World Academy of Science, Engineering and Technology
International Journal of Nutrition and Food Engineering
Vol:14, No:4, 2020

Analytical Authentication of Butter Using Fourier Transform Infrared Spectroscopy Coupled with Chemometrics

M. Bodner, M. Scampicchio

***Abstract*—**Fourier Transform Infrared (FT-IR) spectroscopy coupled with chemometrics was used to distinguish between butter samples and non-butter samples. Further, quantification of the content of margarine in adulterated butter samples was investigated. Fingerprinting region (1400-800 cm^{-1}) was used to develop unsupervised pattern recognition (Principal Component Analysis, PCA), supervised modeling (Soft Independent Modelling by Class Analogy, SIMCA), classification (Partial Least Squares Discriminant Analysis, PLS-DA) and regression (Partial Least Squares Regression, PLS-R) models. PCA of the fingerprinting region shows a clustering of the two sample types. All samples were classified in their rightful class by SIMCA approach; however, nine adulterated samples (between 1% and 30% w/w of margarine) were classified as belonging both at the butter class and at the non-butter one. In the two-class PLS-DA model's (R^2 = 0.73, RMSEP, Root Mean Square Error of Prediction = 0.26% w/w) sensitivity was 71.4% and Positive Predictive Value (PPV) 100%. Its threshold was calculated at 7% w/w of margarine in adulterated butter samples. Finally, PLS-R model (R^2 = 0.84, RMSEP = 16.54%) was developed. PLS-DA was a suitable classification tool and PLS-R a proper quantification approach. Results demonstrate that FT-IR spectroscopy combined with PLS-R can be used as a rapid, simple and safe method to identify pure butter samples from adulterated ones and to determine the grade of adulteration of margarine in butter samples.

***Keywords*—**Adulterated butter, margarine, PCA, PLS-DA, PLS-R, SIMCA.

I. Introduction

THE adulteration of food, addition of substances and ingredients and false claims about origin are of major concern not only for consumers, but also to industries and to control authorities [1], [2]. Adulteration of butter has a long history. Even before the production of margarine, butter was adulterated with cheaper substances, such as chalk and potato starch [1], [3]. Other common adulterants are lard, low priced animal fats e.g. beef and mutton fat and oils [1], [4]. Primary driving force of these kinds of food fraud is and has always been the financial advantage (so called economically motivated adulteration) [5].

Different analytical techniques have been used for the identification of food fraud (food authentication). Besides numerous wet chemical procedures, instrumental analysis is performed, including the use of gas and liquid chromatography coupled to diverse detection methods, isotopic ratio mass spectrometry, mass spectrometry, nuclear magnetic resonance (NMR) spectroscopy, enzyme-linked immunosorbent assay (ELISA) and calorimetry [6]-[8]. As an example, Jin-Man et al. applied gas chromatography for the identification of adulterated milk fat by the quantification of fatty acids, triacylglycerol and cholesterol contents [9]. Lamanna et al. used 1H NMR profiling to evaluate the content of sheep milk in mixtures with cow milk [10]. However, these techniques are often expensive, time-consuming and require specific sample preparation. On the contrary, vibrational spectroscopic techniques offer several advantages over conventional methods, thus they are rapid, easy, economic, non-destructive and with limited risks for the operator's health [11], [12].

In case of mid infrared spectroscopy, samples absorb part of the infrared radiation, leading to the production of a spectral fingerprint. The fingerprints in the mid infrared region (4000-400 cm^{-1}) are the result of stretching, bending and rotating vibrational models of the biomolecules, like lipids, carbohydrates and proteins present in the sample [6], [13]. FT-IR spectroscopy is a powerful and robust technique applied in the field of food science, since it gives information about functional groups and chemical composition of the analyzed matrix [14]. Near infrared (NIR) and mid infrared (MIR) coupled with chemometrics methods have been largely applied to food matrices in the field of food quality and authentication investigations [6], [15], [16]. Several algorithms are available to perform unsupervised classification model, e.g. PCA, supervised not-discriminant modelling model, i.e. SIMCA and supervised discriminant model, such as PLS-DA but also PLS-R, suitable for qualitative and quantitative analysis of spectroscopic or spectrometric data [17]-[23]. Nowadays, chemometrics plays a central role in the authentication of edible fats [24]. Goodacre and Anklam used FT-IR spectroscopy coupled with chemometrics to identify cocoa butter adulterated with other vegetable oils [25]. Yang and collaborators applied FT-IR, NIR and Raman spectroscopy for the discrimination of edible oils and fats [26]. Bassbasi et al. coupled FT-IR and PLS-DA to identify the geographical origin of butter [27].

In the last decade, several studies on the combination of spectroscopic techniques and chemometrics to identify adulterated butter samples have been published. Koca et al. applied PLS-R to FT-IR analysis to quantify the amount of margarine in adulterated butter samples. For this, four calibration models were elaborated, using samples with different contents of margarine in butter as calibration set (0-

Martina Bodner is with the Free University of Bolzano, Italy (e-mail: mbodner@unibz.it).

5%, 0-25%, 20-60% and 0-100% v/v). The prediction ability of these four models were tested against an external validation set composed of adulterated butter samples with an amount of margarine of 2.5%, 13% and 45% (v/v), respectively. The best prediction results were reached using calibration sets of small adulteration ranges, compared to the performance of the calibration set consisting of the whole range of adulteration levels [28]. Mabood et al. investigated several multivariate methods (e.g. PCA, PLS-DA and PLS-R) to determine tallow adulteration in clarified butter using NIR spectroscopy. Here, samples were adulterated in the range between 1% and 20% (w/w). By PLS-DA a discrimination between pure butter samples and the adulterated ones was achieved; while the PLS-R model showed the ability to quantify tallow adulteration at a level lower than 2% (w/w) [29]. Fadzlillah et al. published several studies in which they coupled FT-IR spectroscopy with PLS-R to determine adulterations of butter (with beef fat, mutton fat and lard [4], [30], [31]. Nedelijkovic and Lohumi applied the PLS-R to Raman spectra to quantify and predict the concentration of margarine in adulterated butter samples [32], [33]. Besides the work of Koca [28], there are no other publications, to the authors' knowledge, coupling FT-IR analysis with chemometrics for the detection of butter samples adulterated with margarine.

The aim of this work was to apply classification and discriminant models (PCA, PLS-DA and PLS-R) to FT-IR spectroscopic data to develop a simple screening method for the non-targeted detection of adulterated butter. PLS-R was the better-suited model to correctly identify all butter and non-butter samples. Its ability to detect the amount of adulteration was precise above 7% (w/w) of margarine.

II. Experimental

A. Sample Collection and Preparation

20 butter and 11 margarine samples were purchased in local supermarkets in Bolzano (Italy) and Berlin (Germany) between July and August 2018. Samples were stored under refrigerated (4 °C ± 1 °C) conditions and analyzed within the recommended time of consumption. According to the labels, butter samples were produced exclusively from cow's milk. Specifically, 20 butter samples were purchased, two of them were ghee (or clarified) butter. Samples were chosen as different as possible in order to cover a broad variability of products and therefore chemical dissimilarities. Butter samples were purchased according to their origin of production, type of production (biological or conventional) and type of fat (cream, sour cream or ghee). Margarine samples were bought trying to maximize the differences in terms of composition (type of oils used), production (biological or conventional) and added ingredients (e.g. vitamins, butter, milk) (Table I).

57 adulterated samples were obtained by mixing butter and margarine at concentration ranges of 1%-50% w/w (Table II). Adulterated samples were prepared by randomly choosing eight butter and eight margarine samples with the intent of maximizing the chemical differences and, thus to obtain adulterated samples as different as possible. Pure butter samples were labeled as 0%, pure margarine samples as 100% and the adulterated butter samples were labeled accordingly to their adulteration degree.

TABLE I
List of Butter and Margarine Samples

Products	Origin	Notes
Butter1	Italy	C, D
Butter2	Italy	C, D
Butter3	Germany	C, D
Butter4	Italy	C, D
Butter5	Italy	C, D
Butter6	France	C, D
Butter7	Italy	C, D
Butter8	Italy	C, E
Butter9	Germany	C, E
Butter10	Germany	B, D
Butter11	Italy	B, D
Butter12	Germany	C, D
Butter13	Germany	C, D
Butter14	Turkey	C, F
Butter15	Germany	C, F
Butter16	Germany	B,D
Butter17	France	B, D
Butter18	Germany	B, D
Butter19	Germany	B, E
Butter20	Turkey	B, E
Margarine1	Italy	C, PO, CO, RO
Margarine2	Germany	C, SO, LO, RO, PO, VA, VD, VE
Margarine3	Italy	C, PO, CO, RO, VA, VD
Margarine4	Italy	C, RO, CO, LO, VB
Margarine5	Germany	C, SO, CO, M, BU
Margarine6	Germany	C, PO, RO, VD
Margarine7	France	C, SO, CO, VD
Margarine8	Germany	B, PO, CO, SO, BU
Margarine9	Poland	B, RO, PO, CO, LO
Margarine10	Germany	B, SO, PO, CO
Margarine11	Turkey	C, SO, PO, COS, RO, VA, VD, VE, M

List of butter and margarine samples purchased in Italy and Germany. B = biological production, C = conventional production, D = cream butter, E = sour cream butter, F = ghee (or clarified) butter, BU = butter, M = milk, CO = coconut oil COS = cottonseed oil, L= linseed oil, PO = palm oil, RO = rapeseed oil, SO = sunflower oil, VA = vitamin A, VB = vitamin B1, VD = vitamin D, VE = vitamin E.

TABLE II
List of Butter, Adulterated Butter and Margarine Samples

Set	Butter samples	Adulterated samples	Margarine Samples
Training	3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18	B10M1 1% - 2% - 5% - 10% - 20% - 30% - 50%	1, 2, 3, 5, 7, 10
		B10M5- 1% -5% - 7% - 10% -20% - 30% - 50%	
		B9M1- 1% - 2% - 5%	
		B12M2 - 1% - 2% - 7% - 15% - 40%	
		B17M3- 1% - 2% - 7% - 15% - 40%	
		B2M10 - 1% - 2% - 7% - 15% - 40%	
		B9M1- 10% -20% - 30% - 50%	
		B13M7- 1% - 2% - 7% - 15% - 40%	
Test	1, 2, 11, 19, 20	B1M4 - 1% - 7% - 10% - 30% - 40% - 50%	4, 6, 8, 9, 11
		B1M6- 1% - 7% - 10% - 30 - 40%	
		B19M4 - 1% - 7% - 10% - 30 - 40%	

List of butter, adulterated butter and margarine samples used in the training and test set. Adulterated samples were codenamed using the butter and margarine codes (e.g. B10M1 is the adulterated sample in which Butter10 and Margarine1 were used) and the level of margarine adulteration (% w/w).

B. FT-IR Spectroscopy

FT-IR spectra were collected using a Thermo Nicolet 6700 spectrometer equipped with a Smart Performer Accessory (horizontal attenuated total reflection (HATR) unit, single bounce diamond crystal), a Ge-on-KBr beamsplitter, and a deuterated triglycine sulfate (DTGS) detector. Prior to sample acquisition, a background spectrum of laboratory air was recorded and visually checked for remaining solvent or sample residues. The spectra were collected using OMNIC 7.4 software (Thermo Fisher Scientific Software, Germany). Spectra were acquired in the 4000–525 cm^{-1} range with a spectral resolution of 4 cm^{-1} at 40 °C. The background spectrum was recorded at the same temperature. 32 scans were recorded for each sample. The spectrometer performance was checked periodically using a polystyrene standard to verify wavenumber accuracy and repeatability. After a sample acquisition, the ATR crystal surface was carefully cleaned in a three-step procedure using n-hexane, acetone and methanol (in the given order), and dried with a lint-free tissue. A 1 min waiting period was applied before collecting a new background spectrum. Three replicates of each sample were measured, resulting in three spectra per sample that were exported using the OMNIC software, as previously described by Horn et al. recently [15].

C. Spectra Preprocessing

The initial data matrix consisted of 264 spectra, of which 60 of butters, 33 of margarines and 171 of the prepared adulterated butter samples.

The replicates were averaged and mean centered to have one spectrum for each sample. Thus, the final data matrix consisted of 88 spectra (20 butter samples, 11 margarine samples and 57 prepared adulterated samples).

Spectra preprocessing and statistical analysis were carried out using The Unscrambler X (Camo Software, Norway). Baseline correction and standard normal variate (SNV) were applied to correct both baseline and scatter effect. Afterwards, Savitzky-Golay first derivative (second order of polynomial, 11 points segment) was applied to the corrected spectra to resolve the overlapping bands and enhance the absorbance differences.

Spectral regions that did not provide relevant information (4000-3700 cm^{-1} baseline area, 2799-1800 cm^{-1} absorption of diamond crystal, 682-653 cm^{-1} disturbing absorption band of CO_2) were excluded prior to chemometrics, as previously described [15]. Thus, multivariate analysis was performed on spectra consisted of 1114 data points for the remaining spectral regions 3699-2800 cm^{-1}, 1799-683 cm^{-1}, and 652-525 cm^{-1}. Further analysis was conducted on reduced spectra consisting in 312 data points of the fingerprint region (1400-800 cm^{-1}) only, as previously described [28],[34].

D. Training and Test Sets and Multivariate Analysis

The data set was randomly split into two sub-sets, i.e. training and test set to perform multivariate analysis and verify model performance (Table II). Training set consisted of 62 samples (15 butters, 6 margarines, 41 adulterated butter samples) and test set of 26 samples (5 butters, 5 margarines, 16 adulterated butter samples). Unsupervised multivariate analysis (PCA) as well as supervised classification (SIMCA, PLS-DA) and regression (PLS-R) models were carried out using The Unscrambler X (Camo Software, Norway).

III. Results and Discussion

A. FT-IR Spectra of Butter and Margarine

The FT-IR spectra in the region 4000 – 525 cm^{-1} of a pure commercial butter and margarine sample are shown in Fig. 1.

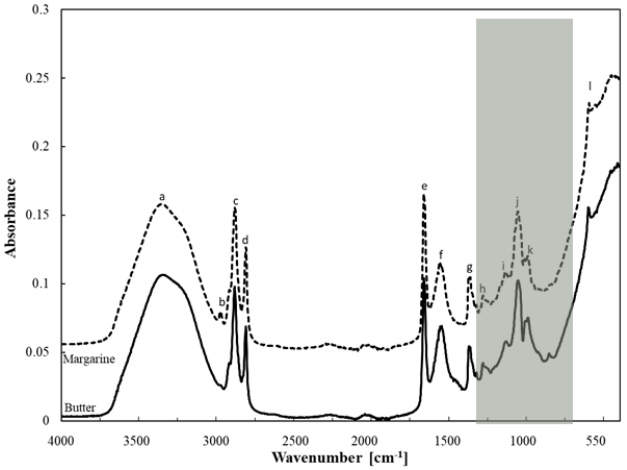

Fig. 1 Spectra butter and margarine sample in the range 4000-525 cm^{-1}. The fingerprint region (1400-800 cm^{-1}) is highlighted in grey

TABLE III
Modes of Vibrations [28], [35]

Assignment	Wavenumber [cm^{-1}]	Functional group vibration
a	3300	O-H symmetric stretching
b	3005	cis C=CH stretching
c, d	2923, 2850	Asymmetric and symmetric stretching vibration of methylene (–CH2) group
e	1746	Carbonyl (C=O) from the ester linkage of triacylglycerol
f	1650	cis C=C
g	1465	Bending vibrations of the CH2 and CH3 aliphatic groups
h	1376	Symmetric bending vibrations of CH3 groups
i, j	1236, 1163	Vibrations of stretching mode from the C–O group in esters
k	1135	−C-O-C stretching vibration
l	725	Overlapping of the methylene (–CH2) rocking vibration and to the out of plane vibration of cis-disubstituted olefins

Strong absorption was observed for both butter and margarine samples in the range of 3000-2800 cm^{-1}. The broad absorption band located at 3300 cm^{-1} (a) is assigned to the hydroxyl stretching vibration of water (O-H). The band related to the stretching of –C=CH groups of cis-unsaturation (b) reached its maximum height at 3004 cm^{-1} in the butter sample and shifted to 3009 cm^{-1} in the margarine sample as previously reported [28]. The bands assigned to stretching, bending and rocking vibration of the methylene group (c, d, g

and l) show little differences, in both intensity and shape, among butter and margarine samples. In the region 1800–1000 cm^{-1} a number of absorption bands related to vibrations of the C-O bond of esters are observed. Differences in the absorbance in butter and margarine samples became evident for the band at 1135 cm^{-1} (k) assigned to the stretching of the C-O-C groups of the bond between glycerol and fatty acid ester carbon in triacylglycerol (Table III).

B. PCA

In a first step, the 88 spectra were subjected to PCA to visualize variations occurring among samples of butter, margarine and different adulterated samples. The unsupervised model was applied to spectra consisted of 1114 data and of 312 data points (fingerprint region), respectively. In both PCA approaches, pure butter samples were well separated from margarine samples on the third principal component. Margarine samples were more spread out than butter samples probably because of the heterogeneity of their compositions, in terms of type and quantity of oils and addition of vitamins. The adulterated samples occupied the space between pure butter and margarine samples. It is interesting to point out that the two ghee butter samples (circled samples) were completed separated from all samples, including the pure butter samples. Since ghee butters are composed mostly of fat and contain almost no water and protein, the FT-IR spectra are quite different from the butter samples (Fig. 2, supplemental material). Ghee samples were thus considered as outlier and PCA and further chemometrics were performed without these two samples (resulting in a data set of 86 samples).

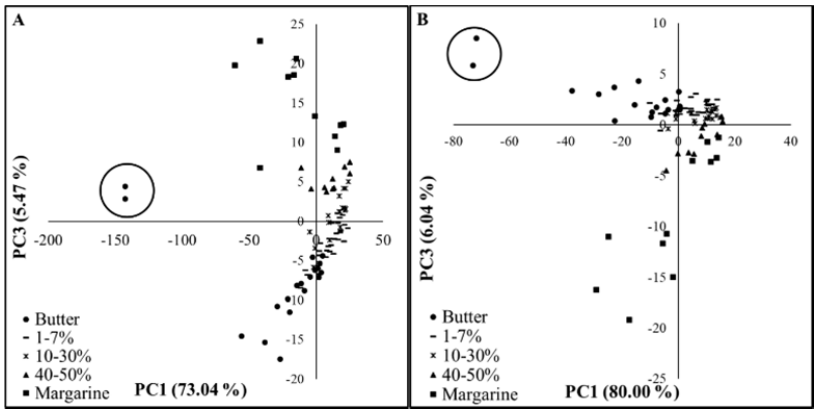

Fig. 2 Score plots PC3 *vs.* PC1 (A = spectra consisted of 1114 data points, B = fingerprint region). Circle samples are the two ghee butter

Fig. 3A shows the results of the PCA performed on the 1114 data points spectra. In Fig. 3C the PCA of the fingerprint region is reported. Pure butter (plain circle) and margarine (dotted circle) samples formed two clear clusters according to the second principal component. Adulterated samples (dashed circle) are located between these two clusters. Regarding the respective two scores plots, the third component (in the case of the 1114 data points PCA) and the second one (in case of the fingerprint PCA) seem to be the one associated with the content of margarine. To better understand the significance of these components, the scores of PC3 (Fig. 3B) and PC2 (Fig. 3D) were plotted *vs.* the amount of margarine adulteration in the analyzed samples. In case of Fig. 3B (1114 data points) the correlation between PC3 scores and the level of adulteration is not evident, whereas in Fig. 3D PC2 shows a stronger correlation (PCA of the fingerprint region). Here, all pure butter samples are characterized by positive values of PC2. By contrast, all pure margarine samples score negative values of PC2. The trend among the adulterated samples noticeably goes in the direction of positive values for lower amount of adulteration (from 0% to 20% w/w), to negative values for higher amount of adulteration (from 30% to 100% w/w). As a result, it is possible to affirm that the samples can be clustered into two groups, i.e. pure butter and low-adulterated sample *vs.* high-adulterated samples and margarine, accordingly to PC2.

In Fig. 4 the PC2-loadings plot of the fingerprint PCA is shown. The band with the highest intensity is the one at 1135 cm^{-1}, inducing the clustering between butter and margarine samples. As previously discussed (Table III), this signal is assigned to the stretching of the C-O-C groups of the bond between glycerol and fatty acid ester carbon in triacylglycerol. The both FT-IR spectra of butter and margarine showed differences in the absorbance in this region. These observations are consisting with results previously reported [28].

C. SIMCA

After selecting the fingerprint region as the most informative for the discrimination among butter, adulterated and margarines samples, and having created the PCA model without the two ghee butters, a supervised modelling method using SIMCA was developed to classify butter and non-butter samples. The dataset was split into two sets: (1) the training set, used to build and internal validate the model and (2) the test set for external validation (Table II). Samples were divided into two classes: butter and non-butter (both adulterated and margarine samples). One model for each class was built performing a PCA. Three latent variables were used for the butter samples and six latent variables for the non-

butter samples. Then, the butter and non-butter samples of the test set were classified accordingly. Table IV summarizes the classification membership of butter and non-butter samples. Samples recognized as member of a class (within the limits on sample-to-model distance and leverage) have a star in the corresponding column.

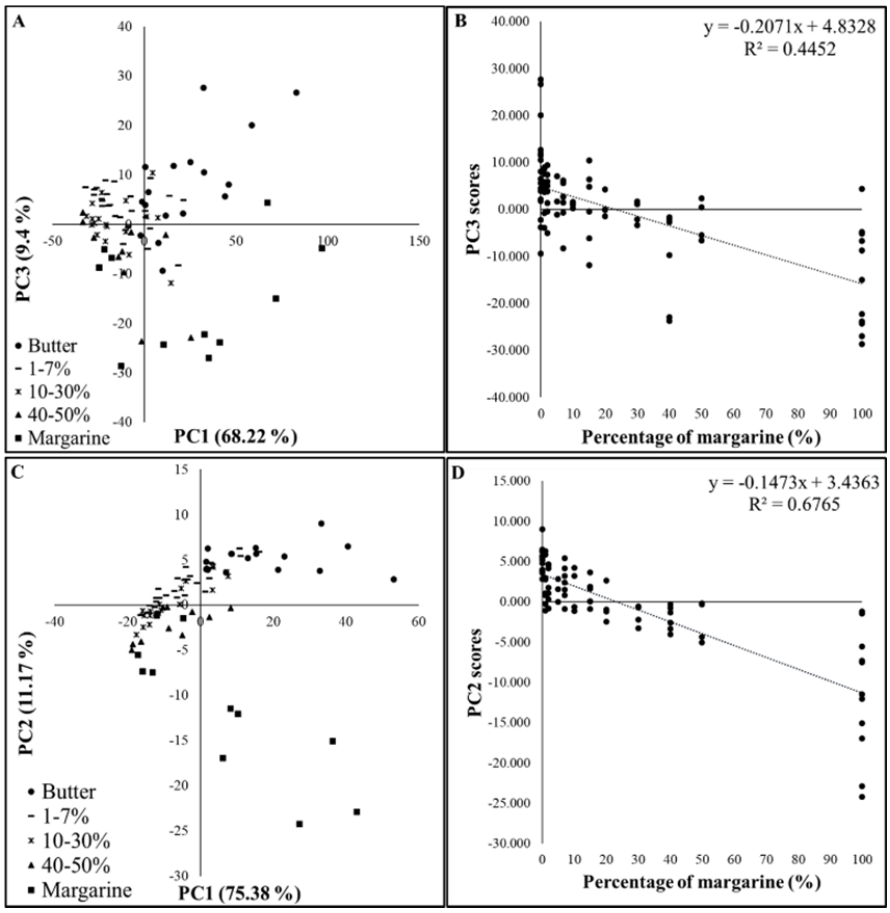

Fig. 3 Score plots PC3 *vs*. PC1 (A) and PC2 *vs*. PC1 (C) and PC3 scores (B and D). A and B are related to spectra consisted of 1114 points; C and D to the fingerprint region. Pure butter samples = plain circle; adulterated samples = dashed circle; margarine samples = dotted circle

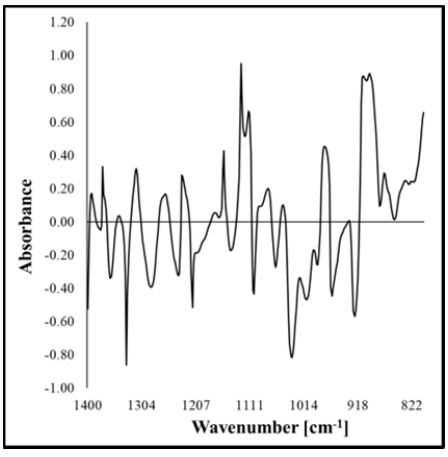

Fig. 4 Loadings plot of the PC2 of the fingerprint PCA

At the 5% significance level, all butter samples are recognized by the rightful class model. Thus, no false positives are present. However, some samples (all low-adulterated samples in the range 1% to 30% w/w) are classified as belonging to two classes and therefore are considered false negatives. Since, SIMCA is a "soft" modeling method; a sample can be classified as belonging to multiple classes when its residual distance from the model is below the statistical limit for different classes. Therefore, pure butter and pure margarine samples are correctly classified, while adulterated butter samples can be rightfully classified only if they contain at least 30% w/w of margarine.

D.PLS-DA

Then, a supervised classification model using PLS-DA was developed to discriminate among pure butter and non-butter samples. Training and test sets were the same used for the SIMCA model. Two latent variables were used to build the model. R^2 and RMSEC (Root Mean Squared Error Calibration) values were 0.79 and 0.15% w/w, while R^2 and RMSEV (Root Mean Squared Error Cross Validation) values were 0.73 and 0.17% w/w. The prediction ability of the model was tested with the test set of samples. Table V summarizes the predicted classification (R^2 = 0.73, RMSEP = 0.26%) of butter and non-butter samples. It is interesting to point out that

the two adulterated samples classified as pure butter contain a concentration of margarine between 1% and 7% w/w. Accordingly, the threshold for a correct classification was calculated at 7% (w/w) of margarine in adulterated samples. Thus, the two-class model was found to be a successful tool in differentiating between pure and adulterated samples.

TABLE IV
CLASSIFICATION MEMBERSHIP OF THE SIMCA MODEL FOR THE TEST SET

		Class	
		Pure Butter	Non-butter
Samples	Butter11	*	
	Butter19	*	
	Butter20	*	
	B1M4 1%	*	*
	B1M6 1%	*	*
	B19M4 1%	*	*
	B1M4 7%	*	*
	B1M6 7%	*	*
	B19M4 7%	*	*
	B1M4 10%	*	*
	B1M6 10%	*	*
	B19M4 10%	*	*
	B1M4 30%	*	*
	B1M6 30%	*	*
	B19M4 30%	*	*
	B1M4 40%		*
	B1M6 40%		*
	B19M4 40%		*
	B1M4 50%		*
	Margarine4		*
	Margarine6		*
	Margarine8		*
	Margarine9		*
	Margarine11		*

TABLE V
CONFUSION MATRIX OF THE PLS-DA MODEL FOR THE TEST SET

		Actual	
		Pure Butter	Non-butter
Predicted	Pure Butter	5	0
	Non-butter	2	19

Sensitivity, specificity, PPV and Negative Predictive Value (NPV) were calculated for SIMCA e PLS-DA models (Table VI). Sensitivity is the ability to classify positive samples to the belonging class, in other words, to detect true positives. Specificity, on the opposite, is the ability to detect true negatives. PPV is a test's probability to correctly identify true positives, avoiding false positives. NPV is a test's probability to correctly identify true negatives, avoiding false negatives [36].

TABLE VI
CONTINGENCY TABLE FOR SIMCA AND PLS-DA MODELS

Classification model	Sensitivity	Specificity	PPV	NPV
SIMCA	29.4%	100%	100%	42.9%
PLS-DA	71.4%	100%	100%	90.5%

Performances of SIMCA and PLS-DA models were similar in terms of specificity and PPV, but differ greatly regarding sensitivity and NPV. A high NPV indicates the test's ability to minimize false negative results [36]. Thus, to minimize the number of pure butter samples classified as non-butter. PLS-DA seems to be a better performing model for the investigation of adulterated butter.

E. PLS-R

Another model was developed using PLS-R considering the fingerprint region and the actual concentration of margarine in butter samples (0%, 1%, 7%, 10%, 30%, 40% and 100% w/w). Training and test sets were the same used for the SIMCA and PLS-DA models. Kernel algorithm was used and the internal validation was performed as k-10 cross-validation. The optimum of the PLS-R model was determined according to RMSEC and RMSEV. Thus, three latent variables were used to build the model. R^2 and RMSEC values were 0.85 and 10.28% w/w, R^2 and RMSEV values were 0.82 and 11.83% w/w. Table VII shows the results for the prediction ($R^2 = 0.84$, RMSEP = 16.54%) of the concentration of margarine (% w/w) of the samples used for the external validation. All pure butter samples were correctly quantified. Regarding the adulterated samples, the prediction of the quantity of margarine was not precise for samples containing less of 7% (w/w) of margarine.

TABLE VII
PREDICTION OF THE CONCENTRATION OF MARGARINE

	Actual [%]	Predicted [%]
Samples	0.0	0.6 ± 0.2
	1.0	3.2 ± 2.1
	7.0	6.1 ± 3.1
	10.0	9.7 ± 1.5
	30.0	28.3 ± 2.4
	40.0	39.4 ± 2.8
	100.0	89.8 ± 12.7

Prediction of the concentration of margarine in the samples of the prediction set. Actual and predicted refer to the actual and predicted concentration of margarine in butter samples (in % w/w). Values are expressed as mean (%) of different predicted samples ± standard deviation (%).

Although the model overestimates the margarine content in low concentration adulterated samples (below 7% w/w), it was capable to correctly identify all pure and adulterated butter samples. As a result, although the model was tested only with 26 samples, it revealed to be appropriate as a quantification tool.

IV. CONCLUSION

FT-IR spectroscopy coupled with chemometrics showed to be a powerful tool to authenticate pure butter samples from adulterated butter samples. Moreover, this approach can be used to quantify the adulteration grade of butter with margarine. Models can be developed using the fingerprint region (1400-800 cm^{-1}). Butter and margarine samples were well separated in the PCA model, even if it was not possible to differentiate adulterated samples. Although all pure butter samples were classified in their rightful class by SIMCA approach, nine adulterated samples (between 1% and 30% w/w of margarine) were classified as belonging both at the butter and non-butter class in the SIMCA approach. Two-class PLS-DA model's ($R^2 = 0.73$, RMSEP = 0.26% w/w) threshold was calculated at 7% w/w of margarine in adulterated butter samples. Finally, PLS-R model identified correctly all butter

and non-butter samples. Regarding its ability to predict the amount of adulteration, it was not precise below 7% (w/w) of margarine. To the authors' knowledge, only one study was published, describing the application of FT-IR spectroscopy coupled with chemometric to estimate the adulteration of butter with margarine [28]. Koca et al. focused on the development of calibration models to predict butter adulterations in samples with adulteration level of 2.5%, 13% and 45% (v/v). The authors did not investigate the ability of the model to identify authentic butter samples [28]. Although the PLS-R model developed in this study is not as precise as Koca's in predicting the adulteration level between 1% and 7% (w/w), it enables the identification of pure butter samples to verify authentic products.

FT-IR spectroscopy requires minimum sample preparation and it is a rapid, sensitive, non-destructive and robust technique for the detection and the quantification of margarine in butter samples. Thus, FT-IR coupled with PLS-DA and PLS-R can be successfully used in quality control analysis for the authentication of pure butter samples from adulterated ones.

ACKNOWLEDGMENT

M.B. Author wants to thank Dr. Susanne Esslinger and Dr. Carsten Fauhl-Hassek for their priceless help and critical suggestions.

REFERENCES

[1] H. Deelstra, D. Thorburn Burns, M. J. Walker, "The adulteration of food, lessons from the past, with reference to butter, margarine and fraud," *Eur Food Res. Technol*., vol. 239, pp. 725-744, 2014.

[2] S. Esslinger, J. Riedl, C. Fauhl-Hassek, "Potential and limitations of non-targeted fingerprinting for authentication of food in official control," *Food Res. Intern*., vol. 60, pp. 189-204, 2014.

[3] F. L. Acar, "Traité des Falsifications des substances médicamenteuses et alimentaires et les moyens de les reconnaître," *Impr de L.J. De Cort*, Fossé-aux-Crapaux, Anvers., pp. 239–240, 1848.

[4] N. A. Fadzlillah, Y. B. Che Man, A. Rohman, I. Amin, M. Shuhaimi, A. Khatib, "Authentication analysis of butter from beef fat using Fourier Transform Infrared (FTIR) spectroscopy coupled with chemometrics," *Int. Food Res. J*., vol. 20, pp. 1383-1388, 2013.

[5] J. Spink, D. C. Moyer, "Defining the Public Health Threat of Food Fraud," *J. Food Sci*., vol. 76, pp. 157-163, 2011.

[6] G. P. Danezis, A. S. Tsagkaris, F. Camin, V. Brusic, C. A. Georgiou, "Food authentication: Techniques, trends & emerging approaches," *Trends in Anal. Chem*., vol. 85, pp. 123-132, 2016.

[7] E. Hong, S. Y. Lee, J. Y. Jeong, J. M. Park, B. H. Kim, K. Kwon, H. S. Chun, "Modern analytical methods for the detection of food fraud and adulteration by food category," *J. Sci. Food Agric*., vol. 97, pp. 3877-3896, 2017.

[8] R. Karoui, J. De Baerdemaeker, "A review of the analytical methods coupled with chemometric tools for the determination of the quality and identity of dairy products," *Food Chem*., vol. 102, pp. 621-640, 2007.

[9] K. Jin-Man, K. Ha-Jung, P. Jung-Min, "Determination of Milk Fat Adulteration with Vegetable Oils and Animal Fats by Gas Chromatographic Analysis," *J. Food Sci*., vol. 80, pp. 1945-1951, 2015.

[10] R. Lamanna, A. Braca, E. Di Paolo, G. Imparato, "Identification of milk mixtures by 1H NMR profiling," *Magn. Resons. Chem*., vol. 49, pp. 22-26, 2011.

[11] M. H. Moh, Y. B. Che Man, F.R. van de Voort, W. J. W. Abdullah, "Determination of peroxide value in thermally oxidized crude palm oil by near infrared spectroscopy," *J. Am. Oil Chem. Soc*., vol. 76, pp. 19-23, 1999.

[12] G. Yildiz, R. L. Wehling, S.L. Cuppett, "Methods for determining oxidation of vegetable oils by near-infrared spectroscopy," *J. Am. Oil Chem. Soc*., vol. 78, pp. 495-502, 2001.

[13] M. Manfredi, E. Robotti, F. Quasso, E. Mazzucco, G. Calabrese, E. Marengo, "Fast classification of hazelnut cultivars through portable infrared spectroscopy and chemometrics," *Spectrochimi. Acta A Mol. Biomol. Spectrosc*., vol. 189, pp. 427-435, 2018.

[14] T. F. Kumosinski, H. M. Farrell, "Determination of the global secondary structure of proteins by Fourier transform infrared (FTIR) spectroscopy," *Trends Food Sci. Technol*., vol. 4, pp. 169-175, 1993.

[15] B. Horn, S. Esslinger, M. Pfister, C. Fauhl-Hassek, J. Riedl, "Non-targeted detection of paprika adulteration using mid-infrared spectroscopy and one-class classification – Is it data preprocessing that makes the performance?," *Food Chem*., vol. 257, pp. 112-119, 2018.

[16] C. A. Nunes, "Vibrational spectroscopy and chemometrics to assess authenticity, adulteration and intrinsic quality parameters of edible oils and fats," *Food Res. Int*., vol. 60, pp. 255-261, 2014.

[17] N. Dupuy, L. Duponchel, J.P. Huvenne, B. Sombret, P. Legrand, "Classification of edible fats and oils by principal component analysis of Fourier transform infrared spectra," *Food Chem*., vol. 51, pp. 245-251, 1996.

[18] M. Estekia, J. Simal-Gandarab, Z. Shahsavaria, S. Zandbaafa, E. Dashtakia, H. Yvan Vander, "A review on the application of chromatographic methods, coupled to chemometrics, for food authentication," *Food control*., vol. 93, pp. 165-182, 2018.

[19] D. Granato, J. S. Santos, G. B. Escher, B. L. Ferreira, R. M. Maggio, "Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective," *Trends in Food Sci. Technol*., vol. 72, pp. 83-90, 2018.

[20] K. Javidnia, M. Parish, S. Karimi, B. Hemmateenejad, "Discrimination of edible oils and fats by combination of multivariate pattern recognition and FT-IR spectroscopy: A comparative study between different modeling methods," *Spectrochimi. Acta A Mol. Biomol. Spectrosc*., vol. 104, pp. 175-181, 2013.

[21] L. Lenhardt, R. Bro, I. Zeković, T. Dramićanin, M. D. Dramićanin, "Fluorescence spectroscopy coupled with PARAFAC and PLS DA for characterization and classification of honey," *Food Chem*., vol. 175, pp. 284-291, 2015.

[22] P. S. Sampaio, A. Soares, A. Castanho, A. S. Almeida, J. Oliveira, C. Brites, "Optimization of rice amylose determination by NIR-spectroscopy using PLS chemometrics algorithm," *Food Chem*., vol. 242, pp. 196-204, 2018.

[23] L. Bertacchi, M. Cocchi, M. Li Vigni, A. Marchetti, E. Salvatore, S. Sighinolfi, M. Silvestri, C. Durante, "The impact of chemometrics on food traceability," *Data Handl. Sci. Technol*., vol. 28, pp. 371-410, 2013.

[24] D. M. A. N. Luykx, S. M. van Ruth, "An overview of analytical methods for determining the geographical origin of food products," *Food Chem*., vol. 107, pp. 897–911, 2008.

[25] R. Goodacre, E. Anklam, "Fourier Transform Infrared Spectroscopy and Chemometrics as a Tool for the Rapid Detection of Other Vegetable Fats Mixed in Cocoa Butter," *J. Am. Oil Chem. Soc*., vol. 78, pp. 993-1000, 2001.

[26] H. Yang, J. Irudayaraj, M. M. Paradkar, "Discriminant analysis of edible oils and fats by FTIR, FT-NIR and FT-Raman spectroscopy," *Food Chem*., vol. 93, pp. 25-32, 2005.

[27] M. Bassbasi, M. De Luca, G Ioele, A. Oussama, G. Ragno, "Prediction of the geographical origin of butters by partial least square discriminant analysis (PLS-DA) applied to infrared spectroscopy (FTIR) data," *J. Food Comp. Anal*., vol. 33, pp. 210-215, 2014.

[28] N. Koca, N. A. Kocaoglu-Vurma, W. P. Harper, L. E. Rodriguez-Saona, "Application of temperature-controlled attenuated total reflectance-mid-infrared (ATR-MIR) spectroscopy for rapid estimation of butter adulteration," *Food Chem*., vol. 121, pp. 778-782, 2010.

[29] F. Mabood, G. Abbas, F. Jabeen, Z. Naureen, A. Al-Harrasi, A. M. Hamaed, J. Hussain, M. Al-Nabhani, M. S. Al Shukaili, A. Khan, S. Manzoor, "Robust new NIRS coupled with multivariate methods for the detection and quantification of tallow adulteration in clarified butter samples," *Food Addit. Contam. Part A*., vol. 35, pp. 404-411, 2018.

[30] N. A. Fadzlillah, Y. B. Che Man, A. Rohman, I. Amin, M. Shuhaimi, A. Khatib, "Application of FTIR-ATR spectroscopy with multivariate analysis for rapid estimation of butter adulteration," *J. Oleo Scie*., vol. 62, pp. 555-562, 2013.

[31] N. A. Fadzlillah, Y. B. Che Man, I. Amin, R. Arieff Salleh, M. Y. Farawahidah, M. Shuhaimi, A. Khatib, "FTIR-ATR Spectroscopy Based Metabolite Fingerprinting as A Direct Determination of Butter

Adulterated With Lard," *Int. J. Food Prop*., vol. 18, pp. 372-379, 2015.
[32] A. Nedeljković, P. Rösch, J. Popp, J. Miočinović, M. Radovanović, P. Pudja, "Raman Spectroscopy as a Rapid Tool for Quantitative Analysis of Butter Adulterated with Margarine," *Food Anal. Methods*, vol. 9, pp. 1315-1320, 2016.
[33] S. Lohumi, H. Lee, M. S. Kim, J. Qin, B. K. Cho, "Through-packaging analysis of butter adulteration using line-scan spatially offset Raman spectroscopy," *Anal. Bioanal. Chem*., vol. 410, pp. 5663-5673, 2018.
[34] A. F. Nurrulhidayah, A. Rohman, I. Amin, M. Shuhaimi, A. Khatib, "Analysis of chicken fat as adulterant in butter using fourier transform infrared spectroscopy and chemometrics," *Grasas y Aceites*, vol. 64, pp. 349-355, 2013.
[35] M. D. Guillén, N. Cabo, "Infrared spectroscopy in the study of edible oils and fats," *J. Sci. Food Agric*., vol. 75, pp. 1-11, 1997.
[36] R. Trevethan, "Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice," *Front. Public Health*, vol. 5, pp. 307, 2017.