Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Food Chemistry 86 (2004) 283–296
Food Chemistry
www.elsevier.com/locate/foodchem

Analytical, Nutritional and Clinical Methods

Chemometric characterization of Italian wines by thin-film multisensors array and artificial neural networks

M. Penza *, G. Cassano

ENEA, C.R. Brindisi, Materials and New Technologies Unit, SS.7, Appia, Km 714, Brindisi 72100, Italy

Received 3 February 2003; received in revised form 10 September 2003; accepted 10 September 2003

Abstract

In the present work, nine samples of Italian wines (three white, three red and three rosè) from different denominations of origin have been analysed by the static headspace sampling method to attempt to classify them by chemometric characterization of the data obtained from a thin-film multisensor array. All wines have also been analysed to measure their ionic conductivity, pH and alcoholic content. An electronic nose comprising four metal oxide semiconductor thin-film sensors has been used to generate a typical chemical fingerprint (pattern) of the volatile compounds present in the wines. Principal component analysis and artificial neural networks were applied to the generated patterns to achieve various classification tasks. The classification performance of nine different pre-processing algorithms has been studied on the basis of three different sensor parameters and three different normalization techniques. The wine patterns generation with array sensor signals and the chemometric treatment are fast and simple by providing a recognition rate and a prediction rate as fairly high as 100% and 78%, respectively. These results can be considered satisfactory and acceptable, with the selected variables useful to differentiate these wines by their class.

Keywords: Chemometric characterization of wines; Multisensors array; Artificial neural networks; Principal component analysis

1. Introduction

Wine composition depends on many factors such as grape varieties, ripening, soil and climate, must-fermentation time, wine-making process, yeasts and oenological microflora and wine aging type. Also, the organoleptic properties for the same kind of wine are different between the various vintage years. Of course, the several grape cultivars are strongly affected by the vineyards located in a typical geographical area. These factors are extremely important for quality wines from specific regions, such as protected designation of origin (PDO) and controlled denomination of origin (CDO) wines. Hence, the wine aroma presents an extremely complex chemical pattern in both qualitative and quantitative terms. It is well known that the flavour of a wine consists of over a thousand of volatile compounds with a wide concentration range from a few ppm to much higher quantities up to 10–15% in weight. Several classes of compounds have been identified in the aroma profile of a wine, the most important may be alcohols, esters, acids, ketones, aldehydes, ethers, terpenes, lactones, sulphur compounds, nitrogen compounds, carbonyl compounds, phenolic compounds, etc. It is obvious that all aroma compounds play a role in the characterization of the flavour pattern of a given wine. In fact, the headspace of a specific wine is the global chemical information intended as a weighted resultant of each volatile compound constituent the wine flavour.

As a general rule, the discrimination of the wines is not an easy task due to the complexity and heterogeneity of its headspace. However, the classification of the wines is very important because of high economic value of the wine-product for some geographical regions, e.g., the various Italian areas such as Apulia and other world-wide areas, to protect the quality wines, to prevent

* Corresponding author. Fax: +39-0831-507674.
E-mail address: michele.penza@brindisi.enea.it (M. Penza).

0308-8146/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodchem.2003.09.027

illegal adulteration of wines, to safeguard human health from wines with high and super alcoholic content, to guarantee the wine quality in import-export market and to control beverage processing. Generally, the sensory analysis based on the trained experts panel test is useful in the wine classification task, but it is not always feasible because of high-cost and is time-consuming and sometimes without any objective estimation. Therefore, it is interesting to use another methods for wine discrimination essentially based on instrumental analytical techniques. In fact, the common methods of chemical analysis such as gas and liquid chromatography, mass spectrometry, nuclear magnetic resonance and spectrophotometry have higher reliability, longer processability, low in situ measurableness and higher costs. In this scheme of analytical methods, it has been proposed to use a multisensors array based on electronic nose-type combined to the multivariate statistical analysis techniques. The electronic nose device has the advantage of high portability for in situ and on-line measurements with lower costs and good reliability.

Recently chemometric classification techniques and pattern recognition analysis methods for wine and other alcoholic beverages have received great attention and have been largely used in the last years (Anklam, Lipp, Radovic, Chiavaro, & Palla, 1998; Di Natale et al., 1995; Di Natale et al., 1996; Frias, Conde, Rodriguez, Dohnal, & Perez-Trujillo, 2002; Guadarrama, Fernandez, Iniguez, Souto, & de Saja, 2000, 2001; Marengo, Aceto, & Maurino, 2001; Peña, Latorre, Garcia, Botana, & Herrero, 1999; Pérez-Magariño, Ortega-Heras, & Gonzales-San José, 2002; Raptis, Siettos, Kiranoudis, & Bafas, 2000). Several authors have proposed methods to discriminate wines, recognising different compounds as markers. The mineral ions have been studied for the classification of Greek wines according to their geographical origin (Kallithraka et al., 2001). The polyphenol content in Spanish wines was determined by HPLC for the differentiation according to their geographical origin using multivariate statistical analysis (Rodriguez-Delgado, Gonzalez-Hernandez, Conde-Gonzalez, & Perez-Trujill, 2002). Also, the discrimination of wines was attempted on the basis of amino acid composition by the use of statistical methods (Soufleros, Bouloumpasi, Tsarchopoulos, & Biliaderis, 2003). Moreover, characterization of the geographical origin of Italian red wines from Apulia region based on nuclear magnetic resonance has been performed by multivariate statistical methods (Brescia et al., 2002).

The aim of this study is to chemometrically classify some Italian wines belonging to various classes (white, red, and rosè) by using a multisensors array and artificial neural networks (ANNs) algorithms.

2. Materials and methods

2.1. Samples

The samples used in the experiments have been bought from local marketers in sealed 750-ml bottles. The samples analysed for this study were nine different Italian wines: three white wines (*Chardonnay del Salento*, *San Severo* and *Salento-Agrivia*); three red wines (*Soleto*, *Chianti*, and Matino); three rosè wines (*Mesagne*, *Castel del Monte* and *Salento-Mottura*) from different denominations of origin and vintage years. All tested wines were from Apulia, a geographical region in the Southern Italy, except one (*Chianti*) from Tuscania, another geographical area in the Central Italy. All Apulia wines were divided into three groups: six wines from Southern Apulia, one from Central Apulia and one from Northern Apulia. The samples of all wines have been analysed by measuring their pH value and ionic conductivity by using a pH-meter (Crison, model GLP21) and a digital conductivity meter (Bicasa, model BE103), respectively; while the alcoholic content of the wines was the labelled value reported on the bottle. The results of the analysis on the nine samples of wines are reported in the Table 1.

Table 1
Physico-chemical characteristics of the wine samples

Sample	Denomination	Colour	Alcoholic content (vol%)	pH	Ionic conductivity ($\mu S\,cm^{-1}$)	Origin/Vintage
Wine 1	Chardonnay del Salento	White	12.5	3.43	2080	S. Apulia/2000
Wine 2	San Severo	White	11.0	3.41	1965	N. Apulia/1998
Wine 3	Salento-Agrivia	White	10.5	3.47	1800	S. Apulia/2000
Wine 4	Soleto	Red	11.5	3.37	2100	S. Apulia/2000
Wine 5	Chianti	Red	12.0	3.35	2110	Tuscania/1999
Wine 6	Matino	Red	12.0	3.32	2080	S. Apulia/1999
Wine 7	Mesagne	Rosè	11.5	3.68	2400	S. Apulia/1999
Wine 8	Castel del Monte	Rosè	12.0	3.05	1730	C. Apulia/2000
Wine 9	Salento-Mottura	Rosè	12.0	3.37	1840	S. Apulia/2000

S., South; N., North and C., Central.

2.2. Multisensors array

A multisensors array-type electronic nose based on four metal oxide (WO_3) semiconducting thin-film sensors has been applied for headspace analysis of the wines. The electronic nose used has been home-fabricated and home-developed for food and flavour analysis purposes (Penza, Cassano, Tortorella, & Zaccaria, 2001). The WO_3 thin films were deposited by PVD systems onto alumina substrates (10 mm × 10 mm × 0.6 mm). The thickness of the WO_3 thin films was 300 nm. Two front-face Al (150 nm) metallic strips (2 mm × 10 mm) were PVD prepared onto the WO_3 thin films to serve as electrical contacts for output of the single sensor signal. A thin layer of metallic catalysts (platinum, gold, palladium and bismuth) was separately PVD-deposited onto the top-surface of the discrete WO_3 thin films sensing elements and between two strip-contacts in an appropriate area (4 mm × 10 mm) in order to specifically modify the sensor surface and to achieve useful cross-sensitivity towards wines headspace. The experimental details of the sensors array preparation have been reported elsewhere (Penza, Martucci, & Cassano, 1998; Penza, Cassano, & Tortorella, 2001). Table 2 shows the technical characteristics of the sensing elements used in the electronic nose.

2.3. Measuring setup and wines headspace sampling

The sensors were located in the test chamber (250 ml inner volume) and thermally contacted to a heated holder. The operating temperature of the sensors was controlled and kept constant at 250 °C during the sensing experiments.

The wine samples (300 ml) were closed for 30 min into distinct graduated bottles (500 ml), bath thermally maintained at 25 °C, for the generation of the headspace before sensing analysis. Then, this generated headspace was transferred into the sensors cell by dry air used as a carrier flowed at a constant rate of 500 ml min^{-1}. The gas flow rate was controlled by a mass flowmeter. The wine sampled headspace and dry air were alternately switched into test cell by a three-way valve according to a determined exposure time of 2 min and a proper relaxation time of the sensors of at least 30 min after exposure to wine headspace. Some no-back valves were needed in order to avoid retrofitting of sampled headspace.

In order to probe and verify the repeatability of the responses of array sensors towards each wine flavour, six samplings of the same headspace were measured. Reference dry air was also used to condition the sensors and to virtually set the array sensors signal to a baseline level.

The electrical characteristics of each WO_3 thin film sensor in the array have been obtained by measuring the electrical current flowing through the films biased by a constant voltage in the format of two-pole probe. A multimeter (HP 34401A) or a programmable electrometer (Keithley 617) was used to measure the d.c. electrical resistance of each sensor depending on the dynamic range of the electrical resistance of the sensors. A multiplexer (home-made), controlled by PC via parallel port, scanned the four sensors of the array at a reading average rate of 4 s per cycle, or equivalently to 1 s per channel. A personal computer, GPIB interfaced with all instrumental equipment under HP-VEE ambient, managed all operations sequence. The data of sensors array are real-time visualized on screen and stored for further analysis in order to perform the pattern recognition techniques by means of a commercial software package (Multi-Variate Statistical Package, MVSP 3.1) for principal component analysis (PCA) and another commercial software package (Qwiknet V.2.22, Craig Jensen, Kirkland, WA, USA) for artificial neural networks (ANNs) algorithms application.

Table 2
Characteristics of the WO_3 thin-film (300 nm) sensors surface-activated with different catalysts used in the sensing array for wine analysis

Sensor	Sensor type	Catalyst	Catalyst thickness (nm)
Sensor 1	WO_3:Pt	Platinum	25
Sensor 2	WO_3:Au	Gold	55
Sensor 3	WO_3:Pd	Palladium	50
Sensor 4	WO_3:Bi	Bismuth	65

2.4. Data analysis

The problem of analysing the data generated by a multisensors array is basically one determining the existing relationships between a set of independent variables (e.g., the output responses from a multisensors array) and another set of dependent variables (i.e., flavour class for the recognition task or chemical components concentration for the quantification task). The numerous multivariate data processing techniques are either linear (i.e., they assume a linear relationship between independent variables and the dependent variables), such as PCA, or nonlinear, such as ANNs (Gardner & Bartlett, 1999). The multivariate statistical methods largely used can be either unsupervised (discrimination of unknown flavour vectors), like PCA, or supervised (the unknown flavour vectors are chemometrically examined by relationships found a priori from a set of known flavour vectors used in an initial calibration, learning or training stage), like ANNs. As a general rule, every multivariate classification scheme is typically dedicated to a specific task, because the selected variables useful for a good discrimination in a given case may not be reliable in a different architecture.

Hence, the multivariate statistical strategies need a trial-and-error chemometric approach.

2.4.1. Principal component analysis

Principal component analysis (PCA) is a chemometric linear, unsupervised and pattern recognition technique used for analysing, classifying and reducing the dimensionality of numerical datasets in a multivariate problem. This method extracts the dominant features from a data matrix in terms of a complementary set of scores and loadings plots. In other words, the PCA procedure can be applied by finding the eigenvectors of the primary matrix of the datapoints, and to form a transformation matrix from these eigenvectors ordered so that the corresponding eigenvalues are in decreasing order. The eigenvalues depend on the normalization procedure applied to the input data prior to processing and type of matrix (covariance or correlation) used for PCA. PCA processes the data matrix by projecting the multidimensional dataset onto a new coordinates base formed by the orthogonal directions with data maximum variance. The eigenvectors of the data matrix are called principal components and they are uncorrelated among them. The magnitude of each eigenvector is expressed by the own eigenvalue, which gives a measure of the variance related to that principal component. As a result of the coordinates change, a data dimensionality reduction to the most significant principal components and an elimination of the less important ones are possible to achieve without any considerable information loss. The main features of PCA are the coordinates of the data in the new base (scores plot) and the contribution to each component of the sensors (loads plot). The score plot is usually used for studying the classification of the data clusters; while the loads plot can be used for giving information on the relative importance of the array sensors to each principal component and their mutual correlation (Penza, Cassano, & Tortorella, 2002).

2.4.2. Artificial neural networks

The most commonly used artificial neural network (ANN) to analyse multisensor array data is the multilayer perceptron (MLP) trained by the error back-propagation algorithm. An ANN consists of a nodes-net of information processing elements called neurones, which are connected together in a given way depending on the net architecture. The strengths of these connections are called weights, which are determined during the training stage for the supervised neural networks. They acquired "knowledge" by the calibration of the net tested by the prediction of unknown input vectors, which are not included in the training set used to learn the net. Generally, a MLP-based ANN is organized into a sequence of layers: the first layer is the input layer with a node for each variable (e.g., one sensor response), the output layer consisting of a node for each variable to be determined (e.g., one flavour class or chemical concentration) and a series of one or more hidden layers, between input and output layer, consisting of a given number of nodes. In an MLP, the nodes of the different layers are feed-forward fully connected; in other terms, the information directly flows from input layer to output layer of the network. Therefore, the signals are propagated from the input layer through the hidden layer(s) to the output layer. A node thus receives signals via connections or weights from other nodes (or from the external world for the nodes of the input layer). The net input for a node j is given by

$$\mathrm{net}_j = \sum_i w_{ji} o_i,$$

where i represents the nodes in a previous layer, w_{ji} is the weight associated with the connection from node i to the node j and o_i is the output of node i. The output of a node is determined by a nonlinear transfer function and the net input of the node. One among the most popular nonlinear transfer function (or activation function) is the sigmoid (or logistic function)

$$o_j = f(\mathrm{net}_j) = \frac{1}{1 + \exp[-(\mathrm{net}_j + \vartheta_j)]},$$

where θ_j is a bias term or threshold value of the node j responsible for accommodating nonzero offsets in the data.

The adequate functioning of a neural network strongly depends on the manner the signals are propagated through the net. The weights play an important role in this propagation and a proper setting of these weights is essential. Usually, this setting is not known a priori and the weights are initially given random and set in a small values range. The process of adapting the weights to an optimum set of values is called training, or learning, or calibration of the net. A representative training set is iteratively presented to the input of the neural network and the difference between the desired solution (target) and the net calculated one (output) is used to adapt the weights step-by-step, according to the learning algorithm. This difference, or error, is back-propagated from output to input of the network for a new iteration to correct the weights until the network error converges to an estimated level initially assigned. In the delta learning rule, a difference vector δ_{jk} is calculated by using the following equations:

$$\delta_{jk} = (t_j - o_k)(1 - o_k),$$

where t_j is the target vector for a known class j and o_k is the calculated output vector. The synaptic weights are modified using this difference vector δ_{jk}. In the popular gradient descent method, on each iteration τ, it is possible to write as follows:

$$\Delta w^{(\tau)} = -\eta \delta_{jk} f(\text{net}_j) + \alpha \Delta w^{(\tau-1)},$$

where η is the learning rate, which determines the rate of convergence of the net to the desired solution of minimum error, while α is the momentum term responsible of the stability of the convergence process. The modified weights $\Delta w^{(\tau)}$ are repeatedly fed back into the net. As previously mentioned, this procedure is repeated for a number of iterations (or epochs) until the network error converges to a suitable level initially set. A measure of the network performance is the total sum squared error E_{tss} defined as follows:2

$$E_{\text{tss}} = \sum_{k=1}^{p} \sum_{j=1}^{m} \delta_{jk}^2,$$

where m represents the number of output neurones (or similarly the output classes); while p is the number of patterns in the input data set.

Many factors influence the performance of a MLP-based ANN, such as the sensors array data pre-processing algorithms, net architecture, weights connectivity and the learning rule (Bishop, 1995; Gardner & Bartlett, 1999).

2.4.3. Pre-processing algorithms

The general approach of using some pre-processing techniques can significantly improve not only the classification performance of a linear technique, but also the recognition ability of the nonlinear techniques. Hence, we have evaluated the classification performance of nine different data pre-processing algorithms by defining three different sensor parameters – referred here as models (x_i) – and we have also studied, for each model, the effects of three different normalization techniques – referred here as methods (z_i). Table 3 lists the models, the methods and the equations used to define them. The no normalization technique, obviously, does not pre-process the data obtained by the models. The vector array normalization sets the sensor values in a range $[0, +1]$, whereas the autoscaling sets the mean value (μ_i) to 0 and the standard deviation (σ_i), or equivalently the variance (v_i), to 1. The vector array normalization divides each sensor value by the sum of all array sensors responses. This means that the flavour concentration dependence of the magnitude of the sensor response has been removed or reduced. Autoscaling is a scaling method that gives equal weighting to each sensor in the array and thus compensates for differences in the magnitude of the sensor signals.

In our study, a total of 45 points were collected: nine different wines sampled five times each. All set of experimental data was of 180 datapoints: 45 points (number of wine samplings) measured by four different sensors. For the chemometric characterization of the wines by PCA pattern recognition, all 45 points have been processed. For the wines classification by the ANNs recognizer, the 45 points have been split into a training set of 36 patterns and a test set of patterns. The data of the first four exposures to wines headspace have been utilized to train the network, while the fifth exposure of each wine sampling was used to test the net.

3. Results and discussion

3.1. PCA-based classification of standard analysis data and headspace of wines

Generally, the wines are characterized by standard analysis consisting in the measurement of different chemico-physical parameters (e.g., pH, total acidity (g l^{-1}), titratable acidity (g l^{-1}), volatile acidity (g l^{-1}), malic acid (g l^{-1}), tartaric acid (g l^{-1}), alcoholic content (vol%), ionic conductivity (μS cm^{-1}), free SO_2 (mg l^{-1}), total SO_2 (mg l^{-1}), residual sugar (g l^{-1}), etc.) by conventional analytical techniques. In our study, we have considered only the routine analysis based on the alcoholic content, pH and ionic conductivity, as reported in the Table 1. These data have been processed by extracting the features by PCA method. Fig. 1 shows the PCA results based on the three measured parameters for the nine samples of Italian wines examined. The PC2–PC3, PC1–PC3 and PC1–PC2 scores plots retain an

Table 3
Definitions of the sensor models and normalization methods used for the chemometric characterization of the wines

Model/Method	Equation	Abbreviation
Relative ratio	$x_i = R_i/R_f$[a]	Ratio
Difference	$x_i = \Delta R = R_i - R_f$	Diff.
Relative difference	$x_i = \Delta R/R_i = (R_i - R_f)/R_i$	Rel. Diff.
No normalization		No norm.
Vector array normalization	$z_i = x_i/(x_1 + x_2 + \cdots + x_n)$	Array norm.
Autoscaling	$z_i = (x_i - \mu_i)/\sigma_i$[b]	Autoscaling

[a] R_f and R_i are the value of the steady state of the electrical resistance of the single sensor in presence of the flavour sample under test at a fixed time t_0 after start of exposure, and in dry air, respectively.
[b] μ_i is the mean; σ_i is the standard deviation.

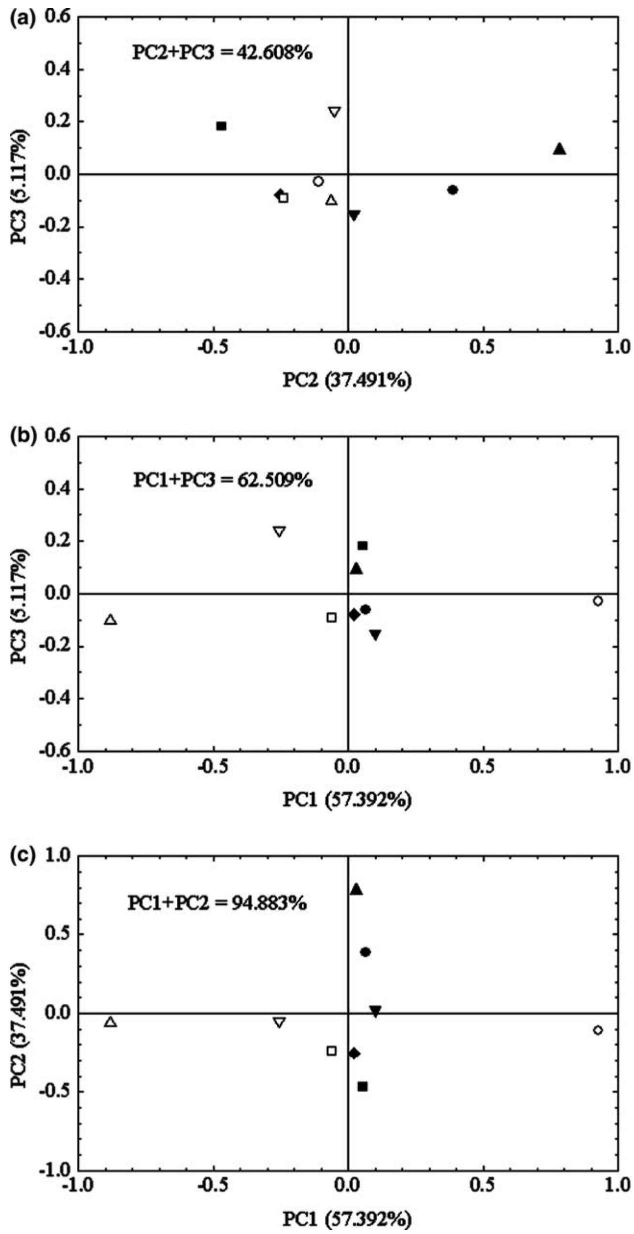

Fig. 1. PCA results of the standard analysis of wines by using data only related to the alcoholic content, pH and ionic conductivity. The scores plots (a) PC2–PC3, (b) PC1–PC3 and (c) PC1–PC2 retain an increasing cumulative variance of 42.608%, 62.509% and 94.883%, respectively. The nine wines under test are *Chardonnay del Salento* (■); *San Severo* (●); *Salento-Agrivia* (▲); *Soleto* (▼); *Chianti* (◆); *Matino* (□); *Mesagne* (○); *Castel del Monte* △ and *Salento-Mottura* (▽).

increasing cumulative variance of 42.608%, 62.509% and 94.883%, respectively. It is completely evident that this standard methodics does not allow a clear distinction of the wines considered. As an example, the datapoints of the two red wines (*Chianti* and *Matino*) are overlapped in the PC2–PC3 plane and they are not well separated even in the PC1–PC2 plane with the greatest cumulative variance. Maybe, PCA method could better discriminate the wines by enhancing the number of processed physico-chemical parameters.

An alternative way to discriminate wines is the headspace analysis by using a multisensors array for determining their chemical fingerprint from sensor signals. Different techniques of volatile components extraction and sampling of the wine headspace have been proposed by various research teams, such as the dynamic headspace (Guadarrama et al., 2000), the solid-phase micro-extraction (SPME) (Guadarrama et al., 2001), the stripping and the liquid–liquid extraction (Ortega-Heras, Gonzales-San José, & Beltran, 2002), the organophilic pervaporation (Pinheiro, Rodrigues, Schafer, & Crespo, 2002) and the static headspace (Di Natale et al., 1996). The static headspace method for the wine analysis is very sensitive to highly or medium volatile compounds present in high concentrations (a few percents), such as ethanol or other interfering alcohols. Hence, this method is scarcely able to detect trace compounds, which are responsible for the determination of wine aroma and provide the typical specificity of the wine headspace. However, the static headspace is a simple, fast and reproducible sampling method with extraction process automated, low-cost, low-toxicity because no solvent is used for extraction. In our measurements, the static headspace sampling has been used to attempt to rapidly evaluate how the aroma compounds present in the wine headspace at minor concentrations (a few ppm) influence the specificity of sampled wine for an possible attempt of characterization of the wines under test.

Fig. 2 shows the typical transient responses of four sensors, operating at 250 °C, exposed towards the headspace of three different wines (one white, one red and one rosè), sampled by the static method with dry air as carrier gas of the volatile components of the wine headspace. The measurements were repeated six times for each wine sample. As it can be noticed, the electrical resistance (sensor signal) of each sensor downshifts upon exposure of examined wine samples and returns to baseline level when dry air is switched again into sensors test cell to recover them. This typical behaviour of the array sensors has been achieved with various kinetics for all wines samples. The sensor responses are fast and reproducible. The data so-obtained from multisensors array have been processed by the PCA technique to investigate the chemometric differentiation of the wines sampled. For data homogeneity, only the last five exposures over total six exposures (the first one off) for all wines samples have been considered for further analysis of classification.

Fig. 3 shows the score plot in the PC1–PC2 plane of the data related to the four-sensor array using the normalized responses (R_i/R_f) as sensor signals for the nine Italian wines (three white, three red and three rosè) considered. Array data from exposure to individual ethanol and methanol (alcohols present in wine headspace) have been also included in the PCA study in or-

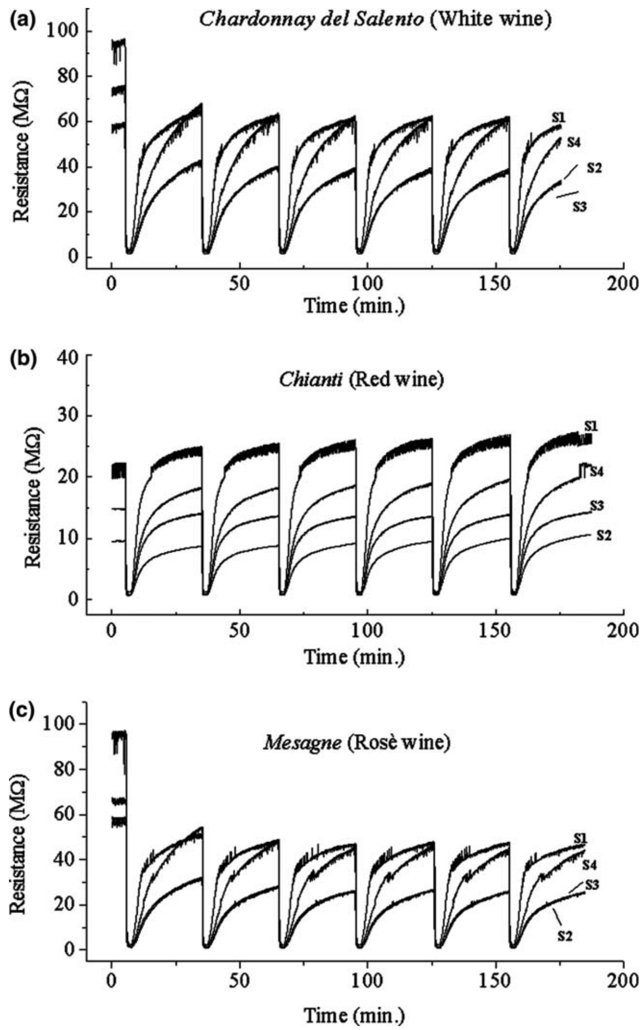

Fig. 2. Typical time responses of the four sensors of array exposed to the headspace of (a) *Chardonnay del Salento* (white wine), (b) *Chianti* (red wine) and (c) *Mesagne* (rosè wine). The exposure to each wine headspace was repeated six times. The exposure time is 2 min. The sensors operated at 250 °C. Dry air was used as reference and carrier gas. The sensors used in the array are WO_3:Pt (S1); WO_3:Au (S2); WO_3:Pd (S3) and WO_3:Bi (S4).

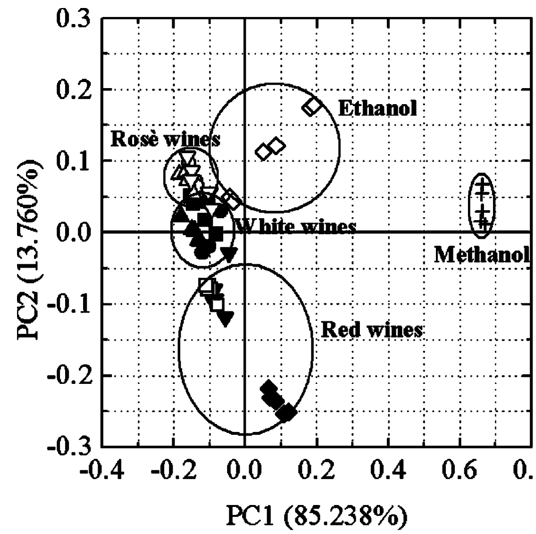

Fig. 3. PC1–PC2 scores plot of nine wines (three white, three red and three rosè), sampled by static headspace, using the data of the normalized responses (R_i/R_f) obtained from a multisensors array of four sensing elements. Five exposures for each wine sample have been evaluated. The wines under test are *Chardonnay del Salento* (■); *San Severo* (●); *Salento-Agrivia* (▲); *Soleto* (▼); *Chianti* (◆); *Matino* (□); *Mesagne* (○); *Castel del Monte* (△); *Salento-Mottura* (▽). Also ethanol (◇) and methanol (+) are indicated in the plot as comparative reference clusters.

Table 4
Variance captured by PCA performed on the correlation matrix of data obtained from the normalized responses (R_i/R_f) of the four sensors in the array for nine Italian wines

PCs	Eigenvalue	Variance (%)	Cumulative variance (%)
1	3.410	85.238	85.238
2	0.550	13.760	98.999
3	0.040	1.001	99.999
4	0.000	0.001	100.000

der to set some reference clusters for comparison to the wine clusters. The results obtained indicate that three different macro-clusters referred to white wines, red wines and rosè wines are retained and two clusters of ethanol and methanol are also discriminated, with methanol better separated from ethanol and other wine clusters. The clusters of the rosè wines and white wines are partially overlapped. However, within each wine macro-cluster, the differentiation among the wines of the same class fails, because the peculiarities (aroma compounds present in low content) of the wine headspace are hidden or cancelled by the interfering compounds (mainly alcohols) present in high amount. As evidenced, the different wines belonging to the same class are misclassified, although the cluster related to *Chianti* (a red wine) is well discriminated; perhaps, due to its geographical origin (Tuscania) which is different from all remaining wines (Apulia). Table 4 shows the percentage variance for this PCA study. Moreover, the PCA analysis, performed on data from a four-sensor array by using the other sensor parameters (ΔR, $\Delta R/R_i$) and normalization techniques (no normalization, autoscaling), shows worse or not better discrimination for the examinated wines.

3.2. ANN-based classification of headspace of wines

The discrimination of the tested wines belonging to the same class (white, red, and rosè) has been tackled with a patterns recognizer based on artificial neural network providing nonlinearity in the multivariate classification performance. A feed-forward fully connected

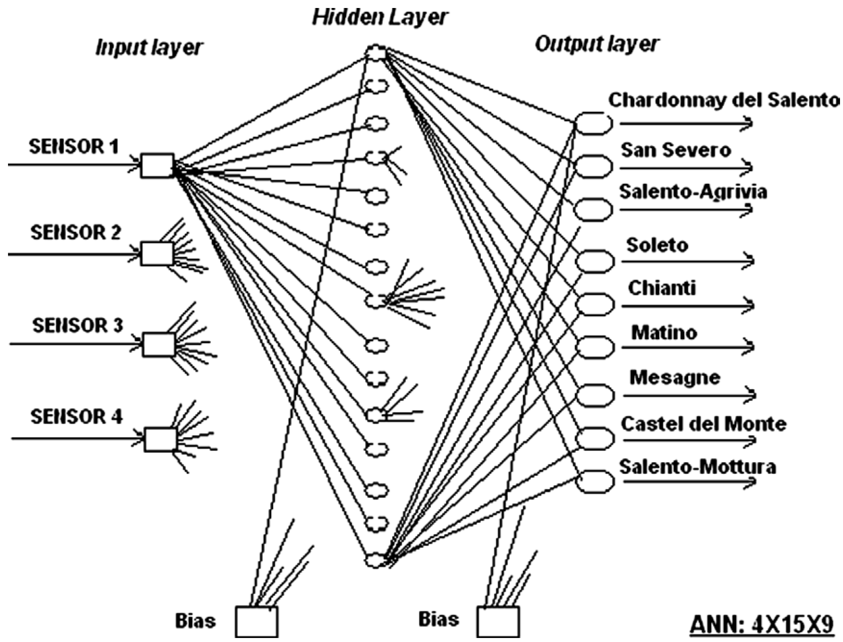

Fig. 4. Three-layered ANN used as patterns classifier for wines identification.

ANN has been trained with the back-propagation learning algorithm. The classifier used, depicted in Fig. 4, was a three-layer net architectured by an input layer with four neurones, a hidden layer with 15 neurones and an output layer with nine neurones. The available set of 45 points has been divided into a 36-point training set and a 9-point testing set.

The strategy used for the classification of the wines is the so-called one-of-many encoding (Di Natale et al., 2001). The output of the network is a multi-dimensional vector with the number of the dimensions equal to the number of the classes (wines) to be determined. Each vectorial dimension is assigned to a class. In the net training file, the class membership of a single data is coded in a numerical format by assigning 1 to the belonging class and 0 to the all others, e.g., *Chardonnay del Salento* code is (10000000); *San Severo* code (010000000) and so on, finally *Salento-Mottura* code (000000001). In the net testing file, the membership of an input data is assigned to the class with greatest net output. The higher is the ratio between the highest output and the second greater output, the higher is the classification score of the input data, the better is the estimation of the assigned membership.

All networks were trained with the learning parameters set to values of learning rate $\eta = 0.1$; momentum term $\alpha = 0.95$; an initial weights range $(-10, +10)$ and a activation function as sigmoid. The training process was stopped before an over-training situation of the net with the stopping criteria of 100% correct percentage of the training set and a proper number of at least 10000 iterative epochs through the net up to a RMS error of 0.01 for each learned pattern. After training step, the test set was presented to the trained ANNs to investigate the identification capability of each net. The prediction rate of a net, defined as the ratio of number of correctly identified patterns to that of total test patterns, has been evaluated. Fig. 5 shows a plot of only the prediction rate for each of the three sensor models examined for wines identification. It appears that the ratio model is better than the others at classifying the wine type, as previously outlined in PCA employing the sensor responses (R_i/R_f) as variables. Similarly, the effect of the choice of the normalization method on the percentage correctly classified has been also investigated. Fig. 6 shows a plot of the prediction rate for

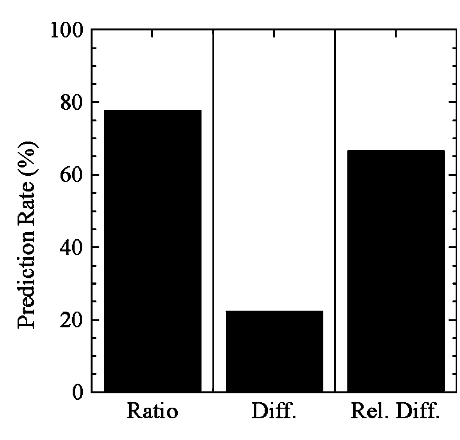

Fig. 5. A bar chart showing the effect of the three sensor models investigated (ratio, difference and relative difference) on the prediction rate of the ANNs used for the task of discrimination of the nine Italian wines under test.

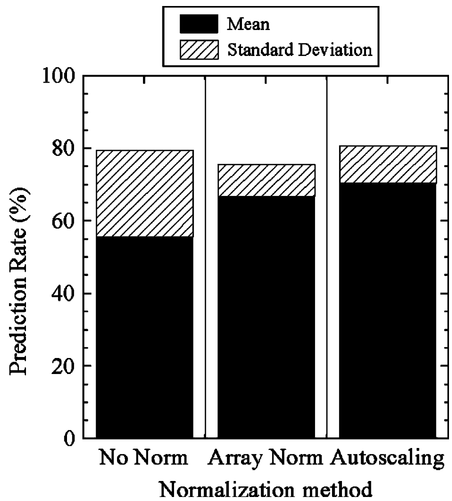

Fig. 6. A bar chart showing the effect of the three normalization methods investigated (no normalization, array normalization and autoscaling) on the prediction rate of the ANNs used for the task of discrimination of the nine Italian wines under test. The error bars indicate the standard deviation of the results averaged over all three sensor models tested.

each of the three normalization methods, as defined in Table 3, and averaged over all the three sensor models (or sensor parameters) investigated. A ranking of the normalization algorithms was achieved from the highest method to the lowest method as follows as: autoscaling, array normalization and finally no normalization. Also, the accuracy of the normalization methods is better than no-normalization algorithm with their standard deviation about two times lower. These results confirm that, usually, the data pre-processing enhances the performance of classification of a neural network with the autoscaling method better than array normalization one.

The overall performance of the predictive recognizer based on ANNs can be appreciated in details by the confusion matrix obtained for each one of the three sensor parameters used. The results are presented in Tables 5–7, averaging the data (a 27-point set generated by 9 wines $\times$ 3 methods) related to each sensor parameter over all the three normalization methods considered. The map of the ANNs prediction perfor-

Table 5
Confusion matrix obtained by using the sensor parameter R_i/R_f, including all three normalization methods investigated (no normalization, array normalization and autoscaling), of the ANNs classifier performing the task of discrimination of the nine Italian wines under test

True/Predicted class	Chardonnay del Salento	San Severo	Salento-Agrivia	Soleto	Chianti	Matino	Mesagne	Castel del Monte	Salento-Mottura
Chardon. Salento	3	0	0	0	0	0	0	0	0
San Severo	0	3	0	0	0	0	0	0	0
Salento-Agrivia	0	0	3	0	0	0	0	0	0
Soleto	0	0	0	2	1	0	0	0	0
Chianti	0	0	0	0	3	0	0	0	0
Matino	0	0	1	0	0	2	0	0	0
Mesagne	0	0	0	0	0	0	2	0	1
Castel del Monte	0	0	0	0	0	0	2	1	0
Salento-Mottura	0	0	0	0	0	0	2	1	0

The total data set to be identified consists of 27 points (9 wines $\times$ 3 methods).

Table 6
Confusion matrix obtained by using the sensor parameter ΔR, including all three normalization methods investigated (no normalization, array normalization, autoscaling), of the ANNs classifier performing the task of discrimination of the nine Italian wines under test

True/Predicted class	Chardonnay del Salento	San Severo	Salento-Agrivia	Soleto	Chianti	Matino	Mesagne	Castel del Monte	Salento-Mottura
Chardon. Salento	1	2	0	0	0	0	0	0	0
San Severo	0	2	1	0	0	0	0	0	0
Salento-Agrivia	1	0	2	0	0	0	0	0	0
Soleto	0	0	0	2	1	0	0	0	0
Chianti	0	0	0	0	3	0	0	0	0
Matino	0	0	1	0	0	2	0	0	0
Mesagne	1	0	1	0	0	0	1	0	0
Castel del Monte	3	0	0	0	0	0	0	0	0
Salento-Mottura	0	0	2	0	0	0	0	0	1

The total data set to be identified consists of 27 points (9 wines $\times$ 3 methods).

Table 7
Confusion matrix obtained by using the sensor parameter, $\Delta R/R_i$, including all three normalization methods investigated (no normalization, array normalization and autoscaling), of the ANNs classifier performing the task of discrimination of the nine Italian wines under test

True/Predicted class	Chardonnay del Salento	San Severo	Salento-Agrivia	Soleto	Chianti	Matino	Mesagne	Castel del Monte	Salento-Mottura
Chardon. Salento	2	0	1	0	0	0	0	0	0
San Severo	0	3	0	0	0	0	0	0	0
Salento-Agrivia	0	0	3	0	0	0	0	0	0
Soleto	0	0	0	3	0	0	0	0	0
Chianti	0	0	0	0	3	0	0	0	0
Matino	0	0	0	1	0	2	0	0	0
Mesagne	0	0	0	0	0	0	3	0	0
Castel del Monte	0	0	0	0	0	0	3	0	0
Salento-Mottura	0	0	0	0	0	0	3	0	0

The total data set to be identified consists of 27 points (9 wines × 3 methods).

mance has been summarized per each wine sample and wine class in Table 8, including all data from Tables 5–7, in order to better outline the wine classification results. As an example, the white wines class has been 100% correctly predicted by the sensor parameter R_i/R_f, whereas in the rosè wines class only 22.2% was correctly predicted by the sensor parameter ΔR. In addition, some wine samples have been 100% correctly predicted (*Chardonnay del Salento*, *San Severo*, *Salento-Agrivia*, *Soleto*, *Chianti* and *Mesagne*) by using the proper sensor parameter as input to ANNs recognizer; but, unfortunately, some other wine samples have not been correctly predicted at all (*Castel del Monte* and *Salento-Mottura*). These two misclassified wines belong to the rosè wines class with the lowest prediction rate. Probably, the prediction rate could be successfully improved by analysing the wines with other sampling methods – different from static headspace – able to better detect the trace volatile compounds of the headspace for a more accurate wines discrimination. In fact, the specificity of the wine headspace is highly influenced by low-content volatile compounds; therefore the interfering high-content compounds have to be removed from the headspace by using more suitable sampling methods to guarantee an enhanced sensing of the aroma-relevant compounds (Pinheiro et al., 2002).

Table 8
Prediction rate of the ANNs classifier by using the different three sensor parameters investigated, including all three normalization methods, for each wine sample and wine class examined

Class/Wine	Prediction rate (%)		
	Sensor parameter averaged over all three normalization methods		
	R_i/R_f	ΔR	$\Delta R/R_i$
Chardonnnay del Salento	100	33.3	66.7
San Severo	100	66.7	100
Salento-Agrivia	100	66.7	100
All white wines mean	100	55.56	88.9
Soleto	66.7	66.7	100
Chianti	100	100	100
Matino	66.7	66.7	66.7
All red wines mean	77.8	77.8	88.9
Mesagne	66.7	33.3	100
Castel del Monte	33.3	0	0
Salento-Mottura	0	33.3	0
All rosè wines mean	33.3	22.2	33.3

Data are from Tables 5–7.

The specific identification results based on the ANNs classifier for the nine Italian wines by using the three no-normalized sensor parameters as inputs to the nets have been illustrated in Figs. 7–9. The predicted classes by ANNs classifier are compared to the true classes by studying the net output and the target value. Assuming a classification threshold of (0.1, 0.7) with the first term as superior limit and the second term as inferior limit for the lower net output and highest net output, respectively, the wine patterns analysis has been evaluated. As it is clearly evident, the resulting classification by the ANNs correctly identifies some wine samples, while others are misclassified with the prediction rate achieved for the three sensor parameters (ratio, difference and relative difference) of 78%, 22% and 67%, respectively.

4. Conclusions

In this paper, the identification of the nine distinct Italian wines has been attempted. The standard analysis

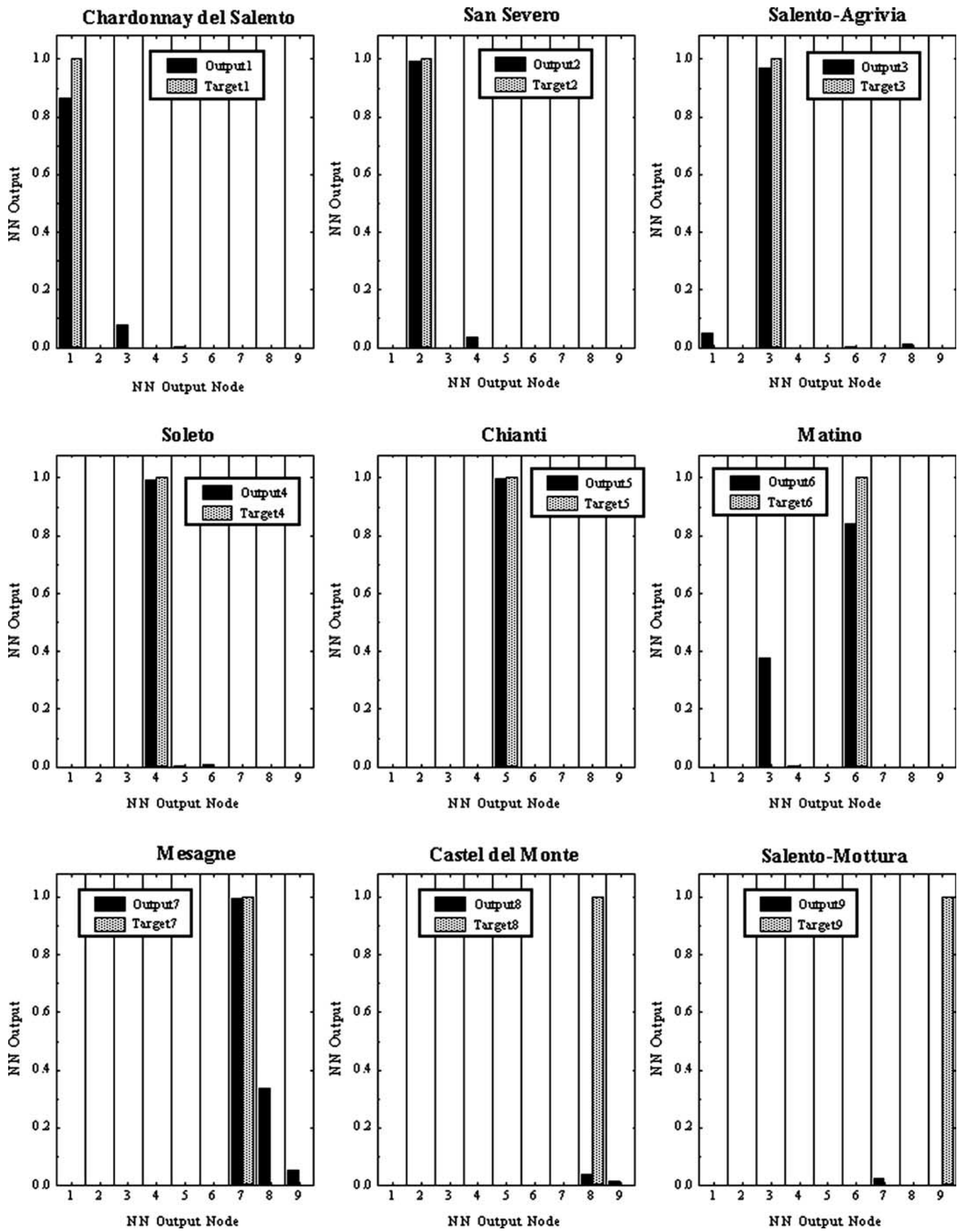

Fig. 7. Patterns of ANNs classification by using the no-normalized sensor parameter R_i/R_f as input to the neural network for the nine Italian wines studied.

data (pH, alcoholic content and ionic conductivity) have been performed by multivariate statistical methods, such as principal component analysis, for classification purposes by showing misclassified wine cases. Mainly, the application of a multisensors array, used as electronic nose-type sensing system, for the chemical analysis of the headspace of nine several Italian wines (three white, three red and three rosè) from different denominations of origin and vintage years is presented. The static headspace sampling of the wines has been employed due to its characteristics of simplicity, rapidity and reproducibility. The chemometric characterization of the sampled wines by means of the data collected from a four-sensor array associated with an artificial neural networks algorithm facilitates the clustering and differentiation of the examined wines to achieve various classification tasks. The classification performance of the neural network has been evaluated by using nine different pre-processing

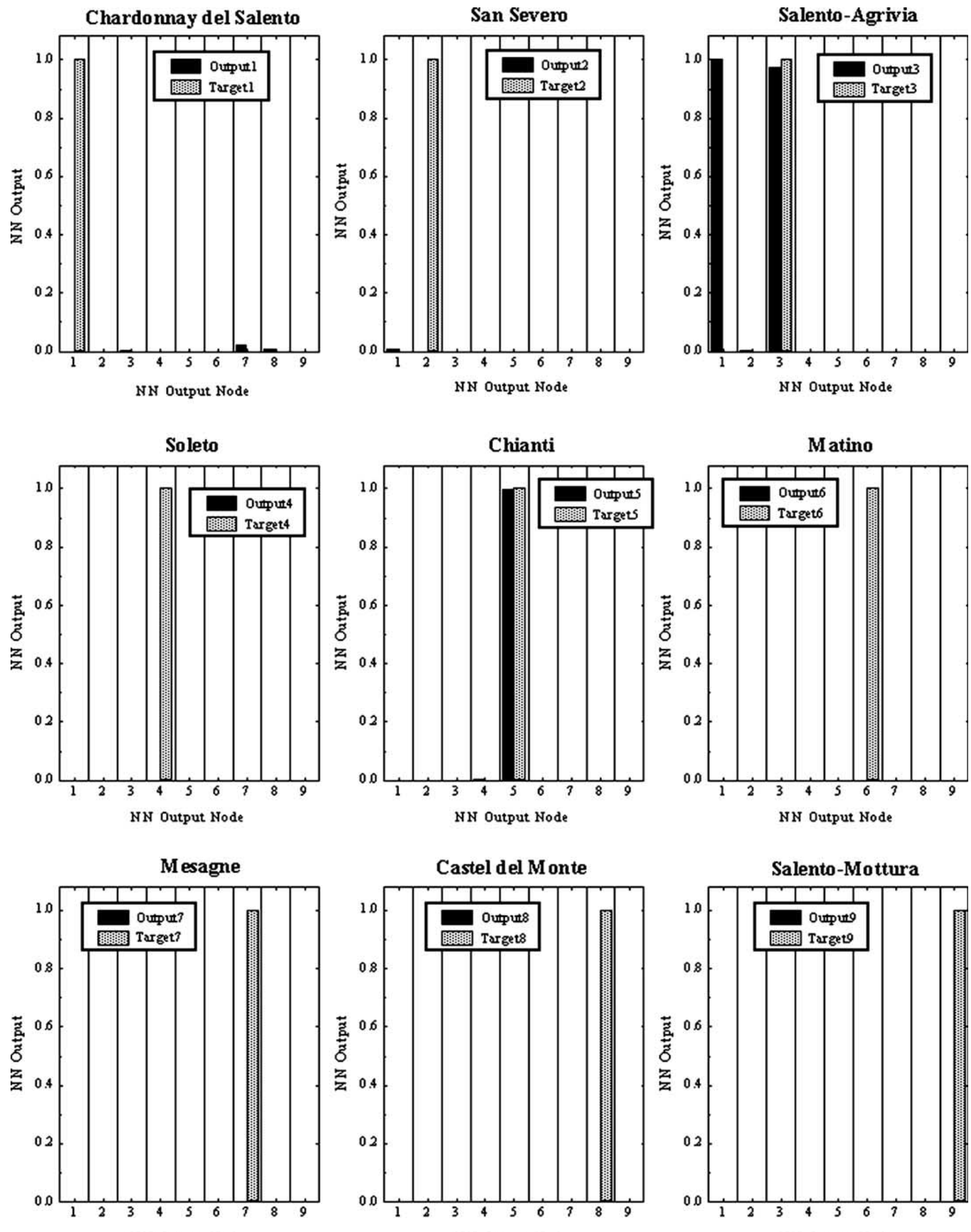

Fig. 8. Patterns of ANNs classification by using the no-normalized sensor parameter ΔR as input to the neural network for the nine Italian wines studied.

algorithms on the basis of three different sensor parameters (ratio, difference and relative difference) and three different normalization techniques (no normalization, array normalization and autoscaling). The best results obtained indicate a recognition rate and a prediction rate as high as 100% and 78%, respectively. These figures of merit are respectable and satisfactory. The net classification performance of the sampled wines could be enhanced by using more selective and sensitive headspace sampling methods responsible for the specific wine aroma in order to better discriminate the samples tested. This work demonstrates the feasibility of an electronic nose for wines discrimination.

Acknowledgements

This work was financially supported by ENEA in the framework "Intesa Enea-Miur" (ex "Intesa Enea-Murst") under project 4335/04.

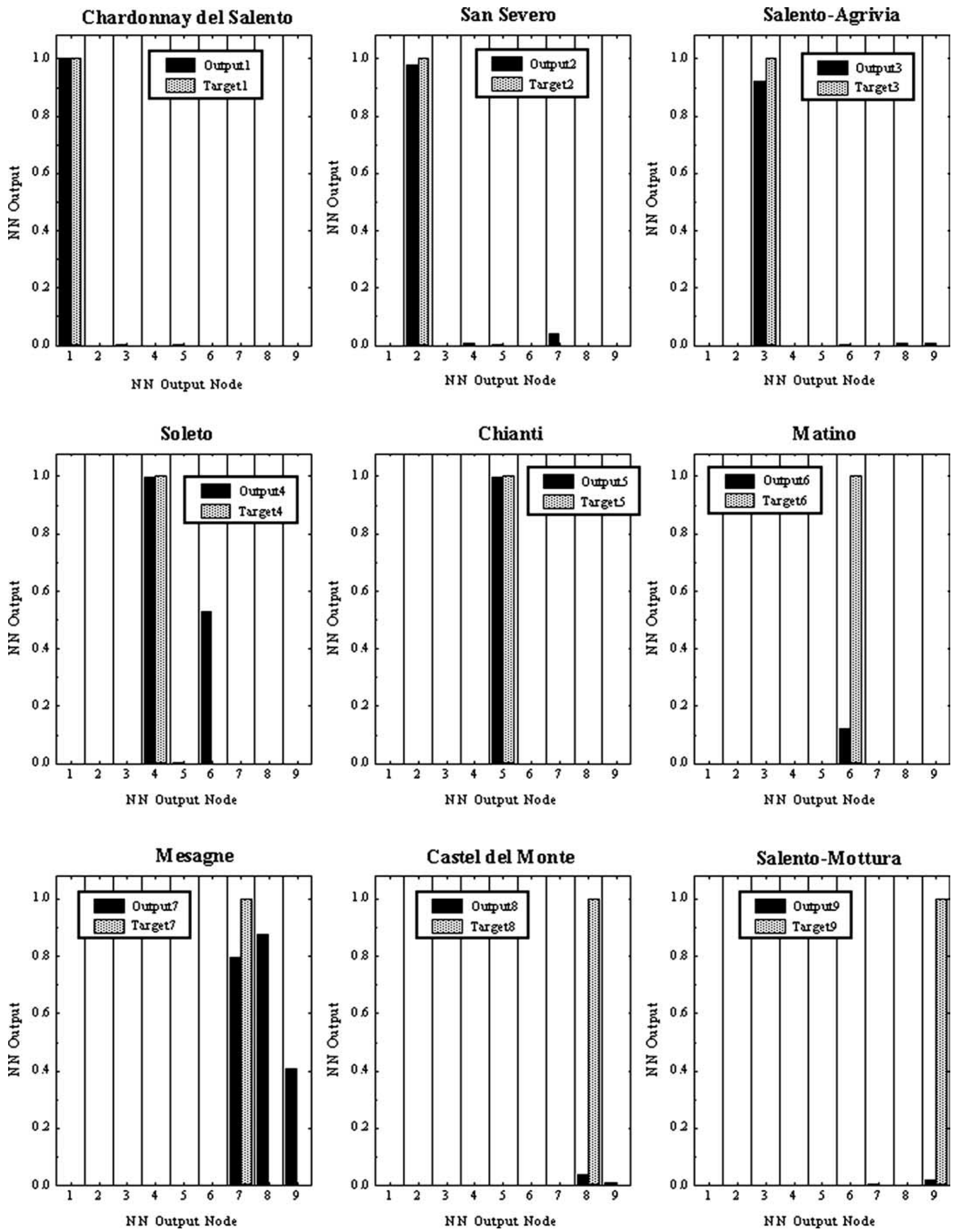

Fig. 9. Patterns of ANNs classification by using the no-normalized sensor parameter $\Delta R/R_i$ as input to the neural network for the nine Italian wines studied.

References

Anklam, E., Lipp, M., Radovic, B., Chiavaro, E., & Palla, G. (1998). Characterisation of Italian vinegar by pyrolysis mass spectrometry and a sensor device ('electronic nose'). *Food Chemistry, 61*, 243–248.

Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford: Oxford University Press.

Brescia, M. A., Caldarola, V., De Giglio, A., Benedetti, D., Fanizzi, F. P., & Sacco, A. (2002). Characterization of the geographical origin of Italian red wines based on traditional and nuclear magnetic resonance spectrometric determinations. *Analytica Chimica Acta, 458*, 177–186.

Di Natale, C., Davide, A. M., D'Amico, A., Sberveglieri, G., Nelli, P., Faglia, G., & Perego, C. (1995). Complex chemical pattern recognition with sensor array: The discrimination of vintage years of wine. *Sensors and Actuators B, 24-25*, 801–804.

Di Natale, C., Davide, A. M., D'Amico, A., Nelli, P., Groppelli, S., & Sberveglieri, G. (1996). An electronic nose for the recognition of the vineyard of a red wine. *Sensors and Actuators B, 33*, 83–88.

Di Natale, C., Macagnano, A., Martinelli, E., Proietti, E., Paolesse, R., Castellari, L., Campani, S., & D'Amico, A. (2001). Electronic nose based investigation of the sensorial properties of peaches and nectarines. *Sensors and Actuators B, 77*, 561–566.

Frias, S., Conde, J. E., Rodriguez, M. A., Dohnal, V., & Perez-Trujillo, J. P. (2002). Metallic content of wines from the Canary Islands (Spain). Application of artificial neural networks to the data analysis. *Nahrung/Food, 46*, 370–375.

Gardner, J. W., & Bartlett, P. N. (1999). *Electronic noses – principles and applications*. Oxford: Oxford University Press.

Guadarrama, A., Fernandez, J. A., Iniguez, M., Souto, J., & de Saja, J. A. (2000). Array of conducting polymer sensors for the characterisation of wines. *Analytica Chimica Acta, 411*, 193–200.

Guadarrama, A., Fernandez, J. A., Iniguez, M., Souto, J., & de Saja, J. A. (2001). Discrimination of wine aroma using an array of conducting polymer sensors in conjunction with solid-phase micro-extraction (SPME) technique. *Sensors and Actuators B, 77*, 401–408.

Kallithraka, S., Arvanitoyannis, I. S., Kefalas, P., El-Zajouli, A., Soufleros, E., & Psarra, E. (2001). Instrumental and sensory analysis of Greek wines; implementation of principal component analysis (PCA) for classification according to geographical origin. *Food Chemistry, 73*, 501–514.

Marengo, E., Aceto, M., & Maurino, V. (2001). Classification of Nebbiolo-based wines from Piedmont (Italy) by means of solid-phase microextraction-gas chromatography-mass spectrometry of volatile compounds. *Journal of Chromatography A, 943*, 123–137.

Ortega-Heras, M., Gonzales-San José, M. L., & Beltran, S. (2002). Aroma composition of wine studied by different extraction methods. *Analytica Chimica Acta, 458*, 85–93.

Peña, R. M., Latorre, M. J., Garcia, S., Botana, A. M., & Herrero, C. (1999). Pattern recognition analysis applied to classification of Galician (NW Spain) wines with certified brand of origin Ribeira Sacra. *Journal of the Science of Food and Agriculture, 79*, 2052–5056.

Penza, M., Martucci, C., & Cassano, G. (1998). NO_x gas sensing characteristics of WO_3 thin films activated by noble metals (Pd, Pt, Au) layers. *Sensors and Actuators B, 50*, 52–59.

Penza, M., Cassano, G., & Tortorella, F. (2001). Gas recognition by activated WO_3 thin-film sensor array. *Sensors and Actuators B, 81*, 115–121.

Penza, M., Cassano, G., Tortorella, F., & Zaccaria, G. (2001). Classification of food, beverages and perfumes by WO_3 thin-film sensors array and pattern recognition techniques. *Sensors and Actuators B, 73*, 76–87.

Penza, M., Cassano, G., & Tortorella, F. (2002). Identification and quantification of individual volatile organic compounds in a binary mixture by SAW multisensor array and pattern recognition analysis. *Measurement Science and Technology, 13*, 846–858.

Pérez-Magariño, S., Ortega-Heras, M., & Gonzales-San José, M. L. (2002). Multivariate classification of rosè wines from different Spanish protected designations of origin. *Analytica Chimica Acta, 458*, 187–190.

Pinheiro, C., Rodrigues, C. M., Schafer, T., & Crespo, J. G. (2002). Monitoring the aroma production during wine-must fermentation with an electronic nose. *Biotechnology and Bioengineering, 77*(6), 632–640.

Pinheiro, C., Serrano, B., Schafer, T., Crespo, J. G., Gregorkiewitz, M., Fort, A., Vignoli, V., Rocchi, S. (2002). Comparison of two electronic nose systems for wine discrimination after selective enrichment with pervaporation. In *Technical digest of the 9th international symposium on olfaction and electronic nose* (pp. 198–199). Rome, September 30–October 2.

Raptis, C. G., Siettos, C. I., Kiranoudis, C. T., & Bafas, G. V. (2000). Classification of aged wine distillates using fuzzy and neural network systems. *Journal of Food Engineering, 46*, 267–275.

Rodriguez-Delgado, M.-A., Gonzalez-Hernandez, G., Conde-Gonzalez, J.-E., & Perez-Trujill, J.-P. (2002). Principal component analysis of the polyphenol content in young red wines. *Food Chemistry, 78*, 523–532.

Soufleros, E. H., Bouloumpasi, E., Tsarchopoulos, C., & Biliaderis, C. G. (2003). Primary amino acid profiles of Greek white wines and their use in classification according to variety, origin and vintage. *Food Chemistry, 80*, 261–273.