Original Article
Food Science and Technology

ISSN 0101-2061 (Print)
ISSN 1678-457X (Online)

DOI: https://doi.org/10.1590/fst.104622

Tracing models for checking beef adulterated with pig blood by Fourier transform near-infrared paired with linear and nonlinear chemometrics

Mingdong LI[1], Joshua Harington AHETO[2], Marwan Mohammed Ahmed RASHED[3], Fangkai HAN[3]*

Abstract

Faced with the aggravating issue of meat fraud caused by the addition of low-cost animal blood food, the present work aims to develop FT-NIR-based tracing models for detecting beef adulterated with pig blood. A total of 210 samples were analyzed, including raw beef, beef adulterated with pig blood-based gel, and pure pig blood-based gel prepared. For spectrum denoising, the first derivative, second derivative, centralization, standard normal variate transform, and multivariate scattering correction algorithms were performed and compared. We built, optimized, and compared partial least squares (PLS), support vector machine (SVM), and extreme learning machine (ELM) models for identifying the adulterated beef and predicting adulteration levels. Results indicated that second derivative was the best preprocessed technique for all chemometrics modeling; ELM model achieved the optimal performance when the sin function was used, achieving 100% accuracy for identifying the adulterated beef, and all root mean square errors were less than 0.16% for predicting adulteration levels in training and test sets. These results suggest the optimal ELM models could be employed for rapidly checking beef adulterated with pig blood-based gel using FT-NIR technology.

Keywords: beef adulteration; FT-NIR; pig blood; machine learning; rapid analysis.

Practical Application: The optimal ELM models for identifying the adulterated beef mixed with pig blood-based gel and predicting adulteration levels were constructed using FT-NIR technology, which can be used to rapidly detect beef adulterated with pig blood, hence preventing illegal mixing and unfair competition.

1 Introduction

Beef has attracted the attention of meat adulterators for centuries due to its high commercial value. Typical case of beef adulteration is the inter-species confounding, and deceives consumers by replacing beef with cheaper alternatives (Han et al., 2020). Therefore, themes for developing analytical techniques reported to measure beef adulteration were mainly focused on the low-cost meats used, such as pork (Kang & Tanaka, 2018), duck (Jiang et al., 2019), chicken (Silva et al., 2020), and other meat (Li et al., 2019). However, a market survey reported in 2019 shown that there was a high possibility for meat fraud with pig blood added in beef (Ma, 2019). The main reason for this type of beef adulteration maybe via using beef soaked in pig blood or added with pig blood-based gel for making meat ball to increase commodity weight. The bovine blood was not used for beef fraud caused by its unpleasant odors. In 2020, the Public Security Bureau of Changchun City, Jilin Province reported the similar case of beef fraud with pig blood used. Although rarely a health hazard, this type of beef adulteration violates consumer's benefit seriously, and hampers the development of the domestic meat industry gravely. The emerging issue of such situation promotes us to propose a practical strategy for rapidly checking beef adulterated with pig blood to fight food fraud.

The specificity analysis technology based on biomarkers such as DNA or protein showed high accuracy for detecting meat adulteration. However, these techniques are usually time-consuming, expensive and require professional skill to operate, hence are not practically suitable for on-site detection (Song et al., 2021). Several rapid analytical techniques for meat adulteration that have been reported, such as electronic nose (Kalinichenko & Arseniyeva, 2020; Sarno et al., 2020; Wang et al., 2019), electronic tongue (Lu et al., 2021; Tian et al., 2019; Zaukuu et al., 2021), hyperspectral imaging (Jiang et al., 2019, 2020; Rady & Adedeji, 2020; Zheng et al., 2019). Compared with the previously mentioned rapid analytical techniques, the Fourier transform near-infrared (FT-NIR) spectroscopy technology is more practical due to its advantages of rapidity, convenience, non-destructive and reliability (Wang et al., 2022a). The FT-NIR spectroscopy is widely used to develop analytical techniques for authenticating meat, for example, analysis of chicken meat authenticity (Parastar et al., 2020), quantitative detection of binary and ternary adulteration of minced beef with pork and duck meat (Leng et al., 2020), fast detection and quantification of pork meat in other meats (Mabood et al., 2020), detection of minced lamb and beef fraud (López-Maestresalas et al., 2019), to name but a few. However, regarding the aggravating issue of meat fraud using low-cost animal blood, there is no research report about the rapid analysis technology using FT-NIR spectroscopy.

Received 06 Oct., 2022
Accepted 05 Dec., 2022
[1]*School of Information and Engineering, Suzhou University, Suzhou, Anhui, China*
[2]*School of Food and Biological Engineering, Jiangsu University, Zhenjiang, Jiangsu, China*
[3]*School of Biological and Food Engineering, Suzhou University, Suzhou, Anhui, China*
*Corresponding author: hanfk11@163.com

FT-NIR technology can be used to measure many attributes of food materials simultaneously. Absorption bands in the NIR region (780–2500 nm) come from overtones and combinations of overtones and/or combinations of fundamental vibrational motions of O-H, N-H, C-H, or S-H, resulting in a complicated relationship between FT-NIR spectral and food compositions. Therefore, while FT-NIR technology is utilized for food measurements, chemometrics models must be constructed and optimized for goals of quality-based classification or prediction of quality indicators. The FT-NIR spectroscopy coupled with the optimized model constitutes a unique analytical strategy for the specific targets and objects (Fodor et al., 2020; Wang et al., 2022b). Although many FT-NIR-based techniques could be standardized, they still worthy of further research to construct reliable and efficient chemometrics models for precise target measurement.

Hence, facing the aggravating issue of meat fraud using cost-effective animal blood, the present work aims to construct tracing models for checking beef adulterated with pig blood by FT-NIR technology. The measurement spectra datasets were collected from raw beef, beef adulterated with pig blood-based gel, and pure pig blood-based gel. Algorithms of the first derivative (1st Der), second derivative (2nd Der), centralization, standard normal variate transform (SNV), and multivariate scattering correction (MSC) were performed in comparison for spectral denoising. Partial least squares (PLS), support vector machine (SVM), and extreme learning machine (ELM) models were constructed, optimized, and compared for identifying the adulterated beef and predicting adulteration levels. The proposed strategy for checking beef adulterated with pig blood was, step 1: acquisition of FT-NIR spectra for an unknown sample; step 2: the identification model used for checking whether beef mixed with pig blood or not; step 3: if the answer is yes, the quantitative model was utilized for predicting the adulteration level.

2 Materials and methods

2.1 Samples prepared

Fresh foreshank and fresh pig blood were purchased at a local slaughterhouse in Suzhou, China. The raw beef and pig blood were transported to the laboratory in an ice-filled container. The fresh pig blood was precipitated and sterilized at 100 °C for 40 minutes. All samples were homogenized and then frozen and kept at a temperature at−20 °C for the FT-NIR measurements.

In total, 210 samples were prepared and used, including 30 samples of raw beef and 30 samples of pure pig blood-based gel, and 150 samples of adulterated beef via beef mixed with pig blood-based gel in a range of 1~5% by weight at 1% steps, 30 samples for each adulteration level.

2.2 FT-NIR measurements

The FT-NIR spectrophotometer (Antaris II, Thermo Electron Company, USA) interfaced with an optical fiber was employed. Twenty-five grams of the sample were prepared for acquiring the FT-NIR spectral. Each sample was measured three times with a spectral resolution of 8.0 cm^{-1} at different positions. Following the operating procedures, each spectrum consisted of an average of 32 scans ranging from 4000–10000 cm^{-1}, yielding 1557 variables for each sample (Teye et al., 2014). Due to the optical fiber used, it is possible to separate the spectrometer from the sample over several meters. Thus, industrial installations with a high degree of flexibility and complete automation are possible (Han et al., 2022c).

2.3 Chemometrics modeling and software

Chemometrics models were constructed and optimized for detecting adulterated beef and predicting adulteration levels. For the purpose of spectrum denoising, the 1st Der, 2nd Der, centralization, SNV, and MSC were all used and compared (Kumar & Chandrakant Karne, 2017). Comparisons were made among models constructed using PLS, SVM, and ELM. PLS and SVM are widely used chemometrics for both classification and regression, with their underlying theory documented in many published works including those by Y.V. Zontov et al. (2020), and Shan Suthaharan (2016).

ELM is a novel algorithm for a single-hidden layer feed-forward neural network that combines the benefits of good generalization performance and extreme fast learning speed. The connection weight between the input and hidden layers, and the hidden layer neuron threshold were generated randomly. Hence, the network structure of ELM is more straightforward than the traditional gradient descent-based artificial neural networks previous widely used. While the numbers of hidden neurons was fixed, the unique optimal solution could be achieved by ELM. For its extreme learning speed and good generalization performance, ELM has attracted the attention of food researchers (Wang et al., 2020). The calculation process of the ELM is described by Equations 1 to 10 below.

Given that the input dataset (x) and the output dataset (y) of the training set are:

$$\mathrm{x}=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1Q} \\ x_{21} & x_{22} & \cdots & x_{2Q} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nQ} \end{bmatrix}_{n\times Q}, y=\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1Q} \\ y_{21} & y_{22} & \cdots & y_{2Q} \\ \vdots & \vdots & & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mQ} \end{bmatrix}_{m\times Q} \quad (1)$$

Where: n, l, and m mean numbers of the input, hidden and output neurons, respectively; Q means the sample size of training set.

Training an ELM model could be divided into the following three steps (Han et al., 2019; Yu, 2015).

Step1: Random generation of the input weight matrix (w) and the bias matrix (b);

$$w=\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & & \vdots \\ w_{l1} & w_{l2} & \cdots & w_{\mathrm{ln}} \end{bmatrix}_{l\times n} \quad (2)$$

Where: w_{ji} means connection weight of the i^{TH} neuron node in the input layer and the j^{TH} neuron node in the hidden layer.

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_l \end{bmatrix}_{l\times 1} \tag{3}$$

Step2: Select a transfer function *g(x)*, such as sigmoidal function, sin function, or hardlim function, which was frequently used in artificial neural networks.

Step3: Calculate the connection weight (β) between hidden and output layers.

$$\beta = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & & \vdots \\ \beta_{l1} & \beta_{l2} & \cdots & \beta_{lm} \end{bmatrix}_{l\times m} \tag{4}$$

Where: β_{jk} means the connection weight between the j^{TH} neuron node in the hidden layer and the k^{TH} neuron node in the output layer.

Given output of the neural network *T* is:

$$T = \left[t_1, t_2, \cdots, t_Q\right]_{m\times Q} \tag{5}$$

$$t_j = \begin{bmatrix} \sum_{i=1}^{l} \beta_{i1} g(w_i x_j + b_i) \\ \sum_{i=1}^{l} \beta_{i2} g(w_i x_j + b_i) \\ \vdots \\ \sum_{i=1}^{l} \beta_{im} g(w_i x_j + b_i) \end{bmatrix}, \quad j = 1, 2, ..., Q \tag{6}$$

Where: $w_i = \left[w_{i1}, w_{i2}, \cdots, w_{in}\right], x_j = \left[x_{1j}, x_{2j}, \cdots, x_{nj}\right]^T$.

The Equation 5 could be expressed as,

$$H\beta = T^T \tag{7}$$

Where: *H* means the output of the hidden layer of the neural network.

$$H = \begin{bmatrix} g(w_1 x_1 + b_1) & g(w_2 x_1 + b_2) & \cdots & g(w_l x_1 + b_l) \\ g(w_1 x_2 + b_1) & g(w_2 x_2 + b_2) & \cdots & g(w_2 x_2 + b_2) \\ \vdots & \vdots & & \vdots \\ g(w_1 x_Q + b_2) & g(w_2 x_Q + b_2) & \cdots & g(w_l x_Q + b_l) \end{bmatrix}_{Q\times l} \tag{8}$$

The following equation could calculate the β according to the theory described by Pro. Guangbin Huang et al. (2006, 2011),

$$\hat{\beta} = H^+ T' \tag{9}$$

For multi-classification, the decision function is,

$$lable(x) = \arg\max T_i(x), i \in (1, 2, \cdots, m) \tag{10}$$

Where: *x* means the unknown sample waiting classified, *m* represents labels of the training samples.

Performances of PLS, SVM, and ELM for identifying the adulterated beef were evaluated by the following Equation 11,

$$R = \frac{N_1}{N_2} \times 100\% \tag{11}$$

Where *R* means the identification accuracy (%) of the training or test set; N_1 means the number of the correctly classified samples; N_2 means the sample size of the training or test set.

Performances of PLS, SVM, and ELM for predicting adulteration levels were evaluated by the root mean square error of cross-validation (*RMSECV*), the root mean squared error of prediction (*RMSEP*), the correlation coefficients (*r*) in the training set (r_t) and prediction set (r_p), the ratio performance to deviation (*RPD*), and the range error ratio (*RER*) (El Orche et al., 2022; Han et al., 2022b; Mahanti et al., 2020). The *RMSECV*, *RMSEP*, *r*, *RPD,* and *RER* were calculated with the following Equations 12 to 16,

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n_1} (\hat{y}_{\backslash i} - y_i)^2}{n_1}} \tag{12}$$

Where n_1 means the sample size of the training set, y_i means the actual adulteration level of the i^{th} sample, and $\hat{y}_{\backslash i}$ means the predicted adulteration level after the i^{th} sample was removed using cross-validation.

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n_2} (y_i - \hat{y}_i)^2}{n_2}} \tag{13}$$

Where n_2 means the sample size of the prediction set, y_i means actual adulteration level of the i^{th} sample, and $\hat{y}_i$ means predicted adulteration level of the i^{th} sample in prediction set.

$$\mathrm{r} = \sqrt{1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \tag{14}$$

Where: $\bar{y}$ means the mean value of actual adulteration levels in training or prediction set.

$$RPD = \frac{\mathrm{Std.}}{RMSECV} \tag{15}$$

Where: *Std.* means standard deviation of the training set samples' actual adulteration levels. For *RPD*, so bigger is better.

An *RPD* above three is considered satisfactory; a value of 5 or higher indicates that the established model can be used for quality control. Regression models with an *RPD* value of 2-3 are considered to perform well enough for fast screening analysis (Wiedemair et al., 2018).

$$RER = \frac{Max - Min}{RMSECV} \quad (16)$$

Where: *Max* and *Min* mean the maximum and the minimum adulteration levels. *RER* above 10 is roughly an indicator of a model with good predictive ability.

All algorithms were implemented in Matlab Version 7.14 (Mathworks, Natick, USA) with windows 10.

3 Results and discussions

3.1 Inputs prepared for modeling

The original FT-NIR spectroscopy profile for raw beef, beef adulterated with pig blood, and pure pig blood-based gel are depicted in Figure 1. As illustrated in Figure 1b, overlapping information existed in region between 8868 cm^{-1} and 10000 cm^{-1}, and hence the region was removed before modeling in order to identify adulterated beef. Afterwards, the 1st Der, 2nd Der, centralization, SNV, and MSC were used independently to denoise the spectra features. Principal component analysis (PCA) was used to reduce the dimensionality of the preprocessed FT-NIR datasets by creating new uncorrelated variables called principal components (PCs). The top three PC scores were utilized as inputs for modeling process because their cumulative contribution rates were all greater than 96.0%.

In terms of modeling for adulteration levels prediction, the original FT-NIR datasets of adulterated beef were preprocessed using 1st Der, 2nd Der, centralization, SNV, and MSC. Afterwards, the competitive adaptive reweighted sampling (CARS) (Li et al., 2009) technique was used to screen the preprocessed FT-NIR spectral for important spectral wavelengths, the results of which are shown in Figure 2. PCA was also implemented on the selected important spectral wavelengths for the purposes of reducing their dimensions and decorrelating their vectors. Regarding the important FT-NIR variables of none preprocessed, 1st Der,

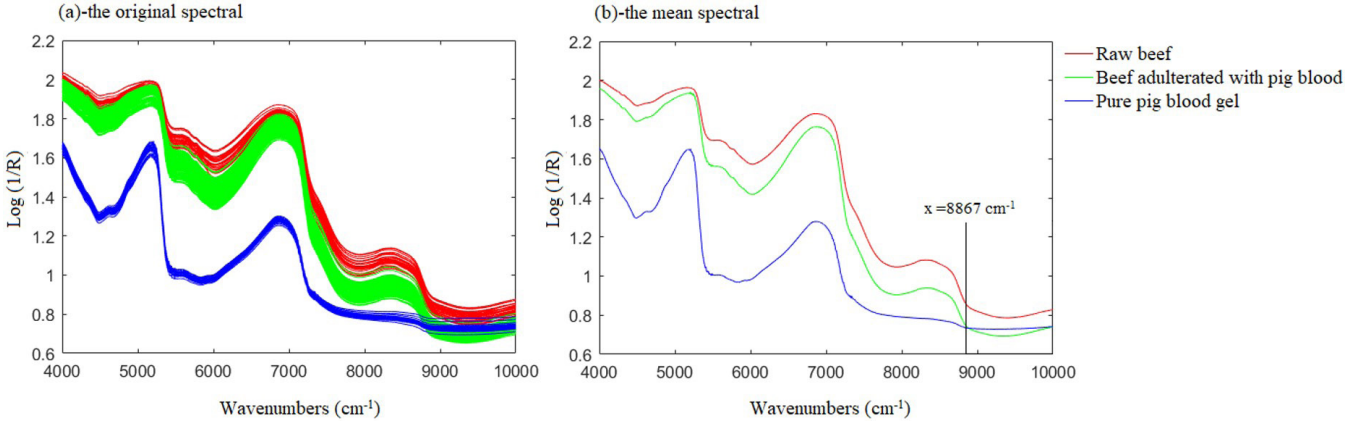


Figure 1. The original and average FT-NIR spectral for raw beef, pure pig blood-based gel, and beef adulterated with pig blood-based gel.

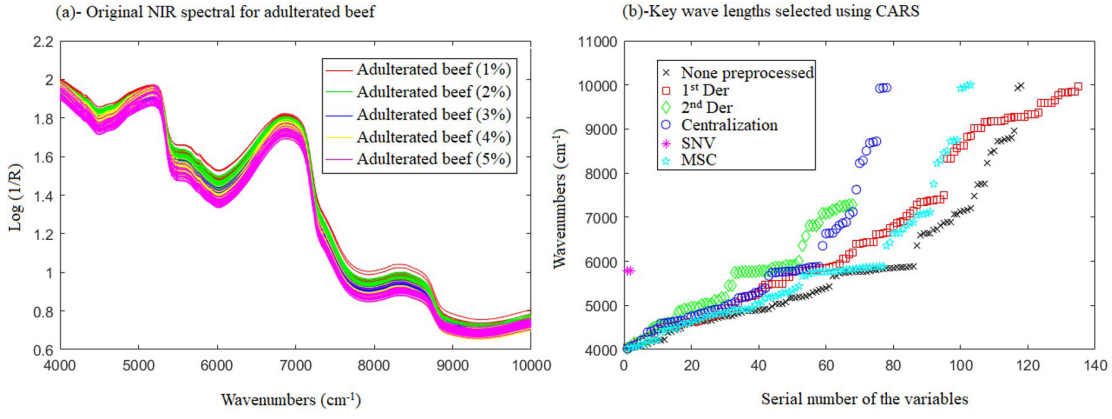


Figure 2. The crucial spectral wavelengths selected for predicting adulteration levels via CARS.

centralization, MSC, and the scores of the top three PCs were utilized as input variables for modeling since their cumulative contribution rates were all over 98.0%. PCA results showed that the cumulative contribution rate of the first 13 PCs of 2nd Der spectral was 90.14%, indicating that the supporting scores of the top 13 PCs could be utilized for modeling; For the SNV spectral, only two wave numbers were selected via CARS (see Figure 2b), therefore all the PCs were used for modeling.

To ensure that the PLS, SVM, and ELM models all used with the same inputs, the Kennard-Stone algorithm was utilized to select one third of the samples in each group as the prediction set and the remaining samples as the training set (Zhang et al., 2017). The groups used in the modeling process to classify the adulterated beef are raw beef, beef adulterated with pig blood, and pure pig blood-based gel; Groups for modeling to predict adulteration levels means beef mixed with pig blood gel in a range of 1~5% by weight at 1% steps.

3.2 Results of PLS models

PLS modeling for identifying the adulterated beef and predicting the adulteration levels were constructed herein. PLS for discriminant analysis (DA), also known as PLSDA was implemented using a MATLAB GUI tool created by Y.V. Zontov and co-authors (Zontov et al., 2020). The results of the PLS models generated are shown in Table 1.

The results of PLS models showed that, when the 2nd Der technique was used for spectral denoising, the optimal PLS models could be obtained, yielding the highest correlation coefficients over 0.95, *RPD* superior to 3.0, and the lowest *RMSEP* of 0.441% for the unknown samples' set; only one samples was misclassified in all training and test sets for identifying the adulterated beef, it is one sample from group of beef adulterated with pig blood was misclassified to the group of raw beef.

3.3 Results of SVM

According to SVM theory, a kernel function is required to accept inputs from low-dimensional spaces, and to calculate the inner product value of vectors in high-dimensional space after a certain transformation, in order to transform the problem of linear inseparability in low-dimensional space into that of linear separability or approximate linear separability in high-dimensional space.

The radial basis function (RBF) is a frequently used kernel functions in SVM modeling, and demonstrated in Formula 17. RBF may be applied to randomly distributed samples after parameter optimization and is regarded as a universal kernel function. Hence, this work adopts RBF function as the kernel function for SVM modeling.

$$K(x\ ,x_i)=\exp(-g\left\|x-x_i\right\|^2),g>0 \tag{17}$$

Parameters of penalty factors *c* and *g* in RBF greatly influence the performance of SVM models. Therefore, the grid division method coupled with mutual verification was utilized to optimize these two parameters achieve the highest possible detection accuracy for adulterated beef and the minimum *RMSECV* for predicting adulteration levels. For the purpose of parameter optimization, the ranges of *c* and *g* were all set to $[2^{-8}, 2^{8}]$ with a step size of 0.5.

Table 2 summarizes the results of the SVM models constructed. It can be seen that the optimal SVM models could be obtained also under the 2nd Der technique was used for spectral denoising, yielding the highest correlation coefficients over 0.95, *RPD* superior to 9.0, and the lowest *RMSEP* of 0.214% for prediction adulteration levels; identical to PLSDA model, there was only one sample misclassified, it is one sample of beef adulterated with pig blood misclassified to raw beef. Figure 3 shows the optimized *c* and *g* for these best SVM models for identifying adulterated beef and predicting adulteration levels.

3.4 Results of ELM

According to ELM theory, the hidden neuron numbers and the activation function of hidden layers should be optimized for a particular pattern recognition problem. For the purpose of optimizing the hidden neuron numbers, as the strategy of cut-and-trial used for artificial neural networks, ranges of the optimal hidden neuron numbers for ELM modeling to identify the adulterated beef and predict adulteration levels herein were all set at [1, 50].

Table 1. Results of PLS models for identifying the adulterated beef and predicting adulteration levels.

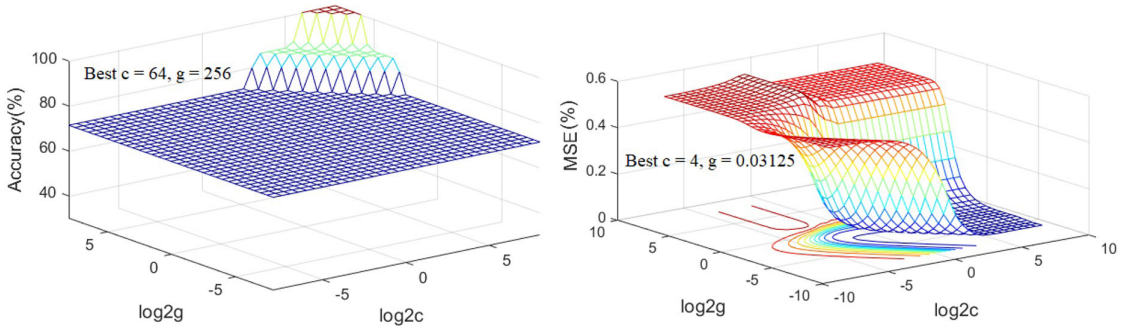
	Identifying adulteration		Predicting adulteration level					
	Training set	Test set	Training set		Test set		*RPD*	*RER*
			r_T	*RMSECV* (%)	r_p	*RMSEP* (%)		
None	100%	98.57%	0.925	0.539	0.855	0.761	2.62	7.42
1st Der	100%	98.57%	0.932	0.512	0.846	0.781	2.76	7.81
2nd Der	**100%**	**98.57%**	**0.962**	**0.387**	**0.951**	**0.441**	**3.65**	**10.3**
Centralization	100%	97.14%	0.937	0.495	0.864	0.739	2.86	8.08
SNV	100%	98.57%	0.912	0.581	0.828	0.856	2.43	6.88
MSC	100%	95.71%	0.949	0.448	0.892	0.658	3.16	8.93

Bold digits represent the optimal performance.

Table 2. Results of SVM models for identifying the adulterated beef and predicting adulteration levels.

	Identifying adulteration		Predicting adulteration level					
	Training set	Test set	Training set		Test set		*RPD*	*RER*
			r_T	*RMSECV* (%)	r_P	*RMSEP*(%)		
None	100%	97.14%	0.963	0.191	0.846	0.387	7.4	20.9
1st Der	100%	95.71%	0.963	0.191	0.859	0.364	7.42	20.9
2nd Der	**100%**	**98.57%**	**0.979**	**0.147**	**0.955**	**0.214**	**9.65**	**27.2**
Centralization	100%	97.14%	0.958	0.204	0.878	0.112	6.93	19.6
SNV	100%	97.14%	0.931	0.259	0.829	0.406	5.47	15.4
MSC	100%	94.29%	0.967	0.180	0.916	0.290	7.88	22.2

Bold digits represent the optimal performance.



(a)-Identifying adulteration (b)-Predicting adulteration levels

Figure 3. The optimized *c* and *g* for these best SVM models for identifying adulterated beef and predicting adulteration levels.

Also, the frequently used activation functions depicted in the following Equations 18 to 20 were performed in comparison for ELM modeling (Han et al., 2022a):

$$\text{Sigmoidal function}: S(x) = \frac{1}{1+e^{-x}} \tag{18}$$

$$\text{Sine function}: S(x) = \sin(x) \tag{19}$$

$$\text{Hardlim function}: S(x) = \begin{cases} 1 & x>0; \\ 0 & x\leq 0. \end{cases} \tag{20}$$

Table 3. summarizes the results of the ELM models constructed. The results demonstrated how the optimal ELM models can be obtained. The 2nd Der technique was used for spectral denoising and Sine was used as the activation function for the hidden layers, yielding the highest correlation coefficients of 0.96, *RPD* superior to 13.0, *RER* better than 38.0, and the lowest *RMSEP* of 0.158% for prediction adulteration levels; Also, all samples in the training and test sets were classified correctly, indicating an accurate identification of the adulterated beef. The optimal neural network structures of ELM models were 3-8-1 and 13-23-1 for identifying the adulterated beef and predicting adulteration levels, respectively.

3.5 General discussions

The FT-NIR with a spectral range of 4000–8867 cm^{-1} was employed to detect samples assessed in this study including raw beef, beef adulterated with pig blood, and pure pig blood. There are differences in basic organic chemicals between the raw beef and pig blood gel used. For example, crude protein content of raw beef used in this study was measured to be 36.5% using the Kjeldahl method (GB 5009.5-2016), which was found significantly higher than the crude protein content of pig blood-based gel (28.67%). Using the Soxhlet extractor method (GB 5009.6-2016), it was revealed that the crude lipids content of beef (4.97%) was also significantly higher than that of the crude lipids content of pig blood gel prepared (0.04%). Variations in the absorption of radiation at different wavelengths are related to the chemical compositions of the samples used (Mendez et al., 2019). Each of the raw beef and pure pig blood-based gel used has a characteristic spectrum (see Figure 1), which allows its identification and differentiation (see Figure 4).

The results of the chemometrics models constructed show that the 2nd Der is the best preprocessed technique for FT-NIR modeling in order to identify adulterated beef made with pig blood-based gel as well as to predict the adulteration levels. Similar findings have been published for predicting verbenalin in *Verbenaofficinalis* using NIR spectroscopy (Schönbichler et al., 2013) and evaluating freshness of pork using NIR hyperspectral imaging (Barbina et al., 2011). FT-NIR spectra tend to have linear baseline increases and these are could removed by 2nd Der technique which have negative peaks where the original had a peak and are thus more readily comprehensible (Benes et al., 2020). The 2nd

Table 3. Results of ELM models for identifying the adulterated beef and predicting adulteration levels.

		Identifying adulteration		Predicting adulteration level					
		Training set	Test set	Training set		Test set		*RPD*	*RER*
				r_T	*RMSECV*(%)	r_p	*RMSEP*(%)		
None	Hardlim	100%	98.57%	0.937	0.245	0.866	0.508	5.78	16.3
	Sin	100%	98.57%	0.944	0.217	0.866	0.533	6.52	18.4
	Sig	100%	98.57%	0.938	0.239	0.877	0.500	5.92	16.7
Dx1	Hardlim	100%	97.14%	0.933	0.260	0.868	0.546	5.44	15.4
	Sin	100%	98.57%	0.949	0.199	0.872	0.528	7.1	20.1
	Sig	100%	98.57%	0.961	0.154	0.875	0.557	9.18	26.0
Dx2	Hardlim	71.43%	71.43%	0.965	0.138	0.930	0.309	10.27	29.0
	Sin	**100%**	**100%**	**0.973**	**0.105**	**0.960**	**0.158**	**13.43**	**38.1**
	Sig	100%	100%	0.968	0.135	0.925	0.289	10.48	29.6
Center	Hardlim	100%	98.57%	0.955	0.176	0.872	0.485	8.02	22.7
	Sin	100%	98.57%	0.962	0.148	0.886	0.483	9.54	27.0
	Sig	95%	84.29%	0.963	0.146	0.885	0.537	9.67	27.4
SNV	Hardlim	100%	100%	0.934	0.257	0.845	0.607	5.51	15.6
	Sin	88.57%	78.57%	0.921	0.304	0.841	0.652	4.66	13.2
	Sig	100%	100%	0.931	0.265	0.836	0.626	5.33	15.1
MSC	Hardlim	100%	98.57%	0.836	0.603	0.813	0.763	2.34	6.63
	Sin	100%	98.57%	0.963	0.147	0.907	0.359	9.63	27.2
	Sig	100%	98.57%	0.971	0.116	0.913	0.344	12.22	34.5

Bold digits represent the optimal performance.

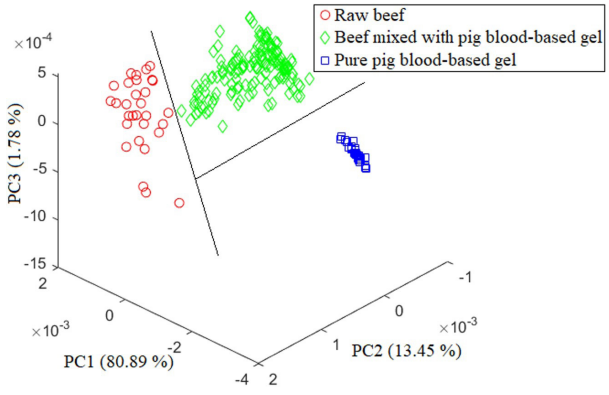


Figure 4. The scatter diagram of all the samples in the space of the top three principal components (PCs) with the second derivative FT-NIR data used.

Der of the absorption FT-NIR spectrum has the advantage of removing both baseline offset and linear slope (Koutsiaris, 2017). For these reasons, the 2nd Der often preferred for FT-NIR original spectral before modeling.

Additionally, as demonstrated by the results of chemometrics models built, ELM outperformed PLS and SVM models in terms of identifying contaminated beef and predicting adulteration levels. The reason for this could be that the relationships between the FT-NIR data matrices of raw beef, beef adulterated with pig blood gel, and pure pig blood-based gel prepared were more complex than linear as a result of the essential characteristic of FT-NIR, which is that the spectra at multiple wave numbers may contain the same chemical information, while the spectra at one wave number may contain different food chemical information. Due to its superior capacity to self-learn and self-adjust, ELM has a major advantage over PLS and SVM for processing nonlinear issues.

4 Conclusions

The aggravating problem of beef contaminated with pig blood led the development of a quick analytical procedure. The combination of FT-NIR and ELM modeling was proposed to accurately identify contaminated beef, with a prediction error of less than 0.2 percent in forecasting adulterated levels. All of the studies demonstrated that using FT-NIR spectroscopy in conjunction with ELM may rapidly detect beef contaminated with pig blood, hence preventing illegal mixing and unfair competition.

Ethical approval

The authors declare that this article does not contain any studies with human or animal subjects.

Conflict of interest

All authors declare no conflicts of interest related to this work.

Funding

Mingdong Li has received research grant from the Key Natural Science Research Project of Suzhou University (2021yzd08); Fangkai Han has received research grants from the back-up of academic and technical leaders of Suzhou University (2020XJHB04), the Natural Science Research Project of colleges and universities in Anhui province (KJ2021ZD0139), the scientific research platform project of Suzhou University (2021XJPT35), and the Support Project for Outstanding Young Talents of colleges and universities in Anhui province (gxyq2022105).

Acknowledgements

This work was sponsored by the Key Natural Science Research Project of Suzhou University (2021yzd08), the back-up of academic and technical leaders of Suzhou University (2020XJHB04), the natural science research project of colleges and universities in Anhui province (KJ2021ZD0139), the scientific research platform project of Suzhou University (2021XJPT35), and the support project for outstanding young talents of colleges and universities in Anhui province (gxyq2022105).

References

Barbina, D. F., ElMasrya, G., Sun, D.-W., & Allen, P. (2011). *Detection of pork freshness using NIR hyperspectral imaging*. In *The 11th International Congress on Engineering and Food (ICEF11), Food Process Engineering in a Challenging World* (pp. 22-26). Paper No. NFP715.

Benes, E., Fodor, M., Kovács, S., & Gere, A. (2020). Application of detrended fluctuation analysis and yield stability index to evaluate near infrared spectra of green and roasted coffee samples. *Processes (Basel, Switzerland)*, 8(8), 913. http://dx.doi.org/10.3390/pr8080913.

El Orche, A., Elhamdaoui, O., Cheikh, A., Zoukeni, B., El Karbane, M., Mbarki, M., & Bouatia, M. (2022). Comparative study of three fingerprint analytical approaches based on spectroscopic sensors and chemometrics for the detection and quantification of argan oil adulteration. *Journal of the Science of Food and Agriculture*, 102(1), 95-104. http://dx.doi.org/10.1002/jsfa.11335. PMid:34032291.

Fodor, M., Mikola, E. E., Geösel, A., Stefanovits-Bányai, É., & Mednyánszky, Z. (2020). Application of near-infrared spectroscopy to investigate some endogenic properties of pleurotus ostreatus cultivars. *Sensors (Basel)*, 20(22), 6632. http://dx.doi.org/10.3390/s20226632. PMid:33228094.

Han, F., Aheto, J. H., Rashed, M. M. A., & Zhang, X. (2022a). Machine-learning assisted modelling of multiple elements for authenticating edible animal blood food. *Food Chemistry: X*, 14, 100280. http://dx.doi.org/10.1016/j.fochx.2022.100280. PMid:35284814.

Han, F., Huang, X., & Teye, E. (2019). Novel prediction of heavy metal residues in fish using a low-cost optical electronic tongue system based on colorimetric sensors array. *Journal of Food Process Engineering*, 42(2), e12983. http://dx.doi.org/10.1111/jfpe.12983.

Han, F., Huang, X., Aheto, J. H., Zhang, D., & Feng, F. (2020). Detection of beef adulterated with pork using a low-cost electronic nose based on colorimetric sensors. *Foods*, 9(2), 193. http://dx.doi.org/10.3390/foods9020193. PMid:32075051.

Han, F., Huang, X., Aheto, J. H., Zhang, X., & Rashed, M. M. A. (2022b). Fusion of a low-cost electronic nose and Fourier transform near-infrared spectroscopy for qualitative and quantitative detection of beef adulterated with duck. *Analytical Methods*, 14(4), 417-426. http://dx.doi.org/10.1039/D1AY01949J. PMid:35014996.

Han, F., Ming, L., Aheto, J. H., Rashed, M. M. A., Zhang, X., & Huang, X. (2022c). Authentication of duck blood tofu binary and ternary adulterated with cow and pig blood-based gel using Fourier transform near-infrared coupled with fast chemometrics. *Frontiers in Nutrition*, 9, 935099. http://dx.doi.org/10.3389/fnut.2022.935099. PMid:36386895.

Huang, G.-B., Wang, D. H., & Lan, Y. (2011). Extreme learning machines: a survey. *International Journal of Machine Learning and Cybernetics*, 2(2), 107-122. http://dx.doi.org/10.1007/s13042-011-0019-y.

Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, 70(1), 489-501. http://dx.doi.org/10.1016/j.neucom.2005.12.126.

Jiang, H., Jiang, X., Ru, Y., Wang, J., Xu, L., & Zhou, H. (2020). Application of hyperspectral imaging for detecting and visualizing leaf lard adulteration in minced pork. *Infrared Physics & Technology*, 110, 103467. http://dx.doi.org/10.1016/j.infrared.2020.103467.

Jiang, H., Wang, W., Zhuang, H., Yoon, S.-C., Yang, Y., & Zhao, X. (2019). Hyperspectral imaging for a rapid detection and visualization of duck meat adulteration in beef. *Food Analytical Methods*, 12(10), 2205-2215. http://dx.doi.org/10.1007/s12161-019-01577-6.

Kalinichenko, A., & Arseniyeva, L. (2020). Electronic nose combined with chemometric approaches to assess authenticity and adulteration of sausages by soy protein. *Sensors and Actuators. B, Chemical*, 303, 127250. http://dx.doi.org/10.1016/j.snb.2019.127250.

Kang, T. S., & Tanaka, T. (2018). Comparison of quantitative methods based on SYBR Green real-time qPCR to estimate pork meat adulteration in processed beef products. *Food Chemistry*, 269, 549-558. http://dx.doi.org/10.1016/j.foodchem.2018.06.141. PMid:30100472.

Koutsiaris, A. G. (2017). Deep tissue near infrared second derivative spectrophotometry for the assessment of claudication in peripheral arterial disease. *Clinical Hemorheology and Microcirculation*, 65(3), 275-284. http://dx.doi.org/10.3233/CH-16181. PMid:27983543.

Kumar, Y., & Chandrakant Karne, S. (2017). Spectral analysis: a rapid tool for species detection in meat products. *Trends in Food Science & Technology*, 62, 59-67. http://dx.doi.org/10.1016/j.tifs.2017.02.008.

Leng, T., Li, F., Xiong, L., Xiong, Q., Zhu, M., & Chen, Y. (2020). Quantitative detection of binary and ternary adulteration of minced beef meat with pork and duck meat by NIR combined with chemometrics. *Food Control*, 113, 107203. http://dx.doi.org/10.1016/j.foodcont.2020.107203.

Li, H., Liang, Y., Xu, Q., & Cao, D. (2009). Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Analytica Chimica Acta*, 648(1), 77-84. http://dx.doi.org/10.1016/j.aca.2009.06.046. PMid:19616692.

Li, X., Gao, X., & Guan, Y. (2019). A novel isothermal amplification method for detecting mouse source component in meat. *Journal of AOAC International*, 102(3), 872-877. http://dx.doi.org/10.5740/jaoacint.18-0325. PMid:30609945.

López-Maestresalas, A., Insausti, K., Jarén, C., Pérez-Roncal, C., Urrutia, O., Beriain, M. J., & Arazuri, S. (2019). Detection of minced lamb and beef fraud using NIR spectroscopy. *Food Control*, 98, 465-473. http://dx.doi.org/10.1016/j.foodcont.2018.12.003.

Lu, B., Han, F., Aheto, J. H., Rashed, M. M. A., & Pan, Z. (2021). Artificial bionic taste sensors coupled with chemometrics for rapid detection

of beef adulteration. *Food Science & Nutrition*, 9(9), 5220-5228. http://dx.doi.org/10.1002/fsn3.2494. PMid:34532030.

Ma, H. (2019). Identification of adulteration in commercial beef and its products by PCR. *China Food Safety Magazine*, 26, 76-77.

Mabood, F., Boqué, R., Alkindi, A. Y., Al-Harrasi, A., Al Amri, I. S., Boukra, S., Jabeen, F., Hussain, J., Abbas, G., Naureen, Z., Haq, Q. M. I., Shah, H. H., Khan, A., Khalaf, S. K., & Kadim, I. (2020). Fast detection and quantification of pork meat in other meats by reflectance FT-NIR spectroscopy and multivariate analysis. *Meat Science*, 163, 108084. http://dx.doi.org/10.1016/j.meatsci.2020.108084. PMid:32062524.

Mahanti, N. K., Chakraborty, S. K., Kotwaliwale, N., & Vishwakarma, A. K. (2020). Chemometric strategies for nondestructive and rapid assessment of nitrate content in harvested spinach using Vis-NIR spectroscopy. *Journal of Food Science*, 85(10), 3653-3662. http://dx.doi.org/10.1111/1750-3841.15420. PMid:32888324.

Mendez, J., Mendoza, L., Cruz-Tirado, J. P., Quevedo, R., & Siche, R. (2019). Trends in application of NIR and hyperspectral imaging for food authentication. *Scientia Agropecuaria*, 10(1), 143-161. http://dx.doi.org/10.17268/sci.agropecu.2019.01.16.

Parastar, H., van Kollenburg, G., Weesepoel, Y., van den Doel, A., Buydens, L., & Jansen, J. (2020). Integration of handheld NIR and machine learning to "Measure & Monitor" chicken meat authenticity. *Food Control*, 112, 107149. http://dx.doi.org/10.1016/j.foodcont.2020.107149.

Rady, A., & Adedeji, A. A. (2020). Application of hyperspectral imaging and machine learning methods to detect and quantify adulterants in minced meats. *Food Analytical Methods*, 13(4), 970-981. http://dx.doi.org/10.1007/s12161-020-01719-1.

Sarno, R., Triyana, K., Sabilla, S. I., Wijaya, D. R., Sunaryono, D., & Fatichah, C. (2020). Detecting pork adulteration in beef for halal authentication using an optimized electronic nose system. *IEEE Access: Practical Innovations, Open Solutions*, 8, 221700-221711. http://dx.doi.org/10.1109/ACCESS.2020.3043394.

Schönbichler, S. A., Bittner, L., Pallua, J., Popp, M., Abel, G., Bonn, G., & Huck, C. (2013). Simultaneous quantification of verbenalin and verbascoside in Verbena officinalis by ATR-IR and NIR spectroscopy. *Journal of Pharmaceutical and Biomedical Analysis*, 84, 97-102. http://dx.doi.org/10.1016/j.jpba.2013.04.038. PMid:23810849.

Silva, L. C. R., Folli, G. S., Santos, L. P., Barros, I. H. A. S., Oliveira, B. G., Borghi, F. T., Santos, F. D., Filgueiras, P. R., & Romão, W. (2020). Quantification of beef, pork, and chicken in ground meat using a portable NIR spectrometer. *Vibrational Spectroscopy*, 111, 103158. http://dx.doi.org/10.1016/j.vibspec.2020.103158.

Song, W., Yun, Y., Wang, H., Hou, Z., & Wang, Z. (2021). Smartphone detection of minced beef adulteration. *Microchemical Journal*, 164, 106088. http://dx.doi.org/10.1016/j.microc.2021.106088.

Suthaharan, S. (2016). Support vector machine. In S. Suthaharan (Ed.), *Machine learning models and algorithms for big data classification: thinking with examples for effective learning* (pp. 207-235). Boston, MA: Springer US. http://dx.doi.org/10.1007/978-1-4899-7641-3_9.

Teye, E., Huang, X., Lei, W., & Dai, H. (2014). Feasibility study on the use of Fourier transform near-infrared spectroscopy together with chemometrics to discriminate and quantify adulteration in cocoa beans. *Food Research International*, 55, 288-293. http://dx.doi.org/10.1016/j.foodres.2013.11.021.

Tian, X., Wang, J., Shen, R., Ma, Z., & Li, M. (2019). Discrimination of pork/chicken adulteration in minced mutton by electronic taste system. *International Journal of Food Science & Technology*, 54(3), 670-678. http://dx.doi.org/10.1111/ijfs.13977.

Wang, J., Wu, X., Zheng, J., & Wu, B. (2022a). Rapid identification of green tea varieties based on FT-NIR spectroscopy and LDA/QR. *Food Science and Technology (Campinas)*, 42, e73022. http://dx.doi.org/10.1590/fst.73022.

Wang, Q., Li, L., Ding, W., Zhang, D., Wang, J., Reed, K., & Zhang, B. (2019). Adulterant identification in mutton by electronic nose and gas chromatography-mass spectrometer. *Food Control*, 98, 431-438. http://dx.doi.org/10.1016/j.foodcont.2018.11.038.

Wang, S., Liu, P., Feng, L., Teng, J., Ye, F., & Gui, A. (2022b). Rapid determination of tea polyphenols content in Qingzhuan tea based on near infrared spectroscopy in conjunction with three different PLS algorithms. *Food Science and Technology (Campinas)*, 42, e94322.

Wang, X., Xing, X., Zhao, M., & Yang, J. (2020). Comparison of multispectral modeling of physiochemical attributes of greengage: Brix and pH values. *Food Science and Technology (Campinas)*, 41(Suppl 2), 611-618. http://dx.doi.org/10.1590/fst.21320.

Wiedemair, V., De Biasio, M., Leitner, R., Balthasar, D., & Huck, C. (2018). Application of design of experiment for detection of meat fraud with a portable near-infrared spectrometer. *Current Analytical Chemistry*, 14(1). http://dx.doi.org/10.2174/1573411013666170207121113.

Yu, L. (2015). *MATLAB intelligent algorithms 30 cases analysis*: Beijing, China: Beijing University of Aeronautics and Astronautics Press.

Zaukuu, J. Z., Gillay, Z., & Kovacs, Z. (2021). Standardized extraction techniques for meat analysis with the electronic tongue: a case study of poultry and red meat adulteration. *Sensors (Basel)*, 21(2), 481. http://dx.doi.org/10.3390/s21020481. PMid:33445458.

Zhang, L., Li, G., Sun, M., Li, H., Wang, Z., Li, Y., & Lin, L. (2017). Kennard-Stone combined with least square support vector machine method for noncontact discriminating human blood species. *Infrared Physics & Technology*, 86, 116-119. http://dx.doi.org/10.1016/j.infrared.2017.08.020.

Zheng, X., Li, Y., Wei, W., & Peng, Y. (2019). Detection of adulteration with duck meat in minced lamb meat by using visible near-infrared hyperspectral imaging. *Meat Science*, 149, 55-62. http://dx.doi.org/10.1016/j.meatsci.2018.11.005. PMid:30463040.

Zontov, Y. V., Rodionova, O. Y., Kucheryavskiy, S. V., & Pomerantsev, A. L. (2020). PLS-DA – A MATLAB GUI tool for hard and soft approaches to partial least squares discriminant analysis. *Chemometrics and Intelligent Laboratory Systems*, 203, 104064. http://dx.doi.org/10.1016/j.chemolab.2020.104064.