



Predictive ability of mid-infrared spectroscopy for major mineral composition and coagulation traits of bovine milk by using the uninformative variable selection algorithm

G. Visentin,¹ M. Penasa, P. Gottardo, M. Cassandro, and M. De Marchi

Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy

ABSTRACT

Milk minerals and coagulation properties are important for both consumers and processors, and they can aid in increasing milk added value. However, large-scale monitoring of these traits is hampered by expensive and time-consuming reference analyses. The objective of the present study was to develop prediction models for major mineral contents (Ca, K, Mg, Na, and P) and milk coagulation properties (MCP: rennet coagulation time, curd-firming time, and curd firmness) using mid-infrared spectroscopy. Individual milk samples ($n = 923$) of Holstein-Friesian, Brown Swiss, Alpine Grey, and Simmental cows were collected from single-breed herds between January and December 2014. Reference analysis for the determination of both mineral contents and MCP was undertaken with standardized methods. For each milk sample, the mid-infrared spectrum in the range from 900 to 5,000 cm^{-1} was stored. Prediction models were calibrated using partial least squares regression coupled with a wavenumber selection technique called uninformative variable elimination, to improve model accuracy, and validated both internally and externally. The average reduction of wavenumbers used in partial least squares regression was 80%, which was accompanied by an average increment of 20% of the explained variance in external validation. The proportion of explained variance in external validation was about 70% for P, K, Ca, and Mg, and it was lower (40%) for Na. Milk coagulation properties prediction models explained between 54% (rennet coagulation time) and 56% (curd-firming time) of the total variance in external validation. The ratio of standard deviation of each trait to the respective root mean square error of prediction, which is an indicator of the predictive abil-

ity of an equation, suggested that the developed models might be effective for screening and collection of milk minerals and coagulation properties at the population level. Although prediction equations were not accurate enough to be proposed for analytic purposes, mid-infrared spectroscopy predictions could be evaluated as phenotypic information to genetically improve milk minerals and MCP on a large scale.

Key words: mid-infrared spectroscopy, dairy cattle, milk mineral, milk coagulation property

INTRODUCTION

Milk quality is crucial to maximize milk's added value and it contributes to increase the profitability of the entire dairy chain. Traditional quality traits have mainly referred to milk chemical composition, particularly protein and fat quantity and concentration, as confirmed by selection indices of several cattle breeds worldwide (Miglior et al., 2005). However, the concept of milk quality is often interpreted differently by processors and consumers. For example, under a processor point of view, adequate milk quality is translated into adequate processing ability, whereas the consumer's perception of milk quality deals more with health aspects.

Minerals represent a relatively small part of cow milk, close to 10 g/L, and they are divided into 2 categories, based on their concentration (Cashman, 2006): macro minerals (normally expressed in mg/kg) and trace elements (normally expressed in $\mu\text{g}/\text{kg}$). Macro minerals include Ca, K, Mg, Na, and P, which are important for the homeostasis of both infants and adults. Indeed, Ca and Mg are involved in bone and tooth health, and in muscular and cardiac contractility (Cashman, 2006; Haug et al., 2007). A deficiency of these minerals is associated with osteoporosis and muscular disorders, and Ca deficiency in the diet might be partially responsible for a greater incidence of hypertension, colon cancer, and obesity (Huth et al., 2006). Potassium is known as the most important intracellular cation, playing a

Received February 19, 2016.

Accepted July 4, 2016.

¹Corresponding author: giulio.visentin@studenti.unipd.it

fundamental role in the maintenance of homeostasis (Young et al., 1995; He and MacGregor, 2008). However, K is found also as an extracellular element and participates in the transmission of nervous impulses, muscle contraction, and regulation of blood pressure (He and MacGregor, 2008). In particular, an increase of dietary K, coupled with a reduction of dietary Na, limits the risk of hypertension (Whelton and He, 2014). Recently, Uribarri and Calvo (2014) claimed that, although considered an essential nutrient, high dietary P intake is a risk factor for bone and cardiovascular diseases. Besides health aspects, Ca and P are essential components of casein micelles and thus they are directly involved in milk coagulation process. Malacarne et al. (2014) reported that high content of inorganic P positively affects casein micelle reactivity to rennet.

Milk coagulation properties (**MCP**); namely, rennet coagulation time (**RCT**, min), curd-firming time (**k₂₀**, min), and curd firmness (**a₃₀**, mm), are currently used to measure milk quality during the cheese-making process, and they are measured by several laboratories for breeding purposes and milk quality payment systems (Tiezzi et al., 2013; Penasa et al., 2015). Several studies demonstrated that milk characterized by short RCT and firm curd results in greater cheese yield and thus increases the efficiency of the entire cheese-making process (Comin et al., 2005; Malacarne et al., 2006; Pretto et al., 2013).

Reference methods normally used to measure milk mineral composition and coagulation traits are expensive and time consuming. Mid-infrared spectroscopy (**MIRS**) is a rapid, nondestructive, and cost-effective laboratory technique that allows the (a posteriori) prediction of innovative phenotypes from milk samples (De Marchi et al., 2014; McParland and Berry, 2016), but the prediction of both MCP (De Marchi et al., 2013; Visentin et al., 2015) and mineral content (Soyeurt et al., 2009; Toffanin et al., 2015) is still a big challenge. To improve the accuracy of MIRS prediction models 2 paths should be considered: (1) to increase the accuracy of reference methods, and (2) to use different statistical approaches coupled with multivariate analyses, including partial least squares (**PLS**) regression. Recently, Gottardo et al. (2015) demonstrated that uninformative variable elimination (**UVE**) can increase the prediction accuracy of MIRS models by reducing the number of uninformative spectral regions. This process is extremely advantageous when models have to be applied subsequently to large spectral data set for the prediction of novel phenotypes, since a lower number of spectral wavelengths used for PLS regression reduces the computational time. Therefore, the aim of the present study was to develop MIRS prediction models for

major mineral contents and MCP using PLS coupled with UVE for the application of these prediction models on spectral data.

MATERIALS AND METHODS

Data

Sample Collection. From January to December 2014, 923 individual cow milk samples were collected in 60 single-breed herds. This data set, subsequently used to develop MIRS prediction models, contained the 4 major cattle breeds reared in the Alpine area of Bolzano province (Italy), where all animals were sampled. Cow breeds considered in the present study were Holstein-Friesian (**HF**, n = 237), Brown Swiss (**BS**, n = 223), Alpine Grey (**AG**, n = 223), and Simmental (**SI**, n = 240). The sampling protocol aimed at covering as much biological variability as possible, for both MCP and major mineral composition. For each cow, two 50-mL aliquots were collected, immediately added with preservative (Bronysolv; ANA.LI.TIK Austria, Vienna, Austria) and kept at refrigerating temperature.

Milk Chemical Composition and Spectra Determination. Both aliquots were transferred to the laboratory of the South Tirol Dairy Association (Bolzano, Italy) and one aliquot was processed the same day of sampling according to the International Committee for Animal Recording (ICAR, 2014) recommendations. For each milk sample, both traditional milk quality traits (pH and contents of protein, casein, fat, lactose, and urea) and MIRS spectra were determined using a MilkoScan FT+ (Foss Electric A/S, Hillerød, Denmark). Each individual spectral information, containing 1,060 infrared transmittance data in the region between 900 and 5,000 cm⁻¹, was stored. Somatic cell count was measured using Fossomatic (Foss Electric A/S) and then converted to SCS through the formula $SCS = 3 + \log_2(SCC/100,000)$. The other 50-mL aliquot was transferred (within 24 h from collection) at refrigerating temperature to the laboratory of the Department of Agronomy, Food, Natural Resources, Animals and Environment of the University of Padova (Legnaro, Italy). This aliquot was subsequently split in 2 sub-aliquots: one was stored at -20°C and the other one was delivered the same day to the laboratory of the Breeders Association of Veneto Region (Padova, Italy).

Reference Analysis of Major Mineral Contents. Milk content of Ca, K, Mg, Na, and P was determined on 251 milk samples in the laboratory of the Department of Agronomy, Food, Natural Resources, Animals and Environment (Legnaro, Italy) and each of the 4 cattle breeds (HF, BS, AG, and SI) was rep-

resented in equal proportion. In the present study, we chose to perform a mineralization of milk samples before their quantification, as recommended by Soyeurt et al. (2009). A nitroperchloric mineralization by a Milestone Start D microwave (1,200 W, Milestone Srl Sorisole, Bergamo, Italy) was used. The microwave contained a SK-10 rotor at high pressure (6.4 MPa) and control systems with temperature probe and software. Milk samples stored at -20°C were thawed in water at 35°C and homogenized before sampling. A total of 2.5 g of sample was introduced in Teflon vessels at high pressure, and 2 mL of H_2O_2 30% and 7 mL of HNO_3 67% were added. The vessel was hermetically sealed and placed in the microwave. The process of mineralization involved 3 steps: (1) the sample was heated to 200°C in 15 min; (2) the sample was kept at 200°C for 18 min; and (3) the sample was cooled down to 35°C . The mineralized sample was added with demineralized water to reach a final volume of 25 mL.

Content of Ca, K, Mg, Na, and P was quantified using inductively coupled plasma optical emission spectrometry, Ciro Vision EOP (Spectro Analytical Instruments GmbH, Kleve, Germany). This method was employed to determine Ca at wavelength 315.887 nm, K at wavelength 766.491 nm, Mg at wavelength 279.079 nm, Na at wavelength 589.592 nm, and P at wavelength 178.287 nm. Instrument operating parameters were optimized for acid solution and calibration standards were matched with nitric acid 5% "suprapure" grade. The elements to be determined were added from single element solutions (Inorganic Ventures, Christiansburg, VA). The concentration range of the calibration solutions was between 0 and 100 mg/kg for all elements. The accuracy and precision of this method were investigated by analyzing the certified reference material BCR-063R skim milk powder (Institute for Reference Materials and Measurements, Geel, Belgium). These standards of skim milk powder were mineralized according to the protocol previously described. The measured and the certified values were in excellent agreement for all the elements (R^2 of calibration curves for all minerals >0.99).

Reference Analysis of Milk Coagulation Properties. Reference analysis was assessed through the Formagraph (Foss Electric A/S), using the method proposed by McMahan and Brown (1982), in the laboratory of the Breeders Association of Veneto Region (Padova, Italy) within 1 d from sample arrival. A 10-mL aliquot of each preserved milk sample was heated to 35°C and then added with 200 μL of a rennet solution (Hansen Standard 160; Pacovis Amrein AG, Bern, Switzerland) diluted to 1.6% (wt/vol) in distilled water. At the start of the analysis, an oscillating loop

pendulum is brought into contact with the milk; when the coagulation begins, the viscosity of milk increases and so a force is back-transmitted to the loop. The output of the Formagraph is therefore a graph of curd firmness against time. Measured traits were (1) RCT, the time required, from the beginning of the test, to induce milk coagulation; (2) k_{20} , the time from RCT to the achievement of a 20 mm width in the graph; and (3) a_{30} , the width of the graph at the end of the test, which lasted 30 min. Milk samples that did not coagulate within the testing time (about 12% of total samples) were classified as noncoagulating and labeled as missing values.

Chemometric Analysis

Data Editing. The normality of each trait and the identification of any possible reference outliers were ascertained both through visual inspection of the histogram/density plot and by carrying out the Shapiro-Wilk normality test in R software (R Development Core Team, 2015). Traits that are not normally distributed are characterized by a low and statistically significant ($P < 0.05$) Shapiro-Wilk value, and they require normalization. In the present study, both milk minerals and MCP were normally distributed, and thus no data transformation was needed. Reference outliers were defined as samples whose values deviated more than 3 SD from the mean of each trait (McDermott et al., 2016a). Based on this definition, none observations for a_{30} and K, 1 observation for RCT, k_{20} , and Mg, 2 observations for Na and P, and 3 observations for Ca were discarded before the development of prediction models.

Spectral data expressed in transmittance were transformed to absorbance by taking the \log_{10} of the reciprocal of the transmittance (McDermott et al., 2016b; Figure 1). Two high noise level spectral regions related to water absorption ($1,600$ to $1,700\text{ cm}^{-1}$, and $3,040$ to $3,660\text{ cm}^{-1}$) were discarded before multivariate analysis (Hewavitharana and van Brakel, 1997; Figure 1). Principal component analysis was carried out on raw spectral data to identify similarities and differences between spectra, to detect any possible spectral outlier. Principal component analysis is a method of data compression that produces a new matrix of uncorrelated variables called principal components; each principal component captures in a descending order as much of the variation of the initial variables (i.e., spectral wavenumbers) as possible. First and second principal component explained 71.09 and 16.01% of total spectral variation, respectively. The score plot of the first versus second principal component did not highlight any specific spectrum outlier, and this was also con-

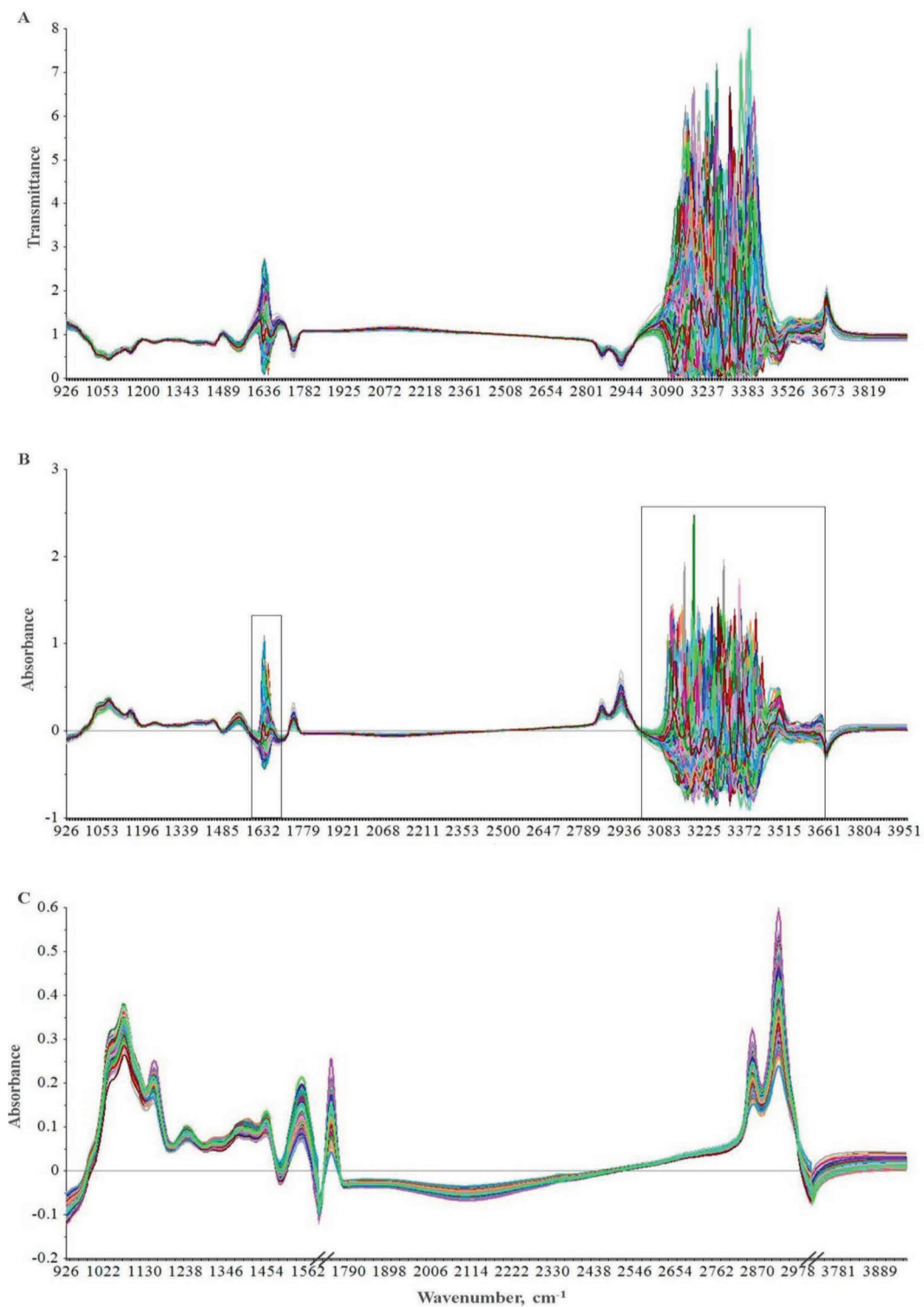


Figure 1. Plot of 923 spectra expressed in (A) transmittance, (B) absorbance with high noise regions included in the box, and (C) absorbance after elimination of high noise regions. Color version available online.

firmed by computing the robust Mahalanobis distance for each sample. Therefore, all spectra were retained for further statistical analysis.

Prediction Models. To develop and to externally validate the prediction models, for each trait of interest the data set was divided randomly into 2 subsets, namely a calibration set (80% of the observations) and a validation set (20% of the observations). The prediction models included each milk mineral and MCP as dependent variable and the edited milk spectra wave-numbers as predictor variables. The calibration set was used to develop the MIRS models through PLS and UVE combined with PLS regression on the edited spectra, and these models were subsequently applied to the validation set. The validation set was used to assess the predictive ability of each prediction model, because the samples included were not considered to build the models. Leave-one-out cross-validation was also performed in the calibration set. This process was repeated 1,000 times for the development of each prediction model, and it was performed both when equations were built through PLS regression on edited spectra and when equations were built using UVE combined with PLS regression (**UVE-PLS**) on edited spectra. Partial least squares regression was carried out using the Chemometrics with R package (Wehrens, 2011), and UVE-PLS was tested using a homemade script in R software (R Development Core Team, 2015), following the approach of Gottardo et al. (2015). Moreover, PLS regression was carried out on spectra expressed in first derivative using a Savitzky-Golay filter with a linear polynomial and a window size of 3 points, but no improvement of predictive accuracy of equations was detected compared with PLS regression performed on untreated spectra.

The optimal number of PLS factors (**#L**) to perform PLS regression was defined as the minimum number of factors to achieve the lowest root mean square error of cross validation (**RMSE_{CV}**). Goodness-of-fit statistics considered were the coefficient of determination of cross validation (**R²_{CV}**) and of external validation (**R²_{EV}**), the **RMSE_{CV}**, the root mean square error of external validation (**RMSE_{EV}**), and the ratio of prediction to deviation (**RPD**), which was calculated as the ratio of the SD of the trait to the **RMSE_{EV}**, and it was used to test the practical utility of the prediction models. In particular, **RPD** values greater than 2 are desired for practical application of the prediction model (Williams, 2007). Results represent the average goodness-of-fit statistics of the 1,000 analyses.

RESULTS

Summary Statistics of Calibration Data Set

Descriptive statistics of milk yield, chemical composition predicted by MilkoScan FT+ (Foss Electric A/S), reference mineral contents, and reference MCP of samples after removal of outlier reference values are reported in Table 1. Milk yield averaged 23.20 kg/d, mean percentages of fat, protein, and casein were 4.07, 3.60, and 2.81, respectively, and SCS averaged 2.82. The most abundant milk mineral was K (1,500.52 mg/kg), followed by Ca (1,348.22 mg/kg) and P (1,010.04 mg/kg). Magnesium content averaged 128.30 mg/kg, and it was the less concentrated in milk. Means of **RCT**, **k₂₀**, and **a₃₀** were 18.57 min, 5.20 min, and 28.07 mm, respectively. Minerals were characterized by appreciable variability, with coefficient of variation

Table 1. Descriptive statistics of milk mineral composition and coagulation properties used for the development of prediction models

Trait ¹	Mean	SD	Range	CV, %
Milk yield, kg/d	23.20	8.42	51.20	36.29
Chemical composition				
Fat, %	4.07	0.72	7.61	17.69
Protein, %	3.60	0.45	3.16	12.50
Casein, %	2.81	0.36	2.67	12.81
SCS	2.82	1.84	12.44	65.25
Mineral composition, mg/kg				
Ca	1,348.22	229.60	1,281.52	17.03
K	1,500.52	228.04	1,108.58	15.20
Mg	128.30	22.10	112.71	17.23
Na	399.18	91.87	448.54	23.05
P	1,010.04	181.04	1,010.95	17.92
Coagulation trait				
RCT , min	18.57	4.61	24.70	24.84
k₂₀ , min	5.20	1.63	11.15	31.30
a₃₀ , mm	28.07	12.37	54.72	44.07

¹**RCT** = rennet coagulation time; **k₂₀** = curd-firming time; **a₃₀** = curd firmness 30 min after rennet addition.

Table 2. Average fitting statistics¹ of prediction models for milk mineral composition and milk coagulation properties in leave-one-out cross-validation obtained from partial least squares (PLS) regression only and from PLS after uninformative variable elimination (UVE-PLS) procedure

Trait ²	PLS				UVE-PLS			
	#L	N ₁	RMSE _{CV}	R ² _{CV} (SD)	N ₂	RMSE _{CV}	R ² _{CV} (SD)	RPD
Mineral composition, mg/kg								
Ca	15	873	146.20	0.60 (0.08)	113	120.00	0.68 (0.06)	1.91
K	15	873	139.11	0.60 (0.05)	173	120.00	0.69 (0.05)	1.90
Mg	20	873	15.62	0.61 (0.05)	93	12.30	0.65 (0.05)	1.80
Na	15	873	90.03	0.29 (0.06)	341	68.80	0.42 (0.08)	1.34
P	20	873	109.77	0.62 (0.07)	210	83.90	0.71 (0.08)	2.16
Coagulation trait								
RCT, min	20	873	3.38	0.46 (0.03)	163	2.86	0.55 (0.03)	1.61
k ₂₀ , min	20	873	1.18	0.46 (0.03)	144	1.00	0.59 (0.03)	1.63
a ₃₀ , mm	20	873	9.26	0.48 (0.05)	110	8.43	0.56 (0.05)	1.47

¹#L = number of model PLS factors; N₁ = number of wavenumbers in PLS regression; RMSE_{CV} = root mean square error in cross validation; R²_{CV} = coefficient of determination in cross-validation; N₂ = number of wavenumbers in PLS regression after UVE procedure; RPD = ratio of prediction to deviation.

²RCT = rennet coagulation time; k₂₀ = curd-firming time; a₃₀ = curd firmness 30 min after rennet addition.

ranging between 15.20% (K) and 23.05% (Na). Milk coagulation properties exhibited greater variability compared with milk chemical composition and mineral contents, with coefficients of variation of 24.84, 31.30, and 44.07% for RCT, k₂₀, and a₃₀, respectively.

Accuracy of Prediction Models

Goodness-of-fit statistics of cross validation and external validation for MIRS prediction models are summarized in Tables 2 and 3, respectively. The optimal #L selected by PLS regression was 15 for Ca, K, and Na, and it was 20 for Mg, P, RCT, k₂₀, and a₃₀. The R²_{CV} (Table 2) of models for milk minerals developed using only PLS regression ranged from 0.29 (Na; RMSE_{CV} of 90.03 mg/kg) and 0.62 (P; RMSE_{CV} of 109.77 mg/

kg). The R²_{CV} (RMSE_{CV}) of models for RCT, k₂₀, and a₃₀ developed using only PLS regression was 0.46 (3.38 min), 0.46 (1.18 min), and 0.48 (9.26 mm), respectively. In external validation (Table 3), R²_{EV} of milk minerals varied from 0.25 (Na) to 0.60 (P), and for MCP it ranged from 0.43 (RCT) to 0.46 (a₃₀).

Uninformative variable elimination reduced the number of wavenumbers used by the software to compute PLS regression and both RMSE_{CV} and RMSE_{EV}, and increased R²_{CV} and R²_{EV} for all prediction models (Tables 2 and 3). The greatest increment of explained variance in external validation was observed for Na prediction, with R²_{EV} that improved from 0.25 to 0.40 (+60%; Table 3), followed by k₂₀ prediction, whose explained variance in external validation increased by 27.3% (R²_{EV} from 0.44 to 0.56). For the other traits, the

Table 3. Average fitting statistics¹ of prediction models for milk mineral composition and milk coagulation properties in external validation obtained from partial least squares (PLS) regression only and from PLS after uninformative variable elimination (UVE-PLS) procedure

Trait ²	PLS				UVE-PLS			
	#L	N ₁	RMSE _{EV}	R ² _{EV} (SD)	N ₂	RMSE _{EV}	R ² _{EV} (SD)	RPD
Mineral composition, mg/kg								
Ca	15	873	153.02	0.58 (0.08)	113	122.00	0.67 (0.06)	1.88
K	15	873	144.23	0.57 (0.08)	173	120.00	0.69 (0.08)	1.90
Mg	20	873	18.22	0.57 (0.06)	93	12.50	0.65 (0.05)	1.77
Na	15	873	91.76	0.25 (0.10)	341	70.00	0.40 (0.08)	1.31
P	20	873	115.43	0.60 (0.08)	210	88.12	0.68 (0.05)	2.05
Coagulation trait								
RCT, min	20	873	4.00	0.43 (0.05)	163	2.90	0.54 (0.05)	1.59
k ₂₀ , min	20	873	1.30	0.44 (0.06)	144	1.22	0.56 (0.06)	1.34
a ₃₀ , mm	20	873	9.30	0.46 (0.03)	110	9.00	0.52 (0.03)	1.37

¹#L = number of model PLS factors; N₁ = number of wavenumbers in PLS regression; RMSE_{EV} = root mean square error in external validation; R²_{EV} = coefficient of determination in external validation; N₂ = number of wavenumbers in PLS regression after UVE procedure; RPD = ratio of prediction to deviation.

²RCT = rennet coagulation time; k₂₀ = curd-firming time; a₃₀ = curd firmness 30 min after rennet addition.

increment of explained variance in external validation was between 13.0 (a_{30}) and 25.6% (RCT; Table 3). The UVE procedure reduced by at least 80% the number of wavenumbers used in PLS regression, with the only exception of Na (60.9%) and P (75.9%) prediction models (Tables 2 and 3).

The RPD was calculated only when models were generated by UVE-PLS, and in external validation it ranged from 1.31 (Na) to 2.05 (P) for milk minerals, and from 1.34 (k_{20}) to 1.59 (RCT) for MCP (Table 3).

DISCUSSION

Data Variation

Mid-infrared spectroscopy prediction models require representative variability in the calibration data set. Regarding milk mineral variation, coefficient of variation for Ca and P of the present study (17.03 and 17.92%, respectively) was greater than coefficient of variation reported by Toffanin et al. (2015) for the same minerals (11.28 and 11.74%, respectively). However, Toffanin et al. (2015) focused their study only on Italian HF cows sampled in a short period (3 mo). On the other hand, Soyeurt et al. (2009) developed MIRS prediction models on a multi-breed data set ($n = 100$) of HF, Jersey, Belgian Blue, Montbéliarde, and Normande cattle breeds over a 1-yr sampling, and coefficients of variation for Ca, K, Mg, Na, and P were comparable to those reported in the current research. Concerning MCP, data variability was similar to that reported by Cassandro et al. (2008) on a large-scale study of 1,071 Holstein-Friesian cows (CV of 27 and 35% for RCT and a_{30} , respectively).

Variable Selection and Accuracy of Prediction Models

Spectral data are characterized by multicollinearity: indeed, some wavenumbers can be rewritten as a linear function of others. Partial least squares regression compresses the initial spectral matrix by creating another matrix containing a smaller amount of variables called PLS factors. Subsequently, these factors are used in the regression equation, and the less relevant and unstable information of the initial spectral matrix is discarded (Næs et al., 2002). Theoretically, variable selection would not be needed because PLS regression already ignores noninformative wavenumbers. However, a large number of wavenumbers and a small number of samples could still have a negative effect on the final accuracy of the developed prediction model, and this is the main reason to use variable selection techniques.

Several variable selection methods are available in the literature, and they have been recently reviewed by Mehmood et al. (2012). Uninformative variable elimination was first introduced by Centner et al. (1996), and it basically creates a stability criterion for each variable; if the criterion is below a specific threshold, this indicates that the specific wavenumber is not essential for the development of the prediction model. By discarding these wavenumbers, the accuracy of prediction models is expected to increase. Recently, UVE has been applied by Gottardo et al. (2015) to 208 milk spectra to develop MIRS prediction models for Ca and titratable acidity. Those authors reported a substantial improvement of the R^2_{EV} , from 0.46 to 0.55 and from 0.72 to 0.80 for Ca and titratable acidity, respectively. Moreover, Niero et al. (2016) applied the UVE method for the prediction of casein and whey fractions, and they obtained an increment of R^2_{CV} from 0.83 to 0.88, 0.36 to 0.60, and 0.66 to 0.74 for α -CN, β -CN, and κ -CN models, respectively. The R^2_{CV} of whey fractions models increased from 0.31 to 0.37 and from 0.31 to 0.47 for α -LA and β -LG, respectively.

To our knowledge only 2 studies have investigated the potential use of MIRS to predict milk mineral content, but none of them used UVE-PLS procedure for the development of calibration models. Soyeurt et al. (2009) developed calibration models for the same milk minerals investigated in the present study, and they reported greater R^2_{CV} for Ca (0.87), Na (0.65), and P (0.85), similar R^2_{CV} for Mg (0.65), and lower R^2_{CV} for K (0.36), compared with findings of the current study. Toffanin et al. (2015) developed calibration equations for Ca and P using 208 milk samples of HF cows through PLS regression, and they obtained R^2_{CV} of 0.56 and 0.70 for Ca and P, respectively, which were lower than those obtained in the present study. Only Soyeurt et al. (2009) quantified the predictive ability of MIRS models on an independent data set, and they reported R^2_{EV} of 0.97, 0.14, and 0.88 for Ca, Na, and P, respectively.

The prediction of MCP by MIRS has been reported in both Italian BS (De Marchi et al., 2009) and HF (De Marchi et al., 2013) cows, and by Visentin et al. (2015) in a multi-breed data set of Irish cattle (HF, Jersey, and Norwegian Red) reared in a grazing dairy system. Results from the present research were comparable with those of Visentin et al. (2015), in terms of fitting statistics of both internal and external validation, but less accurate than those of De Marchi et al. (2013), who obtained R^2_{CV} of 0.76, 0.72, and 0.70 for RCT, k_{20} , and a_{30} , respectively. Nevertheless, the study of De Marchi et al. (2013) was carried out using a quite different reference analysis; indeed, even if they used a Formagraph

instrument, they extended the time of the analysis until 60 min to allow the detection of noncoagulating milk samples (i.e., samples that did not coagulate within 30 min). The wide range of variation in the reference data set improved the overall accuracy of the prediction models.

CONCLUSIONS

Mid-infrared spectroscopy has been used to develop calibration models for milk minerals and coagulation properties. Models explained moderate variance of studied traits and thus they were not accurate enough to be proposed for analytic purposes. However, these models could represent a valid tool for a quick and cost-effective screening or acquiring phenotypes at the population level (or both), and they could be applied for research purposes to spectral data. Finally, the use of the UVE-PLS approach has been validated to be an interesting approach to improve the accuracy of prediction models. Future research will investigate the feasibility of using mid-infrared predictions as indicator traits to genetically improve milk minerals and MCP.

ACKNOWLEDGMENTS

The authors thank Luis Kerschbaumer and Thomas Kerschbamer (Sennereiverband Südtirol, Bolzano, Italy), Christian Plitzner (Vereinigung der Südtiroler Tierzuchtverbände, Bolzano, Italy), Claudio Stecchi (CTS informatica srl, Bolzano, Italy), and Luciano Magro, Massimo Cagnin, and Martina Isaia (Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova) for technical support.

REFERENCES

- Cashman, K. D. 2006. Milk minerals (including trace elements) and bone health. *Int. Dairy J.* 16:1389–1398.
- 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 cows. *J. Dairy Sci.* 91:371–376.
- Centner, V., D.-L. Massart, O. E. de Noord, S. de Jong, B. M. Vandegiste, and C. Sterna. 1996. Elimination of uninformative variables for multivariate calibration. *Anal. Chem.* 68:3851–3858.
- Comin, A., M. Cassandro, M. Povinelli, and G. Bittante. 2005. Genetic aspects of milk coagulation properties in Italian Holstein cows. *Ital. J. Anim. Sci.* 4(Suppl. 2):10–12.
- 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 phenotyping tool for milk traits. *J. Dairy Sci.* 97:1171–1186.
- Gottardo, P., M. De Marchi, M. Cassandro, and M. Penasa. 2015. *Technical note*: Improving the accuracy of mid-infrared prediction models by selecting the most informative wavelengths. *J. Dairy Sci.* 98:4168–4173.
- Haug, A., A. T. Høstmark, and O. M. Harstad. 2007. Bovine milk in human nutrition—A review. *Lipids Health Dis.* 6:25.
- He, F. J., and G. A. MacGregor. 2008. Beneficial effects of potassium on human health. *Physiol. Plant.* 133:725–735.
- 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.
- Huth, P. J., D. B. DiRienzo, and G. D. Miller. 2006. Major scientific advances with dairy foods in nutrition and health. *J. Dairy Sci.* 89:1207–1221.
- ICAR. 2014. International Committee for Animal Recording—Approved by the General Assembly held in Berlin, Germany, on May 2014. http://www.icar.org/wp-content/uploads/2015/11/Guidelines_2014.pdf.
- 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.
- Malacarne, M., A. Summer, E. Fossa, P. Formaggioni, P. Franceschi, M. Pecorari, and P. Mariani. 2006. Composition, coagulation properties and Parmigiano-Reggiano cheese yield of Italian Brown and Italian Friesian herd milks. *J. Dairy Res.* 73:171–177.
- McDermott, A., G. Visentin, M. De Marchi, D. P. Berry, M. A. Fenelon, P. M. O'Connor, O. A. Kenny, and S. McParland. 2016a. Prediction of individual milk proteins including free amino acids in bovine milk using mid-infrared spectroscopy and their correlations with milk processing characteristics. *J. Dairy Sci.* 99:3171–3182.
- McDermott, A., G. Visentin, S. McParland, D. P. Berry, M. A. Fenelon, and M. De Marchi. 2016b. Effectiveness of mid-infrared spectroscopy to predict the color of bovine milk and the relationship between milk color and traditional milk quality traits. *J. Dairy Sci.* 99:3267–3273.
- McMahon, D. J., and R. J. Brown. 1982. Evaluation of Formagraph for comparing rennet solutions. *J. Dairy Sci.* 65:1639–1642.
- McParland, S., and D. P. Berry. 2016. The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. *J. Dairy Sci.* 99:4056–4070.
- Mehmood, T., K. H. Liland, L. Snipen, and S. Sæbø. 2012. A review of variable selection methods in Partial Least Squares Regression. *Chemom. Intell. Lab. Syst.* 118:62–69.
- Miglior, F., B. L. Muir, and B. J. Van Doormaal. 2005. Selection indices in Holstein cattle of various countries. *J. Dairy Sci.* 88:1255–1263.
- Næs, T., T. Isaksson, T. Fearn, and T. Davies. 2002. *A User-Friendly Guide to Multivariate Calibration and Classification*. NIR Publications, Chichester, UK.
- Niero, G., M. Penasa, P. Gottardo, M. Cassandro, and M. De Marchi. 2016. *Short communication*: Selecting the most informative mid-infrared spectra wavenumbers to improve the accuracy of prediction models for detailed milk protein content. *J. Dairy Sci.* 99:1853–1858.
- Penasa, M., M. De Marchi, S. Ton, L. Ancilotto, and M. Cassandro. 2015. Reproducibility and repeatability of milk coagulation properties predicted by mid-infrared spectroscopy. *Int. Dairy J.* 47:1–5.
- 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.
- R Development Core Team. 2015. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Soyeurt, H., D. Bruwier, J.-M. Romnee, N. Gengler, C. Bertozzi, D. Veselko, and P. Dardenne. 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 92:2444–2454.

- 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.
- Uribarri, J., and M. S. Calvo. 2014. Dietary phosphorus intake and health. *Am. J. Clin. Nutr.* 99:247–248.
- Visentin, G., A. McDermott, S. McParland, D. P. Berry, O. A. Kenny, A. Brodtkorb, M. A. Fenelon, and M. De Marchi. 2015. Prediction of bovine milk technological traits from mid-infrared spectroscopy analysis in dairy cows. *J. Dairy Sci.* 98:6620–6629.
- Wehrens, R. 2011. *Chemometrics With R: Multivariate Data Analysis in the Natural Sciences and Life Sciences*. Springer, Heidelberg, Germany.
- Whelton, P. K., and J. He. 2014. Health effects of sodium and potassium in humans. *Curr. Opin. Lipidol.* 25:75–79.
- Williams, P. 2007. *Near-Infrared Technology—Getting the Best Out of Light*. 5.0 ed. PDK Projects Inc., Nanaimo, Canada.
- Young, D. B., H. Lin, and R. D. McCabe. 1995. Potassium's cardiovascular protective mechanisms. *Am. J. Physiol.* 268:R825–R837.