Prediction of sheep bulk milk coagulation properties from mid-infrared spectral data

Carlo Boselli
Experimental Zooprophylactic Institute Lazio and Toscana Mariano Aleandri
Rome, Italy
carlo.boselli@izslt.it

Alberto Guerra
Department of Agronomy, Food, Natural Resources, Animals and Environment University of Padova
Padova, Italy
alberto.guerra@unipd.it

Angela Costa
Department of Veterinary Medical Sciences University of Bologna
Bologna, Italy
angela.costa2@unibo.it

Massimo De Marchi
Department of Agronomy, Food, Natural Resources, Animals and Environment University of Padova
Padova, Italy
massimo.demarchi@unipd.it

***Abstract*—The technological features of milk are essential for cheese manufacturing. This is particularly true for Italy, where most of the milk produced by sheep is intended for cheese production. The possibility to evaluate technological characteristics and coagulation aptitude of milk in advance, before any treatment, is crucial for decision-making at industry level. In the present study, we tested the ability of mid-infrared spectroscopy for prediction of coagulation traits (rennet coagulation time and curd firmness) by using more than 4,000 bulk milk samples of 344 sheep herds. The models developed with a partial least square regression showed that spectral data points can be successfully used to predict the two traits. The coefficient of determination in external validation was 0.42 for rennet coagulation time and 0.28 for curd firmness, indicating that sheep milk delivered to dairies can undergo a preliminary screening only to assess the expected coagulation time. This finding will allow manufacturers to evaluate the milk received from farmers. Further investigation will be need to improve the prediction of rennet coagulation time that can be coupled with composition traits to define premiums or penalties in the payment system.**

Keywords—sheep, cheese, rennet, MIRS

I. INTRODUCTION

Bulk milk technological characteristics are economically relevant for cheese industry. In Italy, approximately the 80% of the milk produced is converted into cheese. Although dairy cattle account for most of the volume produced, species like buffalo, goat and sheep are often linked to high-value Protected Designation of Origin (PDO) cheeses. In 2021, 182,887 lactating ewes were officially monitored in Italy [1], most of them belonging to the Sarda breed (115,455 heads). Due to the limited consumption of raw sheep milk in drinkable form, it is reasonable to expect that the totality of the milk produced by Italian ewes is supplied to dairies for cheese production. This makes the coagulation ability a characteristic of economic interest. There are commercial instruments developed for the analysis of milk coagulation process. In 30 min it is possible to determine the rennet coagulation time (RCT, min) and the curd firmness after 30 min from rennet addition (a_{30}, mm) of the milk. On the other hand, most of the conventional devices allow to analyse only 10 samples per run and are therefore not convenient for large cooperatives or plants for a routine usage on all the milk supplied daily by individual farmers. Nowadays, implementation of rapid methods is advisable to perform screening on the milk supplied by farmers. Mid-infrared spectroscopy (MIRS), for example, has been extensively used for bovine milk analysis and is recognized as a gold standard for some components. In recent years, the scientific community has demonstrated that milk spectra can be used for the prediction of hard-to-be-measured traits, such as cow-related aspects (health, metabolic status, pregnancy, etc) and milk technological aptitude. In minor dairy species, however, the existing prediction models for RCT and a_{30} in most of the cases are based on a small sample size and are characterized by a robustness that is far from being considered sufficient and exploitable. In bovine milk, the coefficient of determination of RCT is reported to vary between 0.55 and 0.62 [2,3]. In the case of a_{30}, the prediction accuracy is generally lower [3]. The possibility to evaluate technological characteristics and coagulation aptitude of bulk sheep milk at the dairy plant gate, before any treatment, becomes crucial and can be of support for industrial decision-making. The aim of the present study was to evaluate the ability of MIRS for the prediction of RCT and a_{30} of sheep bulk milk.

II. MATERIALS AND METHODS

A. Samples

A total of 4,303 bulk milk samples collected from sheep farms were available for analysis at the laboratory of the Istituto Zooprofilattico di Lazio and Toscana "Mariano Aleandri" (Rome, Italy). Samples belonged to farms (n = 344) located in Rome (165), Viterbo (151) and Rieti (28) provinces and were collected between January 2020 and December 2021. The breeds present were mainly Sarda, Lacaune, and Comisana. The Sarda breed accounted for more than half of the data. For each milk sample one aliquot was used for milk coagulation properties determination (10 mL) and the other was intended to MIRS spectrum acquisition.

B. Coagulation properties determination

Tubes used for sampling of bulk milk (about 50 mL) were kept at 4°C from collection until analysis. The lactodynamographic analysis was carried out using Formagraph (MaPe System, Firenze, Italy) by adding for each sample 200 µL of calf rennet solution (75% chymosin and 25% bovine pepsin; 175 international milk clotting units/mL; Caglificio Clerici Spa-Sacco Srl, Cadorago, Italy) diluted to 0.8% (w/w) in distilled water was added to milk. The steps required for the determination of coagulation traits through lactodynamographic analysis have been widely described in Vigolo et al. (2021) [4].

C. MIRS model development

For the milk spectra collection a MilkoScan FT7 (FOSS A/S, Hillerød, Denmark) was used. Aliquots were previously heated (37°C) and underwent homogenization by inversion before the analysis. The 1,060 wavelengths available were expressed in transmittance and were converted to absorbance as: log(1/transmittance). All the reference data were edited before the PLS with the purpose of removing outliers; the minimum and maximum were fixed at 1 and 45 min and at 12 and 91 mm for RCT and a_{30}, respectively. This restriction leaded to a loss of 187 (4.35%) records for RCT and 235 (5.46%) for a_{30}. The models for a_{30} and RCT were developed using partial least square regression (PLS) available in the WinISI II software (Infrasoft International Inc., State College, PA, US). Various combinations of mathematical treatments and spectra transformation were tested. Spectral pre-treatment involved first- and second-order derivatives. The gaps over which derivatives were calculated ranged from 1 to 20 data points and the smooth ranged from 1 to 8. The total database was divided into a calibration set (75% of the data for calibration and cross-validation) and a testing set (25% of the data for external validation).

III. Results

Descriptive statistics of reference data after editing are reported in Table I. The calibration and external validation sets were similar in terms of variability of the traits investigated.

TABLE I. Descriptive Statistics

Trait [a]	n	Mean	CV [b]	Min.	Max.
RCT, min	4,116	16.38	32	2.15	43.45
a_{30}, mm	4,068	52.91	21	12.80	89.02

[a] Rennet coagulation time (RCT) and curd firmness after 30 min from rennet addition (a_{30})

[b] Coefficient of variation (CV, %)

The mean and the standard deviation of RCT were equal to 16.35 and 5.26 min in the calibration set; in the external validation set, the same were 16.48 and 5.18 min. In the case of a_{30}, the calibration set was characterized by a mean of 52.78 mm and a standard deviation of 11.43 mm, whereas the external validation set presented a mean equal to 53.32 mm with a standard deviation of 11.08 mm.

TABLE II. Predictive ability

Trait [a]	Model performance [b]					
	n	R^2_{CV}	SE_{CV}	R^2_{V}	SE_{V}	*RPD*
RCT, min	2,853	0.50	3.41	0.42	3.97	1.31
a_{30}, mm	2,830	0.30	8.39	0.28	9.45	1.17

[a] Rennet coagulation time (RCT) and curd firmness after 30 min from rennet addition (a_{30})

[b] n = number of samples for calibration; R^2_{CV} = coefficient of determination of cross-validation; SE_{CV} = standard error of cross-validation; R^2_{V} = coefficient of determination of external validation; SE_{V} = standard error of external validation; RPD = ratio performance deviation, calculated dividing the standard deviation of reference data by the standard error of predictions (external validation).

The models performance are summarized in Table II. Both the traits showed a slightly greater coefficient of determination in cross-validation than in external validation. Overall, results demonstrate that the predictive ability of MIRS is greater for RCT than a_{30}. The plots of predicted and measured traits (external validation set) are reported in Fig. 1 and Fig. 2 for RCT and a_{30}, respectively.

IV. Discussion

The average milk coagulation ability of the samples of this study were in agreement with a previous study conducted on Sarda and Comisana sheep bulk milk [5] and with the study of [6] on individual milk of Sardinian ewes. Due to the small sample size, the coefficient of determination in external validation in [5] was lower compared to the present study and were far from being considered exploitable, being 0.28 for RCT and 0.02 for a_{30}. In fact, only 117 and 112 samples were available for testing the model of RCT and a_{30} [5]. However, the ratio performance deviation (RPD), which is the ratio of standard deviation of reference data to the standard error of predictions, of both the traits (1.18 for RCT and 1.01 for a_{30}) was similar to that obtained in this study (Table II). Some authors investigated the ability of MIRS for prediction of coagulation properties in individual sheep milk using various spectra pre-treatments [7]. Authors reported an exploitable coefficient of determination for RCT, ranging from 0.49 to 0.59 [7]. On the other hand, as observed in this study, the a_{30} was more difficult to predict compared to RCT. The coefficient of determination, in fact, ranged from 0.31 to 0.42. It is worth to highlight that only 970 samples were used for developing the models [7]. Findings suggest that bulk milk of ewes supplied to dairies can undergo a preliminary screening for its RCT. The possibility of having expected RCT will allow cheese manufacturers to evaluate the coagulation aptitude of the milk received. Together with the milk composition, RCT may be used for payment system, e.g., by introducing premiums or penalties for optimal or suboptimal RCT. Furthermore, whether predictions in individual milk will be available there may opportunity for selective breeding.

References

[1] AIA, 2021. http://bollettino.aia.it/. Accessed: 13/01/2023

[2] M. De Marchi et al., "Prediction of coagulation properties, titratable acidity, and pH of bovine milk using mid-infrared spectroscopy," J. Dairy Sci., vol. 92, p. 423-432, 2009.

[3] G. Visentin et al., "Prediction of bovine milk technological traits from mid-infrared spectroscopy analysis in dairy cows," J. Dairy Sci., vol. 98, p. 6620-6629, 2015.

[4] V. Vigolo, M. Franzoi, M. Penasa, and M. De Marchi, "β-Casein variants differently affect bulk milk mineral content, protein composition, and technological traits," Int. Dairy J., vol. 124, 105221, 2022.

[5] C.L. Manuelian, M. Penasa, G. Giangolini, C. Boselli, S. Currò, and M. De Marchi, "Short communication: Fourier-transform mid-infrared spectroscopy to predict coagulation and acidity traits of sheep bulk milk," J. Dairy Sci. vol 102, p. 1927-1932, 2019.

[6] F. Correddu, G. Gaspa, A. Cesarani, and N.P.P. Macciotta, "Phenotypic and genetic characterization of the occurrence of noncoagulating milk in dairy sheep," J. Dairy Sci., vol. 105, p. 6773-6782, 2022.

[7] M. Cellesi et al., "Prediction of milk coagulation properties and individual cheese yield in sheep using partial least squares regression," Animals, vol. 9, 663, 2019.

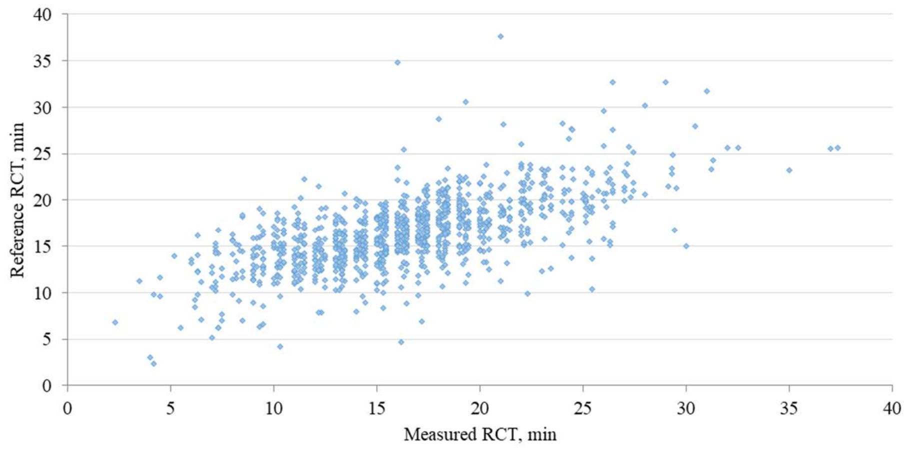

Fig. 1. Plot of predicted (x-axis) and reference data (y-axis) for the bulk milk RCT for the external validation set.

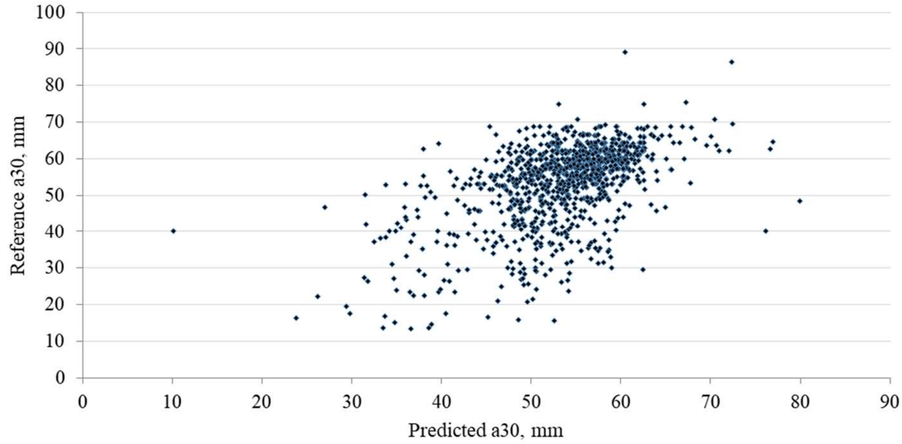

Fig. 2. Plot of predicted (x-axis) and reference data (y-axis) for the bulk milk a_{30} for the external validation set.