Prediction of fatty acid profiles in cow, ewe, and goat milk by mid-infrared spectrometry

M. Ferrand-Calmels, Isabelle Palhiere Palhière, M. Brochard, O. Leray, Jean-Michel Astruc, Marie-Rose Aurel, Sarah Barbey, Frédéric F. Bouvier, P. Brunschwig, Hugues Caillat, et al.

Prediction of fatty acid profiles in cow, ewe, and goat milk by mid-infrared spectrometry

M. Ferrand-Calmels,*[1] I. Palhière,† M. Brochard,* O. Leray,‡ J. M. Astruc,* M. R. Aurel,§ S. Barbey,#
F. Bouvier,‖ P. Brunschwig,* H. Caillat,† M. Douguet,* F. Faucon-Lahalle,*¶ M. Gelé,* G. Thomas,*
J. M. Trommenschlager,** and H. Larroque†
*Institut de l'Elevage, 149 rue de Bercy, F-75595 Paris cedex 12, France
†INRA, UR0631, Station d'Amélioration Génétique des Animaux, F-31326 Castanet-Tolosan cedex, France
‡Actilait, Rue de Versailles, F-39800 Poligny, France
§INRA, UE 0321, Domaine expérimental de La Fage, F-12250 Roquefort-sur-Soulzon, France
#INRA, UE 0326, Domaine expérimental du Pin-au Haras, F-61310 Exmes, France
‖INRA, UE 0332, Domaine expérimental de Bourges-La Sapinière, 18390 Osmoy, France
¶CNIEL, 42 rue de Châteaudun, 75314 Paris cedex 09, France
**INRA, UR 0055, Agro-Systèmes Territoires Ressources, 662 avenue Louis Buffet, F-88500 Mirecourt, France

ABSTRACT

Mid-infrared (MIR) spectrometry was used to estimate the fatty acid (FA) composition in cow, ewe, and goat milk. The objectives were to compare different statistical approaches with wavelength selection to predict the milk FA composition from MIR spectra, and to develop equations for FA in cow, goat, and ewe milk. In total, a set of 349 cow milk samples, 200 ewe milk samples, and 332 goat milk samples were both analyzed by MIR and by gas chromatography, the reference method. A broad FA variability was ensured by using milk from different breeds and feeding systems. The methods studied were partial least squares regression (PLS), first-derivative pretreatment + PLS, genetic algorithm + PLS, wavelets + PLS, least absolute shrinkage and selection operator method (LASSO), and elastic net. The best results were obtained with PLS, genetic algorithm + PLS and first derivative + PLS. The residual standard deviation and the coefficient of determination in external validation were used to characterize the equations and to retain the best for each FA in each species. In all cases, the predictions were of better quality for FA found at medium to high concentrations (i.e., for saturated FA and some mono-unsaturated FA with a coefficient of determination in external validation >0.90). The conversion of the FA expressed in grams per 100 mL of milk to grams per 100 g of FA was possible with a small loss of accuracy for some FA.

Key words: milk, fatty acid, mid-infrared spectrometry

INTRODUCTION

Over the last few years, the FA content and quality of human foods has become a major nutritional topic. Evidence of this can be found in the latest French nutritional guidelines that indicate the recommended intake of individual FA in a diversified diet (ANSES, 2011a). The dairy industry has to face 2 major issues regarding these new recommendations: (1) identifying methods to adapt milk FA composition to the consumers' demands and (2) finding tools to precisely characterize the FA composition of milk.

With regard to the first issue, previous studies have shown that feeding management (Chilliard et al., 2007; Raynal-Ljutovac et al., 2008; Coppa et al., 2013) and genetic selection (Arnould et Soyeurt., 2009; Schennink et al., 2009; Stoop et al., 2009) can be efficient tools to alter the FA composition of milk. However, such strategies can only effectively be organized for entire populations if reliable large-scale measuring techniques are available.

As to finding tools to determine the precise FA composition of milk, mid-infrared (**MIR**) spectrometry can be classically and efficiently used to analyze food products. This technology is already used in dairy farming to measure the fat and protein contents for purposes of milk payment, herd management, and genetic selection. In France, these analyses are performed frequently (once per month for all animals of a herd in dairy cattle or goats managed with the usual milk recording design) at a large scale by laboratories located throughout the country. More recently, this process has been extended to include the analysis of milk FA (Soyeurt et al., 2006, 2011), especially in dairy cattle. Most often, FA composition is predicted using MIR spectra with partial least squares regression (**PLS**). Although it has been reported that preprocessing before PLS regression im-

proves calibration equations (Soyeurt et al., 2011), the possible benefit of wavelength selection before applying PLS regression has not been addressed. It has been suggested that wavelength selection before PLS regression might improve calibration equations and provide good results in various situations (Leardi et al., 1992; Spiegelman et al., 1998). It can be performed using different methods: genetic algorithms (Leardi et al., 1992), wavelet decomposition (Mallat, 2008), or penalization methods such as the elastic net, which is often used on genomic data (Croiseau et al., 2011). However, no comparison of these methods is currently available.

Equations for goat milk were developed using near-infrared spectroscopy (**NIRS**; Andueza et al., 2013), but no calibration equations for predicting milk FA using MIR spectrometry have been published so far for ewe and goat milk. Within the framework of the PhénoFinlait program (Faucon-Lahalle et al., 2009), a large-scale French research and development project aiming at monitoring the composition of cattle, sheep, and goat milk, calibration equations were developed using different mathematical approaches to estimate milk FA composition in French herds using MIR spectrometry.

The objectives of this study were (1) to compare different statistical approaches with wavelength selection for the prediction of milk composition using MIR spectra, (2) to establish whether the use of MIR spectrometry for predicting the FA composition of cow milk could be put into practice in French breeding and feeding systems, and (3) to develop the first referenced equations for FA in goat and ewe milk.

MATERIALS AND METHODS

Experimental Design

Samples were mainly collected from the Institut National de la Recherche Agronomique (**INRA**) experimental farms (Table 1). These experimental farms breed ruminants of the most widespread breeds in France in conditions representative of French breeding systems, with at least 2 different feeding systems depending on the season (winter and summer) for each breed. Besides providing more practical technical conditions for collecting samples, these farms apply controlled breeding conditions and, in some cases, shelter populations with strong genetic variability. When this genetic or feeding variability was not sufficiently representative of the national situation, additional samples were collected from commercial herds. Two samples were collected per animal at each milking. Bronopol was added to one of the samples, which was then analyzed by MIR spectrometry; the other was frozen at −20°C immedi-

ately after collection for future analysis using GC, the reference method.

Cow Milk Samples. A first set of 249 milk samples was collected from 127 Holstein × Normande crossbred dairy cows in 2008 and 2009 at the INRA "Domaine expérimental du Pin" experimental farm. The cows were a part of a QTL detection experiment (Larroque et al., 2002), and were produced after 2 generations of crosses (F2) between Normande and Holstein breeds that display numerous differences, in particular those pertaining to milk fat and protein content.

Milk samples were collected twice during the first lactation, in the winter and summer. The average stage of lactation was 160 DIM in the winter, and 209 DIM in the summer. During the winter (November–April), all cows were given the same diet, formulated to cover their requirements (INRA, 1989) and based on corn silage given ad libitum and completed with grass silage (4 kg of DM/d), beet pulp (1 kg of DM/d), and soybean meal (2 kg of DM/d). During the summer, cows grazed herbage with high nutritive value in a rotational grazing system (35 hm^2 per cow) and were supplemented with corn silage (1.5 kg/d). When the daily milk yield exceeded 21 kg/d in the winter and 23 kg/d in the summer, cows received 1 kg of energy concentrate per 2.5 kg of milk produced over the limit (maximum 7 kg of concentrate in winter and 5 kg in summer).

A second set of 153 milk samples were collected from 42 Montbéliardes and 35 Holsteins (primiparous or multiparous) in 2009 at the INRA experimental farm in Mirecourt (France), which implements an organic dairy farming system. Depending on calving dates, the cows belonged either to a grazing system or to a mixed-crop dairy system without concentrate. Milk samples were collected twice during their lactation: in the winter (March 2009) at 29 and 184 DIM, or in the summer (June 2009) at 120 and 274 DIM, respectively, for the grazing and mixed-crop dairy systems. In the winter, cows from the grazing system received ad libitum hay from the first cut of permanent grassland, which was completed with a second cut of permanent grassland (5.6 kg of DM/d) and a mineral complement. At the same period, cows from the mixed-crop dairy system received a diet dominated by hay from a first cut of alfalfa and orchard grass offered ad libitum and completed by flattened triticale (1.7 kg of DM/d), a second cut of temporary grassland (6.7 kg of DM/d), and a mineral complement. In the summer, cows from the grazing system grazed a new enclosed pasture of permanent grassland (orchard grass, white clover, and fescue). Cows from the mixed-crop dairy system grazed alternatively (night and day) 2 enclosed pastures of permanent grassland (based on ryegrass, white clover, and orchard grass). Milk samples were collected using

Table 1. Production conditions of the collected milk samples

Species	Farm	Breed	Main forage	Season	Year	No. of samples collected	No. of GC analyses
Cow	Pin-Au-Haras	Holstein × Normande	Corn silage	Winter	2008–2009	126	126
			Grazed fresh herbage + corn silage	Summer	2008–2009	123	123
	Mirecourt	Montbéliarde	Hay (permanent grassland)	Winter	2009	10	7
			Hay (alfalfa and orchard grass)	Winter	2009	32	24
		Holstein	Hay (permanent grassland)	Winter	2009	11	6
			Hay (alfalfa and orchard grass)	Winter	2009	24	13
		Montbéliarde	Grazed fresh herbage (orchard grass, white clover, fescue)	Summer	2009	22	16
			Grazed fresh herbage (ryegrass, white clover, orchard grass)	Summer	2009	19	12
		Holstein	Grazed fresh herbage (orchard grass, white clover, fescue)	Summer	2009	17	12
			Grazed fresh herbage (ryegrass, white clover, orchard grass)	Summer	2009	18	10
Ewe	Lafage	Lacaune	Ryegrass silage	Winter	2008	490	75
			Grazed fresh herbage	Spring	2008	493	75
	2 Private herds	Manech Red Faced	Grazed fresh herbage + concentrate	Spring	2009	68	30
	Private herd	Basco-Béarnaise	Grazed fresh herbage	Spring	2009	35	20
Goat	Bourges	Alpine	Hay (grass and legume)	Winter	2008	235	149
				Spring	2008	235	
				Autumn	2008	235	
	Private herd	Saanen	Grazed fresh herbage	Summer	2010	120	49
	Private herd	Saanen	Hay + haylage	Spring	2012	170	41
	Private herd	Alpine	Corn silage + hay + dehydrated alfalfa	Spring	2012	254	48
	Private herd	Alpine and Saanen	Barn dried hay + corn silage	Spring	2012	282	45

the same protocol as described above, but few cows were present at both samplings. Out of the total of 153 milk samples, 100 were selected to be analyzed by GC on their spectral variability.

Goat Milk Samples. A first set of 705 milk samples was collected in 2008 at the INRA experimental farm of Bourges (France) from 235 Alpine dairy goats. These goats were characterized by a diversity of α_{s1}-CN genotypes [ranging from null genotypes (15% OO) to high genotypes (33% AA)], which are known to influence milk FA composition (Chilliard et al., 2006). Samples were collected during the morning milking 3 times during the year (spring, winter, and autumn) and reflected 3 stages of lactation (about 40, 150, and 240 DIM). The goat diet remained unaltered throughout lactation and was based on hay (grass-legume mixture) offered ad libitum (about 1.6 kg of DM/d) and a commercial concentrate mixture (about 1 kg of DM/d). Among the available samples, 149 samples (approximately 50 per stage of lactation) showing a wide variability of spectra were selected to be analyzed for milk FA composition using GC. The second set of samples was collected in July 2010 in a privately owned herd of Saanen dairy goats. The diet of these goats was based on grazed fresh herbage (about 2 kg of DM/d), legume hay (0.4 kg of DM/d), and concentrate (0.8 kg of DM/d). One hundred twenty milk samples were collected during the morning milking. Goats were in the second part of their lactation at a stage from 120 to 350 DIM. Forty-nine samples that were the most representative of the diversity of the MIR spectra observed were chosen for analysis using GC. The last set of samples was collected between March and April 2012 in 3 privately owned herds during the evening milking. Seven hundred three samples were collected (170, 254, and 282 per herd, respectively). The first herd was composed of Saanen goats fed with hay and haylage. The second was composed of Alpine goats fed with corn silage, hay, and dehydrated alfalfa. The third herd was composed of Alpine and Saanen goats fed a diet based on barn dried hay and corn silage. One hundred thirty-four samples (41, 48, and 45 per herd, respectively) were selected to be analyzed by GC.

Ewe Milk Samples. A first set of samples was collected at 2 dates in 2008 from Lacaune dairy ewes at the INRA experimental farm of Lafage (France) from Lacaune dairy ewes. Samples were collected from primiparous and multiparous ewes, during the morning milking, respectively, in March 2008 at 80 DIM for 490 ewes (winter) and in May 2008 at 152 DIM, on average, for 493 ewes (spring). In the winter, all ewes received the same diet of ryegrass silage (1.6 kg of DM/d), hay (0.46 kg of DM/d), barley (0.34 kg/d), and concentrate (0.34 kg/d). In the spring, the ewes grazed 7 h/d (rye-

grass and orchard grass) and received straw. Seventy-five milk samples were selected within each sampling period based on the SCC, fat content, and milk spectra (i.e., a total of 150 frozen milk samples were analyzed using GC to determine FA composition).

A second set of samples was collected, using the same design as described above, from 3 privately owned flocks at the end of April 2009 at an average of 120 DIM with a pasture-based diet. The first herd was composed of Basco-Béarnaise (**BB**) ewes, and the 2 others were composed of Manech Red-Faced (**MRF**) ewes. All ewes grazed 5 h/d (permanent or temporary grassland) and received 1 kg of DM/d of hay (orchard grass, permanent grassland, or alfalfa) and 0.25 to 0.4 kg/d of maize. The MRF ewes also received 0.4 kg/d of concentrate. A total of 103 milk samples, respectively 35 from BB ewes and 68 from MRF ewes, were collected. Fifty of these samples were selected based on fat content, milk spectra, and breed for analysis, using GC (respectively 20 and 30 for the BB and MRF breeds). In all, 200 milk samples from Lacaune ewes (150 samples), BB ewes (20 samples), and MRF ewes (30 samples) for which both MIR spectra and GC results were available were included in the present study.

Milk Analysis

Fourier Transform MIR Measurements. Fresh milk samples were analyzed using MIR spectrometry with defined routine Fourier transform MIR analyses [MilkoScan FT6000 (Foss Electric A/S, Hillerød, Denmark) and Bentley FTS (Bentley Instruments Inc., Chaska, MN)]. Spectra were recorded from 5,012 to 926 cm^{-1}, with a spectral resolution of 3.85 cm^{-1} for the MilkoScan FT6000 analyzer and from 3,998 to 649 cm^{-1} with a spectral resolution of 3.73 cm^{-1} for the Bentley FTS analyzer. Spectra from the MilkoScan FT6000 analyzer were obtained using the calibration mode that allows exporting to comma-separated values (csv) files. Spectra were recorded at the Laboratoire Interprofessionnel Laitier de Normandie (LILANO, Saint-Lô, France; MilkoScan FT6000) and the Marœuil (France) laboratory (Bentley FTS) for cow milk, at the Laboratoire Interprofessionnel Laitier du Massif Central (LIAL MC, Aurillac, France; MilkoScan FT6000) for ewe milk, and at the Laboratoire interprofessionnel laitier du Centre Ouest (LILCO, Surgères, France; MilkoScan FT6000) for goat milk.

Reference Data. The milk FA composition of frozen milk samples was analyzed using GC according to International Organization for Standardization (**ISO**) standards [ISO 14156/IDF 172 for fat content extraction (ISO-IDF, 2001), ISO 15884/FIL 182 for preparation of FA methyl esters from milk fat (ISO-IDF,

2002a), and ISO 15885/FIL 184 for gas-liquid chromatography (ISO-IDF, 2002b)]. Analysis was carried out at the Laboratoire Analyses Alimentaires Recherche Fromagère (LARF, Mamirolle, France). The analysis protocol was identical to that described by Kramer et al. (1997), with the following equipment and protocol: Varian 3800 chromatograph (Varian Inc., Palo Alto, CA); 100-m CP-SIL 88 column (Varian Inc.); hydrogen carrier gas (199.1 kPa); splitless injector 1:50 at 250°C; oven temperature program: 4 min at 70°C, 13°C/min ramp from 70 to 175°C, 27 min at 175°C, 4°C/min ramp from 175 to 215°C, and 31 min at 215°C; and flame ionization detector at 250°C: Varian Star integrator. Approximately 68 individual FA or groups of FA were detected by GC.

MIR Spectrometry Predictions. The quantities expressed in grams per 100 g of FA in CG data were converted to grams per 100 mL using the fat content determined by the spectrometers; outliers were removed according to the Grubbs (1969) test. Approximately 30 ratios and sums were calculated: for example, the elongation index [(C8:0 + C10:0 + C12:0 + C14:0)/(C4:0 + C6:0)], C14 desaturation index [*cis*-9 C14:1/(*cis*-9 C14:1+C14:1)] (Heck et al., 2009), sum of odd FA (C13 + C15 + C17), and the spreadability index (C18:1/C16; Hurtaud et Peyraud, 2007).

Depending on the intended use (for example, for farming management), values need to be provided in grams per 100 g of FA. The fat content predicted by the MIR spectrometers (MilkoScan FT6000 and Bentley FTS), which comply with the IDF standard 141C:2000 (IDF, 2000) and AOAC International official method 972.16 (AOAC International, 2000), was used to convert the data in grams per 100 mL into grams per 100 g of FA.

Statistical Analysis

Equation Computing. For each species, samples were randomly assigned to either the calibration or validation set (cow milk with the MilkoScan FT6000: $n_{calibration} = 245$ and $n_{validation} = 105$; cow milk with the Bentley FTS: $n_{calibration} = 105$ and $n_{validation} = 45$; ewe milk: $n_{calibration} = 140$ and $n_{validation} = 60$; goat milk: $n_{calibration} = 140$ and $n_{validation} = 60$). Random assignment avoided any selection bias but did not ensure total independence between calibration and validation data sets. Indeed, 55 of the 96 cows for which 2 samples were collected had their 2 samples assigned to the same data set. This discrepancy affected only 2 samples for goats and none for ewes. The precision and robustness of the equations obtained with various optimization techniques (genetic algorithms, penalization methods, and denoising and optimization using wavelet transformations) were compared only for cow milk on

MilkoScan FT6000 data. The PLS regression was used as the reference method. In each case, only the informative wavelength bands (i.e., the bands that were not spoiled by water molecules) were retained (representing a total of 446 wavelengths for the MilkoScan FT6000 and 524 wavelengths for the Bentley FTS, according to the manufacturers' recommendations).

Development by PLS Regression and Validation. Equations were developed by univariate PLS regression (Tenenhaus, 2002) with the pls package in R 2.8.1 (Mevik and Wehrens, 2007). This method is suitable when the number of predictors is greater than the number of observations and when a strong collinearity exists between them. Compared with principal component regression, PLS regression maximizes the covariance between X and Y (Tenenhaus, 2002) to construct the new space composed of latent variables. For each equation, an optimal number of latent variables was determined according to the root mean square error of cross-validation.

To compare and assess the equations, several statistical parameters were computed according to the ISO 8196-2/IDF 128-2 standard (ISO-IDF, 2009). These included the mean bias (d), slope bias (b), the Student's t-test for d and b, the standard deviation of accuracy or residual standard deviation ($s_{y,x}$) and the coefficient of determination. These parameters were obtained by linear regression: $y = ax + b$, where x is the prediction obtained by MIR spectrometry and y the reference value obtained by GC. The residual standard deviation was expressed as a relative error = $s_{y,x}$/reference mean (%). To qualify the equations, the relative error and coefficient of determination were used. Estimations were considered sufficiently accurate and robust to be applied routinely when the relative error was under 5% and the coefficient of determination was greater than 0.91 (Coppa et al., 2010). When the relative error ranges from 5 to 10% and coefficient of determination is greater than 0.82, we advise to use these equations only for analytical purposes. If the relative error ranges from 10 to 15% and the coefficient of determination is greater than 70%, use of the concerned equation should be restrained to research and development programs, and results should be interpreted with caution.

Optimization Using Genetic Algorithms. To improve the equations and the quality of the predictions, several authors (Leardi et al., 1992; Spiegelman et al., 1998; Höskuldsson, 2001) have suggested using a selection of variables before PLS regression to limit spectral noise in the equations. The genetic algorithm method has often been applied to spectrometric data in this way, mainly for near-infrared spectra but also for MIR data (Leardi, 2000; Goicoechea and Olivieri, 2003).

The genetic algorithm (**GA**) method is based on evolutionary biology. A population of candidate solutions evolves using genetic operators such as reproduction, mutation, and selection. A solution (represented by a chromosome) is a vector in which each variable (represented by a gene) is coded with either 0 (not selected) or 1 (selected). The initial set of solutions (population) has a predefined number of candidate solutions. Evolution is controlled by a fitness function. In this study, the fitness function is the cross-validated explained variance of PLS regression applied to the selected variables. To breed a new generation (2 new solutions), 2 candidate solutions are selected. During this step of reproduction, crossing-over or mutation could occur. The obtained solutions integrate the population if they appear to be better than the previous solutions. The population size is constant, so the worst solutions are discarded when new solutions are retained. This process is repeated until a set number of generations is reached. To ensure optimal convergence, the GA is run several times.

Wavelengths were selected using GA with MATLAB 7.8. (The MathWorks Inc., Natick, MA; Matlab, R2009a) before PLS regression for cow, ewe, and goat milk (Ferrand et al., 2010). A specific GA for wavelength selection was used (Leardi and Lupiáñez González, 1998). The mutation rate, initial population, and number of variables selected in the solutions of the initial population were set to 1, 30, and 5% respectively. As indicated in the GA-PLS Toolbox that we used, the number of starting variables must not be greater than 200 to avoid the possibility of chance correlations. To meet this requirement, the GA was performed in 2 steps. We first reduced our data to 149 variables using the mean of 3 contiguous wavelengths before performing the algorithm with the selected variables only. Partial least squares regression was applied to both steps.

Optimization Using Penalization Methods. Penalization methods are used to reduce the estimator variance to guarantee the stability of the estimations. Several methods are available such as ridge regression (**RR**), which retains all predictors, the least absolute shrinkage and selection operator (**LASSO**) method in which some coefficients are set to 0, and finally the elastic net (**EN**) method (Zou and Hastie, 2005; Croiseau et al., 2011), which is a compromise between the RR and LASSO methods. Only 1 penalization parameter (λ) is required for RR and LASSO (Equations 1 and 2). In the EN method (Equation 3), a second parameter (α) allows weighting of RR and LASSO penalties:

$$RR : \hat{\beta} = \arg\min\left\{\sum_{i=1}^{n}(y_i - \mathbf{x}_i\boldsymbol{\beta})^2 + \lambda\sum_j \beta_j^2\right\}; \qquad [1]$$

$$\text{LASSO}: \hat{\beta} = \arg\min\left\{\sum_{i=1}^{n}(y_i - \mathbf{x}_i\boldsymbol{\beta})^2 + \lambda\sum_{j}\left|\beta_j\right|\right\}; \quad [2]$$

$$\text{EN}: \hat{\beta} = \arg\min\left\{\sum_{i=1}^{n}(y_i - \mathbf{x}_i\boldsymbol{\beta})^2 + \lambda\left[(1-\alpha)\sum_{j}\beta_j^2 + \alpha\sum_{j}\left|\beta_j\right|\right]\right\},$$
$$[3]$$

where $\boldsymbol{\beta}$ is the vector of the wavelength coefficients β_j, y_i is the FA content of milk i, and $\mathbf{x}_i$ the vector of absorbances (j) of this milk. Similarly to the GA, the EN method selects the most informative wavelengths. The glmnet R package was used to implement the different methods of penalization and the penalization parameters were chosen by cross-validation.

Denoising Using First-Derivative Preprocessing. Use of the first derivative eliminates the uncontrolled variations of the spectra. It is calculated with the following formula:

$$\frac{\mathrm{d}x}{\mathrm{d}\lambda_i} = \frac{x_{i-\delta} - x_{i+\delta}}{\lambda_{i-\delta} + \lambda_{i+\delta}},$$

where $\dfrac{\mathrm{d}x}{\mathrm{d}\lambda_i}$ is the derivative value for the spectral point at wavelength i and $\delta = 2$.

Denoising and Optimization Using Wavelet Transformation and Multiresolution Analysis. Wavelet transformation decomposes a signal into 2 series: a set of approximation coefficients that match the trend (baseline) and the overall variation (noise), and a set of wavelet coefficients that match the fine details of the signal. In multiresolution analysis (Mallat, 1989), wavelet transformation is repeated several times to obtain a fine level of details. To apply multiresolution analysis, MIR spectra were expanded by periodic padding and then decomposed using discrete wavelet transformation with a symmlet-8 wavelet. Wavelet coefficients were extracted at different levels and only coefficients greater than $10^{-3}\sqrt{2\log(\mathrm{N})}$ (Donoho and Johnstone, 1994), where N represents the size of the spectrum and was equal to 512 in the current study, were incorporated into PLS regression analysis as predictor variables. Wavelet analysis was performed using MATLAB 7.8. (R2009a; The MathWorks Inc.).

RESULTS

Descriptive Statistics of GC Data

The descriptive statistics of the different data sets are displayed in Tables 2 and 3 (cow milk), Tables 4 and 5 (ewe milk), and Tables 6 and 7 (goat milk). Cow milk samples contained, on average, 71.5 g of SFA/100 g of FA, 24.3 g of MUFA/100 g of FA, and 2.9 g of PUFA/100 g of FA. Among the SFA, C16:0 was observed to show the highest content (31.7 g/100 g of FA), followed by C14:0 (11.4 g/100 g of FA). The variability of GC data was high, but differed according to the FA, as indicated by the standard deviation, minimum, and maximum. Coefficients of variation ranged from 10% (C4:0) to 46% (C18:3n-3). More specifically, SFA values ranged between a minimum of 58.2 g/100 g of FA and a maximum of 82 g/100 g of FA.

Samples of ewe milk consisted of 74.2 g of SFA/100 g of FA, 20.6 g of MUFA/100 g of FA, and 3.6 g of PUFA/100 g of FA. Coefficients of variation ranged from 8% for SFA (minimum = 59.4 g/100 g of FA and maximum = 83.4 g/100 g of FA) to 50% for total *trans* 18:1 (minimum = 0.9 g/100 g of FA and maximum = 5.5 g/100 g of FA).

In goat milk, SFA represented 72.2 g/100 g of total FA, whereas MUFA and PUFA represented 22.7 and 4.1 g/100 g of total FA, respectively. Among these SFA, the concentration of C16:0 was highest (28.4 g/100 g in the calibration set), and the cumulative proportion of C6:0, C8:0, and C10:0 (specific goat FA) was 16.9 g/100 g of FA (in the calibration set). As indicated by the standard deviation, minimum, and maximum values, the variability of goat GC data was quite considerable, but differed according to the FA. Coefficients of variation ranged from 4% for SFA (minimum = 60.5 g/100 g of FA and maximum = 78.5 g/100 g of FA) to 39% for C18:3n-3.

FA Prediction in Milk (g/100 mL)

Comparison of Methods Used to Calibrate the Equations for Cow Milk. The performances of the various methods (PLS, GA + PLS, penalization methods + PLS, first-derivative preprocessing + PLS, and wavelet decomposition + PLS) are summarized in Table 8 with the $s_{y,x}$ and validation coefficient of determination $\left(R_v^2\right)$. The best equation is the one resulting in the smallest $s_{y,x}$ and the highest R_v^2. Optimal results were generally obtained by combining either GA and PLS (C4:0, C6:0, C8:0, and C10:0 with $R_v^2 \geq 0.93$ and a relative error <6%, and C18:0 with $R_v^2 = 0.89$ and relative error = 9.7%), or first-derivative preprocessing and PLS (C14:0, total *cis* 18:1, and MUFA with $R_v^2 > 0.95$ and a relative error <5%, PUFA with $R_v^2 = 0.87$ and relative error = 6.8%, and *trans* FA with $R_v^2 = 0.90$ and relative error = 11.6%). The reference method (PLS) provided more accurate predictions for C16:0 ($R_v^2 =$

Table 2. Descriptive statistics of GC data (g/100 g of FA) for calibration and validation data sets in cow milk

Trait	Calibration					Validation				
	Mean	SD	Minimum	Maximum	CV	Mean	SD	Minimum	Maximum	CV
C4:0	3.891	0.388	3.02	5.61	9.97	3.825	0.301	3.12	4.68	7.87
C6:0	2.294	0.243	1.57	2.81	10.60	2.284	0.264	0.98	2.72	11.56
C8:0	1.31	0.207	0.40	1.78	15.80	1.311	0.228	0.39	1.71	17.39
C10:0	2.989	0.649	1.58	4.79	21.70	3.000	0.737	0.68	4.73	24.57
C12:0	3.412	0.847	1.74	5.91	24.80	3.465	0.945	0.78	5.78	27.27
C14:0	11.35	1.611	0.28	14.62	14.20	11.41	1.725	4.86	14.61	15.12
C16:0	31.655	4.701	22.62	41.79	14.90	32.148	4.851	21.93	42.45	15.09
C18:0	9.548	2.218	4.32	14.81	23.20	9.165	2.208	5.02	14.14	24.09
Total *trans* 18:1[1]	2.238	0.665	0.89	4.12	29.70	2.164	0.641	0.95	3.73	29.62
cis-9 C18:1	18.46	4.425	10.49	30.14	24.00	18.265	5.037	9.18	35.21	27.58
Total *cis* 18:1	19.138	4.501	11.07	31.08	23.50	19.016	5.126	9.68	36.34	26.96
Total 18:1	21.376	4.768	12.84	33.12	22.30	21.179	5.363	11.23	38.64	25.32
cis-9,*cis*-12 C18:2 (LeA[2])	1.306	0.254	0.14	2.05	19.40	1.399	0.269	0.92	2.16	19.23
cis-9,*trans*-11 C18:2 (CLA[3])	0.536	0.221	0.24	1.40	41.20	0.502	0.187	0.20	1.18	37.25
C18:3n-3 (ALA[4])	0.624	0.287	0.22	1.33	46.00	0.592	0.302	0.20	1.28	51.01
SFA	71.544	5.454	58.25	82.04	7.62	71.671	6.075	53.17	83.90	8.48
Unsaturated FA	27.202	5.196	17.32	40.43	19.10	27.116	5.902	15.34	45.89	21.77
MUFA	24.296	4.951	14.88	36.80	20.40	24.173	5.631	13.20	41.93	23.29
PUFA	2.906	0.464	1.75	4.23	16.00	2.942	0.433	1.93	4.08	14.72
trans FA	2.994	0.897	1.36	5.58	30.00	2.879	0.831	1.40	5.22	28.86
n-3	0.834	0.301	0.37	1.64	36.10	0.802	0.315	0.34	1.45	39.28
n-6	2.025	0.375	0.77	3.21	18.50	2.177	0.475	1.60	5.06	21.82
n-7	2.592	0.548	1.46	4.20	21.15	2.792	2.346	1.46	25.71	84.04
C14 desaturation index[5]	7.414	4.596	1.93	73.33	61.99	7.573	1.699	4.78	12.92	22.43
C18:1/C16:0	0.637	0.240	0.29	1.26	37.71	0.629	0.276	0.23	1.49	43.86
Sum of C13:0 + C15:0 + C17:0	2.156	0.295	1.04	2.92	13.67	2.195	0.305	1.64	3.10	13.92
Elongation index[6]	3.090	0.496	1.38	4.39	16.05	3.139	0.548	1.57	4.49	17.45

[1]The term total is used to designate the sum of a subgroup of FA as detected using GC.

[2]LeA = linoleic acid.

[3]CLA = conjugated linoleic acid.

[4]ALA = α-linolenic acid.

[5]C14 desaturation index = *cis*-9 C14:1/(C14:0 + *cis*-9 C14:1).

[6]Elongation index = (C8:0 + C10:0 + C12:0 + C14:0)/(C4:0 + C6:0).

0.94 and relative error = 5.5%) and C18:3 n-3 (R_v^2 = 0.86 and relative error = 19%). When the wavelet method and PLS were combined, results were very similar to those obtained with first derivative + PLS with R_v^2 >0.95 and a relative error <7% for C6:0, C8:0, C10:0, C12:0, C14:0, *cis*-9 C18:1, total *cis* 18:1, total 18:1, SFA, and MUFA. Penalization methods (EN, LASSO, and RR) gave the worst results for all FA. All models developed by RR resulted in R_v^2 <0.80, except for C4:0, C6:0, and SFA. Whereas other methods all led to R_v^2 >0.90, identical results were observed with LASSO and EN methods with R_v^2 = 0.80 to 0.90.

Performance of the Finalized Equations. The statistical parameters for the predictive equations of milk FA contents are presented in Table 9 (cow milk) and Table 10 (goat and ewe milk). In the 3 species, better estimations were obtained for FA present at medium or high concentrations [notably all SFA (C4:0 to C16:0) and oleic acid (*cis*-9 C18:1 9)] than for FA present at low concentrations. With the data for cow milk

obtained using the Foss analyzer, the relative error ranged from 1.27 to 19.3% and R_v^2 from 0.23 (for the C14:0 desaturation index) to 1 (for SFA). Fourteen FA showed a relative error <7% and R_v^2 >0.90 (C4:0, C6:0, C8:0, C10:0, C12:0, C14:0, C16:0, C18:1 9c, total *cis* 18:1, SFA, MUFA, and unsaturated FA). The estimations for PUFA and *trans* FA were less accurate with, respectively, relative errors of 9.1 and 13.2% and R_v^2 of 0.76 and 0.86. For conjugated linoleic acid and α-linolenic acid equations, appropriate values were obtained for R_v^2 (0.82 and 0.85), but the relative error was >15%. When data for the Bentley analyzer was used, the equations were less accurate with a relative error ranging from 3.1 to 22.5% and R_v^2 from 0.05 to 0.96. The value of R_v^2 for major FA (C4:0, C6:0, C8:0, C10:0, C12:0, C14:0, C16:0, *cis*-9 C18:1, and total *cis* 18:1) ranged from 0.80 to 0.90.

In ewe milk (data from the Foss analyzer), the relative error ranged from 1.0 to 22.3% and R_v^2 from 0.49 (for *cis*-9,*cis*-12 C18:2) to 1 (for SFA). Fourteen FA

Table 3. Descriptive statistics of GC data (converted into g/100 mL) for calibration and validation data sets in cow milk

Trait	Calibration						Validation					
	n[1]	Mean	SD	Minimum	Maximum	CV	n	Mean	SD	Minimum	Maximum	CV
C4:0	242	0.154	0.025	0.10	0.25	16.52	103	0.148	0.024	0.08	0.21	16.55
C6:0	244	0.091	0.017	0.04	0.14	18.37	101	0.087	0.015	0.05	0.12	17.41
C8:0	244	0.052	0.011	0.02	0.09	22.15	103	0.050	0.012	0.02	0.08	23.68
C10:0	244	0.119	0.032	0.05	0.21	27.14	103	0.113	0.034	0.04	0.19	29.81
C12:0	244	0.135	0.041	0.06	0.26	30.14	102	0.129	0.041	0.05	0.23	31.40
C14:0	243	0.452	0.091	0.22	0.77	20.12	105	0.436	0.097	0.21	0.71	22.33
C16:0	244	1.260	0.299	0.62	2.35	23.76	103	1.218	0.272	0.56	1.90	22.37
C18:0	243	0.379	0.113	0.09	0.75	29.88	103	0.348	0.105	0.13	0.60	30.30
Total trans 18:1[2]	243	0.089	0.031	0.02	0.20	35.16	105	0.083	0.030	0.02	0.16	35.60
cis-9 C18:1	242	0.726	0.215	0.22	1.54	29.62	100	0.671	0.193	0.27	1.23	28.81
Total cis 18:1	242	0.753	0.220	0.22	1.59	29.27	101	0.698	0.197	0.28	1.29	28.24
Total 18:1	242	0.842	0.239	0.24	1.70	28.40	101	0.779	0.212	0.30	1.41	27.26
cis-9,cis-12 C18:2 (LeA[3])	242	0.051	0.012	0.01	0.08	22.85	103	0.053	0.012	0.02	0.08	22.89
cis-9,trans-11 C18:2 (CLA[4])	242	0.021	0.010	0.00	0.06	45.20	103	0.019	0.007	0.01	0.04	39.45
C18:3n-3 (ALA[5])	243	0.024	0.012	0.01	0.07	48.64	103	0.022	0.011	0.01	0.05	50.16
SFA	243	2.835	0.525	1.25	4.86	18.54	103	2.724	0.508	1.33	3.70	18.66
Unsaturated FA	242	1.072	0.279	0.33	2.20	26.04	101	1.001	0.246	0.41	1.73	24.61
MUFA	242	0.958	0.260	0.29	1.95	27.11	101	0.891	0.230	0.35	1.58	25.84
PUFA	237	0.113	0.022	0.07	0.18	19.25	102	0.110	0.021	0.06	0.16	18.81
trans FA	243	0.119	0.042	0.02	0.27	35.55	105	0.111	0.039	0.03	0.21	34.88
n-3	243	0.033	0.013	0.01	0.08	40.55	102	0.029	0.011	0.01	0.06	38.55
n-6	244	0.080	0.019	0.02	0.15	24.01	103	0.082	0.019	0.04	0.13	22.76
n-7	243	0.103	0.030	0.04	0.22	29.49	100	0.095	0.027	0.04	0.18	28.67
C14 desaturation index[6]	242	7.106	1.693	1.93	12.92	23.83	103	7.487	1.592	4.78	10.75	21.27
C18:1/C16:0	244	0.712	0.262	0.33	1.42	36.77	103	0.683	0.273	0.27	1.51	40.00
Sum of C13:0 + C15:0 + C17:0	241	0.084	0.017	0.03	0.13	19.56	103	0.083	0.016	0.04	0.12	19.66
Elongation index[7]	244	3.090	0.496	1.38	4.39	16.05	105	3.139	0.548	1.57	4.49	17.45

[1]Number of samples after elimination of outliers using the Grubbs (1969) test.

[2]The term total is used to designate the sum of a subgroup of FA as detected using GC.

[3]LeA = linoleic acid.

[4]CLA = conjugated linoleic acid.

[5]ALA = α-linolenic acid.

[6]C14 desaturation index = cis-9 C14:1/(C14:0 + cis-9 C14:1).

[7]Elongation index = (C8:0 + C10:0 + C12:0 + C14:0)/(C4:0 + C6:0).

Table 4. Descriptive statistics of GC data (g/100 g of FA) for calibration and validation data sets in ewe milk

Trait	Calibration					Validation				
	Mean	SD	Minimum	Maximum	CV	Mean	SD	Minimum	Maximum	CV
C4:0	3.579	0.519	2.23	5.35	14.50	3.525	0.486	2.44	4.57	13.79
C6:0	2.696	0.329	1.68	3.58	12.20	2.717	0.321	2.02	3.32	11.81
C8:0	2.583	0.379	1.51	3.56	14.67	2.672	0.300	1.76	3.33	11.23
C10:0	8.555	1.689	4.28	12.35	19.74	8.793	1.444	5.34	11.04	16.42
C12:0	5.028	1.291	2.46	9.09	25.68	5.183	1.090	3.16	8.03	21.03
C14:0	12.465	1.855	7.65	16.22	14.88	12.506	1.600	9.46	17.45	12.79
C16:0	26.572	3.768	19.52	35.61	14.18	25.371	3.358	19.35	32.03	13.24
C18:0	7.622	1.954	4.00	14.06	25.64	7.739	1.774	3.78	12.06	22.92
Total *trans* 18:1[1]	2.525	1.253	0.89	5.52	49.62	2.702	1.440	0.99	6.87	53.29
cis-9 C18:1	15.843	3.888	10.50	29.58	24.54	16.182	3.922	8.82	27.14	24.24
Total *cis* 18:1	16.371	3.969	10.85	30.13	24.24	16.713	3.990	9.15	27.65	23.87
Total 18:1	18.896	4.663	11.88	32.76	24.68	19.415	4.530	11.15	30.43	23.33
cis-9,*cis*-12 C18:2 (LeA[2])	1.352	0.295	0.79	2.36	21.82	1.340	0.245	0.82	2.03	18.28
cis-9,*trans*-11 C18:2 (CLA[3])	0.892	0.476	0.28	2.13	53.36	0.912	0.505	0.30	1.96	55.37
C18:3n-3 (ALA[4])	0.741	0.154	0.49	1.30	20.78	0.753	0.139	0.44	1.06	18.46
SFA	74.218	5.696	59.49	83.35	7.67	73.696	5.571	61.83	83.40	7.56
Unsaturated FA	24.201	5.375	15.66	38.71	22.21	24.693	5.199	16.130	36.38	21.05
MUFA	20.643	4.684	13.72	34.72	22.69	21.117	4.613	13.29	32.16	21.84
PUFA	3.558	0.876	1.83	5.73	24.62	3.576	0.741	2.16	5.15	20.72
trans FA	3.793	1.879	1.34	8.05	49.54	4.011	2.095	1.43	9.57	52.23
n-3	1.026	0.215	0.64	1.85	20.96	1.036	0.175	0.66	1.48	16.89
n-6	2.057	0.487	1.09	3.48	23.68	2.079	0.383	1.39	2.87	18.42
n-7	2.3544	0.573	1.27	4.26	24.35	2.333	0.602	1.35	3.81	25.82
C14 desaturation index[5]	1.7129	0.534	0.73	3.40	31.17	1.657	0.611	0.76	3.24	36.90
C18:1/C16:0	0.6437	0.234	0.31	1.31	36.34	0.683	0.231	0.30	1.26	33.76
Sum of C13:0 + C15:0 + C17:0	1.9631	0.231	1.37	2.63	11.79	1.973	0.265	1.53	2.67	13.42
Elongation index[6]	4.6278	0.928	2.63	7.76	20.05	4.737	0.837	3.35	6.67	17.68

[1]The term total is used to designate the sum of a subgroup of FA as detected using GC.

[2]LeA = linoleic acid.

[3]CLA = conjugated linoleic acid.

[4]ALA = α-linolenic acid.

[5]C14 desaturation index = *cis*-9 C14:1/(C14:0 + *cis*-9 C14:1).

[6]Elongation index = (C8:0 + C10:0 + C12:0 + C14:0)/(C4:0 + C6:0).

Table 5. Descriptive statistics of GC data (converted into g/100 mL) for calibration and validation data sets in ewe milk

Trait	Calibration						Validation					
	n[1]	Mean	SD	Minimum	Maximum	CV	n	Mean	SD	Minimum	Maximum	CV
C4:0	140	0.240	0.050	0.12	0.38	20.77	56	0.235	0.038	0.18	0.34	16.07
C6:0	140	0.182	0.039	0.09	0.27	21.22	58	0.179	0.034	0.10	0.26	18.86
C8:0	140	0.175	0.042	0.09	0.30	24.27	58	0.177	0.038	0.10	0.27	21.26
C10:0	140	0.581	0.166	0.25	0.99	28.52	58	0.585	0.148	0.32	0.93	25.38
C12:0	140	0.343	0.119	0.14	0.72	34.51	58	0.347	0.105	0.18	0.56	30.29
C14:0	140	0.855	0.246	0.35	1.38	28.72	58	0.840	0.219	0.44	1.24	26.05
C16:0	138	1.779	0.410	0.95	2.78	23.05	58	1.684	0.357	0.97	2.36	21.19
C18:0	140	0.518	0.172	0.22	0.99	33.17	58	0.511	0.145	0.26	0.83	28.37
Total *trans* 18:1[2]	140	0.176	0.101	0.04	0.46	57.05	58	0.179	0.101	0.04	0.38	56.38
cis-9 C18:1	140	1.074	0.336	0.46	1.91	31.24	58	1.075	0.349	0.53	1.93	32.43
Total *cis* 18:1	140	1.110	0.346	0.47	1.96	31.13	58	1.111	0.358	0.55	1.96	32.24
Total 18:1	140	1.287	0.418	0.52	2.29	32.52	58	1.299	0.432	0.59	2.16	33.21
cis-9,*cis*-12 C18:2 (LeA[3])	139	0.090	0.022	0.04	0.14	24.00	56	0.089	0.017	0.06	0.13	18.76
cis-9,*trans*-11 C18:2 (CLA[4])	140	0.063	0.039	0.01	0.19	62.14	58	0.062	0.039	0.01	0.15	62.70
C18:3n-3 (ALA[5])	139	0.050	0.014	0.02	0.09	27.48	58	0.050	0.013	0.02	0.08	26.51
SFA	140	5.039	1.118	2.51	7.79	22.18	58	4.919	1.025	2.80	6.89	20.84
Unsaturated FA	140	1.651	0.519	0.66	2.97	31.41	58	1.657	0.534	0.73	2.68	32.19
MUFA	140	1.408	0.444	0.56	2.47	31.57	58	1.416	0.461	0.63	2.32	32.59
PUFA	140	0.243	0.083	0.10	0.50	34.23	58	0.240	0.077	0.09	0.41	32.00
trans FA	140	0.266	0.154	0.06	0.66	57.83	58	0.269	0.154	0.06	0.62	57.47
n-3	139	0.069	0.020	0.03	0.13	29.08	56	0.071	0.018	0.04	0.12	25.14
n-6	138	0.137	0.040	0.06	0.22	28.91	58	0.138	0.036	0.07	0.22	25.82
n-7	138	0.161	0.058	0.05	0.30	36.06	58	0.158	0.060	0.05	0.28	38.26
C14 desaturation index[6]	138	1.690	0.502	0.73	3.08	29.70	58	1.606	0.556	0.76	2.94	34.59
C18:1/C16:0	140	0.745	0.274	0.34	1.47	36.81	58	0.776	0.250	0.36	1.31	32.24
Sum of C13:0 + C15:0 + C17:0	140	0.133	0.030	0.06	0.21	22.72	58	0.132	0.032	0.07	0.20	24.21
Elongation index[7]	138	4.587	0.869	2.63	6.53	18.95	58	4.675	0.779	3.35	6.29	16.67

[1]Number of samples after elimination of outliers using the Grubbs (1969) test.

[2]The term total is used to designate the sum of a subgroup of FA as detected using GC.

[3]LeA = linoleic acid.

[4]CLA = conjugated linoleic acid.

[5]ALA = α-linolenic acid.

[6]C14 desaturation index = *cis*-9 C14:1/(C14:0 + *cis*-9 C14:1).

[7]Elongation index = (C8:0 + C10:0 + C12:0 + C14:0)/(C4:0 + C6:0).

Table 6. Descriptive statistics of GC data (g/100 g of FA) for calibration and validation data sets in goat milk

Trait	Calibration					Validation				
	Mean	SD	Minimum	Maximum	CV	Mean	SD	Minimum	Maximum	CV
C4:0	2.708	0.371	1.78	4.31	13.70	2.737	0.353	1.92	3.70	12.90
C6:0	2.457	0.264	1.72	3.17	10.74	2.456	0.270	1.72	3.04	10.99
C8:0	2.634	0.377	1.67	3.88	14.31	2.611	0.379	1.50	3.35	14.52
C10:0	9.117	1.569	4.55	13.50	17.21	8.977	1.570	4.08	11.85	17.49
C12:0	4.256	0.891	1.08	7.29	20.94	4.214	0.848	2.29	6.88	20.12
C14:0	10.242	1.523	5.88	14.09	14.87	10.193	1.677	5.80	13.74	16.45
C16:0	28.445	3.047	19.89	44.15	10.71	28.292	2.808	19.40	34.74	9.93
C18:0	8.640	1.909	3.15	15.49	22.09	8.973	1.757	4.20	14.94	19.58
Total trans 18:1[1]	2.264	0.569	0.85	4.62	25.13	2.244	0.572	1.18	3.76	25.49
cis-9 C18:1	18.578	2.645	12.97	29.21	14.24	18.654	2.804	11.84	28.15	15.03
Total cis 18:1	19.093	2.553	13.57	28.10	13.37	19.059	2.54	12.50	26.18	13.33
Total 18:1	21.434	2.682	15.96	32.10	12.51	21.336	2.493	15.84	27.84	11.68
cis-9,cis-12 C18:2 (LeA[2])	2.450	0.488	1.51	3.75	19.92	2.394	0.544	1.49	4.11	22.72
cis-9,trans-11 C18:2 (CLA[3])	0.533	0.139	0.26	0.94	26.08	0.529	0.180	0.18	1.28	34.03
C18:3n-3 (ALA[4])	0.501	0.195	0.03	1.54	38.92	0.525	0.224	0.25	1.33	42.67
SFA	72.211	2.979	60.50	78.45	4.13	72.203	3.039	61.92	78.54	4.21
Unsaturated FA	27.374	5.165	20.47	62.96	18.87	27.463	5.452	20.19	61.80	19.85
MUFA	22.684	2.628	16.85	31.69	11.59	22.596	2.535	16.72	29.16	11.22
PUFA	4.106	0.681	2.51	5.78	16.59	4.001	0.695	2.83	5.68	17.37
trans FA	3.251	0.760	1.39	6.40	23.38	3.175	0.755	1.60	5.31	23.78
n-3	0.699	0.245	0.20	1.99	35.05	0.723	0.283	0.38	1.76	39.14
n-6	3.407	0.626	2.00	5.05	18.37	3.318	0.661	2.13	5.04	19.92
n-7	1.563	0.260	0.80	2.41	16.65	1.567	0.284	1.08	2.49	18.12
C14 desaturation index[5]	1.356	0.625	0.50	6.67	46.11	1.272	0.527	0.65	3.44	41.47
C18:1/C16:0	0.688	0.163	0.35	2.09	23.65	0.697	0.156	0.41	1.55	22.32
Sum of C13:0 + C15:0 + C17:0	1.503	0.231	0.63	2.18	15.39	1.519	0.218	0.83	2.07	14.37
Elongation index[6]	5.167	1.037	2.31	8.40	20.07	5.080	1.019	2.92	7.56	20.06

[1]The term total is used to designate the sum of a subgroup of FA as detected using GC.

[2]LeA = linoleic acid.

[3]CLA = conjugated linoleic acid.

[4]ALA = α-linolenic acid.

[5]C14 desaturation index = cis-9 C14:1/(C14:0 + cis-9 C14:1).

[6]Elongation index = (C8:0 + C10:0 + C12:0 + C14:0)/(C4:0 + C6:0).

Table 7. Descriptive statistics of GC data (converted into g/100 mL) for calibration and validation data sets in goat milk

Trait	Calibration						Validation					
	n[1]	Mean	SD	Minimum	Maximum	CV	n	Mean	SD	Minimum	Maximum	CV
C4:0	232	0.095	0.022	0.05	0.16	22.64	98	0.095	0.020	0.05	0.14	21.08
C6:0	233	0.088	0.023	0.04	0.17	26.53	98	0.086	0.020	0.04	0.13	23.73
C8:0	233	0.095	0.029	0.04	0.20	30.63	98	0.091	0.025	0.04	0.15	27.56
C10:0	233	0.329	0.109	0.12	0.69	33.09	100	0.319	0.100	0.13	0.54	31.22
C12:0	233	0.153	0.051	0.05	0.31	33.51	98	0.147	0.044	0.07	0.26	30.24
C14:0	233	0.365	0.099	0.15	0.68	27.09	100	0.357	0.090	0.19	0.59	25.32
C16:0	233	1.004	0.209	0.53	1.59	20.78	98	0.982	0.206	0.45	1.49	21.00
C18:0	231	0.307	0.102	0.08	0.64	33.12	98	0.314	0.094	0.10	0.53	30.08
Total *trans* 18:1[2]	226	0.079	0.025	0.02	0.16	31.88	98	0.078	0.023	0.03	0.14	30.03
cis-9 C18:1	233	0.660	0.164	0.30	1.25	24.90	98	0.649	0.162	0.34	1.01	24.93
Total *cis* 18:1	232	0.679	0.168	0.31	1.28	24.79	97	0.667	0.165	0.35	1.03	24.67
Total 18:1	229	0.758	0.184	0.34	1.38	24.23	97	0.746	0.177	0.42	1.14	23.69
cis-9,*cis*-12 C18:2 (LeA[3])	233	0.086	0.023	0.04	0.15	27.04	98	0.083	0.023	0.05	0.14	28.11
cis-9,*trans*-11 C18:2 (CLA[4])	233	0.019	0.007	0.01	0.04	35.53	98	0.018	0.006	0.01	0.04	32.14
C18:3n-3 (ALA[5])	233	0.018	0.007	0.00	0.04	40.42	98	0.018	0.006	0.01	0.04	36.75
SFA	233	2.573	0.593	1.29	4.65	23.02	100	2.546	0.560	1.40	3.77	21.99
Unsaturated FA	229	0.950	0.228	0.44	1.71	24.05	96	0.930	0.214	0.54	1.40	22.99
MUFA	229	0.805	0.194	0.38	1.44	24.17	97	0.790	0.186	0.44	1.20	23.56
PUFA	233	0.146	0.041	0.06	0.29	28.30	99	0.141	0.037	0.07	0.22	25.90
trans FA	227	0.115	0.038	0.04	0.23	32.89	97	0.111	0.033	0.05	0.20	29.85
n-3	233	0.025	0.010	0.01	0.06	38.64	98	0.024	0.008	0.01	0.05	34.55
n-6	232	0.121	0.035	0.05	0.24	29.18	99	0.117	0.033	0.06	0.20	28.67
n-7	232	0.056	0.015	0.02	0.10	26.99	98	0.054	0.013	0.03	0.09	24.36
C14 desaturation index[6]	231	1.324	0.502	0.50	3.25	37.92	97	1.217	0.427	0.65	2.69	35.10
C18:1/C16:0	228	0.761	0.137	0.37	1.29	18.02	96	0.756	0.112	0.48	1.09	14.86
Sum of C13:0 + C15:0 + C17:0	232	0.053	0.015	0.02	0.11	27.57	98	0.052	0.011	0.02	0.08	21.08
Elongation index[7]	233	5.167	1.037	2.31	8.40	20.07	100	5.080	1.019	2.92	7.56	20.06

[1]Number of samples after elimination of outliers using the Grubbs (1969) test.

[2]The term total is used to designate the sum of a subgroup of FA as detected using GC.

[3]LeA = linoleic acid.

[4]CLA = conjugated linoleic acid.

[5]ALA = α-linolenic acid.

[6]C14 desaturation index = *cis*-9 C14:1/(C14:0 + *cis*-9 C14:1).

[7]Elongation index = (C8:0 + C10:0 + C12:0 + C14:0)/(C4:0 + C6:0).

Table 8. Comparison of the methods used to develop calibration equations on the MilkoScan FT6000 analyzer (Foss Electric A/S, Hillerød, Denmark) data for FA in cow milk (g/100 mL) on the validation set

FA (g/100 mL of milk)	PLS[1]		AG1PLS[2]		AG2PLS[3]		Elastic net		LASSO[4]		Ridge regression		First derivative + PLS		Wavelet[5] + PLS	
	$s_{y,x}$[6]	R^2	$s_{y,x}$	R^2	$s_{y,x}$	R^2	$s_{y,x}$	R^2	$s_{y,x}$	R^2	$s_{y,x}$	R^2	$s_{y,x}$	R^2	$s_{y,x}$	R^2
C4:0	0.006	0.93	0.006	0.93	0.006	0.93	0.008	0.87	0.008	0.90	0.010	0.84	0.007	0.92	0.006	0.94
C6:0	0.003	0.96	0.003	0.96	0.003	0.97	0.004	0.93	0.004	0.93	0.006	0.84	0.003	0.96	0.003	0.96
C8:0	0.002	0.97	0.002	0.98	0.002	0.97	0.004	0.89	0.004	0.90	0.006	0.75	0.002	0.97	0.002	0.97
C10:0	0.007	0.95	0.007	0.95	0.007	0.96	0.013	0.84	0.012	0.85	0.020	0.62	0.007	0.96	0.007	0.96
C12:0	0.008	0.96	0.008	0.96	0.008	0.96	0.016	0.82	0.016	0.84	0.027	0.48	0.007	0.97	0.007	0.96
C14:0	0.022	0.95	0.022	0.95	0.022	0.95	0.037	0.84	0.037	0.86	0.049	0.74	0.019	0.96	0.021	0.95
C16:0	0.067	0.94	0.074	0.92	0.074	0.93	0.101	0.86	0.098	0.87	0.144	0.71	0.079	0.93	0.077	0.92
C18:0	0.039	0.85	0.034	0.89	0.040	0.85	0.048	0.76	0.048	0.77	0.074	0.46	0.048	0.85	0.038	0.87
Total *trans* 18:1	0.012	0.85	0.012	0.83	0.012	0.83	0.016	0.71	0.016	0.71	0.024	0.30	0.011	0.88	0.011	0.87
cis-9 C18:1	0.034	0.97	0.032	0.97	0.032	0.97	0.068	0.85	0.061	0.90	0.139	0.38	0.030	0.98	0.039	0.96
Total *cis* 18:1	0.035	0.97	0.034	0.97	0.036	0.96	0.071	0.84	0.064	0.90	0.141	0.37	0.030	0.98	0.042	0.95
Total 18:1	0.034	0.97	0.036	0.97	0.042	0.96	0.080	0.84	0.073	0.88	0.164	0.38	0.032	0.98	0.042	0.96
cis-9,*cis*-12 C18:2 (LeA[7])	0.006	0.78	0.006	0.76	0.006	0.75	0.008	0.58	0.007	0.61	0.009	0.45	0.005	0.80	0.005	0.80
cis-9,*trans*-11 C18:2 (CLA[8])	0.003	0.83	0.003	0.83	0.003	0.83	0.004	0.58	0.004	0.71	0.006	0.21	0.003	0.87	0.004	0.78
C18:3n-3 (ALA[9])	0.004	0.86	0.006	0.69	0.008	0.25	0.007	0.54	0.007	0.58	0.008	0.30	0.004	0.82	0.004	0.85
SFA	0.030	1.00	0.036	1.00	0.039	0.99	0.099	0.96	0.098	0.96	0.158	0.91	0.032	1.00	0.036	0.99
MUFA	0.030	0.98	0.034	0.98	0.043	0.97	0.085	0.84	0.076	0.89	0.174	0.41	0.028	0.99	0.040	0.97
PUFA	0.010	0.78	0.009	0.81	0.009	0.81	0.011	0.71	0.011	0.73	0.013	0.53	0.008	0.87	0.009	0.82
trans FA	0.015	0.86	0.015	0.86	0.015	0.86	0.019	0.74	0.020	0.74	0.032	0.26	0.013	0.90	0.013	0.88
n-3	0.005	0.84	0.005	0.79	0.005	0.81	0.006	0.67	0.007	0.65	0.010	0.26	0.005	0.86	0.005	0.84

[1]PLS = univariate partial least squares.

[2]AG1PLS = genetic algorithm 1 step + univariate PLS.

[3]AG2PLS = genetic algorithm 2 steps + univariate PLS.

[4]Least absolute shrinkage and selection operator

[5]Univariate PLS on the wavelet coefficients.

[6]$s_{y,x}$ = residual SD.

[7]LeA = linoleic acid.

[8]CLA = conjugated linoleic acid.

[9]ALA = α-linolenic acid.

Table 9. Statistical parameters in cow milk for the validation set and equations for the MilkoScan FT+ analyzer (Foss Electric A/S, Hillerød, Denmark) and the Bentley FTS analyzer (Bentley Instruments Inc., Chaska, MN)

Trait (g/100 mL of milk)	Cow milk (MilkoScan FT6000)					Cow milk (Bentley FTS)				
	Method[1]	n[2]	$s_{y,x}$[3]	Relative error	R^2	Method	n	$s_{y,x}$	Relative error	R^2
C4:0	PLS	238	0.006	4.4	0.93	PLS	101	0.010	6.72	0.61
C6:0	AG2PLS	238	0.003	3.6	0.96	PLS	101	0.004	4.31	0.86
C8:0	AG1PLS	242	0.002	4.6	0.96	PLS	104	0.003	5.07	0.89
C10:0	AG1PLS	243	0.007	6.5	0.95	PLS	102	0.011	9.21	0.85
C12:0	AG1PLS	241	0.008	6.6	0.95	PLS	102	0.018	12.42	0.82
C14:0	PLS	238	0.024	5.4	0.94	PLS	102	0.030	6.46	0.84
C16:0	PLS	241	0.066	5.4	0.94	PLS	101	0.111	8.56	0.82
C18:0	PLS	243	0.041	11.9	0.84	PLS	100	0.054	14.79	0.49
Total *trans* 18:1	PLS	239	0.012	14.0	0.85	PLS	101	0.014	17.03	0.51
cis-9 C18:1	PLS	238	0.039	5.8	0.96	PLS	101	0.063	9.13	0.86
Total *cis* 18:1	AG1PLS	239	0.041	5.9	0.96	PLS	102	0.063	8.85	0.87
Total 18:1	PLS	239	0.040	5.2	0.96	PLS	101	0.061	7.75	0.89
cis-9,*cis*-12 C18:2 (LeA[4])	AG1PLS	239	0.006	11.0	0.77	PLS	103	0.006	11.26	0.75
cis-9,*trans*-11 C18:2 (CLA[5])	PLS	238	0.003	16.8	0.82	PLS	98	0.003	19.07	0.64
C18:3n-3 (ALA[6])	PLS	239	0.004	19.3	0.85	PLS	101	0.003	17.67	0.81
SFA	PLS	240	0.035	1.3	1.00	PLS	99	0.090	3.07	0.96
Unsaturated FA	PLS	241	0.038	3.8	0.98	PLS	104	0.100	9.59	0.83
MUFA	PLS	240	0.037	4.2	0.97	PLS	102	0.068	7.49	0.89
PUFA	PLS	240	0.010	9.2	0.76	PLS	101	0.010	9.19	0.60
trans FA	AG2PLS	235	0.015	13.2	0.86	PLS	101	0.018	16.54	0.59
n-3	PLS	239	0.004	14.2	0.86	PLS	100	0.004	13.74	0.81
n-6	PLS	240	0.009	10.8	0.78	PLS	100	0.008	10.34	0.68
n-7	PLS	241	0.011	11.6	0.84	PLS	104	0.013	13.53	0.72
C14 desaturation index[7]	PLS	237	1.408	18.8	0.23	PLS	101	1.549	22.47	0.05
C18:1/C16:0	PLS	239	0.088	12.8	0.90	PLS	101	0.145	21.59	0.76
Sum of C13 + C15 + C17	PLS	241	0.008	10.1	0.74	PLS	101	0.009	11.90	0.60
Elongation index[8]	PLS	240	0.212	6.8	0.85	PLS	104	0.310	9.79	0.71

[1]PLS = univariate partial least squares; AG1PLS: genetic algorithm 1 step + univariate PLS; AG2PLS = genetic algorithm 2 steps + univariate PLS.

[2]Number of samples of calibration data set after elimination of outliers.

[3]$s_{y,x}$ = residual SD of validation data set.

[4]LeA = linoleic acid.

[5]CLA = conjugated linoleic acid.

[6]ALA = α-linolenic acid.

[7]C14 desaturation index = *cis*-9 C14:1/(C14:0 + *cis*-9 C14:1).

[8]Elongation index = (C8:0 + C10:0 + C12:0 + C14:0)/(C4:0 + C6:0).

showed a relative error <7% and R^2_v >0.93 (C4:0, C6:0, C8:0, C10:0, C12:0, C14:0, C16:0, *cis*-9 C18:1, total *cis* 18:1, SFA, MUFAs, PUFA, and unsaturated FA). For conjugated linoleic acid and *trans* FA equations, R^2_v values were appropriate (0.91 and 0.93), but the relative error was >15%.

In goat milk (data from the Foss analyzer), the relative error ranged from 1.7 to 18.4% and R^2_v from 0.70 (for the C14 desaturation index) to 0.99 (for SFA). Thirteen FA showed a relative error <7% and R^2_v >0.93 (C4:0, C6:0, C8:0, C10:0, C14:0, C16:0, cis-9 C18:1, total *cis* 18:1, SFA, MUFA, and unsaturated FA). Six FA had a relative error of approximately 8% and R^2_v of approximately 0.9 (C12:0, *cis*-9,*cis*-12 C18:2, PUFA, n-6, and sum of C13:0 + C15:0 + C17:0).

FA Predictions in Fat (g/100 g of FA). The values, in grams per 100 g of FA, can be useful for some purposes such as herd management, milk pay-

ment, and human nutrition. The estimations obtained in grams per 100 mL of milk were, thus, converted into grams per 100 g of FA and compared with GC data. Performances are presented in Table 11. The conversion affected differently the performance of the equation depending on the individual FA and the species. When compared with those observed in grams per 100 mL of milk, relative errors of estimations expressed in grams per 100 g of FA were similar but some coefficient of determination values were lower (0.68 vs. 0.93 for C4:0 and 0.59 vs. 0.76 for PUFA in cow milk, 0.24 vs. 0.70 for n-3 and 0.58 vs. 0.78 for n-6 in ewe milk, and 0.42 vs. 0.76 for *trans* FA in goat milk).

DISCUSSION

Milk FA Composition

Our reference values for cow milk are consistent with the literature [reviewed by ANSES (2011b) for milk and

Table 10. Statistical parameters in ewe and goat milk for the validation set

Trait (g/100 mL of milk)	Ewe milk					Goat milk				
	Method[1]	n[2]	$s_{y,x}$[3]	Relative error	R^2	Method	n	$s_{y,x}$	Relative error	R^2
C4:0	AG1PLS	139	0.010	4.1	0.93	AG2PLS	229	0.004	4.2	0.96
C6:0	AG2PLS	137	0.005	3.0	0.97	AG2PLS	219	0.004	5.2	0.95
C8:0	AG2PLS	139	0.008	4.4	0.96	PLS	225	0.004	4.8	0.97
C10:0	AG2PLS	139	0.041	6.9	0.93	PLS	223	0.013	4.0	0.98
C12:0	AG1PLS	140	0.019	5.6	0.97	AG2PLS	224	0.013	8.8	0.92
C14:0	AG1PLS	137	0.045	5.4	0.96	AG2PLS	227	0.023	6.4	0.93
C16:0	AG1PLS	137	0.091	5.4	0.94	AG2PLS	229	0.042	4.3	0.96
C18:0	PLS	137	0.061	11.9	0.83	AG2PLS	228	0.034	11.1	0.86
Total *trans* 18:1	PLS	139	0.031	17.4	0.91	AG2PLS	223	0.011	14.3	0.76
cis-9 C18:1	AG2PLS	139	0.057	5.4	0.97	AG2PLS	219	0.037	5.7	0.95
Total *cis* 18:1	AG2PLS	139	0.062	5.6	0.97	AG2PLS	224	0.034	5.2	0.96
Total 18:1	AG2PLS	139	0.049	3.8	0.99	AG2PLS	222	0.033	4.5	0.96
cis-9,*cis*-12 C18:2 (LeA[4])	AG1PLS	140	0.012	13.6	0.49	AG2PLS	219	0.007	9.2	0.89
cis-9,*trans*-11 C18:2 (CLA[5])	PLS	135	0.011	19.0	0.91	PLS	224	0.003	16.5	0.71
C18:3n-3 (ALA[6])	PLS	137	0.007	13.4	0.74	AG2PLS	220	0.003	16.2	0.79
SFA	AG2PLS	135	0.049	1.0	1.00	AG2PLS	224	0.043	1.7	0.99
Unsaturated FA	PLS	140	0.048	3.0	0.99	PLS	221	0.039	4.3	0.97
MUFA	AG2PLS	138	0.044	3.1	0.99	PLS	218	0.037	4.7	0.96
PUFA	AG2PLS	140	0.015	6.3	0.96	PLS	220	0.010	7.3	0.92
trans FA	PLS	135	0.041	15.2	0.93	AG2PLS	227	0.016	14.3	0.76
n-3	AG2PLS	140	0.009	13.3	0.70	AG2PLS	222	0.004	14.5	0.82
n-6	AG2PLS	136	0.017	12.3	0.78	AG2PLS	221	0.010	8.8	0.91
n-7	PLS	137	0.016	10.3	0.93	PLS	228	0.005	10.1	0.83
C14 desaturation index[7]	PLS	137	0.355	22.3	0.58	PLS	225	0.221	18.4	0.70
C18:1/C16:0	PLS	137	0.085	11.0	0.89	PLS	219	0.055	7.3	0.75
Sum of C13 + C15 + C17	PLS	139	0.013	9.9	0.84	PLS	218	0.004	7.3	0.88
Elongation index[8]	PLS	138	0.264	5.7	0.89	PLS	227	0.339	6.7	0.89

[1]PLS = univariate partial least squares; AG1PLS = genetic algorithm 1 step + univariate PLS; AG2PLS = genetic algorithm 2 steps + univariate PLS.

[2]Number of samples of calibration data set after elimination of outliers.

[3]$s_{y,x}$ = residual SD for validation data set.

[4]LeA = linoleic acid.

[5]CLA = conjugated linoleic acid.

[6]ALA = α-linolenic acid.

[7]C14 desaturation index = *cis*-9 C14:1/(C14:0 + *cis*-9 C14:1).

[8]Elongation index = (C8:0 + C10:0 + C12:0 + C14:0)/(C4:0 + C6:0).

butter]. The average FA profiles observed are similar to those observed by Chilliard et al. (2007) and Heck et al. (2009). The average SFA content of about 71 ± 4.5 g/100 g of FA in cow milk is in agreement with previous results observed in controlled feeding systems (Hurtaud et al., 2009) and in commercial farms (Heck et al., 2009). In the study of Coppa et al. (2013) about bulk FA composition in 10 European countries, a lower SFA concentration (63.9 g/100 g of FA, minimum = 53.40, maximum = 75.42 g/100 g of FA) and a higher PUFA concentration (3.92 g/100 g of FA, minimum = 2.11, maximum = 7.04 g/100 g of FA) were reported. These differences could be explained by the feeding systems that were mainly based on fresh herbage or conserved forage with grass in the study of Coppa et al. (2013) and the localization of the farms, several of them being located in the mountains. If the average concentrations are different, the range of the values (minimum to

maximum) are nevertheless consistent, except for the PUFA. In Soyeurt et al. (2011) on individual milk, the reported SFA content was lower (67 ± 6.01 g/100 g of fat in the calibration set and 65 ± 4.35 g/100 g of fat in the validation set), but some of the samples used were collected in Ireland where the feeding system was mainly grazed herbs.

For ewe milk, our GC results could be compared with those obtained previously for populations of Lacaune-Sardinian backcross (Carta et al., 2008), Churra ewes (De La Fuente et al., 2009), and the Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail analysis (ANSES, 2011b). The average FA contents obtained in the present study are comparable with those in the literature. However, higher SFA (74 g/100 g of FA), C14:0 (12.5 g/100 g of FA), and C16:0 (26.6 g/100 g of FA) concentrations were observed compared with previous reports, except for C18:0 (7.6 g/100

Table 11. Statistical parameters for the estimation (g/100 g of FA) obtained by dividing the estimation (g/100 mL) by the fat content

FA (g/100 g of FA)	Cow milk		Ewe milk		Goat milk	
	Relative error	R^2	Relative error	R^2	Relative error	R^2
C4:0	4.02	0.68	3.98	0.91	4.27	0.89
C6:0	2.81	0.91	3.07	0.93	4.11	0.84
C8:0	4.00	0.95	3.97	0.88	5.22	0.84
C10:0	6.42	0.93	5.74	0.88	4.83	0.91
C12:0	7.25	0.92	4.62	0.95	8.20	0.81
C14:0	5.13	0.89	5.39	0.82	7.14	0.80
C16:0	4.97	0.88	5.04	0.85	5.29	0.69
C18:0	11.29	0.75	10.91	0.73	11.33	0.55
Total *trans* 18:1	14.18	0.76	16.62	0.87	15.69	0.60
cis-9 C18:1	4.74	0.96	5.45	0.95	6.68	0.75
Total *cis* 18:1	5.00	0.95	5.70	0.94	5.99	0.78
Total 18:1	4.15	0.96	3.91	0.97	4.95	0.79
cis-9,*cis*-12 C18:2 (LeA[1])	11.66	0.61	14.12	0.43	11.62	0.71
cis-9,*trans*-11 C18:2 (CLA[2])	16.28	0.78	19.68	0.86	18.02	0.35
C18:3n-3 (ALA[3])	17.52	0.87	13.44	0.44	19.14	0.78
SFA	1.08	0.98	1.08	0.98	1.75	0.79
MUFA	3.19	0.97	2.92	0.98	5.58	0.74
PUFA	8.89	0.59	6.51	0.90	7.94	0.77
trans FA	12.68	0.80	14.26	0.92	16.01	0.42
n-3	13.48	0.87	12.52	0.24	16.19	0.81
n-6	10.84	0.63	11.96	0.58	8.65	0.75

[1]LeA = linoleic acid.

[2]CLA = conjugated linoleic acid.

[3]ALA = α-linolenic acid.

g of FA), which was actually lower. The concentration of total C18:1 (18.9 g/100 g of FA) was found to be lower than in previous reports and contributed to reducing the MUFA concentrations observed in this study (20.6 g/100 g of FA). These differences can be explained by the breeds that are not the same between the 3 studies. In all cases, content variability was equivalent or higher than that reported in the studies mentioned above.

In goats, the average FA compositions of our samples are in agreement with previous studies (Chilliard et al., 2006; Sanz Ceballos et al., 2009; Andueza et al., 2013) and the range of variation is consistent with the results published for goats from experimental designs that did not receive lipid supplementation (Chilliard et al., 2003) and goats from commercial herds (Alonso et al., 1999; Doyon, 2005).

Comparison of the Mathematical Approaches on Cow Milk

Different mathematical approaches were performed to develop the equations in cow milk. Applying penalization methods before PLS provided less reliable results in almost all cases, whereas applying a GA or first-derivative or wavelet preprocessing improved the results for several FA. The improvement due to first-derivative preprocessing was similar to that ob-

served by Soyeurt et al. (2011), where first-derivative preprocessing improved the accuracy of the equations. Compared with first-derivative preprocessing, the GA has the advantage of applying null coefficients to discarded wavelengths, avoiding spoiling predictions with noisy wavelengths (Leardi and Lupiáñez González, 1998; Spiegelman et al., 1998; Devos and Duponchel, 2011). The wavelet method is an intermediate method that combines preprocessing and wavelength selection. It provided positive results and meant that irrelevant wavelengths could be discarded, but its interpretation was challenging. To the best of our knowledge, it was the first time that wavelet method was used on MIR data to develop a calibration equation. This method seems promising, but further studies should be performed to have a deeper understanding of its limits. Another approach would be to optimize simultaneously the preprocessing and the selection of wavelengths by a parallel GA (Devos and Duponchel, 2011) or at least apply a GA on derivative spectra. Therefore, we chose to retain the best method between GA and PLS for each FA, based on the residual standard error.

Accuracy of the Equations

The results for cow milk on the MilkoScan FT6000 analyzer were comparable in terms of coefficient of

determination and residual standard error with the results reported by Soyeurt et al. (2011), and equivalent or slightly better than those described by De Marchi et al. (2011) and Maurice-Van Eijndhoven et al. (2013) that included different breeds. As previously reported by Soyeurt et al. (2006) and Coppa et al. (2010), estimations were better for the FA present at higher concentrations, except for C18:0, which showed a high relative error. A similar result was observed previously by Maurice-Van Eijndhoven et al. (2013). In cow milk, the results obtained on the Bentley FTS analyzer were less accurate—this was probably due to the lower amounts of data produced (150 instead of 350) and nonoptimization due to the lack of data.

Currently, no specific data are available to compare the results for ewe and goat milk, but the performances are similar to those observed for cow milk, with a more accurate prediction of PUFA. A recent study described the results for goat milk obtained with NIRS (Andueza et al., 2013). The results were expressed in grams per 100 g of FA, and are comparable to those presented in Table 11 for some FA (short- and medium-chain FA). For SFA, MUFA, and PUFA, NIRS seems to provide better results, but this technique requires preliminary dehydration of the samples and would be more difficult to implement on a large scale.

Use of the Equations

These equations could be used routinely for different purposes such as milk payment, breeding, and advice on farming management. To be used for milk payment, highly accurate and precise estimations need to be obtained. The requirements for genetic applications are less stringent and the equations could be used to discriminate milk with different contents even if they are less accurate. Soyeurt et al. (2011) indicated that equations with an R_v^2 greater than 0.95 could be used in milk payment systems and in animal breeding if greater than 0.75. As underlined by Cecchinato et al. (2009) and Maurice-Van Eijndhoven et al. (2013) biased estimations with low R_v^2 can be suitable in breeding programs, particularly when several measurements are done on the same individual. A genetic study based on MIR estimations from our equations (Gion et al., 2011) showed similar heritability to those obtained with GC data and other MIR estimations, which indicates that our estimations are relevant to breeding. Regarding farming management, the PhénoFinlait study confirmed that diet, parity, and stage of lactation had significant effects on the estimated FA composition of milk in the 3 species (Esvan et al., 2010). These results provided evidence of the relevance of our equations to

estimate the FA composition of milk on a large scale and in a large diversity of farming systems. However, 2 application limits had to be underlined: first, a periodic adjustment of the estimations or a spectral standardization is necessary to ensure that the results remain comparable over time and over a large geographical region, to be used for long-term actions on populations spread on a wide territory (Leray et al., 2011; Fernández, 2012); second, to apply these equations to local breeds or feeding systems not included in this set of samples, it would be necessary to validate the equations in a first step and, if necessary, to add new samples to increase the calibration database variability.

For some purposes such as milk payment or farming management, FA composition has to be expressed as a proportion of total fat (g/100 g of FA). As underlined in several studies (Soyeurt et al., 2006; Rutten et al., 2009), estimations of FA contents expressed in grams per 100 g of FA are less accurate than when expressed in grams per 100 mL of milk. The estimations from equations developed here were first obtained in liquid whole milk (g/100 mL of milk) and then converted into grams per 100 g of FA. Our results show that this conversion did not lead a great loss of accuracy for most FA, in agreement with Soyeurt et al. (2011).

CONCLUSIONS

These results show that accurate estimations can be obtained for several FA in individual samples of cow, ewe, and goat milk using MIR spectra with, in the 3 species, an R_v^2 greater than 0.90 for C4:0, C6:0, C8:0, C10:0, C12:0, C14:0, *cis*-9 C18:1, SFA, unsaturated FA, and MUFA. These equations are the first referenced for ewe and goat milk. From a methodological point of view, it seems important to remove spectral noise to improve the robustness of the equations. This can be carried out efficiently by selecting informative wavelengths using GA that apply null coefficients to noninformative wavelengths and, consequently, avoid accumulating the errors linked to spectra.

REFERENCES

Alonso, L., L. Fontecha, L. Lozada, M. J. Fraga, and M. Juarez. 1999. Fatty acid composition of caprine milk: Major, branched chain, and *trans* fatty acids. J. Dairy Sci. 82:878–884.

Andueza, D., J. Rouel, Y. Chilliard, C. Leroux, and A. Ferlay. 2013. Prediction of the goat milk fatty acids near infrared reflectance spectroscopy. Eur. J. Lipid Sci. Technol. 115:612–620.

ANSES (Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail). 2011a. Actualisation des apports nutritionnels conseillés pour les acides gras. Accessed Dec. 3, 2012. http://www.anses.fr/Documents/NUT2006sa0359Ra.pdf.

ANSES (Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail). 2011b. Impact des pratiques en alimentation animale sur la composition en acides gras des produits animaux destinés à l'Homme. Accessed Apr. 04, 2013. http://www.anses.fr/Documents/ALAN2005sa0300Ra.pdf.

AOAC International. 2000. Official Methods of Analysis. 17th ed. AOAC International, Gaithersburg, MD.

Arnould, V. M., and H. Soyeurt. 2009. Genetic variability of milk fatty acids. J. Appl. Genet. 50:29–39.

Carta, A., S. Casu, M. G. Usai, M. Addis, M. Fiori, A. Fraghi, S. Miari, L. Mura, G. Piredda, L. Schibler, T. Sechi, J. M. Elsen, and F. Barillet. 2008. Investigating the genetic component of fatty acid content in sheep milk. Small Rumin. Res. 79:22–28.

Cecchinato, A., M. De Marchi, L. Gallo, G. Bittante, and P. Carnier. 2009. Mid-infrared spectroscopy predictions as indicator traits in breeding programs for enhanced coagulation properties of milk. J. Dairy Sci. 92:5304–5313.

Chilliard, Y., A. Ferlay, J. Rouel, and G. Lamberet. 2003. A review of nutritional and physiological factors affecting goat milk lipid synthesis and lipolysis. J. Dairy Sci. 86:1751–1770.

Chilliard, Y., F. Glasser, A. Ferlay, L. Bernard, J. Rouel, and M. Doreau. 2007. Diet, rumen biohydrogenation and nutritional quality of cow and goat milk fat. Eur. J. Lipid Sci. Technol. 109:828–855.

Chilliard, Y., J. Rouel, and C. Leroux. 2006. Goat's alpha-s1 casein genotype influences its milk fatty acid composition and delta-9 desaturation ratios. Anim. Feed Sci. Technol. 131:474–487.

Coppa, M., A. Ferlay, C. Chassaing, C. Agabriel, F. Glasser, Y. Chilliard, G. Borreani, R. Barcarolo, T. Baars, D. Kusche, O. M. Harstad, J. Verbič, J. Golecký, and B. Martin. 2013. Prediction of bulk milk fatty acid composition based on farming practices collected through on-farm surveys. J. Dairy Sci. 96:4197–4211.

Coppa, M., A. Ferlay, C. Leroux, M. Jestin, Y. Chilliard, B. Martin, and D. Andueza. 2010. Prediction of milk fatty acid composition by near infrared reflectance spectroscopy. Int. Dairy J. 20:182–189.

Croiseau, P., A. Legarra, F. Guillaume, S. Fritz, A. Baur, C. Colombani, C. Robert-Granié, D. Boichard, and V. Ducrocq. 2011. Fine-tuning genomic evaluations in dairy cattle through SNP preselection with the Elastic-Net algorithm. Genet. Res. (Camb.) 93:409–417.

De La Fuente, L. F., E. Barbosa, J. A. Carriedo, C. Gonzalo, R. Arenas, J. M. Fresno, and F. San Primitivo. 2009. Factors influencing variation of fatty acid content in ovine milk. J. Dairy Sci. 92:3791–3799.

De Marchi, M., M. Penasa, A. Cecchinato, M. Mele, P. Secchiari, and G. Bittante. 2011. Effectiveness of mid-infrared spectroscopy to predict fatty acid composition of Brown Swiss bovine milk. Animal 5:1653–1658.

Devos, O., and L. Duponchel. 2011. Parallel genetic algorithm co-optimization of spectral pre-processing and wavelength selection for PLS regression. Chemom. Intell. Lab. Syst. 107:20–58.

Donoho, D. L., and I. M. Johnstone. 1994. Ideal spatial adaptation via wavelet shrinkage. Biometrika 81:425–455.

Doyon, A. 2005. Influence de l'alimentation sur la composition du lait de chèvre: Revue des travaux récents. Colloque sur la chèvre. L'innovation: Un outil de croissance. Centre de référence en agriculture et agroalimentaire du Québec (CRAAQ), Saint-Hyacinthe, Québec, Canada.

Esvan, S., C. Dragan, A. Varenne, J.-M. Astruc, F. Barillet, D. Boichard, P. Brunschwig, A. Dubrulle, F. Faucon-Lahalle, A. Ferlay, G. Lagriffoul, H. Larroque, J. Legarto, I. Palhière, J.-L. Peyraud, R. Rupp, and M. Brochard. 2010. PhénoFinlait, 1ers résultats: Influence de l'alimentation, de l'état physiologique et de la génétique sur la composition en acides gras des laits de vache, brebis et chèvre. 17eme Renc. Rech. Rum., Paris, France.

Faucon-Lahalle, F., M. Brochard, F. Barillet, M. Bolard, P. Brunschwig, K. Duhem, A. Eggen, S. Esvan, M. Ferrand, S. Fritz, P. L. Gastinel, J. L. Guerin, L. Journaux, T. Krychowski, G. Lagriffoul, H. Larroque, C. Lecomte, O. Leray, S. Leverrier, P. Martin, S. Mattalia, G. Miranda, I. Palhière, J. L. Peyraud, and D. Boichard. 2009. PhenoFinLait (LactoScan): French national program for high scale phenotyping and genotyping to detect QTL linked with fine composition of ruminant milk. 6th International Milk Genomics Consortium Symposium, Paris, France.

Fernández, J. A. 2012. Merging spectral datasets from different MIR instruments used in the routine analysis of milk. Pages 55–71 in Proc. Reference Laboratory Network Meeting 2012, Cork, Ireland. ICAR Technical Series No. 16. International Committee for Animal Recording (ICAR), Rome, Italy.

Ferrand, M., B. Huquet, S. Barbey, F. Barillet, F. Faucon, H. Larroque, O. Leray, J. M. Trommenschlager, and M. Brochard. 2010. Determination of fatty acid profile in cow's milk using mid-infrared spectrometry: Interest of applying a variable selection by genetic algorithms before a PLS regression. Chemom. Intell. Lab. Syst. 106:183–189.

Gion, A., H. Larroque, M. Brochard, F. Lahalle, and D. Boichard. 2011. Genetic parameter estimation for milk fatty acids in three French dairy cattle breeds. Interbull, Stavanger, Norway.

Goicoechea, H. C., and A. C. Olivieri. 2003. A new family of genetic algorithms for wavelength interval selection in multivariate analytical spectroscopy. J. Chemometr. 17:338–345.

Grubbs, F. 1969. Procedures for detecting outlying observations in samples . Technometrics 11:1–21.

Heck, J. M. L., H. J. F. van Valenberg, J. Dijkstra, and A. C. M. van Hooijdonk. 2009. Seasonal variation in the Dutch bovine raw milk composition. J. Dairy Sci. 92:4745–4755.

Höskuldsson, A. 2001. Variable and subset selection in PLS regression. Chemom. Intell. Lab. Syst. 55:23–38.

Hurtaud, C., and J. L. Peyraud. 2007. Effects of feeding camelina (seeds or meal) on milk fatty acid composition and butter spreadability. J. Dairy Sci. 90:5134–5145.

Hurtaud, C., J. L. Peyraud, G. Michel, D. Berthelot, and L. Delaby. 2009. Winter feeding systems and dairy cow breed have an impact on milk composition and flavor of two Protected Designation of Origin French cheeses. Animal 3:1327–1338.

IDF (International Dairy Federation). 2000. Standard 141C: Whole milk—Determination of milkfat, protein and lactose content—Guidance on the operation of mid-infrared instruments. IDF, Brussels, Belgium.

INRA (Institut National de la Recherche Agronomique). 1989. Feeding standards for ruminants. Pages 15–22 in Ruminant Nutrition. Recommended Allowances and Feed Table. J. Libbey, ed. Eurotext, London, UK.

ISO-IDF (International Organization for Standardization-International Dairy Federation). 2001. Milk and milk products—Extraction methods for lipids and liposoluble compounds. ISO 14156-IDF 172:2001. IDF, Brussels, Belgium.

ISO-IDF (International Organization for Standardization-International Dairy Federation). 2002a. Milk fat—Preparation of fatty acid methyl esters. ISO 15884-IDF 182:2002. IDF, Brussels, Belgium.

ISO-IDF (International Organization for Standardization-International Dairy Federation). 2002b. Milk fat—Determination of the fatty acid composition by gas-liquid chromatography. ISO 15885-IDF 184. IDF, Brussels, Belgium.

ISO-IDF (International Organization for Standardization-International Dairy Federation). 2009. Milk—Definition and evaluation of the overall accuracy of alternative methods of milk analysis—Part 2: Calibration and quality control in the dairy laboratory. ISO 8196-2-IDF 128-2:2009. IDF, Brussels, Belgium.

Kramer, J. K. G., V. Fellner, M. E. R. Dugan, F. D. Sauer, M. M. Mossoba, and M. P. Yurawecz. 1997. Evaluating acid and base catalysts in the methylation of milk and rumen fatty acids with special emphasis on conjugated dienes and total *trans* fatty acids. Lipids 32:1219–1228.

Larroque, H., Y. Gallard, L. Thaunat, D. Boichard, and J. J. Colleau. 2002. A crossbreeding experiment to detect quantitative trait loci in dairy cattle. Proc. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France. INRA, Castanet-Tolosan, France.

Leardi, R. 2000. Application of genetic algorithm—PLS for feature selection in spectral data sets. J. Chemometr. 14:643–655.

Leardi, R., R. Boggia, and M. Terrile. 1992. Genetic algorithm as a strategy for feature selection. J. Chemometr. 6:267–281.

Leardi, R., and A. L. Lupiáñez González. 1998. Genetic algorithms applied to feature selection in PLS regression: How and when to use them. Chemom. Intell. Lab. Syst. 41:195–207.

Leray, O., M. Ferrand, H. Larroque, J. M. Astruc, M. Douguet, M. Brochard, and K. Duhem. 2011. Harmonisation of milk analysers for fatty acid determination by FTMIR—An essential step prior to collective data use. ICAR Meeting, Bourg-en-Bresse, France. Accessed Nov. 6, 2013. http://www.icar.org/Documents/Bourg-en-Bresse2011/Presentations/session%204%20-%2023%20am/2%20Olivier%20Leray.pdf.

Mallat, S. 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell. 11:674–693.

Mallat, S. 2008. A Wavelet Tour of Signal Processing: The Sparse Way. Academic Press, Burlington, MA.

Maurice-Van Eijndhoven, M. H. T., H. Soyeurt, F. Dehareng, and M. P. L. Calus. 2013. Validation of fatty acid predictions in milk using mid-infrared spectrometry across cattle breeds. Animal 7:348–354.

Mevik, B.-H., and R. Wehrens. 2007. The pls package: Principal component and partial least squares regression in R. J. Stat. Softw. 18:1–24.

Raynal-Ljutovac, K., G. Lagriffoul, P. Paccard, I. Guillet, and Y. Chilliard. 2008. Composition of goat and sheep milk products: An update. Small Rumin. Res. 79:57–72.

Rutten, M. J. M., H. Bovenhuis, K. A. Hettinga, H. J. van Valenberg, and J. A. van Arendonk. 2009. Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. J. Dairy Sci. 92:6202–6209.

Sanz Ceballos, L., E. Ramos Morales, G. de la Torre Adarve, J. Díaz Castro, L. Pérez Martínez, and M. R. Sanz Sampelayo. 2009. Composition of goat and cow milk produced under similar conditions and analyzed by identical methodology. J. Food Compost. Anal. 22:322–329.

Schennink, A., W. M. Stoop, M. H. Visker, J. J. van der Poel, H. Bovenhuis, and J. A. van Arendonk. 2009. Short communication: Genome-wide scan for bovine milk-fat composition. II. Quantitative trait loci for long-chain fatty acids. J. Dairy Sci. 92:4676–4682.

Soyeurt, H., P. Dardenne, F. Dehareng, G. Lognay, G. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, and N. Gengler. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. J. Dairy Sci. 89:3690–3695.

Soyeurt, H., F. Dehareng, N. Gengler, S. McParland, E. Wall, D. P. Berry, M. Coffey, and P. Dardenne. 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. J. Dairy Sci. 94:1657–1667.

Spiegelman, C. H., M. J. McShane, M. J. Goetz, M. Motamedi, Q. L. Yue, and G. L. Coté. 1998. Theoretical justification of wavelength selection in PLS calibration: Development of a new algorithm . Anal. Chem. 70:35–44.

Stoop, W. M., A. Schennink, M. H. Visker, E. Mullaart, J. A. van Arendonk, and H. Bovenhuis. 2009. Genome-wide scan for bovine milk-fat composition. I. Quantitative trait loci for short- and medium-chain fatty acids. J. Dairy Sci. 92:4664–4675.

Tenenhaus, M. 2002. La Regression PLS: Théorie et Pratique. Technip, Lassay-les-Châteaux, France.

Zou, H., and T. Hastie. 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B Stat. Methodol. 67:301–320.