# Fraud detection in hen housing system declared on the eggs' label: An accuracy method based on UV-VIS-NIR spectroscopy and chemometrics

Gema Puertas, and Manuel Vázquez*

Department of Analytical Chemistry, Faculty of Veterinary, University of Santiago de Compostela, 27002-Lugo, Spain

* Corresponding author. Email: manuel.vazquez@usc.es

ABSTRACT

EU regulation classifies egg production in four hen housing systems: organic, free range, barn and cages. However, there are no analytical methods for a complete detection of the housing systems declared on the eggs' label. This work describes a new method for this task. A lipid extract was obtained, and eggs classified in the four housing systems by UV-VIS-NIR spectroscopy and chemometrics. A total of 192 spectra were analysed. Eggs were classified by Support Vector Machine classification and Linear and Quadratic Discriminant Analysis (LDA and QDA, respectively). The cholesterol concentration did not allow to classify correctly the four farming systems. However, the yolk lipid extract successfully classified the eggs of the hen housing system. Results showed 100% accuracy by UV-VIS-NIR spectrum of the yolk lipid extract with the QDA statistical analysis. These results indicate that this lipid extract is a promising tool for analytical verification of the farming system.

Keywords: egg; fraud; cholesterol; farming methods; organic production; chemometrics, spectroscopy

## 1. Introduction

Fraud in food trading has been common since ancient times. Financial reasons or to camouflage the bad appearance and taste of rotten food were the initial motivations. Through the ages, fraudulent practices have been refined and become more sophisticated, following or anticipating advances in the analytical sciences (Oliveri & Downey, 2012). In addition, due to the globalization and expansion of food supply chains, the prevalence and impact of food fraud have increased over recent years (Johnson, Sidwick, Pirgozliev, Edge, & Thompson, 2018). Food fraud is a collective term that includes the deliberate and intentional substitution, addition, tampering, or misrepresentation of food, food ingredients, or food packaging; or false or misleading statements made about a product for economic reason (Spink & Moyer, 2011). In the last year, it has been published several cases of misrepresentation of the hen farming method on the eggs' label (Joint Research Centre, 2018c, 2018b, 2018a). Although several techniques have been developed and used to detect this kind of egg fraud, there is still much to be done regarding its detection.

Eggs are consumed worldwide. They are well known as a source of vitamins, minerals, phospholipids and high-quality proteins. They are included in a healthy diet. For instance, eggs were included in the 87% of the available European food-based dietary guidelines (Montagnese et al., 2015). In the European Union, eggs are produced by hens of the species *Gallus gallus*. The introduction of the Council Directive 1999/74/CE establishes the minimum standards for the welfare protection of laying hens in free range, barn and cage housing systems. In addition, the Commission Directive 2002/4/EC set four codes for the farming systems: 0- for organic production, 1- for free range, 2- for barn and 3- for cages. The Commission Regulation 589/2008 (EC) describes detailed rules on marketing standards for eggs. It sets the four types of eggs that can be found in the European market according to the hen housing system cited above.

Organic eggs are produced according to the requirements of organic production. This method has specific principles that include the observance of a high level of animal welfare, feed ingredients grown without synthetic fertilizers and no products from genetically modified organism (European Council, 2007). Organic systems are a specific form of free-range hens where they also have access to a pasture. They are kept in no more than 6 hens per $m^2$ usable area (European Commission, 2008).

Free-range eggs are from hens that must have continuous daytime access to open-air runs. This space must be mainly covered with vegetation and the maximum stocking density is one hen per 4 $m^2$ at all times.

Barn eggs are produced in systems that required the same conditions as free range except for the open-air runs and the maximum of 9 hens per $m^2$ usable area. They are from hens that are kept in houses that might have multiple aviaries with nest boxes and perches, and the floor is covered with litter material.

Eggs from caged hens are from hens rearing in enriched cages and they must have at least 750 $cm^2$ of cage area per hen. Intensive systems have facilitated the production of animal-based products at relatively low prices.

However, they have been increasingly considered to be responsible for a dramatic reduction in animal welfare, as indicated by the high prevalence of brittle bones in hens and lameness in broilers (Napolitano, Serrapica, & Braghieri, 2013).

Due to differences in stocking density, feed costs and consumption, mortality and productivity, the production costs are lower in cage and barn systems, intermediate in free-range systems and higher in organic systems (Leenstra et al., 2014). Nowadays population is concerned about animal welfare. In most European countries, organic and free range production are increasing compared to other systems due to consumer preferences. Consumers in developed countries are demanding high-quality food, produced under more environmentally friendly conditions. Then, producers face the opposing demands of the need to increase food production while reducing ecological impact of intensive production methods (Abín, Laca, Laca, & Díaz, 2018).

Furthermore, in Switzerland, non-cage systems for layers are obligatory since 1992 and in Spain some big supermarket chains have decided to stop selling eggs from caged hens. These circumstances could force producers to shift to an extensive production method. But, as it was reported at the beginning, it has also claimed the attention of fraudsters.

As the demand of extensive production method has increased, the potential financial gain is higher for fraudsters. However, the risk of getting caught is low as the current identification method is purely administrative. Testing the traceability is the available technique to determine where the egg is from and what kind of hen houses are in the farm of origin. A comprehensive literature search on techniques for verifying egg authenticity showed that there are already described procedures to discriminate organic eggs from conventional eggs. Techniques based on carotenoids profiling (Van Ruth et al., 2011; Van Ruth et al., 2013), fatty acid composition (Samman et al., 2009; Tres, O'Neill, & Van Ruth, 2011) or mineral content (Borges et al., 2015; Küçükyilmaz et al., 2012; Vincevica-Gaile, Gaga, & Klavins, 2013) were developed. However, the current risk is not only fraud with organic eggs. Because organic production has a robust legislation that dissuade fraudsters. Besides, in some countries the consumer prefers free-ranges eggs over organic eggs (Zakowska-Biemans & Tekień, 2017), and these rouse fraudsters' interest on free-range products. Other relevant studies discriminated 1-, 2- , 3- coded eggs through fluorescent patterns on egg surface or stable nitrogen isotope composition (Gregory, Gepp, & Babidge, 2005). Furthermore, eggs from caged hens were differentiated from the others based on physical and chemical egg features such as protein content or foam consistency (Hidalgo, Rossi, Clerici, & Ratti, 2008). Nevertheless, none of these methods allowed a complete classification of the four legal farming methods in UE.

UV-VIS spectroscopy is widely implemented in most food quality control laboratories and it is useful as an electronic eye for correlating visual parameters to specific geographical regions or food varieties. Although this technique has some limitations for characterization and authentication purposes, its combination with NIR and/or MIR spectroscopy can enhance the results (Borràs et al., 2015). NIR spectroscopy is an analytical technique, which has increasing popularity in the food processing industry due to its advantages. It is a non-destructive, environmental friendly and rapid technique capable for on-line application with low running costs (Porep, Kammerer, & Carle, 2015).

In our study, UV-VIS-NIR spectroscopy was applied. This technique has already been used to prevent food fraud. For instance to discriminate wines from controlled designation of origin (Martelo-Vidal, Domínguez-Agis, & Vázquez, 2013), or typify vinegar from protected designation of origin" (Ríos-Reina, García-González, Callejón, & Amigo, 2018).

The cholesterol is contained in the egg yolk. An average of 10.85 mg of cholesterol are obtained from 1 g of egg yolk (Puertas & Vázquez, 2018). Due to its relevance presence in egg and the willing to detect whether there is any influence of the farming system on the amount of cholesterol, the lipid extract with cholesterol was chosen as target in this study.

Thus, the aim of this study was to classify eggs in the four farming systems by UV-VIS-NIR spectroscopy and chemometric tools as a quick method to detect fraud. From our knowledge, this is the first study developing a quick, easy and practical method to classify eggs for the four farming systems. As secondary, it was studied if differences in the cholesterol concentration of eggs due to the hen housing system exist.

## 2. Materials and methods

### 2.1. Raw materials

The eggs (n = 48) were bought from local supermarkets or supplied by Granja Campomayor (Lugo, Spain). For each farming system two different farms were analysed. From each farm, six eggs were employed. Each sample was analysed in duplicate. Fresh eggs yolks were separated manually from egg whites. The albumen residuals were eliminated from the yolk using a blotting paper and the vitelline membrane was removed with a scalpel.

### 2.2. Lipid extraction and cholesterol quantification

All yolk samples were homogenised at 5600 rpm for 1 min using a high performance homogenizer (Ultra Turrax®, IKA, Staufen, Germany). Afterwards, 0.5 mg was weighted for saponification and further quantification. Cholesterol esters hydrolysis was based on previous studies. It was demonstrated that 30 min saponification at 80 ºC is not enough time to hydrolyse sterol esters and released all free cholesterol (Hansen & Wang, 2015). Therefore, it was increased incubation time to 3 h at 80 ºC in a water bath. Besides, the amount of solvents was enlarged to 30 ml of Isopropanol 99% for spectroscopy analysis and 20 ml of 1 M methanolic potassium hydroxide solution. This solution was freshly prepared with methanol and potassium hydroxide (Thermo Fisher Scientific, Waltham, MA, USA). After saponification, samples were cooled at room temperature, filtered through 0.45 μm membrane filter and the clear solution was employed for cholesterol quantification and UV-VIS-NIR spectroscopy.

Cholesterol was quantified by enzymatic kit method (Enzytec™, R-Biopharm AG, Darmstadt, Germany). The enzymatic kit was employed following manufacturer instructions.

### 2.3. Spectral analysis

Samples were directly measured after filtration for UV-VIS-NIR spectroscopy. Spectral measurements were performed in a V-670 spectrophotometer (Jasco Inc., Hachioji, Tokyo, Japan). UV-VIS-NIR spectra were acquired in transmittance mode (T) at 2 nm intervals in the region range of 190 to 2500 nm using a cuvette with a 1 mm path length. The ultraviolet (UV) region was considered from 190 to 380 nm. The visible (VIS) light ranging extended from 380 to 780 nm. And the near infrared (NIR) region covered wavelengths from 780 nm up to 2.5 μm (Porep et al., 2015). Each sample was measured twice. Thus, a total of 192 spectra were obtained. The spectral data were collected with Spectra Manager™ II software (Jasco Inc., Hachioji, Tokyo, Japan).

**2.4. Data processing and statistical analysis**

Tukey HSD test was performed on the cholesterol content data obtained with the enzymatic kit. UV-VIS-NIR spectra were exported from Spectra Manager™ II software to Unscrambler® software Version 10.5 (Camo, Oslo, Norway) for pre-processed treatments and chemometric analysis.

On the spectra, mean centering (MC) and standard deviation scaling (SDS) was performed. In addition, two other pre-treatments were tested: Savitzky-Golay Smoothing (7 points) and second polynomial order (SGS) to reduce random noise and Standard Normal Variate (SNV) to correct for baseline variations due to the different scattering of the samples (Martelo-Vidal et al., 2013; Ríos-Reina et al., 2018).

Principal component analysis (PCA) was carried out prior to any classification approach. It performs a reduction in the data dimensionality to facilitate the visualization of the data, retaining as much as possible the information present in the original data (Tres & Van Ruth, 2011). It is an exploratory analysis of the data to detect outliers, recognize patterns in samples distribution and relationships between variables and classes (Ríos-Reina et al., 2018).

The classification methods used in our work to spectra were Support Vector Machine classification (SVM) and Linear and Quadratic Discriminant Analysis (LDA and QDA). SVM is a pattern recognition method supervised, because the number of categories and samples in each category are previously defined. It defines a function that describes the limit to separate classes by maximising the distance between them. SVM can work with linear and non-linear multivariate analysis. This method employs linear equations instead of quadratic programming to obtain the support vectors. This method is effective for modelling non-linear data and uses an iterative training algorithm to achieve the separation of different classes (Martelo-Vidal et al., 2013). It aims to construct an optimal separating plane by mapping the original data points from the input space into a high-dimensional feature space. The distances from all the data points to the separating plane are minimum. The kernel function, which performs the nonlinear mapping, plays an important role in the procedure. In SVM, radial basis function (RBF) was adopted as the kernel function of SVM classification because of its effectiveness and speed in calibration process (Yang et al., 2018).

LDA and QDA are also supervised classification techniques. LDA maximizes the between-class variance over the within-class variance, in order to create a linear decision boundary between them (Carvalho et al., 2018). The boundary produce by QDA is a quadratic curve, which, unlike LDA, may consist of two separate sections of boundary lines. This means it is thus able to correctly classify samples from a class even if they lie in separate regions in variable space. This is due to the difference in the variance-covariance matrix. LDA uses the pooled variance–covariance matrix and does not take into account different variance structures for the classes, whereas QDA forms a separate variance model for each class (Dixon & Brereton, 2009). In order to reduce the thousands of variables that have been recorded from the spectra and to allow the calculation of the inverse of the variance–covariance matrix for LDA and QDA, PCA is employed. A key feature of this technique is its ability to reduce large multivariate data matrices down to a few orthogonal principal components (PC), which still contain the majority of the information held in the original raw data (Dixon & Brereton, 2009). The number of PC would still need to be less than the number of objects in each class. It permits the determination of the best-fit parameters for classification of samples by a developed model (Martelo-Vidal et al., 2013). The results of this classification methods were validated based on segmented cross validation.

## 3. Results

3.1 Classification by cholesterol concentration in yolk.

As a first approach to detect differences in the hen housing system, cholesterol was quantified for each type of farming system. Figure 1 shows the results expressed on mg of cholesterol per gram of egg yolk. The low average cholesterol content was obtained in the eggs of hens from the housing system 3- cages (13.64 ± 0.650 mg/g) and the highest for the eggs of hens from 2- barn (14.78 ± 1.175 mg/g). Intermediate values were obtained for 0- organic production (14.28 ±1.646 mg/g) and 1- free range (14.09 ± 0.825 mg/g). These results are higher than the reference values but in line with the article on which the sample pretreatment was based on (15.3 mg/g) (Hansen & Wang, 2015) and our previous results (13.92 mg/g) (Puertas, Cazón, & Vázquez, 2018). Although average values were different for the four housing systems, the Tukey HSD test showed only a statistical significant difference in cholesterol between eggs from hens housing system of 2- barn and of 3- cages (p-value < 0.01). These results disagreed with the studies where the cholesterol level in organic eggs was found higher than in conventional eggs (Minelli, Sirri, Folegatti, Meluzzi, & Franchini, 2007; Mugnai et al., 2014). However, they are in concordance with those where cholesterol content of free-range and cage-produced eggs were compared and no significant difference was found (Anderson, 2011). Our results showed that the amount of cholesterol could be a discriminant parameter to detect frauds only on eggs´s labels of barn eggs and eggs from caged hens. Then the amount of cholesterol in egg yolk cannot be a single marker to discriminate between the four farming methods. However, cholesterol extracts can contain other lipophilic compounds.

3.2 Classification by spectral analysis of yolk lipid extract

The whole lipid extract could be a complex marker to discriminate between the four farming methods. The spectral analysis could discriminate between each housing system. Figure 2 shows the flowchart of the proposed procedure to classify by the four housing systems by UV-VIS-NIR spectral analysis.

The average UV-VIS-NIR spectrum of each housing system is shown on Figure 3. Absorbance is higher at the NIR range and greater variability is detected in the VIS spectral data. To clearly visualize it, in Figure 3 it was zoomed from 780 to 380 nm wavelength. Transmittance is slightly minor for free raged eggs. PCA analysis revealed that 93% of the raw data was explained with 2 components. Results were improved up to 96 % when PCA was applied to SNV data.

SVM was firstly applied over the full UV-VIS-NIR spectra raw and pretreated. Table 1 shows the results of the classification. The percentage of properly classified samples is indicated in each classification analysis. The best training and validation accuracies obtained were achieved with MC and SDS pretreatments. Even it did not get 100% accuracy for the four types of samples, it classified properly more than 95 % of the organic and barn eggs. In contrast, with raw data, using SNV or SGS pre-treatments only barn egg samples were classified with 100% accuracy. A spectrum range analysis was performed with MC and SDS pretreatment. Results are shown on Table 1. Using alone the UV range, the accuracy was better than using the NIR or VIS ranges alone. Better results were obtained when the spectra ranges were wider, like UV-VIS or VIS-NIR ranges. Nevertheless, the best method accuracy was achieved using the full UV-VIS-NIR spectra analysis.

Discriminant analysis was also applied. Results from LDA and QDA classification method with raw and pretreated UV-VIS-NIR spectral data are shown in Table 2. PCA is previously applied with the goal of reducing the data matrix to explain the maximum variance with the fewest number of principal component (Dixon & Brereton, 2009). Results are disclosed for 16 PC because it was the minimum number of PC to obtain 100 % accuracy. That accuracy was achieved with QDA. While LDA best result was 94.8 % accuracy with 16 PC and SNV pretreatment.

In order to detect differences between the previous data treatment for QDA, it was studied the minimum number of components to obtain more than 95% accuracy with each. MC & SDS did not achieve 100% accuracy with 16 PC (Table 2). Results on Table 3 includes the percentage of properly classified samples in each farming method. None difference was detected between raw and SGS data. However, SNV data obtained more than 95 % accuracy with only 7 PC. That is in accordance with the literature because SNV is one of the most suitable methods to reduce within-class variability (Wu et al., 1996).

QDA was performed with SNV data at different spectrum ranges and compared with raw data. From each region of the UV-VIS-NIR spectrum, it was obtained the minimum number of components to obtain more than 95% accuracy, the percentage of samples that were properly classify in each farming method and the number of PC needed for 100% accuracy. Results are shown in Table 3. It was observed that the accuracy was improved when reducing the range, in contrast to SVM method. Better results were obtained using the VIS and UV ranges, rather than using NIR range. It was also remarkable that raw data obtained better results than SNV data at VIS and UV ranges. Those results are explained due to the number of variables that each region contains. The whole spectrum contains 2310 variables (waveleghts), NIR range contains 1720 variables, whereas the VIS contains 400 and the UV only 190 variables. As the variables are reduced the data pre-processing changed. Previous data pretreatment with SNV did not produce better results because the non-relevant variability had been also reduced with the decrease of the number of variables (Wu et al., 1996). Raw data obtained slightly better results when analysing the percentage of eggs correctly predicted using the VIS range. Barn eggs were classified with 100 % accuracy using VIS range with raw data and nearly 98 % with SNV data. Succeeding data reduction with PCA needed less PC to preserve the information caused by the main sources of data variability (Wu et al., 1996). Like that when the VIS or UV ranges were employed for classification, only 9 and 12 PC, respectively, were necessary to infer the farming method with 100 % accuracy.

Then VIS range for QDA employed more variables, less PC and obtained higher accuracy. Subsequently, hen farming method can be inferred with 100% accuracy when VIS spectroscopy and QDA, as statistical method, are employed.

## 4. Discussion

Each egg unit has a unique composition, which is distinctive due to the intrinsic metabolism of each hen that originated it (de Oliveira Mendes, Porto, Almeida, Fantini, & Sena, 2019). However, described factors that modify egg size and composition are: diet, hen’s age, strain and other environmental factors (Samman et al., 2009). In this work, hen’s age and strain could not be controlled because it was a consumer survey of products available in supermarkets and, as such, specific characteristics of the laying hens are not known (Samman et al., 2009). Diet was considered through the rearing system. As housing conditions would definitely control feed consumption and conversion. For example, organic and free range hens have access outdoors. They are exposed to sunlight,and they are expected to have access to other food sources (such as insects or seeds from vegetation) (Basmacioğlu & Ergül, 2005; Tres et al., 2011). On the basis of this assumption, housing system would be expected to influence the composition of eggs. In addition, it has been demonstrated that changes in dietary lipids influenced cholesterol and fatty acid composition much more than the strain (González-Muñoz et al., 2009; Washburn, 1979). Then this lipid extract was chosen to infer the housing system because of their capacity to reflect hen feed over strain.

A technique based on carotenoids profiling obtained success classification rates for organic, free range, and barn eggs respectively (n = 48) with right classification of 100 %, 100 %, and 84 %, respectively (Van Ruth et al., 2011). In other study with the same technique, sample size covered Europe (n=69) and non-EU farms (n = 25). Discriminating organic eggs from conventional eggs, there were reached classification rates up to 98% and 72%, respectively (Van Ruth et al., 2013).

When analyzing fatty acid composition, it was detected that organic eggs had higher percentage of palmitic and stearic acids, but no differences in mono and polyunsaturated fatty acids were detected (Samman et al., 2009). A classification technique employing fatty acid fingerprints and chemometric tools was able to identify organic eggs with 92 % accuracy and conventional eggs with 87.5 % (n = 48) (Tres et al., 2011). Same classification, organic *versus* conventional eggs, was improved to 100% accuracy through mineral content fingerprints and chemometrics (n = 60) (Borges et al., 2015). These results are justify with the significant differences in macro and trace element composition found between organic and conventional eggs (Küçükyilmaz et al., 2012; Vincevica-Gaile et al., 2013). Other parameters were studied to differentiate inside the conventional eggs group.

A method based on probabilities in the analysis of fluorescent patterns on egg surface was developed to discriminate 3- coded eggs from 1- and 2-coded. It was detected that if five or more eggs in a sample of 90 eggs had double fluorescent lines, it could be concluded with 99,9% probability that the batch contains some cage-laid egg (Gregory et al., 2005). All these methods only allowed discriminating two or three of the four legal rearing systems in UE.

Scarce previous works tried to find differences between the four hen housing systems by eggs. When physical and chemical egg features were analysed only cage eggs were characterised with 100% accuracy (n = 28) (Hidalgo et al., 2008). Stable isotopes were also study to classify eggs. Nitrogen isotope composition of eggs was reported as a promising indicator to differentiate eggs from caged hens, barn hens and free-range hens (n = 36) (Rogers, 2009). However, theses authors did not report accuracy values.

In our study, the four legal farming methods were correctly classified. The cholesterol concentration was not useful for this goal. However, the UV-VIS-NIR spectra of the lipid extract successfully classified the hen housing system with 100% accuracy. The lipid extract contained more components than cholesterol, for example fatty acids and carotenoids. As it has been studied for human consumption, the lipid matrix of the egg yolk enhanced the bioavailability of valuable carotenoid pigments, including lutein and zeaxanthin (Zaheer, 2017). Furthermore, the profile of carotenoids, fatty acids and cholesterol in egg yolk is highly dependent on the hens' diet (González-Muñoz et al., 2009; Zaheer, 2017). Then their values can significantly change with the hen farm system. Thus, these facts can justify the excellent classification (100% accuracy) of the hen housing system obtained with the UV-VIS-NIR spectra of egg lipid extracts. Further research with eggs from different countries would make this method robust to be regulatorily implemented to avoid fraud in the eggs´ label information following the same strategy of other studies (Van Ruth et al., 2013).

## 5. Conclusion

The method described in this work is the first analytical technique that classifies eggs according to the four legal farming methods in UE by UV-VIS-NIR spectroscopy. Results showed that it is feasible to classify the hen housing system with 100% accuracy employing the lipid extract of egg yolk by UV-VIS-NIR spectrum and QDA statistical analysis.

The best results for reducing PC numbers were obtained using the VIS region. The method proposed is quick, easy and cheap because only minimal laboratory equipment is needed. It holds great potential for classification through reliable analytical system. Further research with eggs from different countries would make this method robust to be regulatorily implemented to avoid fraud in the eggs´ label information.

## 6. Acknowledgements

This work was supported by the Spanish National Plan for Scientific and Technical Research and Innovation. A University Teacher Education grant (FPU 16/05128) by the Spanish Ministry of Education, Culture and Sport to author Gema Puertas is gratefully acknowledged.

## REFERENCES

Abín, R., Laca, A., Laca, A., & Díaz, M. (2018). Environmental assesment of intensive egg production: A Spanish case study. *Journal of Cleaner*

*Production*, *179*, 160–168. https://doi.org/10.1016/j.jclepro.2018.01.067

Anderson, K. E. (2011). Comparison of fatty acid, cholesterol, and vitamin A and E composition in eggs from hens housed in conventional cage and range production facilities. *Poultry Science*, *90*(7), 1600–1608. https://doi.org/10.3382/ps.2010-01289

Basmacioğlu, H., & Ergül, M. (2005). Research on the factors affecting cholesterol content and some other characteristics of eggs in laying hens: The effects of genotype and rearing system. *Turkish Journal of Veterinary and Animal Sciences*, *29*(1), 157–164.

Borges, E. M., Volmer, D. A., Gallimberti, M., Ferreira De Souza, D., Luiz De Souza, E., & Barbosa, F. (2015). Evaluation of macro- and microelement levels for verifying the authenticity of organic eggs by using chemometric techniques. *Analytical Methods*, *7*(6), 2577–2584. https://doi.org/10.1039/c4ay02986k

Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceña, L., & Busto, O. (2015). Data fusion methodologies for food and beverage authentication and quality assessment - A review. *Analytica Chimica Acta*, *891*, 1–14. https://doi.org/10.1016/j.aca.2015.04.042

Carvalho, L. C., Morais, C. L. M., Lima, K. M. G., Leite, G. W. P., Oliveira, G. S., Casagrande, I. P., … Teixeira, G. H. A. (2018). Using Intact Nuts and Near Infrared Spectroscopy to Classify Macadamia Cultivars. *Food Analytical Methods*, *11*(7), 1857–1866. https://doi.org/10.1007/s12161-017-1078-9

de Oliveira Mendes, T., Porto, B. L. S., Almeida, M. R., Fantini, C., & Sena, M. M. (2019). Discrimination between conventional and omega-3 fatty acids enriched eggs by FT-Raman spectroscopy and chemometric tools. *Food*

*Chemistry*, *273*(November), 144–150. https://doi.org/10.1016/j.foodchem.2017.12.084

Dixon, S. J., & Brereton, R. G. (2009). Comparison of performance of five common classifiers represented as boundary methods: Euclidean Distance to Centroids, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Learning Vector Quantization and Support Vector Machines, as dependent on. *Chemometrics and Intelligent Laboratory Systems*, *95*(1), 1–17. https://doi.org/10.1016/j.chemolab.2008.07.010

European Commission. (2008). Commission regulation (EC) no 889/2008 of 5 September 2008 laying down detailed rules for the implementation of council regulation (EC) no 834/2007 on organic production and labelling of organic products with regard to organic production, labelling and co. In *Official Journal of the European Union* (Vol. L 250, pp. 1–84).

European Council. (2007). Council Regulation (EC) No 834/2007 of 28 June 2007 on organic production and labelling of organic products and repealing Regulation (EEC) No 2092/91. *Official Journal of the European Union*, *L 189*, 1–23. https://doi.org/2004R0726 - v.7 of 05.06.2013

González-Muñoz, M. J., Bastida, S., Jiménez, O., Lorenzo de, C., Vergara, G., & Sánchez-Muniz, F. J. (2009). The effect of dietary fat on the fatty acid composition and cholesterol content of Hy-line and Warren hen eggs. *Grasas y Aceites*, *60*(4), 350–359. https://doi.org/10.3989/gya.108208

Gregory, N. G., Gepp, M. J., & Babidge, P. J. (2005). Method for checking label accuracy in barn and free range eggs. *Journal of the Science of Food and Agriculture*, *85*(9), 1421–1426. https://doi.org/10.1002/jsfa.2152

Hansen, H., & Wang, T. (2015). Does the Saponification-GC Method

Underestimate Total Cholesterol Content in Samples Having Considerable Cholesterol Esters? *JAOCS, Journal of the American Oil Chemists' Society*, *92*(11–12), 1735–1738. https://doi.org/10.1007/s11746-015-2736-y

Hidalgo, A., Rossi, M., Clerici, F., & Ratti, S. (2008). A market study on the quality characteristics of eggs from different housing systems. *Food Chemistry*, *106*(3), 1031–1038. https://doi.org/10.1016/j.foodchem.2007.07.019

Johnson, A. E., Sidwick, K. L., Pirgozliev, V. R., Edge, A., & Thompson, D. F. (2018). Metabonomic Profiling of Chicken Eggs during Storage Using High-Performance Liquid Chromatography–Quadrupole Time-of-Flight Mass Spectrometry. *Analytical Chemistry*, *90*(12), 7489–7494. https://doi.org/10.1021/acs.analchem.8b01031

Joint Research Centre. (2018a). January 2018 Summary of Articles on Food Fraud and Adulteration. *JRC Food Fraud Monthly Report*, (January), 1–3.

Joint Research Centre. (2018b). June 2018 Summary of Articles on Food Fraud and Adulteration. *JRC Food Fraud Monthly Report*, (June), 1–3.

Joint Research Centre. (2018c). March 2018 Summary of Articles on Food Fraud and Adulteration. *JRC Food Fraud Monthly Report*, (March), 1–6.

Küçükyilmaz, K., Bozkurt, M., Yamaner, C., Inar, M., Atli, A. U., & Konak, R. (2012). Effect of an organic and conventional rearing system on the mineral content of hen eggs. *Food Chemistry*, *132*(2), 989–992. https://doi.org/10.1016/j.foodchem.2011.11.084

Leenstra, F., Maurer, V., Galea, F., Bestman, M., Amsler-kepalaite, Z., Visscher, J., … Krimpen, V. (2014). Laying hen performance in different production systems ; why do they differ and how to close the gap ? Results of discussions with groups of farmers in The Netherlands , Switzerland and France , benchmarking and model. *European Poultry Science*, *78*, 1–10. https://doi.org/10.1399/eps.2014.53

Martelo-Vidal, M. J., Domínguez-Agis, F., & Vázquez, M. (2013). Ultraviolet/visible/near-infrared spectral analysis and chemometric tools for the discrimination of wines between subzones inside a controlled designation of origin: A case study of Rías Baixas. *Australian Journal of Grape and Wine Research*, *19*(1), 62–67. https://doi.org/10.1111/ajgw.12003

Minelli, G., Sirri, F., Folegatti, E., Meluzzi, A., & Franchini, A. (2007). Egg quality traits of laying hens reared in organic and conventional systems. *Italian Journal of Animal Science*, *6*(SUPPL. 1), 728–730. https://doi.org/10.4081/ijas.2007.1s.728

Montagnese, C., Santarpia, L., Buonifacio, M., Nardelli, A., Caldara, A. R., Silvestri, E., … Pasanisi, F. (2015). European food-based dietary guidelines: A comparison and update. *Nutrition*, *31*(7–8), 908–915. https://doi.org/10.1016/j.nut.2015.01.002

Mugnai, C., Sossidou, E. N., Dal Bosco, A., Ruggeri, S., Mattioli, S., & Castellini, C. (2014). The effects of husbandry system on the grass intake and egg nutritive characteristics of laying hens. *Journal of the Science of Food and Agriculture*, *94*(3), 459–467. https://doi.org/10.1002/jsfa.6269

Napolitano, F., Serrapica, M., & Braghieri, A. (2013). Contrasting attitudes towards animal welfare issues within the food chain. *Animals*, *3*(2), 551–557. https://doi.org/10.3390/ani3020551

Oliveri, P., & Downey, G. (2012). Multivariate class modeling for the verification of food-authenticity claims. *TrAC - Trends in Analytical Chemistry*, *35*, 74–86. https://doi.org/10.1016/j.trac.2012.02.005

Porep, J. U., Kammerer, D. R., & Carle, R. (2015). On-line application of near infrared (NIR) spectroscopy in food production. *Trends in Food Science and Technology*, *46*(2), 211–230. https://doi.org/10.1016/j.tifs.2015.10.002

Puertas, G., Cazón, P., & Vázquez, M. (2018). O-024 Improvements in egg yolk saponification for enzymatic cholesterol quantification. In *The Book of Abstracts of the ISEAC-40 International Conference on Environmental & Food Monitoring* (p. 98).

Puertas, G., & Vázquez, M. (2018). Advances in techniques for reducing cholesterol in egg yolk: a review. *Critical Reviews in Food Science and Nutrition*, *In press*. https://doi.org/10.1080/10408398.2018.1448357

Ríos-Reina, R., García-González, D. L., Callejón, R. M., & Amigo, J. M. (2018). NIR spectroscopy and chemometrics for the typification of Spanish wine vinegars with a protected designation of origin. *Food Control*, *89*, 108–116. https://doi.org/10.1016/j.foodcont.2018.01.031

Rogers, K. M. (2009). Stable isotopes as a tool to differentiate eggs laid by caged, barn, free range, and organic hens. *Journal of Agricultural and Food Chemistry*, *57*(10), 4236–4242. https://doi.org/10.1021/jf803760s

Samman, S., Kung, F. P., Carter, L. M., Foster, M. J., Ahmad, Z. I., Phuyal, J. L., & Petocz, P. (2009). Fatty acid composition of certified organic, conventional and omega-3 eggs. *Food Chemistry*, *116*(4), 911–914. https://doi.org/10.1016/j.foodchem.2009.03.046

Spink, J., & Moyer, D. C. (2011). Defining the Public Health Threat of Food

Fraud. *Journal of Food Science*, *76*(9). https://doi.org/10.1111/j.1750-3841.2011.02417.x

Tres, A., O'Neill, R., & Van Ruth, S. M. (2011). Fingerprinting of fatty acid composition for the verification of the identity of organic eggs. *Lipid Technology*, *23*(2), 40–42. https://doi.org/10.1002/lite.201100084

Tres, A., & Van Ruth, S. M. (2011). Verification of organic feed identity by fatty acid fingerprinting. *Journal of Agricultural and Food Chemistry*, *59*(16), 8816–8821. https://doi.org/10.1021/jf2016682

Van Ruth, S., Alewijn, M., Rogers, K., Newton-Smith, E., Tena, N., Bollen, M., & Koot, A. (2011). Authentication of organic and conventional eggs by carotenoid profiling. *Food Chemistry*, *126*(3), 1299–1305. https://doi.org/10.1016/j.foodchem.2010.11.081

Van Ruth, S. M., Koot, A. H., Brouwer, S. E., Boivin, N., Carcea, M., Zerva, C. N., … Rom, S. (2013). Eggspectation: Organic egg authentication method challenged with produce from ten different countries. *Quality Assurance and Safety of Crops and Foods*, *5*(1), 7–14. https://doi.org/10.3920/QAS2012.0114

Vincevica-Gaile, Z., Gaga, K., & Klavins, M. (2013). Food and Environment: Trace Element Content of Hen Eggs from Different Housing Types. *APCBEE Procedia*, *5*(January), 221–226. https://doi.org/10.1016/j.apcbee.2013.05.038

Washburn, K. W. (1979). Genetic Variation in the Chemical Composition of the Egg. *Poultry Science*, *58*(3), 529–535.

Wu, W., Mallet, Y., Walczak, B., Penninckx, W., Massart, D. L., Heuerding, S., & Erni, F. (1996). Comparison of regularized discrimminant analysis, linear discriminant analysis and quadratic discriminant analysis, applied to NIR data. *Anal. Chim. Acta*, *329*(3), 257–265.

Yang, Y., Wu, Y., Li, W., Liu, X., Zheng, J., Zhang, W., & Chen, Y. (2018). Determination of geographical origin and icariin content of Herba Epimedii using near infrared spectroscopy and chemometrics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *191*, 233–240. https://doi.org/10.1016/j.saa.2017.10.019

Zaheer, K. (2017). Hen egg carotenoids (lutein and zeaxanthin) and nutritional impacts on human health: a review. *CYTA - Journal of Food*, *15*(3), 474–487. https://doi.org/10.1080/19476337.2016.1266033

Zakowska-Biemans, S., & Tekień, A. (2017). Free range, organic? Polish consumers preferences regarding information on farming system and nutritional enhancement of eggs: A discrete choice based experiment. *Sustainability (Switzerland)*, *9*(11). https://doi.org/10.3390/su9111999

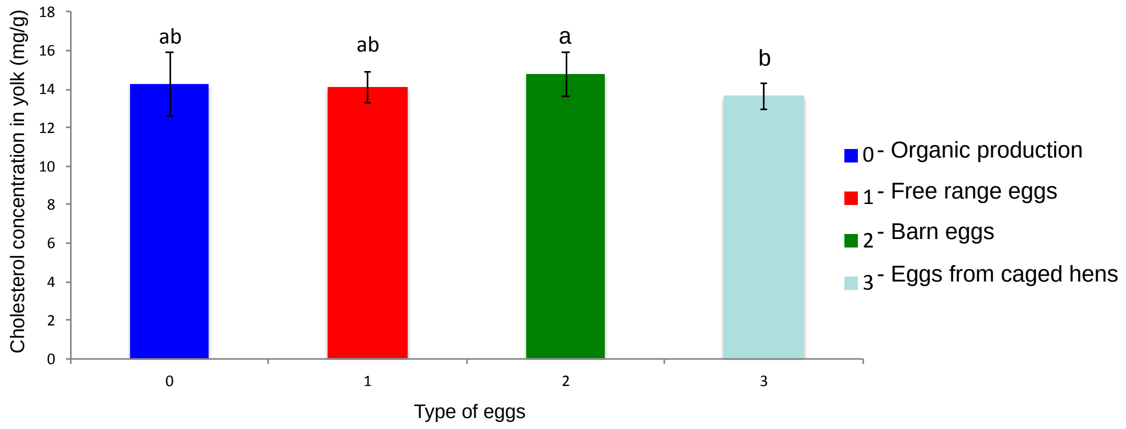


FIGURE 1. Cholesterol results for yolk eggs of the four housing systems (n = 48). Bars with different letters are significantly different ($p < 0.01$).

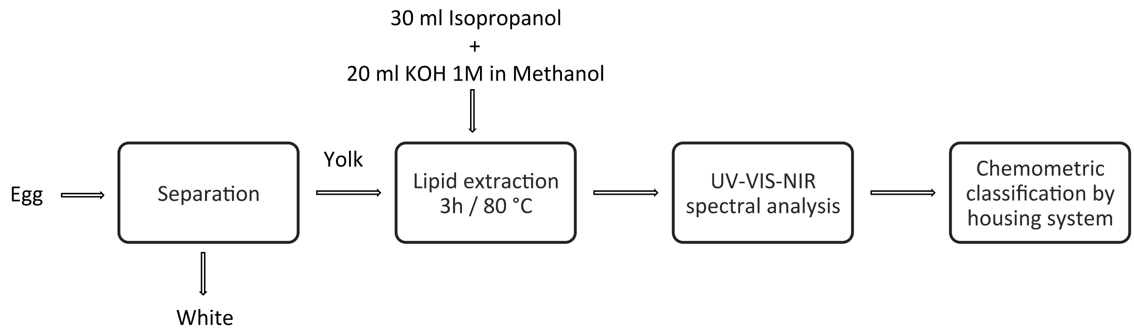


FIGURE 2. Flowchart of the proposed procedure for classification of egg by the four housing systems by UV-VIS-NIR spectral analysis.

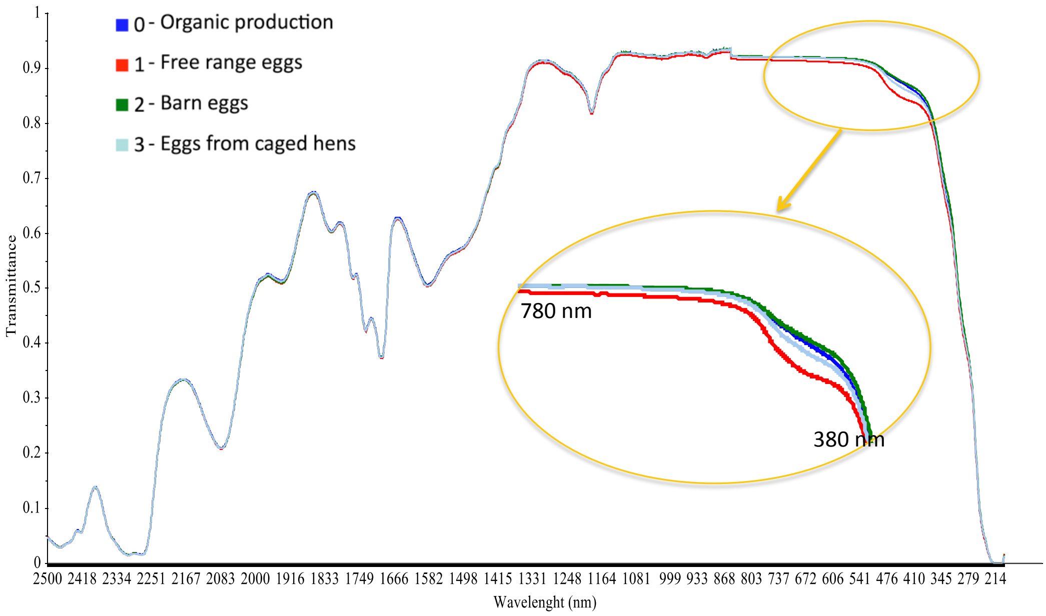


FIGURE 3. Average UV-VIS-NIR spectra of yolk lipid extracts of eggs from the four housing systems.

TABLE 1. SVM results with all the spectrum and at the different spectrum regions with MC & SDS data.

| Data pretreatment | Spectrum region | Accuracy (%) | | Type of eggs correctly predicted (%) | | | |
|---|---|---|---|---|---|---|---|
| | | Training | Validation | 0 | 1 | 2 | 3 |
| Raw | UV-VIS-NIR | 39.1 | 33.3 | 8.3 | 29.2 | 100 | 18.8 |
| MC & SDS | UV-VIS-NIR | 91.2 | 81.8 | 100 | 91.7 | 97.9 | 75.0 |
| SNV | UV-VIS-NIR | 37.5 | 26.0 | 2.1 | 29.2 | 100 | 18.8 |
| SGS | UV-VIS-NIR | 39.1 | 33.3 | 8.3 | 29.2 | 100 | 18.8 |

| Spectrum region | Accuracy (%) | | Type of eggs correctly predicted (%) | | | |
|---|---|---|---|---|---|---|
| | Training | Validation | 0 | 1 | 2 | 3 |
| UV | 86.5 | 80.2 | 89.6 | 87.5 | 91.7 | 77.1 |
| VIS | 65.1 | 58.3 | 58.3 | 85.4 | 79.2 | 37.5 |
| NIR | 72.9 | 64.1 | 93.8 | 50.0 | 72.9 | 75.0 |
| UV-VIS | 87.0 | 80.7 | 97.9 | 81.3 | 100 | 68.8 |
| VIS-NIR | 87.5 | 81.8 | 100 | 83.3 | 87.5 | 79.2 |

Farming method: type 0- organic production, type 1- free range, type 2- barn and type 3- cages. MC is mean centering, SDS is Standard Deviation Scaling, SNV is Standard Normal Variate and SGS is Savitzky-Golay Smoothing 7 points and second polynomial order.

TABLE 2. LDA and QDA results for UV-VIS-NIR spectra and 16 PC.

| Data pretreatment | LDA<br>% Accuracy | QDA<br>% Accuracy |
|---|---|---|
| Raw | 89.1 | 100 |
| SNV | 94.8 | 100 |
| MC & SDS | 74.5 | 98.4 |
| SGS | 88.5 | 100 |

MC is mean centering, SDS is Standard Deviation Scaling, SNV is Standard Normal Variate and SGS is Savitzky-Golay Smoothing 7 points and second polynomial order.

Table 3. QDA results for raw, SNV and SGS pretreatments with all the spectrum and at the different spectrum regions with raw and SNV data.

| **DP** | **No. PC for > 95 % accuracy** | **Accuracy (%)** | **Type of eggs correctly predicted (%)** | | | |
|---|---|---|---|---|---|---|
| | | | **0** | **1** | **2** | **3** |
| Raw | 9 | 97.4 | 100 | 93.8 | 100 | 95.8 |
| SNV | 7 | 95.3 | 100 | 81.3 | 100 | 100 |
| SGS | 9 | 97.4 | 100 | 93.8 | 100 | 95.8 |

| **DP** | **SP** | **PC for > 95 % accuracy** | **Accuracy (%)** | **Type of eggs correctly predicted (%)** | | | | **PC for 100 % accuracy** |
|---|---|---|---|---|---|---|---|---|
| | | | | **0** | **1** | **2** | **3** | |
| Raw | UV | 7 | 96.9 | 95.8 | 95.8 | 100 | 95.8 | 12 |
| SNV | UV | 7 | 96.6 | 91.7 | 93.8 | 100 | 100 | 12 |
| Raw | VIS | 7 | 95.8 | 100 | 83.3 | 100 | 100 | 9 |
| SNV | VIS | 7 | 95.3 | 100 | 83.3 | 100 | 97.9 | 9 |
| Raw | NIR | 16 | 95.3 | 97.9 | 91.7 | 95.8 | 95.8 | > 20 |
| SNV | NIR | 15 | 96.4 | 100 | 91.7 | 97.9 | 95.8 | >20 |

Housing method: type 0- organic production, type 1- free range, type 2- barn and type 3- cages. DP is data pretreatment. SNV is Standard Normal Variate and SGS is Savitzky-Golay Smoothing 7 points and second polynomial order.

***Declaration of Interest Statement**

**Declaration of interests**

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: