Review *Czech Journal of Animal Science, 70, 2025 (1): 1–16*

https://doi.org/10.17221/165/2024-CJAS

Mid-infrared milk screening as a phenotyping tool for feed efficiency in dairy cattle

Ludmila Zavadilová[1]*, *Eva Kašná*[1], *Zuzana Krupová*[1], *Alena Pechová*[2,3], *Petr Fleischer*[3], *Soňa Šlosárková*[3]

[1]*Institute of Animal Science, Prague, Czech Republic*

[2]*Department of Animal Nutrition and Forage Production, Faculty of AgriSciences, Mendel University in Brno, Brno, Czech Republic*

[3]*Veterinary Research Institute, Brno, Czech Republic*

**Corresponding author: zavadilova.ludmila@vuzv.cz*

Citation: Zavadilová L., Kašná E., Krupová Z., Pechová A., Fleischer P., Šlosárková S. (2025): Mid-infrared milk screening as a phenotyping tool for feed efficiency in dairy cattle. Czech J. Anim. Sci., 70: 1–16.

Abstract: Feed efficiency (FE) is one of the most essential traits in dairy cattle, primarily due to the high cost of feed, which constitutes a significant portion of dairy herd expenses. Unfortunately, assessing FE in individual cows requires precise measurement of feed consumption, a labour-intensive and expensive process that is impractical for group-fed cows on production farms. Efforts have been made to predict FE or, more precisely, dry matter intake (DMI), using predictors such as a body weight (BW), milk yield (MY), and milk composition. Recently, Fourier transform mid-infrared (FT-MIR) spectroscopy has been proposed as a tool to enhance the accuracy of DMI prediction. This paper reviews the application of FT-MIR milk spectroscopy for deriving FE phenotype in dairy cattle. FT-MIR is a reliable and widely used method for routine analysis of milk components. In FE phenotyping, predictive equations often incorporate FT-MIR alongside other traits such as BW, MY, milk composition, herd, breed, days in milk, and pregnancy. The most commonly used mathematical approaches are partial least squares (PLS) regression and artificial neural networks (ANN). Prediction accuracy varies across studies, depending on the mathematical method and model employed. Predictions based solely on FT-MIR data have demonstrated moderate accuracy (coefficient of determination), ranging from 0.19 to 0.40. However, integrating all data sources including MY, milk composition, FT-MIR, and near-infrared reflectance spectroscopy (NIR) is crucial and results in higher accuracy, with reported values ranging from 0.03 to 0.81.

Keywords: feed intake; Holstein cows; milk mid-infrared spectroscopy

Feed efficiency (FE) in dairy cows is crucial for dairy farm productivity. The importance of FE lies in the cost of feed, which constitutes a significant portion of dairy herd expenses. A major obstacle to the direct inclusion of feed efficiency in dairy breeding programmes is the inability to regularly and inexpensively acquire information on individual feed consumption on commercial farms. Individual measurement of feed consumption using feed troughs is expensive and labour-intensive

---

Supported by the Ministry of Agriculture, Czech Republic (Grant No.: NAZV QL24010350), and Institutional Support (Support No.: MZe-RO0723).

© The authors. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

and therefore, is only used under experimental conditions. Dairy cows' milk yield (MY), milk composition, body weight (BW) and information on lactation stage and reproductive status can be used to predict dry-matter intake (DMI). These traits, along with FE predictive ability, also include Fourier-transform mid-infrared spectroscopy (FT-MIR), a simple and accurate method widely used to routinely determine the content of milk components, such as fat and protein (Soyeurt 2011). These measurements are subsequently applied to evaluate milk economically or are used in dairy cattle breeding. FT-MIR uses light from the mid-infrared region to scan milk samples and identify the presence of specific chemical bonds. The results are presented as an absorption profile, which consists of absorbance values for individual infrared light wavenumbers across the mid-infrared region. Traits and phenotypes, such as fat, protein, or lactose content, are predicted as a function of the individual FT-MIR wavenumber absorbances (Tiplady 2020). The predicted traits are not limited to phenotypes linked to milk composition and traits, for example, the composition of milk fatty acids (Soyeurt et al. 2011), milk protein composition (Rutten et al. 2011), milk coagulation (Dal Zotto et al. 2008), milk acidity (De Marchi et al. 2009) or cheese yield, curd nutrient recovery (Cecchinato et al. 2015) and ketone bodies (Klein et al. 2012). It turns out that other cattle phenotypes can also be determined based on FT-MIR milk analysis. Currently, it is energy status (Ho et al. 2019), pregnancy status (Toledo-Alvarado et al. 2018), feed and DMI (Shetty et al. 2017), and methane emissions (De Haas et al. 2011).

The relationship between FT-MIR milk absorption profiles and numerous traits or phenotypes is based on molecular and biological mechanisms underlying phenotypic trait expression and is linked to the animal genome (Tiplady 2021). These relatively inexpensively obtained phenotypes could play a crucial role in the genetic and genomic selection of dairy cattle. De Haas et al. (2011) and Wall et al. (2010) reported that FT-MIR milk-predicted traits could be as effective in genetic or genomic selection as directly measured traits, provided there is a close relationship between the actual phenotype and its prediction, and the traits are predicted to a greater extent. Interest is growing increasingly in analysing individual FT-MIR wavenumbers and in utilising FT-MIR data to predict other novel traits relevant to the industry. Furthermore, spectra are already available as a by-product of routine dairy cattle milk recording and testing (Tiplady 2020).

Here, we review the use of FT-MIR to acquire FE phenotype in dairy cattle. We discuss the genetics of FT-MIR data and its genetic relationship to FE in cattle. We also review the studies on FT-MIR-predicted FE, with an emphasis on current research.

What is Fourier transform mid-infrared spectroscopy of milk?

The FT-MIR of milk samples produces an absorption profile, which consists of absorbance values for individual infrared light wavenumbers across the mid-infrared region. According to Bittante and Cecchinato (2013), the resulting milk spectroscopy data can be divided into several regions based on wavelengths: the short-wavelength infrared region (SWIR) from 5 011 to 3 673 cm^{-1}; the transition region between SWIR and the mid-wavelength infrared region (SWIR-MWIR) from 3 669 to 3 052 cm^{-1}; FT-MIR, divided into two mid-infrared regions: MWIR1, from 3 048 to 1 701 cm^{-1}, and MWIR2, from 1 698 to 1 586 cm^{-1}. The transition region between MWIR and the long-wavelength infrared (LWIR) ranges from 1 582 to 925 cm^{-1} (MWIR-LWIR).

The MIR spectrum wave points show different variations between animals, as shown in Figure 1 by Dorea et al. (2018). They used only 361 points from 1 060, with a coefficient of variation greater than 1%.

The heritability of wavelengths in the mid-infrared region of the spectrum was confirmed. For milk spectral data, Rovere et al. (2019) found that the MWIR (3 000–2 500 cm^{-1}) and MWIR-LWIR (1 500–925 cm^{-1}) regions primarily include waves with moderately high (≥ 0.4) heritabilities. Rovere et al. (2019) explained this by associating specific wavenumbers with milk metabolites and components, such as protein and fat. Additionally, most of the heritability peaks were found in regions associated with milk components as well as regions linked to reproductive outcomes, such as conception. The wavelengths with very low heritability showed high absorbance of water molecules. Rovere et al. (2019) also found changes in heritabilities during lactation and between parities. They explained that milk composition changes are due to molecular-biological processes in the cow's body,

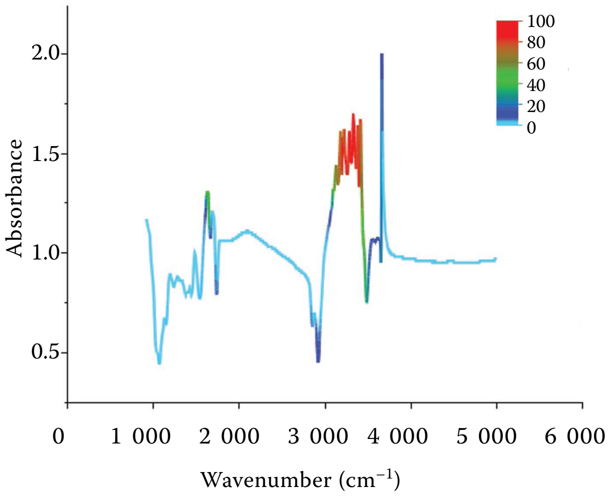

Figure 1. Coefficients of variation (CV) of the light absorbance (vertical axis) for the wavenumbers (horizontal axis).

The CV of absorbance for individual wavenumbers (cm^{-1}) (number of wavenumbers = 1 060) of individual animal milk samples (number of samples = 1 279) obtained through milk Fourier transform mid-infrared spectroscopy; the colour scale of the CV, ranging from light blue (lower CV) to dark red (higher CV), represents increasing CV values; of the 1 060 wavenumbers, only 361 had a CV greater than 1% across animals

Source: Dorea et al. (2018)

such as pregnancy and milk production. These processes have a genetic basis; thus, the genetic background is involved in changes in milk and, subsequently, in changes in spectral absorbance values, as those changes are detected in milk samples.

The wavebands obtained in milk spectral analysis are related to FE in cattle due to the various milk components. These components are influenced by the animal's feed ration and metabolism, as demonstrated in the studies by Shetty et al. (2017) and McParland et al. (2014). To predict DMI and RFI, Shetty et al. (2017) found that the key spectral wave numbers include peaks for fat, protein, and lactose. As shown in Table 1, Shetty et al. (2017) explained the relationship between several wavenumbers of FT-MIR spectra and specific nutrients and chemical bonds in the assessment of dry matter intake. Fat A is a group of milk lipids with strong absorbance bands due to the carbonyl (C=O) group in the region between ~1 700 and 1 800 cm^{-1}, while Fat B exhibits absorbance bands due to the C–H group in the region from ~2 800 to 3 000 cm^{-1} (Inon et al. 2004). The ability of FT-MIR spectral data to predict feed or energy intake can be explained by the fact that the fat-to-protein ratio or milk fatty acid composition is associated with energy balance (McParland et al. 2014). The milk composition detected by FT-MIR data must depend not only on the feed composition but also on the processing and use of the feed by the animal.

Feed efficiency in cattle: A brief introduction to the topic

FE was mentioned among the "new traits" introduced into cattle breeding, alongside milk fatty acid content (Bastin et al. 2011), hoof health (Malchiodi et al. 2017), methane emissions (De Haas et al. 2011) and the use of data from automated milking systems (Jacobs and Siegford 2012). Along with the genetic evaluation of health traits in Scandinavian countries (Heringstad et al. 2000), these are considered milestones in the selection of dairy cattle for sustainable agriculture, focusing on breeding healthy and fertile animals with high productivity and the production of healthy food (Miglior et al. 2017).

Table 1. Important spectral wavenumbers selected by recursive weighted partial least squares regression for dry matter prediction in early lactation, along with the corresponding chemical functional groups, milk components, and feed and nutrient information

Wavenumber (cm^{-1})	Milk components	Nutrients	Feed	Bond	Functional group
3 003, 2 976	fat B	fatty acids, acetic, butyric	fermentable fibre	C–H	alkyl chain
1 789	fat A	fatty acids, acetic, butyric	fermentable fibre	C=O	carbonyl group
1 249	protein (amide III)	amino acids	crude protein	C–N	aromatic amines
1 079, 987, 968, 948, 944	Lactose	propionic (glucose)	sugar starch	O–H	hydroxyl group

Source: adapted from Shetty et al. (2017)

VandeHaar et al. (2016) stated that FE is a complex trait with no simple or singular definition. It can be defined at the animal level, the farm level, or even on a global scale. For animals, FE is often expressed per unit of production. The efficiency of nutrient conversion from feed can be understood as the proportion of feed energy or dry matter captured in production. According to VandeHaar et al. (2016), the fundamental concept of FE is derived from partitioning the gross (GE) energy of feed. GE is divided into components, including energy lost through faeces, gas, urine, and other waste, as well as heat generated during the digestion and metabolism of feed. The remaining portion, known as net energy of feed (NE), is used for maintenance and is partly converted into dairy or meat products.

As expected, the importance of FE lies in the cost of feed, which constitutes a large portion of the expenses for dairy herds. Over the past 100 years, FE has increased in proportion to cattle milk production. FE enhancement through increased milk yield has been achieved through breeding, improved nutrition, and better management. There has been a two-fold increase in FE in the US, and importantly, the share of FE captured in milk has significantly increased compared to the share allocated to a dairy cow's maintenance needs (VandeHaar et al. 2016). For further improvement of FE, breeding should focus directly on FE rather than indirectly on dairy production, as the potential for increasing FE through increasing milk productivity has already been exhausted (VandeHaar et al. 2016). It means selecting the FE trait, which is defined differently from the gross FE (GFE) traits that have strong positive genetic correlations with milk yield (Spurlock et al. 2012).

GFE is one of the fundamental definitions of FE in dairy cattle, expressed as the ratio of milk output to feed input (Connor 2015). Milk yield is typically expressed as solids-corrected or energy-corrected milk yield. Feed input is measured either as DMI or energy intake. As Connor (2015) stated, the heritability of GEF is moderate, ranging from 0.14 to 0.37.

Another term used to assess and define the FE of dairy herds is income over feed cost (IOFC) or return over feed (ROF) (Connor 2015). IOFC derives FE directly from profitability. Generally, it is expressed as the difference between the total revenue from milk sales and the feed costs associated with its production. It is heavily influenced by feed costs and milk prices, both of which fluctuate considerably. Calculating IOFC can be challenging, especially for individual cows, due to the difficulty in obtaining accurate data on individual cow feed intakes. As one of the measures of FE, residual feed intake (RFI) was proposed by Koch et al. (1963). RFI represents the difference between the actual feed or energy intake and predicted feed or energy intake for the animal. RFI is independent of body size and performance level. RFI reflects hereditary variation in key metabolic processes and enables the capture of differences in FE between animals (Connor 2015). How do we define FE as RFI? Suppose FE is expressed as DMI. In that case, RFI is calculated as a residual from a regression model where DMI is the dependent variable, and milk production, weight (maintenance ration), and other feed energy requirements of the animal, such as the stage of pregnancy, are represented as explanatory variables in the model (Pryce et al. 2014; Negussie et al. 2019). In short, RFI is the difference between the actual DMI and the predicted DMI, which is estimated using the explanatory variables (Connor 2015).

However, the first step in the selection process for improving FE is defining and measuring the FE phenotype, specifically individual feed consumption, expressed as DMI. Assessing FE in an individual cow involves measuring feed consumption, which requires individual feed boxes equipped with electronic cow identification and the recording of feed weight. This labour-intensive and costly approach is not feasible for group feeding of cows on production farms (Negussie et al. 2019). Expensive experiments to determine feed consumption are typically limited to a few animals. Note that the FE obtained from a few animals must be used for selection across the entire population.

The next steps in the selection process involve estimating genetic parameters of FE phenotypes, including heritability and correlations with other selection index traits, before applying genomic selection and selection indexes (Pryce et al. 2014; Rovere et al. 2019).

The heritability of RFI is intermediate (0.18–0.40) and varies throughout lactation (0.18 in the middle of lactation) (Negussie et al. 2019). The heritability of FE traits, combined with the use of genomic methods, enables the prediction of breeding values (Pryce et al. 2014).

For a direct approach to estimating breeding values (EBV) for RFI based on the measured phenotype, we refer to the method described by Lidauer

https://doi.org/10.17221/165/2024-CJAS

et al. (2023), where breeding values for RFI and the partial regression coefficients for the energy sink traits are estimated simultaneously:

$$DMI_{hijkl} = HYS_i + c_{hl}\phi_{jl} + \gamma_E \times ECM_{jk} + \\ + \gamma M \times MBW_{jk} + \gamma L \times BWloss_{jk} + \\ + \gamma_G \times BWgain_{jk} + p_k + a_k + e_{hijk} \quad (1)$$

where:

DMI_{hijk} – the DMI observation for cow k, recorded in her contemporary group i and lactation week j;

HYS_i – the fixed effect of herd × year × season contemporary group i;

$c_{hl}\ \phi_{jl}$ – the fixed regression function on lactation week j nested within herd h; ϕ_j is a vector containing the covariates of a third-order Legendre polynomial plus the exponential term $e^{-0.05j}$ for lactation week j;

$\gamma_E, \gamma_M, \gamma_L, \gamma_G$ – partial regression coefficients on a cow k's observations in her lactation week j for energy-corrected milk (ECM), maintain 1 $kg^{0.75}$ of metabolic body weight (MBW, MBW = $BW^{0.75}$), BW loss and BW gain, respectively;

p_k – the random permanent environmental effect of cow k [$p \sim N(0, I\sigma^2_p)$, where I – an identity matrix and σ^2_p is the permanent environmental variance];

a_k – the random additive genetic effect of cow k [$a \sim N(0, A\sigma^2_a)$, where A is the numerator relationship matrix, and σ^2_a is the additive genetic variance];

e_{hijk} – the random residual effect [$e \sim N(0, I\sigma^2_e)$ where I is an identity matrix and σ^2_e is the residual variance.

The EBV for RFI of cow k is the difference between the estimate of the additive genetic effect for cow k, a_k, and the average of all a estimates where a negative estimate of a is equal to the amount of feed saved.

In Australia, Pryce et al. (2015) used two-step genomic methods to predict the genomic breeding values (GEBVs) for feed saved. The feed saved trait combines GEBV for RFI and the maintenance feed requirement based on body weight (BW). RFI was available only for genotyped animals, while BW was known for all animals. The GEBV for RFI was predicted using a multi-trait model with 2 036 animals as the genomic reference population: Australian calf RFI, Australian lactation cows' RFI, and UK and Dutch lactation cows' RFI, all of which had RFI phenotypes. The RFI breeding value for 4 106 genotyped sires without RFI phenotypes was calculated based on the single nucleotide polymorphism effects of genomic BLUP. The sires had BW breeding values. Notably, the feed saved GEBV was calculated using only the BW information for non-genotyped animals.

Negussie et al. (2019) analysed several scenarios of daily DMI recording to determine which scenario is the most advantageous in terms of RFI GEBV accuracy and, at the same time, is not excessively demanding for DMI measurement. They concluded that a higher number of cows per sire is beneficial for achieving higher RFI GEBV accuracy. It is clear that RFI GEBV accuracy decreased with increasing intervals between DMI measurements, but this can be compensated for by increasing the number of monitored cows per bull. What is essential is that the total number of DMI measurements decreases; therefore, the farm's costs for this measurement also decline. Negussie et al. (2019) stressed that the recording of DMI should be limited to specific stages of lactation because DMI in different stages of lactation is only partly related to other stages. Therefore, using measurements throughout the entire lactation leads to lower RFI GEBV accuracy. Similarly, focusing only on a specific part of lactation limits the number of cows monitored and also decreases reliability. This can be compensated for by repeating the recording for more groups of cows.

Lidauer et al. (2023) also suggested a new FE trait, a new RFI-derived metric – regression on expected feed intake (ReFI). This metric is based on regressing DMI on expected DMI using a random regression linear animal model, which calculates expected DMI using energy requirement formulations. The crucial difference from RFI is that this new metric is independent of DMI. Using ReFI for genetic evaluation instead of RFI resulted in the reranking of cows and the selection of the most influential animals. All this is a result of a better-modelled feed intake in the model. The ReFI model (Lidauer et al. 2023):

$$DMI_{ijk} = \beta_i \times eDMI_{jk} + \psi_k \times eDMI_{jk} + \alpha_k \times \\ \times eDMI_{jk} + \varepsilon_{ijk} \quad (2)$$

where:

DMI_{ijk} – a DMI observation for cow k, recorded in her contemporary group i and lactation week j;

β_i – a fixed regression coefficient for the herd × year × season contemporary group i;

ψ_k – a random regression coefficient for the permanent environmental effect of cow k [$\psi \sim N(0, I\sigma^2_\psi)$, where I is an identity matrix and σ^2_ψ is the variance of the permanent environmental regression coefficients];

α_k – a random regression coefficient for the additive genetic effect of cow k [$\alpha \sim N(0, A\sigma^2_\alpha)$, where A is the numerator relationship matrix and σ^2_α is the variance of the additive genetic regression coefficients];

ε_{ijk} – the random residual effect $\varepsilon \sim N(0, I\sigma^2_\varepsilon)$, where I is an identity matrix; $I\sigma^2_\varepsilon$ is the residual variance, and $eDMI_{jk}$ is the expected DMI ($eDMI$), for cow k in lactation week j.

The *eDMI* is calculated using the Equation (3).

$$eDMI = \frac{ER_{jk}}{ED} \tag{3}$$

where:

ER_{jk} – the energy requirement for cow k in lactation week j;

ED – the average energy density, calculated as the average of all ER_{jk} values divided by the average of all DMI_{jk} observations.

In their study, Lidauer et al. (2023) calculated ER_{jk} using coefficients from several dairy cattle nutrition studies, applying the following equation:

$$ER_{jk} = 4.81 \times ECM_{jk} + 0.603 \times MBW_{jk} + 27.6 \times \times BWloss_{jk} + 34.8 \times BWgain_{jk} \tag{4}$$

Where the regression coefficients represent the expected amount of MJ ME needed to produce 1 kg of energy-corrected milk (ECM), maintain 1 $kg^{0.75}$ of metabolic body weight (MBW), the amount of ME made available through the mobilisation of 1 kg of body tissue, i.e. BW loss, and the ME required to gain 1 kg of BW, respectively. Note that BW loss and BW gain have values ≥ 0. The applied regression coefficients in (Equation 4) are averages of estimates based on various energy-feeding systems reported by (Agnew et al. 2003).

The EBV for ReFI for cow k is the difference between the estimate of the additive genetic effect of cow k, α_k (Equation 2) and the fixed regression coefficient estimate β_i (Equation 2), where a negative value equals the percentage of feed saved.

In conclusion, this section states that the essential prerequisite for predicting EBVs for FE traits is the knowledge of individual DMI for some animals in the population.

Mid-infrared spectroscopy in FE phenotyping

Differentiating cows based on FE is not computationally complex. The main obstacle lies in data acquisition. If feed consumption measurements were available for every dairy cow, the FE calculations would be carried out routinely. Alternatively, FE could be predicted during the monthly performance check. The resulting FE phenotype would be entered as a dependent variable in the prediction of EBVs and GEBVs. Some information about the animals and their phenotypes, such as milk yield, lactation stage, age, or origin, is collected regularly. Other phenotypes are more challenging to measure, such as the live weight of animals; they can be estimated using conversion coefficients from body dimensions such as chest girth (Veerkamp and Brotherstone 1997). Obtaining feed consumption is the most costly and labour-intensive. As mentioned, suitable predictors are being sought for the FE phenotype to avoid the need for direct measurement of individual feed consumption.

Therefore, we can conclude from the previous text that it is necessary to conduct an experiment in which data on animal phenotype such as feed consumption, BW data, production, parity and days of lactation, as well as proposed predictors like FT-MIR, are collected. The next step involves subjecting the collected data on feed consumption and predictor traits to statistical analysis to derive a prediction equation. The derived prediction equation includes estimates of regression coefficients for specific predictor traits. What is crucial is that these predictors should be much easier to measure and obtain than feed consumption, preferably for large populations. This is particularly true for spectrometric data routinely measured in dairy cattle milk recording. Figure 2 illustrates the process from the DMI acquisition experiment to the prediction of FE breeding values for cows, without direct measurement of DMI. The process described in Figure 2 can be applied to other feed efficiency indicators, such as those based on energy intake.

https://doi.org/10.17221/165/2024-CJAS

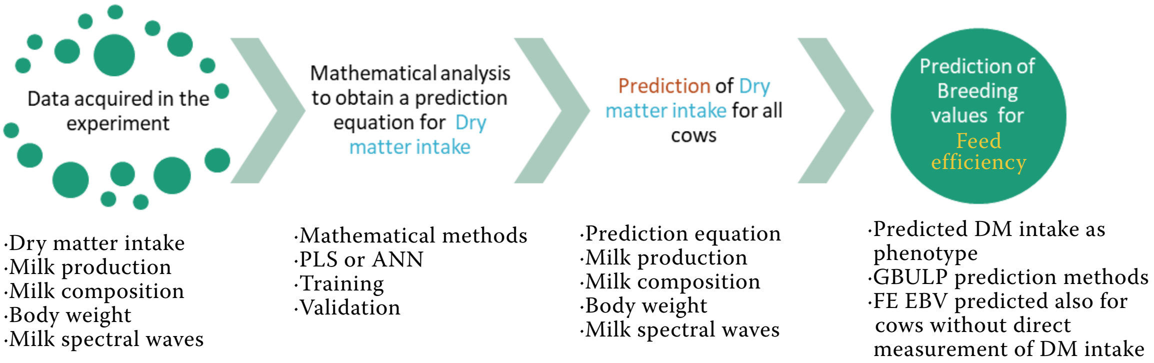

Figure 2. Prediction of breeding values (EBV) for feed efficiency (FE) with milk spectral waves (FT-MIR) included in the prediction of dry matter intake

ANN = artificial neural networks; PLS = partial least squares

In FE phenotyping, the use of mid-infrared spectroscopy defines the FE phenotype in dairy cows as inexpensive, though with limited accuracy. FE can be predicted based on other predictors, such as a cow's weight (BW), milk production (MY), and milk composition (McParland et al. 2014; Shetty et al. 2017; Dorea et al. 2018). However, it is proposed that FT-MIR data derived from milk composition can be used to increase the accuracy of FE prediction. Using a predictor trait, such as phenotypes derived from FT-MIR spectra, may provide a cost-effective indirect measurement of FE, eliminating the need for expensive feed intake measurements for many cows.

Data required for the development of a prediction model

Individual feed consumption. To determine FE in dairy cows, individual feed consumption during lactation must be measured. Then, the DMI or other indicators, such as energy needs, are derived from the obtained individual feed consumption. Obtaining individual feed consumption requires measuring feed intake using feed troughs. Any other approaches for measuring feed consumption, such as using camera systems, are sophisticated estimates. Expensive monitoring of individual feed consumption results in a low number of monitored animals. In the case of dairy cows, the monitored animals should be in approximately the same stage of lactation, have a comparable parity, and belong to the same breed. However, due to the lack of data, information obtained from different trials and breeds is often combined (Dorea et al. 2018; Wallen et al. 2018).

Individual predictive traits, besides FT-MIR waves. Individual predictive traits or predictors are characteristics of animals that can be used to predict feed consumption and energy expenditure. They are divided to traits describing body size or condition, such as BW and BCS, and those related to production, such as MY, fat, protein, and lactose content. Besides this, information on the stage of lactation and other factors that become part of model equations is crucial. This situation arises when the dependent variable of the equation, i.e., the evaluated phenotype, is measured in different environments, such as herds. Shadpour et al. (2022) addressed this situation by including several fixed effects, such as country, season of calving, and lactation number, in the model equation. They obtained the DMI of cows from experiments conducted in Canada, the United States, and Denmark. This fact had to be considered in the formation of the prediction equations. Interestingly, these sets of effects contributed little to the prediction accuracy of the final models (Shadpour et al. 2022).

The BW provides essential information. However, BW data is not readily available on every farm. Therefore, different types of models with varying parameters would offer choices and enable users to predict DMI with reasonable accuracy. Shetty et al. (2017) stated that BW had a crucial influence on the model's predictive ability. As Shadpour et al. (2022) point out, BW can be estimated using, for example, chest width (Veerkamp and Brotherstone 1997). BW is often used as MBW (Dorea et al. 2018; Shadpour et al. 2022).

Milk yield provides useful information on dairy cows' energy needs. It has proven to be one of the primary predictive characteristics for determining FE in dairy cows. To clarify, daily milk yield (MY) refers to the milk production on the control day, linked to the milk sample for FT-MIR analysis. McParland et al. (2014) found that MY, as a predictor variable, improved the accuracy of RFI across lactation but failed to improve the accuracy of predicting effective energy intake (EEI) when added to FT-MIR data.

Milk fat, protein, and lactose content are other milk characteristics linked to energy needs and, therefore, to FE. The key point is that milk composition (McDermott et al. 2016) and fatty acids (Soyeurt et al. 2011) are closely related to feed and are regularly assessed by FT-MIR.

Illustrative experiments using FT-MIR data to predict FE traits

The list of studies on the prediction of FE based on FT-MIR shows that this topic is being continuously investigated to expand on previously established knowledge. For additional information, see the supplemental table. Readers interested in this area of the research are provided with key publications.

McParland et al. (2014) analysed records for energy intake, RFI, and milk MIR spectral data from an Irish research herd across 36 test days, covering 535 lactations in 378 cows. They employed FT-MIR data from a.m. milk samples or/and p.m. samples or/and MY for prediction. All milk samples were analysed using the same MIR spectrometer (Foss MilkoScan FT6000; Foss Electric A/S, Hillerød, Denmark).

Shetty et al. (2017) used milk yield and content data, FT-MIR spectrometric data from Milkoscan FT^{+}, and feed consumption to predict DMI or RFI in Danish Holstein and Danish Jersey cows. Milk yield and feed consumption were measured daily, and milk samples were collected 2 to 6 times a week. The FT-MIR spectrometric data were analysed in milk samples using Milkoscan FT^{+}, corresponding to MY, fat, protein, and lactose contents.

Wallen et al. (2018) also used the MilkoScan FT6000 MIR spectrometer. They predicted DMI and net energy intake (NEI) in Norwegian Red cows. In addition to spectrometric data, they employed a predictive model of milk composition. The highest predictive accuracy of models for both DMI and NEI was achieved when MY, BW, and FT-MIR data were included in the prediction equations.

Dorea et al. (2018) compared two mathematical methods: PLS and ANN. ANN was combined with a Bayesian network (BN) to predict DMI using FT-MIR variables such as MY, BW, and milk composition. They concluded that machine learning is the appropriate course of action when used. The ANN benefiting from milk spectra, including in the procedure, but not PLS, supports ANN as a superior approach for finding helpful information from milk.

Lahart et al. (2019) predicted the DMI of dairy cows using milk FT-MIR combined with faeces NIR. They compared those prediction equations with a benchmark equation based on MY, fat percentage, protein percentage, BW, stage of lactation (SOL), and parity. Milk samples were analysed using a Foss Milkoscan FT6000 spectrometer.

Shadpour et al. (2022) predicted DMI using milk FT-MIR data from three experiments conducted in different countries, employing linear and nonlinear ANN architectures and PLS regression. Weekly milk samples were analysed using a Foss Milkoscan FT6000 spectrometer to predict DMI and (Equation 3) validate the robustness of the developed prediction models.

Ouweltjes et al. (2022) examined the potential of the FT-MIR-predicted FE phenotype for inclusion in the selection index. They used individual feed intake data from feeding trials at the Dairy Campus research farm (Lelystad, the Netherlands) to develop prediction equations. Milk samples were analysed using MilkoScan FT6000 spectrometers. On 676 commercial farms, they predicted the FE phenotype for 78 488 Holstein cows and subsequently calculated EBVs for FE traits. They found low correlations between sires' EBVs for predicted DMI and official DMI estimates. However, higher correlations were observed between sires' EBVs for predicted DMI and official EBVs for milk production (0.33), longevity (0.26), and fertility (–0.27).

Salleh et al. (2023) compared three approaches (PLS, SVM, and RF regression) for their ability to predict DMI in Swedish dairy cattle using milk MIRS data. They found that using the milk MIRS data to predict DMI provided low-to-moderate prediction accuracy $r^2 = 0.07–0.40$. Data were analysed separately for 3–180 days in milk (DIM), early lactation (3–30 DIM), and mid-lactation (30–180 DIM). They preferred PLS over SVM and RF regression. Above all, the mentioned studies demonstrate a progression in the mathematical methods used to develop prediction equations, with increasingly complex mathematical approaches being applied over time (McParland et al. 2014; Dorea et al. 2018; Shadpour et al. 2022). Nevertheless, the original PLS method retains its position as a reliable approach for developing prediction equations (Salleh et al. 2023). In individual studies, data on feed consumption are either obtained through their experiments (Shetty et al. 2017; Ouweltjes et al. 2022), or evaluated from several previously conducted experiments (Dorea et al. 2018; Wallen et al. 2018). One study has progressed from phenotype prediction to breeding value estimation (Ouweltjes et al. 2022).

Development of a predictive equation of FE trait using FT-MIR data

The predicted traits. The primary focus is the prediction of FE traits such as RFI, DMI, NEI, effective energy intake (EEI), and energy balance (EB). Additionally, changes in body condition score (BCS) can also be predicted (McParland et al. 2014). Shetty et al. (2017) utilised DMI and RFI as FE traits. RFI was expressed as the residual from a linear regression model, where DMI was regressed on ECM and weekly MBW. McParland et al. (2014) calculated RFI as the difference between effective energy intake (EEI) based on the equation by Coffey et al. (2001) we repeatedly measured food intake, live weight, milk yield and condition score of Holstein cattle in their first lactation. They were given either a high concentrate or low concentrate diet and were either selected or control animals for genetic merit for kg milk fat plus milk protein. Orthogonal polynomials were used to model each trait over time and random regression techniques allowed curves to vary between animals at both the genetic and the permanent environmental levels. Breeding values for bulls were calculated for each trait for each day of lactation. Estimates of genetic merit for energy balance were calculated from combined breeding values for either (1, and energy demands calculated according to milk, fat, protein and lactose yields, BCS, and EBW and their changes. These calculations were performed for each test-day and categorised into three stages of lactation: 5–60 DIM, 60–180 DIM, and 180–300 DIM. Wallen et al. (2018) developed predictive equations for estimating DMI and net energy intake (NEI) using MY, fat, protein, lactose content, and BW. Rachah et al. (2020) predicted EB, DMI, and EEI in lactating dairy cows × by incorporating MY and concentrate intake, in addition to FT-MIR data, as predictor traits, into their equations. Shadpour et al. (2022) utilised weekly averages of DMI to develop a prediction equation, which was then applied to estimate DMI for animals without direct DMI measurements.

Efforts to accurately predict FE indicators often lead researchers to examine the issue from various angles in their experiments. They analyse several traits derived from data on feed intake and other known indicators related to the animals being studied. One key dilemma they face is whether to predict DMI or RFI or to explore the possibilities of predicting indicators that describe energy consumption. This pursuit also enables them to understand the relationships between different indicators and the changes that occur during lactation.

Preprocessing of FT-MIR data. Mathematical procedures can be applied to adjust raw spectral data before its use in developing prediction equations. This ensures that the adjusted spectral data is thoroughly cleansed of unwanted variance. Shetty et al. (2017) employed a method to enhance the accuracy of weekly averaged FT-IR data by applying Savitzky-Golay first derivatives (Savitzky and Golay 1964). They used a filter width of 7, a polynomial order of 2, and mean centring to achieve this. A similar procedure was applied by Rachah et al. (2020). First-derivative spectra pretreatments are commonly used to enhance resolution and sharpen absorption bands in spectral samples (Savitzky and Golay 1964). Wallen et al. (2018) applied the Savitzky-Golay method to smooth the MIR spectral data using different polynomial degrees and window sizes. Ouweltjes et al. (2022) smoothed 323 selected informative spectral points with a polynomial order of 2 and a filter width of 7, similar to the procedures used by Shetty et al. (2017). Lahart et al. (2019) used rolling averages over segments of 5 data points in length, i.e. boxcar smoothing. First, the spectral data were transformed from transmittance to linear absorbance through log transformation of the reciprocal wavelength values (Soyeurt et al. 2011). Pre-processing FT-MIR

https://doi.org/10.17221/165/2024-CJAS

data helps researchers obtain more reliable results when analysing FT-IR data. In contrast, Shi et al. (2023) did not confirm that preprocessing MIR data would increase prediction accuracy for the traits they studied, including DMI and nitrogen use efficiency (NUE), defined as the ratio of total N output in milk to total N intake from feed. This may have been due to the small sample size of cows (86) they used.

Mathematical methods. The process of deriving prediction equations involves calibration (training) and validation. The calibration output is a prediction equation, the structure of which depends on the model's ability to explain the dependent variable.

Many studies for predicting FE traits using FT-MIR employ PLS regression (McParland et al. 2014; Shetty et al. 2017; Lahart et al. 2019; Rachah et al. 2020; Ouweltjes et al. 2022). In this method, predictors are condensed into a smaller set of uncorrelated orthogonal components called latent variables, and a least squares regression is performed on the latent variables instead of the original data. PLS regression performs well when the number of predictors exceeds the number of observations and when strong collinearity exists between them, as is the case with spectroscopic data. The PLS regression model is currently the most commonly used technique for milk MIR spectra (Grelet et al. 2021).

Dorea et al. (2018) employed artificial neural networks (ANN) and wavelength selection using BN to predict DMI. They found that ANN improved prediction accuracy, particularly when BN was used to select more informative wavelengths. Similarly, Shadpour et al. (2022) employed ANN with positive results. ANN is a complex mathematical procedure that, unlike PLS, is not as straightforward to use, but it produces highly satisfying outcomes. The basic principle of ANN is a self-learning process. Shadpour et al. (2022) used a 2-layer feed-forward perceptron, also known as a single hidden layer feed-forward neural network, to train DMI prediction equations, as shown in Figure 3. This is a commonly applied form of ANN for regression. The ANN training involves an iterative process. The ANN weights and biases for the training step are estimated, and then the prediction errors for the validation set are calculated using the parameters estimated during training. The estimated parameters are used to predict the parameters with the lowest average prediction errors during the iteration process. Levenberg-Marquardt (LM), scaled conjugate gradient (SCG), and Bayesian regularisation (BR) were used as training algorithms in the study by Shadpour et al. (2022).

The BLUP method was used to predict DMI in Norwegian Red dairy cows (Wallen et al. 2018), but the prediction accuracies did not increase substantially. It is clear that using only the content of milk fat, milk protein, and milk lactose, or MIR spectra in the model, does not result in high prediction accuracy for NEI or DMI. In fact, the accuracy of predictions using only the MIR data was consistently lower than that of a model using MY, fat, protein, and lactose concentrations. Based on their results, Wallen et al. (2018) concluded that PLS was more suitable than BLUP for predicting feed intake using the MIR spectra.

In addition to PLS regression, Salleh et al. (2023) investigated the usefulness of support vector machines SVM and random forest regression (RF) for predicting DMI using MIR data. Although these are advanced methods, the coefficient of determination for RF (0.62) and SVM (0.55) was lower than for PLS (0.65). However, Salleh et al. (2023) state that SVM and RF methods are more appropriately applied to predictive classification models than regression.

Thus, several mathematical methods are used to derive equations predicting the FE indicator based on FT-MIR and other predictors, including PLS, machine learning algorithms, and BLUP. The most commonly used mathematical method is PLS, while other methods, such as best linear unbiased prediction (BLUP) (Wallen et al. 2018), are rarely applied. We can conclude that developments in mathematical methods for determining prediction equations have been observed over the past few years, with more complex mathematical approaches being applied over time (McParland et al. 2014; Dorea et al. 2018; Shadpour et al. 2022). However, the original PLS method remains a suitable approach for developing prediction equations (Salleh et al. 2023).

Validation methodology. The calibration (training) step is followed by validation of the prediction model equation using an internal or external data set. In both steps, the similarity between the predicted and measured values of the dependent variable, e.g. DMI, is assessed. The predictive ability of the model refers to its accuracy. The most commonly used statistical quantities for assessing the

https://doi.org/10.17221/165/2024-CJAS

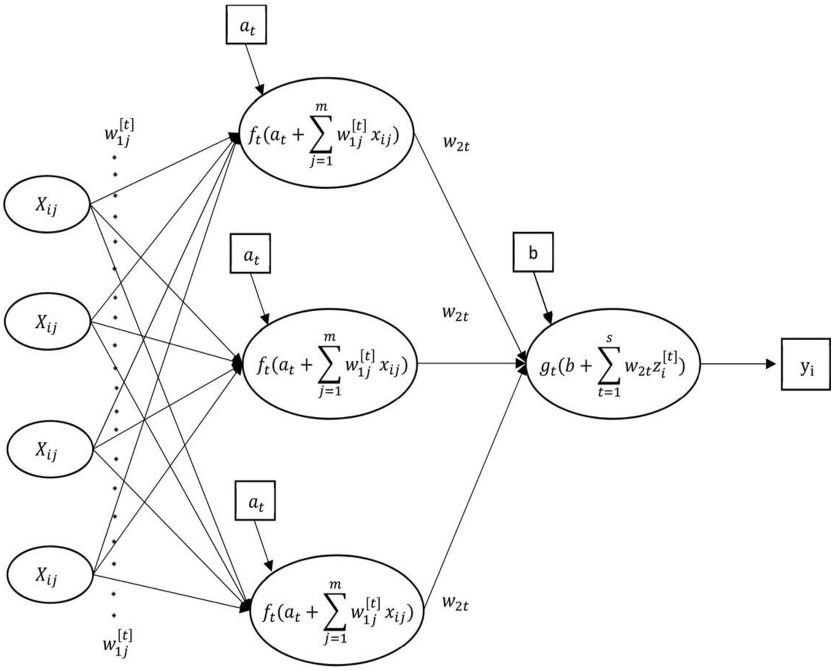

Figure 3. Architecture of a 2-layer feed-forward neural network

y_i = network output (e.g. predicted DMI of individual); where the subscript i refers to the individuals (from 1 to n)

source: Adopted from Shadpour et al. (2022)

predictive ability (reliability or accuracy) of a model are the coefficient of determination (R^2), root mean square error (RMSE), ratio of performance to interquartile distance (RPIQ), and ratio of performance to deviation (RPD). Validating FE trait prediction equations usually involves randomly selecting a proportion of cows or samples from the data set for training and using the remaining data for validation (McParland et al. 2014; Shetty et al. 2017; Wallen et al. 2018; Rachah et al. 2020). Dorea et al. (2018) state that although the potential of the employed methods is demonstrated, the results do not necessarily reflect the predictive accuracy of equations. To increase the credibility of the validation, Lahart et al. (2019) compared validation on a group of cows included in the experiment and a group of cows outside the experiment. They concluded that validation using data obtained outside the actual experiment provides results with greater confidence regarding the valid applicability of the equations. Rachah et al. (2020) created two separate datasets, one for calibration and one for testing, i.e., an external data set. This ensures that the calibration cows were absent from the test set, allowing for correct validation. Cross-validation was based on a split sample (leave-one-cow-out). It must be emphasised that Rachah et al. (2020) used only 63 cows, of which 44 were used for calibration.

Shadpour et al. (2022) employed 10-fold cross-validation in two approaches. The first approach (A) involved deleting 10 percent of the records, and the second approach (B) involved deleting 10 percent of the cows. The testing set was excluded to assess the model's predictive ability for the DMI of Canadian cows only. Therefore, only Canadian cows were included in the validation set. In contrast, the other nine subsets were used as calibration sets to estimate the model parameters. Each of the ten subsets was used once as a testing set during the cross-validation runs. When comparing validation approaches A and B, approach B consistently yielded lower fitting statistics than approach A. This is likely due to inflation caused by including the cow's records in both the calibration and validation sets. Therefore, Shadpour et al. (2022) prefer approach B over A, as it also models the prediction of DMI for new cows, which they likely consider more reliable.

The primary challenge in validating procedures lies in determining whether their results – specifically, high or low accuracy – are consistently reproducible in practice. For instance, Shadpour et al. (2022) aimed to enhance the repeatability of their findings by expanding the validation set to include data from other animals. Similarly, Lahart et al. (2019) chose to use data collected outside of their experiments for validation for the same reason. We can infer that a comprehensive evaluation of data from various experiments could lead to prediction equations that are more widely applicable.

Accuracy of FE trait prediction. The desired outcome of FE prediction is a reliable and accurate estimate of the FE phenotype for animals without measured feed consumption, but for which predictive factors, included in the prediction equation, are available. This accuracy is generally understood as the agreement between the predicted value of the FE trait for an animal and the value that would be found through actual measurement of feed consumption. The published accuracy values achieved in several experiments are shown in Table 2 as the correlation coefficient (r) or R^2. The correlation coefficients between the predicted and actual values of the analysed trait range from 0.21 to 0.73. Accuracy varies between experiments, depending significantly on the mathematical method and model used. From a practical point of view, Ouweltjes et al. (2022) provided the most useful information regarding the application of FT- MIR in FE prediction, with a maximum prediction accuracy of 0.55 (Table 2).

A comparison of published test results shows that FT-MIR alone has low (Shetty et al. 2017) to medium (Dorea et al. 2018; Shadpour et al. 2022) predictive power and must be combined with other predictive factors. As Wallen et al. (2018) point out, prediction equations based solely on MIR values consistently have lower predictive power than those with other predictors. Also, Salleh et al. (2023) found the highest accuracy to be 0.60–0.65 when the prediction of DMI was based on MIR data, supplemented with milk yield and concentrate DMI compared to only 0.19–0.40 when based on MIR alone. Lahart et al. (2019) confirmed earlier findings from Wallen et al. (2018), showing that predictions based on MY, milk composition, and BW were more accurate (r^2 = 0.60) than those based on FT-MIR alone (r^2 = 0.30). However, combining all sources of information, including MY, milk composition, FT-MIR, and NIR, resulted in an r^2 = 0.68.

The question is how much the individual FT-MIR wavenumbers as predictive factors will improve predictive accuracy. Shetty et al. (2017) expressed scepticism about the positive benefits of specific wavenumbers and noted that milk's fat, protein, and lactose content can be derived from milk samples by spectrometry (McDermott et al. 2016); therefore, the spectral data do not provide new information. According to Shetty et al. (2017), the same prediction accuracy can be achieved by using only the standard results from milk recording. In contrast, Dorea et al. (2018) found that using a mathematical method other than PLS – pecifically, ANN with Bayesian networks—and including FT-MIR in the prediction equation improved the prediction accuracy of FE traits over the model containing only milk components (concordance r = 0.80 *vs* 0.72).

Many factors, in addition to the mathematical method used in calibration prediction methods, affect the accuracy of prediction equations for FE traits. In particular, this concerns the agreement between the experimental dataset and the dataset of animals to which the prediction equations are applied. For example, Shi et al. (2023) found that

Table 2. Accuracy of eed efficiency (FE) trait prediction using FT-MIR

Trait	Accuracy (r^2)	Mathematical method	The other predictiv traits	author
EB, EEI, ΔBCS, RFI	0.48–0.78	PLS	MY	McParland et al. (2014)
DMI, RFI	0.29–0.77	PLS	BW, MY	Shetty et al. (2017)
DMI	0.03–0.70	ANN, PLS, BN*	MBW, MY, Mcont	Dorea et al. (2018)
DMI, NEI	0.54–0.65	PLS, NEI	BW, MY,Mcont	Wallen et al. (2018)
DMI	0.16–0.81	PLS, linear regression	BW, MY, Mcont, NIR	Lahart et al. (2019)
DMI, EB, EEI	0.43–0.71*	PLS	MY, CONCTR	Rachah et al. (2020)
DMI	0.47–0.73*	ANN, PLS	MBW, MY, Mcont	Shadpour et al. (2022)
DMI	0.52–0.65	PLS, SVM and RF regression	MY	Salleh et al. (2023)
DMI, rDMI	0.21–0.55*	PLS	–	Ouweltjes et al. (2022)

*Pearson correlation coefficient between observed and predicted DMI values;r^2 = coefficient of determination

ANN = artificial neural networks; BN = Bayesian networks combined with ANN; ΔBCS = daily change in BCS; CONCTR = concentrate intake; DMI = dry matter intake; EB = energy balance; EEI = effective energy intake; Mcont = content of milk fat, protein and lactose; NIR Near-infrared reflectance spectroscopy of faeces; PSL = Partial least squares; rDMI = residual DMI; RFI = residual feed intake changes in feed ratio can reduce the accuracy of the predictive equations, mainly because these changes induce responses in digestibility, milk production, and N intake. McParland and Berry (2016) found that, as expected, FE prediction accuracy is insufficient if the prediction equations are used on a dataset that is completely different from the dataset from which they were derived.

Another perspective is considering the linearity and nonlinearity of the mathematical analysis. Shadpour et al. (2022) concluded that a nonlinear interaction must exist between the weekly average DMI and the predictors used, as nonlinear ANN provided superior predictions compared to linear analysis. This conclusion is supported by the results of Dorea et al. (2018). The validation procedures employed also influence the resulting accuracies. Shadpour et al. (2022) found differences in accuracy, favouring the procedure that deletes records compared to the one that deletes animals. Additionally, Lahart et al. (2019) reported better accuracies for validation using animals within the experiment.

From this, it follows the suggestion that FT-MIR waves should demonstrate good predictive ability for DMI, RFI, or other FE indicators. However, the enhancement of the predictive power of the equation by incorporating FT-MIR waves was affirmed by Dorea et al. (2018), Shadpour et al. (2022), and Salleh et al. (2023), but not by Shetty et al. (2017). A possible explanation is the use of more advanced mathematical procedures by the two authors mentioned above. Both of them utilised ANN. However, this is not the case for Salleh et al. (2023), who achieved better results with PLS than with support vector machine regression (SVM) and random forest regression (RF). Lastly, the effectiveness of the employed mathematical methods depends on additional factors during the derivation of the prediction equations, such as the number of cows involved, the size and length of the experiment, etc.

Breeding applications

The purpose of determining prediction equations for the DMI phenotype is to enable their widespread application in the dairy cattle population. For cows without direct measurements of DMI, the prediction equations are used to estimate DMI values. As discussed earlier, these equations often show a maximum correlation of 0.70 between actual and predicted DMI values. Ouweltjes et al. (2022) investigated the potential use of FT-MIR-predicted DMI values in cattle breeding and reported a maximum correlation of 0.55 between observed and predicted DMI values and 0.34 for rDMI during the validation of prediction equations. Both milk MIR-predicted DMI traits (DMI, rDMI) exhibited genetic variance, with heritability estimates of 0.3 and 0.4 for DMI and rDMI, respectively. These heritability values fall within the range of heritabilities reported for various FE indicators by Jiang et al. (2024). However, the correlations between sire estimates for milk FT-MIR-predicted DMI traits and official breeding values for DMI and rDMI were low (0.14 and 0.03, respectively). Based on those low correlations, Ouweltjes et al. (2022) concluded that the prediction of DMI or rDMI based on FT-MIR is not practical for use. However, they observed that cows consuming more feed tend to produce more milk, live longer, and exhibit poorer fertility, as indicated by the correlations between sire EBVs for milk production (0.33), longevity (0.26), and fertility (–0.27).

According to Wallen et al. (2017), the genomic reference population must increase annually by 4 000 phenotyped and genotyped heifers to improve FE through genomic selection in the target cattle population. The usual approach in selection is that the phenotypes for the directly selected trait are obtained by measuring the offspring of both the dams and sires. However, if phenotyping is more expensive than genotyping, as in the case of measuring feed consumption, then updating the reference population through the genotyping of cows with FE records becomes a viable solution (Wallen et al. 2017). Chesnais et al. (2016) found in real-world data that the accuracy of genomic selection for FE ranged from 0.45 to 0.58. According to Wallen et al. (2017), this level of accuracy could be achieved by incorporating the suggested number of 4 000 cows. Hayes et al. (2009) highlight that increasing the size of the genomic reference population enhances the reliability of selection. Another approach suggested by De Haas et al. (2011), is to use measurements from several countries based on international cooperation.

The use of predicted FT-MIR phenotypes in genomic selection relies primarily on the accuracy of these predicted FE phenotypes. As we have seen, the prediction equations are relatively accurate; the experiment depends on milk production and composition, as well as the cows' body weight. FT-MIR brought improvements in some analyses, but the final prediction accuracy was sometimes as high as in other experiments without FT-MIR. The use of FT-MIRs alone, which is easily achievable, usually does not provide sufficient accuracy. However, one of the primary predictors, BW, is often unavailable. Therefore, the most feasible solution seems to be a combination of genomic selection and various phenotypes, such as the feed-saved trait prediction in Australia (Pryce et al. 2015).

CONCLUSION

Predicting feed efficiency (FE) in dairy cows using FT-MIR data is complex and has intermediate accuracy. Developing robust predictive equations requires extensive experimental validation. FT-MIR data is often combined with other traits that are more crucial for predicting FE. Achieving the same level of prediction accuracy for FE as for milk ingredients is unlikely. Despite this limitation, using FT-MIR to predict FE phenotypes is a promising method due to its simplicity and affordability for the entire population of cows under milk recording. With a sufficient amount of recorded data and an optimised dataset structure in combination with advanced mathematical procedures, the prediction accuracy could be significantly improved, making it practical for breeding of dairy cows.

Conflict of interest

The authors declare no conflict of interest.

REFERENCES

Agnew RE, Yan T, Murphy JJ, Ferris CP, Gordon FJ. Development of maintenance energy requirement and energetic efficiency for lactation from production data of dairy cows. Livest Prod Sci. 2003 Aug;82(2-3):151-62.

Bastin C, Gengler N, Soyeurt H. Phenotypic and genetic variability of production traits and milk fatty acid contents across days in milk for Walloon Holstein first-parity cows. J Dairy Sci. 2011 Aug;94(8):4152-63.

Bittante G, Cecchinato A. Genetic analysis of the Fourier-transform infrared spectra of bovine milk with emphasis on individual wavelengths related to specific chemical bonds. J Dairy Sci. 2013 Sep;96(9):5991-6006.

Cecchinato A, Albera A, Cipolat-Gotet C, Ferragina A, Bittante G. Genetic parameters of cheese yield and curd nutrient recovery or whey loss traits predicted using Fourier-transform infrared spectroscopy of samples collected during milk recording on Holstein, Brown Swiss, and Simmental dairy cows. J Dairy Sci. 2015 Jul;98(7):4914-27.

Coffey MP, Emmans GC, Brotherstone S. Genetic evaluation of dairy bulls for energy balance traits using random regression. Anim Sci. 2001 Apr;73(1):29-40.

Connor EE. Invited review: Improving feed efficiency in dairy production: challenges and possibilities. Animal. 2015;9(3):395-408.

Dal Zotto R, De Marchi M, Cecchinato A, Penasa M, Cassandro M, Carnier P, Gallo L, Bittante G. Reproducibility and repeatability of measures of milk coagulation properties and predictive ability of mid-infrared reflectance spectroscopy. J Dairy Sci. 2008 Oct;91(10):4103-12.

De Haas Y, Windig J J, Calus M P L, Dijkstra J, De Haan M, Bannink A. Genetic parameters for predicted methane production and potential for reducing enteric emissions through genomic selection. J Dairy Sci. 2011 Dec;94 (12):6122-34.

De Marchi M, Fagan CC, O'Donnell CP, Cecchinato A, Dal Zotto R, Cassandro M, Penasa M, Bittante G. Prediction of coagulation properties, titratable acidity, and pH of bovine milk using mid-infrared spectroscopy. J Dairy Sci. 2009 Jan;92(1):423-32.

Dorea JRR, Rosa GJM, Weld KA, Armentano LE. Mining data from milk infrared spectroscopy to improve feed intake predictions in lactating dairy cows. J Dairy Sci. 2018 Jul;101(7):5878-89.

Grelet C, Dardenne P, Soyeurt H, Fernandez JA, Vanlierde A, Stevens F, Gengler N, Dehareng F. Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions. Methods. 2021 Feb;186:97-111

Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. Invited review: Genomic selection in dairy cattle: Progress and challenges. J Dairy Sci. 2009 Feb;92(2):433-43.

Heringstad B, Klemetsdal G, Ruane J. Selection for mastitis resistance in dairy cattle: A review with focus on the situation in the Nordic countries. Livest Prod Sci. 2000 Jun;64(2-3):95-106.

Ho P, Marett L, Wales W, Axford M, Oakes E, Pryce J. Predicting milk fatty acids and energy balance of dairy cows in Australia using milk mid-infrared spectroscopy. Anim Prod Sci. 2019;60(1):164-8.

Chesnais JP, Cooper TA, Wiggans GR, Sargolzaei M, Pryce JE, Miglior F. Using genomics to enhance selection of novel traits in North American dairy cattle. J Dairy Sci. 2016 Mar;99(3):2413-27.

Inon FA, Garrigues S, De La Guardia M. Nutritional parameters of commercially available milk samples by FTIR and chemometric techniques. Anal Chim Acta. 2004 Jun; 513(2):401-12.

Jacobs JA, Siegford JM. Invited review: The impact of automatic milking systems on dairy cow management, behavior, health, and welfare. J Dairy Sci. 2012 May; 95(5):2227-47.

Jiang W, Mooney MH, Shirali M. Unveiling the genetic landscape of feed efficiency in Holstein dairy cows: Insights into heritability, genetic markers, and pathways via meta-analysis. J Anim Sci. 2024;102:1-14.

Klein MS, Buttchereit N, Miemczyk SP, Immervoll AK, Louis C, Wiedemann S, Humpf HU. NMR metabolomic analysis of dairy cows reveals milk glycerophosphocholine to phosphocholine ratio as prognostic biomarker for risk of ketosis. J Proteome Res. 2012 Feb 3;11(2): 1373-81.

Koch RM, Swiger LA, Chambers D, Gregory KE. Efficiency of Feed Use in Beef Cattle. J Anim Sci. 1963 May 1; 22(2):486-94.

Lahart B, McParland S, Kennedy E, Boland T, Condon T, Williams M, McCarthy B. Predicting the dry matter intake of grazing dairy cows using infrared reflectance spectroscopy analysis. J Dairy Sci. 2019 Oct;102(10):8907-18.

Lidauer M.H., Negussie E., Mäntysaari E.A., Mäntysaari P., Kajava S., Kokkonen T., Kallio H. Estimating breeding values for feed efficiency in dairy cattle by regression on expected feed intake. Animal. 2023 Sep;17(9):100917.

Malchiodi F., Koeck A., Mason S., Christen A.M., Kelton D.F., Schenkel F.S., Decker J.E. Genetic parameters for hoof health traits estimated with linear and threshold models using alternative cohorts. J Dairy Sci. 2017 Apr; 100(4):2828-36.

McDermott A., Visentin G., De Marchi M., Berry D., Fenelon M., O'Connor P, McCarthy B. Prediction of individual milk proteins including free amino acids in bovine milk using mid-infrared spectroscopy and their correlations with milk processing characteristics. J Dairy Sci. 2016;99(4):3171-82.

McParland S, Berry D. The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. J Dairy Sci. 2016; 99(5):4056-70.

McParland S, Lewis E, Kennedy E, Moore S, McCarthy B, O'Donovan M, O'Brien B, O'Neill M. Mid-infrared spectrometry of milk as a predictor of energy intake and efficiency in lactating dairy cows. J Dairy Sci. 2014; 97(9):5863-71..

Miglior F, Fleming A, Malchiodi F, Brito LF, Martin P, Baes CF. A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. J Dairy Sci. 2017 Dec;100(12):10251-71.

Negussie E, Mehtio T, Mantysaari P, Lovendahl P, Mantysaari EA, Lidauer MH. Reliability of breeding values for feed intake and feed efficiency traits in dairy cattle: When dry matter intake recordings are sparse under different scenarios. J Dairy Sci. 2019 Aug;102(8):7248-62.

Ouweltjes W, Veerkamp RF, van Burgsteden G, van der Linde R, de Jong G, van Knegsel A, Kooistra S. Correlations of feed intake predicted with milk infrared spectra and breeding values in the Dutch Holstein population. J Dairy Sci. 2022 Jun;105(6):5271-82.

Pryce JE., Gonzalez-Recio O, Nieuwhof G, Wales WJ, Coffey MP, Hayes BJ, Horan B. Hot topic: Definition and implementation of a breeding value for feed efficiency in dairy cows. J Dairy Sci. 2015 Oct;98(10):7340-50.

Pryce JE., Gonzalez-Recio O, Thornhill JB, Marett LC, Wales WJ, Coffey MP, Hayes BJ. Short communication: Validation of genomic breeding value predictions for feed intake and feed efficiency traits. J Dairy Sci. 2014 Jan;97(1):537-42.

Rachah A, Reksen O, Afseth NK., Tafintseva V, Ferneborg S, Martin AD, Heringstad BK. Fourier transform infrared spectroscopy of milk samples as a tool to estimate energy balance, energy and dry matter intake in lactating dairy cows. J Dairy Res. 2020;87(4):436-43.

Rovere G, De Los Campos G, Tempelman RJ, Vazquez AI, Miglior F, Schenkel F, Moore SS. A landscape of the heritability of Fourier-transform infrared spectral wavelengths of milk samples by parity and lactation stage in Holstein cows. J Dairy Sci. 2019 Feb;102(2):1354-63

Rutten MJM, Bovenhuis H, Heck JML, Van Arendonk JAM. Predicting bovine milk protein composition based on Fourier transform infrared spectra. J Dairy Sci. 2011 Nov;94(11):5683-90.

Salleh S, Danielsson R, Kronqvist C. Using machine learning methods to predict dry matter intake from milk mid-infrared spectroscopy data on Swedish dairy cattle. J Dairy Res. 2023 Feb;90(1):5-8.

Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. Anal Chem. 1964 Jul 1;36(8):1627-39.

Shadpour S, Chud T, Hailemariam D, Oliveira H, Plastow G, Stothard P, O'Brien B. Predicting dry matter intake in Canadian Holstein dairy cattle using milk mid-infrared reflectance spectroscopy and other commonly available predictors via artificial neural networks. J Dairy Sci. 2022 Oct;105(10):8257-71.

Shetty N, Lovendahl P, Lund M, Buitenhuis A. Prediction and validation of residual feed intake and dry matter intake in Danish lactating dairy cows using mid-infrared spectroscopy of milk. J Dairy Sci. 2017 Jan;100(1):253-64.

Shi R, Lou WQ, Ducro B, van der Linden A, Mulder HA, Oosting SJ, Li SL, Wang YC. Predicting nitrogen use efficiency, nitrogen loss and dry matter intake of individual dairy cows in late lactation by including mid-infrared spectra of milk samples. J Anim Sci Biotechnol. 2023 Jan 10;14(1).

Soyeurt H, Dehareng F, Gengler N, McParland S, Wall E, Berry DP, O'Donnell M. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. J Dairy Sci. 2011 Apr;94(4):1657-67.

Spurlock DM, Dekkers JCM, Fernando R, Koltes DA, Wolc A. Genetic parameters for energy balance, feed efficiency, and related traits in Holstein cattle. J Dairy Sci. 2012 Sep; 95(9):5393-402.

Tiplady KM, Lopdell TJ, Littlejohn MD, Garrick DJ. The evolving role of Fourier-transform mid-infrared spectroscopy in genetic improvement of dairy cattle. J Animal. Sci Biotechnol 2020;11: 39.

Tiplady KM, Lopdell TJ, Reynolds E, Sherlock RG, Keehan M, Johnson TJJ, Pryce JE, Davis SR, Spelman RJ, Harris BL, Garrick DJ, Littlejohn MD. Sequence-based genome-wide association study of individual milk mid-infrared wavenumbers in mixed-breed dairy cattle. Gen Sel Evol. 2021;53(1):62.

Toledo-Alvarado H, Vazquez AI, de Los Campos G, Tempelman RJ, Bittante G, Cecchinato A. Diagnosing pregnancy status using infrared spectra and milk composition in dairy cows. J Dairy Sci. 2018;101(3):2496-505.

VandeHaar MJ, Armentano LE, Weigel K, Spurlock DM, Tempelman RJ, Veerkamp R. Harnessing the genetics of the modern dairy cow to continue improvements in feed efficiency. J Dairy Sci. 2016 Jun;99(6):4941-54.

Veerkamp RF, Brotherstone S. Genetic correlations between linear type traits, food intake, live weight and condition score in Holstein Friesian dairy cattle. Anim Sci. 1997 Jun;64(3):385-92.

Wall E, Simm G, Moran D. Developing breeding schemes to assist mitigation of greenhouse gas emissions. Animal. 2010;4(3):366-76.

Wallen SE, Lillehammer M, Meuwissen THE. Strategies for implementing genomic selection for feed efficiency in dairy cattle breeding schemes. J Dairy Sci. 2017 Aug; 100(8):6327-36.

Wallen SE, Prestlokken E, Meuwissen THE, McParland S, Berry DP. Milk mid-infrared spectral data as a tool to predict feed intake in lactating Norwegian Red dairy cows. J Dairy Sci. 2018 Jul;101(7):6232-43.

Received: October 11, 2024
Accepted: December 9, 2024
Published online: January 24, 2025