Scientific Reports

https://doi.org/10.1038/s41598-026-55162-z

Article in Press

Machine learning based detection of subacute ruminal acidosis in early lactation dairy cows using multi-sensor behavioral, physiological, and milk production data

Received: 2 January 2026

Accepted: 22 May 2026

Published online: 29 May 2026

Cite this article as: Grigė S., Girdauskaitė A., Grigas O. *et al.* Machine learning based detection of subacute ruminal acidosis in early lactation dairy cows using multi-sensor behavioral, physiological, and milk production data. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-55162-z

Samanta Grigė, Akvilė Girdauskaitė, Ovidijus Grigas, Sigitas Japertas, Dovilė Malašauskienė, Mindaugas Televičius, Mingaudas Urbutis, Karina Džermeikaitė, Justina Krištolaitytė & Ramūnas Antanaitis

Machine learning based detection of subacute ruminal acidosis in early lactation dairy cows using multi-sensor behavioral, physiological, and milk production data

Samanta Grigė[1]*, Akvilė Girdauskaitė [1], Ovidijus Grigas[2], Sigitas Japertas[3], Dovilė Malašauskienė[1], Mindaugas Televičius[1], Mingaudas Urbutis[1], Karina Džermeikaitė[1], Justina Krištolaitytė[1], Ramūnas Antanaitis[1]

[1]Animal Clinic, Veterinary Academy, Lithuania University of Health Sciences, Tilžės Str. 18, LT-47181 Kaunas, Lithuania

[2]Faculty of Informatics, Kaunas University of Technology, 50254 Kaunas, Lithuania

[3]Practical Training and Research Center, Lithuanian University of Health Sciences, Topolių g. 6, LT-54310 Kaunas, Lithuania

*Corresponding author: Samanta.grige1@lsmu.lt

ABSTRACT

Subacute ruminal acidosis (**SARA**) is a common metabolic disorder in early lactation dairy cows that negatively affects rumen function, milk production, and animal welfare. Early identification remains challenging because clinical signs are often subtle and transient. The aim of this study was to evaluate whether multi sensor behavioral, physiological, and milk-production data could be used to identify cows experiencing SARA using machine-learning approaches. The study included early-lactation Holstein cows during the first 100 days in milk. The final dataset comprised 636 cow-day observations, including 134 SARA cases and 502 clinically healthy controls. Cow identification numbers were additionally used as grouping variables during cross-validation to ensure complete separation of individual animals between training and validation subsets. SARA was defined based on continuous ruminal pH measurements, where cows were classified as SARA when ruminal pH remained between 5.2 and 5.8 for at least 180 min per day. Sensor derived variables included rumination time, activity, water intake, reticulorumen temperature, milk yield, and milk composition obtained from intraruminal boluses and an in-line milk analyzer. Six supervised machine learning classifiers were developed to classify SARA status based on combined sensor data. Among the evaluated models, SVM demonstrated the highest discriminatory performance, which achieved an area under the curve (**AUC**) of 0.97, accuracy of 0.95, sensitivity of 0.86, and specificity of 0.98 under repeated cow-level grouped cross-validation. Random forest showed similar performance (AUC = 0.97; accuracy = 0.93 and 0.98, respectively). Across all models, specificity was consistently higher than sensitivity, indicating that healthy cows were classified more accurately than SARA cases. These results demonstrate that integrated behavioral, physiological, and milk production data obtained from automated sensor systems can support classification of cows experiencing SARA under commercial farm conditions. The findings support the potential of multi sensor monitoring systems combined with machine

learning classifiers as a tool for automated detection of rumen health disturbances in precision dairy farming.

Key words: machine learning, dairy cattle, innovations, biosensors.

INTRODUCTION

Subacute ruminal acidosis **(SARA)** remains one of the most common metabolic problems in high-producing dairy cows worldwide. Frequent episodes of moderately low ruminal pH hinder the breakdown of fiber, change fermentation pathways, and increase systemic and local inflammation, all of which lower productivity and increase vulnerability to secondary disorders [1], [2]. Early lactation is considered the period of highest risk, as cows experience a sharp rise in nutrient requirements, variable dry matter intake, and a strong shift toward rapidly fermentable diets [3]. Early identification of SARA is especially crucial since these physiological changes frequently put cows at risk for unstable rumen fermentation.

Diagnosing SARA under commercial conditions remains challenging because many affected cows show no overt clinical signs. Continuous intraruminal pH measurement is acknowledged as the most accurate diagnostic technique, but its practical application is constrained by sensor costs, maintenance needs, and challenges related to long-term deployment in big herds [4]. The limitations of relying on single markers are highlighted by the inconsistent sensitivity and specificity of indirect indicators, such as variations in rumination time, milk fat content, or the fat-to-protein ratio, across herds, feeding systems, and lactation stages [5], [6]. As a result, imperfect combinations of clinical observations, productivity data, and intermittent pH tests remain necessary for herd-level monitoring.

The rapid advancement of precision dairy technologies has made it possible to identify rumen health issues early. Modern precision dairy technologies now provide continuous, high-frequency records of behavioral and physiological data that reflect changes in cow health status, generated through intraruminal boluses, rumination sensors, accelerometers, and in-line milk analyzers [7], [8]. Although these data streams capture subtle deviations in rumination, activity, reticulorumen temperature, and milk composition that often precede clinical disease, most studies have evaluated these indicators separately, and it is still unclear how much multi-sensor data can collectively characterize SARA, especially in early-lactation cows [9].

Despite these advancements, there is still limited research specifically applying machine learning **(ML)** to SARA. Existing studies have limited ability to capture the multifactorial nature of SARA because they often focus on discrete markers, such as rumination behavior or milk content, rather than integrating numerous physiological and behavioral data streams [10], [11]. Since eating behavior, rumen motility, fermentation stability, and systemic metabolism are all impacted by SARA at the same time, its detection is likely improved when combining multiple sensor inputs rather than relying on a single metric [12], [2]. Rumination, activity, intraruminal temperature or pH, and milk composition characteristics can now be continuously measured using multi-sensor platforms. Several studies have demonstrated that incorporating these variables improves the early detection of metabolic disturbances when compared to single indicator approaches [13], [14]. Previous studies have already explored non-invasive and machine-learning based approaches for SARA prediction, but their predictive performance and field applicability remain limited, indicating the need for further integration of behavioural, physiological, and milk-derived data streams [15], [16].

Although wearable and sensor technologies have expanded rapidly in precision dairy farming, several important limitations remain, particularly regarding multisensor integration, interoperability between different monitoring platforms, and the translation of large data streams into practical tools for farmers. Many currently available systems operate as separate data sources, and combining behavioral, physiological, and milk production data into unified analytical frameworks remains a major challenge in precision livestock farming. Furthermore, the practical implementation of these technologies requires not only accurate algorithms but also robust, user oriented systems that can support on farm decision. These challenges have been highlighted in recent reviews of wearable monitoring technologies in dairy cattle, which emphasize that the next step in precision livestock farming is the development of integrated multisensor systems capable of converting complex sensor data into actionable health information [17], [18].

Therefore, the objective of this study was to evaluate whether integrated multi sensor physiological, behavioral, and milk composition data can reliably distinguish early lactation cows experiencing SARA from clinically healthy animals using supervised machine learning models. We hypothesized that models capable of capturing nonlinear relationships would outperform linear models, because SARA is a multifactorial disorder characterized by complex interactions among rumination behavior, milk composition variables, and activity patterns rather than a single linear predictor.

METHODS

Ethical approval

All animal procedures were reviewed and approved by the Institutional Animal Care and Use Committee of the Lithuanian University of Health Sciences (Protocol No. G2-227, approved on 7 March 2025). The study was conducted in accordance with the requirements of EU Directive 2010/63/EU on the protection of animals used for scientific purposes.

Animals and experimental design

This observational study was carried out at the Practical Training and Research Center of the Lithuanian University of Health Sciences and involved primiparous and multiparous Holstein-Friesian cows in early lactation. The trial commenced on 6 October 2023 and continued until 15 August 2025. A total of 636 cow-day observations were classified into two groups based on ruminal pH measurements. Cows were assigned to the

SARA group when ruminal pH remained between 5.2 and 5.8 for at least 180 min/day (n = 134) [1], [19]. Clinically healthy cows served as controls (H) (n = 502), and were frequency-matched to SARA cows by parity and stage of lactation. The final dataset comprised 636 observations, with one observation day retained per cow. For each observation, a single 4-h analytical window was used as model input. Cows exhibiting any concurrent disorder (e.g., mastitis, ketosis, metritis, lameness, or other clinical disease) or lacking complete data were excluded from both groups.

All animals were housed in free-stall barns under consistent management and nutritional conditions. The study period corresponded to early lactation, defined as the first 100 days in milk [20]. Cows were milked using DeLaval milking robots (DeLaval Inc., Tumba, Sweden). The average body weight of the herd was 550 ± 45 kg. Barns were equipped with a DeLaval mechanical ventilation system (DeLaval Inc., Tumba, Sweden). In 2024, the mean annual milk yield per cow was 10,310 kg, with average fat and protein contents of 4.2% and 3.5%, respectively. Throughout the study, cows received a total mixed ration (**TMR**) formulated by a professional nutritionist to meet their physiological and production requirements. Feeding occurred twice daily (08:00 and 16:00), and water was available ad libitum. The detailed composition of the TMR is presented in Table 1.

Table 1. Composition of TMR for lactating dairy cows

TMR Component	Value
Mineral mix	6.0%
Grain concentrate mash	49.0%

Grass silage	10.0%
Corn silage	31.0%
Alfalfa grass hay	4.0%
Chemical composition:	
Dry matter (DM)	50.7%
Neutral detergent fiber	28.3% of DM
Acid detergent fiber	19.8% of DM
Net lactation energy	1.6 Mcal/kg
Crude protein	15.8% of DM
Non-fiber carbohydrates	38.7% of DM

Registration of the parameters

In this study, milk composition (milk yield, fat, protein, lactose contents, and fat-to-protein ratio (**FPR**)) was continuously monitored using the Brolis HerdLine in-line milk analyzer (Brolis Sensor Technology, Vilnius, Lithuania), while rumination time, water intake, reticulorumen temperature, and activity levels were tracked using the SmaXtec monitoring system (SmaXtec Animal Care GmbH, Graz, Austria).

The daily milk composition data for each cow were derived from measurements generated by the Brolis in-line milk analyzer, which operates in the 2100–2400 nm spectral range using a GaSb-based, widely tunable external cavity laser spectrometer. This compact device, mounted directly on milking stalls or robotic milking units, monitored milk flow in transmission mode throughout the entire milking process. It required no reagents or routine maintenance. Milk composition data were captured every five seconds, and final values for fat and protein were calculated as flow-weighted averages across each milking session.

Reticulorumen parameters were recorded using SmaXtec boluses (SmaXtec Animal Care GmbH, Graz, Austria), which provide continuous real-time monitoring. Boluses were administered orally by an experienced veterinarian according to the manufacturer's instructions and linked to each cow's identification number before data collection.

SmaXtec boluses recorded reticulorumen pH, temperature, activity, temperature without drink cycles, and rumination time at 10-min intervals. For cows classified as SARA, all parameters were averaged across the 3-hour period during which pH remained below 5.8 (≥180 min/day). For healthy controls, one 4 h segment was selected from a matched cow-day based on parity and stage of lactation, providing methodologically comparable short term datasets for between-group analyses.

Data transmission was facilitated by antennas connected to the SmaXtec system, equipped with a microprocessor, analog-to-digital converter, and external memory. Data processing and storage were managed through the SmaXtec Messenger software (version 4). Throughout the study, continuous measurements of reticulorumen temperature, rumination, activity, and water intake were collected. A detailed list of monitored variables, measurement units, and recording intervals is provided in Table 2.

Table 2. Description of behavioral, physiological, and milk-production variables recorded by automated monitoring systems and used as model inputs.

Sensor	**Trait**	**Measurement Interval**	**Unit**
	Activity	10 min	Arbitrary units (a.u.) specific to

SmaXtec bolus (SmaXtec Animal Care GmbH, Graz, Austria)			SmaXtec’s internal algorithm.
	Water intake		L/day
	Temperature		°C
	Rumination time		min/day
	Reticulorumen temperature		°C
Brolis Herdline in-line milk analyzer (Brolis Sensor Technology, Vilnius, Lithuania)	Milk yield	Measurements were taken every 5 s throughout each milking.	kg/day
	Protein		%
	Fat		%
	Fat to Protein ratio		
	Lactose		%

Development of classification models and applied analytical techniques

Classification models were developed to distinguish SARA cow-day observations from clinically healthy controls using physiological, behavioral, and milk-production variables. Because analytical windows for SARA cows were selected during periods of depressed ruminal pH, the models should be interpreted as concurrent classification models rather than early prediction tools. All analytical procedures were carried out using Python (version 3.13) with the scikit-learn, statsmodels, SciPy, and matplotlib libraries.

Data from the Brolis HerdLine analyzer and SmaXtec boluses were merged by cow identification number and sampling date. Variables recorded at 10-min intervals were summarized as 4-h mean values. For cows classified as SARA, the analytical window corresponded to the period during which ruminal pH met the case definition for depressed pH. For clinically healthy cows, one comparable 4 h segment was selected from a matched cow-day based on parity and stage of lactation to ensure temporal comparability between groups. For each included cow-day observation, one analytical summary window was used as input for the machine-learning models.

Milk yield and composition parameters were summarized as daily averages generated by the Brolis HerdLine analyzer. Water intake was extracted directly from the SmaXtec bolus system. Records were excluded if any physiological disorder (mastitis, ketosis, metritis, lameness) occurred or if essential variables were missing. Based on previous studies [1], [19], the target variable was defined as a binary classification: SARA = 1 when ruminal pH remained between 5.2 and 5.8 for ≥180 min/day, and SARA = 0 otherwise.

A series of supervised machine-learning classifiers were trained to classify SARA status using the sensor-derived features. Six widely used algorithms were implemented: Logistic Regression **(LogReg)**, Random Forest Classifier **(RF)**, Gradient Boosting Classifier **(GBoost)**, Support Vector Machine **(SVM)**, Decision Tree Classifier **(DT)**, Gaussian Naïve Bayes **(NB).** A detailed description of the models is provided in the Table 3.

Table 3. Summary of machine learning model types, architectures, and key parameters used in this study

Model	Type / Architecture	Key Parameters Used	Notes
Logistic Regression (LogReg)	Linear classifier	max_iter = 1000; penalty = L2; solver = 'lbfgs'	Scaled features; produces calibrated probabilities
Random Forest (RF)	Ensemble (bagging)	n_estimators = 300; random_state = 42	Nonlinear model; robust; no scaling required

Gradient Boosting (GBoost)	Ensemble (boosting)	learning_rate = 0.1; n_estimators = 100; random_state = 42	Sequential tree boosting; moderate regularization
Support Vector Machine (SVM, RBF)	Kernel-based classifier	kernel = 'rbf'; probability = True; C = 1.0; gamma = 'scale'; random_state = 42	Requires scaling; strong discriminative performance
Decision Tree (DT)	CART decision tree	random_state = 42	Interpretable; prone to overfitting
Naive Bayes (NB)	Probabilistic classifier	No tunable hyperparameters	Fast baseline model; assumes independence

Before training, numerical features were scaled using a standardization procedure (mean = 0, STD = 1) for models sensitive to feature magnitude (LogReg, SVM, NB). Tree-based models were trained on unscaled data.

Model evaluation was performed using repeated stratified grouped cross-validation at the cow level. This approach ensured that records from the same cow could not appear simultaneously in training and validation subsets, thereby reducing potential information leakage. Five-fold grouped cross-validation was repeated 10 times using different random partitions.

Model performance was assessed using pooled out-of-fold predictions and evaluated with accuracy, sensitivity, specificity, F1 score, and area under the ROC curve (AUC). To quantify uncertainty in model performance, bootstrap resampling of pooled out-of-fold predictions was performed to estimate 95% confidence intervals for all reported metrics. Model performance was evaluated using the following metrics:

- Accuracy **(ACC)**;
- Sensitivity (**SE**) = TP ÷ (TP + FN);
- Specificity (**SP**) = TN ÷ (TN + FP);
- F1 score (**F1**);
- Area under the ROC curve (**AUC**).

Classifier-specific ROC curves were generated based on predicted probabilities for the positive class (SARA = 1). The DeLong method was applied to compute 95% confidence intervals for AUC estimates. Confusion matrices were calculated for threshold-based metrics.

Statistical analyses and complementary modeling

Additional statistical analyses were performed to support the ML results. Descriptive statistics were calculated for all physiological, behavioral, and milk variables. Group differences were evaluated using Shapiro–Wilk tests for normality followed by Mann–Whitney U tests. Spearman correlation analysis was used to assess associations between variables and SARA status.

All results, including comparative model performance and statistical outputs, were exported to Excel files for transparency and reproducibility.

RESULTS

Classification performance metrics with 95% confidence intervals for all models are presented in Table 4. Confidence intervals were estimated using bootstrap resampling of pooled out-of-fold predictions generated during repeated grouped cross-validation.

Table 4. Performance of machine-learning classifiers for SARA classification using repeated cow-level grouped cross-validation with 95% confidence intervals.

Model	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	F1 (95% CI)
LogReg	0.91 (0.88–0.94)	0.87 (0.83–0.90)	0.58 (0.46–0.71)	0.95 (0.89–0.98)	0.65 (0.57–0.73)
RF	0.97 (0.95–0.98)	0.93 (0.91–0.94)	0.72 (0.61–0.82)	0.98 (0.96–0.99)	0.80 (0.74–0.85)
GBoost	0.96 (0.94–0.97)	0.92 (0.89–0.94)	0.72 (0.59–0.84)	0.97 (0.95–1.00)	0.79 (0.69–0.85)
SVM (RBF)	0.97 (0.95–0.99)	0.95 (0.94–0.97)	0.86 (0.78–0.96)	0.98 (0.95–0.99)	0.89 (0.84–0.94)
DT	0.83 (0.77–0.89)	0.88 (0.85–0.92)	0.75 (0.59–0.84)	0.92 (0.87–0.97)	0.73 (0.65–0.81)
NB	0.84 (0.81–0.88)	0.83 (0.78–0.86)	0.50 (0.42–0.60)	0.91 (0.86–0.96)	0.55 (0.47–0.62)

Performance metrics

Performance metrics for all six supervised classifiers obtained using repeated cow-level grouped cross-validation are summarised below and visualised in Figure 1. Results were calculated from pooled out-of-fold predictions generated across repeated grouped validation folds. Validation folds were additionally checked to confirm that no overlap of cow identification numbers occurred between training and validation subsets.

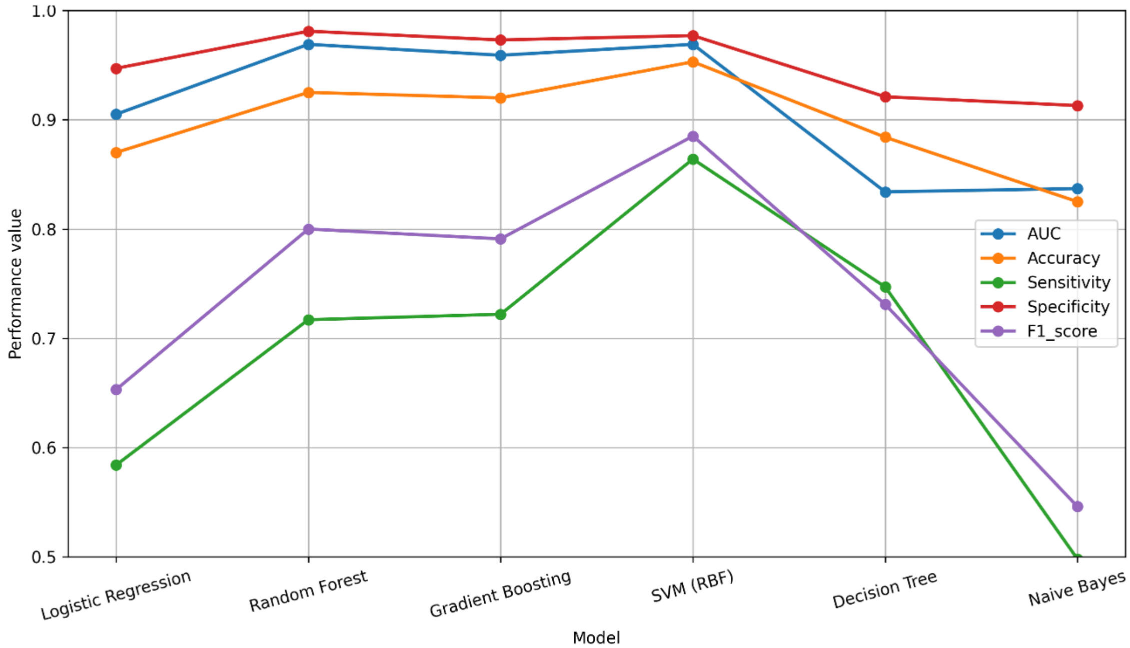

Figure 1. Performance metrics of machine-learning classifiers under repeated grouped cross-validation.

LogReg achieved an AUC of 0.91 and an overall accuracy of 0.87. For SARA cases, specificity was 0.95 and sensitivity was 0.58, resulting in an F1-score of 0.65. The RF classifier showed slightly higher discrimination (AUC 0.97, accuracy 0.93), characterised by a SARA sensitivity of 0.72 and specificity of 0.98 for healthy cows. GBoost performed comparably (AUC 0.96, accuracy 0.92), with a sensitivity of 0.72 and specificity of 0.97. The SVM with an RBF kernel was the best performing model, yielding an AUC of 0.97 and accuracy of 0.95; for SARA, sensitivity reached 0.86, demonstrating a favourable balance between sensitivity and specificity. The DT classifier exhibited lower predictive ability (AUC 0.83, accuracy 0.88), while the NB model performance was (AUC 0.84, accuracy 0.83) with substantially reduced precision and recall for SARA. Overall, the tree based ensemble models (RF, GBoost) and the SVM consistently outperformed other classifiers, each achieving AUC values >0.96 and accuracies of approximately 0.90–0.93. ROC curves for all six classifiers are presented in Figure 2.

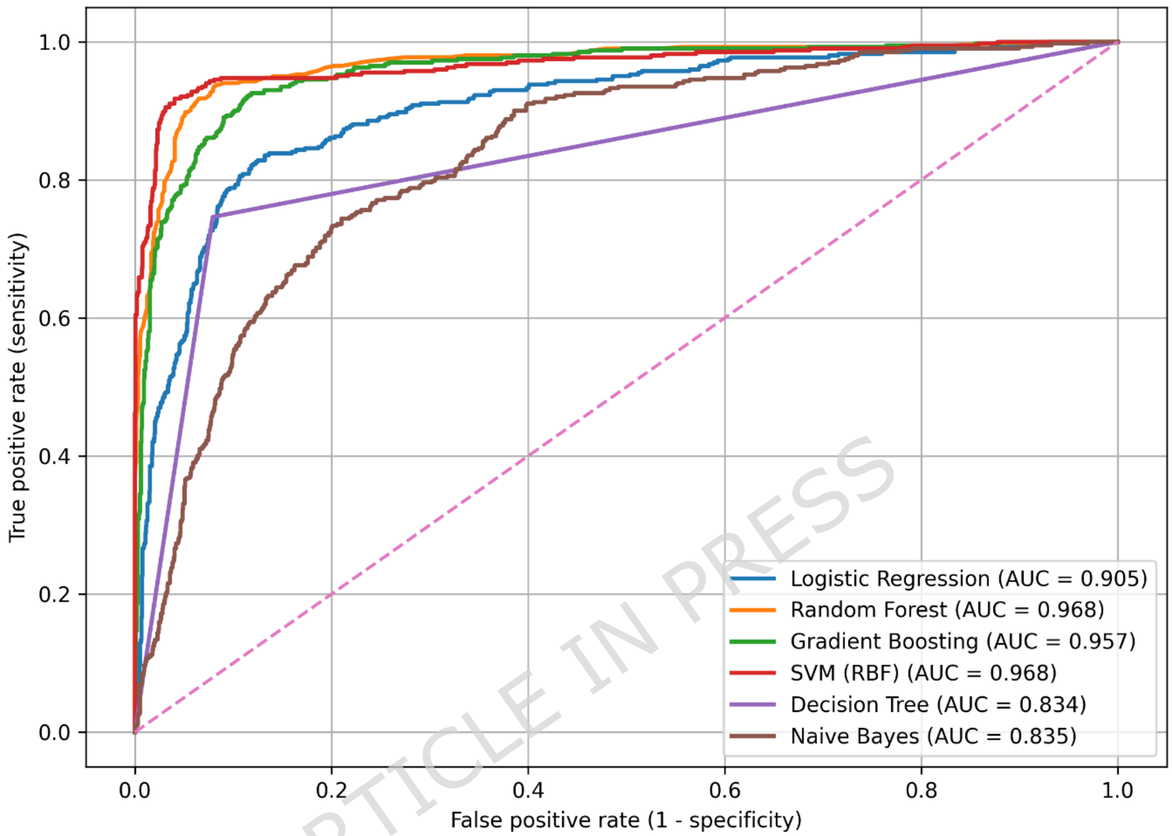

Figure 2. ROC curves obtained from repeated cow level grouped cross-validation.

A detailed comparison of model performance using ACC, SE, SP, and F1 is presented in Figure 3. The SVM classifier showed the strongest overall discrimination, achieving ACC = 0.95, SE = 0.86, SP = 0.98, and F1 = 0.89, together with the highest AUC (0.968). The RF and GBoost models demonstrated similarly high performance (RF: ACC = 0.93, SE = 0.72, SP = 0.98, F1 = 0.80; GBoost: ACC = 0.92, SE = 0.72, SP = 0.97, F1 = 0.79), indicating strong ability to correctly classify both SARA and healthy cows. LogReg provided moderate but balanced performance (ACC = 0.87, SE =

0.58, SP = 0.95, F1 = 0.65; AUC = 0.905). In contrast, DT (ACC = 0.88, SE = 0.75, SP = 0.92, F1 = 0.73) and NB (ACC = 0.83, SE = 0.50, SP = 0.91, F1 = 0.55) showed weaker classification ability, particularly due to reduced SE and F1 for SARA detection.

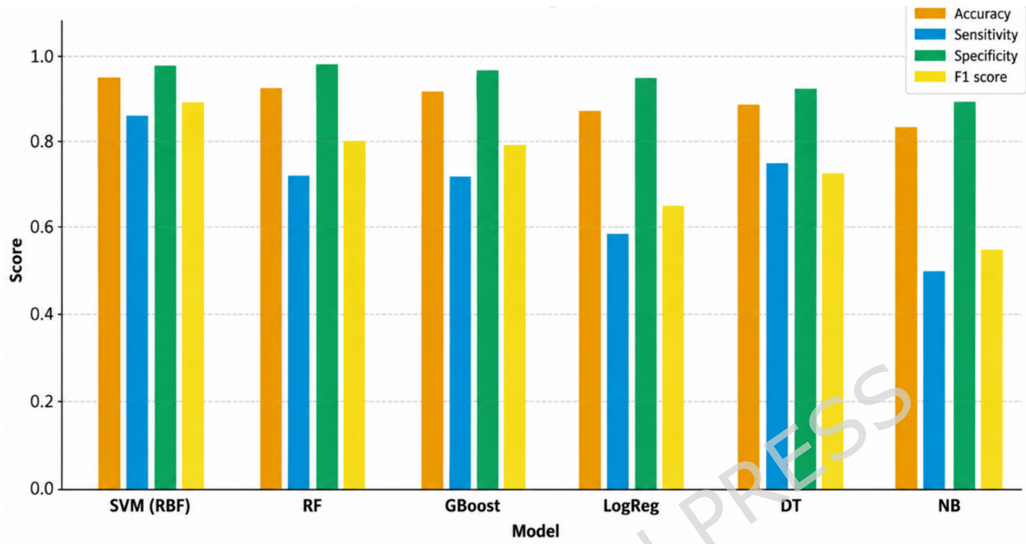

Figure 3. Model performance comparison.

Repeated grouped cross-validation produced consistently high discriminatory performance for the nonlinear models, particularly RF, SVM, and GBoost, which achieved pooled out-of-fold AUC values above 0.95. Logistic Regression showed slightly lower but still strong performance (AUC = 0.90), whereas DT and NB demonstrated comparatively weaker discrimination. The consistency of results across repeated grouped validation folds suggests that model performance remained stable after controlling for potential cow-level information leakage.

Comparison between groups

In addition to model performance, the biological relevance of variables was evaluated using group comparisons and correlation analysis, which showed that rumination time, milk yield, lactose concentration, and temperature parameters were most strongly associated with SARA and therefore likely contributed most to model classification. As expected, daily mean ruminal pH was markedly lower in the SARA group compared with H group (5.61 vs. 6.71; $P < 0.001$). Core temperature measured by the SmaXtec bolus was slightly but significantly reduced in SARA cows (38.85°C vs. 38.92°C; $P = 0.0349$). Rumination time was also lower in SARA cows (395.9 vs. 431.3 min/d; $P = 0.0086$), indicating compromised feeding behavior.

Milk yield differed significantly between groups, with SARA cows producing less milk than healthy cows (9.86 vs. 12.17 kg/d; $P < 0.001$). Milk lactose percentage was also reduced in SARA cows (4.53% vs. 4.79%; $P < 0.001$). In contrast, milk fat percentage tended to be higher in SARA cows (4.48% vs. 4.25%; $P = 0.0896$), although this difference did not reach conventional significance ($0.05 \leq P < 0.10$). Milk protein percentage and the FPR did not differ between groups ($P > 0.50$).

Reticulorumen temperature recorded by the SmaXtec bolus was substantially lower in SARA cows (35.68°C vs. 35.97°C; $P < 0.0001$), consistent with reduced rumination and metabolic activity. Measures of physical activity and water intake did not differ significantly between groups ($P > 0.13$). Temperature-derived variables such as normal temperature and temperature without drink cycles also showed no significant differences. These results are summarized in Table 5.

Table 5. Comparison of physiological and milk parameters between H and SARA groups using Mann–Whitney U test.
*P < 0.05 (significant)

variable	mean H	mean SARA	STD SARA	STD H	P_value
temperature °C	38,92	38,85	0,39	0,29	0,035
Temp Without Drink cycles	39,39	39,34	0,23	0,31	0,085
Normal temperature *	39,55	39,54	0,24	0,23	0,127
Activity	5,94	6,10	4,01	3,12	0,260
Rumination min/24h **	431,25	395,93	129,96	104,23	0,009
Water_intake in gal	94,34	99,40	29,39	31,15	0,136
Total yield***	12,17	9,86	2,47	2,59	< 0,001
Fat	4,25	4,48	1,09	0,81	0,090
Protein	3,22	3,31	0,53	0,28	0,835
Lactose ***	4,79	4,53	0,27	0,21	< 0,001
Milk temperature °C ***	35,97	35,68	0,76	0,82	< 0,001
FPR	1,32	1,36	0,29	0,24	0,510
DIM	54,2	56,8	11,4	12,1	0,218
Parity	2,3	2,4	1,2	1,1	0,412

**P < 0,01 (highly significant)
***P < 0,001 (very highly significant)

Correlation analysis

Spearman rank correlations between groups and all measured variables are presented in Figure 4. Ruminal pH showed a strong negative correlation with SARA (r = -0.85), confirming that lower pH values were tightly associated with the classification of cows as experiencing SARA.

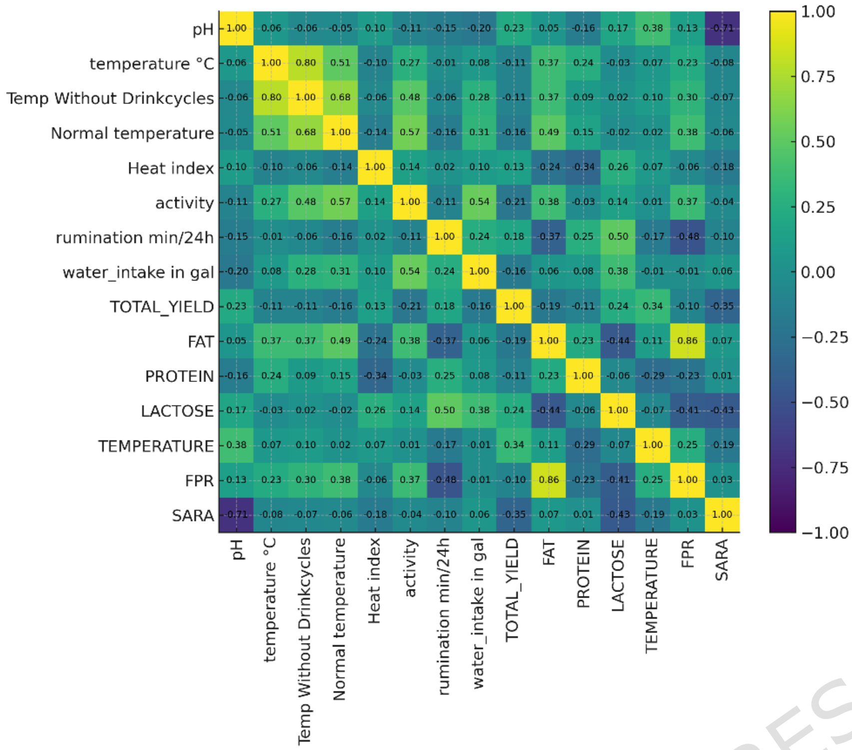

Figure 4. Correlation heatmap.

Among physiological indicators, milk temperature exhibited a moderate negative correlation with SARA (r = - 0.32). Reticulorumen temperature variables, including temperature, normal temperature, and temperature without drink cycles, were only weakly associated with SARA (r < 0.10). Activity demonstrated a minimal correlation (r = + 0.05), indicating limited value as a single predictor.

Behavioral traits showed clearer relationships. Rumination time was moderately and negatively correlated with SARA (r = - 0.32), suggesting that cows with SARA spent less time ruminating compared with healthy cows. Water intake showed a weak-to-moderate positive correlation (r = + 0.18), consistent with increased drinking behavior during periods of lower ruminal pH.

Milk performance traits yielded several notable associations. Daily milk yield was moderately negatively correlated with SARA ($r = -0.42$), and lactose concentration displayed a similar negative association ($r = -0.46$). In contrast, milk fat and protein percentages demonstrated weak positive relationships ($r = +0.10$ to $+0.15$). Consequently, the FPR showed only a weak correlation with SARA ($r = +0.12$), indicating limited diagnostic usefulness in this population of early-lactation cows.

Correlations among variables followed biologically expected patterns, with milk yield negatively correlated with milk fat and positively correlated with lactose, and rumination positively associated with ruminal pH and milk yield. No strong correlations indicative of multicollinearity were observed, as most remained in the low-to-moderate range ($r < 0.50$).

DISCUSSION

This study evaluated the potential of multi-sensor behavioral and physiological data to classify SARA in early lactation dairy cows and compared the performance of six supervised ML classifiers. Using high resolution measurements from in-line milk analyzers and intraruminal boluses, we assessed differences in rumen pH, rumination behavior, temperature-related traits, milk characteristics, and activity between SARA affected and clinically healthy cows.

The high discriminatory performance of SVM (AUC 0.97; ACC 0.95), RF (AUC 0.97; ACC 0.93) and GBoost (AUC 0.96; ACC 0.92) suggests that

SARA in early lactation cows can be differentiated with high accuracy under internal grouped validation when combining multiple behavioural and physiological variables. These algorithms were able to maintain both high overall accuracy and a favourable sensitivity - specificity balance, indicating that SARA is expressed through nonlinear relationships and multivariate patterns rather than single parameter deviations. Such behaviour is typical in precision livestock datasets, where interactions among rumination dynamics, rumen temperature, lactose concentration, and activity require models capable of adapting to complex feature spaces. This observation is consistent with recent reviews highlighting the growing role of advanced machine-learning techniques in dairy health monitoring and the superiority of nonlinear and ensemble algorithms for disease detection [21].

In contrast, the lower performance of LogReg, DT, and particularly NB can be explained by methodological limitations of these algorithms. LogReg assumes linear relationships between predictors and the outcome, which may not fully capture the multifactorial nature of SARA. Decision trees are prone to overfitting, especially with relatively small datasets, which may explain their lower generalisation performance. The reduced sensitivity of NB likely reflects its assumption of independence between predictors, which is unrealistic in biological systems where variables such as rumination, milk yield, and temperature are physiologically related. Similar limitations of linear and probabilistic models have been reported in other disease-prediction studies, where nonlinear and ensemble approaches provide better discrimination in complex biological datasets [22].

Recent research integrating multimodal sensor, production, and biologic data further corroborates the feasibility of early detection of metabolic disturbances in dairy cows using ML-based frameworks. For example, a 2024 study demonstrated successful early detection of metabolic disorders by combining wearable sensor data with phenotypic traits in a cloud-based ML pipeline, achieving strong predictive performance comparable to ensemble classifiers [23]. Similarly, studies focusing on cattle behavioural monitoring report that tree-based algorithms and SVM consistently outperform linear models for classification tasks in commercial herd environments, reinforcing the value of nonlinear approaches under real-world variability [24]. The reduction in ruminal pH in SARA cows (mean 5.61 compared with 6.71 in the H group) confirms a biologically meaningful state of subacute acidosis and demonstrates that the classification threshold successfully captured relevant rumen dysfunction. This difference reflects a substantial disturbance in fermentation stability, consistent with established evidence showing that even moderate depressions in ruminal pH impair digestion and buffering capacity [25]. Correspondingly, SARA cows exhibited shorter rumination times (approximately 36 min/d less), a response that aligns with the documented suppression of chewing and rumination activity under acidic rumen conditions [26]. Thermal responses followed similar patterns. Core temperature decreased by ~0.07 °C, reticulorumen temperature by ~0.29 °C, milk temperature by ~0.31 °C (37.63 °C vs. 37.94 °C), and the individualized normal temperature baseline by ~0.23 °C in SARA cows. Although these temperature shifts were modest, high-resolution sensor technologies readily detect such subtle changes, which may indicate reduced fermentation heat output or altered rumen motility during low-pH periods. The fact that these thermal alterations occurred alongside reduced rumination suggests that both are manifestations of the same underlying metabolic disturbance [27]. Nevertheless, temperature responses to acidosis are not universally consistent: controlled studies have reported the opposite pattern, where decreasing pH is accompanied by increasing reticulorumen and abomasal temperatures that correlate positively with rectal temperature [28]. This variability underscores the complexity of thermal physiology during ruminal acidosis and the potential influence of diet, acidosis severity, or methodological differences.

Milk performance traits were among the strongest indicators of SARA, both in group comparisons and correlation patterns. Daily milk yield showed a moderate negative correlation with SARA (r = – 0.42), consistent with the group-level reduction of more than 2 kg/d in SARA cows (9.86 vs. 12.17 kg/d). Likewise, lactose concentration showed a moderate negative correlation (r = – 0.46) and was lower in SARA cows (4.53% vs. 4.79%), reflecting impaired carbohydrate fermentation and reduced glucose availability during ruminal acidosis [29]; [30]; [31]. In contrast, milk fat and milk protein exhibited only weak positive correlations with SARA (r ≈ +0.10 to +0.15), and the similarly weak correlation of the fat-to-protein ratio (FPR; r = +0.12) mirrored the absence of meaningful group differences. The slight tendency toward higher milk-fat percentage in SARA cows (+0.23), despite no changes in protein or FPR, suggests that milk fat responses are highly variable and influenced by diet dependent or cow specific factors rather than rumen acidity alone [32]. Instead, they highlight the value of sensor integrated, multivariate diagnostic approaches that combine behavioural, physiological, and milk production data rather than relying on single markers [33]. The practical value of automated behavioural monitoring is also supported by recent evidence showing that ear tag derived behavioural data can assist reproductive decision making in commercial dairy herds, reinforcing the broader utility of PLF tools in herd management [34].

Several limitations of this study should be acknowledged. First, ruminal acidosis is a dynamic disorder and not only a threshold-dependent pH event; therefore, classification based on a single pH criterion may only partially capture its biological complexity [34]. Second, although the total number of observations was considerable, the proportion of SARA cases was lower than that of healthy cows, which may have influenced model sensitivity due to class imbalance. Third, sensor-derived variables were summarized over relatively short 3–4 h analytical windows, which may have limited the detection of more gradual behavioural and physiological changes associated with rumen instability. Fourth, all animals originated from a single herd managed under uniform feeding and housing conditions, which may limit the generalisability of the models to other herds and management systems. To reduce potential information leakage, repeated grouped cross-validation at the cow level was implemented in the revised analysis, ensuring that records from the same cow could not appear in both training and validation subsets. Although this approach substantially strengthens internal validation, the models were still developed using data from a single herd under relatively uniform management and feeding conditions. Therefore, external validation across independent herds and production systems remains necessary before broader practical implementation. Finally, although ruminal pH was excluded from the predictor set, analytical windows were selected from periods of already depressed pH; therefore, the models should be interpreted as tools for concurrent classification rather than true predictive models. Future research should focus on developing predictive models using lagged sensor data recorded hours or days before ruminal pH depression, as well as regression or risk-level modelling approaches to predict ruminal pH dynamics and SARA risk in advance.

CONCLUSIONS

This study demonstrated that the integration of multi-sensor behavioral, physiological, and milk-production data can be used to classify cows experiencing subacute ruminal acidosis during early lactation under commercial farm conditions. Among the evaluated models, SVM, RF, and GBoost achieved the highest discriminatory performance, indicating that nonlinear, multivariate relationships best characterize the physiological and production changes associated with SARA. Group comparisons and correlation analyses showed consistent alterations in rumination activity, thermal parameters, milk yield, and lactose concentration, highlighting the multifactorial nature of the disorder. These findings suggest that automated monitoring systems capturing high-frequency data streams may support the identification of rumen instability and provide useful decision-support information for precision dairy management. However, the models were developed using data from a single herd and evaluated using an internal validation approach; therefore, external validation in independent herds and production systems is required before routine practical application. Future research should focus on validating these models across diverse herds and on developing predictive models based on lagged sensor data for early warning of SARA risk.

NOTES

Funding: “This study received no external funding”;

Ethics statement: “All animal procedures were reviewed and approved by the Institutional Animal Care and Use Committee of the Lithuanian

University of Health Sciences (Protocol No. G2-227, approved on 7 March 2025)”.

Conflict of interest statement: ”The authors have not stated any conflicts of interest.”

DATA AVAILABILITY STATEMENT

The datasets used during the current study available from the corresponding author on reasonable request.

Abbreviations:

ACC – Accuracy;

AUC - Area under the ROC curve;

DT- Decision Tree Classifier**;**

F1 - F1 score;

FPR – fat-to-protein ratio;

GBoost - Gradient Boosting Classifier**;**

LogReg - Logistic Regression;

ML - machine-learning

NB - Gaussian Naïve Bayes**;**

RF - Random Forest Classifier**;**

SARA - subacute ruminal acidosis;

SE – Sensitivity;

SP – Specificity;

SVM - Support Vector Machine;

TMR – total mixed ration;

REFERENCES

[1] Humer, E., Khol-Parisini, A., Metzler-Zebeli, B. U., Gruber, L., and Zebeli, Q. (2017). Signals for identifying cows at risk of subacute ruminal acidosis in dairy veterinary practice. *Journal of Animal Physiology and Animal Nutrition.*

[2] Plaizier, J. C., Li, S., Tun, H. M., Khafipour, E., and others (2022). Invited review: Effect of subacute ruminal acidosis on gut health of dairy cows. *Journal of Dairy Science.*

[3] Zhang, R., Zhu, Y., Xue, Y., Chen, G., Wang, Q., Zhao, J., et al. (2022). Responsive changes of rumen microbiome and metabolome in dairy cows with different susceptibility to subacute ruminal acidosis. *Animal Nutrition*, 8, 331–340.

[4] Golder, H. M., Celi, P., Rabiee, A. R., Heuer, C., Bramley, E., and Lean, I. J. (2023). Characterizing ruminal acidosis risk: A multiherd, multicountry study. *Journal of Dairy Science*, 106, 3155–3175.

[5] Zschiesche, M., Stamer, E., Junge, W., and others (2020). The milk fat-to-protein ratio as indicator for ruminal pH parameters in dairy cows: A meta-analysis. *Dairy*, 1, 259–268.

[6] Mulakala, B., Kim, K., Seo, S., Lee, S., and others (2023). Use of milk proteins as biomarkers of changes in the rumen metaproteome of Holstein cows fed low-fiber, high-starch diets. *Journal of Dairy Science*, 106, 9630–9643.

[7] Soriani, N., Panella, G., and Calamari, L. (2013). Rumination time during the summer season and its relationships with metabolic conditions and milk production. *Journal of Dairy Science*, 96, 5082–5094.

[8] Tangorra, F. M., Bassi, E., Calcante, A., and Croce, A. (2024). Internet of Things (IoT): Sensors application in dairy cattle farming. *Animals*, 14, 3071.

[9] Bonmann, N., Firk, R., Stamer, E., Junge, W., and others (2024). Sensor-based and visual behavioral profiling of dry Holstein cows presenting distinct median core body temperatures. *Animals*, 14, 2832.

[10] Wagner, V., Mertens, K. C., et al. (2020). Machine learning to detect behavioural anomalies in dairy cows under subacute ruminal acidosis. *Computers and Electronics in Agriculture*, 170, 105233.

[11] Huot, S., DeVries, T. J., et al. (2024). Predicting subacute ruminal acidosis from milk mid-infrared estimated fatty acids and machine learning on Canadian commercial dairy herds. *Journal of Dairy Science.*

[12] Stauder, A., Neveu, E., et al. (2020). Distinct responses in feed sorting, chewing behavior, and ruminal acidosis risk between primiparous and multiparous Simmental cows fed diets differing in forage and starch levels. *Journal of Dairy Science*, 103, 8467–8481.

[13] Savela, M., Friggens, N. C., et al. (2024). Comparison of marine algae-based rumen buffer with sodium bicarbonate buffer and its role in increasing milk value. *Journal of Dairy Research*, 91, 391–396.

[14] Džermeikaitė, K., Antanaitis, R., et al. (2025). The impact of heat stress on dairy cattle: Effects on milk quality, rumination behaviour, and reticulorumen pH response using machine learning models. *Biosensors*, 15, 608.

[15] Coppa, M., Martin, C., et al. (2023). On-farm evaluation of multiparametric models to predict subacute ruminal acidosis in dairy cows. *Animal.*

[16] Touil, T., Dehareng, F., et al. (2025). Predicting reticuloruminal pH and subacute ruminal acidosis of individual cows using machine learning and Fourier-transform infrared spectroscopy milk analysis. *Journal of Dairy Science.*

[17] Lamanna, M., Bovo, M., and Cavallini, D. (2025). Wearable collar technologies for dairy cows: A systematized review of the current applications and future innovations in precision livestock farming. *Animals.*

[18] Lamanna, M., Bovo, M., Cavallini, D., et al. (2025). Rethinking wearable technology in dairy cows: Challenges and prospects for smart collars. *The Open Agriculture Journal.*

[19] Plaizier, J. C., Krause, D. O., Gozho, G. N., and McBride, B. W. (2008). Subacute ruminal acidosis in dairy cattle: The physiological process and associated health consequences. *Veterinary Clinics of North America: Food Animal Practice*, 24, 269–281.

[20] Kleen, J. L., Hooijer, G. A., Rehage, J., and Noordhuizen, J. P. (2003). Subacute ruminal acidosis (SARA): A review of the epidemiological, pathological and therapeutical aspects. *Journal of Veterinary Medicine Series A.*

[21] Grzesiak, W., Zaborski, D., et al. (2025). The use of selected machine learning methods in dairy cattle farming: A review. *Animals.*

[22] Perez, M. M., et al. (2025). Screening and selection of a machine learning algorithm for development of a model to select cows for clinical examination using automated health monitoring technologies. *Journal of Dairy Science.*

[23] Ferreira, R. E. P., et al. (2024). Multi-modal machine learning for the early detection of metabolic disorder in dairy cows using a cloud computing framework. *Animal Science Journal.*

[24] Hernández, G., et al. (2024). Machine learning-based prediction of cattle activity using sensor-based data. *Sensors.*

[25] Nagaraja, T. G., and Titgemeyer, E. C. (2007). Ruminal acidosis in beef cattle: The current microbiological and nutritional outlook. *Journal of Dairy Science.*

[26] Minami, N. S., Sousa, R. S., and Oliveira, F. L. C. (2020). Subacute ruminal acidosis in Zebu cattle: Clinical and behavioral aspects. *Animals.*

[27] Arlauskaitė, S., Antanaitis, R., et al. (2025). Precision monitoring of rumination and locomotion in relation to milk fat-to-protein ratio in early lactation dairy cattle. *Frontiers in Veterinary Science.*

[28] Antanaitis, R., Žilaitis, V., et al. (2016). Usefulness of acidity and temperature of the rumen and abomasum in diagnosing SARA in dairy cows after calving. *Polish Journal of Veterinary Sciences.*

[29] Johnson, C. (2022). Effect of dietary factors on rumen metabolism and the microbiome in high-yielding dairy cows. Harper Adams University.

[30] Hanušovský, O., et al. (2018). The dynamic of ruminal content pH change and its relationship to milk composition. *Acta Veterinaria Brno.*

[31] Lee, A. R. (2018). An evaluation of physiological and behavioral indicators of postpartum diseases and heat stress in dairy cows. University of Kentucky.

[32] Sandri, E. C., et al. (2020). Transient reductions in milk fat synthesis and their association with ruminal and metabolic profile in dairy cows fed high-starch diets. *Animals.*

[33] Anchitaalagammai, J. V., Kavitha, S., Murali, S., and Velunachiyar, S. (2025). Enhancing dairy farming: IoT and sensor-based technologies for disease detection and health monitoring. *Proceedings of ICICCS.*

[34] Cavallini D, Giammarco M, Buonaiuto G, Vignola G, De Matos Vettori J, Lamanna M, Prasinou P, Colleluori R, Formigoni A and Fusaro I (2025) Two years of precision livestock management: harnessing ear tag device behavioral data for pregnancy detection in free-range dairy cattle on silage/hay-mix ration. Front. Anim. Sci. 6:1547395. doi: 10.3389/fanim.2025.1547395