Regression trees to identify combinations of farming practices that achieve the best overall intrinsic quality of milk

L. Rey-Cadilhac, A. Ferlay, M. Gelé, S. Léger, Claire Laurent

Regression trees to identify combinations of farming practices that achieve the best overall intrinsic quality of milk

L. Rey-Cadilhac,[1] ● **A. Ferlay,[1]** ● **M. Gelé,[2]** ● **S. Léger,[3]** and **C. Laurent[1]*** ●
[1]Université Clermont Auvergne, INRAE, VetAgro Sup, UMR Herbivores, F-63122 Saint-Genès-Champanelle, France
[2]Institut de l'Elevage, F-75012 Paris, France
[3]Université Clermont Auvergne, Laboratoire de Mathématiques Blaise Pascal, UMR6620- CNRS, 63178 Aubière Cedex, France

ABSTRACT

Many studies over the last 30 years have shown the effects of farming practices on milk compounds. Combinations of practices may have antagonistic or synergistic effects on milk compounds, but these combination effects remain underinvestigated. Research needs to focus on overall intrinsic milk quality (including sensory, technological, health, and nutritional dimensions) and identify the combinations that can optimize it. The aim of this study was to identify which combinations of farming practices achieved the best scores for sensory, technological, health, and nutritional dimensions and for overall intrinsic milk quality. Ninety-nine private farms were visited once each to sample their bulk tank milk and survey their farming practices. The surveyed practices concerned herd characteristics, feeding management, housing conditions, and milking and milk storage conditions on the day of test. Analyses of bulk tank milk were designed to evaluate the overall intrinsic quality of the milk for 2 target products: raw milk cheese and semi-skimmed UHT milk. Regression trees were then used to identify the combinations of farming practices that achieved the best scores on each dimension and on overall intrinsic quality of the milk. Breed and diet (type of forage) were the most influential factors for sensory and health dimensions and for technological and nutritional dimension scores, respectively, in the cheese assessment. Overall cheese quality was highly positively correlated with these 4 dimension scores. Therefore, breed and diet emerged as the most influential practices in the regression tree for overall cheese quality. However, the combinations of practices that resulted in the best quality scores differed accord-
ing to dimension studied and product targeted. This suggests that advice on farming practices to improve intrinsic milk quality needs to be adapted according to the end-purpose of the collected milk. This innovative approach combining on-farm data and regression trees provides farm managers with a valuable and practical tool to prioritize practices in terms of their role in shaping milk quality, and to identify the combinations of practices that promote good milk quality and practice thresholds or modalities needed to achieve it.
Key words: dairy cow, bulk tank milk, quality score, dairy products, regression analysis

INTRODUCTION

Over the past 30 yr, there has been ample literature on the effects of farming practices on milk compounds (Baeza-Campone et al., 2020), but the effects of combinations of practices remain under-researched. The studies conducted to date have generally focused on (1) only some compounds, generally considered separately, or (2) the effect of a single practice taken in isolation (often feeding) on one or more compounds. However, certain practices can have a synergistic or antagonistic effect on milk quality. For example, milking with an automatic milking system (**AMS**) instead of in a milk parlor leads to higher milk lipolysis (greater amount of free fatty acids, **FA**) and sometimes higher milk SCC (Hogenboom et al., 2019), both of which are globally detrimental to milk quality. In this case, the AMS has a synergistic effect on milk compounds in terms of quality. Conversely, milk from cows fed fresh grass as their main forage source instead of corn silage is lower in SFA, higher in UFA, and yellower (Nozière et al., 2006; Couvreur et al., 2007), all of which are globally beneficial to milk quality. However, milk from cows fed fresh grass also has a lower fat content (Legarto et al., 2014), which is detrimental to milk quality. In this case, the dominant forage in the cow diet has an antagonist effect on milk compounds in terms of quality.

Research is therefore needed to study the effects of farming practices and their combinations on the overall
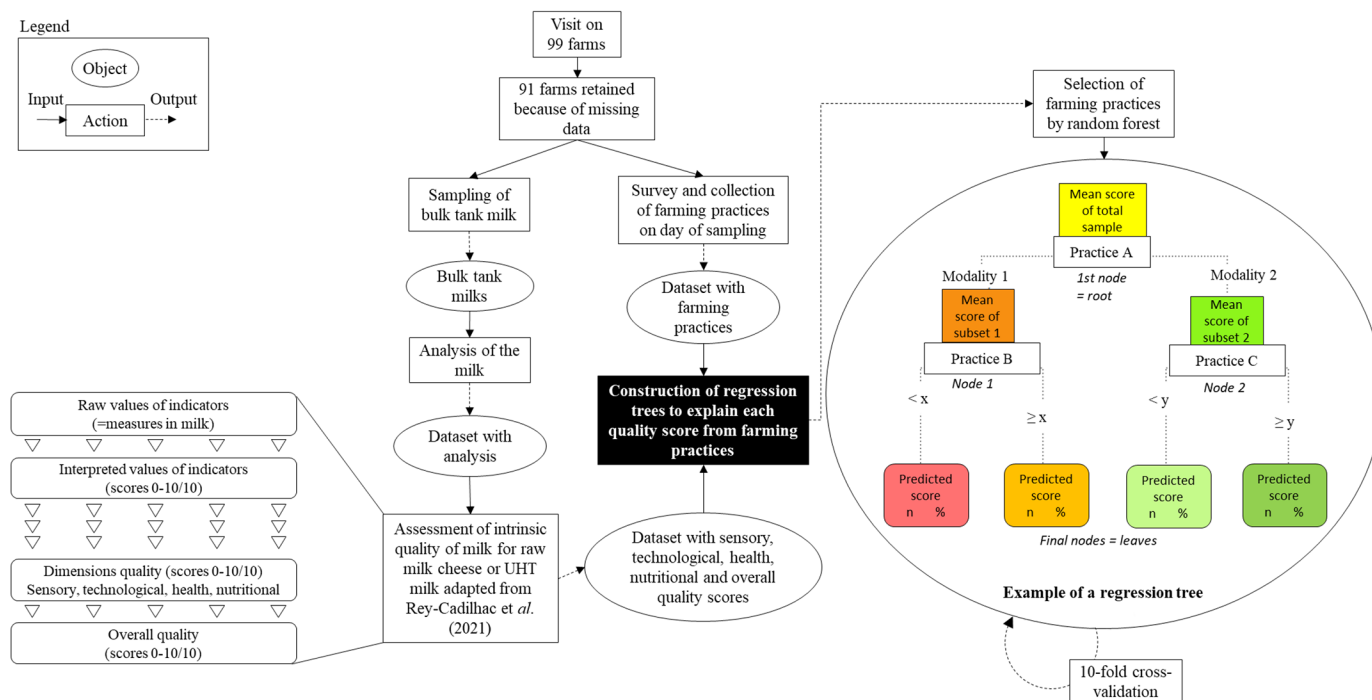
Figure 1. Steps for the construction of regression trees explaining sensory, technological, health, nutritional dimensions, and overall quality scores of the cheese and UHT milk assessments from farming practices. Raw milk cheese = pressed uncooked nonstandardized raw milk cheese, and UHT milk = semi-skimmed standardized UHT milk. In the example regression tree, leaves show predicted score, number of farms within the leaves (n), and % of farms of the total sample. Red and green colors correspond to the worst and best quality scores, with intermediate colors indicating intermediate scores.

intrinsic quality of milk. Rey-Cadilhac et al. (2021) recently reported a multicriteria assessment model of the overall intrinsic quality of milk, taking into account its sensory, technological, health, and nutritional dimensions. They developed 2 separate multicriteria assessments for 2 products based on end-purpose of the milk: pressed uncooked nonstandardized raw milk cheese (cheese assessment) and semi-skimmed standardized UHT milk (UHT milk assessment). The application of the multicriteria assessments allowed us to identify the quality indicators to be optimized to achieve better quality scores of raw milk but it is not possible to explain these scores by farming practices.

To address this gap, the present study was designed to meet 2 objectives: (1) to identify combinations of practices that result in the best dimensions and overall quality scores of bulk tank milk; and (2) to test whether the combinations of practices identified differ according to target product (cheese or UHT milk) and among the target dimensions. To do so, we used the 2 multicriteria assessments of overall intrinsic quality combined with regression trees to identify the combinations of farming practices that achieved the best scores for overall intrinsic quality of raw bulk tank milk.

MATERIALS AND METHODS

Selection of Farms

Only routine milk sampling and farmer surveys were performed during this study, so no ethical approval was needed. The objective was to select 100 private farms that met the following criteria: dominant breed in the herd must be Montbéliarde (**MB**) or Holstein Friesian (**HF**) with <30% of other breeds; milking in a parlor or with an AMS; and lactating cows fed with a dominant forage (corn silage, wet conserved grass, hay, or pasture). The rationale for choosing these criteria was, first, to address the main factors affecting milk quality (breed, milking system, and diet) and, second, to control these factors by having a limited number of modalities for each. The aim was also to have a set of farms that was representative of the diversity of existing systems in France by selecting farms from different milk-producing regions.

In total, 99 private farms were visited once between June and November 2020 in 3 locations in France. During the visit, we collected bulk tank milk samples and surveyed the farming practices applied on the day of sampling (Figure 1).

Survey of Farming Practices

In our study, the farming practices encompassed herd characteristics, feeding management, housing conditions, and milking and milk storage conditions on the day of milk collection. These practices applied on the farms were collected by survey using a questionnaire containing closed-ended questions. This latter was developed through consultation with field and scientific experts in the topics it addresses. It was tested and validated on the experimental farm of INRAE facilities (Herbipôle, Theix, France). It was then applied on the selected farms by 3 trained research personal who surveyed the farm owner(s) or manager(s). The survey was completed using data from milk records (monthly dairy herd improvement testing) or the AMS data collection system for lactation stage, parity, and milk output. The survey also captured observational data recorded by the surveyor during milking; namely, milking equipment conception, milking practices such as teat cleaning, cow dirtiness, and housing conditions. Collected data of farming practices were checked for errors and outliers and for overall consistency between farmer's answers, milk records or AMS, and observational data. We then built aggregated variables, such as teat cleaning scores, complexity of the milking machine, milking machine washing scores, or dominant forage in the diet (Tables 1 and 2 and Supplemental Tables S1, S2, S3, and S4; https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022), to derive variables that were applicable to all farms and could thus serve as comparators. The practices finally tested to implement the trees are described in Table 1 (continuous variables) and Table 2 (categorical variables or continuous variables separated into groups and analyzed as categorical variables due to the frequency of observation and distribution).

Bulk Tank Milk Sampling and Analysis

Bulk tank milk samples were collected once on each farm in a tank containing an even number of milkings (2, 4, or 6) or at least a 12-h milking since the last tank emptying at the farms equipped with an AMS. A 2-L sample of milk was collected from the tank after a 5-min agitation using a sterile sampling rod through the opening, and then directly aliquoted and either kept between 0 and 4°C or directly frozen at −20°C in a portable freezer. The milk compounds measured and method of analysis for each compound are described in Rey-Cadilhac et al. (2021). Most of the analyses (fat, protein, free FA, urea and lactose contents, total bacteria, coliforms, β-glucuronidase-positive *Escherichia coli*, coagulase-positive staphylococci, and *Pseudomonas*

Table 1. Regression tree models input continuous variables characterizing farming practices

Continuous variable	Description of variable[1]	Mean ± SD (range), n = 91	Selected from RF[2]
Herd lactation st	Average lactation stage (d)	183 ± 34 (88–262)	Cheese, UHTmilk, T, H, N
Herd parity	Average number of lactations	2.6 ± 0.4 (1.8–4.4)	S, H,
Herd teat dirt	Teat dirtiness score by the surveyor with scores given for each cow, from 0 = very clean to 2 = very dirty teats. Scores were given on at least 25% of the lactating cows or at least 10 cows and averaged to the herd level.	0.6 ± 0.3 (0.2–1.6)	Cheese, UHTmilk, S, T, H,
Diet concentrates	Percentage of concentrates in the total dry matter intake (%)	26 ± 10 (0–54)	T, N
MilkingEq wash score	Milking machine washing score [0–3] with 0 = nonintensive and 3 = intensive milking machine washing routines (Supplemental Table S1[3])	1.8 ± 0.8 (0.0–3.0)	
Milking teat cleaning score	Teat cleaning score [0–10], adapted from Monsallier et al. (2012) with 0 = nonintensive and 10 = very intensive cleaning routines (Supplemental Table S2[3])	6.2 ± 1.9 (0.0–9.5)	S
Tankstorage time	Milk storage time (h) since first milk arrived in the tank	23 ± 11 (11–63)	UHTmilk, S, T, N

[1]Average value of the herd of lactating cows or information about the lactating cows.
[2]Variables were selected from the variable importance graphs of random forest (RF; Supplemental Figures S1 and S2 (https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022) to be implemented in the overall UHT milk quality (UHTmilk), overall cheese quality (Cheese), and cheese technological (T), sensory (S), health (H), and nutritional (N) dimension regression trees.
[3]Supplemental Tables S1 and S2: https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022.

Table 2. Regression tree models input categorical variables characterizing farming practices

Categorical variable	Description of variable[1]	Category	% total (n/N), N = 91	Selected from RF[2]
Herd breed	Dominant breed in the herd (>80% of cows)	MB = Montbéliarde	41 (37/91)	Cheese, UHTmilk, T, N
		HF = Holstein Friesian	48 (44/91)	
		Other = mixed herds with pure HF and MB cows or MB × HF crossbreds (any other animals were always <30% of the herd)	11 (10/91)	
Diet forage	Dominant forage in the dry matter forages intake (description in Supplemental Table S4[3])	WetConservedGrass = grass silage or haylage	26 (24/91)	Cheese, UHTmilk, S, H, N
		CornSilage = corn or sorghum silage	30 (27/91)	
		Hay	19 (17/91)	
		Pasture = pasture or fresh grass	25 (23/91)	
Diet corn silage pres	Presence of corn or sorghum silage in the diet	Yes	54 (49/91)	Cheese, H
		No	46 (42/91)	
Diet grass silage pres	Presence of grass silage or haylage in the diet	Yes	56 (51/91)	
		No	44 (40/91)	
Diet hay pres	Presence of hay in the diet	Yes	59 (54/91)	
		No	41 (37/91)	
Diet pasture pres	Presence of pasture in the diet	Yes	45 (41/91)	
		No	55 (50/91)	
Housing pres	Time spent in the barn outside of milking time	Yes	77 (70/91)	
		No	23 (21/91)	
Housing type	Type of housing	Bedded_area = strawbedded area	16.5 (15/91)	
		Cubicles = cubicles (any kind of litter)	60.5 (55/91)	
		100%Pasture = cows all the time on pasture (not concerned)	23 (21/91)	
Housing stocking density	Stocking density in the barn (number of cows per place[2])	Inf1 = 1 or <1 cow/place	59 (54/91)	
		Sup1 = >1 cow/place	18 (16/91)	
		100%Pasture = cows on pasture all the time (not concerned)	23 (21/91)	
MilkingEq system	Milking machine	CMS = conventional milking system	75 (68/91)	
		AMS = automatic milking system	25 (23/91)	
MilkingEq complexity	Complexity of the milking pipeline	Modalities defined in Supplemental Table S3[3]: Short_simple = short and simple	23 (21/91)	
		Intermed = intermediate	30 (27/91)	
		Large_complex = large and complex	22 (20/91)	
		AMS (not concerned)	25 (23/91)	
MilkingEq milk filtration	Use of a milk filter before arrival to the tank (to remove particulates from milk)	Yes	46 (42/91)	
		No	54 (49/91)	
MilkingEq milk fall	Milk fall of >2 m at arrival in the tank	Yes	38 (35/91)	
		No	62 (56/91)	
Milking cup drop	Cup drop during milking	Yes	27.5 (25/91)	
		No	47 (43/91)	
		AMS (not concerned)	25.5 (23/91)	
Milking cup air intake	Cup air intake during the milking	Rarely	44 (40/91)	
		Often	31 (28/91)	
		AMS	25 (23/91)	
Tank storage temp	Milk storage temperature in the empty tank	~4°C	91 (83/91)	
		~12°C	9 (8/91)	

[1]Average value of the herd of lactating cows or information about the lactating cows.

[2]Variables selected from the variable importance graphs of random forest (RF; Supplemental Figures S1 and S2; https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022) to be implemented in the overall UHT milk quality (UHTmilk), overall cheese quality (Cheese), and cheese technological (T), sensory (S), health (H), and nutritional (N) dimension regression trees.

[3]Supplemental Tables S3 and S4: https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022.

counts, SCC, lactofermentation test, and pH variation during the test, pH, Dornic acidity, and Ramsdell and alcohol tests) were performed on fresh milk within 24 h, whereas other analyses were performed on fresh milk within 48 h (gel firmness and rennet coagulation time) or on frozen milk samples (mesophilic lactic acid bacteria counts, calcium, phosphorus, caseins, soluble proteins, FA and vitamin B_2 content, and b color index).

Calculation of Milk Quality Scores

Rey-Cadilhac et al. (2021) developed 2 assessments of the intrinsic quality of raw milk designed to score the overall quality of a farm's bulk tank milk on a scale of 0 to 10 in relation to its end-purpose; that is, either pressed, uncooked, nonstandardized raw milk cheese or standardized semi-skimmed UHT milk. These assessments were developed through a participative approach by consulting experts in different fields. They built the hierarchical structure of the assessment from the indicators measured in milk to the final score. They also defined the thresholds to interpret the indicators and the ponderation (weight given to each indicator in percentage) and compensation rules to aggregate the information from one level to another to the final overall scores via several aggregation steps (Rey-Cadilhac et al., 2021).

The multicriteria assessments use values measured via milk analyses (i.e., raw values of indicators) to calculate the scores of the 4 milk quality dimensions (sensory, technological, health, and nutritional scores) as well as an overall quality score for the 2 target end-products (Figure 1). The indicators used in each dimension of the cheese assessment are described in Supplemental Table S5 (https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022). These 2 assessments were adapted and applied here to the bulk tank milks of the 99 farms based on their results of analysis. Adaptations from the initial assessments were made with input from the same expert group that created them in Rey-Cadilhac et al. (2021), and are described in Supplemental File S1 (https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022).

Statistical Analysis

All data statistical analysis was performed using R software (version 4.0.5; https://www.r-project.org/).

Imputation of Missing Analytical Data. When analytical data were missing (raw values of some indicators for some bulk tank milks: alcohol test, Ramsdell test, Dornic acidity, fat content, and pH variation), they were approximated using the method of multivariate imputation by chained equations by random forest imputation with the "MICE" package (van Buuren and Groothuis-Oudshoorn, 2011) in R (https://www.r-project.org/). This multiple imputation method led to 5 imputed data sets that were then averaged to obtain a final analysis data set with a better stability of estimated values. Quality scores were then calculated from this averaged analysis data set.

Correlations Among Milk Quality Scores. Correlation coefficients among milk quality scores were calculated as follows: (1) for each dimension of each product, among the dimension scores and the scores of the indicators characterizing that dimension (interpreted values of measures in milk, Figure 1); (2) for each product among the dimension scores and overall quality scores; and (3) among the overall quality scores of the 2 targeted end-products. When data did not fit the normality hypothesis, a Spearman coefficient was calculated for correlations (1) and (2), and a Pearson correlation coefficient was calculated for (3). For (1), we retained the indicators that were best correlated (r ≥ 0.4) to the dimension they characterized.

Control of Confounding Effects Among Farming Practice Variables. Having confounding variables is not a problem when applying the regression tree method. However, this fact should be kept in mind when interpreting the results. The intensity of the relations among quantitative variables was assessed by the coefficient of correlation. The intensity of the relations among quantitative and qualitative variables was assessed by the correlation ratio (**corRatio**) and the intensity of the relations among qualitative variables was evaluated by Cramer's V coefficient. The threshold to consider confounding effects among variables was 0.4.

Regression Trees. Regression trees were constructed to identify the combinations of farming practices determining the scores obtained with the assessments for cheese and UHT milk. The trees were constructed by the Classification and Regression Tree (**CART**) method in R using the "rpart" package (Figure 1). The CART method was selected for its ability to deal with either categorical or continuous data without conditions on the data such as normality or non-multicollinearity. Moreover, it allowed us to identify the combinations of farming practices that best explained the quality scores and to rank the farming practices in terms of their explanatory power.

To balance the potential instability of the trees, practices were preselected using the random forest "increase in node purity" measure of variable importance (Breiman et al., 1984) in R using the "randomForest" package (with ntree = 500 trees and mtry = 15 tested practices). Practices were thus ranked according to their importance values for overall quality and for each quality dimension (Supplemental Figure S1 for cheese quality scores and Supplemental Figure S2 for UHT milk quality scores; https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022), and top 5 practices (i.e., with the 5 highest importance values) were retained for implementation in the regression trees (see the last column of Tables 1 and 2).

Within the CART method, practices are recursively selected at a determined numerical or categorical value (e.g., the amount of concentrate in the herd diet >30%

of feed intake) that splits individuals into 2 mutually exclusive subsets that are as homogeneous as possible in terms of the response variable (i.e., quality scores) by minimizing the sums of squares inside the subsets (Breiman et al., 1984). The regression trees therefore split the whole source set into 2 branches, giving 2 subsets called nodes. The process is repeated until a stop criterion is reached: in our case, we chose to have a minimum of 10 individuals in the final nodes, called leaves, to avoid subsets with small numbers of individuals that would not be reproducible. We then pruned the trees based on a 10-fold cross-validation to obtain optimal-sized trees without overfitting to have valid trees on new data. The coefficient of determination, R^2, was calculated for each tree to compare the different models. The averages of the final leaves of each tree were compared using ANOVA followed by a Tukey test. Significance was declared at $P \leq 0.05$, and trends were considered at $0.05 < P \leq 0.10$.

RESULTS

This article focuses on the results obtained for milk intended for cheese processing. For milk intended for processing into UHT milk, only the results on overall quality are presented and discussed here (other results are presented in Supplemental File S2 for correlations among the quality dimensions, and Supplemental Figure S3 for quality-dimension regression trees; https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022).

Surveyed Farming Practices

Of the 99 surveyed and milk-sampled farms, 5 were excluded as they had missing data on farming practices, and 3 were excluded as they had a similar amount of pasture and corn silage in the diets, leaving a total of 91 farms for analysis (Figure 1). The number of lactating cows ranged from 22 to 200, with an average of 67 ± 33 cows (mean $\pm$ SD), and milk output ranged from 8.3 to 34.5 L/d per cow, with an average of 23.4 ± 5.0 L/d per cow. Averages and frequencies of the farming-practice modalities are presented in Tables 1 and 2. Lactating herds were, on average, at 183 ± 34 DIM and 2.6 ± 0.4 parities. Diets contained $26 \pm 10\%$ of concentrate in total DMI. Diets were based on wet conserved grass, corn silage, hay, and pasture on 26, 30, 19, and 25% of the farms, respectively. Supplemental Table S4 reports the diet groups. The dominant breed was HF on 48% of farms and MB on 41%. The remaining 11% had mixed herds with pure HF and MB cows or MB × HF crossbreds (but any other animals were always <30% of the herd). Furthermore, 23% of farms were on full-day

Table 3. Overall quality and dimensions scores[1] (0–10) obtained on the 91 bulk tank milks, calculated from an adapted version of the assessments published by Rey-Cadilhac et al. (2021; see Supplemental File S1, https://data.mendeley.com/datasets/3xdpk76v8b/1, Rey-Cadilhac et al., 2022)

Dimension	Cheese assessment	UHT milk assessment
Sensory	4.0 ± 2.0 (0.1–8.7)	7.7 ± 3.8 (0.0–10.0)
Technological	4.1 ± 1.8 (0.9–8.2)	6.0 ± 2.6 (0.0–9.7)
Health	4.9 ± 2.3 (0.6–8.6)	2.8 ± 1.8 (0.0–7.1)
Nutritional	3.0 ± 2.0 (0.5–8.2)	3.4 ± 1.8 (0.1–7.3)
Overall quality	4.0 ± 1.4 (1.6–6.8)	4.2 ± 1.4 (0.7–6.6)

[1]Mean ± SD (range).

grazing, 16.5% had a straw-bedded area, and 60.5% had cubicles (with a different sort of litter). Finally, 75% had a conventional milking parlor and 25% has an AMS. Intensiveness of the teat-cleaning method evaluated on a 0 to 10 scale (where 0 = nonintensive to 10 = very intensive) averaged 6.2 ± 1.9.

There were no confounding variables among quantitative farming variables (r > 0.4, Supplemental Table S6; https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022) or among quantitative and qualitative farming variables (corRatio > 0.4, Supplemental Table S7; https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022). A few effects confounded between qualitative variables were demonstrated (Cramer's V coefficient >0.4, Supplemental Table S8; https://data.mendeley.com/datasets/3xdpk76v8b/1; Rey-Cadilhac et al., 2022) and were related to the construction of the variables (e.g., Housing_type and Housing_stocking density have "100%Pasture" as a common modality; Table 2) or to the systems studied (e.g., the absence of AMS in farms with pasture as dominant forage).

Factors Explaining Cheese Scores on Sensory, Technological, Health, and Nutritional Dimensions

Correlation Between Dimension Scores and Overall Quality and Indicator Scores. Scores for the sensory, technological, health, and nutritional dimensions in the cheese assessment measured on a 0 to 10 scale averaged 4.0 ± 2.0, 4.1 ± 1.8, 4.9 ± 2.3, and 3.0 ± 2.0, respectively (Table 3). Overall cheese quality score was positively correlated with cheese technological and health scores (r = 0.65 and 0.56, respectively; $P < 0.001$) and highly positively correlated with cheese sensory and nutritional scores (r = 0.79 and 0.78, respectively; $P < 0.001$).

Concerning the relation between dimension scores and scores of the indicators characterizing them in the multicriteria assessment, cheese sensory score was mostly correlated with *Pseudomonas*, coliform, and

total bacterial counts and yellow color index scores (r = 0.57, 0.48, 0.43, and 0.44, respectively). Cheese technological score was mostly correlated with casein content, gel firmness, fat:protein ratio, and calcium content scores (r = 0.78, 0.77, 0.53, and 0.44, respectively). Cheese health score was mostly correlated with FA profile scores [r = −0.47, 0.60, 0.68, 0.64, and 0.68 for sum of C6:0 + C8:0 + C10:0, total SFA, C16:0 content, α-linolenic acid (**ALA**) content, and n-6:n-3 ratio scores, respectively] and the *E. coli* count score (r = 0.57). Cheese nutritional score was mostly correlated with caseins, ALA, and calcium content scores (r = 0.70, 0.72, and 0.46, respectively).

Combinations of Practices Explaining Dimension Scores. In the regression tree of the cheese sensory dimension (Figure 2A), the best sensory score was obtained for systems using grass-based diets with a mean parity ≥2.4 lactations and with a teat-cleaning routine score <5.8 (score of 5.3/10). The lowest score was obtained for systems with cows fed corn silage–based diets (score of 2.4/10). Other combinations of practices, including diet forage, herd parity, teat-cleaning routine score, and tank storage time variables, led to intermediate scores (from 3.4 to 5.0) that were not significantly different from these 2 systems.

In the regression tree of the cheese technological dimension (Figure 2B), the best technological score was obtained for systems with MB herds distributing concentrates at 31% or more of DMI (score of 6.1/10), followed by systems with MB herds fed <31% concentrates (4.8/10), followed by systems with HF and mixed/crossbred herds, regardless of how long the milk was stored (2.8 to 3.7/10).

In the regression tree of the cheese health dimension (Figure 2C), the best health score was obtained for systems without corn silage in the diet combined with low teat dirtiness (<0.69) (score of 7.3/10), followed by systems with no corn silage but higher teat dirtiness (≥0.69) (5.8/10), followed by systems with corn silage in the diet regardless of the stage of lactation of the herd (2.6 to 3.8).

In the regression tree of the cheese nutritional dimension (Figure 2D), the best score was obtained with MB herds (score of 4.8/10), whereas HF and mixed/crossbred herds, regardless of the stage of lactation of these herds, obtained lower scores (1.7 to 2.6/10).

Combinations of Practices Explaining Overall Quality Scores. Overall quality scores assessed for cheese and UHT milk were comparable, at 4.0 ± 1.4 and 4.2 ± 1.4 out of 10, respectively (Table 3). There was a strong positive correlation between overall quality of cheese and UHT milk (r = 0.70, *P* < 0.001).

In the regression tree of cheese overall quality (Figure 3A), the best overall cheese quality score was obtained for systems with MB and mixed/crossbred herds fed without corn silage, regardless of how dirty the teats were (scores from 4.8 to 5.6/10), whereas cheese overall quality score was lowest in systems with HF cows fed corn silage, regardless of how dirty the teats were (2.5 to 3.1/10). Systems with MB and mixed/crossbred herds fed corn silage had comparable scores with systems with HF herds fed without corn silage, and their scores were intermediate between the first 2 systems (4.2 to 4.5/10).

In the regression tree of UHT milk overall quality (Figure 3B), the best UHT milk quality score was obtained for systems with grass-based diets combined with a mean lactation stage of <168 DIM (score of 5.2/10), whereas the lowest UHT milk quality score was obtained for systems with corn silage–based diets (3.1/10). Systems using grass-based diets with a mean lactation stage ≥168 DIM had intermediary and not significantly different scores (3.8 to 4.8/10).

DISCUSSION

The results of this study demonstrated that it is possible to calculate intrinsic quality scores of bulk tank milks and to use regression trees to identify combinations of farming practices that result in the best quality scores. These combinations of practices were not the same in the sensory, technological, health, and nutritional dimension trees. A scan of the literature found no data on the association between overall milk quality and the practices that influence it. We therefore focus the first part of the discussion on relationships among scores for the indicators (that are at the base of the assessment, Figure 1) that best correlated with dimension scores and practices in the regression trees of these dimension scores. The second part of the discussion deals with the utility and limits of these tools for translation into the field.

Farming Practices Influenced Milk Compounds and Dimension Scores

The regression trees ranked farming practices in terms of their importance in explaining the quality scores. These trees form a pyramidal structure with practices closer to the top (the root) being the most influential and the practices underneath (on nodes) being less influential. Breed and dominant forage in the diets were the practices that had the greatest influence on quality scores. Moreover, Supplemental Figure S1 shows that when diet forage or breed was the most important practice, they were far ahead of other practices in most cases, meaning there was no competition from other practices for the place of number 1 splitter.
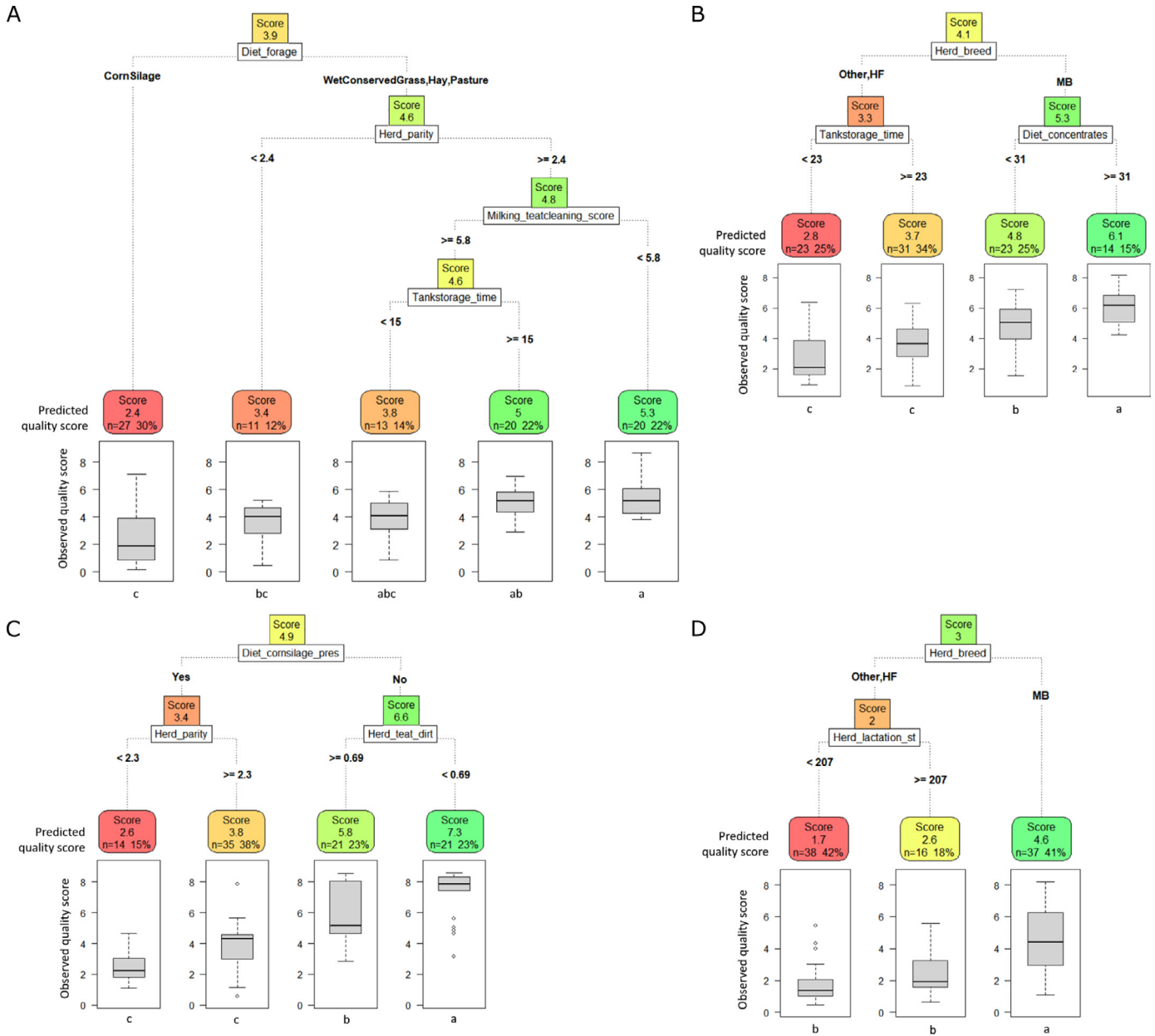
Figure 2. Regression trees explaining (A) cheese sensory ($R^2 = 0.38$, $RMSE_{cv} = 1.65$), (B) technological ($R^2 = 0.38$, $RMSE_{cv} = 1.50$), (C) health ($R^2 = 0.53$, $RMSE_{cv} = 1.77$), and (D) nutritional ($R^2 = 0.42$, $RMSE_{cv} = 1.63$) dimensions from farming practices. Variables are described in Table 2. Values in final leaves represent the predicted score value, number of individuals in the leaf (n), and percentage of individuals of the total sample. Means of the final leaves with different letters (a–c) are significantly different ($P < 0.05$). HF = Holstein Friesian; MB = Montbéliarde; $RMSE_{cv}$ = root mean square error obtained from the 10-fold cross-validation. The box represents the first (Q1) to third (Q3) quartile, the line represents the median, the whiskers represent the minimum and maximum (excluding outliers), and the points represent the outliers [greater than Q3 + 1.5 × (Q3 − Q1) or less than Q1 − 1.5 × (Q3 − Q1)].

Breed was the main driver of both technological and nutritional scores. This can be explained by some common indicators (i.e., casein and calcium contents) being the most strongly correlated with both these dimensions. Indeed, cow breed has an important effect on milk calcium (Baeza-Campone et al., 2020): milk from MB cows has a higher calcium content than milk from HF cows (Macheboeuf et al., 1993; Gaignon et al., 2018). Milk from MB cows is also richer in caseins than HF cow milk (Macheboeuf et al., 1993) and has a casein profile more conducive to gel firmness (with a higher content of κ-caseins and the B-variant of κ- and β-caseins; Macheboeuf et al., 1993; Martin and Coulon, 1995). The gel firmness indicator was also highly cor-
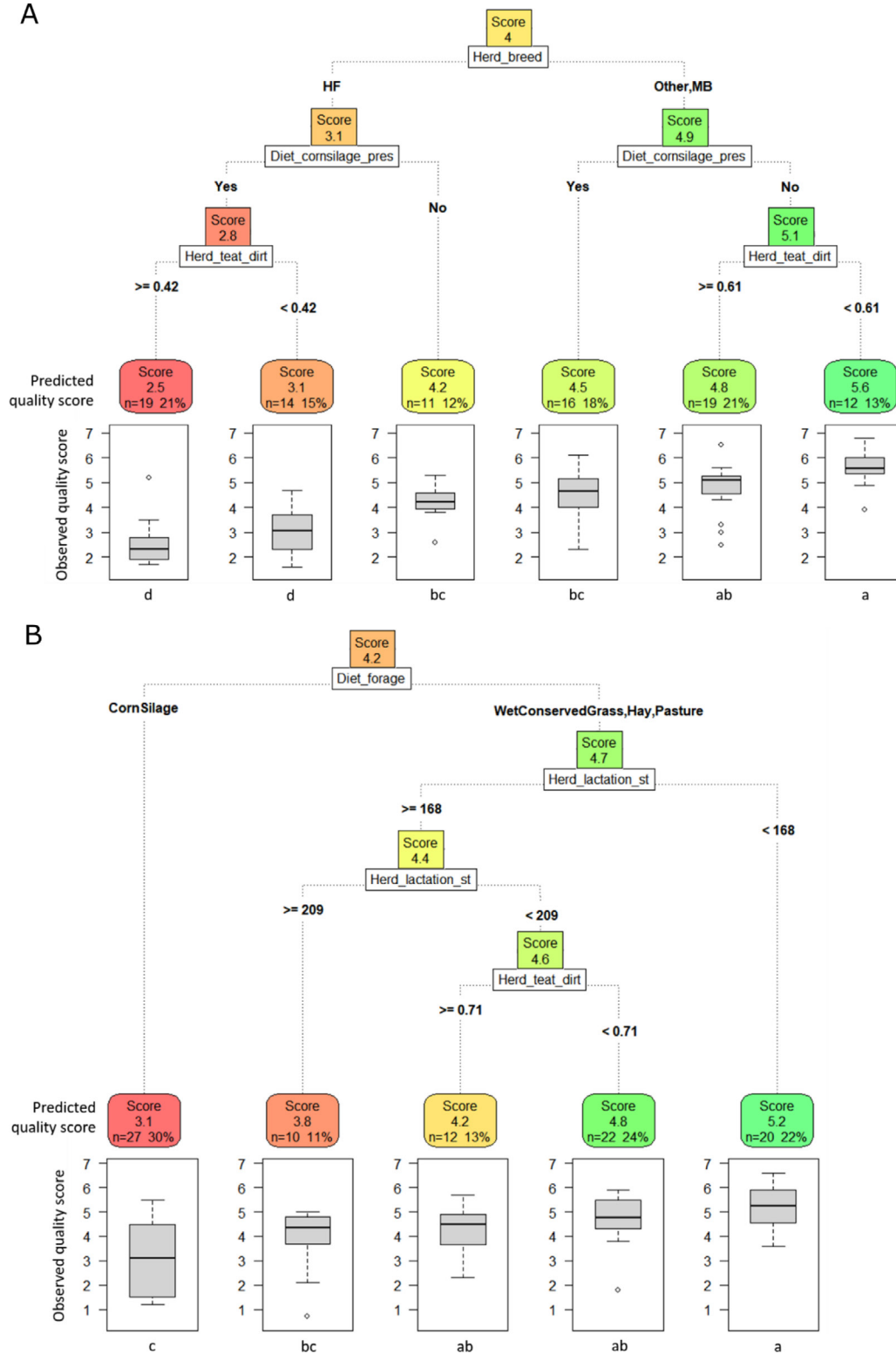
Figure 3. Regression trees explaining (A) overall cheese ($R^2 = 0.58$, $RMSE_{cv} = 0.97$) and (B) UHT milk ($R^2 = 0.34$, $RMSE_{cv} = 1.21$) qualities from farming practices. Variables are described in Table 2. Values in final leaves represent the predicted score value, number of individuals in the leaf (n), and percentage of individuals of the total sample. Means of the final leaves with different letters (a–d) are significantly different ($P < 0.05$). HF = Holstein Friesian; MB = Montbéliarde; $RMSE_{cv}$ = root mean square error obtained from the 10-fold cross-validation. The box represents the first (Q1) to third (Q3) quartile, the line represents the median, the whiskers represent the minimum and maximum (excluding outliers), and the points represent the outliers [greater than $Q3 + 1.5 \times (Q3 - Q1)$ or less than $Q1 - 1.5 \times (Q3 - Q1)$].

related with the technological dimension of the assessment.

For the technological score, when the herd had MB cows, a higher amount of concentrate ($\geq$31% of total DMI) led to higher technological scores, which is likely related to the casein content indicator. Increasing the amount of concentrates often increases the energy density of the diet, which leads to a higher milk protein content, a higher casein content (Coulon and Rémond, 1991), and better gel firmness.

Type of forage was the most influential factor for sensory and health dimensions, the first splitter being either main forage in the diet or presence of corn silage in the diet, with higher sensory and health scores in case of less corn silage or the absence of it, respectively. This can be explained by the best correlated indicators, such as the yellow color index of milk in the sensory dimension tree, and indicators related to milk FA profile (ALA, C16:0, SFA content, and n-6:n-3 ratio) in the health dimension tree. Indeed, FA profile is strongly influenced by type of forage, with larger differences among pasture- and corn silage–based diets, because pasture-feeding results in milk that is richer in UFA such as ALA, poorer in SFA (notably C16:0), and has a lower n-6:n-3 ratio (Dewhurst et al., 2006; Chilliard et al., 2007; Ferlay et al., 2013, 2017). Moreover, the yellow color of milk comes mainly from its carotenoid contents (notably β-carotenes), which is highly dependent on the carotenoid content of ingested forages (Nozière et al., 2006). Thus, milks from grazing cows are the richest in β-carotenes and therefore yellower, followed by milks from cows fed grass forage, and finally milk from cows fed corn silage, which contains much less carotenoids (Nozière et al., 2006).

The other practices in the sensory regression tree (herd parity, teat cleaning score, and tank storage time) appeared to be linked to the microbial profile of the bulk tank milks, as the best correlated indicators were *Pseudomonas*, coliforms, and total bacteria counts. The effects of farming practices on milk microbial quality are complex and unclear, with sometimes contradictory results. Effect of parity on microbial counts in milk can be explained by an effect of parity on teat skin microbiota. Total microbiota on teat skin is assumed to increase with age (Monsallier et al., 2012), and teat skin is known to be a major source of microorganisms in raw milk (Vacheyrou et al., 2011; Frétin et al., 2018). Bacic et al. (1968) found that cows in third and higher parities had higher bacterial counts in milk than cows from first and second parities. This could be due to progressive establishment of the teat flora over time or a loss of integrity of teat skin with age (Monsallier et al., 2012), which would support the idea that, in this study, herds with older cows (mean parity $\geq$2.4) led to better sensory scores due to higher total bacterial counts in their milk.

Over the past few decades, there has been an evolution in milking materials and global farming practices toward more hygienic production conditions, together with pressure to enhance teat cleaning that has prompted farmers to take measures against pathogenic and spoilage bacteria. These practices have led to milks that are increasingly poor in terms of bacterial and, more generally, microbial levels and diversity needed for traditional cheese-making (Verdier-Metz et al., 2009; Montel et al., 2014; Baeza-Campone et al., 2020). This trend was corroborated here, with a median of 11,000 cfu/mL for total flora. In addition, our measures showed that 60% and 55% of the milks had *Pseudomonas* and coliform counts $\leq$100 and $\leq$10 cfu/mL, respectively. In the multicriteria assessment of the milk for cheese, these levels correspond to high quality scores for spoilage flora (quality score of 10/10 if <100 and <10 cfu/mL for *Pseudomonas* and coliform counts, respectively; Rey-Cadilhac et al., 2021) and low scores for total flora (a score of 10/10 corresponds to a total flora count of 45,000 cfu/mL and a score of 0 corresponds to total flora counts $\leq$5,000 or $\geq$100,000 cfu/mL; Rey-Cadilhac et al., 2021). On average in our study, the quality scores for *Pseudomonas*, coliforms, and total flora counts were therefore 7.7 $\pm$ 3.8, 8.2 $\pm$ 3.3 and 2.0 $\pm$ 2.6 out of 10, respectively. Less intensive teat-cleaning routines result in higher total bacterial levels and higher *Pseudomonas* and coliform levels in milk [Zucali et al., 2011; Monsallier et al., 2012 (with the same teat-cleaning score grid)]. As most milks in our study had low *Pseudomonas* and coliforms counts, these practices led globally to numerically better sensory scores.

When the teat-cleaning routine was more intensive (teat-cleaning score $\geq$5.8), the next splitter was time of in-tank storage >15 h, which gave numerically better sensory scores. This result is seemingly contradictory to previous studies because it is generally reported that growth of microorganisms is slowed by low-temperature in-tank storage (<6°C), whereas psychrotrophic bacteria such as *Pseudomonas* continue to proliferate (Perin et al., 2019; Skeie et al., 2019; Baeza-Campone et al., 2020). However, some studies reported no growth of psychrotrophic bacteria during refrigerated storage <6°C (Malacarne et al., 2013; O'Connell et al., 2016). Both of these studies had low levels of psychrotrophic bacteria at baseline, which may explain why they developed less during storage. The bulk tank milks studied here also had low levels of *Pseudomonas* compared with other stud-

ies. *Pseudomonas* counts were 2.27 ± 1.17 log cfu/ mL of milk compared with 3.61 log cfu/mL after 24 h and 4.05 log cfu/mL after 48 h in storage at 4°C in Malacarne et al. (2013), and 3.00 to 3.14 log cfu/mL in O'Connell et al. (2016). The total bacterial count found here was similar to that of previous studies. In the sensory dimension tree, quality scores from the subsets with more or less intensive teat cleaning were thus not significantly different (from 3.8 to 5.3).

Concerning the health dimension tree, when no corn silage was distributed, quality scores were higher when the teats were cleaner, which may be explained by the *E. coli* count indicator. Indeed, this finding is consistent with Zucali et al. (2011), who reported that herds with a higher number of cleaned cows gave bulk tank milks that tended to have lower *E. coli* counts.

Regarding overall cheese quality, the regression tree identified combinations of system-level practices to optimize scores: farms with MB or mixed/crossbred herds receiving grass-based diets in adequate housing conditions to maintain very good teat cleanliness were able to produce milk of good quality for processing into pressed, uncooked, nonstandardized raw milk cheese. Indeed, breed was the most important practice in the cheese technological and nutritional trees and in the overall cheese quality tree. Quality scores were better for milks from MB than from HF herds. The next most influential practices were tied to corn silage in the diets: presence or absence of corn silage or whether corn silage was the main dietary forage. These practices were present in the cheese sensory and health trees and in the cheese overall quality tree. Quality score was better when diets contained little or no corn silage. The third practice in the cheese health tree and the overall quality tree was cow teat dirtiness. Better scores were obtained with more intensively cleaned cows. These influential practices present in both dimensions and overall quality trees could explain why overall quality score was strongly and positively correlated (r >0.5, $P < 0.001$) with sensory, technological, health, and nutritional scores.

Field Applications: Utility and Limits of On-Farm Data and Regression Trees

This innovative study found encouraging results, because, for the first time, it was possible using real on-farm data to correctly explain the overall intrinsic quality of milk for cheese processing based on combinations of farming practices.

Regression trees are useful management tools (in this case for milk quality) for 5 main reasons. First, they allow users to prioritize farming practices in terms of importance in explaining milk quality from the root to the leaves of the tree. Second, they can identify combinations of practices that result in good milk quality. Third, compared with linear regression, regression trees clearly show the possibility to obtain the same quality of milk from different combinations of practices (e.g., systems with HF herds not fed corn silage had similar overall quality of milk for cheese as systems with MB and mixed/crossbred herds fed corn silage). Fourth, they can identify actionable drivers to improve the quality of milk (threshold values for quantitative practices and modalities for qualitative practices). For example, <31% or ≥31% of concentrates in the diet separated medium from high scores on the technological dimension. Finally, they are easy to understand and interpret, as reported by De'ath and Fabricius (2000).

Working with on-farm data allowed us to study a wide range of practices from different milk-producing regions of France in terms of breed, diet, housing, and milking conditions applied in those farms. In our sample, the diversity of practices may be somewhat limited by the fact that the livestock systems sometimes fell into patterns, such as intensive farming with HF cows, corn silage-based diets, and an AMS, or more extensive farming such as grassland-based systems with MB. The major practices (forage system, breed, and milking machine) thus sometimes partly overlapped. However, the regression tree method allows us to mitigate this issue by, for example, separating the data set first by breed (MB vs. HF; Figure 2A) and then evaluating within-node data without the breed factor interfering with the next practices (intra-breed subsets). The importance measure of the variables obtained by the random forest method also allows us to identify whether certain practices could have similar power to explain the variation in the quality scores. Here, main forage in diet, dominant breed, and the type of milking machine were always separate. Therefore, even if one of these practices proved the most influential in the trees, it did not hide the others. However, the structure of our database means that subsets created by nodes of the tree cannot, in turn, be subdivided. We chose to have a minimum of 10 individuals in the final leaves as a parameter for the construction of the trees. Therefore, any node with <20 individuals cannot be subdivided. For example, farms with HF herds that did not distribute corn silage at all were limited in number (n = 11). This explains why in the overall cheese quality tree (Figure 3A), the algorithm could not integrate new influential practices to explain their scores in finer detail. Conversely, the farms with HF herds that fed corn silage were greater in number (n = 33), and thus we were able to integrate an additional influential practice (herd teat dirtiness) in the tree.

More globally, a sample of a hundred or so farms is a limitation for studying more complex combinations of practices. In an effort to avoid overfitting, the trees had to be pruned quite short, and so they contained combinations of 2 to 4 farming practices. Therefore, the goodness of fit, R^2, of some trees indicated that a large part of the variability of the quality scores remained to be explained. This preliminary study should be consolidated in the future. For that, sampling more farms would make it possible to study more complex combinations of practices without overfitting, to improve the accuracy of the results and develop more generic regression trees. For that purpose, the database could be completed in 2 directions in order of importance: (1) by a wider spectrum of intra-breed practices; for example, farms with grazing HF cows or farms with MB cows fed corn silage–based diets, and intra-milking machine type; for example, with farms feeding pasture- or hay-based diets and an AMS; and (2) by a wider spectrum of practices encompassing other breeds, more contrasted diets (e.g., transition diets without one dominant forage), and other housing and milking configurations.

Data collected from farmer surveys can be less accurate then data measured on experimental farms. However, to mitigate this bias, some of the information recorded here, such as milking practices, housing conditions, and animal dirtiness, was observed and measured rather than being collected (by survey) from farmers. This approach also helped make the regression trees easily usable in the field, because the information needed to apply them is readily accessible directly on-farm.

The regression trees obtained in this study cannot be extrapolated to all dairy products. Indeed, the assessments used to calculate quality scores focused on 2 specific products: pressed uncooked nonstandardized raw milk cheese and semi-skimmed standardized UHT milk (Rey-Cadilhac et al., 2021). Therefore, any recommendations that could be made based on the regression trees presented here are applicable only to these 2 products. The practices found to be most important for explaining quality scores were not the same for cheese as for UHT milk (Figure 3); breed was the main driver in the overall cheese quality tree, whereas dominant forage was the main driver in the overall UHT milk quality tree. Therefore, the combinations of practices that will achieve good raw milk for UHT milk are not the same as those to obtain a good raw milk for cheese. Practices adopted and advice given need to be adjusted and adapted according to the end-purpose of the milk. Thus, it may prove complicated to advise farmers who deliver milk for multiple processes or who do not know the fate of their milk.

CONCLUSIONS

In this study, we found that breed and diet are the most influential practices in terms of shaping all sensory, technological, health, and nutritional dimensions and overall quality scores of milk quality for cheese. However, these preliminary results need to be consolidated with larger and more complex databases. The main contribution of this study is the use of the regression trees to identify combinations of practices that affect milk quality. The focus on cheese assessment showed that the combinations of practices that had the biggest influence on quality scores were different between different dimensions of quality. Moreover, comparison of the cheese and UHT overall quality trees showed that the combinations of practices that achieve the best overall quality scores were different according to the end-purpose of the milk. Consequently, advice on farming practices to improve the intrinsic quality of milk needs to be adjusted and adapted according to end-purpose of the collected milk.

ACKNOWLEDGMENTS

REFERENCES

Bacic, B., H. Jackson, and L. F. L. Clegg. 1968. Distribution of bacteria in milk drawn directly from the cow's udder. J. Dairy Sci. 51:47–49. https://doi.org/10.3168/jds.S0022-0302(68)86918-8.

Baeza-Campone, E., S. Prache, A. Clinquart, C. Feidt, J. Gautron, L. Guiller, B. Lebret, F. Lefèvre, and B. Martin. 2020. Variabilité des propriétés des produits animaux selon les conditions d'élevage et de transformation - les determinants. Pages 183–532 in La Qualité Des Aliments d'origine Animale Selon Les Conditions de Production et de Transformation. Rapport de l'expertise Scientifique Collective. S. Prache and V. Santé-Lhoutelier, ed. INRAE.

Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen. 1984. Classification and Regression Trees. Wadsworth.

Chilliard, Y., F. Glasser, A. Ferlay, L. Bernard, J. Rouel, and M. Doreau. 2007. Diet, rumen biohydrogenation and nutritional quality of cow and goat milk fat. Eur. J. Lipid Sci. Technol. 109:828–855. https://doi.org/10.1002/ejlt.200700080.

Coulon, J. B., and B. Rémond. 1991. Variations in milk output and milk protein content in response to the level of energy supply to the dairy cow: A review. Livest. Prod. Sci. 29:31–47. https://doi.org/10.1016/0301-6226(91)90118-A.

Couvreur, S., C. Hurtaud, P. G. Marnet, P. Faverdin, and J. L. Peyraud. 2007. Composition of milk fat from cows selected for milk fat globule size and offered either fresh pasture or a corn silage-based diet. J. Dairy Sci. 90:392–403. https://doi.org/10.3168/jds.S0022-0302(07)72640-1.

De'ath, G., and K. E. Fabricius. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. Ecology 81:3178–3192. https://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2.

Dewhurst, R. J., K. J. Shingfield, M. R. F. Lee, and N. D. Scollan. 2006. Increasing the concentrations of beneficial polyunsaturated fatty acids in milk produced by dairy cows in high-forage systems. Anim. Feed Sci. Technol. 131:168–206. https://doi.org/10.1016/j.anifeedsci.2006.04.016.

Ferlay, A., L. Bernard, A. Meynadier, and C. Malpuech-Brugère. 2017. Production of trans and conjugated fatty acids in dairy ruminants and their putative effects on human health: A review. Biochimie 141:107–120. https://doi.org/10.1016/j.biochi.2017.08.006.

Ferlay, A., B. Graulet, and Y. Chilliard. 2013. Maîtrise par l'alimentation des teneurs en acides gras et en composés vitaminiques du lait de vache. INRA Prod. Anim. 26:177–192. https://doi.org/10.20870/productions-animales.2013.26.2.3146.

Frétin, M., B. Martin, E. Rifa, V.-M. Isabelle, D. Pomiès, A. Ferlay, M.-C. Montel, and C. Delbès. 2018. Bacterial community assembly from cow teat skin to ripened cheeses is influenced by grazing systems. Sci. Rep. 8:200. https://doi.org/10.1038/s41598-017-18447-y.

Gaignon, P., M. Gelé, C. Hurtaud, and A. Boudon. 2018. Characterization of the nongenetic causes of variation in the calcium content of bovine milk on French farms. J. Dairy Sci. 101:4554–4569. https://doi.org/10.3168/jds.2017-14043.

Hogenboom, J. A., L. Pellegrino, A. Sandrucci, V. Rosi, and P. D'Incecco. 2019. Invited review: Hygienic quality, composition, and technological performance of raw milk obtained by robotic milking of cows. J. Dairy Sci. 102:7640–7654. https://doi.org/10.3168/jds.2018-16013.

Legarto, J., M. Gelé, A. Ferlay, C. Hurtaud, G. Lagriffoul, I. Palhière, J. L. Peyraud, B. Rouillé, and P. Brunschwig. 2014. Effets des conduites d'élevage sur la production de lait, les taux butyreux et protéique et la composition en acides gras du lait de vache, chèvre et brebis évaluée par spectrométrie dans le moyen infrarouge. INRA Prod. Anim. 27:269–282. https://doi.org/10.20870/productions-animales.2014.27.4.3073.

Macheboeuf, D. D., J. B. Coulon, and P. D'hour. 1993. Aptitude à la coagulation du lait de vache. Influence de la race, des variants génétiques des lactoprotéines du lait, de l'alimentation et du numéro de lactation. INRA Prod. Anim. 6:333–344. https://doi.org/10.20870/productions-animales.1993.6.5.4214.

Malacarne, M., A. Summer, P. Franceschi, P. Formaggioni, M. Pecorari, G. Panari, P. Vecchia, S. Sandri, E. Fossa, C. Scotti, and P. Mariani. 2013. Effects of storage conditions on physico-chemical characteristics, salt equilibria, processing properties and microbial development of raw milk. Int. Dairy J. 29:36–41. https://doi.org/10.1016/j.idairyj.2012.10.005.

Martin, B., and J.-B. Coulon. 1995. Facteurs de production du lait et caractéristiques des fromages. I. Influence des facteurs de production sur l'aptitude à la coagulation des laits de troupeaux. Lait 75:61–80. https://doi.org/10.1051/lait:199515.

Monsallier, F., I. Verdier-Metz, C. Agabriel, B. Martin, and M.-C. Montel. 2012. Variability of microbial teat skin flora in relation to farming practices and individual dairy cow characteristics. Dairy Sci. Technol. 92:265–278. https://doi.org/10.1007/s13594-012-0064-7.

Montel, M.-C., S. Buchin, A. Mallet, C. Delbes-Paus, D. A. Vuitton, N. Desmasures, and F. Berthier. 2014. Traditional cheeses: Rich and diverse microbiota with associated benefits. Int. J. Food Microbiol. 177:136–154. https://doi.org/10.1016/j.ijfoodmicro.2014.02.019.

Nozière, P., B. Graulet, A. Lucas, B. Martin, P. Grolier, and M. Doreau. 2006. Carotenoids for ruminants: From forages to dairy products. Anim. Feed Sci. Technol. 131:418–450. https://doi.org/10.1016/j.anifeedsci.2006.06.018.

O'Connell, A., P. Ruegg, K. Jordan, B. O'Brien, and D. Gleeson. 2016. The effect of storage temperature and duration on the microbial quality of bulk tank milk. J. Dairy Sci. 99:3367–3374. https://doi.org/10.3168/jds.2015-10495.

Perin, L. M., J. G. Pereira, L. S. Bersot, and L. A. Nero. 2019. The microbiology of raw milk. Pages 45–64 in Raw Milk. L. A. Nero and A. F. De Carvalho, ed. Academic Press. https://doi.org/10.1016/B978-0-12-810530-6.00003-1

Rey-Cadilhac, L., R. Botreau, A. Ferlay, S. Hulin, C. Hurtaud, R. Lardy, B. Martin, and C. Laurent. 2021. Co-construction of a method for evaluating the intrinsic quality of bovine milk in relation to its fate. Animal 15:100264. https://doi.org/10.1016/j.animal.2021.100264.

Rey-Cadilhac, L., A. Ferlay, M. Gelé, S. Léger, and C. Laurent. 2022. Supplementary Material – Regression trees to identify combinations of farming practices that afford the best overall intrinsic quality of milk. Mendeley Data, V1. https://doi.org/10.17632/3xdpk76v8b.1.

Skeie, S. B., M. Håland, I. M. Thorsen, J. Narvhus, and D. Porcellato. 2019. Bulk tank raw milk microbiota differs within and between farms: A moving goalpost challenging quality control. J. Dairy Sci. 102:1959–1971. https://doi.org/10.3168/jds.2017-14083.

Vacheyrou, M., A.-C. Normand, P. Guyot, C. Cassagne, R. Piarroux, and Y. Bouton. 2011. Cultivable microbial communities in raw cow milk and potential transfers from stables of sixteen French farms. Int. J. Food Microbiol. 146:253–262. https://doi.org/10.1016/j.ijfoodmicro.2011.02.033.

van Buuren, S., and K. Groothuis-Oudshoorn. 2011. mice: Multivariate imputation by chained equations in R. J. Stat. Softw. 45:1–67. https://doi.org/10.18637/jss.v045.i03.

Verdier-Metz, I., V. Michel, C. Delbès, and M.-C. Montel. 2009. Do milking practices influence the bacterial diversity of raw milk? Food Microbiol. 26:305–310. https://doi.org/10.1016/j.fm.2008.12.005.

Zucali, M., L. Bava, A. Tamburini, M. Brasca, L. Vanoni, and A. Sandrucci. 2011. Effects of season, milking routine and cow cleanliness on bacterial and somatic cell counts of bulk tank milk. J. Dairy Res. 78:436–441. https://doi.org/10.1017/S0022029911000598.

ORCIDS

L. Rey-Cadilhac ● https://orcid.org/0000-0002-4554-7662

A. Ferlay ● https://orcid.org/0000-0002-0651-792X

M. Gelé ● https://orcid.org/0000-0002-8186-1851

C. Laurent ● https://orcid.org/0000-0002-8350-8496