



# Adulteration Detection in Goat Dairy Beverage Through NIR Spectroscopy and DD-SIMCA

José Luan Da Paixão Teixeira<sup>1</sup> · Elem Tamirys Dos Santos Caramês<sup>1</sup> · Débora Parra Baptista<sup>2</sup> · Mirna Lúcia Gigante<sup>2</sup> · Juliana Azevedo Lima Pallone<sup>1</sup>

Received: 29 January 2021 / Accepted: 26 October 2021 / Published online: 6 November 2021  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

Bovine cheese whey presents physico-chemical similarity with goat cheese whey, making the occurrence of adulteration of goat dairy beverage by addition of bovine whey possible. This study aimed to develop a green analytical method, using NIR spectroscopy and Data Driven–Soft Independent Modelling of Class Analogies (DD-SIMCA), to verify the authenticity of goat dairy beverage and to detect adulteration with bovine whey. For this, 60 authentic samples of goat dairy beverage were used and 180 samples were adulterated with 10, 25, 50 and 100% bovine whey. The classification model built based on the full spectra (10,000–4000  $\text{cm}^{-1}$ ) showed 95 and 100% of right assignments for authentic and adulterated samples, respectively. After the variable selection process (5500–4000  $\text{cm}^{-1}$ ), the results indicated 100% of classification correct for the samples. In this sense, the combination of NIR and class modeling has proved to be efficient for controlling the authenticity of goat dairy beverages.

**Keywords** Cheese whey · Co-product · Near-infrared · Classification · Class-modeling

## Introduction

Traditionally, goat's milk is used for the production of fermented dairy products, especially cheese. Compared to the cow milk, goat milk exhibits smaller fat globules and a higher percentage of short- and medium-chain fatty acids (Haenlein and Anke 2011). This contributes to goat dairy products having greater digestibility and a healthier lipid metabolism when compared to cow milk products.

The cheese whey, which represents between 85 and 90% of the volume of milk (obtained after the cheese has been drained), has relevant nutritional characteristics; it contains approximately 55% of milk nutrients (lactose 4.5–5% w/v), soluble proteins (0.6–0.8% w/v), lipids (0.4–0.5% w/v) and salts (8–10% dry extract), and is also a source of lactic acid

(0.05% w/v) and hydrophilic vitamins (González-Martínez et al. 2002).

Due to these characteristics, the whey is considered an important co-product of the dairy industry and is used for the manufacturing of whey powder, whey protein concentrates and isolates, isolated protein fractions, bioactive peptides, lactose, whey butter and components of the milk fat globule membrane (Costa et al. 2010; Foegeding et al. 2011).

The production of dairy beverages represents an interesting use of cheese whey that does not require pre-processing. It results from the mixture of milk and whey, fermented or not, with or without the addition of other ingredients, and displays a milk base of at least 51% (w/w) of the total ingredients used (Brazil 2005). Bovine whey can easily and fraudulently replace goat whey due to similarities in its physicochemical properties. This adulteration is of particular concern, as many consumers looking for this product are typically allergic to cow milk proteins, leading to increased risk of numerous gastrointestinal disorders (Haenlein 2004).

Generally, the traditional methods used to control the physical–chemical quality of food are time consuming, require the use of many chemical reagents, generate toxic waste, destroy samples and are often unable to detect fraud in lower concentrations in other food matrices (de Sousa

✉ Juliana Azevedo Lima Pallone  
jpallone@unicamp.br

<sup>1</sup> Department of Food Science, School of Food Engineering, State University of Campinas, Street Monteiro Lobato, 80, Campinas, São Paulo CEP: 13.083-862, Brazil

<sup>2</sup> Department of Food Technology, School of Food Engineering, State University of Campinas, Street Monteiro Lobato, 80, Campinas, São Paulo CEP: 13.083-862, Brazil

Lobato et al. 2018). In contrast, the near-infrared spectroscopy (NIR) technique associated with chemometric tools has been successfully used to detect food fraud. The main positive points are that this technique requires minimal or no sample preparation, does not require chemical reagents and, after the construction of classification models or multivariate calibration, can provide results quickly and efficiently (Hein et al. 2016; Hwang et al. 2016).

NIR spectra present bands that correspond mainly to overtones and combinations of fundamental vibration of groups containing hydrogen (Chen et al. 2017). Therefore, the use of chemometric tools is essential to obtain information about the chemical composition of the samples and allows for food analysis, and is a green alternative to traditional methods (Cevoli et al. 2013; Ferrão et al. 2007; Xu et al. 2013).

Mathematical models of supervised pattern recognition that seek to assess the feasibility of using NIR to detect food fraud can be divided into discrimination methods and class modeling techniques (CMTs) (Xu et al. 2012). The discrimination methods are based on fitting a limit between two or more classes, to classify future samples in one of these classes (Chen et al. 2020). On the other hand, CMTs describe the boundary around a class and not between classes, using samples from only one specific class. Brereton (2011) reports that class modeling creates an acceptance area around the class of interest, rather than determining limits between classes such as the discriminant model approach; based on the created acceptance area, a future unknown sample can be identified according to the within/outside the acceptance limit criteria.

Through the use of CMTs, the target class can be modeled separately without changing the original model, resulting in more effective models when compared to traditional discriminant methods used to detect adulterations or to assure food authenticity (Oliveri et al. 2020; Rodionova et al. 2016b). There are several class modeling algorithms available, including Unequal Class Modeling (UNEQ) (Derde and Massart 1986), Soft Independent Modeling of Class Analogies (SIMCA) (Yang et al. 2013), and more recently, One-Class Partial Least Squares (OC-PLS) (Xu et al. 2013) and Data Driven–Soft Independent Modelling of Class Analogies (DD-SIMCA) (Pomerantsev and Rodionova 2013).

Studies using class modeling to detect fraud in various food matrices, including dairy products, have already been published (Chen et al. 2017; López et al. 2014; Xu et al. 2012). However, the use of NIR coupled to DD-SIMCA to detect adulterations in goat dairy beverages is an unprecedented study. In this context, the present work aims to study the feasibility of combining NIR spectroscopy and class modeling (DD-SIMCA algorithm) for identifying bovine whey in goat dairy beverage after the processing of Minas Frescal cheese.

## Material and Methods

### Samples and Obtaining the Goat and Cow Dairy Beverage

Authentic goat and cow milk were obtained directly from the producers. The milk samples were transported to the pilot plant immediately after milking, submitted to heat treatment (65 °C/30 min), cooled to 4 °C and stored in a cold chamber ( $4 \pm 1$  °C) until processing.

Minas Frescal cheese was manufactured in triplicate to obtain the goat and cow whey according to the protocol described by Diamantino et al. (2014), with modifications. The milk was heated to 35 °C and 250 ppm of 50% calcium chloride, 85% lactic acid (0.25 mL per L of milk, in 10% aqueous solution) and a coagulant (CHY-MAX Powder Extra NB, Chr. Hansen, Hoersholm, Denmark) was added in sufficient quantity to coagulate the milk was after 35 min. After coagulation, the gel was cut, with the aid of horizontal and vertical liras, into cubes of 1.5–2.0 cm. After resting for 5 min, slow stirring was performed for 30 min. The curd was then kept at rest for 10 min then draining of the curd commenced. The whey was collected, immediately cooled and kept refrigerated until dairy beverages were prepared.

Pasteurized goat milk was mixed with the obtained goat whey after the manufacture of Minas Frescal cheese in the proportions of 51% milk and 49% whey, following the Brazilian regulation about the composition of dairy beverage (Brazil 2005). A total of 60 (sixty) authentic samples (51% goat milk + 49% goat whey) of goat dairy beverage and 120 (one hundred and twenty) adulterated samples with 10, 25, 50 and 100% of cow cheese whey as goat whey replacement were produced for the study. Moreover, samples of cow dairy beverage (51% cow milk + 49% cow whey) were produced and used in the construction of the models, since they could be labeled/marketed as goat milk beverages, representing a health risk to consumers. The general scheme for obtaining the samples can be seen in Fig. 1.

The levels of adulteration were determined to test the sensitivity and specificity of the method applied, considering real situations in which counterfeiters substitute goat whey with cow whey to obtain profit without visual/sensory detection.

### Spectrum Acquisition on FT-NIR

Authentic goat dairy beverage samples were considered as positive controls, whereas all adulterated samples, in addition to cow dairy beverage, were part of the negative control.

To obtain the FT-NIR spectra, about 2 mL of the dairy beverage samples was placed in a transmittance accessory,

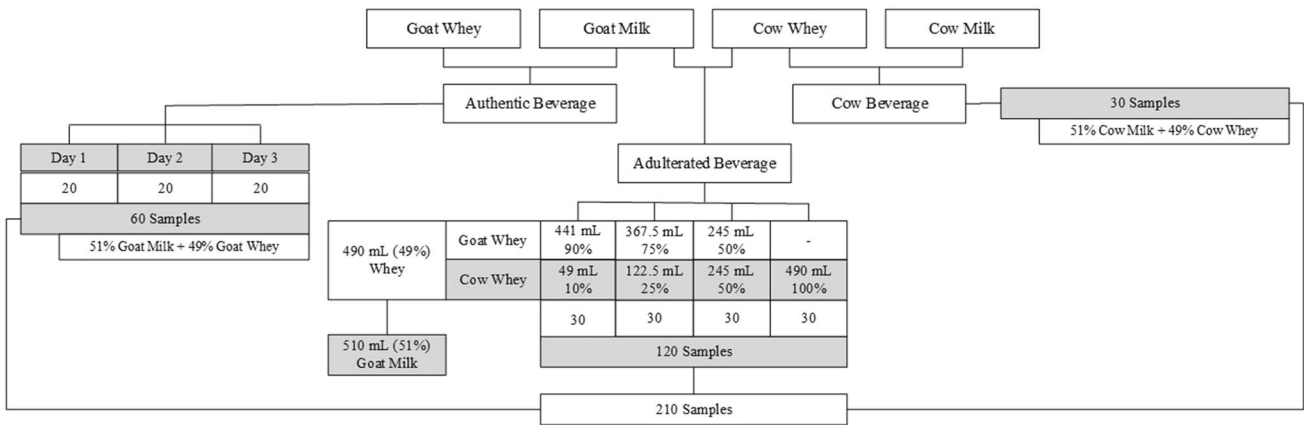


Fig. 1 Flowchart for obtaining authentic and adulterated goat dairy beverage samples

consisting of a Petri dish and the liquid reflector (PerkinElmer, Waltham, EUA, part number L118-0503). To collect the spectra, the FT-NIR spectrometer was used (PerkinElmer, Waltham, EUA, model Spectrum 100 N) with a spectral range between 10,000 and 4000  $\text{cm}^{-1}$ , with resolution of 4  $\text{cm}^{-1}$  being used, and 32 scans obtained from spectra.

### Chemometric

The chemometric analysis was performed using the MATLAB R2018b software coupled to the PLS-toolbox version 8.6 (Eingevector Research Inc. 2010). Initially, an unsupervised method of principal component analysis (PCA) was used for all authentic and adulterated samples, in order to carry out an exploratory analysis of the spectra information obtained.

In this work, we used an advanced algorithm data-oriented SIMCA class modeling (DD-SIMCA) (Zontov et al. 2017). According to these authors, this algorithm theoretically allows to calculate incorrect classification rates, based on a principal component analysis (PCA) of the predictive matrix  $X (I \times J)$  of the target class, where  $I$  and  $J$  denote the number of samples and variables, respectively.

The DD-SIMCA offers the possibility to estimate distribution parameters oriented to the data, making it possible to create a decision rule/acceptance area for a given type I  $\alpha$  error, that is, the portion of false negative decisions. In addition, a second cut-off point is defined to limit the outer edge to a given  $y$  value, which specifies the possibility that at least one regular sample is mistakenly considered an extreme value. Detailed formulas related to the calculation of the acceptance area for the parameter values provided can be found in Rodionova et al. (2016a).

For the construction of the DD-SIMCA models, the 210 samples of authentic and adulterated dairy beverages were

randomly divided into two (2) sets: calibration and external validation. Thus, the training set was composed by 40 randomly selected authentic samples (2/3 of the authentic samples), while the test set, for external validation, was composed of 20 authentic samples (1/3 of the authentic samples), 120 of adulterated dairy beverage samples and 30 cow dairy beverages. The addition of authentic samples in the non-directed class during the validation step is important to test the possibility of false positives by the proposed model, in other words, to test the capacity of the model to distinguish between an unknown group of samples, the ones belonging to the target class or not.

To assess the quality of the models obtained, sensitivity and specificity values were determined according to Eqs. 4 and 5 (Barra et al. 2019).

$$Sensitivity(\%) = \frac{TP}{TP + FN} * 100 \tag{1}$$

$$Specificity(\%) = \frac{TN}{TN + FP} * 100 \tag{2}$$

where TP means true positive, TN means true negative, FN is false negative and FP is false positive.

The data, previously mean-centered, were pre-processed using the methods Multiplicative Scatter Correction (MSC), Standard Normative Variance (SNV), 1stDV and 2ndDV of Savitzky-Golay and/or Smoothing, applied in order to correct radiation scattering and baseline deviation.

According to Shrestha et al. (2017), the selection and incorporation of the most relevant spectral regions improve the performance of the models. Thus, in addition to analyzing the spectra obtained in 10,000–4000  $\text{cm}^{-1}$ , the selection of variables was also carried out, considering the loadings obtained after performing the PCA, capable of revealing the main wavelengths responsible for

discrimination between authentic and adulterated samples of goat dairy beverage with cow whey.

## Results and Discussion

### Analysis of FT-NIR Spectral Data

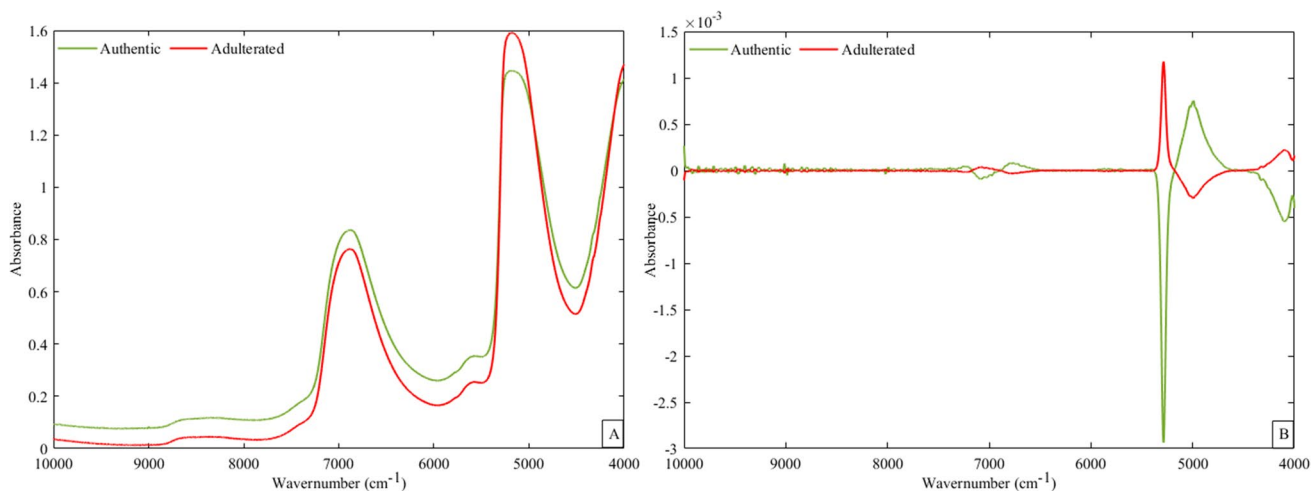
Observing the obtained spectra (Fig. 2A), two bands stand out, 6900 and 5200  $\text{cm}^{-1}$ , which are associated with the O–H connections of the water present in the samples. Da Paixão Teixeira et al. (2020) and Xu et al. (2013) also used the FT-NIR resources combined with chemometrics and correlated these same spectral bands (6900 and 5200  $\text{cm}^{-1}$ ) associated with OH bonds in the analysis of authentic and adulterated samples of goat milk and in the survey of spectral profiles of cow yogurt Chinese, respectively. This spectral profile can be easily found in different matrices with high water concentrations, such as in frozen guava and yellow passion fruit pulp, as well as in grape juice (Alamar et al. 2016; Musingarabwi et al. 2016).

The spectra showed dispersion of electromagnetic radiation and deviation from baselines, typically found in NIR. Thus, to reduce these effects, several pre-processing techniques were tested and the best performance was obtained through the smoothing of the first derivative of Savitzky-Golay (1stDV). The effect of using 1stDV on the authentic and adulterated spectra of goat dairy beverage can be seen in Fig. 2B. Comparing the average spectra of authentic and adulterated samples after the application of pre-processing, it can be seen that the numbers of wavelengths 7070, 6753, 5284, 4989 and 4087  $\text{cm}^{-1}$  were mainly responsible for the clustering observed between authentic and adulterated samples of goat dairy beverage (Fig. 2B).

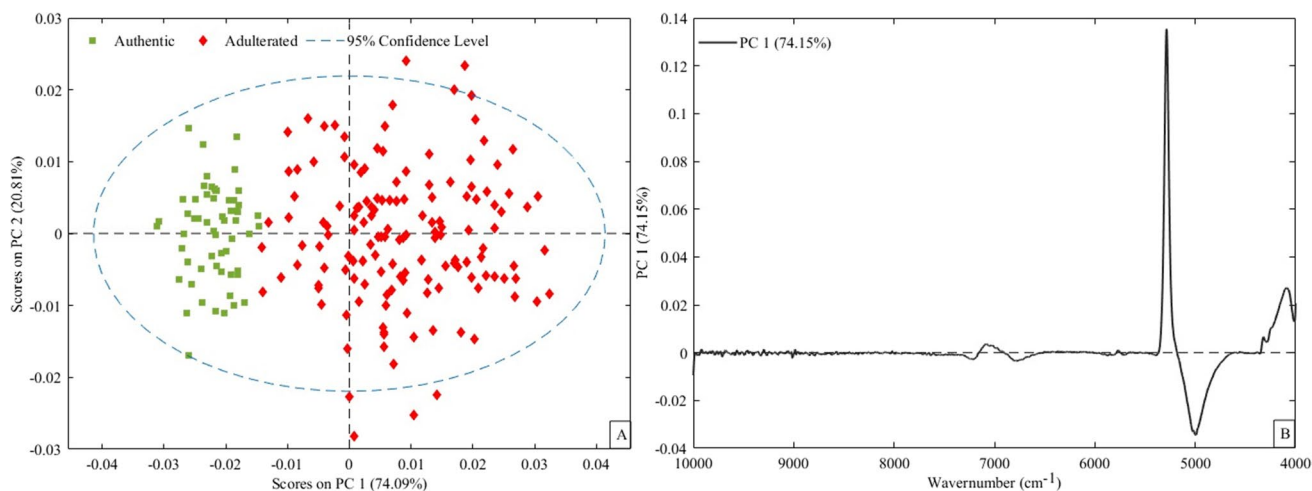
The bands at 6753 and 5284  $\text{cm}^{-1}$  are related to the OH bonds and the water content present in the samples and after applying the pre-processing, it was also possible to verify differences in the 4989  $\text{cm}^{-1}$  bands that can be associated with the protein content, such as amides (dos Santos Pereira et al. 2020). The 7070  $\text{cm}^{-1}$  region that corresponds to the first O–H and N–H overtone and the 4087  $\text{cm}^{-1}$  regions are related to the C–H + C–C combination region (Osborn et al. 1993; dos Santos Pereira et al. 2020).

Figure 3A shows the PCA using pre-processed spectra based at 1stDV of Savitzky-Golay. The PCA model was established with four (4) PCs. In Fig. 3A, the PC1xPC2 scores are found which represented 94.9% of the total accumulated variance. As can be seen in Fig. 3A, authentic and adulterated samples of goat dairy beverage showed a tendency to cluster, indicating that the PCA presented relevant information to identify the similarities responsible for the clustering between the sample classes.

When looking at the loadings presented by PC1 (Fig. 3B), we noticed again, as expected, that the wavelengths at 5284, 4989 and 4087  $\text{cm}^{-1}$  were the most important for discrimination between authentic and adulterated samples, corroborating the result previously presented after the use of 1stDV as pre-processing of the spectra (Fig. 2B). The band close to 5200  $\text{cm}^{-1}$  is in the first OH overtone region and in cases of matrices with a high amount of water (as is the case of goat and cow milk samples), we can correlate it to this compound; but this spectral region can also be attributed to the first overtone of the bond CONH2-R (Eldin 2011; Osborn et al. 1993). The spectral region close to 5000  $\text{cm}^{-1}$  can be associated with the combination of N–H groups and the region of the band close to 4087  $\text{cm}^{-1}$  is related to the combinations of  $\text{CH}_2$



**Fig. 2** Authentic and adulterated medium FT-NIR spectra of goat dairy beverage (A) and the smoothing effect (B) of the first Savitzky-Golay derivative (1stDV) on the goat dairy beverage spectra



**Fig. 3** Principal component analysis (PCA) (A) and loadings of PC1 after sample pre-processing (B)

bonds that have already been associated with chain groups linked to proteins (Núñez-Sánchez et al. 2016).

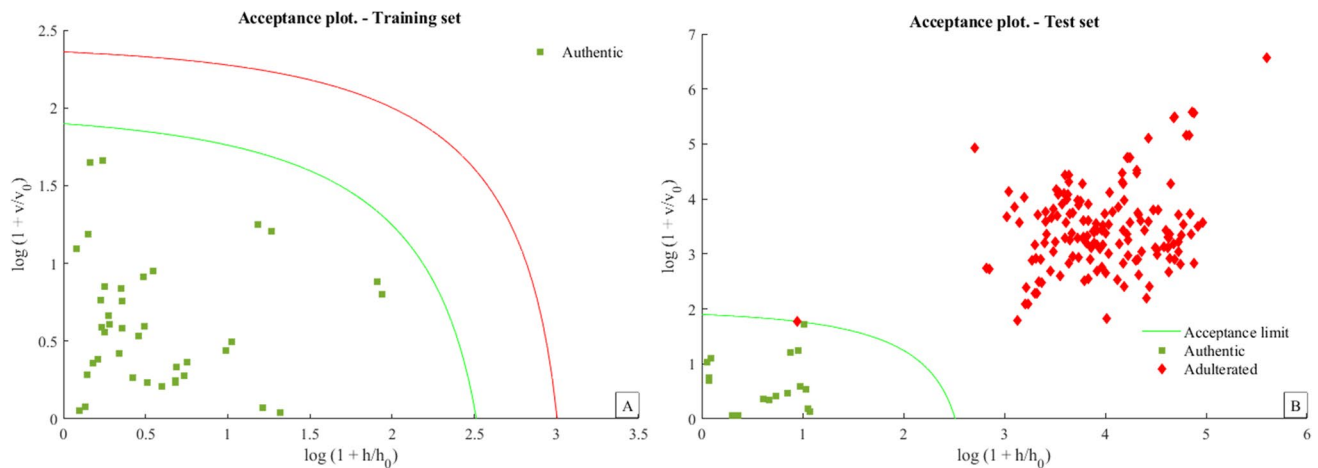
Although the total protein content presented by goat's milk is similar to that reported by cow's milk, differences related to the lower level of  $\alpha_{s1}$ -casein and higher level  $\beta$ -casein in goat's milk when compared to cow's milk can be verified (Amigo & Fontecha 2011; Clark and Mora García 2017; Harish Kumar et al. 2016; Picon et al. 2019). Harish Kumar et al. (2016) also reinforced that what differentiates goat milk protein from cow milk is the structure of the casein micelle, the wide variety of bioactive peptides in these fractions and the presence of non-protein nitrogen compounds such as amino acids, nucleotides and nucleosides.

Silanikove et al. (2010) and Haenlein (2004) also reported that the casein micelles found in goat's milk are larger than those in cow's milk (in general, they vary from 100 to 200 nm compared to 60 to 80 nm) and have different rates of sedimentation, solubilization, thermal stability, and contents of calcium and phosphorus. Besides the differences in casein fraction compositions and micelle structures, different protein compositions presented by goat and cow whey also contribute to the differences observed, since whey corresponded to 49% of the dairy beverage's formulations evaluated in the present study. Goat  $\beta$ -lactoglobulins ( $\beta$ -LG) have lower net charge when compared to  $\beta$ -LGs present in cow's milk. Similarly, even though  $\alpha$ -lactalbumins ( $\alpha$ -LA) in goat's and cow's milk are closely homologous,  $\alpha$ -LAs from the two species also differ in the composition of charged amino acids (Amigo & Fontecha 2011). These differences in protein composition between goat and cow milk proteins are characterized as probably the main factor responsible for the differentiation between authentic and adulterated samples of goat dairy beverage with cow's serum, according to the trend verified in the PCA.

### Modeling of Single Class DD-SIMCA

Figure 4A shows the chi-square acceptance area of the DD-SIMCA model of the target class for the data set, where the full spectral range was considered (10,000–4000  $\text{cm}^{-1}$ ). The green curve limits the acceptance area, while the red curve limits the external area. In this plotting, each sample in the training set is marked as a green dot on the acceptance plot and is considered a “regular” sample; conversely, the sample marked by a dot between the green and red lines will be considered as an “extreme” sample (Chen et al. 2020). In our study, extreme samples were not detected. It should be emphasized that the red curve is the outer edge of the given  $y$  value, which controls the likelihood that at least one regular sample will be erroneously classified as external.

On the other hand, Fig. 4B provides the predictive performance of the DD-SIMCA model for the set of tests considering the full spectral range. Analyzing the 170 samples used as a non-segmented class (20 authentic samples, 120 adulterated samples and 30 samples of cow dairy beverage), we can observe that only 1 sample of authentic dairy beverage was incorrectly classified as an adulterated sample (non-segmented class). Thus, the percentage of correct answers was 95% for authentic samples and 100% of adulterated samples were correctly rejected (considered non-authentic) by the model of a class. In our work, the DD-SIMCA model was built using 2 PCs ( $\alpha=0.01$ ) and the applied beta value was 0.39%. Rodionova et al. (2016a) discriminated the authenticity of olive samples and used 6 PCs to provide 98% specificity ( $\alpha=0.05$ ) in class modeling. On the other hand, Mazivila et al. (2020) used 30 authentic samples of cow milk powder (20 in the training set and 10 in the test set) and 20 samples adulterated with sucrose (in a concentration range of 1 to 3%) and melanin (in a concentration range of 0.8 to 2%) (non-segmented class), reporting that the best performance



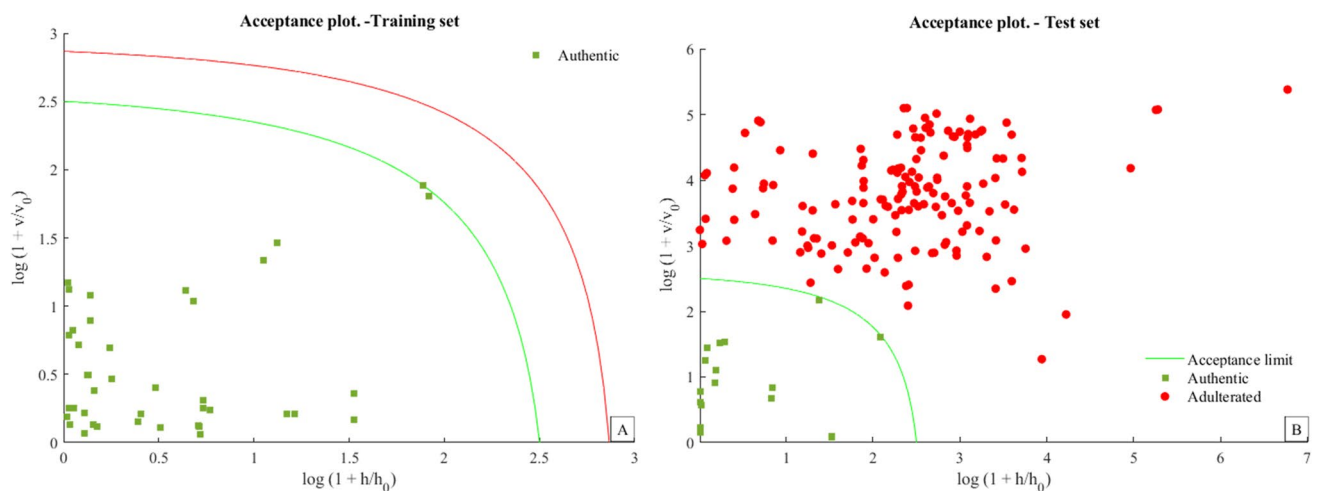
**Fig. 4** Acceptance area of the target class chi-square (A) and predictive performance in the test set (B) of the DD-SIMCA model using the entire spectral region

of the DD-SIMCA model was obtained with the use of 3 PCs. These authors even reported that the modeling of a class was successful, since no extreme or external samples were found in the training set and all 10 samples in the authentic milk powder test set were also correctly classified as members of the target class; additionally, all 20 samples adulterated with sucrose and melanin (non-segmented class) were correctly classified as not belonging to the target class.

In order to improve the performance of the classification model, selection of variables was performed, according to the loadings of PC1, obtained during the PCA analysis. In this context, only the region between 5500 and 4000  $\text{cm}^{-1}$  was selected, as it is the region responsible for discriminating authentic and adulterated samples of goat dairy beverage, carrying out a new class modeling by DD-SIMCA

using the same data set. Figure 5A shows the chi-square acceptance area of the DD-SIMCA model of the target class, for the data set with variable selection. In this case, the model was created with only 1 principal component, the  $\alpha$  value established was 0.01 and no sample was classified as extreme. Analyzing the predictive performance of the DD-SIMCA model with variable selection (Fig. 5B), it is also observed that no sample of the segmented class was incorrectly classified as a non-segmented class and all samples of the non-segmented class were correctly classified, with a beta value 0.13%.

Table 1 provides the confusion matrix for the prediction of the DD-SIMCA models using the full spectral range (10,000–4000  $\text{cm}^{-1}$ ) and with the selected variables (5500–4000  $\text{cm}^{-1}$ ). The sensitivity and specificity values



**Fig. 5** Acceptance area of the target class chi-square (A) and predictive performance in the test set (B) of the DD-SIMCA model obtained after variable selection

**Table 1** Confusion matrices of the results of the prediction of the DD-SIMCA models

		Authentic	Adulterated
10,000—4000 cm <sup>-1</sup>	Prediction as authentic	19	0
	Prediction as adulterated	1	150
	Sensitivity (%)	95	100
	Specificity (%)	100	95
5500—4000 cm <sup>-1</sup>	Prediction as authentic	20	0
	Prediction as adulterated	0	150
	Sensitivity (%)	100	100
	Specificity (%)	100	100

obtained by the DD-SIMCA model using the full spectral range and with selected variables showed excellent predictive performance in the classification of authentic and adulterated samples of goat dairy beverage, with 95 and 100%, respectively.

Therefore, the performance of the model with selected variables was superior when compared to the model previously presented in Fig. 4 (A and B), using full spectral data; although more information is available in full range spectra, it also has more noise and information that could not be relevant to the studied objective. Chen et al. (2020) indicated that the combination of NIR spectroscopy with variable selection and class modeling was feasible to identify frauds in a drug adulterated with *Sophora flavescens* in powder and corn flour. Maximum sensitivity and specificity values of 92 and 100% were reported, respectively, related to the correct identification of the most important spectral regions present in their global spectra.

## Conclusions

The results obtained during this work showed an excellent potential of NIR spectroscopy for the creation of DD-SIMCA models with efficient discriminatory power between authentic samples of goat dairy beverage and samples adulterated with cow whey. The variable selection was carried out through the use of the information obtained by the loadings graph of the PCA, where the improvement in the final performance of the proposed DD-SIMCA model was observed, since the reduction of the spectral range also led to a reduction the presence of noise and/or irrelevant or redundant information for the desired objective. As a suggestion, the results obtained through the use of a reduced spectral range represent a possibility to carry out future studies that aim to use a portable, cheaper and accessible NIR device,

capable of detecting adulterations in goat dairy beverages. In this sense, this procedure is a promising alternative to the traditional methods used to detect fraud in foods that are usually laborious, time-consuming and expensive.

**Author Contribution** José Luan da Paixão Teixeira performed the experimental work and wrote the paper. Elem Tamirys dos Santos Caramês performed the chemometric analysis and wrote part of the paper. Débora Parra Baptista performed chemical analysis to assure the quality of all samples used during the paper and helped experimental work. Juliana Azevedo Lima Pallone and Mirna Lúcia Gigante revised and edited the paper and supervised the experimental work.

**Funding** The authors would like to thank the São Paulo Research Foundation (FAPESP) for the scholarship grant of PhD student José Teixeira and funding of the research project (scholarship no. 2018/08864–8 and financial support—regular no. 2018/09759–3). They would also like to thank the Coordination for the Improvement of Higher Education Personnel (CAPES) (Financial Code 001) and the National Council for Scientific and Technological Development (CNPq) for the doctoral scholarship of Elem Caramês (scholarship no. 142414/2016–6) and Débora Baptista (scholarship no. 140739/2016–5).

## Declarations

**Informed Consent** Not applicable.

**Conflict of Interest** José Luan da Paixão Teixeira declares that he has no conflict of interest. Elem Tamirys dos Santos Caramês declares that no conflict of interest. Débora Parra Baptista declares that she has no conflict of interest. Mirna Lúcia Gigante declares that she has no conflict of interest. Juliana Azevedo Lima Pallone declares that she has no conflict of interest.

## References

- Alamar PD, Caramês ETS, Poppi RJ, Pallone JAL (2016) Quality evaluation of frozen guava and yellow passion fruit pulps by NIR spectroscopy and chemometrics. *Food Res Int* 85:209–214. <https://doi.org/10.1016/j.foodres.2016.04.027>
- Amigo L, Fontecha J (2011) Goat Milk. In: Fuquay JW, Fox PF, McSweeney PLH (eds) *Encyclopedia of dairy sciences*, 2nd edn. Elsevier Academic Press, London, pp 484–493
- Barra I, Mansouri MA, Cherrah Y, Kharbach M, Bouklouze A (2019) FTIR fingerprints associated to a PLS-DA model for rapid detection of smuggled non-compliant diesel marketed in Morocco. *Vib Spectrosc* 101:40–45. <https://doi.org/10.1016/j.vibspec.2019.02.001>
- Brazil (2005) Ministério da Agricultura, Pecuária e Abastecimento – MAPA – Instrução Normativa n° 16, de 23 de agosto de 2005, Regulamento Técnico de Identidade e Qualidade de Bebida Láctea. *Diário Oficial da União*. Brasília, Distrito Federal
- Brereton RG (2011) One-class classifiers. *J Chemom* 25:225–246. <https://doi.org/10.1002/cem.1397>
- Cevoli C, Gori A, Nocetti M, Cuiibus L, Caboni MF, Fabbri A (2013) FT-NIR and FT-MIR spectroscopy to discriminate competitors, non compliance and compliance grated Parmigiano Reggiano cheese. *Food Res Int* 52:214–220. <https://doi.org/10.1016/j.foodres.2013.03.016>

- Chen H, Tan C, Lin Z, Wu T (2017) Detection of melamine adulteration in milk by near-infrared spectroscopy and one-class partial least squares. *Spectrochim Acta A Mol Biomol Spectrosc* 173:832–836. <https://doi.org/10.1016/j.saa.2016.10.051>
- Chen H, Tan C, Li H (2020) Untargeted identification of adulterated Sanqi powder by near-infrared spectroscopy and one-class model. *J Food Compos Anal* 88:103450. <https://doi.org/10.1016/j.jfca.2020.103450>
- Clark S, Mora García MB (2017) A 100-Year Review: Advances in goat milk research. *J Dairy Sci* 100:10026–10044. <https://doi.org/10.3168/jds.2017-13287>
- Costa MR, Elias-Argote XE, Jiménez-Flores R, Gigante ML (2010) Use of ultrafiltration and supercritical fluid extraction to obtain a whey buttermilk powder enriched in milk fat globule membrane phospholipids. *Int Dairy J* 20:598–502. <https://doi.org/10.1016/j.idairyj.2010.03.006>
- Da Paixão Teixeira JL, dos Caramês ET, S, Baptista DP, et al (2020) Vibrational spectroscopy and chemometrics tools for authenticity and improvement the safety control in goat milk. *Food Control* 112:1–8. <https://doi.org/10.1016/j.foodcont.2020.107105>
- de Sousa Lobato KB, Alamar PD, dos Santos Caramês ET, Pallone JAL (2018) Authenticity of freeze-dried açai pulp by near-infrared spectroscopy. *J Food Eng* 224:105–111. <https://doi.org/10.1016/j.jfoodeng.2017.12.019>
- Derde MP, Massart DL (1986) UNEQ: a disjoint modelling technique for pattern recognition based on normal distribution. *Anal Chim Acta* 184:33–51. [https://doi.org/10.1016/S0003-2670\(00\)86468-5](https://doi.org/10.1016/S0003-2670(00)86468-5)
- Diamantino VR, Beraldo FA, Sunakozawa TN, Penna ALB (2014) Effect of octenyl succinylated waxy starch as a fat mimetic on texture, microstructure and physicochemical properties of Minas fresh cheese. *LWT - Food Sci Technol* 56:356–362. <https://doi.org/10.1016/j.lwt.2013.12.001>
- dos Santos Pereira EV, de Sousa Fernandes DD, de Araújo MCU, Diniz PHGD, Maciel MIS (2020) Simultaneous determination of goat milk adulteration with cow milk and their fat and protein contents using NIR spectroscopy and PLS algorithms. *LWT* 127:109427. <https://doi.org/10.1016/j.lwt.2020.109427>
- Eingevector Research Inc. PLS Toolbox 8.6. Manson, USA
- Eldin AB (2011) Near infra-red spectroscopy. In: *Wide Spectra of Quality Control*. InTech, pp 237–248. <https://doi.org/10.5772/24208>
- Ferrão MF, Mello C, Borin A, Maretto DA, Poppi RJ (2007) LS-SVM: Uma nova ferramenta quimiométrica para regressão multivariada. Comparação de modelos de regressão LS-SVM e PLS na quantificação de adulterantes em leite em pó empregando NIR. *Quim Nova* 30:852–859. <https://doi.org/10.1590/S0100-4042007000400018>
- Foegeding EA, Luck P, Vardhanabhati B (2011) Milk protein products | whey protein products. In: Fuquay JW, Fox PF, McSweeney PLH (eds) *Encyclopedia of dairy sciences*, 2nd edn. Elsevier Academic Press, London, pp 873–878
- González-Martínez C, Becerra M, Cháfer M, Albors A, Carot JM, Chiralt A (2002) Influence of substituting milk powder for whey powder on yoghurt quality. *Trends Food Sci Technol* 13:334–340. [https://doi.org/10.1016/S0924-2244\(02\)00160-7](https://doi.org/10.1016/S0924-2244(02)00160-7)
- Haenlein GFW (2004) Goat milk in human nutrition. *Small Rumin Res* 51:155–163. <https://doi.org/10.1016/j.smallrumres.2003.08.010>
- Haenlein GFW, Anke M (2011) Mineral and trace element research in goats: a review. *Small Rumin Res* 95:2–19. <https://doi.org/10.1016/j.smallrumres.2010.11.007>
- Hein PRG, Chaix G, Clair B, Brancheriau L, Gril J (2016) Spatial variation of wood density, stiffness and microfibril angle along Eucalyptus trunks grown under contrasting growth conditions. *Trees Struct Funct* 30:871–882. <https://doi.org/10.1007/s00468-015-1327-8>
- Hwang SW, Horikawa Y, Lee WH, Sugiyama J (2016) Identification of Pinus species related to historic architecture in Korea using NIR chemometric approaches. *J Wood Sci* 62:156–167. <https://doi.org/10.1007/s10086-016-1540-0>
- Kumar H, Yadav D, Naveen Kumar RS, AKG, (2016) Nutritional and nutraceutical properties of millets: a review. *Indian J Dairy Sci* 5:513–518
- López MI, Trullols E, Callao MP, Ruisánchez I (2014) Multivariate screening in food adulteration: untargeted versus targeted modelling. *Food Chem* 147:177–181. <https://doi.org/10.1016/j.foodchem.2013.09.139>
- Mazivila SJ, Páscoa RNMJ, Castro RC, Ribeiro DSM, Santos JLM (2020) Detection of melamine and sucrose as adulterants in milk powder using near-infrared spectroscopy with DD-SIMCA as one-class classifier and MCR-ALS as a means to provide pure profiles of milk and of both adulterants with forensic evidence: a short communic. *Talanta* 216:120937. <https://doi.org/10.1016/j.talanta.2020.120937>
- Musingarabwi DM, Nieuwoudt HH, Young PR, Eyéghè-Bickong HA, Vivier MA (2016) A rapid qualitative and quantitative evaluation of grape berries at various stages of development using Fourier-transform infrared spectroscopy and multivariate data analysis. *Food Chem* 190:253–262. <https://doi.org/10.1016/j.foodchem.2015.05.080>
- Núñez-Sánchez N, Martínez-Marín AL, Polvillo O, Fernández-Cabanás VM, Carrizosa J, Urrutia B, Serradilla JM (2016) Near infrared spectroscopy (NIRS) for the determination of the milk fat fatty acid profile of goats. *Food Chem* 190:244–252. <https://doi.org/10.1016/j.foodchem.2015.05.083>
- Oliveri P, Malegori C, Mustorgi E, Casale M (2020) Qualitative pattern recognition in chemistry: theoretical background and practical guidelines. *Microchem J* 162:105725. <https://doi.org/10.1016/j.microc.2020.105725>
- Osborn BG, Fearn T, Hindle PH (1993) *Theory of near infrared spectroscopy Practical NIR spectroscopy with applications in food and beverage analysis* (2nd ed.). Longman Singapore Publishers (Pte) Ltd, London
- Picon A, López-Pérez O, Torres E, Garde S, Nuñez M (2019) Contribution of autochthonous lactic acid bacteria to the typical flavour of raw goat milk cheeses. *Int J Food Microbiol* 299:8–22. <https://doi.org/10.1016/j.ijfoodmicro.2019.03.011>
- Pomerantsev AL, Rodionova OY (2013) Concept and role of extreme objects in PCA/SIMCA. *J Chemom* 28:429–438. <https://doi.org/10.1002/cem.2506>
- Rodionova OY, Oliveri P, Pomerantsev AL (2016a) Rigorous and compliant approaches to one-class classification. *Chemom Intell Lab Syst* 159:89–96. <https://doi.org/10.1016/j.chemolab.2016.10.002>
- Rodionova OY, Titova AV, Pomerantsev AL (2016b) Discriminant analysis is an inappropriate method of authentication. *TrAC - Trends Anal Chem* 78:17–22. <https://doi.org/10.1016/j.trac.2016.01.010>
- Shrestha S, Deleuran LC, Gislum R (2017) Separation of viable and non-viable tomato (*Solanum lycopersicum* L.) seeds using single seed near-infrared spectroscopy. *Comput Electron Agric* 142:348–355. <https://doi.org/10.1016/j.compag.2017.09.004>
- Silanikove N, Leitner G, Merin U, Prosser CG (2010) Recent advances in exploiting goat's milk: quality, safety and production aspects. *Small Rumin Res* 89:110–124. <https://doi.org/10.1016/J.SMALLRUMRES.2009.12.033>
- Xu L, Ye ZH, Yan SM, Shi PT, Cui HF, Fu XS, Yu XP (2012) Combining local wavelength information and ensemble learning to enhance the specificity of class modeling techniques: identification of food geographical origins and adulteration. *Anal Chim Acta* 754:31–38. <https://doi.org/10.1016/j.aca.2012.10.011>



- Xu L, Yan SM, Cai CB, Wang ZJ, Yu XP (2013) The feasibility of using near-infrared spectroscopy and chemometrics for untargeted detection of protein adulteration in yogurt: Removing unwanted variations in pure yogurt. *J Anal Methods Chem* 2013. <https://doi.org/10.1155/2013/201873>
- Yang IC, Tsai CY, Hsieh KW et al (2013) Integration of SIMCA and near-infrared spectroscopy for rapid and precise identification of herbal medicines. *J Food Drug Anal* 21:268–278. <https://doi.org/10.1016/j.jfda.2013.07.008>
- Zontov YV, Rodionova OY, Kucheryavskiy SV, Pomerantsev AL (2017) DD-SIMCA – A MATLAB GUI tool for data driven

SIMCA approach. *Chemom Intell Lab Syst* 167:23–28. <https://doi.org/10.1016/j.chemolab.2017.05.010>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.