J. Dairy Sci. 104
https://doi.org/10.3168/jds.2020-19587

Goat farm variability affects milk Fourier-transform infrared spectra used for predicting coagulation properties

Christos Dadousis,[1] Claudio Cipolat Gotet,[1] Giorgia Stocco,[1]* Alessandro Ferragina,[2] Maria L. Dettori,[3] Michele Pazzola,[3] Adriano Henrique do Nascimento Rangel,[4] and Giuseppe M. Vacca[3]

[1]Department of Veterinary Science, University of Parma, 43126 Parma, Italy
[2]Food Quality and Sensory Science Department, Teagasc Food Research Centre, D15 KN3K, Ireland
[3]Department of Veterinary Medicine, University of Sassari, 07100 Sassari, Italy
[4]Agricultural School of Jundiaì, Federal University of Rio Grande do Norte, 59078970 Natal, Brazil

ABSTRACT

Driven by the large amount of goat milk destined for cheese production, and to pioneer the goat cheese industry, the objective of this study was to assess the effect of farm in predicting goat milk-coagulation and curd-firmness traits via Fourier-transform infrared spectroscopy. Spectra from 452 Sarda goats belonging to 14 farms in central and southeast Sardinia (Italy) were collected. A Bayesian linear regression model was used, estimating all spectral wavelengths' effects simultaneously. Three traditional milk-coagulation properties [rennet coagulation time (min), time to curd firmness of 20 mm (min), and curd firmness 30 min after rennet addition (mm)] and 3 curd-firmness measures modeled over time [rennet coagulation time estimated according to curd firmness change over time (RCT_{eq}), instant curd-firming rate constant, and asymptotical curd firmness] were considered. A stratified cross validation (SCV) was assigned, evaluating each farm separately (validation set; VAL) and keeping the remaining farms to train (calibration set) the statistical model. Moreover, a SCV, where 20% of the goats randomly taken (10 replicates per farm) from the VAL farm entered the calibration set, was also considered (SCV_{80}). To assess model performance, coefficient of determination (R^2_{VAL}) and the root mean squared error of validation were recorded. The R^2_{VAL} varied between 0.14 and 0.45 (instant curd-firming rate constant and RCT_{eq}, respectively), albeit the standard deviation was approximating half of the mean for all the traits. Although average results of the 2 SCV procedures were similar, in SCV_{80}, the maximum R^2_{VAL} increased at about 15% across traits, with the highest observed for time to curd firmness of 20 mm (20%) and the lowest for RCT_{eq} (6%). Further investigation evidenced important variability among farms, with R^2_{VAL} for some of them being close to 0. Our work outlined the importance of considering the effect of farm when developing Fourier-transform infrared spectroscopy prediction equations for coagulation and curd-firmness traits in goats.

Key words: goat, coagulation, curd firmness, farm, infrared spectra

INTRODUCTION

A large proportion of the world's goat milk is destined to cheese production, especially in those countries in the Mediterranean basin (FAOSTAT, 2018). This region is characterized by adverse weather and environmental conditions in which autochthonous goat breeds are well adapted and usually managed in extensive or semi-extensive management types (Di Trana et al., 2015; Stella et al., 2018). It has been shown that the farming system used represents a very large source of variation (ranging from 16–70% of the total variability) in milk composition and milk processing characteristics, such as the coagulation properties (Pazzola et al., 2018b). These values are greater compared with those of bovine (from 9–16%; Bittante et al., 2015) and ovine (from 16–43%; Vacca et al., 2015) farming methods. Indeed, a great variability of goat farming has been reported (Usai et al., 2006). The importance of the type of farming system is related to the destination of the milk produced and the genetics of the animals (Pazzola et al., 2018b). For instance, indigenous breeds are more suitable for harsh environments and extreme extensive management (Di Trana et al., 2015), and they are able to produce a milk characterized by better composition (e.g., high milk fat and protein) and technological characteristics than that from cosmopolitan breeds (Čermak et al., 2013; Paschino et al., 2020).

Among the milk technological characteristics, traditional milk-coagulation properties (**MCP**) are widely

Received September 4, 2020.
Accepted November 13, 2020.
*Corresponding author: giorgia.stocco@unipr.it

used to describe the complex process of cheesemaking. Moreover, the extension of MCP through the calibration of the curd firmness (**CF**) as a function of time ($\mathbf{CF_t}$) provides a more complete overview of the coagulation process (Bittante, 2011). There is extensive and well-documented literature on the importance and relevance of MCP, mainly in cattle (Bittante et al., 2012; Stocco et al., 2017; Nilsson et al., 2019), but also in sheep (Caballero-Villalobos et al., 2018; Cipolat-Gotet et al., 2018) and, to a lesser extent, in goats (Vacca et al., 2020). In addition, MCP show heritability estimates between 0.15 and 0.27 in cattle (Dadousis et al., 2016) and between 0.09 and 0.19 in sheep (Bittante et al., 2017). Hence, directional selection on desirable MCP characteristics is applicable. This could be of particular interest in goats, especially for those breeds (e.g., Alpine, Toggenburg) characterized by weak or nonexpressing alleles (e.g., F, N allele) of α_{S1}-casein, associated with unfavorable coagulation process (Maga et al., 2009; Devold et al., 2011). It worth noting that genetic pattern at a specific locus might change over time because of the selective pressure, as recently evidenced in casein genes for the Murciano-Granadina goat breed (Pizarro et al., 2020). However, high MCP analysis costs and logistics pose restriction for their wide-scale application.

Currently, a potential solution to overcome those limitations can be derived via Fourier-transform infrared (**FTIR**) spectroscopy. Indeed, there is an increasing interest in the dairy sector on the usefulness of FTIR information for the prediction of a variety of phenotypes (Tiplady et al., 2019), either directly measurable in milk (e.g., fatty acids; Soyeurt et al., 2006) or related to the milk processing characteristics (e.g., cheesemaking traits, MCP; Ferragina et al., 2013; Visentin et al., 2015) and the animal condition (e.g., energy efficiency, lameness; McParland and Berry, 2016; Bonfatti et al., 2020). In dairy cattle, recent advanced research made MCP predictions via FTIR spectroscopy applicable in the milk payment system of some Protected Designation of Origin (PDO) cheese consortia to reward or penalize dairy farmers (e.g., Trentigrana PDO cheese; Benedet et al., 2018). In the case of small ruminants, the practical use of the FTIR predictions along the dairy chain is still lacking. Although there is ongoing research in sheep on the use of FTIR spectroscopy for the prediction of MCP and CF_t parameters (Ferragina et al., 2017; Correddu et al., 2019), up to present, there are no data available in goats.

An important factor to consider when developing prediction equations via milk FTIR spectra is the structure of the data, especially for traits not directly measurable in milk (e.g., technological traits, animal health, environment). In bovine milk, it has been shown that a random cross validation might overestimate the prediction accuracy of methane emission traits (Wang and Bovenhuis, 2019). Rather, a stratified cross validation (**SCV**) where, for example, each farm is evaluated separately, might provide a more realistic model assessment (Wang and Bovenhuis, 2019). In previous studies, great variability was observed in different goat farming systems (Usai et al., 2006) and in MCP and CF_t parameters among individual farms (Pazzola et al., 2018b; Vacca et al., 2018). Hence, the type of goat farm is a factor that should be assessed, and its effect should be quantified for FTIR prediction models for MCP and CF_t parameters.

Altogether, the economic importance of MCP and CF_t parameters in the dairy sector justifies further investigation on the practical application at a wide scale of milk FTIR spectroscopy to predict MCP and CF_t parameters because these applications could pioneer the entire goat cheese industry at the farm, breeding, and dairy plant levels. To this purpose, our objectives were to (1) investigate the potential of milk FTIR spectroscopy for the prediction of MCP and CF_t parameters in goats and (2) quantify the effect of the farm variability on the prediction accuracy of MCP and CF_t parameters using individual Sarda goat milk samples.

MATERIALS AND METHODS

Farm Characteristics, Milk Sampling, and Analyses

The study involved 452 Sarda goats reared in 14 farms (**F01**–**F14**) distributed across the island of Sardinia (Italy). Sampled farms were officially registered in the flock book and recording system of provincial associations of goat breeders. Farm characteristics are summarized in Table 1. In brief, the extensive system consisted of family-managed farms with pasture feeding, natural mating, and milking on the return of goats from pasture; the semi-extensive system was characterized by cultivated grasslands, control of estrus, and control of kidding season.

Individual milk samples (100 mL/goat) were collected during the afternoon milking (1 sampling day for each farm). Milk was sampled from the recorder jar under each stall in mechanical milking systems, and from the stainless steel graduated pails in the hand-milked systems, over the entire milking of each goat. Milk samples were then stored at 4°C and analyzed within 24 h after collection. For each individual milk sample, 2 measurements of MCP were performed using a lactodynamograph (Formagraph; Foss Electric A/S, Hillerød, Denmark) during a 30 min test analysis, following the

Table 1. Characteristics of sampled farms (n = 14)

Item	Management system[1]	
	Extensive	Semi-extensive
Farms, number	6	8
Goats, number	183	269
Flock size, number of farms		
Small (<100 goats)	1	1
Medium (100–200 goats)	3	5
Large (>200 goats)	2	2
Altitude, number of farms		
Plain (<200 m asl[2])	3	2
Hill (200–500 m asl)	2	4
Mountain (>500 m asl)	1	2
Milking, number of farms		
Mechanical	3	4
Hand-milked	3	4
Milk quality, mean ± SD		
Fat, %	5.01 ± 0.98	5.33 ± 1.32
Protein, %	3.97 ± 0.52	3.87 ± 0.51
SCS[3]	6.58 ± 1.64	6.75 ± 1.68
LBC[4]	1.80 ± 0.91	1.71 ± 0.86

[1]Extensive system: family-managed farms, feeding at pasture, natural mating, milking when goats are back from pasture; semi-extensive system: cultivated grasslands, control of estrus and kidding season.

[2]asl = above sea level.

[3]SCS = $\log_2$ (SCC × 10^{-5}) + 3.

[4]LBC = logarithmic total bacterial count = $\log_{10}$ (total bacterial count/1,000).

procedure reported by Pazzola et al. (2018b). In brief, 10 mL of milk (run in duplicate) for each sample were heated to 35°C for 15 min and then mixed with 200 μL of the rennet solution [Hansen Naturen Plus 215 (Pacovis Amrein AG, Bern, Switzerland), with 80 ± 5% chymosin and 20 ± 5% pepsin; 215 international milk clotting units/mL; diluted to 1.2% (wt/vol) in distilled water to reach the final value of 0.0513 international milk clotting units/mL of milk]. Coagulation process occurred at 35°C. The MCP recorded were as follows: rennet coagulation time (**RCT**, min), time to CF of 20 mm ($\mathbf{k_{20}}$, min), and CF 30 min after rennet addition ($\mathbf{a_{30}}$, mm).

During lactodynamographic analysis, the Formagraph instrument recorded every 15 s the width (mm) of the oscillatory graph designed by the pendula immerged in the milk samples after rennet addition. Consequently, 120 CF observations were recorded for each individual milk sample. The 30 min test analysis allowed us to use the following 3-parameter model (Bittante, 2011):

$$CF_t = CF_P \times \left(1 - e^{-k_{CF}\left(t - RCT_{eq}\right)}\right),$$

where CF_t = curd firmness at time t (mm); CF_P = the asymptotical potential value of CF at an infinite time in absence of syneresis (mm); k_{CF} = the curd-firming instant rate constant (%/min); and RCT_{eq} = RCT estimated by CF_t equation on the basis of all data points (min). Values of the aforementioned traits outside of the interval of the mean ± 3 standard deviations (**SD**) were considered outliers and excluded from further analysis.

For each milk sample, a FTIR spectrophotometer (MilkoScan FT6000; Foss Electric A/S) was used to assess milk composition (fat and protein; ISO-IDF 2013) and to collect the spectrum over the range from wavenumber 5,011 to 925 × cm^{-1}. Spectra were stored as absorbance (A) using the transformation A = log(1/T), where T is the transmission. Two spectral acquisitions were performed for each sample, and the results were averaged before data analysis.

Somatic cell count was determined by Fossomatic 5000 (Foss Electric A/S) according to ISO-IDF (2006) standards, and later transformed into the logarithmic SCS [SCS = $\log_2$(SCC × 10^{-5}) + 3; Ali and Shook, 1980]. Total bacterial count was determined using a BactoScan FC150 analyzer (Foss Electric A/S) according to ISO-IDF (2004) standards, and transformed into the logarithmic bacterial count [logarithmic bacterial count = $\log_{10}$ (total bacterial count/1,000)].

Statistical Analysis and FTIR Spectra

Modeling and Repeatability of Coagulation Traits. Files containing the 120 CF values for each milk sample were processed fitting a curvilinear regression with the PROC NLIN procedure (SAS Institute Inc., Cary, NC). The parameters of each individual equation were estimated with the Marquardt iterative method (350 iterations and 10^{-5} level of convergence).

To estimate the coefficient of repeatability (%), MCP and CF_t parameters (2 replicates per goat), were analyzed using a MIXED procedure (SAS Institute Inc., Cary, NC) that included the random effects of farm, animal, pendulum (measuring unit of the Formagraph instrument), as well as the the residual. The coefficient of repeatability (**REP**, %) for MCP and CF_t parameters was then calculated as the ratio of the sum of the variances of the random effects of farm, animal, and pendulum to the total variance.

Spectra Editing and Chemometric Model. Before spectra analysis, the absorbance values of every wavelength in the FTIR spectra of the milk samples were centered and standardized to a null mean and a unit sample variance. To detect outliers, Mahalanobis distances were calculated by means of the Mahalanobis function implemented in R software (R Core Team, 2013). No samples were discarded because all spectra

presented a distance value lower than the mean ± 3 SD. The spectra were not subjected to any other mathematical pretreatment.

A Bayesian linear regression was used to predict the RCT, k_{20}, a_{30}, RCT_{eq}, k_{CF}, and CF_P. All phenotypes were regressed to 1,060 spectra under the following model: $y = \mu + \sum_{j=1}^{1,060} x_{ij}\beta_j + e_i$, where y is the analyzed trait, μ = the overall mean, x_{ij} = the FTIR wavelengths of the ith sample (j = 1 to 1,060), β_j = the regression coefficients, and e_i = residual with $iid \sim N\left(0, \sigma_e^2\right)$. The BayesB model implemented in the *BGLR* R package was adopted (de los Campos and Perez-Rodriguez, 2015) as described in Ferragina et al. (2017).

Stratified Cross Validation Procedures. An external SCV scheme was used to assess the model's predictive ability, where 1 farm at a time consisted of the validation set (**VAL**). Goats from the remaining farms comprised the calibration (**CAL**) set. The procedure was repeated 14 times, such that all farms were evaluated. In addition, to assess the importance of shared variability between CAL and VAL, a SCV where 20% of the goats from 1 farm to be validated was included in CAL, and the VAL set consisted of the remaining 80% of the goats from the evaluated farm, was considered (referred to as $\mathbf{SCV_{80}}$ hereafter). To account for individual sampling variability, the 20% of the goats for the SCV_{80} was sampled at random, and the procedure was repeated 10 times per farm. Results from SCV were averaged across the 14 farms and, in the SCV_{80}, over the 10 replicates per farm. For all calibrations, model performance was measured using the coefficient of determination ($\mathbf{R^2}$), the root mean squared error, and the SD of both CAL and VAL sets.

RESULTS AND DISCUSSION

Prediction Accuracy of Coagulation Traits in Goat Milk

Descriptive statistics and prediction results of the SCV are presented in Table 2. Mean values were consistent with those reported in the Sarda goat milk literature (Pazzola et al., 2018a). Repeatability of coagulation traits ranged from 98% (for RCT and RCT_{eq}) to 84% (for k_{CF} and CF_P). The CF measurements (a_{30} and CF_P traits) are generally characterized by a reduced instrumental repeatability and reproducibility in later time after rennet addition, which is more profound after gelation (Ferragina et al., 2017). Compared with other species, repeatability values of goat RCT, RCT_{eq}, and CF_P traits were similar to that of bovine (Stocco et al., 2017) and ovine (Ferragina et al., 2017). Goat milk is generally characterized by slower increase of CF, weaker casein network forming after gelation, and earlier syneresis compared with bovine and ovine milk (Inglingstad et al., 2014; Pazzola et al., 2018b; Roy et al., 2020). Because of these characteristics of the goat coagulation process and because the traditional lactodynamograph set up for analysis of bovine milk was designed to explore primarily the coagulation and the first part of curd-firming process, not syneresis, a slight decrease of repeatability of CF measurements after RCT was expected. For this reason, REP is commonly very high for the first traits measured (e.g., RCT and RCT_{eq}) and tends to decrease over time both in the case of traditional and modeled coagulation traits (Stocco et al., 2015). This phenomenon is explained by the fact that, during the test, the variation related to the curd firming and syneresis tends to accumulate over time. In the present study, only a_{30} showed a higher REP value than those reported for bovine (Stocco et al., 2017) and ovine milk (Ferragina et al., 2017). This could be due to the fact that milk from Sarda goats of the present study is characterized by very good milk quality (e.g., high fat and protein contents; Table 1) and better coagulative aptitude, faster gelation and curd firming times, and firmer coagulum than other dairy goat breeds (e.g., Alpine, Saanen; Vacca et al., 2018). Among the factors influencing the reliability of the FTIR predictions, the goodness (repeatability and accuracy) of the reference values is very important (Caredda et al., 2016). Indeed, it is interesting to notice that the prediction accuracy decreased with progressed coagulation (e.g., higher for RCT and lower for a_{30}), along with decreasing REP values (Table 2).

Regarding SCV predictions (Table 2), RCT and RCT_{eq} showed the highest R^2_{CAL} (0.64 and 0.61, respectively), followed by CF_P ($R^2_{CAL} = 0.50$). The remaining traits had $R^2_{CAL} < 0.50$, and the lowest was observed for k_{CF} (0.37). In general, results in the CAL set were comparable to those reported in ovine milk (Ferragina et al., 2017), in particular for the traits directly related to CF (a_{30} and CF_P). In the VAL set, the R^2_{VAL} was lower and the root mean squared error was higher, albeit with much higher SD for both parameters compared with CAL, and the ranking among traits was analogous to the CAL. Because this was the first study to investigate the effect of farm on the prediction accuracy of MCP and CF_t parameters in goat milk via FTIR spectroscopy, comparison with literature was restricted. However, a recent study (Stocco et al., 2021) assessing the goat breed (4 breeds considered) effect on the prediction of MCP and CF_t parameters via FTIR spectroscopy, by using a random 5-fold cross validation

Table 2. Descriptive statistics and repeatability (REP) of traditional milk-coagulation properties (MCP) and curd firmness over time (CF_t) model parameters and results from stratified cross validation (SCV) calibrations using mid-infrared spectra of individual goat milk samples

	Descriptive statistics[2]				Prediction statistics[3]							
					Calibration (CAL)				Validation (VAL)				
Item[1]	n	Mean	SD	REP	n	SD_{CAL}	R^2_{CAL}	$RMSE_{CAL}$	SD_{VAL}	R^2_{VAL}	$R^2_{interval}$	$RMSE_{VAL}$	$RMSE_{VAL}^{interval}$
Traditional MCP													
RCT, min	892	12.9	4.42	97.7	416	4.49 ± 0.18	0.64 ± 0.04	2.69 ± 0.16	3.95 ± 0.92	0.42 ± 0.21	0.00–0.75	3.69 ± 1.24	2.23–6.29
k_{20}, min	839	3.5	1.12	85.6	397	1.12 ± 0.01	0.49 ± 0.05	0.80 ± 0.04	0.99 ± 0.26	0.29 ± 0.21	0.00–0.61	0.91 ± 0.16	0.72–1.27
a_{30}, mm	901	37.5	11.0	87.2	422	11.0 ± 0.34	0.47 ± 0.03	8.07 ± 0.33	10.2 ± 3.27	0.27 ± 0.17	0.02–0.63	9.63 ± 3.49	6.47–17.83
CF_t parameters													
RCT_{eq}, min	892	13.6	4.06	97.7	415	4.06 ± 0.14	0.61 ± 0.05	2.55 ± 0.15	3.65 ± 0.80	0.45 ± 0.20	0.00–0.80	3.30 ± 1.04	2.10–5.32
k_{CF}, %/min	867	22.9	8.46	84.2	408	8.21 ± 0.17	0.37 ± 0.07	6.60 ± 0.37	7.91 ± 2.00	0.14 ± 0.16	0.00–0.56	8.28 ± 2.16	5.01–12.69
CF_p, mm	873	42.7	9.31	83.9	409	8.91 ± 0.15	0.50 ± 0.01	6.31 ± 0.14	7.98 ± 1.68	0.32 ± 0.19	0.00–0.69	7.19 ± 1.84	4.90–11.04

[1]Traditional milk-coagulation properties: RCT = rennet coagulation time; k_{20} = curd-firming time; a_{30} = curd firmness 30 min after rennet addition. CF_t model parameters according to 3-parameter model: RCT_{eq} = RCT estimated according to curd firm change over time modeling; k_{CF} = instant curd-firming rate constant; CF_p = asymptotical curd firmness.

[2]$REP, \% = \frac{\sigma^2_{Farm} + \sigma^2_{Animal} + \sigma^2_{Pendulum}}{\sigma^2_{Farm} + \sigma^2_{Animal} + \sigma^2_{Pendulum} + \sigma^2_e} \times 100.$

[3]Average ± SD from the SCV calibrations. For R^2_{VAL} and $RMSE^2_{VAL}$, intervals of validations were also included. Results were averaged over the 14 runs (1 per farm). RMSE = root mean squared error.

procedure, reported R^2_{VAL} from 0.42 to 0.68 for MCP (RCT and a_{60}, respectively) and from 0.14 to 0.60 for CF_t parameters (syneresis rate and CF_P, respectively). The study also confirmed decreased prediction accuracies in a SCV scenario (using 3 breeds as CAL, and the remaining breed as VAL set), suggesting the importance of considering the breed of goats when developing FTIR calibrations. Similar to those results, our study showed the importance of considering the differences among farms on the prediction accuracy of MCP and CF_t parameters. This variability was evident when observing the high SD of both R^2_{VAL} and $RMSE_{VAL}$ (Table 2), which were higher compared with a previous study on the same traits and statistical methodology in sheep (Ferragina et al., 2017).

Effect of Farm Variability on the Prediction Accuracy of Coagulation Traits

By including 20% of the VAL farm in the training set (SCV_{80}), our expectation was to increase R^2_{VAL} because important variation was included in the model training, and also because by using this approach, CAL and VAL data sets were not completely independent (Figure 1). On average, R^2_{VAL} remained the same as the SCV procedure, and was of 0.45, 0.32, 0.29, 0.44, 0.17, and 0.33 for RCT, k_{20}, a_{30}, RCT_{eq}, k_{CF}, and CF_P, respectively, with similar SD to the SCV (data not shown). However, although the minimum R^2_{VAL} was again close to 0, the maximum obtained R^2_{VAL} values were increased (0.87, 0.73, 0.73, 0.85, 0.65, and 0.79 for RCT, k_{20}, a_{30}, RCT_{eq}, k_{CF}, and CF_P, respectively), representing an increase of ~20% for k_{20}, ~16% each for RCT, a_{30}, and k_{CF}, and ~14% for CF_P, with the minimum (~0.06%) for the RCT_{eq}. On average, R^2_{VAL} results for each coagulation trait among farms presented in Figure 1 were analogous to the SCV, albeit with no repetitions per farm in that case. A considerable R^2_{VAL} variation among farms was observed (Figure 1). Interaction between farm and trait was also present. More precisely, across the traits, we observed the following: (1) farms with either low or high variability of prediction model performance (e.g., F02 and F11 for RCT, respectively), (2) consistent high or low R^2_{VAL} values relative to the remaining farms across the traits (e.g., F02 vs. F12), (3) different R^2_{VAL} patterns, showing either high or low R^2_{VAL} (e.g., F01 and F10 comparing k_{CF} to all the rest of the traits), (4) general low predictability of k_{CF} trait, with 3 farms (F01, F04, and F08) showing R^2_{VAL} close to 0, (5) similar variation patterns across farms of RCT and RCT_{eq} traits, and interestingly, (6) R^2_{VAL} close to 0 across all traits in F12. The overall model performance presented in Table 2 and Figure 1 was clearly improved (data not shown) when excluding this specific farm

(F12). It is important to consider that the region where milk samples were collected has been characterized for decades by extensive and semi-extensive goat farming management, which is highly variable among areas of the island (Usai et al., 2006). As aforementioned, the variability of farms affects both composition and coagulation ability of goat milk (Pazzola et al., 2018b; Vacca et al., 2018). Hence, variability of R^2_{VAL} among farms was, up to an extent, expected. In particular, 2 of the farms (F11 and F12) were located in an area with high altitude and adverse environmental conditions. Those factors, together with the lower hygienic control practiced by the farmers over the goats (the flocks were let free to graze without supervision in extensive farms), represent a source of milk quality variation (Pazzola et al., 2018b) that further influences the processing characteristics. For example, changes occurring at the milk composition and coagulation level often caused by bacterial or SCC are well documented in goats (Barrón-Bravo et al., 2013; Stocco et al., 2019). In addition, the high genetic variability characterizing the Sarda breed (Dettori et al., 2015; Pazzola et al., 2018a), as well as other nongenetic factors (e.g., parity, days in milk), might have caused the large differences in the R^2_{VAL} values among farms. It is important to consider that the cross validation procedure is usually used to evaluate the performance of prediction equations where data are split randomly into a CAL and

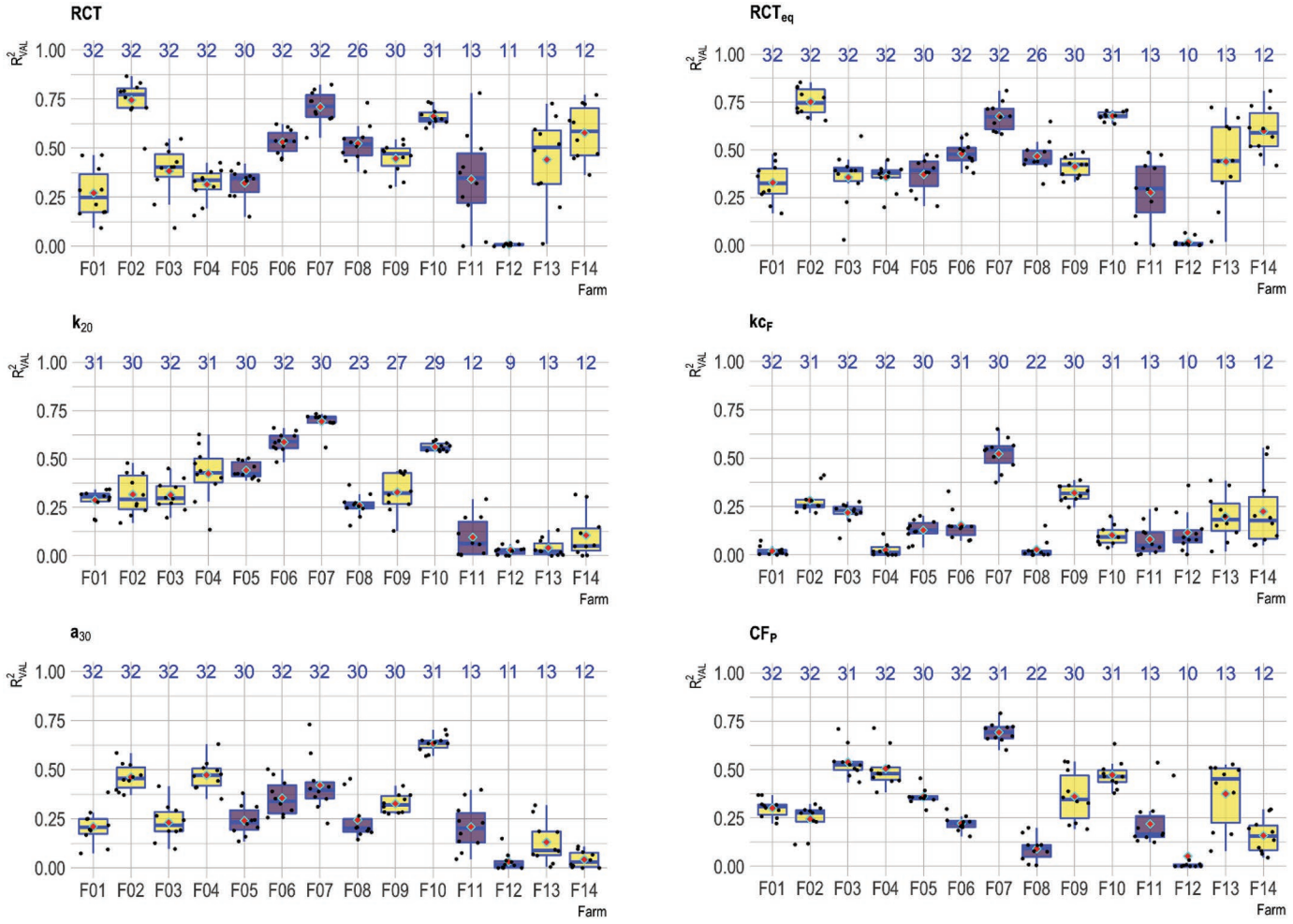

Figure 1. Coefficient of determination of validation (R^2_{VAL}) results per farm (F01 to F14; purple boxes refer to extensive farms; yellow boxes refer to semi-extensive farms) of traditional milk-coagulation properties (MCP) and curd firmness over time (CF_t) model parameters using mid-infrared spectra of individual goat milk samples in the second stratified cross validation scenario (SCV_{80}). Traditional MCP: RCT = rennet coagulation time; k_{20} = curd-firming time; a_{30} = curd firmness 30 min after rennet addition. CF_t model parameters according to 3-parameter model: RCT_{eq} = RCT estimated according to curd firm change over time modeling; k_{CF} = instant curd-firming rate constant; CF_p = asymptotical curd firmness. Each farm was evaluated separately with 20% of the farm included in the calibration set. The procedure was repeated 10 times per farm (black dots); vertical lines within each boxplot represent the median, and red rhombus is the mean of the 10 replicates per farm; blue numbers on top refer to the number of goats in validation per farm.

a VAL set. However, it has been demonstrated that when there are dependence structures in the data, cross validation may overestimate prediction accuracies (Roberts et al., 2017). In particular, Qin et al. (2016) indicated that random cross validation underestimates the error of the prediction equation when traits to be predicted are analyzed in batches that have systematic differences among them. In our case, because of the differences among farms within farming systems (Table 1), we chose to build calibration equations directly at the farm level to consider the differences in milk-coagulation traits (and therefore in the milk spectra) that came from the differences among farms. Wang and Bovenhuis (2019) investigated the feasibility of bovine milk infrared spectra to predict methane emissions by comparing random and block cross validation (using farms as blocks) procedures. They showed R^2_{VAL} values of 0.49 and 0.01, respectively, for random and block cross validation. They suggested that the difference in the prediction accuracy between the 2 procedures could have been due to the confounding effect of farm and date of milk infrared collection, and especially to the breath sensors used to measure methane emissions, which largely differed among farms.

CONCLUSIONS

Overall, our work evidenced the feasibility of using FTIR spectroscopy to predict MCP and CF_t parameters in goat milk. Despite this, a great variability was observed among farms and traits. The generally low R^2_{VAL} do not justify the practical application, at present, of the predicted coagulation traits. However, among traits, RCT and RCT_{eq} showed the highest accuracies, and k_{CF} showed the lowest accuracies. Moreover, our results demonstrated the importance of farm variability in relation to coagulation traits, which should be considered when developing FTIR calibrations to avoid misleading results. Future studies with other farming systems, statistical models, and increased sample sizes are expected to show improvements in the model performance. A further investigation on the predictive performance of FTIR on individual cheese yield traits would be interesting.

ACKNOWLEDGMENTS

This research was supported by the "Fondo di Ateneo per la Ricerca 2019" (Finanziamento straordinario una tantum per la ricerca, one-time extraordinary research grant) University of Sassari (Sassari, Italy). The authors thank the farmers for giving access to their flocks; the Provincial Farmers Associations (Associazioni Interprovinciali e Provinciali degli Allevatori, AIPA/APA) of Cagliari, Nuoro, Sassari, and Oristano (Italy) for their support in sample collection; and the Regional Farmer Association of Sardinia (Associazione Regionale degli Allevatori, ARA Sardegna, Cagliari, Italy) for support in milk analysis. The authors have not stated any conflicts of interest.

REFERENCES

Ali, A. K. A., and G. E. Shook. 1980. An optimum transformation for somatic cell concentration in milk. J. Dairy Sci. 63:487–490.

Barrón-Bravo, O. G., A. J. Gutiérrez-Chávez, C. A. Ángel-Sahagún, H. H. Montaldo, L. Shepard, and M. Valencia-Posadas. 2013. Losses in milk yield, fat and protein contents according to different levels of somatic cell count in dairy goats. Small Rumin. Res. 113:421–431. https://doi.org/10.1016/j.smallrumres.2013.04.003.

Benedet, A., C. L. Manuelian, M. Penasa, M. Cassandro, F. Righi, M. Sternieri, P. Galimberti, A. V. Zambrini, and M. De Marchi. 2018. Factors associated with herd bulk milk composition and technological traits in the Italian dairy industry. J. Dairy Sci. 101:934–943. https://doi.org/10.3168/jds.2017-12717.

Bittante, G. 2011. Modeling rennet coagulation time and curd firmness of milk. J. Dairy Sci. 94:5821–5832. https://doi.org/10.3168/jds.2011-4514.

Bittante, G., C. Cipolat-Gotet, F. Malchiodi, E. Sturaro, F. Tagliapietra, S. Schiavon, and A. Cecchinato. 2015. Effect of dairy farming system, herd, season, parity and days in milk on modeling of the coagulation, curd firming and syneresis of bovine milk. J. Dairy Sci. 98:2759–2774. https://doi.org/10.3168/jds.2014-8909.

Bittante, G., C. Cipolat-Gotet, M. Pazzola, M. L. Dettori, G. M. Vacca, and A. Cecchinato. 2017. Genetic analysis of coagulation properties, curd firming modeling, milk yield, composition and acidity in Sarda dairy sheep. J. Dairy Sci. 100:385–394. https://doi.org/10.3168/jds.2016-11212.

Bittante, G., M. Penasa, and A. Cecchinato. 2012. Invited review: Genetics and modeling of milk coagulation properties. J. Dairy Sci. 95:6843–6870. https://doi.org/10.3168/jds.2012-5507.

Bonfatti, V., P. N. Ho, and J. E. Pryce. 2020. Usefulness of milk mid-infrared spectroscopy for predicting lameness score in dairy cows. J. Dairy Sci. 103:2534–2544. https://doi.org/10.3168/jds.2019-17551.

Caballero-Villalobos, J., A. Figueroa, K. Xibrraku, E. Angón, J. M. Perea, and A. Garzón. 2018. Multivariate analysis of the milk coagulation process in ovine breeds from Spain. J. Dairy Sci. 101:10733–10742. https://doi.org/10.3168/jds.2018-14752.

Caredda, M., M. Addis, I. Ibba, R. Leardi, M. F. Scintu, G. Piredda, and G. Sanna. 2016. Prediction of fatty acid in sheep milk by mid-infrared spectrometry with a selection of wavelengths by genetic algorithms. Lebensm. Wiss. Technol. 65:503–510. https://doi.org/10.1016/j.lwt.2015.08.048.

Čermak, B., V. Kral, J. Frelich, L. Bohačova, B. Vondraškova, J. Špička, E. Samkova, M. Podsedniček, A. Węglarz, J. Makulska, and P. Zapletal. 2013. Quality of goat pasture in less-favoured areas (LFA) of the Czech Republic and its effect on fatty acid content of goat milk and cheese. Anim. Sci. Pap. Rep. 31:331–346.

Cipolat-Gotet, C., M. Pazzola, A. Ferragina, A. Cecchinato, M. L. Dettori, and G. M. Vacca. 2018. Technical note: Improving modeling of coagulation, curd firming, and syneresis of sheep milk. J. Dairy Sci. 101:5832–5837. https://doi.org/10.3168/jds.2017-14256.

Correddu, F., M. Cellesi, J. Serdino, M. G. Manca, M. Contu, M. C. Dimauro, I. Ibba, and N. P. P. Macciotta. 2019. Genetic parameters of milk fatty acid profile in sheep: Comparison between gas chromatographic measurements and Fourier-transform IR spectroscopy predictions. Animal 13:469–476. https://doi.org/10.1017/S1751731118001659.

Dadousis, C., S. Biffani, C. Cipolat-Gotet, E. L. Nicolazzi, A. Rossoni, E. Santus, G. Bittante, and A. Cecchinato. 2016. Genome-wide association of coagulation properties, curd firmness modeling, pro-

tein percentage, and acidity in milk from Brown Swiss cows. J. Dairy Sci. 99:3654–3666. https://doi.org/10.3168/jds.2015-10078.

de los Campos, G., and P. Perez Rodriguez. 2015. BGLR: Bayesian Generalized Linear Regression. R package version 1.0.4. Accessed May 25, 2019. http://CRAN.R-project.org/package=BGLR.

Dettori, M. L., M. Pazzola, P. Paschino, M. G. Pira, and G. M. Vacca. 2015. Variability of the caprine whey protein genes and their association with milk yield, composition and renneting properties in the Sarda breed. 1. The LALBA gene. J. Dairy Res. 82:434–441. https://doi.org/10.1017/S0022029915000461.

Devold, T. G., R. Nordbø, T. Langsrud, C. Svenning, M. J. Brovold, E. S. Sørensen, B. Christensen, T. Ådnøy, and G. E. Vegarud. 2011. Extreme frequencies of the α_{S1}-casein "null" variant in milk from Norwegian dairy goats - Implications for milk composition, micellar size and renneting properties. Dairy Sci. Technol. 91:39–51.

Di Trana, A., L. Sepe, P. Di Gregorio, M. A. Di Napoli, D. Giorgio, A. R. Caputa, and S. Claps. 2015. The role of local sheep and goat breeds and their products as a tool for sustainability and safeguard of the Mediterranean environment. Pages 77–112 in The Sustainability of Agro-Food and Natural Resources Systems in the Mediterranean Basin. Springer International Publishing AG, Cham, Switzerland.

FAOSTAT (Food and Agriculture Organization of the United Nations Statistics Division). 2018. Statistical Database of the Food and Agriculture Organization of the United Nations. Accessed July 20, 2018. http://www.fao.org/faostat/en/#data/QL.

Ferragina, A., C. Cipolat-Gotet, A. Cecchinato, and G. Bittante. 2013. The use of Fourier-transform infrared spectroscopy to predict cheese yield and nutrient recovery or whey loss traits from unprocessed bovine milk samples. J. Dairy Sci. 96:7980–7990. https://doi.org/10.3168/jds.2013-7036.

Ferragina, A., C. Cipolat-Gotet, A. Cecchinato, M. Pazzola, M. L. Dettori, G. M. Vacca, and G. Bittante. 2017. Prediction and repeatability of milk coagulation properties and curd-firming modeling parameters of ovine milk using Fourier-transform infrared spectroscopy and Bayesian models. J. Dairy Sci. 100:3526–3538. https://doi.org/10.3168/jds.2016-12226.

Inglingstad, R. A., H. Steinshamn, B. S. Dagnachew, B. Valenti, A. Criscione, E. O. Rukke, T. G. Devold, S. B. Skeie, and G. E. Vegarud. 2014. Grazing season and forage type influence goat milk composition and rennet coagulation properties. J. Dairy Sci. 97:3800–3814. https://doi.org/10.3168/jds.2013-7542.

ISO-IDF. 2004. International Organization for Standardization and International Dairy Federation. Milk: Quantitative determination of bacteriological quality - Guidance for establishing and verifying a conversion relationship between routine method results and anchor method results. International Standard ISO 21187 and IDF 196:2004. 2004. ISO, Geneva, Switzerland, and IDF, Brussels, Belgium.

ISO-IDF. 2006. International Organization for Standardization and International Dairy Federation. Milk: Enumeration of somatic cells - Part 2: Guidance on the operation of fluoro-opto-electronic counters. International Standard ISO 13366–2 and IDF IDF 148–2:2006. 2006. ISO, Geneva, Switzerland, and IDF, Brussels, Belgium.

ISO-IDF. 2013. International Organization for Standardization and International Dairy Federation. Milk and liquid milk products: Determination of fat, protein, casein, lactose and pH content. International Standard ISO 9622 and IDF 141:2013. 2013. ISO, Geneva, Switzerland, and IDF, Brussels, Belgium.

Maga, E. A., P. Daftari, D. Kültz, and M. C. T. Penedo. 2009. Prevalence of α_{S1}-casein genotypes in American dairy goats. J. Anim. Sci. 87:3464–3469. https://doi.org/10.2527/jas.2009-1854.

McParland, S., and D. P. Berry. 2016. The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. J. Dairy Sci. 99:4056–4070. https://doi.org/10.3168/jds.2015-10051.

Nilsson, K., H. Stålhammar, M. Stenholdt Hansen, H. Lindmark-Månsson, S. Duchemin, F. Fikse, D. J. de Koning, M. Paulsson, and M. Glantz. 2019. Characterisation of non-coagulating milk and effects of milk composition and physical properties on rennet-induced coagulation in Swedish Red Dairy Cattle. Int. Dairy J. 95:50–57. https://doi.org/10.1016/j.idairyj.2019.03.006.

Paschino, P., G. Stocco, M. L. Dettori, M. Pazzola, M. L. Marongiu, C. E. Pilo, C. Cipolat-Gotet, and G. M. Vacca. 2020. Characterization of milk composition, coagulation properties and cheesemaking ability of goats reared in extensive farms. J. Dairy Sci. 103:5830–5843. https://doi.org/10.3168/jds.2019-17805.

Pazzola, M., M. L. Dettori, and G. M. Vacca. 2018a. The Sarda goat, A resource for the extensive exploitation in the Mediterranean environment. In Sustainable Goat Production in Adverse Environments: Volume II. J. Simoes and C. Gutierrez, ed. Springer, Cham, Switzerland.

Pazzola, M., G. Stocco, P. Paschino, M. L. Dettori, C. Cipolat-Gotet, G. Bittante, and G. M. Vacca. 2018b. Modeling of coagulation, curd firming, and syneresis of goat milk from 6 breeds. J. Dairy Sci. 101:7027–7039. https://doi.org/10.3168/jds.2018-14397.

Pizarro, M. G., V. Landi, F. J. Navas, J. M. León, A. Martínez, J. Fernández, and J. V. Delgado. 2020. Nonparametric analysis of casein complex genes' epistasis and their effects on phenotypic expression of milk yield and composition in Murciano-Granadina goats. J. Dairy Sci. 103:8274–8291. https://doi.org/10.3168/jds.2019-17833.

Qin, L. X., H. C. Huang, and C. B. Begg. 2016. Cautionary note on using cross-validation for molecular classification. J. Clin. Oncol. 34:3931–3938. https://doi.org/10.1200/JCO.2016.68.1031.

R Core Team. 2013. R: A language and environment for statistical computing. http://www.R-project.org/.

Roberts, D. R., V. Bahn, S. Ciuti, M. S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J. J. Lahoz-Monfort, B. Schröder, W. Thuiller, D. I. Warton, B. A. Wintle, F. Hartig, and C. F. Dormann. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40:913–929. https://doi.org/10.1111/ecog.02881.

Roy, D., A. Ye, P. J. Moughan, and H. Singh. 2020. Gelation of milks of different species (dairy cattle, goat, sheep, red deer, and water buffalo) using glucono-δ-lactone and pepsin. J. Dairy Sci. 103:5844–5862. https://doi.org/10.3168/jds.2019-17571.

Soyeurt, H., P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, and N. Gengler. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. J. Dairy Sci. 89:3690–3695. https://doi.org/10.3168/jds.S0022-0302(06)72409-2.

Stella, A., E. L. Nicolazzi, C. P. Van Tassell, M. F. Rothschild, L. Colli, B. D. Rosen, T. S. Sonstegard, P. Crepaldi, G. Tosser-Klopp, and S. Joost. 2018. AdaptMap: exploring goat diversity and adaptation. Genet. Sel. Evol. 50:61. https://doi.org/10.1186/s12711-018-0427-5.

Stocco, G., C. Cipolat-Gotet, T. Bobbo, A. Cecchinato, and G. Bittante. 2017. Breed of cow and herd productivity affect milk composition and modeling of coagulation, curd firming, and syneresis. J. Dairy Sci. 100:129–145. https://doi.org/10.3168/jds.2016-11662.

Stocco, G., C. Cipolat-Gotet, A. Cecchinato, L. Calamari, and G. Bittante. 2015. Milk skimming, heating, acidification, lysozyme, and rennet affect the pattern, repeatability, and predictability of milk coagulation properties and of curd-firming model parameters: A case study of Grana Padano. J. Dairy Sci. 98:5052–5067. https://doi.org/10.3168/jds.2014-9146.

Stocco, G., C. Dadousis, G. M. Vacca, M. Pazzola, P. Paschino, M. L. Dettori, A. Ferragina, and C. Cipolat-Gotet. 2021. Breed of goat affects the prediction accuracy of milk coagulation properties using Fourier-transform infrared spectroscopy. J. Dairy Sci. 104. https://doi.org/10.3168/jds.2020-19491.

Stocco, G., M. Pazzola, M. L. Dettori, C. Cipolat-Gotet, A. Summer, and G. M. Vacca. 2019. Variation in caprine milk composition and coagulation as affected by udder health indicators. Int. Dairy J. 98:9–16. https://doi.org/10.1016/j.idairyj.2019.06.005.

Tiplady, K. M., R. G. Sherlock, M. D. Littlejohn, J. E. Pryce, S. R. Davis, D. J. Garrick, R. J. Spelman, and B. L. Harris. 2019. Strategies for noise reduction and standardization of milk mid-infrared spectra from dairy cattle. J. Dairy Sci. 102:6357–6372. https://doi.org/10.3168/jds.2018-16144.

Usai, M. G., S. Casu, G. Molle, M. Decandia, S. Ligios, and A. Carta. 2006. Using cluster analysis to characterize the goat farming system in Sardinia. Livest. Sci. 104:63–76. https://doi.org/10.1016/j.livsci.2006.03.013.

Vacca, G. M., M. Pazzola, M. L. Dettori, E. Pira, F. Malchiodi, C. Cipolat-Gotet, A. Cecchinato, and G. Bittante. 2015. Modeling of coagulation, curd firming and syneresis of milk from Sarda. J. Dairy Sci. 98:2245–2259. https://doi.org/10.3168/jds.2014-8902.

Vacca, G. M., G. Stocco, M. L. Dettori, G. Bittante, and M. Pazzola. 2020. Goat cheese yield and recovery of fat, protein, and total solids in curd are affected by milk coagulation properties. J. Dairy Sci. 103:1352–1365. https://doi.org/10.3168/jds.2019-16424.

Vacca, G. M., G. Stocco, M. L. Dettori, E. Pira, G. Bittante, and M. Pazzola. 2018. Milk yield, quality and coagulation properties of 6 breeds of goats: Environmental and individual variability. J. Dairy Sci. 101:7236–7247. https://doi.org/10.3168/jds.2017-14111.

Visentin, G., A. McDermott, S. McParland, D. P. Berry, O. A. Kenny, A. Brodkorb, 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. https://doi.org/10.3168/jds.2015-9323.

Wang, Q., and H. Bovenhuis. 2019. Validation strategy can result in an overoptimistic view of the ability of milk infrared spectra to predict methane emission of dairy cattle. J. Dairy Sci. 102:6288–6295. https://doi.org/10.3168/jds.2018-15684.

ORCIDS

Christos Dadousis https://orcid.org/0000-0002-7689-6443
Claudio Cipolat Gotet https://orcid.org/0000-0002-2318-4231
Giorgia Stocco https://orcid.org/0000-0002-6786-9806
Alessandro Ferragina https://orcid.org/0000-0001-9533-1221
Maria L. Dettori https://orcid.org/0000-0001-5413-0335
Michele Pazzola https://orcid.org/0000-0001-5685-0416
Adriano Henrique do Nascimento Rangel https://orcid.org/0000-0002-2835-4156
Giuseppe M. Vacca https://orcid.org/0000-0003-3110-5211