LWT - Food Science and Technology 100 (2019) 20–27

Contents lists available at ScienceDirect

LWT - Food Science and Technology

journal homepage: www.elsevier.com/locate/lwt

Quantification and identification of adulteration in the fat content of chicken hamburgers using digital images and chemometric tools

David Douglas de Sousa Fernandes[a], Florencia Romeo[b], Gabriela Krepper[b], María Susana Di Nezio[b], Marcelo Fabián Pistonesi[b], María Eugenia Centurión[b,**], Mário César Ugulino de Araújo[a], Paulo Henrique Gonçalves Dias Diniz[c,*]

[a] Universidade Federal da Paraíba, Departamento de Química, Laboratório de Automação e Instrumentação Em Química Analítica/Quimiometria (LAQA), Caixa Postal 5093, 58051-970 João Pessoa, PB, Brazil
[b] Departamento de Quimica, Universidad Nacional Del Sur, INQUISUR (UNS-CONICET), Av. Alem 1253 (B8000CPB), Bahía Blanca, Buenos Aires, Argentina
[c] Universidade Federal Do Oeste da Bahia, Campus Reitor Edgard Santos, Centro Das Ciências Exatas e Das Tecnologias, Rua Bertioga, 892, Bairro Morada Nobre I, Barreiras-BA, CEP: 47.810-059, Brazil

ARTICLE INFO

Keywords:
Meat
Fat content
Adulteration
Food quality
Chemometrics

ABSTRACT

In this work, we developed an eco-friendly methodology for quantification and identification of adulteration in the fat content of chicken hamburgers by combining color histograms (in RGB, HSI, and Grayscale channels) obtained from digital images and chemometric tools. For this, 74 samples of chicken hamburgers with a fat content of 14.27–47.55% (w w^{-1}) were studied, taking into account adulterations with a fat content higher than 20% (w w^{-1}), as limited by Argentinean legislation. In both quantitative and qualitative approaches, chemometric models containing HSI histograms achieved the best results, because this is very suitable in situations where there is a need to separate the chromaticity from the intensity. In other words, the opacity of the sample surfaces increases with increasing fat content. PLS/HSI achieved the best quantification result with a R^2 of 0.95, RMSEP of 2.01% w w^{-1}, REP of 7.26% w w^{-1} and RPD of 4.47 in the prediction set, while SPA-LDA/Grayscale + HSI reached the most satisfactory in the test set with only one misclassified sample. Therefore, the proposed methodologies represent excellent alternatives to conventional Soxhlet extraction method, since they follow the primary principles of Green Analytical Chemistry, avoiding waste generation, besides not using either chemical reagents or solvents.

1. Introduction

Meat intake is important for a balanced human nutrition, especially on developing countries whose diets are based on cereals and other crops. Meat and meat products are a great source of high biological value proteins, lipids, B group vitamins, especially vitamin B12, and minerals like iron and zinc. These macronutrients are essential for growth and body functions (Baltic & Boskovic, 2015; Bender, 1992; Pereira & Vicente, 2013).

As a result of the expensive price of red meats, the consumption of chicken meat increased in last years. In a developing country like Argentina, chicken consumption rises to 43 kg per capita in 2015 (USDA, 2015). Within the meat products, meat hamburgers and chicken are widely consumed. In this context, fast food consumption has been increasing as they are economical and easy to prepare. On the other hand, a serious public health problem of obesity in children and young people has been associated with hamburger intake (Drewnowski & Specter, 2004). Furthermore, the intake of this food product with a high content of saturated fatty acids can be directly related to increased risk for some diseases like hypercholesterolemia and colon, breast and prostate cancers (Chizzolini, Zanardi, Dorigoni, & Ghidini, 1999; Santé-Lhoutellier, 2014).

To elaborate hamburgers, minced meat is blended with fat, generating a product with high values of lipids and cholesterol. For this reason, fat content in these kind of food products must be regulated. In Argentina, the food regulation by the Argentinean Food Codex establishes that the amount of fat in hamburgers should not exceed 20% (w w^{-1}) (Código Alimentario Argentino, 2017). The official method recommended by the AOAC to determine fat content in meat and meat products is the Soxhlet extraction method (AOAC, 1990). This

* Corresponding author.
** Corresponding author.
E-mail addresses: mecentur@criba.edu.ar (M.E. Centurión), paulo.diniz@ufob.edu.br, paulodiniz.quimica@gmail.com (P.H.G.D. Diniz).

https://doi.org/10.1016/j.lwt.2018.10.034
Received 26 May 2018; Received in revised form 23 September 2018; Accepted 13 October 2018
Available online 13 October 2018
0023-6438/ © 2018 Elsevier Ltd. All rights reserved.

technique has plenty of disadvantages, not only being laborious and time consuming, but also using large amounts of sample and harmful solvents such as chloroform, methanol or ether (Krepper et al., 2018; Pérez-Palacios, Ruiz, Martín, Muriel, & Antequera, 2008). To overcome these drawbacks, Near Infrared (NIR) Spectroscopy and Nuclear Magnetic Resonance (NMR) have been successfully carried out as an alternative for this purpose (Krepper et al., 2018; Sørland, Larsen, Lundby, Rudi, & Guiheneuf, 2004; Weeranantanaphan, Downey, Allen, & Sun, 2011; Zamora-Rojas, Garrido-Varo, de Pedro-Sanz, Guerrero-Ginel, & Pérez-Marín, 2011). However, these techniques require expensive instrumentation, personal skills, previous knowledge and several procedures of sample pretreatment, although they are noninvasive and nondestructive.

For illustration, in our previous work (Krepper et al., 2018) NIR spectroscopy has been employed for quantification of fat content in chicken hamburgers, achieving a coefficient of correlation of 0.94, RMSEP of 1.59 mg kg^{-1}, relative error of prediction (REP) of 7.69% and ratio performance to deviation (RPD) of 3.02. This performance for the prediction set was obtained for the first derivative Savitzky-Golay smoothing with a second-order polynomial and window size of 19 points coupled with Interval Selection by the Successive Projections Algorithm for Partial Least Squares (*i*SPA-PLS). Other preprocessing techniques (baseline correction, standard normal variation and multiplicative scattering correction) and multivariate calibration methods (full-spectrum PLS and the Interval PLS (*i*PLS)) have been also used for comparison.

On the other hand, digital image-based food analysis has gained great prominence in the literature as a good alternative and the number of publications has increased considerably in recent years. This is due to the easiness of the use of digital images as a new tool to obtain analytical data, using scanners, webcams, digital cameras and cell phones (Byrne, Barker, Pennarun-Thomas, Diamond, & Edwards, 2000; Capitán-Vallveya, López-Ruiz, Martínez-Olmos, Erenas, & Palma, 2015; Grudpan, Kolev, Lapanantnopakhun, McKelvie, & Wongwilai, 2015). For example, digital image analysis were already been employed to estimate some quality parameters in meat products, such as color in beef (Larraín, Schaefer, & Reed, 2008) and pre-sliced hams (Valous, Mendoza, Sun, & Allen, 2009), moisture and tenderness in large cooked beef joints (Zheng, Sun, & Zheng, 2006a,b) and fat content in poultry meat (Chmiel, Słowiński, & Dasiewicz, 2011), cold meats (Zapotoczny, Szczypiński, & Daszkiewicz, 2016) and cured meat products (Cruz-Fernández, Luque-Cobija, Cervera, Morales-Rubio, & de la Guardia, 2017). However, these approaches generally involved the use of computer vision systems associated with several steps for image processing (before chemometric or statistical analysis), such as conversion of color spaces, image segmentation and texture analysis.

In such scenario, the use of color histograms (describing the frequency distribution of the pixels as a function of the recorded color component) obtained from simple digital images has been successfully demonstrated for evaluating the quality of food products, such as teas (Diniz et al., 2012), edible vegetable oils (Milanez & Pontes, 2014), honeys (Domínguez & Centurión, 2015; Domínguez, Diniz, Di Nezio, Araújo, & Centurión, 2014), coffees (Souto et al., 2015), extra virgin olive oils (Milanez & Pontes, 2015), propolis (Pierini et al., 2016) and hard candies (Botelho, Dantas, & Sena, 2017). The main advantage of the use of color histograms coupled with chemometric tools is the rapid acquisition of the analytical information using a common image acquisition device, without need for any image processing, other than extraction of the histograms.

Therefore, the aim of this work was the development of a simple, inexpensive and reliable methodology for quantification and identification of adulteration in the fat content of chicken hamburgers by combining color histograms obtained from digital images and suitable techniques of multivariate calibration and classification, respectively. For quantification purpose, Partial Least Squares (PLS) was employed to construct multivariate calibration models using color histograms in the Grayscale, RGB (Red-Green-Blue), HSI (Hue-Saturation-Intensity) channels and their combinations as analytical information. Considering the disadvantages of the Soxhlet extraction method and taking into account that merely a binary response (yes/no; i.e., adulterated or not) is required to assess the fraud in terms of the fat content of chicken hamburgers, we also proposed a qualitative approach based only in the color histograms obtained from the digital images to identify adulteration in chicken hamburger samples with a fat content higher than 20% (w w^{-1}) as recommended by Argentinean legislation. For this, an exploratory analysis of the data was initially performed by using Principal Component Analysis for screening the natural variability of the samples. Following, multivariate classification models were developed using Partial Least Squares Discriminant Analysis (PLS-DA) and the Successive Projections Algorithm for variable selection coupled with Linear Discriminant Analysis (SPA-LDA).

2. Materials and methods

2.1. Samples

In this work, 74 chicken hamburger samples were studied, being 54 samples prepared in the laboratory using a meat processor, with different fat and minced chicken meat contents provided by different local butcher shops, and 20 commercial samples acquired in local supermarkets located at Bahía Blanca City, Buenos Aires Province, Argentina. To prepare the samples in the laboratory, a stainless-steel meat mincer machine Spar Mixer (model UH-22MEC-B) was employed to mince the chicken meat and fat, which were passed 5 times in the mincer in order to simulate the same conditions of commercial hamburger preparation. The mincer was appropriately sanitized before the sample preparation. All samples were kept frozen at $-4\,^{\circ}C$ until analysis.

As minced chicken meat contains fat, it has been mixed with chicken fat and then the fat content of the mixture has been determined by the official AOAC method using Soxhlet extraction. The fat content in the 74 chicken hamburger samples ranged from 14.27 to 47.55% (w w^{-1}), being 27 samples with a fat content equal to or less than 20% (w w^{-1}), and other 47 samples greater than 20% (w w^{-1}) (Fig. 1a). This range was selected considering possible adulterations with a fat content higher than 20% (w w^{-1}), as limited by Argentinean legislation. Fat content in commercial hamburgers was determined using the official Soxhlet extraction method recommended by the Association of Official Analytical Chemists (AOAC, 1990). To illustrate the highest variation in fat content of the studied chicken hamburger samples, twelve selected photos are shown in Fig. 1b. For image acquisition, all samples were quartered, weighted (1.5 g) and then placed into a ring-shaped rubber support (sample holder) of 1.3 cm of internal diameter and 0.9 cm of height to maintain the same dimensions (Fig. 1c). All 74 samples were prepared in authentic triplicate and then three images from each replicate were captured. Only average values for each sample were used for calculations.

2.2. Instrumentation and software

The apparatus for image capturing was composed of a mechanical support, an Olympus® digital camera (model SP-510 UZ, with a resolution of 7.1 Mpixels), a circular fluorescent lamp and a polytetrafluoroethylene (PTFE) sample holder, as described by Diniz et al. (2012). The sample holder was set 12 cm under the digital camera lens and 10 cm under the center of a circular fluorescent lamp. This configuration maintains fixed positioning, luminosity, focus and sample-to-camera distance to ensure reproducibility of the measurements. Moreover, the PTFE sample holder avoids light scattering and fluorescence effects, which could cause effects on image color histograms. All measurements were carried out at room temperature (25 ± 1 °C).

Data analysis was performed using Matlab® 2010a (Mathworks Inc.), and images were converted into color histograms with the free

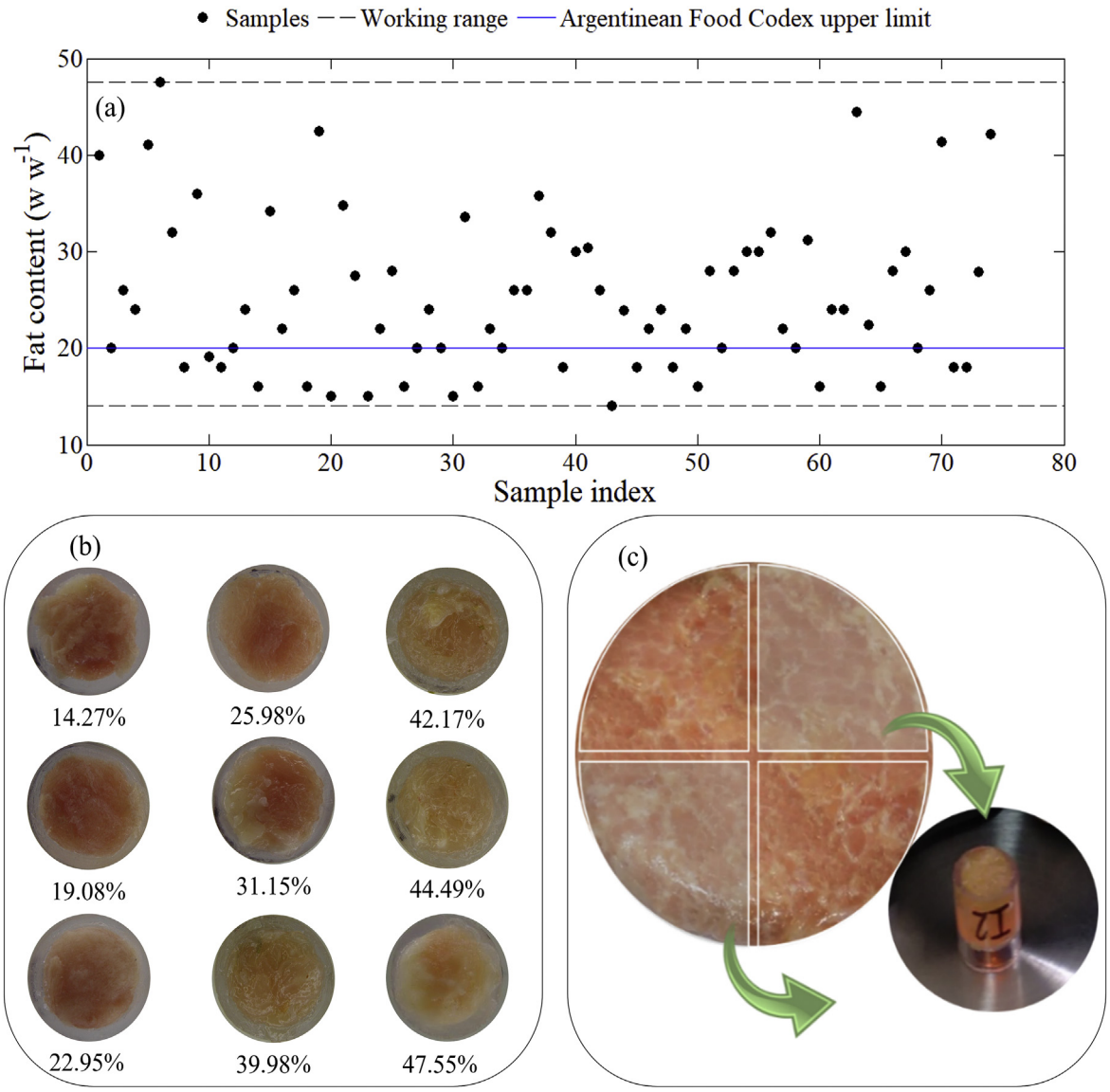

Fig. 1. (a) Fat content in the 74 studied chicken hamburger samples ranging from 14.27 to 47.55% (w w^{-1}), in which (b) the highest variation is illustrated in twelve selected photos. For image acquisition, (c) all samples were quartered and then placed into a sample holder of 1.5 g capacity.

Matlab interface "Imagens_gui", downloaded at http://www.laqa.quimica.ufpb.br/donwload.

2.3. Color histograms and multivariate data analysis

Digital images were recorded in triplicate for each replicate, in JPEG format with a resolution of 2880 × 1620 pixels. For the present study, a Region of Interest (ROI) corresponding to 40% of the original image was selected and its RGB (Red, Green and Blue), GS (Grayscale) and HSI (Hue, Saturation and Intensity) histograms were obtained. Mean histograms from each sample were used for calculations. Then, before using the chemometric tools, color levels with pixel frequency values equal to zero simultaneously in all samples were eliminated to avoid compromising results.

2.3.1. Chemometric procedure

2.3.1.1. Quantitative approach. In order to quantify the fat content in chicken hamburger samples, PLS was employed to construct multivariate calibration models using the color histograms obtained from the digital images as analytical information. For this purpose, calibration (50 samples) and prediction (24 samples) sets were initially selected by using the Kennard-Stone (KS) algorithm, taking into account the $\boldsymbol{X}$ and $\boldsymbol{y}$ distances simultaneously, in order that the calibration set contain samples with the smallest and largest values of $\boldsymbol{y}$, thus avoiding extrapolation problems (Kennard & Stone, 1969). This procedure includes the most representative samples into the calibration set, while prediction samples were not included for the model construction, being only used for the final data evaluation. The number of latent variables (LV) was selected based on the root mean square error of cross-validation (*RMSECV*) in the calibration set. The performance of the cross-validated calibration models was evaluated in terms of coefficient of determination (R^2_{CV}) and *RMSECV*. The predictive ability of the final models was evaluated in terms of coefficient of determination (R^2_{Pred}), root mean square error of prediction (*RMSEP*), relative error of prediction (*REP*) and ratio performance to deviation (*RPD*) in the external prediction set. R^2_{CV}, *RMSECV*, R^2_{Pred}, *RMSEP*, *REP* and *RPD* are calculated as follows:

$$R^2_{CV} = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2}$$

Table 1
Results obtained by PLS with differing color histograms in the Grayscale, RGB, HSI channels and their combinations for determination of the fat content in chicken hamburger samples.

Color histograms[a]	Calibration		Prediction			
	R^2	RMSECV(% w w^{-1})	R^2	RMSEP (% w w^{-1})	REP (%)	RPD
Grayscale (4)	0.25	7.49	0.49	5.05	20.3	1.40
RGB (3)	0.31	6.77	0.73	4.74	18.1	1.92
HSI (12)	0.85	2.88	0.95	2.01	7.26	4.47
Grayscale + RGB (3)	0.33	6.92	0.72	4.48	17.8	1.89
Grayscale + HSI (12)	0.80	3.24	0.94	2.12	7.85	4.08

[a] The number of latent variables used to construct the PLS model is indicated in parenthesis.

$$RMSECV = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$$

$$R^2_{Pred} = 1 - \frac{\sum_{j=1}^{m}(\hat{y}_j - y_j)^2}{\sum_{j=1}^{m}(y_j - \bar{y}_j)^2}$$

$$RMSEP = \sqrt{\frac{1}{m}\sum_{j=1}^{m}(\hat{y}_j - y_j)^2}$$

$$REP = \left(\frac{RMSEP}{\frac{1}{n}\times\sum_{i=1}^{n} y_i}\right)\times 100$$

$$RPD = \frac{Sd_{Pred}}{RMSEP}$$

where n and m are sample sizes of the calibration (i) and prediction (j) sets, respectively; $\hat{y}_i$ and $\hat{y}_j$ are model-predicted values of the samples; y_i and y_j are reference values of the samples; $\bar{y}_i$ and $\bar{y}_j$ are reference mean values of the samples; Sd_{Pred} is the standard deviation of the prediction samples.

Elliptical joint confidence region (EJCR) test was also applied to the best PLS result in the prediction set in order to confirm verify the absence of bias, i.e. to evaluate the accuracy of the constructed model. In this case, the results obtained by the chemometric modeling are compared with their respective reference values. For this purpose, an ordinary least squares (OLS) fitting was then obtained, and the estimated intercept (a) and slope (b) were compared with their ideal values of 0 and 1, respectively (Diniz, Pistonesi, Araújo, & Band, 2013; Krepper et al., 2018).

2.3.1.2. Qualitative approach. Initially, an exploratory analysis of the data was performed by using Principal Component Analysis (PCA) for screening the natural variability of the samples. Following, multivariate classification models were developed by using PLS-DA and SPA-LDA. For this, dataset was also divided into training (50 samples) and test (24 samples) sets by applying the KS algorithm taking into account only the $\boldsymbol{X}$ distances in this case. The classification models were constructed and then validated using a full cross-validation step. Test samples were only used for the final data evaluation and comparison of the performance of the multivariate classifiers, which was evaluated in terms of accuracy, sensitivity and specificity (Costa et al., 2015; Lavine, 2009; Massart, Vandeginste, Buydens, Lewi, & Smeyers-Verbeke, 1998).

In the discussion, the terms positive and negative will be used, which refer to the fat content permitted (named as Permitted) and not-permitted (named as Adulterated) by the Argentinean legislation, i.e. under and above 20% w w^{-1}, respectively. In other words, a false positive indicates a sample with a fat content equal or lower than 20% w w^{-1} classified as adulterated, while a false negative indicates an adulterated sample classified as one containing less than 20% w w^{-1} of fat content. Accuracy rate was calculated as the number of correct classifications divided by the total number of samples in the set under consideration (training or test set). Sensitivity rate was calculated as the number of correct positive decisions divided by the total number of known positive cases. Finally, the specificity rate was calculated as the number of correct negative decisions divided by the total number of known negative cases (Lavine, 2009; Massart et al., 1998).

All algorithms used in both quantitative and qualitative approaches were performed with Matlab® 2009b (Mathworks Inc.) software. PLS-DA was calculated using the Classification toolbox for Matlab® (version 4.0) released by Milano Chemometrics and QSAR Research Group (Ballabio & Consonni, 2013).

3. Results and discussion

3.1. Quantification of the fat content in chicken hamburgers

To quantify the fat content in chicken hamburger samples PLS models were constructed using different color histograms in the Grayscale, RGB, HSI channels and their combinations as analytical information. These results are summarized in Table 1. As can be seen, relative errors of prediction (REP) lower than 8% were obtained when HSI is used alone or combined with Grayscale (i.e. Grayscale + HSI) histograms. This occurs because HSI is very suitable in situations where there is a need to separate the chromaticity (defined as the degree of color purity and is related to both hue and saturation) from the intensity/luminosity (Gonzalez & Woods, 1992; Souto et al., 2015). This avoids problems related to variations in the distribution of light during the image acquisition process, because the opacity of the sample surface increases with increasing fat content, affecting the results since it depends only on the statistical distribution of the pixels (color histograms) as a function of the recorded color component in a digital image (Diniz et al., 2012). This finding is illustrated in the PLS regression coefficients obtained for HSI histograms (Fig. 2a), in which Hue and Intensity channels clearly contain the analytical information responsible by the differentiation between the fat contents in the chicken hamburger samples. As can be seen, Hue has been most prominent in the first latent variables, while Intensity has been highlighted in the last ones. This occurs because the robust nature of Hue parameter affords superior precision, as it is stable, simple to calculate, easily obtained from commercial image capturing devices and continuous over the entire color gamut. In comparison with R, G, B intensities and RGB absorbance, Hue has been demonstrated to be 2 to 3 times superior as a quantitative analytical parameter (Cantrell, Erenas, de Orbe-Payá, & Capitán-Vallvey, 2010). Thus, the best performance for PLS modeling in the prediction set was then obtained by using HSI histograms, achieving R^2 of 0.95, RMSEP of 2.01% w w^{-1}, REP of 7.26% w w^{-1} and RPD of 4.47. As shown in the predicted vs reference plot in Fig. 2b, the prediction samples are randomly distributed on both sides of the bisecting line, indicating the absence of systematic error. In Fig. 2c, EJCR plot of the prediction set for the PLS/HSI model confirms the absence of biased-results, since they contained the ideal theoretical point (Granato, Calado, & Jarvis, 2014). It is worth to highlight that the PLS/Grayscale + HSI model has also been shown to be suitable for quantification of fat content in chicken burger samples, although the results are slightly lower than those obtained by PLS/HSI; i.e. PLS/Grayscale + HSI achieved R^2 of 0.94, RMSEP of 2.12% w w^{-1}, REP of 7.85% w w^{-1} and RPD of 4.08 in the prediction set.

Regarding the literature, Cruz-Fernández et al. (2017) demonstrated that their results were better in terms of coefficient of correlation and prediction errors than those published by Chmiel et al. (2011) and Zapotoczny et al. (2016). Despite this, Cruz-Fernández et al. (2017) obtained prediction errors of the order of 20% for the determination of fat content in *salchichón*, *salami* and cured ham, while in our work the prediction error was less than 7.26%. This result confirms the superiority of the use of color histograms as analytical information instead

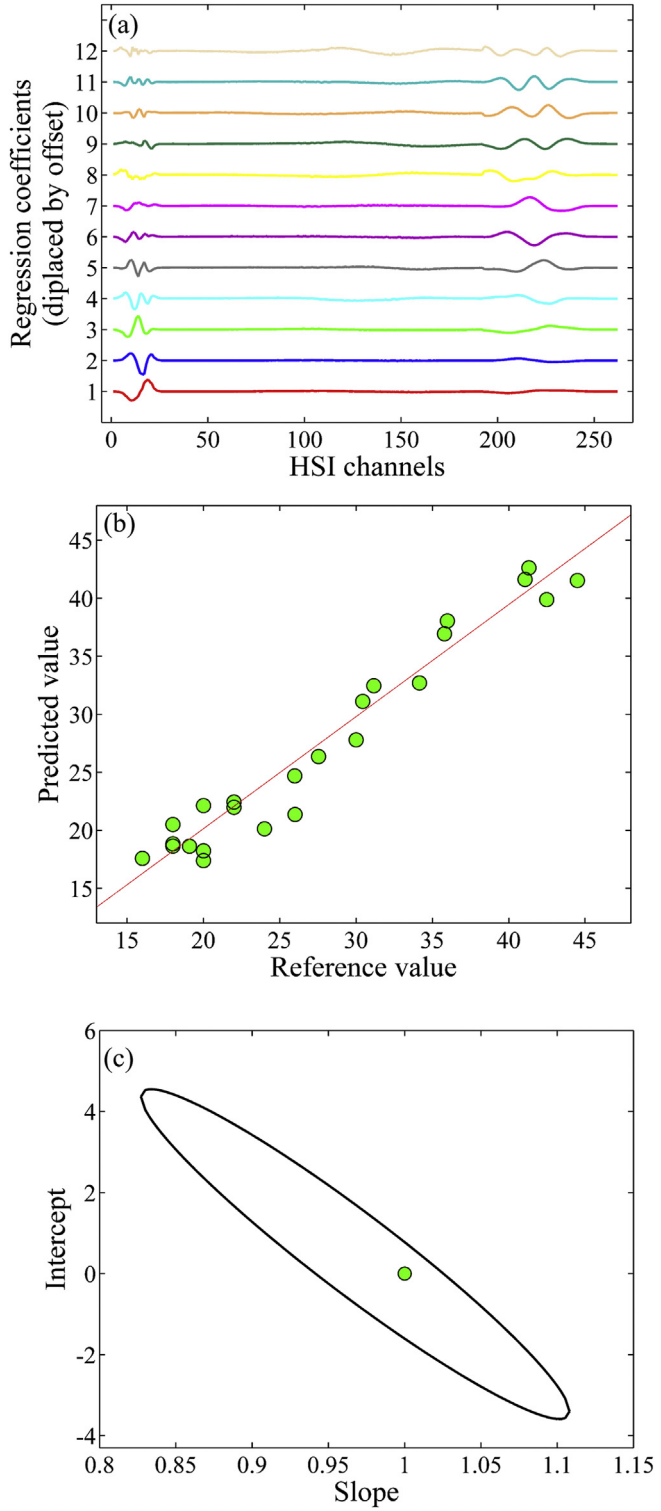

Fig. 2. Plots of (a) Regression coefficients, (b) predicted versus reference values and (a) elliptical joint confidence region containing the ideal theoretical point for the prediction set obtained for the PLS model constructed with the HSI histograms for quantification of the fat content in the chicken hamburger samples. The ideal result (predicted = reference) is indicated by a red straight line in (b), which corresponds to the bisecting line of the plots. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

color descriptors/attributes, because they are most informative and less correlated. Moreover, it is worth to highlight that the methodology here proposed improves the results obtained in our previous work employing NIR spectroscopy (Krepper et al., 2018), which obtained a REP of 7.69% and RPD of 3.02.

3.2. Identification of adulteration in the fat content in chicken hamburgers

Considering that fat quantification presents intrinsic disadvantages such as high consumption of time and organic solvents, the development of a simple method requiring only a binary response (adulterated or not) would be more economical and less laborious to identify adulterations in chicken hamburger samples. In other words, it is only necessary to identify if the samples contain more or less than 20% (w w^{-1}) fat. Since response has been achieved from multiple non-specific signals (color histograms obtained from digital images), a multivariate classification approach using pattern recognition techniques is required. This strategy is also referred in literature as non-target analysis, because the data set is used as a fingerprint of the sample (Callao & Ruisánchez, 2018).

Mean histograms of the chicken hamburger samples with fat content under and above 20% w w^{-1} are named as 'Permitted' and 'Adulterated', respectively, as shown in Fig. 3a. As can be seen, the profiles of the mean histograms for both Permitted and Adulterated classes are very similar, which justifies the use of chemometric tools, since simple visual inspection cannot evaluate adulterations of the fat content in these samples. For this, an exploratory analysis of all studied chicken hamburger samples was then performed by PCA using Grayscale, RGB, HSI, Grayscale + RGB, and Grayscale + HSI histograms. PCA score plots for these color spaces are visualized in Fig. 3b–f. As observed, there is a high overlap between the studied classes (Permitted and Adulterated) for all obtained results. Therefore, PCA cannot provide the identification of the chicken hamburger samples adulterated with a fat content higher than 20% (w w^{-1}), which requires the use of suitable supervised pattern recognition techniques such as PLS-DA and SPA-LDA.

Table 2 summarizes the results for identification of adulteration of the fat content in chicken hamburger samples in terms of classification accuracy rate obtained for the different color histograms in the Grayscale, RGB, HSI, Grayscale + RGB and Grayscale + HSI channels using PLS-DA and SPA-LDA in both training and test sets. Best results were obtained using Grayscale + HSI histograms for both multivariate classifiers, where PLS-DA and SPA-LDA achieved a correct classification rate of 85.7 and 93.4%, respectively. Table 3 presents exclusively the confusion matrix, with the accuracy, sensitivity and specificity, in order to compare the performance of these best classification results obtained by PLS-DA and SPA-LDA using Grayscale + HSI histograms. Here again, similar to what happened in the quantitative approach, Grayscale + HSI histograms proved to be most suitable as containing the relevant analytical information for identifying the adulterations in the fat content of chicken hamburger samples.

In general, SPA-LDA coupled with Grayscale + HSI histograms reached the most satisfactory result among all constructed classification models, attaining 92.0 and 95.8% of correct classifications in both training and test sets, respectively, while PLS-DA/Grayscale + HSI achieved classification accuracy rates of only 84.0 and 87.5%. In this case, the optimal number of variables selected by SPA to construct the LDA model was 16, with a lowest validation *G* cost value of 0.7166. Sensitivities of 94.1 and 90.0%, and specificities of 90.9 and 100% were obtained for both training and test sets, respectively. The selected variables by SPA-LDA in the Grayscale + HSI channels and its respective Fisher's discriminant function plot are indicated in Fig. 4a and b, respectively. On the other hand, PLS-DA/Grayscale + HSI model was constructed with an optimal number of 16 latent variables, providing sensitivities of only 88.2 and 80.0%, and specificities of 80.0 and 92.8% for both training and test sets, respectively.

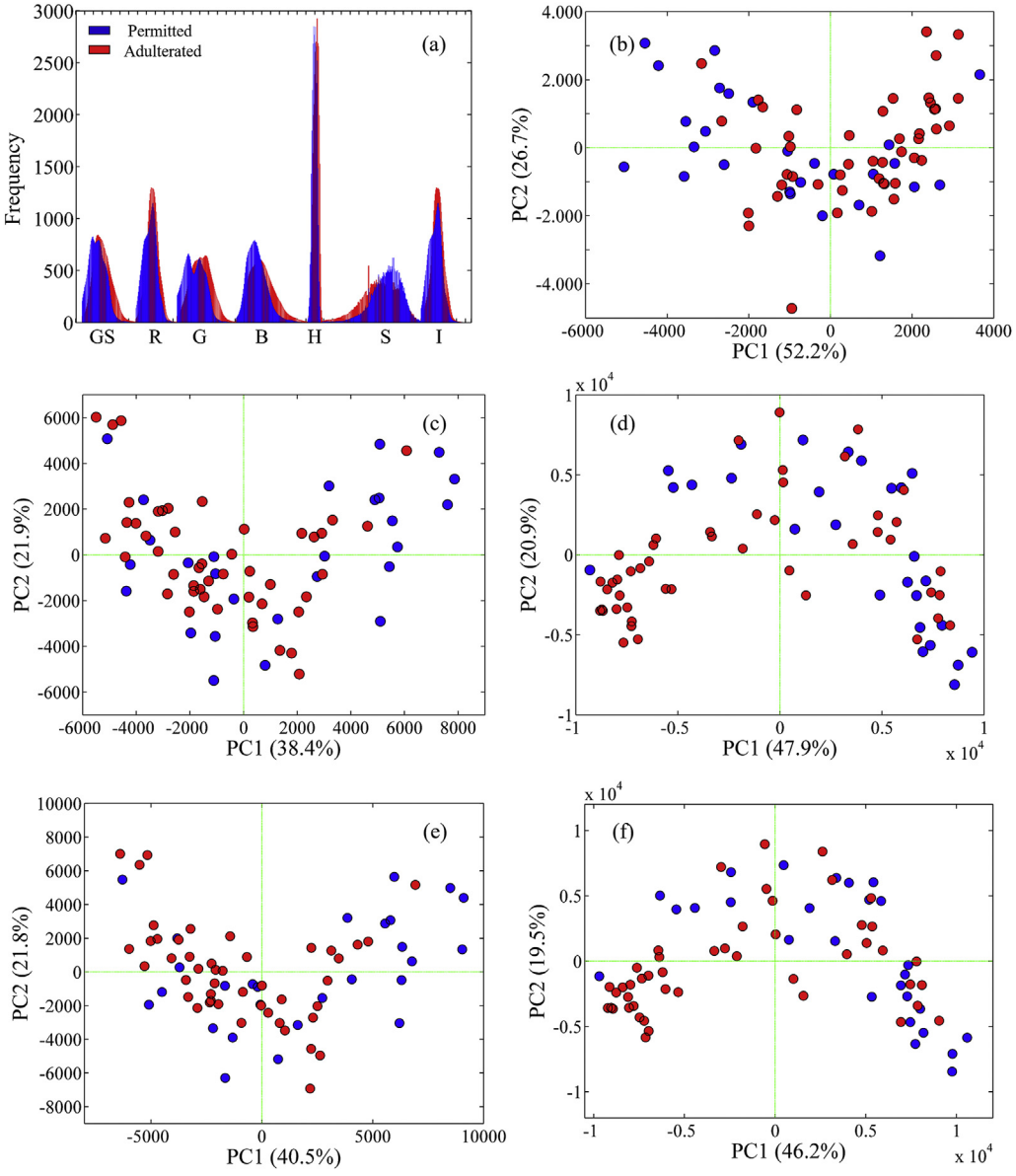

Fig. 3. (a) Mean histograms in the Grayscale, RGB and HSI channels for the studied chicken hamburger samples, and the respective PCA score plots obtained using the (b) Grayscale, (c) RGB, (d) HSI, (e) Grayscale + RGB, and (f) Grayscale + HSI histograms. Samples in the 'Permitted' and 'Adulterated' classes are indicated in blue and red, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 2
Classification accuracy obtained for the differing color histograms in the Grayscale, RGB, HSI channels and their combinations using PLS-DA and SPA-LDA.

Color histograms	Classification Accuracy (%)			
	PLS-DA		SPA-LDA	
	Training[a]	Test	Training[a]	Test
Grayscale	(3) 70.0	70.8	(19) 84.0	75.0
RGB	(7) 74.0	79.2	(15) 92.0	70.8
HSI	(15) 76.8	95.8	(27) 100	83.3
Grayscale + RGB	(13) 88.0	79.2	(17) 94.0	70.8
Grayscale + HSI	(16) 84.0	87.5	(16) 92.0	95.8

[a] The number of latent variables/selected variables is indicated in parenthesis.

Comparing the overall performance of the multivariate classifiers, it is worth to highlight that PLS-DA classified incorrectly 4 Permitted samples as Adulterated, and also 7 Adulterated samples as Permitted, totalizing 11 classification errors. On the other hand, SPA-LDA misclassified only 5 samples, being 2 Permitted samples incorrectly classified as Adulterated, and 3 Adulterated samples as Permitted. In order to evaluate the predictive ability of these multivariate classifiers, it was verified that the classification errors of PLS-DA in the test set were attributed to two samples with a fat content of 20.00% and other one with 35.98% w w^{-1}. Oppositely, the unique classification error in the test set for SPA-LDA is due to a sample with a fat content of 19.09% w w^{-1}. In other words, most of classification errors were found very close to the decision border of 20% w w^{-1}. This occurs because the constructed models depend on the values obtained by the Sohxlet extraction reference method, which naturally involves a lot of sample manipulation, which can lead to analyte losses, directly affecting the procedure of analytical measurement. In this sense, the good predictive ability of constructed SPA-LDA model (with only one misclassification in the test

Table 3
Confusion matrix, with the accuracy, sensitivity and specificity, obtained by PLS-DA and SPA-LDA using Grayscale + HSI histograms.

	PLS-DA (16)[a]				SPA-LDA (16)[a]			
	Training		Test		Training		Test	
	Permitted	Adulterated	Permitted	Adulterated	Permitted	Adulterated	Permitted	Adulterated
Permitted	15	2	8	2	16	1	9	1
Adulterated	6	27	1	13	3	30	0	14
Sensitivity (%)	88.2		80.0		94.1		90.0	
Specificity (%)	81.8		92.8		90.9		100	
Accuracy (%)	84.0		87.5		92.0		95.8	

[a] The number of latent variables/selected variables is indicated in parenthesis.

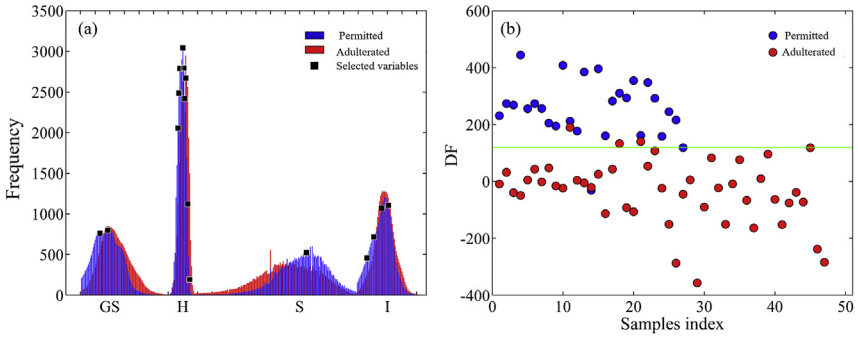

Fig. 4. (a) Selected variables by SPA-LDA in the Grayscale + HSI channels and (b) its respective Fisher's discriminant function plot for the identification of adulteration of the fat content in the studied chicken hamburger samples. Samples in the 'Permitted' and 'Adulterated' classes are indicated in blue and red, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

set) is due to the selected variables included in the model, which is more stable and lead to the lowest classification error, since SPA is an iterative forward selection method that solves collinearity problems by selecting such variables as whose information content is minimally redundant (Soares, Gomes,Galvão Filho, Araújo, & Galvão, 2013).

These results of the proposed methodology are very useful from the experimental point of view, because they provide a fast tool for screening adulteration in the fat content of chicken hamburgers, since only samples with levels close to the limit established by the Argentine legislation should be sent for quantitative analysis. This directly impacts on the reduction of the use of reagents and solvents, besides saving time of analysis, which follows, therefore, the principles of Green Analytical Chemistry.

4. Conclusion

In this work, the quality of chicken hamburgers in terms of their fat content was evaluated by combining color histograms obtained from digital images and suitable techniques of multivariate calibration and classification (pattern recognition). In both quantitative and qualitative approaches, chemometric models containing HSI histograms achieved the best results, because it is very suitable in situations where there is a need to separate the chromaticity from the intensity/luminosity. In our case, the opacity of the chicken hamburger sample surfaces increases with increasing fat content. This analytical information is easily translated by the color histograms since it depends only on the statistical distribution of the pixels as a function of the recorded color component in a digital image. Therefore, the proposed digital image-based methodology is a promising eco-friendly tool to monitor (identify and quantify) adulteration in the fat content of chicken hamburger samples. Additionally, it represents an excellent alternative to conventional Soxhlet extraction method recommended by AOAC, since it follows the primary principles of Green Analytical Chemistry, avoiding waste generation, besides not using either chemical reagents or solvents. However, a larger and more varied testing of meat types (bovine, swine, poultry, etc.) must be implemented to guarantee any generalization of the proposed methodology.

Acknowledgements

Authors acknowledge the financial support of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES; Brazil) and Ministerio de Ciencia, Tecnología e Innovación Productiva (MINCYT; Argentina) (Project CAPES/MINCYT REDE No. 015/14). Argentinian authors acknowledge financial support from Universidad Nacional del Sur (Dpto. De Química-INQUISUR) and Proyecto Grupo de Investigación (PGI 24/Q067) granted for Secretaría General de Ciencia y Tecnología (Argentina). G. Krepper and D.D.S Fernandes thanks to Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, Argentina) y CAPES for the scholarships, respectively. M.F.P. is also grateful to CIC (Comisión de Investigaciones Científicas de la Provincia de Buenos Aires).

References

AOAC (Association of Official Analytical Chemists) (1990). *Official methods of analysis* (15th ed.). (Washington DC).

Ballabio, D., & Consonni, V. (2013). Classification tools in chemistry. Part 1: Linear models. PLS-DA. *Anal Methods, 5*, 3790–3798. https://doi.org/10.1039/C3AY40582F.

Baltic, M. Z., & Boskovic, M. (2015). When man met meat: Meat in human nutrition from ancient times till today. *Procedia Food Science, 5*, 6–9. https://doi.org/10.1016/j.profoo.2015.09.002.

Bender, A. (1992). *Meat and meat products in human nutrition in developing countries. FAO food and nutrition paper 53*. Rome: Meat and health.

Botelho, B. G., Dantas, K. C. F., & Sena, M. M. (2017). Determination of allura red dye in hard candies by using digital images obtained with a mobile phone and N-PLS. *Chemom Intell Lab Syst, 167*, 44–49. https://doi.org/10.1016/j.chemolab.2017.05.004.

Byrne, L., Barker, J., Pennarun-Thomas, G., Diamond, D., & Edwards, S. (2000). Digital imaging as a detector for generic analytical measurements. *Trends in Analytical Chemistry, 19*, 517–522. https://doi.org/10.1016/S0165-9936(00)00019-4.

Callao, M. P., & Ruisánchez, I. (2018). An overview of multivariate qualitative methods for food fraud detection. *Food Control, 86*, 283–293. https://doi.org/10.1016/j.foodcont.2017.11.034.

Cantrell, K., Erenas, M. M., de Orbe-Payá, I., & Capitán-Vallvey, L. F. (2010). Use of the hue parameter of the hue, saturation, value color space as a quantitative analytical parameter for bitonal optical sensors. *Analytical Chemistry, 82*, 531–542. https://doi.org/10.1021/ac901753c.

Capitán-Vallveya, L. F., López-Ruiz, N., Martínez-Olmos, A., Erenas, M. M., & Palma, A. J. (2015). Recent developments in computer vision-based analytical chemistry: A tutorial review. *Analytica Chimica Acta, 899*, 23–56. https://doi.org/10.1016/j.aca.

2015.10.009.
Chizzolini, R., Zanardi, E., Dorigoni, V., & Ghidini, S. (1999). Calorific value and cholesterol content of normal and low-fat meat and meat products. *Trends in Food Science & Technology, 10*, 119–128. https://doi.org/10.1016/S0924-2244(99)00034-5.
Chmiel, M., Słowiński, M., & Dasiewicz, K. (2011). Application of computer vision systems for estimation of fat content in poultry meat. *Food Control, 22*, 1424–1427. https://doi.org/10.1016/j.foodcont.2011.03.002.
Código Alimentario Argentino (2017). *Capítulo VI: Alimentos cárneos Y afines, artículo 330*. http://www.anmat.gov.ar/alimentos/codigoa/Capitulo_VI_2017.pdf, Accessed date: 13 March 2018.
Costa, G. B., Fernandes, D. D. S., Almeida, V. E., Araújo, T. S. P., Melo, J. P., Diniz, P. H. G. D., et al. (2015). Digital image-based classification of biodiesel. *Talanta, 139*, 50–55. https://doi.org/10.1016/j.talanta.2015.02.043.
Cruz-Fernández, M., Luque-Cobija, M. J., Cervera, M. L., Morales-Rubio, A., & de la Guardia, M. (2017). Smartphone determination of fat in cured meat products. *Microchemical Journal, 132*, 8–14. https://doi.org/10.1016/j.microc.2016.12.020.
Diniz, P. H. G. D., Dantas, H. V., Melo, K. D. T., Barbosa, M. F., Harding, D. P., Nascimento, E. C. L., et al. (2012). Using a simple digital camera and SPA-LDA modeling to screen teas. *Anal Methods, 4*, 2648–2652. https://doi.org/10.1039/C2AY25481F.
Diniz, P. H. G. D., Pistonesi, M. F., Araújo, M. C. U., & Band, B. S. F. (2013). Eco-friendly sonoluminescent determination of free glycerol in biodiesel samples. *Talanta, 114*, 38–42. https://doi.org/10.1016/j.talanta.2013.04.009.
Domínguez, M. A., & Centurión, M. E. (2015). Application of digital images to determine color in honey samples from Argentina. *Microchemical Journal, 118*, 110–114. https://doi.org/10.1016/j.microc.2014.08.002.
Domínguez, M. A., Diniz, P. H. G. D., Di Nezio, M. S., Araújo, M. C. U., & Centurión, M. E. (2014). Geographical origin classification of Argentinean honeys using a digital image-based flow-batch system. *Microchemical Journal, 112*, 104–108. https://doi.org/10.1016/j.microc.2013.09.008.
Drewnowski, A., & Specter, S. (2004). Poverty and obesity: The role of energy density and energy costs. *American Journal of Clinical Nutrition, 79*, 6–16. https://doi.org/10.1093/ajcn/79.1.6.
Gonzalez, R. C., & Woods, R. E. (1992). *Digital image processing* (3rd ed.). Boston: Addison-Wesley Publishing.
Granato, D., Calado, V. M. A., & Jarvis, B. (2014). Observations on the use of statistical methods in food science and technology. *Food Research International, 55*, 137–149. https://doi.org/10.1016/j.foodres.2013.10.024.
Grudpan, K., Kolev, S. D., Lapanantnopakhun, S., McKelvie, I. D., & Wongwilai, W. (2015). Applications of everyday it and communications devices in modern analytical chemistry: A review. *Talanta, 136*, 84–94. https://doi.org/10.1016/j.talanta.2014.12.042.
Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics, 11*, 137–148. http://www.jstor.org/stable/1266770.
Krepper, G., Romeo, F., Fernandes, D. D. S., Diniz, P. H. G. D., de Araújo, M. C. U., Di Nezio, M. S., et al. (2018). Determination of fat content in chicken hamburgers using NIR spectroscopy and the Successive Projections Algorithm for interval selection in PLS regression (*i*SPA-PLS). *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 189*, 300–306. https://doi.org/10.1016/j.saa.2017.08.046.
Larraín, R. E., Schaefer, D. M., & Reed, J. D. (2008). Use of digital images to estimate CIE color coordinates of beef. *Food Research International, 41*, 380–385. https://doi.org/10.1016/j.foodres.2008.01.002.
Lavine, B. K. (2009). Validation of classifiers. In B. Walczak, R. Tauler, & S. Brown (Vol. Eds.), *Comprehensive chemometrics: Vol. 3*, (pp. 587–599). Oxford: Elsevier.
Massart, D. L., Vandeginste, B. G. M., Buydens, S. J., Lewi, P. J., & Smeyers-Verbeke, J. (1998). *Handbook of chemometrics and qualimetrics.* Amsterdam: Part A. Elsevier.
Milanez, K. D. T. M., & Pontes, M. J. C. (2014). Classification of edible vegetable oil using digital image and pattern recognition techniques. *Microchemical Journal, 113*, 10–16. https://doi.org/10.1016/j.microc.2013.10.011.
Milanez, K. D. T. M., & Pontes, M. J. C. (2015). Classification of extra virgin olive oil and verification of adulteration using digital images and discriminant analysis. *Anal Methods, 7*, 8839–8846. https://doi.org/10.1039/C5AY01765C.
Pereira, P. M. C. C., & Vicente, A. F. R. B. (2013). Meat nutritional composition and nutritive role in the human diet. *Meat Science, 93*, 586–592. https://doi.org/10.1016/j.meatsci.2012.09.018.
Pérez-Palacios, T., Ruiz, J., Martín, D., Muriel, E., & Antequera, T. (2008). Comparison of different methods for total lipid quantification in meat and meat products. *Food Chemistry, 110*, 1025–1029. https://doi.org/10.1016/j.foodchem.2008.03.026.
Pierini, G. D., Fernandes, D. D. S., Diniz, P. H. G. D., Araújo, M. C. U., Di Nezio, M. S., & Centurión, M. E. (2016). A digital image-based traceability tool of the geographical origins of Argentine propolis. *Microchemical Journal, 128*, 62–67. https://doi.org/10.1016/j.microc.2016.04.015.
Santé-Lhoutellier, V. (2014). Meat and human diet: Facts and myths. *Encyclopedia of meat sciences. Reference module in food science* (pp. 118–123). (2nd ed.). . https://doi.org/10.1016/B978-0-12-384731-7.00178-1.
Soares, S. F. C., Gomes, A. A., Galvão Filho, A. R., Araújo, M. C. U., & Galvão, R. K. H. (2013). The successive projections algorithm. *Trends in Analytical Chemistry, 42*, 84–98. https://doi.org/10.1016/j.trac.2012.09.006.
Sørland, G. H., Larsen, P. M., Lundby, F., Rudi, A., & Guiheneuf, T. (2004). Determination of total fat and moisture content in meat using low field NMR. *Meat Science, 66*, 543–550. https://doi.org/10.1016/S0309-1740(03)00157-8.
Souto, U. T. C. P., Barbosa, M. F., Dantas, H. V., Pontes, A. S., Lyra, W. S., Diniz, P. H. G. D., et al. (2015). Screening for coffee adulteration using digital images and SPA–LDA. *Food Anal Methods, 8*, 515–1521. https://doi.org/10.1007/s12161-014-0020-7.
USDA (United States Department of Agriculture) (2015). *Argentina poultry & products annual.* https://gain.fas.usda.gov/Recent%20GAIN%20Publications/Poultry%20and%20Products%20Annual_Buenos%20Aires_Argentina_9-18-2015.pdf, Accessed date: 13 March 2018.
Valous, N. A., Mendoza, F., Sun, D., & Allen, P. (2009). Colour calibration of a laboratory computer vision system for quality evaluation of pre-sliced hams. *Meat Science, 81*, 132–141. https://doi.org/10.1016/j.meatsci.2008.07.009.
Weeranantanaphan, J., Downey, G., Allen, P., & Sun, D. (2011). A review of near infrared spectroscopy in muscle food analysis: 2005–2010. *Journal of Near Infrared Spectroscopy, 19*, 61–104. https://doi.org/10.1255/jnirs.924.
Zamora-Rojas, E., Garrido-Varo, A., de Pedro-Sanz, E., Guerrero-Ginel, J. E., & Pérez-Marín, D. (2011). Monitoring NIRS calibrations for use in routine meat analysis as part of Iberian pig-breeding programs. *Food Chemistry, 129*, 1889–1897. https://doi.org/10.1016/j.foodchem.2011.05.139.
Zapotoczny, P., Szczypiński, P. M., & Daszkiewicz, T. (2016). Evaluation of the quality of cold meats by computer-assisted image analysis. *Lebensmittel-Wissenschaft und -Technologie- Food Science and Technology, 67*, 37–49. https://doi.org/10.1016/j.lwt.2015.11.042.
Zheng, C., Sun, D. W., & Zheng, L. (2006a). Correlating colour to moisture content of large cooked beef joints by computer vision. *Journal of Food Engineering, 77*, 858–863. https://doi.org/10.1016/j.jfoodeng.2005.08.013.
Zheng, C., Sun, D. W., & Zheng, L. (2006b). Classification of tenderness of large cooked beef joints using wavelet and Gabor textural features. *Transactions of the American Society of Agricultural and Biological Engineers, 49*, 1447–1454. https://doi.org/10.13031/2013.2202.