Combining analysis of variance and three-way factor analysis methods for studying additive and multiplicative effects in sensory analysis

Journal:	*Journal of Chemometrics*
Manuscript ID:	Draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Romano, Rosaria; University of Calabria, Economics, Statistics and Finance Næs, Tormod; Nofima Mat, Brockhoff, Per; Technical University of Denmark, DTU Compute
Keyword:	descriptive sensory analysis, scaling effects, assessor model, three-way analysis

Combining analysis of variance and three-way factor analysis methods for studying additive and multiplicative effects in sensory panel data

Rosaria Romano[1,*], Tormod Næs[2,3], Per Bruun Brockhoff[4]

rosaria.romano@unical.it, tormod.naes@nofima.no, pbb@imm.dtu.dk

[1]Department of Economics, Statistics and Finance, University of Calabria, 87036 Cosenza, Italy

[2]Nofima Mat AS, Osloveien 1, 1430 Ås, Norway

[3] Department of Food Science, University of Copenhagen, 1958 Frederiksberg C, Denmark

[4]DTU Compute, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

[*]Correspondence: Rosaria Romano, Department of Economics, Statistics and Finance, University of Calabria, 87036 Arcavacata di Rende, Cosenza, Italy.
E-mail: rosaria.romano@unical.it

Abstract:

Data from descriptive sensory analysis are essentially three-way data with assessors, samples and attributes as the three ways in the data set. Because of this there are several ways that the data can be analysed. The paper focuses on the analysis of sensory characteristics of products while taking into account the individual differences among assessors. In particular we will be interested in considering the multiplicative *assessor model* which explicitly models the different usage of scale. A multivariate generalization of the model will be proposed which allows to analyse the differences in the use of the scale with reference to the existing structure of relationships between sensory descriptors. The *multivariate assessor model* will be tested on a data set from milk. Relations between the proposed model and other multiplicative models like PARAFAC and ANOVA will be clarified.

Keywords: Descriptive sensory analysis, Scaling effects, Assessor model, Three-way analysis.

1. Introduction

In descriptive sensory analysis a group of trained assessors, the sensory panel, gives scores on a continuous scale for a certain number of sensory attributes for all products in the study. Besides studying variation in products/samples, which is usually the main objective of the analysis, differences between assessors as well as relationships among descriptors should be taken into account in order to understand better the system under investigation.

A number of methods have been proposed and used for the purpose of analysing the different aspects separately and all three aspects simultaneously taking the three-way structure of the data into account. The methods applied are often modifications or combinations of Analysis of Variance (ANOVA), Principal Components Analysis (PCA) and Three-Way Factor Analysis (TWFA) models, depending on the focus of the study. Important examples of methods which have a solid basis in sensory analysis are general and used in many areas of statistics [1-6] while others are closely related specifically to the effects that are specific for sensory analysis [7-17]. Most of these methods have been illustrated in a recent book [18].

One of the approaches given special attention in this paper is the so-called multiplicative *assessor model* [9] which explicitly models the product effect and the product × assessor interaction effects by a joint multiplicative term. The model focuses on differences in the different use of the scale between assessors and it is based on the assumption that these effects are linearly related to the main effects of products. It has been shown [16] that scaling differences may considerably affects results of the analysis. Therefore, the information about the differences between the assessors in the use of the scale plays a crucial role. First of all because it is common practice in descriptive sensory analysis to calculate and analyse the average of individual judgments. Correcting for scaling differences before averaging may simplify and improve analysis. Secondly, the information on these differences could be used to perform a pre-processing of data in which any distortions could be resized. Furthermore, this information could be used to improve the performance of panels in the future. In recent work, [19] and [20], the original assessor model from [9] has been extended to the ANOVA mixed model framework, the Mixed Assessor Model (MAM), which is typically needed to obtain the proper univariate statistical inference for attribute-wise analysis of sensory data, see also [18]. In [19] the focus is on how to obtain the proper analysis of the product information and it is shown in a big meta study of thousands of sensory attributes that it clearly improves the attribute-wise statistical power. In [20] it is shown how it is possible to simultaneously obtain univariate assessor performance focussed analysis within the same mixed model framework. The model, as it stands now, however, is essentially still a univariate model and must be utilised for each sensory attribute separately. For this reason and since this new work on the MAM is likely to increase the future use of this approach for the analysis of sensory data, we aim in this paper to bridge the gap between the univariate assessor model approach and the generic multivariate structure of sensory data.

The main purpose of the present paper is to extend the multiplicative *assessor model* to comprise several attributes. A new model named *multivariate assessor model* is proposed, which explicitly models the product effect and the product × assessor interaction effects taking the multivariate structure of the sensory descriptors into account. Here too, as in the univariate case, focus is on scaling effects linearly related to the main effects of products. As with all data compression models, the basic assumption of the *multivariate assessor* *model* is that there is a reduced number of latent variables that summarize the relationships between sensory descriptors and that allow to analyse sensory similarities and differences between the products. Furthermore, the model assumes that assessors scale the sensory attributes (manifest variables) in a different way. The latter hypothesis is typical of the *multivariate assessor model* and differs from other models proposed in the literature that suggest that assessors perceive the same underlying sensory dimensions (latent variables), but using these in different ways [8, 15].

Different versions of the *multivariate assessor model* will on one hand be presented for the purpose of theoretically clarifying the characteristics of the various models and the relations to existing models and approaches. Some of them permit to demonstrate important relations between a number of methods frequently used for modelling sensory data. Some of the methods that will be involved in the discussion are PARAFAC [5], regular factorial ANOVA, ASCA [21] and PCA. On the other hand, a particular version of the model will be presented as a valid analysis tool for sensory data. This version of the *multivariate assessor model* permits to analyse the product-space, with the key consideration on the identification of scaling differences among the assessors. It is based on a multivariate component decomposition of the product effects and scaling effects separately. Such a separate decomposition, allows to get information on the sensory differences and similarities between the products, which is the main objective of any sensory analysis, as well as information on the differences in the use of the scale among the assessors considering the set of sensory variables simultaneously. This additional information provided by the model may be used to perform a pre-processing of data before continuing the analysis with the classical statistical methods. This will improve the results of the analysis. In addition this type of information on the assessors performance is a great potential for a panel leader to improve the future panel performance. Such a specific multivariate extension of the assessor model for the analysis of multivariate sensory data is a novelty. It gives insight into the communality among the multiplicative effects that is not obvious if each variable is treated separately. Furthermore, it links more directly to a multivariate analysis of the product effects than if each variables is transformed individually.

How to interpret, validate and estimate the model will be discussed and visualised using an example from sensory analysis of milk.

2. Methods

2.1. Univariate Assessor Model

Let Y^k_{ijm} denote the score of assessor i (i=1, ..., I) on attribute k (k=1, ..., K) of the mth replicate (m=1, ..., M) of the jth product (j=1, ..., J). A model accounting explicitly for all individual differences, apart from the so-called *disagreement* (see below) is the multiplicative *assessor model* [9, 11]. The model can be formulated for each attribute k as:

$$Y^k_{ijm} = \alpha^k_i + \beta^k_i v_j + e^k_{ijm}, \qquad \text{where } e^k_{ijm} \sim N(0, \sigma^2_i) \tag{1}$$

As can be seen, the model includes assessor main effects α_i and multiplicative interaction effects $\beta_i v_j$, which are simply the product of the scaling effect β_i with the product effect v_j. Assessors with large β_i use a larger portion of the scale than the average assessor. Note that differently from a classical two-way ANOVA with assessor, product and assessor × product interaction, the model only treats the part of the interaction effects connected to the usage of the scale without considering all the other non-additive assessor differences generally called *disagreement.* In addition, the error variance (σ_i^2) here allows for different assessors' variability. Specifically the error terms e_{ijm} include all systematic interaction effects not accounted for by the multiplicative terms and individual differences between the sensory replicates. Although both differences are important for determining panel reliability incorporating this aspect in a multivariate setting is beyond the scope of the present paper. For the rest of this section, when considering univariate models the k superscript will be omitted.

The *assessor model* can also be written as:

$$Y_{ijm} = \alpha_i + v_j + \beta^*_i v_j + e_{ijm}, \qquad (2)$$

i.e. a model that also incorporates the main effects for product with $\beta = 1+\beta^*$. Note that the model can equivalently be formulated with a general mean μ, but for the multiplicative model it is usually omitted.

In [9] and [11] formal model fit hypothesis tests are suggested as a way of investigating the validity of the *assessor model*.

For simplicity and without loss of generality we will in the rest of this paper unless otherwise stated, subtract the assessor means from the data ending up with the model:

$$X_{ijm} = v_j + \beta^*_i v_j + e_{ijm} = \beta_i v_j + e_{ijm} \qquad (3)$$

This correspond to correcting data by removing differences between assessors in location (level effect).

The estimation of the model parameters is achieved by an iterative algorithm described by the authors in their original paper [9].

Note that the *assessor model* is closely related to the model proposed by Mandel in 1971 [22] which consists in the use of a multiplicative model based on PCA for modelling of interactions:

$$Y_{ijm} = \mu + \alpha_i + v_j + \sum_a t_{ia} p_{ja} + e_{ijm} \tag{4}$$

Here a is the number of reduced dimensions in the interactions. If one in the *Mandel model* assumes that a = 1 and that $v_j = p_j$ (or better proportional to each other), one ends up with the multiplicative *assessor model* (2).

2.2. Multivariate Assessor Model

If the data (averaged over replications) are corrected for the assessor-wise attribute averages the *assessor model* (3) formulated for all attributes simultaneously (without assuming any common structure among the attributes or samples) can be written as:

$$X_{ijk} = \beta_{ik} v_{jk} + e_{ijk} \tag{5}$$

Then, it is clear from a data compression perspective that the *multivariate assessor model* (MUAM) is simply the attribute-wise 1-component PCA of the assessors-by-products matrix without correcting for (removing) product effects. But as can be noted this model incorporates no link between the attributes, i.e. there is no modelling involved associated with the relation between the attributes.

As sensory data rarely vary in the full dimensional space of all attributes a dimension reduction approach will often enhance the stability and interpretability of the results. Hence, it is likely to expect that similarly the scaling differences would also benefit in the same way by a dimension reduction.

At this aim we suggest some possible restrictions that can be used to connect the attributes with each other. The main idea is to assume a multivariate component decomposition of the products effects and scaling effects separately, i.e.

- a L-component product-by-attribute structure: $\hat{v}_{jk} = \sum_{l=1}^{L} t_{jl} p_{lk}$

and

- a H-component assessor-by-attribute structure: $\hat{\beta}_{ik} = \sum_{h=1}^{H} c_{ih} d_{hk}$

The factor models can in principle be defined in different ways, but here we confine ourselves to PCA models. The former component model is possibly the most obvious since products usually vary in a low dimensional sensory space. Note that for the maximum number of components this is exactly the model (5), so these assumptions represent a true restriction.

The full model using these restrictions and then called *restricted multivariate assessor model* (RMUAM) can be formulated as:

$$X_{ijk} = \left(\sum_{h=1}^{H} c_{ih} d_{hk} \right) \left(\sum_{l=1}^{L} t_{jl} p_{lk} \right) + e_{ijk} \qquad (6)$$

The parameters of the RMUAM are estimated by two independent PCA's. The algorithm for the estimation of the model parameters is described in the Appendix.

As discussed above, the model in (6) imposes restrictions on both the product effects and the scaling constants, but in a rather flexible way. The extreme variant of this is to set the number of components in both models equal to 1. This model will here be called the *one-component* RMUAM and can be written as:

$$X_{ijk} = c_{i1} d_{1k} t_{j1} p_{1k} + e_{ijk} \qquad (7)$$

Note that model (7) represents a very strict assessor model, it essentially assumes that the use of scale for a single assessor is identical except for a multiplicative effect and that this effect is the same for each attribute. In addition, this version of the model is closely related to the one-dimensional PARAFAC model. This connection will be discussed in a later section.

3. Assessing the adequacy of the Multivariate Assessor Model

The multivariate extensions of the *assessor model* lead to a hierarchy of models as discussed above: the most flexible full MUAM (5) with no assumptions on the relations between attributes nor samples, the RMUAM (6) which reduces the dimensionality of both the products and the scalings structures, and the *one-component* RMUAM (7). Assessing this hierarchy of models in a practical data analysis situation requires a strategy based on different model comparisons with different focus.

The test of adequacy of the MUAM (5) can be done by simply checking for uni-dimensionality of the individual PCA models. If some attributes follow the multiplicative structure and others do not, it may be possible to continue further investigations with the former group only. Or one may choose to use the MUAM whether it fits the data completely or not, knowing that in this way the generic scaling effect has now been separated from the real perceptual disagreement effects. The latter, represented by the residuals from the MUAM could then be subjected to further multivariate analysis to study this information.

Then the RMUAM (6) should be evaluated for different number of components in each of the two modes (product and scaling). The question is whether there is a link between the attributes that can be adequately modelled by a reduced factor model in at least one of the two modes.

Another way of assessing the validity of the RMUAM consists in evaluating how much variability is explained by ignoring information on scaling differences among assessors. This can be done by calculating the *model explained variance* but replacing the predicted $\hat{X}_{ijk}$ values with the product by attribute averaged data for each i^{th} assessor slice, i.e. by replicating the v_{jk} matrix I times. If the explained variance computed in this way is close to the ones obtained from the RMUAM then this is an indication of a poor performance of the RMUAM. In other words, a multivariate model accounting also for the scaling effects does not provide further insights in the analysis of data.

Implicit in the assessment of RMUAM (6) is the check of the validity of the *one-component* RMUAM (7), since one component is one of the models that take part in this comparison.

4. Assessing the quality of the Restricted Multivariate Assessor Model

The model in (6) provides three different type of explained variance by combining the predicted values from the two PCA fits into predicted values for the full model:

- the *model explained variance* relative to the total variation

$$1-\frac{\sum_{ijk}\left(X_{ijk}-\hat{X}_{ijk}\right)^2}{\sum_{ijk}X_{ijk}^2} \tag{8}$$

- the *product explained variance* relative to the total product variation

$$1-\frac{\sum_{jk}\left(v_{jk}-\hat{v}_{jk}\right)^2}{\sum_{jk}v_{jk}^2} \tag{9}$$

- the *scaling explained variance* relative to the total scaling variation

$$1-\frac{\sum_{ik}\left(\beta_{ik}-\hat{\beta}_{ik}\right)^2}{\sum_{ik}\beta_{ik}^2} \tag{10}$$

The *model explained variance* can be calculated for all the combinations of possible number of components in the two separated PCA's. This is possible because the two decompositions are independent from each other, that means the product variation and the scaling variation can be decomposed by a proper number of components according to their respective structure. The optimal number of components for the full model comes up from the combination of the two PCA's producing the highest *model explained variance*.

Besides different explained variances, the RMUAM also provides different sets of loadings which can be plotted in order to visually detect both how products differ with respect to the sensory descriptors and how these descriptors are scaled differently by assessors.

5. Relations of the Multivariate Assessor Model with ANOVA and 3-way Factor Analysis

Another interesting way of combining PCA and ANOVA was proposed by Smilde and co-workers [21]. The method is called ASCA and is based on using standard multivariate ANOVA for estimating the effects and then using PCA for each of the effect matrices separately. The PCA can, however, also run on matrices that are composed of combinations of for instance main effects and interactions matrices. Note that the estimated interaction matrix for multivariate response data can be matricized/unfolded before PCA in different ways according to the three dimensions/ways of the data set. Alternatively, the cube of interactions can be directly investigated by three-way methods. This is the strategy behind PARAFASCA model [23], which uses PARAFAC as a three-way method to explore interactions.

In the following we will consider the multivariate ANOVA on data averaged over replicates and with the assessor effects subtracted, i.e.:

$$X_{ij}{}^{k} = v_j^k + \gamma_{ij}^k + e_{ij}^k \tag{11}$$

Again the mean is subtracted from each assessor and attribute combination and therefore no assessor or average effect is needed. The ν_j^k's are the sample main effects, the γ_{ij}^k's are the interactions between assessors and attributes and the e_{ij}^k's are the error terms. For this model, an ASCA/PARAFASCA approach provides LS estimates of the effects matrices $\mathbf{\Lambda} = \{\nu_j^k\}$ and $\mathbf{\Gamma} = \{\gamma_{\mathrm{ij}}^{\mathrm{k}}\}$, and then analyses the two matrices separately by PCA or PARAFAC. If we in addition assume that both the matrices can adequately be fitted by one-dimensional PCA and PARAFAC models, we end up with:

$$\mathbf{\Lambda} = t_j p_k \tag{12}$$

and

$$\mathbf{\Gamma} = a_i s_j r_k \tag{13}$$

for the terms. Assuming further that the j (product) dimensions are identical (t_j=s_j), and also that the k (attribute) dimensions are identical (p_k=r_k), we end up with the following model for X_{ijk}:

$$X_{ijk} = (1 + a_i)t_j p_k = a_i^* t_j p_k \tag{14}$$

which would have been identical to *one-component* RMUAM in (6) if a_i=$c_i d_k$, that is if assessors had presented the same scalings for the different latent dimensions. In other words, the estimated ANOVA model with a PCA/PARAFAC decomposition of each of the effects accompanied with restrictions discussed above, leads us to the same restricted model as was obtained in (7) using a totally different approach.

On the other hand the *one-component* RMUAM is also strictly related to PARAFAC model:

$$X_{ijk} = c_{i1}d_{1k}t_{j1}p_{1k} + e_{ijk} = c_i d_k t_j + e_{ijk} \quad (15)$$

In fact, it corresponds to a 1-component PARAFAC model of the matrix **X** which is centred for each assessor and attribute combination.

The two models are not exactly the same because of the constraint in the RMUAM model that the scalings average to 1. However, this difference can easily be removed if the 1-term is introduced in the scaling part:

$$X_{ijk} = \left(1 + \sum_{h=1}^{H} c_{ih}d_{hk}\right)\left(\sum_{l=1}^{L} t_{jl}p_{lk}\right) + e_{ijk} \quad (16)$$

The *one-component* RMUAM can then be written as:

$$X_{ijk} = (1 + c_{i1}d_{1k})t_{j1}p_{1k} + e_{ijk} \quad (17)$$

At this point, the comparison between the fit of the model (17), which corresponds to a 1-component PARAFAC model, and of the RMUAM (6) selecting 1-component for both the scaling and the product mode, will produce the same results.

6. Results

6.1 Data Description

Six varieties of milk with respect to two dairy cow breeds (Holstein Friesian (HF) and Jersey (JE)) and 6 different farms (UGJ, HM, EMC, OA, JP, KI) were profiled by a panel of 10 assessors over 9 descriptors (green and feed odor, yellow and grey appearance, creamy, boiled milk, sweet, bitter and sourness flavor). The samples were evaluated in 3 replicates, randomized within the full experiment, according to a continuous scale anchored at 0 and 15. The data were collected in a three-way table (samples *x* assessors *x* attributes) with the $J \times M$ products ($J=6$ products in $M=3$ replicates) as the first *way,* the $I=10$ assessors as the second *way*, and the $K=9$ attributes as the third *way*.

The MATLAB® software has been used for implementing multivariate data analysis and making plots. All analyses for multi-way models were performed in MATLAB® (Mathworks, Inc.) using the PLS_Toolbox version 4.0 (Eigenvector Research, Manson, WA). Additional in-house made routines using the R free software were used for implementing the *multivariate assessor model*.

First a two-way ANOVA with assessor and interaction as random effects is run on the raw data. Results in Figure 1 show the attribute-wise (1 - p-value) for all the effects in the model. As it can be seen, there is a significant assessor effect (1 - p-value > 0.95) for all attributes but *feed odor* and *sourness flavour*. There are significant differences among products for the four attributes *yellow* and *grey appearance* and *creamy* and *sourness flavour*. Finally, there are significant assessor × product interaction effects for all the attributes apart from *sweet flavour*.

Results from PCA on raw data averaged across assessors and replicates can be seen in Figure 2. The biplot shows a strong separation among samples with respect to the cows' breed on the first principal component, which explains most of the variation (77.4% Exp. Var.). Specifically JE milk is described as *yellow*, *creamy* and *sweet* milk, whereas the HF samples are characterized by the sensory attributes *grey appearance*, *bitter* and *boiled milk flavor*. The second principal component (16.7% Exp. Var.) discriminates samples within the same race. In particular the UGJ-JE milk presents higher values on the attributes *sourness flavor* and *feed odor*.

In the following only data corrected for the assessor level effect will be used.

6.2. Assessing the Multivariate Assessor model

As discussed in section (3) assessing the appropriateness of the *multivariate assessor model* is a multi-step procedure according to the MUAM models hierarchy.

6.2.1. Testing the full multivariate assessor model

The first step consists in checking for attribute-wise 1-component PCA of the assessor-by-product matrix corrected for the assessor effects. Figure 3 shows that the uni-dimensionality assumption is satisfied in more than an half of the cases. The first principal component explains most of the variability for all the sensory attributes except for *grey appearance*, *creamy flavour* and *feed odor* where the second component also plays an important role. Figure 3, also shows that the amount of variance explained by the first principal components of the different PCA is very high as compared to the variability explained by the remaining components.

6.2.2. Testing the restricted multivariate assessor model

The second step of the models comparison strategy consists in testing the adequacy of the *restricted multivariate assessor model* to focus on how much of the variation in the interaction structure is explained by the model. This model comparison is done by modelling only the six variables that have passed the first test, i.e. the variables presenting a uni-dimensional structure of the assessor-by-product matrix (green odor, yellow appearance, and boiled milk, sweet, bitter and sourness flavor).

At this point the RMUAM (6) is computed for each combination of components in the two separated PCA models. Note that, the testing of the *one-component* RMUAM (7) will be part of this when the one-component structure for the two separated PCA models is taken into account.

Results in Table I show how the best model is the one with two components in the product structure and two components in the scaling structure since it explains more variability (61.8%), while it seems that beyond the third dimension the increase is modest. Here the explained variance is computed with respect to the total variability of data with the main effect of the assessors removed. Since the algorithm behind the model does not provide a global fit but just refit the scalings taking the product structure fixed throughout, the explained variances in the table increase as the number of components in the scaling structure increases but could decrease when the number of components in the product structure increases. A RMUAM solution with 2 components in the product structure and 2 components in the scaling describes 98.5% of the product variation (Table I third column), 76.4% of the scaling variation (Table I third raw) and 61.8% of the total variation (Table I second column). The product variation corresponds to the explained variance in the two PC-component of the product information, i.e. the PCA on the averaged assessors corrected data (Table I third column). The scaling variation corresponds to the explained variance in the two PC-component of the scaling information estimated by the RMUAM (Table I last row). The total variation corresponds to the variation explained by the RMUAM relative to the total variance, i.e. the total variance of the full cube of assessor corrected values. This solution explains almost the same variation (69.2%) as the full model with (5, 8) components, respectively. This underlines the advantage of the RMUAM in explaining a major part of the information by using a reduced number of components. The RMUAM fit can also be compared to the situation where only the average product configuration is used as the model for all the assessors. Results in Table I highlights the good performance of the RMUAM since the explained variances of all its possible combinations of the two dimension are always higher than those obtained by considering only the average product configuration (Table I last column).

The loading plot of the 2-component model fitting the scaling structure is shown in Figure 4. This plot allows visualizing and exploring the relationships between assessors and attributes concerning the scaling effects. Specifically Figure 4 shows how assessors use the scale differently for each attribute. In fact it can be seen that assessors 4 and 6 utilize a large range of the scale for attributes *green odor* and *boiled milk flavor*, which are situated in the positive direction of the first component. The assessors 10 and 2 have also high scalings, but for the *sourness flavor* attribute located in the opposite direction. Finally, the assessors who are at the far ends of the second principal component are those which show differences in range for the attributes *bitter* and *sweet flavor*. There are no substantial differences for the attribute *yellow appearance* located at the origin of the axes.

In order to have a feedback on the results of the RMUAM, the standardized deviations of each assessor with respect to each attribute are shown in Figure 5. As can be seen all assessors have the same mean equal zero (denoted by the 'x' markers) since the data were corrected in order to remove the individual differences in location. The graphs for individual attributes confirm that present a higher range, assessors 2 and 10 on *sourness flavor*, assessor 3 on *sweet flavor*, and assessors 4 and 6 on *green odor* and *boiled milk flavor*. In addition, assessors 8, 5 and 1 have very small range on the attribute *bitter flavour*. In fact these are the same assessors who were located on the opposite side of this variable in the loading plot of the RMUAM. Furthermore, the detailed information of Figure 5 confirms the absence of differences in use of scale for the attribute *yellow appearance*.

As discussed in section 3 implicit in the assessment of RMUAM is the validation of the *one-component* RMUAM. Results in Table I show that a model with 1 component in both the product and the scaling structures explains 46% of the total variability which is quite low compared to the model with two components (61.8%). The *one-component* RMUAM is then inappropriate for the milk data.

7. Discussion and Conclusions

In this paper we have discussed the problem of analysing the sensory data as three way data by taking into account all the three ways of information: products, assessors and attributes. We have emphasized the importance of considering the individual differences among assessors in the use of the scale in a multivariate perspective that takes into account the relationship between the sensory variables.

As a first contribution we have extended the univariate *assessor model* to comprise several attributes. In its more general version we have shown how the MUAM is simply the attribute-wise 1-component PCA of the assessors-by-products matrix corrected for the assessor effects. Thus, considering the uni-dimensionality of the milk data, we have found out that it is appropriate for a restricted group of variables.

Since the MUAM does not take into account relations among the different sensory attributes a restricted version of it defined RMUAM has been presented, which can be used to connect the attributes among them. It is based on a principal component decomposition of both the product and the scaling effects. In its first version (not shown in the paper), the proposed algorithm for the estimation of model parameters consisted in an iterative procedure that calculated recursively the PCA on the product effects, the assessor-attribute wise scalings and the PCA on the estimated scalings. However, the algorithm set up this way did not produce any reasonable results out of trying to actually optimize jointly the model. Thus, the alternative was to consider two separate PCA's: one to decompose the product structure and one to decompose the scaling structure. The predicted values from the two PCA's are then combined in order to get predicted values for the RMUAM. Note that this two-step procedure based on simple separated PCA allows to estimate the RMUAM without fitting it globally as a truly multiway model. We have tested this model on milk data for each combination of components in the two separated PCA models. Results have highlighted the advantage of the model in explaining a major part of the information by using a reduced number of components. It has also been shown that the RMUAM provides a better understanding of the data since it explains more information compared to the situation where only the average product configuration is used as the model for all the assessors. Note that the RMUAM also provides graphical outputs to visualize and explore relations between assessors and attributes concerning the scaling structure. This is a great potential of the model since with a few simple graphics (loading plots from the two separated PCA) gives you information on the sensory differences and similarities between the products, and the differences in the use of the scale among the assessors considering the set of sensory variables simultaneously.

The extreme variant of the RMUAM consists in setting the number of components in the two separated PCA equal to 1. This type of model has been called *one-component* RMUAM. It has been theoretically compared with other methods frequently used for modelling sensory data: ANOVA and PARAFAC. In the first case, we have shown that the estimated ANOVA model with a PCA decomposition of the product effects and a PARAFAC decomposition of the interaction effects under some assumptions leads to the *one component* RMUAM model. In the second case we have discussed how the *one component* RMUAM is very close to a 1-component PARAFAC model on a matrix centred by subtraction of the main effects and interactions between assessors and attributes. Specifically, *one component* RMUAM is comparable with PARAFAC but results cannot be exactly the same due to the constant 1-term in the scaling part.

An apparent limitation of the *multivariate assessor model* in all of its versions is that it only looks at scaling effects, so it does not fit the entire data when there are attributes present with real perceptual disagreements. However it can still play the important role of separating the generically present scaling part of the interaction prior to subsequent multivariate methods, to make sure that the scaling effect is not mistaken for any other effect in the data.

Acknowledgements

The authors would like to acknowledge Wender L.P. Bredie, and Jannie S. Vestergaard, University of Copenhagen, Denmark, for the use of the milk data.

Appendix

Algorithm for the estimates of the parameters in the *Restricted Multivariate Assessor Model*

The decomposition of the product structure is based on a PCA of the matrix of product by attribute averaged data:

$$v_{jk} = \sum_{l=1}^{L} t_{jl} p_{lk} + e_{jk}$$

The decomposition of the scaling structure is based on a PCA of the assessor by attribute matrix holding the attribute-wise scalings:

$$\beta_{ik} = \sum_{h=1}^{H} c_{ih} d_{hk} + e_{ik}$$

The attribute-wise scalings β_{ik} are estimated by assessor-wise *least squares* linear regressions of the observations X_{ijk} on the given product values v_{jk}. The *least squares* criterion can be written as:

$$\sum_{ijk} \left(X_{ijk} - \beta_{ik} v_{jk} \right)^2$$

References

1. Bro R, Jakobsen M. Exploring complex interactions in designed data using GEMANOVA. Colour changes in fresh beef during storage. *J. Chemom.* 2002; **16**: 294-304.

2. Dijksterhuis GB, Gower JC. The interpretation of generalized procrustes analysis and allied methods. *Food Qual. Pref.* 1991; **3**: 68-87.

3. Kroonenberg PM, De Leeuw J. Principal component analysis of three mode data by means of alternating least squares algorithm. *Psychometrika* 1980; **45**: 69–97.

4. Mardia KV, Kent JT, Bibby JM. Multivariate Analysis. Academic Press: London, 1979.

5. Smilde AK, Bro R, Geladi P. Multi-way Analysis. Applications in the chemical sciences. Wiley PL New York, 2004.

6. Stahle L, Wold S. Multivariate analysis of variance (MANOVA). *Chemom. Intell. Lab. Syst.* 1990; **9**: 127-141.

7. Arnold GM, Williams AA. The use of Generalized Procrustes Analysis in sensory analysis. In Statistical Procedures in Food Research, Piggott JR (ed). Elsevier: Amsterdam, 1986; 233-253.

8. Bro R, Quannari EM, Kiers HAL, Næs T, Frost MB. Multi-way models for sensory profiling data. *J. Chemom.* 2008; **22**: 36-45.

9. Brockhoff PB, Skovgaard I. Modelling individual differences between assessors in sensory evaluation. *Food Qual.Prefer.* 1994; **5**: 215–224.

10. Brockhoff PB, Hirst D, Næs T. Analysing individual profiles by three-way factor analysis. In Multivariate Analysis of Data in Sensory Science, Næs T, Risvik E (eds). Elsevier: Amsterdam, 1996; 307-342.

11. Brockhoff PB. Statistical testing of individual differences in sensory profiling. *Food Qual. Pref.* 2003; **14**; 425–434.

12. Luciano G, Næs T. Interpreting sensory data by combining principal component analysis and analysis of variance. *Food Qual. Prefer.* 2009; **20**: 167-175.

13. Næs T. Handling individual differences between assessors in sensory profiling. *Food Qual. Pref.* 1990; **2**: 187–199.

14. Næs T, Langsrud O. Fixed or random assessors in sensory profiling? *Food Qual. Prefer.* 1998; **9**: 145–152.

15. Quannari EM, Meyners M. Identifying Assessor Differences in weighting the underlying sensory dimensions. *J. Sens. Stud.* 2001;**16**: 505-515.

16. Romano R, Brockhoff PB, Hersleth M, Tomic O, Næs T. Correcting for different use of the scale and the need for further analysis of individual differences in sensory analysis. *Food Qual. Pref.* 2008; **19**: 197-209.

17. Schlich P. Defining and validating assessor compromises about product distances and attribute correlations. In Multivariate Analysis of Data in Sensory Science, Næs T, Risvik E (eds). Elsevier: Amsterdam, 1996; 259-306.

18. Næs T, Brockhoff PB, Tomic O. Statistics for sensory and consumer science. Wiley, 2010.

19 Brockhoff PB, Schlich P., Skovgaard I. Taking individual scaling differences into account by analyzing profile data with the Mixed Assessor Model. *Revised version Submitted to Food Quality and Preference* 2014.

20. Peltier C, Brockhoff PB, Visalli M, Schlich P. The MAM-CAP table: A new tool for monitoring panel performances. *Food Quality and Preference* 2014; **32** (Part A): 24-27.

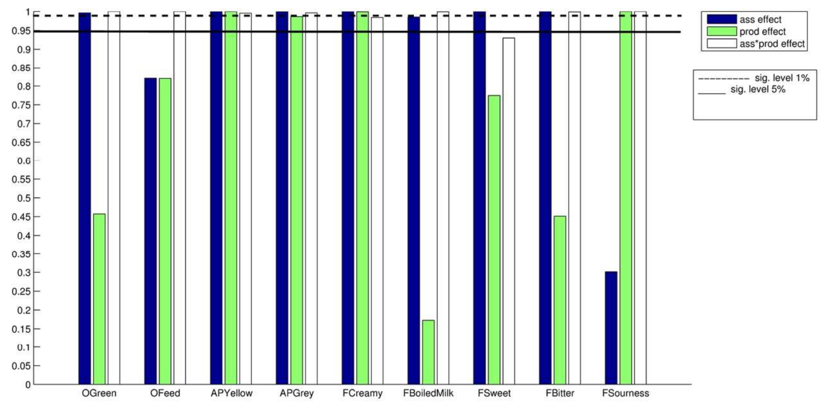
21. Jansen JJ, Hoefsloot HCJ, Van der Greef J, Timmerman ME, Westerhuis JA, Smilde A. ASCA: Analysis of multivariate data obtained from an experimental design. *J. Chemom.* 2005; **19**: 469-481.

22. Mandel J. A New Analysis of Variance Model for Non-Additive Data. *Technometrics* 1971; **13**: 1-18.

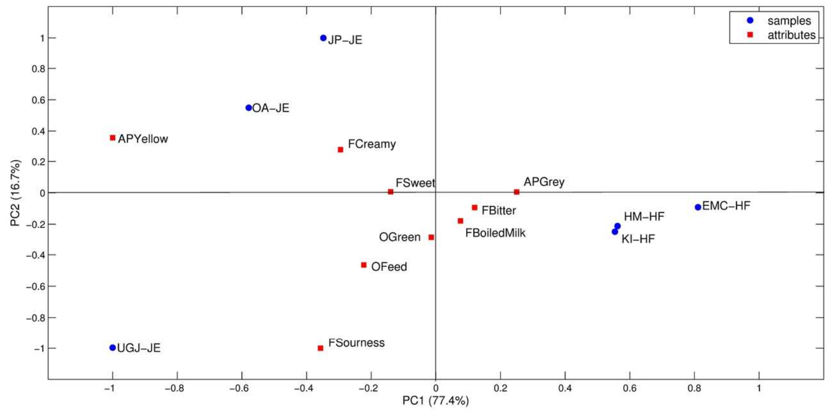
23. Jansen JJ, Bro R, Hoefsloot HCJ, van den Berg F, Westerhuis JA, Smilde AK. PARAFASCA: ASCA combined with PARAFAC for the analysis of metabolic fingerprinting data. *J. Chemom.* 2008; **22**: 114-121.

Table I. Explained Variance from the RMUAM for each combination of components in product and scaling structures.

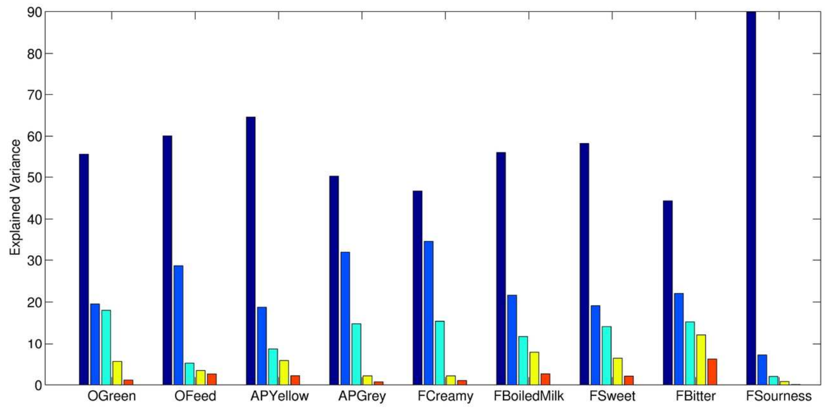
		scaling structure						*PCA on product averages*	*Product PCA as model for full data matrix*
		PC1	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**		
product structure	**PC1**	46.0	51.3	51.8	52.4	54.1	55.6	81.7	41.6
	PC2	55.5	61.8	63.3	63.9	66.9	68.4	98.5	50.2
	PC3	56.3	62.1	63.5	64.1	67.2	68.6	99.5	50.7
	PC4	56.7	62.6	63.9	64.4	67.5	69.0	99.9	51.0
	PC5	56.9	62.9	64.1	64.6	67.7	69.2	100.0	51.1
PCA on scaling values		44.5	76.4	89.4	95.6	99.5	100.0		



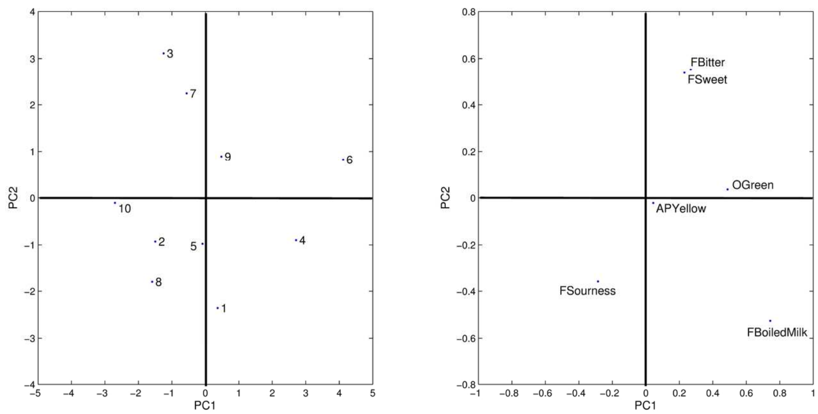
Two-way ANOVA results (1 - p-value) on the raw data for each attribute
166x76mm (300 x 300 DPI)



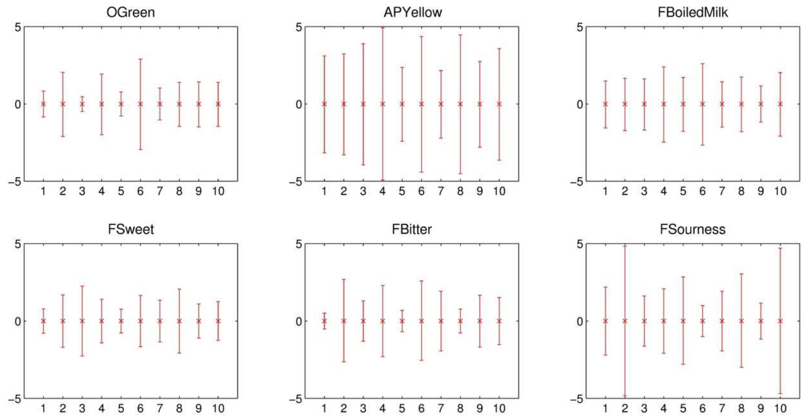
Biplot from PCA on averaged data
166x76mm (300 x 300 DPI)



Explained Variances from attribute-wise PCA
166x76mm (300 x 300 DPI)



Loading plot of the assessor by attribute structure from 2-component RMUAM
166x76mm (300 x 300 DPI)



Assessors raw scalings (standard deviation of the assessor level corrected scores)
166x76mm (300 x 300 DPI)