Interest, limitations and perspectives of using milk MIR infrared spectra for management and breeding of dairy cattle

H. Soyeurt[1], N. Gengler[1], F. Dehareng[2]

[1] *University of Liège, Gembloux Agro-Bio Tech, TERRA Teaching and Research Centre, Passage des déportés 2, Gembloux, Belgium*
[2] *Walloon Research Centre, Valorisation of agricultural products, Chaussée de Charleroi 24, Gembloux, Belgium*

Abstract

The increasing consumer's concern about the relationship between food and human health as well as the environmental impact of food production combined with the increase of production cost and the variability of milk price must be take into account by dairy herd improvement organisations to develop future management and breeding tools for dairy farmers. In this context, the use of Fourier transform mid-infrared (**FTMIR**) spectrometry is relevant to record phenotypes related to the nutritional quality and process ability of milk, the cow health and well-being and the environment. Those new phenotypes obtained from different equations have not the same accuracy and reliability due to the use of different calibration sets and/or a partial uncovering of the spectral variability. So, even if the computational facility to implement mid-infrared (**MIR**) prediction equations is a reality, some points of attention must be considered to ensure high quality MIR predictions and spectra. This is related to a well-designed calibration set often reached through international collaborations, an appropriate standardization of the spectral data used to perform the prediction and an open communication about the equation accuracy and the representation of the spectral variability of a milk sample to be analysed compared to the used calibration set. More and more investigations are conducted currently to use directly the spectral data to limit the prediction bias. However, although a confirmation of the genetic background related to those particular spectral phenotypes, additional researches must be conducted to define the best methodology to include those records in breeding indices.

Keywords:mid-infrared, milk

Paper

The increasing consumer's concern about the relationship between food and human health as well as the environmental impact of food production require considering the analysis of new characteristics of milk composition. Moreover, the increase of production cost and the variability of milk price force the dairy farmers to optimize their production. Dairy herd improvement (DHI) organisations must face with this new context to develop management and breeding tools to help their farmers. To achieve this objective, the recording of a large number of phenotypes covering the nutritional quality and process ability of milk, the cow health and well-being and the environment is needed. Those phenotypes must be easy to record at individual scale and at low cost. The use of Fourier transform mid-infrared (FTMIR) spectrometry of milk fulfils those conditions. The merging of these novel phenotypes with other information like pedigree information and cow identification and characteristics must be also easy to organize.

Mid-infrared is an infrared radiation located between 3 and 50 µm defined using wavenumbers. The MIR spectrum reflects the absorption of fundamental vibrations related to specific chemical bounds within a molecule (Smith, 1996). So, the milk MIR spectrum is representative of the global milk composition. To quantify a specific milk component, the most informative spectral data points must be combined together. So, the coefficients estimated for a specific prediction equation are applied on the spectral data in order to provide one value representing the estimated quantification of the desired trait, namely called prediction. The first developed equations concerned the measurement of fat and protein as these traits were used for the milk payment and were also useful to appreciate the nutritional quality of the milk produced. Equations related to the prediction of lactose and urea content in milk appeared some years after. Currently the MIR technology is used by all milk laboratories all around the world for milk payment (i.e., farm level) and for milk recording (i.e., cow level) to develop management and breeding tools. Fifteen years ago, research

conducted by Gembloux researchers showed the interest to extend the use of MIR to predict additional phenotypes related to milk quality like the detection of abnormal milk (Pierna et al., 2016), the quantification of fatty acids or minerals (Soyeurt et al. 2006 and 2009). Those first findings have fostered the emergence of new MIR research on milk by different research teams. Currently, many equations were developed by researchers to assess the environmental fingerprint of milk production (e.g., Vanlierde et al., 2015, 2016 and 2018), the nutritional quality of milk (Soyeurt et al., 2009, 2011, 2012), the cow health status (e.g., deRoos et al., 2007; McParland et al., 2011; Grelet et al., 2018), the feed efficiency (McParland et al., 2015) and the technical properties of milk (De Marchi et al., 2009). So, without any doubt, the use of FTMIR spectrometry in milk is of interest to provide new phenotypes useful for the development of the future management and breeding tools.

The use of MIR spectral data is computationally easy but some care must be given to the obtaining of reliable MIR predictions and high quality spectra. The obtaining of reliable predictions is related to two aspects: first, the structure of the calibration set (i.e., reference data used to estimate the coefficients of the established prediction equations) and second the accuracy of the prediction estimated from one or several sets of independent samples. Indeed, the calibration set used must be representative of the dairy cattle population where the prediction equation will be applied. If a part of the variability of the spectral data and/or the studied trait is missing, this leads to extrapolate the prediction and so this increases the chance to decrease the expected accuracy. This expected accuracy must be assessed using independent sets of samples. However, this independency is not always easy to fulfil due to practical or financial issues. Therefore, it is always required to know the methodology used to estimate the accuracy of a prediction given by an equation. Unfortunately, currently the information about the calibration structure and, sometimes, the overall accuracy is missing. So, DHI organizations can apply prediction equations with honesty thinking that the prediction is reliable and accurate but nothing is currently available to confirm this statement. More than the equation itself, the accuracy of the prediction depends also on the quality of the MIR spectral data provided by the spectrometer. Some routines are implemented in milk laboratories to ensure the quality of spectral data by analyzing a stable reference substance through time. Unfortunately, the composition of this substance does not present the same level of milk composition complexity. This can lead to a differential effect of this standardization on milk MIR spectrum per wavenumber.

So the perspectives of the MIR are related to those constraints. There is a need to establish a guide of good practices to predict MIR phenotypes and to ensure a high quality of MIR data. Some investigations and solutions exist to solve those problems. The accuracy of the developed equation can be communicated by the manufacturer or by the team responsible of this equation. This is largely done currently. However, to know if the spectral variability of the analyzed milk sample is included in the calibration set, the standardized Mahalanobis distance of the spectrum to be predicted to calibration dataset must be estimated. Through a principal component analysis, this could be possible by using eigenvectors calculated from the calibration used by the manufacturer. This is not possible currently if you have not an access to the calibration set. However, this could be the best way to avoid extrapolation during the prediction process based on uncovered spectral variability. Moreover, the current prediction equations used by milk laboratories can be developed from different brands of spectrometers and therefore from different calibration sets. This lead potentially to a variable accuracy of the prediction which can impact the development of management or breeding tools. The creation of uniformed prediction equation is therefore of interest. The standardization of MIR data as proposed by Grelet et al. (2015 and 2017) using milk samples with known milk composition allows ensuring high quality MIR spectral data between milk laboratories and also the standardization of MIR spectra between different brands. This opens therefore the possibility to promote a limited number of prediction equations for a specific traits used by all dairy stakeholders.

Based on our experience, the creation of prediction equations by research teams with the collaboration of many farms located in different place could 1) increase the variability of the studied traits and the spectral data improving the robustness of the developed equation and, 2) facilitate the communication about the accuracy of the prediction and the estimation of standardized Mahalanobis distance. The creation of online platform such as the European milk recording (www.milkrecording.eu) opens this perspective.

We cannot finish this paper without mentioning that all of those new phenotypes were finally related to the same raw information, the spectral data. The combination of multiple phenotypes coming from the same origin can increase potentially the prediction bias. So, more and more initiatives are emerging to try to use directly the MIR spectral data especially for breeding purposes (Soyeurt et al., 2010; Bittante et al., 2013; Wang et al., 2016). Those studies confirmed the genetic background related to the wavenumbers of the milk MIR spectra. This was expected as the spectral data represent the global milk composition whose components have a known genetic variability. However, the direct use of milk MIR spectral data base breeding values is not easy as each individual wavenumber is not a direct representative of any specific trait contained in the breeding goal. Additional researches are still needed to solve this problem.

In conclusion, the use of FTMIR spectrometry in milk to enhance the number of phenotypes to consider different aspect of the milk production is of interest. But this emerging intensive use of MIR does not hide the complexity behind the obtaining of predictions and spectra to ensure the highest possible quality.

REFERENCES

Bittante G and Cecchinato A (2013) Genetic analysis of the Fourier-transform infrared spectra of bovine milk with emphasis on individual wavelengths related to specific chemical bonds. *Journal of Dairy Science* 96(9): 5991–6006. DOI: 10.3168/jds.2013-6583.

De Marchi M, Fagan CC, O'Donnell CP, et al. (2009). Prediction of coagulation properties, titratable acidity, and pH of bovine milk using mid-infrared spectroscopy. *Journal of Dairy Science* 92(1):423-432. DOI:10.3168/jds.2008-1163.

de Roos APW, van den Bijgaart HJCM, Hørlyk J, et al. (2007) Screening for subclinical ketosis in dairy cattle by fourier transform infrared spectrometry. *Journal of Dairy Science* 90(4): 1761–1766. DOI: 10.3168/jds.2006-203.

Grelet C, Fernández Pierna JA, Dardenne P, et al. (2015) Standardization of milk mid-infrared spectra from a European dairy network. *Journal of Dairy Science* 98(4): 2150–2160. DOI: 10.3168/jds.2014-8764.

Grelet C, Pierna JAF, Dardenne P, et al. (2017) Standardization of milk mid-infrared spectrometers for the transfer and use of multiple models. *Journal of Dairy Science* 100(10): 7910–7921. DOI: 10.3168/jds.2017-12720.

Grelet C, Vanlierde A, Hostens M, et al. (2018) Potential of milk mid-IR spectra to predict metabolic status of cows through blood components and an innovative clustering approach. *Animal*: 1–10. DOI: 10.1017/S1751731118001751.

McParland S, Banos G, Wall E, et al. (2011) The use of mid-infrared spectrometry to predict body energy status of Holstein cows1. *Journal of Dairy Science* 94(7): 3651–3661. DOI: 10.3168/jds.2010-3965.

McParland S, Kennedy E, Lewis E, et al. (2015) Genetic parameters of dairy cow energy intake and body energy status predicted using mid-infrared spectrometry of milk. *Journal of Dairy Science* 98(2): 1310–1320. DOI: 10.3168/jds.2014-8892.

J.A. Fernández Pierna, D. Vincke, V. Baeten, et al. (2016) Use of a multivariate moving window PCA for the untargeted detection of contaminants in agro-food products, as exemplified by the detection of melamine levels in milk using vibrational spectroscopy, *Chemometrics and Intelligent Laboratory Systems*, 152:157-162, DOI:10.1016/j.chemolab.2015.10.016.

Smith BC (1996) *Fundamentals of Fourier Transform Infrared Spectroscopy. Florida: CRC Press LLC*. CRC Press.

Soyeurt H, Dardenne P, Dehareng F, et al. (2006) Estimating fatty acid content in cow milk using mid-infrared spectrometry. *Journal of Dairy Science* 89(9): 3690–3695. DOI: 10.3168/jds.S0022-0302(06)72409-2.

Soyeurt H, Bruwier D, Romnee J-M, et al. (2009) Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *Journal of Dairy Science* 92(6): 2444–2454. DOI: 10.3168/jds.2008-1734.

Soyeurt H, Misztal I and Gengler N (2010) Genetic variability of milk components based on mid-infrared spectral data. *Journal of Dairy Science* 93(4): 1722–1728. DOI: 10.3168/jds.2009-2614.

Soyeurt H, Dehareng F, Gengler N, et al. (2011) Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *Journal of Dairy Science* 94(4): 1657–1667. DOI: 10.3168/jds.2010-3408.

Soyeurt H, Bastin C, Colinet FG, et al. (2012) Mid-infrared prediction of lactoferrin content in bovine milk: potential indicator of mastitis. *animal* 6(11). Cambridge University Press: 1830–1838. DOI: 10.1017/S1751731112000791.

Vanlierde A, Vanrobays M-L, Dehareng F, et al. (2015) Hot topic: Innovative lactation-stage-dependent prediction of methane emissions from milk mid-infrared spectra. *Journal of Dairy Science* 98(8): 5740–5747. DOI: 10.3168/jds.2014-8436.

Vanlierde A, Vanrobays M-LM-L, Gengler N, et al. (2016) Milk mid-infrared spectra enable prediction of lactation-stage-dependent methane emissions of dairy cattle within routine population-scale milk recording schemes. *Animal Production Science* 56(3): 258–264. DOI: 10.1071/AN15590.

Vanlierde A, Soyeurt H, Gengler N, et al. (2018) Short communication: Development of an equation for estimating methane emissions of dairy cows from milk Fourier transform mid-infrared spectra by using reference data obtained exclusively from respiration chambers. *Journal of Dairy Science* 101(8): 7618–7624. DOI: 10.3168/JDS.2018-14472.

Wang Q, Hulzebosch A and Bovenhuis H (2016) Genetic and environmental variation in bovine milk infrared spectra. *Journal of Dairy Science* 99(8): 6793–6803. DOI: 10.3168/jds.2015-10488.