Adv. Hort. Sci., 2012 26(3-4): 187-192

Discrimination of grapevine varieties cultivated in the Czech Republic by Artificial Neural Networks

E. Svobodová*[(1)], C. Pandolfi, P. Hlásná Čepková*, S. Mancuso****

* *Department of Crop Sciences and Agroforestry in Tropics and Subtropics, Faculty of Tropical Agri-Science, Czech University of Life Sciences Prague, Czech Republic.*

** *Dipartimento di Scienze delle Produzioni Agroalimentari e dell'Ambiente, Università degli Studi di Firenze, Viale delle Idee, 30, 50019 Sesto Fiorentino (FI), Italy.*

Key words: ampelography, phyllometry, *Vitis vinifera*, variety identification.

Abstract: An artificial neural network approach, based on fractal leaf parameters, and classical ampelography were used to identify nine grapevine varieties cultivated at the St. Claire's vineyard, Prague Botanic Garden. Fifty healthy, fully-expanded leaves were collected for each variety, scanned using an optical scanner and then elaborated by computer programs. Fourteen phyllometric parameters were qualitatively and quantitatively analysed by the digital image analysis. Comparative frames were constructed for each variety and the relationships among varieties were assessed using artificial neural networks. Results were then compared with the outcome from traditional ampelographic analysis. The Artificial Neural Network technique appears to be a complementary approach to the traditional ampelography methods commonly used for cultivar discrimination, since the equipment necessary for this analysis is very inexpensive and available. Application of the technique led to the distinction of nine selected varieties of *Vitis vinifera*.

1. Introduction

Ampelography is a traditional morphological method used for the identification and discrimination of varieties of grapevine (*Vitis vinifera* L.). It has also been found to be very useful in detecting and describing variability among chosen potential clone candidates (Poljuha *et al.*, 2006). Nowadays, a link between historic descriptions and the molecular fingerprints is also very important, especially in cases of possible homonymy and synonymy of autochthonous varieties (Cervera *et al.*, 2001). Official descriptors have been published to be used together with the analysis of germplasm material (Dettweiler, 1993; Ortiz *et al.*, 2004).

All the organs used in ampelography (leaves, grapes, shoots etc.) usually change their aspect during phenologic phases, hence it is important to find the characteristics able to discriminate between varieties according to these phases: shoots in the early stage, leaves in the moment when they are mature, and later also the grapes and mature berries (Cancellier, 2007).

The use of fractal-based measurements of digitally-acquired images eliminates problems with different phenologic phases as well as problems with subjectivity. This allows defining the good shape measure that can be effectively applied to leaf shapes, so they can be compared and analysed by meaningful and objective criteria (Mancuso, 1999). Using fractal parameters and phyllometric outputs, an artificial neural network can be constructed and effectively used to differentiate varieties and accessions. It is an easy method that requires low-cost equipment such as an optical scanner, personal computer and free software (Mugnai *et al.*, 2008).

St. Claire's Vineyard is an historical area located close to the centre of Prague, and where a long tradition of grapevine cultivation dates back to the 13th century. Following a period of decline, the vineyard is today regaining its former importance. New varieties are being planted and no ampelographic observations have yet been carried out.

The main aim of the present work was to discriminate the grapevine varieties cultivated under the climatic conditions of Prague (Czech Republic) using either the subjective method according to the international descriptors published by IPGRI (1997), or an objective computing method which consists of analysis using an artificial neural network. These two approaches are also compared in this paper to assess differences between subjective and objective analysis.

(1) Corresponding author: evasvobodova@its.czu.cz.

Received for publication 3 September 2012
Accepted for publication 6 March 2013

2. Materials and Methods

Plant material

Leaf samples for the analysis of nine varieties were collected at the St. Claire's Vineyard, Prague Botanic Garden. Nine varieties (Table 1), cultivated in a sufficient quantity to provide a sufficient source of samples, were chosen for analysis. Berry and grape samples were collected at the time of the harvest, which was done according to the further utilization of the grapes.

Table 1 - List of selected varieties

	Variety
1	'Müller Thurgau'
2	'Gewürztraminer'
3	'Rhine Riesling'
4	'Italian Riesling'
5	'Moravian Muscat'
6	'Sauvignon'
7	'Blue Portugal'
8	'White Chasselas'
9	'Red Chasselas'

Morphological and phenological characterization

For the classical ampelography, the following method was applied. In the period 2006-2008, 22 morphological and six phenological traits were evaluated in 10 randomly chosen plants for each variety, according to the Descriptors for Grapevine (IPRGI, 1997). All parameters chosen for the study are listed in Table 2 with the codes of updated descriptors (OIV, 2009).

Digital image analysis of leaves, phyllometric parameters and fractal analysis

For each variety, 50 healthy, fully-expanded leaves were collected from ten randomly selected plants in late spring 2008. Leaf images were acquired (200 dpi, 256 greyscale) using an optical scanner. Based on the method described by Mancuso and Nicese (1999), 14 phyllometric parameters (Table 3) were determined for each leaf using image-analysis software (UTHSCSA Image Tool Program 3.0).

Table 2 - List of characteristics selected for the description of observed cultivars. The codes are in line with OIV guidelines (OIV, 2009)

Character code	Description
001	Young shoot: aperture of tip
003	Young shoot: intensity of anthocyanin colouration on prostrate hairs of tip
006	Shoot: attitude
007	Shoot: colour of dorsal side of internode
008	Shoot: colour of ventral side of internode
051	Young leaf: colour of the upper side of blade (4th leaf)
065	Mature leaf: size of blade
067	Mature leaf: shape of blade
068	Mature leaf: number of lobes
070	Mature leaf: area of anthocyanin colouration of main veins on the upper side of blade
076	Mature leaf: shape of teeth
079	Mature leaf: degree of opening/ overlapping of petiole sinus
102	Woody shoot: structure of surface
103	Woody shoot: main colour
202	Bunch: length (peduncle excluded)
204	Bunch: density
220	Berry: length
223	Berry: shape
241	Berry: formation of seeds
225	Berry: colour of skin
236	Berry: particularity of flavour
303	Time of beginning of berry ripening (véraison)
233	Berry: must yield
502	Bunch: weight of a single bunch
503	Single berry weight
505	Sugar content of must

The codes in bold letters indicate the characteristics which were evaluated by ANOVA.

Table 3 - Fourteen phyllometric parameters measured by image analysis software

	Parameter	Definition
1	Area	The area of the leaf
2	Perimeter	The perimeter of the leaf
3	Major axis length	The length of the longest line that can be drawn through the leaf
4	Minor axis length	The length of the longest line that can be drawn through the leaf perpendicular to the major axis
5	Roundness	Computed as: $(4*\pi*\text{area}) / \text{perimeter}^2$
6	Elongation	The ratio of the length of the major axis to the length of the minor axis
7	Feret diameter	The diameter of a circle having the same area as the leaf
8	Compactness	Computed as: $\text{sqrt}(4*\text{area}/\pi)$ / major axis length
9	Integrated density	Computed as the product of the mean grey level and the number of pixels in the image of the leaf
10	Minimum grey level	Minimum grey level of the leaf
11	Mean grey level	Mean grey level of the leaf
12	Median grey level	Median grey level of the leaf
13	Mode grey level	Mode grey level of the leaf
14	Maximum grey level	Maximum grey level of the leaf

The fractal spectrum of the leaves was obtained using fractal image analysis software (HarFA, Harmonic and Fractal Image Analyzer 4.9.1,), according to the method described by Mancuso (2002). Briefly, greyscale image of each leaf was thresholded for a grey value between 0 and 255 and the fractal dimension for each grey value was then assessed using the box counting method. The implementation of these methods has been described in detail by Mancuso *et al*. (1999). After drawing the baseline (fractal dimension = 1) which separates the fractal (>1) from the non-fractal (<1) zone of the spectrum, five fractal parameters (First X, Peak X, Last X, Peak Y and Total Peak Area) were calculated (Fig.1 A).

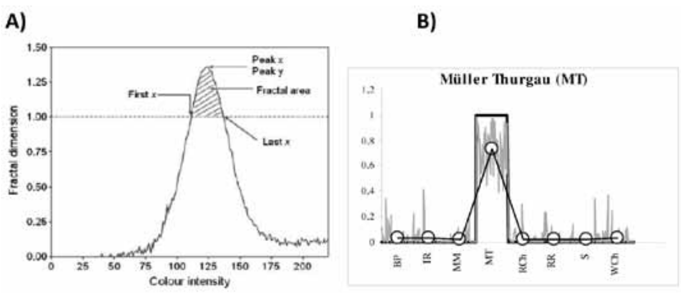

Fig. 1 - Fractal spectrum of one leaf and identification of five parameters to be implemented in the ANN. Graphical representation of Fractal parameters: Theoretical situation (A); and a real value (B).

Artificial Neural Network analysis

An Artificial Neural Network (ANN) was constructed as previously described in Pandolfi *et al*. (2009 a). Fourteen phyllometric parameters and five fractal parameters were used as input layers, and the nine grapevine accessions represented the output. To optimize the neural network activity, the number of hidden neurons and the number of iterations was modified. With regard to the hidden layer, many factors (such as learning scheme, numbers of nodes of the output and input and connections between them) play an important role for the determination of the best configuration (Zurada and Malinowsji, 1994). The ANN outputs are represented by a XY-graph for each accession, with the accession names on the x-axis, and the y-axis representing the output. Each graph aims to show how the ANN was able to discriminate the selected accession in comparison with the others. The level of similarity is expressed by number, which ranges between 0 (false) and 1 (true) (Fig. 1 B). Due to the natural variability among leaves, the output of the expected class tends to report a value close to 1, but less than 1, while the others should be close to 0 (Mugnai *et al*., 2008).

Statistical analysis

Statistical analysis of the selected characteristics was performed by analysis of variance (ANOVA) and a comparision was done using the Turkey HSD Test (program Statistica 7.0 CZ).

The graphical presentation of variability of the tested varieties was carried out using Principal Component Analysis (PCA) (software Statistica 7.0 CZ).

NTSYS 2.1 was used to investigate neural network outputs performing a cluster analysis by Unweighted Pair Group Method Analysis (UPGMA) based on the similarity matrix calculated using the cosine function (Eq. 1).

Equation 1:

$$COSINE_{(x,y)} = \frac{\sum_i (x_i y_i)}{\sqrt{\left(\sum_i x_i^2\right)\left(\sum_i y_i^2\right)}}$$

3. Results

Morphological data

Seven characteristics (Table 2), as evaluated by ANOVA, showed differences between the varieties during the three consecutive years of sampling, as well as variability within the variety. The number of lobes observed in the mature leaves varied during the study period in three varieties ('Müller Thurgau', 'Rhine Riesling', and 'Blue Portugal').

A significant variation was observed both in the weigh of a bunch and the weight of berries, as well as in the size of berries within one variety.

With regard to must yield and its sugar content, there was no significant variance among the varieties, however all the varieties demonstrated a significant correlation between the sugar content of the must and the time of berry ripening (r=0.74). The highest sugar content was measured in 2006, when the mean temperatures remained quite high during whole vegetative period (in July as high as 25°C), whilst the total precipitation was about 5 mm in July.

For most of the 19 characteristics that were not included in the ANOVA analysis, there was considerable stability over the course of the years. The greatest variability was shown in the size of a blade of a mature leaf; it remained stable only in 'Rhine Riesling'. In the other eight varieties, the size of the leaves varied significantly.

Correlations between morphological and phenological characters were also assessed. The output obtained from both morphologic and phenological characteristics after three years of observations (Fig. 2) was a score plot cre-

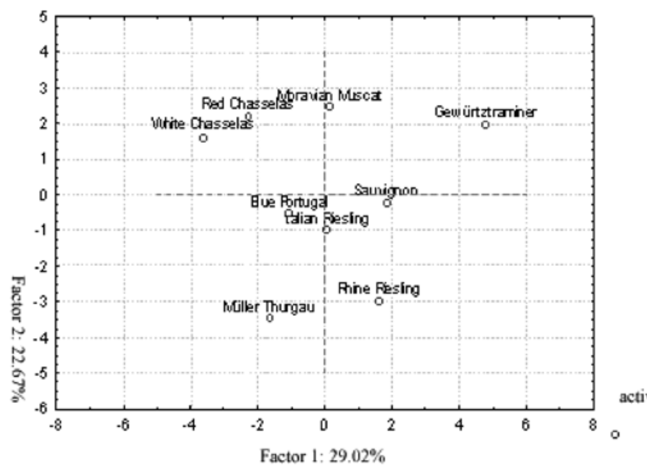

Fig. 2 - PCA representation.

ated by PCA analysis. According to the graph, factor 1 explained 29.02% of the total variation, whereas factor 2 expressed 22.67% of the total. Evaluating both approaches (morphological and phenological), ‘Red Chasselas’ and ‘White Chasselas’ were identified as being quite close, which indicated their relationship to one another. The second group was formed in the very centre of the graph and includes ‘Blue Portugal’, ‘Italian Riesling’ and ‘Sauvignon’. More distant from the centre group and also from each other were ‘Müller Thurgau’ and ‘Rhine Riesling’. The only variety which remained completely separate was ‘Gewürztraminer’.

Neural network analysis

In all eight cases, the artificial neural network was able to recognise all the accessions presented in the learning phase. Throughout there was a clearly defined peak for the selected accession; therefore the varieties were well-separated one from another. The highest level of similarity was present in ‘Müller Thurgau’ (0.810). However, all other accessions showed quite high average output in a range from 0.541 (‘Rhine Riesling’) to 0.657 (‘Moravian Muscat’) (Fig. 3).

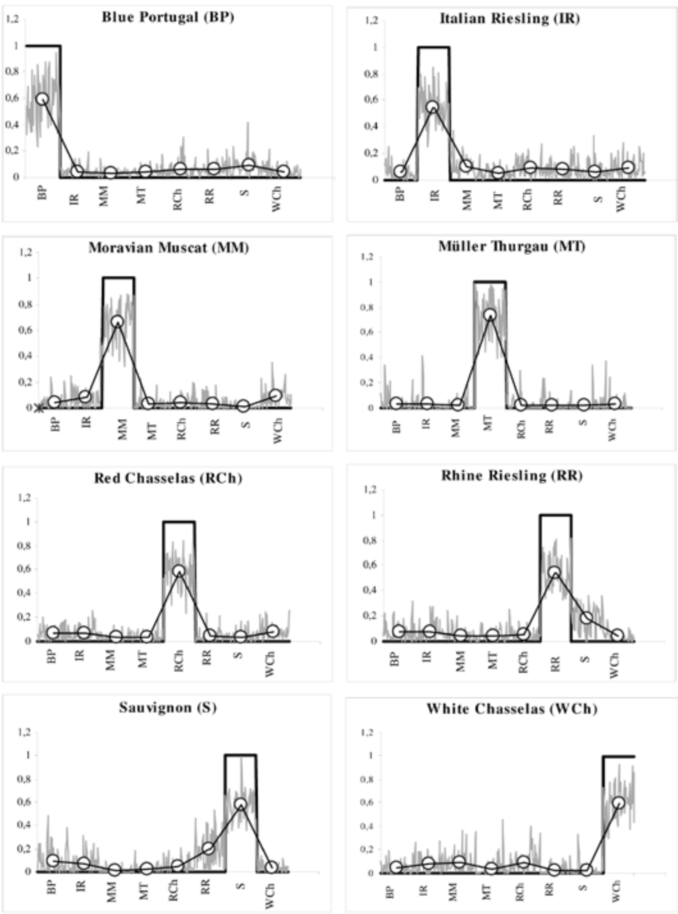

Fig. 3 - Neural networks of eight chosen varieties.

‘Müller Thurgau’ also showed the lowest degree of similarity with the other varieties and this was confirmed by the UPGMA diagram, where ‘Müller Thurgau’ remained completely separated from all other varieties.

Another interesting result was the position of both ‘Chasselas’ in the neural network diagram and the final dendrogram. In the dendrogram, ‘Red Chasselas’ was more closely related to ‘Italian Riesling’ than to ‘White Chasselas’.

Using the data from the neural network, Euclidean distances were calculated and a dendrogram based on the distance matrix data, by applying an Unweighted Pair Group Method with Average Mean (UPGMA) cluster analysis, was constructed (Fig. 4).

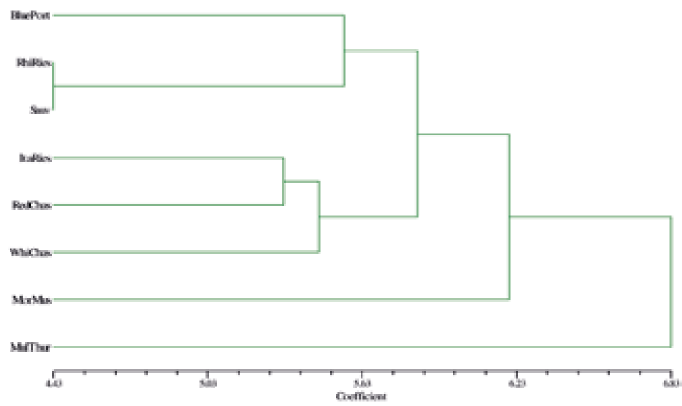

Fig. 4 - UPGMA dendrogram.

The dendrogram formed two clusters, and left two varieties more (‘Müller Thurgau’) or less (‘Moravian Muscat’) separated from the others. Two other varieties, ‘Rhine Riesling’ and ‘Sauvignon’, were indicated at the same level. Even though they were the two most related samples, they still remained quite distant, with the similarity coefficient at 4.43. The results showed a closer relationship between ‘Italian Riesling’ and ‘Red Chasselas’ than the relationship between both ‘Chasselas’, even though they are considered to be colour variations of the same variety.

4. Discussion and Conclusions

In the present investigation all general characteristics which are common to *Vitis vinifera*, without regarding the differences between the varieties (IPGRI, 1997), were confirmed with regard to morphology for the varieties under study.

There was only a slight variation in the number of lobes observed in the mature leaves, which is due to the genotype fixation of this trait, as proposed by Poljuha *et al.* (2006). In three varieties (‘Müller Thurgau’, ‘Rhine Riesling’ and ‘Blue Portugal’) the number of lobes varied during the observation period. This way have occurred due to the unisindentification of the depth of the status. (Cancellier, 2007).

The size of a berry was strongly correlated either with the weight of a berry, or with the weight of a bunch. The smallest berries were observed in an important must variety

‘Rhine Riesling’. This result corresponded with the findings of Fregoni (2005), who indicated that varieties which produce wines of excellent quality generally have smaller berries with higher amount of skin and lower amount of pulp, in comparision with the table varieties. Ojeda *et al.* (2001) pointed out that the size of a berry is an environment-dependent trait. Water deficit that occurs in the early phase from flowering to véraison causes an irreversible decrease in cell volume. Due to changes in water import by xylem, this may lead to a decrease of mesocarp cell turgor (Thomas *et al*., 2006). Under water stress, especially in the early growth stages, the division and elongation of cells decreases (Williams, 2000), thus berries remain smaller. In 2008, when precipitation events were regularly distributed throughout the year, the berries of all nine varieties were larger compared to other observation years.

Sugar content, and thus the quality of the wine, is significantly influenced by the weather, as proven by Grifoni *et al.* (2006) from observations of Italian wines. A surprisingly low sugar content was found in ‘Moravian Muscat’, despite it being a variety bred in former Czechoslovakia (Pavloušek 1999), and thus adapted to its climatic conditions. However, in this case the sugar content was influenced by the early harvest date, since the grapes were used for the production of federweisser (a freshly fermented must with low alcohol content).

The must yield was also strongly correlated to the size of the berry. According to Fregoni (2005) table varieties have lower must yields than must varieties. However, ‘White Chasselas’, despite being a table variety, gave the highest must yield during the three study years, while the must varieties, like for instance ‘Blue Portugal’, produced significantly lower quantities the must (57.00±6.25 ml). This may be the influence of rains right before harvest.

In all eight cases, the artificial neural network was able to recognise all the accessions presented in the learning phase, as published for example by Mancuso (2002) or Mugnai *et al*. (2008). An interesting result was the position of ‘Müller Thurgau’, which remained separated from all other accessions. Dettweiller *et al.* (2000) confirmed that ‘Müller Thurgau’ is the descendent of a crossing between ‘Madeleine Royal’ and ‘Rhine Riesling’. Nevertheless, in our results ‘Rhine Riesling’ remained distant from ‘Müller Thurgau’, and was assessed as close to ‘Sauvignon’, corresponding with the data published by Moravcovà *et al*. (2004). As reported in their work (Moravcová *et al*., 2004), ‘Red Chasselas’ is only a mutation of berry colour of ‘White Chasselas’. In the dendrogram, however, ‘Red Chasselas’ was more closely related to ‘Italian Riesling’ than to ‘White Chasselas’.

Comparing the dendrogram with the PCA score plot, the results were slightly different. Both ‘Chasselas’ varieties were the closest, which corresponded with the results of Moravcovà *et al*. (2004). But also in this case, there was no evident relationship between ‘Rhine Riesling’ and ‘Müller Thurgau’, since they occurred in different quadrants.

Unfortunately, the position of ‘Gewürztraminer’ in the PCA graph could not be compared with the ANN output. However, the score plot put this variety completely separate from other varieties, e.g. ‘Rhine Riesling’, as already described by Imazio *et al*. (2002). The great distance of ‘Gewürztraminer’ from other varieties might also be due to its closer relationship to *Vitis silvestris* (Regner *et al*., 2000; Lacombe *et al*., 2003).

When using the ANN approach it is very important to choose well-developed, healthy leaves (Mugnai *et al*., 2008). An ANN needs a suitable set of sample data to achieve excellent quality in the training set, as this operation can be usually defined as the key point of all the ANN building process (Pandolfi *et al*., 2009 a). Unfortunately, the Euclidean distance did not permit a clear clustering and thus in our case a PCA graph could be considered a more precise tool, even though the data were mostly obtained from subjective observations. However, the possibility of a primary misidentification while still in the vineyard must always be considered: in early spring, when the plants do not have mature leaves or even grapes, an identification error may occur, especially when the varieties are not clearly divided.

Both approaches - ampelography or artificial neural networks – are important methods based on morphological traits, which complement each other to give clear classification of the varieties. The ANN technique represents an more economic alternative to genetic methods commonly used for cultivar discrimination since the equipment necessary for this analysis is inexpensive and commonly available (Pandolfi *et al*., 2009 a).

In this study, application of ANNs led to the distinction of nine selected varieties of *Vitis vinifera*. For the further uses, the ANN profiles obtained from our analysis may be kept in databases for breeding programs, as well as for the description of new cultivars. The data obtained might also be helpful for rapid and reliable identification and description of still unknown varieties, as shown by Pandolfi *et al.* (2009 b).

Acknowledgements

The authors would like to thank St. Clair’s Vineyard (Prague Botanic Garden), especially Martin Beránek, Ph.D., for his support and assistance during the experiment. We would also like to acknowledge an unknown reviewer for his/her kind and useful suggestions.

References

CANCELLIER S., 2007 - *Proposta di una guida per il riconoscimento dei vitigni. Esempio: i vitigni di Veneto.* - Rivista di Viticoltura e di Enologia, 60(1): 3-16.

CERVERA M.T., RODRIGUEZ I., CABEZAS J.A., CHAVEZ J., MARTÍNEZ-ZAPATER J.M., CABELLO F., 2001 - *Morphological and molecular characterization of grapevine accessions known as ‘Albillo’.* - Amer. J. of Enol. and Vitic., 52(2): 127-135.

DETTWEILER E., 1993 - *Evaluation of breeding characteristics in Vitis*.- Vitis, 32: 249-253.

DETTWEILER E., JUNG A., ZYPRIAN E., TÖPFER R., 2000 - *Grapevine variety 'Müller-Thurgau' and its true-to-type descent*. - Vitis, 39(2): 63-65.

FREGONI M., 2005 - *Viticoltura di qualità*. - 2nd edition, Phytoline s.r.l., Affi, pp. 819.

GRIFONI D., MANCINI M., MARACCHI G., ORLANDINI S., ZIPOLI G., 2006 - *Analysis of Italian Wine Quality Using Freely Available Meteorological Information*. - Amer. J. of Enology and Viticulture, 57(3): 339-346.

IMAZIO S., LABRA M., GRASSI F., WINFIELD M., BARDINI M., SCIENZA A., 2002 - *Molecular tools for clone identification: the case study of the grapevine variety 'Traminer'*. - Plant Breeding, 121: 531-535.

IPGRI, 1997 - *Descriptors for Grapevine* (Vitis *spp*.). - International Union for the Protection of New Varieties of Plants, Geneva, Switzerland; Office International de la Vigne et du Vin, Paris, France and International Plant Genetic Resources Institute, Rome, Italy.

LACOMBE T., LAUCOU V., DI VECCHI M., BORDENAVE L., BOURSE T., SIRET R., DAVID J., BOURSIQUOT J.M., BRONNER A., MERDONOGLU D., THIS P., 2003 - *Contribution à la caractérisation et à la protection in situ des populations de* Vitis vinifera *L. ssp. Silvestris (Gmelin) Hegi, en France*. - Les Actes du BRG, 4: 383-404.

MANCUSO S., 1999 - *Fractal geometry-based image analysis of grapevine leaves using the box-counting algorithm*. - Vitis, 38(3): 97-100.

MANCUSO S., 2002 - *Discrimination of grapevine (*Vitis vinifera *L.) leaf shape by fractal spectrum*. - Vitis, 41(3): 137-142.

MANCUSO S., NICESE F.P., 1999 - *Identifying olive (*Olea europaea*) varieties using artificial neural networks*. - J. Amer. Soc. Hort. Sci., 124(5): 527-531.

MORAVCOVÁ K., BARÁNEK M., PIDRA M., 2004 - *The use of RAPD markers for differentiation of grapevine varieties registered in the Czech Republic*. - Horticultural Science (Prague), 31 (3): 96-101.

MUGNAI S., PANDOLFI C., AZZARELLO E., MASI E., MANCUSO S., 2008 - Camelia japonica *L. genotypes identified by an artificial neural network based on phyllometric and fractal parameters.* - Plant Systematics and Evolution, 270: 95-108.

OIV, 2009 - *The 2nd edition of the OIV descriptors list for grape varieties and* Vitis *species*. - Office International de la Vigne et du Vin, Paris, France.

OJEDA H., DELOIRE A., CARBONNEAU A., 2001 - *Influence of water deficits on grape berry growth*. - Vitis, 40(3): 141-145.

ORTIZ J.M., MARTÍN J.P., BORREGO J., CHÁVEZ J., RODRÍGUEZ I., MUÑOZ G., CABELLO F., 2004 - *Molecular and morphological characterization of a Vitis gene bank for the establishment of a base collection*. - Genetic Resources and Crop Evolution, 51: 403-409.

PANDOLFI C., MESSINA G., MUGNAI S., AZZARELLO E., MASI E., DIXON K, MANCUSO S., 2009 b - *Discrimination and identification of morphotypes of* Banksia integrifolia *(Proteaceae) by an Artificial Neural Network (ANN), based on morphological and fractal parameters of leaves and flowers.* - Taxon, 58(3): 925-933.

PANDOLFI C., MUGNAI S., AZZARELLO E., BERGAMASCO S., MASI E., MANCUSO S., 2009 a - *Artificial neural networks as a tool for plant identification: a case study on Vietnamese tea accessions*. - Euphytica, 166: 411-421.

PAVLOUŠEK P., 1999 - *Vinohradnictví-odrůdy révy vinné*. 1. vydání. - Mendelova lesnická a zemědělská univerzita v Brně, Brno, pp. 122.

POLJUHA D., SLADONJA B., PERŠURIĆ D., 2006 - *Evaluation of intravarietal variability of cv. 'Istrian Malvasia' by the ampelographic parameters*. - Rivista di Viticoltura e di Enologia, 2(3): 13-18.

REGNER F., STADLBAUER A., EISENHELD C., KASERER H., 2000 - *Considerations about the evolution of grapevine and the role of traminer*. - Acta Horticulturae, 528: 179-184.

THOMAS T.R., MATTHEWS M.A., SHACKEL K.A., 2006 - *Direct in situ measurement of cell turgor in grape (*Vitis vinifera *L.) berries during development and in response to plant water deficits*. - Plant, Cell and Environment, 29: 993-1001.

WILLIAMS L.E., 2000 - *Grapevine water relations*, pp. 121-126. - In: CHRISTENSEN L.P. (ed.), *Raisin production manual*. University of California, Division of Agriculture and Natural Resources, Publication 3393, Oakland, CA.

ZURADA J.M., MALINOWSJI A., 1994 - *Multilayer perceptron networks: selected aspects of training optimization*. - International Journal of Applied Mathematics and Computer Science, 4: 281-307.