

Full Length Research Paper

Detection and quantification of cow milk adulteration using portable near-infrared spectroscopy combined with chemometrics

Mohammed Abdallah Musa^{1,2} and Shuming Yang^{1*}

¹State Key Laboratory of Agro-Product Quality and Safety, Institute of Quality Standards and Testing Technology for Agro-products, Chinese Academy of Agricultural Sciences, Beijing 100081, P. R. China.

²Department of Dairy Production, Faculty of Animal Production, University of Sinnar, Abu Naama, Sinnar State, Sudan.

Received 8 November, 2020; Accepted 4 December, 2020

Milk adulteration is a common phenomenon in many countries, which draws extensive attention from humans due to health hazards that might result in some fatal diseases. In this study, a portable near-infrared (NIR) spectrometer combined with multivariate analysis was used to detect and quantify milk adulteration. Fresh cow milk samples were collected from eight dairy farms in Beijing and Hebei province of China. Water, urea, starch and goat milk were used to adulterate milk at 11 different concentrations. The data driven soft independent modeling of class analogy (DD-SIMCA) method was employed for qualitative analysis. Partial least squares regression (PLSR) was applied for statistical analysis of the obtained NIR spectral data. The results showed that the DD-SIMCA approach achieved satisfactory classification. By the PLSR model, standard error of prediction (SEP) values of 4.35, 0.34, 4.74 and 5.56 g/L were obtained for water, urea, starch and goat milk, respectively. These results demonstrated the feasibility and reliability of NIR spectroscopy combined with multivariate analysis in the prediction of the total contents of the investigated adulterants in cow milk.

Key words: Portable near-infrared (NIR)-spectroscopy, milk adulteration, DD-SIMCA, Partial least squares (PLS) regression.

INTRODUCTION

Milk adulteration is usually conducted to meet the regulatory requirements while lowering the milk quality by substitution of cheap substances, admixture or extraction of valuable milk components (Poonia et al., 2017). It poses serious threats to human health and becomes a global concern, particularly in developing countries. The

possible reasons for milk adulteration might be a high demand of milk by all ages, easy adulteration operations and lack of feasible and accurate detection tools (Kamthania et al., 2014).

As a nutritionally balanced mixture and perishable food, milk has attracted the interest of many researchers. It has

*Corresponding author. E-mail: yangshuming@caas.cn.

Table 1. Percentages of different adulterants added to milk samples.

Level No.	Adulterants			
	Water (%)	Goat milk (%)	Starch (%)	Urea (%)
1	0.0	0.0	0.0	0.0
2	10	10	5	0.5
3	20	20	10	0.7
4	30	30	15	0.9
5	40	40	20	1
6	50	50	25	2
7	5	5	30	1.5
8	15	15	35	1.7
9	25	25	40	1.9
10	35	35	45	2.5
11	45	45	50	2.7

been characterized as a valuable source of fat, protein, carbohydrates, vitamins, and minerals (Faraz et al., 2013). The relatively low costs and high nutritional values contribute a significant part to human diets for milk. However, increased demand makes it a highly possible target for adulteration (Salih et al., 2017). According to Moore et al. (2012), milk is among the seven common food items most susceptible to adulteration. Many additives and food ingredients such as starch, rice flour, skim milk powder, whey powder, reconstituted milk, salt, vegetable oil, animal fat, glucose, melamine and urea have been used as milk adulterants to increase the thickness and viscosity of milk and to maintain the composition of fat, protein and carbohydrates (Campos Motta et al., 2014; Singuluri and Sukumaran, 2014; Soomro et al., 2014).

Due to the common practice of milk adulteration and its consequences on human health, it is necessitated to develop a suitable method for adulteration detection. So far, many types of techniques have been investigated for quantitative detection of adulterants in milk. These techniques are challenged by the complexity, time inefficiency and high costs. Near-infrared (NIR) spectroscopy is one of the techniques that have recently been used for food quality control. It has been reported for the prediction of quality parameters in oranges (Cayuela and Weiland, 2010), monitoring of oil oxidative stability (Allendorf et al., 2012), quantification of transfat in edible oils (Birkel and Rodriguez-Saona, 2011) and mineral fortification in whole-grain cornmeal (Hassel and Rodriguez-Saona, 2012). This approach has been used to determine several milk components including protein, fat and lactose (Etzion et al., 2004; Kawasaki et al., 2008); however, no study describes its use in monitoring milk authenticity. NIR spectroscopy is distinguished by its simplicity and lower costs, which make it an ideal tool for rapid screening. Recently, industrial companies of NIR spectroscopy have developed portable NIR systems, which are fast, simple and precise (Kim et al., 2008).

To our knowledge, no literature has been found concerning a portable NIR analyzer in detecting adulterants in milk. This study therefore aims to investigate the performance and evaluate the feasibility of portable NIR spectrometers combined with multivariate analysis. Water, starch, goat milk and urea were used as adulterants for the investigation. Chemometrics, utilized as a Microsoft Excel add-in, was employed as a novel method with unique and powerful features, which was enabled to perform the most types of multivariate analysis in chemistry to help understand data and build models. Qualitative and quantitative models were developed using pattern recognition techniques including data driven soft independent modeling of class analogy (DD-SIMCA) and partial least squares regression (PLSR), respectively.

MATERIALS AND METHODS

Sample preparation

Fresh cow milk samples were collected from eight dairy farms in Beijing and Hebei province of China. A total of 44 samples from the farms were adulterated by water, starch, goat milk and urea at different levels (Table 1). All the samples were used together and divided into two sets for PLSR, including a training set (75% of the samples) and a test set for validation (25% of the samples).

Portable NIR spectroscopic analysis

All samples were scanned using MicroNIR 1700, a miniature NIR spectrometer developed and manufactured by JDSU (JDS Uniphase Corporation, California, USA, www.jdsu.com). It features a small size (45 mm diameter × 42 mm height), low weight (60 g) and short analysis time (a few seconds). This technology can be controlled and operated on a tablet, smartphone or portable computer. The MicroNIR spectrometer is an ultra-compact instrument designed for diffuse reflection, transfection or transmission mode. It uses a linear variable filter (LVF) mounted over a diode array detector (DAD) that separates incoming light into different wavelengths. The spectrometer integrates light sources

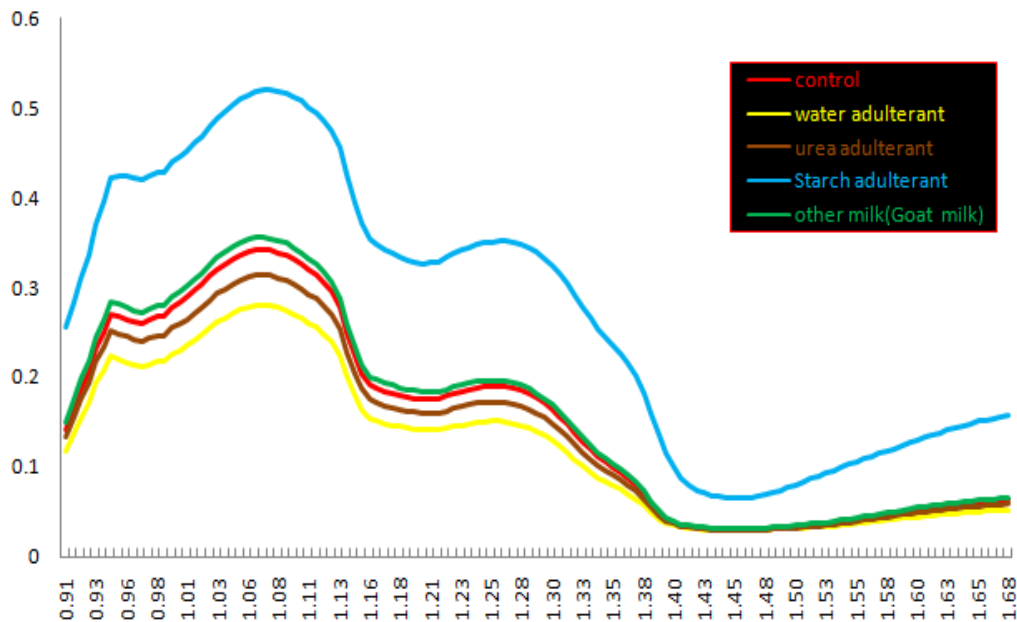


Figure 1. Five mean NIR spectra for natural (pure) milk and adulterated milk.

with readout electronics in a small construct.

White reference measurement was obtained using a white black (or dark) reference was obtained at a fixed place in the room. Absorbance values were calculated as $-\log(\text{Sample-Black/White-Black})$ with data sent to the host computer as .csv files. Three replicate spectra were acquired from all samples. Replicas were averaged for data analysis in the following chemometric analysis.

Statistical analysis

Multivariate analysis was used to develop qualitative and quantitative models with NIR data sets. Discrimination and quantification of adulterated milk samples were evaluated using soft independent modeling of class analogy (SIMCA) and PLSR, respectively. Chemometrics was performed in Microsoft Excel (Pomerantsev A.L., 2014). The Excel file SIMCA template.xlsx is a pattern for performing DD-SIMCA, an advanced one-class classification method for class modeling (Pomerantsev, 2008; Pomerantsev and Rodionova, 2013), which allows theoretical computing of misclassification errors. This technique is based on principal component analysis (PCA) applied to the training data of a target class. The PLSR template.xlsx file presents the forms for measuring the performance of the projection to latent spaces/structures (PLS) regression. PLS regression is an innovative method for regression analysis using Excel¹⁶, which allows development of a relationship between the matrix of predictors X and the vector of responses Y . This technique is based on projection to latent structures.

RESULTS AND DISCUSSION

Spectral characteristics

Figure 1 shows the normalization of five mean NIR spectra in the region of 950 to 1650 nm for pure milk and

milk adulterated with water, starch, goat milk and urea. The obtained spectra showed two prominent absorption bands centered at around 1070 and 1270 nm, respectively. These two bands were associated with strong O=H stretch overtone/combination of water (Laporte et al., 1999; Rodriguez et al., 2006), which was attributed to a short spectral range (from 950 to 1650 nm) on the spectrometer. Samples adulterated with urea demonstrated strong absorption at 1063 nm associated with NH_4^+ deformation, indicating the decomposition of urea, while C=O was absorbed at 1255 nm (Mortland, 1996). The spectra of milk adulterated with goat milk did not exhibit any visual difference from the control, which could be ascribed to close similarity of the components in this substance with the control milk.

Nevertheless, those mean NIR spectra were extremely similar to each other, making it difficult to classify the spectra. Therefore, the DD-SIMCA method was applied in this study to distinguish between natural (pure) and adulterated milk samples.

DD-SIMCA classification

The application of DD-SIMCA using a training dataset demonstrated the performance of this method in absence of adulterated samples. For this purpose, we verified the model established for natural or pure milk, which was directly collected from the cows in the farms. First, the test set included natural or pure milk samples collected from the cows in the same farms, but not included in the previous training set. Second, we adulterated the samples

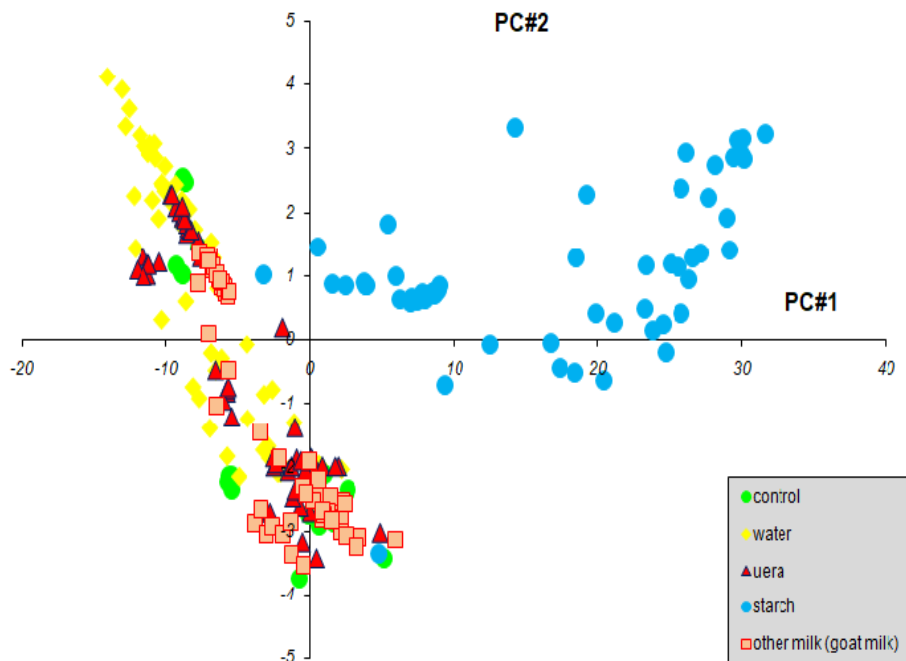


Figure 2. Explorative analysis milk data PCA using the whole dataset1 (training dataset) scores plot PC1 vs PC2.

with water, starch, goat milk and urea. They helped verify the performance of the model regarding outliers. Milk samples that were essentially dissimilar to the target class (control sample) were taken as 'low quality' adulterated samples, whereas milk samples close to the target class (control sample) were regarded as 'high quality' adulterated samples. The latter helped analyze the most challenging cases and assess the values of type II errors.

Preliminary PCA performed on the whole data revealed both big similarities and differences between pure and adulterated milk samples that were collected from a variety of dairy farms (Figure 2). Though all milk samples contained the same components, these samples were adulterated with various levels of adulterants, resulting in spectrum discrepancies. Spectral absorbance reflected the concentrations of the components.

Individual decision rules were developed for each adulterant and the control milk sample using the DD-SIMCA method. For this purpose, several milk samples of the target class were collected as the training set. The test set comprised of many samples of the target class and adulterated samples. To illustrate the result, pure milk samples were chosen as the target class and other adulterated samples were analyzed against a model developed for this class. A proportion of 75% of both the control and adulterated samples were collected in the training set whereas the test set contained 25%.

The PCA model with four principal components (PCs) explained 74% of the total variance. The decision rules

were constructed for the chi-squared distributions, with $N_h = 2$ and $N_v = 6$ DoFs. For $\alpha = 0.01$ (left panel of Figure 3), all training objects were located inside the acceptance area. As for the test set (right panel of Figure 3), the samples originated from all adulterated samples were located far from the acceptance area and could easily be classified as outliers. All test milk samples from the target class control milk (pure) were classified properly, but two samples originated from the subset of milk adulterated with urea (0.5%) and milk adulterated with water (5%) were wrongly accepted. This could be attributed to the low concentrations of the adulterants.

SIMCA classification of the other classes was done in a similar way. The final results are as shown in Figures 4 to 7.

In this study, we demonstrated that the PCA model with four PCs reliably separated the target class (control (pure) milk samples) from all other classes (adulterated milk samples). The proposed classification method recognized such alien objects successfully.

Similar classifications were repeated for the cases where each adulterated milk sample was selected as a target class. The results provided individual decision rules with specific values of α and β errors.

Quantification of adulterants in milk

The adulterant levels in milk were quantified using PLSR analysis to generate calibration models. Individual models

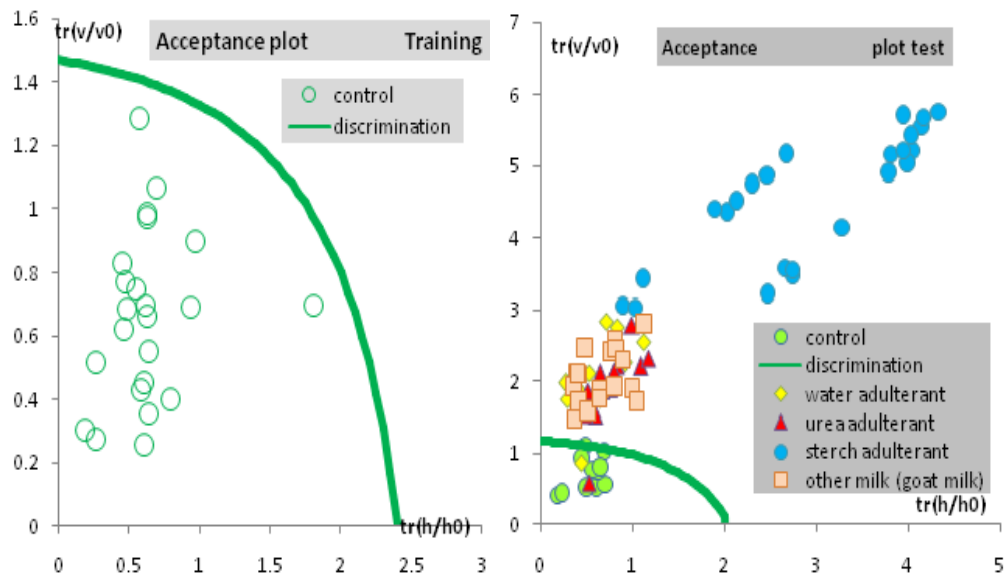


Figure 3. The results of SIMCA classification of class 1 (control samples).

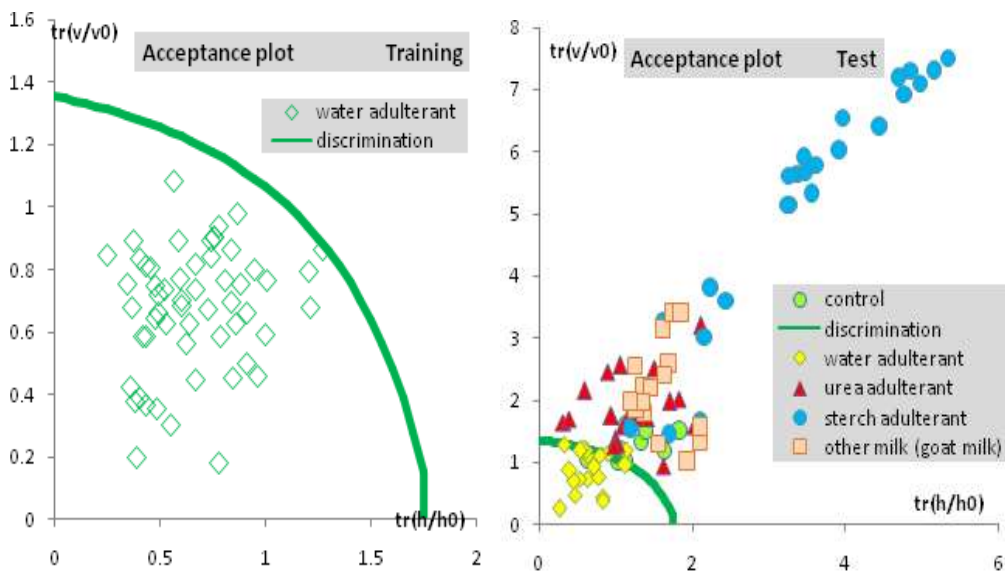


Figure 4. The results of SIMCA classification of class 2 (milk samples adulterated with water).

were developed using the same spectral regions identified by classification analysis (Table 2). This PLSR calibration model was then used to test the prediction ability for each independent 25% test sample set, as shown in Figures 8, 9, 10 and 11. This was an external validation part of the PLSR model because the spectral data used for the test set were not utilized in building the PLSR model. It can be seen from these figures that the PLSR model has very good prediction ability according to the root mean square error of prediction (RMSEP), which might be caused by the fact that the 25% test samples

were not used in building the PLSR calibration model. RMSEP is a statistical measure of how well the model predicts new samples (not used when building a model). RMSEP expresses the average error to be expected in future predictions when a calibration model is applied to unknown samples.

PLSR models suggested a good correlation between infrared estimated concentrations and spiked adulterant levels in the milk (Table 2).

The PLSR scatter plots (Figures 8 to 11) illustrated a good fit between the reference levels and NIR predicted

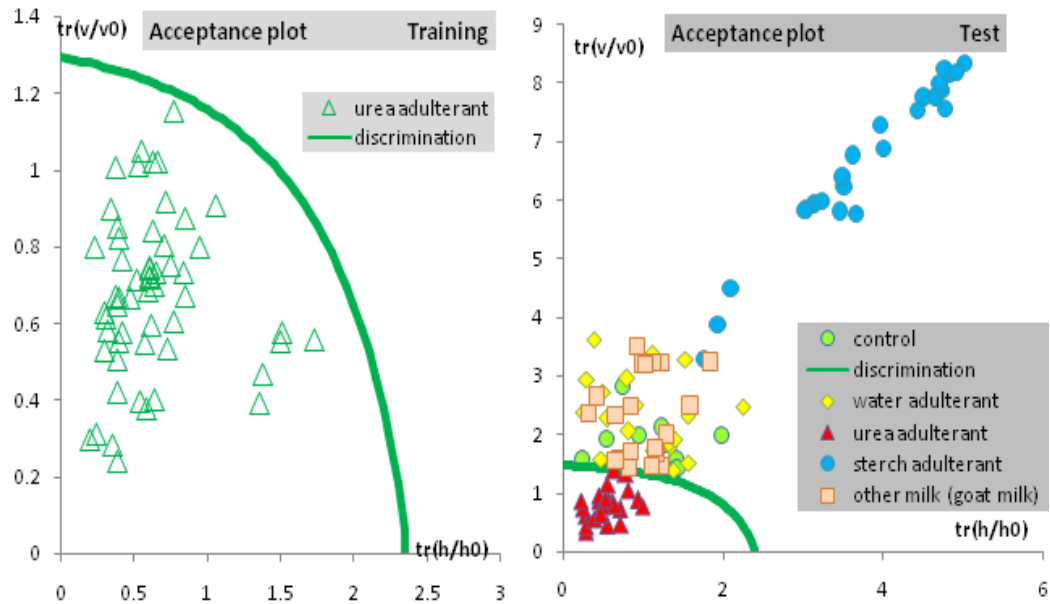


Figure 5. The results of SIMCA classification of class 3 (milk samples adulterated with urea).

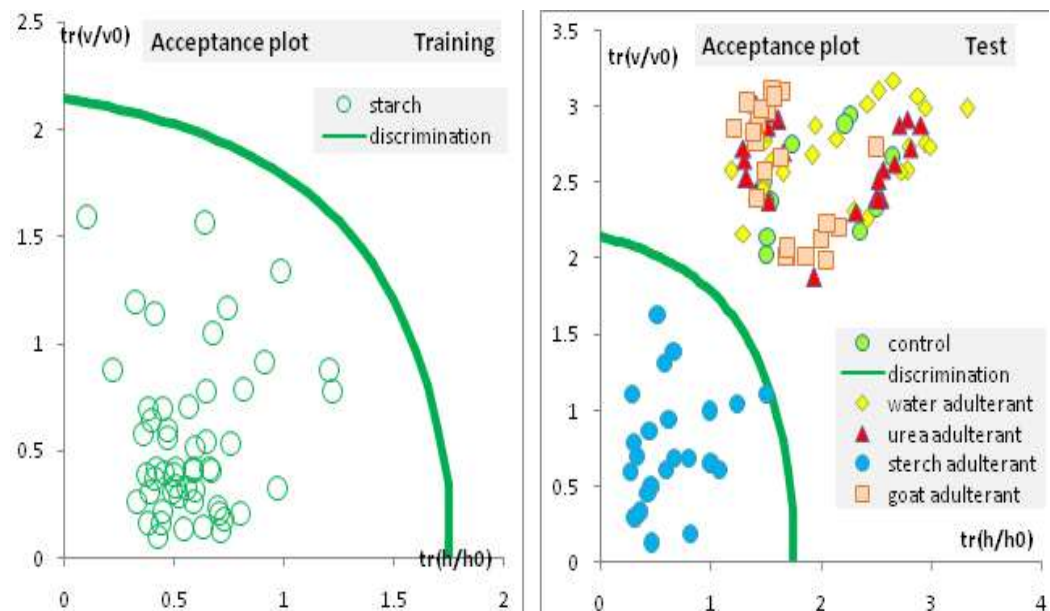


Figure 6. The results of SIMCA classification of class 4 (milk samples adulterated with starch).

values for milk adulteration. Models were developed using 5, 10, and 13 factors and explained more than 97% of the variance in the multispectral data set with R^2 val ranging from 0.87 to 0.94, revealing prediction robustness of the calibration models. The SEP values, an estimate of the error in predicting the adulteration level in an unknown sample, were 4.35, 0.34, 4.74 and 5.56 g/L for estimation of the levels of adulteration with water, urea, starch and goat milk

respectively (Table 2).

In Figure 8, the statistical result revealed the lowest RMSEP value 4.35% (v/v) with a model factor of 13. This region contained bands arising from the combination of OH symmetric and asymmetric stretching modes of water (Maeda, 1995). It was noted that the error limit of $\pm 4.35\%$ (v/v) was slightly higher than the minimum water content contained in a milk sample set. Therefore, when milk was adulterated by adding a very small volume of water within

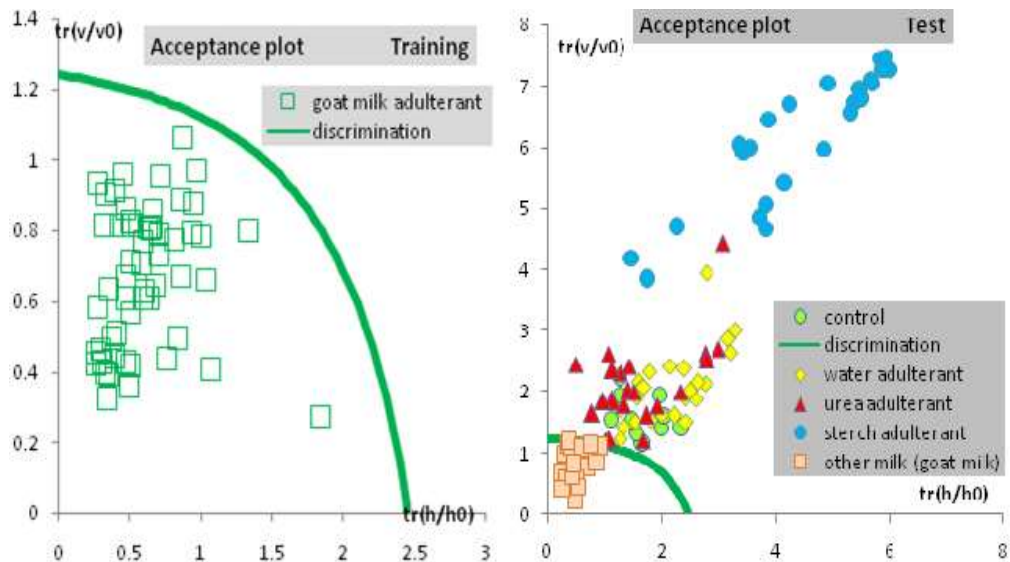


Figure 7. The results of SIMCA classification of class 5 (milk samples adulterated with goat milk).

Table 2. Calibration and cross-validation results of multivariate models developed using portable spectrometers.

Adulterant	Analysis result of PLSR						
	Factor	SEC	SECV	SEP	R ² cal	R ² val	R ² pd
Water	13	3.3	4.25	4.35	0.97	0.94	0.93
Urea	10	0.28	0.34	0.34	0.91	0.87	0.87
Starch	5	3.28	4.74	4.74	0.96	0.93	0.93
Goat milk	10	5.44	6.37	5.56	0.91	0.89	0.91

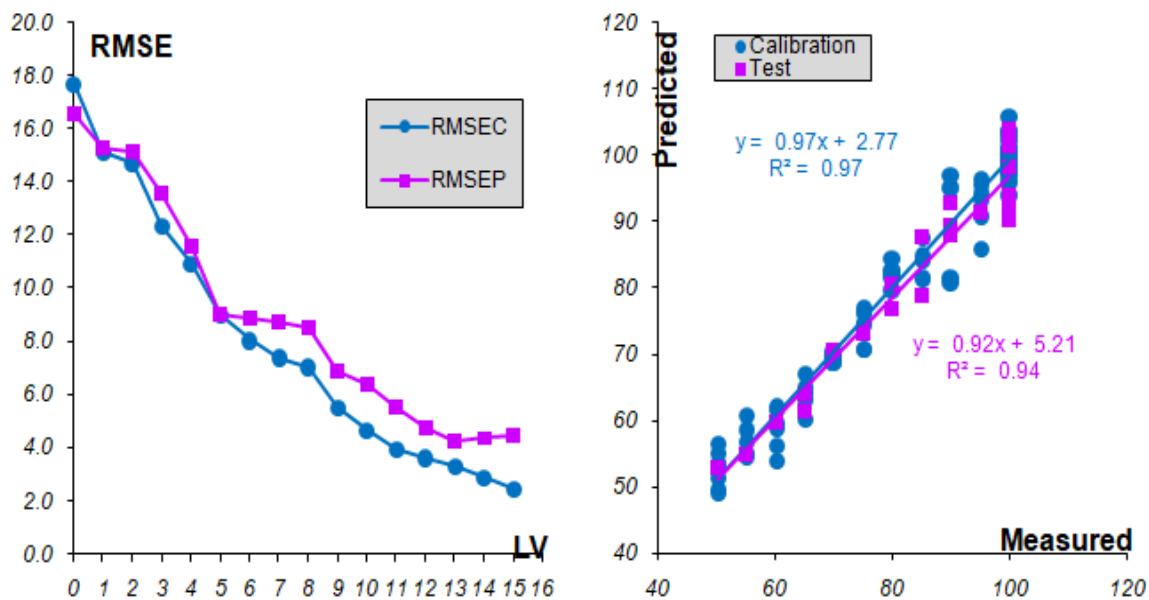


Figure 8. Root mean square errors (RMSE) and partial least squares regression (PLSR) curves for training (circle) and validation (square) data sets of milk adulterated with water.

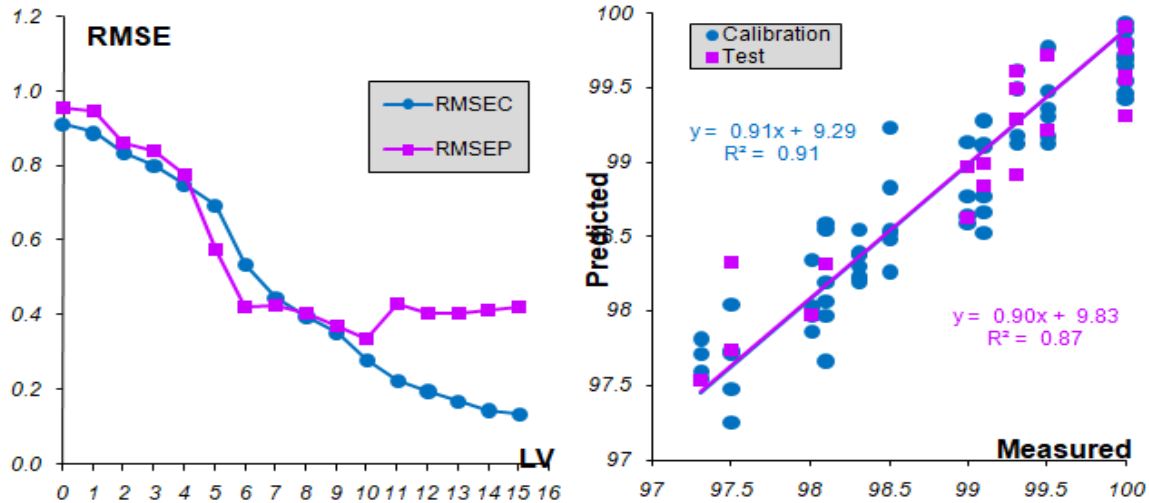


Figure 9. Root mean square errors (RMSE) and partial least squares regression (PLSR) curves for training (circle) and validation (square) data sets of milk adulterated with urea.

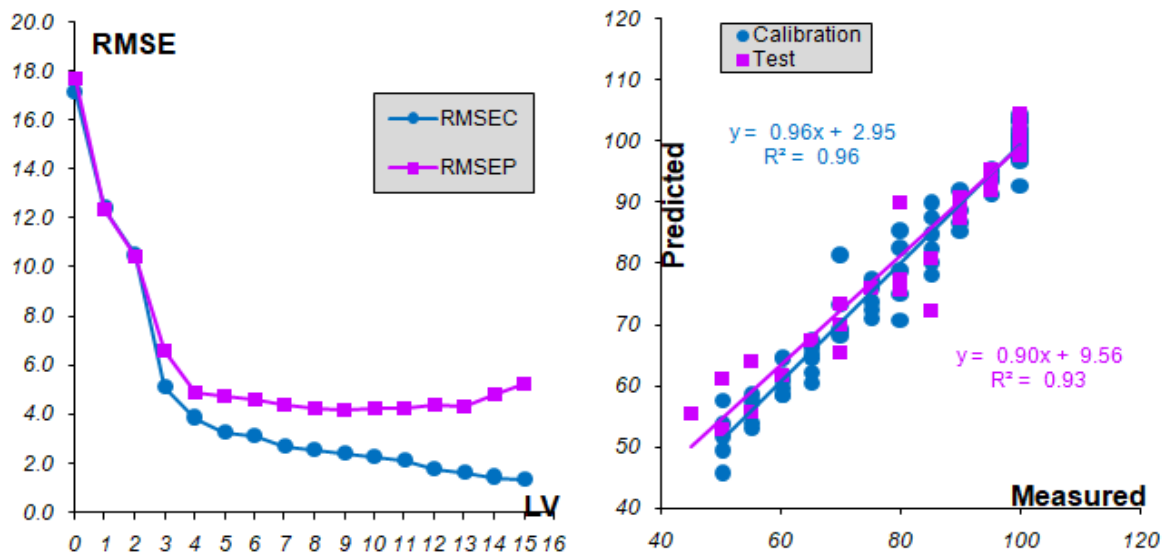


Figure 10. Root mean square errors (RMSE) and partial least squares regression (PLSR) curves for training (circle) and validation (square) data sets of milk adulterated with starch.

the error limit, it was difficult to determine the content of water in such a milk sample. This may be attributed to the fact that typical cow milk normally contains water as a main component. Thus, low contamination of water in milk was not obvious in the NIR spectrum.

These low SEP values can be associated with the presence of distinct and specific absorption signals for each adulterant used in the spiking process. Our regression modeling showed the same prediction abilities as those described in the earlier literatures (Laporte and Paquin, 1999; Rodriguez et al., 2006) and had minor differences with that reported in Santos et al. (2013)

work.

Conclusion

This study examined the applicability of a portable near-infrared (NIR) spectrometer (MicroNIR 1700) combined with a multivariate analysis method chemometrics in the monitoring of adulteration in cow milk. The results revealed that the method is unique and suitable for detection of water, urea, starch and goat milk. Further study on the feasibility validation of this technique used in

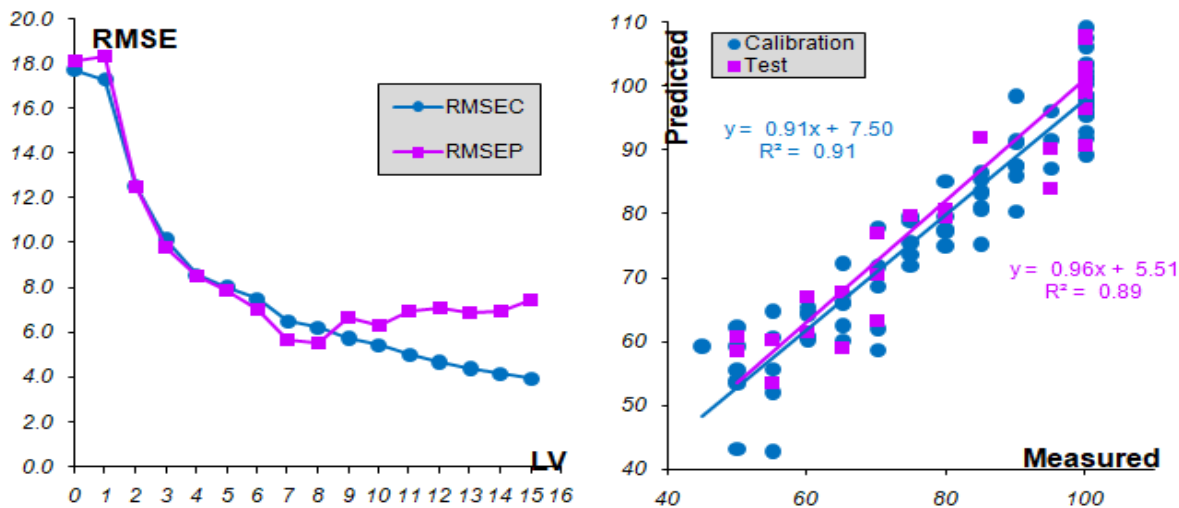


Figure 11. Root mean square errors (RMSE) and partial least squares regression (PLSR) curves for training (circle) and validation (square) data sets of milk adulterated with other milk (goat milk).

monitoring adulterants in cow milk is recommended.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

Authors are highly grateful to IAEA CRP project No. (D52040-22238) for his financially supported for conducting this study.

ABBREVIATIONS

NIR, Near-infrared spectroscopy; **PCA**, principal component analysis; **DD-SIMCA**, data driven soft independent modeling of class analogy; **PLSR**, partial least squares regression; **RMSE**, root mean square error; **RMSECV**, root mean square error of cross validation; **RMSEP**, root mean square error of prediction.

REFERENCES

- Allendorf M, Subramanian A, Rodriguez-Saona L (2012). Application of a handheld portable mid-infrared sensor for monitoring oil oxidative stability. *Journal of the American Oil Chemists' Society* 89:79-88
- Birkel E, Rodriguez-Saona L (2011). Application of a portable handheld infrared spectrometer for quantitation of trans fat in edible oils. *Journal of the American Oil Chemists' Society* 88(10):1477-1483.
- Campos Motta TM, Hoff RB, Barreto F, Andrade RBS, Lorenzini DM, Meneghini LZ, Pizzolato TM. (2014). Detection and confirmation of milk adulteration with cheese whey using proteomic-like sample preparation and liquid chromatography–electrospray–tandem mass spectrometry analysis. *Talanta* 120:498-505
- Cayuela JA, Weiland C (2010). Intact orange quality prediction with two portable NIR spectrometers. *Postharvest Biology and Technology* 58:113-120.
- Etzion Y, Linker R, Cogan U, Shmulevich I (2004). Determination of protein concentration in raw milk by mid-infrared Fourier transform infrared/attenuated total reflectance spectroscopy. *Journal of Dairy Science* 87:2779-2788
- Faraz A, Laleef M, Mustaf MI, Akhtar P, Yagoob M, Rehman S (2013). Detection of adulteration, chemical composition and hygienic status of milk supplied to various canteens of educational institutes and public places in Faisal Abad. *Journal of Animal and Plant Science* 23:119-124.
- Hassel S, Rodriguez-Saona LE (2012). Application of a handheld infrared sensor for monitoring the mineral fortification in whole grain cornmeal. *Food Analytical Methods* 5:571-57
- Kamthania M, Saxena J, Saxena K, Sharma DK (2014). Milk adulteration: Methods of detection & remedial measures. *International Journal of Engineering and Technology Research* 1:15-20.
- Kawasaki M, Kawamura S, Tsukahara M, Morita S, Komiya M, Natsuga M (2008). Near-infrared spectroscopic sensing system for on-line milk quality assessment in a milking robot. *Computers and Electronics in Agriculture* 63:22-27.
- Kim SS, Young C, Mizaikoff B (2008). Miniaturized mid-infrared sensor technologies. *Analytical and Bioanalytical Chemistry* 390:231-237.
- Laporte MF, Paquin P (1999). Near-infrared analysis of fat, protein, and casein in cow's milk. *Journal of Agricultural and Food Chemistry* 47:2600-2605.
- Maeda H, Ozaki Y, Tanaka M, Hayashi N, Kojima T (1995). Near Infrared Spectroscopy and Chemometrics Studies of Temperature-Dependent Spectral Variations of Water: Relationship between Spectral Changes and Hydrogen Bonds. *Journal of Near Infrared Spectroscopy* 3:19.
- Moore JC, Spink J, Lipp M (2012). Development and application of a database of food ingredient fraud and economically motivated adulteration from 1980 to 2010. *Journal of Food Science* 77:R118-R126
- Mortland MM (1996). Urea complexes with montmorillonite an infrared absorption study. *Clay Minerals* 6:143-156.
- Pomerantsev AL (2014). *Chemometrics in Excel*, Wiley & Sons, Hoboken NJ.
- Pomerantsev AL (2008). Acceptance areas for multivariate classification derived by projection methods. *Journal of Chemometrics* 22:601-609.
- Pomerantsev AL, Rodionova OY (2013). Concept and role of extreme objects in PCA/SIMCA. *Journal of Chemometrics* 28:429-438.

- Poonia A, Jha A, Sharma R, Singh HB, Rai AK, Sharma N (2017). Detection of adulteration in milk: A review. *International Journal of Dairy Technology* 70(1):23-42.
- Rodriguez-Saona LE, Koca N, Harper WJ, Alvarez VB (2006). Rapid determination of swiss cheese composition by Fourier transform infrared/attenuated total reflectance spectroscopy. *Journal of Dairy Science* 89:1407-1412.
- Salih MAM, Yang S (2017) Common Milk Adulteration in Developing Countries Cases Study in China and Sudan: A Review. *Advances in Dairy Research* 5:192.
- Santos PM, Pereira-Filho ER, Rodriguez-Saona LE (2013). Application of hand-held and portable infrared spectrometers in bovine milk analysis. *Journal of Agricultural and Food Chemistry* 61:1205-1211
- Singuluri H, Sukumaran MK (2014) Milk adulteration in Hyderabad, India-a comparative study on the levels of different adulterants present in milk. *Journal of Chromatography and Separation Techniques* 5:1-3.
- Soomro AA, Khaskheli M, Memon MA, Barham GS, Haq IU, Fazlani SN, Khan IA, Lochi GM, Soomro RN (2014). Study on adulteration and composition of milk sold at Badin. *International Journal of Research in Applied, Natural and Social Sciences* 2(9):57-70.