# Prediction of evaluated energy balance (NEL and ME) in dairy cows by milk mid-infrared (MIR) spectra

*L.M. Dale[1], A. Werner[1], H. Spiekers[2], P. Hertel-Böhnke[2], E. Stamer[3], F. Gollé-Leidreiter[1], M. Au[4] and F. Onken[4]*

*[1]Regional association for performance testing in livestock breeding of Baden-Wuerttemberg (LKV - Baden-Wuerttemberg), Stuttgart, Germany*
*Corresponding Author: ldale@lkvbw.de*
*[2]Institute for Animal Nutrition and Feed Management, Bavarian State Research Centre for Agriculture (LfL), Grub, Germany*
*[3]TiDa - Animal and Data GmbH, Westensee, Germany*
*[4]German Association for Quality and Performance Testing e.V., Bonn, Germany*

## Abstract

With the help of milk mid-infrared spectra (MIR) a wealth of information can be obtained by establishing relationships with so-called reference methods. Well-known and established is the determination of the milk ingredients fat, protein, lactose and urea. Additional to being used in herd management, fat and protein content are also used to determine the "milk price". The current research focus is the detection of indirect quantities, such as: pregnancy, mastitis, ruminal acidosis, lameness, energy balance, ketosis or methane emissions. The objective of this study was to build global spectrometric equations for energy balance calculated by the two evaluation systems net energy lactation (NEL) and metabolisable energy (ME), in order to determine the cow energy status. The application may be used in the field of herd management and also as a possible factor in breeding selection.

The present work is part of the collaborative project optiKuh, funded by the German Federal Ministry of Food and Agriculture. 12 research farms from different German states such as Baden Württemberg, Nordrhein-Westfalen, Bayern, Schleswig-Holstein, Rheinland-Pfalz, Niedersachsen and Mecklenburg-Vorpommern provided NEL and ME in between 2014 and 2017. The German Association for Quality and Performance Testing e.V. (DLQ), in which the regional MROs are united, provided approx. 40,000 milk samples available with energy balance information for the calibration equation establishment for the two energy balances. The energy balance models are built on standardised and unstandardized spectral data. To detect the outliers a difference between fat determined by the FOSS or Bentley spectrometers and the RobustMilk equation was calculated and no more than 2% difference was accepted. Out of the complete data set, 30% were considered outliers. To identify animal variables that were positively or negatively associated with cow energy status, the spectral data set was first pre-processed by Savitzky-Golay first derivative to remove the offset differences between samples for baseline correction, before Legendre polynomial modelling. Then the data was submitted to ridge regression using the "glmnet" package in R. For “glmnet” model, the sampling moment, lactation stage and important breeds such as Holstein (HOL), Red Holstein (HOL), Brown Swiss (BSW) and Simmental (SIM) and together with the Legendre polynomial data based on days in milk correction for the 212 OptiMIR selected wavenumbers of spectral data were input variables for modelling.

There were three models performed. The 1st model was based on spectral data random selection, the 2nd on random selection of animals and the 3rd, a global model with a cross-validation trial. The global energy balances, NEL and ME calibration models showed high coefficients of determination ($R^2$ = 0.75, respectively $R^2$ = 0.85%) and very poor RPD 2.00, respectively a poor RPD 2.50. The RPD is the ratio of standard deviation to standard error of cross validation. It has been underlined that a very poor model means that it allows to compare groups of cows, or to distinguish high or low values while the poor model allows rough screening. Standardised spectra showed a better robustness compared to unstandardized spectra (RPD: NEL 2.50 vs. 2.00, ME 3.04 vs. 2.47).

*Keywords: energy balance, NEL, ME, MIR milk spectral data, dairy cow, cow health.*

## Introduction

Mid infrared spectroscopy (MIR) is using the infrared light from the electromagnetic spectrum which shows specific absorption patterns when sent through a milk sample caused by frequency dependent interactions with the chemical bonds of the chemical milk components. With the help of milk MIR spectra a wealth of information can be obtained by establishing relationships with reference methods. This technique is routinely used by milk laboratories and recording organisation (MRO) to determine the concentration of the main milk components fat, protein, lactose and urea. In the last years MIR became an increasingly applied technique that provides different molecular signatures in the dairy cattle industry. Since 2006, Soyeurt *et al.* performed different calibrations models for bovine milk fatty acids (FAs) and their milk MIR spectral predictions are used currently to generate multivariate prediction equations for over 30 FAs, these equations are routinely updated, with accuracies ranging from 68% to 100% (Grelet *et al.* 2014). The current research focus is the detection of indirect quantities, such as: methane emissions (Dehareng *et al.*, 2012), ketone bodies (Grelet *et al.*, 2016), milk protein composition (Rutten *et al.*, 2011), assessing the effect of pregnancy stage (Laine *et al.*, 2017), mastitis status (Dale *et al.*, 2019), lameness detection (Mineur *et al.*, 2017), body energy traits (McParland *et al.*, 2011), etc. The balance between energy intake and sum of energy for maintenance and production in dairy cows could be an important target for cow selection in modern breeding goals and herd management. There are few studies that are using dairy cow energy balance (NEL and ME) traits and MIR spectral data with the aim to calibrate and to generate prediction tools for use in commercial dairy herds. Since 2011, in UK prediction equations for energy balance and intake are applied to the largest spectral datasets (McParland *et al.*, 2011, Smith *et al.*, 2018) and MIR-based energy trait predictions from routinely collected national data has been used in genetic improvement of livestock to obtain sustainable energy profiles. The objective of this study was to build global spectrometric equations for energy balance calculated by the two evaluation systems net energy lactation (NEL) and metabolisable energy (ME), in order to determine the cow energy status. The application may be used in the field of herd management and also as a possible factor in breeding selection.

## Material and methods

The present work is part of the collaborative project optiKuh, funded by the German Federal Ministry of Food and Agriculture, where 12 research farms from different German states such as Baden Württemberg, North Rhine-Westphalia, Bavaria, Schleswig-Holstein, Rhineland-Palatinate, Lower Saxony and Mecklenburg-Western Pomerania provided all information for the calculation of both energy balances (NEL,

ME) in between 2014 and 2017. The energy balances based on NEL were calculated in accordance with GfE (2001), and for the energy balances based on ME, the calculation was based on Susenbeth (2018). The 12 research farms collected approximately 40,000 milk samples from important breeds such as Holstein (HOL), Red Holstein (HOL), Brown Swiss (BSW) and Simmental (SIM) linked to weekly averages of individually recorded energy balance information. Local MROs and associated milk laboratories, organized in the "German Association for Quality and Performance Testing e.V." (DLQ), provided milk recording results and standardised as well as non-standardised MIR spectral data from FOSS and Bentley FTIR analysers for the calibration equation establishment of the two energy balances.

The spectral data are weekly registered at MROs and combined with energy balance data. The spectral absorbance data set was first pre-processed by Savitzky-Golay first derivative to remove the offset differences between samples for baseline correction, before Legendre polynomial transformation based on DIM was applied. To detect the outliers, the difference between the official fat content provided by the laboratories and the fat content derived from the RobustMilk MIR equation was calculated and no more than 4% for standardised spectra and 1% for non-standardised spectra of relative difference was accepted. Out of the complete data set, approximately 30% of spectral data were considered outliers. Then the data was submitted to ridge regression using the “glmnet” package in R.

The input matrix consisted of sampling moment (mixed, morning, evening), lactation stage and breeds serving as fixed effects and the 212 OptiMIR selected wavenumbers, subset of the pre-processed spectral data, were input variables for “glmnet” model. From the statistical point of view it was necessary to perform different validations models to understand better if the calibration model could be applied to different animals or just for the target animals included in calibration. Therefore there were three models performed. The 1st model was based on spectral data, a random selection of 70% of the data in the calibration model and 30% of the data in the validation model, the 2nd model was based on a random selection of animals, 25% of animals from each research farm are part of calibration model and the 3rd model, a global model with a cross-validation trial (Table 1.).

## Results and discussions

Nearly 26,000 energy balances - on NEL basis and nearly 29,000 energy balances on ME basis of the 12 experimental farms have been used for the new equations. The different calibration models that are presented in Table 1, it shows the composition of the models. The 1st model was based on spectral data random selection and around 1,468 animals were part of the calibration model while around 1,145 were in the validation model. For the 2nd model, based on random selection of animals, 25% of

*Table 1. Energy balance (NEL, ME) calibration and validation records.*

| Models | | Calibration-Set Records | | Validation-Set Records | |
|---|---|---|---|---|---|
| | | NEL | ME | NEL | ME |
| Model 1 | Spectra | 24,159 | 21,705 | 8,040 | 7,221 |
| | Animals | 1,468 | 1,411 | 1,243 | 1,145 |
| Model 2 | Spectra | 22,380 | 19,839 | 9,819 | 9,087 |
| | Animals | 1,110 | 1,085 | 396 | 365 |
| Model 3 | | Cross-Validation - NEL | | Cross-Validation - ME | |
| | Spectra | 26,138 | | 28,926 | |
| | Animals | 1,511 | | 1,450 | |

*Table 2. Energy balance (NEL, ME) – Model 3 calibration and validation statistics.*

| Model | Unit | #LV | Mean | SD | SEC | R²c | SECV | R²cv | RPDcv |
|---|---|---|---|---|---|---|---|---|---|
| EB - NEL | [MJ/d] | 12 | 2.47 | 17.29 | 8.27 | 0.75 | 8.27 | 0.75 | 2.001 |
| Standardised | | | | | | | 7.53 | 0.84 | 2.502 |
| NonStandardised | | | | | | | 8.08 | 0.76 | 2.007 |
| EB - ME | [MJ/d] | 12 | 0.08 | 23.54 | 8.99 | 0.85 | 8.94 | 0.85 | 2.580 |
| Standardised | | | | | | | 8.41 | 0.89 | 3.049 |
| NonStandardised | | | | | | | 9.06 | 0.84 | 2.475 |

#LV = number of terms (latent variable)
SEC = standard error of calibration
R²c = calibration coefficient of determination
SECV = standard error of cross-validation
R²cv = cross-validation coefficient of determination
RPD = ratio of SD to SECV. See RPD class

animals from each research farm, it can be pointed out that 1,110 animals were used for calibration and around 396 for validation. Regarding the 3rd model, a global model with a cross-validation trial, all research farms animals (1,511) were used for final model.

It can be emphasized that there is a difference of 0.5 in RPD between standardized and non-standardized devices. But for a better variability and a better robustness of the models, was combined the standardized and non-standardized spectra and this model will be used for validation in commercial farms. It has to be pointed out that for the non-standardized spectral data only a maximum of 1% relative deviation between the official fat content provided by the laboratories and the fat content derived from the RobustMilk MIR equation was accepted. Therefore the equation quality is better and the RPD is higher as 2.

## Conclusions

The global energy balances, calculated by the two evaluation systems NEL and ME calibration models showed high coefficients of determination ($R^2$ = 0.75, respectively $R^2$ = 0.85) and very poor RPD 2.00, respectively a poor RPD 2.50. The RPD is the ratio of standard deviation to standard error of cross validation. It has been underlined before that a very poor model means that it allows to compare groups of cows, or to distinguish high or low values while the poor model allows rough screening. Standardised spectra showed a better robustness compared to unstandardized spectra (RPD: NEL 2.50 vs. 2.00, ME 3.04 vs. 2.47).

## List of references

**Dale, L.M., A. Werner and F. Gollé-Leidreiter**. 2019. "MastiMIR" - A mastitis early warning system based on MIR spectra. ICAR 2019

**Dehareng, F., C. Delfosse, E. Froidmont, H. Soyeurt, C. Martin, N. Gengler, A. Vanlierde and P. Dardenne**. 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows. Animal, 6(10): 1694-1701.

**GfE**. 2001. Empfehlungen zur Energie- und Naehrstoffversorgung der Milchkuehe und Aufzuchtrinder. Heft 8, DLG-Verlag Frankfurt a.M.

**Grelet, C., J.A. Fernández-Pierna, H. Soyeurt, F. Dehareng, N. Gengler, and P. Dardenne**. 2014. Creation of universal MIR calibration by standardization of milk spectra: example of fatty acids. EAAP - 65th Annual Meeting, Copenhagen. p.108.

**Grelet, C., C. Bastin, M. Gelé, J.B. Davière, M. Johan, A. Werner, R. Reding, , J.A. Fernandez-Pierna, F.G. Colinet, P. Dardenne, and N. Gengler**. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. J. Dairy Sci., 99(6): 4816-4825.

**Lainé, A., C. Bastin, C. Grelet, H. Hammami, F.G. Colinet, L.M. Dale, A. Gillon, J. Vandenplas, F. Dehareng, and N. Gengler**. 2017. Assessing the effect of pregnancy stage on milk composition of dairy cows using mid-infrared spectra. J. Dairy Sci. 100:2863-2876.

**Mineur, A., A. Köck, C. Grelet, N. Gengler, C. Egger-Danner and J. Sölkner**. 2017. First results in the use of milk mid-infrared spectra in the detection of lameness in Austrian dairy cows. Agric Conspec Sci, 82(2): 163-166.

**McParland, S., G. Banos, E. Wall, M.P. Coffey, H. Soyeurt, R.F. Veerkamp, and D.P. Berry**. 2011. The use of mid-infrared spectrometry to predict body energy status of Holstein cows. J. Dairy Sci. 94:1222-1239.

**R Core Team.** 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

**Rutten, M.J.M., H. Bovenhuis, K.A. Hettinga, H.J.F. van Valenberg, and J.A.M. van Arendonk**. 2009. Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. J. Dairy Sci. 92:6202-6209.

**Smith, S. L., M. P. Coffey, and E. Wall**. 2018. Association of milk MIR-derived body energy traits with fertility parameters in cows. Proc. 11th WCGALP: 1121.

**Soyeurt, H., P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, and N. Gengler**. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. J Dairy Sci. 89(9): 3690-3695.

**Susenbeth, A**. 2018. Der Energiebedarf von Milchkühen heutiger Rassen. Tagungsband optiKuh, LfL-Schriftenreihe (2/2018): 40- 43