Journal of Current Science and Technology, April-June 2026 Vol. 16 No. 2, Article 181

ISSN 2630-0656 (Online)

Cite this article: Phuphaphud, A., Tonmitr, N., Suntara, C., Tanusilp, S., Hanpinitsak, P., & Katanyukul, T. (2026). Detection of estrus in dairy cows: A proof-of-concept near-infrared milk sensing and machine-learning study. *Journal of Current Science and Technology*, *16*(2), Article 181. https://doi.org/10.59796/jcst.V16N2.2026.181

Journal of Current Science and Technology

Journal homepage: https://jcst.rsu.ac.th

Detection of Estrus in Dairy Cows: A Proof-of-Concept Near-Infrared Milk Sensing and Machine-Learning Study

Arthit Phuphaphud[1], Norrawit Tonmitr[1,*], Chanon Suntara[2], Sora-at Tanusilp[1], Panawit Hanpinitsak[1], and Tatpong Katanyukul[1]

[1]Faculty of Engineering, Khon Kaen University, Khon Kaen 40002, Thailand
[2]Faculty of Agriculture, Khon Kaen University, Khon Kaen 40002, Thailand

*Corresponding author; E-mail: norrawit@kku.ac.th

Received 10 November 2025; Revised 3 December 2025; Accepted 6 January 2026; Published online 25 March 2026

Abstract

Accurate estrus detection is critical for reproductive management in dairy herds, yet current methods are either labor-intensive or require dedicated hardware. This proof-of-concept study investigated whether near-infrared (NIR) milk spectra, combined with routine milk-composition data, can be used to detect estrus in dairy cows. Five clinically healthy Thai milking cows were monitored for 21 days each (total 593 milk samples), with estrus labels assigned based on experienced stockperson observations confirmed by behavioral signs and tail-paint rubbing. For each sample, inline NIR transmission spectra (860–1754 nm) and milk composition (fat, protein, lactose, solids-not-fat, density, pH, daily yield) were acquired. During estrus, milk composition showed modest but consistent shifts: protein increased by approximately 0.29 percentage points and lactose by 0.41 percentage points, while daily milk yield decreased by about 7.4 ± 2.7 kg/day relative to non-estrus days. A leakage-aware, leave-one-cow-out cross-validation framework was used to compare six classifiers. Logistic regression, gradient boosting, and decision-tree-based ensembles achieved internal accuracies of between 0.99 and 1.00, with Extra Trees and random forest yielding F1-scores of 0.99 and 0.92, respectively. Feature-importance analysis indicated that specific NIR bands in the water and protein-related regions, together with milk yield and protein percentage, contributed most strongly to estrus discrimination, whereas density and pH had minimal influence. These proof-of-concept results (N = 5, single site) demonstrate technical feasibility but require multi-site validation in substantially larger cohorts (N ≥ 50) before any clinical or on-farm adoption can be recommended.

Keywords**:** *dairy cows; estrus detection; near infrared spectroscopy (NIR); reproductive management*

1. Introduction

Accurate estrus identification is critical for achieving successful artificial insemination and sustaining high reproductive efficiency in dairy farming (Agutu et al., 2024; Denis-Robichaud et al., 2024; Mahendran et al., 2023; McDougall, 2006). Timely detection enables insemination within the narrow fertile window of about 18–24 hours during the 21-day estrus cycle in dairy cows (Folman et al., 1984; Islam, 2011; Yizengaw, 2017). Failure to recognize estrus promptly lengthens calving intervals, lowers herd fertility, and can delay pregnancy confirmation for several months (Paisley et al., 1978; Szenci, 2021).

Traditional estrus-detection methods rely on visual observation of behavioral signs such as mounting, chin resting, and increased locomotor activity (Sayid et al., 2024; Palmer et al., 2010; Reith & Hoy, 2018; Schilkowsky et al., 2021; Sveberg et al., 2015). Although these signs are physiologically reliable, manual observation in large herds is labor-intensive and prone to human error, often resulting in

missed silent estrus events. Consequently, research has focused on automated, noninvasive approaches to improve detection precision and consistency, including behavior monitoring and computer vision systems (da Silva Santos et al., 2023; Arıkan et al., 2023). Milk-composition analysis has been proposed as an alternative indicator because variations in fat and progesterone concentrations are correlated with reproductive status (Petit et al., 2001; Tenghe et al., 2015; Toledo-Alvarado et al., 2018). However, laboratory-based assays such as chromatography and immunoassays remain time-consuming, costly, and invasive.

Near-infrared (NIR) spectroscopy offers a rapid, reagent-free technique for quantifying milk constituents and tracking hormonal changes associated with the estrus cycle (Behdad et al., 2024; Džermeikaitė et al., 2023; Takemura et al., 2015). NIR spectroscopy has been successfully applied to agricultural quality control, including milk analysis and adulteration detection (Holroyd, 2013; Kawamura et al., 2007; Kleanthous et al., 2022; Tsenkova et al., 2006; Wu et al., 2008). Yet, its real-time use for estrus detection in dairy herds remains largely unexplored. NIR spectra capture subtle chemical and structural changes linked to progesterone fluctuations. Therefore, this method represents a promising foundation for automated estrus monitoring. However, previous NIR-based estrus studies have typically involved small, homogeneous cohorts and have not always used standardized on-farm acquisition protocols or rigorous, leakage-aware cross-validation, which may limit their generalizability and inflate apparent performance.

In parallel, machine learning (ML) has become a powerful framework for pattern recognition and classification in agricultural systems (Benos et al., 2021; Evangelista et al., 2021; Hempstalk et al., 2015; Le et al., 2024; Liakos et al., 2018; Sharma et al., 2020; Shine & Murphy, 2021; Wang et al., 2022b). Integrating NIR spectral data with ML algorithms can enable data-driven estrus-identification systems that surpass conventional behavioral observation (Ho et al., 2019; Ho & Pryce, 2020). Nevertheless, challenges remain in spectral preprocessing, class imbalance, and validation design, all of which influence model reliability.

Therefore, this proof-of-concept study aims to standardize an on-farm NIR milk-acquisition protocol and evaluate a machine-learning-based pipeline for estrus detection in dairy cows. The proposed framework integrates spectral preprocessing and class-imbalance handling to classify estrus and non-estrus stages using NIR spectra combined with milk-composition parameters (fat, protein, lactose, and solids-not-fat) in a small (N = 5), single-site cohort. The effectiveness of multiple classifiers is evaluated under cow-wise cross-validation, providing an internally valid assessment of technical feasibility while recognizing that larger multi-site studies will be required to establish population-level performance and clinical utility.

2. Objectives

1) To design and implement an on-farm near-infrared (NIR) milk-sensing system and standardized acquisition protocol suitable for routine milking in a Thai dairy herd.

2) To characterize within-cow changes in milk composition between estrus and non-estrus days.

3) To develop and preliminarily evaluate a leakage-aware machine-learning pipeline that combines NIR spectra with milk-composition features for classifying estrus versus non-estrus at the individual-cow level.

3. Materials and Methods

3.1 Study Design and Population

Milk samples were collected from five Thai Milking Zebu cows (IDs 55, 60, 1755, 5929, 6311) housed at the experimental farm of the Faculty of Agriculture, Khon Kaen University, Thailand. The cows were crossbreds comprising 75% Holstein Friesian and 25% other breeds selected for milk yield and reproductive performance under tropical conditions. For each cow, milk was sampled twice daily at routine morning and afternoon milkings. From each milking, a single composite milk sample was drawn and mixed, and three technical replicate NIR scans were recorded. This resulted in six spectra per cow per day (3 replicates × 2 milkings), of which poor-quality scans were removed during subsequent quality control. Sampling spanned a 21-day observation window within a two-month period aligned with individual estrous cycles. Data collection was conducted over approximately eight weeks, from early January 2024 to 28 February 2024. Each cow was monitored for a contiguous 21-day window within this period, and the last milk samples were obtained on 28 February 2024. At each milking, approximately 200 mL of milk was collected into dry, sterile containers, gently inverted several times to ensure homogeneity, and kept at 4–8 °C during transport to the measurement station. For each

acquisition (time point), the spectrometer recorded 32 raw scans that were averaged to a single spectrum; outlier checks were applied at the scan and time point levels. This yielded six averaged spectra per cow per day. Near-infrared (NIR) spectroscopy and a digital milk-composition analyzer were used to obtain milk spectra and routine composition parameters (fat, protein, lactose, solids-not-fat, and density), and milk yield per milking and sample pH were recorded as additional physiological indicators for exploratory analysis. Figure 1 summarizes the workflow of this study, from sampling and measurement through preprocessing and data alignment to subject-wise modeling for estrus prediction. All spectral and composition measurements were completed within 2 h after milking; prior to scanning, samples were equilibrated to minimize temperature-induced variability. All procedures were approved by the Institutional Animal Care and Use Committee of Khon Kaen University (IACUC-KKU-141/66; reference 660201.2.11/746 (180)).

3.1.1 Sample Size Justification

This investigation was designed as a proof-of-concept pilot while adhering to the 3R principles of animal experimentation (Replacement, Reduction, and Refinement). During protocol development for Institutional Animal Care and Use Committee (IACUC) approval (IACUC-KKU-141/66), a preliminary variance assessment based on farm records indicated that a cohort of five cows would be sufficient to demonstrate the technical feasibility of the NIR sensing and DSP/ML pipeline, enable comparison of alternative preprocessing and classification algorithms, and characterize within-cow changes across a single estrous cycle while minimizing the number of animals used. Similar exploratory estrus-detection studies using milk spectroscopy have employed comparably small cohorts (Takemura et al., 2015 reported N = 6), supporting the use of small-N designs in early-stage work. In this context, N = 5 was judged appropriate for a pilot study and was approved by the ethics committee as the minimum number that could still provide informative internal cross-validation. Despite these precedents, a cohort of N = 5 cows remains statistically underpowered for estimating population-level effect sizes or clinically relevant sensitivity and specificity. We therefore treat this study purely as a technical proof-of-concept and not as a definitive evaluation of diagnostic performance.

3.1.2 Estrus Confirmation and Labeling

Estrus status for each cow was determined by daily monitoring of behavioral signs by an experienced herd manager and veterinarian. Cows were observed at least twice per day for classical estrous behaviors, including restlessness, mounting or standing to be mounted, increased locomotor activity, and clear vulval mucous discharge. When these behavioral signs suggested imminent or ongoing estrus, a confirmatory transrectal ultrasound examination of the ovaries was performed that same morning immediately after milking. Because of cost and handling considerations, ultrasound was carried out only on days with suspected estrus and was not performed routinely on all days. An estrus event was defined as a day on which both behavioral signs and ultrasound findings were consistent with pre-ovulatory follicular activity. For labeling, milk samples collected at the morning and afternoon milkings on the confirmed estrus day (i.e., within approximately 12 h of the ultrasound examination) were assigned to the "estrus" class, whereas all other samples in the 21-day observation window were labeled "non-estrus."

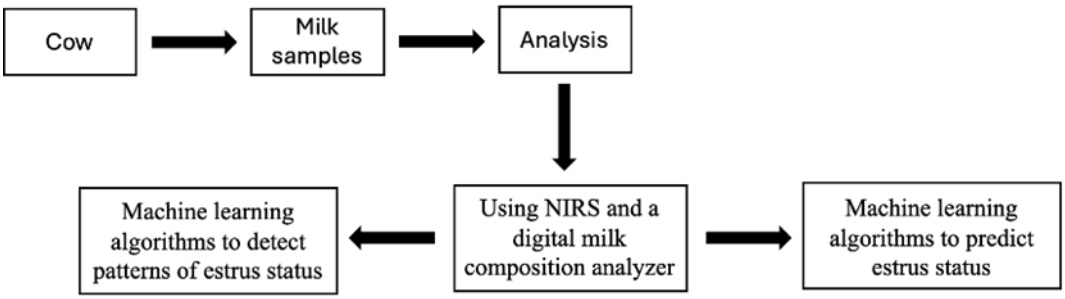

Figure 1 Block diagram of the data pipeline from sampling to estrus prediction

3.2 Spectral Acquisitions

Spectral data were collected with a fiber-coupled near-infrared (NIR) spectrometer operating in the 800–1700 nm region and equipped with an indium–gallium–arsenide (InGaAs) detector array. Illumination was provided by a stabilized 25 W tungsten–halogen source coupled through bifurcated fiber optics to a fixed-geometry sampling probe. To standardize the optical path and minimize operator variability, the probe was positioned at the bottom of the sample container using a 3D-printed standoff that fixed both working distance and incidence angle. Reflected light from the milk returned through the collection fiber to the spectrometer, which digitized the signal at 16-bit resolution. The system was warmed for 30 min prior to each session to stabilize lamp output and detector response, and samples were equilibrated before scanning to reduce temperature-induced spectral variability.

At each milking time point, the spectrometer used an integration time of 30 ms. For each physical sample, three technical scans were acquired without moving the cuvette between scans, screened for outliers (saturation or abnormally low counts), or obvious artifacts such as spikes due to air bubbles or transient probe misplacement, and then averaged to one spectrum to improve signal-to-noise while limiting exposure time. Scans that failed these quality-control criteria were immediately remeasured or discarded. All remaining valid scans were retained as separate spectra (technical replicates) for subsequent analysis, yielding two to three accepted spectra per milking after quality control. All acquisition settings (integration time, number of scans, QC thresholds, and referencing schedule) were held constant for all cows and days to ensure comparability across the dataset. The probe/container assembly was enclosed by an opaque fabric shroud with a simple light baffle to suppress stray light.

White and dark reference scans were recorded at the start of each session and whenever the ambient temperature changed by more than 2 °C. These white (PTFE) and dark (EVA) measurements were used to calibrate the spectrometer response before acquiring any milk spectra in that session, and an additional calibration cycle was performed approximately every 30 min during milking or whenever ambient conditions changed noticeably. A 20 mm polytetrafluoroethylene (PTFE) plate served as the white reference, and a 5 mm matte-black ethylene-vinyl-acetate (EVA) sheet served as the dark reference. Relative reflectance spectra were computed as

$$R(\lambda) = \frac{I(\lambda) - D(\lambda)}{W(\lambda) - D(\lambda)} \quad (1)$$

where I, W, and D denote the sample, white, and dark signals, respectively. Quality Control (QC) checks were applied to each scan; spectra with more than 0.5 % saturated pixels or abnormally low total counts were flagged and reacquired. The acquisition software logged lamp runtime, ambient temperature, and reference timestamps to support maintenance scheduling and drift analysis. The resulting NIR reflectance spectra provide indirect information on milk constituents such as fat, protein, lactose, and water through overtone and combination bands in the 800–1700 nm region (Baeten & Dardenne, 2021). These spectra, together with routine milk-composition parameters, formed the input to the digital signal-processing and machine-learning (DSP/ML) pipeline used to classify estrus and non-estrus stages, with the subsequent digital preprocessing steps described in Section 3.4.

3.3 Milk Composition Measurement

A digital milk analyzer (LactoStar 3510, Funke-Gerber, Germany) was used immediately after each milking to quantify routine composition parameters. The instrument reported fat, protein, lactose, solids-not-fat (SNF), and density in its native units. The analyzer was warmed up according to the manufacturer's procedure, then zeroed and verified at the start of each session. Before measurement, each sample was gently inverted several times to ensure homogeneity and brought within the admissible inlet-temperature range. In parallel with composition analysis, milk pH was measured on an aliquot of the same sample using a calibrated handheld digital pH meter, which was checked daily with pH 4.0 and 7.0 buffer solutions. Milk yield at each milking was recorded and used to compute daily yield as the sum of morning and afternoon milkings. Immediately after milking, samples were carried in insulated containers to the measurement area adjacent to the parlor. NIR scanning and composition analysis were initiated as soon as mixing was completed and were typically finished within 10 minutes of sample collection; all samples were processed within 30 minutes.

Each sample was measured at least three times. Triplicates were acquired back-to-back without removing the sample from the inlet. If the maximum relative difference among the triplicate readings exceeded 2%, one or two additional replicates were taken, and the median of the three most consistent

readings was recorded. Between samples, the fluid path was rinsed and cleaned following the manufacturer's sequence.

All composition values were time-stamped and linked to the corresponding NIR spectra via a common sample identifier so that both data sources referenced the same physical sample and time window. Milk pH and milk yield were recorded for the same milking events and associated with the corresponding composition–spectral records. Composition measurements were scheduled immediately before or after NIR acquisition for that time point to minimize temporal mismatch. Composition variables were later combined with spectral features in the digital signal-processing and machine-learning pipeline, whereas pH and milk yield were analyzed as additional exploratory indicators of estrus-related physiological change.

3.4 Spectral Data Pretreatment

Raw spectra were first converted to relative reflectance using the session white and dark references described in Section 3.2. Spectral end regions with a low signal-to-noise ratio were trimmed, and scans with saturation or abnormally low total counts were excluded by a quality-control screen. Where indicated, reflectance was transformed to absorbance to improve linearity with concentration:

$$A(\lambda) = -\log_{10} R(\lambda) \tag{2}$$

baseline variation was removed using asymmetric least-squares (AsLS) regression (Aernouts et al., 2015; Ramirez-Morales et al., 2021; Caponigro et al., 2023). Given a spectrum y and a second-order difference matrix D, the baseline z was estimated by solving:

$$\min_{z} \sum_i W_i (y_i - z_i)^2 + \lambda \|Dz\|_2^2 \tag{3}$$

where $w_i \geq 0$ are adaptive weights updated iteratively to suppress negative residuals, and λ is the smoothness parameter. Zero-phase Savitzky–Golay filtering was then applied for gentle denoising; the window length and polynomial order were chosen on a small development subset and held fixed for all subsequent analyses (Aernouts et al., 2015; Ramirez-Morales et al., 2021; Caponigro et al., 2023).

To correct multiplicative scatter and path-length effects, the standard normal variate (SNV) was applied to each spectrum x:

$$x' = \frac{[x - \mu(x)]}{\sigma(x)} \tag{4}$$

where $\mu(x)$ and $\sigma(x)$ are the mean and standard deviation across wavelengths (Aernouts et al., 2011; Yang et al., 2022; Lanjewar et al., 2024). First- and second-order derivatives were computed using the Savitzky–Golay operator to enhance small but consistent features related to milk composition (Aernouts et al., 2015; Ramirez-Morales et al., 2021; Caponigro et al., 2023). After trimming and preprocessing, each sample was represented by NIR absorbance values at 224 wavelength points between 860.6 and 1754.1 nm. These spectral features were then concatenated with seven scalar milk-related variables (fat, protein, lactose, solids-not-fat, density, pH, and daily milk yield), yielding a 231-dimensional feature vector for each observation. Feature vectors were finally mean-centered and scaled to unit variance.

All pretreatment parameters, including trimming limits, AsLS hyperparameters, filter settings, and scaling statistics, were estimated exclusively on the training folds within each cross-validation split and then applied unchanged to validation/test data to prevent information leakage. The entire workflow was implemented in Python (NumPy, SciPy, and scikit-learn) within a Google Colab environment, using a fixed random seed of 42 for reproducibility.

3.5 Data Analysis

3.5.1 Unsupervised Visualization

Principal component analysis (PCA) was applied to a feature matrix comprising the preprocessed NIR spectra and routine milk-composition parameters. PCA reduces dimensionality by projecting data onto directions of maximal variance, aiding structure discovery and visualization (Yeung & Ruzzo, 2001; Karamizadeh et al., 2013; Wu et al., 2017). Before PCA, all features were mean-centered and scaled to unit variance within the model-training loop to avoid leakage. To prevent information leakage, PCA was fit only on the training portion within each cross-validation split (GroupKFold with cow ID as the grouping variable, k = 5), and the learned transformation was applied unchanged to the held-out cow.

Two-dimensional and three-dimensional score plots were inspected for qualitative separation between predefined estrus and non-estrus groups. Axes reported the percent variance explained by each

component, and 95% confidence ellipses were overlaid for visualization only. PCA was used strictly for exploratory visualization and interpretation; group labels informed only the plot coloring, were not used to fit the PCA model, and no clustering algorithm was used to define groups. Interpretation emphasized composition-related structure rather than definitive class separation, consistent with the small cohort size.

3.5.2 Classification

Six supervised classifiers were evaluated to predict estrus status from a feature set comprising preprocessed NIR spectra and routine milk-composition parameters: logistic regression, decision tree, random forest, Extra Trees, k-nearest neighbors (k-NN), and gradient boosting. To limit overfitting with small *N*, we prioritized regularized, low-variance models (logistic regression) and shallow tree ensembles (Extra Trees, random forest) as primary models; k-NN, decision tree, and gradient boosting were retained as sensitivity analyses. All features were mean-centered and scaled to unit variance within each training fold; the fitted scaler was then applied to the corresponding validation fold. Although tree-based models are relatively scale-invariant, a common scaler ensured a consistent pipeline across models.

To address class imbalance, weights and decision-threshold tuning were used by default. The Synthetic Minority Over-sampling Technique (SMOTE) was reserved for sensitivity checks and, when used, was applied only within training folds; no synthetic samples were added to validation/test data. Where supported, class weights were enabled to further reduce bias toward the majority class. Primary analyses are therefore based on models trained without SMOTE, with SMOTE-treated models reported only as secondary sensitivity analyses.

Model selection and evaluation used stratified group k-fold cross-validation with groups defined by cow identification to prevent subject leakage. Within each fold, the training partition was processed in the following order: scaling, SMOTE, and model fitting. Hyperparameters were tuned by grid search on the training partition; the fitted preprocessing and model were then applied unchanged to the held-out partition. Reported scores are the mean and standard deviation across folds.

Performance was summarized from the confusion-matrix counts of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) (Recio et al., 2013; Phuphaphud et al., 2022). Accuracy, precision, recall, and F1 score were computed as follows (Sokolova & Lapalme, 2009; Wang et al., 2022a):

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{5}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{6}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{7}$$

$$\text{F1 score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8}$$

confusion matrices and precision–recall curves were examined with estrus treated as the positive class to characterize class-specific behavior.

4. Results

4.1 NIR Spectra of Milk Samples

Figure 2 shows representative unpretreated absorbance spectra of milk acquired in the 800–1700 nm region (integration time 30 ms). Across samples, the overall shape was consistent and reflects the major milk constituents. Prominent water bands appear near 970 nm and 1450 nm, commonly assigned to O–H overtone and combination vibrations (Workman Jr, 1996; Golic et al., 2003; Weyer & Lo, 2006). Additional features attributed to C–H overtones in lipids and proteins are visible around 1150–1250 nm, with a shoulder in the 1650–1690 nm region (Kaddour et al., 2006; Muncan et al., 2021). Wavelengths beyond 1700 nm lie outside the instrument's calibrated range and were not used for analysis; the plotting window in Figure 2 is extended to 1760 nm for completeness.

Progesterone does not have a distinct, isolated absorption band in this near-infrared window; any hormonal influence on the spectra is expected to be indirect via correlated changes in milk composition, particularly fat and protein. Consequently, the subsequent analysis relies on multivariate patterns across the spectra rather than a single diagnostic peak, consistent with prior NIR-based estrus studies (Vance et al., 2016; Wenz, 2021; Chen et al., 2023). The spectra in Figure 2 were subsequently pretreated as described in section 3.4 to improve the signal-to-noise ratio and stabilize path-length and scatter effects prior to visualization and classification.

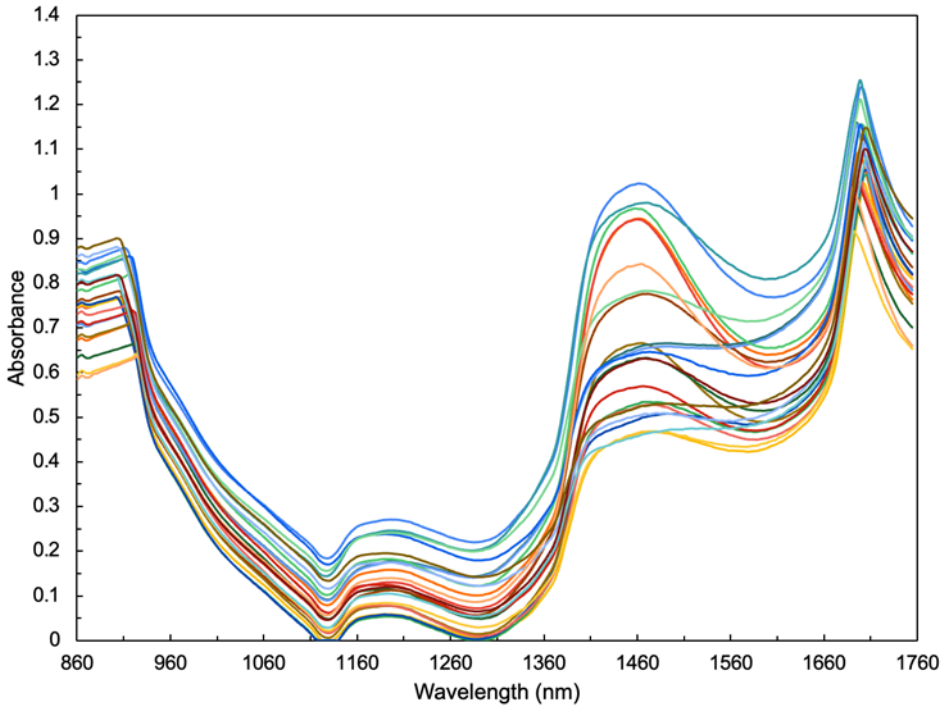

Figure 2 Raw absorbance spectra of milk samples measured using near-infrared spectroscopy

4.2 Milk Composition Analysis

Per-cow mean milk composition, daily milk yield, and pH during estrus and non-estrus are summarized in Figures 3 and 4. Each panel shows, for every cow, the mean value of fat, SNF, protein, lactose, density, or daily milk yield across all sampling days within each physiological state; vertical lines denote the within-cow standard deviation. Values for fat, SNF, protein, lactose, and density were obtained from the LactoStar 3510 analyzer, and daily milk yield was calculated from the recorded morning and afternoon volumes. For statistical comparison, morning and afternoon measurements were averaged within estrus and non-estrus days for each cow. Within-cow differences were then computed as non-estrus minus estrus and evaluated across the five cows using paired *t*-tests, with Wilcoxon signed-rank tests as a non-parametric sensitivity analysis. The resulting mean differences, standard deviations, and p-values are reported in Table 1. Because these paired comparisons are based on only five cows, the resulting p-values should be interpreted with caution; they provide a rough indication of the direction and consistency of within-cow changes rather than precise estimates of population-level statistical significance. Across cows, fat, SNF, protein, and lactose were all slightly higher during estrus than during non-estrus, corresponding to negative mean differences (non-estrus − estrus) of −0.34, −0.88, −0.29, and −0.41 percentage points, respectively. The increases in protein and lactose reached conventional significance (paired *t*-test $p = 0.016$ for both), whereas fat and SNF showed similar upward trends that did not reach significance in this small cohort ($p \approx 0.07$). Density and pH did not differ systematically between estrus and non-estrus (paired *t*-test $p > 0.25$ and $p \approx 0.96$, respectively), suggesting that these variables contribute little discriminatory power on their own. In contrast, milk yield was consistently lower during estrus for all cows, with an average reduction of 7.37 kg/day relative to non-estrus (paired *t*-test $p = 0.0038$). This pattern is evident in the milk-yield panel of Figure 3, where estrus bars are uniformly lower than non-estrus bars despite overlapping error bars for individual animals. These composition and yield shifts are modest in absolute terms and should be interpreted as exploratory given the small sample size, but they are consistent with an estrus-associated redistribution of milk constituents and reduced output. In the subsequent classification analyses, fat, SNF, protein, lactose, density, and milk-yield parameters were used as auxiliary features alongside NIR spectral features, whereas pH was excluded because it showed no systematic difference between stages.

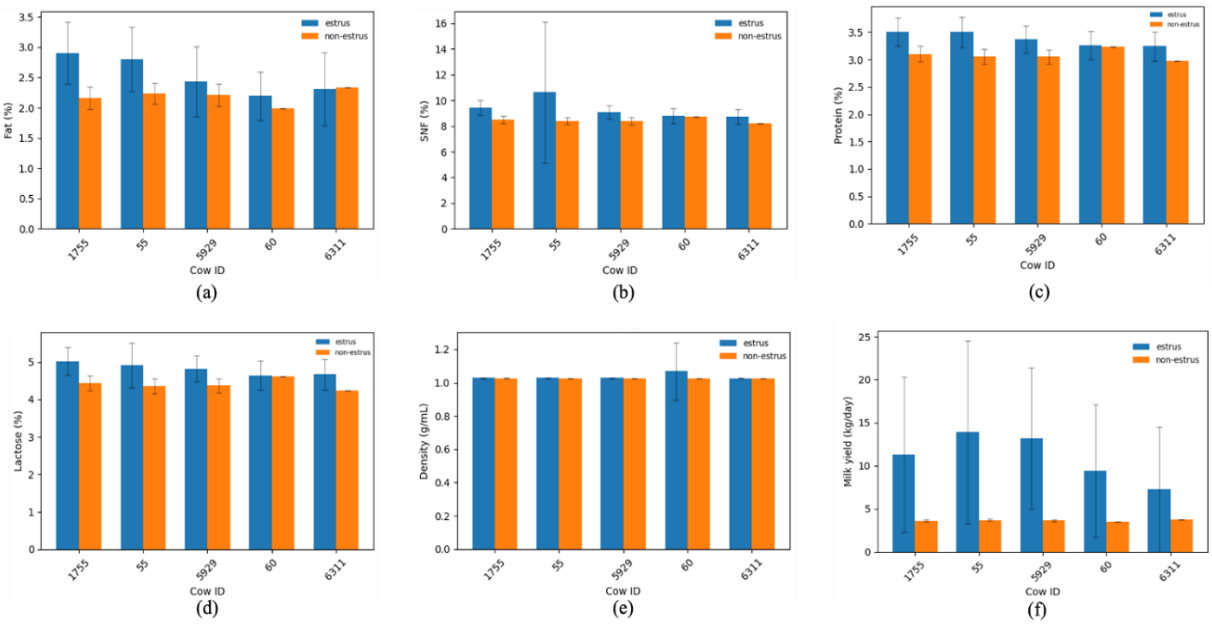

Figure 3 Per-cow mean milk composition and yield during estrus and non-estrus: (a) fat (%), (b) solids-not-fat (SNF, %), (c) protein (%), (d) lactose (%), (e) density (g/mL), and (f) milk yield (kg/day)

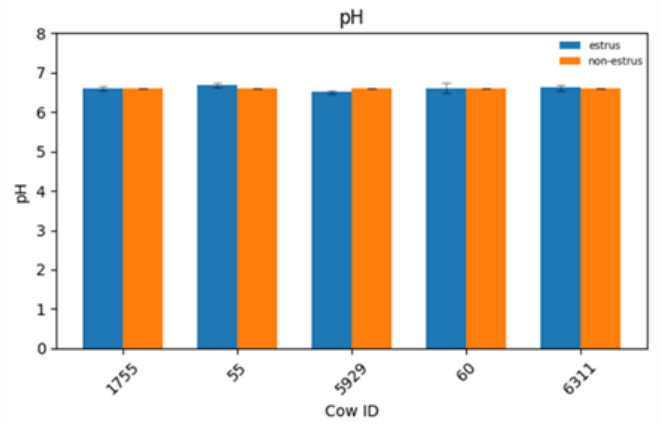

Figure 4 Per-cow mean milk pH during estrus and non-estrus days.

Table 1 Within-cow differences in milk composition, milk yield, and pH between non-estrus and estrus days

Variable	Mean difference (non-estrus - estrus)	SD of difference	Paired *t*-test (*p*)	Wilcoxon (*p*)
Fat (%)	–0.34	0.31	0.071	0.125
SNF (%)	–0.88	0.82	0.076	0.063
Protein (%)	–0.29	0.16	0.016	0.063
Lactose (%)	–0.41	0.23	0.016	0.063
Density (g/mL)	–0.01	0.02	0.260	0.063
pH	0.00	0.06	0.959	1.000
Milk yield (kg/day)	–7.37	2.72	0.0038	0.063

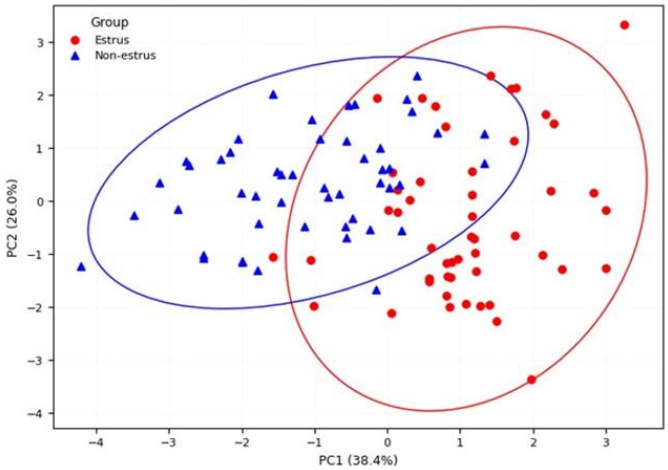

Figure 5 PCA score plot of preprocessed near-infrared spectra concatenated with milk-composition features

4.3 PCA Based Visualization of Estrus Stages

PCA was applied to the preprocessed NIR spectra concatenated with milk-composition parameters to reduce dimensionality and visualize dominant patterns in the feature space. All variables were mean-centered and scaled to unit variance and the PCA model was always fit on the training folds and then applied unchanged to the held-out cow to avoid any information leakage. PCA was used strictly for visualization and interpretation; group labels were not used to fit the model.

Figure 5 shows the PC1–PC2 score plot obtained by pooling predictions across all Leave-One-Cow-Out folds. PC1 explains 38.4% of the variance and PC2 explains 26.0%, so the first two components together capture 64.4% of the total variance. Samples labeled estrus tend to have higher PC1 scores, whereas non-estrus samples are concentrated at lower PC1 scores. For visualization in Figure 5, 95% confidence ellipses were drawn around the group centroids in the PC1–PC2 space, assuming a multivariate normal distribution of scores and, using the sample covariance within each group. Given the small and imbalanced dataset, these ellipses are intended only as qualitative guides to cluster separation rather than as formal inferential intervals. The ellipses indicate a centroid shift primarily along PC1 with partial overlap between groups. Dispersion along PC2 is comparable across groups and reflects within-group variability rather than clear stage separation. This pattern suggests that estrus is associated with a modest, largely composition-related shift in the multivariate feature space, but that unsupervised visualization alone cannot provide a reliable decision boundary because substantial overlap remains, especially given the small proof-of-concept cohort (N = 5 cows). Accordingly, we treat PCA as an exploratory tool only and rely on supervised classifiers under subject-wise cross-validation for quantitative performance assessment rather than any single spectral or compositional threshold.

4.4 Estrus Prediction using Machine-Learning Models

Six supervised classifiers were evaluated using leave-one-cow-out cross-validation (LOCO-CV) to assess how well the combined NIR-plus-milk-composition feature set discriminated estrus from non-estrus. The final dataset comprised 593 spectral–composition feature vectors from five cows, of which 31 were labeled estrus and 562 non-estrus, corresponding to an approximate 1:18 class imbalance. To mitigate this imbalance during training, we used class weighting for logistic regression and all tree-based models so that misclassifying an estrus sample incurred a higher loss than misclassifying a non-estrus sample. For each LOCO-CV fold, all preprocessing (standardization and, in the sensitivity runs, SMOTE oversampling) was fit only on the four training cows and then applied unchanged to the held-out cow to avoid information leakage, while the natural class imbalance in the held-out cow was preserved. SMOTE was reserved for a secondary analysis.

Each classifier operated on the full 231-dimensional feature space (224 NIR wavelengths plus seven scalar milk-composition variables, as described in Section 3.4), and no prior feature selection was applied, so all models were trained on the same information. Table 2 summarizes the LOCO-CV performance on the original, imbalanced dataset (no

SMOTE). Logistic regression, decision tree, gradient boosting, and k-nearest neighbors all achieved apparent perfect performance within this dataset (mean accuracy, precision, recall, and F1 = 1.00, SD = 0.00). The tree-ensemble models also performed strongly. Extra Trees reached a mean accuracy of 0.998 and F1 of 0.985, with estrus-class recall around 0.97 (SD ≈ 0.06). Random forest obtained a mean accuracy of 0.988 and F1 of 0.918, but with lower mean precision (≈ 0.88, SD ≈ 0.20), indicating a small number of false positive estrus predictions relative to the other models.

Because only five cows were available, we further examined per-cow metrics for three representative models (logistic regression, random forest, and Extra Trees). After quality control, the number of samples per cow used in the LOCO-CV evaluation ranged from 115 to 121 feature vectors (121 each for cows 55, 1755, and 5929, and 115 for cows 60 and 6311), so the per-cow performance patterns are not driven by extreme differences in data volume. The per-cow accuracies, precisions, recalls, and F1 scores are listed in Table 3 and visualized as F1 scores in Figure 6. Logistic regression classified all cows perfectly (F1 = 1.00 for every animal). Extra Trees also showed consistently high performance, with F1 ≥ 0.92 for all cows (the lowest value occurring for cow 60). Random forest performed similarly well for cows 5929, 1755, and 6311 (F1 = 1.00) and slightly lower for cow 55 (F1 ≈ 0.92), but it showed a clear drop for cow 60 (precision ≈ 0.55, F1 ≈ 0.67), reflecting a tendency to over-predict estrus for that individual. Thus, the per-cow breakdown reveals that random-forest errors are concentrated in one cow rather than uniformly distributed across the cohort.

Aggregated confusion matrices for these three models are shown in Figure 7. Logistic regression correctly classified all 593 samples (562 non-estrus and 31 estrus), yielding no false positives or false negatives. Extra Trees also achieved excellent discrimination, with all non-estrus samples correctly identified and only one estrus sample misclassified as non-estrus (562 true negatives, 30 true positives, 1 false negative, 0 false positives). In contrast, the random-forest confusion matrix shows six false positive estrus predictions and one false negative (556 true negatives, 30 true positives), consistent with its lower precision and reduced F1 for cow 60. Taken together, these near-perfect internal results demonstrate that estrus and non-estrus samples are highly separable within this small, single-herd cohort, but they should be interpreted as optimistic upper bounds on performance; substantially lower accuracy and F1 scores are anticipated once the method is applied to larger, multi-farm populations with greater biological and management variability.

For completeness, LOCO-CV was also repeated with SMOTE oversampling applied only to the training folds, and the resulting metrics are summarized in Table 4. As expected for such a small dataset, many SMOTE-based models achieved almost perfect scores (frequently 1.00 for accuracy, precision, recall, and F1), which we interpret as clear evidence of overfitting to synthetic patterns rather than genuine robustness. Consequently, the SMOTE runs are reported purely as a sensitivity check to illustrate how the pipeline behaves under artificially balanced classes, and our primary performance estimates for this proof-of-concept study are those obtained without SMOTE.

Table 2 Leave-one-cow-out cross-validated estrus-classification performance without SMOTE

Model	Accuracy	Precision	Recall	F1 score
Decision tree	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
Extra Trees	0.998 ± 0.004	1.000 ± 0.000	0.971 ± 0.064	0.985 ± 0.034
Gradient boosting	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
Logistic regression	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
Random forest	0.998 ± 0.023	0.881 ± 0.197	0.971 ± 0.064	0.918 ± 0.144
k-nearest neighbors	0.998 ± 0.004	1.000 ± 0.000	0.971 ± 0.064	0.985 ± 0.034

Table 3 Per-cow leave-one-cow-out F1 scores for selected classifiers (no SMOTE)

Cow ID	Logistic regression	Random forest	Extra Trees
55	1.000	0.923	1.000
60	1.000	0.667	0.923
1755	1.000	1.000	1.000
5929	1.000	1.000	1.000
6311	1.000	1.000	1.000

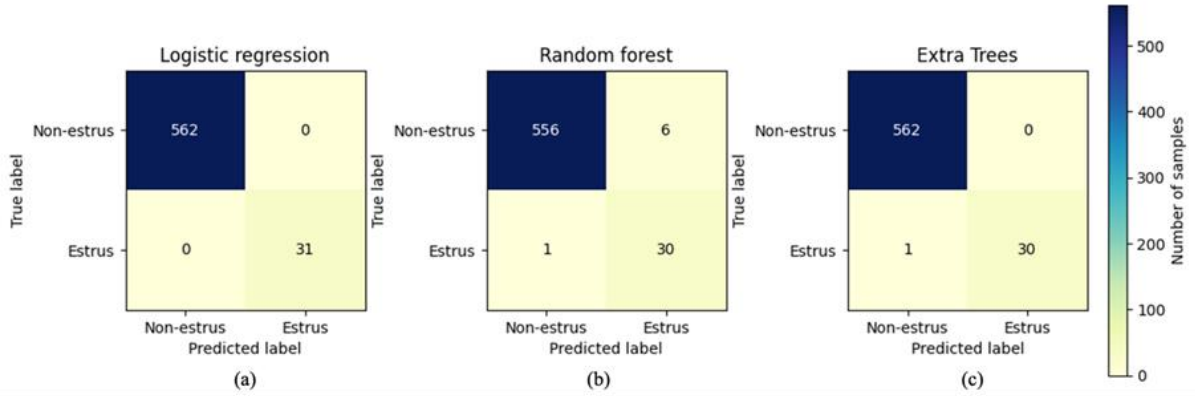

Figure 6 Per-cow F1-scores under leave-one-cow-out cross-validation for three representative classifiers: logistic regression, random forest, and Extra Trees

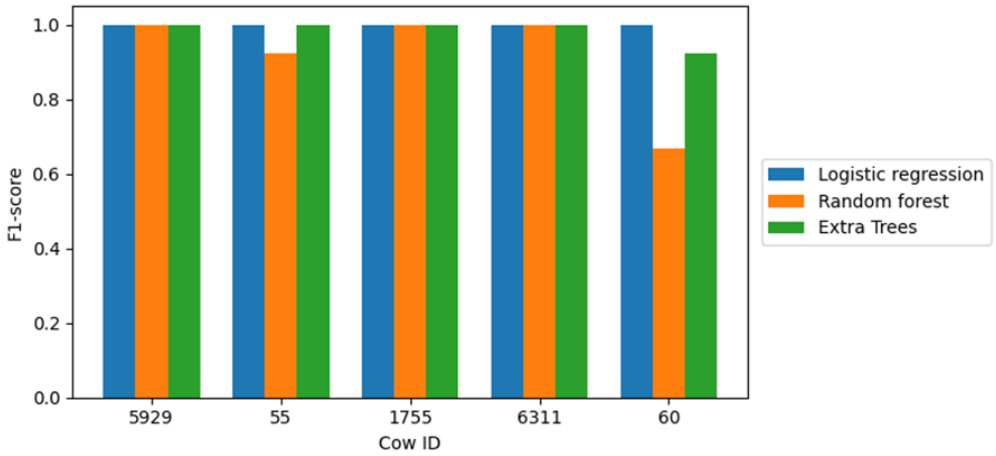

Figure 7 Aggregated 2×2 confusion matrices under leave-one-cow-out cross-validation for (a) logistic regression, (b) random forest, and (c) Extra Trees, with non-estrus as the negative class and estrus as the positive class

Table 4 Leave-one-cow-out cross-validated estrus-classification performance with SMOTE applied to training folds only

Model	Accuracy	Precision	Recall	F1 score
Decision tree	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
Extra Trees	0.998 ± 0.004	1.000 ± 0.000	0.971 ± 0.064	0.985 ± 0.034
Gradient boosting	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
Logistic regression	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
Random forest	0.986 ± 0.023	0.859 ± 0.197	0.971 ± 0.064	0.905 ± 0.144
k-nearest neighbors	0.998 ± 0.004	0.975 ± 0.000	1.000 ± 0.064	0.987 ± 0.034

5. Discussion

This proof-of-concept indicates that NIR milk spectra, when combined with routine milk-composition data and supervised learning, can distinguish estrus from non-estrus under controlled, single-farm conditions. Across the five cows, estrus days were characterized by small but consistent within-cow shifts: protein and lactose increased by about 0.3–0.4 percentage points, while daily milk yield decreased by roughly 7 kg/cow. These trends are biologically plausible, as hormonal changes around estrus can alter feed intake and energy partitioning, leading to transient reductions in yield and subtle changes in osmotic balance and casein synthesis that manifest as slightly higher protein and lactose concentrations even when absolute output is lower. Density and pH showed no systematic change between stages, suggesting limited discriminatory power on their own.

On this combined spectral–compositional feature set, several classifiers achieved very strong internal performance under LOCO-CV. Without SMOTE, logistic regression, decision tree, gradient boosting, and k-nearest neighbors all reached apparent perfect accuracy, precision, recall, and F1 within this dataset, while the Extra Trees ensemble misclassified only one estrus sample. Per-cow analysis confirmed that logistic regression and Extra Trees performed consistently well across animals, whereas random forest showed concentrated errors in a single cow. Aggregated confusion matrices show very few false positives and false negatives overall. At the same time, the fact that several distinct model families achieved near-perfect LOCO-CV metrics on only five cows strongly suggests that these results represent optimistic upper bounds rather than realistic field performance, and they should be interpreted as hypothesis-generating evidence of technical feasibility rather than definitive validation.

Compared with earlier work, the internal accuracies reported here are encouraging but must be viewed in context. Takemura et al. (2015), using NIR spectroscopy on raw milk from six cows, reported about 89% accuracy for estrus detection. Our leakage-aware LOCO-CV framework, which integrates both spectral and composition features and uses modern ensemble methods, yields internal accuracies around 98–100%, roughly a 10 percentage-point improvement. However, both studies rely on very small cohorts, and differences in herd, management, and data-processing pipelines mean that this apparent gain may be dataset-specific or partly driven by overfitting. Larger, multi-site studies with consistent validation protocols are needed before any robust comparison of absolute performance can be made.

Feature-importance patterns from the tree-based models provide some mechanistic insight. The highest-ranked predictors were NIR bands associated with water and protein absorption, together with milk yield, protein, and lactose, which is consistent with the univariate composition results. In contrast, density and pH had negligible importance in these models and showed no meaningful differences between estrus and non-estrus in Table 1. We nevertheless retained them in the feature set to maintain alignment with the full LactoStar output and to allow future, larger studies to confirm whether they remain uninformative; if their low importance is replicated, they can be safely omitted from simplified models.

Several limitations of the present study are therefore critical. The most severe are the small sample size (N = 5 cows) and the single-farm setting, which prevent reliable estimation of population-level performance and limit generalizability across breeds, environments, and management systems. All evaluations are internal, based on leave-one-cow-out cross-validation, with no independent external test set or multi-farm replication. Estrus labels were derived from expert visual observation supplemented by targeted ultrasound rather than a biochemical or device-based gold standard such as progesterone assays or automated activity monitors. The observation period covered only one 21-day cycle per cow, so seasonal and longer-term effects could not be assessed. Finally, the combination of high feature dimensionality and strong internal metrics, especially under SMOTE, indicates a real risk of overfitting; the SMOTE-based results in particular should be regarded as illustrative of pipeline behavior under artificial class balance rather than as deployable performance estimates.

Despite these constraints, the study offers several contributions that justify publication as a pilot. It establishes a practical on-parlour protocol for NIR milk acquisition and integrated composition measurement, coupled with a leakage-aware LOCO-CV framework that can be reused in larger trials. It shows that, in principle, a realistic electro-optical chain combined with routinely available milk-recording data can support estrus detection during normal milking. The observed effect sizes in composition, yield, and model performance provide starting points for power calculations and study design in future work. The logical next step is a multi-site validation study involving at least 50 cows across multiple farms over 6–12 months, with blinded comparison to progesterone assays and validated activity-monitoring systems, explicit external test sets and calibration-transfer procedures, and a cost–benefit analysis relative to existing estrus-detection methods. Addressing these points will be essential to move from the current proof-of-concept to a robust, generalizable tool for dairy herd reproductive management.

6. Conclusion

This study presents a proof-of-concept implementation of an on-farm NIR milk-sensing system combined with routine milk-composition measurements and a leakage-aware machine-learning pipeline for estrus detection in dairy cows. Within a single Thai herd of five cows, modest but consistent shifts in milk composition, including higher protein

and lactose levels and reduced daily milk yield were observed during estrus, and several classifiers achieved near-perfect leave-one-cow-out performance when supplied with both NIR and composition features. At the same time, the very small, single-site cohort, reliance on expert observation with targeted ultrasound rather than a biochemical gold standard, and the absence of any external validation mean that these accuracy estimates are almost certainly optimistic and cannot be generalized to wider populations. SMOTE-based analyses were used strictly as exploratory sensitivity checks and are likely to overestimate real-world performance. Overall, the present work should be interpreted as demonstrating technical feasibility and providing a methodological template, rather than as evidence that the system is ready for clinical or commercial use. Future multi-farm studies with substantially larger cow numbers (on the order of more than 50 animals), progesterone-confirmed or device-based reference labels, explicit external test sets, and calibration-transfer procedures will be essential to obtain realistic performance estimates and to determine whether NIR-based estrus detection can offer a robust and cost-effective complement to existing monitoring methods.

7. Acknowledgements

This work was supported by the Young Researcher Development Project, Khon Kaen University (2023). We thank the Faculty of Agriculture, Khon Kaen University, for assistance with sample collection and estrus verification.

8. Abbreviations

Abbreviation	Full Term
3R	Replacement, Reduction, and Refinement
AsLS	Asymmetric Least-Squares
DSP/ML	Digital Signal-Processing and Machine-Learning
VA	Ethylene-Vinyl-Acetate (Dark reference material)
F1	F1-score (harmonic mean of precision and recall)
FN	False Negatives
FP	False Positives
IACUC	Institutional Animal Care and Use Committee
InGaAs	Indium–Gallium–Arsenide
k-NN	k-Nearest Neighbors
LOCO-CV	Leave-One-Cow-Out Cross-Validation
ML	Machine Learning
NIR	Near-Infrared
NIRS	Near-Infrared Spectroscopy
PC1 / PC2	Principal Component 1 / Principal Component 2
PCA	Principal Component Analysis
PTFE	Polytetrafluoroethylene (White reference material)
QC	Quality-Control
SMOTE	Synthetic Minority Over-sampling Technique
SNF	Solids-Not-Fat
SNV	Standard Normal Variate
TN	True Negatives
TP	True Positives

9. CRediT Statement:

Norrawit Tonmitr: Conceptualization, Methodology, Software, Validation, Visualization, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Funding Acquisition
Arthit Phuphaphud: Conceptualization, Methodology, Formal Analysis, Investigation, Resources, Writing – Original Draft
Chanon Suntara: Methodology, Validation, Investigation, Resources, Writing – Review & Editing, Supervision
Sora-at Tanusilp: Validation, Resources, Data Curation, Visualization
Panawit Hanpinitsak: Software, Validation, Formal Analysis
Tatpong Katanyukul: Software, Validation, Formal Analysis, Supervision

10. References

Aernouts, B., Polshin, E., Saeys, W., & Lammertyn, J. (2011). Mid-infrared spectrometry of milk for dairy metabolomics: A comparison of two sampling techniques and effect of homogenization. *Analytica Chimica Acta*, *705*(1-2), 88-97. https://doi.org/10.1016/j.aca.2011.04.018

Aernouts, B., Van Beers, R., Watté, R., Huybrechts, T., Lammertyn, J., & Saeys, W. (2015).

Visible and near-infrared bulk optical properties of raw milk. *Journal of Dairy Science*, *98*(10), 6727-6738. https://doi.org/10.3168/jds.2015-9630

Agutu, F. O., Mbuku, S. M., Ondiek, J. O., & Bebe, B. O. (2024). Economic viability of using OvSynch and fixed timed artificial insemination protocol in breeding improvement of pastoral herds in the rangelands. *Tropical Animal Health and Production*, *56*(2), Article 68. https://doi.org/10.1007/s11250-024-03907-1

Arıkan, İ., Ayav, T., Seçkin, A. Ç., & Soygazi, F. (2023). Estrus detection and dairy cow identification with cascade deep learning for augmented reality-ready livestock farming. *Sensors*, *23*(24), Article 9795. https://doi.org/10.3390/s23249795

Baeten, V., & Dardenne, P. (2021). Application of NIR in agriculture. In Y. Ozaki, C. Huck, S. Tsuchikawa, & S. B. Engelsen (Eds.), *Near-infrared spectroscopy: Theory, spectral analysis, instrumentation, and applications* (pp. 331–345). Springer. https://doi.org/10.1007/978-981-15-8648-4_14

Behdad, S., Pakdel, A., & Massudi, R. (2024). A novel diagnostic approach to Paratuberculosis in dairy cattle using near-infrared spectroscopy and aquaphotomics. *Frontiers in Cellular and Infection Microbiology*, *14*, Article 1374560. https://doi.org/10.3389/fcimb.2024.1374560

Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. *Sensors*, *21*(11), Article 3758. https://doi.org/10.3390/s21113758

Caponigro, V., Marini, F., Scannell, A. G., & Gowen, A. A. (2023). Single-drop technique for lactose prediction in dry milk on metallic surfaces: Comparison of Raman, FT–NIR, and FT–MIR spectral imaging. *Food Control*, *144*, Article 109351. https://doi.org/10.1016/j.foodcont.2022.109351

Chen, Y., Wang, S., & Zhang, F. (2023). Near-infrared luminescence high-contrast in vivo biomedical imaging. *Nature Reviews Bioengineering*, *1*(1), 60-78. https://doi.org/10.1038/s44222-022-00002-8

da Silva Santos, A., de Medeiros, V. W. C., & Gonçalves, G. E. (2023). Monitoring and classification of cattle behavior: A survey. *Smart Agricultural Technology*, *3*, Article 100091. https://doi.org/10.1016/j.atech.2022.100091

Denis-Robichaud, J., Oliveira, A. P., Sica, A., Soriano, S., Araújo, R. L., Pereira, M. H. C., ... & Vasconcelos, J. L. M. (2024). Is prolonged luteal phase a problem in lactating Holstein cows?. *Journal of Dairy Science*, *107*(10), 8582-8591. https://doi.org/10.3168/jds.2024-24792

Džermeikaitė, K., Bačėninaitė, D., & Antanaitis, R. (2023). Innovations in cattle farming: application of innovative technologies and sensors in the diagnosis of diseases. *Animals*, *13*(5), Article 780. https://doi.org/10.3390/ani13050780

Evangelista, C., Basiricò, L., & Bernabucci, U. (2021). An overview on the use of near infrared spectroscopy (NIRS) on farms for the management of dairy cows. *Agriculture*, *11*(4), Article 296. https://doi.org/10.3390/agriculture11040296

Folman, Y., Kaim, M., Herz, Z., & Rosenberg, M. (1984). Reproductive management of dairy cattle based on synchronization of estrous cycles. *Journal of Dairy Science*, *67*(1), 153-160. https://doi.org/10.3168/jds.S0022-0302(84)81279-5

Golic, M., Walsh, K., & Lawson, P. (2003). Short-wavelength near-infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature. *Applied Spectroscopy*, *57*(2), 139-145. https://doi.org/10.1366/000370203321535033

Hempstalk, K., McParland, S., & Berry, D. P. (2015). Machine learning algorithms for the prediction of conception success to a given insemination in lactating dairy cows. *Journal of Dairy Science*, *98*(8), 5262-5273. https://doi.org/10.3168/jds.2014-8984

Ho, P. N., & Pryce, J. E. (2020). Predicting the likelihood of conception to first insemination of dairy cows using milk mid-infrared spectroscopy. *Journal of Dairy Science*, *103*(12), 11535-11544. https://doi.org/10.3168/jds.2020-18589

Ho, P. N., Bonfatti, V., Luke, T. D. W., & Pryce, J. E. (2019). Classifying the fertility of dairy cows using milk mid-infrared spectroscopy. *Journal of Dairy Science*, *102*(11), 10460-10470. https://doi.org/10.3168/jds.2019-16412

Holroyd, S. E. (2013). The use of near infrared spectroscopy on milk and milk products. *Journal of Near Infrared Spectroscopy*, *21*(5), 311-322. https://doi.org/10.1255/jnirs.1055

Islam, R. (2011). Synchronization of estrus in cattle: a review. *Veterinary World*, *4*(3), 136-141. https://doi.org/10.5455/vetworld.2011.136-141

Kaddour, A. A., Grand, E., Barouh, N., Baréa, B., Villeneuve, P., & Cuq, B. (2006). Near-infrared spectroscopy for the determination of lipid oxidation in cereal food products. *European Journal of Lipid Science and Technology*, *108*(12), 1037-1046. https://doi.org/10.1002/ejlt.200600132

Karamizadeh, S., Abdullah, S. M., Manaf, A. A., Zamani, M., & Hooman, A. (2013). An overview of principal component analysis. *Journal of Signal and Information Processing*, *4*(3B), 173-175. https://doi.org/10.4236/jsip.2013.43B031

Kawamura, S., Kawasaki, M., Nakatsuji, H., & Natsuga, M. (2007). Near-infrared spectroscopic sensing system for online monitoring of milk quality during milking. *Sensing and Instrumentation for Food Quality and Safety*, *1*(1), 37-43. https://doi.org/10.1007/s11694-006-9001-x

Kleanthous, N., Hussain, A. J., Khan, W., Sneddon, J., Al-Shamma'a, A., & Liatsis, P. (2022). A survey of machine learning approaches in animal behaviour. *Neurocomputing*, *491*, 442-463. https://doi.org/10.1016/j.neucom.2021.10.126

Lanjewar, M. G., Parab, J. S., & Kamat, R. K. (2024). Machine learning based technique to predict the water adulterant in milk using portable near infrared spectroscopy. *Journal of Food Composition and Analysis*, *131*, Article 106270. https://doi.org/10.1016/j.jfca.2024.106270

Le, A. T., Shakiba, M., Ardekani, I., & Abdulla, W. H. (2024). Optimizing plant disease classification with hybrid convolutional neural network–recurrent neural network and liquid time-constant network. *Applied Sciences*, *14*(19), Article 9118. https://doi.org/10.3390/app14199118

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, *18*(8), Article 2674. https://doi.org/10.3390/s18082674

Mahendran, S. A., Wathes, D. C., Booth, R. E., & Blackie, N. (2023). Effects of the individual and pair housing of calves on long-term heifer production on a UK commercial dairy farm. *Animals*, *14*(1), Article 125. https://doi.org/10.3390/ani14010125

McDougall, S. (2006). Reproduction performance and management of dairy cattle. *Journal of Reproduction and Development*, *52*(1), 185-194. https://doi.org/10.1262/jrd.17091

Muncan, J., Tei, K., & Tsenkova, R. (2021). Real-time monitoring of yogurt fermentation process by aquaphotomics near-infrared spectroscopy. *Sensors*, *21*(1), Article 177. https://doi.org/10.3390/s21010177

Paisley, L. G., Mickelsen, W. D., & Frost, O. L. (1978). A survey of the incidence of prenatal mortality in cattle following pregnancy diagnosis by rectal palpation. *Theriogenology*, *9*(6), 481-491. https://doi.org/10.1016/0093-691X(78)90113-9

Palmer, M. A., Olmos, G., Boyle, L. A., & Mee, J. F. (2010). Estrus detection and estrus characteristics in housed and pastured Holstein–Friesian cows. *Theriogenology*, *74*(2), 255-264. https://doi.org/10.1016/j.theriogenology.2010.02.009

Petit, H. V., Dewhurst, R. J., Proulx, J. G., Khalid, M., Haresign, W., & Twagiramungu, H. (2001). Milk production, milk composition, and reproductive function of dairy cows fed different fats. *Canadian Journal of Animal Science*, *81*(2), 263-271. https://doi.org/10.4141/A00-096

Phuphaphud, A., Saengprachatanarug, K., Posom, J., Taira, E., & Panduangnate, L. (2022). Prediction and classification of energy content in growing cane stalks for breeding programmes using visible and shortwave near infrared. *Sugar Tech*, *24*(5), 1497-1509. https://doi.org/10.1007/s12355-021-01075-2

Ramirez-Morales, I., Aguilar, L., Fernandez-Blanco, E., Rivero, D., Perez, J., & Pazos, A. (2021). Detection of bovine mastitis in raw milk, using a low-cost NIR spectrometer and k-NN algorithm. *Applied Sciences*, *11*(22), Article 10751. https://doi.org/10.3390/app112210751

Recio, J. A., Hermosilla, T., Ruiz, L. A., & Palomar, J. (2013). Automated extraction of tree and plot-based parameters in citrus orchards from aerial images. *Computers and Electronics in*

Agriculture, *90*, 24-34. https://doi.org/10.1016/j.compag.2012.10.005

Reith, S., & Hoy, S. (2018). Review: Behavioral signs of estrus and the potential of fully automated systems for detection of estrus in dairy cattle. *Animal*, *12*(2), 398–407. https://doi.org/10.1017/S1751731117001975

Sayid, A., Mosissa, D., Asnaku, F., Tamrat, D., Ayda, M., & Asmarech, Y. (2024). Behavioral and physical signs of estrous manifested in hormonally estrus-induced boran and holstein friesian dairy cattle. *EAS Journal of Veterinary Medical Science*, *6*, 1-8. https://doi.org/10.36349/easjvms.2024.v06i01.001

Schilkowsky, E. M., Granados, G. E., Sitko, E. M., Masello, M., Perez, M. M., & Giordano, J. O. (2021). Evaluation and characterization of estrus alerts and behavioral parameters generated by an ear-attached accelerometer-based system for automated detection of estrus. *Journal of Dairy Science*, *104*(5), 6222-6237. https://doi.org/10.3168/jds.2020-19667

Sharma, A., Jain, A., Gupta, P., & Chowdary, V. (2020). Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access*, *9*, 4843-4873. https://doi.org/10.1109/ACCESS.2020.3048415

Shine, P., & Murphy, M. D. (2021). Over 20 years of machine learning applications on dairy farms: A comprehensive mapping study. *Sensors*, *22*(1), Article 52. https://doi.org/10.3390/s22010052

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427-437. https://doi.org/10.1016/j.ipm.2009.03.002

Sveberg, G., Rogers, G. W., Cooper, J., Refsdal, A. O., Erhard, H. W., Kommisrud, E., ... & Ropstad, E. (2015). Comparison of Holstein-Friesian and Norwegian Red dairy cattle for estrus length and estrous signs. *Journal of Dairy Science*, *98*(4), 2450-2461. https://doi.org/10.3168/jds.2014-8905

Szenci, O. (2021). Recent possibilities for the diagnosis of early pregnancy and embryonic mortality in dairy cows. *Animals*, *11*(6), Article 1666. https://doi.org/10.3390/ani11061666

Takemura, G., Bázár, G., Ikuta, K., Yamaguchi, E., Ishikawa, S., Furukawa, A., ... & Tsenkova, R. (2015). Aquagrams of raw milk for oestrus detection in dairy cows. *Reproduction in Domestic Animals*, *50*(3), 522-525. https://doi.org/10.1111/rda.12504

Tenghe, A. M. M., Bouwman, A. C., Berglund, B., Strandberg, E., Blom, J. Y., & Veerkamp, R. F. (2015). Estimating genetic parameters for fertility in dairy cows from in-line milk progesterone profiles. *Journal of Dairy Science*, *98*(8), 5763-5773. https://doi.org/10.3168/jds.2014-8732

Toledo-Alvarado, H., Vazquez, A. I., De Los Campos, G., Tempelman, R. J., Gabai, G., Cecchinato, A., & Bittante, G. (2018). Changes in milk characteristics and fatty acid profile during the estrous cycle in dairy cows. *Journal of Dairy Science*, *101*(10), 9135-9153. https://doi.org/10.3168/jds.2018-14480

Tsenkova, R., Atanassova, S., Morita, H., Ikuta, K., Toyoda, K., Iordanova, I. K., & Hakogi, E. (2006). Near infrared spectra of cows' milk for milk quality evaluation: Disease diagnosis and pathogen identification. *Journal of Near Infrared Spectroscopy*, *14*(6), 363-370. https://doi.org/10.1255/jnirs.661

Vance, C. K., Tolleson, D. R., Kinoshita, K., Rodriguez, J., & Foley, W. J. (2016). Near infrared spectroscopy in wildlife and biodiversity. *Journal of Near Infrared Spectroscopy*, *24*(1), 1-25. https://doi.org/10.1255/jnirs.1199

Wang, J., Zhang, Y., Bell, M., & Liu, G. (2022a). Potential of an activity index combining acceleration and location for automated estrus detection in dairy cows. *Information Processing in Agriculture*, *9*(2), 288-299. https://doi.org/10.1016/j.inpa.2021.04.003

Wang, Y., Kang, X., He, Z., Feng, Y., & Liu, G. (2022b). Accurate detection of dairy cow mastitis with deep learning technology: A new and comprehensive detection method based on infrared thermal images. *Animal*, *16*(10), Article 100646. https://doi.org/10.1016/j.animal.2022.100646

Wenz, J. J. (2021). Influence of steroids on hydrogen bonds in membranes assessed by near infrared spectroscopy. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1863*(4), Article 183553. https://doi.org/10.1016/j.bbamem.2021.183553

Weyer, L. G., & Lo, S.-C. (2006). Spectra–structure correlations in the near-infrared. In J. M.

Chalmers & P. R. Griffiths (Eds.), *Handbook of Vibrational Spectroscopy*. John Wiley & Sons, Ltd. https://doi.org/10.1002/0470027320.s4102

Workman Jr, J. J. (1996). Interpretive spectroscopy for near infrared. *Applied Spectroscopy Reviews*, *31*(3), 251-320. https://doi.org/10.1080/05704929608000571

Wu, D., He, Y., & Feng, S. (2008). Short-wave near-infrared spectroscopy analysis of major compounds in milk powder and wavelength assignment. *Analytica Chimica Acta*, *610*(2), 232-242. https://doi.org/10.1016/j.aca.2008.01.056

Wu, X., Wu, B., Sun, J., & Yang, N. (2017). Classification of apple varieties using near infrared reflectance spectroscopy and fuzzy discriminant c-means clustering model. *Journal of Food Process Engineering*, *40*(2), Article e12355. https://doi.org/10.1111/jfpe.12355

Yang, K., An, C., Zhu, J., Guo, W., Lu, C., & Zhu, X. (2022). Comparison of near-infrared and dielectric spectra for quantitative identification of bovine colostrum adulterated with mature milk. *Journal of Dairy Science*, *105*(11), 8638-8649. https://doi.org/10.3168/jds.2022-21969

Yeung, K. Y., & Ruzzo, W. L. (2001). Principal component analysis for clustering gene expression data. *Bioinformatics*, *17*(9), 763-774. https://doi.org/10.1093/bioinformatics/17.9.763

Yizengaw, L. (2017). Review on estrus synchronization and its application in cattle. *International Journal of Advanced Research in Biological Sciences*, *4*(4), 67-76. http://doi.org/10.22192/ijarbs.2017.04.04.010