



Classifying the fertility of dairy cows using milk mid-infrared spectroscopy

P. N. Ho,^{1*} V. Bonfatti,² T. D. W. Luke,^{1,3} and J. E. Pryce^{1,3}

¹Agriculture Victoria, AgriBio, Centre for AgriBioscience, Bundoora, Victoria 3083, Australia

²Department of Comparative Biomedicine and Food Science, University of Padova, Legnaro 35020, Italy

³School of Applied Systems Biology, La Trobe University, Bundoora, Victoria 3083, Australia

ABSTRACT

The objective of this study was to investigate the potential of milk mid-infrared (MIR) spectroscopy, MIR-derived traits including milk composition, milk fatty acids, and blood metabolic profiles (fatty acids, β -hydroxybutyrate, and urea), and other on-farm data for discriminating cows of good versus poor likelihood of conception to first insemination (i.e., pregnant vs. open). A total of 6,488 spectral and milk production records of 2,987 cows from 19 commercial dairy herds across 3 Australian states were used. Seven models, comprising different explanatory variables, were examined. Model 1 included milk production; concentrations of fat, protein, and lactose; somatic cell count; age at calving; days in milk at herd test; and days from calving to insemination. Model 2 included, in addition to the variables in model 1, milk fatty acids and blood metabolic profiles. The MIR spectrum collected before first insemination was added to model 2 to form model 3. Fat, protein, and lactose percentages, milk fatty acids, and blood metabolic profiles were removed from model 3 to create model 4. Model 5 and model 6 comprised model 4 and either fertility genomic estimated breeding value or principal components obtained from a genomic relationship matrix derived using animal genotypes, respectively. In model 7, all previously described sources of information, but not MIR-derived traits, were used. The models were developed using partial least squares discriminant analysis. The performance of each model was evaluated in 2 ways: 10-fold random cross-validation and herd-by-herd external validation. The accuracy measures were sensitivity (i.e., the proportion of pregnant cows that were correctly classified), specificity (i.e., the proportion of open cows that were correctly classified), and area under the curve (AUC) for the receiver operating curve. The results showed that in all models, prediction accuracy obtained through 10-fold random cross-validation was higher than that of

herd-by-herd external validation, with the difference in AUC ranging between 0.01 and 0.09. In the herd-by-herd external validation, using basic on-farm information (model 1) was not sufficient to classify good- and poor-fertility cows; the sensitivity, specificity, and AUC were around 0.66. Compared with model 1, adding milk fatty acids and blood metabolic profiles (model 2) increased the sensitivity, specificity, and AUC by 0.01, 0.02, and 0.02 unit, respectively (i.e., 0.65, 0.63, and 0.678). Incorporating MIR spectra into model 2 resulted in sensitivity, specificity, and AUC values of 0.73, 0.63, and 0.72, respectively (model 3). The comparable prediction accuracies observed for models 3 and 4 mean that useful information from MIR-derived traits is already included in the spectra. Adding the fertility genomic estimated breeding value and animal genotypes (model 7) produced the highest prediction accuracy, with sensitivity, specificity, and AUC values of 0.75, 0.66, and 0.75, respectively. However, removing either the fertility estimated breeding value or animal genotype from model 7 resulted in a reduction of the prediction accuracy of only 0.01 and 0.02, respectively. In conclusion, this study indicates that MIR and other on-farm data could be used to classify cows of good and poor likelihood of conception with promising accuracy. **Key words:** likelihood of conception, discriminant analysis, prediction accuracy

INTRODUCTION

Good fertility is a key driver of profit in dairy farming because it ensures better control of culling, milk sales, and number of replacements (Kaniyamattam et al., 2016). This is arguably of greater importance in seasonal calving systems compared with nonseasonal systems because good fertility is a prerequisite to maintain a compact calving period and match the high energy requirements of the cow in early lactation with peak pasture growth rate (Armstrong et al., 2010; Shaloo et al., 2014). Multiple factors have been reported to be associated with variation in conception rate. Nongenetic factors include quality and quantity of bull semen (DeJarnette et al., 2004), age, body con-

Received January 31, 2019.

Accepted July 23, 2019.

*Corresponding author: phuong.ho@agriculture.vic.gov.au

dition, energy balance, RUP, milk yield, milk protein concentration, and health status of the cow (Roche et al., 2007; Shorten et al., 2015; Rodney et al., 2016; Leroy et al., 2017; Morton et al., 2017; Rearte et al., 2018), days postcalving (Clay et al., 2004), heat stress (Morton et al., 2007), lameness (Alawneh et al., 2011), and insemination season (Cornwell et al., 2006). Berry et al. (2011) indicated that about 2.3% of the phenotypic variation in conception rate could be attributed to additive genetic effects.

Despite the large efforts that have been made to investigate factors related to conception rate, comparatively few authors have attempted to predict the outcome of an individual insemination event (i.e., pregnant vs. open). Prior knowledge of how likely a cow is to get pregnant, once inseminated, would enable farmers to optimize breeding decisions. For example, sexed or premium bull semen could be used for cows predicted to have a high likelihood of conception, whereas cows with predicted poor fertility could be mated using semen from beef bulls, multiple doses, or semen from bulls of known high genetic merit for fertility. To our knowledge, Grzesiak et al. (2010), Shahinfar et al. (2014), and Hempstalk et al. (2015) are the only authors who have reported the ability of some on-farm data to predict likelihood of conception to a given insemination of dairy cows. The value of the prediction accuracy ranged between 0.66 (Hempstalk et al., 2015) and 0.91 (Grzesiak et al., 2010). Unfortunately, some of the important variables used in these studies might be difficult to obtain on-farm (e.g., BCS and BW) or cannot be predicted a priori (e.g., year).

Mid-infrared (MIR) spectroscopy is routinely used by worldwide milk recording organizations to quantify the concentration of fat, protein, and lactose in milk samples. Moreover, an increasing number of studies have reported a good capability of MIR to predict traits related to cow metabolic status, such as energy balance, milk and serum fatty acids, and BHB [see reviews by De Marchi et al. (2014) and Gengler et al. (2016)], which are known to affect fertility (Ribeiro et al., 2013). Also, Toledo-Alvarado et al. (2018b) reported that milk fatty acid profiles predicted by MIR could be used to identify cows in proestrus or estrus. Alternatively, the MIR spectra were also used directly as indicators of health and fertility status of the cows. Toledo-Alvarado et al. (2018a) showed that pregnant versus open cows postinsemination could be discriminated with promising accuracy using MIR spectra, parity, and DIM. Hempstalk et al. (2015), however, concluded that the inclusion of MIR spectra did not improve the accuracy of predicting the likelihood of conception to an insemination compared with the same model but without MIR spectra.

The objective of this study was to investigate the potential of milk MIR spectra, together with other variables, for classifying cows of good and poor likelihood of conception at first insemination. The additional variables evaluated were milk fatty acids and blood metabolic profiles (fatty acids, BHB, and urea), milk production and composition, DIM at herd test, days from calving to insemination, calving age, fertility genomic EBV, and SNP genotypes of the cow.

MATERIALS AND METHODS

Animal Data

Records of insemination and date of calving were available for 8,064 spring-calving cows from 19 commercial dairy herds located in Victoria, Tasmania, and New South Wales of Australia in 2016 and 2017. The cows were between first and sixth parity and predominantly Holstein-Friesian (74.3%), but the data set also included 8.2% purebred Jersey and 17.5% crossbred animals. Other data available included DIM at herd test, days from calving to insemination, age at calving, previous-lactation milk yield, milk fat yield, and milk protein yield, all expressed on a 305-d basis; current-lactation herd test-day milk yield, fat, protein, and lactose percentages; SCC; milk and serum fatty acids, BHB, and urea; fertility genomic EBV (GEBV); genotype of the cow; and MIR spectra. Milk fatty acids and blood metabolic profiles were predicted from MIR using the equations developed by Ho et al. (2019) and Luke et al. (2019), respectively. Milk production, milk composition, insemination and calving records, fertility GEBV, and 47,162 SNP genotypes (BovineSNP50 BeadChip), edited for the routine genomic evaluations, were obtained from DataGene (Bundoora, Victoria, Australia). To incorporate the genotype data into the prediction model, a genomic relationship matrix (GRM; a matrix of $8,604 \times 8,604$ estimating the fraction of total DNA that 2 individuals share) was first derived using the method of Yang et al. (2010). A principal component analysis was then applied on the GRM using the R function “prcomp” to determine the optimal number of GRM components to be included in future analyses, a model (i.e., model 7 as described later) that included MIR spectra, previous-lactation 305-d milk yield, milk fat yield, and milk protein yield, current-lactation herd test-day milk yield, DIM at herd test, days from calving to insemination, calving age, and fertility GEBV was iteratively run with a descending order of size of eigenvalue. The preliminary analysis showed that the first 84 components (explaining 84.6% of the total variation of the GRM) produced the greatest contribution to

the prediction accuracy and thus were used for model development.

Spectral Data

In this data set, all cows were milked twice daily in accordance with the standard commercial practices of herd-testing organizations in Australia. Milk samples (either a.m. or p.m.) were collected and sent to TasHerd Pty Ltd. (Hadspen, Tasmania, Australia) to be analyzed for fat, protein, and lactose concentrations and SCC using a NexGen Series FTS Combi machine (Bentley Instruments, Chaska, MN), and the corresponding spectra were obtained for this study. Each cow had 2 to 8 records. A recorded spectrum includes 899 data points, with each point representing the absorption of infrared light through the milk sample at a particular wavenumber in the 649 to 3,999 cm^{-1} region.

Data Manipulation

The main objective of this study was to examine the potential of MIR spectra and other on-farm data for classifying cows of good and poor likelihood of conception at first insemination. Thus, we first divided the cows in the data set into 3 groups as shown in Table 1: good (cows that conceived at first insemination), average (cows that conceived following 2 or more inseminations and cows that did not conceive but where number of inseminations was >1), and poor (cows with no conception event recorded and that had only 1 insemination). Conception was confirmed by a calving in the subsequent year and was coded binarily as 1 (pregnant) or 0 (open). Mating records that resulted in abortions were removed from the data. The conception event was assumed to result from the last recorded insemination.

It was hypothesized that cows in the good and poor groups might have significantly different metabolic conditions and, consequently, different likelihood to conceive, whereas the metabolic condition of cows in the average group could be similar to that of cows in the other 2 groups. By focusing on the good and poor groups, the differences would be magnified and may help improve the predictability of the model. Second, to be predictive, only spectral records obtained before the first insemination were retained, which reduced the data to 6,488 records of 2,897 cows for final analyses. The mean \pm standard deviation of the number of days between milk sampling for spectral collection and first insemination was 43.4 ± 25.1 . Although there were multiple spectra per cow (i.e., 2.2 on average), we considered each spectrum to be unique because of the large differences in terms of, for example, diet, lactation

stage, and management at the time each observation was recorded, which is a common practice in many MIR studies (Soyeurt et al., 2011; McParland et al., 2014; van Gastelen et al., 2018).

Pretreatments were also applied to the raw spectra. First, spectral regions (2,998–3,998 cm^{-1} , 1,615–1,652 cm^{-1} , and 649–925 cm^{-1}) characterized by low signal: noise ratio, which is the consequence of high water absorption, were removed before chemometric analyses (Hewavitharana and van Brakel, 1997). This resulted in 536 wavenumbers available for model development. Second, to discard the spectra that are potentially outliers, a standardized Mahalanobis distance (often known as global H distance; Shenk and Westerhaus, 1995) between each spectrum and the population average was calculated. Then, spectra with a global distance greater than 3 ($n = 24$) were considered to be outliers and eliminated as recommended by Williams (2004). Finally, extended multiplicative correction (Kohler et al., 2009) and first-order Savitzky–Golay derivative (Savitzky and Golay, 1964) were applied to the reduced spectra.

The prediction equations of Ho et al. (2019) and Luke et al. (2019) were applied on the preprocessed spectra to derive milk fatty acids (C4:0, C6:0, C8:0, C10:0, C12:0, C14:0, C16:0, C17:0, C18:0, C18:1 *cis*-9, and C20:0) and the concentrations in sera of fatty acids, BHB, and urea, respectively.

Model Development and Evaluation of Performance

Discriminant models to classify cows that conceived at first insemination and did not conceive within the breeding season were developed using partial least squares discriminant analysis (PLS-DA) and implemented with the mixOmics R package of Lê Cao et al. (2011). Partial least squares discriminant analysis is a variant of partial least squares regression when the response variable is categorical that is used to find the relationship between 2 matrices. It is one of the most well-known classification methods in chemometrics, metabolomics, and proteomics and has the ability to analyze highly collinear data, which is often a problem with conventional regression methods, such as logistic regression (Gromski et al., 2015). The predictors were scaled using an option in the package (i.e., each variable was standardized by dividing itself by the standard deviation). Each model's performance was evaluated in 2 ways: 10-fold random cross-validation and herd-by-herd external validation. In the 10-fold random cross-validation, the data set was randomly split into 10 parts that were balanced in terms of the ratio of pregnant and open cows using the groupdata2 R package (Olsen,

2017). One part was reserved for validation, whereas the remaining data were used for model training. This process was repeated 10 times until each part of the data had been validated once. In the herd-by-herd external validation, the data of a given herd was excluded and used as a validation of the model trained with the data of the other 18 herds. The process continued until every herd had been validated once (i.e., 19 times, as there were 19 herds in this study).

The accuracy of each discriminant model was evaluated by producing the receiver operating characteristic curves and calculating the area under the curve (AUC) through the 2 validation processes described previously. The optimal cut-off value for each test variable was

defined as the point where the sum between sensitivity and specificity was at a maximum (i.e., equal weighing of false-positive and false-negative test results), where sensitivity is the proportion of pregnant cows that were correctly classified and specificity is the proportion of open cows that were correctly classified. The PLS-DA method used in the mixOmics package already uses a prediction threshold based on distances that optimally determine class membership of the samples tested. Therefore, according to Lê Cao et al. (2011), AUC and receiver operating characteristic curves are not needed to estimate the performance of the model and are provided only as complementary performance measures. The estimated *P*-values from Wilcoxon tests between

Table 1. Description (mean \pm SD) of explanatory variables used, besides infrared spectra, for classifying cows as having good, average, and poor likelihood of conception at first insemination

Item	Likelihood of conception at first insemination ¹			<i>P</i> -value ²
	Good (n = 4,123)	Average (n = 2,356)	Poor (n = 2,618)	
DIM at herd test (d)	62.6 \pm 56.9	69.0 \pm 58.5	57.9 \pm 49.9	***
Days from calving to insemination (d)	106.3 \pm 59.2	144.4 \pm 91.7	96.2 \pm 49.9	***
Age at calving (mo)	48.6 \pm 24.6	56.8 \pm 30.6	48.4 \pm 24.5	***
Traits of previous lactation ³ (305-d kg)				
Milk yield	6,901 \pm 1,734	7,185 \pm 1,759	7,319 \pm 1,813	***
Fat yield	280.5 \pm 61.9	279.1 \pm 61.3	293.7 \pm 67.1	***
Protein yield	236.0 \pm 55.8	240.8 \pm 55.8	248.1 \pm 59.1	***
Lactose yield	324.3 \pm 82.0	324.8 \pm 84.1	345.8 \pm 84.9	***
Traits of current lactation per herd test day				
Milk yield (kg/d)	27.6 \pm 7.8	28.9 \pm 8.6	28.8 \pm 9.0	***
Fat (%)	3.65 \pm 0.83	3.49 \pm 0.82	3.76 \pm 1.09	***
Protein (%)	3.35 \pm 0.40	3.22 \pm 0.43	3.28 \pm 0.42	***
Lactose (%)	5.11 \pm 0.19	5.10 \pm 0.21	5.09 \pm 0.21	***
SCC	135 \pm 523	166 \pm 590	110 \pm 377	*
Milk fatty acids (g/100 g of milk)				
C4:0	0.096 \pm 0.044	0.086 \pm 0.044	0.101 \pm 0.051	***
C6:0	0.048 \pm 0.027	0.042 \pm 0.027	0.051 \pm 0.033	***
C8:0	0.031 \pm 0.017	0.027 \pm 0.016	0.033 \pm 0.020	***
C10:0	0.066 \pm 0.041	0.051 \pm 0.040	0.065 \pm 0.049	***
C12:0	0.066 \pm 0.049	0.056 \pm 0.048	0.071 \pm 0.058	***
C14:0	0.294 \pm 0.121	0.272 \pm 0.121	0.031 \pm 0.150	***
C16:0	1.250 \pm 0.399	1.175 \pm 0.426	1.292 \pm 0.459	***
C17:0	0.039 \pm 0.007	0.038 \pm 0.007	0.039 \pm 0.008	***
C18:0	0.223 \pm 0.103	0.208 \pm 0.109	0.229 \pm 0.114	***
C18:1 <i>cis</i> -9	0.659 \pm 0.205	0.630 \pm 0.213	0.681 \pm 0.227	***
C20:0	0.004 \pm 0.002	0.004 \pm 0.002	0.004 \pm 0.002	NS
Short-chain fatty acids	0.232 \pm 0.125	0.203 \pm 0.125	0.248 \pm 0.151	***
Medium-chain fatty acids	1.713 \pm 0.524	1.611 \pm 0.548	1.771 \pm 0.624	***
Long-chain fatty acids	0.885 \pm 0.309	0.839 \pm 0.324	0.916 \pm 0.349	***
De novo fatty acids	1.256 \pm 0.544	1.161 \pm 0.462	1.311 \pm 0.551	***
Blood metabolic profiles (mmol/L of blood)				
Fatty acids	0.445 \pm 0.172	0.410 \pm 0.184	0.484 \pm 0.167	***
BHB	0.427 \pm 0.168	0.475 \pm 0.156	0.382 \pm 0.172	***
Urea	0.676 \pm 0.168	0.649 \pm 0.182	0.692 \pm 0.161	***
Fertility genomic EBV	103.5 \pm 4.5	102.6 \pm 4.2	103.2 \pm 4.6	***

¹Good = cows that conceived at first insemination; average = cows that conceived following 2 or more inseminations; poor = cows with no conception event recorded. n = number of records.

²*P*-values were obtained from 1-way ANOVA tests with pairwise comparisons.

³Estimate yield in kilograms by a standard period of 305 d.

P* < 0.05; **P* < 0.0005; NS = *P* \geq 0.05.

the predicted scores of one class versus the other were also obtained, but because they were all statistically significant, they are not reported here.

In this study, 7 models comprising different explanatory variables were tested for their capability in classifying cows of good and poor likelihood of conception (Table 2). Model 1 included features that are always available on farms that adhere to the herd-testing program, such as milk production, milk composition, DIM at herd test, and days from calving to insemination. Models 2 and 3 aimed to compare the additional value of milk fatty acids and blood metabolic profiles versus the MIR spectrum when being incorporated into the basic model, respectively. Fat, protein, and lactose percentages, milk fatty acids, and blood metabolic profiles were removed from model 3 to create model 4. Preliminary results showed that adding MIR spectra produced prediction accuracy that was comparable (model 4) with the model using both MIR-derived traits and the spectra (model 3); thus, MIR-derived traits were not considered in future models. Accordingly, models 5, 6, and 7 were used to investigate the contribution of adding the fertility GEBV and animal genotypes on top of the predictors in model 4 to the model performance. The statistical measures of performance of the 7 models were compared using a 1-way ANOVA test in R with pairwise comparisons. Noticeably, in order for the 7 models to be developed using PLS-DA and subsequently have statistically fair comparisons, a random noise matrix with dimensions of $n \times p$, where $n = 536$ is the number of wavenumbers in the reduced spectra and p is the number of records of the validation set, was generated from a uniform distribution in the interval 0.0 to 1.0 and multiplied by a very small constant of 10^{-10} . Such a matrix was then used in models 1 and 2 to represent the spectral wavenumbers. This method has been proposed previously to identify the uninformative MIR wavenumbers by Gottardo et al. (2016). All analyses in the present study were performed using R statistical software version 3.4.4 (R Development Core Team, 2018).

RESULTS AND DISCUSSION

The ability to accurately predict the outcome of an individual insemination event given to a cow (i.e., pregnant vs. open) would allow farmers to implement strategies to optimize breeding decisions. For instance, sexed semen could be used to breed cows with a high likelihood of conception, whereas beef semen or semen from bulls of known high genetic merit of fertility could be used for cows predicted to have poor likelihood of conception. Additionally, farmers might adjust feeding or management strategies to help predicted poor cows

Table 2. Predictors included in each model for classifying cows as having good or poor likelihood of conception at first insemination¹

Model	MIR	DIM	DAI	Calving age	Previous lactation information	MY	Fat	Protein	Lactose	SCC	Milk fatty acids	MP	Fertility GEBV	Genotype
1		x	x	x	x	x	x	x	x	x				
2		x	x	x	x	x	x	x	x	x	x			
3		x	x	x	x	x	x	x	x	x	x	x		
4	x	x	x	x	x	x								
5	x	x	x	x	x	x							x	
6	x	x	x	x	x	x							x	x
7	x	x	x	x	x	x							x	x

¹MIR = milk mid-infrared spectroscopy; DIM = DIM at herd test; DAI = days from calving to insemination; Previous lactation information = 305-d milk yield (kg), 305-d fat yield (kg), and 305-d protein yield (kg); MY = milk yield on herd test day (kg/d); Fat = fat (%); Protein = protein (%); Lactose = lactose (%); Calving age = age at current calving (mo); Milk fatty acids (g/100 g of milk) = C4:0, C6:0, C8:0, C10:0, C12:0, C14:0, C16:0, C18:0, C18:1 *cis*-9, and C20:0, predicted by Ho et al. (2019); MP = blood metabolic profiles (mmol/L of blood; fatty acids, BHB, and urea), predicted by Luke et al. (2019); Fertility GEBV = genomic EBV for fertility; Genotype = first 84 principal components of the genomic relationship matrix.

improve their physiological conditions and potential probability of conception. In this study, we found that MIR data obtained from herd testing in early lactation can be used to predict cows that are divergent in probability of conception.

In this study, data on 2,987 cows from 19 commercial Australian herds were used to classify cows that contrasted in likelihood of conception to first insemination. The herds were distributed in different regions (mainly in the state of Victoria) to make sure that the data were sufficiently representative. This is important because the Australian dairy industry is well recognized to have diverse feeding systems, which range from grazed pasture to TMR (Dairy Australia, 2016a). Differences in feeding and genetics have been reported to significantly affect milk composition and thus MIR spectra (Jenkins and McGuire, 2006; Gottardo et al., 2017; Tomassini et al., 2019). Figure 1 presents the conception rate to first service of the herds used in this study. The conception rate ranged from 0.22 to 0.54 with an average of 0.38. These results are comparable with those reported by Dairy Australia (2016b), where the conception rate to first service ranged between 0.22 and 0.61 with an average of 0.39.

One of the important steps in data editing was splitting the cows into 3 fertility groups: good (those that conceived to 1 insemination), average (those that conceived following 2 or more inseminations and those that did not conceive but where number of inseminations was >1), and poor (those with no conception event

recorded and that had only 1 insemination). The hypothesis behind this was that cows in the good and poor groups are more likely to differ in their metabolic status, which would result in different reproductive performance (Oikonomou et al., 2008; Pryce et al., 2016). Such differences in metabolic status are expected to be captured by MIR spectra (Belay et al., 2017; Grelet et al., 2018; Pralle et al., 2018; Luke et al., 2019). The metabolic characteristics of the cows in the average group were hypothesized to be similar to those of the other 2 groups, which consequently makes them difficult to differentiate.

As shown in Table 1, the means of the predictors for the cows in the good and poor groups seemed to differ from each other more often, whereas average cows were similar to those in the other 2 groups. Cows in the poor group produced significantly more milk and had higher yields of fat, protein, and lactose (305-d yield, kg) compared with cows in the good group (7,319 vs. 6,901, 293.7 vs. 280.5, 248.1 vs. 236.0, and 345.8 vs. 324.3, respectively). Milk, fat, and protein yields of cows in the average group were between the yields of cows in the other 2 groups. Conversely, the results for several of the other traits in our analysis were not consistent; for example, the average cows had higher BHB but lower serum fatty acids compared with the good cows (0.475 vs. 0.427 and 0.410 vs. 0.445, respectively). The imperfect prediction accuracy of BHB ($R^2 \approx 0.48$) and serum fatty acids ($R^2 \approx 0.61$) could be an explanation for this result (Luke et al., 2019). Although differences in the means of predictors of cows in the average group were statistically significant from those of cows in the good and poor groups, the pattern was not consistent and therefore makes interpretation difficult. Indeed, we attempted to train the models, using the same explanatory variables, to classify pregnant versus open cows in the entire data set (i.e., 3 categories instead of 2), and the prediction accuracy was around 50% (data not shown), which can be achieved just by random chance (Chollet and Allaire, 2018). The idea of creating extreme groups to improve model performance has been proposed in previous classification studies (Preacher et al., 2005; Grzesiak et al., 2010). For example, Grzesiak et al. (2010) assigned cows that conceived with 2 or fewer inseminations to the good group and the remaining animals to poor groups.

Table 3 shows the classification accuracy of the 7 models obtained through 10-fold random cross-validation and the herd-by-herd external validation. The prediction accuracy of all the models obtained through the random cross-validation was consistently higher than that of the herd-by-herd external validation, with the differences in AUC ranging from 0.01 to 0.09. This is understandable because in the first validation ap-

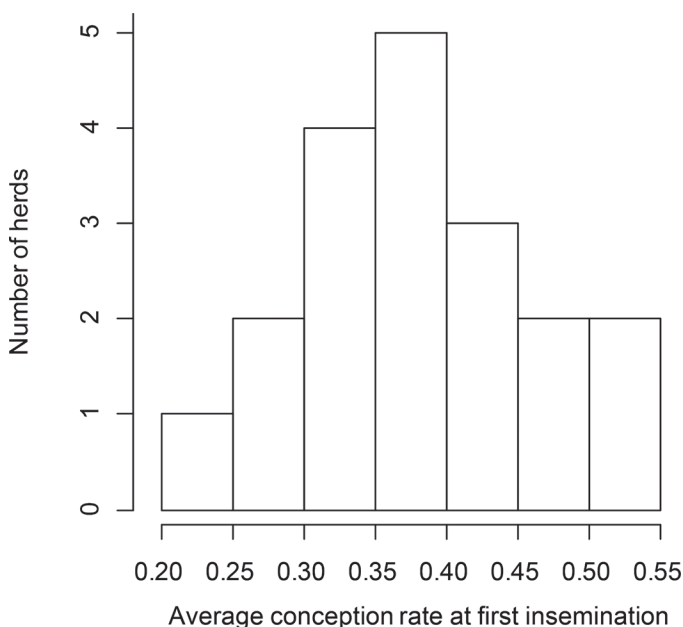


Figure 1. Distribution of conception rate at first insemination across the 19 herds (unedited data) used in this study.

proach, the data were first pooled together and then partitioned randomly into 10 parts without any consideration of cows or their herds. As a result, records from the same herd might have appeared in both the training and validation sets. It should, however, be noted that this is the most common approach used in the majority of MIR prediction studies to evaluate model performance. The small size of the reference data is probably the most likely reason for not being able to perform an external validation. A reduction in prediction accuracy in external validation compared with that in random cross-validation has been reported by several authors. Luke et al. (2019) observed that the values of coefficients of determination decreased by 0.07, 0.11, and 0.55 for external validation compared with random cross-validation for models predicting serum concentrations of BHB, fatty acids, and urea in Australian dairy cows, respectively. McParland et al. (2012) indicated that the model for predicting energy balance developed using data from the Scotland's Rural College research farm did not work when applied to the data from the Teagasc Animal and Grassland Research and Innovation Center in Moorepark, Ireland, with the correlation coefficient decreasing from 0.7 to 0.1. However, the standard deviation of prediction accuracy obtained from herd-by-herd external validation varied more greatly than that obtained from random cross-validation.

Interestingly, the average classification accuracy of the best model (model 7) in our study remained consistently high even in the herd-by-herd external validation with an average sensitivity, specificity, and AUC of 0.75, 0.66, and 0.75, respectively. According to Šimundić (2009), the model diagnostic accuracy is good if the value of the AUC is between 0.7 and 0.8. The standard deviations of prediction accuracy obtained

from herd-by-herd external validation were, however, higher than random cross-validation (Table 3), which indicates that the results are more variable and thus a larger data set is needed to improve the robustness of the model. It was also noticeable that the 3 herds with the lowest prediction accuracy had a mean conception rate at first AI of 0.46, whereas this was 0.35 for the 3 herds with the highest prediction accuracy (data not reported). On the one hand, it may imply that the model worked better on herds with low fertility, but this requires further research to confirm. On the other hand, it may simply indicate that we need more data to improve the robustness of the model.

Using random cross-validation as a reference, the results from our study are slightly lower than those obtained by Grzesiak et al. (2010) but higher than those of Shahinfar et al. (2014) and Hempstalk et al. (2015). Although Grzesiak et al. (2010) showed that their model could differentiate between cows that conceived within 2 inseminations and the rest, with a sensitivity and specificity of around 0.85, the model was validated only once by splitting the data into training set (768 cows) and validation set (150 cows). Last, the variables used in their study such as BW and BCS might be difficult to collect, especially on extensively managed, pasture-based dairy production systems. However, investment in farm automation, including walk-over weighers and cameras for automated body condition scoring, may change this in the future. Shahinfar et al. (2014) and Hempstalk et al. (2015) reported an AUC value of around 0.67 for predicting the likelihood of conception to any given insemination, which is 0.16 lower than our result of 0.83 (model 7). The low prediction accuracy could be due to the fact that they did not create extreme groups of cows as in this study and Grzesiak et al. (2010) but rather considered

Table 3. Validation accuracy (mean \pm SD) of the partial least squares discriminant analysis models for classifying cows as having good or poor likelihood of conception at first insemination^{1,2}

Model ³	10-fold random cross-validation				Herd-by-herd external validation			
	LV (no.)	Sensitivity	Specificity	AUC	LV (no.)	Sensitivity	Specificity	AUC
1	24	0.65 \pm 0.05 ^a	0.54 \pm 0.04 ^a	0.66 \pm 0.02 ^a	13	0.64 \pm 0.15 ^a	0.61 \pm 0.20	0.66 \pm 0.14 ^a
2	24	0.72 \pm 0.02 ^b	0.62 \pm 0.03 ^b	0.71 \pm 0.02 ^b	13	0.65 \pm 0.16 ^a	0.63 \pm 0.20	0.68 \pm 0.14 ^a
3	24	0.80 \pm 0.02 ^c	0.68 \pm 0.03 ^c	0.81 \pm 0.02 ^c	10	0.73 \pm 0.20 ^{ab}	0.63 \pm 0.26	0.72 \pm 0.13 ^{ab}
4	20	0.79 \pm 0.03 ^c	0.68 \pm 0.03 ^c	0.80 \pm 0.02 ^c	11	0.74 \pm 0.20 ^{ab}	0.62 \pm 0.26	0.72 \pm 0.15 ^{ab}
5	22	0.81 \pm 0.02 ^c	0.69 \pm 0.02 ^c	0.81 \pm 0.02 ^c	11	0.74 \pm 0.18 ^{ab}	0.64 \pm 0.23	0.74 \pm 0.13 ^{ab}
6	21	0.80 \pm 0.02 ^c	0.71 \pm 0.03 ^{cd}	0.82 \pm 0.02 ^{ce}	13	0.75 \pm 0.16 ^{bc}	0.62 \pm 0.21	0.73 \pm 0.12 ^{ab}
7	21	0.80 \pm 0.02 ^c	0.72 \pm 0.03 ^d	0.83 \pm 0.02 ^e	13	0.75 \pm 0.16 ^{bc}	0.66 \pm 0.20	0.75 \pm 0.11 ^{bc}

^{a-e}Values within a column with different superscripts are significantly different ($P < 0.05$).

¹Good = cows that conceived at first insemination; poor = cows with no conception event recorded.

²LV = number of latent variables included in the model; sensitivity = proportion of pregnant cows that were correctly classified; Specificity = proportion of open cows that were correctly classified; AUC = area under the curve of the receiver operating curve.

³See Table 2 for model descriptions.

only pregnant versus open cows at any given insemination. The imperfect heat detection and unknown effects of other factors such as herd, year, male fertility, abortion, and insemination technician capability were claimed to contribute to such poor results. This could further be complicated by synchronization programs; for example, cows that calve late in the seasonal calving system are often synchronized and undergo timed AI without a need to observe the signs of estrus (Herlihy et al., 2011). Hempstalk et al. (2015) also concluded that including MIR spectra did not improve prediction accuracy, which disagrees with our findings.

In the current study, the inclusion of milk MIR information, either indirectly via milk composition, milk fatty acids, or blood metabolic profiles or directly via MIR wavenumbers, significantly improved the model performance compared with the model including only milk production, milk composition, SCC, DIM at herd test, days from calving to insemination, and age at calving. The improvement in prediction accuracy was between 0.02 and 0.15 for both validation methods. The results presented in Table 3 imply that using only basic on-farm information (model 1) was not sufficient to classify cows into 2 extreme groups. Adding milk fatty acids and blood metabolic profiles predicted using the MIR equations developed by Ho et al. (2019) and Luke et al. (2019) raised the classification accuracy by 0.02 to 0.05 (model 2). Interestingly, we further improved the prediction accuracy of the model by between 0.04 and 0.10 by incorporating the MIR spectra (model 3), implying that MIR spectra capture variation in fertility beyond milk fatty acids and blood metabolic profiles. Using milk metabolomic or proteomic approaches may elucidate some of these compounds (Goldansaz et al., 2017; Ceciliani et al., 2018; Xu et al., 2018; Greenwood and Honan, 2019). The removal of MIR-derived traits from model 3 did not change prediction accuracy, which means that the useful information obtained from the MIR prediction equations of milk fatty acids, blood metabolic profiles, and milk composition is already included in the MIR spectra. These results agree well with the report of Mineur (2017), who showed that adding MIR-predicted fatty acids and metabolic profiles into a model that already has MIR spectra did not improve the prediction accuracy of lame cows. Grelet et al. (2015) stated that using the spectra directly as a reflection of animal health and metabolic status would be a better option than the intermediate traits.

Fertility of dairy cows has been reported to be heritable, with estimates ranging from 0.01 to 0.13 depending on the component trait (Haile-Mariam et al., 2003; Liu et al., 2008; Berry et al., 2014). In Australia, the fertility breeding value includes calving interval, lactation length, calving to first service interval, first service

nonreturn rate, and pregnancy rate (Haile-Mariam et al., 2013). The incorporation of fertility GEBV and the animal genotypes (derived from the first 84 principal components of the genomic relationship matrix) would therefore be expected to improve the performance of the model. Although the difference was not statistically significant, a 1 to 4% increase in sensitivity, specificity, and AUC was observed in models 5, 6, and 7 compared with model 4. Compared with the performance of model 7, discarding fertility GEBV (model 5) and animal genotype (model 6) reduced the prediction accuracy by 0.01 and 0.02, respectively.

Although we have shown that the top models (5, 6, and 7) could correctly classify approximately 74% of cows of good and poor likelihood of conception at first insemination, it is important to explore how the models would perform when applied to a random population (i.e., a population that also includes cows from the average group; Table 2). Accordingly, model 7 was chosen for this test. Briefly, we repeated the process of herd-by-herd external validation for model 7 and observed the proportion of correct classification for good, average, and poor groups. The prediction accuracy remained the same for the good and poor cows (i.e., 0.75; Table 3), whereas it was only 0.49 for the average group. In other words, the model predicted half of the average group to be pregnant and the other half to be open after first insemination. The cows predicted to be poor needed on average 138 d to have their first service given, whereas this was 112 d for the cows predicted to be good. Although imperfect efficiency of heat detection could partly explain this, negative energy balance may be the most common cause. Butler (2003) indicated that negative energy balance suppresses the pulsatility of LH and reduces the responsiveness of the ovary to LH stimulation. Further, during a period of negative energy balance, plasma glucose, insulin, and IGF-I are reduced (Spicer et al., 1993), which consequently shifts postpartum ovarian activity and strongly affects the resumption of the ovarian cycles (Senatore et al., 1996). Leroy et al. (2008) also reported an inferior oocyte quality in cows with negative energy balance. Importantly, our finding confirms that the model worked as expected to classify cows of good and poor fertility but applied only to first insemination and not to any insemination, as presented in Shahinfar et al. (2014) and Hempstalk et al. (2015). Further work is required to develop models that can predict the likelihood of conception for any insemination.

With the average accuracy (i.e., AUC) obtained through random cross-validation and herd-by-herd external validation of 0.83 and 0.75, respectively, the model could be used to rank animals in a herd into high versus low likelihood of conception to first ser-

vice groups. This ranking could further be refined by combining with other information—for example, serum metabolic profiles derived using the equations of Luke et al. (2019) and breeding values. Subsequently, farmers may use this information to decide which semen type to give to those groups of cows or whether any other management actions are needed. Moreover, the model might also be used to generate a large number of fertility traits for cows that have MIR records. The MIR-predicted fertility phenotypes could be used for genomic analyses (Gengler et al., 2018). Finally, for application of this research, one way to share models from PLS-DA is through executable files in which the parameters have been embedded. This is often done because the number of parameters of a PLS-DA model is large (e.g., 546 for model 7).

CONCLUSIONS

In this study, we have shown that MIR spectroscopy of milk samples collected in early lactation together with other on-farm data could be used to classify cows that conceived at first insemination and did not conceive within the breeding season with reasonably good accuracy. The calibration models were externally validated with reliable results. Such information could be useful in decision support tools to help farmers optimize their breeding decisions. The model might also be used to generate, on a large scale, fertility phenotypes for genomic evaluation.

ACKNOWLEDGMENTS

The authors thank DairyBio, jointly funded by Dairy Australia (Melbourne, Australia) and Agriculture Victoria (Melbourne, Australia), for funding this project. This work was also part of the project “MIRforProfit: Integrating very large genomic and milk mid-infrared data to improve profitability of dairy cows,” funded by the Australian Government Department of Agriculture as part of the Rural R&D for Profit program. The staff of DataGene (Bundoora, Victoria, Australia) are gratefully acknowledged for providing the data used in this study. Special thanks are given to Peter Nish from TasHerd Pty Ltd. (Hadspen, Tasmania, Australia) for providing MIR spectral data and Hico Pty Ltd. (Maffra, Victoria, Australia) for collecting milk samples and the farmers whose data were used.

REFERENCES

Alawneh, J. I., R. A. Laven, and M. A. Stevenson. 2011. The effect of lameness on the fertility of dairy cattle in a seasonally breeding pasture-based system. *J. Dairy Sci.* 94:5487–5493.

- Armstrong, D. P., K. A. Tarrant, C. K. M. Ho, L. R. Malcolm, and W. J. Wales. 2010. Evaluating development options for a rain-fed dairy farm in Gippsland. *Anim. Prod. Sci.* 50:363–370.
- Belay, T. K., B. S. Dagnachew, Z. M. Kowalski, and T. Ådnøy. 2017. An attempt at predicting blood β -hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle. *J. Dairy Sci.* 100:6312–6326.
- Berry, D. P., R. D. Evans, and S. McParland. 2011. Evaluation of bull fertility in dairy and beef cattle using cow field data. *Theriogenology* 75:172–181.
- Berry, D. P., E. Wall, and J. E. Pryce. 2014. Genetics and genomics of reproductive performance in dairy and beef cattle. *Animal* 8(Suppl. 1):105–121.
- Butler, W. R. 2003. Energy balance relationships with follicular development, ovulation and fertility in postpartum dairy cows. *Livest. Prod. Sci.* 83:211–218.
- Ceciliani, F., C. Lecchi, C. Urh, and H. Sauerwein. 2018. Proteomics and metabolomics characterizing the pathophysiology of adaptive reactions to the metabolic challenges during the transition from late pregnancy to early lactation in dairy cows. *J. Proteomics* 178:92–106.
- Chollet, F., and J. J. Allaire. 2018. *Deep Learning with R*. Manning Publications, Shelter Island, NY.
- Clay, J. S., B. T. McDaniel, and C. H. Brown. 2004. Variances of and correlations among progeny tests for reproductive traits of cows sired by AI bulls. *J. Dairy Sci.* 87:2307–2313.
- Cornwell, J. M., M. L. McGilliard, R. Kasimanickam, and R. L. Nebel. 2006. Effect of sire fertility and timing of artificial insemination in a Presynch + Ovsynch protocol on first-service pregnancy rates. *J. Dairy Sci.* 89:2473–2478.
- Dairy Australia. 2016a. Australia's 5 main feeding systems. Accessed Apr. 20, 2019. <http://www.dairyaustralia.com.au/~media/Documents/Animal%20management/Feed%20and%20nutrition/Feeding%20Systems%20latest/Aus%20five%20main%20feeding%20systems.pdf>.
- Dairy Australia. 2016b. The InCalf Fertility Data Project 2011. Accessed Apr. 20, 2019. <http://www.dairyaustralia.com.au/Animal-management/Fertility/About-InCalf.aspx>.
- 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.
- DeJarnette, J. M., C. E. Marshall, R. W. Lenz, D. R. Monke, W. H. Ayars, and C. G. Sattler. 2004. Sustaining the fertility of artificially inseminated dairy cattle: The role of the artificial insemination industry. *J. Dairy Sci.* 87:E93–E104.
- Gengler, N., A. Mineur, S. Vanderick, H. Hammami, and G. Consortium. 2018. Milk mid-infrared spectra based biomarkers contributing to genetic improvement for udder health, fertility and longevity. *ICAR Tech. Ser.* 23:221–227.
- Gengler, N., H. Soyeurt, F. Dehareng, C. Bastin, F. Colinet, H. Hammami, M. L. Vanrobays, A. Lainé, S. Vanderick, C. Grelet, A. Vanlierde, E. Froidmont, and P. Dardenne. 2016. Capitalizing on fine milk composition for breeding and management of dairy cows. *J. Dairy Sci.* 99:4071–4079.
- Goldansaz, S. A., A. C. Guo, T. Sajed, M. A. Steele, G. S. Plastow, and D. S. Wishart. 2017. Livestock metabolomics and the livestock metabolome: A systematic review. *PLoS One* 12:e0177675.
- Gottardo, P., M. Penasa, N. Lopez-Villalobos, and M. De Marchi. 2016. Variable selection procedures before partial least squares regression enhance the accuracy of milk fatty acid composition predicted by mid-infrared spectroscopy. *J. Dairy Sci.* 99:7782–7790.
- Gottardo, P., M. Penasa, F. Righi, N. Lopez-Villalobos, M. Cassandro, and M. De Marchi. 2017. Fatty acid composition of milk from Holstein-Friesian, Brown Swiss, Simmental and Alpine Grey cows predicted by mid-infrared spectroscopy. *Ital. J. Anim. Sci.* 16:380–389.
- Greenwood, S. L., and M. C. Honan. 2019. Symposium review: Characterization of the bovine milk protein profile using proteomic techniques. *J. Dairy Sci.* 102:2796–2806.

- 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.
- Grelet, C., A. Vanlierde, M. Hostens, L. Foldager, M. Salavati, K. L. Ingvarsen, M. Crowe, M. T. Sorensen, E. Froidmont, C. P. Ferris, C. Marchitelli, F. Becker, T. Larsen, F. Carter, and F. Dehareng. 2018. Potential of milk mid-IR spectra to predict metabolic status of cows through blood components and an innovative clustering approach. *Animal* 13:649–658.
- Gromski, P. S., H. Muhamadali, D. I. Ellis, Y. Xu, E. Correa, M. L. Turner, and R. Goodacre. 2015. A tutorial review: Metabolomics and partial least squares-discriminant analysis—A marriage of convenience or a shotgun wedding. *Anal. Chim. Acta* 879:10–23.
- Grzesiak, W., D. Zaborski, P. Sablik, A. Żukiewicz, A. Dybus, and I. Szatkowska. 2010. Detection of cows with insemination problems using selected classification models. *Comput. Electron. Agric.* 74:265–273.
- Haile-Mariam, M., P. J. Bowman, and J. E. Pryce. 2013. Genetic analyses of fertility and predictor traits in Holstein herds with low and high mean calving intervals and in Jersey herds. *J. Dairy Sci.* 96:655–667.
- Haile-Mariam, M., J. M. Morton, and M. E. Goddard. 2003. Estimates of genetic parameters for fertility traits of Australian Holstein-Friesian cattle. *Anim. Sci.* 76:35–42.
- Hempstalk, K., S. McParland, and D. P. Berry. 2015. Machine learning algorithms for the prediction of conception success to a given insemination in lactating dairy cows. *J. Dairy Sci.* 98:5262–5273.
- Herlihy, M. M., D. P. Berry, M. A. Crowe, M. G. Diskin, and S. T. Butler. 2011. Evaluation of protocols to synchronize estrus and ovulation in seasonal calving pasture-based dairy production systems. *J. Dairy Sci.* 94:4488–4501.
- Hewavitharana, A. K., and B. van Brakel. 1997. Fourier transform infrared spectrometric method for the rapid determination of casein in raw milk. *Analyst (Lond.)* 122:701–704.
- Ho, P. N., L. C. Maret, B. Wales, M. Axford, E. M. Oakes, and J. E. Pryce. 2019. Predicting milk fatty acids and energy balance of dairy cows in Australia using milk mid-infrared spectroscopy. *Anim. Prod. Sci.* <https://doi.org/10.1071/AN18532>.
- Jenkins, T. C., and M. A. McGuire. 2006. Major advances in nutrition: Impact on milk composition. *J. Dairy Sci.* 89:1302–1310.
- Kaniyamattam, K., M. A. Elzo, J. B. Cole, and A. De Vries. 2016. Stochastic dynamic simulation modeling including multitrait genetics to estimate genetic, technical, and financial consequences of dairy farm reproduction and selection strategies. *J. Dairy Sci.* 99:8187–8202.
- Kohler, A., M. Zimonja, V. Segtnan, and H. Martens. 2009. Standard normal variate, multiplicative signal correction and extended multiplicative signal correction preprocessing in biospectroscopy. Pages 139–162 in *Comprehensive Chemometrics*. S. D. Brown, R. Tauler, and B. Walczak, ed. Elsevier, Oxford, UK.
- Lê Cao, K.-A., S. Boitard, and P. Besse. 2011. Sparse PLS discriminant analysis: Biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics* 12:253.
- Leroy, J. L., G. Opsomer, A. Van Soom, I. G. F. Goovaerts, and P. E. J. Bols. 2008. Reduced fertility in high-yielding dairy cows: Are the oocyte and embryo in danger? Part I. The importance of negative energy balance and altered corpus luteum function to the reduction of oocyte and embryo quality in high-yielding dairy cows. *Reprod. Domest. Anim.* 43:612–622.
- Leroy, J. L., J. De Bie, L. Jordaens, K. Desmet, A. Smits, W. F. Marei, P. E. Bols, and V. Van Hoek. 2017. Negative energy balance and metabolic stress in relation to oocyte and embryo quality: An update on possible pathways reducing fertility in dairy cows. *Anim. Reprod.* 14:497–506.
- Liu, Z., J. Jaitner, F. Reinhardt, E. Pasma, S. Rensing, and R. Reents. 2008. Genetic evaluation of fertility traits of dairy cattle using a multiple-trait animal model. *J. Dairy Sci.* 91:4333–4343.
- Luke, T. D. W., S. Rochford, W. J. Wales, V. Bonfatti, L. Maret, and J. E. Pryce. 2019. Metabolic profiling of early-lactation dairy cows using milk mid-infrared spectra. *J. Dairy Sci.* 102:1747–1760.
- 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.
- McParland, S., E. Lewis, E. Kennedy, S. G. Moore, B. McCarthy, M. O'Donovan, S. T. Butler, J. E. Pryce, and D. P. Berry. 2014. Mid-infrared spectrometry of milk as a predictor of energy intake and efficiency in lactating dairy cows. *J. Dairy Sci.* 97:5863–5871.
- Mineur, A. 2017. Use of MIR spectral data of milk in the detection and prevention of lameness in dairy cows. MS Thesis. The University of Liège, Belgium.
- Morton, J. M., M. J. Auldist, M. L. Douglas, and K. L. Macmillan. 2017. Milk protein concentration, estimated breeding value for fertility, and reproductive performance in lactating dairy cows. *J. Dairy Sci.* 100:5850–5862.
- Morton, J. M., W. P. Tranter, D. G. Mayer, and N. N. Jonsson. 2007. Effects of environmental heat on conception rates in lactating dairy cows: Critical periods of exposure. *J. Dairy Sci.* 90:2271–2278.
- Oikonomou, G., G. Arsenos, G. E. Valergakis, A. Tsiaras, D. Zygouannis, and G. Banos. 2008. Genetic relationship of body energy and blood metabolites with reproduction in Holstein cows. *J. Dairy Sci.* 91:4323–4332.
- Olsen, R. L. 2017. Subsetting methods for balanced cross-validation, time series windowing, and general grouping and splitting of data. Accessed Dec. 17, 2018. <https://github.com/LudvigOlsen/groupdata2>.
- Pralle, R. S., K. W. Weigel, and H. M. White. 2018. Predicting blood β -hydroxybutyrate using milk Fourier transform infrared spectrum, milk composition, and producer-reported variables with multiple linear regression, partial least squares regression, and artificial neural network. *J. Dairy Sci.* 101:4378–4387.
- Preacher, K. J., D. D. Rucker, R. C. MacCallum, and W. A. Nicewander. 2005. Use of the extreme groups approach: A critical reexamination and new recommendations. *Psychol. Methods* 10:178–192.
- Pryce, J. E., K. L. Parker Gaddis, A. Koeck, C. Bastin, M. Abdelsayed, N. Gengler, F. Miglior, B. Heringstad, C. Egger-Danner, K. F. Stock, A. J. Bradley, and J. B. Cole. 2016. Invited review: Opportunities for genetic improvement of metabolic diseases. *J. Dairy Sci.* 99:6855–6873.
- R Development Core Team. 2018. The GNU Project. The R Project for Statistical Computing. Accessed Nov. 4, 2018. <http://www.rproject.org/>.
- Rearte, R., S. J. LeBlanc, S. G. Corva, R. L. de la Sota, I. M. Lacau-Mengido, and M. J. Giuliadori. 2018. Effect of milk production on reproductive performance in dairy herds. *J. Dairy Sci.* 101:7575–7584.
- Ribeiro, E. S., F. S. Lima, L. F. Greco, R. S. Bisinotto, A. P. A. Monteiro, M. Favoreto, H. Ayres, R. S. Marsola, N. Martinez, W. W. Thatcher, and J. E. P. Santos. 2013. Prevalence of periparturient diseases and effects on fertility of seasonally calving grazing dairy cows supplemented with concentrates. *J. Dairy Sci.* 96:5682–5697.
- Roche, J. R., K. A. Macdonald, C. R. Burke, J. M. Lee, and D. P. Berry. 2007. Associations among body condition score, body weight, and reproductive performance in seasonal-calving dairy cattle. *J. Dairy Sci.* 90:376–391.
- Rodney, R. M., J. K. Hall, C. T. Westwood, P. Celi, and I. J. Lean. 2016. Precalving and early lactation factors that predict milk casein and fertility in the transition dairy cow. *J. Dairy Sci.* 99:7554–7567.
- Savitzky, A., and M. J. Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36:1627–1639.
- Senatore, E., W. Butler, and P. Oltenacu. 1996. Relationships between energy balance and post-partum ovarian activity and fertility in first lactation dairy cows. *Anim. Sci.* 62:17–23.
- Shahinfar, S., D. Page, J. Guenther, V. Cabrera, P. Fricke, and K. Weigel. 2014. Prediction of insemination outcomes in Holstein dairy cattle using alternative machine learning algorithms. *J. Dairy Sci.* 97:731–742.

- Shalloo, L., A. Cromie, and N. McHugh. 2014. Effect of fertility on the economics of pasture-based dairy systems. *Animal* 8(Suppl. 1):222–231.
- Shenk, J. S., and M. O. Westerhaus. 1995. Forage analysis by near infrared spectroscopy. Pages 111–120 in *Forages*. Vol. II. The Science of Grassland Agriculture. 5th ed. R. F. Barnes, D. A. Miller, and C. J. Nelson, ed. Iowa State University Press, Ames.
- Shorten, P. R., C. Morris, and N. Cullen. 2015. The effects of age, weight, and sire on pregnancy rate in cattle. *J. Anim. Sci.* 93:1535–1545.
- Šimundić, A.-M. 2009. Measures of diagnostic accuracy: Basic definitions. *EJIFCC* 19:203–211.
- Soyeurt, H., F. Dehareng, N. Gengler, S. McParland, E. Wall, D. P. Berry, M. 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.
- Spicer, L. J., E. Alpizar, and S. Echtenkamp. 1993. Effects of insulin, insulin-like growth factor I, and gonadotropins on bovine granulosa cell proliferation, progesterone production, estradiol production, and (or) insulin-like growth factor I production in vitro. *J. Anim. Sci.* 71:1232–1241.
- Toledo-Alvarado, H., A. I. Vazquez, G. de los Campos, R. J. Tempelman, G. Bittante, and A. Cecchinato. 2018a. Diagnosing pregnancy status using infrared spectra and milk composition in dairy cows. *J. Dairy Sci.* 101:2496–2505.
- Toledo-Alvarado, H., A. I. Vazquez, G. de los Campos, R. J. Tempelman, G. Gabai, A. Cecchinato, and G. Bittante. 2018b. Changes in milk characteristics and fatty acid profile during the estrous cycle in dairy cows. *J. Dairy Sci.* 101:9135–9153.
- Tomassini, A., G. Curone, M. Solè, G. Capuani, F. Sciubba, G. Conata, A. Miccheli, and D. Vigo. 2019. NMR-based metabolomics to evaluate the milk composition from Friesian and autochthonous cows of Northern Italy at different lactation times. *Nat. Prod. Res.* 33:1085–1091.
- van Gastelen, S., H. Mollenhorst, E. C. Antunes-Fernandes, K. A. Hetinga, G. G. van Burgsteden, J. Dijkstra, and J. L. W. Rademaker. 2018. Predicting enteric methane emission of dairy cows with milk Fourier-transform infrared spectra and gas chromatography-based milk fatty acid profiles. *J. Dairy Sci.* 101:5582–5598.
- Williams, P. 2004. *Near-Infrared Technology: Getting the Best Out of Light: A Short Course in the Practical Implementation of Near-Infrared Spectroscopy for the User*. PDK Projects Inc., Nanaimo, BC, Canada.
- Xu, W., J. Vervoort, E. Saccenti, R. van Hoeij, B. Kemp, and A. van Knegsel. 2018. Milk metabolomics data reveal the energy balance of individual dairy cows in early lactation. *Sci. Rep.* 8:15828.
- Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. E. Goddard, and P. M. Visscher. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42:565–569.