J Sci Food Agric 1998, **78**, 417–422

Characterisation of White Vinegars of Different Sources with Artificial Neural Networks

Vincenzo Gerbi,[1]* Giuseppe Zeppa,[1] Riccardo Beltramo,[2] Alberta Carnacini[3] and Andrea Antonelli[4]

[1] Dipartimento Valorizzazione e Protezione Risorse Agroforestali—Settore Microbiologia e Industrie agrarie, Università di Torino, Via P Giuria 15, 10126 Torino, Italy
[2] Dipartimento Scienze Merceologiche, Università di Torino, P.zza Arbarello 8, 10122 Torino, Italy
[3] Istituto Microbiologia e Tecnologia agraria e forestale—Università di Reggio Calabria, P.zza S. Francesco Gallina 4, 89100 Reggio, Calabria, Italy
[4] Istituto Industrie agrarie, Università di Bologna, Via S. Giacomo 7, 40126 Bologna, Italy

(Received 11 July 1996; revised version received 8 December 1997; accepted 15 March 1998)

Abstract: Wine and cider vinegars currently attract growing interest from consumers, giving rise to a consequent increase in supply. A full appreciation of their quality is only possible, however, through recognition of their superior quality when compared with vinegars produced from raw materials of inferior quality. Therefore, it is necessary to pinpoint the parameters that define the quality of these products. Chemico-physical and sensory analysis has been used to draw up artificial neural networks (ANNs), on the basis of a vast sampling of vinegars from various countries, produced from a variety of raw materials, that was already subjected to multivariate statistical analysis. Among the chemical parameters, polyalcohols and other elements such as pH, tartaric acid and proline proved to be highly reliable, whereas other volatile substances and the results of sensory analysis were not very discriminating and could not be used to re-classify samples of unknown origin. The positive results obtained indicate that ANNs are a powerful mathematical tool, since they can be used to construct models that predict the botanical origin of the product and to re-classify samples of unknown origin, without any initial restrictive hypothesis. © 1998 Society of Chemical Industry.

J Sci Food Agric **78**, 417–422 (1998)

Key words: vinegars; artificial neural networks; sensory analysis; analytical parameters; gas chromatography

INTRODUCTION

Over-production in some agricultural sectors clearly represents one of the major concerns of agriculture in the European Union. Genetic improvements and changes in dietary habits, as well as errors in political planning, have led to the periodic accumulation of excesses that must also be periodically destroyed to avoid price collapses, with disastrous consequences for the agricultural sector. There is, therefore, a great need to avoid the creation of future excesses and to diminish those that have already been produced, while, at the same time, to keep reducing the social cost of the mere destruction of excess produce.

Research in the agricultural food producing sector is called for, therefore, to assume a primary role aimed at the development of new uses for raw materials and at improving the quality of conventional products, so as to avoid a market reasoning largely based on quantity. An increase in the quality of agricultural food products is only possible, however, if adequate chemico-physical and/or sensory parameters are established to correctly

Contract/grant sponsor: CNR—RAISA Project.
Contract/grant number: Subproject 4.
* To whom correspondence should be addressed.

 J Sci Food Agric 0022–5142/98/$17.50. Printed in Great Britain

characterise such products objectively. In the case of vinegars, in particular, these difficulties are considerably increased by the fact that it is a product of a double series of more of less drastic micro-biological and technical operations on raw materials from many origins and of various characteristics.

Of the over 4·5 million hectolitres of vinegar, with 10% acidity, produced in the European Union, only 0·9 million hectolitres are produced from wine, with 0·4 million produced from malt or cider and 3·2 million from alcohol. Since Italy produces about 540 000 hl of vinegar per year from the acetic bioxydation of wine, it is obvious that, together with Spain, another major producer of vinegar manufactured only from wine, Italy is particularly interested in characterising this product, produced from the bioxydation of very complex and commercially valuable raw material. This policy is justified in Italy and other Mediterranean countries, such as Spain, by economic and cultural considerations aimed at protecting consumers and typical Mediterranean products. These products can, in fact, only be truly protected by official recognition of their superior quality based on chemico-physical and/or sensorial parameters derived from studies aimed at characterising such products by analysing the variability of their characterisation in function of the place of origin, production technology and commercial type.

The number of publications aimed at characterising wine vinegars is quite limited compared to literature dedicated to other food products, and it is totally insufficient for the purposes of qualifying its superior quality when compared to the far inferior but much more widespread alcohol vinegars. The present research was not, therefore, aimed at finding protection parameters for legal controls that already exist in current norms (Ministero Agricoltura e Foreste 1965; Mecca and Vicario 1971; Sakata *et al* 1991) but at isolating parameters capable of defining the quality, origin and worth of commercially distributed vinegars.

The results from chemical, physical and sensory analysis of a large number of wine, alcohol, cider, malt and honey vinegars, elaborated through multivariate statistical analysis, led to the identification of discriminant parameters that could be correlated to the sensory quality of the product (Carnacini and Gerbi 1992; Antonelli *et al* 1993; Gerbi *et al* 1994; Carnacini *et al* 1995; Antonelli *et al* 1994b). This approach does not allow the transfer of results to unknown samples not utilised for the initial mathematical model. In fact, only linear discriminant analysis (LDA) can be used to reclassify new samples. However, mathematical requirements, and the need to include a high number of samples in the validation set, make it difficult to apply LDA.

A possible solution to these problems may reside in the use of artificial neural networks (ANNs) that have already been applied in the field of nutritional sciences (Romeo 1992, 1993; Margarita *et al* 1994), since they constitute an efficient tool for characterising and discriminating food products. ANNs have been shown to be efficient in learning, or pinpointing, data regularities and in generalising such regularities to provide coherent results, even with cases not taken into account during learning.

MATERIALS AND METHODS

Sixty-five vinegar samples acquired on Italian, French, Spanish and Swiss markets were analysed. The samples were divided into three categories according to raw matter: wine (49 samples), cider (12 samples) and alcohol (4 samples). Each sample represented a different brand name, resulting in an imbalance in the number of samples in each category, since cider and, more especially, alcohol vinegars are not easily available on the market. Unlike multivariate analysis, the construction of ANNs is not significantly affected by this imbalance, since samples are attributed to the training set and the validation set at random, thereby ensuring that the various categories are represented in both data sets.

The main analytical parameters (density, total acidity, volatile acidity, fixed acidity, dry matter, ash and their alkalinity, pH) were determined according to official Italian methods (Ministero Agricoltura e Foreste 1965). Ethanol was determined by packed gas chromatography after neutralisation of the sample with Na_2CO_3 (Antonelli 1994). Tartaric, malic, lactic, citric and succinic acids and glycerol were determined by HPLC using a ion exchange column. Total polyphenols were determined by spectrophotometry (Singleton and Rossi 1965), while fractionation of tannic and non-tannic phenols was carried out according to Peri and Pompei (1980). Procyanidins and catechins were determined according to Margheri and Falcieri (1972). Metals were determined by atomic absorption spectrophotometry. Prolin was determined, after purification (Adams 1974), according to Pirini *et al* (1992). Gas chromatography was carried out by directly injecting the neutralised sample at pH 7 in a packed column with Carbowax 20M (66 g kg^{-1}) on Carbopack B AW (80-120 mesh).

The testing panel was made up of 20 trained wine assessors (Gerbi *et al* 1990). Preliminary tests were used to devise methods of treating the samples to limit the taste and smell aggressiveness (e.g. dilution with cold or hot water, neutralisation with alkalines), in keeping with the recommendations of Nieto *et al* (1993). However, it was found that such procedures tended to distort aroma or attenuate differences between samples. In order not to excessively tire the assessors, therefore, only six vinegars were examined at each tasting session. Sensory analysis was carried out twice a week, for three consecutive weeks. Each day, 12 vinegars were tasted. Samples were served in normal tasting glasses, and the taste was

evaluated using a glass rod or stainless-steel teaspoon to limit the quantity of vinegar ingested. Sensory attributes in these tests were described previously (Gerbi *et al* 1997), and a 'wheel' card with unstructured scales was used.

After sensory analysis, the panel expressed a degree of liking for the vinegars, with a score between 0 and 100. During tasting sessions, the assessors were only told the acidity of the product, without being given any information on the manufacturer's name or the source (wine, apple, etc) of the vinegars. The cards were read and the evaluations of the panel were then transformed into numeric data using a graphic digitiser and special software (Zeppa and Gerbi 1995). The data were processed using SPSS for Windows (SPSS 1993). The values obtained from the cards were normalised to the maximum score for each assessor.

The ANNs were constructed with NeuroShell 2 (Ward Systems Group, Frederick, MD, USA) software using the analytical and sensory variables as input neurons and the three categories of vinegars as output neurons. The architecture of the networks used was of the three-layer, fully interconnected, feed-forward type. They were constructed with a learning rate of 0·1 and a momentum of 0·1. The ANNs provide for each parameter employed in the network construction a 'contribution factor', which is a number which represent the relative importance of the parameters in predicting the network's output. The higher the number, the more the variable is contributing to the prediction or classification (Ward Systems Group 1993).

To avoid overtraining the networks, NET-PERFECT™ was used, which is a procedure implemented in NeuroShell 2 that creates an entirely separate set of data, called test-set, and uses it to evaluated how well the network is predicting. NET-PERFECT™ was used to compute the optimum point to save the network when it is able to generalise well on new data. Particularly, the learning of the networks was carried out with a limit of 20 000 events after the minimum mean value of re-classification error of the test set was reached.

RESULTS AND DISCUSSION

Table 1 shows the mean values and the standard errors calculated for each analytical parameter for the vinegar categories. It was not always possible to carry out all the desired analytical and sensory tests on all the samples, since the number of samples available was insufficient for some tests. The number of vinegar samples used for the construction of the various ANNs was lower, therefore, than the samples acquired and was based on the analytical parameters used. For simplicity, these parameters were grouped in uniform sets to pinpoint the ones that were most efficient in re-classifying the samples.

Main volatile substances

The 65 samples were randomly subdivided, with about 20% of the total number attributed to the validation set. The training set, therefore, was made up of 51 samples and the validation set of 14. The distribution percentage of the samples between the two data sets was chosen empirically and represents a compromise between the need to have the maximum number of samples in the training set while at the same time representing all the three product categories in the validation set.

Since the use of all the volatile substances determinable by direct gas chromatographic analysis of the samples as input neurons provided a ANN with an average learning level of 87%, variables were chosen on the basis of their *contribution factor* in order to increase this percentage. The process of variable elimination using the lowest contribution factor was repeated several times until a new ANN was generated. This new ANN had an average learning level of 90% using only six variables (acetoin, 2-methyl-1-butanol, 3-methyl-1-butanol, ethyl acetate, iso-butanol and 1-propanol) (see Table 2a). The ANN construction process, from the extract of two data sets through to learning, was repeated five times. Table 2 presents the average learning network generalisation levels, while Table 4 presents the related contribution factor. Besides using six input neurons, the ANN was formed with 14 hidden neurons. Variations in the number of hidden neurons did not influence the network learning level.

All classification errors were related to mistaking wine vinegars for cider vinegars and vice versa, while there was no classification error for alcohol vinegars (Table 3). More specifically, only wine vinegars with 6% acidity, and never those with 7% acidity, were mistaken for cider vinegars, underlining, therefore, that wine vinegars at 7% acidity were well discriminated. This difficulty in classification by ANN confirms the present authors' previous findings using LDA (Antonelli *et al* 1994a), as well as the earlier findings of Nieto *et al* (1993), that it is impossible to classify products purely on the basis of volatile substances despite the high contribution factor of acetoine, except for those that are clearly characterised by their composition, such as alcohol vinegars.

Polyalcohols

The 64 samples examined were sub-divided randomly 80% of the samples attributed to training set (50 products) and 20% to the validation set (14 products).

TABLE 1
Mean values (X) and standard errors (SE) of analytical parameters for vinegar categories

	Wine		*Alcohol*		*Cider*	
	X	*SE*	*X*	*SE*	*X*	*SE*
Density	1·0103	0·0011	1·011	0·0014	1·0136	0·002
Alcohol (% vol)	0·53	1·01	0·44	0·46	0·21	0·13
Total acidity (g 100 ml^{-1})	6·60	0·81	7·35	1·08	5·40	0·45
Volatile acidity (g 100 ml^{-1})	6·19	0·94	6·79	1·85	4·78	0·54
Fixed acidity (g 100 ml^{-1})	0·48	0·61	0·94	1·72	0·85	0·53
Extract (g $litre^{-1}$)	14·21	4·37	3·63	4·01	16·38	4·31
Ash (g $litre^{-1}$)	2·03	0·58	0·33	0·20	2·25	0·42
Ash alcalinity (meq $litre^{-1}$)	17·51	5·82	2·60	2·27	24·80	5·48
Glycerol (g $litre^{-1}$)	3·33	1·29	0·21	0·31	1·67	0·90
Proline (mg $litre^{-1}$)	286	14	8	10	12	10
pH	2·78	0·17	2·36	0·05	3·00	0·10
Tartaric acid (g $litre^{-1}$)	1·53	0·60	0·08	0·14	0·02	0·04
Malic acid (g $litre^{-1}$)	0·36	0·26	0·06	0·13	0·94	1·01
Lactic acid (g $litre^{-1}$)	0·49	0·36	0·02	0·05	0·72	0·32
Citric acid (g $litre^{-1}$)	0·21	0·14	0·31	0·96	0·26	0·29
Total phenols (mg $litre^{-1}$)	197	110	8	11	551	633
Tannic phenols (mg $litre^{-1}$)	96	98	8	11	143	85
Not tannic phenols (mg $litre^{-1}$)	108	71	0	0	408	608
Catechins (mg $litre^{-1}$)	11	11	3	2	44	57
Procyanidins (mg $litre^{-1}$)	64	59	9	7	203	173
OD 420 nm	0·32	0·59	0·21	0·27	0·35	0·15
Lightness (%)	0·82	0·15	0·69	0·18	0·77	0·08
Saturation (%)	14·00	16·00	12·27	18·20	24·79	9·38
Chroma (nm)	578	9	580	2	576	1
Sodium (mg $litre^{-1}$)	69	44	18	9	38	33
Calcium (mg $litre^{-1}$)	115	72	98	49	112	50
Potassium (mg $litre^{-1}$)	682	227	45	41	90	247
Magnesium (mg $litre^{-1}$)	60	24	21	10	47	16
Erythritol (mg $litre^{-1}$)	54	27	1	3	25	11
Xylitol (mg $litre^{-1}$)	4	5	0	0	38	27
Arabitol (mg $litre^{-1}$)	134	130	17	21	159	135
Mannitol (mg $litre^{-1}$)	110	71	33	28	125	93
Sorbitol (mg $litre^{-1}$)	49	52	11	30	3408	1556
scyllo-Inositol (mg $litre^{-1}$)	29	26	1	2	4	4
myo-Inositol (mg $litre^{-1}$)	148	76	11	18	94	63

The ANN used was made up of seven input neurons and 12 hidden neurons, with a particularly high learning level and a 100% re-classification percentage for the validation set (see Table 2). Unlike the results obtained with main volatile substances, the reduction of input neurons in this case only had the effect of lowering the re-classification percentage, giving rise to less than optimal ANNs.

Findings obtained using multivariate analysis (Antonelli *et al* 1994b) were confirmed, ie polyalcohols were found to be necessary for characterising vinegars of various botanical origins, since they maintain the profile of the original raw material even after alcoholic fermentation and acetic bioxydation (Gerbi *et al* 1995). ANNs attributes a different degree of importance to polyalcohols when compared to the results of multivariate analysis. More specifically, the importance that discriminant analysis attributes to mannitol and xilitol is not confirmed, since, generally, ANN places them in fifth and fourth place, respectively, in the discriminant factor scale (Table 4).

Other chemico-physical parameters

The 55 samples available were subdivided randomly, 80% of the cases attributed to the training set (44 products) and 20% to the validation set (11 products). All the chemico-physical parameters not yet used were then used to construct the ANN. The values of these parameters gave rise to a ANN made up of 27 input neurons and 21 hidden neurons with a 100% learning

TABLE 2
Average learning and artificial neural network generalisation levels

			Total number of cases	*Correctly classified cases*	*Wrongly classified cases*
Volatile substances	Training set	Number	51	46	5
		%		90	10
	Validation set	Number	14	12	2
		%		85	15
Polyalcohols	Training set	Number	50	49	1
		%		99	1
	Validation set	Number	14	14	0
		%		100	0
Sensory describers	Training set	Number	50	39	11
		%		78	22
	Validation set	Number	14	11	3
		%		80	20

level. Since the high number of input neurons tends to render practical use difficult, variables were selected on the basis of their contribution factor. The elimination process for variables with low contribution factors was repeated until an ANN with an average learning level of 100% with only three variables (pH, tartaric acid and proline) was obtained. Similar results were previously obtained using LDA (Antonelli *et al* 1993), thereby confirming the importance of tartaric acid and proline, both present only in wine vinegars, for the discrimination of wine vinegars from alcohol and cider vinegars.

TABLE 3
Average distribution on five cycles of re-classification

Actual group	*Total number of cases*	*Predicted group*		
		Wine	*Cider*	*Alcohol*
Wine	49	47	1	1
Cider	13	5	8	0
Alcohol	4	0	0	4

TABLE 4
Relative importance of artificial neural network input (contribution factor)

Volatile substances	Acetoin	9·8
	2-Methyl-1-butanol	9
	3-Methyl-1-butanol	7·1
	Ethyl acetate	6·8
	iso-Butanol	5·5
	1-Propanol	4·2
Polyalcohols	Erythritol	23·1
	Xylitol	12·1
	Arabitol	5·5
	Mannitol	7·7
	Sorbitol	18·9
	scyllo-Inositol	17·3
	myo-Inositol	8·2

Sensory analysis

Of the 64 vinegars examined, 50 were used for the training set and 14 for the test set. The architecture of the ANN was made up of 16 input neurons, corresponding to the sensorial describers used, and 17 hidden neurons. Learning and generalisation levels (Table 2) were very low. Even the elimination of variables with low contribution factors did not increasing the learning level, which, on the contrary, fell even lower when less than 8–10 input neurons were used. The ANN, therefore, presented difficulties in self-configuration and in pinpointing data regularities because of excessive variability. This variability, confirmed by the lack of systematic error re-classification and by the differences in the five cycles of contribution factor simulation (not reported here for reasons of brevity), is caused both by an insufficient or too general training programme of tasters and by the use of cards with sensory describers with insufficient characterisation capacity. Discriminant analysis, used on a greater number of comparison categories (Gerbi *et al* 1994), confirms the ANN results, by underlining, in particular, that taste describers are not very suited to characterising vinegars, probably because of the tasters' discomfort provoked by the high aggressivity of the product.

CONCLUSIONS

The application of ANN to the results of the chemico-physical and sensorial analyses carried out on commercial vinegars underlines the variability of classification capacity of the parameters used and the potential of ANN. In fact, while the volatile substances determined by direct injection in gas chromatography and sensory analysis proved to be of little discriminating value, polyalcohols were shown to be excellent chemical parameters for the discriminant characterisation of vinegars.

The results obtained largely confirm the results of multivariable analysis and, more specifically, LDA, but they are still at the preliminary stage and require further in-depth study. Research must be extended with a greater number of samples so as to increase the learning levels of the various networks in order to improve their generalisation level, and, at the same time, the number of comparison categories must be increased in order to check, for example, the classification of ANNs using samples obtained by mixing, either overtly or covertly, wine and alcohol vinegars.

ACKNOWLEDGEMENTS

Research was carried out with contributions from CNR–RAISA Project–Subproject 4.

REFERENCES

Adams R F 1974 Determination of amino acids profiles in biological samples by gas chromatography. *J Chromatogr* **95** 189–212.

Antonelli A 1994 Ethanol determination by packed glc: a quick method with small sample amount and high sensitivity. *Wine Wissen* **49** 553–555.

Antonelli A, Zeppa G, Gerbi V, Carnacini A, Natali N 1993 Importance of the quality control of vinegar for valorization of typical product. In: *Proceedings of Seventh European Conference on Food Chemistry* (Vol 2). IATA, Valencia, Spain, pp 416–423.

Antonelli A, Zeppa G, Gerbi V, Natali N, Carancini A 1994a Caratterizzazione degli aceti di origine diversa mediante i componenti volatili determinabili per iniezione diretta in gascromatografia. In: *Proceedings of the RAISA Meeting of Sarteano* (*SI*) *26–27 October*. CNR-RAISA, Rome, Italy.

Antonelli A, Zeppa G, Gerbi V, Carnacini A 1994b I polialcoli negli aceti di diversa origine botanica e geografica. In: *Proceedings of the RAISA Meeting of Sarteano* (*SI*) *26–27 October*. CNR-RAISA, Rome, Italy.

Carnacini A, Gerbi V 1992 Le vinaigre de vin: un produit mediterraneen. *Wien Wissen* **47** 216–225.

Carnacini A, Gerbi V, Zeppa G, Antonelli A 1995 Parametri chimico-fisici caratterizzanti gli aceti di vino e loro relazione con il giudizio organolettico. In: *Proceedings of the 2nd Congresso Nazionale di Chimica degli Alimenti*. Società Chimica Italiana, Naxos (ME), Italy, pp 595–604.

Gerbi V, Ubigli M, Zeppa G 1990 Problemi di analisi sensoriale dell'aceto. *Quad Vitic Enol Univ Torino* **14** 79–92.

Gerbi V, Zeppa G, Antonelli A, Carnacini A 1994 Caratterizzazione degli aceti di vino mediante analisi sensoriale. In: *Proceedings of the RAISA Meeting of Sarteano* (*SI*) *26–27 October*. CNR-RAISA, Rome, Italy.

Gerbi V, Zeppa G, Antonelli A, Natali N, Carnacini A 1995 Evoluzione dei costituenti principali del vino e del sidro nel corso dell'acetificazione. *Indust Bevande* **24** 241–246.

Gerbi V, Zeppa G, Antonelli A, Carnacini A 1997 Sensory characterisation of wine vinegars. *Food Qual Pref* **6** (4) 27–34.

Margarita S, Beltramo R, Giomo A 1994 Le reti neurali quale strumento di classificazione: l'esempio del formaggio Montasio. In: *Proceedings of the XVI Congresso Nazionale di Merceologia*. SIME, Pavia, Italy, pp 426–433.

Margheri G, Falcieri E 1972 Importanza dell'evoluzione delle sostanze polifenoliche nei vini rossi di qualità durante l'invecchiamento. *Vin Ital* **81** 501–511.

Mecca F, Vicario G 1971 Determinazione dell'acido acetico non biogenico negli aceti mediante misura della radioattività naturale del radiocarbonio. *Chim Indust* **51** 985–986.

Ministero Agricoltura e Foreste 1965 *Metodi ufficiali di analisi per i mosti, i vini e gli aceti*. Poligrafico dello Stato, Rome, Italy.

Nieto J, Gonzalez-Vinas M A, Barba P, Martin Alvarez P J, Aldave L, Garcia Romero E, Cabezudo M D 1993 Recent progress in wine winegar R&D and some indicators for the future. In: *Food Flavors, Ingredient and Composition*, ed Charalambous G. Elsevier Science, New York, USA, pp 469–500.

Peri C, Pompei C 1980 Fractionament des composés phénoliques present dans le vins. *FV OIV* no 726.

Pirini A, Conte L S, Francioso O, Lercker G 1992 Capillary gas chromatographic determination of free amino acids in honey as a means of discrimination between different botanical sources. *J High Res Chromatogr* **15** 165–170.

Romeo F 1992 Valutazione delle caratteristiche qualitative delle merci. Simulazione di giudizi soggettivi mediante reti neuronali. *Riv Merceol* **31** (1) 5–24.

Romeo F 1993 Simulazione di giudizi soggettivi su oli di oliva mediante reti neuronali. *Riv Merceol* **32** (2) 109–121.

Sakata K, Kawai S, Yagi A, Ina K, Kawamura Y 1991 Carbon-13-NMR spectroscopic analysis of vinegar. *J Jpn Food Sci Technol* **38** (9) 765–769.

Singleton V L, Rossi J A 1965 Colorimetry of total phenolics with phosphomolibdic-phosphotungsting acid reagent. *Am J Enol Vitic* **16** 144–158.

SPSS 1993 *Statistical Package for Social Science* (Vers 5.0.2). SPSS Inc, Chicago, IL, USA.

Ward Systems Group 1993 *NeuroShell 2—User's Manual*. Ward Systems Group Inc, Frederick, MD, USA.

Zeppa G, Gerbi V 1995 Impiego di una tavoletta digitalizzatrice nella lettura di schede non strutturate per l'analisi sensoriale dei vini. *Vignevini* **22** (4) 7–11.