ATR-MIR spectroscopy to predict commercial milk major components: a comparison between a handheld and a benchtop instrument

Giulia Gorla[a], Montserrat Mestres[b], Ricard Boqué[c], Jordi Riu[c], Davide Spanu[a], Barbara Giussani[a]

[a] Dipartimento di Scienza e Alta Tecnologia. Università degli Studi dell'Insubria. Via Valleggio, 9. 22100 Como, Italy

[b] Instrumental Sensometry (iSens), Department of Analytical Chemistry and Organic Chemistry, Universitat Rovira i Virgili, Campus Sescelades, 43007 Tarragona, Spain

[c] Chemometrics, Qualimetrics and Nanosensors Group, Department of Analytical Chemistry and Organic Chemistry, Universitat Rovira i Virgili, Campus Sescelades, 43007 Tarragona, Spain

corresponding author:
Barbara Giussani
Dipartimento di Scienza e Alta Tecnologia
Università degli Studi dell'Insubria
via Valleggio, 9 - 22100 Como – Italy
tel: +39 031 234 6434
email: barbara.giussani@uninsubria.it

Highlights:

- ATR-FT-MIR coupled with multivariate regression allowed to predict fat, carbohydrate and protein contents in commercial milk samples
- ATR-FT-MIR coupled with chemometrics allowed to classify commercial milk samples according to the presence or absence of lactose
- Portable ATR-FT-MIR instrument are a good alternative to benchtop instrument in the quantitative analysis of macrocomponents in commercial milk samples
- Portable ATR-FT-MIR instrument coupled with chemometrics could be used in quality control of commercial milk samples

ABSTRACT

There is a growing need of measurement technologies that can be used close to the sample source and optical spectroscopy is an excellent example of this genre of technology: from the lab to the field. This study investigates the possibility to quantify the major components and to detect the presence or absence of lactose in commercial milks with ATR-MIR spectroscopy. We explored the possibility to use a portable and economical ATR-MIR instrument, comparing the results with a benchtop system. Commercial milk samples from Italy, Switzerland and Spain were chosen covering the maximum range of variation for protein, carbohydrate and fat content. The analytical protocol was optimized to make it as fast and useable as possible for both instruments, from the sample pretreatment to the instrumental parameters. Multivariate calibration was used to correlate the recorded spectra to the content of the major milk components, while a classification was done in order to classify samples with or without lactose. A comparison was performed between the predictive capabilities of the models built with different data pretreatments, different variable selection methods and different validation systems to obtain the best results and to assure robust models.

Keywords: ATR-MIR, milk, milk major constituents, fatty acids determination in milk, proteins determination in milk, carbohydrates determination in milk

1. INTRODUCTION

Quality requirements in the food industry are increasingly promoting the emergence of techniques enabling real-time monitoring of processes and products, even on the field or in the farm. For these reasons, portable and hand-held equipment represents an ideal tool for food quality control.

Among the available technologies, infrared spectroscopy (IR) has proven to be a rapid, high-throughput and non-destructive solution. Since each sample provides a unique and characteristic infrared spectrum, this spectrum can be considered as the sample fingerprint [1]. Moreover, when coupled with chemometric techniques like multivariate regression[2], spectra can be used for the simultaneous prediction of several chemical and/or physical properties of samples.

Attenuated total reflectance in the mid-infrared (ATR-FT-MIR) spectroscopy requires a very small quantity of sample (a drop in the case of liquids) and none or little sample pretreatment is needed, allowing the analysis of food samples in their natural state and being a "reagent free" technique. Moreover, it is excellent for strong absorptive samples like aqueous solutions, which are difficult to measure with other FT-MIR methods due to the strong absorption bands of water [3] as occurs when dealing with milk, which has already been analyzed by using ATR-FT-MIR.

Among food products, milk, and more specifically cow milk, is one of the most important ones because of its nutritional properties. This is the main reason for its wide consumption worldwide [4], although it should be noted that its lactose content is a serious problem for people with intolerance to the intake of this sugar [5].

In all cases, rapid, low-cost, accurate and quantitative methods for the analysis of both raw and commercial milk are increasingly required to establish milk value for consumer information and for quality control. Thus, for example Bassbasi et al. [6] determined solid non-fat in raw milk, Etzion et al. [3] quantified proteins and Soyeurt et al. estimated fatty acids [7]. The good results obtained in the different applications have led to the development of portable instruments, which gather the advantages of ATR-FT-MIR technology but, moreover, allow performing the analysis directly where it is needed, either on the field or in the production lines. In the literature there are already successful applications of this technique, for example to control the wine vinification process [8], to identify hazelnut varieties [9], to determine the protein content in microencapsulated fish oil supplement [10], to quantify trans fats of edible oils [11] and fast food [12] and to evaluate tomato juice quality parameters [13].

In this study, the use of a portable ATR-FT-MIR spectrometer was evaluated as a fast technique for determining the major constituents in commercially milk samples purchased in local supermarkets from Italy, Switzerland and Spain and to classify samples according to the presence or absence of lactose (as declared by the producer). The results obtained using the hand-held device were compared to those obtained when analyzing the same samples with a conventional benchtop spectrometer, either with no sample pretreatment or after applying a slight homogenization, with the aim to compare the performances of both instruments and to prove if the portable technique could be a valid alternative to the benchtop one.

2. EXPERIMENTAL

46 cow milk samples were purchased in Italian (17 milks), Swiss (7 milks) and Spanish (22 milks) local supermarkets. For a robust modeling and to emulate the whole commercial milk population, the sample set was selected covering the maximum variation range in protein, fat and carbohydrate contents. Thus, the samples were chosen with low, medium and high-fats content according to the commercial labels, 11 milks for early childhood (with a high concentration of proteins and carbohydrates – some of them fortified with fibers and omega3 and omega6 fatty acids), 11 and 35 whole milks with and without lactose, respectively, and 5 UHT (Ultra-High Temperature processing) fresh milks. The samples ranged between 1.3 and 7.0 g/100 mL for proteins, between 0.1 and 3.9 g/100 mL for fat and between 3.4 and 9.3 g/100 mL for carbohydrates. Complete information about the selected samples can be found in the Supplementary Material section. Samples were stored at room temperature (at 4°C in the case of fresh samples) until their analysis. Regarding sample preparation, prior to the analysis the milk pack was gently shaken and then 120 µL were taken and transferred into a glass vial to be thermostated at 25° C or 40°C depending on the chosen procedure. Finally, the samples were sonicated with a BRANSONIC® CPX3800H-E under temperature control.

2.1 ATR-FT-MIR analysis

ATR-FT-MIR spectra were collected with a portable 4100 EXOSCAN Spectrometer (Agilent Technologies) equipped with a diamond crystal window. Samples were analyzed by placing one milk drop on top of the crystal and spectra were recorded in the region between 4000 and 650 cm^{-1} with 32 scans and at 4 cm^{-1} resolution.

The same samples were also analyzed with a benchtop instrument, a Nicolet™ iS10 FTIR Spectrometer (Thermo Scientific™) and the spectra were collected in absorbance mode from 4000 to 600 cm^{-1} with 32 scans and at 2 cm^{-1} resolution.

Samples were randomly collected and analyzed in triplicate. An air background was acquired before each measurement and the sample holder was carefully cleaned after each measurement by following a cleaning sequence with four solvents of increasing polarity to remove any residue from the equipment: pentane, dichloromethane, acetone and ethanol. Before to apply the next solvent, and specially before sample measurement, the crystal was thoroughly dried with cotton wipes. This analytical procedure was applied to both instruments in order to compare the results.

Absorbance data were used for the multivariate modeling. The average spectrum was calculated and several preprocessing methods (and combination of them) were tested in order to remove unwanted variations in the spectra due to noise or baseline shifts not related with the chemical information.

2.2 Chemometric tools and data processing

Calculations were made in R environment, using the R Studio interface [14] and several packages. ChemoSpec [15] and ChemoSpecUtils [16] were used for spectra visualization, manipulation and data pretreatments. Exploratory analysis and prediction models were performed with Chemometrics [17] and ChemometricswithR [18]. The final models, both for the prediction and for the classification were built with

the package mdatools [19]. The package plsVarSel [20] was employed to select variables, and prospectr [21] was used to develop different validation strategies. Some evaluations were carried out using the CAT (Chemometric Agile Tool) software [22] and The Unscrambler X (CAMO Analytics). PLS-DA models were calculated using the PLS Toolbox 8.7 (R8.7.1) (Eigenvector Research Inc., Wenatchee, WA, USA) for Matlab R2019a (Mathworks Inc, Natick, MA, USA).

PLS and PLS-DA models

Partial Least Squares-Regression (PLS-R) [23] was used to predict major components in milk samples. PLS-DA [24] was chosen to classify samples according to the lactose content (with or without lactose, as stated by the producers).

Signal pretreatment methods

Different pretreatment methods and combinations of them were evaluated in order to obtain the best results in terms of prediction error and number of latent variables used.

To avoid ATR spectral noise, smoothing with different window widths (from 5 points to 13) and polynomial order (p=1, 2, 3) and different combinations with derivative Savitsky-Golay filter (d=1, 2) were tested. Moreover, Standard Normal Variate (SNV) and Multiplicative Scatter Correction (MSC) were evaluated in order to remove the light scattering influence if present. Spectra were mean centered before modelling.

Variable selection methods

In order to improve the performances and to simplify the models, different variable selection methods were evaluated [20]. Results from loading weights (LW), regression coefficients (RC), Variable Importance in Projection (VIP) and the combination of the three were considered and compared. Backward variable elimination PLS (BVE-PLS) [25], Iterative predictor weighting PLS (IPW-PLS) [26] and Interval PLS (iPLS) [27] were also used. For iPLS, a combination of different intervals was evaluated in forward and backward arrangements in one-step procedures.

Model validation

Different validation methods were used to assess the model robustness. First, leave-one-out (LOO) and leave-more-out validations were applied to get a first fast comparison between different model performances and to check the presence of outlier samples.

Final models were validated with an external validation strategy. Samples were divided into calibration (cal) and validation (val) sets using two selection algorithms: CADEX and HONIGS. The CADEX [28] algorithm builds the training set with the most different samples by selecting them using a distance maximin criterion [29]. The HONIGS [30] one is optimized for the application on absorbance values. It selects the sample with the maximum absorbance, inserts it into the calibration set and removes the absolute value from all the other spectra. The procedure continues until the desired number of samples is reached in the calibration model.

Different proportions between calibration and validation sets were compared: ½ cal - ½ val, 1/3 cal – 2/3 val, 2/3 cal – 1/3 val.

3. RESULTS AND DISCUSSION

3.1 Spectroscopic data

Different acquisition parameters were tested in order to obtain suitable quality signals. For both instruments, resolutions of 2, 4 and 8 cm^{-1} and scan numbers of 32, 64 and 128 were tested. Preliminary models suggested that 32 scans and 2 cm^{-1} resolution for the benchtop instrument and 4 cm^{-1} for the portable one showed the best compromise between analysis speed and information collected. The models also allowed ascertaining that data at wavenumbers lower than 700 cm^{-1} were very noisy. Therefore, we decided to not consider this spectral region in the following modeling. Final models were built from samples with 899 wavenumbers for the handheld system and 1389 wavenumbers for the benchtop one.

Spectra from the two instruments (Figure 1) were visually compared and, as expected, those from the portable spectrometer were noisier.

The main IR absorbances, for both instruments, can be identified as those typical of aqueous samples: the O-H stretching around 3300 cm^{-1} and the H_2O bending at 1640 cm^{-1}. Other important absorbance bands are the CH_2 symmetric stretching at 2860 cm^{-1}, the C=0 stretching at 1743 cm^{-1} and the C–N and N–N stretching around 1550 cm^{-1}. Between 1045 cm^{-1} and 1250 cm^{-1} there are some slightly covered peaks due to C–O, C–C, C–H stretching and C–O–C ether stretching.[31]

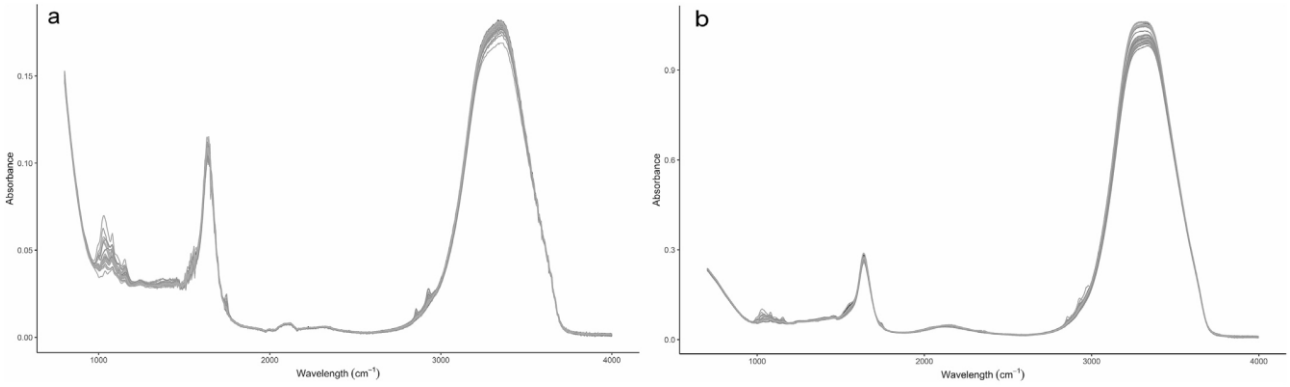

Figure 1 – Spectra of the 46 milk samples. a) portable instrument; b) benchtop instrument

According to literature, the water signal can be removed with different methods [8,9]. However, we observed that there was not a significant change in the model results (prediction error and number of used factors) when removing that region. Therefore, we decide to work with the whole spectra as the input for the PLS and PLS-DA algorithms.

3.2 Prediction of macronutrients

3.2.1 Portable ATR-FT-MIR

Spectra recorded in triplicate for each milk sample were averaged and a PLS regression model was developed and optimized for each of the investigated macronutrient. Concerning the data pretreatments, it is worthwhile mentioning that the application of SNV and MSC did not gave best results in any of the models

in terms of prediction error and number of factors used in the prediction. The best model for each macronutrient was chosen as the best compromise between model complexity, number of selected variables and prediction error. An outlier was eliminated (L23) from all the models. This sample contained dissolved cereals and for this reason, it showed spectroscopic characteristics different from the other investigated samples.

A Savitsky Golay filter (second-order derivative, window size 15 points, polynomial order 2) gave the best results as the smoothing pretreatment when predicting the protein content. Several variable selection methods were tested (Table 1), returning models with RMSECV values ranging between 0.21 and 0.28 g/100 mL, with the number of factors varying between 3 and 4 and the number of selected variables between 77 and 516. The best results were achieved with 3 factors and 94 variables selected with the RC method. The selection includes the spectral zone related to the signals of proteins (1500-1700 cm^{-1}). The model used an extra factor just to model two samples (L20 and L29), which are enriched in proteins. We decided not to eliminate these samples to consider the greatest possible variability of commercial milks. The elimination of them led to a model with one factor less and an error of prediction not considerably different as the one obtained with all the samples.

The best model for the fat content prediction was achieved again when applying a Savitsky Golay filter (second-order derivate window size 11 points, polynomial order 2). After the application of the different variable selection methods, the RMSECV value was between 0.68 and 0.86 g/100mL, with one factor and the number selected variables fluctuating between 85 and 602. A good compromise could be achieved by using 258 variables in the areas of CH_2 symmetric stretching and OH stretching. In this case, samples L7, L8 and L15 were identified as outliers (they emerged as outliers in the influence plot - Q^2 value - and they were badly predicted by the model), showing the ability of the models to detect milk samples fortified with vegetable oils such as sunflower and olive, or oily fish oil.

To predict carbohydrates the best results were obtained by employing a Savitsky Golay filter (second-order derivate, window size 15 points, polynomial order 2). RMSECV values ranged from 0.40 to 0.52 g/100mL, with 2, 3 or 4 factors depending on the variable selection method. The number of selected variables changed between 10 and 344. The model with the best predictive ability was the one with 10 variables selected in the spectroscopic region between 800 and 2250 cm^{-1} and 2 factors. Samples with enrichment in fibers (L7, L8, L15, L16) were detected as outliers as they emerged from the influence plot for both T^2 and Q values. The results are summarized in Table 1.

	Proteins			**Fats**			**Carbohydrates**		
Variable selection method	***Nº selected variables***	***RMSECV (Factors)***	***R^2***	***Nº selected variables***	***RMSECV (Factors)***	***R^2***	***Nº selected variables***	***RMSECV (Factors)***	***R^2***
None	859	0.25(5)	0.9034	859	0.78(1)	0.5918	859	0.49(4)	0.8799
RC	94*	0.22(3) *	0.9563*	124	0.77(1)	0.5986	410	0.46 (4)	0.8930
LW	141	0.26(4)	0.8977	103	0.78(1)	0.5950	72	0.47 (4)	0.8887
VIP	132	0.25(4)	0.9008	118	0.77(1)	0.5996	79	0.46 (4)	0.8947
filter	34	0.25(3)	0.9090	103	0.78(1)	0.5950	40	0.44 (4)	0.9029

BVE	164	0.23(4)	0.9512	158	0.65(1)	0.5893	124	0.48 (4)	0.8868
IPW	90	0.23(4)	0.9534	85	0.80 (2)	0.5089	10*	0.40 (2) *	0.9195*
IPLS.b	516	0.28(3)	0.8810	602	0.78 (1)	0.7290	344	0.47 (4)	0.8907
IPLS.f	172	0.23(3)	0.9176	258*	0.63 (3) *	0.6913*	172	0.52 (3)	0.8627

Table 1 - Results for the different variable selection methods used to improve the prediction ability of milk major components with the portable system. Best models are indicated with *. LW = loading weights, RC = regression coefficients , VIP = Variable Importance in Projection, filter = combination of LW, RC and VIP, BVE = Backward variable elimination PLS , IPW = Iterative predictor weighting PLS, iPLS.b = Interval PLS backward, iPLS.f = Interval PLS forward.

The plots 'predicted vs reference values' of the best models (indicated with * in Table 1) for proteins, fats and carbohydrates built with spectra recorded with the portable ATR instrument are depicted in Figure 2.

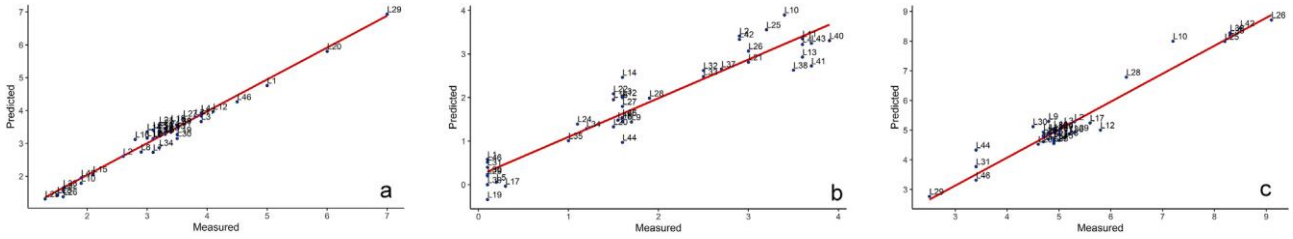

Figure 2 – Predicted vs measured values for the best models with portable ATR-FT-MIR: a) proteins, b) fats, c) carbohydrates

The definitive model validation was carried out with an external validation approach. Samples were then divided into calibration and validation sets using the Kennard-Stone and HONIGS algorithms and trying different proportions of samples between the two sets. The results are summarized in Table 2.

Validation method	Proteins		Fats		Carbohydrates	
	RMSEP (Factors)	R^2 in prediction	RMSEP (Factors)	R^2 in prediction	RMSEP (Factors)	R^2 in prediction
Full cross-validation (LOO)^	0.22 (3)	0.9563	0.63 (3)	0.6913	0.40 (2)	0.9195
CV random^	0.22 (3)	0.9526	0.66 (3)	0.7006	0.40 (2)	0.9167
CAL -TEST (1/2 cal - 1/2 val) – KS	0.18 (3)	0.8502	0.75 (1)	0.5438	0.33 (2)	0.8283
CAL -TEST (2/3 cal - 1/3 val) – KS	0.16 (3)	0.8908	0.52 (3)	0.8173	0.34 (2)	0.8608
CAL -TEST (1/3 cal - 2/3 val) – KS	0.23 (3)	0.8677	0.88 (1)	0.4675	0.40 (2)	0.8924
CAL -TEST (1/2 cal - 1/2 val) – HO	0.20 (3)	0.9135	0.80 (3)	0.5385	0.43 (2)	0.8791
CAL -TEST (2/3 cal - 1/3 val) – HO	0.18 (3)	0.9267	0.77 (3)	0.6875	0.43 (2)	0.9055
CAL -TEST (1/3 cal - 2/3 val) – HO	0.25 (3)	0.9234	0.88 (3)	0.4212	0.46 (4)	0.9015

Table 2 - Results in prediction for different validation methods from portable ATR-MIR spectra after variable selection. KS= CADEX algorithm by Kennard & Stone. HO = HONIGS algorithm.

The results were not significantly different in terms of prediction errors, confirming the robustness of these models. The best compromise for all the considered milk components is the ratio 2/3 cal – 1/3 val, as showed in the table, but it is possible to state that a fraction of 1/3 of the analyzed samples can be considered representative of the commercial milk evaluating an appropriate subset for each type of nutrient. These error values are very satisfactory if considering that the nutrition declaration of nutrient on foods according to Regulation 1169/2011/EC accepts a tolerance of 2g/100g for proteins and carbohydrates and 1.5g/100g for

fat [33]. The prediction errors obtained are thus suitable for quality control purposes of commercial milk samples. Therefore, it is worthwhile to note that a contribution in the prediction errors may be due to the allowed tolerance values declared in the milk labels.

3.2.2 Benchtop ATR-FT-MIR

Using the benchtop instrument milk analyses were conducted following two different strategies. The first strategy was carried out to make an immediate comparison between the portable system and the benchtop one. In this case, samples were analyzed with no pretreatment. However, even if the literature reported the homogeneity of commercial samples [34], in the first analyzed samples some scattering effects seemed to occur, probably ascribable to undissolved milk components like fats. A second data analysis strategy was then optimized.

Milk samples were pretreated with the aim to homogenize them as much as possible without any chemical damage. Heating and ultrasonication [35] were considered for this purpose and combinations of different temperatures and different sonication times were studied. Two temperature values were considered: the ambient one (25°C), and 40 °C to make our results comparable to those found in the literature [16, 17]. The higher temperature value tested corresponds to that of melting for milk fat but does not deteriorate proteins, which denature above 70° C. The ultrasonication time was investigated up to a maximum of 15 minutes so as not to exceed the reasonable limits of the method usability because of the time that the pretreatment takes. Ultrasonication was carried out in a thermostatic bath allowing maintaining the temperature controlled at 25°C or 40°C.

Pretreatment-free milk samples

Different signal pretreatments were tested for each of the macronutrients. The best results were obtained applying second-order Savitsky Golay derivatives with a second-order polynomial and a window width of 11 for all the major components. SNV and MSC were tested but they did not significantly change the results. PLS models obtained applying different variable selection methods returned the results summarized in Table 3.

	Proteins			**Fats**			**Carbohydrates**		
Variables selection method	***Nº selected variables***	***RMSECV (Factors)***	***R^2***	***Nº selected variables***	***RMSECV (Factors)***	***R^2***	***Nº selected variables***	***RMSECV (Factors)***	***R^2***
None	1369	0.36 (5)	0.8857	1369	0.48 (1)	0.8527	1369	0.57 (2)	0.8864
RC	115	0.26 (3)	0.9417	44	0.50 (1)	0.8429	90*	0.29 (4)*	0.9488*
LW	327	0.28 (4)	0.9298	191	0.48 (1)	0.8563	349	0.35 (5)	0.9118
VIP	81	0.23 (5)	0.9515	44	0.50 (1)	0.8429	191	0.36 (5)	0.9378
filter	33 *	0.24 (3) *	0.9501*	44	0.50 (1)	0.8429	46	0.34 (4)	0.9458
BVE	82	0.28 (4)	0.9322	61	0.49 (1)	0.8459	212	0.46 (4)	0.8803
IPW	10	0.54 (3)	0.7474	10*	0.47 (1)*	0.8532*	10	0.71 (3)	0.7207
IPLS.b	274	0.22 (3)	0.9579	411	0.41 (3)	0.9000	411	0.42 (4)	0.9010
IPLS.f	411	0.21 (4)	0.9610	685	0.40 (4)	0.9000	137	0.36 (4)	0.9290

Table 3 - Results for different variable selection methods to improve the prediction ability of milk major components with the benchtop system. Best models are indicated with *. LW = loading weights, RC = regression coefficients, VIP = Variable Importance in Projection, filter = combination of LW, RC and VIP, BVE = Backward variable elimination PLS , IPW = Iterative predictor weighting PLS, iPLS.b = Interval PLS backward, iPLS.f = Interval PLS forward.

It could be noticed that models built on data recorded with the benchtop instrument used less variables than the ones built on the portable equipment. This could be ascribed to a different quality of the signal (e.g. more sensitivity, less noise), which is better for the benchtop spectrometer, as expected.

The PLS models for carbohydrates prediction gave us further information. The interpretation of the influence plot allowed identifying different samples of milk (L8, L15, L16) containing vegetable oils such as olive and sunflower oil (plot not shown). Moreover, milk samples containing vegetable fibers showed a residual value higher than the others in the validation step (data not shown). Thus, the model allowed identifying some peculiar characteristics in milk samples.

The plots 'predicted vs reference values' of the best models (indicated with * in Table 3) for proteins, fats and carbohydrates built with spectra recorded on samples without any pretreatment with the benchtop ATR instrument are depicted in Figure 3.

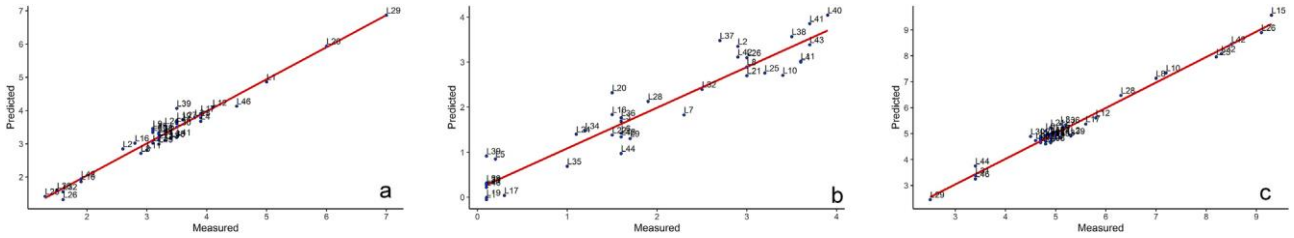

Figure 3 – Predicted vs measured values for the best models with benchtop ATR-FT-MIR on samples without any pretreatment: a) proteins, b) fats, c) carbohydrates

The models that gave the best results in terms of low prediction error and a smaller number of factors (indicated in Table 3 with *) were then externally validated as previously described. Results are reported in Table 4.

Validation method	Proteins		Fats		Carbohydrates	
	RMSEP (Factors)	R^2 in prediction	RMSEP (Factors)	R^2 in prediction	RMSEP (Factors)	R^2 in prediction
Full cross-validation (LOO)^	0.24 (3)	0.9501	0.47 (1)	0.8532	0.30 (4)	0.9488
CV random^	0.23 (3)	0.9503	0.40 (1)	0.8963	0.30 (4)	0.9501
CAL -TEST (1/2 cal - 1/2 val) – KS	0.22 (3)	0.8664	0.41 (1)	0.9023	0.24 (4)	0.7118
CAL -TEST (2/3 cal - 1/3 val) – KS	0.24 (3)	0.8571	0.41 (1)	0.9443	0.26 (4)	0.7306
CAL -TEST (1/3 cal - 2/3 val) – KS	0.29 (2)	0.7995	0.47 (1)	0.8735	0.25 (4)	0.9181
CAL -TEST (1/2 cal - 1/2 val) – HO	0.25 (3)	0.9448	0.45 (1)	0.8982	0.26 (4)	0.8959
CAL -TEST (2/3 cal - 1/3 val) – HO	0.19 (4)	0.9838	0.47 (1)	0.8662	0.28 (4)	0.8026
CAL -TEST (1/3 cal - 2/3 val) – HO	0.27 (3)	0.9441	0.42 (1)	0.8986	0.43 (4)	0.8197

Table 4 - Results in prediction for different validation methods from benchtop ATR-MIR spectra after variable selection. KS= CADEX algorithm by Kennard & Stone. HO = HONIGS algorithm.

The different validation methods essayed did not lead to substantial differences in the prediction results, supporting the idea of robustness of these models. Prediction model performances for proteins and fats varied less than those of carbohydrates. It is worthwhile noting that the selection of 1/2, 2/3 or 1/3 of samples for the calibration set did not affect significantly the prediction errors for all the major components, as observed for the models built with the data obtained by the hand-held spectrometer. Considering that the purchased milk samples were chosen covering almost all the variability in commercially available Italian, Spanish and Swiss samples, models suggested that most probably only a fraction of these samples is necessary to explain all the major variations in commercial milk samples. In other words, we could be confident that a good representativeness of the commercial milk was achieved. Also in this case, results are very satisfactory with respect to the tolerances accepted by the European regulation about nutrient values.

Pretreated milk samples

5 milk samples with heterogeneous characteristics were chosen to optimize pretreatment conditions, that is, heating and sonication time. Times of 1, 5, 8 and 15 minutes and temperatures of 25°C and 40°C were considered, and all the combination of these conditions were tested.

Spectra were acquired and PLS models were then built on the milk samples, one for each combination of pretreatment conditions (8 PLS models were built).

After that, RMSECV using venetian blind algorithm (putting into the segment the replicates of the same milk) were calculated. The average errors (expressed as percentage) obtained ranged from 9.3% to 32% for fats, from 4% to 7.3% for proteins, and from 1.9% to 4.5% for carbohydrates. Considering this error ranges, we focused the attention on fats prediction as it showed the higher variability.

An MLR (Multiple Linear Regression) model was then calculated using temperature, time (scaled between -1 and +1), and their higher terms as **X** matrix and RMSECV errors (expresses as percentage) as the **y** vector. The resulting MLR equation follows:

$$\mathbf{y} = 23.39 - 1.47 \times \text{Time} - 3.82 \times \text{Temperature} - 1.35 \times \text{Time} \times \text{Temperature} - 6.04 \times \text{Time}^2$$

The results are displayed on the response surface (**Error! No s'ha trobat l'origen de la referència.**). An interaction between the two investigated factor is clearly visible. The best results seemed to arise for a sonication time of 15 minutes and a temperature of 40 °C.

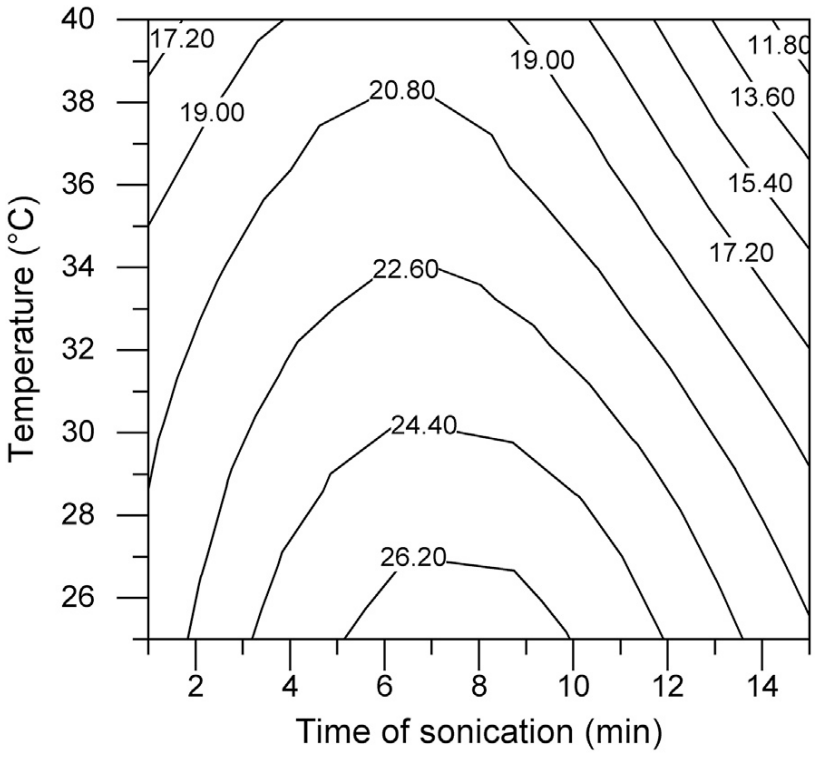

Figure 4 – Response surface: MLR model built with y =average percentage error (RMSECV) for fats prediction.

After being pretreated under these conditions, all the milk samples (46) were then analyzed, and PLS models were calculated. The best models were obtained with second-order Savitsky Golay derivatives with a second-order polynomial and a window width of 11 as data pretreatment for all the investigated properties. Different variable selection methods were employed, and the best results are summarized in Table 5.

	Proteins			Fats			Carbohydrates		
Variables selection method	***Nº selected variables***	***RMSECV (Factors)***	***R^2***	***Nº selected variables***	***RMSECV (Factors)***	***R^2***	***Nº selected variables***	***RMSECV (Factors)***	***R^2***
None	1369	0.41(5)	0.8005	1369	0.39(1)	0.8940	1369	0.44 (5)	0.8915
RC	49*	0.18 (2)*	0.9693*	73	0.37 (1)	0.9017	247	0.29(4)	0.9545
LW	136	0.40 (5)	0.8807	70*	0.37 (1)*	0.9020*	372	0.44 (4)	0.8946
VIP	139	0.30 (5)	0.9464	119	0.37 (1)	0.9003	195	0.33 (5)	0.9399
filter	12	0.32 (3)	0.9271	70	0.37 (1)	0.9020	42*	0.29 (4)*	0.9515*
BVE	324	0.30 (5)	0.9386	45	0.38 (1)	0.9012	396	0.35 (5)	0.9026
IPW	78	0.38 (5)	0.8683	156	0.37 (2)	0.9015	10	0.76 (1)	0.6923
IPLS.b	274	0.17 (5)	0.9760	822	0.37(2)	0.8995	1096	0.29 (5)	0.9520
IPLS.f	274	0.17 (5)	0.9770	548	0.37 (1)	0.8980	548	0.30 (5)	0.9590

Table 5 - Results for different variable selection methods to improve the prediction ability of milk major components with the benchtop system. Best models are indicated with *. LW = loading weights, RC = regression coefficients , VIP = Variable Importance in Projection, filter = combination of LW, RC and VIP, BVE = Backward variable elimination PLS , IPW = Iterative predictor weighting PLS, iPLS.b = Interval PLS backward, iPLS.f = Interval PLS forward.

The plots 'predicted vs reference responses' of the best models (indicated with * in Table 5) for proteins, fats and carbohydrates built with spectra recorded on pretreated milk samples with the benchtop ATR instrument are depicted in Figure 5.

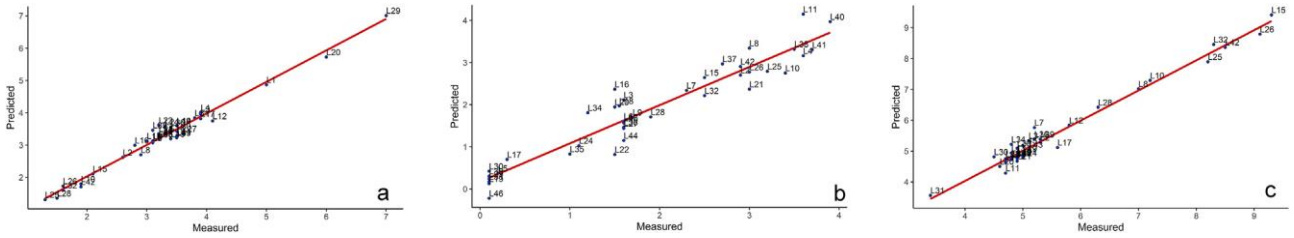

Figure 5 – Predicted vs measured values for the best models with benchtop ATR-FT-MIR on pretreated samples: a) proteins, b) fats, c) carbohydrates

The best models (labeled in Table 5 with *), considered as the ones with the best compromise between complexity, predictive capacity, and number of variables, where then validated with external validation considering again different proportions between calibration and validation sets. Results are reported in Table 6.

Validation method	Proteins		Fats		Carbohydrates	
	RMSEP (Factors)	R^2 in prediction	RMSEP (Factors)	R^2 in prediction	RMSEP (Factors)	R^2 in prediction
Full cross-validation (LOO)^	0.18 (2)	0.9693	0.37 (1)	0.9020	0.28 (4)	0.9515
CV random^	0.18 (2)	0.9705	0.38 (1)	0.8937	0.28 (4)	0 .9568
CAL -TEST (1/2 cal - 1/2 val) – KS	0.21 (2)	0.9374	0.34 (1)	0.9580	0.29 (4)	0.9089
CAL -TEST (2/3 cal - 1/3 val) – KS	0.23 (2)	0.9424	0.31 (1)	0.9484	0.28 (4)	0.9527
CAL -TEST (1/3 cal - 2/3 val) – KS	0.22 (2)	0.9222	0.38 (1)	0.9300	0.29 (4)	0.9034
CAL -TEST (1/2 cal - 1/2 val) – HO	0.20 (2)	0.9523	0.38 (1)	0.9170	0.32 (4)	0.9725
CAL -TEST (2/3 cal - 1/3 val) – HO	0.18 (2)	0.9517	0. 28(1)	0.9541	0.28 (4)	0.9772
CAL -TEST (1/3 cal - 2/3 val) – HO	0.20 (2)	0.9382	0.44 (1)	0.9032	0.40 (4)	0.9572

Table 6 - Results in prediction for different validation methods from benchtop ATR-MIR spectra after variable selection. KS= CADEX algorithm by Kennard & Stone. HO = HONIGS algorithm.

Once again, there are no noticeable differences in predictive abilities when changing the ratio of the test set: therefore, one-third of the samples could be considered representative of the commercial milk samples. And as in the previous cases, results are very satisfactory if compared with the tolerances admitted by the European regulation about nutrient values.

Heating and sonication pretreatments, as expected, allowed to obtain better results in the case of proteins (models required one factor less to obtain slightly lower prediction errors) and fats, while it seemed that did not affect the prediction of carbohydrates. This is probably because heating and sonication help to homogenize fat micelles and protein globules. Since carbohydrates do not organize in large macrocompounds, the effect of this pre-treatments is negligible. The results obtained with the benchtop instrument for all macronutrients in this work are numerically comparable to those obtained in a previous work[34], in which performances are about the same but with a higher number of factors.

The outliers identified in the previous models (L16 and L23, L7, L8 and L15 - sample with vegetable oils) were not detected in these models but it is it is worthwhile to note that samples enriched in vitamin B6 and B12 (L46), showed a Q value higher than the others.

Comparison between the results obtained by portable and benchtop ATR-FT-MIR

To check if the results obtained by the two ATR-FT-MIR instruments (with and without sample pretreatment in the case of the benchtop instrument) were comparable, a linear regression approach was used. The theoretical regression line obtained if the results in comparison were identical should have intercept=0 and slope=1. Therefore, if the two methods in comparison provide comparable results, the joint confidence interval of the intercept and the slope of the regression line between the two methods has to include the theoretical point (0,1).[36]

We performed the comparison between the macronutrient values obtained with the portable and the benchtop ATR-FT-MIR instruments, both with and without sample pretreatment. The results are comparable for all the combinations. As an example, Figure 6 shows the comparison between the portable ATR-FT-MIR data and the benchtop ones obtained on milk without sample pretreatment. Since the theoretical point (0,1) is within the limits of the joint confidence region for an α significance value of 5% in all cases, we can conclude that the results are comparable.

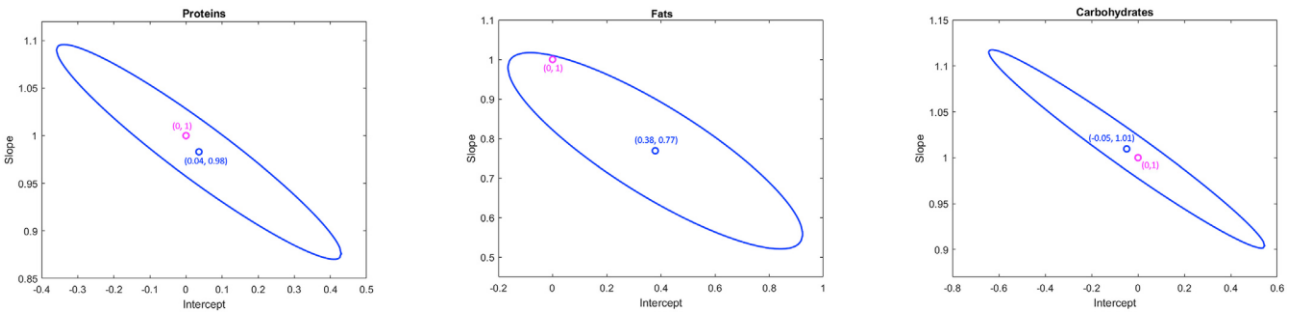

Figure 6 - Joint confidence region (α=5%) of the intercept and the slope of the regression line for the data obtained by the portable and the benchtop (without sample pretreatment) ATR-FT-MIR instruments

Comparison between the results obtained by ATR-FT-MIR and the reference methods

To compare the precision between our method and the reference methods, 5 milk samples with heterogeneous characteristics were measured in triplicate with the portable and the benchtop (with and without sample pretreatment) ATR-FT-MIR instruments. Prediction was carried out to evaluate the standard deviation of the results (HONIGS algorithm, 2/3 cal – 1/3 val model). The results, expressed as the average of the 5 standard deviations obtained, are summarized in table TOT. These results were compared to those reported in the ISO 8968-1:2004(E) [RIF] for proteins, ISO1211:2010(E) [RIF] for fats, and ISO/DIS 22184:2019(E) for carbohydrates contents (expressed as sugars) [RIF] (repeatability values in all cases), obtained in all cases by interlaboratory circuits.

Method	Proteins	Fats	Carbohydrates
Reference	0.038%	0.031-0.043% (depending on the fat content)	0.026% and 0.15% (depending on the type of sugar)
Portable ATR-FT-MIR	0.18%	0.36%	0.40%
Benchtop ATR-FT-MIR without sample pretreatment	0.24%	0.072%	0.050%
-Benchtop ATR-FT-MIR with sample pretreatment	0.039%	0.058%	0.15%

Table 7 - Comparison between the results obtained by ATR-FT-MIR instruments and the reference methods

Considering that the results reported for the reference methods are repeatability values obtained through interlaboratory circuits, the standard deviation obtained by the portable ATR-FT-MIR can be considered acceptable. Better results are obtained by the benchtop instrument, especially for carbohydrates when the samples were not pretreated, and for proteins and fats in the case of pretreated samples.

3.3 Qualitative prediction of lactose

In order to predict the presence or absence of lactose in milk samples, PLS-DA models were built with the spectra acquired with the three different configurations (portable instrument, benchtop instrument with and without milk sample pretreatment). A dummy vector **y** was built assigning 0 to lactose free milks (as declared by the producer) and 1 to the remaining milk samples. A Savitsky-Golay filter (second-degree derivative with a window width of 11 and a second-degree polynomial) was chosen as the optimal data pretreatment for all the classification models. Sensitivity, specificity and the total amount of classification errors were considered in order to choose the best classification model. PLS-DA models were calculated with PLS Toolbox 8.7 (R8.7.1) and the threshold was determined as implemented in the algorithm.

3.3.1 Portable ATR-FT-MIR

Considering all the spectroscopic signal, specificity and selectivity of 1 for both classes were obtained using a model with 3 factors. However, it was possible to obtain the same good result even with a small number of variables and with less complex models with 2 latent variables. The VIP algorithm allowed to select 120 variables while the filter algorithm selected only 25 variables around 1050 cm^{-1}.

3.3.2 Benchtop ATR-FT-MIR

Without Pretreatment milk samples

Different models on benchtop instrument spectra were calculated to classify milk samples according to the absence of lactose declared by the producer. Several variable selection methods were tested, and the filter methods based on regression coefficients (RC) allowed to obtain specificity and sensitivity of 1 for both

classes with 1 factor and using the spectroscopic region between 1000 and 1250 cm^{-1}. The model identified as outlier the sample containing dissolved fibers.

Pretreated milk samples

The use of the whole spectra in the PLS-DA model allowed selectivity and specificity values of 1 for classification with 2 factors. With the RC filter method, it was possible to reach the same performance with 58 selected variables.

Two outliers emerged from the model and were subsequently removed: one infant milk mixed with cereals (L23) and one enriched with magnesium and vitamins B_6 and B_{12}(L46).

4. CONCLUSIONS

From the results obtained in this study, we concluded that the handheld ATR-FT-MIR instrument provided comparable errors in prediction (external validation) to the ones obtained by the benchtop instrument regardless of whether the sample has been pretreated or not. Moreover, both instruments allowed a correct sample classification regarding the presence/absence of lactose.

Models built on the benchtop spectrometer data allowed identifying samples quite different from the common ones. Regarding fats, the pretreatment with heating and sonication lead to better results, indicating that commercial milk is not so homogeneous as declared in the literature.

The handheld ATR-FT-MIR instrument proved to be a very good alternative to the benchtop instrument. Prediction errors obtained for all the macronutrients are very satisfactory if compared with the tolerances accepted by the European nutrition declaration of nutrient on foods, making ATR-FT-MIR spectroscopy coupled with chemometrics a good tool for commercial milk quality control purposes. Moreover, the analytical protocol developed and optimized in this project allowed to reach the best results for the investigated macronutrients ever published in the literature in our knowledge for commercially available milk sample analyzed with ATR-FT-MIR spectroscopy.

Aknowledgements

We thank the Ministerio de Economía y Competitvidad and Fondo Europeo de Desarrollo Regional (FEDER) (project AGL2015-70106-R, AEI-FEDER,UE) and the Ministerio de Ciencia, Innovación e Universidades and European Regional Development Fund (Project CTQ2016-77128-R) for the financial support given.

REFERENCES

[1] D.I. Ellis, V.L. Brewster, W.B. Dunn, J.W. Allwood, A.P. Golovanov, R. Goodacre, Fingerprinting food: Current technologies for the detection of food adulteration and contamination, Chem. Soc. Rev.

41 (2012) 5706–5727. https://doi.org/10.1039/c2cs35138b.

[2] J.J. Roberts, D. Cozzolino, An Overview on the Application of Chemometrics in Food Science and Technology—An Approach to Quantitative Data Analysis, Food Anal. Methods. 9 (2016) 3258–3267. https://doi.org/10.1007/s12161-016-0574-7.

[3] Y. Etzion, R. Linker, U. Cogan, I. Shmulevich, Determination of Protein Concentration in Raw Milk by Mid-Infrared Fourier Transform Infrared/Attenuated Total Reflectance Spectroscopy, J. Dairy Sci. 87 (2004) 2779–2788. https://doi.org/10.3168/jds.S0022-0302(04)73405-0.

[4] 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. 102 (2007) 621–640. https://doi.org/10.1016/j.foodchem.2006.05.042.

[5] M. Li, J. Chen, J. Xu, S. Fu, H. Gong, Determination of Lactose in Milk by Raman Spectroscopy, Anal. Lett. 48 (2015) 1333–1340. https://doi.org/10.1080/00032719.2014.979358.

[6] M. Bassbasi, S. Platikanov, R. Tauler, A. Oussama, FTIR-ATR determination of solid non fat (SNF) in raw milk using PLS and SVM chemometric methods, Food Chem. 146 (2014) 250–254. https://doi.org/10.1016/j.foodchem.2013.09.044.

[7] H. Soyeurt, P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, N. Gengler, Estimating fatty acid content in cow milk using mid-infrared spectrometry, J. Dairy Sci. 89 (2006) 3690–3695. https://doi.org/10.3168/jds.S0022-0302(06)72409-2.

[8] J. Cavaglia, B. Giussani, M. Mestres, M. Puxeu, O. Busto, J. Ferré, R. Boqué, Early detection of undesirable deviations in must fermentation using a portable FTIR-ATR instrument and multivariate analysis, J. Chemom. (2019) 1–11. https://doi.org/10.1002/cem.3162.

[9] M. Manfredi, E. Robotti, F. Quasso, E. Mazzucco, G. Calabrese, E. Marengo, Fast classification of hazelnut cultivars through portable infrared spectroscopy and chemometrics, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 189 (2018) 427–435. https://doi.org/10.1016/j.saa.2017.08.050.

[10] J. Vongsvivut, P. Heraud, W. Zhang, J.A. Kralovec, D. McNaughton, C.J. Barrow, Rapid Determination of Protein Contents in Microencapsulated Fish Oil Supplements by ATR-FTIR Spectroscopy and Partial Least Square Regression (PLSR) Analysis, Food Bioprocess Technol. 7 (2014) 265–277. https://doi.org/10.1007/s11947-013-1122-8.

[11] E. Birkel, L. Rodriguez-Saona, Application of a portable handheld infrared spectrometer for quantitation of trans fat in edible oils, JAOCS, J. Am. Oil Chem. Soc. 88 (2011) 1477–1483. https://doi.org/10.1007/s11746-011-1814-z.

[12] M.M. Mossoba, C.T. Srigley, S. Farris, J.K.G. Kramer, S. Chirtel, J. Rader, Evaluation of the Performance of a Portable Mid-Infrared Analyzer for the Rapid Determination of Total Trans Fat in

Fast Food, JAOCS, J. Am. Oil Chem. Soc. 91 (2014) 1651–1663. https://doi.org/10.1007/s11746-014-2521-3.

[13] H. Ayvaz, A. Sierra-Cadavid, D.P. Aykas, B. Mulqueeney, S. Sullivan, L.E. Rodriguez-Saona, Monitoring multicomponent quality traits in tomato juice using portable mid-infrared (MIR) spectroscopy and multivariate analysis, Food Control. 66 (2016) 79–86. https://doi.org/10.1016/j.foodcont.2016.01.031.

[14] RStudio Team, RStudio: Integrated Development for R., (2018). http://www.rstudio.com/.

[15] B.A. Hanson, ChemoSpec: Exploratory Chemometrics for Spectroscopy, (2019) R package version 5.0.229. https://cran.r-project.org/package=ChemoSpec.

[16] Bryan A. Hanson, ChemoSpecUtils: Functions Supporting Packages ChemoSpec and ChemoSpec2D, (2019) R package version 0.2.211. https://cran.r-project.org/package=ChemoSpecUtils.

[17] Peter Filzmoser and Kurt Varmuza, chemometrics: Multivariate Statistical Analysis in Chemometrics, (2017) R package version 1.4.2. https://cran.r-project.org/package=chemometrics.

[18] Ron Wehrens, Chemometrics With R: Multivariate Data Analysis in the Natural Sciences and Life Sciences, (2011). https://doi.org/10.1007/978-3-642-17841-2.

[19] S. Kucheryavskiy, mdatools: Multivariate Data Analysis for Chemometrics, (2018) R package version 0.9.1. https://cran.r-project.org/package=mdatools.

[20] T. Mehmood, K.H. Liland, L. Snipen, S. Sæbø, A review of variable selection methods in Partial Least Squares Regression, Chemom. Intell. Lab. Syst. 118 (2012) 62–69. https://doi.org/10.1016/j.chemolab.2012.07.010.

[21] A.S. and L. Ramirez-Lopez, An introduction to the prospectr package, in: R Packag. Vignette, 2013: p. R package version 0.1.3.

[22] Chemometric Agile Tools. http://www.gruppochemiometria.it/index.php/software/19-download-the-r-based-chemometric-software, (n.d.). http://www.gruppochemiometria.it/index.php/software/19-download-the-r-based-chemometric-software.

[23] S. Wold, M. Sjostrom, L. Eriksson, S. Sweden˚˚, PLS-regression, a basic tool of chemometrics, Chemom. Intell. Lab. Syst. 58 (2001) 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1.

[24] R.G. Brereton, G.R. Lloyd, Partial least squares discriminant analysis: Taking the magic away, J. Chemom. 28 (2014) 213–225. https://doi.org/10.1002/cem.2609.

[25] J.A. Fernández Pierna, O. Abbas, V. Baeten, P. Dardenne, A Backward Variable Selection method for PLS regression (BVSPLS), Anal. Chim. Acta. 642 (2009) 89–93. https://doi.org/10.1016/j.aca.2008.12.002.

[26] M. Forina, C. Casolino, C. Pizarro Millan, Iterative predictor weighting (IPW) PLS: A technique for the elimination of useless predictors in regression problems, J. Chemom. 13 (1999) 165–184. https://doi.org/10.1002/(SICI)1099-128X(199903/04)13:2<165::AID-CEM535>3.0.CO;2-Y.

[27] L. Nørgaard, A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, S.B. Engelsen, Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy, Appl. Spectrosc. 54 (2000) 413–419. https://doi.org/10.1366/0003702001949500.

[28] R.W.K. and L.A. Stone, Computer Aided Design of Experiments, Technometrics. 11 (1969) 137–148. https://doi.org/10.1007/978-1-349-95810-8_109.

[29] F. Westad, F. Marini, Validation of chemometric models - A tutorial, Anal. Chim. Acta. 893 (2015) 14–24. https://doi.org/10.1016/j.aca.2015.06.056.

[30] D.E. Honigs, G.M. Hieftje, H.L. Mark, T.B. Hirschfeld, Unique-Sample Selection via Near- Infrared Spectral Subtraction, (1985) 2299–2303. https://doi.org/10.1021/ac00289a029.

[31] C. Grelet, J.A. Fernández Pierna, P. Dardenne, V. Baeten, F. Dehareng, Standardization of milk mid-infrared spectra from a European dairy network, J. Dairy Sci. 98 (2015) 2150–2160. https://doi.org/10.3168/jds.2014-8764.

[32] H.J. Luinge, E. Hop, E.T.G. Lutz, J.A. van Hemert, E.A.M. de Jong, Determination of the fat, protein and lactose content of milk using Fourier transform infrared spectrometry, Anal. Chim. Acta. 284 (1993) 419–433. https://doi.org/10.1016/0003-2670(93)85328-H.

[33] E. Commission, I. Disclaimer, December 2012 GUIDANCE DOCUMENT FOR COMPETENT AUTHORITIES FOR THE CONTROL OF COMPLIANCE WITH EU LEGISLATION ON : Regulation (EU) No 1169 / 2011 of the European Parliament and of the Council of 25 October 2011 on the provision of food information to cons, Eur. Comm. 1 (2012) 1–15. https://www.fsai.ie/uploadedFiles/guidance_tolerances_december_2012.pdf.34

[34] F.A. Iñón, S. Garrigues, M. De La Guardia, Nutritional parameters of commercially available milk samples by FTIR and chemometric techniques, Anal. Chim. Acta. 513 (2004) 401–412. https://doi.org/10.1016/j.aca.2004.03.014.

[35] I. Bosiljkov, T.*, Tripalo, B., Brnčić, M., Ježek, D., Karlović, S. and Jagušt, Influence of high intensity ultrasound with different probe diameter on the degree of homogenization (variance) and physical properties of cow milk, African J. Biotechnol. 10 (2011) 34–41. https://doi.org/10.5897/AJB10.887.

[36] J. Mandel, F.J. Linnig, Study of Accuracy in Chemical Analysis Using Linear Calibration Curves, Anal. Chem. 29 (1957) 743–749. https://doi.org/10.1021/ac60125a002.

SUPPLEMENTARY MATERIALS

Milk samples and their nutritional information:

Labels	Proteins (g/100 mL)	Fats (g/100 mL)	Carbohydrates (g/100 mL)	With Lactose	For childhood
L1	5	0.1	4.7	NO	NO
L2	2.6	2.9	5.3	YES	YES
L3	3.9	1.6	5.1	YES	NO
L4	3.9	3.6	4.7	YES	NO
L5	3.3	0.2	4.8	YES	NO
L6	3.1	1.6	4.8	YES	NO
L7	3.1	2.3	5.2	YES	NO
L8	2.9	3	7	YES	YES
L9	3.1	1.7	4.8	YES	NO
L10	1.9	3.4	7.2	YES	YES

L11	3	3.6	4.7	YES	NO
L12	4.1	1.6	5.8	NO	NO
L13	3.3	3.6	4.8	YES	NO
L14	3	1.6	4.9	YES	NO
L15	2.1	2.5	9.3	YES	YES
L16	2.8	1.5	4.6	YES	NO
L17	3.9	0.3	5.6	YES	NO
L18	3.4	1.55	4.9	YES	NO
L19	3.5	0.1	5	YES	NO
L20	6	1.5	4.6	YES	NO
L21	3.2	3	4.9	NO	NO
L22	3.2	1.5	4.7	YES	NO
L23	2.9	2.8	13.5	YES	YES
L24	3.3	1.1	5	NO	NO
L25	1.3	3.2	8.2	YES	YES
L26	1.6	3	9.1	YES	YES
L27	3.6	1.6	4.9	NO	NO
L28	1.5	1.9	6.3	YES	YES
L29	7	0.1	2.5	NO	NO
L30	3.5	0.1	4.5	YES	NO
L31	3.8	0.1	3.4	NO	NO
L32	1.6	2.5	8.3	NO	YES
L33	1.6	2.5	8.3	NO	YES
L34	3.2	1.2	4.8	NO	NO
L35	3.1	1	4.9	NO	NO
L36	3.4	1.6	5.2	YES	NO
L37	3.2	2.7	4.9	YES	NO
L38	3.2	3.5	4.9	YES	NO
L39	3.5	0.1	5.3	YES	NO
L40	3.2	3.9	4.9	YES	NO
L41	3.5	3.7	5	YES	NO
L42	1.9	2.9	8.5	YES	YES
L43	3.4	3.7	5.1	YES	NO
L44	3.5	1.6	3.4	NO	NO
L45	3.2	1.6	5	YES	NO
L46	4.5	0.1	3.4	NO	NO

514

515