Electronic Journal of Applied Statistical Analysis
EJASA (2011), Electron. J. App. Stat. Anal., Vol. 4, Issue 2, 245 – 254
e-ISSN 2070-5948, DOI 10.1285/i20705948v4n2p245

USE OF GENETIC ALGORITHM ON MID-INFRARED SPECTROMETRIC DATA: APPLICATION TO ESTIMATE THE FATTY ACIDS PROFILE OF GOAT MILK

Marion Ferrand[*(1)], Bérénice Huquet[(1)], Frédéric Bouvier[(2)], Hugues Caillat[(3)], Francis Barillet[(3)], Félicie Faucon - Lahalle[(1,4)], Hélène Larroque[(3)], Olivier Leray[(5)], Isabelle Palhiere[(3)], Mickaël Brochard[(1)]

[(1)]Institut de l'Elevage, France
[(2)]INRA - UE Bourges, France
[(3)]INRA - SAGA, France
[(4)]CNIEL, France
[(5)]Actilait, France

Received 01 October 2010; Accepted 28 April 2011
Available online 14 October 2011

Abstract: *To know and to control the fine milk composition is an important concern in the dairy industry. The mid-infrared (MIR) spectrometry method appears to be a good, fast and cheap method for assessing milk fatty acid profile with accuracy. Although partial least squares (PLS) regression is a very useful and powerful method to determine fine milk composition from spectra, the estimations are often less accurate on new samples coming from different spectrometers. Therefore a genetic algorithm (GA) combined with a PLS was used to produce models with a reduced number of wavelengths and a better accuracy. Number of wavelengths to consider is reduced substantially by 5 or 10 according the number of steps in the genetic algorithm. The accuracy is increased on average by 9% for fatty acids of interest.*

Keywords: *Mid-infrared (MIR) spectrometry, goat milk, fatty acid, genetic algorithms, Partial Least Squares (PLS) regression.*

[*] Corresponding author. E-mail: marion.ferrand@inst-elevage.asso.fr

1. Introduction

Milk is a complex product with a large number of components such as proteins, fatty acids, lactose, minerals in variable concentrations. For many years, milk has been considered as a raw material. At the farm level, the target was to produce milk with high overall protein content and a given fat content. Measurement and selection procedures have been developed and implemented in this way. More recently, emphasis has been put on milk elementary components since many of them have confirmed effects on human health. Levers to adapt products to the changing demands of the market are mainly genetics, feeding and food technology. One of the present limitations for answering to these demands is the lack of fast, low cost and sensitive phenotyping techniques. In this context, scientific (INRA, Institut de l'Elevage, Actilait) and economic stakeholders, from milk production (milk recording and DHI organizations, milk testing laboratories, breeding organizations, artificial insemination organizations, extension services) to milk processing (federation of dairy factories) gathered in the PhenoFinLait project. The aim of this vast program is to develop a cheap and large scale phenotyping procedure for individual milk components (fatty acids and proteins) and to apply this procedure on a specific design in farms allowing an analysis of the genetic and the environmental factors involved in the milk's composition. The expected result is a new and highly innovative way to drive milk composition, both by animal selection and herd management to fulfill the requirements of human nutritional needs.

In goat, the milk fatty acid composition differs from cow, and is characterized by a higher concentration of short and medium chain fatty acids and a lower level of palmitic acid (C16:0) [1]. As in other dairy species, fatty acid composition of goat milk is highly dependent on the diet and more particularly forages and lipid supplementation [2, 3]. However, food seems to have specific effects on milk's fatty acid composition of goat compared with that of dairy cattle or ewe [4]. Goat is also particular because of its polymorphism at the αs1 casein gene which is responsible for quantitative variations of milk protein content and also milk fat content and its fatty acid composition [5]. In this context, the PhenoFinLait project represents an opportunity to study, at a large scale, the fatty acid composition of goat milk and to better describe the factors affecting it. At the same time, it will allow to detect, for the first time in goat, QTL or gene responsible for the variation in milk fatty acid composition.

The study presented in this article deals with the development of a reliable, cheap and easy-to-use method for individual milk fatty acid content measurement.

Soyeurt et al. [6,7] has shown the possibility to estimate cow milk fatty acid content from Mid Infra-Red (MIR) spectra currently measured for milk fat and protein contents determination by milk testing laboratories in milk recording schemes. The method used to develop the equations is the Partial Least Squares (PLS) regression. Spiegelman et al. [11] showed the theoretical interest of eliminating noisy wavelengths before applying a PLS regression and several authors [8, 9, 10] suggested the use of genetic algorithms (GA) to select the useful wavelengths. In this optic, in a previous study we used a genetic algorithm to improve the accuracy of the estimation in cow milk (Ferrand et al., in press). The number of wavelengths to consider was reduced substantially by 4 and accuracy was increased on average by 15%. The aim of this new study is to check whether or not it is also possible to improve the estimation of fatty acid profile in goat milk. With PLS regression, the estimation quality in goat milk is not as good as the estimations in cow milk. It may be linked to the number of cells in milk or to the lower content of fatty acids in goat

milk. By using a genetic algorithm, it could be possible to select the informative wavelengths only and to improve the estimations of fatty acid profile in goat milk [11].

2. Wavelengths selection by genetic algorithm

Genetic algorithms (GA) are often used to solve optimization problems where we search a pool of solutions among the best. This method is based on evolutionary biology [12, 13]. A population of candidate solutions evolves using genetic operators like reproduction, mutation and selection. A solution, so-called chromosome, is a vector where each variable, so-called gene, is coded with 0 (not-selected) or 1 (selected). Initial population has a predefined number of candidate solutions. The evolution is controlled by a fitness function. To breed a new generation (two new solutions), two candidate solutions are selected. During this step of reproduction, crossing-over (2 candidates solutions are mixed to create 2 new ones) or mutation (a gene coded 1 mutates and is coded 0; and inversely) could occur. The obtained solutions integrate the population if they appear better than the previous solutions. The population is constant, so the worst solutions are discarded when new solutions integrate the population. This process is repeated until the fixed number of generations is reached. To ensure an optimal convergence GA is run several times. For more details, it is possible to consult the book from Haupt [14] and the article from Leardi [8].

3. Material and methods

3.1 Milk samples

705 milk samples from 235 Alpine dairy goats were collected in 2008 at the INRA experimental farm of Bourges at three stage of lactation (about 40, 150 and 240 days). The goat's diet was almost similar throughout lactation and was based on grass hay offered ad libitum and a commercial concentrate mixture. These samples were collected in tubes containing a preservative (Bronopol).

For each goat, one sample was analyzed by MIR spectrometry, and one other was frozen at -20°C. Among them, 149 samples (about 50 per stage of lactation) with a large variability of spectra were selected to be analyzed for milk fatty acid composition by the referenced method.

In complementary, 50 milk samples were collected in 2010 in private farm and were used as independent dataset.

3.2 MIR spectra

After a transport at 4°C to the laboratory (LILCO of Surgères), fresh milk samples were analyzed for milk spectra extraction using MIR spectrometry with defined routine FT-MIR analyzers (Milkoscan FT6000). Spectra have been recorded from 5012 to 926 cm-1. According to Foss [15], only informative wavelength bands, i.e. bands not spoiled by water molecule, were kept (representing a total of 446 wavelengths). No pre-treatments were applied as suggested by Soyeurt and al. [6].

3.3 Fatty acid composition

Frozen milk samples were analyzed for milk fatty acid composition using gas chromatography according to ISO standards (ISO 14156 | IDF 172 for the fat content extraction, ISO 15884 | FIL 182 for the preparation of fatty acid methyl esters from milk fat and ISO 15885 | FIL 184 for the gas-liquid chromatography). Analyses were realized by the laboratory LARF (Mamirolle). The analysis protocol is the same as described by Kramer [16] with the following material : Chromatograph VARIAN 3800 ; column CP-SIL 88 de 100 m ; hydrogen carrier gaz (29 PSI); splitless injector 1:50 at 250°C ; oven temperature program : 4 mn at 70°C, 13°C/mn between 70-175°C, 27 mn at 175°C, 4°C/min between 175-215°C, 31 mn at 215°C ; flam ionization detector at 250°C : integrator STAR from VARIAN.

Around 75 individual or groups of fatty acids were detected by CPG. The quantities expressed in fat percentage were converted in g/100mL by using the fat content given by the spectrometer MilkoScan FT6000.

3.4 Calculation of calibration equations

MIR spectra and milk fatty acid composition of samples presenting a large variability in their composition were retained to calculate the equations.

3.4.1 Reference method

These equations were developed by univariate and multivariate PLS regression [17], data being centered but not reduced according to Bertrand et al. [18]. For each equation, optimal number of latent variables was chosen according to root mean square error of cross-validation (RMSEPcv). In first approach a leave-one-out cross-validation was used to validate the equations. In a confirmation step, an independent dataset was used.

PLS regressions were performed with the package PLS [19] on R 2.8.1. [20]

3.4.2 Genetic algorithm combined with reference method

The algorithm used in this paper is the algorithm developed by Leardi [8] which is specific to wavelengths selection. The same levels of parameters were kept. In our previous work, we have checked three parameters (mutation rate, initial population size and number of variables selected in the solution of initial population), for 6 fatty acids [21] and finally we retained the same parameters that Leardi [8]. Mutation rate, initial population, and number of variables selected in the solution of initial population were fixed to 1%, 30 and 5 respectively.

To reduce the risk of overfitting we have performed the algorithm in two steps as described by Leardi and 5 independent runs were carried out:

- First step: on the average of 3 contiguous wavelengths
- Second step: on wavelengths selected in first step

We performed the algorithm on standardized data. The cross-validation used for this step was a stratified 5-fold cross-validation.

Following the variables selections, PLS regressions were applied as described before. For each fatty acid, univariate and multivariate PLS were tested. We kept the best method according to the leave-one-out cross-validation results.

GA were performed with MATLAB 7.8 [22]

3.4.3 Comparison of analysis methods

To compare and to assess the equations for each method, several summary statistics were computed: mean, standard deviation (Sd), standard error of cross-validation (SE_{CV}), and cross-validation coefficient of determination (R^2_{CV}). We considered that an estimation was precise enough and robust to be applied in routine, when R^2_{CV} was upper than 0.80. For R^2_{CV} in the range of 0.70 to 0.80, we advise to use these equations with caution. The accuracy was checked according to SE_{CV} criteria. To evaluate the performance of genetic algorithm, we compared the SE_{CV} of the model issued of PLS regression on variables selected by GA with the SE_{CV} of the multivariate PLS regression without variable selection.

The equations retained were tested on a new dataset to confirm their relevance. The validation was done according to the ISO norm 8196 (ISO 8196|IDF 128:2009 Milk - Definition and evaluation of the overall accuracy of alternative methods of milk analysis.)

4. Results and discussion

For the fatty acids of interest, GA selected on average 51 variables out of 446 in the form of wavelength bands. The number of selected wavelengths varied from 10 to 140 according to the fatty acids, because the optimal selection (the lowest SE_{CV}) is, for some fatty acids, after the first step (GA carried out on the average of 3 contiguous wavelengths) and not after the second step (on wavelengths selected in the first step). The 2272-1792 cm^{-1} band was rarely selected, while the 2970-2434 cm^{-1} and the 1344-926 cm^{-1} bands were selected for most fatty acids (Figure 1). These spectra areas are associated to chemical bonds of methyl and methylene groups of fat content.

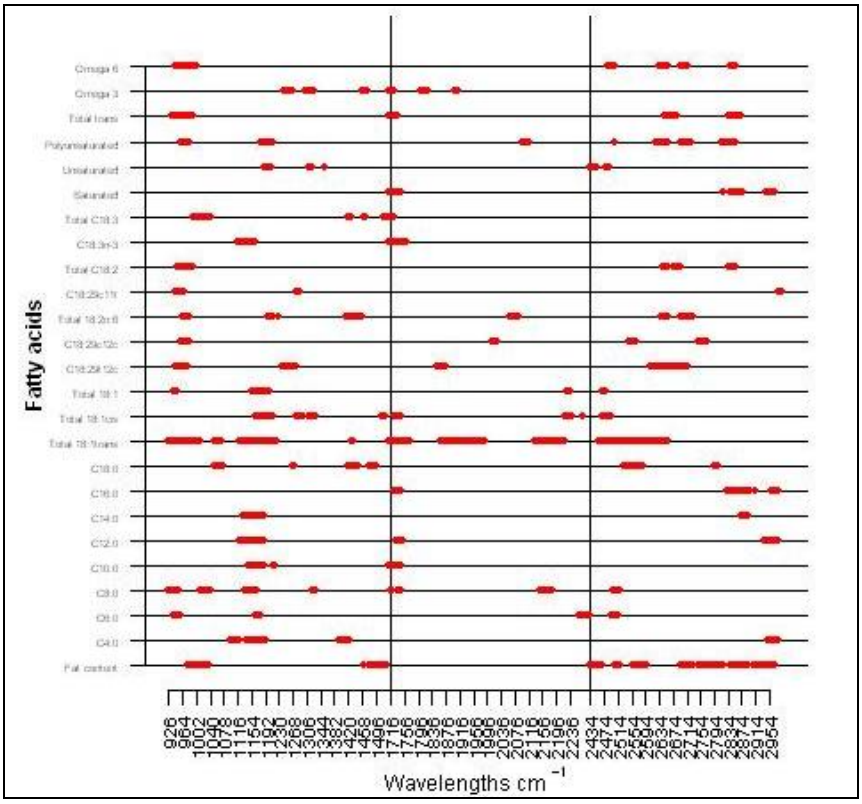

Figure 1. Selected wavelengths by genetic algorithm (second step).

Use PLS2 regressions provided good estimations for 9 fatty acids ($R2_{CV}$>80%) and correct estimations for 8 FA (70<$R2_{CV}$<80%). Using genetic algorithms before PLS regression, we got good estimations for 9 fatty acids ($R2_{CV}$) >80%) and correct estimations for 12 FA (70<$R2_{CV}$<80%). These results are not as good as those in cow and sheep milk, but the fatty acid composition of goat milk is different from that of cow milk. For our samples, the average fat content was 3,98 g/100 ml in cow milk whereas it was 3.31 g/100 ml with a lower variability for some fatty acids in goat milk. The smallest variability of fat content in goat milk is one explanation of the lower estimation quality (lower R^2_{CV}). Other parameters could explain this too. For instance, the proteolysis activity or the cell number could impact the spectrum and make their estimations more difficult.

Nevertheless the results show that the equation accuracy is improved by the genetic algorithm use. By comparing the SE_{CV} of the model using GA+PLS with the model using PLS only, the accuracy is increased on average by 9%. As a remark, the best method to apply after genetic algorithm for the majority of fatty acids is the multivariate PLS.

More specifically, the accuracy increased by 9% for linoleic acid (C18:2 9c12C), by 13
% for palmitic acid (C16:0) and by 14% for myristic acid (C14:0) (Table 1). These fatty acids are of a wide interest regarding nutrition so that such an accuracy gain is really significant regarding dairy industry. The accuracy improvement concerned all classes: fatty acids with a R^2_{CV} below or upper 70%. Improvement level was not linked with estimation quality or fatty acid families. However for some fatty acids there was not a real improvement. For omega 3 fatty acids the estimations remain more accurate with PLS regression on 446 wavelengths.

Comparing to PLS regression, GA+PLS regression have null coefficients for unselected wavelengths. Thus, we can expect that estimations will be less influenced by a change on the spectra resulting from variation of outside factors (temperature, chemical preservative (bronopol) in milk). However this hypothesis must be verified. For instance, it is important to be able to perform accurate estimation of alpha-linolenic acid C18:3 n-3 in the future since there are important challenges on this fatty acid in the forthcoming years. Up to now, the relative error of this fatty acid is rather high, about 22%. The use of genetic algorithms does not decrease the relative error. However, it leads to use only 27 out of 446 wavelengths and we can suppose that the estimations will be more stable over time if fewer wavelengths are used.

Table 1. Statistical parameters for each calibration equation in goat milk (PLS regression only or genetic algorithm (GA) + PLS regression).

			PLS2			GA + PLS					
Fatty acids	**Mean**	**Sd**	**SECV**	**R. Error**	**R^2_{CV}**	**Kind of PLS**	**number of var.**	**SECV**	**R. Error**	**R^2_{CV}**	**Improvement**
Fat content	3,31	0,666	0,024	1%	1,00	PLS1	138	0,016	0,5%	1,00	33%
C4:0	0,092	0,025	0,011	12%	0,79	PLS2	35	0,009	10%	0,88	22%
C6:0	0,078	0,02	0,007	9%	0,87	PLS2	25	0,007	10%	0,88	2%
C8:0	0,08	0,022	0,010	13%	0,78	PLS2	28	0,010	12%	0,79	1%
C10:0	0,264	0,071	0,037	14%	0,73	PLS1	21	0,034	13%	0,79	9%
C12:0	0,134	0,041	0,023	17%	0,69	PLS2	36	0,018	14%	0,82	21%
C14:0	0,307	0,077	0,034	11%	0,82	PLS2	22	0,029	9%	0,88	14%
C16:0	0,996	0,197	0,059	6%	0,92	PLS2	31	0,051	5%	0,94	13%
C18:0	0,282	0,099	0,052	18%	0,75	PLS2	21	0,050	18%	0,78	4%
Total 18:1 trans	0,074	0,026	0,018	24%	0,53	PLS2	15	0,016	22%	0,62	9%
Total 18:1cis	0,681	0,164	0,070	10%	0,83	PLS2	51	0,067	10%	0,84	4%
Total 18:1	0,756	0,176	0,072	9%	0,84	PLS2	26	0,065	9%	0,87	9%
C18:29t12c	0,006	0,002	0,001	22%	0,42	PLS2	56	0,001	19%	0,59	26%
C18:29c12c	0,086	0,02	0,012	14%	0,67	PLS2	26	0,011	13%	0,72	9%
Total 18:2n-6	0,092	0,021	0,013	14%	0,66	PLS2	31	0,011	13%	0,73	15%
C18:29c11t	0,017	0,005	0,004	22%	0,45	PLS2	9	0,003	20%	0,55	20%
Total C18:2	0,109	0,024	0,015	14%	0,62	PLS2	18	0,013	12%	0,72	15%
Total C18:3 n-3	0,013	0,004	0,003	24%	0,41	PLS2	27	0,003	22%	0,49	3%
Total C18:3	0,014	0,004	0,003	23%	0,45	PLS2	23	0,003	21%	0,48	5%
Saturated	2,351	0,485	0,087	4%	0,97	PLS2	31	0,082	3%	0,97	6%
Monounsaturated	0,798	0,184	0,074	9%	0,85	PLS2	20	0,071	9%	0,86	4%
Polyunsaturated	0,128	0,028	0,018	14%	0,63	PLS2	45	0,016	12%	0,70	9%
Trans	0,1	0,031	0,021	21%	0,53	PLS2	22	0,020	20%	0,60	5%
Omega 3	0,018	0,005	0,004	20%	0,46	PLS1	115	0,004	20%	0,45	-11%
Omega 6	0,109	0,027	0,016	14%	0,67	PLS2	25	0,014	13%	0,75	10%

The retained equations were validated on a new dataset. The results are presented in table 2. For the majority of fatty acids we confirmed the interest of genetic algorithms (increase of 12% for linoleic acid (C18:2 9c12C), of 15% for palmitic acid (C16:0) and of 22% for myristic acid (C14:0)). However for some fatty acids, in particular for the C18 family, the two methods give equivalent results.

Table 2. Statistical parameters for each equation in goat milk on an independent dataset (PLS regression only or genetic algorithm (GA) + PLS regression). Validation according to ISO norm 8196.

			PLS2			GA + PLS			
Fatty acids	**Mean**	**Sd**	**$S_{y,x}$**	**R. Error**	**R^2**	**$S_{y,x}$**	**R. Error**	**R^2**	**Improvement**
Fat content	3,040	0,474	0,028	1%	1,00	0,021	1%	1,00	23%
C4:0	0,084	0,012	0,008	9%	0,56	0,005	6%	0,80	34%
C6:0	0,078	0,013	0,005	7%	0,83	0,004	5%	0,91	26%
C8:0	0,083	0,016	0,006	8%	0,84	0,005	6%	0,90	22%
C10:0	0,298	0,058	0,026	9%	0,80	0,019	6%	0,90	29%
C12:0	0,127	0,027	0,016	13%	0,64	0,013	10%	0,79	24%
C14:0	0,341	0,046	0,026	8%	0,67	0,021	6%	0,80	22%
C16:0	0,883	0,146	0,066	7%	0,80	0,056	6%	0,86	15%
C18:0	0,258	0,052	0,047	18%	0,21	0,035	14%	0,56	25%
Total 18:1 trans	0,064	0,019	0,018	28%	0,09	0,019	29%	0,00	-5%
Total 18:1cis	0,491	0,099	0,057	12%	0,68	0,061	12%	0,64	-7%
Total 18:1	0,559	0,101	0,052	9%	0,74	0,049	9%	0,77	5%
C18:29t12c	0,008	0,002	0,002	24%	0,28	0,002	27%	0,07	-13%
C18:29c12c	0,061	0,009	0,007	12%	0,31	0,006	11%	0,46	12%
Total 18:2n-6	0,069	0,010	0,008	12%	0,34	0,007	10%	0,53	16%
C18:29c11t	0,015	0,006	0,006	42%	0,05	0,006	43%	0,00	-2%
Total C18:2	0,084	0,014	0,011	13%	0,40	0,011	13%	0,35	-4%
Total C18:3 n-3	0,025	0,006	0,006	23%	0,03	0,006	23%	0,06	1%
Total C18:3	0,026	0,006	0,006	23%	0,00	0,006	23%	0,02	1%
Saturated	2,279	0,357	0,075	3%	0,96	0,054	2%	0,98	29%
Monounsaturated	0,592	0,108	0,054	9%	0,76	0,056	9%	0,74	-4%
Polyunsaturated	0,118	0,019	0,015	13%	0,36	0,015	12%	0,42	5%
Trans	0,089	0,026	0,025	28%	0,07	0,026	29%	0,00	-4%
Omega 3	0,034	0,007	0,007	21%	0,12	0,007	20%	0,16	2%
Omega 6	0,087	0,013	0,010	11%	0,42	0,009	10%	0,51	8%

5. Conclusion

The mid-Infrared Spectrometry is of a strong interest for the estimation of fatty acid content in cow and small ruminant milk. It is possible to obtain estimation equations rather quickly (with traditional PLS method), although in goat milk estimations are not as accurate as in cow and sheep milk. Before using the mid-infrared spectrometry in routine to estimate the fatty acid profile, it is necessary to have established equations providing sufficient accurate estimations. As the authors cited in the introduction, the use of a genetic algorithm to select informative wavelengths allows us to improve the quality of estimations and to stabilize the equations over the time. Since null coefficients are applied on discarded wavelengths, we avoid spoiling predictions with extra background noise not related to milk composition. On a new dataset, the improvement is confirmed for the saturated fatty acids but no completely for the unsaturated fatty acids.

Future researches will focus on other pre-treatment procedures, while increasing simultaneously the initial sample size to get more accurate estimation equations of milk fatty acid profile. Another important point to study will be the standardization of the results between different spectrometers.

The advancements of the PhénoFinLait program are available on http://www.phenofinlait.fr/

Acknowledgements

The authors thank INRA experimental farms for the technical support and steering committee of Phénofinlait for the constructive discussions. This study received financial support from Apis-Gène, French Ministry of Agriculture and France Génétique Elevage.

References

[1]. Tomotake, H., Okuyama, R., Katagiri, M., Fuzita, M., Yamato M., Ota, F. (2006). Comparison between holstein cow's milk and Japanese-Saanen goat's milk in fatty acid composition, lipid digestibility and protein profile. *Bioscience, biotechnology, and biochemistry*, 70, 2771-2774.

[2]. Chilliard, Y., Ferlay, A., Rouel, J., Lamberet, G. (2003). A Review of Nutritional and Physiological Factors Affecting Goat Milk Lipid Synthesis and Lipolysis. *J. Dairy Sci.*, 86, 1751–1770.

[3]. Raynal-Ljutovac, K., Lagriffoul, G., Paccard, P., Guillet, I., Chilliard, Y. (2008). 'Composition of goat and sheep milk products: An update. *Small Ruminant Research*, 79, 57-72.

[4]. Bernard, L.,Bonnet, M., Leroux, C., Shingfield, K. J., Chilliard, Y. (2009). Effect of sunflower-seed oil and linseed oil on tissue lipid metabolism, gene expression, and milk fatty acid secretion in alpine goats fed maize silage–based diets. *J. Dairy Sci.*, 92, 6083–6094.

[5]. Mahé, M.F., Manfredi, E., Ricordeau, G., Piacère, A., Grosclaude, F. (1993). Effets du polymorphisme de la caséine αs1 caprine sur les performances laitières: analyse intradescendance de boucs de race Alpin. *Genet. Sel. Evol.*, 26, 151-157.

[6]. Soyeurt, H., Dardenne, P., Dehareng, F., Lognay, G., Veselko, G., Marlier, M., Bertozzi, C., Mayeres, P., Gengler, N. (2006). Estimating Fatty Acid Content in Cow Milk Using Mid-Infrared Spectrometry. *J. Dairy Sci.*, 89, 3690-3695.

[7]. Soyeurt, H., Colinet, F.G., Arnould, V. M.-R., Dardenne, P., Bertozzi, C., Renaville, R., Portetelle, D., Gengler, N. (2007). Genetic Variability of Lactoferrin Content Estimated by Mid-Infrared Spectrometry in Bovine Milk. *J. Dairy Sci*., 90, 4443-4450.

[8]. Leardi, R., Lupiañez, G. (1998). Genetic algorithms applied to feature selection in PLS regression: how and when to use them. *Chemometr. Intell. Lab. Syst*., 41, 195-208.

[9]. Bangalore, A. S., Shaffer, R. E., Small, G.W. (1996). Genetic Algorithm-Based Method for Selecting Wavelengths and Model Size for Use with Partial Least-Squares Regression: Application to Near-Infrared Spectroscopy. *Anal. Chem*., 68, 4200-4212.

[10]. Roger, J. M., Bellon-Maurel, V. (2000). Using Genetic Algorithms to Select Wavelengths in Near-Infrared Spectra: Application to Sugar Content Prediction in Cherries. *Appl. Spectrosc.*, 54, 1313-1320.

[11]. Spiegelman, C.H., McShane, M.J., Goetz, M.J., Motamedi, M., Yue, Q.L., Coté, G.L. (1998). Theoretical Justification of Wavelength Selection in PLS Calibration: Development of a New Algorithm. *Anal. Chem.*, 70, 35–44.

[12]. Holland, J. (1992). Les algorithmes génétiques. *Pour la Science*, 179, 44-51.

[13]. Hoskuldsson, A. (2001). Variable and subset selection in PLS regression. *Chemometr. Intell. Lab. Syst.*, 55, 23-38.
[14]. Haupt, R.L., Haupt, S.E. (2004). *Practical Genetic Algorithms*, New Jersey: Wiley.
[15]. Foss (1998), *Reference Manual of Milkoscan FT120 (Type 71200)*, Denmark: Foss.
[16]. Kramer, J.K.G., Fellner, V., Dugan, M.E.R., Sauer F.D., Mossoba1, M.M., Yurawecz, M.P. (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.
[17]. Tenenhaus, M. (2002). *La regression PLS*, Lassay-les-Chateaux: Technip.
[18]. Bertrand, D., Dufour, E. (2006). *La spectrométrie infrarouge et ses applications analytiques*, Paris: Tec&Doc Lavoisier.
[19]. *The R Foudation for statistical computing*. R version 2.8.1 (2008-12-22).
[20]. Mevik, B.H., Wehrens, R. (2007). The PLS Package: Principal Component and Partial Least Squares Regression in R. *Journal of Statistical Software*, 18, 1-24.
[21]. Ferrand, M., Huquet, B., Barbey, S., Barillet, F., Brochard, M., Faucon, F., Larroque, H., Leray, O. (2010). Determination of fatty acid profile in cow milk using Mid-Infrared spectrometry: interest of applying a variable selection by genetic algorithms before a PLS regression. *Chemometr. Intell. Lab. Syst.*, 106, 183-189.
[22]. *The MathWorks*, Inc. Matlab version7.8.0.347 (R2009a).