Preference study using a latent class approach. Analysis of European preferences for smoked salmon

Michel Séménou[1*], Philippe Courcoux[1], Mireille Cardinal[2],
Huguette Nicod[3], Alexandra Ouisse[3]

[1] Unité Mixte de Recherche de Sensométrie et Chimiométrie – Ecole Nationale d'Ingénieurs des Techniques Agricoles et Alimentaires UMR ENITIAA/ INRA, Rue de la Géraudière, BP 82225, 44322 Nantes Cedex 3, France
[2] IFREMER, Département de Sciences et Techniques Alimentaires Marines, rue de l'Ile d'Yeu, BP 21105, 44311 Nantes Cedex 3, France
[3] ADRIANT, rue P. Bobierre, BP 62303, 44323 Nantes Cedex 3, France

*: Corresponding author : semenou@enitiaa-nantes.fr

Abstract:

We present in this paper the advantages of the latent class vector model approach in the analysis of preference ratings. We illustrate the practical application of this methodology on the basis of a preference study of smoked salmons in European countries. A set of 30 samples was selected spanning the characteristic variability of cold-smoked salmon available in six European countries (Belgium, Denmark, France, Germany, Italy and the United Kingdom). Consumers from the main countries of consumption (Belgium, France, Italy, Germany and the United Kingdom) were asked to rate the acceptability of each product. A quantitative descriptive analysis with a trained panel and physical and chemical measurements of smoked salmons were also conducted. Using a latent class vector model approach, it was possible to identify different classes of homogeneous preferences. Each of these classes was characterised by the country of origin of the consumers, sensory descriptions and the physical and chemical properties of the products.

Keywords: Sensory analysis; Consumer preferences; Smoked salmon; Preference mapping; Segmentation; Latent class vector model

1. Introduction

Preference evaluation is generally assessed by using a hedonic scale, assigning an absolute liking score to each product to be tested. Internal preference mapping on hedonic ratings is a classic way to study consumer preferences. The MDPREF analysis (Caroll, 1972, 1980; Greenhoff & MacFie, 1994, Schiffman, Reynolds & Young, 1981) is a technique consisting of a singular value decomposition of the data matrix (centred by consumer). It leads to a graphical interpretation of individual preferences using a biplot representation. Unfortunately, such a method does not always make clear the existence of clusters among panellists. Moreover, the number of dimensions to consider is not easy to define because of the regularity of decrease of singular values. One way to overcome these difficulties is to consider a latent class vector model approach (De Soete & Winsberg, 1993). This method allows the practitioner to cluster consumers into a small number of segments and to obtain a simultaneous representation of stimuli (products) and classes on the basis of a vector model (as in the MDPREF approach). Based on a probabilistic assumption, this technique enables the implementation of statistical significance tests on the number of classes and dimensions to explore.

The objective of this study is to identify homogeneous classes in consumer panels according to their overall preferences using a latent class vector model methodology. We apply the latent class vector model approach to a real data set resulting from the European project *EUROSALMON* (Courcoux et al. 2006).

Moreover, we relate these segments to product characteristics obtained by sensory profiling experiments and physical and chemical measurements. In addition, we examine relationship between the consumer's country of origin on the different classes.

2. Materials and methods

2.1. Materials

Samples of commercial smoked salmon products were collected directly from producers in six different European countries; Belgium, Denmark, France, Germany, Italy and the United Kingdom. These products were selected according to the procedure described by Cardinal et al. (2004) with a first sampling in supermarkets of the same countries, in order to cover the range of smoked salmon characteristics available on the European market today. Among these products, 6 originated from Belgium, 7 from Denmark, 21 from France, 6 from Germany, 4 from Italy and 13 from the United Kingdom. On the basis of the results obtained after a sensory evaluation of these samples (Cardinal et al. 2004), 30 products were selected to analyse the preferences of consumers from the main countries of consumption. The aim was to select 30 samples reflecting the observed variance in the sensory space. Following an external preference mapping methodology, a cubic linear regression model was postulated linking consumers' acceptability of salmon to characterization measurements. We have considered the four first principal components of sensory data.

If we denote by y the vector of preferences for products and by $\mathbf{X}$ the matrix of the measurement effects, the regression model, classically used in preference mapping, is defined by:

$$y = \mathbf{X}\beta + \varepsilon$$

where β is the vector of the unknown parameters and ε is the residual random vector with a multivariate normal distribution. The least square estimator of β is classically obtained by:

$$\hat{\beta} = (\mathbf{X'X})^{-1}\mathbf{X'}y$$

The optimality criterion of an experimental design is directly linked to the covariance matrix of this estimator. Our selection is based on a D-optimality criterion that consists of

selecting the 30 products among the 57 candidates such that $\det((\mathbf{X'X})^{-1})$ is minimal, which is equivalent to minimising the generalised variance of the estimator (Atkinson & Donev, 1992). In order to achieve this selection according to the D-optimality criterion, an iterative procedure based on the Fedorov exchange algorithm was used (Fedorov, 1972).

2.2. Methods

2.2.1. Sensory methods

Descriptive sensory analysis was performed on the 30 smoked salmons by a trained panel of 24 assessors, with a wide experience in smoked salmon evaluation, from the staff of IFREMER. The experimental design used was the same as for the study of 102 products and is described by Cardinal et al. (2004). Regarding the 30 smoked salmon samples, an experimental design was constructed in order to balance various parameters, such as the sensory characteristics of products presented within a session and their country of origin. All the samples offered to the panellists were produced two weeks before their assessment. Trained panellists assessed five different samples of smoked salmon in each session, using a presentation design balanced for order and carry-over effects. Evaluations lasted for three weeks with two sessions per week.

For each sample, panellists received one slice of cold-smoked salmon from a packet of 100 or 200g, stored at +4°C until analysis. Before the evaluation session, the packets of smoked salmon were opened and kept at ambient temperature (18-20°C). Then each slice was served in aluminium foil and products were blind-coded with 3-digit numbers. Mineral water and bread were provided as palate neutralisers between samples.

Attributes of smoked salmon were chosen during training sessions and related to odour, appearance, flavour and texture. A summary of the sensory attributes used is given in Table 1.

Sessions were carried out in individual partitioned booths. A computerised system (Fizz, Biosystèmes, Dijon, France) allowed panellists to rate sensory attributes on a continuous unstructured linear scale from low intensity (score 0) to high intensity (score 10). Table 2 summarises the main sensory characteristics of the 30 products presented to consumers according to the groups identified by Cardinal et al. (2004).

2.2.2. Methods for characterisation of smoked salmon

Several measurements were made in order to characterise smoked salmon. Details concerning this part can be found in Cardinal et al. (2004). A summary of the microbiological, physical and chemical attributes is given in Table 3.

2.2.3. Consumer methodology

A total of 1063 subjects (501 male and 562 female), coming from five different countries of Western Europe, participated in the study. The number of consumers recruited in each country was approximately the same: 211 in France, 200 in Belgium, 230 in Italy, 207 in Germany and 215 in the United Kingdom. They were living in 2 or 3 different cities in each country. Participants were between 18 and 65 years old, with about one third aged from 18 to 34, one third from 35 to 49 and one third from 50 to 65 years. Consumers were pre-recruited for the test following a screener adapted for each country by considering the frequency of consumption of smoked salmon: five to six times a year for Belgium, France and Italy, once a month for Germany and once to twice a year for the United Kingdom. Assessment of smoked salmon took place over six visits to centrally located venues. At each session, the consumers

were presented with 5 different samples of smoked salmon. Presentation and assessment were sequential monadic. The same experimental design was used as that for the trained panel in order to balance sensory characteristics and the origin of salmon in the different sessions. Order and carry-over effects were balanced and bread or crackers and mineral water were provided to subjects between samples.

Consumers were asked to rate the acceptability of each product on a 9-point structured scale (dislike extremely - like extremely).

2.2.3. Statistical analysis

In order to obtain a segmentation of the panel into clusters of homogeneous preferences, a latent class vector model approach (De Soete & Winsberg, 1993; Courcoux & Chavanne, 2001) was used. This technique aims to determine segments from the preference data, using an Expectation–Maximisation algorithm (Dempster, Lair & Rubin, 1977; McLachlan & Basford, 1988).

Let us define by $Y = ((y_{ij}))$, $i = 1$ to N, j $= 1$ to M, the NxM matrix of the centred preference ratings assessed by the N consumers for the M products. We assume that there are T latent classes of homogeneous consumers. We want to obtain a simultaneous representation of the M products as points and the T classes as vectors emanating from the origin in an R-dimensional space. We consider that, for the individual i in the class t, the random vector of ratings $y_i = (y_{ij})$, $j = 1$ to M, is normally distributed with mean vector u_t and covariance matrix $\sigma^2 I$ (where I represents the identity matrix in R^M):

$$y_i \sim \mathrm{N}\left(u_t, \sigma^2 I\right) \text{ for the individual } i \text{ in the segment } t.$$

$u_t = (u_{tj})$, $j = 1$ to M, is a specific mean vector for the class t and σ^2 is common to all latent classes. u_t represents the mean utility for each product in the class t. We define by b_{jr}

and a_{tr}^{*}, respectively, the coordinate of the product j and the coordinate of the vector terminus for the class t on dimension r. We assume that the utility, u_{tj}, of product j for the latent class t is a linear function of the orthogonal projection of the object point j onto the vector representing class t:

$$u_{tj} = \alpha_t \frac{\sum\limits_{r=1}^{R} a_{tr}^{*} b_{jr}}{\sqrt{\sum\limits_{r=1}^{R} a_{tr}^{*2}}} \tag{1}$$

which can be written:

$$u_{tj} = \sum\limits_{r=1}^{R} a_{tr} b_{jr} \tag{2}$$

with:

$$a_{tr} = \frac{\alpha_t a_{tr}^{*}}{\sqrt{\sum\limits_{r=1}^{R} a_{tr}^{*2}}} .$$

In the class t, the probability density function of the random vector y_i is defined by:

$$f(y_i / a_t, B, \sigma^2) = \left(\sigma\sqrt{2\pi}\right)^{-M} \exp\left(-\frac{(y_i - u_t)'(y_i - u_t)}{2\sigma^2}\right) \tag{3}$$

with $a_t = (a_{tr})$, $r = 1$ to R, $B = ((b_{jr}))$, $j = 1$ to M, $r = 1$ to R and $u_t = (u_{tj})$, $j = 1$ to M.

Thus, the unconditional probability density function of y_i becomes a finite mixture of multivariate normal densities:

$$g(y_i / A, B, \sigma^2, \lambda) = \sum\limits_{t=1}^{T} \lambda_t f(y_i / a_t, B, \sigma^2) \tag{4}$$

with $A = ((a_{tr}))$, $t = 1$ to T, $r = 1$ to R and $\lambda = (\lambda_t)$, $t = 1$ to T, denoting the weights vector of the different classes under the constraint $\lambda_t > 0$ and $\sum\limits_{t=1}^{T} \lambda_t = 1$.

Assuming independence between the random variables y_i, the likelihood function associated with all the observations $y = (y_i)$, $i = 1$ to N, is defined by:

$$L(A,B,\sigma^2,\lambda/y) = \prod_{i=1}^{N} g(y_i/A,B,\sigma^2,\lambda) = \prod_{i=1}^{N} \sum_{t=1}^{T} \lambda_t f(y_i/a_t,B,\sigma^2) \quad (5)$$

Maximum likelihood estimates of parameters A, B, σ^2 and class membership probabilities λ can be obtained by maximising (5). This can be solved iteratively by using an EM algorithm (Dempster et al., 1977). More details concerning the estimation procedure can be found in De Soete and Winsberg (1993).

Thanks to a probabilistic assumption, the latent class vector model methodology provides a likelihood ratio test, which can help the practitioner to choose an adequate number, T, of classes and the space dimensionality, R. As in most mixture problems, the sampling distribution of this test is unknown and can be drawn using a Monte Carlo simulation procedure (Aitkin, Anderson & Hinde, 1981). One can first determine the number of classes by comparing the log likelihoods obtained for T and $T+1$ clusters and then define the dimensionality of the model for the "right" number, T, of classes using the same technique for R and T dimensions (with $R \leq T$). A study concerning the efficiency of this significance test procedure can be found in De Soete and Winsberg (1993) on the basis of simulated data.

3. Results

3.1. Internal preference mapping

First of all, we performed an internal preference mapping (MDPREF) in order to visualize the individual preferences and the mean preference (as vectors) and the common products configuration (as points) in a multidimensional space.

In the case of the salmon study, MDPREF shows a high variability of individual preferences and a main direction preference which is coherent with the overall liking scores,

the most appreciated products 9, 36, 8, 2, 28 are on the main direction of preference, contrary to products 34, 32, 19). The individual preferences are centred, so that only the direction of preference can be interpreted. It clearly appears that there exists a great heterogeneity in term of products' perception in the consumers' panel, leading us to study the existence of different clusters according to the consumers' preferences.

3.2. Segmentation of the panel

The application of the latent class vector model methodology to the preferences of 1063 consumers for the 30 smoked salmons gives the following results.

First, we determine the number of classes T by comparing the log likelihoods obtained for T classes (T dimensions), L_T, and $T+1$ classes ($T+1$ dimensions), L_{T+1}, evaluating the Likelihood ratio statistic defined by $2(L_{T+1}-L_T)$. The significant level is given by the P-value obtained on the basis of Monte Carlo simulations (Table 4).

Secondly, the number T of classes being fixed, we determine the number R of dimensions by comparing the log likelihoods obtained for T dimensions, L_T, and R dimensions, L_R, evaluating the Likelihood ratio statistic defined by $2(L_R-L_T)$. The significant level is given by the P-value obtained on the basis of Monte Carlo simulations (Table 5).

We retain the solution with six classes and five dimensions. The weights of the different classes are, respectively, 0.115, 0.076, 0.315, 0.084, 0.367 and 0.043. The three main classes (C1, C3 and C5) are discriminated on the first three preference dimensions (dimensions 1 and 3 for C1 and C3, dimension 2 for C5). Classes 2, 4 and 6 seem to be more marginal. Configurations of products and classes are plotted in Figures 2 and 3. The length of the latent class vectors is a function of the discrimination among the products in the different classes.

Most consumers strongly reject products 34 and 41 but there are differences between the preferences expressed in the different classes. By performing multiple comparisons tests for the products on the basis of an analysis of variance by class with a product and consumer effect, we observe a high discrimination of the smoked salmons for classes 1, 3, 4, 5 and 6, while this segregation is weaker in class 2. Table 6 gives the most rejected and the most appreciated products for each class. If we consider the perception of products in classes 1, 3 and 5, product 34 is disliked in these classes. Salmons 31, 30 and 7 are highly rejected in class 1 but not in segments 3 and 5. Product 41 is rejected in C3 but not in C1 and C5. Products 49 and 2 are the most appreciated in C1 and C5 and moderately accepted in C3. Consumers belonging to C1 and C3 appreciate product 8 but this smoked salmon does not appear as preferred in C5. Product 9 is well appreciated in C5 but is not preferred as much in C1 and C3.

3.2. Correlation with physical and chemical measurements and sensory attributes

By correlating the physical and chemical measurements and sensory attributes with product appreciation in the different classes, the reasons for the acceptance or rejection of smoked salmon by consumers in the different clusters can be characterised (Figure 4).

In classes 1, 3 and 4, products associated with spoilage indicators (TVBN and TMA for chemical measurement, Lactic, Yeast and Total for microbiological determination, Famin, Orancid, Oamine for sensory evaluation) are strongly rejected. Moreover, individuals belonging to class 1 dislike highly-smoked salmons (Phenol, Fcold, Frubb, Orubb) and products with a strong taste or smell (Fglob, Oglob). In contrast, they appreciate the flavour and odour of fish (Ffish, Ofish). Products characterised by the flavour of fish and the odour of tuna are well appreciated in class 2 while there is a rejection of highly-smoked salmons

(Frubb, Fcold, Ocoldash, Orubb). The characteristics of products preferred by consumers of classes 1 and 2 are quite similar; the two classes like products with a low intensity of smoking treatment. However, in class 1, consumers reject any spoilage note whereas in class 2, it is not a criterion of rejection, either because these consumers do not recognise this characteristic as a spoilage note or they are not sensitive to this criterion. Consumers in class 3 reject products with a pasty texture (Tpast) and appreciate salty samples, a smoked taste (Salt, Fsalt, Fwood, Owwod) and fish characteristics (Ffish, Ofish). Preferences in cluster 4 are not very far from those of class 3 and are also associated with salt (Salt, Fsalt) but, in this class, the correlation between preference and wood smoke flavour (Fwood) is less strong. In class 5, preferences are linked to the appearance of products (Atrans, a*) and consumers dislike both salty products and the flavour of bacon (Fsalt, Fbacon). Individuals in class 6 attach great importance to the look of smoked salmon with a preference for orange products, homogeneity of appearance and no tearing of the slices (Aoran, Ahomog, Anotear, a*). They reject products characterised by the odour of tuna (Otuna). Unlike class 5, this group seems less sensitive to a salty perception.

The description of the different classes of consumer shows roughly three main characteristics of preference; people who like low-treated products (low-salted and low-smoked) (classes 1 and 2), those who prefer salted and smoked products (classes 3 and 4) and the consumers who choose according to the appearance of the product (orange colour, translucence) (classes 5 and 6).

The relation of physical and chemical measurements and sensory attributes with product appreciation in the different classes allows the identification of the product range to be developed by the industry to respond to consumer preferences. Processing parameters are known to determine to a large extent the final characteristics of cold-smoked salmon (Birkeland, Rora, Skara, & Bjerkeng, 2004; Serot et al., 2004). The salting method used, such

as injection of brine, brine salting or dry salting, leads to products with different colour characteristics, levels of salt in the flesh and texture properties (Sigurgisladottir & al, 2000; Birkeland & al, 2003). Dry salting is likely to satisfy consumers in classes 5 and 6 as it is a more suitable technique for preserving the orange colour and the homogeneous and translucent appearance of the smoked salmon slice as well as its integrity (no tearing). Regarding the level of treatment that discriminates classes 1 and 2 from classes 3 and 4, the choice of processing conditions, like the kind of smoke generator (friction, thermostated hotplate or autocombustion), the type of smokehouse (closed or open air conditioning equipment), smoking temperature (18°C to 26°C), hygrometry and air velocity in the smokehouse and, of course, the time of smoking, leads to either a product with a low smoke note and a perceptible raw fish note or a product with a more intense smoked odour and flavour. The choice of these parameters and the control of the hygienic conditions in the plant explain the smoked salmon characteristics observed and the product diversity offered to consumers (Knockaert, 1990, Cardinal et al., 2001).

3.3. Characterisation of the classes by country of origin

In order to study a possible relation between the country of origin of the consumers and the composition of the different classes, we perform a Chi square test, the largest posterior probability being used to allocate each consumer into one specific class. Table 7 shows the breakdown of the different panellists according to their country of origin in the different classes.

The distribution of the consumers in the different classes according to their country of origin is not balanced. Figure 5 illustrates this breakdown. In some classes, we observe a high proportion of some nationalities and a low proportion of others. Compared to what we should

have obtained in the case of independence between classes and country of origin, the results show:

- Class 1: a high proportion of Belgian, French and German people, and a low number of Italians and English people;

- Class 2: more than 66% Belgians or Italians with a low proportion of German and English consumers.

- Class 3: a high number of French and English people and a lower proportion of Belgians and Germans.

- Class 4: a high proportion of German consumers and an absence of Belgians.

- Class 5: Italians are well represented while French people are in the lowest proportion.

- Class 6 (the smallest): mostly English people and very few French consumers.

Belgian consumers make a large contribution to segments 1 and 2, characterised by a good appreciation of low-smoked salmons. Their opinions about spoiled products are divided (rejection in class 1 and indifference in class 2). Their low number in class 3 confirms their rejection of highly-smoked salmons. Previously, a study of the characteristics of smoked salmon sold on the European market (Cardinal et al., 2004) showed that Belgian products, as well as Italian ones, had the lowest phenol content compared to the products processed in other European countries. These low values of phenol generally indicate a light smoking treatment of products. Therefore, it is possible that the habit of consumption of low-smoked products explains, to some extent, the number of Belgian consumers in classes 1 and 2.

French people are well represented in clusters 1 and 3, corresponding to a great acceptance of products with the flavour and odour of fish and a rejection of spoiled salmons. Their appreciation concerning highly-smoked salmons is divided (rejection in cluster 1 and acceptance in class 3). We observe a low proportion of these consumers in class 5, which is

composed of individuals who dislike salty products and the flavour of bacon but who accord great importance to appearance characteristics.

The relation between the general characteristics of smoked salmon sold on a national market and the preferences of the population seems less obvious in the case of German consumers and perhaps too simple. We have previously shown a high phenol content in German products but the preferences of consumers are not so clear regarding highly-smoked salmon. Indeed, Germans are present in high proportions in clusters 1 and 4, corresponding to individuals having a different sensitivity towards highly-smoked salmons (rejection in class 1 and appreciation of the flavour of wood smoke in class 4).

We find Italian consumers in a high proportion in clusters 2 and 5, which correspond to a preference for salmons with a flavour of fish and an odour of tuna (class 2) and a great importance placed on product appearance in class 5 (orange and translucent). For consumers in these two clusters, spoilage indicators are not really important in the appreciation of the products.

More than 75% of English consumers belong to classes 3 and 5 in similar proportions. Their opinions about salty salmons are divided (consumers in class 3 appreciate them but those in class 5 dislike them). Among the individuals belonging to segment 1 and segment 2, there is a low proportion of English people, denoting that a highly-smoked taste is not really a drawback in terms of preference for English consumers. Considering the characteristics of smoked salmon analysed from the English market (Cardinal et al. 2004), a possible explanation could be the smoke level of the products. Indeed, the average phenol content in British products shows that consumers are used to finding smoked salmon with a rather high smoke note and therefore may prefer to eat this kind of product. The presence of English consumers in classes 5 and 6 denotes the importance of the appearance of smoked salmon (translucent aspect, orange colour, homogeneity).

Among the overall tests of association between class membership and variables characterizing consumers, only the country of origin has a significant effect (P-value lower than 5%).

4. Conclusion

The latent class vector model approach is an interesting method for the analysis of preference ratings. The clustering into a small number of classes and representation in a low dimension space provides an economical model for preference data. Moreover, thanks to a probabilistic assumption, the great advantage of this technique is the implementation of statistical significance tests on the number of classes and dimensions to explore, something that is not possible with the usual MDPREF analysis. The simultaneous representation of product configurations and latent classes gives an exhaustive view of the differences between the expressed preferences in the different clusters.

Concerning the study of smoked salmon preferences, we demonstrate the existence of 6 classes denoting differences in the appreciation of smoked salmon by European consumers. A global analysis leads to three largest classes of consumers (C1, C3 and C5) with different preferences according to the characteristics of the products. The study of the correlations between external data on the products and preferences in the different classes provides the main links between consumer perceptions and the physical and chemical measurements and sensory evaluation of the smoked salmons in order to understand the reasons behind the consumer's choice. Thus, it is possible to relate preferences in each class to product characteristics. We also identify a high correspondence between physical and chemical measurements and the results of sensory profiling of smoked salmons. Among the three clusters, the first one groups people who prefer products with a less smoked note and, above all, with a raw fish note (with two sub-groups according to their sensitivity to a spoilage note).

In the second cluster, consumers prefer a more smoked and salty product (with two sub-groups of different size according to their sensitivity to a pasty texture and the level of smoking) while in the third one, consumers like products with an attractive appearance (with two sub-groups according to their sensitivity to a salty taste).

The study of the relation between the country of origin of the consumers and the composition of the different clusters shows that all the different countries are represented in the different clusters but not in the same proportions. The link between the preferences of a population from one country and the characteristics of the products sold on its national market does not seem so clear. The explanation of the preference only by "they prefer what they generally eat" is certainly too simple. This study shows that European consumers, in general, like different kinds of smoked salmon and select their products according to the intensity of smoking, the intensity of salting and the product appearance (orange colour, tearing). In order to help the consumer choose their preferred product, one suggestion is to put more information on the packaging, for example the level of smoking and salt, parameters that are not usually known by the consumer before they open and eat the product.

From a commercial point of view, the knowledge of consumer preferences should allow the characteristics of smoked salmon to be adapted to meet consumer demand.

Note

References

Aitkin, M., Anderson, D., & Hinde, J. (1981). Statistical modelling data on teaching style. *Journal of the Royal Statistical Society*, *Series A*, Vol. 144, 419-461

Atkinson A. C., & Donev A. N. (1992). Optimum experimental designs. Oxford Statistical Science Series. Clarendon Press, Oxford.

Birkeland, S., Skara, T., Bjerkeng, B., & Rora, A. M. B. (2003) Product yield and gaping in cold-smoked Atlantic salmon (*Salmo salar*) fillets as influenced by different injection salting techniques. *Journal of Food Science*, 68(5), 1743-1748.

Birkeland, S., Rora, A. M. B., Skara, T., & Bjerkeng, B. (2004) Effects of cold smoking procedures and raw material characteristics on product yield and quality parameters of cold-smoked Atlantic salmon (*Salmo salar L.*) fillets. *Food Research International*, 37, 273-286.

Cardinal, M., Knockaert, C., Torrissen, O., Sigurgisladottir, S., Morkore, T., Thomassen, M. S., & Vallet, J.-L. (2001) Relation of smoking parameters to the yield, colour and sensory quality of smoked Atlantic salmon *(Salmo salar)*. *Food Research International*, 34, 537-550.

Cardinal, M., Gunnlaugsdottir, H., Bjoernevik, M., Ouisse, A., Vallet, J. L., & Leroi, F. (2004). Sensory characteristics of cold-smoked Atlantic salmon (*Salmo salar*) from European market and relationships with chemical, physical and microbiological measurements. *Food Research International*, 37, 181-193.

Caroll, J.D (1972). Individual preferences and multidimensional scaling In R. N. Shepard, A. K. Rommey & S. B. Nerlove, Multidimensional scaling theory and applications in behavioural sciences (Vo1. 1, pp 105-155). New York: Seminar Press

Caroll, J.D (1980). Models and methods for multidimensional analysis of preferential choice (or other dominance data) In X. Lagermann & H. Feger, Similarity and choice (pp. 234-289) Bern: Huber

Courcoux P. & Chavanne C. (2001) Preference mapping using a latent vector model. *Food Quality and Preference* 12, pp. 369-372.

Courcoux P., Qannari E.M. & Schlich P. (2006) Sensory workshop: Segmentation of consumers and characterization of cross-cultural differences. *Food Quality and Preference* 17, pp. 658-668.

De Soete G. & Winsberg S. (1993). A Latent Class Vector Model for Preference Ratings. *Journal of Classification* Volume 10 Number pp. 195-218.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation for incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, vol. 39, pp. 1-38

Fedorov V. V. (1972). Theory of optimal experiments. Academic Press, New York.

Greenhoff K. & MacFie H.J.H. (1994). Preference Mapping in Practice in H.J.H. MacFie and D.M.H. Thomson (Eds.) *Measurement of Food Preferences*, Blackie Academic & Professional, Glasgow, pp. 137-166.

Knockaert, C. (1990) *Le fumage du poisson*. Brest: Service de documentation et des publications, IFREMER.

McLachlan, G. J., & Basford, K. E. (1988). Mixture models. Inference and applications to clustering New York: Marcel Dekker.

Schiffman, S. S., Reynold, M. L., & Young, F. W. (1981). Introduction to multidimensional scaling. Theory, methods and applications. Orlando: Academic Press.

Sérot, T., Baron, R., Knockaert, C., & Vallet, J. L. (2004) Effect of smoking processes on the contents of 10 major phenolic compounds in smoked fillets of herring (*Clupea harengus*). *Food Chemistry*, 85, 111-120.

Sigurgisladottir, S., Sigurdardottir, M.S., Torrissen, O., Vallet, J.L., & Hafsteinsson, H. (2000) Effect of different salting and smoking processes on the microstructure, the texture and yield of Atlantic salmon (*Salmo salar*) fillets. *Food Research International*, 33, 847-855.

Figure captions

Figure 1: Plot of products' configuration (as points), individual preferences and mean preference (as vectors) on dimensions 1 and 2.

Figure 2: Representation of the latent class solution on dimensions 1 and 2.

Figure 3: Representation of the latent class solution on dimensions 1 and 3.

Figure 4: Correlations between physical and chemical variables and sensory attributes and the mean preferences expressed in each cluster of the latent class vector model solution.

Figure 5: Breakdown of the consumers in the different classes according to their country of origin.

Table captions

Table 1: List of sensory attributes.

Table 2: Main sensory characteristics of the 30 smoked salmons.

Table 3: List of physical and chemical attributes.

Table 4: Significance test for the number T of classes.

Table 5: Significance test for the number R of dimensions.

Table 6: Preferences in each class.

Table 7: Composition of the different classes in terms of country of origin.

Figure 1: Plot of products' configuration (as points), individual preferences and mean preference (as vectors) on dimensions 1 and 2.
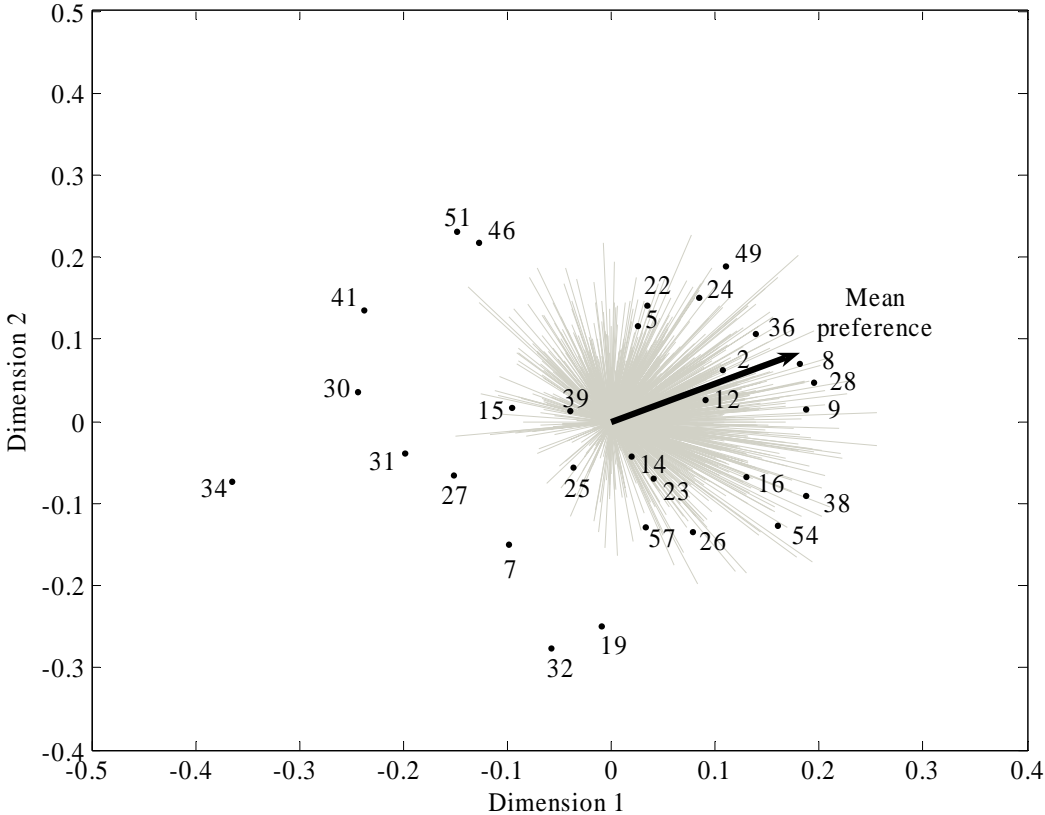
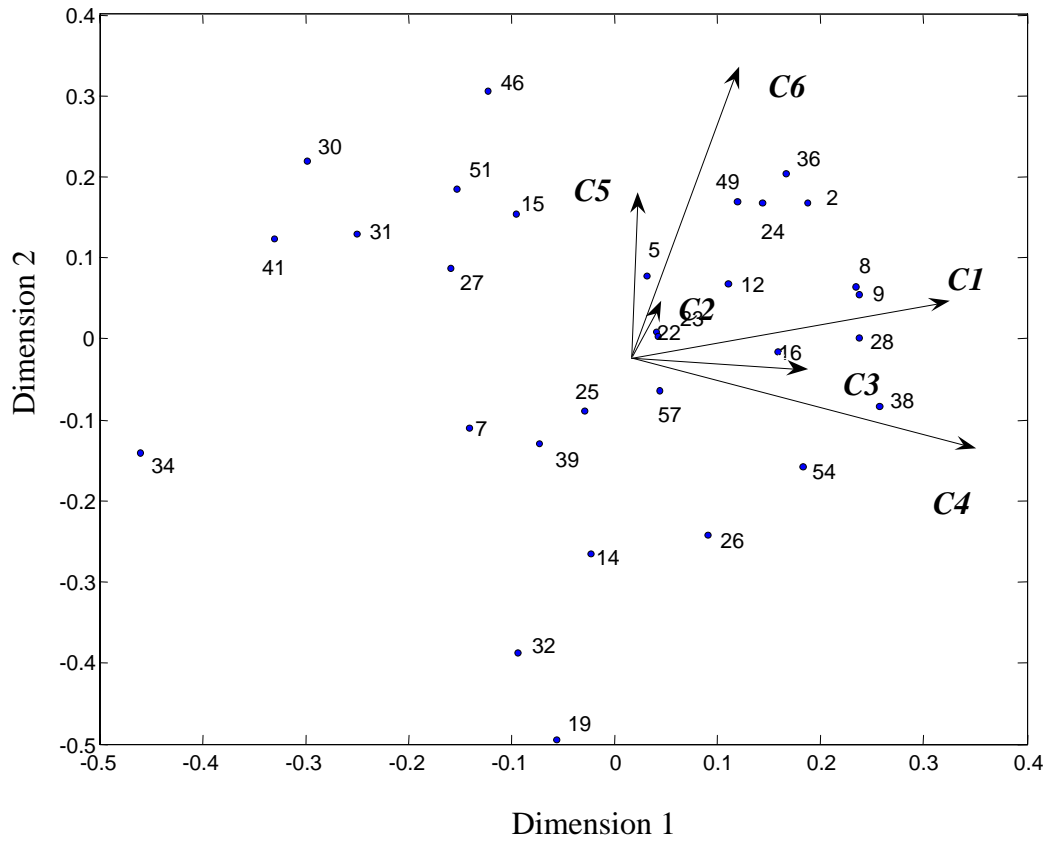Figure 2 : Representation of the latent class solution on dimensions 1 and 2.

Figure 3: Representation of the latent class solution on dimensions 1 and 3.
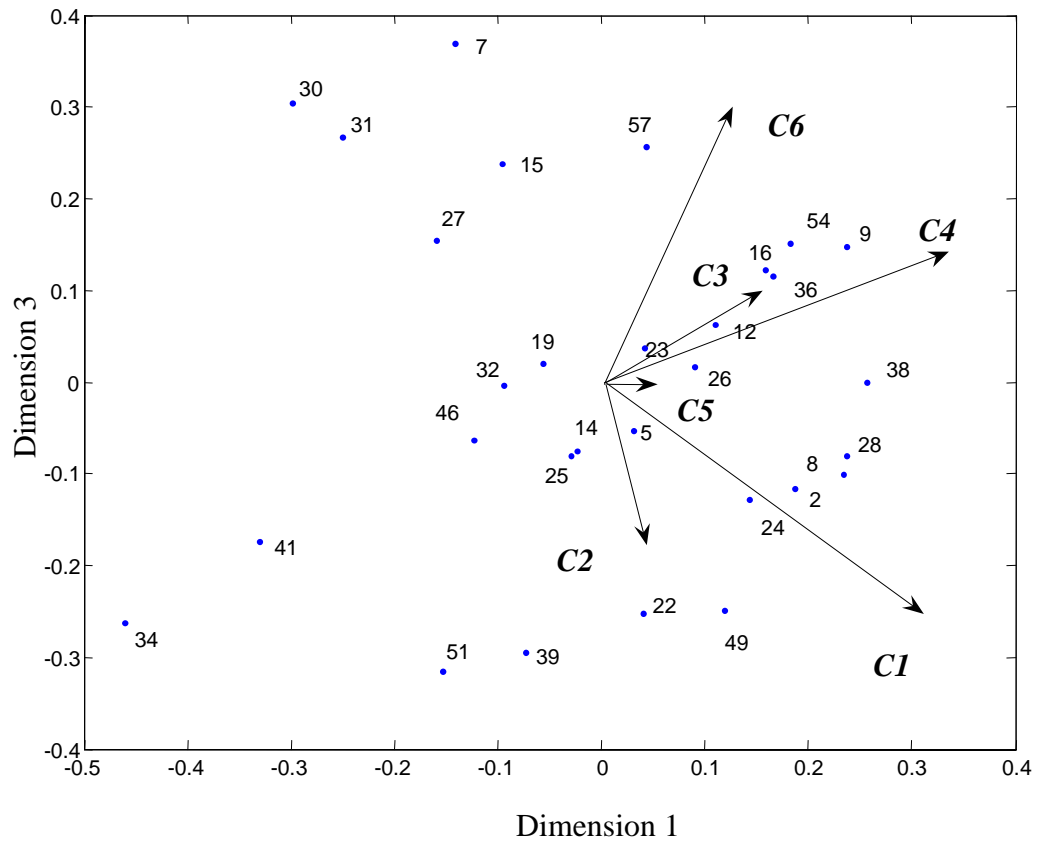
Figure 4: Correlations between physical and chemical variables and sensory attributes and the mean preferences expressed in each cluster of the latent class vector model solution.
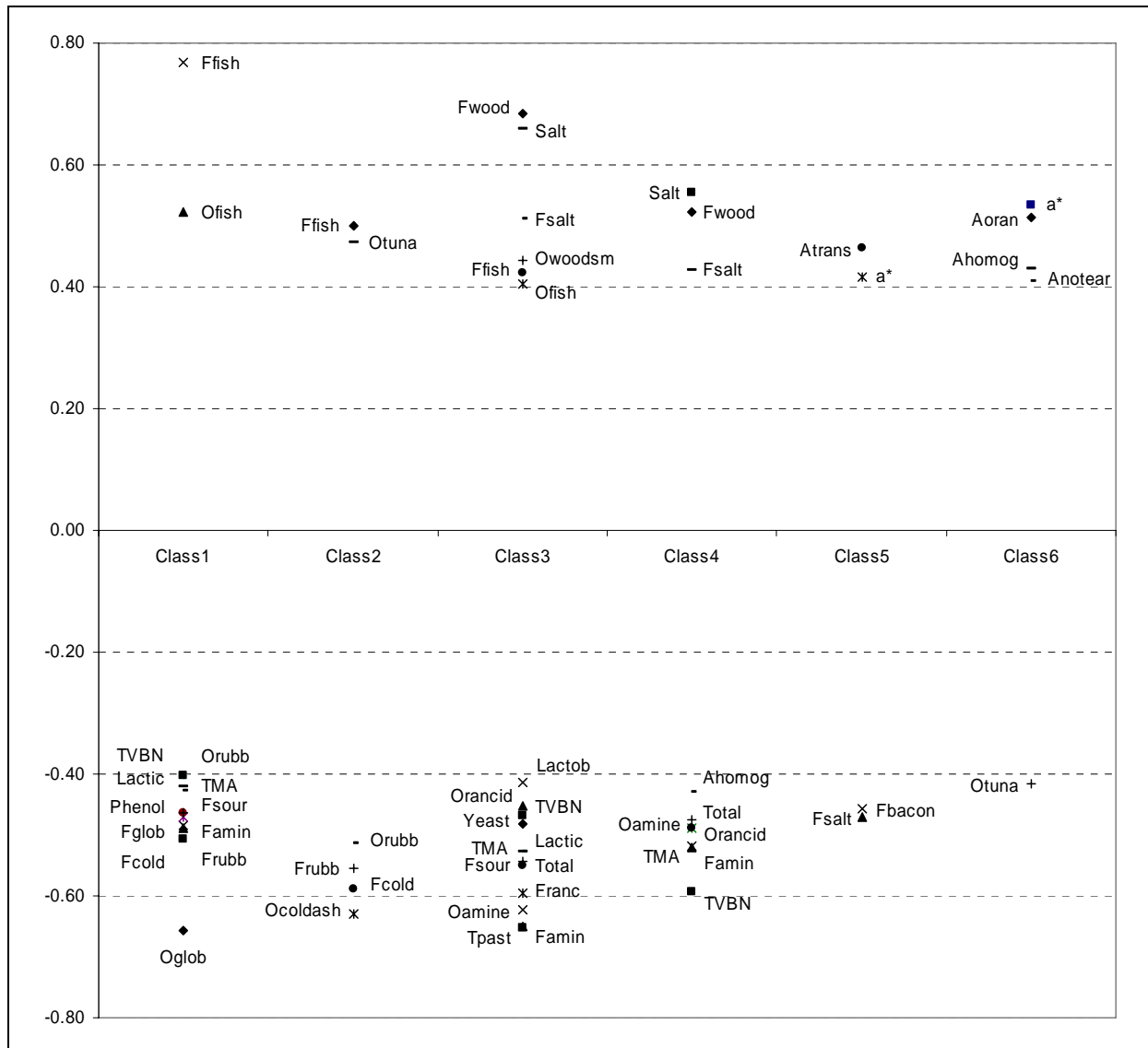
Figure 5: Breakdown of the consumers in the different classes according to their country of origin.
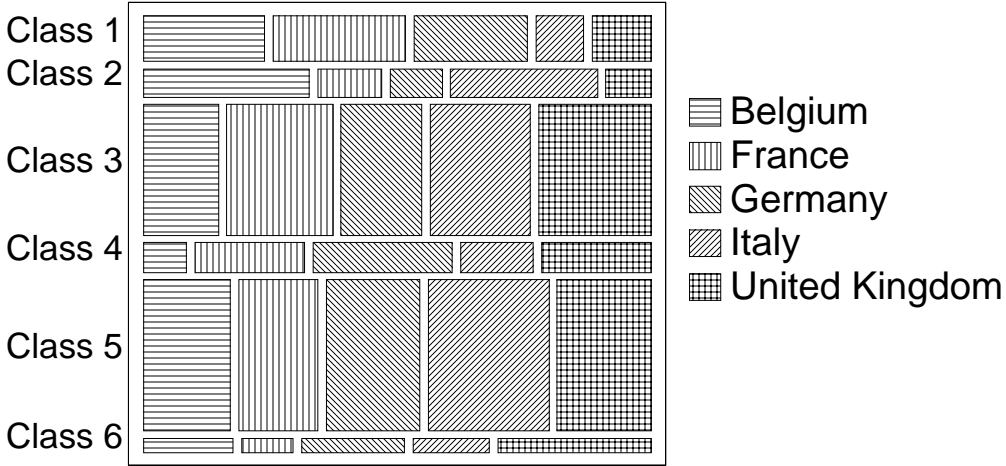
Table 1: List of sensory attributes.

Odour	Oglob: global intensity of the odour
	Owoodsm: odour of wood smoke
	Ocoldash: odour of ashes once the fire is out
	Obacon: odour of bacon
	Otuna: odour of tuna
	Ofish: odour of raw salmon
	Oherr: odour of herring
	Orancid: odour of rancid, like oxidised fish oil
	Oamine: odour of amine, like trimethylamine solution
	Orubb: odour of rubber, like burnt tyres
Appearance	Apink: pink colour
	Aoran: orange colour
	Ahomog: homogeneity of colour
	Anotear: Evaluation of the degree of tearing of the slices (brittleness)
	High degree of tearing (score 0) no tearing (score 10)
	Atrans: translucent appearance
	Afatty: fatty aspect
	Awhstr: white stripes
Flavour	Fglob: global intensity of the flavour
	Fwood: wood smoke
	Fcold: cold ashes
	Fbacon: bacon
	Ffish: raw salmon
	Fherr: herring
	Franc: rancid
	Fsalt: salty taste
	Fsour: sour
	Famin: amine
	Frubb: rubber
Texture	Tfirm: firm
	Tcrunch: crunchy
	Tmelt: melting
	Tfatty: fatty
	Tpast: pasty

Table 2: Main sensory characteristics of the 30 smoked salmons.

Group of products	Product code	Sensory description of each class of products identified by clustering analysis
Group 1	32 - 39 - 25	very pink colour, no translucent appearance, tearing texture, intense global odour and flavour, *herring* or *cold ashes note*, low firmness, salty taste (very high level)
Group 2	23 - 27 - 16	pink colour, no translucent appearance, tearing texture, intense global odour and flavour, *wood fire* and *cold ashes note*, low firmness, salty taste (high), fatty texture
Group 3	38 - 26	intermediate colour between orange and pink, tearing slices, *light wood fire note* and *raw fish note*, low firmness, salty taste, pasty texture, melting texture
Group4	41 - 51	orange and homogeneous colour, translucent appearance and visible stria, intense global odour and flavour, *intense amine note*, *low smoked and low salted*, pasty texture
Group 5	57 - 34	intermediate colour between orange and pink (or grey), non homogeneous colour, intense global odour and flavour, *amine note*, *rubber note*, low salted, pasty texture
Group 6	5 - 30 - 36	orange, translucent appearance, fatty appearance, visible stria, intense global odour, *medium wood fire smoke odour, amine note*, *low level of salt*
Group 7	7 - 15 - 31 - 54	orange and homogeneous colour, no tearing slices, *intense global odour and flavour, cold ash note* or *rubber note*, firm and crunchy texture, salty
Group 8	22 - 24 - 46 - 49	Intense orange and homogeneous colour, translucent appearance, no tearing slices, *low odour, raw fish note* and *wood fire note*, firm and crunchy texture
Group 9	12 - 14 - 19	orange and homogeneous colour, medium intensity of odour, *wood fire smoke odour* and *flavour*, *light cold ash*, firm texture, medium level of salt
Group 10	2 - 28	orange, light to medium odour and flavour, *raw fish note*, *light wood fire note*, *melting texture and fatty texture*
Group 11	8 - 9	orange and homogeneous colour, fatty and translucent appearance, visible stria, no tearing slices, *wood fire smoke odour and flavour*, *fish note*, firm texture

Table 3: List of physical and chemical attributes.

Chemical analysis	TVBN: Total Volatile Basic Nitrogen content
	TMA: Trimethylamine content
	Lipid: Total fat content
	Salt: Salt content
	Phenol: Total phenol content
	Water: Water content
	pH: pH
Physical measurements	L*: Lightness
	a*: Hue parameter (red)
	b*: Hue parameter (yellow)
Microbiological characterisation	Total: Total psychrotrophic count
	Lactic: Total lactic acid bacteria count
	Lactob: Lactobacilli count
	Enterob: Enterobacteriaceae count
	Brocho: Brochothrix count
	Yeast: Yeast count

Table 4: Significance test for the number T of classes.

T (number of classes)	R (number of dimensions)	Log likelihood L_T	Likelihood ratio statistic $2(L_{T+1}-L_T)$ (T vs. $T+1$ classes)	P-value
2	2	- 66772	290	0.00
3	3	- 66627	148	0.00
4	4	- 66553	124	0.01
5	5	- 66491	126	0.00
6	6	- 66428	92	0.06

Table 5: Significance test for the number R of dimensions.

R (number of dimensions) $T = 6$ classes	Log likelihood for T=6 classes and R dimensions L_R	Likelihood ratio statistic $2(L_R\text{-}L_6)$ (6 vs. R dimensions)	P-value
6	- 66428		
5	- 66448	40	0.130
4	- 66504	152	0.00

Table 6: Preferences in each class.

Classes	Disliked products	Preferred products
1	31, 30, 7, 34	38, 22, 2, 49, 8, 28, 24
2	27, 7, 19, 15, 30,	39, 31, 38, 51, 9, 34, 16, 2, 8, 36, 25, 41, 28
3	34, 41, 27, 51, 46	12, 28, 26, 8, 38, 54, 16, 9
4	34, 51, 30, 31, 46, 41	54, 9, 38, 36, 57
5	19, 32, 14, 26, 34	9, 27, 15, 2, 49, 46, 36
6	34, 32, 39, 19, 14	36, 46, 15, 31, 12, 30

Table 7: Composition of the different classes in terms of country of origin.

		Belgium	France	Germany	Italy	United Kingdom	Total
Classes	1	31	34	29	12	15	121
	2	26	10	8	23	7	74
	3	55	78	59	74	83	349
	4	7	18	23	12	18	78
	5	74	67	80	103	80	404
	6	7	4	8	6	12	37
	Total	200	211	207	230	215	1063