

**Review:****The Chemometrics Techniques in Combination with Instrumental Analytical Methods Applied in Halal Authentication Analysis****Abdul Rohman<sup>1,2,\*</sup> and Anggita Rosiana Putri<sup>1</sup>**<sup>1</sup>Faculty of Pharmacy, Universitas Gadjah Mada, Sekip Utara, Yogyakarta, 55281, Indonesia<sup>2</sup>Research Center of Halal Products, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia**\* Corresponding author:**

tel: +62-87838445216

email: [abdul\\_kimfar@ugm.ac.id](mailto:abdul_kimfar@ugm.ac.id)

Received: October 3, 2017

Accepted: March 29, 2018

DOI: 10.22146/ijc.28721

**Abstract:** Halal food is taken into account as any food permitted to be consumed by Muslim according to Syariah law. Due to the development of science and technology in which some new food components such as food additives have been synthesized and produced, some industries used non-halal components such as pig derivatives in food products to reduce the production cost. Non-halal components added in food products are difficult to detect visually due to the close similarity between non-halal ingredients and components present in food. As a consequence, some scientists developed and proposed some instrumental techniques like spectroscopy, chromatography and molecular biology-based methods for identification of non-halal components. Food matrix is very complex to be analyzed. Therefore, the signals obtained during chemical and biological analyses are very complex which are difficult to interpret. Fortunately, a statistical technique called with chemometrics can be used an alternative method to handle the complex data met during analysis of non-halal components. Chemometrics has been widely used in many aspects of analysis in many types of the sector. In this review, some chemometrics techniques used to treat responses obtained from instrumental measurements intended for analysis of non-halal components in food matrix were highlighted.

**Keywords:** chemometrics; non-halal component; authentication; classification; multivariate calibration

**■ INTRODUCTION**

Halal food analysis is intended to identify the presence of non-halal components present in food. Halal food is Islamic term used to describe any foods and drinks permissible to be consumed by Muslim according to *Shariah* law [1]. Halal food must not contain any components or derivatives coming from human, non-halal animals such as pig and dog, animals which are not slaughtered according to Islamic means and *najs* (dirty) components. Besides, halal food must be safe for consumption and is not prepared, processed, and manufactured using equipment contaminated with *najs*. Among non-halal components present in food products, pig derivatives such as pork, lard, and porcine gelatin are frequently meet [2].

The presence of non-halal components in food products can be either low or high concentrations. Low levels of pig derivatives become a problem. Therefore sensitive techniques are required to detect them. Some analytical techniques are proposed and developed by halal scientist to analyse pig derivatives including infrared spectroscopy [3], nuclear magnetic resonance spectroscopy [4], chromatographic-based methods [5-6], electronic nose [7], differential scanning calorimetry [8], real-time polymerase chain reaction [9], and enzyme-linked immunosorbent assay (ELISA) [10-11]. Some of these methods result in complex signals which are difficult to interpret. Fortunately, a tools knowns as chemometrics can handle these difficulties.

In this review, some chemometrics techniques

used for pattern recognition including principal component analysis, cluster analysis, a soft independent model of class analogy and discriminant analysis have been explored for assisting chemical data treatment obtained from numerous instrumental techniques intended for identification of non-halal components in food. In addition, multivariate calibrations were exploited for developing the prediction models in the quantitative analysis of non-halal components.

## ■ CHEMOMETRICS

Chemometrics means performing statistical and mathematical calculations on measurements of chemical data. The measurement covered anything from measuring of pH of hydrogen ion activity to computing a Fourier transform interpolation of a spectrum [12-13]. The International Chemometrics Society defines chemometrics as the science of relating measurements made on a chemical system or process to the state of the system via application of mathematical or statistical methods [14].

The chemometrics enables food scientist to solve the calibration problems for the analysis of spectral data. Some chemometrics methods are being simple to understand, but others are very complex to understand and implement [15]. Among chemometrics techniques widely used in food science including halal food analysis are experimental design, data processing such as smoothing and derivatization, explanatory analysis like factor analysis and principal component analysis, multivariate calibration like partial least square, multivariate curve resolution analysis, artificial intelligence (neural networks and genetic algorithms), stepwise multiple linear regression, and principal component regression [16-17].

### Multivariate Calibration

Calibration can be understood as the mathematical relationship between the predictor and response variables. Based on the variable number, calibration can be grouped into two main categories, namely (1) univariate calibration in which one variable is measured, and (2) multivariate calibration which uses several variables such as applying absorbance values at some

wavelengths in spectroscopic measurement. Chemometrics related to multivariate calibration [18-19]. The main advantage of multivariate calibration over univariate calibration is the reliability of prediction results for unknown samples obtained. The chemometrics of multivariate calibration is intended to develop a model which correlates the information in the set of known measurements to the desired property with numerous variables assessed. The multivariate calibration regression is extensively used in making decisions relating to product quality during online monitoring and quality control [20].

Various multivariate calibrations existed, namely classical least square (CLS), stepwise multiple linear regression (SMLR), principal component regression (PCR), partial least square (PLS), artificial neural network (ANN), and genetic algorithm (GA). Among them, SMLR, PCR, and PCR are widely used for quantifying non-halal components in food products. Some statistical parameters are used during performing multivariate calibrations, namely coefficient of determination ( $R^2$ ) for the relationship between two variables, root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP), and predicted residual error of sum square (PRESS) [17,20].

### Multiple Linear Regression (MLR)

MLR is one of the inverse calibration models in which concentration (y-axis) is modeled toward responses (x-axis). MLR involves finding regression equations in the form:

$$C_t = b_{0t} + b_{1t}A_1 + b_{2t}A_2 + \dots + b_{6t}A_6 \quad (1)$$

$C_t$  is a concentration of analyte target,  $b_0, b_1, \dots, b_6$  are coefficients, and  $A$  is predictor variables (for example absorbances). To carry out MLR, the number of calibration specimens must be greater than the number of predictors (variables) [17]. There are three methods for MLR, namely (1) forward selection, (2) backward selection, and (3) stepwise selection [21].

The forward selection starts with an empty equation with constant term only, without any predictor variables. The variables are added to the test equation

one at a time as the procedure progresses. The next variable is added, and the new equation is observed. If the second variables offer the better equation than the first variable, the second variable is used and the second variable is removed if the new equation is worse with that with one variable. The procedure proceeds in this manner until no further variate has a significant effect on the fitted equation [17]. The backward selection regression is identical with those obtained from the forward regression analysis. But, the equation is started using all variables. Then, each variable is removed from the last variable, and the new model is evaluated. The process is continued till all variables are evaluated, and variables offering the significant effect is included in the model, and that which are not significant is removed from the model. The stepwise regression is modified forward selection and is often available in the commercial statistical software. As with forwarding selection, the procedure increases the number of variables in the equation at each step, but at each stage, the possibility of deleting a previously included variable is considered [21]. It is very difficult to generalize the superiority of one technique over another technique of MLR because the relative performance of the MLR methods is frequently influenced by data set used [22].

SMLR has been used as calibration technique for quantifying of lard as an adulterant in VCO using predictor variables of the temperature peaks,  $T_o$ ,  $T_e$  and the melting and cooling range ( $Tr$ ) of thermal profiles using differential scanning calorimetry (DSC). The regression models obtained are as follows:

$$\text{Step (1) \% lard} = 295.6 - 11.72 \times T_e A \quad (2)$$

with  $R^2$  adjusted 0.9467

$$\text{Step (2) \% lard} = 293.1 - 11.36 \times T_e A - 2.17 \times Tr D \quad (3)$$

with  $R^2$  adjusted 0.9582

$T_e A$  is variable of endset peak A, while  $Tr D$  is variable of the temperature range for peak D [23]. SMLR is successfully used for modeling between the actual value of lard and predicted the value of lard using predictor variables of  $T_e A$  and  $Tr D$ . The SMLR equation for quantification of lard as an adulterant in butter is also reported by Nurrulhidayah et al. [4] using thermal profiles of DSC. The equation of SMLR obtained was:

$$\% \text{ lard} = 293.1 - 11.36 \times T_e A - 2.17 \times Tr D \quad (4)$$

where  $T_e A$  is the endset of peak A and  $Tr D$  is the range of thermal transition for peak D. These parameters can serve as a good measurement parameter in detecting lard adulteration in butter. SMLR using DSC parameter based on heating thermogram is successfully used for prediction of lard in canola oil with model prediction equation of:

$$\% \text{ lard} = 8.2792A + 9.2597 \quad (5)$$

with  $R^2$  adjusted of 0.9777 and  $P < 0.0002$ . A is peak area of the heating thermal profile of lard [24].

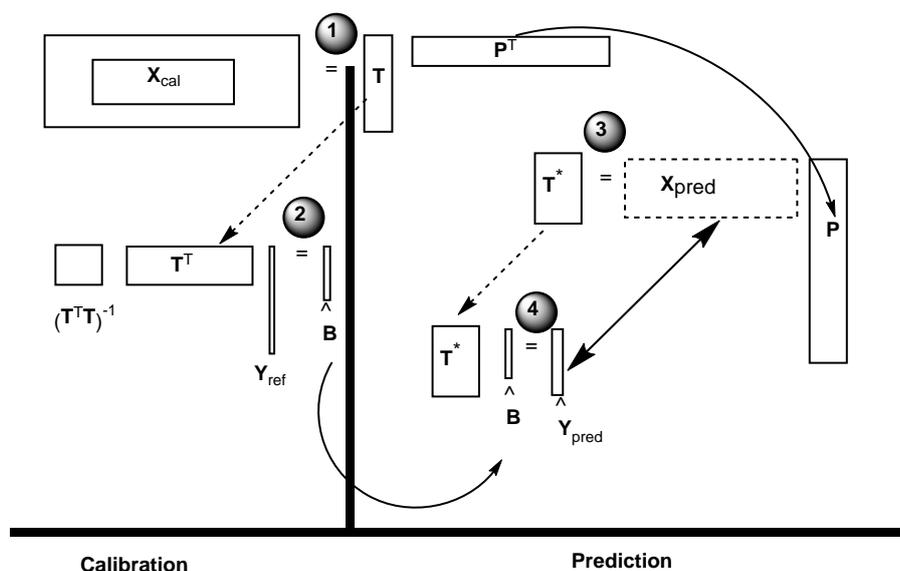
Lard as an adulterant in Argentinean milk fats has been quantified using SMLR with 16 fatty acid compositions as predictor variables. Three SMLR techniques (backward, stepwise and forward) were used and compared [25]. In calibration model, among three MLR techniques, backward selection offered the best calibration model for correlation between the actual value of lard and predicted values using fatty acid profiles with  $R^2$  adjusted of 0.943 over stepwise ( $R^2$  adjusted 0.915) and forward ( $R^2$  adjusted 0.916). However, when two calibration models selected were applied during the prediction set ( $m = 260$ ), the lowest RMSEP value (1.77%) was obtained using stepwise model. Therefore, stepwise was chosen for quantifying lard.

### Principal Component Regression (PCR) and Partial Least Square (PLS)

Principal component regression (PCR) is one of inverse calibration in which the concentration (x-axis) is used as a predictor, while response such as absorbance at several wavelengths is located in the y-axis. PCR performs multiple inverse calibrations of predictor variables against the scores (known as principal components) rather than original variables [17,19]. PCR uses regression to converts scores or principal components (PCs) into concentration. PCs are obtained during principal component analysis (PCA).

PCA is data reduction technique which reduces the data set of matrix  $X$  ( $K \times N$ ) into much smaller variables  $A$  (known as PC or factor). The mathematical model is:

$$X = TP^T + E \quad (6)$$



**Fig 1.** Schema for the calibration of a PCR model and the prediction of new samples. Steps: (1) Decomposition of calibration data matrix in scores and loadings matrix. (2) Calculation of regression vector. (3) Projection of prediction spectra into calibration space. (4) Prediction of the property of the new samples. Adapted from Ref. [12]

$X$  is a matrix containing  $A$  scores for PCs ( $N \times A$ ),  $P$  is a matrix containing  $A$  loadings for PCs ( $K \times A$ ), and  $E$  is a residual matrix of the model ( $K \times N$ ).

The next step is to resolve matrix  $X$  into its PCs as below:

$$X = TP^T + E = \sum_{a=1}^A t_a p_a^T + E \quad (7)$$

$X$  can be represented by matrix scores ( $T$ )

$$T = XP \quad (8)$$

Matrix scores ( $T$ ) and loading matrix ( $P$ ) are obtained from matrix  $X$ . Matrix  $Y$  is calculated by regressing  $Y$  against  $T$ .

$$Y = TB + E \quad (9)$$

$B$  is regression matrix and is calculated using least square regression provided the matrix value  $Y$  is known. The calibration and prediction model using PCR is depicted in Fig. 1.

Partial least square regression (PLSR) is introduced by Wold in 1975. This calibration technique is the most reported multivariate calibration for quantitative analysis of non-halal components. PLSR is a way to solve problems involving high collinearity or the need to calculate correlated  $Y$  variables [26]. Since its original formulation, PLSR has been associated with other mathematical methods and algorithms. The algorithms most widely used to implement PLSR are non-linear iterative partial

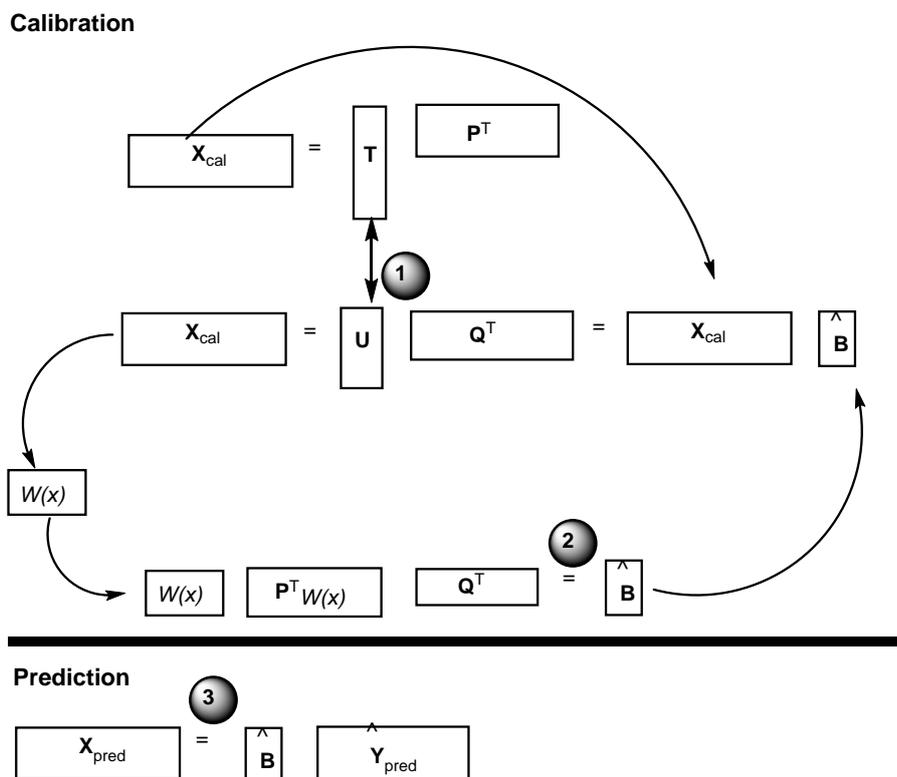
least squares (NIPALS) and SIMPLS. The variables of concentration ( $y$ -axis) are related to the predictor variables ( $x$ -axis) through auxiliary variables known as latent variables (factors or components), which are linear combinations of the variables  $x_1, x_2, \dots, x_K$ . These components are highly similar to the PCs calculated by PCA and used for PCR. As in PCA, matrix  $X$  and  $Y$  are subjected to auto-scaling.  $X$  is used for extracting few latent variables ( $a = 1, 2, \dots, A$ ). Therefore, each matrix is resolved into the combination of  $A$  factors ( $A \leq K$ ), allowing to calculate:

$$X = TP^T + E = \sum_{a=1}^A t_a p_a^T + E \quad (10)$$

$$Y = UQ^T + F = \sum_{a=1}^A u_a q_a^T + F \quad (11)$$

In which matrix  $T$  ( $M \times A$ ) and  $U$  ( $M \times A$ ) are score matrices for blocks  $X$  and  $Y$ , respectively with  $M$  samples,  $A$  factors,  $K$  variables and  $P$  analytes.  $PT$  ( $A \times K$ ) and  $Q^T$  ( $A \times P$ ) are the loading matrices for blocks  $X$  and  $Y$ , while  $E$  and  $F$  are the residual matrices for  $X$  and  $Y$ , respectively [26-27]. Schematically, the calibration and prediction model using PLSR is depicted in Fig. 2.

The calibration process is initiated by computing loading weights  $W(X)$ , extracted from variables in  $X$  matrix, and the desired the number of latent variables is stored in score matrix  $T$  which used for modeling the



**Fig 2.** Schema for the calibration of a PLSR model and the prediction of new samples. Steps: (1) Simultaneous decomposition of  $X$  and  $Y$  matrix and calculation of weights matrix. (2) Calculation of regression matrix. (3) Prediction of new samples. Adapted from Ref. [12]

variables in matrices of  $X$  and  $Y$ . In addition, variables in  $Y$  can be modeled from variables in  $X$  using regression coefficient matrix  $B$ . The coefficients in  $B$  can be estimated as a function of loadings  $X$  and  $Y$  as well as  $P$  and  $Q$  [19].

$$\hat{B} = W(P^T W)^{-1} Q^T \quad (12)$$

For calibration, the regression matrix  $\hat{B}$  which allow samples to be predicted without any resolving  $B$  into scores and loadings is calculated. For example, if the spectrum of certain analytes is defined by vector  $x_i$ , the concentrations of the analytes  $y$  can be calculated from:

$$\hat{y}_i = x_i^T \hat{B} \quad (13)$$

PLSR finds the directions of greatest variability by comparing both spectral and target property information with the new axes, called PLSR components or PLSR factors. The principal difference between PCR and PLSR is that the first PC or factor in PCR represents the highest variations in the responses, whereas the first  $PC_1$  in PLSR it represents the most relevant variations showing the best

correlation with the target property values [17]. The optimum number of factors used to build PLSR and PCR models depend on analytical target and properties of samples [28]. PCR and PLSR can be used in specific responses or the whole responses, for example, all absorbance of FTIR spectra. Therefore they allow more information to be included in PCR and PLS calibration models [29]. Our group and other researchers typically optimize PCR and PLSR as calibration technique for analysis of a non-halal component in food products. In general, PLSR was better than PCR for predicting non-halal components, so that PLS was finally used for calibration model.

PLSR was used for quantification of lard as an adulterant in CLO using absorbance at frequency region of  $1500\text{--}1030\text{ cm}^{-1}$ . The relationship between the actual value of lard ( $x$ -axis) and FTIR predicted value ( $y$ -axis) with equation of  $y = 0.872x - 0.392$  [30]. PLSR is also used as calibration technique for quantification of lard

in the mixture with other animal fats of chicken fat, beef fat, and mutton fat. As variables, the absorbance of FTIR spectra at 1500–900  $\text{cm}^{-1}$  are exploited for quantification of lard in the mixture with animal fats. The equations obtained are  $y = 0.995x + 0.098$ ,  $y = 0.999x + 0.016$  and  $1.002x + 0.034$  for lard in mutton, beef and chicken, respectively with RMSEC of 0.98, 0.61 and 0.73% [31]. PLSR using absorbance at 1500–1000  $\text{cm}^{-1}$  either with normal or first derivative spectra were suitable for quantification of lard in selected vegetable oils [32].

PLSR was used to develop a calibration model for analysis of pork in beef meatball using absorbance at frequency region of 1200–1000  $\text{cm}^{-1}$ . The equation obtained for the relationship between actual PF value and FTIR predicted values was  $y = 0.999x + 0.004$ , with  $R^2$  of 0.999 and RMSEC of 0.442%. The PLS calibration model using 4 principal components obtained RMSEP of 0.742% [5]. Lard in meatball broth was also quantitatively determined using PLSR at frequency region of 1018–1284  $\text{cm}^{-1}$ . The values of  $R^2$  and RMSEC obtained were 0.9975 and 1.34% (v/v), respectively [33].

### Classification Analysis

To assign to predefined categories to samples and predicting an unknown sample as belonging to one of several distinct groups, chemometrics of classification is used. A classification model is used to predict a sample's class based on closest examples [17]. Two types of classification are introduced in chemometrics, namely (1) unsupervised pattern recognition such as principal component analysis (PCA) and cluster analysis, and (2) supervised pattern recognition technique such as discriminant analysis and K-nearest neighbor (k-NN). These classification techniques are primarily used in halal food analysis.

### Principal Component Analysis

Principal component analysis (PCA) is probably the most popular chemometrics technique, and it is used by almost all scientific disciplines. PCA is also likely to be the oldest multivariate analysis technique [34]. PCA is an important part of chemometrics and provides the most compact representation of all the variation in a data table. Exploratory algorithms such as principal component

analysis (PCA) are designed to reduce large complex data sets into a series of optimized and interpretable size [15]. PCA is trying to find out factors or principal components  $PC_1, PC_2, \dots, PC_n$ , which are linear combinations of the original variables describing each object,  $X_1, X_2, \dots, X_n$ , i.e., If there are 14 variables, there would be 14 principal components ( $PC_1, PC_2, \dots, PC_{14}$ ). Each PC would be a linear combination of  $X_1, X_2, \dots, X_{14}$ . In this approach, the new variables (PC) may seem a pointless exercise since we obtain  $n$  new variables in place of the  $n$  original ones, and hence no reduction in the amount of data. However, the principal components are also chosen so that the first principal component ( $PC_1$ ) corresponds to the most variation in the data set, the second principal components ( $PC_2$ ), accounts for the next largest variation and so on. Therefore, when significant correlation takes place, the number of significant PCs is much less than the number of original variables [17].

There are 4 parts of PCA, namely: data, score, loading, and residuals. The scores of PCs are sometimes called as hidden or latent variables. Objects or samples which have same scores of PCs can be understood as the same object, and the object that has similar scores of  $PC_2$  are similar. Thus, PCA can be used for classification among objects [35]. Che Man et al. [36] have used PCA for classification of lard and other 16 edible fats and oils using absorbance of FTIR spectra at 16 different wavenumbers as variables. Using 4 principal components ( $PC_1, PC_2, PC_3$ , and  $PC_4$ ), an about 90% variation are described.  $PC_1$  accounted for 44.1% of the variation, while  $PC_2$  described 30.2% of the variation. PCA can classify lard and other edible fats and oils using the plot of  $PC_1$  and  $PC_2$ . Extra virgin olive oil, chicken fat and rice bran oil can be classified as one group due to close similarity of PCs scores. From loading plot of PCA, absorbance at wavenumbers of 2853, 2922 and 1465  $\text{cm}^{-1}$  are more responsible for classification. PCA is used for selecting animals fats having similar to lard to be used as an animal fat model to be adulterated with lard. Using HPLC chromatogram as variables, chicken fat is chosen as animal fat to be adulterated due to its similarity in score two first PCs to lard [37]. PCA was also successfully

**Table 1.** Application of chemometrics techniques in combination with instrumental analysis for halal authentication

Issues of halal authentication	Chemometrics used	Instrumental methods	Ref.
Classification of lard from other animal fats and vegetable oils	PCA, CA	FT-MIR spectroscopy	[36]
Differentiation of lard from other animal fats	PCA	HPLC-refractive index detector	[37]
Classification of pure ghee and that adulterated with lard	PCA	FT-MIR spectroscopy	[48]
Differentiation between porcine gelatin and bovine gelatin	PCA, DA	FT-MIR spectroscopy	[38]
Classification of porcine and bovine gelatin	PCA	Electrophoresis	[39-40]
Differentiation between porcine gelatin and bovine gelatin in food and capsule shells	PCA	HPLC-fluorescence detector	[41-43]
Classification of porcine gelatin, bovine gelatin, and fish gelatin	CA	FT-MIR spectroscopy	[45]
Classification of lard from other animal fats	CA	FT-Raman	[46]
Analysis of lard as adulterant in palm oil	SIMCA	FT-NIR	[49]
Discrimination between lard and other animal fats and vegetable oils	DA	FT-MIR spectroscopy	[32]
Authentication of cod liver oil from lard	DA	FT-MIR spectroscopy	[30]
Classification of virgin coconut oil (VCO) and VCO added with lard	DA	FT-MIR spectroscopy	[7]
Classification of biscuit with lard and without lard	DA	FT-MIR spectroscopy	[47]

used for classification between lard in pure ghee using FTIR spectra absorbances at optimized wavenumbers of 3030–2785, 1786–1680, and 1490–919  $\text{cm}^{-1}$  with 5% of lard in pure ghee could be detected [48].  $\text{PC}_1$  and  $\text{PC}_2$  contributed to 92 and 8% variances, respectively.

Porcine gelatin (non-halal) and bovine gelatin (halal) are also classified using PCA using variables from several analytical signals. Hashim et al. [38] classified bovine and porcine gelatins using FTIR spectra absorbance at 3290–3280 and 1660–1200  $\text{cm}^{-1}$  based on a second derivative study. PCA using these variables has been used to identify gelatin source for applications in the food and pharmaceutical industries. Classification of bovine and porcine gelatins in processed food products is carried out using Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) and PCA. As variables used are molecular weight. Based on  $\text{PC}_1$  and  $\text{PC}_2$ , two groups of processed foods containing porcine gelatin and bovine gelatin were classified.  $\text{PC}_1$  and  $\text{PC}_2$  accounted for 46 and 29% of the variation, respectively; thus, 78% of the variance was accounted for the first 2 PCs [39]. PCA was also used for authentication

of bovine porcine from porcine gelatin using molecular weight as variables as determined SDS-PAGE. The results showed that PCA that 5% porcine gelatin in bovine gelatin could be being detected. The PCA provided good differentiation of samples with 71% of the variation accounted by the first two PCs [40].

PCA intended for classification of porcine gelatin and bovine gelatin from several producers using amino acid composition as determined by HPLC with fluorescence detector as variables are carried out by Nemati et al. [41]. Using  $\text{PC}_1$  and  $\text{PC}_2$ , gelatin sources can be classified and differentiated successfully. The amino acid composition was also used as variables in PCA for classification of bovine and porcine in laboratory prepared gummy [42], and laboratory prepared capsule shells [43].

### Cluster Analysis

Cluster analysis is one type of unsupervised pattern recognition technique frequently used for classification of objects [36]. Cluster analysis (CA) is a method for dividing a group of objects into classes so that similar

objects fall in the same class. As in other unsupervised pattern recognition techniques, the object groups are not known previously to the mathematical analysis and no assumptions are made about the distribution of the variables [17].

CA searches for objects which are close together in the variable space. The distance,  $d$ , between two points in  $n$ -dimensional space with coordinates  $(x_1, x_2, x_3, \dots, x_n)$  and  $(y_1, y_2, y_3, \dots, y_n)$  is usually taken as the Euclidian distance defined as:

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (14)$$

In CA, the similar objects (having the similar Mahalanobis distance) will be clustered together. The main advantage of CA over PCA is that CA can provide numerical values of similarity among the objects evaluated. Therefore, the information obtained is more objective. In addition, CA can reduce dimensionality while retaining the required information [44].

CA is used for classification of lard and other animal fats and vegetable oils based on Mahalanobis distance using absorbance of FTIR spectra at 16 wavenumbers as variables. CA is capable of clustering lard and others as represented by dendrogram illustrating the linkage steps of analyzed fats and oils. Based on dendrogram, Group 1 contained lard, cod liver oil, corn, soybean, rice bran, sesame, sunflower, extra virgin olive, pumpkin seed, walnut, and canola oils. Group 2 comprises chicken fat and palm oil. Furthermore, beef fat, mutton fat, grapeseed oil, and virgin coconut oil can be considered as group 3, 4, 5, and 6, respectively [36].

In order to classify porcine gelatin, bovine gelatin, and fish gelatin as well as their mixture, CA using FTIR spectra at optimized specific frequency regions was used. The spectra within frequency regions of 1722–1487  $\text{cm}^{-1}$  are treated with first derivatization and vector normalization as pre-processing of data. CA can classify these gelatins [45]. CA was also used for classification of lard in binary mixture with other animal fats including beef tallow, chicken fat, duck oil using Raman intensities at wavenumbers of 2000–700  $\text{cm}^{-1}$  [46]. Some independent modelling of class analogy (SIMCA), chemometrics used for unsupervised pattern classification, has been used by

Basri et al. [49] for authentication of palm oil from lard by classifying between pure palm oil and palm oil adulterated with lard at concentrations of 0.5, 1, 3, 5, 10, 15, 20, 30, 40, 50% of lard. SIMCA could successfully classify between palm oil and that adulterated using absorbances of near spectra with accuracy levels of 99%.

### Discriminant Analysis

Unlike PCA and cluster analysis, discriminant analysis (DA) is one of the supervised pattern recognition technique. The starting point of linear discriminant analysis (LDA) is to find a linear discriminant function (LDF),  $Y$ , which is a linear combination of the original measured variables:

$$Y = a_1X_1 + a_2X_2 + \dots + a_nX_n \quad (15)$$

The original  $n$  measurements for each object are combined into a single value of  $Y$ , so the data have been reduced from  $n$  dimensions to one dimension. The coefficients of the terms are chosen in such a way that  $Y$  reflects the difference between groups as much as possible: objects in the same group will have similar values of  $Y$  and objects in different groups will have very different values of  $Y$ . Thus the LDF provides a means of discriminating between the two groups [17].

Discriminant analysis (DA) is carried out for classification of lard as an adulterant in high value edible oil of cod liver oil. As predictor variables, absorbance of FTIR spectra at 1500–1030  $\text{cm}^{-1}$  are used. DA can classify the samples as lard as an adulterant in CLO and pure CLO [30]. DA is also successfully used for classification of lard and lard mixed with other animal fats namely chicken fat, mutton fat and beef fat using variables of absorbance of FTIR spectra at 3300–650  $\text{cm}^{-1}$ . DA using variables of absorbance at 1500–1000  $\text{cm}^{-1}$  and visualized as Mahalanobis distance can classify lard as adulterant mixed with selected vegetable oils (canola oil, corn oil, extra virgin olive oil, soybean oil, and sunflower oil) and pure vegetable oils [32]. Mansor et al. [7] used DA for classification of lard as an adulterant in virgin coconut oil (VCO) using FTIR spectroscopic technique at combined wavenumbers of 3020–3000 and 1120–1000  $\text{cm}^{-1}$ .

In order to classify among lard and other animal fats as well as vegetable oils in biscuit, DA at combined frequency regions of 3050–2800, 1800–1600, and 1500–650  $\text{cm}^{-1}$  was used. The Coomans plot revealed that all vegetable fats/oils and animal fats, including lard, are clustered in a distinct group. DA was also successfully employed to assign real food samples, in particular biscuit, into one of two groups (halal and non-halal groups). DA is a very useful means for Halal screening technique to enhance the Halal authentication process [47].

Hashim et al. [38] used DA for classification of porcine and bovine gelatins using frequency regions of 3290–3280 and 1660–1200  $\text{cm}^{-1}$ . These frequency regions were selected due to its capability to provide valuable information about the origin of the gelatin. The Cooman's plot proved to be useful for visualizing DA for classification. DA using Cooman's plot can provide a rapid determination of gelatin source for application in food, pharmaceutical, and photographic applications.

## ■ CONCLUSION

The development of chemometrics software coupled with selected instrument techniques such as spectrophotometer and chromatograph allow halal scientist to identify and to quantify non-halal components in food and pharmaceutical products. The multivariate calibration of partial least square is the most reported calibration technique during quantification of lard and pork in the products. Meanwhile, principal component analysis, cluster analysis, and discriminant analysis are commonly used to make classification between halal products and non-halal products.

## ■ ACKNOWLEDGMENTS

I acknowledged to the Ministry of Research and Higher Education, Republic of Indonesia for financial assistance in conducting research activities in the halal analysis which allow me to prepare this review manuscript.

## ■ REFERENCES

- [1] Mursyidi, A., 2013, The role of analytical chemistry in Halal certification, *J. Food Pharm. Sci.*, 1 (1), 1–4.
- [2] Rohman, A., and Che Man, Y.B., 2012, Analysis of pig derivatives for halal authentication studies, *Food Rev. Int.*, 28 (1), 97–112.
- [3] Guntarti, A., Martono, S., Yuswanto, A., and Rohman, A., 2015, FTIR spectroscopy in combination with chemometrics for analysis of wild boar meat in meatball formulation, *Asian J. Biochem.*, 10 (4), 165–172.
- [4] Nurrulhidayah, A.F., Che Man, Y.B., Rohman, A., Rosman, A.S., Ismail, A., Mustafa, S., and Khatib, A., 2015, Detection of butter adulteration with lard by employing  $^1\text{H-NMR}$  spectroscopy and multivariate data analysis, *J. Oleo Sci.*, 64 (7), 697–703.
- [5] Indrasti, D., Che Man, Y.B., Mustafa, S., and Hashim, D.M., 2010, Lard detection based on fatty acids profile using comprehensive gas chromatography hyphenated with time-of-flight mass spectrometry, *Food Chem.*, 122 (4), 1273–1277.
- [6] Rohman, A., Triyana, K., Sismindari, and Erwanto, Y., 2012, Differentiation of lard and other animal fats based on triacylglycerols composition and principal component analysis, *Int. Food Res. J.*, 19 (2), 475–479.
- [7] Mansor, T.S.T., Che Man, Y.B., and Rohman, A., 2011, Application of fast gas chromatography and Fourier transform infrared spectroscopy for analysis of lard adulteration in virgin coconut oil, *Food Anal. Methods*, 4 (3), 365–372.
- [8] Nurrulhidayah, A.F., Arief, S.R., Rohman, A., Amin, I., Shuhaimi, M., and Khatib, A., 2015, Detection of butter adulteration with lard using differential scanning calorimetry, *Int. Food Res. J.*, 22 (2), 832–839.
- [9] Sudjadi, Wardani, H.S., Sepminarti, T., and Rohman, A., 2016, Analysis of porcine gelatin DNA in commercial capsule shell using real-time polymerase chain reaction for halal authentication, *Int. J. Food Prop.*, 19 (9), 2127–2134.
- [10] Masiri, J., Benoit, L., Barrios-Lopez, B., Thienes, C., Meshgi, M., Agapov, A., Dobritsa, A., Nadala, C., and Samadpour, M., 2016, Development and

- validation of a rapid test system for detection of pork meat and collagen residues, *Meat Sci.*, 121, 397–401.
- [11] Perestam, A.T., Fujisaki, K.K., Nava, O., and Hellberg, R.S., 2017, Comparison of real-time PCR and ELISA-based methods for the detection of beef and pork in processed meat products, *Food Control*, 71, 346–352.
- [12] Bro, R., Jerome, J.W., Paul, R.M., and Kowalski, B.R., 1997, Review of chemometrics applied to spectroscopy: 1985–95, Part 3—Multi-way analysis, *Appl. Spectros. Rev.*, 32 (3), 237–261.
- [13] Lavine, B.K., 1998, Chemometrics, *Anal. Chem.*, 70 (12), 209–228.
- [14] Gemperline, P., 2006, *Practical Guide to Chemometrics*, CRC Press, Florida.
- [15] CAMO, 2016, *Chemometrics*, <http://www.camo.com/rt/Resources/chemometrics.html>, accessed on 17 July 2016
- [16] Hemmateenejad, B., 2006, Chemometrics in Iran, *Chemom. Intell. Lab. Syst.*, 81, 202–208.
- [17] Miller, J.N., and Miller, J.C., 2010, *Statistics and Chemometrics for Analytical Chemistry*, Prentice Hall, England.
- [18] Feudale, R.N., Woody, N.A., Tan, H., Myles, A.J., Brown, S.D., and Ferré, J., 2002, Transfer of multivariate calibration models: A review, *Chemom. Intell. Lab. Syst.*, 64 (2), 181–192.
- [19] Romia, M.B., and Bernardez, M.A., 2009, “Multivariate calibration for quantitative analysis” in *Infrared Spectroscopy for Food Quality analysis and control*, Sun, D.W. Eds., Elsevier, Amsterdam, 51–79.
- [20] Rohman, A., 2012, Application of FTIR spectroscopy for quality control in pharmaceutical products: a review, *Indonesian J. Pharm.*, 23 (1), 1–8.
- [21] Adam, M.J., 2004, *Chemometrics in Analytical Spectroscopy*, 2<sup>nd</sup> ed., The Royal Society of Chemistry, Cambridge.
- [22] Ragno, G., Ioele, G., and Risoli, A., 2004, Multivariate calibration techniques applied to the spectrophotometric analysis of one-to-four component systems, *Anal. Chim. Acta*, 512 (1), 173–180.
- [23] Mansor, T.S.T., Che Man, Y.B., and Shuhaimi, M., 2012, Employment of differential scanning calorimetry in detecting lard adulteration in virgin coconut oil, *J. Am. Oil Chem. Soc.*, 89 (3), 485–496.
- [24] Marikkar, J.M.N., Ghazali, H.M., Che Man, Y.B., and Lai, O.M., 2002, The use of cooling and heating thermograms for monitoring of tallow, lard and chicken fat adulterations in canola oil, *Food Res. Int.*, 35 (10), 1007–1014.
- [25] Rebecchi, 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 Chem.*, 192, 1025–1032.
- [26] Naes, T., Isaksson, T., Fearn, T., and Davis, T., 2002, *Multivariate Calibration and Classification*, NIR Publications, Chichester.
- [27] Martens, H., and Martens, M., 2001, *Multivariate Analysis of Quality, an Introduction*, John Wiley & Sons, Chichester.
- [28] Geladi P., 2003, Chemometrics in spectroscopy. Part 1. Classical chemometrics, *Spectrochim. Acta, Part B*, 58 (5), 767–782.
- [29] Sádecká, J., and Tóthová, J., 2007, Fluorescence spectroscopy and chemometrics in the food classification—a review, *Czech J. Food Sci.*, 25 (4), 159–173.
- [30] Rohman, A., and Che Man, Y.B., 2009, Analysis of cod-liver oil adulteration using Fourier transform infrared (FTIR) spectroscopy, *J. Am. Oil Chem. Soc.*, 86, 1149–1153.
- [31] Rohman A., and Che Man, Y.B., 2010, FTIR spectroscopy combined with chemometrics for analysis of lard in the mixtures with body fats of lamb, cow, and chicken, *Int. Food Res. J.*, 17, 519–527.
- [32] Rohman, A., Che Man, Y.B., Hashim, P., and Ismail, A., 2011, FTIR spectroscopy combined with chemometrics for analysis of lard adulteration in some vegetable oils, *CyTA J. Food*, 9 (2), 96–101.
- [33] Kurniawati, E., Rohman, A., and Triyana, K., 2014, Analysis of lard in meatball broth using Fourier transform infrared spectroscopy and chemometrics, *Meat Sci.*, 96 (1), 94–98.

- [34] Abdi, H., and Lynne, J.W., 2010, Principal Component Analysis, *Wiley Interdiscip. Rev. Comput. Stat.*, 2 (4), 433–459.
- [35] Ballabio, D., Skov, T., Leardi, R., and Bro, R., 2008, Classification of GC-MS measurements of wines by combining data dimension reduction and variable selection techniques, *J. Chemom.*, 22 (8), 457–463.
- [36] Che Man, Y.B., Rohman, A., and Mansor, T.S.T., 2011, Differentiation of lard from other edible oils by means of Fourier transform infrared spectroscopy and chemometrics, *J. Am. Oil Chem. Soc.*, 88 (2), 187–192.
- [37] Rohman, A., and Che Man, Y.B., 2011, Authentication analysis of cod liver oil from beef fat using fatty acid composition and FTIR spectra, *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess.*, 28 (11), 1469–1474.
- [38] Hashim, D.M., Che Man, Y.B., Norakhasa, R., Shuhaimi, M., Salmah, Y., and Syahariza, Z.A., 2010, Potential use of Fourier transform infrared spectroscopy for differentiation of bovine and porcine gelatins, *Food Chem.*, 118 (3), 856–860.
- [39] Nur Azira, T., Amin, I., and Che Man, Y.B., 2012, Differentiation of bovine and porcine gelatins in processed products via sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and principal component analysis (PCA) techniques, *Int. Food Res. J.*, 19 (3), 1175–1180.
- [40] Nur Azira, T., Che Man, Y.B., Raja Mohd Hafidz, R.N., Aina, M.A., and Amin, I., 2014, Use of principal component analysis for differentiation of gelatin sources based on polypeptide molecular weights, *Food Chem.*, 151, 286–292.
- [41] Nemati, M., Oveisi, M.R., Abdollahi, H., and Sabzevari, O., 2004, Differentiation of bovine and porcine gelatins using principal component analysis, *J. Pharm. Biomed. Anal.*, 34 (3), 485–492.
- [42] Raraswati, M.A., Triyana, K., Wahyudi, T., and Rohman, A., 2014, Differentiation of bovine and porcine gelatins in soft candy based on amino acid profiles and chemometrics, *J. Food Pharm. Sci.*, 2, 1–6.
- [43] Widyaninggar, A., Wahyudi, T., Triyana, K., and Rohman, A., 2012, Differentiation between porcine and bovine gelatins in commercial capsule shells based on amino acid profiles and principal component analysis, *Indonesian J. Pharm.*, 23 (2), 104–109.
- [44] Guimet, F., Boqué, R., and Ferré, J., 2004, Cluster analysis applied to the exploratory analysis of commercial Spanish olive oils by means of excitation-emission fluorescence spectroscopy, *J. Agric. Food. Chem.*, 52 (22), 6673–6679.
- [45] Cebi, N., Durak, M.Z., Toker, O.S., Sagdic, O., and Arici, M., 2016, An evaluation of Fourier transforms infrared spectroscopy method for the classification and discrimination of bovine, porcine and fish gelatins, *Food Chem.*, 190, 1109–1115.
- [46] Lee, J.Y., Park, J.H., Mun, H., Shim, W.B., and Lim, S.H., 2018, Quantitative analysis of lard in animal fat mixture using visible Raman spectroscopy, *Food Chem.*, 254, 109–114.
- [47] Che Man, Y.B., Syahariza, Z.A., and Rohman, A., 2011, Discriminant analysis of fats and oils using FTIR spectroscopy for Halal analysis, *Food Anal. Methods*, 4 (3), 404–409.
- [48] Upadhyay, N., Jaiswal, P., and Narayan Jha, S., 2018, Application of attenuated total reflectance Fourier Transform Infrared spectroscopy (ATReFTIR) in MIR range coupled with chemometrics for detection of pig body fat in pure ghee (heat clarified milk fat), *J. Mol. Struct.*, 1153, 275–281.
- [49] Basri, K.N., Hussain, M.N., Bakar, J., Sharif, Z., Khir, M.F.A., and Zoolfakar, A.S., 2017, Classification and quantification of palm oil adulteration via portable NIR spectroscopy, *Spectrochim. Acta, Part A*, 173, 335–342.