



Energy profiling of dairy cows from routine milk mid-infrared analysis

S. L. Smith,¹ S. J. Denholm,^{1*} M. P. Coffey, and E. Wall¹
Scotland's Rural College (SRUC), Edinburgh EH9 3JG, UK

ABSTRACT

The balance of body energy within and across lactations can have health and fertility consequences for the dairy cow. This study aimed to create a large calibration data set of dairy cow body energy traits across the cow's productive life, with concurrent milk mid-infrared (MIR) spectral data, to generate a prediction tool for use in commercial dairy herds. Detailed phenotypic data from 1,101 Holstein Friesian cows from the Langhill research herd (SRUC, Scotland) were used to generate energy balance (EB) and effective energy intake (EI), both in megajoules per day. Pretreatment of spectral data involved standardization to account for drift over time and machine. Body energy estimates were aligned with their spectral data to generate a prediction of these traits based on milk MIR spectroscopy. After data edits, partial least squares analysis generated prediction equations with a coefficient of determination from split sample 10-fold cross validation of 0.77 and 0.75 for EB and EI, respectively. These prediction equations were applied to national milk MIR spectra on over 11 million animal test dates (January 2013 to December 2016) from 4,453 farms. The predictions generated from these were subject to phenotypic analyses with a fixed regression model highlighting differences between the main dairy breeds in terms of energy traits. Genetic analyses generated heritability estimates for EB and EI ranging from 0.12 to 0.17 and 0.13 to 0.15, respectively. This study shows that MIR-based predictions from routinely collected national data can be used to generate predictions of dairy cow energy turnover profiles for both animal management and genetic improvement of such difficult and expensive-to-record traits.

Key words: dairy cow, energy balance, mid-infrared spectroscopy, genetics

INTRODUCTION

The process of lactogenesis is energetically demanding for dairy cows. If the cow does not consume enough

energy to cover her energy requirements, she has to mobilize her own energy reserves to meet the deficit and can enter a state called negative energy balance (NEB; Coffey et al., 2004). Most cows will experience NEB and depth of NEB in early lactation (Coffey et al., 2004), but the duration of time that a cow spends in NEB can be detrimental to health and fertility (Koenen et al., 2001; Garnsworthy et al., 2008). This is because, in addition to the energy requirements for lactation, maintenance, and mobility, energy intake is also a metabolic and hormonal function that underlies reproductive and immune performance (Reist et al., 2003; Banos and Coffey, 2010). Banos and Coffey (2010) demonstrated a significant negative genetic correlation between days to first observed estrous and energy balance (-0.16 to -0.28). Therefore, the use of particular cow energy traits for inclusion in selection criteria to improve resumption of cycling in subsequent lactations could be advocated. However, calculating energy traits frequently enough to monitor NEB levels and loss of condition, at both the herd and the individual cow level across time, for cows in commercial milking herds has been an inherently difficult thing to do in practice. This is due to the fact that recording feed intake for individual cows is impractical in terms of both time and expense (Banos and Coffey, 2010). Using the traditional approach of subjective BCS is reactive in the sense that, by the time the change in phenotype can be recorded, the damage may already have been done. Prediction equations for energy balance have previously been developed based directly on mid-infrared (MIR) spectral data, a by-product generated from routine milk recording, and using these experimental data to validate calculations (Reist et al., 2002; Friggens et al., 2007; McParland et al., 2011). Monitoring NEB and feed intake on a regular basis using MIR analyses of the milk would have distinct advantages: (1) spectral data are already gathered for predicting milk components, so additional predicted traits have only a small marginal cost; (2) any problems can be identified quickly and mitigation strategies deployed immediately.

Mid-infrared spectroscopy uses infrared light from the electromagnetic spectrum and works because the particular frequency absorbed by a molecule is determined by its shape and chemical bonds. The

Received December 6, 2018.

Accepted July 24, 2019.

*Corresponding author: Scott.Denholm@sruc.ac.uk

technique is used routinely for assessing the fat and protein content of milk as well as lactose, urea, and casein content (De Marchi et al., 2014). It is becoming increasingly apparent that different physiological processes can leave molecular signatures in the milk of a dairy cow that could potentially be detected by this technique. Soyeurt et al. (2011) used a reference set of bovine milk fatty acids (**FA**) and their milk MIR spectra to generate multivariate prediction equations for over 30 FA, which have been routinely updated as part of an EU-funded project, Optimir (www.optimir.eu), with accuracies ranging from 68% to 100% (Grelet et al., 2014). Furthermore, De Marchi et al. (2009) used reference measurements and concurrent milk spectra to develop prediction equations for milk acidity, McParland et al. (2011) to predict body energy traits, Rutten et al. (2011) milk protein composition, Dehareng et al. (2012) methane emissions, and Grelet et al. (2016) ketone bodies.

The objective of this study was to use reference phenotypic data obtained from the dairy research herd at Scotland's Rural College (**SRUC**) and combine it with concurrent MIR spectral data to generate prediction equations for energy balance (**EB**) and energy intake (**EI**) such that predictions of these traits can be generated for cows in the national herd as part of routine milk recording and genetic analyses performed on MIR-derived predictions.

MATERIALS AND METHODS

Sample Data and Generation of Energy Phenotypes

Phenotypic data were collected from the Langhill lines of Holstein Friesian (**HOL**) dairy cows ($n = 1,101$) housed at SRUC Dairy Research Centre at Crichton Royal Farm in Dumfries, Scotland, and recorded between 2005 and 2018.

Dairy cattle in this herd have been subject to a long-term 2×2 factorial experiment for divergent feeding systems (high- and low-concentrate) and genetic lines ("Control" cows selected to be in line with the UK average genetic merit for kg of fat and protein, and "Select" cows bred to maximize these traits; Roberts and March, 2013). At various stages over the course of the experiment, the components of the diets may change to account for changes in management system, yield levels, and, in the instance of one of these diet changes (home-grown vs. by-products), main ingredient types. However, these experiments have always run for 7 or more years, so are long-term in nature as well. Further, even if ingredient types have changed, we

still aim to differentiate the diets by energy content. A summary of the diets is presented in the Supplemental Information (<https://doi.org/10.3168/jds.2018-16112>). Langhill cows are extensively monitored for production, health, and fertility, and data relevant to the present study were extracted for analysis. Traits of interest included daily milk yield (**MY**, kg; the sum of morning, midday, and evening milkings), daily BW (kg; the average of 3 daily measurements), weekly fat (**FC**) and protein (**PROT**) contents (%), weekly BCS (scored 0 to 5, inclusive, in increments of 0.25, with higher scores indicating fatter animals; Lowman et al., 1976), and daily DMI (kg) for 3 consecutive days a week. Feed composition data included CP (g/kg DM), OM (g/kg DM), organic matter digestibility (**OMD**, g/kg DM), and ME (MJ/kg DM). Records with DIM between 5 and 305 [i.e., with week in milk (**WIM**) 1 to 44] were retained for analysis.

Not all observations were collected or available at a daily level; thus, daily estimates for MY, FC, PROT, DMI, BW, and BCS were generated using the random regression model in Equation [1]. This ensured a complete longitudinal data set for each cow, lactation, and DIM (Banos and Coffey, 2010; McParland et al., 2011; Banos et al., 2012):

$$y_{itk} = \mu + F_{itk} + \left[\sum_{n=0}^4 \beta_n P_n(t) \right]_k + \left[\sum_{n=0}^4 \lambda_{in} P_n(t) \right]_k + \varepsilon_{itk}, [1]$$

where y_{itk} is the trait observation of cow i on DIM t ($t = 5, \dots, 305$) of lactation k ($k = 1, \dots, 5$); μ is the overall mean; F_{itk} are the fixed effects of diet group (diet rations, both high- and low-concentrate, were modified 2 times between 2005 and 2018 and thus fitted as 6 separate diets, or 6 classes); genetic group (2 classes, i.e., Control or Select line cows); calving age (3 classes: early-, mid-, and late-calving cows); year-by-season of calving interaction (2 calving seasons per year, i.e., 28 classes); year-by-month of record interaction (168 classes); $\sum_n \beta_n P_n(t)$ represents the fixed regression of the overall lactation curve (β_n are fixed regression coefficients) described by fourth-order orthogonal polynomials (P_n) on DIM (t); $\sum_n \lambda_{in} P_n(t)$ represents the random regression of the individual animal effect (λ_{in} are random regression coefficients) described by fourth-order orthogonal polynomials (P_n) on DIM (t) and describes the deviation from the overall lactation curve of the i th individual cow ($i = 1, \dots, 1,101$); and ε_{itk} is the random residual effect. All effects were fitted within each individual lactation k . All statistical analyses were conducted using the software package ASReml, ver-

sion 3 (Gilmour et al., 2009). This generated a total complete longitudinal data set of 691,232 daily animal test-date records.

The body energy traits of interest in this study, as highlighted above, were effective EI (MJ/d), constituting the daily energy acquired through feeding and calculated using OM intake, CP, and ME; and EB (MJ/d), representing the difference between the amount of energy taken in minus that spent or contributed as a function of MY, FC, PROT, DMI, BW, and BCS. Formulae for the calculation of these energy traits are summarized in Equations [2] and [3]:

$$EI = 0.001 \times (DMI \times OM) \times EEC \quad [2]$$

$$EB = EI - EE_{out}, \quad [3]$$

where $DMI \times OM$ is a cow's daily intake of OM (kg); EEC is effective energy content of feed per kilogram of OM; and EE_{out} is the energy used by a cow for milk production, maintenance, and general activity. Full details of these formulae and their underlying components can be found in Emmans (1994) and Banos and Coffey (2010). These body energy traits constitute the reference phenotypic values used to calibrate concurrent spectral data described subsequently.

Mid-Infrared Data

Milk Sampling and Analysis. Mid-infrared analysis of milk samples obtained from dairy cows at SRUC's dairy research herd was conducted by National Milk Records and tested on their suite of Foss MilkoScan machines (Foss Analytics, Hillerød, Denmark), based at National Milk Laboratories (Hillington and Wolverhampton), for routine milk recording services. For the purposes of this study, only the MIR analysis of evening samples from Langhill recording were used, following previous analyses carried out on the same herd that highlighted higher accuracies of prediction when using evening compared with morning and midday milk samples (McParland et al., 2011). The Langhill herd at Crichton Royal Farm is milk recorded weekly; in contrast, the UK milk recording system takes place on a monthly basis.

Pretreatment and Standardization of Mid-Infrared Data. Following MIR analysis of the milk sample, a spectrum of 1,060 data points is generated; each point represents the absorption of mid-infrared light by the milk sample at a particular wavelength within the range of 900 cm^{-1} to $5,000\text{ cm}^{-1}$ (Lainé et al., 2017). Before use in generating prediction equations,

the spectral data were subject to several pretreatments. First, the transmittance data were converted to a linear absorbance scale using the transformation $\log_{10}^{-0.5}$. Second, spectral data were standardized against a master spectrometer (Grelet et al., 2015), to account for drift incurred by collection of spectral data from different MIR instruments and across time (Grelet et al., 2014), ensuring that resultant prediction tools developed using such spectral data can be applied to other machines within the same standardization network over time. Finally, a first-derivative calculation was applied to the data by subtracting from each wavepoint value the immediately preceding wavepoint value, similar to the process described in McParland et al. (2011).

Alignment of Spectra with Trait to Create Reference Data Set. Mid-infrared spectral data from the Langhill research herd were collected from early 2012 onward. Daily modeled energy traits were then aligned with pretreated spectral data based on animal ID and date of sampling (within ± 1 d of milk sampling), such that only those spectra with a corresponding phenotypic record for both energy traits were retained. The final reference (calibration) data set consisted of over 30,594 animal test-date records of body energy trait measures, combined with concurrent pretreated MIR spectra, for lactations 1 to 5.

Data Preprocessing. Biological edits were applied to the base phenotypes used in the calculation of energy traits (MY, FC, PROT, DMI, BW, and BCS), retaining only those records with values within 3 standard deviations (SD) of their respective annual means. Regarding the spectra, and as a form of quality assurance, the Mahalanobis distance was calculated for each composite of spectral points, and only those with a distance less than 20 were retained (Grelet et al., 2016).

Further preprocessing of the calibration data (via data thinning) was required to ensure a balanced representation between low and high reference energy values (Grelet et al., 2016). Thinning was carried out following Grelet et al. (2016), such that data were divided across the densest parts of their distribution into equal-sized windows with a proportion (%) of total records selected randomly from each window. Defining the densest area on the bell-shaped distribution within which to apply thinning was a balance between representing as much of the range as possible and not restricting each interval to too few values. Thinning was found to be optimal when selecting 50% of records from each of 10 (15-MJ/d) intervals across the distributions of EB and EI within the ranges $-75 \leq x \leq 75$ and $75 \leq x \leq 225$, respectively. This resulted in a reduced data set that spanned the same range of energy values as the full

Table 1. Summary statistics of the energy balance (EB) and energy intake (EI) calibration data sets

Trait ¹	EB calibration data set n = 11,526				EI calibration data set n = 11,941			
	Mean	SD	Maximum	Minimum	Mean	SD	Maximum	Minimum
MY (kg/d)	29.9	8.2	59.2	4.0	29.7	7.9	56.6	3.0
FC (%)	3.7	0.6	6.0	1.7	3.8	0.6	6.0	1.7
PROT (%)	3.2	0.3	4.7	2.4	3.2	0.3	4.7	2.4
DMI (kg/d)	17.3	5.2	32.5	1.6	17.6	5.0	30.4	4.2
BW (kg)	592.5	67.0	919.8	394.7	597.8	67.3	919.8	401.6
BCS (0–5)	2.0	0.3	4.0	0.9	2.1	0.4	3.8	0.9
CP (g/kg of DM)	159.7	29.9	209.0	53.0	158.9	30.3	209.0	53.0
OM (g/kg of DM)	924.8	9.7	992.0	901.0	925.0	9.5	992.0	901.0
OMD (g/kg of DM)	753.4	66.3	865.0	581.0	755.8	65.0	865.0	581.0
ME (MJ/kg of DM)	11.5	0.8	12.9	9.3	11.5	0.8	12.9	9.3
EB (MJ/d)	1.57	37.78	101.27	-130.54	—	—	—	—
EI (MJ/d)	—	—	—	—	169.55	44.91	285.54	41.95

¹MY = milk yield; FC = milk fat content; PROT = milk protein; OMD = organic matter digestibility.

data set but with an increased balance between low and high values. Window size, frequency, and proportion of records retained were all varied to optimize the accuracy of the calibration data set.

Finally, records were considered outliers and removed from each of the prediction data sets if their predicted residual sum of squares (**PRESS**), obtained from the analysis of all reference data, exceeded 1 SD of the mean PRESS value. This process highlighted data from 2 sample dates in particular, where the majority of the predictions (79 and 80%, respectively) generated from spectra on these dates were greater than 3 SD of the mean PRESS; thus, these data were removed. This generated final calibration data sets for EB (n = 11,526) and EI (n = 11,941). A summary of the data is presented in Table 1.

Calibration, Validation, and Tool Development

Partial least squares (**PLS**) regression was used to calibrate the spectral data, using modeled energy phenotypes to generate prediction equations for EB and EI, respectively, based on a subset of the most informative wavepoints; namely 212 specific wavepoints in the range 968 to 2,966 cm⁻¹ (Lainé et al., 2017). Calibration of the spectral data with reference phenotypes was conducted by PLS using scripts written in Python version 3.5 (van Rossum, 1995) and Microsoft SQL (Microsoft Corp., Redmond, WA), using the NIPALS method. The Python script was created to assess the optimal number of latent factors to use in the PLS model, up to a maximum of 20; that is, the number of factors (or variables) that maximize the variation explained while minimizing the prediction error.

Accuracy was measured using the coefficient of determination (R^2), generated from the regression of

predicted values on actual values by internal split-sample 10-fold cross-validation and external 4-way cross-validation; whereby the data set was split randomly into quarters and the PLS run 4 times, each time one of the quarters serving as the validation file and the remaining quarters as the calibration file. To investigate the effects of having repeated observations per cow, a second external 4-way cross-validation was carried out masking cows (and their corresponding data) instead of records. Accuracy of prediction was assessed using the R^2 from validation methods, in addition to the root mean square error and the ratio of SD of the reference values to the standard error of prediction (**RPD**; a model is deemed to not predict the reference values well if the RPD value is less than or equal to 1.5; Nicolaï et al., 2007) from the split-sample validation only. The optimal calibration data set generated for each of the 2 energy traits with their concurrent spectra generated linear prediction equations, which were then applied to a data set of national spectra.

Application of Tool to National Data: Phenotypic and Genetic Analysis

Since 2013, National Milk Records has routinely collected spectral data from over 360 commercial farms in the UK as part of an Innovate UK-funded project. From 2016 National Milk Records has retained spectral data generated by all relevant Foss machines; these data encompass MIR-analyzed milk samples from 4,822 herds, 1,484,377 animals, and over 20 million test dates. Since the start of 2018, it has been routine for SRUC to receive almost 15 million spectral data points per day. These data points are fed through a bespoke software platform, which ensures these data are processed (transformed, standardized, first deriva-

tive calculations, according to protocols) and subject to prediction equations from the calibration data sets discussed above, which generate a series of MIR-derived phenotypes, such as the energy traits described here. Subsequently, predictions for each of the 2 traits were scaled according to the mean and SD of the reference data set from which they are derived (Langhill), to account for the fact that data from this research herd represents such a broad range of phenotypes due to the nature of the 2×2 long-term factorial experiment performed at SRUC on diet and genetic line. To meet the objective of demonstrating the current tool's application and performing genetic analyses, MIR-derived predictions were obtained for 11,717,032 animal test-dates from almost a 4-year period (2013 to 2016).

Phenotypic Analyses. A fixed regression model was used to generate fitted curves for each predicted energy trait for each of the 5 main dairy breeds represented in the national data. These analyses focused on records from HOL, Shorthorn (**SHO**), Ayrshire (**AYR**), Jersey (**JER**), and Guernsey (**GUE**) cows. To ensure robustness of phenotypes (no attempt was made here to partition the animal genetic component from the permanent environment at this stage), the following edits were applied to the MIR-derived predictions before analysis: records were retained from DIM between 5 and 305, from cows whose age at first calving was between 16 and 48 mo, and at least 5 records per animal per lactation from the first 5 lactations were used. Further, energy predictions within ± 3 SD of the overall mean were retained. The final data set totaled 6,888,665 test-date records on 626,237 animals from 4,198 herds. Fitted curves using predicted phenotypes were estimated using the following fixed regression model for each breed individually:

$$y_{iot} = \mu + F_{iot} + \sum_{n=0}^3 \beta_n P_n(t) + \sum_{n=0}^3 \beta_n P_n(t) \times E_o + a_i + \varepsilon_{iot}, \quad [4]$$

where y_{iot} is the predicted trait of cow i belonging to breed o at WIM t ($t = 1, \dots, 44$); μ is the overall mean; F_{iot} are fixed effects of month and year of phenotypic record, lactation number, age at first calving (month), herd-year-season of calving, and breed (HOL, SHO, AYR, JER, or GUE); $\sum_n \beta_n P_n(t)$ represents the fixed regression of the overall lactation curve (β_n are fixed regression coefficients) described by third-order Legendre polynomials (P_n) on WIM (t); $\sum_n \beta_n P_n(t) \times E_o$ represents the interaction between the fixed regression of the overall lactation curve and breed and describes the deviation from the lactation curve of the o th individual breed ($o = 1, \dots, 5$); a_i is the random additive

genetic effect of the i th individual cow; and ε_{iot} is the random residual effect. All analyses were carried out using ASReml version 3 (Gilmour et al., 2009).

Genetic Analysis. In addition to the data edits highlighted above, HOL records obtained from the national data were further preprocessed by applying the following constraints, such that all cows used for the purposes of the present study (1) were from sires who had sired at least 10 daughters, (2) were the first 250 daughters from each sire, and (3) had at least 8 records per animal per lactation from the first 4 lactations only (defined as lactations 1, 2, and 3 and above). These additional constraints were applied to help represent the population, ensuring that well-proven bulls did not dominate the data set and removing any potential selection bias between young and proven bulls. We did not have as much latitude, in terms of record numbers or overdominance of bulls in the other breeds, to require such additional edits above those applied to all breeds (i.e., number of test records per cow, minimum number of records in a contemporary group, minimum and maximum trait values). These data were linked to a pedigree consisting of over 422,000 individuals. The final data set totaled 1,079,160 test-date records on 93,735 animals from 2,977 herds. Variance components were estimated for each of the 5 breeds via the univariate repeated measures model in Equation [5], with respective pedigree relationship matrices fitted to account for the genetic relationships between individuals within breed:

$$y_{it} = \mu + F_{it} + \sum_{n=0}^5 \beta_n P_n(t) + a_i + p_i + \varepsilon_{it}, \quad [5]$$

where y_{it} is the predicted trait at time t WIM ($t = 1, \dots, 44$); μ is the overall mean; F_{it} are the fixed effects of month and year of phenotypic record, lactation number, age at first calving (month), and herd-year-season of calving; $\sum_n \beta_n P_n(t)$ represents the fixed regression of the overall lactation curve (β_n are fixed regression coefficients) described by fifth-order Legendre polynomials (P_n) on WIM (t); a_i and p_i are the random additive genetic and permanent environmental effects of the i th individual cow, respectively; and ε_{it} is the random residual effect. All analyses were carried out using ASReml version 3 (Gilmour et al., 2009).

RESULTS AND DISCUSSION

Calibration, Validation, and Tool Development

Regarding the final data set (reduced, with outliers removed) the R^2 obtained from internal split-sample

Table 2. Calibration statistics for energy balance (EB) and energy intake (EI) from 3 independent data sets (full, reduced, reduced with outliers removed)¹

Trait	N _r	N _f	R ^{2 2}	R ^{2 3}	R ^{2 4}	RMSE ²	RPD ³
Full data set							
EB	30,429	20	0.42	0.31	0.30	32.8	1.32
EI	30,565	20	0.39	0.24	0.24	38.48	1.28
Reduced data set ⁵							
EB	16,145	20	0.48	0.36	0.36	33.84	1.39
EI	17,219	20	0.45	0.29	0.29	40.31	1.35
Reduced data set ⁵ with outliers removed ⁶							
EB	11,525	20	0.77	0.60	0.60	17.96	2.1
EI	11,940	20	0.75	0.53	0.52	22.35	2.01

¹N_r = number of records in the data set; N_f = number of factors used in prediction equation; RMSE = root mean square error; RPD = ratio of SD of the reference values to the SE of prediction.

²Obtained using internal split-sample 10-fold cross-validation.

³Obtained using external 4-way cross-validation (on records).

⁴Obtained using external 4-way cross-validation (on cows).

⁵Covering the same range of values as the full data set with a balanced representation between low and high reference values.

⁶Outliers were values with a predicted residual sum of squares (PRESS) exceeding ± 1 SD of the mean PRESS value.

10-fold cross-validation were 0.77 and 0.75 for EB and EI, respectively. Further, from the external 4-way cross-validation, the R² were 0.60 and 0.53 for EB and EI, respectively (with external 4-way cross-validation = 0.60 and 0.52 for EB and EI, respectively). The optimal numbers of latent factors selected in these analyses were 20 for each trait (the maximum set). These values, along with additional information regarding the calibration and validation, are presented in Table 2.

The RPD from split-sample validation for EB and EI was greater than 2, suggesting value in these prediction models at discriminating low from high values (Nicolaï et al., 2007). Based on the internal split-sample 10-fold cross-validation, our results for EB are in the bracket of those calculated in McParland et al. (2011, 2012) but slightly improved on the bracket for EI (0.45 to 0.60 in McParland et al., 2011, 2012). This is perhaps unsurprising, given common use of some of the Langhill research herd data between both studies.

In this study we have applied a MIR-based prediction for energy traits from HOL cows. In reality, the prediction data set uses divergent lines of HOL for milk production traits interacting with 2 contrasting planes of nutrition (2 × 2), with the nutritional planes altering during the data window (6 diet rations in total). Although this does not reflect the other breeds, it is, by design, meant to cover a wide range of production potentials and feeding systems, and, as such, we expect divergent ranges of energy profiles within and between animals and experimental units. In the development of the model, the number of factors was tested until the minimum PRESS was attained, to choose the optimal number of PLS factors to diminish the possibility of

overfitting (Haaland and Thomas, 1988). Although a relatively high number of factors were used, the cross-validation error was not excessively high compared with a lower number of factors. Furthermore, it should be noted that the R² could be, in part, influenced by the fact that EI and EB traits were generated from random regression modeling on the component trait elements (i.e., MY, DMI, BW, BCS) from within the research herd. In the development of a MIR spectral predictive tool, we explored the difference in R² for MIR predictions of repeated records of BCS, a component trait of EB. When comparing MIR predictions obtained using actual BCS data during calibration with predictions obtained using BCS solutions from random regression modeling, the influence was minimal. This suggested that R² generated from the compiled EB and EI traits may not necessarily be overinflated. Ideally the cross-validation would be performed on an independent data set, but unfortunately we did not have access to an external set of data to test MIR predictions with our 4-fold validations. The wavepoints studied here are those more informative for the major FA in milk, and changes in the fat profile have shown to be an indicator of EB profiles (de Vries and Veerkamp, 2000; Løvendahl et al., 2010).

As we continue to assimilate spectra (now over 20 million animal test-dates worth of spectra and MIR-derived phenotypes) and collect and improve the accuracy of energy phenotypes (e.g., the potential use of calorimetric chambers), we should be able to improve system-specific equations and, by incorporating data from across systems, countries, and climates, improve the development of multicountry, multisystem tools (McParland et al., 2012).

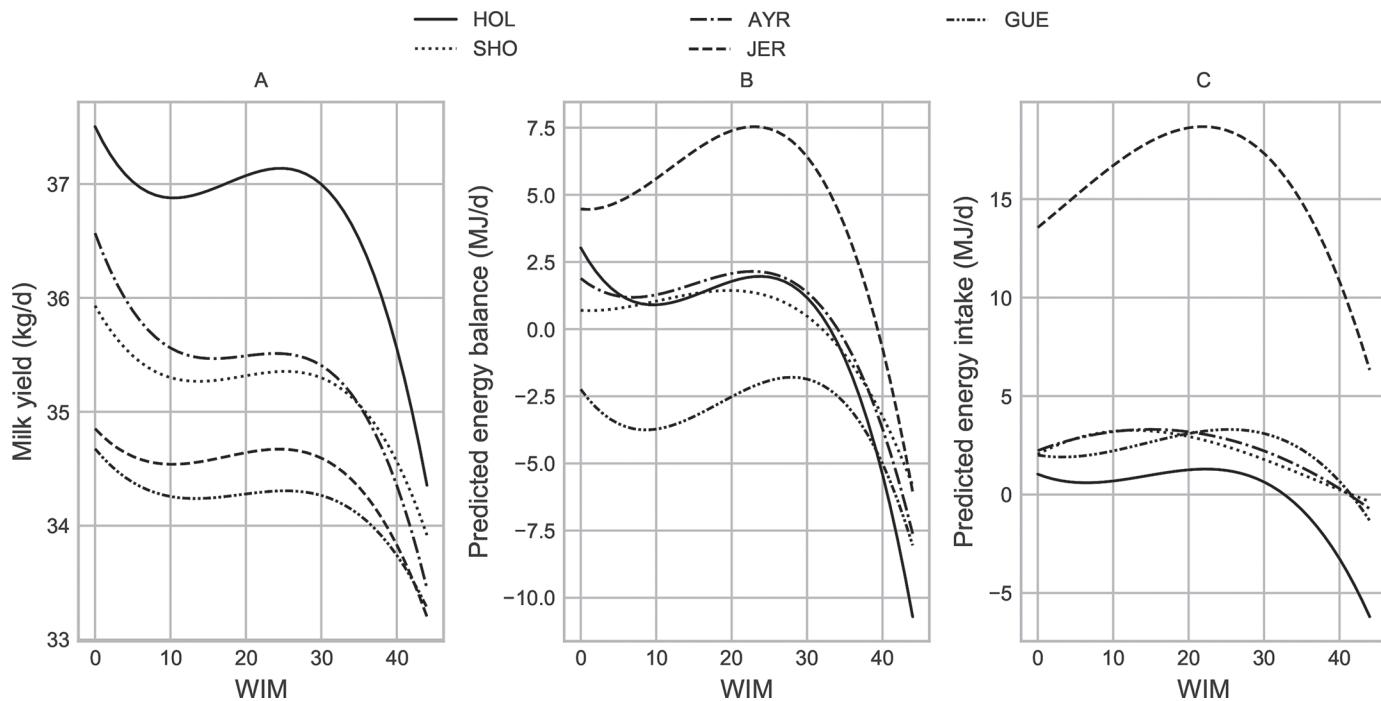


Figure 1. Predicted milk yield and energy traits from a fixed regression model, plotted by week in milk (WIM) for (A) milk yield, (B) energy balance, and (C) energy intake for first-lactation Holstein Friesian (HOL), Shorthorn (SHO), Ayrshire (AYR), Jersey (JER), and Guernsey (GUE) cows.

Application of Tool to Utilize National Data: Phenotypic and Genetic Analysis

Application of the calibration data set to the national spectra generated 11,529,162 and 11,501,992 animal test-dates worth of predictions for EB and EI, respectively. Records were lost (1.6% and 1.84% for EB and EI, respectively) due to quality assurance, applied to ensure that predictions considered outside the ability of these tools to predict usefully were not reported.

Phenotypic Analysis. Figure 1 depicts the trend of EB, EI, and MY across WIM. The profiles of EB and EI generated from fixed phenotypic regression curves are an attempt to model the balance between energy acquired from both feed and body fat mobilization with the energy demands of lactation, growth, and maintenance, with all other fixed effects considered.

Across all breeds, the trend in EB showed an initial recovery from a loss of energy in early lactation, reaching a peak toward the end of mid-lactation and subsequent decline before second calving. The difference in EI profiles between breeds was greatest between JER and HOL cows, suggesting that, at 1 WIM, JER were consuming 12.5 MJ/wk more than HOL. Moreover, at the greatest disparity (22 WIM), JER were consuming 17.4 MJ/wk more than HOL. Overall, JER had the most positive EB profile and the highest EI profile, de-

spite producing the second-to-lowest average MY (34.4 kg/wk). This is in contrast to studies that have suggested JER cows to be more energy efficient than HOL (Grainger and Goddard, 2004; Kristensen et al., 2015). It has been shown, however, that JER cows eat more per unit of BW than HOL and, hence, have a higher intake capacity, likely due to their relatively larger digestive tract (Grainger and Goddard, 2004; Prendiville et al., 2009). Moreover, JER cows tend to produce milk with higher FC, which is energy dense, highlighting the fact that methods of measuring efficiency can be different. In contrast to JER, GUE cows showed the most negative EB profile across first lactation and the second-lowest EI profile in early lactation, consistent with their delivery of the lowest MY among all breeds (34.2 kg/wk).

At the onset of lactation, EB levels in JER cows were 6.7 MJ/wk higher than in GUE cows. Jersey and SHO animals tended to experience their nadir in EB at the start of lactation (1 WIM), much earlier than AYR (7 WIM), GUE (9 WIM), and HOL cows (10 WIM). The initial rate of EB loss in early lactation was fastest for HOL (-0.19 MJ/wk) and slowest for SHO cows (-0.004 MJ/wk). Energy intake was observed increasing from the start of lactation to a maximum at wk 13 (SHO), 15 (AYR), 22 (HOL and JER), and 25 (GUE). Despite the same duration to maximum EI

of 22 wk observed in JER and HOL, JER showed the fastest rate of gain from start of lactation (0.22 MJ/wk) and HOL the lowest (0.01 MJ/wk). Despite the fact that HOL cows displayed the fastest decline of EB to nadir in early lactation and had the lowest rate of gain in EI to its maximum, they produced on average the greatest amount of milk of all 5 breeds (36.7 kg/wk). Moreover, HOL cows reached their peak EI in the same amount of time as JER cows reached their peak MY (22 WIM). The duration between peak MY and peak EB was second-lowest for HOL at 24 wk; strong selection for yield may have predisposed this breed to more readily utilize body energy through mobilization of their fat reserves to support lactation (Coffey et al., 2003), supporting the lower EI profile for HOL shown here despite highest MY.

Recovery from the initial nadir of EB was fastest in JER animals (0.13 MJ/wk), followed by HOL (0.07 MJ/wk), and lowest for SHO (0.04 MJ/wk). When looking at both energy trends across lactation, it is apparent that the peak in EI occurred 1 wk before a peak in EB for JER animals, 2 wk before for HOL, 3 wk before for GUE, 7 wk before for SHO, and 8 wk before for AYR. Furthermore, peak MY occurred 1 wk before the nadir of EB in early lactation for SHO and JER, 7 wk before in AYR, 9 wk before in GUE, and 10 wk before in HOL. The most positive EB across all breeds was achieved in mid-lactation. Maximum EB levels were observed at wk 20 for SHO, 23 for AYR and JER, 24 for HOL, and 28 for GUE. Shorthorn and AYR cows showed similar energy profiles and had similar average MY (35.2 kg/wk and 35.3 kg/wk for SHO and AYR, respectively). Cows belonging to these 2 breeds tended to be larger and heavier than their counterparts from the other breeds, and as such it may be that a greater proportion of their available energy goes into processes such as growth, at the expense of production. Such is consistent with the negative relationship between body size and efficiency (Dickinson et al., 1969).

Following the peaks observed in mid-lactation, the rate of EI decline leading to subsequent calving was fastest for GUE (-0.62 MJ/wk), followed by HOL (-0.6 MJ/wk), and slowest for SHO (-0.29 MJ/wk). This tail-off toward the end of the milk and energy distributions for all breeds may reflect the influence of fewer records available toward the end of lactation in the random regression rather than biological differences in the rate of decline between breeds. Alternatively, it may reinforce the idea that EB represents a different biological trait at different stages of lactation with differing efficiencies (Emmans, 1994; Løvendahl et al., 2010), such that prediction tools may need to be lactation-stage specific. With the reference data and

concurrent spectra available from the Langhill herd, developing and testing such tools is highly plausible. It is important to bear in mind, however, that the tools established in the present study, and any future tools developed using the Langhill research herd, are based on one breed only (HOL), and, as such, validation data from other breeds would be necessary to assess the accuracy of tools applied across breeds before any deployment in the field.

Genetic Analysis. Estimates of additive genetic and permanent environmental variance, heritability, and repeatability of EB and EI for the different breeds are reported in Table 3. The heritability of EB for all breeds in the present study is within the range of other estimates from the literature (Berry et al., 2007, $h^2 = 0.06$ to 0.29; Hüttmann et al., 2009, $h^2 = 0.04$ to 0.34; Banos and Coffey, 2010, $h^2 = 0.10$ to 0.34). However, unlike previous studies, the estimates here are based on MIR-derived predictions over all lactation, which have tended to be lower (McParland and Berry, 2016). Moreover, our h^2 estimates for both EB and EI were similar to previous estimates obtained from MIR-derived predictions by McParland et al. (2015), who estimated the h^2 of EB and EI as 0.1 and 0.2, respectively. Further, our h^2 estimates of EB and EI are still greater than the estimates of some functional traits incorporated in national and international selection criteria and, therefore, may be more effective for use in selection criteria due to their significant genetic correlation with such traits (Banos and Coffey, 2010). For example, EB has previously been associated with days to first estrus ($r = -0.16$ to -0.28) and days to conception ($r = -0.28$ to -0.64 ; Banos and Coffey, 2010).

Estimated additive genetic variation for these traits, across all breeds, highlights that energy partitioning is under at least some genetic control and, therefore, can facilitate the selection of animals with appropriate energy profiles that could contribute to the genetic improvement of fertility (Banos and Coffey, 2010), could be combined with measures such as SCC to improve health (Banos et al., 2006), or could be predicted alongside important blood metabolites to optimize identification of ketosis (Grelet et al., 2016). It has been shown the heritability of these traits (or their components) can vary across lactation (Coffey et al., 2001); as such, going forward the data should be modeled using test-day or random regression models and accounting for fewer records and error variance in early and late lactation when cows are drying off.

Lastly, optimization via genetic selection can only go so far; management of appropriate nutrition and levels of concentrate (see Coffey et al., 2004, for greater effects of low-concentrate diet on body energy loss) and

Table 3. Genetic parameters of energy balance (EB) and energy intake (EI) of the 5 breeds, with corresponding SE in parentheses¹

Breed	Cows	Herds	Trait	n	σ_a^2	σ_{pe}^2	\hbar^2	c^2	rep
HOL	599,772	2,802	EB	1,050	252	128.90* (3.50)	32.60* (2.50)	0.14* (0.00)	0.03* (0.00)
			EI	1,052	266	196.20* (5.04)	39.90* (3.61)	0.15* (0.00)	0.03* (0.00)
SHO	3,713	298	EB	37,322	107.40* (15.90)	63.50* (12.90)	0.12* (0.02)	0.07* (0.01)	0.19* (0.01)
			EI	37,522	197.10* (23.70)	38.70 (18.50)	0.15* (0.02)	0.03 (0.01)	0.18* (0.01)
AYR	6,933	491	EB	67,199	132.50* (13.30)	48.70* (10.20)	0.15* (0.01)	0.05* (0.01)	0.20* (0.01)
			EI	67,514	197.10* (17.2)	37.30* (12.90)	0.15* (0.01)	0.03* (0.01)	0.18* (0.01)
JER	13,699	523	EB	154,480	183.70* (9.40)	22.80* (6.40)	0.15* (0.01)	0.02* (0.01)	0.16* (0.00)
			EI	151,903	235.00* (12.10)	13.20 (8.20)	0.14* (0.01)	0.01 (0.01)	0.15* (0.00)
GUE	2,120	84	EB	20,511	183.10* (25.20)	29.50 (17.90)	0.17* (0.03)	0.03 (0.02)	0.20* (0.01)
			EI	20,694	184.10* (29.00)	46.50 (21.50)	0.13* (0.02)	0.03 (0.01)	0.16* (0.01)

¹HOL = Holstein Friesian; SHO = Shorthorn; AYR = Ayrshire; JER = Jersey; GUE = Guernsey; σ_a^2 = additive genetic variance; σ_{pe}^2 = permanent environmental variance; \hbar^2 = heritability; c^2 = proportion of total phenotypic variance due to permanent environmental variance; rep = repeatability.

*P < 0.05.

their cost will also determine the effectiveness and profitability of cows with sustainable energy profiles.

The prediction of energy traits from milk in this study relies on several assumptions: that the research herd from which reference data were sourced is representative of the herds in which it is intended to predict (McParland et al., 2012); that the reference data—itself modeled by random regression (Banos and Coffey, 2010)—is accurate; that changes in the energy profiles of dairy cows leave some sort of molecular signal in their milk; and that we are able to detect this signal and calibrate it. We have demonstrated that by exploiting the relationship between the MIR absorption spectra (ultimately, relative milk fat and protein molecular concentrations) and concurrent phenotypes, a robust tool for energy traits can be applied to herds in the UK with cost and labor advantages over more manual (e.g., measurement of DMI) or subjective methods (e.g., BCS) or those that do not accurately reflect underlying energy stores (e.g., BW).

CONCLUSIONS

The automated prediction of energy traits deployed via routine milk recording and as part of DHI schemes could provide farmers with a low-cost and, more importantly, noninvasive tool to guide nutritional and transition cow management within herds. This would enable farmers to maximize number of lactations and reduce early culling, as well as optimize feed efficiency (due to the strong genetic correlation of EB with residual feed intake; McParland and Berry, 2016; Hurley et al., 2017) while ameliorating the effects of health and fertility issues associated with NEB. This could be particularly effective if contemporary selection objectives (primarily for yield) favor cows that mobilize their body tissue and show compromised fertility (Coffey et al., 2001).

ACKNOWLEDGMENTS

The authors gratefully acknowledge collaboration with National Milk Records (Martin Busfield, Victoria Hicks; Harrogate, UK) and Marks and Spencer (London, UK); funding from Innovate UK (previously TSB; Swindon, UK; grant no: TP12483-81287) and Biotechnology and Biological Sciences Research Council (BBSRC, Swindon, UK; grant no: BB/K002260/1); and cooperation with the participating farms. Ian Archibald (SRUC, Edinburgh, UK) is acknowledged for curating the Langhill data, Jenni Flockhart (SRUC) for details of the Langhill diets, and all staff at Crichton farm for collecting samples and managing animals. The Langhill experiment at Crichton Dairy Research Centre and authors MC, SJD, and EW are supported by the

Scottish Government Rural Affairs, Food and the Environment (RAFE, Edinburgh, UK) Strategic Research Portfolio 2016–2021.

REFERENCES

- Banos, G., and M. P. Coffey. 2010. Genetic association between body energy measured throughout lactation and fertility in dairy cattle. *Animal* 4:189–199. <https://doi.org/10.1017/S1751731109991182>.
- Banos, G., M. P. Coffey, R. F. Veerkamp, D. P. Berry, and E. Wall. 2012. Merging and characterising phenotypic data on conventional and rare traits from dairy cattle experimental resources in three countries. *Animal* 6:1040–1048. <https://doi.org/10.1017/S1751731111002655>.
- Banos, G., M. P. Coffey, E. Wall, and S. Brotherstone. 2006. Genetic relationship between first-lactation body energy and later-lifeudder health in dairy cattle. *J. Dairy Sci.* 89:2222–2232. [https://doi.org/10.3168/jds.S0022-0302\(06\)72293-7](https://doi.org/10.3168/jds.S0022-0302(06)72293-7).
- Berry, D. P., B. Horan, M. O'Donovan, F. Buckley, E. Kennedy, M. McEvoy, and P. Dillon. 2007. Genetics of grass dry matter intake, energy balance, and digestibility in grazing Irish dairy cows. *J. Dairy Sci.* 90:4835–4845. <https://doi.org/10.3168/jds.2007-0116>.
- Coffey, M. P., G. C. Emmans, and S. Brotherstone. 2001. Genetic evaluation of dairy bulls for energy balance traits using random regression. *Anim. Sci.* 73:29–40. <https://doi.org/10.1017/S135772980058021>.
- Coffey, M. P., G. Simm, W. G. Hill, and S. Brotherstone. 2003. Genetic evaluations of dairy bulls for daughter energy balance profiles using linear type scores and body condition score analyzed using random regression. *J. Dairy Sci.* 86:2205–2212. [https://doi.org/10.3168/jds.S0022-0302\(03\)73810-7](https://doi.org/10.3168/jds.S0022-0302(03)73810-7).
- Coffey, M. P., G. Simm, J. D. Oldham, W. G. Hill, and S. Brotherstone. 2004. Genotype and diet effects on energy balance in the first three lactations of dairy cows. *J. Dairy Sci.* 87:4318–4326. [https://doi.org/10.3168/jds.S0022-0302\(04\)73577-8](https://doi.org/10.3168/jds.S0022-0302(04)73577-8).
- De Marchi, M., C. C. Fagan, C. P. O'Donnell, A. Cecchinato, R. Dal Zotto, M. Cassandro, M. Penasa, and G. Bittante. 2009. Prediction of coagulation properties, titratable acidity, and pH of bovine milk using mid-infrared spectroscopy. *J. Dairy Sci.* 92:423–432. <https://doi.org/10.3168/jds.2008-1163>.
- De Marchi, M., V. Toffanin, M. Cassandro, and M. Penasa. 2014. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits. *J. Dairy Sci.* 97:1171–1186. <https://doi.org/10.3168/jds.2013-6799>.
- de Vries, M. J., and R. F. Veerkamp. 2000. Energy balance of dairy cattle in relation to milk production variables and fertility. *J. Dairy Sci.* 83:62–69. [https://doi.org/10.3168/jds.S0022-0302\(00\)74856-9](https://doi.org/10.3168/jds.S0022-0302(00)74856-9).
- Dehareng, F., C. Delfosse, E. Froidmont, H. Soyeurt, C. Martin, N. Gengler, A. Vanlierde, and P. Dardenne. 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows. *Animal* 6:1694–1701. <https://doi.org/10.1017/S175173112000456>.
- Dickinson, F. N., B. T. McDaniel, and R. E. McDowell. 1969. Comparative efficiency of feed utilization during first lactation of Ayrshire, Brown Swiss, and Holstein cows. *J. Dairy Sci.* 52:489–497. [https://doi.org/10.3168/jds.S0022-0302\(69\)86593-8](https://doi.org/10.3168/jds.S0022-0302(69)86593-8).
- Emmans, G. C. 1994. Effective energy: A concept of energy utilization applied across species. *Br. J. Nutr.* 71:801–821. <https://doi.org/10.1079/bjn19940188>.
- Friggens, N. C., C. Ridder, and P. Løvendahl. 2007. On the use of milk composition measures to predict the energy balance of dairy cows. *J. Dairy Sci.* 90:5453–5467. <https://doi.org/10.3168/jds.2006-821>.
- Garnsworthy, P. C., K. D. Sinclair, and R. Webb. 2008. Integration of physiological mechanisms that influence fertility in dairy cows. *Animal* 2:1144–1152. <https://doi.org/10.1017/S1751731108002358>.
- Gilmour, A. R., B. Gogel, B. Cullis, and R. Thompson. 2009. ASReml User Guide Release 3.0. VSN Int. Ltd., Hemel Hempstead, UK.
- Grainger, C., and M. E. Goddard. 2004. A review of the effects of dairy breed on feed conversion efficiency—An opportunity lost? *Anim. Prod. Aust.* 25:77–80. <https://doi.org/10.1071/SA0401020>.
- Grelet, C., C. Bastin, M. Gelé, J.-B. Davière, M. Johan, A. Werner, R. Reding, J. A. Fernandez Pierna, F. Colinet, P. Dardenne, N. Gengler, H. Soyeurt, and F. Dehareng. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. *J. Dairy Sci.* 99:4816–4825. <https://doi.org/10.3168/jds.2015-10477>.
- Grelet, C., J. A. Fernández Pierna, P. Dardenne, V. Baeten, and F. Dehareng. 2015. Standardization of milk mid-infrared spectra from a European dairy network. *J. Dairy Sci.* 98:2150–2160. <https://doi.org/10.3168/jds.2014-8764>.
- Grelet, C., J. A. Fernández Pierna, H. Soyeurt, F. Dehareng, N. Gengler, and P. Dardenne. 2014. Creation of universal MIR calibrations by standardization of milk spectra: Example of fatty acids. Page 108 in Book of abstracts of the 65th annual meeting of the European Federation of Animal Science. Wageningen Academic Publishers, Copenhagen, Denmark.
- Haaland, D. M., and E. V. Thomas. 1988. Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Anal. Chem.* 60:1193–1202. <https://doi.org/10.1021/ac00162a020>.
- Hurley, A. M., N. López-Villalobos, S. McParland, E. Lewis, E. Kennedy, M. O'Donovan, J. L. Burke, and D. P. Berry. 2017. Genetics of alternative definitions of feed efficiency in grazing lactating dairy cows. *J. Dairy Sci.* 100:5501–5514. <https://doi.org/10.3168/jds.2016-12314>.
- Hüttmann, H., E. Stamer, W. Junge, G. Thaller, and E. Kalm. 2009. Analysis of feed intake and energy balance of high-yielding first lactating Holstein cows with fixed and random regression models. *Animal* 3:181–188. <https://doi.org/10.1017/S175173110800325X>.
- Koenen, E. P. C., R. F. Veerkamp, P. Dobbelaar, and G. De Jong. 2001. Genetic analysis of body condition score of lactating Dutch Holstein and Red-and-White heifers. *J. Dairy Sci.* 84:1265–1270. [https://doi.org/10.3168/jds.S0022-0302\(01\)74588-2](https://doi.org/10.3168/jds.S0022-0302(01)74588-2).
- Kristensen, T., C. Jensen, S. Østergaard, M. R. Weisbjerg, O. Aaes, and N. I. Nielsen. 2015. Feeding, production, and efficiency of Holstein-Friesian, Jersey, and mixed-breed lactating dairy cows in commercial Danish herds. *J. Dairy Sci.* 98:263–274. <https://doi.org/10.3168/jds.2014-8532>.
- Lainé, A., C. Bastin, C. Grelet, H. Hammami, F. G. Colinet, L. M. Dale, A. Gillon, J. Vandendries, 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. <https://doi.org/10.3168/jds.2016-11736>.
- Løvendahl, P., C. Ridder, and N. C. Friggens. 2010. Limits to prediction of energy balance from milk composition measures at individual cow level. *J. Dairy Sci.* 93:1998–2006. <https://doi.org/10.3168/jds.2009-2739>.
- Lowman, B. G., N. A. Scott, and S. H. Somerville. 1976. Condition scoring of cattle. Bulletin No. 6. East of Scotland College of Agriculture, Edinburgh, UK.
- McParland, S., G. Banos, B. McCarthy, E. Lewis, M. P. Coffey, B. O'Neill, M. O'Donovan, E. Wall, and D. P. Berry. 2012. Validation of mid-infrared spectrometry in milk for predicting body energy status in Holstein-Friesian cows. *J. Dairy Sci.* 95:7225–7235. <https://doi.org/10.3168/jds.2012-5406>.
- McParland, S., G. Banos, E. Wall, M. P. Coffey, H. Soyeurt, R. F. Veerkamp, and D. P. Berry. 2011. The use of mid-infrared spectrometry to predict body energy status of Holstein cows. *J. Dairy Sci.* 94:3651–3661. <https://doi.org/10.3168/jds.2010-3965>.
- McParland, S., and D. P. Berry. 2016. The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. *J. Dairy Sci.* 99:4056–4070. <https://doi.org/10.3168/jds.2015-10051>.
- McParland, S., E. Kennedy, E. Lewis, S. G. Moore, B. McCarthy, M. O'Donovan, and D. P. Berry. 2015. Genetic parameters of dairy cow energy intake and body energy status predicted using mid-infrared spectrometry of milk. *J. Dairy Sci.* 98:1310–1320. <https://doi.org/10.3168/jds.2014-8892>.
- Nicolaï, B. M., K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K. I. Theron, and J. Lammertyn. 2007. Nondestructive measurement

- of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.* 46:99–118. <https://doi.org/10.1016/j.postharvbio.2007.06.024>.
- Prendiville, R., K. M. Pierce, and F. Buckley. 2009. An evaluation of production efficiencies among lactating Holstein-Friesian, Jersey, and Jersey × Holstein-Friesian cows at pasture. *J. Dairy Sci.* 92:6176–6185. <https://doi.org/10.3168/jds.2009-2292>.
- Reist, M., D. Erdin, D. von Euw, K. Tschüemperlin, H. Leuenberger, Y. Chilliard, H. M. Hammon, C. Morel, C. Philipona, Y. Zbinden, N. Kuenzi, and J. W. Blum. 2002. Estimation of energy balance at the individual and herd level using blood and milk traits in high-yielding dairy cows. *J. Dairy Sci.* 85:3314–3327. [https://doi.org/10.3168/jds.S0022-0302\(02\)74420-2](https://doi.org/10.3168/jds.S0022-0302(02)74420-2).
- Reist, M., D. K. Erdin, D. von Euw, K. M. Tschüemperlin, H. Leuenberger, H. M. Hammon, C. Morel, C. Philipona, Y. Zbinden, N. Künzi, and J. W. Blum. 2003. Postpartum reproductive function: Association with energy, metabolic and endocrine status in high yielding dairy cows. *Theriogenology* 59:1707–1723. [https://doi.org/10.1016/S0093-691X\(02\)01238-4](https://doi.org/10.1016/S0093-691X(02)01238-4).
- Roberts, D. J., and M. D. March. 2013. Feeding systems for dairy cows. Homegrown versus by-products feeds. P. C. Garnsworthy and J. Wiseman, ed. Context Products Ltd., Leicestershire, UK.
- Rutten, M. J. M., H. Bovenhuis, J. M. L. Heck, and J. A. M. van Arendonk. 2011. Predicting bovine milk protein composition based on Fourier transform infrared spectra. *J. Dairy Sci.* 94:5683–5690. <https://doi.org/10.3168/jds.2011-4520>.
- Soyeurt, H., F. Dehareng, N. Gengler, S. McParland, E. Wall, D. P. Berry, M. P. Coffey, and P. Dardenne. 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* 94:1657–1667. <https://doi.org/10.3168/jds.2010-3408>.
- van Rossum, G. 1995. Python tutorial, Technical Report CS-R9526. Accessed Mar. 7, 2019. www.python.org. Centrum voor Wiskunde en Informatica (CWI), Amsterdam, the Netherlands.

ORCIDS

- S. L. Smith  <https://orcid.org/0000-0002-0521-2956>
- S. J. Denholm  <https://orcid.org/0000-0002-6291-8269>
- E. Wall  <https://orcid.org/0000-0002-7072-5758>