Food Research 6 (4) : 337 - 344 (August 2022)
Journal homepage: https://www.myfoodresearch.com

FOOD RESEARCH

FULL PAPER

The use of FTIR spectroscopy combined with pattern recognition and multivariate calibration for rapid authentication of Etawa goat milk from soymilk

[1]Windarsih, A., [2,3,*]Rohman, A. and [4]Irnawati

[1]*Research Center for Food Technology and Processing (PRTPP), National Research and Innovation Agency (BRIN), Yogyakarta, 55861, Indonesia*
[2]*Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia*
[3]*Center of Excellence Institute for Halal Industry and Systems (IHIS), Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia*
[4]*Faculty of Pharmacy, Universitas Halu Oleo, Kendari, Sulawesi Tenggara, 93232, Indonesia*

Article history:
Received: 5 July 2021
Received in revised form: 23 August 2021
Accepted: 6 December 2021
Available Online: 18 August 2022

Keywords:
Etawa goat milk,
Soymilk,
FTIR spectroscopy,
Authentication,
Chemometrics

DOI:
https://doi.org/10.26656/fr.2017.6(4).491

Abstract

Etawa goat milk (EGM) is valuable milk with a higher price compared to other kinds of milk. It is susceptible to being adulterated with low-priced milk for economical reasons. Therefore, the development of an analytical method for rapid authentication of EGM from soymilk (SM) is highly required. Fourier transform infrared (FTIR) spectroscopy combined with chemometrics was evaluated for rapid qualitative and quantitative analysis of SM in raw fresh EGM samples. Samples were prepared by mixing goat milk with soymilk in several concentration levels ranging from 0-100% (v/v). FTIR spectra revealed the differences between authentic and adulterated EGM samples in certain wavenumber regions. Linear discriminant analysis (LDA) at wavenumbers of 3000-750 cm^{-1} was successfully used for discrimination between authentic and adulterated EGM samples. The Cooman plot showed a clear classification between pure and adulterated goat milk samples. Partial least square (PLS) using wavenumber combination of 3000-2800 cm^{-1} and 1800-650 cm^{-1} as well as principal component regression (PCR) using wavenumber of 3000-750 cm^{-1} has been successfully used for quantification of SM in EGM. However, PLS offered a better model for quantification with higher values of R^2 calibration (0.9997) and validation (0.9992) and lower values of RMSEC (0.848) and RMSEP (1.19). It can be concluded that this method is promising for the rapid authentication of goat milk samples.

1. Introduction

Goat milk is considered to be highly nutritious milk because it contains a lot of important nutrition such as lipid-soluble vitamins, fats, protein and minerals which are beneficial for health (Arief *et al*., 2018). Etawa goat milk (EGM) is produced by the Etawa goat species, a goat from India which is also called Jamnapari goat which provides high-quality goat milk and has been consumed in some countries in the world (Pereira *et al*., 2020). The concentration of total solid in goat milk is approximately 13.2% consisting of fat (4.5%), protein (3.6%), lactose (4.5%), and minerals (0.8%) (Amigo and Fontecha, 2011; Di Pinto *et al*., 2017). Most of the fatty acids in goat milk are short and medium fatty acids such as caproic acid, caprylic acid, capric acid, and lauric acid accounting for about 15-18%. These kinds of fatty acids are easy to digest by the lipase enzyme and be absorbed because the shorter fatty acid chain is easier to digest (Raynal-Ljutovac *et al*., 2008; Silanikove *et al*., 2010). Goat milk also contains important proteins which are divided into two categories, namely casein and whey protein. αs1-casein, αs2-casein, β-casein, and κ-casein are the main casein constituent while the major components of whey protein are β-lactoglobulin and α-lactalbumin (Amigo and Fontecha, 2011). EGM also contains some important minerals and vitamins such as calcium, sodium, magnesium, chlorine, phosphorus, potassium, vitamin A, and vitamin D (Raynal-Ljutovac *et al*., 2008; Kalyankar *et al*., 2015).

From an economic point of view, EGM possesses a higher price compared to other milk such as cow milk and buffalo milk. The price of high-quality EGM can be two or three times higher than cow milk (Arief *et al*.,

*Corresponding author.
Email: *abdulkimfar@gmail.com*

eISSN: 2550-2166 / © 2022 The Authors.
Published by Rynnye Lyan Resources

2018). As a consequence, it is vulnerable to adulteration with other substances to gain more economic benefits. Adulteration in milk is often practised by an unethical player using lower-priced substances and it can be categorized as a fraudulent practice (Azad and Ahmed, 2016; Rebechi *et al*., 2016; Poonia *et al*., 2017). Soymilk (SM) is not categorized as milk. However, it is potential and often used as an adulterant in high price milk because it is widely available, easy to obtain, and has a lower price than any other type of milk (Sharma *et al*., 2010). Due to its physical appearance such as its white colour, it is difficult when is mix with goat milk and makes it a potential adulterant in other milk. As a consequence, the development of an analytical method for the detection of adulterants including SM in EGM is inevitable.

Several analytical methods have been developed for milk authentication analysis such as high-performance liquid chromatography (HPLC), gas chromatography (GC), gas chromatography-mass spectrometry (GCMS), liquid chromatography-mass spectrometry (LCMS), immunological test, electrophoretic technique, and DNA -based method (Abernethy and Higgs, 2013; Scano *et al*., 2014; Di Pinto *et al*., 2017; Ke *et al*., 2017). HPLC and GC are the main methods commonly used for the analysis of triacylglycerol and fatty acids composition in milk authentication, respectively (Scano *et al*., 2014; Bernardi *et al*., 2015; Cossignani *et al*., 2019). However, these methods have some drawbacks including requiring several steps of sample preparation, time-consuming, need a highly skilled analyst, destructive, and quite expensive. Therefore, a rapid and simple analytical method for goat milk authentication is necessary.

The vibrational spectroscopy method namely Fourier transform infrared (FTIR) spectroscopy has emerged as a fingerprint analytical technique that is interesting and promising for the detection of adulteration in some types of food. It offers fast analysis, non-destructive, simple sample preparation, reliable, and reproducible analytical results (Rohman *et al*., 2016). Moreover, it has a lower cost compared to the other methods. Using the attenuated total reflectance (ATR) technique for sample measurement allows for direct measuring samples without any sample preparation steps that obviously shorten the analysis time (Jawaid *et al*., 2013; Feng *et al*., 2019). Combined with chemometrics which is capable of analyzing the complex data generated from FTIR measurement, it becomes a powerful analytical method for food authentication analysis (Worley and Powers, 2013; Erwanto *et al*., 2016). Several chemometrics techniques have been widely used in food authentication analysis such as principal component analysis (PCA) and linear discriminant analysis (LDA) for qualitative analysis as well as partial least square (PLS) and principal component regression (PCR) for quantitative analysis. PCA and LDA are used for sample differentiation and it has been successfully used for the differentiation of authentic and adulterated food and pharmaceutical products. Moreover, PLS and PCR have been used for authentication by predicting the concentration of adulterant samples (Valdés *et al*., 2018; Caballero *et al*., 2020). FTIR spectroscopy combined with chemometrics of PCA and PLS has been used for analyzing the presence of formalin in cow milk (Balan *et al*., 2020). The presence of cow milk in goat milk was also successfully analyzed using FTIR spectroscopy combined with chemometrics of PLS regression and SIMCA (soft independent class modelling analogy) (Yaman, 2020). Moreover, FTIR spectroscopy in combination with chemometrics of pattern recognition and multivariate calibration has also been used for the detection of foreign substances in milk such as melamine, cyromazine, urea, sucrose, and formalin (Poonia *et al*., 2017).

To the best of our knowledge, the use of FTIR spectroscopy in combination with chemometrics of multivariate analysis for rapid authentication of Etawa goat milk adulterated with soymilk has not been reported. Moreover, there is no report about the comparative study on PLS and PCR for the quantification of soymilk in Etawa goat milk. Therefore, the objective of this research was to develop a rapid and reliable analytical method using FTIR spectroscopy and chemometrics of classification and multivariate calibration for direct and rapid authentication of Etawa goat milk adulterated with soymilk.

2. Materials and methods

2.1 Sample collection

Raw fresh Etawa goat milk (EGM) samples were obtained from a local farm in Yogyakarta, Indonesia. Fresh milk was taken from a female goat in the morning. The adulterant used was soymilk (SM) which was obtained from a local market in Yogyakarta, Indonesia. Fresh milk samples were directly used for sample preparation to be used for FTIR measurement. To obtain uniformity, samples were homogenized using a vortex mixer for 1 min. Samples homogenization was also aimed to prevent scattering during spectra measurement and analysis.

2.2 Sample preparation

The authentic sample of raw fresh EGM was prepared in glass vials as well as the pure SM sample. The adulterated samples of EGM were prepared by mixing EGM with SM in binary mixtures with

 © 2022 The Authors. Published by Rynnye Lyan Resources

concentrations of adulterant (SM) ranging from 0% to 100 % v/v. The samples were prepared in a total volume of 2 mL and each series was placed in a glass vial with a clear label. All samples were directly measured using FTIR-ATR spectrophotometer after being prepared by placing milk samples on the ATR crystal.

2.3 FTIR measurement

FTIR measurement was performed according to Windarsih *et al*. (2020). FTIR spectra acquisition was carried out using a Thermo Nicholet iS10 FTIR spectrophotometer (Thermo Scientific, USA). Pure and adulterated milk samples were measured using an FTIR spectrophotometer employing the attenuated total reflectance (ATR) technique. FTIR spectrophotometer was equipped with KBr (potassium bromide) beam splitter and DTGS (deuterated triglycine sulphate) detector. Samples were directly placed by dropping the milk samples on ATR crystal and measured at a wavenumber of 4000-650 cm^{-1}. The measurement was performed using a number of scans 32 and a resolution of 8 cm^{-1}. Samples were measured in absorbance mode and each sample was measured in three replicates. The background spectra were measured prior to each sample measurement. To remove trace, ATR crystal was cleaned using n-hexane and acetone after each sample measurement.

2.4 FTIR spectra analysis

The FTIR spectra were analyzed to figure out the differences between FTIR spectra of EGM and SM using OMNIC software (Thermo Scientific, USA). FTIR spectra of each sample were treated with preprocessing techniques such as smoothing and baseline correction. The functional groups were identified from the FTIR spectra of EGM and SM. Spectra derivatization was also performed to deeply investigate the differences or unique patterns specific to each sample. Derivatization was carried out either the first derivatization or the second derivatization using the Savitsky-Golay technique.

2.5 Chemometrics analysis

Chemometrics was performed to manage the huge variables resulted from FTIR measurements. Linear discriminant analysis (LDA) was used for sample classification whereas chemometrics of partial least square (PLS) and principal component regression (PCR) was used for quantitative analysis by predicting adulterants concentration. All samples were analyzed using TQ Analyst software (Thermo Scientific, USA). Samples were divided into two categories, namely the calibration dataset consisting of 42 datasets and the validation dataset consisting of 20 datasets for external validation.

2.5.1 Linear discriminant analysis

Linear discriminant analysis (LDA) was performed in authentic samples and adulterated samples of EGM with SM. The variables used were absorbance values extracted from FTIR spectra. Data were processed using the mean centering technique prior analysis and optimization of wavenumber region for creating LDA model was carried out in the range of 4000-650 cm^{-1} to obtain optimum region for classification. The discrimination result was evaluated using the Coomans plot to justify the classification. Moreover, the Mahalanobis distance was observed to know the distance between authentic and adulterated samples.

2.5.2 Partial least square and principal component regression

Chemometrics of Partial least square (PLS) and Principal component regression (PCR) was used to create a model for quantitative analysis of SM in EGM. The variables used for PLS and PCR models were the absorbance values from FTIR spectra. Preprocessing data was carried out prior to multivariate analysis such as standard normal variate (SNV) and mean centering technique (MCT). The optimized models were obtained from wavenumber optimization in the range of 4000-650 cm^{-1} employing either normal or derivative (first and second derivative) spectra. The models were evaluated using the coefficient of determination (R^2) both in calibration and validation models as well as the RMSEC (root mean square error of calibration) and RMSEP (root mean square error of prediction) for evaluating model error in calibration and validation models, respectively.

3. Results and discussion

3.1 FTIR spectra analysis

FTIR is known as a fingerprint technique that is beneficial to be used for sample differentiation, especially for authentication purposes. The spectra pattern of EGM is clearly different from SM as can be seen in Figure 1. EGM showed two major spectra patterns in the range of 3500-3000 cm^{-1} which correspond to the water spectra and in the range of 3000-2800 cm^{-1} and 1800-950 cm^{-1} which belong to the spectra of lipid/fat (Balan *et al*., 2020). These results support that the main components in raw fresh EGM are water and fats. The spectra identification including wavenumbers, functional groups, and vibration modes are shown in Table 1. On the other hand, the spectra of SM were only dominated by water spectra at the wavenumber of 3500-3000 cm^{-1}. It is suggested that the major compound in SM is water. The strong and broad peak of -OH stretching vibration in wavenumber of 3500 -3000 cm^{-1} was caused by hydrogen bonding of water

eISSN: 2550-2166 © 2022 The Authors. Published by Rynnye Lyan Resources

FULL PAPER

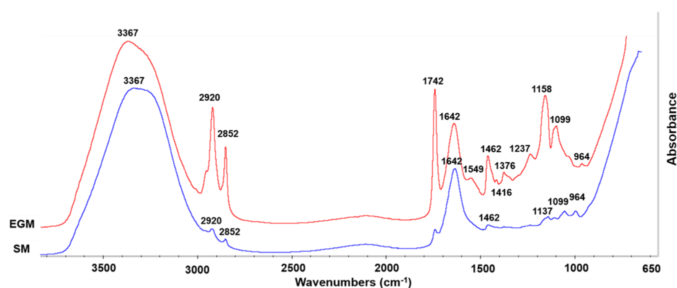

Figure 1. FTIR spectra of Etawa goat milk (EGM) and soymilk (SM) measured at 4000-650 cm^{-1}

Table 1. Functional groups and vibration modes of Etawa goat milk and soymilk (Balan *et al.*, 2020; Yaman, 2020)

Wavenumber (cm^{-1})	Functional groups and vibration modes
3367	-OH stretching
2920	$-CH_3$, $-CH_2$, and -CH asymmetric stretching
2852	$-CH_3$, $-CH_2$, and -CH symmetric stretching
1742	Carbonyl (C=O) stretching
1642	C=C aromatic stretching -NH bending vibration
1549	C=C alkene stretching vibration
1462	$-CH_3$-O- symmetric deformation vibration
1416	$-CH_3$, $-CH_2$, and -CH bending vibration
1376	$-CH_3$, $-CH_2$, and -CH bending vibration
1237	$-CH_3$, $-CH_2$, and -CH bending vibration
1158	$-CH_3$, $-CH_2$, and -CH bending vibration
1137	$-CH_3$, $-CH_2$, and -CH bending vibration
1099	-C-O stretching vibration
964	$-CH_3$, $-CH_2$, and -CH bending vibration

molecules. Peaks at 2920 cm^{-1} and 2852 cm^{-1} correspond to the stretching vibration of $-CH_3$, $-CH_2$, and -CH from fat compounds. Other peaks related to fat compounds could be observed at 1742 cm^{-1} (C=O/carbonyl), 1099 cm^{-1} (C-O), and some peaks related to bending vibration of $-CH_3$, $-CH_2$, and -CH. The peak at 1549 cm^{-1} (C=C alkene) supported the presence of unsaturated fats that come from unsaturated fatty acid vibrations. The peak at 1642 cm^{-1} corresponds to the protein molecules. This peak corresponds to Amide I which comes from the vibration of C=O and C-N in peptide linkages (Kong and Yu, 2007). Meanwhile, the carbohydrate compounds of lactose could be observed in a region of 1100-100 cm^{-1} with a maximum peak found at 1099 cm^{-1} (Balan *et al.*, 2020).

Even though EGM and SM showed different spectra patterns, adulterated EGM samples possessed very similar spectra patterns to authentic EGM. However, deep investigation in certain regions found that there were changes in certain regions as the adulterant concentration increased. The differences could be observed in the wavenumber region of 2980-2820 cm^{-1} ($-CH_3$, $-CH_2$, and -CH alkane stretching vibration) and 1770-1730 cm^{-1} (carbonyl/C=O stretching vibration). There was a decrease in absorbance value when the concentration of adulterant (SM) increased (Figure 2). EGM sample mixed with 5% of SM still had high absorbance values in these two regions while EGM with SM concentration of 95% showed very low absorbance values. It could be explained because SM only had a very weak and small peak in the region of 2980-2820 cm^{-1} and no peak was observed in the carbonyl region. Therefore, these two regions are the potential to be used as markers for adulteration detection in raw fresh EGM samples. The more the SM concentration in EGM, the lower the absorbance values at 2980-2820 cm^{-1} and 1770 -1730 cm^{-1}. However, in an unknown sample, it is difficult to distinguish whether the sample is authentic or adulterated when there is no information regarding the sample. Therefore, the use of advanced statistical methods such as chemometrics is highly required to overcome these problems.

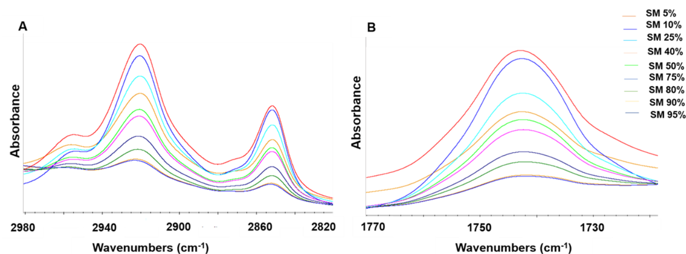

Figure 2. FTIR spectra of adulterated Etawa goat milk (EGM) with soymilk (SM) using soymilk level of 5-95% observed at region of 2980-2820 cm^{-1} and 1770-1730 cm^{-1}

3.2 Chemometrics analysis

3.2.1 Linear discriminant analysis for samples classification

LDA is one of the supervised pattern recognition techniques which is widely used for sample classification. It is often used in food and pharmaceutical product analysis, especially for sample discrimination. LDA was successfully used for classification between authentic EGM and adulterated EGM with SM. All adulterated samples of EGM with SM were perfectly separated from authentic EGM samples as can be seen in the Coomans plot (Figure 3). Pure raw fresh EGM samples appeared on the left side of the plot while all adulterated samples of EGM with SM appeared on the right side. LDA was created using normal FTIR spectra, without derivatization treatment at the wavenumber of 3000-750 cm^{-1}. The algorithm in LDA provided a good technique for retrieving data from variables and recognizing samples. As a consequence, it was successfully identified adulterated samples containing lower SM concentrations even in 5% of SM in EGM. LDA is one of the supervised pattern recognitions which is widely used for sample classification and it is often used in food analysis because it is simple, easy to perform, and easy to interpret the results. The distance in

© 2022 The Authors. Published by Rynnye Lyan Resources

the x-axis and y-axis, known as Mahalanobis distance, shows the distance between the authentic/pure EGM samples and adulterated EGM with SM. EGM and SM were well separated in greater Mahalanobis distance. Greater Mahalanobis distance values indicated better sample discrimination performed by LDA. The result indicated that a combination of FTIR and chemometrics of LDA could be used for discrimination between authentic and adulterated goat milk samples.

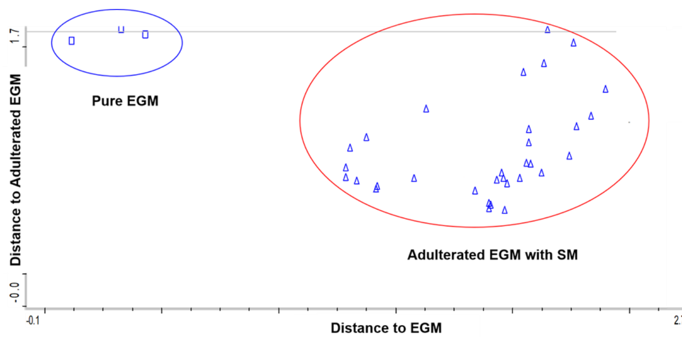

Figure 3. Coomans plot from linear discriminant analysis of pure Etawa goat milk (EGM) and adulterated EGM with soymilk (SM)

3.2.2 Quantitative analysis using PLS and PCR models

For quantification of SM in EGM, chemometrics of PLS and PCR was used to predict the concentration of SM in EGM. PLS and PCR are categorized as chemometrics of multivariate calibration and are often used for predicting the concentration of target samples. However, they have different algorithms for creating a model, therefore it is necessary to compare the performance of PLS and PCR for quantification of SM in EGM to evaluate each model's performance. Optimization of the PLS and the PCR model was carried out in several wavenumber regions not only at individual wavenumber ranges but also at wavenumber combinations to obtain the best model. The selection of wavenumber was aimed to obtain the best region for creating the model because whole wavenumber regions do not always give the best model. Several wavenumbers sometimes do not play an important role in creating PLS and PCR models because of the high noise. The selection of the best PLS and PCR model was based on the R^2, RMSEC, and RMSEP values.

For the PLS model, it searches latent variables within the huge numbers of variables for creating a model. The latent variables were used to search the relationship between actual and calculated/predicted values. Using 10 factors, the PLS model performed at a wavenumber combination of 3000-2800 cm^{-1} and 1800-650 cm^{-1} employing normal spectra demonstrated a good calibration model for detection and quantification of SM in EGM. Factors were created from the variables, and the number of factors used was based on the lowest PRESS (predicted residual error sum of squares) value. The PRESS value obtained from 10 PLS factors was 7.43. PRESS value indicated the error occurred during creating the PLS model. Therefore, the model with the lowest value of PRESS is expected to have the lowest error. The lower the PRESS value, the lower the RMSEC value of the model (Worley and Powers, 2013). The PLS model was also built not only using normal spectra but also compared to the Savitsky-Golay derivative spectra of the first derivative and the second derivative (Table 2). Derivatization could affect the spectra resolution and sometimes it could result in a better model than normal

Table 2. PLS of calibration and validation model for quantification of SM in EGM using normal, first derivative, and second derivative

Wavenumber (cm^{-1})	Number of factors	Spectral treatment	Calibration		Validation	
			R^2	RMSEC	R^2	RMSEP
3700-650	6	Normal	0.9938	3.80	0.9769	8.08
	5	Derivative 1	0.9965	2.83	0.9857	6.43
	2	Derivative 2	0.8170	19.60	0.8328	19.70
3000-750	10	Normal	0.9997	0.90	0.9992	1.59
	5	Derivative 1	0.9973	2.52	0.9911	5.07
	6	Derivative 2	0.9987	1.73	0.9945	4.13
3000-2800	5	Normal	0.9666	8.72	0.9459	14.20
	2	Derivative 1	0.7763	21.50	0.8424	19.20
	2	Derivative 2	0.7191	23.60	0.8248	20.30
1800-650	10	Normal	0.9997	0.83	0.9995	1.30
	6	Derivative 1	0.9975	2.40	0.9951	3.79
	2	Derivative 2	0.8169	19.60	0.8284	20.00
3000-2800 and 1800-650	**10**	**Normal**	**0.9997**	**0.85**	**0.9995**	**1.19**
	5	Derivative 1	0.9960	3.06	0.9866	6.21
	7	Derivative 2	0.9985	1.84	0.9927	4.63

*The selected model was marked in **bold**

© 2022 The Authors. Published by Rynnye Lyan Resources

spectra. The selected PLS for quantification of SM in EGM had a high value of R^2 (0.9997) and a low value of RMSEC (0.848) (Figure 4). The model was validated using external validation to confirm the performance of the calibration model. External validation was carried out using external validation samples. It was aimed to test the model validity using external validation samples. The result of external validation is similar to the calibration model, It means that the model has good validity. The obtained R^2 of the validation model was 0.9992 with the RMSEP value of 1.19. The high value of R^2 (close to 1) indicated high accuracy as well as a good fitness of the PLS model whereas low values of RMSEC and RMSEP represented low error in models. RMSEC and RMSEP values also correspond to the precision of the model. The lower the RMSEC and RMSEP values, the higher the precision of the model. Moreover, the difference between RMSEC and RMSEP values also indicated the good fitness of the model. The model with a lower difference between RMSEC and RMSEP is considered to be a good model.

For the PCR model, it creates a principal component (PC) to reduce variables from hundreds of original variables used to several PC values. PC represents the variations in original variables and it is used to create a regression model using several selected PCs. Optimization of PCR model was also carried out in the wavenumber region of 4000-650 cm^{-1} employing either normal spectra or derivative spectra (first and second derivative). The optimization results are shown in Table 3. PCR model performed at the wavenumber region of 3000-750 cm^{-1} using normal spectra was selected as the best model for quantification of SM in EGM. The selected PCR model used 8 principal components for creating regression in the PCR model. These 8 principal components are capable of defining hundreds of original variables. The obtained R^2 and RMSEC in the calibration model were 0.9981 and 2.11, respectively whereas the value of R^2 and RMSEP in the validation model was 0.9957 and 3.42, respectively. The R^2 values indicated good fitness of the PCR model while the RMSEC and RMSEP values correlated to good precision. There was also only a slight difference between RMSEC and RMSEP which supports the good model of PCR for SM quantification in EGM.

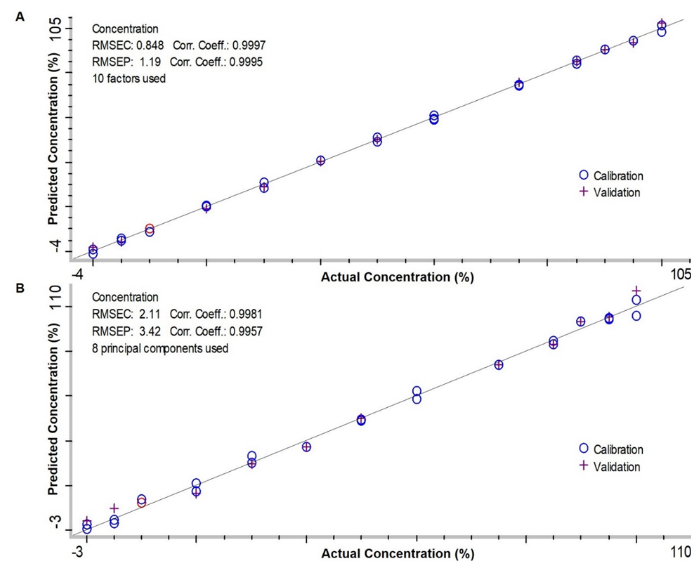

Figure 4. Partial least square (A) and principal component regression (B) models for quantification of soymilk (SM) in Etawa goat milk (EGM)

The residual values of calibration and validation models were also used as one of the parameters for PLS and PCR model evaluation. Residual values provided

Table 3. PCR of calibration and validation model for quantification of SM in EGM using normal, derivative 1, and derivative 2 spectra

Wavenumber (cm^{-1})	Number of factors	Spectral treatment	Calibration		Validation	
			R^2	RMSEC	R^2	RMSEP
3700-650	6	Normal	0.9919	4.33	0.9702	9.18
	6	Derivative 1	0.9950	3.41	0.9847	6.57
	6	Derivative 2	0.9943	3.62	0.9910	5.37
3000-750	**6**	**Normal**	**0.9981**	**2.11**	**0.9957**	**3.42**
	6	Derivative 1	0.9964	2.88	0.9938	4.14
	6	Derivative 2	0.9964	2.87	0.9941	4.19
3000-2800	4	Normal	0.8789	16.2	0.8944	16.5
	4	Derivative 1	0.8842	15.9	0.9082	14.9
	4	Derivative 2	0.7746	21.5	0.8624	18.7
1800-650	6	Normal	0.9952	3.34	0.9811	7.51
	6	Derivative 1	0.9963	2.92	0.9940	4.12
	6	Derivative 2	0.9935	3.88	0.9932	4.32
3000-2800 and 1800-650	7	Normal	0.9929	4.05	0.9844	6.72
	7	Derivative 1	0.9954	3.26	0.9870	6.07
	7	Derivative 2	0.9952	3.34	0.9932	4.40

*The selected model was marked in **bold**

eISSN: 2550-2166 © 2022 The Authors. Published by Rynnye Lyan Resources

information about the difference between actual concentration values and predicted concentration values measured using FTIR spectroscopy of the targeted analyte. The lower residual values are always expected when creating PLS and PCR models because it indicates that the models are capable of predicting the target analyte concentration accurately. Moreover, the plot of residual values also could be used to evaluate the error that occurred in models. There are two types of error in creating a calibration model, namely systematic error and random error. Good models are free from systematic error. The results of this study suggested that only random error occurs in both PLS and PCR models indicated by random scattering of each residual value.

From the above results, both PLS and PCR showed good multivariate calibration models for the quantification of SM in EGM. However, the PLS model provided higher R^2 values both in calibration and validation models and lower values of RMSEC and RMSEP. It suggested that the PLS model offered a better chemometrics model for the detection and quantification of SM as an adulterant in EGM. This model could be applied to predict the concentration of SM in an unknown EGM sample adulterated with SM. The FTIR spectra of the unknown sample are measured then the absorbance value is used to predict the SM concentration through the developed model.

4. Conclusion

FTIR spectroscopy provided rapid and reliable analysis for authentication of Etawa goat milk (EGM) adulterated with soymilk (SM). Combination with chemometrics of classification and regression demonstrated a good analytical technique for detection and quantification of the presence of soymilk in Etawa goat milk. FTIR spectroscopy is a fingerprint technique that is beneficial for the differentiation of each sample. Combined with chemometrics of pattern recognition and multivariate calibration, it can be used as a good analytical tool for sample classification and quantification not only in fresh goat milk but also in any kind of milks as well as dairy products. Therefore, it can be concluded that FTIR spectroscopy combined with chemometrics is a promising method for future authentication of milk and dairy products.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The author thanks to the technicians from Integrated Research and Testing Laboratory (LPPT) Universitas Gadjah Mada.

References

Abernethy, G. and Higgs, K. (2013). Rapid detection of economic adulterants in fresh milk by liquid chromatography–tandem mass spectrometry. *Journal of Chromatography A*, 1288(2013), 10–20. https://doi.org/10.1016/j.chroma.2013.02.022

Amigo, L. and Fontecha, J. (2011). Milk. Goat Milk. In Fuquay, J.W. (Ed.) Encyclopedia of Dairy Sciences 2nd ed., p. 484–493. USA: Academic Press. https://doi.org/10.1016/B978-0-12-374407-4.00313-7

Arief, Jamarun, N., Pazla, R. and Satria, B. (2018). Milk quality of ETAWA crossbred dairy goat fed by product of palm oil industry. *International Journal of Dairy Science*, 13(1), 15–21. https://doi.org/10.3923/ijds.2018.15.21

Azad, T. and Ahmed, S. (2016). Common milk adulteration and their detection techniques. *International Journal of Food Contamination* 3(22), 1-9. https://doi.org/10.1186/s40550-016-0045-3

Balan, B., Dhaulaniya, A.S., Jamwal, R., Amit, Sodhi, K.K., Kelly, S., Cannavan, A. and Singh, D.K. (2020). Application of Attenuated Total Reflectance-Fourier Transform Infrared (ATR-FTIR) spectroscopy coupled with chemometrics for detection and quantification of formalin in cow milk. *Vibrational Spectroscopy*, 107(103033), 1-7. https://doi.org/10.1016/j.vibspec.2020.103033

Bernardi, N., Benetti, G., Haouet, N.M., Sergi, M., Grotta, L., Marchetti, S., Castellani, F. and Martino, G. (2015). A rapid high-performance liquid chromatography-tandem mass spectrometry assay for unambiguous detection of different milk species employed in cheese manufacturing. *Journal of Dairy Science*, 98(12), 8405–8413. https://doi.org/10.3168/jds.2015-9769

Caballero, D., Ríos-Reina, R. and Amigo, J.M. (2020). Chemometrics and Food Traceability. In Reference Module in Food Science. Elsevier E-Book. https://doi.org/10.1016/B978-0-08-100596-5.22859-X

Cossignani, L., Pollini, L. and Blasi, F. (2019). Invited review: Authentication of milk by direct and indirect analysis of triacylglycerol molecular species. *Journal of Dairy Science*, 102(7), 5871–5882. https://doi.org/10.3168/jds.2019-16318

Di Pinto, A., Terio, V., Marchetti, P., Bottaro, M., Mottola, A., Bozzo, G., Bonerba, E., Ceci, E. and Tantillo, G. (2017). DNA-based approach for species identification of goat-milk products. *Food Chemistry*, 229, 93–97. https://doi.org/10.1016/j.foodchem.2017.02.067

 © 2022 The Authors. Published by Rynnye Lyan Resources

Erwanto, Y., Muttaqien, A.T., Sugiyono, Sismindari and Rohman, A. (2016). Use of Fourier Transform Infrared (FTIR) Spectroscopy and Chemometrics for Analysis of Lard Adulteration in "Rambak" Crackers. *International Journal of Food Properties*, 19(12), 2718–2725. https://doi.org/10.1080/10942912.2016.1143839

Feng, L., Zhu, S., Chen, S., Bao, Y. and He, Y. (2019). Combining Fourier transform Mid-infrared spectroscopy with chemometric methods to detect adulterations in milk powder. *Sensors*, 19(13), 2934. https://doi.org/10.3390/s19132934

Jawaid, S., Talpur, F.N., Sherazi, S.T.H., Nizamani, S.M. and Khaskheli, A.A. (2013). Rapid detection of melamine adulteration in dairy milk by SB-ATR-Fourier transform infrared spectroscopy. *Food Chemistry*, 141(3), 3066–3071. https://doi.org/10.1016/j.foodchem.2013.05.106

Kalyankar, S.D., Khedkar, C.D. and Patil, A.M. (2015). Goat: Milk. Encyclopedia of Food and Health. 1st ed. Elsevier E-Book. https://doi.org/10.1016/B978-0-12-384947-2.00358-5

Ke, X., Zhang, J., Lai, S., Chen, Q., Zhang, Y., Jiang, Y., Mo, W. and Ren, Y. (2017). Quantitative analysis of cow whole milk and whey powder adulteration percentage in goat and sheep milk products by isotopic dilution-ultra-high performance liquid chromatography-tandem mass spectrometry. *Analytical and Bioanalytical Chemistry*, 409(1), 213–224. https://doi.org/10.1007/s00216-016-9987-9

Kong, J. and Yu, S. (2007). Fourier transform infrared spectroscopic analysis of protein secondary structures. *Acta Biochimica et Biophysica Sinica*, 39 (8), 549–559. https://doi.org/10.1111/j.1745-7270.2007.00320.x

Pereira, E.V. dos S., Fernandes, D.D. de S., de Araújo, M.C.U., Diniz, P.H.G.D. and Maciel, M.I.S. (2020). Simultaneous determination of goat milk adulteration with cow milk and their fat and protein contents using NIR spectroscopy and PLS algorithms. *LWT-Food Science and Technology*, 127, 109427. https://doi.org/10.1016/j.lwt.2020.109427

Poonia, A., Jha, A., Sharma, R., Singh, H.B., Rai, A.K. and Sharma, N. (2017). Detection of adulteration in milk: A review. *International Journal of Dairy Technology*, 70(1), 23–42. https://doi.org/10.1111/1471-0307.12274

Raynal-Ljutovac, K., Lagriffoul, G., Paccard, P., Guillet, I. and Chilliard, Y. (2008). Composition of goat and sheep milk products: An update. *Small Ruminant Research*, 79(1), 57–72. https://doi.org/10.1016/j.smallrumres.2008.07.009

Rebechi, S.R., Vélez, M.A., Vaira, S. and Perotti, M.C. (2016). Adulteration of Argentinean milk fats with animal fats: Detection by fatty acids analysis and multivariate regression techniques. *Food Chemistry*, 192, 1025–1032. https://doi.org/10.1016/j.foodchem.2015.07.107

Rohman, A., Windarsih, A., Riyanto, S., Sudjadi, Shuhel Ahmad, S.A., Rosman, A.S. and Yusoff, F.M. (2016). Fourier transform infrared spectroscopy combined with multivariate calibrations for the authentication of avocado oil. *International Journal of Food Properties*, 19(3), 680–687. https://doi.org/10.1080/10942912.2015.1039029

Scano, P., Murgia, A., Pirisi, F.M. and Caboni, P. (2014). A gas chromatography-mass spectrometry-based metabolomic approach for the characterization of goat milk compared with cow milk. *Journal of Dairy Science*, 97(10), 6057–6066. https://doi.org/10.3168/jds.2014-8247

Sharma, R., Poonam and Rajput, Y.S. (2010). Methods for detection of soymilk adulteration in milk. *Milchwissenschaft*, 65(2), 157–160.

Silanikove, N., Leitner, G., Merin, U. and Prosser, C.G. (2010). Recent advances in exploiting goat's milk: Quality, safety and production aspects. *Small Ruminant Research*, 89(2–3), 110–124. https://doi.org/10.1016/j.smallrumres.2009.12.033

Valdés, A., Beltrán, A., Mellinas, C., Jiménez, A., and Garrigós, M.C. (2018). Analytical methods combined with multivariate analysis for authentication of animal and vegetable food products with high fat content. *Trends in Food Science and Technology*, 77, 120–130. https://doi.org/10.1016/j.tifs.2018.05.014

Windarsih, A., Irnawati and Rohman, A. (2020). Application of FTIR - ATR spectroscopy and chemometrics for the detection and quantification of lard oil in bovine milk fat. *Food Research*, 4(5), 1732–1738. https://doi.org/10.26656/fr.2017.4(5).087

Worley, B. and Powers, R. (2013). Multivariate analysis in metabolomics. *Current Metabolomics*, 1(1), 92–107. https://doi.org/10.2174/2213235X11301010092

Yaman, H. (2020). A rapid method for detection adulteration in goat milk by using vibrational spectroscopy in combination with chemometric methods. *Journal of Food Science and Technology*, 57(8), 3091–3098. https://doi.org/10.1007/s13197-020-04342-4

 © 2022 The Authors. Published by Rynnye Lyan Resources