- Wilderjans, T. F., & Cariou, V. (2016). CLV3W: A clustering around latent variables approach to detect panel disagreement in three-way conventional sensory profiling data. *Food Quality and Preference*, 47, 45-53.