

# Classification of Polish wines by application of ultra-fast gas chromatography

Tomasz Majchrzak, Wojciech Wojnowski, Justyna Płotka-Wasyłka\*

*Department of Analytical Chemistry, Faculty of Chemistry, Gdańsk University of Technology, 11/12 Narutowicza Street, 80-233 Gdańsk, Poland*

Corresponding author: [juswasyl@pg.edu.pl](mailto:juswasyl@pg.edu.pl); [plotkajustyna@gmail.com](mailto:plotkajustyna@gmail.com)

## Abstract

The potential of ultra-fast gas chromatography (GC) combined with chemometric analysis for classification of wine originating from Poland according to the variety of grape used for production was investigated. A total of 44 Polish wine samples differing in the type of grape (and grape growth region) used for the production as well as parameters of the fermentation process, alcohol content, sweetness and others which characterize wine samples were analysed. The selected features coming from ultra-fast GC analysis were subsequently used as inputs for both principal component analysis (PCA) and supervised machine learning. Using the proposed classification algorithm, it was possible to classify white and red wines according to the variety of grape used for production with a 98.7% and 98.2% accuracy, respectively. The model was characterised by good recall and area under receiver operating characteristic which was 1.000 for white wines and 0.992 for red wines. Cuveé wines (made from various types of grapes) were also successfully classified which leads to the conclusion that the proposed classification method can be used not only to differentiate between wines made from different grapes but also to detect possible adulterations, provided known, non-adulterated samples are available as a reference. The model was also used to classify wine samples based on other features, such as the geographic region in which the vineyard is situated, type of yeast used, the temperature of fermentation, sweetness, etc. In all cases, a high classification accuracy (in most cases >90%) was achieved. The obtained results could be applied in the wine industry.

## Keywords

Wine; classification; ultra-fast gas chromatography; chemometric analysis; principal component analysis; support vector machines

## 1. Introduction

Nowadays, wine identification, as well as classification, has gained increasing attention as a means to detect mislabeling, taking into account the great variability of the sale price depending on wine age, vintage year, varietal, or geographical origin (Yu et al 2014). In fact, counterfeiting of food and alcoholic beverages including wine is recently one of the risks relevant for producers, distributors, consumers, and national governments from many points of view such as economic (price), health (allergens), and religious reasons (Gliszczyńska-Świągło & Chmielewski 2017). As a result of the above-mentioned issues, the Food and Drug Administration created the term ‘economically motivated adulteration’ (EMA) as a subcategory of food fraud (Everstine et al 2013). Food and Drug Administration proposes a working definition of EMA as the fraudulent, intentional substitution or addition of a substance in a product for the purpose of increasing the apparent value of the product or reducing the cost of its production, i.e., for economic gain (Sotirchos et al 2017).

36 The most prevalent fraudulent procedure is the partial or complete substitution of an authentic ingredient or  
37 material with a cheaper and easily available component (Hrbek et al 2015) which results in a worse quality product  
38 usually without a substantial effect on human health.

39 The authenticity of various products such as wine is often associated with a geographical area of production  
40 and/or specific processing technology. Therefore, to protect food products specific to a given area or  
41 manufactured using a particular process from imitation and to safeguard their authenticity indexes such as the  
42 Symbols of Protected Geographical Indication, Traditional Speciality Guaranteed (TSG) and Protected  
43 Designation of Origin (PDO) have been created and introduced by the European Union.

44 In regard to wine, several parameters can be monitored to establish wine age, vintage year, the origin of the wine,  
45 etc. Analytical techniques are often used to determine the profile of such compounds like phenols and polyphenols  
46 (Hernandez et al 2006), flavonoids (Fang et al 2007), amino acids (Shen et al 2011), as well as pigment  
47 composition (Alcalde-Eon et al 2006). Moreover, the profile of volatile compounds is of high importance when  
48 ensuring correct identification and authenticity due to the fact that the flavour of food and beverages, including  
49 wine, is one of the key indicators of their quality (Yu et al 2014; Gliszczyńska-Świgło & Chmielewski 2017).  
50 This is because the volatile compounds which characterize every food product and drink are diverse and  
51 originate from raw materials and/or are generated during production, maturation and storage. Thus, such aroma  
52 markers could be identified to confirm their authenticity (Pillonel et al 2003).

53 Several methods and techniques are used to evaluate the quality or detect adulteration in food. They comprise  
54 chromatographic methods such as high-performance liquid chromatography (HPLC) and gas chromatography  
55 (GC) as well as spectroscopic techniques, e.g. ultraviolet-visible (UV-Vis) spectroscopy, mass spectrometry  
56 (MS) and fluorescence spectroscopy, Fourier transform infrared (FT-IR) spectroscopy and nuclear magnetic  
57 resonance (NMR) spectroscopy. The analysis is often supplemented with the use of chemometric techniques  
58 (Hristov et al 2016; Nedyalkova et al 2017; Szczepańska et al 2017; Wiczerzak et al 2016) providing  
59 satisfactory results for quality control or determination of food authenticity. Despite the fact that the above-  
60 mentioned techniques are usually the most specific and sensitive, they require the use of expensive equipment  
61 and high-degree technical expertise (Gliszczyńska-Świgło & Chmielewski 2017). Moreover, they cannot be used  
62 for the determination of volatile compounds without a derivatization step (excluding GC). Gas chromatography  
63 coupled to mass spectrometry (GC-MS) has been applied to ensure identity and authenticity of wines (Hernanz et  
64 al 2009; Pereira et al 2011).

65 On the other hand, determination of the changes in the composition of the volatile fraction of products may  
66 sometimes be insufficient to confirm their authenticity. In addition, the analytical procedure based on the  
67 application of chromatographic techniques to determine the indicators of food authenticity consists of many  
68 steps such as sample preparation (extraction, derivatization), separation, and identification of compounds which  
69 are usually labour- and time-consuming and generate extra costs. Thus, the best solution to solve these problems  
70 is the application of a tool which is characterized by direct, rapid, and effective determination of the authenticity  
71 of a product based on its aroma.

72 One of the choices for these purposes is the application of sensory evaluation which is another reliable method for  
73 a vintage year or wine age determination. However, the sensory evaluation has its own deficiencies with the most  
74 important being that it is a subjective method and only trained and experienced panellists can reliably evaluate  
75 flavour of the product such as wine (Yu et al 2014). Another example of a technique which can be used for

76 monitoring the authenticity of wine is the electronic nose (e-nose) which is a very rapid, robust and cost-effective.  
77 Furthermore, the use of e-noses involves no special sample preparation to determine the aroma of a wine. Due to  
78 these advantages, e-noses are becoming increasingly popular as objective and automated techniques to characterize  
79 food flavours (Yu et al 2014). Similar advantages are presented by another analytical technique, ultra-fast gas  
80 chromatography which also can be used for aroma profiling of wine. Although several parameters have to be  
81 adjusted to increase the separation speed (e.g. carrier gas flow rate, the temperature-program heating rates, the  
82 column length, etc.), there is virtually no sample preparation required which shortens the analysis time.  
83 The most common approach in the analysis of the volatile fraction of wine samples using gas chromatography is  
84 the identification and determination of the headspace constituents (De la Calle García et al 1998). Such an approach  
85 is not viable when using ultra-fast gas chromatography due to the short length of chromatographic columns and  
86 steep temperature ramps. Instead, a holistic, “fingerprinting” approach could be used.  
87 In this work, ultra-fast GC was applied to determine the aroma differences among the Polish wines originating  
88 from a variety of grapes and characterized by different fermentation parameters. A multivariate statistical analysis  
89 of obtained data was performed to classify wine samples based on the variety of grapes and other variables. It  
90 needs to be noted that presented technique requires no sample preparation and is very rapid and furthermore does  
91 not require the use of solvents which follows the guidelines of Green Analytical Chemistry. Therefore, such  
92 technique may be a very useful tool for the purposes of the food industry.

93

## 94 **2. Materials and methods**

### 95 *2.1. Wine samples*

96 A total of 44 samples (red, white and rosé) originating from Polish vineyards located in different parts of Poland  
97 were collected. All samples were stored at room temperature (21°C) and protected from light. Information  
98 regarding the samples is presented in Table 1.

Table 1. Information on wine samples analysed.

No <sub>sample</sub>	Production year	Region	Location m a.s.l.	Grape variety	Alcohol content [%]	Sweetness	Type of yeast	Additives	Filtration (yes/no)	Fermentation processes (1 or 2)	Fermentation temperature	Color
1W	2016	West Pomeranian	125	Solaris	13.6	Extra Dry	Uclm325	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	22	W
2W	2016	West Pomeranian	125	Solaris	12.9	Dry	Uclm325	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	22	W
1R	2016	West Pomeranian	125	Allegro	12.9	Dry	Murvinb	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	17	R
2R	2016	West Pomeranian	125	Regent, Rondo	12.1	Dry	Wild&Pur	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	20	R
3W	2016	West Pomeranian	125	Seyval Blanc	9.5	Semi-Sweet	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	18	W
4W	2016	West Pomeranian	125	Seyval Blanc	10.1	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	18	W
3R	2016	Kuyavian-Pomeranian	74	Rondo	13.5	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	NO	1	17	R
4R	2016	Kuyavian-Pomeranian	74	Regent	13.5	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	NO	1	17	R
5W	2016	Kuyavian-Pomeranian	74	Bianca	12	Semi-Sweet	Cks102	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	NO	1	12	W
6W	2016	Kuyavian-Pomeranian	74	Solaris	17	Dry	Cks102	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	NO	1	12	W
5R	2016	Pomeranian	120	Regent	12	Dry	Saccharomyces Cerevisiae Bs 7 Fruity –Regent, Merlot	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	19	R
7W	2016	Pomeranian	120	Aurora, Bianca, Hiberna, Muscat	12	Dry	Saccharomyces Bayanus Bs-11	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Ascorbic Acid	YES	1	23	W
8W	2016	Pomeranian	120	Aurora, Bianca	12	Dry	Saccharomyces Bayanus Bs-11	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Ascorbic Acid	YES	1	23	W
6R	2016	Kuyavian	92	Pinot Noir, Pinot Gris	12	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub>	NO	1	16.5	R
1Re	2016	Kuyavian	92	Regent, Rondo	11	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub>	NO	1	16.5	Ro
9W	2016	Kuyavian	92	Hiberna, Bianca, Muller Thurgau	12.5	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Ascorbic Acid	NO	1	16.5	W
7R	2014	Subcarpathian	320	Rondo	12	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	17	R
8R	2015	Subcarpathian	320	Regent	12	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	17	R
2Re	2014	Subcarpathian	320	Rondo Rose	11.5	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	17	Ro
W	2015	Subcarpathian	320	Bianca	12.5	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	17	W
W	2015	Subcarpathian	320	Hiberna	12.5	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	17	W
W	2012	Subcarpathian	320	Hiberna	17	Sweet	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	17	W
W	2016	Subcarpathian	296	Hiberna	23	Semi-Sweet	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	NO	1	17	W
W	2016	Lesser Poland		Marechal Foch, Leon Millot	11.5	Dry	Lalvin71b	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	NO	1	17	R
W	2015	Lesser Poland	335	Hiberna	13	Dry	Ck S102	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> ,	YES	1	12	W
W	2015	Masovian	151	Jutrzenka	10	Semi-Sweet	Enartis Ferm Aroma White	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	YES	1	17	W

16W	2014	Masovian	151	Jutrzenka	11	Dry	Enartis Ferm Aroma White	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	YES	1	17	W
17W	2015	Masovian	151	Aurora, Bianca	10	Dry	Fermivin Pdm. Bio L1	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	YES	1	17	W
18W	2016	Masovian	151	Aurora, Bianca	12	Semi-Dry	Oenoferm Inter Dry F3	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	YES	1	16	W
19W	2014	Masovian	151	La Crescent	11.5	Dry	Enovi	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	W
20W	2015	Masovian	151	La Crescent, St. Pepin	12	Dry	Fermivin	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	W
21W	2015	Masovian	151	Andalamina, Kristally, Prarie Star	10	Dry	Fermivin Pdm. Bio L1	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	W
22W	2016	Masovian	151	Andalamina, Kristally	11	Dry	Oenoferm Inter Dry F3	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	YES	1	17	W
23W	2015	Masovian	151	Seywal Blanc	13	Semi-Dry	Enartis Ferm Aroma White	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	YES	1	17	W
24W	2016	Masovian	151	St. Pepin, La Crescent	16	Sweet	Oenoferm Bouquet F3	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Alcohol Addition	YES	1	17	W
10R	2015	Masovian	151	Frontenac	13	Semi-Sweet	Enartis Ferm Red Fruit	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	2	17	R
11R	2016	Masovian	151	Frontenac	13	Dry	Oenoferm Clolor F3	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	R
12R	2015	Masovian	151	Regent	12	Dry	Aromatic Wine Complex Yeast Est. 2005 Spittferm	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	R
13R	2015	Masovian	151	Regent, Frontenac	11.5	Semi-Dry	Aromatic Wine Complex Yeast Est. 2005 Spittferm	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	R
14R	2016	Masovian	151	Heridian	11	Dry	Oenoferm Color F3	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	R
15R	2015	Masovian	151	Leon Millot, Marechal Foch, Regent	12	Semi-Dry	Red Fruit, Aromatic Wine Complex Premium Yease Est. 2005 Spiritferm	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	R
R	2016	Masovian	151	Leon Millot, Marechal Foch,	12	Semi-Dry	Oenoferm Color F3	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	R
R	2014	Masovian	151	Marechal Foch, St. Croix	11.5	Dry	Oenoferm Color F3	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	R
R	2016	Masovian	151	St. Croix, Sabrevois	10	Dry	Oenoferm Color F3	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	R
)"	2012	Masovian	151	Cherry	14	Sweet	Cherry	K <sub>2</sub> S <sub>2</sub> O <sub>5</sub> , Chaptalisation	NO	1	17	R

R. red; ro. rosé; w. white

Sweetness (gram of sugar per litre): Extra dry: 0 g/l; dry: up to 4 g/l; semi-dry: up to 12 g/l; semi-sweet: up to 45 g/l; sweet: more than 45 g/l



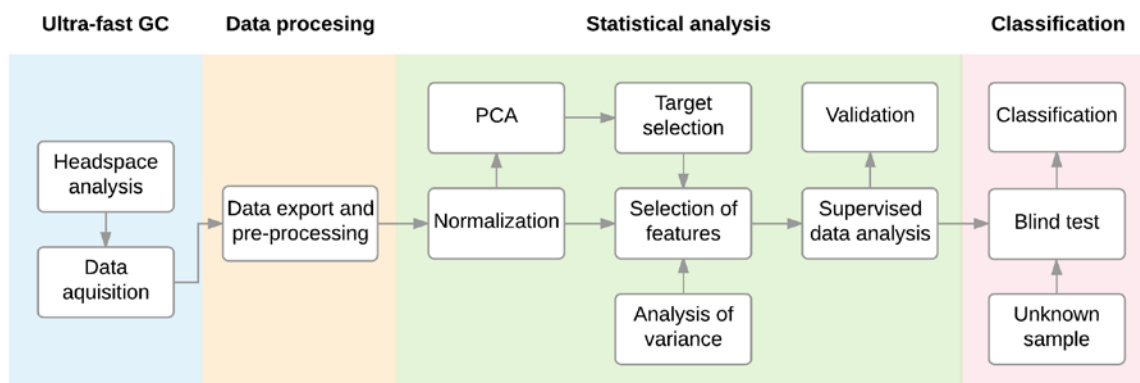
93           2.2. *Headspace analysis*

94   Prior to the analysis, the wine was stored at room temperature without the access of light. Samples of 5 ml were  
95   poured into 20 ml glass headspace vials and sealed with caps lined with a silicon-PTFE membrane and incubated  
96   for 10 min at 40°C in order to facilitate the transfer of analytes to the sample's volatile fraction. During incubation,  
97   the samples were stirred at 500 rpm. Static headspace analysis was performed using the Heracles II ultra-fast gas  
98   chromatography device equipped with the HS100 autosampler (Alpha M.O.S., Toulouse, France). The device was  
99   fitted with two parallel 10-m columns packed with the MXT-5 and MXT-1701 stationary phases (Restek,  
100   Bellefonte, PA, USA), respectively. Each column was coupled to a flame-ionization detector ( $\mu$ FID), and hydrogen  
101   of 6N purity delivered using the Precision Hydrogen Trace 250 generator (Peak Scientific Instruments, Inchinnan,  
102   UK) was used as carrier gas. The implemented parameters were based on a previously reported method developed  
103   for the analysis of alcoholic beverages (Wiśniewska et al 2016a; Wiśniewska et al 2016b). The static  
104   headspace sampling volume was 2.5 ml at 0.25 ml/s. The injector temperature was set to 200°C and the injection  
105   time was 15 s. During this time the analytes were trapped on a Tenax® TA sorptive material at 40°C and held for  
106   20 s and then purged into chromatographic columns through thermal desorption at 240°C. The oven was ramped  
107   from 70°C to 270°C at 2°C/min, and the acquisition duration was set to 100 s.

108           2.3. *Data processing and statistical analysis*

109   Data from the ultra-fast GC analysis was exported and processed using a Visual Basic-based macro. Statistical  
110   data analysis was performed using Orange v. 3.7 machine learning toolkit (Demšar et al 2013). Normalized features  
111   (chromatographic peak areas) with the highest impact on the classification outcome were then selected based on  
112   the result of the one-way analysis of variance (ANOVA). The selected features were then used as inputs for both  
113   principal component analysis (PCA) and supervised machine learning. PCA is an unsupervised multivariate  
114   statistical method which is used primarily to reduce the dimensionality of a data set and also verify the validity of  
115   data and to visualise it (Majchrzak et al 2017). Twenty features with the highest impact on the classification based  
116   on the ANOVA were used as inputs for the analysis. The number of features was reduced in order to avoid the so-  
117   called 'voodoo correlations', that is coincidental correlations which occur when the ratio of the number of  
118   measurements to the number of independent variables is low (Amann et al 2014). The supervised machine  
119   learning was conducted using the support vector machines method with regression loss ( $\epsilon$ ) of 0.10 and RBF kernel  
120   (Boser et al 1992). The method was validated using stratified 10-fold cross-validation and subsequently evaluated

121 in a blind test with 66% of data selected through random sampling used as a training set and the remaining 34%  
122 of data used for testing. Schematic representation of the described process is shown in Figure 1.

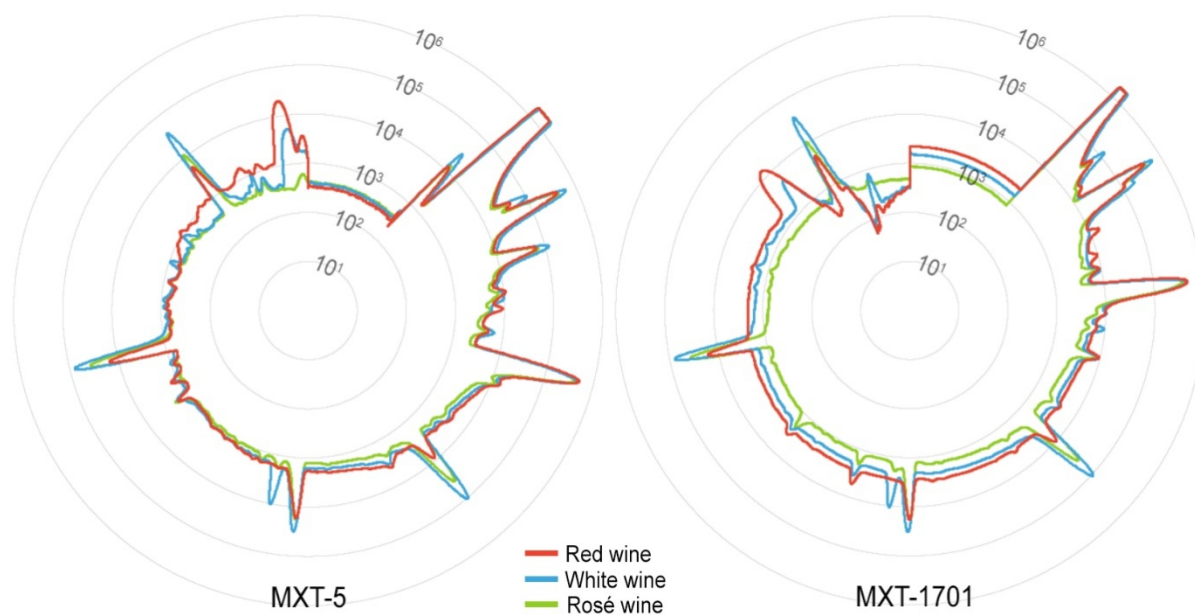


123

124 Figure 1. Schematic representation of the described process

### 125 3. Results and discussion

126 Radar plots of the obtained chromatograms are depicted in Figure 2. The resolution is relatively low and only a  
127 fraction of peaks is separated at the base. However, using ultra-fast GC it was possible to reduce the time of a  
128 single analysis to 100 seconds, which would pose a significant challenge when using classical gas chromatography.  
129 Moreover, the obtained chromatograms should be viewed as the sample's 'smellprint', which can be used for  
130 holistic analysis and classification. However, the differences between the composition of white, red and rosé wines  
131 evident in the chromatograms cannot on their own be considered a basis for classification, and even less so when  
132 discrimination between different grapes of the same colour is attempted. For this reason, it is necessary to use data  
133 analysis techniques such as PCA.



134

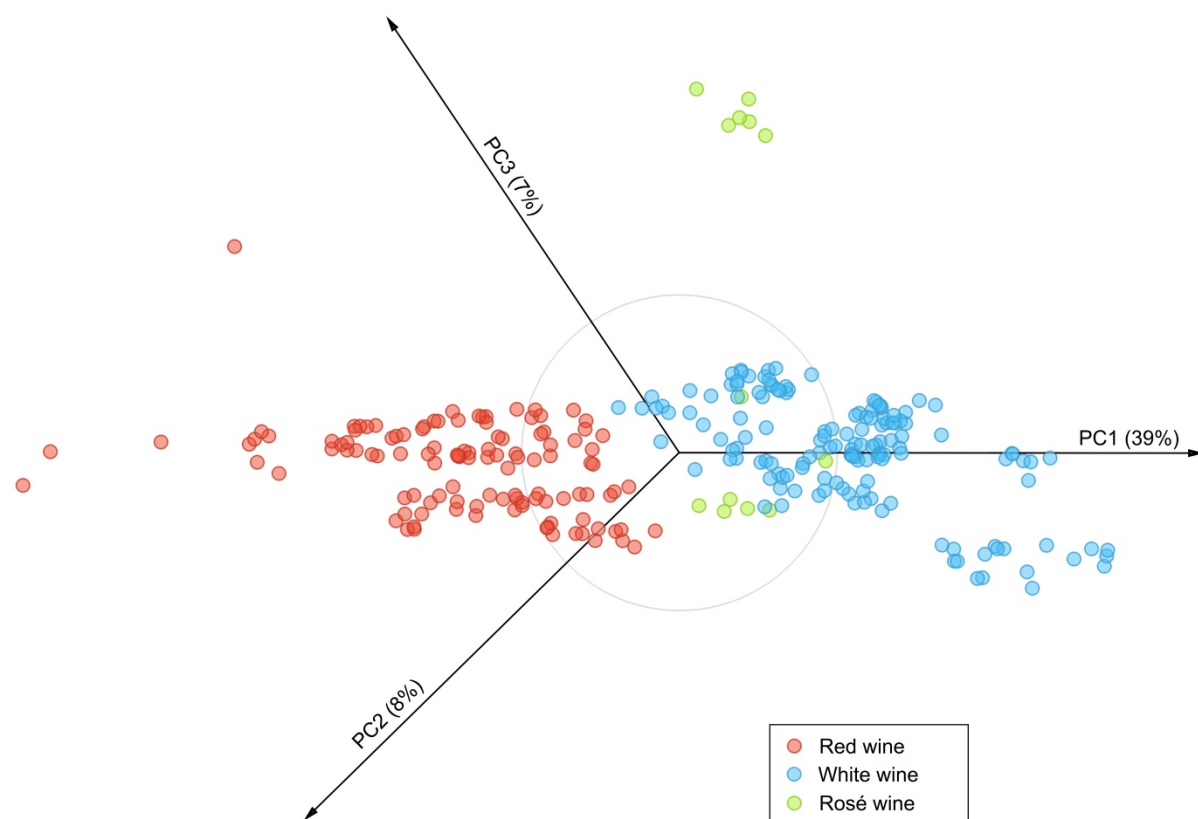
135 Figure 2. Aroma profile (chromatograms) of different wine types obtained with ultra-fast GC; the circumference  
 136 of the plots denotes retention time (100 s in total), and the radius denotes abundance (signal of FID detectors) in a  
 137 logarithmic scale.

138 *3.1. Principal component analysis*

139 The first six principal components covered 71% of total variance of the dataset. A FreeViz projection (in which  
 140 the data points remain immobile, however, the position of dimensional axes is optimized in order to provide the  
 141 most informative projection) of the principal component analysis of the entire data set is depicted in Figure 3.

142 Based on the results it was concluded that in the further statistical analysis of red and white wines will be conducted  
 143 separately, as their headspace composition is evidently distinct. Conversely, samples of rosé wine were discarded  
 144 from further analysis as their number was much lower than that of red and white wines and the results would be  
 145 difficult to compare directly.





146

147 Figure 3. Projection of the result of principal component analysis of red, white and rosé wine samples

148 *3.2. Classification using support vector machines*

149 Using the proposed classification algorithm, it was possible to classify white wines according to the variety of  
 150 grape used with a 98.7% and 98.2% accuracy in the case of red wines. The model was characterised by good recall  
 151 and area under receiver operating characteristic which was 1.000 and 0.992 for white and red wines, respectively.  
 152 In a blind test, it was possible to correctly classify 100% of analysed samples according to the variety, including  
 153 cuveé wines. Cuveé wines (wines produced from a mixture of several grape varieties) were also successfully  
 154 classified (100% successful classification). Since it was possible to differentiate between wines made from a  
 155 particular grape cultivar and cuveé wines made from the same grape cultivar alongside others, it would be  
 156 furthermore possible to detect admixtures of wines made from other grape varieties.

157 The model was also used to classify wine samples based on other features, namely the geographic region in which  
 158 the vineyard is situated, alcohol content, type of yeast used, the temperature of fermentation, sweetness and post-  
 159 fermentation treatment (Table 1). The classification evaluation results for these scenarios are listed in Table 2 and  
 160 Table 3. In the case of white wines, the best accuracy was achieved when classifying according to grape varieties  
 161 (98.7%), the temperature of fermentation (97.3%) and geographic region (96.7%). The worst classification

162 accuracy was achieved in the case of alcohol content (80.4%). It should be noted though that since the wine samples  
 163 were not diluted prior to the analysis the chromatographic peak corresponding to ethanol was in each case  
 164 overloaded, and so the feature was automatically discarded based on ANOVA. Because of that the classification  
 165 according to alcohol content is explicitly not based on the actual alcohol concentration in the samples.

166 In the case of red wines, the best accuracy was achieved for classification based on the type of yeast and  
 167 temperature of fermentation – 100% in both cases. The lowest accuracy was achieved in the case of classification  
 168 according to the grape variety. However, in a blind test, the samples were discriminated with 100% accuracy  
 169 according to all the features besides alcohol content and the temperature of fermentation.

170 Table 2. Classification parameters of various features of white wine samples

Feature	Classification Accuracy	Area under ROC	Precision	Recall	Blind test accuracy
Grape variety	98.7%	1.000	1.000	1.000	100%
Geographic region	96.7%	1.000	1.000	1.000	100%
Sugar content	87.7%	1.000	1.000	1.000	88.2%
Alcohol content	80.4%	0.960	0.629	0.900	84.3%
Yeast used	92.7%	1.000	0.895	0.850	96.1%
Post-fermentation treatment	93.7%	0.993	0.921	0.948	92.2%
Temperature of fermentation	97.3%	1.000	1.000	1.000	100%

171

172 Table 3. Classification parameters of various features of red wine samples

Feature	Classification Accuracy	Area under ROC	Precision	Recall	Blind test accuracy
Grape variety	98.2%	0.992	0.921	1.00	100%
Geographic region	99.7%	1.00	0.994	1.00	100%
Sugar content	96.7%	1.000	1.000	1.000	100%
Alcohol content	99.4%	1.000	1.000	1.000	87.1%
Yeast used	100%	1.000	1.000	1.000	100%
Post-fermentation treatment	100%	1.000	1.000	1.000	100%
Temperature of fermentation	98.8%	0.998	0.987	1.000	97.1%

173

#### 174 4. Summary

175 A total of 44 Polish wines differing in many features were analysed by application of ultra-fast gas  
 176 chromatography. The results have been furthermore subjected to a chemometric analysis. Using ultra-fast GC it  
 177 was possible to reduce the time of a single analysis to 100 seconds which would pose a significant challenge when  
 178 using classical GC. Based on the chemometrics results it was concluded that the further statistical analysis of red  
 179 and white wines will be conducted separately, as their headspace composition is evidently distinct. Conversely,  
 180 samples of rosé wine were discarded from further analysis as their number was much lower than that of red and  
 181 white wines and the results would be difficult to compare directly.

182 Using the proposed classification algorithm, it was possible to classify white and red wines according to the variety  
183 of grape used for production with a 98.7% and 98.2% accuracy, respectively. The model was characterised by  
184 good recall and area under receiver operating characteristic. Moreover, cuveé wines were also successfully  
185 classified which leads to the conclusion that the proposed classification method can be used not only to differentiate  
186 between wines made from different grapes but also to detect possible adulterations. In addition, the model was  
187 used to classify wine samples based on other features, namely the geographic region in which the vineyard is  
188 situated, alcohol content, type of yeast used, the temperature of fermentation, sweetness and post-fermentation  
189 treatment with satisfying results. Due to such advantages as no sample preparation, short time of analysis and no  
190 waste production (only wine taken into consideration), ultra-fast GC combined with chemometric analysis could  
191 be a reliable tool for detection of adulteration in the wine industry.

## 192 **Compliance with Ethical Standards**

193 **Funding:** This study was funded by the Faculty of Chemistry, Gdańsk University of Technology for financial  
194 support within the minigrant program (Decision no. 4914/E-359/M/2017).

195 **Conflict of Interest:** Justyna Płotka-Wasyłka has received research mini-grant from by the Faculty of Chemistry,  
196 Gdańsk University of Technology and she declares no conflict of interest. Tomasz Majchrzak declares that he has  
197 no conflict of interest. Wojciech Wojnowski declares that he has no conflict of interest.

198 **Ethical approval:** This article does not contain any studies with human participants or animals performed by any  
199 of the authors.

200 **Informed consent:** Not applicable.

## 201 **References**

- 202 Alcalde-Eon C, Escribano-Bailon MT, Santos-Buelga C, Rivas-Gonzalo JC (2006) Changes in the detailed  
203 pigment composition of red wine during maturity and ageing. A comprehensive study. *Anal Chim Acta* 563(1–  
204 2):238–254.
- 205 Amann A, Costello BDL, Miekisch W, Schubert J, Buszewski B, Pleil J, Ratcliffe N, Risby T (2014) The human  
206 volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *J*  
207 *Breath Res* 8: 34001.
- 208 Boser BE, Guyon IM, Vapnik VN (1992) A training algorithm for optimal margin classifiers. In: *Proceedings of*  
209 *the fifth annual workshop on Computational learning theory - COLT '92*, pp. 144–152.
- 210 De la Calle García D, Reichenbacher M, Danzer K, Hurlbeck C, Bartsch C, Feller KH (1998) Analysis of Wine  
211 Bouquet Components Using Headspace Solid-Phase Microextraction-Capillary Gas Chromatography. *J High*  
212 *Resolut Chromatogr* 21: 373–377.
- 213 Demšar J, Curk T, Erjavec A, Hočevar T, Milutinovič M, Možina M, Polajnar M, Toplak M, Starič A, Stajdohar  
214 M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013) Orange: Data Mining Toolbox in Python. *J Mach*  
215 *Learn Res* 14: 2349–2353.
- 216 Everstine K, Spink J, Kennedy S (2013) Economically motivated adulteration (EMA) of food: common  
217 characteristics of EMA incidents. *J Food Prot* 76: 723–35.
- 218 Fang F, Li JM, Pan QH, Huang WD (2007) Determination of red wine flavonoids by HPLC and effect of aging.  
219 *Food Chem* 101(1):428–433.

- 220 Gliszczyńska-Świągło A, Chmielewski J (2017) Electronic Nose as a Tool for Monitoring the Authenticity of  
221 Food. A Review. *Food Anal Meth* 10:1800–1816.
- 222 Hernandez T, Estrella I, Carlavilla D, Martin-Alvarez PJ, Moreno-Arribas MV (2006) Phenolic compounds in  
223 red wine subjected to industrial malolactic fermentation and ageing on lees. *Anal Chim Acta* 563(1–2):116–125.
- 224 Hernanz D, Gallo V, Recamales AF, Melendez-Martínez AJ, GonzalezMiret ML, Heredia FJ (2009) Effect of  
225 storage on the phenolic content, volatile composition and colour of white wines from the varieties Zalema and  
226 Colombard. *Food Chem* 113(2):530–537.
- 227 Hrbek V, Vaclavik L, Elich O, Hajslova J (2014) Authentication of milk and milk-based foods by direct analysis  
228 in real time ionization-high resolution mass spectrometry (DART-HRMS) technique: a critical assessment.  
229 *Food Contr* 36:138–145.
- 230 Hristov H, Nedyalkova M, Madurga S, Simeonov V (2017) Boron Oxide Glasses and  
231 Nanocomposites: Synthetic, Structural and Statistical Approach. *J. Mat. Sci. & Technol.* 33: 535 – 540.  
232
- 233 Majchrzak T, Wojnowski W, Dymerski T, Gębicki J, Namieśnik J (2017) Electronic noses in classification and  
234 quality control of edible oils: A review. *Food Chem* 246:192–201.
- 235 Nedyalkova M, Donkova B, Simeonov V (2017) Chemometrics Expertise in the Links between  
236 Ecotoxicity and Physicochemical Features of Silver Nanoparticles: Environmental Aspects. *J. AOAC*  
237 *Int.*, 100(2): 359 – 364.
- 238 Pereira AC, Reis MS, Saraiva PM, Marques JC (2011) Madeira wine ageing prediction based on different  
239 analytical techniques: UV–vis, GC-MS, HPLC-DAD. *Chemometr Intell Lab* 105(1):43–55.
- 240 Pillonel L, Ampuero S, Tabacchi R, Bosset JO (2003) Analytical methods for the determination of the  
241 geographic origin of Emmental cheese: volatile compounds by GC/MS-FID and electronic nose. *Eur Food Res*  
242 *Technol* 216:179–183.
- 243 Shen F, Ying Y, Li B, Zheng Y, Zhuge Q (2011) Multivariate classification of rice wines according to ageing  
244 time and brand based on amino acid profiles. *Food Chem* 129(2):565–569.
- 245 Sotirchos DG, Danezis GP, Georgiou CA (2017) Introduction, Definitions and Legislation. In: Georgiou CA,  
246 Danezis GP, (Eds.) *Food Authentication: Management, Analysis and Regulation*, Wiley-Blackwell, Greece, pp.  
247 3-18.
- 248 Szczepanska N, Kudlak B, Nedyalkova M, Simeonov V, Namiesnik J (2017) Application of  
249 Chemometric Techniques in Studying of Toxicity of Selected Commercially Available Products for  
250 Infants and Children. *Environ. Monit. Assess.* 189: 309.
- 251 Wiczerzak M, Kudlak B, Yotova G, Nedyalkova M, Tsakovski S, Simeonov V, Namiesnik J (2016)  
252 Modeling of pharmaceuticals mixtures toxicity with deviation ratio and best-fit functions models. *Sci.*  
253 *Tot. Environ.*, 571: 259-268.
- 254 Wiśniewska P, Śliwińska M, Namieśnik J, Wardencki W, Dymerski T (2016a) The Verification of the  
255 Usefulness of Electronic Nose Based on Ultra-Fast Gas Chromatography and Four Different  
256 Chemometric Methods for Rapid Analysis of Spirit Beverages. *J. Anal. Methods Chem.* 1–12.
- 257 Wiśniewska P, Śliwińska M, Dymerski T, Wardencki W, Namieśnik J (2016) Differentiation Between  
258 Spirits According to Their Botanical Origin. *Food Anal. Methods* 9: 1029–1035.  
259
- 260 Yu H, Dai X, Yao G, g Xiao Z (2014) Application of Gas Chromatography-Based Electronic Nose for  
261 Classification of Chinese Rice Wine by Wine Age. *Food Anal Meth* 7:1489–1497.
- 262
- 263

264

265

266

267