Agricultural Produces: Synopsis of Employed Quality Control Methods for the Authentication of Foods and Application of Chemometrics for the Classification of Foods According to Their Variety or Geographical Origin

N. E. Tzouros and I. S. Arvanitoyannis*  
Dept. of Agriculture Crop and Animal Production, School of Technological Sciences, University of Thessaly, Fytokou Str., Nea Ionia Magnesias, 38446 Volos, Hellas (Greece)

Referee: Dr. Ramon Aparicio, Consejo Superior de Investigaciones Científicas (CSIC), Avenida Padra Garcia Tejero 4, 41012 Seville, Spain

* To whom all correspondence should be addressed: I. S. Arvanitoyannis (Dr. Ph.D.), Asst. Professor, Dept. of Agriculture, University of Thessaly, Fytokou Str., Nea Ionia Magnesias, 38446 Volos, Hellas (Greece), e-mail: parmenion33@hotmail.com, Fax: +3042193144, Tel: 3042193104

ABSTRACT: A review of quality control methods and applications of multivariate statistical techniques on the authentication and classification of agricultural products is presented. The products reported within the frame of this article were vegetables, fruits, juices, jams, wines, cereals, bakery products, oils, tea, coffee, honey, sugar-syrups, salad dressings, and gums. The perspective of multivariate statistics as a promising tool to authenticate and classify these food products according to their geographical origin or variety was demonstrated. Several representative figures and informative synoptical tables for agricultural food products were provided both for the quality control methods employed and the multivariate analyses implemented.

KEY WORDS: chemometrics, multivariate analysis, quality control, authentication, PCA, cluster analysis, discriminant analysis.

ABBREVIATIONS: ACF, autocorrelation and autocovariance functions; Acid, acidity; AMT, angle measure technique; ANN, artificial neural network; ANNS, artificial neural network simulation; ANOVA, analysis of variance; CDA, canonical discriminant analysis; CVA, Canonical Variance Analysis; DA, discriminant analysis; DHS-GC, dynamic headspace-gas chromatography; DryMat, dry matter; FTIR, fourier transform infrared; GLC, gas liquid chromatography; GPA, generalized procrustes analysis; GS-SP, gas chromatography-sniffing port analysis; HPAEC, high-performance anion exchange chromatography; HPLC, high-performance liquid chromatography; HRGC, high-resolution gas chromatography; IR, infrared; Ithick, instron thickness; Katak, katalase activity; LDA, linear discriminant analysis; LPO, lipoygenase; MAP, modified atmosphere packaging; MIRS, mid-infrared spectroscopy; Moist, expelled moisture; MS, mass spectrometry; NIR, near infrared; NMR, nuclear magnetic resonance; NN, neural networks; PC, principal component(s); PCA, principal component analysis; PCR, principal component regression; Pgreen, Pred, penetrometer at green and red sides of apple, respectively; PLS, partial least square; PDO, peroxidase; QDA, quantitative descriptive analysis; RPCA, Regression Principal Component Analysis; RSM, response surface methodology; SDD, size and distance distribution; SVD, singular value decomposition; TI, time-intensity.
I. INTRODUCTION

A. Multivariate Techniques

Thirty years have already passed since chemometrics resumed its modern form, but there has been no consensus regarding its definition yet. There are two prevailing trends in all branches of technology and science vis-à-vis chemometrics: (1) experimentation is demanding more and more resources, becoming more and more expensive, while in most cases the number of 'objects' (observations, cases, or samples) is fairly small, and tends to decrease further with time and (2) more and more properties, variables, spectra, chromatograms, electrophoretic samples, etc., on the samples or experimental runs (Wold and Sjöström, 1998).

Multivariate analytical techniques are being widely applied in industry, government, and university-related centers (Hair et al., 1995). It is anticipated that multivariate analysis methods will predominate in the future and will result in drastic changes in the manner in which researchers think about problems and design their research (Hardyck and Petrinovich, 1976). PCA targets to the reduction of the dimensionality of a data set in which there is a large number of intercorrelated variables, while retaining as much as possible of the information present in the original data. The reduction is achieved through a linear transformation to a new set of uncorrelated variables (PC scores) in which the first few ones will express most of the variation of the original variables. Linear discriminant analysis (LDA) is usually applied to a subset of scores as variables. Principal component regression (PCR) can be carried out on a dummy variable or variables that indicate class membership. PLS models the relationship between the data sets using a series of local least-squares fits. Similar to PCA, an axis rotation method is occasionally applied (Kemsley, 1996). It is mainly employed to relate blocks of variables measured on sets of objects (Geladi, 1988).

Factor analysis transforms a n-dimensional data structure to another with considerably less dimensions, like PCA, but gives the opportunity to the researcher to select between an orthogonal or an oblique rotation of the factors. A orthogonal rotation should lead to uncorrelated factors, whereas an oblique one might result in correlated factors. Correspondence factor analysis has the advantage over other techniques that it relates the variables to those objects for which they are particularly meaningful, and around a point in the graph representing a variable one finds the objects scoring high with regard to this particular variable. Nonlinear mapping (also called multidimensional scaling) is a dimension reducing method that attempts to retain the distances between data points as well as possible. Clustering techniques aim at grouping variables or objects in terms of their similarities/properties and are divided into hierarchical and nonhierarchical ones (Massart et al., 1988). Procrustes Analysis was first introduced as a method for matching two configurations when two assessors (in the case of sensory evaluation) scored the same set of samples. The analysis involves the mathematical operations of transformation to a common origin, rotation of axes, and possibly an isotropic scale change, by which one configuration is made to approach the other as nearly as possible (Arnold and Williams, 1998).

The current chemometric research is directed toward more sophisticated technology, which is a direct result of tackling more difficult data analysis problems. Feature selection, use of neural networks, and pattern recognition methods and validating predictions conducted by soft modeling methods are the topics that are continuously gaining interest (Lavine, 1998). Chemometrics can be visualized as the implementation of mathematical and statistical tools to the interpretation of patterns in multivariate data. They are fast and precise in analyzing instrumental results. Chemometrics are applied in process monitoring and control, determination of geographical origin, detecting adulterated products and sensory evaluation (Bailey and Rohback, 1994). The applications of chemometric methods to the identification, authentication or quality determination of agricultural and minimally processed agricultural products are given, synoptically, in Table 1.

B. Quality Control Methods Employed for Food Classification

The need to determine the origin of an oil has become mandatory after the introduction of the
<table>
<thead>
<tr>
<th>Statistical method</th>
<th>Products</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correspondence Analysis</td>
<td>Olive oil</td>
<td>Nielsen et al., 1998</td>
</tr>
<tr>
<td>Canonical Discriminant Analysis</td>
<td>Olive oil</td>
<td>Stefanoudaki et al., 1997</td>
</tr>
<tr>
<td>Canonical Variate Analysis</td>
<td>Potatoes</td>
<td>Anderson et al., 1999</td>
</tr>
<tr>
<td>Partial Least Squares</td>
<td>Peppermint oil</td>
<td>Chialva et al., 1993</td>
</tr>
<tr>
<td>Partial Least Squares</td>
<td>Patchouli and amyris oils</td>
<td>Peppard, 1994</td>
</tr>
<tr>
<td>Partial Least Squares</td>
<td>Mayonnaise</td>
<td>Indahl et al., 1999</td>
</tr>
<tr>
<td>Partial Least Squares</td>
<td>Tomatoes</td>
<td>Lee et al., 1999</td>
</tr>
<tr>
<td>Partial Least Squares</td>
<td>Wheats</td>
<td>Dowell, 1998</td>
</tr>
<tr>
<td>Partial Least Squares</td>
<td>Rice</td>
<td>Meullinenet, 1999</td>
</tr>
<tr>
<td>Partial Least Squares</td>
<td>Soy sauce</td>
<td>Iizuka &amp; Aishima, 1997</td>
</tr>
<tr>
<td>Partial Least Squares</td>
<td>Wheat flour</td>
<td>Delwiche &amp; Weaver, 1994</td>
</tr>
<tr>
<td>Partial Least Squares</td>
<td>Potatoes</td>
<td>Lavine, 1998</td>
</tr>
<tr>
<td>Principal Components Regression</td>
<td>Olive oil</td>
<td>Aparicio et al., 1992</td>
</tr>
<tr>
<td>Principal Components Regression</td>
<td>Mayonnaise</td>
<td>Indahl et al., 1999</td>
</tr>
<tr>
<td>Principal Components Regression</td>
<td>Wheat flour</td>
<td>Meullinenet, 1999</td>
</tr>
<tr>
<td>Stepwise Discriminant Analysis</td>
<td>Olive oil</td>
<td>Morales et al., 1994</td>
</tr>
<tr>
<td>Stepwise Discriminant Analysis</td>
<td>Pasta</td>
<td>Fardet et al., 1998</td>
</tr>
<tr>
<td>Cluster Analysis</td>
<td>Almonds</td>
<td>Prats-Moya et al., 1997</td>
</tr>
<tr>
<td>Cluster Analysis</td>
<td>Peel oils from acid citrus</td>
<td>Tamura et al., 1993</td>
</tr>
<tr>
<td>Cluster Analysis</td>
<td>Black, oolong and green tea</td>
<td>Kobayashi et al., 1993</td>
</tr>
<tr>
<td>Cluster Analysis</td>
<td>Olive oil</td>
<td>Morales et al., 1995</td>
</tr>
<tr>
<td>Cluster Analysis</td>
<td>Tomatoes</td>
<td>Lee et al., 1999</td>
</tr>
<tr>
<td>Linear Discriminant Analysis</td>
<td>Mayonnaise</td>
<td>Indahl et al., 1999</td>
</tr>
<tr>
<td>Linear Discriminant Analysis</td>
<td>Honey</td>
<td>Sanz et al., 1995</td>
</tr>
<tr>
<td>Multidimensional Scaling</td>
<td>Olive oil</td>
<td>Morales et al., 1995</td>
</tr>
<tr>
<td>Generalised Procrustes Analysis</td>
<td>Walnut fruit</td>
<td>Sinesio &amp; Moneta, 1997</td>
</tr>
<tr>
<td>Generalised Procrustes Analysis</td>
<td>Corn-soy breakfast cereals</td>
<td>Faller et al., 1998</td>
</tr>
<tr>
<td>Factorial Discriminant Analysis</td>
<td>Orange juice</td>
<td>Twomey et al., 1995</td>
</tr>
</tbody>
</table>
‘Appelation of origin’ on olive oils and is based on the well-defined physicochemical attributes of each cultivar (Stefanoudaki et al., 1998). In the case of virgin olive oil, authenticity issues concern the geographical origin, method of processing, and the variety (Aparicio et al., 1997). Recent international regulations have established analytical criteria to define olive oil authenticity (detection of adulterations with seed oils or other solvent-extracted oils) and quality grade (extra virgin, virgin, ‘lampante’, ‘refined’, ‘pure’ etc.) (Sacchi et al., 1998).

Physical and chemical tests proposed to detect adulterants in olive oil include ultraviolet spectroscopy based on absorbance in the regions of 208 to 210 and 310 to 320 nm for virgin olive oil adulterated with refined olive oil, gas chromatography, HPLC analysis of the triglycerides and fatty acid composition, NMR, and spectrophotofluorometry (Lai et al., 1994). NIR can be used on-line and in-line to monitor production processes in the olive oil industry (Wesley et al., 1996). The detection of the adulteration of olive oil with oils of a similar composition or in grading olive oils is possible through the evaluation of nonacylglycerol components, such as sterols, fatty alcohols, and their esters, and/or the examination of the stereospecific structure of the triglycerols (Li-Chan, 1994).

Instant coffee is more susceptible to adulteration than green coffee because of its powdered form. The most frequently occurring coffee adulterations are as follows: (1) adulteration with coffee substitutes, (2) mixing of two species (cheaper robusta with pure arabica), and (3) mixing of expensive coffee beans from one region with cheaper ones coming from other areas. The detection of instant coffee adulteration is based on the determination and analysis of mineral or caffeine content, carbohydrate contents (measurement of the proportion of adulterant, i.e., chicory, barley, figs, and caramel), and total xylose, total glucose, free fructose, and free mannitol (their presence in high levels indicates adulteration). High-performance anion-exchange chromatography (HPAEC) is commonly employed for the determination of carbohydrate profile of instant coffee as a tool for detecting any possible adulteration. Another method successfully employed for the detection of adulteration is infrared spectroscopy (IR) (Briandet et al., 1996). Scanning NIR spectrophotometry was shown to contribute to the determination of flour quality (Delwiche and Weaver, 1994). NIR is also employed in the determinations of adulterated orange juice (Twomey et al., 1995).

The geographic origin of a honey can be determined from the microscopic analysis of pollen sediment, concentrated by centrifuging the honey dissolved in water (Sanz et al., 1995). The main criteria for the characterization of unifloral honeys are sugar content, electrical conductivity, and pH analysis, complemented by pollen analysis, and in some cases thixotropy and α-amylase activity. The compounds used as indicators of citrus honey quality are aroma compounds (e.g., methyl anthranilate), amino acids and their degradation products, aromatic acids and their esters, aromatic and flavonoids such as hesperetine, which is a degraded carotenoid (Mateo and Bosch-Reig, 1998).

The quality and external appearance of apples are important for the consumer. These quality attributes can be described by color, texture, flavor (volatile compounds) and taste in addition to physical attributes such as size, shape, and firmness (Karlsen, 1999). Quality attributes such as aroma, intensity and acceptability, sweetness, acidity, and ripe fruit flavor were shown to be significantly affected by the firmness of kiwifruit (Esti et al., 1998). Implementation of Factorial Discriminant Analysis of data obtained from the analysis of flavonone glycosides led to the differentiation of varieties of grapefruits and sweet oranges (Sanz et al., 1995).

Various chromatographic methods were used for the quantification of anthocyanins in fruit juices and purees. Gas chromatography was also used to quantify the main sugars and acids in fruit and vegetables (Defernez et al., 1995). Clear beverages (e.g., apple juice) are intended to remain transparent until their consumption, whereas cloudy beverages (e.g., orange juice) should have a pleasant appearance that is stable throughout their shelf life. Turbidity has been used to study the intensity of light scattered form nonhomogeneities in apple juice and wine (Carrasco and Siebert, 1999).
II. APPLICATIONS OF CHEMOMETRICS TO AGRICULTURAL PRODUCTS

A. Vegetables

Statistical and artificial neural network (ANN) pattern recognition techniques applied to NIR spectra of soy sauce samples collected in Japan led to correct classification ratios of 84.2% for PLS2 and 76.3% for ANN, suggesting typical quality differences in soy sauce among three regions of Japan (Iizuk and Aishina, 1997).

Natural, frozen, and canned legumes were analyzed by flame photometry to determine their composition in water, protein, ash, carbohydrates, fat, and saline. Factor analysis was applied to all data set leading to a clustering of samples according to their processing method. In the resulting graphic presentation (Figure 1), which included 92.71% of the global information, factor 1 expressed the ‘nutrients’ in such a way that the samples situated to the left had greater nutrient content (Barrado et al., 1994).

Application of House of Quality in translation of consumer needs into sensory attributes for peas showed that the most important attributes are pea taste, the absence of bitter taste, juiciness, crispiness, and the absence of tough skin (Bech et al., 1997).

The impact of storage conditions on the flavor of stored French beans after rehydration was evaluated by gas chromatography — sniffing port analysis (GS-SP) and GC-mass spectrometry of volatile compounds, by quantitative descriptive analysis (QDA), and hedonic sensory evaluation. The main observations were (1) ‘French bean’ and ‘sweet’ attributes intensities decreased at elevated temperature and \(a_w\) range 0.3 to 0.5, while the attributes for ‘chemical’ and ‘burned’, ‘musty’, and ‘bitter’ increased, (2) exposure to light at \(a_w\) 0.1 decreased ‘French bean’ scores and increased scores for ‘chemical’ and ‘burned’, (3) volatiles are responsible for the change in flavor and appreciation of dried French beans by storage conditions (van Ruth et al., 1995).

Two legumes (kidney beans and chickpeas) and two nut kernels (almonds and hazelnuts) were compressed with a Universal testing machine at
<table>
<thead>
<tr>
<th>Quality Analysis Method</th>
<th>Products</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>GC</td>
<td>Olive oil</td>
<td>Morales et al., 1994</td>
</tr>
<tr>
<td></td>
<td>Peel oils from acid citrus</td>
<td>Tamura et al., 1993</td>
</tr>
<tr>
<td></td>
<td>Essential oils</td>
<td>Perry et al., 1999</td>
</tr>
<tr>
<td></td>
<td>Roselle tea</td>
<td>Chen et al., 1998</td>
</tr>
<tr>
<td></td>
<td>Patchouli and amyriss oils</td>
<td>Peppard, 1994</td>
</tr>
<tr>
<td></td>
<td>Orange juice</td>
<td>Elmore et al., 1994, Muccarone et al., 1998</td>
</tr>
<tr>
<td>Reversed Phase HPLC</td>
<td>Black, oolong, green tea</td>
<td>Kobayashi et al., 1993</td>
</tr>
<tr>
<td>HPLC</td>
<td>Olive oil</td>
<td>Stefanoudaki et al., 1997</td>
</tr>
<tr>
<td>1H-NMR</td>
<td>Olive oil</td>
<td>Sacchi et al., 1998</td>
</tr>
<tr>
<td></td>
<td>Almond oil</td>
<td>Remaud et al., 1997</td>
</tr>
<tr>
<td></td>
<td>Cinnamon oil</td>
<td>Remaud et al., 1997</td>
</tr>
<tr>
<td>Mid-IR</td>
<td>Seed oils</td>
<td>Dupuy et al., 1996</td>
</tr>
<tr>
<td>18O,818O</td>
<td>Olive oil</td>
<td>Angerosa et al., 1999</td>
</tr>
<tr>
<td>HRGC</td>
<td>Peppermint oil</td>
<td>Chialva et al., 1993</td>
</tr>
<tr>
<td></td>
<td>Olive oil</td>
<td>Aparicio &amp; Morales, 1998</td>
</tr>
<tr>
<td></td>
<td>Strawberry jam</td>
<td>Lesscheave, 1991</td>
</tr>
<tr>
<td>SNIF</td>
<td>Almond-cinnamon oils</td>
<td>Remaud et al., 1997</td>
</tr>
<tr>
<td>FTIR</td>
<td>Olive oil</td>
<td>Lai et al., 1994</td>
</tr>
<tr>
<td></td>
<td>Instant coffee</td>
<td>Downey &amp; Boussion, 1996</td>
</tr>
<tr>
<td>NIR</td>
<td>Vegetable oils</td>
<td>Sato, 1994</td>
</tr>
<tr>
<td></td>
<td>Olive oil</td>
<td>Wesley et al., 1994</td>
</tr>
<tr>
<td></td>
<td>Mayonnaise</td>
<td>Indahl et al., 1999</td>
</tr>
<tr>
<td></td>
<td>Soy sauce</td>
<td>Iizuka &amp; Aishima, 1997</td>
</tr>
<tr>
<td></td>
<td>Orange juice</td>
<td>Tworney et al., 1995</td>
</tr>
<tr>
<td></td>
<td>Wheat</td>
<td>Dowell, 1998, Delwiche &amp; Weaver, 1994</td>
</tr>
<tr>
<td></td>
<td>Potatoes</td>
<td>Laving, 1998</td>
</tr>
<tr>
<td>Positive-ion fast atom bombardment</td>
<td>Olive oil</td>
<td>Lamberto &amp; Saitta, 1995</td>
</tr>
<tr>
<td>Thermogravimetric techniques</td>
<td>Rapeseed oils</td>
<td>Wesolowski &amp; Ercinska, 1998</td>
</tr>
<tr>
<td>C13/C12</td>
<td>Olive oil</td>
<td>Angerosa et al., 1997</td>
</tr>
<tr>
<td>Sensory evaluation</td>
<td>Olive oil</td>
<td>Morales et al., 1995</td>
</tr>
<tr>
<td></td>
<td>Tomatoes</td>
<td>Johansson et al., 1999</td>
</tr>
<tr>
<td></td>
<td>Walnut fruit</td>
<td>Sinesio &amp; Moneta, 1997</td>
</tr>
<tr>
<td></td>
<td>Blueberry soups</td>
<td>Rivi et al., 1997</td>
</tr>
<tr>
<td></td>
<td>Wines</td>
<td>Heymann &amp; Noble, 1989</td>
</tr>
<tr>
<td></td>
<td>Rice</td>
<td>Meullenet, 1999</td>
</tr>
<tr>
<td></td>
<td>Peas</td>
<td>Poulsen et al., 1997</td>
</tr>
<tr>
<td>HPLS</td>
<td>Blueberries</td>
<td>Ehlenfeldt et al., 1994</td>
</tr>
<tr>
<td>HSGC</td>
<td>Mandarin juice</td>
<td>Moshonas &amp; Shaw, 1997</td>
</tr>
<tr>
<td>Texture analysis</td>
<td>Rice</td>
<td>Meullenet, 1999</td>
</tr>
<tr>
<td>Image analysis</td>
<td>Barley</td>
<td>Moral et al., 1998</td>
</tr>
<tr>
<td></td>
<td>Pasta</td>
<td>Fardet et al., 1998</td>
</tr>
<tr>
<td>Inductively coupled plasma for organic elements</td>
<td>Almonds</td>
<td>Prats-Moya et al., 1997</td>
</tr>
<tr>
<td>GC-MS</td>
<td>Roselle tea</td>
<td>Chen et al., 1998</td>
</tr>
<tr>
<td>Conductivity, pH, chemical analysis of constituents</td>
<td>Honey</td>
<td>Sanz et al., 1998</td>
</tr>
<tr>
<td>Artificial neural networks</td>
<td>Ricebran</td>
<td>Hussain et al., 1996</td>
</tr>
<tr>
<td></td>
<td>Mustard</td>
<td>Hussain et al., 1996</td>
</tr>
<tr>
<td></td>
<td>Safflower</td>
<td>Hussain et al., 1996</td>
</tr>
<tr>
<td></td>
<td>Palmolin</td>
<td>Hussain et al., 1996</td>
</tr>
<tr>
<td></td>
<td>Sunflower</td>
<td>Hussain et al., 1996</td>
</tr>
<tr>
<td></td>
<td>Patchouli and amyriss oils</td>
<td>Peppard, 1994</td>
</tr>
<tr>
<td></td>
<td>Soy sauce</td>
<td>Iizuka &amp; Aishima, 1997</td>
</tr>
<tr>
<td></td>
<td>Instant coffee</td>
<td>Downey &amp; Boussion, 1996</td>
</tr>
<tr>
<td></td>
<td>Bread</td>
<td>Horimoto et al., 1995</td>
</tr>
<tr>
<td>Atomic Emission Spectrometry</td>
<td>Potatoes</td>
<td>Anderson et al., 1999</td>
</tr>
</tbody>
</table>
FIGURE 1. Representation of the first two factors of the factor analysis (factor 1: nutrients; factor 2: saline). Cp: chick-pea for natural and pea for canned legumes; cb: coloured bean; fgb: green bean; fp: pea; cgb: green bean; fbb: broad bean; and wb: white bean (Barrado et al., 1994).
various $a_w$ levels, at 22°C. The nut kernels demonstrated a much smoother force-displacement curve due to their high oil content, while the effect of absorbed moisture on their texture was more moderate than on legumes (Borges and Peleg, 1997). Sixteen frozen pea samples were sensory evaluated by employing 14 descriptors. The samples were separated on the PC-plot (PC1 vs. PC4) according to the following attributes: juiciness, sweetness, crispiness, mealiness, and bitterness along the first principal component; hardness, pod odor, and pod taste along PC4 (Poulser et al., 1997).

Preference to ecologically when compared with conventionally grown tomatoes was also investigated. PCA on sensory analysis data led to score plots, where conventional Gitana was expressed by PC2 as being high in sweetness and tomato-taste intensity, while conventional Aromata and ecological Jamaica were mirrored by PC1, the first one as firm and the second as high in red color intensity. It was concluded that varietal differences play a more important role on the variability of tomatoes than the growing conditions (Johanson et al., 1999). The quality of commercial diced tomatoes was described by physicochemical and sensorial parameters. Analysis of these parameters with PCA showed that quality could be satisfactorily correlated to d- and L-lactic acid, total acidity and volatile acidity, contents (Porretta, 1994). The texture of aseptic, cold-fill, and hot-fill tomatoes was assessed both with descriptive analysis and instrumental measurements. Sensory and instrumental data were compared and correlated using correlation analysis, PCA, cluster analysis, and PLS. The biplot of PLS1 vs. PLS2 showed that samples could be satisfactorily separated according to their processing method. Cold-fill tomatoes were correlated to firmness variables, whereas hot-filled ones to cohesiveness, chewy, and metallic variables (Lee et al., 1999).

Ecologically and conventionally grown carrots of 2 consecutive years were evaluated sensorially. A correlation analysis was performed to investigate whether there is an association, if any, among the sensory attributes. In the first year, conventionally grown carrots were sweeter and crunchier, while ecological ones were harder.

However, in the second year, ecological ones had a stronger aftertaste than conventional ones (Haglund et al., 1999).

Flavor volatiles of carrots blanched from 0 to 300 s were identified by gas chromatography and mass spectrometry, while sensory evaluation was applied to the vegetables by a trained panel. Ratings on quality attributes of color, texture, raw carrot aroma, sweetness, flavor, and overall impression decreased with blanching time, whereas cooked carrot aroma increased. Correlations among blanching times, flavor volatiles, and sensory attributes were found (Shamaila et al., 1996).

The quality of leafy vegetables both frozen and sterilized was evaluated with instrumental and sensory analysis. For sterilized spinach two PCs explained 56% of variance in the data. Steam-blanching sterilized spinach had more spinach-like aroma, radiation blanched had a strange mouthfeel and a bitter taste, while water-blanching spinach was soft and flat in taste (Ponne et al., 1994).

The aroma of fresh and hot-air dried bell peppers (Capsicum annuum) was evaluated by sensory and instrumental methods. Almost 75% of variation was effectively described by the first three PCs. Sample scores on the first two PCs showed that the unripe green peppers were separated from the ripe yellow and ripe red peppers on PC1, while white fruits were separated from the others on PC2. Sample scores on the first two PCs showed that the green and white peppers were separated from the red and yellow fruits on PC1. On PC2 (Figure 2) the green and white, and the red and yellow ones were separated from each other (Luning, 1995).

Sweet potatoes, after having been blanched for various time periods and canned in syrup, were analyzed with sensory evaluation and texture analyzer (uniaxial compression). Data were treated with ANOVA. Blanching time and temperature affected firmness significantly. Sensory texture and overall acceptability reached their highest values for samples blanched at 62°C for 30 or 45 min prior to canning (Truong, 1998). Potatoes were classified by variety using isoelectrophoretic focusing patterns and a feed-forward neural net. NIR was employed in conjunction with PLS to determine the total reducing content of sliced potatoes (Lavine, 1998).
FIGURE 2. (a) PCA sample scores of fresh green and red bell peppers cv. Mazurka (GM, RM), white cv. Blondy (WB), and yellow cv. Kelvin (YK); (b) PCA sample scores of hot-air dried green and red bell peppers cv. Mazurka, white cv. Blondy, and yellow cv. Kelvin (Luning et al., 1995).
B. Fruits

Strawberry, raspberry, and apple purees were classified according to fruit type by using Fourier transform infrared spectroscopy in conjunction with discriminant analysis. One hundred and forty-nine spectra of purees were classified with 100% success, 98.3% success for classifying strawberry, and 75% for classifying raspberry, according to whether the fruits were fresh or freeze thawed (Defernez et al., 1995). The mid-infrared spectra of raspberry and strawberry purees were acquired on two Fourier-transform spectrometers. The application of discriminant analysis by PCA showed a rather low percentage (55.5%) of successful classification of two fruit species (Figure 3) (Holland et al., 1997).

The evaluation of combined GC/MS/FTIR data sets of strawberry aroma was also performed. Cluster analysis showed that lactones, esters, alcohols, and acids identified in strawberry aroma can be separated into different clusters for a chromatogram section of a strawberry Lickens-Nickerson pentane extract (Nikiforov et al., 1994).

Sources of variation affecting assessment of whole apple fruit quality were identified and multivariate relationships between analytical and sensory characteristics were examined. Highly colored ‘McIntosh’ apples were significantly heavier, sweeter, and had a higher pH than the less colored ones, while highly colored ‘Jonagold’ were sweeter and more fruity than the rest of the same variety (Dever et al., 1995).

The apples were subjected to instrumental and sensory analysis followed by analysis. PCA analysis of sensory data identified an odor/flavor-factor and a texture factor along the first and second principal component, respectively. Sensory hardness, chewiness, and mushiness were well correlated to instrumentally measured force and work required for flesh penetration. A similar observation was reported for sensory (odor and flavor) attributes with volatile composition and texture data (Karlsen, 1999). The sensory evaluation of peeled and unpeeled apples from 12 varieties was performed. Internal Preference Mapping indicated four major dimensions of preference, with the most important dimension highly correlated with texture attributes. Further-

more, clusters of the investigated varieties were formed successfully (Daillant-Spinler et al., 1996).

Seventy-two apples were artificially manipulated to have different characteristics. They were sensory evaluated in terms of ‘mealiness’ and ‘firmness’ by examining six sensory variables. Seven more physical and chemical variables, instrumentally determined, were included. To facilitate the analysis, the data were grouped in two sets and studied separately by PCA.

Highbush and Rabbiteye blueberries cultivars were analyzed with HPLC for organic acids. A PCA based on the relative percentages of citric, succinic, and quinic acid disclosed that Highbush (V. corymbosum) blueberries are taxonomically different from Rabbiteye (V. ashei) blueberries (Ehlenfeldt et al., 1994). Sensory quality of modified atmosphere packaged (MAP) Highbush blueberries was studied in relation to storage temperature, wrapping film type, and initial high oxygen atmosphere in conjunction with PCA. The storage temperature emerged as the most important factor followed by the film type. High storage temperature was correlated with bitter taste and storage flavor, while low storage temperature with acidic taste and blueberry flavor (Rosenfeld et al., 1999).

Sensory evaluation of walnut fruit from five European countries was performed and PCA and GPA (Generalized Procrustes Analysis) were applied to the sensory data. From the sensory attributes only ‘shell whiteness’ showed significant cultivar country interaction and variations both among cultivars and countries. PCA and GPA plot showed a differentiation of samples from different countries (Figure 4) (Sinesio and Moneta, 1997).

The essential oils from the steam distillation of the leaves of avocados of three racial lineages were assayed by the use of capillary column gas-liquid chromatography and mass spectrometry. In a three-dimensional plot of the three principal factors, the cluster consisting of five Mexican cultivars formed a well-separated group. West Indian and Guatemalan races have less differences between them, based on chemical data, than from the Mexican race (King and Knight, 1992).

The perception of fruit as a snack by women was examined. The treatment of sensory analysis data with PCA resulted in two factors accounting
FIGURE 3. Two-dimensional score plot for the first two principal components defined by the corrected spectra of the raspberry and strawberry samples (Holland et al., 1997).
FIGURE 4. (a) attribute and sample space derived from PCA for walnut fruits. (Repeated samples are connected with arrows); (b) consensus coordinates for samples and attributes derived from GPA for dimensions 1 and 2. (Sinesio and Moneta, 1997.)
for 59 and 13% of the variance, respectively. Factor 1 showed a clear separation of the apple and banana from the orange and kiwi. Bananas and apples were perceived as being suitable to eat on trips, because they caused no mess and were convenient. Oranges and kiwis were regarded as more refreshing, fashionable, and appropriate for summer consumption. Grapes were well separated from the others, as being more expensive (Jack et al., 1997).

The aroma quality of six acid citrus fruit was characterized by instrumental and objective methods such as the GC-sniff test with flavor wheel and logarithmic values of odor units. Monoterpenes aldehydes, alcohols, and their esters were shown to contribute significantly to the characteristic aroma of lemon and lime: citronellal, citronellol, and carvone to sudachi; citronellol, 2-decenal, geranyl acetate to kabosu; geraniol, linalyl acetate, geranyl acetate, and linalool to kabosu; unsaturated aliphatic aldehydes to yuzu. Cluster analysis on odor quality showed that lemon and lime were most similar, while the rest apart from kabosu formed another cluster (Tamura et al., 1993).

The instability of spectrometers affects PCA accuracy, because PCA selects the variables with the largest variance. The relative effects of sources of instrumental instability using a model developed for fruit puree classification were investigated. The factors contributing to the instability were identified as seasonal variability in the fruit composition and slight changes in band shape and frequencies of absorption. Moreover, it should be ensured that the multivariate analyses performed with mid-infrared spectra are valid in the long term (Defornez and Wilson, 1997).

C. Juices

The GC headspace profile of commercial brands of orange juice was correlated with their aroma. PCA showed that scores for the juices and extracts are significantly different on PC1 (64% of variance) and on PC2 (14% of variance). ‘Muddy’ and metallic attributes were more associated with the extracts, while ‘bitter’ and ‘powdery’ were associated with the second axis (Elmore et al., 1994). Nineteen flavor components of Italian orange juices were used to discriminate blond and blood juices. PCA showed that blond and blood juices cluster separately, and the same was observed among the cultivars of each group, too (Maccarone et al., 1998). NIR spectra of orange juice adulterated at a 100 g Kg⁻¹ or 50 g Kg⁻¹ level were treated with PCA, and factorial discriminant analysis concluded that the detection of adulteration was possible with an average accuracy of 90%. The adulterants were fruit products or a synthetic sugar/acid mixture. At this point it is important to report that no adulterated samples were classified as authentic (Twomey et al., 1995).

Sixteen unpasteurized juices and one pasteurized juice from mandarin and mandarin hybrid fruit were analyzed with headspace gas chromatography (HSGC), and 42 volatile constituents were quantified per sample. Most of the mandarin and mandarin hybrid samples were clustered in a small area (Figure 5), while orange juice samples were widely dispersed and clearly separated from the others (Moshonas and Shaw, 1997).

The flavor of methylene chloride extracts from processed grapefruit was analyzed with GC. The juice samples were classified into low-, medium-, and high-quality products. PCA demonstrated that highest quality juices were tightly clustered. Eighty-two percent of the samples were classified in the correct preference category using only myrcene, caryophyllene, linalool, and °Brix. The most preferred juices were strongly associated with low myrcene, low linalool, and intermediate levels of caryophyllene (Jella et al., 1998).

Sensory analysis was carried out on seven blueberry soups. Their sensory profiling was completed by employing a trained panel and projective mapping. A PCA on the conventional profiling data revealed an underlying four-dimensional perceptual structure expressing 99% of the total variation (Risvik et al., 1997).

The effects of several processing methods on varietal apple juice character were investigated with analytical and sensory evaluation. Canonical discriminant analysis (CDA) of analytical data led to correct allocation of all juices to both process type and variety. CDA of sensory data was
FIGURE 5. PCA (PC1-PC3) involving quantities of 42 volatile constituents in samples of 17 mandarin and mandarin hybrid juices (+) and 22 orange juices (black squares): point a = Murcott and point b = commercial mandarin juice samples.
very effective in the separation of the processes, but considerably less with cultivars (Figure 6) (Cliff et al., 1991). Seven sensory attributes of eight drink products were treated with Thurstonian scaling and correspondence analysis to find any correlation to the thirst-quenching characteristics of the products. The selected drinks were a carbonated lemon drink, orange juice, orange squash, cola, an isotonic drink, sparkling mineral water, diet cola, and a strawberry-flavored milk. PLS analysis showed that the orange juice and lemon drink were associated with "thirst-quenching" and "acceptability," while the strawberry milk with the attributes "sweet" and "thick." A thirst-quenching drink gathered the following characteristics: acidity, astrigency, fruity and strong flavor, not thick, not sweet, and not carbonated (McEwan and Colwill, 1996).

Spherical polymer beads were suspended in clear, yellow, and red liquids. The samples were measured by turbidimetry and assessed by panelists. The application of PCA to the samples showed that in factor PCI dark was negatively correlated to turbid, opaque, and glowing, while homogeneous to particulate. PC2 was expressed by homogenous and particulate. The first two PCs explained 99% of the total variance. The haze perception of model systems resembled to juices and other cloudy products (Carrasco and Siebert, 1999).

The identities and amount of adulterants in fruit juices can be detected by looking at the distribution of elements or organic molecules in the product. The fingerprint of sugars or nonvolatile acids in apple juice can detect adulteration as less as 20% by volume. Liquid chromatography applied to orange juice has been effective in detecting low levels of augmentation by pulp wash or beet sugar (Bailey and Rohrbach, 1994). Sweetness and fruitiness were evaluated by time-intensity (TI) procedures in solutions varying in concentrations of glucose and peach essence at two temperatures. Out of the 11 TI parameters studied, 10 varied significantly in terms of sweetness and 7 of fruitiness. An increase in glucose or peach essence concentration resulted in maximum sweetness intensity. Similarly, an increase in peach flavor concentration also led to maximum fruitiness intensity (Cliff and Noble, 1990). NIR was used in conjunction with PLS to determine the concentration of sugars and acids in orange juice. LDA, K-NN, and SIMCA analysis of the furanic and phenolic constituents of cider brandies showed that these spirits could be correctly classified by age (Lavine, 1998).

D. Jams

Fresh strawberry flavor and off-flavor attributes changed during storage, especially at higher temperatures, while the attributes "strawberry jam, green and sweetness" did not exhibit any alteration. A positive correlation of the volatiles 2,5-dimethyl-4-methoxy-3(2H) furanone, 2,5-dimethyl-4-hydroxy-3(2H) furanone with fresh strawberry and a negative with off-flavor was found. α-terpinol was positively correlated with off-flavor and negatively with fresh strawberry. These three volatiles express more than 90% of sensory attributes (Golaszewski, 1998).

The volatile constituents of strawberry jam were identified and their olfactory impact was estimated by HRGC effluent sniffing. Cooker design largely influenced flavor losses at low pressures. Condensation of vapors during cooking resulted in a more flavorful product. PC plots showed a satisfactory separation of samples processed under different cooking conditions (at atmospheric pressure, at 8 KPa, at 40 KPa) (Lesschaeve, 1991).

Peach jam samples were studied during aging employing an Instron machine. Jam texture ranged from very firm (100% isoglucose syrup) to very soft (100% maltose syrup). Three weeks aging was required for the stabilization of the jam texture. PCA showed that the jams could be effectively classified according to their mechanical and textural attributes (Raphaelides et al., 1996).

E. Wines

Cabernet Sauvignon wines from four regions and Chardonnay wines from three vintages were evaluated by descriptive analysis. Canonical variance analysis (CVA) led to separation of wines according to their origin and based on their sen-
FIGURE 6. Canonical discriminant analysis of analytical (Figure a) and sensory (Figure b) data using juice extraction process and cultivar for class differentiation of apple juice. S: ‘Spartan’; M: ‘McIntosh’; G: ‘Golden Delicious’; D: ‘Delicious’; L: ‘Liquefaction’; N: ‘Non-oxidative’; O: ‘Oxidative’; D: ‘Diffusion’ (Cliff et al., 1991).
sory attributes dominated by bell pepper aroma and phenol. This separated the wines of the Southern region from the rest. The most important term on the nonsignificant second axis was vanilla. CVA also separated different vintages based on sensory attributes (Figure 7). Only bell pepper aroma varied more among regions than within regions. The 1982 and 1983 wines were separated along the first dimensions, with the 1981 wines located between them. The 1981 vintage wines were found to be more bitter and less fruity and floral than the rest vintages. The 1983 wines were more floral and less citrus than the 1982 wines. Only the term “flora” differed significantly among the vintages (Heymann and Noble, 1989).

Six varietal red wines produced in Murcia (Spain) were analyzed for selected oenological and color parameters (31 variables in total). Discriminant analysis allowed a clear separation of the 80.7% of the cases analyzed by a selection of only three variables \((L^*, a^*, C^*)\). The employment of 10 variables made possible the characterization of 95% of the analyzed cases, but 100% in Monastrell-I, Tempranillo, Graciano, and Cinsault Noir wines (Almela et al., 1996). NIR data were analyzed with PLS to determine the concentration of ethanol, glycerol, and sugars in botrytized grape wines. An ANN was used to classify wines by geographic region based on their trace metal content (Lavine, 1998).

F. Cereals

To investigate the involvement of corn germ enzymes in off-odor formation, crude enzyme and purified lipoxygenase (LPO) and peroxidase (PDO) extracts were prepared and added to the homogenates of blanched corn. Statistical analysis of sensory data was performed with ANOVA and PCA. The addition of LPO increased ‘painty’ and ‘stale/oxidized’ off-odor descriptors and lowered ‘sweet’ and ‘corn’ descriptors. POD was not important in off-odor formation, and LPO seems to be a potential blanching indicator (Theerakulkait, 1995).

The effects of sugar and feed moisture contents on the sensory characteristics were investigated. The GPA results of determined sensory and physical properties showed that the breakfast attributes were mainly affected by the sugar-addition level, while feed moisture affected significantly the sensory character of the samples that included 5% sugar. Finally, increased heat and mechanical work in the extruder resulted in higher sensory residual soy flavor and off-flavor intensities (Faller et al., 1998).

NIR spectral information was used to classify red and white wheats in order to minimize subjectivity in class determinations. The application of PLS to these data demonstrated more than 99% correct classification for single kernels when visible and NIR regions were used. Single kernel classifications had an average of 100% correct classifications of bulk specimens (Dowell, 1998).

NIR analysis was applied to cooked rice and sensory and instrumental texture analyses. The application of PLS to the data sets showed that the sensory texture attribute ‘roughness’ was related to protein content, ‘springiness’, ‘hardness’, ‘guminess’, and ‘chewiness’. The location of samples scores were in accordance with main chemical variation among the samples. Varieties grown in California were higher in amylose and lower in protein than those grown in Arkansas (Windham et al., 1997). The effect of various postharvest processing treatments on sensory characteristics of cooked rice was investigated by applying sensory descriptive methods. Cooked rice quality was affected by rough rice wet holding, drying temperature, storage temperature and storage duration (Meullenet, 1999). Rice plant and panicle morphology in relation to grain cracking were studied in 17 varieties and at four maturity stages. The treatment of data with PCA resulted in four new factors that explained 72% of the total variance. PC1 was dominated by yield and biomass related attributes, while PC2 by percentage cracked grains and grain moisture. Predictive models explained more than 96% of cracking of rice grains and the agronomic practices in controlling the grain cracking were addressed.

The possibility of predicting the malting quality of barley grain, indicated by malt extract yield, by characteristics measured either on plants in bloom or in mature dry gain by image analysis, was investigated. An analysis of the results with PCA suggested that malt extract production could
FIGURE 7. Projection of Cabernet Sauvignon wine sensory data on canonical variates 1 and 2 for class variable = regions. Attribute loadings and factor scores for Napa wines (black squares), Soloma (black triangles), Lake (circles) and Southern wines (white squares) surrounded by circles including 95% confidence intervals (Heymann & Noble, 1989).
be predicted when the protein content was determined (Moral et al., 1998).

Ten inorganic elements were determined in the kernel of 19 almond cultivars grown in the same field and year, but of different origins. Both cluster analysis and PCA distinguished the American cultivars from the rest, most of them from the Mediterranean area, except for an American cultivar (Titan), which was very similar to a Spanish one (Atocha). Ca and Fe could characterize three American cultivars, while Cu and Zn defined the Mediterranean almonds as a group. The mineral fraction was a characteristic feature affected by soil and climatic factors as well (Prars-Moya et al., 1997).

The concentration of selected mineral and trace elements in potatoes was used to differentiate potatoes grown in Idaho and outside of Idaho. The content of selected minerals is a reflection of the soil type and, importantly, the environmental growing conditions. Although PCA revealed groupings between Idaho and non-Idaho potatoes, the use of neural network models in conjunction with discriminate function analysis or bagging strategy resulted in the higher rates of correct classifications (98 to 99%) (Anderson et al., 1999).

G. Bakery Products

Flours from hard red spring and hard red winter wheat were tested for water absorption, dough mixing tolerance, loaf height, internal grain appearance, and overall bake score. Only water absorption could be modeled with NIR, due to its ability to determine protein level and/or starch damage. PCR and PLS analyses were applied to the spectral data set leading to promising discrimination grouping (Delwiche and Weaver, 1994).

Twenty flours from 16 different varieties cultivated in 1990 and 1992, and a Swedish reference flour, were fermented to sourdoughs. Barley breads from each flour type were baked with and without an admixture of barley sourdough. PCA showed that breads with a sourdough admixture scored higher for total and acidulous taste than breads without sourdough (Marklinder et al., 1996).

Five different methods of extracting features from textural images in food by multivariate modeling of the sensory porosity of wheat baguettes were compared. The employed methods were angle measure technique (AMT), the singular value decomposition (SVD), the autocorrelation and autocovariance functions (ACF), and the size and distance distribution (SDD) method. The results from PCR and PLS analyses showed that all the methods are suited to extract sensory porosity, but the AMT method proved to be the most effective in this case (Kraal et al., 1998). A trained neural network (NN) was created and used to predict loaf volume of breads made from different wheat cultivars. The NN was more accurate, faster, and easier than Principal Component Regression Analysis (Horimoto et al., 1995).

H. Pasta

U.S. hard white wheat varieties were employed for the production of several patented Asian noodles (16 Taiwanese, 18 Thai, and 18 Malaysian) with the same machine. PCA of sensory data and texture data showed that the Taiwanese raw noodles were the smoothest, springiest, and were characterized by a higher integrity of noodles. Thai bamee noodles were the hardest, most cohesive ones, with the highest density and were more easily stuck between teeth and toothpull due to their high starch content. Malaysian noodles were the softest, least dense, cohesive, and the least sticky. The specific texture characteristics of each country’s style noodle led to a clustering of the samples on a PC-plot reflecting their origin (Figure 8) (Janto et al., 1998).

Six different pasta types were investigated. Digital images of the protein network in cross sections of the samples were obtained followed by textural image analysis. Similarities and differences in protein network texture were assessed by principal component, stepwise discriminant, and variance analyses. The method of cooking was shown to have a strong influence on the texture. In general, texture, protein enrichment, botanical origin, and fractionation were shown to considerably influence the final pasta product (Fradet et al., 1998).
FIGURE 8. Texture characteristics of Taiwanese (w), Thai (t), and Malaysian (m) fresh noodles. Numbers in parentheses indicate the loadings of corresponding descriptors obtained from the eigenvectors of the PCA (Janto et al., 1998).
I. Oils

1. Olive Oils

Two Cretan olive cultivars, Koroneiki and Mastoidis, were studied in terms of their triglyceride composition as a function of their maturity time. Isocratic HPLC was applied and PCA and Canonical Discrimination Analysis (CDA) were employed to group olive oil samples within each variety according to their geographical origin, as shown in Figure 9 (Stefanoudaki et al., 1997).

$^1$H-NMR spectroscopy was applied to the analysis of 55 extra virgin olive oil samples from four Italian regions (Campania, Lazio, Sicily, and Umbria) and obtained from different olive varieties. The oil samples in their crushing majority (96%) were correctly classified in terms of $\beta$-sitosterol, n-alkenals, and other volatiles. Four groups of samples are demonstrated clustered according to their geographical origin on a PCA plot. Two Spanish samples from two typical olive varieties of Andalusia were also studied and found to be significantly different compared with the 55 Italian samples (Sachi et al., 1998).

Chemometrics showed that the geographical and varietal origin of olive oil samples could be related to chemical compounds such as fatty acids, sterols, aliphatic and triterpenic alcohols, phenolic substances, tocopherols, phytol, triacylglycerols, and volatiles. Stable isotope ratio analysis can verify authenticity and origin of foodstuffs, like wine and maize oils. Cluster analysis of oil samples, produced in various Mediterranean countries, based on determination of $^{13}$C and $^{18}$O led to the following conclusions: Italian oils show a greater variability of $^{13}$C values, while samples coming from Morocco, Tunisia, and Spain are more enriched in $^{18}$O because of geoclimatic parameters. Oils from Greece and Turkey show similar $^{13}$C and $^{18}$O data with lower isotopic values due to temperate climate and nearness to the sea. PCA analysis showed that PC2 can rather satisfactorily discriminate oils according to their origin, whereas PC1 could also discriminate some Italian oils (Angerosa et al., 1999).

Fourier transform infrared spectroscopy (FTIR) in conjunction with PCA and DA (discriminant analysis) led to the correct classification of 93% of samples in a calibration set and 100% in an independent validation set of extra virgin and refined olive oils. Seed oils (olive, corn, etc.) were satisfactorily clustered according to plant species (Lai et al., 1994).

Extra virgin olive oils adulterated with sunflower, rapeseed, and soybean oil in a range of 5 to 95% could be identified correctly in 90% of cases with the help of discriminant analysis equation developed for the treatment of NIR data with PCA. The accuracy of the level of adulteration was approximately 0.9% (Wesley et al., 1996).

Positive-ion fast atom bombardment-mass spectrometry, with $m$-nitrobenzyl alcohol as matrix and a methanolic sodium iodide (NaI) solution, was used to characterize vegetable oils by their triacylglycerol composition. Olive oils always resulted in negative PC1 scores, although most of the seed oils had positive PC1 scores. Peanut oils differentiated from the others by being close to the extra virgin oils, but with no overlapping. Corn, soybean, and sunflower oils seemed to be quite clearly apart from the other oils (Lamberto and Saitta, 1995).

Stepwise multiple regression and regression on principal components were used in deriving equations that relate profile to overall grading of 57 samples of Spanish virgin olive oil. RP-PCA correctly classified 87.5% of 40 new samples, against 95% by stepwise multiple regression (Aparicio et al., 1992).

The aroma compounds of virgin olive oil were detected by means of HRGC. The samples originated from two Spanish, one Greek, and one Italian olive varieties picked at three different stages of ripeness. The stages were characterized with cluster analysis on the volatile compounds data set. Alcohols produced from linoleic acid, hexanal, and hexyl acetate were shown to characterize ripeness (Aparicio and Morales, 1998).

Dynamic headspace sampling methods prior to capillary gas chromatography were used for the determination of volatiles occurring in virgin olive oil. The most important compounds were identified with mass spectrometry as well. Stepwise linear discriminant analysis was applied to volatiles, and 90.5% of correct classifications on the basis of their origin were obtained by employing only three volatiles: 1,3-hexadien-5-yn, 2-
FIGURE 9. Canonical representation for the discrimination of olive oil samples according to cultivar, geographic origin and sampling date. The samples at the top of the plot belong to Koroneiki cultivar, while the ones below the dashed line to Mastoidis (Stefanoudaki et al., 1997).
methyl-1-propanol, and 3-hexenyl acetate. Samples from Greece, Italy, and Spain were successfully classified in separate groups, as can be seen in Figure 10 (Morales et al., 1994).

The identification of olive oil, hazelnut oil, and mixtures of both oils (85:15) may be possible through the determination of presence/absence of (E)-5-methylhept-2-en-4-one (fildertone) (Blanch et al., 1998). No volatile appears to be responsible, on its own, for the sensory attributes evaluated in the virgin olive oil samples. On the contrary, several groups of them gave rise to different green sensory notes (Morales et al., 1996).

The Koroneiki variety showed the maximum concentration of esters responsible for the green (grass) perception, while Arbequina had the maximum concentration of furans, responsible for sweet (ripe fruit) perception. Finally, the total concentration of furans is able to distinguish the Picual and Coroneiki varieties from Coratina and Arbequina (Aparicio et al., 1997). Quantitative analysis of isoprenoids and methylsterols isolated from each grade of olive oil falls in agreement with the decreasing negative values of 13C/12C ratios corresponding to the best grades of olive oil. These results have turned out to be a good tool for detecting adulteration of olive oil with the cheaper pomace oil, even down to a 5% level (Angerosa et al., 1997).

Thirty-two samples of six European varieties of virgin olive oil were characterized by 55 chemical compounds and 55 sensory attributes evaluated by six different panels. Multidimensional scaling was used to bring out inter-intra dissimilarities from datasets of sensory attributes and volatile compounds. Cluster analysis based on volatiles showed that the Koroneiki and unripe Arbequina samples are located quite apart from the other varieties. Another great group consisted of two subgroups — one clusters Italian samples and the other basically Spanish varieties that appear separated in two groups (Morales et al., 1995).

PCA and PLS were used to treat variables of olive oil samples to classify extra virgin and ordinary olive oil samples. PLS showed a tendency to overfit, but it compressed the most relevant information into the first few scores, whereas PCA did not. In addition, PLS introduced groupings into the scores from data where there was no inherent class structure. Finally, PLS resulted in higher prediction success rates (Kemsley, 1996).

Data on consumer preference, perception, and purchase motives concerning vegetable oil were collected in Denmark, England, and France. Correspondence analysis of the CGS kind interpreted attributes close to a product, as typical of that particular product and the results are demonstrated in two-dimensional plots. At the attribute level country of origin is mostly a Danish attribute, whereas odor and hints at sunshine are mostly a French one. Strong, characteristic taste and odor of virgin olive oil is related to good cooking results and good taste in France, whereas in England it is also related to poor taste and poor cooking results severely affecting the consumer preferences (Nielsen et al., 1998).

2. Other Oils

Blends of ricebran, mustard, sunflower, safflower, and palm olein subjected to GLC analysis with an artificial neural network (ANN) model resulted in the identification of the constituent oils with 100% accuracy. Models based on multiple linear regression successfully predicted the composition of the blends (Husain et al., 1996).

Fifteen commercial oil samples were studied with the mid-IR optical fiber method, and the results were analyzed with PCA. The second derivative of the spectral data had to be obtained and the first PC separated sunflower oil from olive and peanut extracts based on the different concentration of linoleic acid (Dupuy et al., 1996).

Ninety-nine samples of Italian pepermint oil were analyzed with HRGC. PCA showed that oil composition is strongly influenced by the geographic origin of plants, their harvesting date, and the time elapsed since their latest transplantation. Linear PLS disclosed the variables involved in the maturation and aging process, but no reliable classification model on oils of different origin could be developed (Chialva et al., 1993).

The detection of adulterated bitter almond and cinnamon oils can be achieved by SNIF-NMR using benzaldehyde as a molecular probe, based on the determination of the origin of specific deuterium contents of benzaldehyde: syn-
FIGURE 10. Two-dimensional plot of canonical variables showing the discrimination among Greek, Spanish, and Italian olive oils according to their geographical origin (Morales et al., 1994).
thetic (ex-toluene and ex-benzal chloride), natural (ex-kernels from apricots, peaches and cherries, and ex-bitter almond), and semisynthetic (ex-cinnamaldehyde extracted from cinnamon). PCA plots provide a clear classification of these samples (Remaud et al., 1997).

NIR and PCA were employed for the classification of various vegetable oils. The score plot of PC1 vs. PC2 shows that along the X-axis the order of the samples corresponds to the degree of saturation; the left side is rich in unsaturated fatty acids. The Y-axis mirrors the amount of C18:1 saturation in decreasing order, as can be seen in Figure 11 (Sato, 1994).

FTIR spectroscopy with attenuated total reflectance was applied for characterizing edible oils and fats. PC scores 1 and 4 accounted for 89.75 of total sum of squares. Axis 4 was related to the degree of unsaturation of fatty acids starting with caprylic acid (saturated) in the lower area and ending with linolenic acid in the upper area (trisaturated). A similar grouping was observed for edible fats too, where 72.1% polyunsaturated fatty acids were placed opposite to low-fat (LF) butters (0.4% polyunsaturated fatty acids) and oils (with the highest proportion of oleic acid) (Safar et al., 1994).

Chemical and thermoanalytical data on refined rapeseed oils were correlated, and the thermogravimetric techniques were proved to be useful in defining the quality of rapeseed oils. The plot of the PC1 vs. PC2 explained more than 84% of the variation. Strong relations were apparent between the chemical variables and PC1, except for the iodine number, but no relations were found for PC2, while thermoanalytical variables were strongly correlated with PC1 (Wesolowski and Erecinska, 1998).

NIR was utilized for the determination of oil type in full-fat mayonnaise as well as the detection of minor variations in the oil content and variations in the content of individual fatty acids. It was shown that for these particular data LDA was superior to the regression analysis based on PLS2 and PCR (Indahl et al., 1999).

Patchouli oil and amyrin oil samples were analyzed by gas chromatography. Data sets were analyzed with PLS and artificial neural network simulation (ANNS). ANNS functioned satisfac-

torily in distinguishing the sensory acceptance of the oils, but a relatively large amount of data with a high samples/variables ratio was required (Pepard, 1994).

GC, IR, and GC-MS were employed for grouping the phenolic (thymol was predominant over carvacrol) from the nonphenolic chemotypes of essential oils of Thymus hyemalis Lange of Spanish origin. Principal component analysis and cluster analysis of chemical variability enabled the satisfactory separation of closely related phenolic from nonphenolic groups. Multidimensional scaling analysis of essential oil components suggested a relative incompatibility within plant groups and populations (Saez, 1998).

The employment of GC analysis of essential oils from Dalmatian Sage (Salvia officinalis L.) disclosed the presence of three chemotypes with different proportions of alpha- and beta-thujone. Flowering parts of Salvia officinalis L. had higher oil contents and beta-pipene levels than leaves and lower thujone levels. PC analysis showed that most of the individual flowering plant oils can be grouped in three qualitative thujone chemotypes (Perry et al., 1999).

J. Tea

Thermally generated volatiles from roselle tea were collected and analyzed with GC and GC-MS. More than 37 compounds were characterized and classified into four groups: fatty acid derivatives, sugar derivatives, phenolic derivatives, and terpenes. PCA led to the observation that a combination of the terpene derivatives with fragrance notes and the sugar derivatives with a caramel-like odor are responsible for the roselle aroma (Chen et al., 1998).

Seventy-seven peaks of gas chromatograms of typical black, oolong, and green tea volatiles and the ratios of their areas were treated with cluster analysis. Black tea has two semiclusters containing mainly Assam and Darjeeling teas, respectively. Green tea also has two major semiclusters, and one of them is grouped very tightly. Oolong tea has three semiclusters, two being Taiwan teas and one a product from the mainland of China. Green tea aroma is similar to that of commercially available
FIGURE 11. Score plot of PC1 vs. PC2 for vegetable oils. The samples are clustered according to the plant species (Sato, 1994).
products. Oolong tea has a characteristic aroma due to the various conditions of the manufacturing process (Kobayashi et al., 1993). Green and black teas were correctly classified using amino acids, caffeine, theobromine, and theophylline as features (Lavine, 1998).

K. Coffee

NIR spectroscopy has substantially contributed to a very successful discrimination between beverages produced from pure Arabica, pure Robusta and Arabica:Robusta blends. Correct classification rates on lyophilized and vacuum-dried laboratory prepared beverages were 87 and 95%. These data confirmed previous results on coffee beans and reinforce utility of NIR for coffee authentication. This discrimination process is based on the determination of caffeine and/or others alkaloids content(s) (Downey and Boussion, 1996).

FTIR spectroscopy is a promising method with great potential for determining the carbohydrate profile of instant coffee. PCA and linear discriminant analysis (LDA) resulted in a 98% successful classification rate. ANN gave 100% successful discrimination of pure coffee from adulterated one (Briandlet et al., 1996).

The relationships among 13 aroma, flavor, mouthfeel, and appearance variables for 18 soluble coffees were analyzed employing flavor profiling. PCA led to the formation of clusters; for strong coffees, all the variables surpass the general mean, while for moderate coffees only aroma and mouthfeel attributes do. Finally, for light coffees all attributes are lower than the average and for samples with poor sensory character, mouthfeel and aroma had small values, while appearance variables had very high ones (Zamora and Calvino, 1996).

Twenty judges performed a variety of chemosensory tasks so that the ones with the best scores would form a panel for coffee evaluation. An average of correct responses (P%), one-way ANOVA and PCA were compared. The panel was finally composed of nine panelists selected by the three methods. It was concluded that the three methods should be simultaneously used to achieve a higher security degree in the acceptance or rejection of panelists (Zamora and Calvino, 1996).

L. Honey

Forty-eight honeys from the La Rioja region of Spain were analyzed in terms of 14, legally required, quality control parameters to establish their quality. Samples originated from two geographic areas of La Rioja: Valley and Sierra. The parameters determined were conductivity, pH, ash, free acidity, hydroxy-methyl-furfural (HMF), diastatic activity, various sugars, moisture fructose, ratio fructose/glucose, and optical rotation. The application of factor analysis and linear discriminant analysis resulted in the classification of samples according to their geographic origin (Valley or Sierra) with 83% accuracy (Sanz et al., 1995).

Fifteen parameters of seven types of Spanish commercial unifloral honeys were measured and stepwise discriminant analysis was performed. The most discriminant variables were electrical conductivity, color (X,Y,L), water content, fructose, and sucrose. All sunflower, eucalyptus, and forest honey samples were correctly classified into their a priori established honey types. Rosemary, citrus, lavender, and heather honeys accounted for 92.3, 93.3, 91.7, and 91.7% correct classifications, respectively. The overall percentage of correct classification amounted to 95.7%. Microscopical analysis for honey classification proved not to be adequate on its own and had to be complemented with consistent chemical or physicochemical determinations (Mateo Bosch-Reig, 1998).

M. Sugar-Syrups

Thick beet sugar juices samples were collected from five sugar factories in Denmark and analyzed with a fluorospectrometer. Four external parameters for 47 samples, two levels of dilution, 20 excitation wavelengths, and 311 emission wavelengths were obtained. The plot of scores for the PCs 2 and 3 showed a clear-cut classification
to the producing factories (a, b, d, e, and f) and with a clear tendency of timing within each cluster from below to above, ranging from the early to the late samples (taking on account the time period they were processed in the factory as a function of the entire campaign period). Fluorescence spectra of finished sugar and beet samples were also analyzed with multidimensional techniques. Finally, it was concluded that process irregularities as well as chemical species can also be detected and validated by multiway chemometric techniques (Munck et al., 1998).

Regression analysis applied to sensory evaluated syrups showed that the triplet ‘viscous’, ‘smooth’, and ‘slippery’ gave the best average coefficient of determination in the mouth. The triplet ‘watery’, ‘even’, and ‘drippy’ gave the best one for pouring; and the triplet ‘viscous’, ‘even’, and ‘smooth’ the best one for spreading (Elejalde and Kokini, 1992).

O. Gums

Binding of Na⁺ in aqueous gum systems as determined by ²²Na-NMR spectroscopy and its relations to perceived saltiness were examined. Nonmetric KYST-2A multidimensional scaling (MDS) and ANOVA were applied to the data. As Na⁺ increased in both ionic and nonionic systems, NMR transverse relaxation rates converged and perceived saltiness equalized. Food components that bind Na⁺ may suppress saltiness perception, which could be important in low-sodium foods (Rosee et al., 1994).

III. CONCLUSIONS

The employment of several multivariate methods has become a prerequisite for several applications related primarily to food quality control both in terms of authentication and classification according to variety and/or geographical origin. In fact, it has been repeatedly shown that employment of methods like principal component analysis, canonical analysis, linear discriminant analysis, cluster analysis, partial least squares, and surface response methodology simplify substantially the classification/grouping task. Over the last 2 decades, these methods have enjoyed a universal approval in several fields of food science. In particular, agricultural produces represent one of the easiest to adulterate products, and meticulous implementation of multivariate methods was shown to be of great importance.

REFERENCES


science — a demonstration of the feasibility of a high exploratory, inductive evaluation strategy of fundamental scientific significance. Chemometrics and Intelligent Laboratory Systems, 44, 31–60.


90. Rosee, T.R., Shirley, L., Schmidt, S.J., and Klein, B.P. (1994) Na+ binding as measured by 23Na nuclear magnetic resonance spectrometry influences the per-


