

## Rapid classification and quantification of mixture containing Sumbawa's horse milk and soybean milk using Near Infrared spectroscopy combined with chemometrics

<sup>1,\*</sup> Arifah, M.F., <sup>1</sup>Pratama, N.P., <sup>2</sup>Lestari, D., <sup>3</sup>Windarsih, A. and <sup>4</sup>Rohman, A.

<sup>1</sup>Faculty of Health, Universitas Jenderal Achmad Yani Yogyakarta, 55294, Indonesia

<sup>2</sup>Faculty of Pharmacy, Universitas Muhammadiyah Kalimantan Timur, 75124, Indonesia

<sup>3</sup>The National Research and Innovation Agency (BRIN), Yogyakarta, 55861, Indonesia

<sup>4</sup>Department of Chemistry, Faculty of Pharmacy, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia

### Article history:

Received: 18 June 2024

Received in revised form: 1 October 2024

Accepted: 11 January 2025

Available Online: 19

December 2025

### Keywords:

Animal milk,  
Vegetable milk,  
NIR spectroscopy,  
Chemometrics,  
Food adulteration

### DOI:

[https://doi.org/10.26656/fr.2017.9\(6\).142](https://doi.org/10.26656/fr.2017.9(6).142)

### Abstract

Horse milk (HM) can be described as a distinctive dairy product with increased nutritional content. Soybean milk (SM) is widely mixed with animal milk due to its lower price than other milk. The adulteration of HM with SM could pose economic fraud. This study aims to develop a rapid and reliable approach for detecting HM adulteration with SM using Near Infrared (NIR) spectroscopy combined with chemometrics. The binary mixtures of HM and SM were prepared to develop calibration and validation models. The classification (with PCA and PLS-DA) and quantification analysis (using PLS regression) were applied to interpret NIR spectral data. The PCA model assessed seven principal components and no outlier data based on Summary Fit and Hotelling's T2 Range column plot. PCA and PLS-DA were utilized to distinguish between pure HM and adulterated SM in score plot and biplot. The VIP scores were generated at wavelengths 1908-1952, 2554, and 2557-2592 nm using the PLS-DA model. A combination of OH-bound, CH-bound, and NH functional groups was categorized as the wavenumber selected in VIP scores. The regression result was obtained with  $y = x - 0.0000001026$  with  $R^2 = 0.9985$ . The NIR-PLSR model could successfully distinguish HM from SM adulteration with an accuracy of  $R^2 > 0.999$  and  $Q^2 > 0.80$  and precision of  $RMSEE = 0.0164$  and  $RMSECV = 0.2556$  with good prediction. A dairy product authenticity study could be performed using NIR spectra combined with chemometric techniques.

## 1. Introduction

Horse milk (HM) is a valuable and unique dairy product compared to other types of milk. HM casein is primarily composed of  $\beta$ -casein (78.5%), with other caseins and whey proteins present in horse milk such as  $\beta$ -lactoglobulin,  $\alpha$ -lactalbumin, serum albumin, lactoferrin, and others (Czyżak-Runowska *et al.*, 2021; Kondybayev *et al.*, 2021). In Indonesia, fermented horse milk has been found to exhibit antibacterial activity (Hermawati *et al.*, 2004). Sumbawa's horse milk made in Indonesia has also been recognized as a high-priced dairy product, which may be a source of economic fraud.

Soybean (*Glycine max*) is a widely imported commodity in total global production. The main soybean consumer in the world is China, the USA, the EU, Brazil, India, and Indonesia in 2020 (Voora *et al.*, 2023). The composition of soybean contains 40% protein, 35% carbohydrate, 20% oil, ash, and other substances.

Soybean amino acids are a source of protein for humans and are employed in a variety of fermented products (Medic *et al.*, 2014). The vegetable protein of soybean is commonly a mixture of milk based on their similar functional properties and lower cost than other milk (Berhow *et al.*, 2020).

Milk fraud occurs when inferior and potentially hazardous ingredients are combined or substituted (Yulyana *et al.*, 2024). The most prevalent type of milk fraud tries to boost profits by adding other dairy (Dos Santos Pereira *et al.*, 2020; Ji *et al.*, 2022), water (Kasemsumran *et al.*, 2007), formalin (Balan *et al.*, 2020), cheese whey (Vinciguerra *et al.*, 2019), and non-dairy protein such as soybean, pea and whey protein (Du *et al.*, 2018). A distinctive dairy product with increased nutritional content and potential health advantages is horse milk (HM). Soybean milk (SM) is a type of vegetable milk that is frequently blended with milk since

\*Corresponding author.

Email: [mitsalina.fildzah.arifah@gmail.com](mailto:mitsalina.fildzah.arifah@gmail.com)

it has similar beneficial properties to milk and is less expensive than other milk. Therefore, the development of a reliable and rapid method for the authentication of horse milk with adulteration to ensure the quality of this unique dairy product is crucial (Arifah *et al.*, 2022).

Near Infrared (NIR) spectroscopy is a non-destructive and dependable method for determining chemical composition based on the interaction of electromagnetic radiation pull-through substances. NIR spectrum data is frequently used for chemical, pharmaceutical, foods, and agricultural product analysis and validation. The wavelength range of near-infrared (NIR) spectroscopy is mainly between 2500-800 nm, while the wavelength range of mid-range infrared is utilized at longer wavelength between 4000-400 nm. The NIR instrument provided similar prediction accuracy compared with the MIR instrument (Shi *et al.*, 2010). The IR spectral model determined the fatty acid (FA) composition of milk products through the cis-unsaturation absorption band and the fatty acid carbon chain around the 1600-1800 nm region. Due to detection limitations, the IR original spectrum couldn't detect FA composition in diverse samples (Lintvedt *et al.*, 2023; Sato and Kawano, 1991).

NIR spectroscopy combined with chemometrics is a rapid detection method to evaluate milk quality from ingredients of raw samples, intermediate until the final product in the food industry (Mabood *et al.*, 2017). The rapid NIR approach reflected vegetable protein content from soybean in milk powder mixtures at wavelength range of 1850-2000 nm (Maraboli *et al.*, 2002). The previous study discovered that IR spectroscopy integrated with chemometrics could distinguish horse milk from any other milks (Arifah *et al.*, 2022). Protein and amino acids were measured using IR spectra and chemometric analysis of SM (Balastreri *et al.*, 2016). However, the use of NIR spectroscopy for the authentication of HM from SM is quite limited. Therefore, the current study aims to use NIR spectroscopy in conjunction with chemometrics to authenticate HM adulteration with SM.

## 2. Materials and methods

### 2.1 Materials

Horse milk (HM) samples were collected from a horse farm in Sumbawa, West Nusa Tenggara, Indonesia. Horse milk samples were collected using sterile and insulated containers with ice boxes throughout the shipping process. Soybean milk (SM) was made by macerating dry soybeans with water and grinding them until they were smooth. Afterward, SM was collected by evaporating and filtering them using filter paper. HM

was applied to SM samples randomly in 19 various concentrations (ranging from 0% to 100% v/v in Table 1. NIR spectra were obtained for every sample using Near Infrared (NIR) Analyzer.

Table 1. Concentration of Horse Milk (HM), Soybean Milk (SM) and Mixture Milk (MM) for chemometrics modelling.

No. Samples	Horse milk (HM) (v/v)	Soybean milk (SM) (v/v)
H100%	100%	0%
H95%	95%	5%
H90%	90%	10%
H85%	85%	15%
H75%	75%	25%
H70%	70%	30%
H65%	65%	35%
H60%	60%	40%
H55%	55%	45%
H50%	50%	50%
H45%	45%	55%
H40%	40%	60%
H35%	35%	65%
H30%	30%	70%
H25%	25%	75%
H20%	20%	80%
H15%	15%	85%
H10%	10%	90%
H0%	0%	100%

### 2.2 NIR instrumentation

Spectral data were acquired in reflectance mode and recorded as absorbance using Spectra Star™ XT Near Infrared (NIR) Analyzer (Westborough, United States) at 1 nm intervals from 680 to 2600 nm. As a cleaning solvent for NIR instruments, n-Hexane solvent was used. Before each analysis, the analytical surface of the sample cell was brushed, washed with n-Hexane solvent (Merck), rinsed with distilled water, and dried with soft paper tissue. To evaluate the optimal sample preparation process, spectra were taken from milk samples. Soft Independent Modeling of Class Analogy (SIMCA) software version 18 was used to perform classification and quantification analysis for NIR spectra models such as Principal Component Analysis (PCA), Partial Least Squares-Discriminant Analysis (PLS-DA), and Partial Least Squares Regression (PLSR) (Balastreri *et al.*, 2016; Galvan *et al.*, 2022).

### 2.3 Chemometrics modelling

SIMCA is a statistical tool to compress spectral data sets and select the optimal predictive information. Principal Component Analysis (PCA) is a statistical technique used to investigate and reduce variations in

data between pure and adulterated HM fingerprints. PCA analysis interprets several key components including Score plots, a Summary of fit, Hotelling's T<sup>2</sup> Range Column Plot, and Biplot (Cui *et al.*, 2021). Score plots identified a pattern difference of HM and adulterated milk. The summary fit determined the selected principal components to explain the majority of the variance in the data. Hotelling's T<sup>2</sup> Range Column Plot is a multivariate outlier approach, and Biplot visualizes the loading plot and score plot. Furthermore, the potential of PLS-DA and PLS technique to automatically distinguish between pure milk and contaminated milk was assessed. To create prediction models from spectral data, multivariate calibration techniques such as partial least squares (PLS) are used (Rohman *et al.*, 2020). These techniques function determined the variable importance of the projections (VIP) score connected with the target variable. Score plots, summary fit, Hotelling's T<sup>2</sup> column plot, Biplot, R<sup>2</sup> values, Q<sup>2</sup> values, R<sup>2</sup>X, R<sup>2</sup>Y, and VIP scores were the evaluation parameters obtained for chemometrics analysis (Worley and Powers, 2016).

### 3. Results and discussion

#### 3.1 Near infrared fingerprinting of milk samples

Pure HM, pure SM, and representative mixture milk (HM 95%, HM 75%, and HM 50%) appeared in NIR spectra fingerprinting with a variation of absorbance from 0 to 0.3 at the wavelength region of 400-1800 nm in Figure 1. Absorbance spectra of milk were presented to be the highest at the region of 1200-1400 and 1600-1900 nm, whereas the lower absorbance was found at the region range of 600-900 nm. The soybean spectrum targeted the -OH functional group at 1016, 1645, and

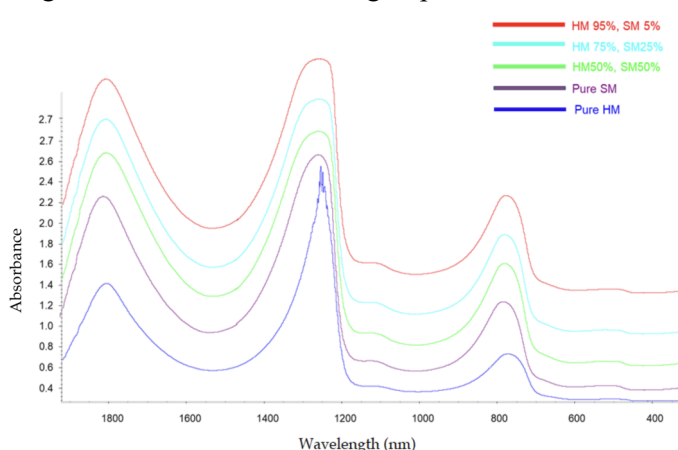


Figure 1. Pure HM, pure SM, and representative mixture milk (HM 95%, HM 75%, and HM 50%) correspond to the high variation of absorbance at the wavelength of 400-1800 nm.

1940 nm (Balastreri *et al.*, 2016). NIR represents the second overtone region at 1160 and 1445 nm, which interacts with CH<sub>3</sub>, CH<sub>2</sub>, CH, CONH<sub>2</sub>, ROH, and H<sub>2</sub>O bonds related essentially to the milk's lipid, amino acid, and water, which vary depending on milk type and adulteration levels (Galvan *et al.*, 2022). It was challenging to distinguish between the three groups since the NIR spectra of HM, SM, and MM were identical. Therefore, NIR spectroscopic data gathering using chemometrics is a required method that is appropriate for food screening.

#### 3.2 Classification analysis

The mixture of HM adulterated with SM was prepared with various concentrations. The supervised pattern of chemometrics analysis was applied to measure various concentrations of HM and SM in a range of 0-100% v/v (Table 1). The goal of this study was to implement NIR spectroscopy coupled with PCA and PLS-DA for the authentication of various milk. The PLS analysis assessed data variations based on Summary Fit, Hotelling's T<sup>2</sup> Range column plot, Biplot, R<sup>2</sup>X values and Q<sup>2</sup> values. Summary fit selected the seven principal components (PCs) to explain the majority of the variance in the data (Figure 2a). Hotelling's T<sup>2</sup> Range column plot was susceptible to the overall distance from the model center with low T<sup>2</sup> values. The low T<sup>2</sup> values of this data were well-represented by the PCA model. At 95% and 99% confidence levels, Hotelling's T<sup>2</sup> column plot identified all samples without any outliers observed in Figure 2b. Biplot also visualized the combination of the loading plot and score plot (Figure 2c).

Principal Component Analysis (PCA) is a statistical modeling tool for reducing data dimension. PCA is a practical approach to Partial Least Squares-Discriminant Analysis (PLS-DA) for reliability model. Two statistical methods quietly induce separation between experimental groups. When PCA regularly fails to cluster their groups, an alternative method to use PLS-DA model (Worley and Powers, 2016; Galvan *et al.*, 2022). Seven PCs could be extracted from PC1 (52.5%) and PC2 (36.7%) of the total variance. The PCA and PLS-DA offered accurate techniques to visualize and quantify HM-SM in score plots. Score plots also identified a 95% confidence level for all experimental groups. Score plots for PCA and PLS-DA derived from NIR spectra fingerprinting presented into four classes, such as class 1 (pure HM), class 2 (HM 0-35% v/v), class 3 (40-50% v/v), and class

Table 2. The prediction of PCA and multivariate analysis data using NIR spectra.

No.	Model	A	R <sup>2</sup> X	R <sup>2</sup> Y	Q <sup>2</sup>	RMSEE	RMSECV
1.	PCA	7	0.998	-	0.995	-	-
2.	PLS-DA	8	0.998	0.981	0.856	0.0132	0.1139
3.	PLS	9	0.999	0.999	0.860	0.0164	0.2556

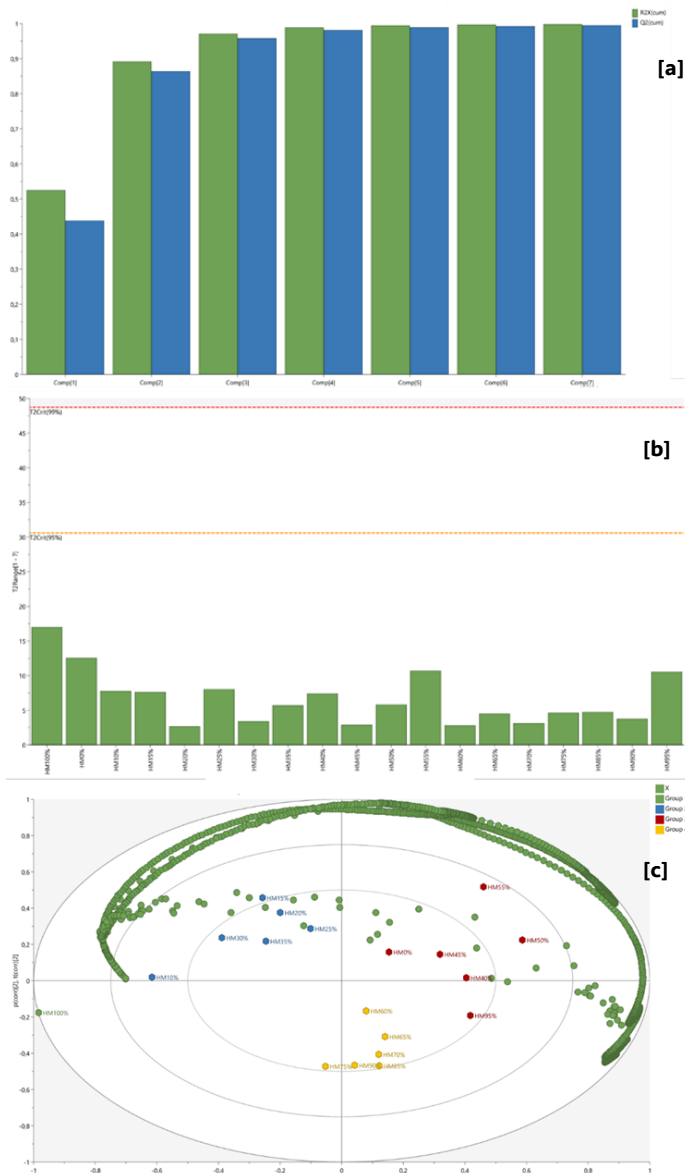


Figure 2. PCA analysis was obtained with summary of fit [a] and Hotelling's T2 Range Column Plot [b], and Biplot [c].

4 (60–95% v/v) in Figure 3. However, HM10% was less distinct from pure HM in one quadrant of the PLS-DA Figure 1. Pure HM, pure SM, and representative mixture milk (HM 95%, HM 75%, and HM 50%) correspond to the high variation of absorbance at the wavelength of 400-1800 nm.

model. NIR spectra of pure HM and adulterated milk was successfully built as a model of PCA with  $R^2X = 0.998$  and  $Q^2 = 0.995$  using three replicates of every sample. Table 2 summarizes the PLS-DA model with high accuracy and precision:  $R^2X = 0.998$ ,  $R^2Y = 0.981$ , and  $Q^2 = 0.856$ . The result of the classification analysis provided a good prediction for distinguishing pure HM from adulterated milk.

The evaluation of the permutation test with 100 permutations confirmed the PLS-DA model. Permutation test validated the PLS-DA model using intercepts of  $R^2 = 0.0$ ,  $0.565$  and  $Q^2 = 0.0$ ,  $-0.559$ . PLS-DA was evaluated regression result of original  $R^2X = 0.998$ ,  $R^2Y = 0.981$ , and  $Q^2 = 0.856$ . The  $Q^2$  value of the original model is higher than permuted models ( $Q^2 = 0.0$ , -

0,559). The original  $R^2$  values ( $R^2X = 0.998$ ,  $R^2Y = 0.981$ ) were compared to the values in the permuted models ( $R^2 = 0.0$ ,  $0.565$ ) to evaluate the permutation result in Figure 4. Hence, it can be found that the outcome of this study predicted the PLS-DA model with

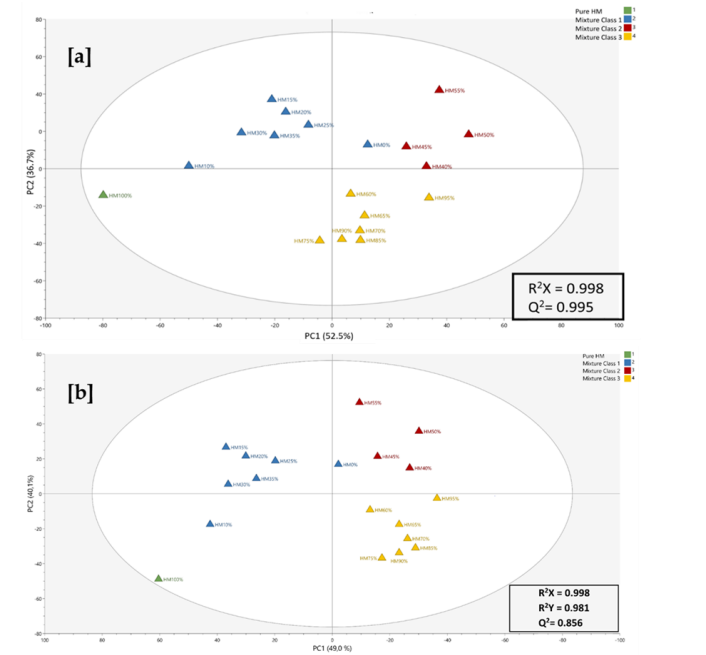


Figure 3. PCA [a] and PLS-DA [b] score plot authenticated pure HM, pure SM, and mixture HM-SM based on NIR spectra fingerprinting.

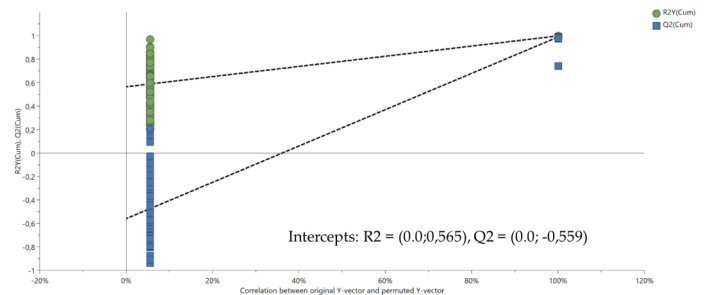


Figure 4. Permutation test determined  $R^2$  and  $Q^2$  using PLS model.

good reliability.

The variable importance in the projection (VIP) score quantifies the result of individual X variables on the PLS model, indicating which X variables contribute most significantly to the y variation (Farrés *et al.*, 2015). The VIP method selected 80 variables with VIP scores greater than 1.5 to discriminate milk classification, as shown in Figure 5. The selected VIP scores could indicate critical variables for milk classification (Galvan *et al.*, 2022). VIP scores were correlated with the target variables at 80 wavelength regions of 2554, 2557-2592, 1908, 1909, 1910, 1911, 1912, 1913, 1914, 1915, 1916, 1917, 1918, 1919, 1920, 1921, 1922, 1923, 1924 until 1952 nm (in Table 3). In another study, NIR spectra of SM were applied in the wavelength range of 1850–2000 nm (Maraboli *et al.*, 2002). Wavelengths of 2554 and 2557–2592 nm could also be categorized as CH functional group (Nandiyanto *et al.*, 2019). Wavelengths

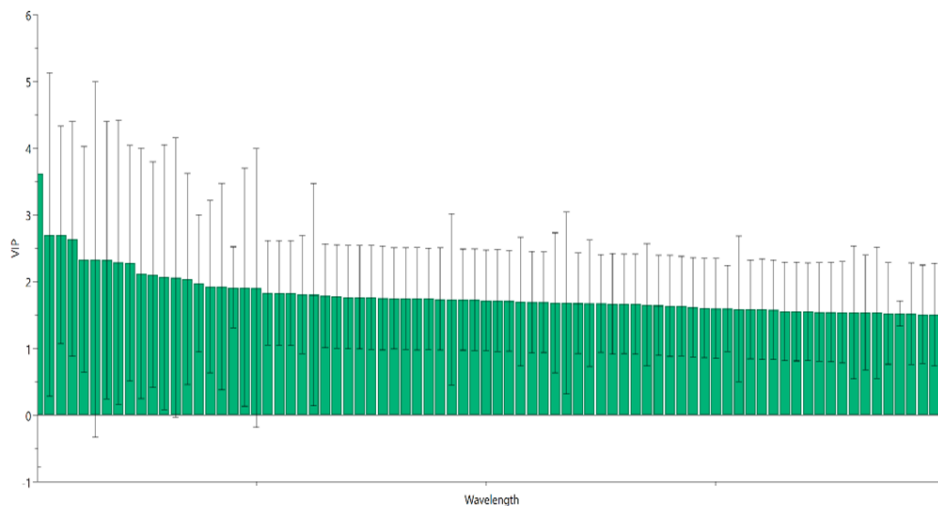


Figure 5. The variable importance in projection (VIP) revealed the presence of selected wavelengths (at 1908-1952, 2554, and 2557-2592 nm).

Table 3. Variable importance of projections (VIP) score was selected more than 1.5 score.

Region	Wavelength (nm)	VIP Score	Functional group	Ref.
Water (Soybean spectrum)	1908-1952 1940	>1.5	OH stretching and combination (OH bound to protein)	(Maraboli <i>et al.</i> , 2002) (Balastreri <i>et al.</i> , 2016)
Water (Carbohydrate)	2554	>1.5	O-H stretching bound to carbohydrate content	(Masilamani <i>et al.</i> , 2020)
Amino acid (Protein bound water)	1910-1952	>1.5	NH functional groups of the protein	(Masilamani <i>et al.</i> , 2020)
Alkana	2554,2557-2580, 2582,2584-2592	>1.5	CH functional groups	(Nandiyanto <i>et al.</i> , 2019)

of 1910 to 1980 nm indicated a combination of OH bound to protein or NH functional groups of the protein. An OH stretching of carbohydrate content was revealed at the region of 2254 nm (Masilamani *et al.*, 2020).

### 3.3 Quantification analysis

The use of NIR spectroscopy in conjunction with the Partial Least Square Regression (PLSR) model to assess and quantify HM adulteration with SM was examined. The selection of the PLS regression model of pure HM, pure SM, and MM performed accuracy and precision analyses to be validated. Seventeen concentrations of distinct samples, pure HM, and pure SM were used in the accuracy and precision studies. The PLS of accuracy and

precision criteria are listed in Table 2. The parameter validation of PLS evaluated  $R^2X = 0.999$ ,  $R^2Y = 0.999$ , and  $Q^2 = 0.86$  with a good prediction model. The regression results explained a reliable correlation between the observed and predicted values in Figure 6. The prediction model was obtained with  $y = x - 0.00000001026$  with  $R^2 = 0.9985$ . The lower error of root means square error of estimate (RMSEE) = 0.0164 and root mean square error of cross-validation (RMSECV) = 0.2556 to avoid overprediction. Both models achieved low values for RMSEE and RMSECV, indicating good precision (Windarsih *et al.*, 2023). The goal of this research is to provide a reliable and fast analysis technique for dairy products. The presented

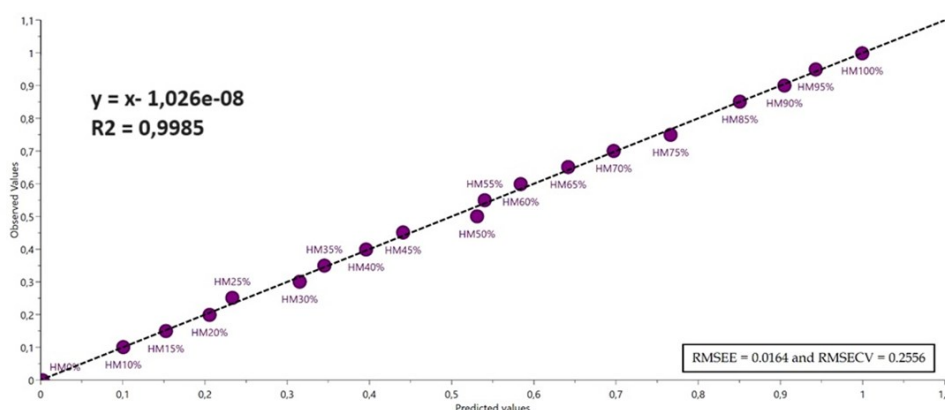


Figure 6. Regression linear of PLS modelling showed good linearity in HM and adulteration concentrations (0-100% v/v) without any outlier.

approach of PLS successfully separated and distinguished HM from contaminated SM. This PLS method was successfully implemented to detect adulteration in milk, which could be applied to consumers in future. PLS approach could be developed for quality control or milk authenticated using accessible portable devices or testing kits. Therefore, The PLS model could ensure milk authenticity, safety, and prevention of milk fraud (Grassi et al., 2022).

#### 4. Conclusion

The combination of NIR spectroscopy with chemometrics provides a reliable and accurate method for the authentication of horse milk (HM) adulteration with soybean milk (SM). Chemometric and validation of PLS were both capable of detecting SM as adulterant at all concentration levels (5-100%). The developed NIR spectra with the PLSR model could distinguish HM from SM adulteration with a high degree of accuracy and precision. The outcomes of this study demonstrated that combining NIR spectra with chemometric analysis is an effective method for screening specific wavelengths (at regions of 2554, 2557-2592, and 1908-1952 nm) to evaluate the authenticity of HM and adulteration.

#### Conflict of interest

The authors declare no conflict of interest.

#### Acknowledgments

The authors acknowledge the funding support from the Penelitian Dosen Pemula (PDP) 2025 grant under the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia (Contract No. 126/C3/DT.05.00/PL/2025). The authors also extend their appreciation to the National Research and Innovation Agency (BRIN) and the Faculty of Health, Universitas Jenderal Achmad Yani Yogyakarta, for their support and facilitation throughout this research.

#### References

- Arifah, M.F., Irnawati, Ruslin, Nisa, K., Windarsih, A. and Rohman, A. (2022). The application of FTIR spectroscopy and chemometrics for the authentication analysis of horse milk. *International Journal of Food Science*, 2022(1), 7643959. <https://doi.org/10.1155/2022/7643959>.
- 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. <https://doi.org/10.1016/j.vibspec.2020.103033>.
- Balastri, C., Baretta, D. and Paulino, A.T. (2016). Near-infrared spectroscopy and multivariate analysis for the determination of nutritional value of soybean meal and maize bran. *Analytical Letters*, 49(10), 1548–1563. <https://doi.org/10.1080/00032719.2015.1118483>.
- Berhow, M.A., Singh, M., Bowman, M.J., Price, N.P.J., Vaughn, S.F. and Liu, S.X. (2020). Quantitative NIR determination of isoflavone and saponin content of ground soybeans. *Food Chemistry*, 317, 126373. <https://doi.org/10.1016/j.foodchem.2020.126373>.
- Cui, F., Kim, M., Park, C., Kim, D., Mo, K. and Kim, M. (2021). Application of principal component analysis (PCA) to the assessment of parameter correlations in the partial-nitrification process using aerobic granular sludge. *Journal of Environmental Management*, 288, 112408. <https://doi.org/10.1016/j.jenvman.2021.112408>.
- Czyżak-Runowska, G., Wójtowski, J.A., Danków, R. and Stanisławski, D. (2021). Mare's milk from a small polish specialized farm—basic chemical composition, fatty acid profile, and healthy lipid indices. *Animals*, 11(6), 1590. <https://doi.org/10.3390/ani11061590>.
- Dos Santos Pereira, E.V., de Sousa Fernandes, D.D., 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*, 127, 109427. <https://doi.org/10.1016/j.lwt.2020.109427>.
- Du, L., Lu, W., Cai, Z., Bao, L., Hartmann, C., Gao, B. and Yu, L. (2018). Rapid detection of milk adulteration using intact protein flow injection mass spectrometric fingerprints combined with chemometrics. *Food Chemistry*, 240, 573–578. <https://doi.org/10.1016/j.foodchem.2017.07.107>.
- Farrés, M., Platikanov, S., Tsakovski, S. and Tauler, R. (2015). Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *Journal of Chemometrics*, 29(10), 528–536. <https://doi.org/10.1002/cem.2736>.
- Galvan, D., Lelis, C.A., Effting, L., Melquiades, F.L., Bona, E. and Conte-Junior, C.A. (2022). Low-cost spectroscopic devices with multivariate analysis applied to milk authenticity. *Microchemical Journal*, 181, 107746. <https://doi.org/10.1016/j.microc.2022.107746>.
- Grassi, S., Tarapoulouzi, M., D'alessandro, A.,

- Agriopoulou, S., Strani, L. and Varzakas, T. (2022). How Chemometrics Can Fight Milk Adulteration. *Foods*, 12(1), 139. <https://doi.org/10.3390/foods12010139>.
- Hermawati, D., Sudarwanto, M., Soekarto, S.T., Zakaria, F.R., Sudradjat, S. and Tjatu, R.F.S. (2004). Aktivitas antimikroba pada susu kuda sumbawa. *Teknologi dan Industri Pangan*, 15(1), 47–53.
- Ji, Z., Zhang, J., Deng, C., Hu, Z. and Du, Q. (2022). Identification of mare milk adulteration with cow milk by liquid chromatography-high resolution mass spectrometry based on proteomics and metabolomics approaches. *Food Chemistry*, 405, 134901. <https://doi.org/10.1016/j.foodchem.2022.134901>.
- Kasemsumran, S., Thanapase, W. and Kiatsoonthon, A. (2007). Feasibility of near-infrared spectroscopy to detect and to quantify adulterants in cow milk. *Analytical Sciences*, 23(7), 907–910. <https://doi.org/10.2116/analsci.23.907>.
- Kondybayev, A., Loiseau, G., Achir, N., Mestres, C. and Konuspayeva, G. (2021). Fermented mare milk product (Qymyz, Koumiss). *International Dairy Journal*, 119, 105065. <https://doi.org/10.1016/j.idairyj.2021.105065>.
- Lintvedt, T.A., Andersen, P.V., Afseth, N.K., Heia, K., Lindberg, S.K. and Wold, J.P. (2023). Raman spectroscopy and NIR hyperspectral imaging for in-line estimation of fatty acid features in salmon fillets. *Talanta*, 254, 124113. <https://doi.org/10.1016/J.TALANTA.2022.124113>.
- Mabood, F., Jabeen, F., Ahmed, M., Hussain, J., Mashaykhi, S.A.A. Al Rubaiey, Z.M.A. Al Farooq, S., Boqué, R., Ali, L., Hussain, Z., Al-Harrasi, A., Khan, A. L., Naureen, Z., Idrees, M. and Manzoor, S. (2017). Development of new NIR-spectroscopy method combined with multivariate analysis for detection of adulteration in camel milk with goat milk. *Food Chemistry*, 221, 746–750. <https://doi.org/10.1016/j.foodchem.2016.11.109>.
- Maraboli, A., Cattaneo, T.M.P., Giangiacomo, R., Maraboli, A., Maria, T., Cattaneo, P. and Giangiacomo, R. (2002). Detection of vegetable proteins by NIR spectroscopy detection of vegetable proteins from soy, pea and wheat isolates in milk powder by near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 10(1), 63–69.
- Masilamani, P., Venkatesan, S., Janaki, P., Evera, T., Sundareswaran, S. and Rajkumar, P. (2020). Role of near-infrared spectroscopy in seed quality evaluation: a review. *Agricultural Reviews*, 41(2), 106-115. <https://doi.org/10.18805/ag.r-1960>.
- Medic, J., Atkinson, C. and Hurburgh, C.R. (2014). Current knowledge in soybean composition. *Journal of the American Oil Chemists' Society*, 91(3), 363–384. <https://doi.org/10.1007/s11746-013-2407-9>.
- Nandiyanto, A.B.D., Oktiani, R. and Ragadhita, R. (2019). How to read and interpret FTIR spectroscopy of organic material. *Indonesian Journal of Science and Technology*, 4(1), 97-118.
- Rohman, A., Ghazali, M.A.I.B., Windarsih, A., Irnawati, Riyanto, S., Yusof, F.M. and Mustafa, S. (2020). Comprehensive review on application of FTIR spectroscopy coupled with chemometrics for authentication analysis of fats and oils in the food products. *Molecules*, 25(22), 5485. <https://doi.org/10.3390/molecules25225485>.
- Sato, T. and Kawano, S. (1991). Near infrared spectral patterns of fatty acid analysis from fats and oils, *Journal of the American Oil Chemists Society*, 68, 827–833. <https://doi.org/https://doi.org/10.1007/BF02660596>.
- Shi, G., Han, L., Yang, Z., Chen, L. and Liuab, X. (2010). Near infrared spectroscopy calibration transfer for quantitative analysis of fish meal mixed with soybean meal. *Journal of Near Infrared Spectroscopy*, 18(3), 217–223. <https://doi.org/10.1255/jnirs.878>.
- Vinciguerra, L.L., Marcelo, M.C.A., Motta, T.M.C., Meneghini, L.Z., Bergold, A.M. and Ferrão, M.F. (2019). Chemometric tools and FTIR-ATR spectroscopy applied in milk adulterated with cheese whey. *Quimica Nova*, 42(3), 249–254. <https://doi.org/10.21577/0100-4042.20170337>.
- Voora, V., Larrea, C. and Bermúdez, S. (2023). Global market report: soybeans. Retrieved from website: <https://www.iisd.org/system/files/2024-02/2024-global-market-report-soybean.pdf>
- Windarsih, A., Rohman, A., Irnawati and Riyanto, S. (2021). The combination of vibrational spectroscopy and chemometrics for analysis of milk products adulteration. *International Journal of Food Science*, 2021(1), 8853358, <https://doi.org/10.1155/2021/8853358>.
- Windarsih, A., Arifah, M.F., Suratno and Rohman, A. (2023). The application of untargeted metabolomics using UHPLC-HRMS and chemometrics for authentication of horse milk adulterated with cow milk. *Food Analytical Methods*, 16(2), 401-412. <https://doi.org/10.1007/s12161-022-02426-9>.
- Worley, B. and Powers, R. (2016). PCA as a practical indicator of OPLS-DA model reliability. *Current Metabolomics*, 4(2), 97–103. <https://doi.org/10.2174/2213235x04666160613122429>.
- Yulyana, A., Rohman, A., Arifah, M.F., Windarsih, A.

and Nisa, K. (2024). Authentication analysis of goat milk from cow milk using Fourier transform infrared spectroscopy and chemometrics. *Food Research*, 8 (2), 459–466. [https://doi.org/10.26656/fr.2017.8\(2\).1025](https://doi.org/10.26656/fr.2017.8(2).1025).

Zain, S.M., Behkami, S., Bakirdere, S. and Koki, I.B. (2016). Milk authentication and discrimination via metal content clustering-A case of comparing milk from Malaysia and selected countries of the world. *Food Control*, 66, 306–314. <https://doi.org/10.1016/j.foodcont.2016.02.015>.