Prediction of bovine milk technological traits from mid-infrared spectroscopy analysis in dairy cows

G. Visentin,*† A. McDermott,*† S. McParland,*[1] D. P. Berry,* O. A. Kenny,‡ A. Brodkorb,‡ M. A. Fenelon,‡ and M. De Marchi†

*Animal & Grassland Research and Innovation Center, Teagasc, Moorepark, Fermoy, Co. Cork, Ireland
†Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy
‡Teagasc Food Research Center, Moorepark, Fermoy, Co. Cork, Ireland

ABSTRACT

Rapid, cost-effective monitoring of milk technological traits is a significant challenge for dairy industries specialized in cheese manufacturing. The objective of the present study was to investigate the ability of mid-infrared spectroscopy to predict rennet coagulation time, curd-firming time, curd firmness at 30 and 60 min after rennet addition, heat coagulation time, casein micelle size, and pH in cow milk samples, and to quantify associations between these milk technological traits and conventional milk quality traits. Samples (n = 713) were collected from 605 cows from multiple herds; the samples represented multiple breeds, stages of lactation, parities, and milking times. Reference analyses were undertaken in accordance with standardized methods, and mid-infrared spectra in the range of 900 to 5,000 cm^{-1} were available for all samples. Prediction models were developed using partial least squares regression, and prediction accuracy was based on both cross and external validation. The proportion of variance explained by the prediction models in external validation was greatest for pH (71%), followed by rennet coagulation time (55%) and milk heat coagulation time (46%). Models to predict curd firmness 60 min from rennet addition and casein micelle size, however, were poor, explaining only 25 and 13%, respectively, of the total variance in each trait within external validation. On average, all prediction models tended to be unbiased. The linear regression coefficient of the reference value on the predicted value varied from 0.17 (casein micelle size regression model) to 0.83 (pH regression model) but all differed from 1. The ratio performance deviation of 1.07 (casein micelle size prediction model) to 1.79 (pH prediction model) for all prediction models in the external validation was <2, suggesting that none of the prediction models could be used for analytical purposes. With the exception of casein micelle size and curd firmness at 60 min after rennet addition, the developed prediction models may be useful as a screening method, because the concordance correlation coefficient ranged from 0.63 (heat coagulation time prediction model) to 0.84 (pH prediction model) in the external validation.
Key words: milk coagulation properties, milk heat stability, casein micelle size, milk acidity, grass

INTRODUCTION

World dairy food production from cow milk has increased consistently by 2% annually (FAOSTAT, 2014). Rapid and low-cost tools to quantify milk quality processability can provide the dairy industry with a more effective monitoring strategy, to aid in segregating milk before manufacturing. Rennet coagulation time (**RCT**), curd-firming time ($\mathbf{k_{20}}$), and curd firmness after 30 or 60 min from rennet addition ($\mathbf{a_{30}}$ and $\mathbf{a_{60}}$, respectively) reflect milk coagulation properties, which are milk characteristics that describe the reactivity of milk after rennet addition. Shorter coagulation time and greater curd-firming capacity are correlated with greater cheese yield and improved sensory properties (Aleandri et al., 1989; Martin et al., 1997; Pretto et al., 2013) and quality (Ng-Kwai-Hang et al., 1989; Malacarne et al., 2014). Other milk technological traits of interest for cheese production include casein micelle size (**CMS**), expressed as the diameter of the colloidal particles (Fox and Brodkorb, 2008), heat coagulation time (**HCT**), defined as the time required to induce milk coagulation at a given temperature (Davies and White, 1966), and both pH and titratable acidity.

Mid-infrared spectroscopy (**MIRS**) is a rapid, low-cost technique that measures the interaction between physical matter and electromagnetic radiation in the region between 900 and 5,000 cm^{-1}. Mid-infrared spectroscopy is currently used routinely to quantify milk components such as protein, fat, casein, and lactose

in individual cow milk aliquots sampled at herd milk testing. Recent research has documented the ability of MIRS to predict novel milk quality traits (Soyeurt et al., 2011; De Marchi et al. 2014) and animal characteristics (McParland et al., 2011; De Marchi et al., 2014). The ability to predict milk technological traits with MIRS could be advantageous, because all individual cow milk samples are subjected to MIRS analysis and thus the marginal cost in predicting these new traits is negligible. Few studies, however, have reported the ability of MIRS to predict milk coagulation properties in dairy cows and these studies have been largely confined to Italian Holstein Friesian (Dal Zotto et al., 2008; De Marchi et al., 2013) and Brown Swiss (De Marchi et al., 2009) cows. To our knowledge, no study has attempted to predict either CMS or HCT from milk MIRS; only De Marchi et al. (2009) has investigated the potential of MIRS to predict milk pH.

The objectives of the present study were (1) to evaluate the accuracy of MIRS in predicting milk coagulation properties, CMS, HCT, and pH using a data set of milk samples from multiple cow breeds fed predominantly a grazed grass diet, and (2) to quantify the associations between these milk technological traits and traditional milk quality traits.

MATERIALS AND METHODS

Data

A total of 713 individual milk samples from 605 cows were collected weekly between August 2013 and August 2014. Samples originated from 7 different Irish research herds and consisted of Holstein-Friesian, Jersey, and Norwegian Red cows, as well as their respective crosses. Milk samples were from different stages of lactation, different parities, and both morning and evening milkings. Cows were on different experimental treatments based on a grazed-grass basal diet but were also, at times, supplemented with concentrate and grass silage. Milk aliquots (50 mL) were stored at 4°C after sampling. Within 48 h of collection, samples were analyzed in the laboratory of Teagasc Animal and Grassland Research and Innovation Center (Moorepark, Fermoy, Co. Cork, Ireland) for milk chemical composition (protein, fat, casein, urea, TS), determined using a MilkoScan FT6000 (Foss Electronic A/S, Hillerød, Denmark). The resulting spectrum, containing 1,060 transmittance data points in the mid-infrared region between 900 and 5,000 cm^{-1}, was stored. All samples were then preserved with Broad Spectrum Microtab II, containing 8 mg of broponol and 0.3 mg of natamycin (D&F Control System Inc., Norwood, MA), and kept at refrigerator

temperature for further analysis. For logistical reasons, not all technological milk quality parameters were quantified on every sample.

Gold-Standard Analysis for Milk Coagulation Properties

Milk coagulation properties were quantified using a Formagraph (Foss Electronic A/S, Hillerød, Denmark) on the preserved individual samples within 5 d of collection. The Formagraph is a mechanical instrument in which an aliquot of milk, following inclusion of raw rennet diluted in distilled water, is brought into contact with an oscillating loop pendulum. Initially, no force is transmitted from the uncoagulated milk to the pendulum. When coagulation begins, the viscosity of the milk increases, resulting in a force applied to the pendulum. The output of the Formagraph is a graph of the value of curd firmness against time. A 10-mL aliquot of milk was warmed to 35°C. Subsequently, 1 mL of raw rennet (Chy-Max Plus, 190 international milk clotting units/mL, Chr. Hansen A/S, Hørsholm, Denmark) was diluted in 20 mL of distilled water (1:20 vol/vol). The level of coagulant added to milk samples was adjusted based on protein concentration of the milk sample (O'Callaghan et al., 1999, 2001) and testing began immediately after rennet addition.

The measured traits were (1) RCT, defined as the time (min) taken from rennet addition to the beginning of coagulation, (2) k_{20}, the time from the gel development to a width of 20 mm in the graph, and (3) a_{30}, measured as the width of the graph 30 min after rennet addition. To facilitate the determination of RCT for noncoagulating milk samples (i.e., samples whose rennet coagulation time was longer than 30 min; Ikonen et al., 2004), the duration of the analysis was extended from 30 to 60 min, as suggested by De Marchi et al. (2013). Thus, curd firmness after 60 min from rennet addition (a_{60}, mm) was determined as the width of the graph at the end of analysis for all samples.

Gold-Standard Analysis for CMS, Heat Stability, and pH

The hydrodynamic diameter of the casein micelles was determined using a Zetasizer Nano system (Malvern Instruments Inc., Worcester, UK). The measurements were carried out at 25°C using the noninvasive backscatter optics at 173°. Samples were analyzed within 24 h of collection and diluted with MilliQ water (Millipore, Billerica, MA) before measurement (approximately 1:30), which is sufficient for dynamic light scattering measurements (attenuator of 4 to 5). The

cumulative method was used to determine the mean particle size (diameter) that corresponded to the mean of the volume distribution of the main peak. The viscosity of the solvent was assumed to be the same as that of water, given the low concentration of protein.

Heat coagulation time was tested using the method outlined by Davies and White (1966) within 48 h of sample collection, and all analyses were undertaken by the same operator. Briefly, an aliquot of 3.4 g of each milk sample was transferred into an individual glass tube suitable for the Elbanton BV (Kerkdriel, the Netherlands) hot oil bath. The oil temperature was set at 140°C with an oscillator speed of 8 rpm. Heat coagulation time was recorded as the time when each sample started to flocculate based on visual assessment. The test time was set to 30 min (Davies and White, 1966); therefore, samples that did not coagulate within this period (n = 15) did not provide any information about their heat susceptibility. Their inclusion in the calibration data set was not feasible, because regression models of reference values on predicted values cannot be performed on observations with missing true values of HCT. Therefore, these samples were classified as heat noncoagulating and discarded from further statistical analysis. Repeatability information of the present methods is reported by Davies and White (1966).

Milk pH was assessed, within 24 h of sample collection, with a SevenCompact pH meter S220 (Mettler Toledo AG, Greifensee, Switzerland).

Development of MIRS Prediction Models

Development of the prediction models was undertaken using SAS software (ver. 9.3; SAS Institute Inc., Cary, NC). Spectral data, expressed in transmittance, were transformed to absorbance by taking the $\log_{10}$ of the reciprocal of the transmittance. The prediction models were developed using only spectra that did not include high-noise-level regions (Hewavitharana and Brakel, 1997), which are part of the spectra related to water absorption (Heuer et al., 2001). Consequently, spectral regions between 1,580 and 1,710 cm^{-1} and between 2,990 and 3,690 cm^{-1} were discarded before chemometric analysis. Principal component analysis (PROC PRIN COMP; SAS Institute Inc.) was performed on the raw spectra, providing principal components that were used to detect similarities and differences among individual spectra and to identify spectral outliers. The first 2 principal components explained 66.38 and 18.78% of the total spectral variation, respectively. Both individual sample Mahalanobis distances and a visual inspection of the plot of the first principal component against the second principal component did not indicate the presence of outliers and therefore no samples were discarded.

The distribution of each trait, as well as the identification of outlier reference values, was determined using PROC UNIVARIATE (SAS Institute Inc.). If the distribution of a trait was non-normal, based on visual inspection and the Shapiro-Wilk test statistic, the reference values were transformed using a natural logarithm transformation. Consequently, HCT and k_{20} were log-transformed. Outliers were defined as reference values >3 standard deviations from the mean. Based on this definition, 1 observation for RCT, 2 observations for log-transformed HCT, 6 observations for log-transformed k_{20} and pH, and 34 observations for CMS were identified as outliers and subsequently discarded. The relatively high number of outlier values for CMS existed mainly for 2 reasons: (1) error in laboratory processing (n = 17), and (2) CMS values deviating more than 3 standard deviations from the mean (n = 17). Further examination of these data points did not indicate any trend in date of measurement and characteristics (e.g., parity, breed, stage of lactation) of the cows that contributed the samples.

Partial least squares (**PLS**) regression (PROC PLS; SAS Institute Inc.) was undertaken to generate the prediction models. Before performing the analysis, the initial data set was divided randomly into 2 different subsets for each trait separately, representing a model calibration data set (80% of the entire data set) and a validation data set (the remaining 20% of the entire data set). The calibration data set was used to generate the prediction models. The validation data set was considered as an independent data set, because the samples were not used to calibrate the prediction models. Developed prediction models were applied to the validation data set to quantify their predictive ability. The mean, standard deviation, and range of the milk technological traits in the validation and calibration data set were similar. Mathematical pretreatments (Savitzky-Golay first and second derivatives) were applied to the raw spectra but no improvement on the prediction models accuracy was detected; therefore, the untreated spectra were used to generate the models.

The optimal number of PLS factors was defined as the minimum number of factors to achieve the lowest root mean predicted residual sum of squares. The goodness-of-fit statistics considered in the present study were the coefficient of determination in the cross-validation (i.e., one-at-a-time cross-validation) and external validation ($\mathbf{R^2_C}$ and $\mathbf{R^2_V}$, respectively) data sets and the standard error of prediction in the cross-validation and external validation ($\mathbf{SEP_C}$ and $\mathbf{SEP_V}$, respectively) data sets. To evaluate the practical utility of the prediction

models, the ratio performance deviation (**RPD**) and concordance correlation coefficient (**CCC**) were calculated. The RPD was calculated as the ratio of the standard deviation of the trait and the standard error of prediction from the prediction model (Williams, 2007). The CCC was calculated as (Lin, 1989):

$$CCC = \frac{2 \times COV\left(\hat{Y}; Y\right)}{\sigma_{\hat{Y}}^2 + \sigma_{Y}^2 + \left(\mu_{\hat{Y}} - \mu_{Y}\right)^2},$$

where $COV(\hat{Y}; Y)$ is the covariance between the reference (Y) and predicted values ($\hat{Y}$), σ_{Y}^2 is the variance of the reference values, $\sigma_{\hat{Y}}^2$ is the variance of the predicted values, and μ_{Y} and $\mu_{\hat{Y}}$ represent the mean of the reference and predicted values, respectively. Bias for each prediction model was calculated as the average of the difference between the reference value and the respective predicted value for each sample; a *t*-test was used to determine if this was significantly different from zero. Reference values for each prediction model were also linearly regressed on the predicted values to obtain the linear regression coefficient; a *t*-test was used to determine if the linear regression coefficient differed from 1.

Pearson correlations were calculated between both reference and predicted values for the technological traits and other measures of milk quality traits (i.e., protein, casein, fat, total solids, and urea concentration).

RESULTS

Summary Statistics of Reference Traits

The DIM when milk samples were taken varied from 5 to 375; 19.27% of milk samples (n = 138) were from the first 60 DIM and 50% (i.e., n = 356) were from between 61 and 200 DIM. Milk samples were available from cows in parities 1 to 11. Morning milking samples represented almost 70% of the total samples collected (n = 497). A total of 146 samples were from 2 herds with autumn-calving cows, whose diet was partially grass-silage based; the remainder of the data originated from 6 herds with spring-calving cows that were fed predominantly grazed grass.

The most represented breed was Holstein-Friesian (n = 443), followed by Jersey (n = 109) and Norwegian Red (n = 17); 144 samples were from crossbreed cows. Noncoagulating milk samples represented 15.71% of the total data set (n = 88), of which the majority (n = 62) were from Holstein-Friesian cows.

Descriptive statistics of milk technological traits are in Table 1. The coefficient of variation of all traits, with the exception of pH, varied from 18.13 (CMS) to 70.76% (HCT); the coefficient of variation for pH was only 1.18%. The Shapiro-Wilk value for both k_{20} and HCT suggested that neither trait was normally distributed but, following transformation with the natural logarithm, the Shapiro-Wilk value increased from 0.87 to 0.97 for k_{20}, and from 0.87 to 0.98 for HCT. This transformation reduced the coefficient of variation from 64.26 to 37.50% for k_{20} and transformed k_{20}, respectively, and from 70.76 to 35.59% for HCT and transformed HCT, respectively.

Prediction Model Accuracy

The goodness-of-fit statistics for MIRS prediction models are summarized in Table 2. The number of factors included in the prediction models varied from 14 (HCT and CMS) to 17 (pH), with the exception of the a_{60} prediction model, in which 7 factors were used. The most accurate prediction model was for pH (R^2_C and SEP_C of 0.73 and 0.06, respectively), followed by RCT (R^2_C and SEP_C of 0.61 and 5.64 min, respectively). The accuracy of predicting a_{60} and CMS was poor, with respective R^2_C of 0.26 and 0.23, and SEP_C of 10.33 mm

Table 1. Number of samples (n), mean, SD, range, CV, and Shapiro-Wilk value (W) for normality distribution test of reference values for rennet coagulation time (RCT), curd-firming time (k_{20}), curd firmness at 30 and 60 min (a_{30}, a_{60}), heat coagulation time (HCT), casein micelle size (CMS), and pH

Trait	n	Mean	SD	Range	CV	W
RCT, min	560	20.70	8.93	47.25	43.15	0.98
k_{20}, min	523	6.12	3.93	18.50	64.26	0.87
k_{20}, $\log_e$ min	523	1.63	0.61	2.76	37.50	0.97
a_{30}, mm	467	32.17	15.61	72.88	48.54	0.98
a_{60}, mm	559	31.54	12.23	70.68	38.79	0.95
HCT, min	492	8.75	6.19	27.02	70.76	0.87
HCT, $\log_e$ min	492	1.94	0.69	3.16	35.59	0.98
CMS, nm	654	175.36	31.79	231.08	18.13	0.98
pH	702	6.68	0.11	0.66	1.63	0.99

Table 2. Fitting statistics[1] of prediction models in cross and external validation for rennet coagulation time (RCT), curd-firming time (k_{20}), curd firmness at 30 and 60 min (a_{30}, a_{60}), heat coagulation time (HCT), casein micelle size (CMS), and pH

Trait[2]	Cross validation						External validation						
	n	#L	SEP_C	R^2_C	RPD	CCC	n	Bias	Slope (SE)	SEP_V	R^2_V	RPD	CCC
RCT, min	450	16	5.64	0.61	1.59	0.76	110	−0.10	0.55 (0.05)	5.85	0.55	1.49	0.71
k_{20}, $\log_e$ min	414	15	0.39	0.59	1.56	0.74	109	−0.08	0.52 (0.05)	0.42	0.51	1.43	0.67
a_{30}, mm	378	15	11.32	0.50	1.41	0.66	89	−0.85	0.53 (0.06)	10.38	0.46	1.35	0.66
a_{60}, mm	458	7	10.33	0.26	1.16	0.41	101	1.23	0.23 (0.04)	11.49	0.25	1.15	0.37
HCT, $\log_e$ min	389	14	0.46	0.55	1.48	0.71	103	0.04	0.47 (0.05)	0.51	0.46	1.36	0.63
CMS, nm	535	14	28.11	0.23	1.14	0.37	119	0.75	0.17 (0.04)	28.75	0.13	1.07	0.28
pH	553	17	0.06	0.73	1.92	0.84	149	0.00	0.83 (0.04)	0.06	0.71	1.79	0.84

[1]n = number of samples; #L = number of model factors; SEP_C = standard error of prediction in cross validation; R^2_C = coefficient of determination in cross validation; RPD = ratio performance deviation; CCC = concordance correlation coefficient; Bias = average difference between the reference value and the respective predicted value; Slope = linear regression coefficient of reference values on predicted reference values; SEP_V = standard error of prediction in external validation; R^2_V = coefficient of determination in external validation.

[2]Prediction models and results for curd-firming time and heat coagulation time are natural logarithm values.

and 28.11 nm. The RPD values in external validation ranged from 1.07 (CMS) to 1.79 (pH), whereas the CCC in the external validation data set was between 0.28 (CMS) and 0.84 (pH). The mean bias of prediction (i.e., the average of the difference between the gold standard and predicted values for each individual sample) was not different ($P > 0.05$) from zero in external validation. The slope of the linear regression of the reference on the predicted values of each technological trait in the external validation varied from 0.17 (CMS) to 0.83 (pH); all were different from unity ($P < 0.05$).

Model predictive ability was also evaluated on back-transformed values for both HCT and CMS. The model fit statistics on the back-transformed predicted HCT and CMS were less accurate than the fit statistics using natural logarithm-transformed HCT and CMS. For example, the coefficient of determination in external validation was 0.29 for HCT and 0.43 for k_{20}, whereas the SEP_V was 3.01 min for k_{20} and 5.64 for HCT. The slope of the linear regression coefficient of the original (i.e., untransformed) reference values on the back-transformed predicted values for both models was different from unity ($P < 0.05$), whereas the bias was different from zero only for HCT (−1.31 min). Neither prediction model was useful for analytic purposes (RPD of 1.17 and 1.32 for HCT and k_{20}, respectively), and the models were characterized as having moderate predicting ability (CCC of 0.49 and 0.59 for HCT and k_{20}, respectively).

Correlations with Milk Composition Traits

Pearson correlations between technological traits and milk composition traits are in Table 3. Correlations between the reference technological traits and milk chemical composition traits were similar to those between the predicted traits and milk chemical com-

position traits. In general, the greater the accuracy of the MIRS prediction model, the more comparable the correlations with the milk composition traits for either the reference or predicted technological traits. Rennet coagulation time had a strong positive correlation with k_{20} (correlations of 0.77 and 0.86 for the reference and MIRS-predicted values, respectively) and was strongly negatively correlated (−0.73 to −0.82 for the reference and MIRS-predicted values, respectively) with a_{30}. Rennet coagulation time was associated with milk composition, particularly with protein and casein concentrations. Correlations between RCT and protein were −0.46 (reference RCT and protein percentage) and −0.50 (MIRS-predicted RCT and protein percentage), whereas correlations between RCT and casein were −0.44 (reference RCT and casein percentage) and −0.49 (MIRS-predicted RCT and casein percentage). Milk coagulation properties were weakly correlated with other milk technological traits, with the exception of RCT and pH (correlations between 0.63 and 0.74 for either the reference or MIRS-predicted values). Casein micelle size was not associated with milk composition; correlations ranged from −0.01 to 0.08 and were not different from zero ($P > 0.05$). A weak correlation existed between reference CMS and reference RCT (0.23). Predicted CMS was weakly correlated with predicted a_{60} (0.31), predicted pH (−0.07), and milk protein (0.20) and casein (0.15) levels. Heat coagulation time had the strongest correlations with milk urea (0.48) and protein (0.22) concentration. Heat coagulation time was also weakly associated with milk coagulation properties, with correlations ranging from −0.22 (k_{20}) to 0.34 (a_{30}).

DISCUSSION

Fundamental to the successful development of an accurate prediction model is the incorporation of ex-

Table 3. Pearson correlations[1] between reference (below diagonal) and mid-infrared spectroscopy (MIRS)-predicted (above diagonal) rennet coagulation time (RCT), curd-firming time (k_{20}), curd firmness at 30 and 60 min (a_{30}, a_{60}), heat coagulation time (HCT), casein micelle size (CMS), pH, and MIRS-predicted milk protein (PRT), fat, TS, urea, and casein concentrations

	RCT	k_{20}	a_{30}	a_{60}	HCT	CMS	pH	PRT	Fat	TS	Urea	Casein
RCT	—	0.86	−0.82	−0.40	−0.32	0.09	0.74	−0.50	−0.46	−0.42	−0.49	−0.49
k_{20}	0.77	—	−0.90	−0.55	−0.38	−0.05	0.67	−0.64	−0.47	−0.46	−0.51	−0.65
a_{30}	−0.73	−0.79	—	0.62	0.46	0.09	−0.60	0.62	0.38	0.37	0.50	0.62
a_{60}	−0.29	−0.38	0.50	—	0.65	0.31	−0.19	0.60	0.01	0.04	0.55	0.58
HCT	−0.21	−0.22	0.34	0.28	—	0.02	−0.12	0.29	0.00	−0.02	0.55	0.23
CMS	0.23	0.21	−0.07	−0.01	−0.11	—	−0.07	0.20	0.02	0.04	−0.02	0.15
pH	0.63	0.51	−0.39	−0.09	−0.09	0.04	—	−0.41	−0.49	−0.47	−0.33	−0.42
PRT	−0.46	−0.54	0.52	0.36	0.22	0.08	−0.34	—	0.42	0.54	0.42	0.90
Fat	−0.36	−0.35	0.27	0.01	0.01	0.01	−0.43	0.42	—	0.93	0.12	0.44
TS	−0.33	−0.35	0.28	0.05	−0.01	0.02	−0.40	0.54	0.93	—	0.08	0.60
Urea	−0.38	−0.37	0.38	0.27	0.48	−0.01	−0.31	0.41	0.11	0.07	—	0.40
Casein	−0.44	−0.56	0.53	0.37	0.18	0.05	−0.36	0.90	0.44	0.60	0.39	—

[1]Correlations <|0.11| were not different from zero ($P > 0.05$). Correlations within both curd-firming time and HCT are based on back-transformed values.

tensive variability in the calibration data set, representative of the population in which the model will be used (De Marchi et al., 2014). In the present study, this was achieved. Comparison of the results of the present study with previous international studies is, nonetheless, difficult because of differences in production system as well as potential differences in methodologies of analysis (i.e., different gold-standard methods). Previously documented coefficients of variation for milk coagulation properties in international populations of between 37% (RCT) and 45% (a_{30}; Ikonen et al., 2004; Cassandro et al. 2008) were lower than observed in the multi-breed population of cows used in the present study. Moreover, compared with the present study, less variation in milk coagulation properties was detected by O'Brien et al. (1999) over a 12-mo study period of Irish herd bulk milk samples. Few studies (Auldist et al., 2004; De Marchi et al., 2013) have considered a_{60} in their analysis because, during cheese manufacture, curd is normally cut after 30 min from rennet addition (Cassandro et al., 2008). Auldist et al. (2004), studying coagulation ability of New Zealand Holstein and Jersey cows, and De Marchi et al. (2013), studying coagulation ability of Italian Holstein-Friesian cows, reported a coefficient of variation for a_{60} of 10 and 62%, respectively. The variability in casein micelle size and pH documented by Chen et al. (2014) using milk samples from 550 lactating Holstein cows in the UK was low but comparable to that observed in the present study. On average, HCT values of the present study were consistent with those of Holt et al. (1978) based on bulk milk samples over a period of 15 mo in southwest Scotland; moreover, as expected, results from our study had greater variability for this trait. Irish Holstein-Friesian was the breed with the greatest prevalence of noncoagulating milk samples (13.99%). Similar proportions were documented by Ikonen et

al. (2004) on Finnish Ayrshire and by Poulsen et al. (2013) on Swedish Red, although the latter authors defined noncoagulating milk samples as those that did not coagulate within 60 min. The high prevalence of noncoagulating milk samples is a serious concern as it negatively affects dairy industry profitability.

Practical Utility of the Prediction Model

For a prediction model to be considered useful for analytic purposes, an RPD value >2 is desired (Williams, 2007). No prediction model in the present study had an RPD in external validation that surpassed this threshold. Nonetheless, the CCC of the prediction equations in external validation demonstrated that some models could at least be used for screening purposes. A CCC of between 0.21 and 0.40 indicates fair predictive ability, between 0.41 and 0.60 indicates moderate predictive ability, between 0.61 and 0.80 indicates substantial predictive ability, and between 0.81 and 1.00 indicates accurate predictive ability (Lin, 1989; De Marchi et al., 2014). Of the traits investigated in the present study, the pH prediction model (CCC = 0.84) had accurate predictive ability, whereas the prediction models for RCT, transformed k_{20}, a_{30}, and transformed HCT were characterized by substantial predictive ability (CCC of 0.71, 0.67, 0.66, and 0.63 for RCT, k_{20}, a_{30}, and HCT, respectively). Poor predictive ability existed for CMS and a_{60}. No prediction model in the present study, on average, significantly over- or underestimated the reference values for any of the traits investigated. Although not necessarily important for breeding purposes, significant bias in estimation could affect milk price if a milk pricing system was based on milk technological traits. To our knowledge, this is the first study that has attempted to quantify this bias in coagulation traits as well as acidity traits. The less-than-unity linear regres-

sion coefficient of the reference value on the predicted value suggested a rescaling in variance between the reference and predicted values, and therefore the extent of the difference between reference samples is not fully reflected in differences in their predicted value. This could have implications for breeding programs, for example, where the true variance in milk coagulation properties may actually be underestimated using predicted rather than reference values; thus, the breeding goal trait should be the gold-standard trait with its respective variance and the mid-infrared predictor trait included in the selection criterion. Nonetheless, if the number of records in the calibration data set used to develop the prediction models, as well as the variability in the reference values, could be increased, then the prediction accuracy of the models might also improve.

Prediction of Rennet Coagulation Traits

Although previous studies have highlighted the effectiveness of MIRS in predicting milk coagulation properties (Dal Zotto et al., 2008; De Marchi et al. 2009), only De Marchi et al. (2013) included noncoagulating milk samples in the calibration data set and used the same instrument (i.e., Formagraph) as used in the present study. The identification of noncoagulating samples could be an extremely useful tool because extended RCT has repercussions on the profitability and efficiency of the dairy industry. Indeed, milk characterized by poor reactivity to addition of rennet increases the time of the entire cheese-making process. For instance, the implementation of MIRS prediction models at the dairy processor level could be advantageous to segregate milk for use in either cheese production (good reactivity to rennet addition, i.e., short RCT) or fluid consumption (poor reactivity to rennet addition, i.e., long RCT).

To our knowledge, no study has investigated the ability of MIRS to predict milk coagulation properties in a grazing production system, such as exists in Ireland, New Zealand, Southern Australia, and elsewhere. Nevertheless, comparison with other studies is difficult because of the different reference procedures adopted. In fact, even if the reference instrument for measuring milk coagulation properties was the Formagraph, as was the case in both the present study and that of De Marchi et al. (2013), the methodology used in the present study adjusted the level of added rennet based on the protein concentration of each individual milk sample. Therefore, the amount of coagulant was not the same for samples with different protein levels. In contrast, De Marchi et al. (2013) kept the quantity of rennet at a constant level (200 μL/10 mL of milk sample) when determining milk coagulation properties. Moreover, previous research undertaken in Italian herds

of Holstein and Brown Swiss cows (Dal Zotto et al., 2008; De Marchi et al., 2009) used the Computerized Renneting Meter (Polo Trade, Monselice, Italy) as the reference analysis instrument. As described by Pretto et al. (2011), differences in gold-standard methodologies have to be considered when comparing the results for milk coagulation properties, because different instruments and methodologies can yield different milk coagulation properties values on the same samples.

Overall, prediction models in the present study were more accurate than reported by Dal Zotto et al. (2008) and De Marchi et al. (2009), but less accurate than reported by De Marchi et al. (2013). In particular, Dal Zotto et al. (2008), who attempted to evaluate the effectiveness of milk MIRS to predict milk coagulation properties using 83 samples collected from Italian Holstein Friesian cows, obtained R^2_C ranging between 0.29 and 0.31 for RCT and a_{30}, respectively. De Marchi et al. (2009), using 1,200 Brown Swiss milk samples from 37 Italian herds, obtained R^2_C of 0.62 and 0.37 for RCT and a_{30}, respectively. Finally, based on milk samples from 335 Italian Holstein Friesian cows, De Marchi et al. (2013) developed robust MIRS prediction models for RCT ($R^2_C = 0.76$), k_{20} ($R^2_C = 0.72$), and a_{30} ($R^2_C = 0.70$). These prediction models are currently used in several milk laboratories in Italy to routinely provide milk coagulation phenotypes for genetic evaluations (Tiezzi et al., 2013; Penasa et al., 2014).

The prediction of some innovative milk traits is quite difficult, especially when these traits are not strictly related to the mid-infrared energy absorption of specific chemical functional bounds (i.e., N–H bonds in protein or C–H bonds in lipids). This is the case for milk coagulation traits, where several milk constituents affect MIRS predictions. As reported by De Marchi et al. (2009), loading variation across spectra demonstrated that several parts of the spectra contributed to the MIRS prediction models. Indeed, specific peaks across spectra, related to mid-infrared energy absorption of protein as well as lipid chemical bonds, were observed both in the present study and in De Marchi et al. (2009). The difficulty of MIRS in predicting the technological traits considered in the present study is also confirmed by the optimum number of model factors included in the PLS prediction model, which were greater in the prediction models for milk coagulation properties compared with other milk quality traits (i.e., casein fractions, FA composition; De Marchi et al., 2014).

Prediction of HCT, CMS, and pH

The present study is the first, to our knowledge, to quantify the effectiveness of MIRS to predict both HCT

and CMS. The observed predictive ability of HCT in the present study is very promising, despite the subjective nature of the HCT reference measurement. Heat coagulation time is very important for the dairy industry because all milk intended for human consumption is subjected to a heat treatment, and milk with high heat susceptibility (i.e., low HCT) is not suitable for milk processability, especially for the production of milk powder (i.e., mechanical obstruction of machinery for milk powder). In any case, MIRS has been demonstrated to be effective in predicting β-LG genotypes (Rutten et al., 2011), and β-LG A was demonstrated to be more susceptible to heat treatment (Hill et al., 1997). The predictive ability of CMS was poor and unsatisfactory; only 23 and 13% of the total variance in CMS was explained by the prediction model in cross and external validation, respectively. Although the repeatability of dynamic light scattering measurements with the reference instrument used in the present study was high (error ± 2%; Malvern, 2014), the poor predictive ability of CMS could be due to a limitation of the MIRS instrumentation in detecting dimensions of small particles. These results need further investigation in an independent population.

Milk acidity is usually measured as pH and titratable acidity (expressed in Soxhlet-Henkel degrees) in the dairy processing plant and it affects the aggregation rate of paracasein micelles, the reactivity of rennet, and the rate of syneresis (Toffanin et al., 2015). With the exclusion of hyperacidic milk types, milk with high acidity is considered more favorable for cheese making; indeed, the relationship between milk acidity and coagulation traits has been demonstrated in several studies (Cassandro et al., 2008; Pretto et al., 2013). Even if the measurement of pH during the cheese-making process is easy, the ability to predict milk pH on individual cow milk samples using MIRS, combined with the existence of potential genetic variation in this characteristic (Ikonen et al., 2004), could facilitate genetic improvement at the population level. Results from our study demonstrated that MIRS could be used to predict milk pH, and the accuracy of prediction in the present study was greater than that reported in the only other study that has evaluated the potential of milk MIRS to predict milk pH (De Marchi et al., 2009; R^2_C of 0.59 vs. 0.73).

Correlations Between Technological Traits and Milk Composition

The similarity in correlations with milk composition traits irrespective of whether reference values or predicted values of the technological traits were considered was somewhat expected, given the relatively good MIRS prediction accuracy of these traits. Milk technological traits were correlated with several traditional milk traits, particularly protein, casein, and urea concentrations. Previous research (Ikonen et al., 2004; Cassandro et al., 2008), evaluating both phenotypic and genetic correlations between coagulation traits and milk chemical composition, reported weak correlations between RCT with protein and casein levels. Strong correlations between milk coagulation traits and protein and casein contents were, nonetheless, expected in the present study, considering that rennet correction on protein concentration was used for the determination of milk coagulation traits. Therefore, comparison with previous literature is not completely valid because of differences in the gold-standard methods used; this was supported by O'Callaghan et al. (1999), who demonstrated that different concentrations of coagulant yielded different RCT values.

The present study demonstrated that, of all milk constituents, the nitrogen fraction (i.e., protein and casein concentrations) had the strongest correlation with RCT. This conclusion was expected because casein is the only milk constituent susceptible to addition of rennet. Therefore, curd formation occurs because of the aggregation of paracasein micelles, which, depending on their rate of aggregation, can subsequently incorporate other milk constituents, such as milk fat.

Lower milk pH values were associated with more favorable milk coagulation properties, which is consistent with several previous studies (Ikonen et al., 2004; Cassandro et al., 2008) that considered different measurements of acidity (e.g., pH, titratable acidity). The lack of correlation between CMS and protein is in direct contrast with Chen et al. (2014), who analyzed 550 individual milk samples for Holstein cows in the UK. Those authors, however, had different laboratory methodologies for the determination of protein concentration and for the determination of CSM.

The positive correlation between RCT and k_{20} and the negative correlation between RCT and a_{30} were expected because if coagulation begins within a short period, paracasein micelles have more time available to aggregate and to create a stronger curd, when measurement of curd firmness is detected at 30 min from rennet addition. In the same way, milk samples with short RCT achieved 20 mm of curd firmness (k_{20}) earlier relative to those relative to those characterized by poor reactivity to rennet addition (i.e., long RCT). The positive correlation between CMS and RCT is because major mineral constituents (i.e., Ca and P concentrations) influence both CMS and RCT (Chen et al., 2014; Gustavsson et al., 2014).

CONCLUSIONS

The combination of MIRS with PLS regression on untreated spectra could be a useful screening tool to acquire innovative milk technological phenotypes rapidly and at low cost. Such advancements could be important for the dairy industry to monitor milk quality before processing, and to discriminate more accurately milk that is better adapted for cheese manufacturing, whether for direct human consumption or stocking (milk powder production). Technological quality traits predicted by MIRS could be an important source of routine information on a very large population of animals from which to generate estimated breeding values for milk technological traits for exploitation in breeding programs. More effort, however, should be made to expand the number of samples and their variability in the calibration data set to increase the robustness of the prediction models.

ACKNOWLEDGMENTS

REFERENCES

Aleandri, R., J. C. Schneider, and L. G. Buttazzoni. 1989. Evaluation of milk for cheese production based on milk characteristics and Formagraph measures. J. Dairy Sci. 72:1967–1975.

Auldist, M. J., K. A. Johnston, N. J. White, W. P. Fitzsimons, and M. J. Boland. 2004. A comparison of the composition, coagulation characteristics and cheesemaking capacity of milk from Friesian and Jersey dairy cows. J. Dairy Res. 71:51–57.

Cassandro, M., A. Comin, M. Ojala, R. Dal Zotto, M. De Marchi, L. Gallo, P. Carnier, and G. Bittante. 2008. Genetic parameters of milk coagulation properties and their relationships with milk yield and quality traits in Italian Holstein Friesian cows. J. Dairy Sci. 91:371–376.

Chen, B., M. J. Lewis, and A. S. Grandison. 2014. Effect of seasonal variation on the composition and properties of raw milk destined for processing in the UK. Food Chem. 158:216–223.

Dal Zotto, R., M. De Marchi, A. Cecchinato, M. Penasa, M. Cassandro, P. Carnier, L. Gallo, and G. Bittante. 2008. Reproducibility and repeatability of measures of milk coagulation properties and predictive ability of mid-infrared reflectance spectroscopy. J. Dairy Sci. 91:4103–4112.

Davies, D. T., and J. C. D. White. 1966. The stability of milk protein to heat: I. Subjective measurement of heat stability of milk. J. Dairy Res. 33:67–81.

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.

De Marchi, M., V. Toffanin, M. Cassandro, and M. Penasa. 2013. Prediction of coagulating and noncoagulating milk samples using mid-infrared spectroscopy. J. Dairy Sci. 96:4707–4715.

De Marchi, M., V. Toffanin, M. Cassandro, and M. Penasa. 2014. Invited review: Mid-infrared spectroscopy as a phenotyping tool for milk traits. J. Dairy Sci. 97:1171–1186.

FAOSTAT. 2014. Metadata. Accessed June 23, 2014. http://faostat3.fao.org/faostat-gateway/go/to/download/Q/*/E.

Fox, P. F., and A. Brodkorb. 2008. The casein micelle: Historical aspects, current concepts and significance. Int. Dairy J. 18:677–684.

Gustavsson, F., M. Glantz, A. J. Buitenhuis, H. Lindmark-Månsson, H. Stålhammar, A. Andrén, and M. Paulsson. 2014. Factors influencing chymosin-induced gelation of milk from individual dairy cows: Major effects of casein micelle size and calcium. Int. Dairy J. 39:201–208.

Heuer, C., H. J. Luinge, E. T. G. Lutz, Y. H. Schukken, J. H. van der Maas, H. Wilmink, and J. P. T. M. Noordhuizen. 2001. Determination of acetone in cow milk by Fourier transform infrared spectroscopy for the detection of subclinical ketosis. J. Dairy Sci. 84:575–582.

Hewavitharana, A. K., and B. V. Brakel. 1997. Fourier transform infrared spectrometric method for rapid determination of casein in raw milk. Analyst (Lond.) 122:701–704.

Hill, J. P., M. J. Boland, and A. F. Smith. 1997. VI. 2—The effect of β-lactoglobulin variants on milk powder manufacture and properties. Bull. Int. Dairy Fed. 2:372–394.

Holt, C., D. D. Muir, and W. M. Sweetsur. 1978. Seasonal changes in the heat stability of milk from creamery silos in south-west Scotland. J. Dairy Res. 45:183–190.

Ikonen, T., S. Morri, A.-M. Tyrisevä, O. Ruottinen, and M. Ojala. 2004. Genetic and Phenotypic Correlations Between Milk Coagulation Properties, Milk Production Traits, Somatic Cell Count, Casein Content, and pH of Milk. J. Dairy Sci. 87:458–467.

Lin, L. I.-K. 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics 45:255–268.

Malacarne, M., P. Franceschi, P. Formaggioni, S. Sandri, P. Mariani, and A. Summer. 2014. Influence of micellar calcium and phosphorus on rennet coagulation properties of cows milk. J. Dairy Res. 81:129–136.

Malvern. 2014. Zetasizer Nano ZS. Accessed Nov. 25, 2014. http://www.malvern.com/en/products/product-range/zetasizer-range/zetasizer-nano-range/zetasizer-nano-zs/.

Martin, B., J.-F. Chamba, J.-B. Coulon, and E. Perreard. 1997. Effect of milk chemical composition and clotting characteristics on chemical and sensory properties of Reblochon de Savoie cheese. J. Dairy Res. 64:157–162.

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.

Ng-Kwai-Hang, K. F., I. Politis, R. I. Cue, and A. S. Marziali. 1989. Correlations between coagulation properties of milk and cheese yielding capacity and cheese composition. Can. Inst. Food Sci. Technol. J. 22:291–294.

O'Brien, B., R. Mehra, J. F. Connolly, and D. Harrington. 1999. Seasonal variation in the composition of Irish manufacturing and retail milks: 1. Chemical composition and renneting properties. Ir. J. Agric. Food Res. 38:53–64.

O'Callaghan, D. J., E. P. Mulholland, A. P. Duffy, C. P. O'Donnell, and F. A. Payne. 2001. Evaluation of hot wire and optical sensors for on line monitoring of curd firmness during milk coagulation. Ir. J. Agric. Food Res. 40:227–238.

O'Callaghan, D. J., C. P. O'Donnell, and F. A. Payne. 1999. A comparison of on-line techniques for determination of curd setting time using cheesemilks under different rates of coagulation. J. Food Eng. 41:43–54.

Penasa, M., F. Tiezzi, A. Sturaro, M. Cassandro, and M. De Marchi. 2014. A comparison of the predicted coagulation characteristics and composition of milk from multi-breed herds of Holstein-Friesian, Brown Swiss and Simmental cows. Int. Dairy J. 36:6–10.

Poulsen, N. A., H. P. Bertelsen, H. B. Jensen, F. Gustavsson, M. Glantz, H. Lindmark Månsson, A. Andrén, M. Paulsson, C. Bendixen, A. J. Buitenhuis, and L. B. Larsen. 2013. The occurrence of noncoagulating milk and the association of bovine milk coagula-

tion properties with genetic variants of the caseins in 3 Scandinavian dairy breed. J. Dairy Sci. 96:4830–4842.

Pretto, D., M. De Marchi, M. Penasa, and M. Cassandro. 2013. Effect of milk composition and coagulation traits on Grana Padano cheese yield under field conditions. J. Dairy Res. 80:1–5.

Pretto, D., T. Kaart, M. Vallas, I. Jõudu, M. Henno, L. Ancilotto, M. Cassandro, and E. Pärna. 2011. Relationships between milk coagulation property traits analyzed with different methodologies. J. Dairy Sci. 94:4336–4346.

Rutten, M. J. M., H. Bovenhuis, J. M. L. Heck, and J. A. M. van Arendonk. 2011. Prediction of β-lactoglobulin genotypes based on milk Fourier transform infrared spectra. J. Dairy Sci. 94:4183–4188.

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 system, and countries. J. Dairy Sci. 94:1657–1667.

Tiezzi, F., D. Pretto, M. De Marchi, M. Penasa, and M. Cassandro. 2013. Heritability and repeatability of milk coagulation properties predicted by mid-infrared spectroscopy during routine data recording, and their relationships with milk yield and quality traits. Animal 7:1592–1599.

Toffanin, V., M. De Marchi, N. Lopez-Villalobos, and M. Cassandro. 2015. Effectiveness of mid-infrared spectroscopy for prediction of the contents of calcium and phosphorus, and titratable acidity of milk and their relationship with milk quality and coagulation properties. Int. Dairy J. 41:68–73.

Williams, P. 2007. Statistical terms for evaluation of accuracy and precision. Pages 5–1 to 5–17 in Near Infrared Technology—Getting the Best Out of Light. 5.0 ed., PDK Grain, Nanaimo, BC, and Winnipeg, MB, Canada.