PREDICTION OF WINE SENSORIAL QUALITY BY ROUTINELY MEASURED CHEMICAL PROPERTIES

ADRIÁNA BEDNÁROVÁ[1], ROMAN KRANVOGL[2],
DARINKA BRODNJAK-VONČINA[2], TJAŠA JUG[3]

[1]*Department of Chemistry, Faculty of Natural Sciences, University of SS Cyril and Methodius in Trnava, Nám. J. Herdu 2, Trnava, SK-917 01, Slovak Republic (adriana.bednarova@ucm.sk)*

[2] *Faculty of Chemistry and Chemical Engineering, University of Maribor, Smetanova 17, 2000 Maribor, Slovenia*

[3] *Chamber of Agriculture and Forestry of Slovenia, Institute for Agriculture and Forestry, Pri hrastu 18, 5000 Nova Gorica, Slovenia*

Abstract: The determination of the sensorial quality of wines is of great interest for wine consumers and producers since it declares the quality in most of the cases. The sensorial assays carried out by a group of experts are time-consuming and expensive especially when dealing with large batches of wines. Therefore, an attempt was made to assess the possibility of estimating the wine sensorial quality with using routinely measured chemical descriptors as predictors. For this purpose, 131 Slovenian red wine samples of different varieties and years of production were analysed and correlation and principal component analysis were applied to find inter-relations between the studied oenological descriptors. The method of artificial neural networks (ANNs) was utilised as the prediction tool for estimating overall sensorial quality of red wines. Each model was rigorously validated and sensitivity analysis was applied as a method for selecting the most important predictors. Consequently, acceptable results were obtained, when data representing only one year of production were included in the analysis. In this case, the coefficient of determination (R^2) associated with training data was 0.95 and that for validation data was 0.90. When estimating sensorial quality in categorical form, 94 % and 85 % of correctly classified samples were achieved for training and validation subset, respectively.

Key words: overall sensorial quality, prediction, Slovenian wine, artificial neural networks, multivariate data analysis

1. Introduction

Wine is a complex mixture of organic as well as inorganic compounds determining the final organoleptic properties of wine. The factors influencing the quality of wine are related to winemaking environment including ground, climate and the variety of the employed oenological practice (POHL, 2007). There are numerous studies showing the connection of chemical compounds, especially phenolic and volatile as well as some non-volatile compounds with human perception (SÁENZ-NAVAJAS *et al.,* 2012; LORRAIN *et al.,* 2013). Sensory analysis is the most common method of assessing the biological characteristics of foodstuffs and therefore its quality. Typical tests to estimate wine quality rely on sensorial assays carried out by a group of experts. Although these methods are probably the best option for evaluating a small number of samples, they are time-consuming and expensive since an extensive training of panellists is necessary for reproducible results and even the trained panel can perform only a limited number of analyses per day due to fatigue or environmental

DE GRUYTER OPEN

interferences. Successful application of conventional analytical techniques to the evaluation of sensory properties of foodstuffs is very difficult, primarily because the relationship between chemical composition and flavour is not properly known and also due to high complexity and variability of food composition (LEGIN *et al.,* 2003). Though, it was reported that the global signal provided by sophisticated analytical methods can be considered the fingerprint of sample flavour and used as the analytical signal for characterising the commodity, especially, when chemometrics is used to obtain the most detailed information from the sample (LÓPEZ-FERIA *et al*., 2008). In addition, using equipment such as electronic nose and electronic tongue connected with appropriate predictive modelling techniques could improve the performance of estimating sensory properties (LEGIN *et al.,* 2003; BURATTI *et al*., 2007). Methods of multivariate data analysis (MVA) are frequently used for investigation of relations and interactions inside a large data table. For wine classification the MVA methods are especially useful when different attributes of wine authenticity are predicted utilising various types of analytes (ARVANITOYANNIS *et al*., 1999; COZZOLINO *et al.,* 2009; SAURINA, 2010; KRUZLICOVA *et al*., 2013). However, since institutions controlling wine quality or wineries do not commonly possess modern analytical equipment, it would be interesting to assess the feasibility of estimating the sensorial quality of wine only by simple chemical properties such as alcoholic grade, density, pH, content of extract or SO_2, that are routinely measured at the mentioned institutions or wineries.

Thus, the goal of present study was to examine the possibility of predicting the overall sensorial quality of Slovenian red wines based on the results of routinely chemical analyses performed in the control institution. For this purpose, artificial neural networks (ANNs), specifically multilayer perceptron, was utilised. In addition, the interrelations among the studied oenological descriptors were discovered in detail.

2. Materials and methods

2.1 Wine samples and analytical procedures

Altogether 131 red wine samples originated in Primorska winegrowing region in Slovenia were analysed during five years of production (2002 – 2006). More specifically, the analysed wine samples represented varieties such as Cabernet Sauvignon (51 samples), Merlot (46 samples), and a blend of several red wines varieties (34 samples). Sampling was made by the Chamber of Agriculture and Forestry of Slovenia in Nova Gorica. The following wine descriptors were determined: relative density at 20 °C, contents of total extract (g/dm^3), reducing sugars (g/dm^3), ash (g/dm^3), free SO_2 (mg/dm^3), ethanol (%), total acidity (g/dm^3), volatile acidity (g/dm^3), and pH. In addition, the sensorial quality was obtained by group of experts evaluating the wine properties (colour, aroma, taste and harmony) using a twenty point scale in total. All analytical methods were accomplished according to Official Gazette Republic of Slovenia No. 43/01, and the sensorial analysis was performed according to Official Gazette Republic of Slovenia No. 32/00. Sensory evaluation was performed in a group of 5 experts, representatives of consumers, producers and wine experts. They

were allowed to evaluate up to 40 samples per day. In every evaluation they give up to 2 points for clarity, 2 for colour, up to 4 for odour, up to 6 for taste and 6 for harmony. The WineScan FT 120 instrument utilising FTIR was employed for simultaneous determination of mentioned wine descriptors. The analytical procedures are described in detail in (BEDNÁROVÁ *et al.,* 2013).

2.2 Statistical analyses

In addition to the mentioned oenological characteristics, i.e. relative density (denoted as *Density*), total extract (*Extract*), reducing sugars (*RedSug*), ash (*Ash*), free SO$_2$ (*SO2Free*), ethanol (*Ethanol*), total acidity (*TotAcid*), volatile acidity (*VolAc*), and pH (*pH*), further descriptors, namely non-volatile acidity (g/dm^3) and reduced extract (g/dm^3) were calculated as follows:

$$\text{non-volatile acidity } (NVolAc) = TotAcid - VolAc$$

$$\text{reduced extract } (SFE) = Extract - (RedSug - 1)$$

Furthermore, ratios *VolAc*/*TotAcid* (*VA_TA*) and *NVolAc*/*TotAcid* (*NVA_TA*) were added to the list of analysed variables. After performing exploratory analysis of the available data, i.e. checking normality and presence of outliers, the interrelations among all oenological descriptors were discovered by correlation analysis (CA) and principal component analysis (PCA). However, the main regard was taken on the significant correlations of the target variable – overall sensorial quality (denoted as *Quality*) with the studied oenological descriptors. Secondly, the target variable *Quality* was transformed into categorical form and nonparametric tests were applied to perform detailed insight to the connection of sensorial quality with wine chemical descriptors.

The main objective of this work was to discover the possibility of estimating the wine sensorial quality by routinely measured chemical properties. For this purpose, two approaches were used: (1) classification of samples into three created groups of *Quality_cat* by artificial neural networks (ANNs) and linear discriminant analysis (LDA) and (2) prediction of the continual target variable *Quality* by ANNs. In all model building procedures, a method for input variables selection was performed as well as proper validation of the created model. To assess the predictive ability (predictivity) of the model, i.e. the measure of how well the model can predict values of new data, the data not utilised in model calculations were employed. External validation, when performed judiciously, is generally regarded as the most rigorous assessment of predictivity, since predictions are made for samples not used in the model development in comparison to e.g. cross-validation methods.

The results of the predictions were evaluated by the coefficient of determination (R^2) between the actual and predicted values of *Quality* and root mean square error (*RMSE*) as a measure of the difference between values predicted by a model and the values actually observed. The *RMSE* of a model prediction with respect to the estimated variable X_{model} is defined as the square root of the mean squared error:

$$RMSE = \sqrt{\frac{\sum (X_{obs,i} - X_{model,i})^2}{n}}$$

where X_{obs} is set of observed values and X_{model} is set of modelled values at time/place i. Regarding the classification of wine samples according to the *Quality_cat*, the success of classification was evaluated by the percentage of correctly classified objects into the categories.

2.3 ANN model optimisation

Learning the input/output relationship through the training process is the key feature of artificial neural networks (ZUPAN and GASTEIGER, 1993; BISHOP, 1995; HAYKIN, 1999). A prescribed set of well-defined rules for the solution of a learning problem is called a learning algorithm. In our case, the BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm was applied as one of the most recommended techniques in neural networks. ANN models have been built by employing the Automated Network Search included in the STATISTICA software (Statsoft, Tulsa, USA) where different network architectures were automatically tested and the best alternatives were determined. To verify the generalisation (prediction) of a model, the validation data (subset of samples not used in model calculation) were applied to demonstrate whether the input-output relationship computed by network utilising training data was correct. The complexity of the model in terms of the selection of the input (predictor) variables was examined with the aid of so-called sensitivity analysis indicating the importance of input variables by particular neural network (HUNTER *et al.,* 2000). Thus, the results of sensitivity analysis were used for informative purposes, but also to perform input pruning. Also, the number of neurons in the hidden layer was optimised due to avoiding the over-training the network. Therefore, a third subset of data was selected for computing prediction error periodically during training to avoid over-training. Consequently, together three subsets of data were created by random dividing the whole data set into three parts: (1) training subset, (2) test subset used to determine if over-training has occurred, and (3) external validation subset for evaluation of the predictivity of the model.

3. Results and discussion

3.1. Correlation analysis

The direct inter-relations between the studied descriptors and the target overall sensorial quality were analysed by the correlation analysis. Moreover, the information about the correlations among all descriptors was important in the further model building part, especially in the step of selection of input variables subset. All data with exception of two outliers were included ($n = 129$) in the correlation analysis, containing all varieties and all vintages. Since departures from normal distribution were identified by Shapiro-Wilk test ($p < 0.05$) in case of several studied variables, the Spearman correlation analysis was employed as a non-parametric alternative to Pearson correlation (MILLER and MILLER, 2010). Another detail of importance - due to a large size of data (129 samples) also relatively small values of the correlation coefficient meant a statistically significant correlation. For the sake of differentiation, a blue colour is used for the highest significant correlations ($p < 10^{-6}$) in Table 1. The

less important correlations are marked in black. As expected, the target variable *Quality* correlated highly to *Ethanol* ($r_S = 0.42$) as shown in Table 1. In addition, positive correlations of *Quality* with *SFE*, *Extract*, *pH* as well as *VolAc* and negative correlations of the target variable with *TotAcid* and *NVolAc* were found. This means, that high content of non-volatile acids could affect decreasing the overall sensorial quality of wine, which is in agreement with oenological practice: wines with higher acidity are usually rougher. On the other hand, great level of volatile acids and especially of ethanol indicates high wine quality.

Table 1. Correlation table with Spearman correlation coefficients ($n = 129$). Statistically significant ($p < 0.05$) correlation coefficients are bold; highly significant ($p < 10^{-6}$) correlation coefficients are denoted by blue colour.

	Density	*Ethanol*	*Extract*	*TotAcid*	*NVolAc*	*VolAc*	*Ash*	*SO2Free*	*RedSug*	*SFE*	*pH*
Ethanol	**-0.38**										
Extract	**0.71**	**0.31**									
TotAcid	**0.32**	**-0.21**	0.13								
NVolAc	**0.40**	**-0.30**	0.14	**0.94**							
VolAc	**-0.32**	**0.41**	-0.04	**-0.25**	**-0.50**						
Ash	0.04	0.07	0.16	**-0.43**	**-0.47**	0.15					
SO2Free	**0.25**	-0.10	**0.17**	**-0.22**	**-0.17**	-0.13	0.13				
RedSug	**0.43**	**0.21**	**0.59**	0.01	0.03	0.11	-0.03	0.13			
SFE	**0.68**	**0.29**	**0.95**	**0.18**	**0.18**	-0.06	**0.18**	0.14	**0.38**		
pH	**-0.17**	**0.19**	0.00	**-0.71**	**-0.71**	**0.22**	**0.67**	0.06	**-0.19**	0.03	
Quality	-0.03	**0.42**	**0.28**	**-0.31**	**-0.30**	**0.22**	0.10	0.00	0.12	**0.28**	**0.22**

The highly significant correlations between variables *TotAcid* and *NVolAc* (0.94) as well as *Extract* and *SFE* (0.95) corresponded with fact, that variables *NVolAc* and *SFE* were calculated from measured descriptors *TotAcid* and *Extract*, respectively. Thus, correlations of *TotAcid* and *NVolAc* as well as other linearly dependent variables to other descriptors are analogous. Further significant correlations were proved; for instance, the correlation of *pH* with *TotAcid* and *NVolAc* (-0.71) was logical. However, the significant positive correlation of *pH* with *Ash* (0.67) was not expected, although higher mineral content means more cations ant therefore higher pH. Additionally, *Ethanol* negatively correlated with *Density* (-0.38), *NVolAc* (-0.30), *TotAcid* (-0.21) and positively with *VolAc* (0.41), *pH* (0.19) and a group of mutually positively correlated variables *Extract*, *SFE* and *RedSug*.

In the study (SÁENZ-NAVAJAS *et al.,* 2012), similar conventional oenological descriptors of red wines were determined and their correlations between the attributes of in-mouth sensory perception evaluated. The authors found out that the titratable acidity positively correlated with the perceived acidity of wine, and the ethanol content was an active contributor to the astringency perception. Interestingly, the content of reducing sugars was not significantly correlated with sensory sweetness. However, the correlations of total sensorial quality were not included in the mentioned study.

3.2. Nonparametric tests

An important issue in this research was also the determination of factors affecting the variability of data. It is well known, that different wine varieties could be distinguished by their different chemical properties (BELTRÁN *et al.,* 2006; CÂMARA *et al.,* 2006; KRUZLICOVA *et al.,* 2009). Similarly, the vintage, i.e. the year of production as a factor covering especially the climatic conditions of wine production, contributes significantly to the differentiation of the wine chemical composition and hence also sensory characteristics. Indeed, according to the Kruskal-Wallis test, statistically significant differences ($p < 0.01$) among categories of *Vintage* used as a factor were found in case of descriptors *Ethanol*, *VolAc*, *SFE*, *Extract* and more importantly, also in case of target variable *Quality*. When testing *Variety* as a factor, statistically significant differences between varieties were proved only by *Ash* and *pH*. However, the aim of this paper was to build a model capable to predict the overall sensorial quality with minimum limitation of domain of application, i.e. including data originated from various wine varieties and vintages.

Furthermore, non-parametric tests were applied to see the differences in oenological descriptors when three categories of wine samples were built according to their *Quality* value: (1) premium quality wines with *Quality* value greater than 18.0 ($n = 45$); (2) quality wines with *Quality* greater than 17.0 but less than or equal to 18.0 ($n = 45$) and (3) quality and country wines with *Quality* less than or equal to 17.0 ($n = 39$). The results are summarised in Table 2. To examine statistically significant differences between particular categories, Mann-Whitney test was performed.

Table 2. Summary of investigated oenological descriptors with the median values for each category of *Quality_cat*. The statistically significant results of Kruskal-Wallis test are denoted bold.

Descriptor	*p*-value	Category of *Quality_cat* / category code		
		1/ [a]	2/ [b]	3/ [c]
Ethanol	**5.10[-07]**	12.8 [b,c]	12.1 [a]	12.2 [a]
Density	0.665	0.994	0.994	0.994
TotAcid	**0.001**	5.47 [c]	5.67 [c]	6.18 [a,b]
NVolAc	**0.003**	4.75 [c]	5.05 [c]	5.48 [a,b]
VolAc	**0.044**	0.590 [b]	0.470 [a]	0.530
pH	0.086	3.49 [c]	3.49	3.46 [a]
Ash	**0.011**	2.65	2.79 [c]	2.61 [b]
Extract	**0.012**	27.6 [b,c]	25.8 [a]	26.1 [a]
RedSug	0.527	1.90	1.90	1.70
SFE	**0.006**	26.6 [b,c]	24.6 [a]	25.6 [a]
SO2Free	0.79	24.0	26.0	26.0
NVA_TA	**0.015**	0.863 [b,c]	0.890 [a]	0.893 [a]
VA_TA	**0.017**	0.110 [b,c]	0.089 [a]	0.085 [a]

Descriptor units are introduced in Materials and Methods.
[a,b,c] The median value of the given category is significantly different from the mean value of the coded category (tested by Mann-Whitney test).

Our results differed from the study (ŠNUDERL *et al.,* 2009) where analysis of variance was used for discovering significant differences between three groups of sensorial quality. Statistically significant differences were proved only in case of ash content and density in the mentioned study. However, the disagreement with our study was obvious since different varieties were considered.

3.2. Principal component analysis

Principal component analysis (PCA) is the most widely used multivariate data analysis method for transforming the original measurement variables into new, uncorrelated variables called principal components (VARMUZA and FILZMOSER, 2009). Using this procedure, a set of orthogonal axes that represents the direction of greatest variance in the data is found (GEMPERLINE, 2006). Usually, only two or three principal components are necessary to explain a significant fraction of the information present in multivariate data when the original (measurement) variables are inter-correlated. In addition, PCA graphical output provides displaying inter-relations among original variables in the space of newly calculated PCs together with detecting possible natural grouping of samples.

The first two PCs calculated from the studied descriptors accounted for 56.61 % of the total data variability. As it was obvious in the biplot (Fig. 1), the target variable *Quality* was closely located to descriptors *Ash*, *Ethanol*, *pH* and *VolAc* representing their positive correlation.
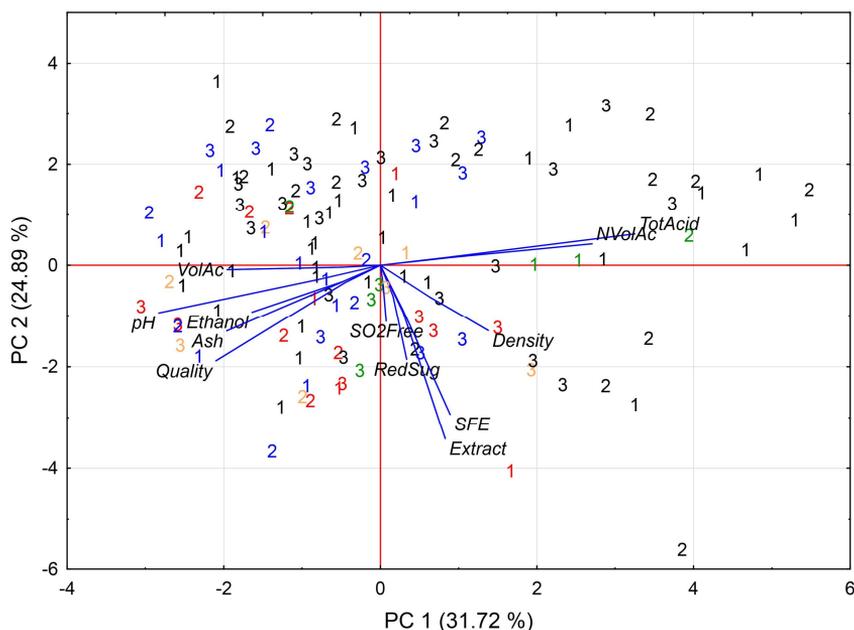

Fig. 1. Biplot for all red wine samples labelled by numbers according to their variety (1 – Cabernet Sauvignon, 2 – mixture of varieties, 3 – Merlot) and by colour according to their year of production (yellow – 2002, red – 2003, blue – 2004, black – 2005, green – 2006).

On the contrary, *Quality* was oppositely situated to adjacent descriptors *NVolAc* and *TotAcid*, i.e. *Quality* was in negative and significant correlation to them. The mutual position of the studied descriptors was in good agreement with the results of the correlation analysis.

In addition, the objects (wine samples) labelled according to their variety and vintage were mixed in the plane of PC 2 vs. PC 1 (Fig. 1). This means that no natural grouping of samples according to their variety and vintage occurred. Though, it is worth noting that the right half of graph (related to the highest positive values of PC 1) contained mostly objects representing only one vintage (2005). On the contrary, objects representing all other vintages were located mostly by negative values of PC 1. This is important due to the distribution of samples in relation to the axis PC 1 representing the original descriptors *NVolAc*, *TotAcid*, *pH*, *Ash*, *Ethanol* and *Quality* and *VolAc* following their loadings. Accordingly, wine samples with high values of *NVolAc* and *TotAcid* and lower values of *pH*, *Ash*, *Ethanol* and *Quality* originated mostly from vintage 2005. This agreed with results of Kruskal-Wallis test proving the statistically significant differences in *Vintage* categories for variable *Quality*.

3.3. Prediction of sensorial quality using multilayer perceptron

Artificial neural networks (ANNs) do not assume any initial mathematical relationship between the input and output variables, so they are particularly useful when the underlying mathematical model is unknown or uncertain (MILLER and MILLER, 2010). Finding the proper network design (number of layers and neurons) is usually a trial and error procedure (ZUPAN and GASTEIGER, 1993). Nevertheless, the key decision on the number of hidden layers and neurons was accomplished by Automated Network Search (ANS) involved in software STATISTICA, which was helpful for creating a variety of different network architectures and for choosing the network with the best performance. It was found in all cases that the best network type was the three-layer perceptron (3-MLP), i.e. the architecture with one hidden layer. The optimal number of hidden neurons was sought by examining several types of the 3-MLP with regard to the performance of the network. To avoid over-training of the network, a subset of data selected for monitoring the error periodically during the training was used (so-called test subset). Thus, the whole data set was randomly divided into three subsets: (1) training (65 % of data), (2) test (10 %) and (3) validation data (25 %) which were not included in the training and utilised to assess the predictive (generalisation) ability of developed model. According to the sensitivity analysis, the predictors (input variables) with repeatedly low value of sensitivity were sequentially removed from the set of inputs.

Firstly, all data were included in analyses (*n* =129) and only oenological descriptors were utilised as continual input variables. After the selection of predictors, the optimum architecture for networks utilising eleven inputs (*SFE*, *Extract*, *TotAcid*, *NVolAc*, *pH*, *VA_TA*, *SO2Free*, *Ethanol*, *Ash*, *VolAc* and *NVA_TA*) contained six hidden units. The predictive ability decreased gradually when more neurons were added, i.e. with higher number of hidden neurons the network became over-trained. On the other hand, when less hidden neurons were included in the model calculation,

the performance of model decreased considerably. The criterion used to select the adequate ANN model consisted of selecting the number of neurons, which gave a minimum final error in a minimal number of iterations during the training phase. The best performance was obtained when a network utilised the BFGS algorithm for 63 epochs and when the hyperbolic tangent in the hidden layer and logistic function in the output layer were used as activation functions. However, after optimisation of number of input and hidden neurons, the final performances of networks achieved only unsatisfactory results – the maximum value of coefficient of determination (R^2) related to training data was 0.64 and that for validation data was 0.42.

Table 3. The results of comparing predicted vs. actual values of *Quality* for selected 3-MLP models in case of all samples ($n = 129$) and samples representing only one vintage ($n = 66$) with using different inputs.

Number of objects	Input variables	R^2			RMSE		
		Training	Test	Validation	Training	Test	Validation
129	SFE, Extract, TotAcid, NVolAc, pH, VA_TA, SO2Free, Ethanol, Ash, VolAc, NVA_TA	0.637	0.514	0.421	0.430	0.582	0.686
129	Vintage, Variety, TotAcid, NVolAc, SFE, NVA_TA, VolAc, pH, Ethanol, Ash, VA_TA, SO2Free, Extract, Density	0.830	0.671	0.616	0.294	0.475	0.561
66	NVolAc, TotAcid, Ash, SO2Free, Ethanol, RedSug, SFE, pH, NVA_TA, Density, VolAc	0.947	0.936	0.901	0.186	0.274	0.229

Note: The input variables are ordered by decreasing importance according to the sensitivity analysis

To improve the performance of neural networks, factors *Vintage* and *Variety* were included in analyses as categorical input variables. This led to decreasing error in computations and consequently to improving results (Table 3). The BFGS algorithm for 37 epochs was selected as optimum with using hyperbolic tangent and logistic function as activation functions in hidden and output layer, respectively. Six hidden neurons were selected to get acceptable training and testing error. The resulting sensitivity values indicated the importance of predictors in following order: *Vintage, Variety, TotAcid, NVolAc, SFE, NVA_TA, VolAc, pH, Ethanol, Ash, VA_TA, SO2Free, Extract* and *Density*. This means that factors *Vintage* and *Variety* significantly influenced the recognition and extraction of data information for *Quality* prediction. The results for comparing predicted vs. actual target descriptor *Quality* are summarised in Table 3 showing increased values of R^2 and decreasing *RMSE* values in all data subsets.

Furthermore, as it was found by the Kruskal-Wallis test (Section 3.2.), there were proved statistically significant differences between categories of factor *Vintage* regarding few oenological descriptors as well as the target descriptor *Quality*.

Accordingly, the next step was to discover the possibility of predicting the overall sensorial quality when only samples originated from one year of production were involved in the analysis. Since the highest number of samples was obtained in vintage 2005 ($n = 66$), further model building was accomplished with data characterising only wine samples originated from 2005. In this case, the random distribution of objects into training, test and validation subsets was following: (1) 70 % of data used as training subset, (2) 10 % as test subset and (3) 20 % of data for validation. The results markedly improved as is shown in Table 3. The best 3-MLP contained 6 hidden neurons and again, hyperbolic tangent in hidden and logistic function in output layer as activation functions, and BFGS algorithm for 57 epochs were employed. The performance of the best ANN model is presented in Fig. 2.

Fig. 2. Correlation between the predicted and actual values of *Quality* for the 3-MLP achieving the best performance. The objects characterising wine samples originating only from vintage 2005 are divided into three subsets: (1) 47 objects used as training subset, labelled by blue squares (□), (2) 6 samples (○) used as test subset and (3) 13 samples (△) used for external validation.

Accordingly, the routinely measured chemical descriptors are strongly connected with the year of wine production and adding data representing more vintages makes the predictions more demanding.

To compare the predictive performance of the developed ANN models, other studies dealing with estimating sensorial characteristics are briefly discussed here. For instance, the overall sensory quality of 15 wine samples was estimated by means of measurements based on electronic nose, amperometric electronic tongue and

spectrophotometric determinations in the study of (BURATTI *et al*., 2007) where genetic algorithms were utilised for variables selection. Consequently, the response of carbon electrode, contents of total flavonoids, non-anthocyanin flavonoids and total anthocyanins turned out to be the most important and resulted in predictive power of 66 % determined by bootstrap procedure. Furthermore, different parameters of taste and flavour of wine were predicted by back-propagation neural network in the work of (LEGIN *et al.,* 2003) on the basis of measurements by electronic tongue. The mean relative error was within 4 – 27 % for external validation data. In the study of (SÁENZ-NAVAJAS *et al.,* 2012), the sensory astringency was predicted by partial least squares models for 34 red wine samples. Among different non-volatile compounds as well as other analytes, only few were selected the most important, such as contents of ethanol, protocatechuic acid ethyl ester, protein-precipitable proanthocyanidins, *trans-p*-coumaric, *cis*-aconitic, and *cis*-caftaric acids. It is noteworthy that the ethanol content was found the most important contributor to the astringency estimation. Resulting R^2 associated with validation of model was 0.629 and the *RMSE* was 0.441. Thus, we could conclude that the predictive ability of routinely measured oenological descriptors for sensorial quality estimation tested in our study was comparable with studies mentioned previously.

3.4. Classification according to the sensorial quality

Furthermore, an alternative approach was applied to find a proper model for prediction of overall wine sensorial quality with using routinely measured oenological descriptors. In this case, the categorical form of target variable – *Quality_cat* was employed as a categorical dependent variable. For this purpose, the wine samples were categorised into three groups according to their *Quality* value: (1) premium quality wines with *Quality* value greater than 18.0 (n = 45); (2) quality wines with *Quality* value greater than 17.0 but less than or equal to 18.0 (n = 45) and (3) quality and country wines with *Quality* less than or equal to 17.0 (n = 39). Regarding the distribution of the objects into training, test and validation subsets, similar way was used as in previous analyses when all vintages were included. Each MLP was trained using the BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm for a maximum of 200 training cycles. The networks were optimised against the cross-entropy error function (BISHOP, 1995) and again, the sensitivity analysis was utilised to eliminate predictors with poor discrimination ability. Several configurations (one hidden layer with four to seven neurons) of the ANN model network have been tried out to establish the optimum network architecture for the best prediction performance. It was observed that the network predictive ability (performance for the validation subset) decreased with adding more than six neurons of hidden layer due to over-training. The perceptron with best performance achieved considerable results (Table 4) when six neurons and hyperbolic tangent as activation function were employed in the hidden layer.

Similarly as in Section 3.3, factors *Vintage* and *Variety* involved as predictors improved the results of classification. In this case, six hidden neurons and logistic activation function in hidden layer were appropriate. In training subset, 79 from 84

objects were correctly classified. Regarding validation data, 8 objects from 32 were misclassified.

Table 4. The results of predicting categorical *Quality_cat* for selected 3-MLP models in case of all samples (*n* = 129) and samples representing only one vintage (*n* = 66) with using different inputs.

Number of objects	Input variables	Correctly classified objects (%)		
		Training	Test	Validation
129	*SFE, TotAcid, NVolAc, Ethanol, Density, Ash, pH, VolAc, NVA_TA, Extract, SO2Free*	86.9	66.7	71.9
129	*Vintage, Variety, NVA_TA, pH, NVolAc, TotAcid, VolAc, SFE, Extract, Ethanol, Ash, SO2Free*	94.1	75.0	75.0
66	*NVolAc, TotAcid, pH, Density, SFE, VA_TA, Ethanol, VolAc, Ash, RedSug, SO2Free*	93.6	66.6	84.6

Note: The input variables are ordered by decreasing importance according to the sensitivity analysis

Finally, the prediction of *Quality_cat* was performed including only data characterising vintage 2005 (*n* = 66) in the next step. Consequently, significantly higher ANN performances were achieved regarding classification results (Table 4). In this case, lower number of samples was incorporated and the optimum number of hidden units was 5 with logistic activation function in hidden layer. Briefly, 70 % of data were selected as training subset (*n* = 47), 10 % as test subset (*n* = 6) and 20 % of data were used for validation (*n* = 13). The classification success achieved 94 % for training set, i.e. 3 objects were misclassified; and the prediction ability evaluated by classification success for validation subset was 85 %, this means 11 objects from total 13 were correctly classified (Table 4). The predictors were selected by sensitivity analysis according to their importance in following order: *NVolAc, TotAcid, pH, Density, SFE, VA_TA, Ethanol, VolAc, Ash, RedSug* and *SO2Free*.

In addition, linear discriminant analysis (LDA) was accomplished to compare the performance of ANN with other classification technique. In LDA, the highest percentage of correctly classified objects for training set was 71.2 % and 65.2 % for leave-one-out validation. In this classification model, the stepwise selection procedure was applied to optimise the set of input variables and as a result, only four descriptors sufficed, concretely *Density*, *Ash*, *pH* and *TotAcid*. Still, the classification success of LDA was markedly lower in comparison to results of ANN.

It should be stressed that the number of papers dealing with the classification of wines according to their overall sensorial quality is limited. Nevertheless, in (ŠNUDERL *et al.,* 2009), LDA was applied to classify the wine samples into three categories of their sensorial quality. The best classification success (78.3 %) was obtained with nine selected input variables, concretely contents of free and total SO_2, total extract, reducing sugars, ash, polyphenols, and lactic, tartaric and citric acids. To our knowledge, the application of ANN for classification of wine samples according to the sensorial quality was not yet reported.

4. Conclusions

Since the sophisticated and modern analytical equipment is often not available to the most wine producers, the utilisation of routinely determined chemical properties as predictors of wine quality would be of high importance for practical applications. This work showed the possibility of estimating the overall sensorial quality of Slovenian red wines by simple oenological descriptors. Artificial neural networks (ANNs), specifically three-layer perceptrons (3-MLPs) with the BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm, were employed as prediction tool for achieving considerable results.

The exploratory data analysis was accomplished to discover the relations among the studied descriptors and the target sensorial quality using correlation analysis, principal component analysis as well as non-parametric tests. Numerous statistically significant correlations were proved and a good agreement amongst the applied computational methods has been observed. The highest positive correlation of sensorial quality was found with content of ethanol and the highest negative correlation with total acidity.

The importance of vintage as a factor was also obvious in the prediction performances of developed ANN models since considerably higher performances of ANN models were reached when only data representing one vintage were included in predictions. Two approaches of sensorial quality prediction were performed in the terms of using target variable in continual as well as categorical form. In both, the set of input variables was optimised and the resulting models were tested by external validation subset of data. Consequently, the predictive ability of the best 3-MLP model for estimation of sensorial quality in continual form achieved R^2 value of 0.62 in case of including all data, and 0.90 when wine samples originated only from one vintage were selected. In addition, the predictive power of 3-MLP models used in classification of wine samples into three categories was 75 % for all data and 85 % for data representing only vintage 2005. Accordingly, markedly better results were achieved in case of predictions including only data representing one vintage in comparison of predictions covering data of all vintages. Hence, involvement of more data related to wines of different origins in predictions demands more complex characterisation of samples in terms of incorporating various types of analytes in analyses.

References

ARVANITOYANNIS, I.S., KATSOTA, M.N., PSARRA, E.P., SOUFLEROS, E.H., KALLITHRAKAY, S.: Application of quality control methods for assessing wine authenticity: Use of multivariate analysis (chemometrics). Trends Food Sci. Tech., 10, 1999, 321-336.

BEDNÁROVÁ, A., KRANVOGL, R., BRODNJAK-VONČINA, D., JUG, T., BEINROHR, E.: Characterization of Slovenian Wines Using Multidimensional

Data Analysis from Simple Enological Descriptors. Acta Chim. Slov., 60, 2013, 274-286.

BELTRÁN, N.H., DUARTE-MERMOUD, M.A., BUSTOS, M.A., SALAH, S.A., LOYOLA, E.A., PEÑA-NEIRA, A.I., JALOCHA, J.W.: Feature extraction and classification of Chilean wines. J. Food Eng., 75, 2006, 1-10.

BURATTI, S., BALLABIO, D., BENEDETTI, S., COSIO, M.S.: Prediction of Italian red wine sensorial descriptors from electronic nose, electronic tongue and spectrophotometric measurements by means of Genetic Algorithm regression models. Food Chem., 100, 2007, 211-218.

BISHOP, C.M.: Neural Networks for Pattern Recognition, Clarendon Press, Oxford, 1995, 482 pp.

CÂMARA, J.S., ALVES, M.A., MARQUES, J.C.: Multivariate analysis for the classification and differentiation of Madeira wines according to main grape varieties. Talanta, 68, 2006, 1512-1521.

COZZOLINO, D., CYNKAR, W.U., SHAH N., DAMBERGS R.G., SMITH P.A.: A brief introduction to multivariate methods in grape and wine analysis. Int. J. Wine Res., 1, 2009, 123-130.

GEMPERLINE, P.: Practical Guide to Chemometrics, CRC Press, Boca Raton, 2006, 520 pp.

HAYKIN, S.: Neural Networks: A comprehensive Foundation, Pearson Education, Dehli, 1999, 823 pp.

HUNTER, A., KENNEDY, L., HENRY, J., FERGUSON, I.: Application of neural networks and sensitivity analysis to improved prediction of trauma survival. Comput. Meth. Prog. Bio., 62, 2000, 11-19.

KRUZLICOVA, D., MOCAK, J., BALLA, B., PETKA, J., FARKOVA, M., HAVEL, J.: Classification of Slovak white wines using artificial neural networks and discriminant analysis. Food Chem., 112, 2009, 1046-1052.

KRUZLICOVA, D., FIKET, Ž., KNIEWALD, G.: Classification of Croatian wine varieties using multivariate analysis of data obtained by high resolution ICP-MS analysis. Food Res. Int., 54, 2013, 621-626.

LEGIN, A., RUDNITSKAYA, A., LVOVA, L., VLASOV, Y., DI NATALE, C., D'AMICO, A.: Evaluation of Italian wine by the electronic tongue recognition, quantitative analysis and correlation with human sensory perception. Anal. Chim. Acta, 484, 2003, 33-44.

LÓPEZ-FERIA, S., CÁRDENAS, S., VALCÁRCEL, M.: Simplifying chromatographic analysis of the volatile fraction of foods. Trends Anal. Chem., 27, 2008, 794-803.

LORRAIN, B., TEMPERE, S., ITURMENDI, N., MOINE, V., DE REVEL, G., TEISSEDRE, P.-L.: Influence of phenolic compounds on the sensorial perception and volatility of red wine esters in model solution: An insight at the molecular level. Food Chem., 140, 2013, 76-82.

MILLER, J.N., MILLER, J.C.: Statistics and Chemometrics for Analytical Chemistry, Pearson Education Limited, Harlow, 2010, 278 pp.

SÁENZ-NAVAJAS, M.-P., AVIZCURI, J.-M., FERREIRA, V., FERNÁNDEZ-ZURBANO, P.: Insights on the chemical basis of the astringency of Spanish red wines. Food Chem., 134, 2012, 1484-1493.

POHL, P.: What do metals tell us about wine? Trends Anal. Chem., 26, 2007, 941-949.

SAURINA, J.: Characterization of wines using compositional profiles and chemometrics. Trends Anal. Chem., 29, 2010, 234-245.

ŠNUDERL, K., MOCAK, J., BRODNJAK-VONČINA, D., SEDLÁČKOVÁ, B.: Classification of white varietal wines using chemical analysis and sensorial evaluations. Acta Chim. Slov., 56, 2009, 765-772.

VARMUZA, K., FILZMOSER, P.: Introduction to Multivariate Statistical Analysis in Chemometrics, CRC Press, Boca Raton, 2009, 321 pp.

ZUPAN, J., GASTEIGER, J.: Neural Networks for Chemists: An introduction, VCH, New York, 1993, 305 pp.