ANNALS OF “DUNAREA DE JOS” UNIVERSITY OF GALATI
MATHEMATICS, PHYSICS, THEORETICAL MECHANICS
FASCICLE II, YEAR XIV (XLV) 2022, No. 2
DOI: https://doi.org/10.35219/ann-ugal-math-phys-mec.2022.2.10

Rapid detection of milk adulteration using Raman spectroscopy and statistical modelling

Anca Becze [1]*, Dorina Simedru[1]

[1] *INCDO-INOE2000, Research Institute for Analytical Instrumentation, ICIA Cluj-Napoca Subsidiary, 400293 Cluj-Napoca, Romania*
* *Corresponding author: anca.naghiu@icia.ro*

Abstract
Food adulteration has become a concern for consumers and food safety authorities. Milk is a commune adulterated food product, like melamine adulteration, which resulted in devastating effects, especially on young children. Because of the current fast paste economy, it is essential to develop equally fast analysis methods to ensure reliable and sensitive results quickly with little to no sample preparation. For that purpose, a Raman method was developed and Partial least squares regression (PLS) was applied in order to develop a model for adulterated goat milk detection. Minitab 17 software was used for the statistical modeling of data. Validation matrices were constructed using unadulterated goat milk and goat milk adulterated with cow milk in different proportions (0-50%). The prediction model had a correlation coefficient of 99.8 %.
Keywords: milk, food adulteration, Raman spectroscopy, Minitab, statistical modelling, PLS

1. INTRODUCTION

While cow's milk is the leading milk used by dairy factories in the EU, in several Member States, other milk contributes significantly to milk production. In 2018, Spain produced 1.0 million tons of milk from sheep and goats, while Greece and France both produced 0.8 million tons. Italy has also produced 0.7 million tons of milk from animals other than cows, including almost all EU buffalo milk production. A majority (57.1%) of milk delivered to dairy factories in Greece came from sheep and goats in 2018. Just over one-fifth (22.5%) of milk delivered to dairy factories in Cyprus comes from sheep and goats [1,2].

There is no harmonized definition in the EU for the "food fraud" term ("fraudulent or deceptive practices" mentioned in Reg 178:2002 art 8). However, it is widely accepted that an intentional breach of EU law for economic gain by deceiving the consumer is food fraud. Adulteration is a legal term used to define foods that do not meet legal requirements. Adulterated foods are a risk to consumers through several side effects on the human organism and have a negative impact on the agro-food market [3-5].

Milk and dairy adulteration has come to global attention after the discovery of melamine contamination in dairy products in China in 2008. However, the history of milk adulteration is ancient. In 1850, a scandal involving the adulteration of milk was reported, which resulted in the deaths of 8,000 children in New York alone [3].

Due to the increasing incidence of food adulteration cases, developing rapid, robust, and selective analysis methods is essential to ensure the practical testing of many samples. Various analytical techniques have been used to detect milk and milk product adulteration, each of them having advantages and disadvantages. These techniques usually focus on the determination of fat of foreign substances in the milk and usually need long and laborious methods for sample preparation [4-8], like fat determination using chromatography [9].

This study, which is part of the national project “*Development of methods for determining food adulteration*” aims to develop quick and easy methods to determine the adulteration of goat milk using portable Raman equipment and the PLS statistic method for developing the model equation (18N/08.02.2019),

2. EXPERIMENTAL

2.1 Milk samples

In order to streamline the method and due to the differences that the milk composition presents, it is important to have representative samples. This is why cow and goat milk samples were obtained from 3 different farms in the Cluj County area and 2 farms from Alba County. The 5 samples were analyzed in pure form and then whey were analyzed after adulterated with cow's milk, in proportion of 0-50%.

2.2 Raman analysis

A PROGENY portable Raman spectrometer, from Rigaku, Japan, was used for the analysis of the milk samples.

No sample preparation was needed prior to Raman analysis. The samples were placed in a 4 ml bottle which was fixed in the special vial holder of the spectrometer that allows the proximity of the sample to the laser source to set between 0-5 mm (Fig. 1).

Fig. 1 Raman milk sample analysis

The analysis method has the following working parameters: • Laser frequency: 1064 nm; • Laser power: 490 mW; • Exposure time: 6000 ms;• Spectral range: 200-2000 cm-1 •Detector: InGaAs type with cooling; • Accessories: conical tip with adjusted focus and bottle adapter of different sizes.

2.3 PLS regression

Partial least squares regression (PLS) was used to model the equation for the concentration of adulterated goat milk. To perform PLS, Minitab software uses the partially non-linear iterative algorithm (NIPALS) developed by Herman Wold. The algorithm reduces the number of predictors using a technique similar to principal component analysis to extract a set of components that have a maximum correlation between predictors and response variables. For modeling the equation, 4 components have been used, namely the peak areas from ~818 cm^{-1}, ~1300 cm^{-1}, ~1441 cm^{-1}, and ~1905 cm^{-1}, based on data previously obtained [10].

3. RESULTS AND DISCUSSION

Table 1 shows the average results obtained for the 5 goat milk sample that was adulterated with cow milk in different proportions.

Table 2 shows the variation of the experimental data. The degree of freedom (DF) represents the number of independent comparisons between the elements of an observation quantity or the number of values that can be chosen arbitrarily within a specification. DF is used to estimate the values of the parameters of the unknown population. Adjusted amounts of squares (Adj SS) are measures of variation for different components of the model. Adjusted Mean Squares (Adj MS) measure how much variation a term or pattern explains, assuming that all other terms are in the pattern, regardless of the order in which they were entered. The value F is used to determine if the term is associated with the answer [11]. Minitab uses the value of F to calculate the value of p, which is used to make a decision about the statistical significance of terms and model. The p-value is equal to the area under the distribution curve, depending on the statistical value of the test. Values should be less than 0.05.

Table 1. Raman analysis

Crt. No.	**Adulteration percentage (%)**	**Maximum intensity (A.U.)**			
		818 cm^{-1}	**1300 cm^{-1}**	**1441 cm^{-1}**	**1905 cm^{-1}**
1.	0.0	426.12	1,101.10	2,105.53	603.56
2.	2.5	483.48	931.24	1,823.15	587.05
3.	5.0	517.65	844.17	1,760.62	551.13
4.	10.0	535.26	730.16	1,583.97	430.74
5.	25.0	649.19	641.50	1,484.27	400.31
6.	50.0	701.58	637.68	1,160.52	379.46
7.	100.0	958.33	622.79	1,021.08	316.23

Table 2. Analysis of the variation of the adulteration percentage

Crt. no.	**Source**	**Degree of freedom**	**The sum of the squares (Adj SS)**	**The value of the average squares (Adj MS)**	**F Distribution**	**P value**
1.	Regression	40	0.796061	0.199015	2107.96	0.00012
2.	Residual error	2	0.000189	0.000094		
3.	**Total**	**6**	**3333.33**			

The relationships between the answer and the variables in the model are significant from a statistical point of view because the p value is 0.00012, which is below the established significance level of 0.05.

R2 is the percentage variation of the response that is explained by the model. It is calculated as 1 minus the ratio between the sum of the errors of the squares and the total sum of the squares, practically R2 indicates how well a model fits with the data [11]. To vary the model, R2 was calculated. (Table 3). The predicted R2 is calculated by systematically removing each observation from the data set, estimating the regression equation and determining how well/correctly the model predicts the removed observation. The best prediction was observed for model that used at least 2 components for the prediction equation. The degree of error is lower the more components (peaks at different wavelengths) are used.

Table 3. PLS model selection and validation

Component number	**X variation**	**Error**	**R^2**	**Predicted residual sum of squares**	**Predicted R^2**
1	0.92543	0.149599	0.812121	0.419309	0.473395
2	0.98154	0.001669	0.997904	0.009041	0.988645
3	0.98875	0.000239	0.999700	0.004723	0.994068
4	1.00000	0.000189	0.999763	0.002643	0.996681

The 4 components of the PLS models are the 4 peaks at different wave lengths (~818 cm^{-1}, ~1300 cm^{-1}, ~1441 cm^{-1}, and ~1905 cm^{-1}). The correlation between the different variables regarding the adulteration of goat's milk is presented in Figure 2.

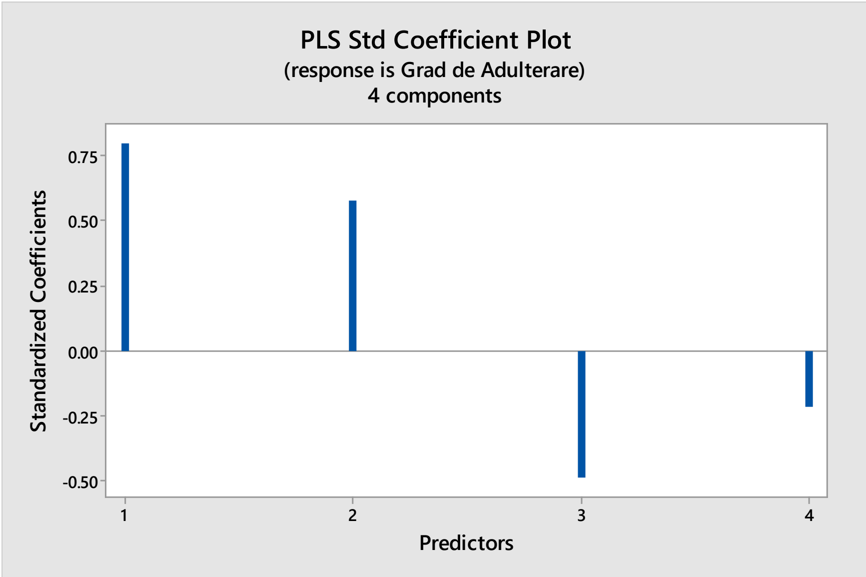


Figure 2. Scheme of non-standard regression coefficients for adulteration of goat's milk

It can be observed that the peaks at ~818 cm^{-1}, ~1300 cm^{-1} have a positive correlation with the degree of adulteration, while the peaks at ~1441 cm^{-1}, and ~1905 cm^{-1} have a negative correlation. It can be also observed that the peak at 818 cm^{-1} has that highest relationship of predictors and responses (0.75) while the ~1905 cm^{-1} has the lowest (-0.20).

The evaluation of the PLS model obtained after data processing using the Minitab 17 software are presented in Figure 3. Small differences in fitted and cross-validated values identify a high correlation of the model.

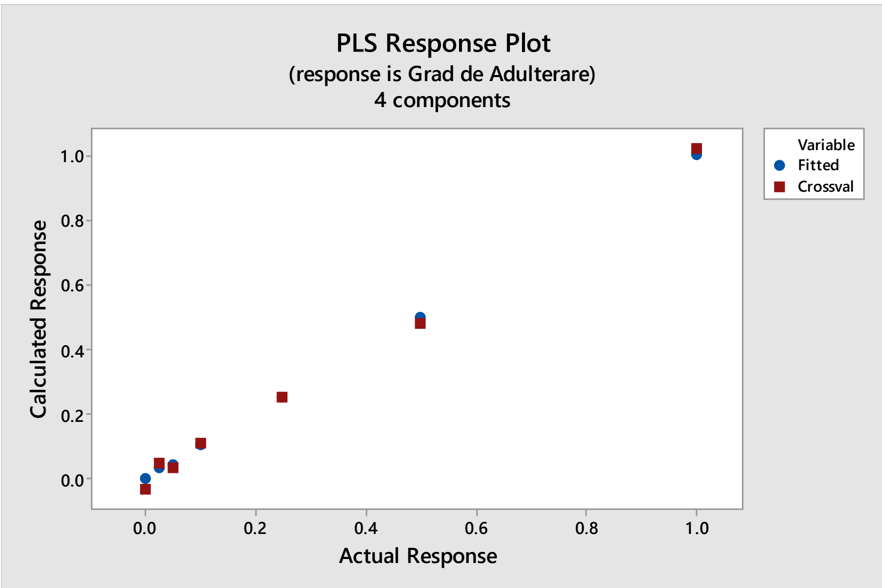
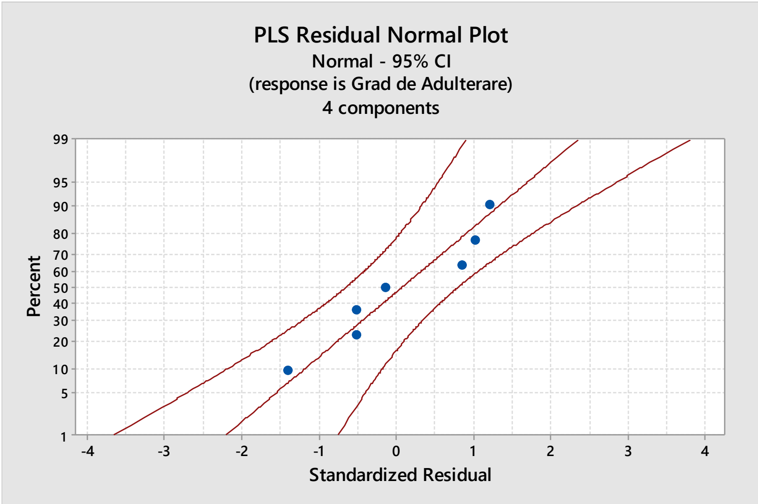


Figure 3. PLS response graphic for adulteration percentage

The normal probability scheme of the residues displays the residues compared to their expected values when the distribution is normal (Fig. 4). The normality of the residuals is proven because the points follow a straight line.

For the realization of the model equation only the values of the maxima from 818 cm^{-1} and 1441 cm^{-1} were used (Fig. 4).



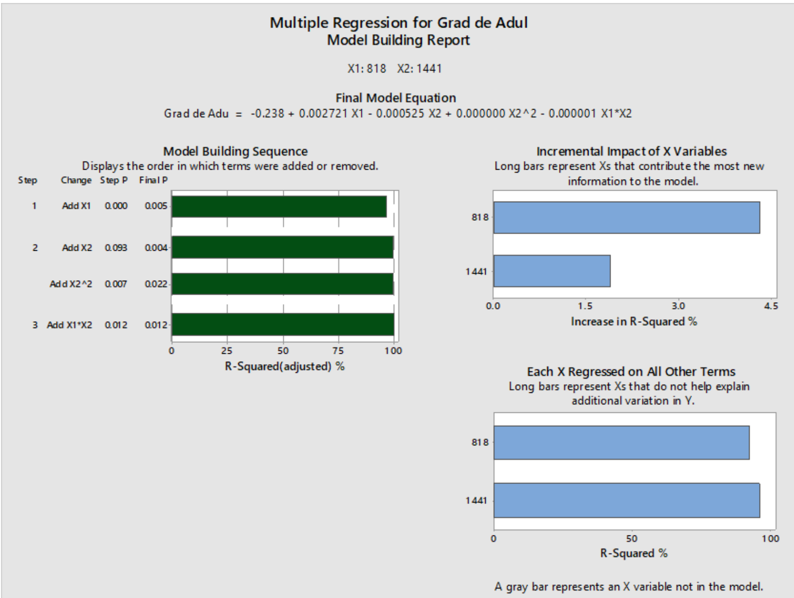


Fig. 4. The report on the creation of the equation for predicting the degree of goat milk adulteration

The final equation of the model is:

$$DA\ \%= -0.238+0.002721\cdot X1-0.000525\cdot X2-0.000001\cdot X1\cdot X2 \quad (1)$$

Where:
DA – degree of adulteration (%);
X1 – area of the peak at 818 cm^{-1} (AU);
X2- area of the peak at 1441 cm^{-1}(AU).

4. CONCLUSIONS

PLS demonstrated that 2 factors account for most of the variation in the response. The cross-validation analysis confirms with the PRESS values being the lowest 0.009.

Multiple regression analysis showed that the maximum from 818 cm^{-1}, has the greatest impact on the variation of the degree of adulteration of goat's milk. The relationships between the response and the variables in the model are statistically significant (p value <0.05). Multiple regression analysis showed that the maxima of 818 and 1441 cm^{-1} have the greatest impact on the variation of the degree of adulteration of goat's milk. The variation of the Y answer can be explained in a very large percentage by the chosen variables.

The model fits well with the data, which shows that the equation can be used to predict the degree of adulteration of goat's milk with cow's milk.

ACKNOWLEDGMENTS

This research was funded by the Romanian Ministry of Research and Innovation, CCCDI-UEFISCDI, project number-18N/08.02.2019/PNCDI III.

References

1. Eurostat, https://ec.europa.eu/eurostat/statisticsexplained/index.php/Milk_and_milk_product_statistics# Milk_production, accessed on 16.11.2022;
2. Eurostat, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Agriculture,_forestry_and_fishery_statistics; accessed on 30.05.2020;
3. Balan B., Dhaulaniya S. A., Jamwal et R.al., "Application of Attenuated Total Reflectance-Fourier Transform Infrared (ATR- FTIR) spectroscopy coupled with chemometrics for detection and quantification of formalin in cow milk," Vibrational Spectroscopy, vol. 107, article 103033, (2020).
4. Elainy Virginia dos Santos Pereira, David Douglas de Sousa Fernandes, Mário César Ugulino de Araújo, Paulo Henrique Gonçalves Dias Diniz, Maria Inês Sucupira Maciel, Simultaneous determination of goat milk adulteration with cow milk and their fat and protein contents using NIR spectroscopy and PLS algorithms, LWT, Volume 127 (2020) 109427, ISSN 0023-6438, https://doi.org/10.1016/j.lwt.2020.109427.
5. Elainy Virgínia dos Santos Pereira, David Douglas de Sousa Fernandes, Mário César Ugulino de Araújo, Paulo Henrique Gonçalves Dias Diniz, Maria Inês Sucupira Maciel, In-situ authentication of goat milk in terms of its adulteration with cow milk using a low-cost portable NIR spectrophotometer, Microchemical Journal, Volume 163, (2021) 105885.
6. Piras, C., Hale, O.J., Reynolds, C.K. et al. Speciation and milk adulteration analysis by rapid ambient liquid MALDI mass spectrometry profiling using machine learning. Sci Rep 11, 3305 (2021). https://doi.org/10.1038/s41598-021-82846-5.
7. Lingyu Li, Jun Wang, Minjing Li, Yingchun Yang, Zhixuan Wang, Jing Miao, Zhongkai Zhao, Jie Yang, Detection of the adulteration of camel milk powder with cow milk by ultra-high performance liquid chromatography (UPLC), International Dairy Journal, 121,(2021) 105117, ISSN 0958-6946.
8. Guangzao Huang, Lei-ming Yuan, Wen Shi, Xi Chen, Xiaojing Chen, Using one-class autoencoder for adulteration detection of milk powder by infrared spectrum, Food Chemistry, 372, (2022) 131219, ISSN 0308-8146, https://doi.org/10.1016/j.foodchem.2021.131219.
9. Rayana A. Costa, Camilo L.M. Morais, Thalles R. Rosa, Paulo R. Filgueiras, Monike S. Mendonça, Isabelly E.S. Pereira, Bruno V. Vittorazzi, Marisa B. Lyra, Kássio M.G. Lima, Wanderson Romão, Quantification of milk adulterants (starch, H2O2, and NaClO) using colorimetric assays coupled to smartphone image analysis, Microchemical Journal, Volume 156, (2020), 104968, ISSN 0026-265X,https://doi.org/10.1016/j.microc.2020.104968.
10. Becze A., Naghiu M. O., Babalau L. V. Fuss, Cepoi L., Simedru D., AGRIFA Int. Conf., 2020, 11. https://icia.ro/wp-content/uploads/2020/10/Poster-8.pdf
11. MINITAB, Available online: https://support.minitab.com/en-us/minitab-express/1/help-and-how -to/modeling-statistics/anova/how-to/two-way-anova/interpret-the-results/all-statistics-and-graphs/