Standardization of milk infrared spectra for the retroactive application of calibration models

V. Bonfatti,*[1] A. Fleming,† A. Koeck,† and F. Miglior†‡
*Department of Comparative Biomedicine and Food Science, University of Padova, 35020, Legnaro, Italy
†Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, ON, N1G 2W1, Canada
‡Canadian Dairy Network, Guelph, ON, N1K 1E5, Canada

ABSTRACT

The objective of this study was to standardize the infrared spectra obtained over time and across 2 milk laboratories of Canada to create a uniform historical database and allow (1) the retroactive application of calibration models for prediction of fine milk composition; and (2) the direct use of spectral information for the development of indicators of animal health and efficiency. Spectral variation across laboratories and over time was inspected by principal components analysis (PCA). Shifts in the PCA scores were detected over time, leading to the definition of different subsets of spectra having homogeneous infrared signal. To evaluate the possibility of using common equations on spectra collected by the 2 instruments and over time, we developed a standardization (STD) method. For each subset of data having homogeneous infrared signal, a total of 99 spectra corresponding to the percentiles of the distribution of the absorbance at each wavenumber were created and used to build the STD matrices. Equations predicting contents of saturated fatty acids, short-chain fatty acids, and C18:0 were created and applied on different subsets of spectra, before and after STD. After STD, bias and root mean squared error of prediction decreased by 66% and 32%, respectively. When calibration equations were applied to the historical nonstandardized database of spectra, shifts in the predictions could be observed over time for all investigated traits. Shifts in the distribution of the predictions over time corresponded to the shifts identified by the inspection of the PCA scores. After STD, shifts in the predicted fatty acid contents were greatly reduced. Standardization reduced spectral variability between instruments and over time, allowing the merging of milk spectra data from different instruments into a common database, the retroactive use of calibrations equations, or the direct use of the spectral data without restrictions.

Key words: infrared spectra, standardization, retroactive calibration models

INTRODUCTION

In the last decade, Fourier transform mid-infrared (**IR**) spectroscopy calibration equations have been applied for the determination of several novel milk components and functional traits (Egger-Danner et al., 2015). Given the considerable effort in building calibration models, it would be beneficial if these models could be applied over different instruments and over time, or to apply them retroactively to historical spectral databases (e.g., for benchmark trends and breeding purposes). However, the transferability of a calibration equation over instruments and over time is not always guaranteed (Wang et al., 1991). Because of variations in the instrumental response among different spectrometers, the unrestricted use of the spectra obtained with an instrument different from the one used for calibration might lead to biased and inaccurate predictions (Grelet et al., 2015).

Likewise, when an instrument response is measured over time, changes in the signal intensity due to electronic drift or detector instability can occur. If this happens after (or before) the calibration of an instrument, the subsequent (or retroactive) use of the calibration model might lead to incorrect predictions, depending upon the magnitude of the drift.

To overcome the problem of spectral variation among instruments and over time, adjustments of the predicted data to reference standards (i.e., slope and bias corrections) can be made. However, these corrections cannot be applied over traits for which there are no reference samples; for example, body energy status or methane emission of cows, for which calibration models have also been developed (McParland et al., 2011; Dehareng et al., 2012). In addition, the number of traits to be predicted is rapidly increasing, as is the interest in

using the spectral information directly as an indicator of animal health and efficiency (Dehareng et al., 2012), or in combining different spectral databases to build more robust calibration models (Grelet et al., 2016). Thus, the adjustment of spectra instead of predictions is the most appropriate approach (Grelet et al., 2015).

Analysis of identical samples is commonly used to standardize each machine to overcome instrument-to-instrument variations. The standardization (**STD**) process consists of applying a transformation function to the new spectra, converting the signal to that which would be obtained on the original instrument. Recently, this approach has been successfully used to create a network of standardized IR spectrometers across different European countries (Grelet et al., 2015). However, when a calibration needs to be applied retroactively on nonstandardized spectra, this approach does not guarantee the transferability of the equation to the spectra collected previously because spectra might have shifted over time. For the same reason, the direct use of historical spectra collected by nonstandardized instruments is not feasible.

The objective of this study was to develop a STD method to harmonize the IR predictions obtained over time and across 2 milk laboratories of Canada, to create a uniform historical database of spectra allowing (1) the retroactive application of calibration models, and (2) the direct use of the full spectrum without restriction for the development of indicators of animal health and efficiency.

MATERIALS AND METHODS

Fourier-Transform Mid-Infrared Spectra

A proportion of milk spectra collected from IR spectrometers at the Canadian DHI organizations CanWest DHI (Guelph, Ontario) and Valacta (Sainte-Anne-de-Bellevue, Québec) was transferred to the Canadian Dairy Network (Guelph, Ontario) and has been stored in a database since January 2013. Roughly 60,000 milk spectra are added every month. The database used for this study included 2,213,318 spectra records, of which 1,246,308 were from Valacta and 967,010 from CanWest DHI. Spectra acquisition was performed in each of the 2 laboratories using a MilkoScan FT6000 (Foss Electric A/S, Hillerød, Denmark).

Exploratory Principal Component Analysis

To monitor spectra variation over time, a principal component analysis (**PCA**) was performed on the spectra collected by each of the laboratories. The scores of the first 10 principal components (**PC**) were plotted over time and inspected. Due to the interference of water absorption, the O–H bending and stretching regions of the spectra (between 1,628 and 1,658 cm^{-1} and between 3,105 cm^{-1} and 3,444 cm^{-1}, respectively) do not contain useful chemical information and were removed before the PCA. As a consequence, 862 over a total of 1,060 spectral variables available were used.

Figure 1 shows the scores of the main PC obtained for the spectra collected at Valacta and CanWest DHI labs from January 2013 to July 2015 and from March 2013 to July 2015, respectively. For the spectra collected at the Valacta laboratory, 2 and 3 major shifts were detected in PC2 and PC10, respectively, leading to the separation of the spectra population in a total of 6 subsets having homogeneous signal. For CanWest DHI spectra, one small shift affecting both PC 2 and 4 occurred, leading to the separation of the population into 2 subsets. Within subset 2, other minor shifts were visible in the scores of the second PC, but they were not clearly identifiable and they were not used to create additional subsets.

Consequently, a total of 8 different subsets of spectra were defined, with sample size ranging from 77,753 to 668,827 records. The details about time intervals and number of spectra records for each subset are reported in Table 1.

Creation of Spectral Data for Standardization

Commonly, STD of 2 machines is performed through the acquisition of spectra on common samples to link the response signal of the instrument used to create the calibration equation (master) to the signal produced by the other (slave). Inter-laboratory tests are organized and several sets of identical samples are distributed to the different laboratories to build the STD model. Usually, samples for STD consist of aliquots of raw milk with large variations in fat and protein. These STD samples are measured on the master instrument and on the slave instrument, leading to response matrices that are then used to develop the STD model. The model works by "converting" the spectra from a new slave spectrometer to match the spectra obtained by the master instrument. The STD coefficients need to be updated periodically due to physical wear and perturbation on apparatus, which can result in spectral deviations in time. Results of Grelet et al. (2015) showed that STD coefficients become less effective in reducing differences between master and slave spectra starting 1 mo after their creation. Thus, this approach might not be effective to standardize spectra collected before the development of the STD procedure. If calibration equations are
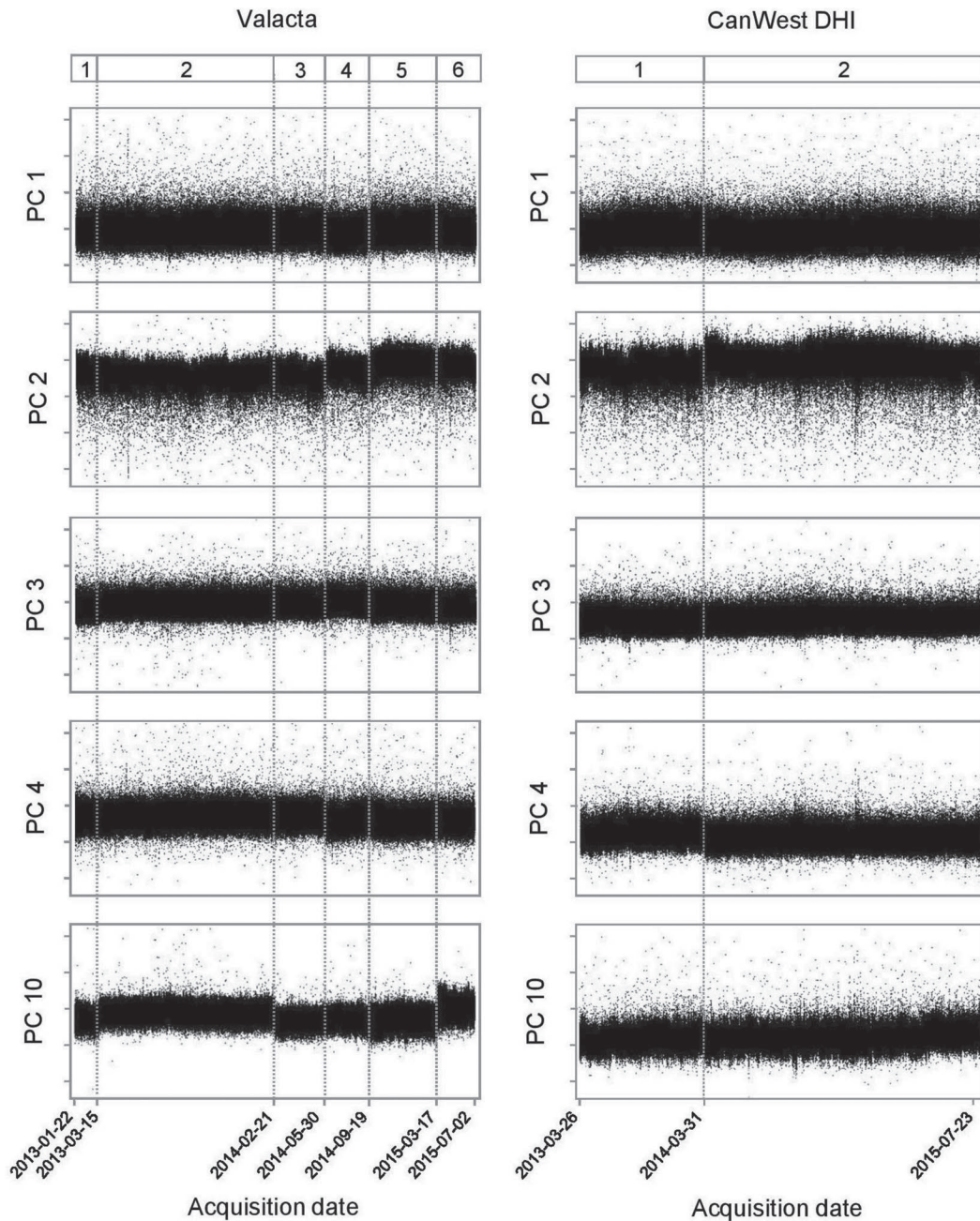
Figure 1. Distribution over time of the scores of principal components (PC) for spectra collected at Valacta (Sainte-Anne-de-Bellevue, QC, Canada) and CanWest DHI (Guelph, ON, Canada) milk laboratories. The numbers identify the subsets of spectra having homogeneous signal. Date is year–month–day.

applied retroactively on spectra that had been collected on nonstandardized instruments, spectra acquired on common samples are not available. However, it can be assumed that, for an infinite number of samples, in the absence of shifting and of instrument-to-instrument variation, the distribution of the absorbance at each wavenumber (as well as the distribution of the PCA scores) in a certain time interval or laboratory would be the same as that of spectra collected in another time interval or by a different laboratory. Hence, it is possible to create a set of "artificial" STD spectra using the percentiles of the distribution of the absorbance at each wavenumber obtained from the master and slave instruments. Following this approach, in every subset

Table 1. Descriptive statistics of subsets of data having homogeneous spectral signal

Laboratory[1]	Subset	No. of records	Date[2]
Valacta	1	77,753	2013–01–22
Valacta	2	548,325	2013–03–15
Valacta	3	153,203	2014–02–21
Valacta	4	138,593	2014–05–30
Valacta	5	211,028	2014–09–19
Valacta	6	117,406	2015–03–17
CanWest DHI	1	298,183	2013–03–26
CanWest DHI	2	668,827	2014–03–31

[1]Valacta (Sainte-Anne-de-Bellevue, QC, Canada) and CanWest DHI (Guelph, ON, Canada).

[2]First date of spectra acquisition (year–month–day).

of homogeneous spectra identified by the exploratory PCA, the Mahalanobis distance between each record and the average spectra was computed iteratively. At every iteration, spectra with a global distance value >3 were considered outliers (Shenk and Westerhaus, 1995) and removed, to exclude samples that could affect the distributions of the absorbance values. After outlier elimination, the percentiles of the distribution of the absorbance at each wavenumber were retained. For each percentile, the values obtained for the 862 wavenumbers were used to originate a new spectrum. Thus, we created a total of 99 spectra records (standardization set; **STDset**) for each subset of homogeneous spectra identified by the exploratory PCA. These artificial spectra were used to build the response matrices used to convert the signal produced from an instrument in a certain period to that produced by another instrument or by the same instrument in another period.

Master and slave spectra were defined upon the spectra populations having homogeneous IR signal identified by the PCA. The spectra produced by the instrument of Valacta laboratory between September 2014 and March 2015 (subset 5, Figure 1) were defined as the master spectra population. The spectra collected by Valacta in other time intervals, as well as the spectra collected by CanWest DHI, were considered slave spectra.

Because a STD model needs to be designed for every master–slave combination, the existence of 7 slave spectra populations led to the development of 7 different STD models.

Development of Standardization Models

For a given master–slave combination, the STDset from the master and from the slave lead to response matrices **M** and **S**. The response signal $\mathbf{m}_j$ obtained on the master at wavenumber j was assumed to be linearly related to the signal (**s**) of wavenumber j of the slave $\mathbf{s}_j$.

The signal at the jth wavenumber of the master spectra was regressed on the signal at the corresponding wavenumber $\mathbf{s}_j$ of the slave, so that

$$\mathbf{m}_j = \mathbf{s}_j b_j + a_j,$$

where b_j is the regression coefficient and a_j is the offset term for the jth wavenumber. Each time a new sample is measured on the slave instrument, the new spectrum **W** can be standardized into $\mathbf{W_{std}}$ as follows:

$$\mathbf{W_{std}} = \mathbf{W} \times \mathbf{B} + \mathbf{a},$$

where **B** is a diagonal matrix containing the regression coefficients for each wavenumber, and **a** is the vector of offset terms.

Calibration Equations

A total of 912 spectra collected between February 2014 and April 2015 also contained information on fatty acid composition. Of these spectra, 370 were collected by Valacta between September 2014 and March 2015 and belonged to the master spectra population (subset 5, Figure 1); 191 and 54 belonged, respectively, to subsets 3 and 4 of the spectra collected at the same laboratory, whereas 297 belonged to subset 2 of CanWest DHI. The samples for which fatty acid composition was available belonged to 290 animals in 40 herds.

Calibration models were developed for the contents of SFA, short-chain fatty acids (**SCFA**), and C18:0, using the 370 records belonging to the master spectra population. All calibration models were developed by partial least squares (**PLS**) regression using the "plsr" function implemented in the PLS package (Mevik and Wehrens, 2007) of the R software (R Development Core Team, 2013) and a 10-fold cross-validation procedure. The number of PLS terms was chosen to minimize the root mean squared error of prediction (**RMSEP**).

A gap-segment derivative of first order with a filter length (i.e., the spacing between points over which the derivative is computed) of 3 and a segment size (i.e., the range over which the points are averaged) of 2 were applied to the spectra before regression models. The gap-segment derivative mathematical treatment was performed using the PROSPECTR package (Stevens and Ramirez-Lopez, 2013) of the R software. As for PCA analysis, spectral regions corresponding to water absorption were discarded and regression models were performed on a total of 862 spectra variables. The RMSEP and R^2 in cross-validation were calculated for each calibration model. Calibration equations were applied to the historical spectral database to obtain predictions

of SFA, SCFA, and C18:0 from all the available spectra, before and after STD.

Validation of the Standardization Method

Validation of the STD procedure was carried out in 2 steps: (1) comparing predictions of fatty acid contents obtained, before and after STD, by the master spectra and the slave spectra of the STDset; and (2) comparing the prediction accuracy of calibration models for fatty acid composition in external validation, before and after STD.

For the first validation experiment, fatty acid predictions were obtained, before and after STD, from all combinations of master and slave spectra belonging to the STDset. Then, for each master–slave combination, regressions between predictions were performed. The RMSEP, R^2, bias, and slope (deviation from 1) of regressions were calculated and compared, before and after STD, by ANOVA (SAS 9.3; SAS Institute Inc., Cary, NC).

To estimate the effect of spectra STD on prediction accuracy of calibration models, calibration equations developed for SFA, SCFA, and C18:0 were tested in external validation on the 3 subsets of slave spectra (subsets 3 and 4 of Valacta and subset 2 of CanWest DHI) for which reference data were also available. Validation subsets 3 and 4 of Valacta included, respectively, 40 and 156 samples deriving from animals that also provided samples belonging to the master spectra population. Conversely, there was no relationship between the samples of the master spectra population and the validation set of CanWest DHI.

The relationship between the measured values and the predicted values obtained before and after spectra STD was evaluated. The effect of spectra STD on the performance parameters in validation (RMSEP, R^2, bias, and slope) was assessed by ANOVA (SAS 9.2; SAS Institute Inc.).

RESULTS AND DISCUSSION

Effect of Standardization on Spectra

Examples of the effect of STD model on the IR signal of the slave spectra belonging to the STDset are shown in Figure 2. As the instruments used by the 2 laboratories were of the same brand and model, variations in the IR signal obtained across labs were limited. This was also observed by Grelet at al. (2015). However, differences in the shape of the spectra were visible. In particular, in the spectral region between 1,680 and 2,080 cm^{-1}, shifts in absorbance peaks were observed

in spectra collected by the same instrument in different time intervals. After STD, differences between the signal of the slave spectra and that of the master spectra were greatly reduced.

Even though we used instruments of the same brand, the approach presented in this study could be applied to instruments of different brands. As instruments of different brands have different spectral ranges, the number of wavenumbers has to be harmonized (e.g., by interpolation) before STD.

Validation of the Standardization Procedure

Values of RMSEP and R^2 in cross-validation calculated for SFA, SCFA, and C18:0 calibration models are reported in Table 2, and are in agreement with those obtained in previous studies (Soyeurt et al., 2006; Bonfatti et al., 2016).

The relationship between fatty acid predictions obtained before STD from each combination of master and slave spectra belonging to the STDset is presented in Figure 3. Bias and slope deviations between predictions of fatty acids obtained from master spectra and from slave spectra were observed in most of the subsets. The largest biases in SFA predictions were obtained for CanWest DHI spectra. This was expected because those spectra were collected by an instrument different from that used to build the calibration models. However, for SCFA and C18:0, CanWest DHI spectra led to prediction accuracies similar to those obtained for spectra collected by Valacta in different periods. For SCFA, biases were observed in 4 out of 5 Valacta subsets, and predictions were more affected by variations of spectra signal within the same instrument over time than by variations of signal across instruments. For C18:0, biased predictions were obtained in particular for subset 1 and 2 of Valacta and for subset 1 of CanWest DHI spectra.

Slope deviations from 1 were also observed. In particular, a deviation greater than 0.1 affecting all the investigated traits was detected in subset 1 of Valacta spectra. Slope deviations greater than 0.1 were also observed in subset 3 of Valacta spectra, for SFA, and in most of the subsets for SCFA.

Results indicate that within each subset of data having homogeneous signal, the severity of bias and slope deviation was not consistent across traits. The magnitude of the effect exerted by spectral shifting is expected to be variable across traits, according to the spectral regions affected by the shifts. Variations occurring in specific spectral regions might not affect a particular trait but have a considerable effect on another one. As a consequence, spectra providing accurate predictions
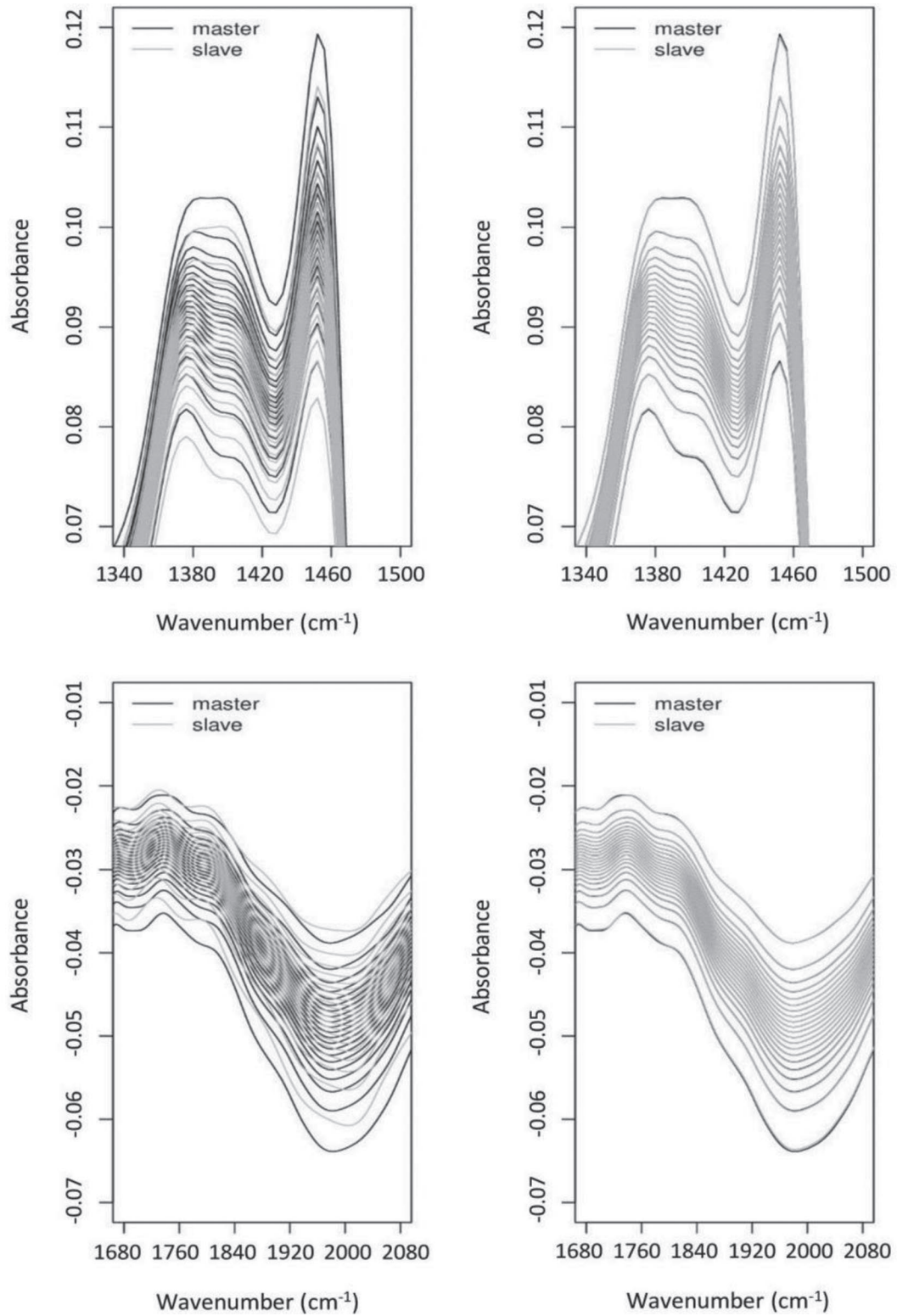
Figure 2. Absorbance in 2 spectral regions obtained from the same instrument in 2 time intervals before (left) and after (right) standardization.

Table 2. Performance parameters of the calibration equations developed by partial least square regression (PLS) for the content of SFA, short-chain fatty acids (SCFA), and C18:0[1]

Trait	Mean	SD	PLS terms	RMSEP	R^2 cross-validation
SFA, g/dL	2.985	0.601	11	0.125	0.950
SCFA, g/dL	0.479	0.117	16	0.046	0.791
C18:0, g/dL	0.148	0.042	16	0.017	0.764

[1]Calibration models were developed on the spectra collected by Valacta in the period between September 2014 and March 2015; RMSEP = root mean squared error of prediction.

for a trait might produce inaccurate predictions for another trait. This underlines the need to standardize spectra independently on the results obtained for a single or for a few traits. This is particularly clear when comparing the results obtained for the 3 traits in subset 4 of Valacta spectra (Figure 3). Before STD, the bias produced by spectra belonging to this subset was very close to zero for C18:0 but substantial for SFA and SCFA. In addition, a slope deviation greater than 0.1 was observed for SCFA predictions.

Calibration equations developed for fatty acid composition using spectra collected by Valacta in a specific period were not always transferable to the spectra collected in the same laboratory in other periods. Biases in SFA and C18:0 predictions were found for Valacta spectra, corresponding to the shift in the scores of PC10, which was used to define subset 6. This confirms that calibration models can take advantage of relatively small changes in spectra and that instrument-to-instrument differences, as well as inherent instrument variability, can lead to inaccurate predictions.

Regressions between predictions of fatty acid contents produced by the master and slave spectra belonging to the STDset were compared before and after STD. The effect of the STD procedure on RMSEP, bias, R^2, and slope (deviation from 1) is reported in Table 3. After STD, for all the investigated traits, values of RMSEP and bias decreased toward zero, as did the deviation of the slope from 1. Conversely, R^2 was not affected by the application of STD models and was always greater than 0.99.

Effect of Spectra Standardization on Prediction Accuracy

To test the effect of STD on prediction accuracy of calibration models, calibration equations for SFA, SCFA, and C18:0 were applied on 3 validation subsets before and after STD. The least squares means of the effect of STD on performance parameters of calibration models in validation are reported in Table 4. Overall, reductions of 66% ($P < 0.05$) and 32% ($P < 0.05$) in bias and RMSEP, respectively, were observed after spectra STD, whereas R^2 and slope remained unchanged. This confirmed that STD is beneficial for accuracy of predictions, even when they are obtained from instruments of the same brand and model.

Variation over Time in Predicted Fatty Acids

When calibration equations were applied to the historical nonstandardized database of spectra, shifts in the predictions could be observed over time for all the investigated traits. As an example, the distribution of the predicted C18:0 content obtained over time from spectra collected by the 2 laboratories, before and after STD, is reported in Figure 4. For spectra collected by Valacta before STD, shifts in the distribution of the predictions over time corresponded to the shifts identified by the inspection of the PCA scores. After STD, shifts in the predicted C18:0 content were greatly reduced. The bias between values of C18:0 predicted at CanWest DHI and Valacta decreased, but shifts over time in the

Table 3. Least squares means of parameters for the regression between fatty acid predictions obtained from the master and slave spectra belonging to the standardization set, before and after spectral standardization (STD)[1]

Parameter	SFA			Short-chain fatty acids			C18:0		
	Before STD	After STD	*P*-value	Before STD	After STD	*P*-value	Before STD	After STD	*P*-value
R^2	0.997	0.997	0.937	0.999	0.999	0.629	0.999	0.999	0.292
RMSEP	0.153	0.000	<0.01	0.005	0.000	<0.05	0.001	0.000	<0.05
Bias	0.373	0.000	<0.001	0.060	0.000	<0.01	0.032	0.000	<0.01
Slope	0.082	0.007	<0.05	0.109	0.003	<0.01	0.060	0.000	<0.05

[1]RMSEP = root mean squared error of prediction; bias was expressed as deviation from zero; slope was expressed as deviation from one.
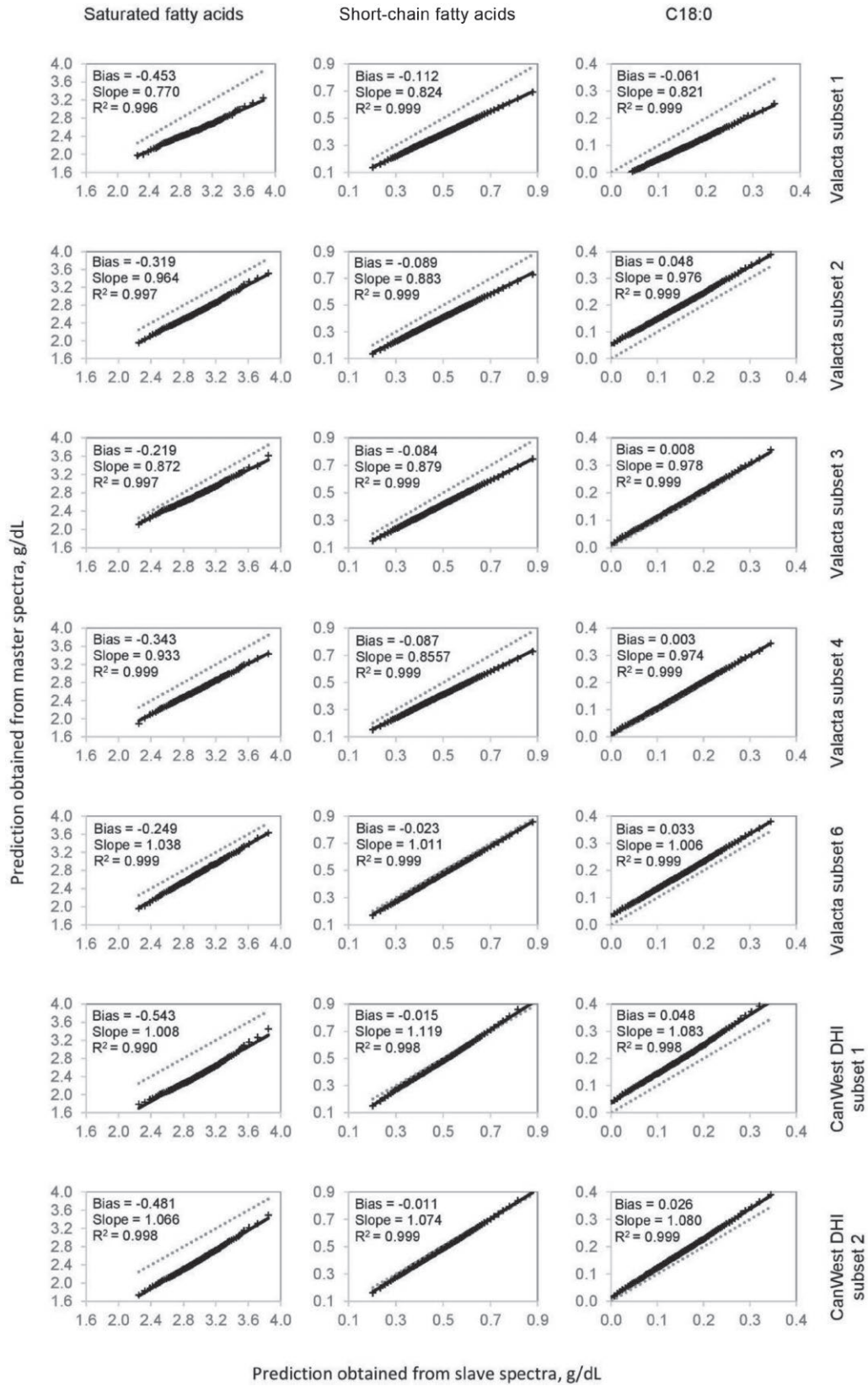
Figure 3. Comparison between prediction of fatty acid contents obtained, before standardization, from the master and slave spectra used to build standardization models.
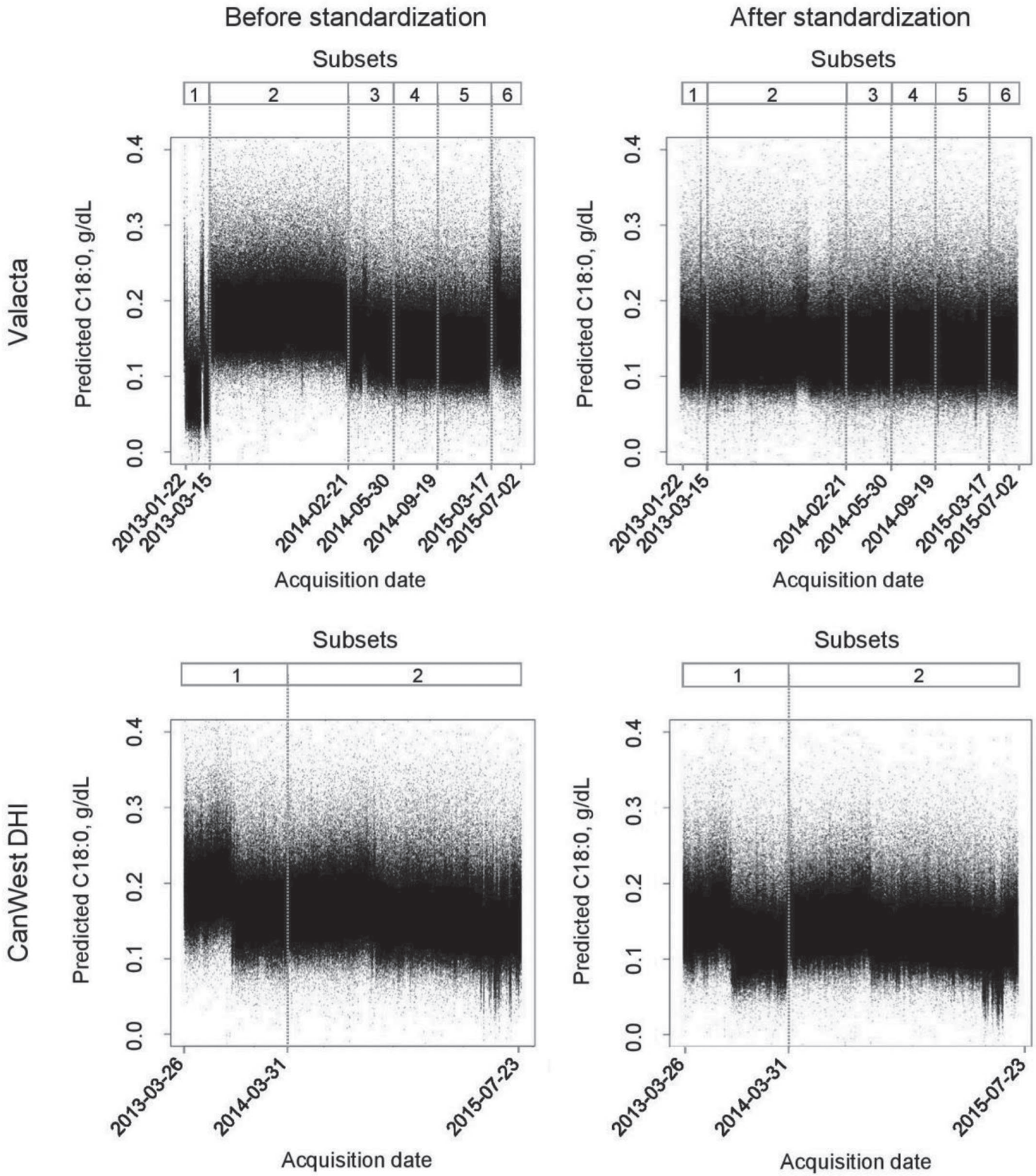
Figure 4. Predictions of C18:0 obtained over time from the spectra collected at Valacta (Sainte-Anne-de-Bellevue, QC, Canada) and CanWest DHI (Guelph, ON, Canada) milk laboratories before and after standardization. Date = year–month–day.

Table 4. Least squares means of the effect of spectra standardization (STD) on the validation performance of calibration models for fatty acid composition[1]

Parameter	Before STD	After STD	P-value
R^2	0.629	0.611	0.565
RMSEP	0.194	0.133	<0.05
Bias	0.151	0.052	<0.05
Slope	0.242	0.285	0.376

[1]RMSEP = root mean squared error of prediction; bias was expressed as deviation from zero; slope was expressed as deviation from one.

predictions of C18:0 from CanWest DHI spectra remained. Those shifts corresponded to minor variations occurring in the scores of the second PC (Figure 1) that were not taken into account when building the STD models. This confirms that very small variations in the IR signal might lead to inaccurate predictions and that minor spectral shifts should be taken into account when building STD models.

CONCLUSIONS

This study presented a standardization procedure for the harmonization of historical databases of spectra collected by nonstandardized instruments to allow the retroactive use of calibration models or the direct use of the spectral variables. Spectra variability across laboratories and over time was greatly reduced by standardization, allowing the merging of milk spectral data from different instruments into a common database and the retroactive use of calibration equations or the direct use of the spectral information without restrictions. The standardization procedure increased prediction accuracy of calibration models. This procedure was developed specifically for the standardization of spectra belonging to historical databases created when instruments were not yet standardized and it has to be considered complementary to the classical procedures for standardization, which make use of spectra acquired on common reference samples. To guarantee the correct application of the calibration models on a running spectrometer, the instrument should be standardized using the traditional standardization methods, the robustness of the calibration models should be increased by periodic updates, and the variability of the spectral signal over time should be monitored carefully. The traditional standardization method might also be used in the future to further verify the proposed procedure.

REFERENCES

Bonfatti, V., L. Degano, A. Menegoz, and P. Carnier. 2016. Short communication: Mid-infrared spectroscopy prediction of fine milk composition and technological properties in Italian Simmental. J. Dairy Sci. 99:8216–8221.

Dehareng, F., C. Delfosse, E. Froidmont, H. Soyeurt, C. Martin, N. Gengler, A. Vanlierde, and P. Dardenne. 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows. Animal 6:1694–1701.

Egger-Danner, C., J. B. Cole, J. E. Pryce, N. Gengler, B. Heringstad, A. Bradley, and K. F. Stock. 2015. Invited review: Overview of new traits and phenotyping strategies in dairy cattle with a focus on functional traits. Animal 9:191–207.

Grelet, C., C. Bastin, M. Gelé, J.-B. Davière, M. Johan, A. Werner, R. Reding, J. A. Fernandez Pierna, F. G. Colinet, P. Dardenne, N. Gengler, H. Soyeurt, and F. Dehareng. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. J. Dairy Sci. 99:4816–4825.

Grelet, C., J. A. Fernández Pierna, P. Dardenne, V. Baeten, and F. Dehareng. 2015. Standardization of milk mid-infrared spectra from a European dairy network. J. Dairy Sci. 98:2150–2160.

McParland, S., G. Banos, E. Wall, M. P. Coffey, H. Soyeurt, R. F. Veerkamp, and D. P. Berry. 2011. The use of mid-infrared spectrometry to predict body energy status of Holstein cows. J. Dairy Sci. 94:3651–3661.

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.

R Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Shenk, J. S., and M. O. Westerhaus. 1995. Forage analysis by near infrared spectroscopy. Pages 111–120 in Forages. Vol. II. The Science of Grassland Agriculture. 5th ed. R. F. Barnes, D. A. Miller, and C. J. Nelson, ed. Iowa State University Press, Ames.

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

Stevens, A., and L. Ramirez-Lopez. 2013. An introduction to the prospectr package. R package Vignette R package version 0.1.3. Accessed Dec. 10, 2016. https://cran.r-project.org/web/packages/prospectr/prospectr.pdf.

Wang, Y., D. J. Veltkamp, and B. R. Kowalsky. 1991. Multivariate instrument standardisation. Anal. Chem. 63:2750–2756.