Standardization of milk mid-infrared spectra from a European dairy network

C. Grelet,1 J. A. Fernández Pierna,1 P. Dardenne, V. Baeten, and F. Dehareng2
Walloon Agricultural Research Center, Valorisation of Agricultural Products Department, 24 Chaussée de Namur, 5030 Gembloux, Belgium

ABSTRACT

The goal of this study was to find a procedure to standardize dairy milk mid-infrared spectra from different Fourier transform mid-infrared spectrophotometers (different brands or models) inside a European dairy network to create new farm-management indicators (e.g., fertility, health, feed, environmental impact) based on milk infrared spectra. This step is necessary to create common spectral databases, allowing the building of statistical tools, to be used by all instruments of the network. The method used was piecewise direct standardization (PDS), which matches slave-instrument spectra on master-instrument spectra. To evaluate the possibility of using common equations on different instruments, the PDS method was tested on a set of milk samples measured on each machine, and an equation predicting fat content of milk is applied on all. Regressions were performed between master and slaves fat predictions, before and after PDS. Bias and root mean square error between predictions were decreased after PDS, respectively, from 0.3781 to 0.0000 and from 0.4609 to 0.0156 (g of fat/100 mL of milk). The stability over time of these results was confirmed by an application of the coefficients created by PDS 1 mo later on the slave spectra. These preliminary results showed that the PDS method permits a reduction of the inherent spectral variability between instruments, allowing the merging of Fourier transform mid-infrared milk spectra from different instruments into a common database, the creation of new types of dairy farm management indicators, and the use of these common calibrations for all Fourier transform mid-infrared instruments of the European dairy network.

Key words: Fourier transform mid-infrared spectrometry, standardization, dairy milk, piecewise direct standardization

INTRODUCTION

This work is the first step of a project aiming to develop innovative farm-management web applications based on the use of Fourier transform mid-infrared (FT-MIR) spectrometry analysis of milk to enable a sustainable and profitable management of the milk production. Fourier transform mid-infrared spectrometry is the worldwide method of choice for composition and quality controls during routine liquid milk testing. It allows a fast, nondestructive quantification of milk chemical properties to avoid reference methods, which are usually tedious, expensive, and time consuming. In 1961, a patent application for a FT-MIR method determining fat, protein, and lactose in milk was introduced (Goulden, 1964). The first apparatus, an IRMA (Infrared Milk Analyzer, Grubb Parsons, Newcastle upon Tyne, UK) using a monochromator, was based on the principle of measuring direct absorption of the infrared energy at specific frequencies by carbonyl groups in the ester linkages of the fat molecules, by peptide linkages between amino acids of protein molecules, and by the O-H groups in lactose molecules. A second generation of infrared instrumentation has adopted the change from wavenumber selection by diffraction grating to optical filters (Grappin and Jeunet, 1976) and was largely used by Central milk laboratory testing, where both tank milk and individual-cow samples were tested. Fourier transform mid-infrared supplies complementary chemical information and allows a high throughput with high sensitivity in a short response time from a very small quantity of sample (Ghosh and Jayas, 2009). In 1993, the first purpose-built FT-MIR instrument based on the Fourier transform infrared (FT-MIR) technology was marketed (Anadis MI-200; Asselain et al., 1996). With the introduction of the FT-MIR, new applications have been developed because of the use of the full spectrum of the sample. In this way, FT-MIR has been applied for the determination of more and more milk components such as proteins composition (Bonfatti et al., 2011), minerals (Soyeurt et al., 2009), ketone bodies (van Knegsel et al., 2010), lactoferrin (Soyeurt et al., 2007), and fatty acid profile (Rutten et al., 2009; Soyeurt et al., 2011). Then recent studies were performed using these milk components predicted by FT-MIR to
predict physiological indicators of the animal (Friggens et al., 2007; Mohammed et al., 2011).

In the context of this research project the FT-MIR spectrum is directly considered as a reflection of the state of the cows, avoiding the step of milk composition, to obtain indicators concerning fertility, health, environment, and feeding among others. Until now, only a few studies have been performed to show the potential of the entire FT-MIR spectra as an indicator of those parameters. Only recent studies have shown that predictions based on direct spectra are much more global, sensitive, and accurate than those based on milk components when they are predicted from FT-MIR. Dehareng et al. (2012) have shown that enteric methane was better predicted when directly working with FT-MIR spectra than the results based on fatty acid predictions. Also recently, the FT-MIR spectrum of milk was shown to be a good indicator of body energy status (McParland et al., 2011), energy intake and efficiency (McParland et al., 2014), and fertility diagnosis (Laine et al., 2013) in dairy cattle.

This innovative approach of using FT-MIR spectroscopy needs the support of important spectral databases associated with reference values for each of the properties to be studied. For this reason, the OptiMIR project was built; it is a European Interreg project involving 6 countries and focuses on the development of prediction tools directly based on FT-MIR spectra. In this work, a large number of commercially available mid-infrared spectrometers (21) from different manufacturers (3) installed in different laboratories (10) located in different countries (3) were used. Because of differences of the instrumental responses between different FT-MIR spectrometers, spectra obtained on one instrument cannot readily be compared with a library acquired on a different instrument. Moreover, the use of calibration models developed on an instrument with FT-MIR spectra obtained on another instrument will usually lead to an increased uncertainty of the prediction model. This is a drawback when recalibrating an instrument or using a historical database. Therefore, spectral corrections adapted to each instrument (standardization procedures) are needed (Rodriguez et al., 2011). One of the most common techniques for instrument standardization is the piecewise direct standardization (PDS) proposed by Wang et al. (1991). However, in previous studies it was mainly used with near-infrared spectra (Bouveresse and Massart, 1996) and was not tested for milk spectra. Then, the objective of this work was to demonstrate and validate the use of the PDS to standardize spectra from different models and manufacturers of instruments, by reducing the inherent instrument-to-instrument variability, within the dairy network, such that the milk spectra from all spectrometers (the slaves) can be compared with the milk spectra of a standard instrument (the master). In such a way, it should be possible to create and maintain international databases containing data collected by all the FT-MIR instruments and to relate them to chemical characteristics of the milk (e.g., protein, fat, fatty acids content among others) and to animal physiology (fertility, nutrition, health, and environment).

MATERIALS AND METHODS

Standardization

Standardization samples are measured on a master instrument and on a slave instrument, leading to response matrices M and S. The PDS method is based on the fact that the variation of spectroscopic data is limited to small spectral regions. In PDS, the response mj measured at wavenumber j on the master instrument is related to the wavenumbers located in a small window (sj) of size n around j (neighboring) measured on the slave instrument (Figure 1). The window (sj) was composed of 5 wavenumbers and was the same for all instruments:

\[ s_j = [S_{(j-n)}, \ldots, S_j, \ldots, S_{(j+n)}]. \]

A regression using the principal component regression method is calculated between each spectral response on the master at wavenumber j and the corresponding window sj on the slave. Vector bj is the vector of transformation coefficients for the jth wavenumber, and b0j is the offset term:

\[ m_j = s_j b_j + b_{0j}. \]

The F matrix contains the bj coefficient transformation vectors for all wavenumbers. This way of calculating the bj parameter using a moving spectral window leads to a banded diagonal matrix. The b0 vector contains the offset terms for all wavenumbers. Each time a new sample is measured on the slave instrument, the obtained spectra Snew can be standardize into \((S_{\text{new}})_{\text{std}}\) using F and b0:

\[ (S_{\text{new}})_{\text{std}} = S_{\text{new}} F + b_0. \]

The standardization model for every master–slave combination needs to be designed, describing the shift between each slave instrument and the master instrument.
Instrumentation

Table 1 shows a summary of the 21 different instruments, located in 10 laboratories, and their characteristics used in this study. The different machine types are FT 6000 and FT+ (Foss, Hillerød, Denmark), FTS (Bentley, Chaska, MN) and Standard Lactoscope FT-MIR automatic (Delta Instruments, Drachten, the Netherlands). All the instruments of this study are located in Germany (11), France (7), and Belgium (3). The wavenumber ranges of the different brand are 925.66 to 5,010.15 cm\(^{-1}\) for Foss FT 6000 and FT+ instruments, 649.03 to 3,998.59 cm\(^{-1}\) for Bentley instruments and 397.31 to 4,000 cm\(^{-1}\) for Delta instruments. The resolution used was 8 cm\(^{-1}\) for Delta and Bentley instruments and unknown for Foss instruments.

Milk Samples

Analysis of identical samples is needed to standardize each machine. To achieve this, several interlaboratory studies were organized; 21 sets of identical samples were distributed to the different laboratories to standardize
the spectra. These sets of samples were produced according to the IDF (International Dairy Federation) norm 141, ISO/DIS 9622 IDF 141 (IDF, 2012). All the sets created consisted of 10 samples of raw milk with large variations in fat (between 1 and 5% mass/vol) and protein (between 2.9 and 5% mass/vol). In all cases, once received, milk samples were homogenized and analyzed at 40 ± 2°C in triplicate in each laboratory following a predefined protocol.

Methodology

An FT 6000 instrument (Foss) was defined as the master instrument. Figure 2 shows a typical milk spectrum measured using the master; as for all milk spectra, the water spectrum was subtracted. Different characteristic peaks can be clearly characterized, as well as a noisy area induced by H₂O absorption. Figure 3 shows the same master spectrum after removing the noisy areas; that is, 1,600 to 1,689 cm⁻¹ and 3,008 to 5,010 cm⁻¹.

The main components of dairy milk can be linked with characteristic bands on milk FT-MIR spectra because of their chemical composition and chemical bonds absorbing light at specific wavenumbers. Lactose induces a response around 1,045 cm⁻¹ with C–O stretching vibration of alcohols functions, 1,076 cm⁻¹ with C–O, C–C, and C–H stretching vibration, and

<table>
<thead>
<tr>
<th>Brand</th>
<th>Type</th>
<th>Number of instruments</th>
<th>Frequency reported by constructors (cm⁻¹)</th>
<th>Number of wavenumbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foss Electric A/S (Hillerød, Denmark)</td>
<td>FT 6000</td>
<td>8</td>
<td>926–5010</td>
<td>1,060</td>
</tr>
<tr>
<td>Foss Electric A/S</td>
<td>FT+</td>
<td>5</td>
<td>926–5010</td>
<td>1,060</td>
</tr>
<tr>
<td>Bentley (Chaska, MN)</td>
<td>FTS</td>
<td>7</td>
<td>649–3999</td>
<td>899</td>
</tr>
<tr>
<td>Delta Instruments (Drachten, the Netherlands)</td>
<td>Lactoscope</td>
<td>1</td>
<td>397–4000</td>
<td>935</td>
</tr>
</tbody>
</table>

Figure 2. Raw milk spectrum from the master instrument. T = transmittance.
1,157 and 1,250 cm$^{-1}$ with C–O–C ether stretching. Proteins appear around 1,550 cm$^{-1}$ with peaks of C–N and N–N stretching. Fat chains appear around 1,390 and 1,454 cm$^{-1}$ with C–H bending of $-\text{CH}_3$ and $-\text{CH}_2$, and around 2,862 and 2,927 cm$^{-1}$ with C–H stretching of $-\text{CH}_3$ and $-\text{CH}_2$. Fat also appears around 1,743 cm$^{-1}$ because of the C=O ester stretching (Socrates, 1980).

In a first step, an interlaboratory study was organized with all the laboratories and apparatus, and the standardization coefficients $F$ and $b_0$ were created as previously explained, using the milk set measured in all the instruments. To validate the standardization method, a fat-prediction model developed on the master instrument was applied on all slave instruments before and after the standardization procedure. The fat model was developed by partial least squares method and based on the whole milk spectrum. Then all the slave predictions were compared with the prediction obtained by the master instrument. In a second step and to really validate the method, the coefficients obtained during the first interlaboratory study were applied to spectra obtained during a second interlaboratory study, realized 1 mo later. All the results were expressed in terms of $R^2$ (determination coefficient), root mean square error ($\text{RMSE}$), slope (deviation from 1), and bias between the slave and the master predictions. Effects of instrument brands, PDS, and interlaboratory studies on these results were assessed by an ANOVA. All computations, chemometric analysis, and graphics were carried out with programs developed in Matlab v7.5.0 (The Mathworks Inc., Natick, MA) and the PLS toolbox v. 4.11 (Eigenvector Research Inc., Wenatchee, WA). For ANOVA the Minitab Statistical Software (Minitab Inc., State College, PA) was used.

**RESULTS AND DISCUSSION**

**Standardization**

To be able to perform the standardization, a first step to harmonize the number of wavenumbers was applied. A linear interpolation was performed on all slave spectra. In the case of the Bentley instruments, spectra...
were interpolated from 649 to 3,999 cm\(^{-1}\) to 926 to 3,008 cm\(^{-1}\). Spectra in transmittance values are transformed into absorbance by a log10. Then the PDS procedure was applied to each of the slave instruments. Figures 4, 5, and 6 illustrate the effect of standardization on spectra from instruments of the 3 brands, on the differences between the absorbance values of master and slave instruments before and after standardization.
and these slave spectra, and on fat predictions from an equation created on the master. For the 3 brands, slave spectra perfectly matched master spectra after PDS, and differences were strongly reduced after standardization of the slaves except in noisy areas that had to be discarded because they were induced by water response. The fat predictions from the slaves were compared with fat predictions from the master before and after standardization. For the 3 brands, the PDS method allowed reduction in the differences between

Figure 5. Effect of standardization of a Bentley (Chaska, MN) slave spectra on differences with master spectra and on fat predictions. (A) Spectra of the master and a Bentley slave instrument before (left) and after standardization (right), from 926 to 3,008 cm⁻¹. (B) Differences between spectra of the master and a Bentley slave instrument before (left) and after standardization (right), from 926 to 3,008 cm⁻¹. (C) Comparison of fat predictions by the master and a Bentley slave before (left) and after (right) standardization. RMSE = root mean square error. T = transmittance. Color version available online.
the slaves and the master predictions. Bias for Delta, Bentley, and Foss instruments, respectively, decreased from −0.5287 to 0.0000, from −0.8654 to 0.0000, and from 0.1557 to 0.0000. The RMSE also decreased for the 3 instruments, respectively, from 0.6114 to 0.0271, from 0.9979 to 0.0127, and from 0.1693 to 0.0056 g of fat/100 mL of milk, showing that the fat prediction developed on the master can be used on instruments from these 3 brands with limited error.

Figure 7 illustrates the RMSE between fat predictions of the master and each slave instrument, before and after standardization. As expected, when comparing fat
predictions before standardization of the slaves with the master, larger differences were found with Delta and Bentley slave instruments than with Foss slaves. For all instruments before standardization, RMSE ranged between 0.0083 and 1.2074. However, after PDS, these values decreased and ranged between 0.0066 and 0.0466 for all slave instruments, allowing a clear use of the fat prediction developed on the master in all the slave instruments included in the interlaboratory study.

The R², slope, bias, and RMSE between master and slave fat predictions before and after PDS were calculated for each instrument and averaged (Table 2) to get an overview of the PDS effect. The R² between master and slaves predictions did not change after standardization ($P = 0.512$) and were higher than 0.999. Slope deviation between predictions was greatly reduced, on average from 0.1379 to 0.0035 ($P = 0.000$). Bias and RMSE between predictions decreased after standardization, from 0.3781 to 0.0000 for bias ($P = 0.000$) and from 0.4609 to 0.0156 on average for RMSE ($P = 0.000$).

Table 3 shows the average of RMSE by brands, before and after PDS. Before standardization the RMSE mean values were different for each brand ($P = 0.001$), i.e., 0.0404, 0.6114, and 1.1001 for Foss, Delta, and Bentley, respectively. After PDS, the averaged RMSE decreased to 0.0133 for Foss, 0.0271 for Delta, and 0.0179 for Bentley. Analysis of variance indicated that these values were not significantly different regarding the 3 brands ($P = 0.279$). For all brands, the PDS method allowed great reduction in the predictions errors with the master predictions.

### Validation of the Standardization Coefficients

The coefficients obtained from the first interlaboratory study were applied to the spectra of a second interlaboratory study 1 mo later. Table 4 shows the results of these standardizations with the first interlaboratory study coefficients. The R² was still higher than 0.999, and no significant differences were observed ($P = 0.283$) with or without PDS.

Slope, bias, and RMSE still decreased greatly after applications of first interlaboratory study coefficients on spectra from the second interlaboratory study. This clearly indicates the stability of the model as well as the PDS procedure between instruments along time. In average, slope, bias, and RMSE were, respectively, reduced from 0.1304 to 0.0093 ($P = 0.000$), from 0.0093 to 0.0035 ($P = 0.000$), and from 0.4118 to 0.0350 ($P = 0.000$), and from 0.4458 to 0.0393 ($P = 0.000$). These results show that standardization coefficients seem to be less adapted to reduce differences between master and slaves spectra 1 mo later after their creation.

This study, through several interlaboratory studies, has also shown that the standardization coefficients can be used during time, even if the error is slightly increasing, which can be explained by physical wear and perturbation on apparatus resulting in spectral deviations in time. Further monthly interlaboratory study should bring more robustness to the standardization coefficients, making the standardization step stable for a long time and allowing a harmonization of the predictions of the network for important farm management indicators (health, fertility, feeding, environmental impact).

### Table 2. Means of statistical results from regressions between fat predictions by the master and the slaves, before and after standardization of interlaboratory study 1 ($n = 20$)¹

<table>
<thead>
<tr>
<th>Item</th>
<th>Before PDS</th>
<th>After PDS</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.512</td>
</tr>
<tr>
<td>Slope</td>
<td>0.1379</td>
<td>0.0035</td>
<td>0.000</td>
</tr>
<tr>
<td>Bias</td>
<td>0.3781</td>
<td>0.0000</td>
<td>0.000</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.4609</td>
<td>0.0156</td>
<td>0.000</td>
</tr>
</tbody>
</table>

¹PDS = piecewise direct standardization; slope = deviation according to 1; RMSE = root mean square error.

### Table 3. Root mean square error (RMSE) between fat predictions of master and slaves, averaged by brand, before and after PDS ($n = 20$)¹

<table>
<thead>
<tr>
<th>Item</th>
<th>Bentley (n = 7)</th>
<th>Delta (n = 1)</th>
<th>Foss (n = 12)</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before PDS</td>
<td>1.1001</td>
<td>0.6114</td>
<td>0.0404</td>
<td>0.001</td>
</tr>
<tr>
<td>After PDS</td>
<td>0.0179</td>
<td>0.0271</td>
<td>0.0133</td>
<td>0.279</td>
</tr>
</tbody>
</table>

¹PDS = piecewise direct standardization. 2Bentley (Chaska, MN), Delta (Drachten, the Netherlands), and Foss (Hillerød, Denmark).
The methodology proposed here will allow a complete correction of physical variation within and between different instruments and brands, allowing the construction of a FT-MIR international database and models. In this work this database has been successfully related to an important chemical characteristic of the milk (fat content). Further work is needed to validate the extension of these results to common calibrations related to animal characteristics such as fertility, nutrition, health, and environment.

CONCLUSIONS

The study described here illustrates that a network of FT-MIR instruments can be standardized; this is illustrated by the reduction of differences between fat predictions of different instruments and brands. The methodology used consisted of a simple interpolation followed by the chemometric PDS tool as a standardization technique. This is an important first step to build a common transnational database with spectra coming from different FT-MIR instruments, including different brands, which should allow the creation of new common indicators for farm management. Robust models built on historical data sets collected by a single instrument or master over several years can then be applied or transferred to other instruments from the same or different brands that are working in the same or different wavenumber ranges.

ACKNOWLEDGMENTS

The authors thank INTERREG and the Walloon Region for their financial support in this project; the laboratories of the OptiMIR project for their collaboration and participation in interlaboratory studies; and Comité du lait (Belgium), AWE (Walloon Breeding Association, Belgium), CRA-W (Walloon Agricultural Research Center, Belgium), LKV-BW (Germany), LKV-NRW (Germany), LACOLAIT (France), CLASEL (France), ADECL62 (France), and GxABT (Belgium) for their help on organization, spectra collection and transfer. The authors acknowledge Ouissam Abbas (CRA-W) for spectral bands interpretation and Claire Darimont for her valuable work on interlaboratory studies organization.
REFERENCES


Socrates, G. 1980. Infrared Characteristic Group Frequencies. 2nd ed. Wiley & Sons Ltd., Chichester, UK.


