Iranian Society of Animal Science

Journal of Livestock Science and Technologies

Shahid Bahonar University of Kerman

ISSN: 2322-3553 (Print) ISSN: 2322-374X (Online)

Paper type: Original Research

Predicting blood beta-hydroxybutyric acid in dairy cow herds through machine learning-based feature selection: On-farm data basis

Daniel Moodi[1], Amin Khezri[1], Abbas Ali Naserian[2], Mostafa Ghazizadeh-Ahsae [3], Omid Dayani[1], Vahid Bahrampour[4]*

[1]*Department of Animal Science, College of Agriculture, Shahid Bahonar University of Kerman, Iran*
[2]*Department of Animal Science, College of Agriculture, Ferdowsi University of Mashhad, Iran*
[3]*Department of Computer Engineering, Shahid Bahonar University of Kerman, Iran*
[4]*Department of Agricultural Engineering, National University of Skills. Tehran. Iran*

*Corresponding author,
E-mail address:
akhezri@uk.ac.ir
aminkhezri@gmail.com

Received: 18 Dec 2024,
Received in revised form: 07 Jan 2025,
Accepted: 03 Feb 2025,
Published online: 05 Feb 2025,

ORCID
Daniel Moodi
0009-0007-5634-8902
Amin Khezri
0000-0002-5371-9831
Abbas Ali Naserian
0000-0003-1179-6262
Mostafa Ghazizadeh-Ahsaee
0000-0002-3482-7642
Omid Dayani
0000-0002-7067-8242
Vahid Bahrampour
0000-0001-9500-9873

Abstract In dairy industry, high producing fresh dairy cows commonly experience adipose tissue mobilization to support their energy requirements. Precise prediction of blood beta-hydroxybutyric acid (BHBA) concentration could significantly enhance the cow health and welfare, therefore, this study aimed to identify the key factors influencing BHBA levels and develop predictive models based on nutritional and performance data in fresh dairy cows. In this trial, four years data from 325 fresh Holstein cows were collected and analyzed. Various machine learning algorithms, including decision trees, random forests, Lasso and ridge regression models, as well as boosting and bagging techniques, were employed to estimate BHBA levels and identify the influential factors. These algorithms were assessed using the coefficient of determination (R^2). The random forest model demonstrated the lowest error, with a mean absolute error of 0.02, while the linear model exhibited the highest error, with a mean absolute error of 1.25. It was found that factors including milk production, previous lactation days in milk (DIM), sampling day, body weight change, BCS at parturition, and the amount and type of dietary fat, as well as overall diet quality were highly significant for estimating blood BHBA levels ($P<0.05$). Notably, the results indicated that cows with a BCS of 3 or lower, as well as those with a score of 3.75, are crucial categories for predicting BHBA. Additionally, the level and type of fatty acids in the diet, particularly lauric (C12:0), palmitic (C16:0), linolenic (C18:3), and oleic acids (C18:1), had significant influence on BHBA in fresh cows ($P<0.05$). These findings highlight the importance of integrating these critical factors into predictive models to enhance metabolic health monitoring and improve dairy herd management practices.

Keywords: beta-hydroxybutyric acid, machine learning algorithms, dairy cows

Introduction

Management of metabolic health in dairy herds is essential for optimizing milk production and ensuring animal welfare. A key metabolic indicator of health in dairy cows is the level of blood beta-hydroxybutyric acid (BHBA), a primary ketone body, which is produced during tissue fatty acid metabolism (Khezri et al., 2009). High blood BHBA concentrations as a sign of ketosis in early lactation, are associated with increased risk of displaced abomasum, higher culling rates, and reduced reproductive success (McArt et al., 2013). Additionally, metabolic issues in early lactation have been linked to altered lactation curves, characterized by a lower peak milk yield and decreased 305-day milk production (Hut et al., 2021). However, traditional metabolic monitoring methods are often invasive and time-intensive. Consequently, there Is increasing interest in developing predictive models that use on-farm data to estimate BHBA levels in a non-invasive

2025, 13 (1): 67-78 DOI: 10.22103/jlst.2025.24546.1589

and cost-effective way (Jansen et al., 2021).

Data generation in agriculture and biotechnology has greatly increased in recent years due to the very rapid development of high-performance technologies (Mohammadabadi et al., 2024). These data are obtained by studying products, foods, and biological molecules to understand the role of different aspects of agriculture in determining the structure, function, and dynamics of living systems (Hamidi et al., 2017). Artificial neural networks have been proposed to alleviate the limitation of traditional methods and can be used to handle nonlinear and complex data, even when the data are imprecise and noisy (Hamidi et al., 2017; Ghotbaldini et al., 2019).

Machine learning (ML) is a valuable tool for understanding the complex biological systems and can also highlight the gaps in scientific knowledge. Furthermore, this technique has opened new possibilities for predictive analytics in various fields such as dairy herd management. A crucial step in creating robust ML models is feature selection, which involves identifying the most relevant variables to enhance prediction accuracy. Effective feature selection improves model performance, reduces computational demands, and increases interpretability (Guyon and Elisseeff, 2003). In the context of BHBA prediction, feature selection can provide insights into key indicators of metabolic health, supporting targeted intervention strategies (Shine and Murphy, 2021).

Despite the potential of ML-based models, their implementation in real-world dairy herd management is not without challenges. One significant limitation is the variability and quality of on-farm data, which can be inconsistent due to differences in recording practices, environmental conditions, and management systems. To overcome this, standardizing the data collection protocols and integrating automated systems for accurate and consistent data logging is essential. Another challenge lies in the interpretability of complex ML algorithms, which can be improved by employing simpler algorithms where possible or integrating explainable AI techniques (Schillings et al., 2021). Additionally, the computational demands of ML models, particularly for large datasets, may limit their applicability to smaller farms or in regions with limited technological infrastructure.

Recently, application of ML in dairy science has gained attention, providing valuable insights into herd management and production, because of successfully predicted traits relevant to dairy cow health and performance, including mastitis incidence (Kamphuis et al., 2010; Ebrahimie et al., 2018) and milk yield (Gianola et al., 2011; Wallén et al., 2018). Machine learning algorithms, capable of handling large datasets and uncovering complex variable relationships, offer promising tools for predicting BHBA levels using accessible on-farm data. Several studies have attempted to estimate blood BHBA levels based on behavioral data, rumination time and locomotion (Monshouwer, 2020), milk components (Satoła and Bauer, 2021), breed, parity, DIM, daily milk yield, and milk fatty acid profiles (Mandujano Reyes et al., 2021), and near-infrared analysis of milk of cows (Giannuzzi et al., 2022), but to authors' knowledge no research has yet explored the application of ML for predicting BHBA levels in fresh dairy cows based on diet analysis. Therefore, this study aimed to develop an ML-based model to predict blood BHBA levels in dairy herds using on-farm data. Moreover, the objectives were to identify the most significant predictors through rigorous feature selection and to validate the model's predictive performance. By integrating new ML techniques with comprehensive on-farm data, this research also seeks to provide a practical tool for the early detection and management of ketosis through blood BHBA prediction, ultimately enhancing dairy herd health and productivity.

Materials and methods

Animals and diets

Data were collected from 325 fresh Holstein dairy cows at the Khazaei dairy farm (35.63° N latitude and 59.49° E longitude) in Jolge Rokh, Khorasan Razavi Province, over a four-year period (2018–2021). Cows were randomly selected post-calving without considering factors such as parity, BCS, or other relevant characteristics and weighed at two intervals, including calving and 21-d of lactation. The selection process ensured inclusion of cows from all parties and body condition score ranges, representing a diverse cross-section of the herd to enhance reproducibility and applicability of the findings. The data were categorized into three primary sections: blood parameters, performance metrics, and nutritional information. After calving, the cows were housed in a dedicated fresh-cow barn, where they were provided a carefully formulated diet to meet the nutritional demands of a cow producing 40 kilograms of milk per day. The average forage and chemical composition of the diets was as follows: forage (43.85 ± 1.62 %); net energy for lactation (NEL, 1.73 ± 0.06 Mcal/kg); crude protein (CP, 17.05 ± 0.35 %); rumen-degradable protein (RDP, 11.65 ± 0.45 % of CP); rumen-undegradable protein (RUP, 5.41 ± 0.27 % of CP); fat (3.7 ± 0.22 %); neutral detergent fiber (NDF, 29.9 ± 0.84 %); acid detergent fiber (ADF, 19.75 ± 0.76 %); organic matter (OM, 91.5 ± 0.42 %); calcium (Ca, 0.86 ± 0.07 %) and phosphorus (P, 0.6 ± 0.05 %).

Sampling and data collection

The BHBA levels were measured using a Novavet ketone meter (Sensivity: 95 %, WD1624, England) and based on validation of BHBA test-strips for venous samples. Blood samples were collected into EDTA blood tubes by jugular venipuncture from all examined cows two hours after morning feeding on days 1 and 21 post-calving. Milk yield was recorded twice monthly (d-15 and 30) until the end of lactation period. Nutritional requirements were continuously monitored, and diets were adjusted as necessary.

Ration information, including the diet summaries, energy, fatty acids (C12:0, C16:0, C18:0, C18:1, C18:2, C18:3 and …), carbohydrates (WSC, starch and …), crude protein, organic matter, NDF, ADF, and minerals (Na, Cl, Ca, S and P) were extracted from the feed report of the National Academies of Sciences and Medicine's software (NASM, 2021). All relevant data were recorded in Excel for further analysis. To ensure data integrity, cross-checks were performed by comparing feed composition results with values provided in the software database.

The performance data used in the modeling included: lactation period, calving weight, weight at 21 days of lactation, weight difference between calving and 21 days of lactation, percentage of weight difference, blood BHBA concentration, sampling day, month of parturition, calving status, days in milk, total milk yield, standardized milk yield for 305 days (standardized to kilograms of fat- and protein-corrected milk, FPCM), mean milk yield, previous lactation DIM, previous milk yield, standardized milk yield for 305 days in previous lactation, average of milk yield in previous lactation, BCS at parturition, BCS in the first and second months postpartum, BCS at dry-off (BCS was measured on a scale of 1 to 5 by three skilled individuals independently), daily milk yield at 15, 30, and 45 days of lactation. This comprehensive dataset facilitated the modeling of BHBA concentrations and other performance metrics to predict and mitigate potential metabolic issues in dairy cows.

Data preprocessing

To prepare the data for modeling, various standard data preprocessing techniques were employed. First, data imputation was utilized to address any missing values. Second, outliers were detected using a combination of the interquartile range (IQR) method and visual inspection of scatter plots. Data points falling beyond 1.5 times the IQR above the third quartile or below the first quartile were flagged as potential outliers. Identified outliers were evaluated for biological plausibility, and those deemed erroneous were excluded or replaced using mean substitution. Finally, normalization techniques were used to ensure that different attributes of the model were assigned equal weights by Box-Cox method (log-transformed). The preprocessing procedures were carried out using several Python libraries, including NumPy and scikit-learn, among others. These tools facilitated the efficient handling and preparation of the data, ensuring data suitability for subsequent modeling and analysis (Gareth at al., 2021).

Machine learning algorithms

Machine learning algorithms were evaluated for their effectiveness in predicting blood BHBA concentrations and the metabolic status of dairy cows. The implemented models fell within the category of regression models, including various linear models. Each model comprised 260 variables, which included factors such as body weight and its normalized and scaled values, changes in body weight, pre- and post-partum body condition scores, milk production during the first 45 days, calving condition score, parity, open days, BHBA sampling day, days in milk before parturition, average pre-parturition milk yield, standardized milk yield, total pre-parturition milk yield, normalized and scaled values, and dietary information.

The choice of specific algorithms, such as ridge and Lasso regression, was guided by their ability to handle multicollinearity and high-dimensional data efficiently. Ridge regression was selected for its effectiveness in scenarios with many correlated predictors, while Lasso regression was chosen for its capacity to perform variable selection, resulting in more interpretable models. Decision trees and random forests were included due to their robustness in handling nonlinear relationships and their effectiveness in ensemble learning approaches. Other potential algorithms, such as support vector machines and neural networks, were not considered due to their higher computational demands and potential overfitting risks with relatively small datasets (Schapire, 2003).

Evaluating multiple algorithms within a specific domain is standard in machine learning, as algorithm performance can be influenced by various factors such as variables, sample sizes, and data characteristics. This study explored several linear models, including simple linear models, multiple linear models, ridge and Lasso regression models, decision trees, and random forests. The K-fold cross-validation method was employed to partition the data into training and test datasets at 80:20 ratio. Models were trained on the training data, and their accuracy was subsequently assessed using the test data. This approach ensured a robust evaluation of model performance and predictive accuracy (Pedregosa et al., 2011).

In this study, the linear model was employed with a focus on its predictive utility rather than inferential interpretation (Kutner et al., 2005). The methodologies utilized in this investigation are further elaborated as follows.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \qquad i = 1,2,\ldots,n$$

The primary objective of the ridge and Lasso linear models in the current study was to minimize the sum of squared errors (SSE; Hastie et al., 2009). The formula for SSE is given by:

$$\mathrm{SSE} = \sum_{i=1}^{n}(y_i - \hat{y})^2 = \sum_{i=1}^{n}\Big(y_i - (\hat{\beta}_0 + \sum_{j=1}^{m}\hat{\beta}_m x_{ij})\Big)^2$$

In ridge regression, an additional penalty term was introduced to the objective function to be minimized:

$$\sum_{i=1}^{n}(y_i - (\hat{\beta}_0 + \sum_{j=1}^{m}\hat{\beta}_m x_{ij}))^2 + \lambda\sum_{j=1}^{m}\hat{\beta}_m{}^2 ,\lambda \geq 0$$

Ridge regression is particularly useful when the number of predictive variables exceeds the number of observations, or when multicollinearity exists among the variables. These challenges arise when the model fits the data too closely using all available variables, which can lead to overfitting and poor generalizability to new data. By adding a penalty term, ridge regression helped to reduce the model's parameters, thereby decreasing the number of significant independent variables.

Models generated by the Lasso linear model are generally more straightforward to interpret compared to those produced by the ridge linear model (Tibshirani, 1996). This makes the following Lasso model an attractive alternative for situations where model interpretability is paramount.

$$\sum_{i=1}^{n}(y_i - (\hat{\beta}_0 + \sum_{j=1}^{m}\hat{\beta}_m x_{ij}))^2 + \lambda\sum_{j=1}^{m}|\hat{\beta}_m| ,\lambda \geq 0$$

The decision tree algorithm was used to serves as a decision support model, represented by a tree-like diagram that illustrates decisions and their potential outcomes at each node. In this algorithm sub-trees (lower levels of the tree) were constructed through the recursive partitioning of each sample and the objective of selecting each feature at each node was to minimize the residual sum of squares, thereby optimizing the model's predictive accuracy as follows:

$$Rss = \sum_{j=1}^{J}\sum_{i\epsilon R_j}(y_i - \hat{y}_{R_j})^2$$

The Random Forest algorithm was used to leverage multiple decision trees to construct a highly effective predictive model and multiple decision trees were generated from the training samples. During the construction of these trees, a random subset of predictors was selected as candidates for each split, rather than considering the full set of predictors. Typically, the number of predictors considered at each split was the square root of the total number of predictors. This random sampling of predictors helped in enhancing the model's robustness and accuracy.

The bagging technique was employed to reduce the variance in decision trees by averaging the predictions from multiple trees. The concept is based on the principle that the averaging multiple observations reduces the overall variance. Specifically, the variance of the mean of n observations with variance σ2 is σ2/n. This principle was effectively applied to decision trees in the bagging process. By drawing multiple training sets from the population, constructing separate decision tree models for each set, and then averaging their predictions, a single, low-variance decision tree model was obtained. This process involves calculating $\hat{f}1$ (x), $\hat{f}2$(x), ..., $\hat{f}B$(x) using B separate training sets and averaging their predictions (Breiman, 1996). This ensemble method produced a more stable and accurate model by mitigating the variability inherent in individual decision trees.

Boosting technique was also used to enhance decision tree predictions through combining multiple decision trees ($\hat{f}1$, ..., $\hat{f}B$). Each new tree was trained to focus on the residual errors of the preceding trees, thereby enhancing the model's overall accuracy. By leveraging information from previous trees, boosting ensures that each subsequent tree is dependent on its predecessors, which stands in contrast to bagging's approach of constructing independent trees (Schapire, 2003).

Model evaluation

The employed models, trained with six machine learning algorithms, were evaluated using 10-fold cross-validation. In this method, the dataset was partitioned into ten equal, mutually exclusive subsets. The evaluation process iterates ten times, with each iteration using a different 10% of the data as the test set while the remaining 90% serves as the training set. To ensure that no cow appears in both the training and test sets simultaneously, each cow was included only once in each cross-validation loop. This systematic partitioning guarantees that different subsets of data were utilized for training and testing across the iterations, making this method an effective resampling technique for estimating the model's error.

Results and Discussion

Correlation analyses

Figure 1 presents the scatterplots depicting the relationships between blood BHBA levels and various performance parameters with Pearson correlation coefficients (r) provided in each panel, the presence of patterns, and outliers which we intended to estimate. The correlation between BHBA concentration and calving weight was 0.23 (P =0.00038; Figure 1). Though modest, this correlation is meaningful when considering the diverse management and environmental factors affecting dairy cows, including feed type, milk production, and climatic conditions. Despite the low correlation, the significance of the P value (below 0.05) is notable, especially in studies with large sample sizes or multiple influencing factors (Thomas and Krebs, 1997). In the current study, a correlation coefficient of 0.60 was observed between lactation period and cow weight at calving (Figure 2), consistent with other research showing similar increases in weight and age (Erb and Ashworth, 1961). As cows mature, their weight and milk production generally increase; however, weight gains are more notable at younger ages, tending to stabilize as cows reach maturity (NASM, 2021).

The effect of cow weight on milk production was strongest in the first month (r =0.44), decreasing slightly in the second month (r =0.41) (Figure 2). Heavier cows also tended to lose more weight in the first month, as indicated by a correlation of 0.33 (Figure 2). Similar findings were reported by Roche et al. (2009), who noted a positive association between body weight and milk yield, indicating that heavier cows typically produce more milk, although this relationship can vary depending on factors like breed, diet, and management practices (Ebrahimi et al., 2015; Martens, 2023).

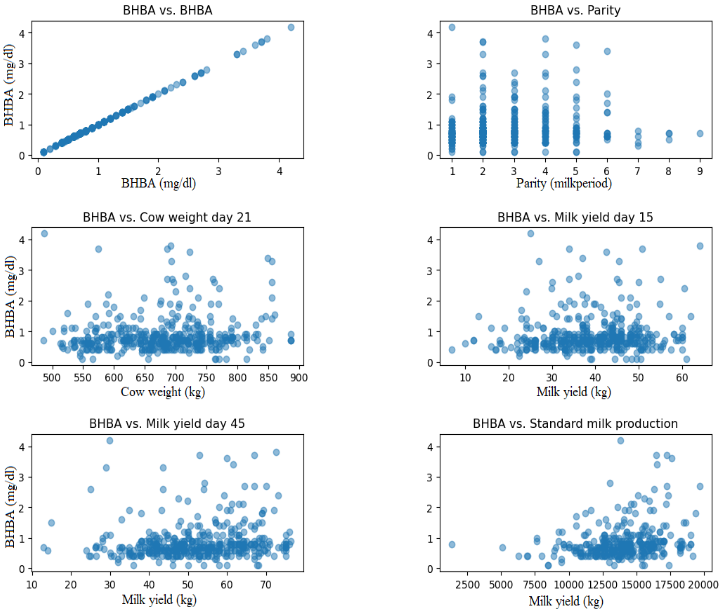
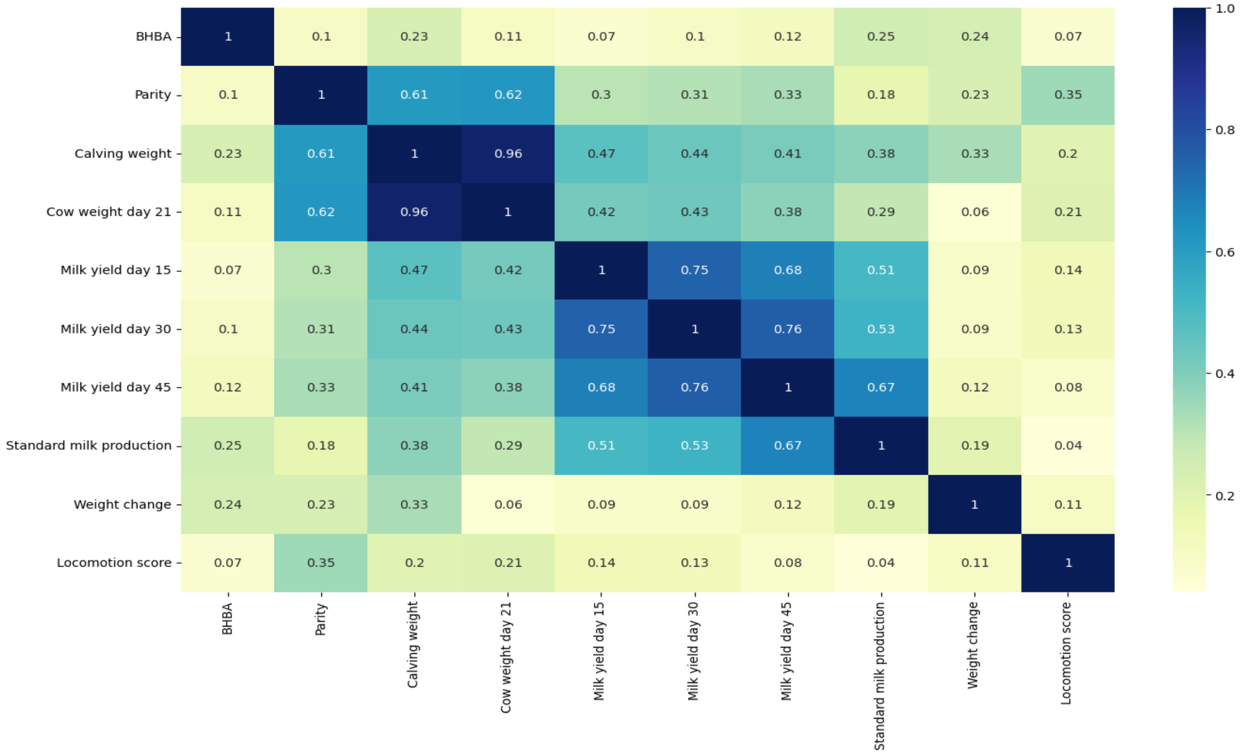


Figure 1. Scatterplot of blood beta-hydroxybutyric acid (BHBA) based on various performance factors in dairy cows



	BHBA	Parity	Calving weight	Cow weight day 21	Milk yield day 15	Milk yield day 30	Milk yield day 45	Standard milk production	Weight change	Locomotion score
BHBA	1	0.1	0.23	0.11	0.07	0.1	0.12	0.25	0.24	0.07
Parity	0.1	1	0.61	0.62	0.3	0.31	0.33	0.18	0.23	0.35
Calving weight	0.23	0.61	1	0.96	0.47	0.44	0.41	0.38	0.33	0.2
Cow weight day 21	0.11	0.62	0.96	1	0.42	0.43	0.38	0.29	0.06	0.21
Milk yield day 15	0.07	0.3	0.47	0.42	1	0.75	0.68	0.51	0.09	0.14
Milk yield day 30	0.1	0.31	0.44	0.43	0.75	1	0.76	0.53	0.09	0.13
Milk yield day 45	0.12	0.33	0.41	0.38	0.68	0.76	1	0.67	0.12	0.08
Standard milk production	0.25	0.18	0.38	0.29	0.51	0.53	0.67	1	0.19	0.04
Weight change	0.24	0.23	0.33	0.06	0.09	0.09	0.12	0.19	1	0.11
Locomotion score	0.07	0.35	0.2	0.21	0.14	0.13	0.08	0.04	0.11	1

Figure 2. Correlation between beta-hydroxybutyric acid (BHBA) and performance factors in dairy cows

The correlations between blood BHBA concentration and standard milk production and parturition weight were 0.25 and 0.23, respectively (Figure 2). Additionally, a modest association was found between weight change and BHBA concentration ($r = 0.24$, $P = 0.00018$). Research by Ospina et al. (2010) supports that elevated BHBA levels are linked to subclinical ketosis, which negatively impacts milk yield and overall health, underscoring BHBA's significance as a metabolic status indicator affecting milk production.

Analysis of data in the current study showed that in the early postpartum months, milk production was significantly influenced by factors such as parity (0.3), cow weight (0.47), and locomotion score (0.14). However, no specific relationship was observed between milk production and weight loss during this period. Over time, the correlation between milk production and the metabolic status of the animals increased, ranging from 0.07 to 0.12 (Figure 2). These are in agreement with the findings of Dechow et al. (2002), indicating that older cows (higher parity) and cows with better body condition scores and mobility tend to have higher milk yields. However, the exact correlation values may differ due to variations in herd management and environmental conditions. In the current study, present study, milk production during early lactation showed to be a reliable predictor of future milk yield, demonstrating a high and accurate coefficient. Consistent with these results, Coffey et al. (2002) also showed that milk production in the initial weeks of lactation was a reliable indicator of overall lactation performance. This predictive relationship can guide dairy management decisions related to feeding and breeding.

Model selection

Traditional statistical analysis of data with numerous dependent variables can be challenging, but machine learning (ML) algorithms, which better handle large datasets, can properly manage extensive farm data with highly correlated variables. In this study, we assessed body weight change (%) in the first 21 days of lactation as a factor in a linear model hypothesized to represent the energy balance. However, using only this factor proved insufficient for precise prediction. Similar to findings by Grzesiak et al. (2003), in our study, ML methods demonstrated superior predictive accuracy over traditional statistical models in estimating milk yield and body condition score changes, particularly with complex, non-linear relationships among multiple variables. Roche et al. (2009) emphasized that BCS and weight changes were critical metabolic indicators but highlighted the necessity of multifactorial approaches that incorporate additional variables, such as feed intake and milk yield, for more accurate predictions.

In the current study, inclusion of all predictive factors in one linear equation did not also yield optimal results, but applying feature selection methods improved the accuracy (Table 1). Sahin et al. (2012) found that feature selection enhanced the regression model performance for milk yield prediction, highlighting that selecting relevant predictors rather than including all available variables is critical to enhance the predictive accuracy. This approach aligns with the current findings that optimal predictive factors should be carefully selected. Furthermore, skewed linear model error distributions in our study further limited the effectiveness of linear approaches, aligning with De Vries et al. (2011), who recommended the ML models for skewed data due to their capacity to handle non-normal distributions.

In our study, the forward selection method was also used, starting with an empty model and adding variables iteratively to minimize residual sum of squares (RSS). The final model was evaluated using R^2, AIC, and BIC to select the most appropriate predictive factors. By employing AIC and BIC, the model's prediction error for unseen data was reduced (Table 1). Grzesiak et al. (2006) also demonstrated the effectiveness of stepwise selection for complex data, emphasizing the importance of combining multiple criteria, like AIC and BIC, for model selection to avoid overfitting and ensuring model robustness.

Table 1. Performance of machine learning algorithms in linear regression prediction of blood beta-hydroxybutyric acid (BHBA) concentration

Model type	Mean absolute error	SD of MSE[1]	Minimum of MSE	Maximum of MSE
The linear model t-test	1.258620	4.158575	0.004027	33.497567
The forward selection model R^2	0.028770	0.025506	0.000183	0.112742
Forward selection model AIC	0.027620	0.023730	0.000439	0.101792
Forward selection model BIC	0.023440	0.015959	0.001279	0.058719
Ridge linear model	0.021858	0.015900	0.000159	0.059979
Lasso linear model	0.026066	0.020277	0.005266	0.088457
Decision tree model	0.031359	0.029421	0.000000	0.127274
Bagging model	0.020820	0.017289	0.000262	0.067841
Random forest model	0.020634	0.017324	0.000603	0.071668
Gradient boosting model	0.023740	0.018940	0.000041	0.088894

[1] Standard deviation of the mean squared errors

In our study, R^2 favored 53 variables in the model (Figure 3a) with highest adjusted R^2 (Table 2 shows 53 variables), while AIC and BIC recommended fewer predictors for improved accuracy, with AIC selecting 49 variables (Figure 3b) and BIC only 5 variables (Figure 3c) with lowest AIC and BIC. Burnham and Anderson (2004) affirmed the superiority of AIC and BIC for model selection, as models chosen with these criteria typically offer better predictive performance and are less prone to overfitting, supporting our findings.

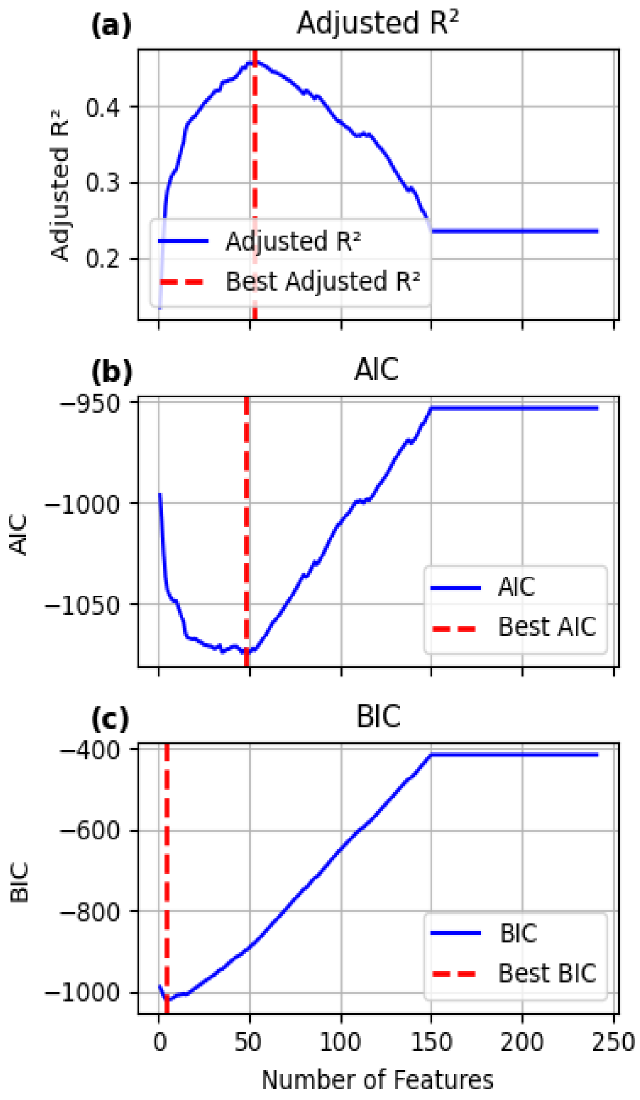


Figure 3. Model Selection Based on R2(a), Akaike information criterion (AIC) (b) and Bayesian information criterion (BIC) (c)

Ridge and random forest regressions resulted in the best performance for estimating BHBA concentration (Table 1), demonstrating the lowest errors by adding penalty terms or using ensemble methods to handle data complexity and prevent overfitting (Hut et al., 2021). These results are in agreement with DeVries et al. (2020), who reported that advanced regression techniques, including ridge regression and random forest regression, are effective in improving the prediction accuracy and managing the multicollinearity in dairy cow data.

In the current study, the random forest model exhibited superior performance (Table 1), which is comparable to studies on milk yield estimation (Grzesiak et al., 2006; Gianola et al., 2011) and mastitis detection (Cavero et al., 2008; Sun et al., 2010). Moreover, ridge regression models also provided valuable parameters and weights, and feature selection was facilitated by both forward and backward linear models. The model with the lowest error, based on squared errors, included 53 features. The ten most significant predictive variables, as determined by the linear and random forest models, are presented in Table 3.

Feature selection

This study aimed to predict plasma BHBA concentrations using herd performance and nutritional parameters, particularly through dietary factors readily accessible in dairy herds. Notably, features like dietary sodium and chlorine, though less commonly reported, emerged as important contributors to the dietary cation-anion difference (DCAD), which influences milk production and fat content. DCAD modulates systemic acid-base balance, which can directly affect metabolic processes, including energy partitioning and utilization in early lactation. Hu et al. (2007) showed that higher dietary DCAD levels can enhance milk production and alter milk composition. However, our study found that the optimal DCAD concentration remains uncertain, consistent with the variability reported in other studies (Lean et al., 2006). Notably, features like dietary sodium and chlorine, though less commonly reported, emerged as important contributors to the DCAD with influence on milk production and fat content. Our models demonstrated that since DCAD can impact milk production, fat content, and feed intake (Iwaniuk et al., 2015), it may also influence the energy status of cows in early lactation and potentially affect the level of blood BHBA, consistent with the variability reported in other studies (Lean et al., 2006; Duffield et al., 2009; Ospina et al., 2010, 2013; Iwaniuk et al., 2015).

The effect of NaCl on rumen function and milk yield in our study was partially supported by studies indicating that salt intake can affect rumen contraction rate and acetate to propionate ratio (White et al., 2019). Acetate serves as a primary precursor for milk fat synthesis, while propionate supports gluconeogenesis. Inadequate gluconeogenic precursors can elevate BHBA levels as cows mobilize body fat to meet their energy deficits.

Our study confirmed that weight loss, as an indicator of negative energy balance, can be managed through targeted nutrition. These findings agree with Roche et al. (2009) and Hut et al. (2021), who showed that nutritional strategies can have crucial role in minimizing the weight loss of dairy cows during early lactation. The correlation between greater weight loss and higher negative energy balance also concurs with the existing literature (Castañón et al., 2023; Mekuriaw, 2023).

BCS provides a precise assessment of the energy status, though it may be less applicable in extremely obese or underweight cows (Roche et al., 2015). BCS at calving is particularly crucial, as it influences the dry matter intake (DMI) in early lactation, postpartum BCS decline, milk yield, immune strength, and fertility (Roche et al., 2015; 2023). These researchers discussed how higher BCS levels were associated with decreased DMI and increased metabolic disorders.

Adipose tissue mobilization in over-conditioned cows releases non-esterified fatty acids (NEFAs) into the bloodstream, which are partially oxidized into BHBA in the liver thus increasing the nutrient intake challenges and fat mobilization issues during the transition from dry periods to lactation. BCS at calving affects feed intake capacity and is inversely related to DMI, although the underlying mechanisms remain unclear (Richards et al., 2020). Our study found an inverse relationship between BCS and DMI as shown by Poczynek et al. (2023). Leaner cows typically demonstrate lower milk production and longer intervals between calving and estrus, suggesting potential discomfort (Roche et al., 2015; Hut et al., 2021; Mekuriaw, 2023; Roche et al., 2023).

Table 2. The variables of R^2 adjusted model selection

N	Features	Coefficient	N	Features	Coefficient
1	Constant[1]	211.51	28	Diet 1 sulfur	5.615189
2	C18:3 diet 1	0.002713	29	Diet 1 C18:0	-0.00019
3	Body weight loss square[2]	0.000205	30	Lameness	0.002521
4	Diet sodium (Na)	4.473716	31	Diet 1 magnesium (Mg)	3.351681
5	Previous lactation days	0.000118	32	Number of inseminations previous milk peri	-6.7E-05
6	Body condition score less than 3	-0.00976	33	Previous average milk scaled and normaliz	-0.02386
7	Diet 1 DCAD[3]	0.001844	34	Previous standard milk scaled and normaliz	0.028914
8	Body condition= 3.75	-0.02116	35	Diet 1 ash	-0.1236
9	Diet 2 WSC[4]	-0.02424	36	Week 1 of sampling[6]	-0.0058
10	Diet 2 chlorine (CL)	-3.03436	37	Diet 2 calcium (Ca)	0.483042
11	C18:1 trans	0.144041	38	C12:0	0.14207
12	C14:0	-0.09207	39	Diet 1 fat C140	0.021879
13	Diet chlorine (CL)	1.11041	40	Others fatty acids[7]	0.006961
14	Diet 1 C18:1 trans	-0.05108	41	C18:1 trans	-0.09994
15	Magnesium (Mg)	-4.12864	42	C18:2	0.001566
16	Days used diet 1	-0.00288	43	Forage NDF	0.042128
17	Diet 2 NDF[5]	0.08139	44	Dietary 1 water intake	0.003859
18	BCS 4 and higher	0.002839	45	Days of first Insemination in previous lactation period	0.001435
19	Average milk normalized	115.7359	46	Lactation period	-0.00775
20	Total milk previous lactation scaled and normalized	-0.06639	47	Lactation period = 4& more	0.025597
21	Lactation period 3	0.018561	48	Phosphorus	-1.11492
22	NDF[5]	-0.12329	49	NEL[8]	-0.81461
23	Diet 2 sodium (Na)	3.999608	50	Average milk production in previous lactatio period	-126.72
24	Days used diet 2	0.002106	51	Average milk production scaled and normalized	487.0803
25	WSC[4]	-0.02191	52	Diet 2 DCAD	0.010128
26	Diet 1 sodium (Na)	-3.99446	53	Percentage of lost weight	0.008476
27	Dystocia	0.002568	54	Total saturated fatty acids	-0.00157

Due to dietary adjustments made during the experiment, the feeding period was divided into diet 1 and diet 2

[1] Constant: Y-intercept

[2] Square of weight loss in fresh dairy cows: (Lost weight of fresh cows)^2

[3] DCAD: Dietary cation-anion difference

[4] WSC: Water soluble carbohydrates

[5] NDF: Neutral detergent fiber

[6] Week 1 of sampling: Blood sampling of BHBA in first 7 days

[7] Others fatty acids: Fatty acids except C12:0, C14:0, C16:0, C16:1, C18:0, C18:1 trans, C18:1 cis, C18:2, C18:3 based on NRC (2021)

[8] Net energy for lactation

Table 3. Important variables based on different models

Selected features based on Random Forest	Selected features based on BIC[1]	Selected features based on AIC[2] and R^2
Previous days in milk	Fat C18:3	Fat C18:3
Sampling day	Body weight loss squared	Body weight loss squared
12-Carbon saturated fatty acids	Dietary sodium	Dietary sodium (Na)
First-feed days	Pre-calving lactation days	Previous lactation days
Second-feed days	Body condition score less than 3	Body condition score less than 3
Total saturated fatty acids		Dietary DCAD
Palmitic acid		Body condition score 3.75
Body weight loss squared		Water soluble carbohydrates
Weight change		Dietary chlorine (CL)
Milk production day 45		Fat C18:1

[1] Bayesian Information Criterion

[2] Akaike Information Criterion

In the present study, BCS levels of 3.75 and below 3 were identified as critical scores for further investigation. Some studies noted that cows with a BCS below 3 produced more milk than those with higher scores (Roche et al., 2009), resulting in a more pronounced negative energy balance and elevated blood BHBA concentrations. Conversely, Berry et al. (2007) found that cows with a BCS of 3 or lower produced less milk throughout lactation, aligning with our findings. In our herd case study, a target BCS of 3.5 at calving was associated with higher instances of lameness and illness among cows with a BCS below 3.

Our study also identified dietary content of water-soluble carbohydrates (WSC) as a significant predictor of blood BHBA levels. WSC showed a negative correlation with BHBA concentrations, supporting rumen fermentation and milk production benefits. These findings align with previous studies (Lee et al., 2003; Wei and Liu, 2018; Ravelo et al., 2022), which highlight improved nitrogen utilization and rumen conditions in response to dietary WSC. According to NASEM (2021), WSC facilitates microbial fermentation in the rumen, enhancing the production of acetate and butyrate, which are vital for milk synthesis and energy supply.

Diets with up to 5% WSC of dry matter have shown positive impacts on rumen fermentation and milk production (Klevenhusen et al., 2019; Ravelo et al., 2022). Moreover, some studies reported increased feed intake, milk yield, and milk protein content with reduced milk fat in lactating goats (Schmidely et al.,1999; Tajaddini et al., 2021). They also noted decreased non-esterified fatty acids, blood BHBA, blood urea, and increased blood insulin, aligning with our results on blood BHBA levels.

In the current study, fatty acids, including linolenic, oleic, lauric, and palmitic acids, were also included in the models and analyzed to estimate blood BHBA concentrations. Fatty acids provide critical energy for high-producing dairy cows and play a crucial role in hepatic nutrient partitioning, which is governed by key metabolic pathways, fatty acids and other substrates (Weld et al., 2020). An imbalance in hepatic fatty acid metabolism can contribute to metabolic disorders such as hyperketonemia and hepatic lipidosis, generating ketone bodies like BHBA (White, 2015). Our findings on stearic and oleic acids in triglyceride storage and milk production correspond with Ibeagha-Awemu et al. (2016). Although the regulatory mechanisms in the liver remain somewhat obscure, results of some studies shows that high fatty acid levels impact genes like pyruvate carboxylase, enhancing tricarboxylic acid (TCA) cycle activity and potentially increasing ketone production and glycerol accumulation, which can lead to hyperketonemia and hepatic lipidosis (Reynolds et al., 2003; White, 2015). Furthermore, it has been shown that inclusion of linseed and sunflower oils in diets of dairy cows can reduce milk fat by 30% and influence mammary gland gene expression, potentially ameliorating the negative energy balance (Ibeagha-Awemu et al., 2016).

In our study, previous lactation DIM data was used to estimate blood BHBA levels. Lean et al. (2023) found that ketosis risk increased from 0.1% in primiparous cows to 13% in cows with over four lactations, consistent with other findings in this field (Benedet et al., 2019). Evaluating the relationship between milk yield and blood BHBA concentrations in the current study showed that higher-producing cows experienced longer periods of negative energy balance, thus increasing risk of ketosis and aligns with the findings of Kaufman et al. (2016).

Conclusions

In this study, we aimed to predict blood BHBA levels in dairy cows using machine learning algorithms based on nutritional and performance data. Our results highlighted the key predictors, including dietary fats, weight changes post-calving, previous lactation DIM, and body condition score. Saturated and unsaturated fatty acids, such as palmitic, lauric, oleic, and linoleic acids, along with body condition score at parturition, play critical roles in the health and performance of dairy cows during the transition period. These findings emphasize the need for targeted nutritional and management strategies to optimize cow health and production. Dairy farmers are encouraged to manage body condition scores, monitor weight changes post-calving, and optimize dietary fatty acid composition to reduce ketosis risks. Adjusting dietary cation-anion difference (DCAD) is also recommended for improving energy balance. Future research should explore additional predictors such as rumen microbiota and refine machine learning models to improve predictive accuracy, enabling more precise and actionable insights for dairy herd management.

Disclosure statement

The authors declare no conflict of interest in regards to this article.

References

Amaral, D.M., James Croom, W., Rakes, A.H., Leonard, E.S., Linnerud, A.C., 1985. Increased concentration of sodium chloride on milk production of cows fed low fiber diets. *Journal of Dairy Science* 68 (11), 2940-2947.

Benedet, A., Manuelian, C. L., Zidi, A., Penasa, M., De Marchi, M., 2019. Invited review: β-hydroxybutyrate concentration in blood and milk and its associations with cow performance. *Animal* 13(8), 1676-1689.

Breiman, L., 1996. Bagging predictors. *Machine Learning* 24(2), 123-140.

Burnham K.P., Anderson, D.R., 2004. Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research* 33(2), 261-304.

Castañón, M., Vargas-Bello-Pérez, E., Martens, H., 2023. Invited Review: Increasing milk yield and negative energy balance: A gordian knot for dairy cows? *Animals* (13), 3097-3115.

Cavero, D., Tölle, K.H., Henze, C., Buxadé, C., Krieter, J., 2008. Mastitis detection in dairy cows by application of neural networks. *Livestock Science* 114(2–3), 280-286.

Coffey, M.P., Emmans, G.C., Brotherstone, S., 2002. Genetic evaluation of dairy bulls for energy balance traits using random regression. *Animal Science* 74(1), 29-40.

Dechow, C.D., Rogers, G.W., Clay, J.S., 2002. Heritability and correlations among body condition score loss, body condition score, production and reproductive performance. *Journal of Dairy Science* 85(11), 3062-3070.

De Vries, A., Marcondes, M.I., 2020. Review: Overview of factors affecting productive lifespan of dairy cows. *Animal* (S1), s155-s164.

Duffield, T.F., Lissemore, K.D., McBride, B.W., Leslie, K.E., 2009. Impact of hyperketonemia in early lactation dairy cows on health and production. *Journal of Dairy Science* 92(2), 571-580.

Ebrahimie, E., Ebrahimi, F., Ebrahimi, M., Tomlinson, S., Petrovski, K.R., 2018. Hierarchical pattern recognition in milking parameters predicts mastitis prevalence. *Computers and Electronics in Agriculture* 147, 6-11.

Ebrahimi, Z., Mohammadabadi, M.R., Esmailizadeh, A.K., Khezri, A. Najmi Noori, A., 2015. Association of PIT1 gene with milk fat percentage in Holstein cattle. *Iranian Journal of Applied Animal Science* 5, 575-582.

Erb, R. E., Ashworth, U.S., 1961. relationships between age, body weight, and yield of dairy cows. *Journal of Dairy Science* 44(3), 515-523.

Gareth, J., Witten, D., Hastie, T., Tibshirani, R. 2021. An Introduction to Statistical Learning: With Applications in R. 1st Ed. New York, NY. Springer

Ghotbaldini, H., Mohammadabadi, M., Nezamabadi-Pour, H., Babenko, O. I., Bushtruk, M. V., Tkachenko, S. V. 2019. Predicting breeding value of body weight at 6-month age using Artificial Neural Networks in Kermani sheep breed. *Acta Scientiarum. Animal Sciences* 41(1), e45282.

Giannuzzi, D., Mota, L. F. M., Pegolo, S., Gallo, L., Schiavon, S., Tagliapietra, F., Katz, G., Fainboym, D., Minuti, A., Trevisi, E., Cecchinato, A. 2022. In-line near-infrared analysis of milk coupled with machine learning methods for the daily prediction of blood metabolic profile in dairy cattle. *Scientific Reports* 12(1), 1-13.

Gianola, D., Okut, H., Weigel, K.A., Rosa, G.J., 2011. Predicting complex quantitative traits with Bayesian neural networks: A case study with Jersey cows and wheat. *BMC Medical Genetics* 12 87-99.

Grzesiak, W., Błaszczyk, P., Lacroix, R., 2006. Methods of predicting milk yield in dairy cows -Predictive capabilities of Wood's lactation curve and artificial neural networks (ANNs). *Computers and Electronics in Agriculture* 54(2), 69-83.

Grzesiak, W., Lacroix, R., Wójcik, J., Blaszczyk, P., 2003. A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records. *Canadian Journal of Animal Science* 86(1), 3-10.

Guyon, I.M., André, E., 2003. An introduction to variable and feature selection. *Journal of Machine Learning Research* 3, 1157-1182.

Hamidi, S. P., Mohammadabadi, M. R., Foozi, M. A., Nezamabadi-Pour, H. 2017. Prediction of breeding values for the milk production trait in Iranian Holstein cows applying artificial neural networks. *Journal of Livestock Science and Technologies* 5(2) 53-61.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning: Data mining, inference, and prediction. 2nd Ed. New York, NY. Springer.

Hu, W., Murphy, M.R., Constable, P.D., Block, E., 2007. Dietary cation-anion difference and dietary protein effects on performance and acid-base status of dairy cows in early lactation. *Journal of Dairy Science* 90(7), 3355-66.

Hut, P.R., Hostens, M.M., Beijaard, M.J., van Eerdenburg, F.J.C.M., Hulsen, J.H.J.L., Hooijer, G. A., Stassen, E.N., Nielen, M., 2021. Associations between body condition score, locomotion score, and sensor-based time budgets of dairy cattle during the dry period and early lactation. *Journal of Dairy Science* 104(4), 4746-4763.

Ibeagha-Awemu, E.M., Li, R., Ammah, A.A., Dudemaine, P.L., Bissonnette, N., Benchaar, C., Zhao, X., 2016. Transcriptome adaptation of the bovine mammary gland to diets rich in unsaturated fatty acids shows greater impact of linseed oil over safflower oil on gene expression and metabolic pathways. *BMC Genomics* 17, 104-127.

Iwaniuk, M. E., Weidman, A. E., & Erdman, R. A. 2015. The effect of dietary cation-anion difference concentration and cation source on milk production and feed efficiency in lactating dairy cows. *Journal of Dairy Science* 98(3), 1950-1960.

Jansen, H.M., Zschiesche, M., Albers, D., Wemheuer, W., Sharifi, A.R., Hummel, J., 2021. Accuracy of subclinical ketosis detection with rapid test methods for BHBA in blood in commercial dairy farms. *Dairy* 671-683.

Kamphuis, C., Mollenhorst, H., Heesterbeek, J., Hogeveen, H., 2010. Detection of clinical mastitis with sensor data from automatic milking systems is improved by using decision-tree induction. *Journal of Dairy Science* 93, 3616-3627.

Kaufman, E.I., LeBlanc, S.J., McBride, B.W., Duffield, T.F., DeVries, T.J., 2016. Association of rumination time with subclinical ketosis in transition dairy cows. *Journal of Dairy Science* 99(7), 5604-5618.

Khezri, A., Rezayazdi, K., Mesgaran, M.D., Moradi-Sharbabk, M. 2009. Effect of different rumen-degradable carbohydrates on rumen fermentation, nitrogen metabolism and lactation performance of

Holstein dairy cows. *Asian-Australasian Journal of Animal Sciences* 22, 651-658.

Kutner, M.H., Nachtsheim, C. J., Neter, J., Li, W., 2005. Applied Linear Statistical Models. 5th Ed, McGraw-Hill, Irwin, New York.

Lean, I.J., DeGaris, P.J., McNeil, D.M., Block, E., 2006. Hypocalcemia in dairy cows: Meta-analysis and dietary cation- anion difference theory revisited. *Journal of Dairy Science* 89(2), 669-684.

Lean, I.J., LeBlanc, S.J., Sheedy, D.B., Duffield, T., Santos, J.E.P., Golder, H.M., 2023. Associations of parity with health disorders and blood metabolite concentrations in Holstein cows in different production systems. *Journal of Dairy Science* 106(1), 500-518.

Lee, M.R.F., Merry, R.J., Davies, D.R., Moorby, J.M., Humphreys, M.O., Theodorou, M.K., MacRae, J.C., Scollan, N.D., 2003. Effect of increasing availability of water-soluble carbohydrates on *in vitro* rumen fermentation. *Animal Feed Science and Technology* 104(1-4), 59-70.

Mandujano Reyes, J. F., Walleser, E., Hachenberg, S., Gruber, S., Kammer, M., Baumgartner, C., Mansfeld, R., Anklam, K., Döpfer, D. 2021. Full model selection using regression trees for numeric predictions of biomarkers for metabolic challenges in dairy cows. *Preventive Veterinary Medicine* 193,105422

Martens, H. 2023. Invited Review: Increasing milk yield and negative energy balance: A Gordian Knot for dairy cows? *Animals* 13(19), 3097.

McArt, J.A.A., Nydam, D.V., Oetzel, G.R., 2013. Dry period and parturient predictors of early lactation hyperketonemia in dairy cattle. *Journal of Dairy Science* 96(1), 198-209.

Mekuriaw, Y., 2023. Negative energy balance and its implication on productive and reproductive performance of early lactating dairy cows: review paper. *Journal of Applied Animal Research* 51(1), 220-229.

Michalski, R.S., Carbonell, J.G., Mitchell, T.M., 2013. Machine learning: An artificial intelligence approach. *Artificial Intelligence* 25(2), 236-238.

Mohammadabadi, M., Kheyrodin, H., Afanasenko, V., Babenko Ivanivna, O., Klopenko, N., Kalashnyk, O., Ievstafiieva, Y., Buchkovska, V. 2024. The role of artificial intelligence in genomics. *Agricultural Biotechnology Journal* 16(2), 195-279.

Monshouwer, R. 2020. Detection of subclinical ketosis in dairy cows using behaviour sensor data. M.Sc. Thesis, University of Twente. Netherlands.

NASM, National Academies of Sciences and Medicine., 2021. Nutrient Requirements of Dairy Cattle: 8th Ed. Washington, D.C., USA. The National Academies Press.

Ospina, P.A., McArt, J. A., Overton, T. R., Stokol, T., Nydam, D.V., 2013. Using nonesterified fatty acids and β-hydroxybutyrate concentrations during the transition period for herd-level monitoring of increased risk of disease and decreased reproductive and milking performance. *Veterinary Clinics of North America: Food Animal Practice* 29(2), 387-412.

Ospina, P.A., Nydam, D.V., Stokol, T., Overton, T.R., 2010. Association between the proportion of sampled transition cows with increased non-esterified fatty acids and β-hydroxybutyrate and disease incidence, pregnancy rate, and milk production at the herd level. *Journal of Dairy Science* 93(8), 3595-3601.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12, 2825-2830.

Poczynek, M., Nogueira, L.S., Carrari, I.F., Carneiro, J.H., Almeida, R., 2023. Associations of body condition score at calving, parity, and calving season on the performance of dairy cows and their offspring. *Animals* 13(4), 596-604.

Ravelo, A.D., Vyas, D., Ferraretto, L.F., Faciola, A., 2022. Effects of sucrose and lactose as partial replacement to corn in lactating dairy cow diets: a review. *Translational Animal Science* 6(2).1-8.

Richards, B.F., Janovick, N.A., Moyes, K.M., Beever, D. E., Drackley, J.K., 2020. Comparison of prepartum low-energy or high-energy diets with a 2-diet far-off and close-up strategy for multiparous and primiparous cows. *Journal of Dairy Science* 103(10), 9067-9080.

Roche, J.R., Friggens, N.C., Kay, J.K., Fisher, M.W., Stafford, K.J., Berry, D.P., 2009. Invited review: Body condition score and its association with dairy cow productivity, health, and welfare. *Journal of Dairy Science* 92(12), 5769-5801.

Roche, J.R., Meier, S., Heiser, A., Mitchell, M.D., Walker, C.G., Crookenden, M.A., Riboni, M. V., Loor, J.J., Kay, J.K., 2015. Effects of precalving body condition score and prepartum feeding level on production, reproduction, and health parameters in pasture-based transition dairy cows. *Journal of Dairy Science* 98(10), 7164-7182.

Roche, S.M., Ross, J.A., Schatz, C., Beaugrand, K., Zuidhof, S., Ralston, B., Allan, N., Olson, M. 2023. Impact of dystocia on milk production, somatic cell count, reproduction and culling in Holstein dairy cows. *Animals* 13(3), 346-354.

Sahin, A., Ulutaş, Z., Adkinson, A., Adkinson, R., Bölümü, Z., Üniversitesi, G., Turkey, T., 2012. Genetic and environmental parameters and trends for milk production of Holstein cattle in Turkey. *Italian Journal of Animal Science* 11, 44-51.

Satoła, A., Bauer, E.A. 2021. Predicting Subclinical Ketosis in Dairy Cows Using Machine Learning Techniques. *Animals* 11(7), 2131.

Schapire, R.E., 2003. The Boosting Approach to Machine Learning: An Overview. MSRI Workshop on Nonlinear Estimation and Classification, Springer, New York.

Schillings, J., Bennet, R., Rose, C.R., 2021. Exploring the potential of precision livestock farming technologies to help address farm animal welfare. *Frontiers in Animal Science* 2, 639678.

Shine, P., Murphy, M.D., 2021. Over 20 Years of Machine Learning Applications on Dairy Farms: A Comprehensive Mapping Study. *Sensors* 22(1), 52-87.

Sun, Z., Samarasinghe, S., Jago, J., 2010. Detection of mastitis and its stage of progression by automatic milking systems using artificial neural networks. *Journal of Dairy Research* 77(2), 168-175.

Tajaddini, M.A., Dayani, O., Khezri, A., Tahmasbi, R., Sharifi-Hoseini, M.M. 2021. Production efficiency, milk yield, and milk composition and fatty acids profile of lactating goats feeding formaldehyde-treated canola meal in two levels of dietary crude protein. *Small Ruminant Research* 204, 106519.

Thomas, L., Krebs, C.J., 1997. A review of statistical power analysis software. *Bulletin of the Ecological Society of America* 78(2), 128-139.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. *Journal of Royal Statistical Society* 58, 267-288.

Wallén, S.E., Prestløkken, E., Meuwissen, T.H.E., McParland, S., Berry, D.P., 2018. Milk mid-infrared spectral data as a tool to predict feed intake in lactating Norwegian Red dairy cows. *Journal of Dairy Science* 101, 6232-6243.

Wei, Z., Zhang, B., Liu, J., 2018. Effects of the dietary nonfiber carbohydrate content on lactation performance, rumen fermentation, and nitrogen utilization in mid-lactation dairy cows receiving corn stover. *Journal of Animal Science and Biotechnology* 14(9), 20-26.

Weld, K.A., Oliveira, R.C., Bertics, S.J., Erb, S.J., White, H.M., 2020. Hepatic pyruvate carboxylase expression differed prior to hyperketonemia onset in transition dairy cows. *PLoS ONE* 15, 11-30.

White, C.H., Davis, N.G., Van Emon, M.L Wyffels, S.A., DelCurto, T., 2019. Impacts of increasing levels of salt on intake, digestion, and rumen fermentation with beef cattle consuming low-quality forages. *Translational Animal Science* (Suppl 1), 1818-1821.

White, H.M., 2015. The role of TCA cycle anaplerosis in ketosis and fatty liver in periparturient dairy cows. *Animals* 5(3), 793-802.