JAST (Journal of Animal Science and Technology) TITLE PAGE
Upload this completed form to website with submission

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Research article
Article Title (within 20 words without abbreviations)	Comparative Performance of Multiple Linear Regression and Machine Learning Models for Predicting Methane Emissions in Korean Holstein Cows
Running Title (within 10 words)	Prediction of Methane Emissions in Holstein Cows
Author	Inpyo Hong1, Kyunghwan Oh3, Chungsu Lee1, Jaedon Oh1,2, Hongsik Kong1,2, Sungdae Lee4, Eunjoong Kim5, Sanggun Roh6, Wonseob Kim7, Honggu Lee3, Yoonseok Lee1,2
Affiliation	1 School of Biotechnology, Hankyong National University, Anseong 17579, Republic of Korea 2 Center for Genetic Information, Hankyong National University, Anseong 17579, Republic of Korea 3 Department of Animal Science and Technology, Sanghuh College of Life Sciences, Konkuk University, Seoul 05029, Republic of Korea 4 Animal Nutrition and Physiology Division, National Institute of Animal Science, Rural Development Administration, Wanju 55365, Republic of Korea 5 Department of Animal Science, Kyungpook National University, Sangju 37224, Republic of Korea 6 Graduate School of Agricultural Science, Tohoku University, Sendai 980-8572, Japan 7 Department of Animal Science, College of Health and Bio Science, Daegu University, 201 Daehak-ro, Jinryang-eup, Gyeongsan-si, Gyeongsangbuk-do, 38453, Republic of Korea
ORCID (for more information, please visit https://orcid.org)	Inpyo Hong (0009-0008-8013-6317), Kyunghwan Oh (0009-0001-9413-053X) Hong-gu Lee (0000-0002-0679-5663) Yoonseok Lee (0000-0003-4803-5192)
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	This work was supported by the Rural Development Administration (RDA) of the Korean Government (Grant No. G24002318892).
Acknowledgements	This research was supported by the Rural Development Agency (RDA) of the Korean Government (Central Executive Agency) (No. G24002318892).
Availability of data and material	All the data generated in this study have been included in the manuscript. For further access, please contact the corresponding author, Hong-gu Lee.
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Lee YS Supervision: Lee YS Writing – original draft: Hong IP† Formal analysis: Hong IP†, Oh KH†, Lee CS Methodology: Lee CS, Oh KH† Data curation: Oh KH† Investigation: Oh KH†, Lee CS Writing – review & editing: Kim WS, Oh JD, Kong HS, Lee SD, Kim EJ, Roh SG Corresponding authors: Lee HG*, Lee YS*

	† Hong IP and Oh KH contributed equally to this work. * Lee HG and Lee YS contributed equally to this work.
Ethics approval and consent to participate	All experimental procedures in this study were approved by the Institutional Animal Care and Use Committee (IACUC) of Konkuk University (Permit numbers: KU23156, KU23234)

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	**Fill in information in each box below**
First name, middle initial, last name	Yoonseok Lee, Honggu Lee
Email address – this is where your proofs will be sent	yoonseok95@hknu.ac.kr, hglee66@konkuk.ac.kr
Secondary Email address	-
Address	Hankyong National University, 327 Jungang-ro, Anseong-si, Gyeonggi-do 17579, Republic of Korea, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Republic of Korea
Cell phone number	+82-10-3814-8660, +82-10-3307-4387
Office phone number	+82-31-670-5332 +82-2-450-0523
Fax number	Not applicable

Abstract

The linear mixed-effects models demonstrated moderate predictive performance, with adjusted R^2 values ranging from 0.811 to 0.820 and CCC values around 0.900. The inclusion of animal ID as a random effect appropriately accounted for the repeated-measures structure of the dataset and explained additional between-animal variation, as indicated by the estimated variance components and ICC values. The relatively small difference between marginal and conditional R^2 suggests that most of the predictive ability originated from the fixed effects, particularly the CH4/CO2 ratio, whereas repeated measurements within animals contributed a smaller but non-negligible proportion of variance. This study aimed to develop and compare multiple linear regression (MLR) and machine learning (ML) models, including scikit-learn linear regression and artificial neural networks (ANN), to predict methane emissions from Holstein dairy cows in Korea. Methane emissions and associated variables, including body weight, dry matter intake, energy-corrected milk, and CH_4/CO_2 ratio, were measured using the GreenFeed system. The MLR models demonstrated moderate predictive performances, with adjusted R^2 values ranging from 0.811 to 0.820. Machine-learning models, particularly ANN models, outperformed MLR, achieving higher adjusted R^2 values (up to 0.839), lower root mean square errors (RMSE), and higher concordance correlation coefficients (CCC). Among the ANN models developed Models 4 and 5 showed the best predictive performance, with CCC values of 0.925 and 0.929, respectively. These results highlight the potential of artificial intelligence approaches to predict enteric methane emissions in dairy cattle accurately. Future studies should focus on expanding the datasets and validating the models in diverse production environments to enhance their practical applicability in sustainable livestock management and greenhouse gas mitigation strategies.

Keywords : Methane emissions, GreenFeed system, Multiple linear regression, Artificial neural network, Korean Holstein cattle

Introduction

Atmospheric greenhouse gases, including carbon dioxide (CO_2), chlorofluorocarbons (CFCs), methane (CH_4), tropospheric ozone, and nitrous oxide (N_2O) have contributed significantly to climate change over the past half-century[1]. The beef and dairy industries account for approximately 14.5% of global greenhouse gas (GHG) emissions, with enteric methane (CH_4) emissions contributing nearly 40% of this total[2]. Thus, understanding and reducing livestock methane emissions is essential to meeting the GHGs mitigation targets outlined by the UN Framework Convention on Climate Change. In Korea, recent Net Zero Emissions by 2050 (NZE Scenario) have emphasized the importance of reducing livestock methane emissions[3,4].

Among livestock, Ruminal methane emissions are a normal process that occurs during the microbial fermentation of complex carbohydrates in the rumen. The first step in reducing these emissions is to accurately measure methane output from dairy cows. Methane emissions have been evaluated using both direct measurement and modeling approaches. Among direct measurement, there are two methods to direct measurement techniques: respiration chambers and the GreenFeed (GF) system[5].

Respiration chambers provide accurately and precisely measurement of CH_4 and CO_2 by enabling continuous monitoring under a controlled environment. They eliminate external influences, making them useful for studying the effect of reducing methane on feed additives or nutrition. However, these methods are expensive, labor-intensive, and impractical for large-scale use. In contrast, GF system, developed by C-Lock Inc. provide a portable and less intrusive alternative. It measures ruminant methane emissions during natural behavior, allowing scalability and flexibility for field trials[6,7]. By collecting and analyzing exhaled breath, the GF system provide reliable and repeatable measurements in field conditions[8,9,10]. Compared with respiration chambers, the GF system is more scalable and suitable for long-term and large-scale monitoring of methane emissions in various environments[11,12].

Model-based approaches rely on equations incorporating feed components, feed intake, and animal information. Ghilardelli (2022) developed a equations for predicting CH_4 emissions with volatile fatty acid production derived from fermentation tests of in vitro system based on the rumen fluid[13]. Other equation have used mid-infrared vibrational spectroscopy or large dataset to estimate methane emission[14,15,16].

Several studies have developed equations to estimate methane emissions based on multiple linear regression (MLR) and meta-analyses. These equations incorporate factors such as nutrient intake, digestibility, gross energy in-take (GEI), live weight (LW), energy-corrected milk (ECM), the CH_4/CO_2 ratio, and dry matter intake (DMI)[17,18]. Generally, ruminal CH_4 is directly produced by methanogens in the rumen, while CO_2 is a major by-product of fermentation, generated in proportion to volatile fatty acid production. Because CH_4 and CO_2 are produced simultaneously in fermentation pathways, their ratio reflects the metabolic balance in the rumen. This ratio normalizes methane output against total gas exchange, making it less sensitive to short-term fluctuations. In addition, Suzuki et al(2021) have shown that CH_4 and CO_2 is a reliable predictor of methane emissions in dairy cattle.

Multiple linear regression (MLR) analysis is a widely used modeling approach for evaluating CH_4 emissions. A challenge in implementing the MLR technique is that these assumptions may not always be fulfilled, which may lead to biased results and failure to provide satisfactory predictions. However, ML algorithms are beneficial when handling non-linear and complex datasets without any prior assumptions, even if the datasets are noisy and imprecise. This includes ML algorithms such as artificial neural networks (ANN).

Therefore, the aim of this study is to construct multiple linear regression (MLR) and machine learning (ML) models based on CH_4/CO_2 ratio, and to evaluate their predictive performance by comparing predicted methane emission estimate from traditional MLR with those from ML models, using GF system data collected from Holstein cows in Korea.

Materials and Methods

Measurement of methane and carbon dioxide emissions

Methane (CH_4) and carbon dioxide (CO_2) emissions from dairy cows were individually measured using a GF system (C-Lock Inc., Rapid City, SD), installed in a free-stall barn. The GF system is designed to collect and measure gases emitted by individual animals, which are identified using radio frequency identification (RFID) tags during animal visits. Prior to gas measurements, adaptation training was conducted to familiarize the animals with the GF system, specifically to maintain proper head positioning within the GF system. The dust filters were replaced when the airflow rate decreased to 27 L/s. The accuracy of the system was calibrated before the experiment using a CO_2 recovery test, with an average recovery rate of 104.2 ± 5.99%.

Gas measurements were conducted by luring the cows into the site using pelleted concentrate as bait. The exhaled gases were collected and filtered for dust, and the airflow rate was measured. The collected gas samples were then analyzed using a non-dispersive infrared sensor to determine the concentrations of CH_4 and CO_2. Gases exhaled by individual animals were sampled at 9-hr intervals over a four-day period, resulting in a total of eight measurements per animal. During each measurement, 100 to 150 g of the concentrate was pre-measured and placed in the feed bin to facilitate the entry of individual animals. Upon entry, the RFID tag of each animal was identified. Each gas measurement lasted for a minimum of 5 min per animal and the entry and exit times for all animals were recorded. A waiting period of 2-3 min was provided between measurements to avoid data overlapping. The remaining feed was weighed after each measurement to adjust dry matter intake. Calculated methane and CO_2 emissions (g/d) data were transferred to the C-Lock server, where the data were processed and presented to the laboratory.

All experimental procedures in this study were approved by the Institutional Animal Care and Use Committee (IACUC) of Konkuk University (Permit numbers: KU23156, KU23234). All methods were carried out in accordance with the relevant guidelines and regulations. This study is reported in accordance with the ARRIVE guidelines (https://arriveguidelines.org).

Data description and processing

Individual feed intake was measured over a 4-day period. The total mixed ration (TMR) offered and refusals were weighed and recorded daily to estimate individual feed intake. The total dry matter (DM) intake was calculated by adding the concentrate supplied during gas measurements with the GF system, along with the additional concentrate provided by the TMR. The DM content was corrected by oven drying (60℃ for 48h) samples of TMR and concentrate. The TMR was offered after morning milking, and a concentrate was provided during each gas measurement. Feed refusals were weighed individually on the day after the feed was offered.

The cows were milked twice daily (07:00 and 16:00) in a tandem milking parlor, and individual milk yields were recorded during each milking. Milk samples were obtained by mixing milk from the morning milking (50%) and evening milking (50%), resulting in a total of 40 mL of milk sampled daily. After mixing, samples were stored at 4 ℃ with the addition of milk preservative (Broad Spectrum Microtabs IITM, Advanced Instruments, MA, USA). Milk component analysis was carried out using an automatic milk analyzer (CombiFoss 7 DC, FOSS, Hillerød, Denmark), including fat and protein percentages obtained. Energy-corrected milk (kg/d) was calculated using Equation 1 (Tyrrell & Reid, 1965) as follows:

$$\text{ECM} = \text{Milk yield (kg/day)} \times (376 \times \text{Milk Fat (\%)} + 209 \times \text{Milk Crude protein (\%)} + 948) / 3138 \qquad (1)$$

The data used to develop the model for estimating cow methane emissions were obtained from the GF system, and various physiological indicators, including daily building intake and body weight, were collected during the measurement period. Following examination of the descriptive statistics of the collected data, histograms and normality tests were conducted to identify the characteristics of the dataset. In addition, outliers were selected for removal. Because the range of features (variables) in the dataset can vary, which can result in suboptimal model performance, it is recommended that the data be normalized to enhance the learning efficiency of the ANN model. In this study, all data points in the ANN model were normalized to the interval (0, 1) using the min-max normalization expression. Normalization was performed by applying min-max normalization using Equation 2.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (2)$$

Here, X_{norm} represents the normalized value, X signifies the existing value, X_{min} denotes the minimum value of the dataset, and X_{max} indicates the maximum value of the dataset. A normalized dataset was employed to train the model.

Feature selection is a crucial phase in building a predictive model, as it can significantly impact its predictive capacity. Feature selection was conducted by examining the correlation between variables and implementing a variance inflation factor (VIF) analysis. The correlation between variables was evaluated using Pearson's correlation coefficient and represented as a heat map using the Python Seaborn Library. In instances where the correlation between variables was high, less important variables were removed to minimize the adverse effects on model performance. Furthermore, a VIF analysis was conducted to identify instances of multicollinearity and ascertain the extent of the correlation between the variables. For each feature, a VIF score was calculated, and features with values exceeding a predefined threshold were excluded. The threshold for the VIF analysis was set to 10, and variables with VIF scores below this threshold were selected. The VIF scores were calculated using statsModel in python.

Multiple linear regression and machine learning

Because the dataset consisted of repeated measurements collected from the same animals, methane prediction equations were developed using linear mixed-effects models (LMMs) to account for within-animal correlation. Fixed effects included CH4/CO2 ratio, BW, ECM, and DMI according to the feature selection results, whereas animal ID was included as a random intercept effect. Models were fitted using the lme4 package in R. Missing observations were excluded prior to model fitting. Model performance was evaluated using adjusted marginal R^2, adjusted conditional R^2, root mean square error (RMSE), and concordance correlation coefficient (CCC). In addition, random-effect variance, residual variance, and the intraclass correlation coefficient (ICC) were estimated to quantify the contribution of between-animal variation.

We developed an estimation equation based on a mixed model to estimate methane emissions. The fixed effects used in the model were ECM, DMI, CH_4/CO_2 ratio, and BW, which introduced the results of the feature selection process. The random effect was determined as an individual. The Lme4 module in R was used to develop the mixed model.

Artificial intelligence models have been employed to enhance models developed using conventional statistical techniques. These models were developed using MLR and ANN models. The data were split into training and test sets to train the AI models for the selected variables. The data were divided into two distinct sets: a training set and a test set. The training set was further divided into ten subsets that were used to establish the validation set. The ratio of the training set to the test set was 8:2. train_test_split module in scikit-learn was employed to partition the data, and the scikit-learn k-fold function was used for a 10-fold cross-validation. The AI model was calibrated using MLR in scikit learn and Keras in Tensorflow, and hyperparameter optimization was conducted prior to model fitting.

Hyperparameter tuning of the ANN model was conducted using Kerastuner in TensorFlow and a candidate-based validation method. The numbers of dense layers, nodes per layer, epochs, and batch sizes were adjusted. In Kerastuner, the number of dense layers was set between one and six, and the number of nodes per layer was set between 16 and 512 in increments of 16. The number of epochs was set between five and 500, and the batch size was set between five and 35 in increments of five.

Model evaluation

The adjusted R^2, root mean squared error (RMSE), and concordance correlation coefficient (CCC) were employed as the model evaluation metrics for the MLR, ML, and ANN models. Adjusted R^2 is an indicator used to evaluate the explanatory power of a model in a regression analysis. The performance of the model was evaluated by considering the number of variables. The RMSE was calculated by squaring the prediction errors, averaging them, and taking the square root. A smaller RMSE value indicates that the model prediction is closer to the actual value. The CCC is a statistical metric that evaluates the agreement between two variables and is used to assess how well a dataset of predicted and measured values matches. By comparing the three metrics, the performances of the MLR, ML, and ANN models were evaluated.

Results

We used data from the automatic milking system (AMS) to construct MLR and ML models in this study. Similarly, Suzuki et al(2021) developed MLR model by using CH_4/CO_2 ratio, body weight (BW), dry matter intake (DMI), and energy-corrected milk (ECM) collected from AMS in dairy cattle in Japan. Based on this study, we used four features as input variables, then constructed MLR and ML models to evaluate their predictive performance for methane emissions in Holstein cows in Korea.

The first step was to develop methane prediction equations using linear mixed-effects models implemented in the lme4 package in R, followed by the construction of machine-learning models using scikit-learn and artificial neural network libraries.

Schematic overview of the analysis workflow for the development of MLR and ML models is presented in Figure 1. To compare the model performance, predicted equations for both MLR and ML models were developed. The first step is to develop the predicted equation of MLR using the 'lme4' option of the R program, and the second is to construct the model using Scikit-learn and the artificial neural network library for machine learning

Descriptive statistics and feature selection

A knowledge-based approach was used to relevant features. As reported by Suzuki et al. (2021), BW, DMI, ECM, and CH_4/CO_2 ratio were effective predictor of methane emissions in lactating Holstein cows. Accordingly, these four features were adopted for model construction in this study.

A total of 352 observations (44 cows with eight repeated measurements per animal) were collected using the GF system. After excluding observations containing missing values, 346 observations were used for mixed-model analyses.

Descriptive statistics of the CH_4/CO_2 ratio, body weight (BW), DMI, and ECM are presented in Table 1. Total three hundred fifty-two data of each features were collected for each feature using the GF system. As shown in Table 1, The mean ± SD value were 343.484 ± 94.618 g/day for CH4 emissions, 0.023 ± 0.005 for CH_4/CO_2 ratio, 633.496 ± 49.660 kg for BW, 25.732 ± 2.402 kg/day for DMI, and 22.981 ± 5.045 for ECM.

Based on previous study reported by Suzuki et al (2021), BW, DMI, ECM, and CH_4/CO_2 ratio features were used in this study to develop the MLR and machine learning models. The correlation coefficients and variance inflation factors of selected four features (VIF) are shown in Figure 2. As shown in Figure 2A, significant ($p < 0.05$) correlations were observed between BW and DMI ($r = 0.432$), BW and ECM ($r = 0.526$), and DMI and ECM ($r = 0.341$). The VIF of the four features are presented in Figure 2B. VIF value are an evaluation index used to determine the reliability of the model as a measure of multicollinearity in MLR model development. While feature with VIF < 5 are generally preferred, we applied a threshold of < 10 to ensure stability and interpretability of the models. As shown in Figure 2B, the CH_4/CO_2 ratio had the lowest VIF score (3.19) among those features, while BW, DMI, and ECM had VIF values of 5.53, 5.36, and 6.39, respectively. Accordingly, we selected these four features for developing the MLR and ML models.

Development of MLR for prediction of methane emissions

We first developed MLR model to compare the performances of the MLR and ML, the initial step was to construct an MLR model with BW, DMI, CH_4/CO_2, and ECM. The regression equations derived from BW, ECM, DMI, and CH_4/CO_2 ratio to predict methane emissions in Holstein cows, measure with the GF system, are shown in Table 2.

Model performance was evaluated using adjusted marginal and conditional R^2, RMSE, and concordance correlation coefficient (CCC), with particular emphasis on CCC as an indicator of agreement between predicted and observed values. When BW, ECM, and CH4/CO2 ratio were used as independent variables, the model had an adjusted R^2 of 0.811, an RMSE of 38.7, and a CCC of 0.906. When DMI was added as a fourth variable, the adjusted R^2 increased to 0.820 and RMSE decreased to 37.7. However, the CCC value declined slightly to 0.900. These results indicate that accounting for repeated measurements at the animal level provided additional explanatory power while maintaining high predictive agreement between observed and predicted methane emissions.

Model performance was evaluated using Adjusted R^2, RMSE, and concordance correlation coefficient (CCC), with particular emphasis on CCC as an indicator of agreement between predicted and ovserved values. When BW, ECM, and CH_4/CO_2 ratio were used as independent variables, the model had and Adjusted R^2 of 0.811, an RMSE of 38.7, and a CCC of 0.906. When DMI was added as a fourth variable, the Adjusted R^2 value increased 0.820, the RMSE value decreased 37.7. However, the CCC value declined slightly to 0.900. indicating weaker despite improvement in Adjusted R^2 and RMSE. This result highlights the importance of CCC as a robust criteria for assessing predictive reliability, beyond conventional fit statistic.

Development of ML for prediction of methane emissions

Next, we developed ML models to evaluate whether they could outperform traditional MLR model. These models included scikit-learn linear regression and artificial neural network algorithms trained with three and four input features. The scikit-learn model was implemented using ‘LinearRegression’ option, while ANN model were trained using the same feature and adjusted parameters through hyperparameter tuning (learning rate, batch size, number of folds for cross-validation, regularization strength (alpha), and number of layers).

Models performance was evaluated using the adjusted R^2, RMSE, and CCC values, with particular emphasis on CCC as an indicator of agreement between predicted and observed methane emissions. As shown in Table 3, changes in the adjusted parameters influenced the adjusted R^2, RMSE, and CCC values of the ANN models. In particular, Model 1 exhibits the highest adjusted R^2 and RMSE values, whereas

Model 4 exhibits the lowest RMSE value. With three features, ANN model 1 consisted of one input layer (3), two hidden layers (49-49), and one output layer (1), trained with 10-fold cross-validation, a regularization strength of 0.001, a learning rate of 0.1, and a batch size of 64. With four features, ANN model4 consisted of one input layer (4), one hidden layer (1), and one output layer (1), trained with 5-fold cross-validation, a regularization strength of 0.01, a learning rate of 0.1, and a batch size of 64.

Among the ANN models, if the highest adjusted R^2 was prioritized, Models 2 and 3 were optimal. However, when RMSE and CCC were considered, Model 4 and 5 performed best. Model 4 achieved the lowest RMSE (35.310) and slightly higher CCC (0.925), while Model 5 had a slightly higher RMSE (35.810) but the highest (0.929). Although Models 2 and 3 (BW, ECM, CH_4/CO_2, and DMI) showed the highest adjusted R2, Models 4 and 5 were considered as optimal because their lower RMSE and higher CCC indicated greater prediction accuracy and consistency. The comparative performances of the Scikit-learn and ANN models for predicting methane emissions using three and four features are summarized in Table 4.

For the three-feature model, the scikit-learn model had an adjusted R^2 of 0.853, an RMSE of 33.523, and a CCC of 0.926, while The ANN Model 1 showed a slightly lower adjusted R^2 of 0.847 and a higher RMSE of 38.827, but a comparable CCC of 0.928. These results suggest that although scikit-learn provided better model fit and lower prediction error, the ANN model exhibited slightly stronger agreement with observed methane emissions.

For the four-feature models, The Scikit-learn model achieved an adjusted R^2 of 0.820, RMSE of 35.980, and CCC of 0.913. ANN Model 4 improved the performance with an adjusted R^2 of 0.839, RMSE of 35.310, and CCC of 0.925, while ANN Model 5 showed an adjusted R^2 of 0.830, RMSE of 35.810, and the highest CCC among the four feature models (0.929).

Taken together, the inclusion of an additional feature (DMI) slightly improved the ANN model performance, particularly in terms of CCC. while scikit-learn regression performed better with three , ANN models consistently showed higher CCC values across both feature sets, indicating their robustness in predicting methane emissions.

Comparison of predicted performance between MLR and ML models

When using three features (BW, ECM, and CH_4/CO_2), the Scikit-learn ML model showed stronger agreement with observed methane emissions than MLR, with a higher CCC (0.926 vs. 0.900), along with improved Adjusted R^2 (0.853 vs. 0.811) and lower RMSE (33.523 vs. 38.738). The ANN Model1 demonstrated a strong performance, achieved a comparable Adjusted R^2 of 0.847 and the highest CCC value of 0.928 among the three-feature models.

When DMI was added as a fourth feature, slight improvements in prediction accuracy were observed across all models. The ANN Model4 had the lowest RMSE (35.310) and highest CCC (0.925), while ANN Model5 showed the highest CCC (0.929) among all the models tested.

Taken together, these results indicate that ML models, particularly Artificial Neural Networks, outperformed MLR in predicting methane emissions and that the inclusion of DMI as an additional feature contributed to enhanced model performance.

Discussion

In the present study, we compared the predictive performance of MLR and machine learning (ML) models, including scikit-learn and ANN algorithms, for estimating methane emissions from Holstein dairy cows in Korea. Based on results reported by Suzuki et al. (2021), methane emissions were measured using the GF system, and body weight (BW), dry matter intake (DMI), energy-corrected milk (ECM), and the CH_4/CO_2 ratio were used as input data for model development.

The linear mixed-effects models demonstrated moderate predictive performance, with adjusted R^2 values ranging from 0.811 to 0.820 and CCC values around 0.900. The inclusion of animal ID as a random effect appropriately accounted for the repeated-measures structure of the dataset and explained additional between-animal variation, as indicated by the estimated variance components and ICC values. The relatively small difference between marginal and conditional R^2 suggests that most of the predictive ability originated from the fixed effects, particularly the CH4/CO2 ratio, whereas repeated measurements within animals contributed a smaller but non-negligible proportion of variance.

The MLR models demonstrated moderate predictive performance, with adjusted R^2 values of 0.811-0.820 and CCC values around 0.900. These results are consistent with previous studies[16], which reported similar predictive performances using comparable variables. Notably, the CH_4/CO_2 ratio proved to be an indicator, reflecting its strong physiological relationship with enteric methane production and its ability to normalize short-term fluctuations in gas exchange[19]. However, because the CH4/CO2 ratio is directly associated with methane production during rumen fermentation, models including this variable should primarily be interpreted as calibration models based on breath gas composition rather than fully independent methane prediction models. Therefore, caution should be taken when interpreting model performance based on this variable. In contrast, relatively weak correlations were observed between CH4 emissions and BW, DMI, and ECM, suggesting that multivariate models are necessary to capture complex interactions affecting methane production.

The machine learning models developed with the scikit-learn linear regression algorithm outperformed MLR, particularly in terms of CCC. With three features (BW, ECM, and CH_4/CO_2 ratio), the Scikit-learn model had an adjusted R^2 of 0.853 and RMSE of 33.523, outperforming the MLR model (adjusted R^2 = 0.811, RMSE = 38.738). Similarly, with four features (BW, ECM, DMI, and CH_4/CO_2 ratio), the Scikit-learn model maintained an adjusted R^2 of 0.820 and slightly lower RMSE compared to MLR, suggesting that even with simple linear modeling, ML frameworks can optimize predictive capacity through improved data handling and regularization techniques inherent to modern libraries.

Among all the evaluated models, the ANN model demonstrated the best overall predictive performance. The ANN models had adjusted R^2 values of up to 0.839 and CCC values of up to 0.929. ANN Model 4, which was developed with four input datasets, showed the lowest RMSE (35.310) and ANN Model 5 had the highest CCC (0.929), indicating excellent agreement between the predicted and observed values. This superior performance is likely due to the ability of ANN models to capture complex nonlinear relationships between predictor variables and methane emissions that conventional linear models may not fully represent. However, it should also be noted that the improvement in predictive performance between the ANN and MLR models was relatively modest. For example, the increase in adjusted R^2 from the best MLR model (0.820) to the best ANN model (0.839) was approximately 1.9%. Although ANN models showed slightly higher predictive accuracy, simpler and more interpretable models such as MLR may still be advantageous for practical on-farm applications where model transparency and ease of implementation are important considerations.

Additionally, hyperparameter tuning—such as adjusting the number of dense layers, nodes per layer, batch size, and learning rate—plays a crucial role in optimizing ANN models. As demonstrated in this study, hyperparameter optimization was crucial for improving model generalization and minimizing overfitting or underfitting, which is consistent with previous reports[20,21].

Because the dataset consisted of repeated measurements obtained from 44 individual cows, linear mixed-effects models with animal ID as a random effect were implemented throughout the regression analyses to appropriately account for within-animal correlation. Nevertheless, future studies using larger and more diverse populations are needed to improve model generalizability and external validity.

Taken together, these findings emphasize that while MLR and Scikit-learn models offer straightforward and interpretable predictive tools with moderate accuracy, ANN models provide a higher level of predictive precision by effectively capturing complex patterns in the data. One important limitation of the present study relates to the structure of the dataset. The dataset consisted of repeated measurements collected from 44 individual cows, and machine-learning models were trained using a random train–test split. As a result, observations from the same cow could potentially appear in both training and testing datasets, which may lead to optimistic estimates of model performance. To partially address this issue, we also implemented a linear mixed-effects model that explicitly accounted for individual animal variation by including animal ID as a random effect. Nevertheless, future studies using larger datasets should apply cow-level data partitioning strategies to further improve model generalizability. This highlights the potential of artificial intelligence-based approaches to meaningfully contribute to methane mitigation strategies and sustainable livestock management.

Future studies could benefit from incorporating additional biological and environmental features, such as rumen microbial profiles, feed composition, and animal bioactivity, to further enhance model robustness. Moreover, applying the developed models across different breeds, production systems, and geographic conditions is essential for validating their applicability on a larger scale.

The present results demonstrate that both MLR and machine learning models can effectively predict methane emissions in Holstein dairy cows using variables such as BW, DMI, ECM, and CH4/CO2 ratio. The MLR models showed reasonable predictive performance (adjusted R^2 of 0.811–0.820), providing a useful basis for methane prediction under field conditions. However, the ML models, particularly the ANN models, showed superior predictive ability, achieving higher adjusted R^2 values (up to 0.839), lower RMSE, and higher CCC compared to both traditional MLR and Scikit-learn models. These findings highlight the potential of ANN-based approaches for better capturing the complex nonlinear relationships govern ruminal methane emissions in dairy cattle.

Among the developed ANN models, Model 4 and Model 5 demonstrated the best overall predictive performance. ANN Model 4, using BW, DMI, ECM, and CH_4/CO_2 ratio as inputs, achieved an adjusted R^2 of 0.839, an RMSE of 35.310, and a CCC of 0.925. ANN Model 5 showed a slightly lower adjusted R^2 (0.830), but achieved the highest CCC value (0.929), indicating excellent agreement between the predicted and observed methane emissions. These results confirm that the ANN models, especially Models 4 and 5, offer significantly improved predictive accuracy compared with traditional linear methods.

Overall, the results suggest that while MLR models can serve as useful and interpretable tools for methane emissions prediction, ML models, particularly ANN models, provide greater accuracy, robustness, and potential applicability for prediction of ruminal methane emissions. However, given the relatively limited dataset size and the repeated-measures structure of the data, the results of the machine-learning models should be interpreted cautiously. Further validation using larger and more diverse datasets will be necessary to confirm the robustness and generalizability of these models. The successful application of artificial intelligence models in this study demonstrates their valuable role in developing advanced strategies for predicting and mitigating methane emissions in dairy cattle.

Future research should focus on expanding the datasets, incorporating additional biological and environmental variables, and validating the model performance across different breeds, production systems, and geographic locations to enhance the model generalizability and real-world applicability.

Acknowledgments

This research was supported by the Rural Development Agency (RDA) of the Korean Government (Central Executive Agency) (No. G24002318892).

References

1. Forster P, Ramaswamy V, Artaxo P, Berntsen T, Betts R, Fahey DW, et al. Changes in atmospheric constituents and in radiative forcing. In: Solomon S, Qin D, Manning M, editors. Climate change 2007: the physical science basis. Cambridge: Cambridge University Press; 2007. p. 747-845.
2. Gerber P, Steinfeld H, Henderson B, Mottet A, Opio C, Dijkman J, et al. Tackling climate change through livestock: a global assessment of emissions and mitigation opportunities. Rome: Food and Agriculture Organization of the United Nations; 2013.
3. Ministry for the Environment. Actions on agricultural emissions: a discussion document on proposals to address greenhouse gas emissions from agriculture [Internet]. Wellington (NZ): Ministry for the Environment; 2020 [cited 2025 Sep 25]. Available from: https://environment.govt.nz/publications/action-on-agricultural-emissions-a-discussion-document-on-proposals-to-address-greenhouse-gas-emissions-from-agriculture/
4. Ministry for the Environment. Climate change response (Zero Carbon) amendment act [Internet]. Wellington (NZ): Ministry for the Environment; 2019 [cited 2025 Sep 25]. Available from: https://environment.govt.nz/acts-and-regulations/acts/climate-change-response-amendment-act-2019/
5. Ghassemi Nejad J, Ataallahi M, Park KH, Sung KI, Lee HG. Advances in methane emission estimation in livestock: a review of data collection methods, model development and the role of AI technologies. Animals. 2024;14:435. https://doi.org/10.3390/ani14030435
6. Hristov AN, Oh J, Giallongo F, Frederick T, Harper MT, Weeks HL, et al. The use of an automated system (GreenFeed) to monitor enteric methane and carbon dioxide emissions from ruminant animals. J Vis Exp. 2015;103:52904. https://doi.org/10.3791/52904
7. Hammond KJ, Humphries DJ, Crompton LA, Bannink A, Dijkstra J, Yanez-Ruiz DR, et al. Methane emissions from cattle: estimates from short-term measurements using a GreenFeed system compared with respiration chambers or sulphur hexafluoride tracer. Anim Feed Sci Technol. 2015;203:41-52. https://doi.org/10.1016/j.anifeedsci.2015.02.008
8. Huhtanen P, Ramin M, Cabezas-Garcia EH. Repeatability and between-cow variability of enteric CH4 and total CO2 emissions. Adv Anim Biosci. 2013;4:588. https://doi.org/10.1017/S2040470013000485
9. Hammond K, Crompton LA, Bannink A, Dijkstra J, Yanez-Ruiz DR, OKiely P, et al. Methane (GreenFeed) compared emissions from growing dairy heifers estimated using an automated head chamber to respiration chambers or SF6 techniques. Adv Anim Biosci. 2013;4:447. https://doi.org/10.1017/S2040470013000345

10. Hammond KJ, Crompton LA, Bannink A, Dijkstra J, Yanez-Ruiz DR, OKiely P, et al. The GreenFeed system for measurement of enteric methane emission from cattle. Anim Prod Sci. 2016;56:181-189. https://doi.org/10.1071/AN15380
11. Huhtanen P, Cabezas-Garcia EH, Utsumi S, Zimmerman S. Enteric methane emission can be reliably measured by the GreenFeed monitoring unit. Livest Sci. 2019;222:31-40. https://doi.org/10.1016/j.livsci.2019.02.012
12. Ghilardelli F, Ferragina A, Morandi S, Cremonesi P, Daghio M, Fracassetti D, et al. Near-infrared calibration models for estimating volatile fatty acids and methane production from in vitro rumen fermentation of different total mixed rations. JDS Commun. 2022;3:19-25. https://doi.org/10.3168/jdsc.2021-0205
13. Shadpour S, Sargolzaei M, Miglior F, Schenkel FS, Baes CF, Fleming A. Predicting methane emission in Canadian Holstein dairy cattle using milk mid-infrared reflectance. J Dairy Sci. 2022;105:8272-8285. https://doi.org/10.3168/jds.2021-21464
14. Hemple S, Krause A, Ammon C, Amon B, Amon T. How the selection of training data and modeling approach affects the estimation of ammonia emissions from a naturally ventilated dairy barn classical statistics versus machine learning. Sustainability. 2020;12:1030. https://doi.org/10.3390/su12031030
15. Bi H, Neethirajan S. Mapping methane the impact of dairy farm practices on emissions through satellite data and machine learning. Climate. 2024;12:223. https://doi.org/10.3390/cli12020223
16. Suzuki T, Oshita T, Kobayashi N, Nonaka K, Sato S. Prediction of enteric methane emissions from lactating cows using methane to carbon dioxide ratio in the breath. Anim Sci J. 2021;92:13637. https://doi.org/10.1111/asj.13637
17. Suzuki T, Oshita T, Kobayashi N, Nonaka K, Sato S. Prediction of enteric methane emission from beef cattle in Southeast Asia. Anim Sci J. 2018;89:1287-1295. https://doi.org/10.1111/asj.13063
18. Chen X, Li C, Zhang Y, Liu H, Li J, Bao J, et al. Can machine learning algorithms perform better than multiple linear regression in predicting nitrogen excretion from lactating dairy cows. Sci Rep. 2022;12:12478. https://doi.org/10.1038/s41598-022-16666-5
19. Brask M, Weisbjerg MR, Hellwing ALF, Bannink A, Lund P. Methane production and diurnal variation measured in dairy cows and predicted from fermentation pattern and nutrient or carbon flow. Animal. 2015;9:1795-1806. https://doi.org/10.1017/S1751731115001184
20. Higaki S, Takahashi H, Kuroda T, Obuchi S, Nishizaki H, Suzuki T. Estrous detection by continuous measurements of vaginal temperature and conductivity with supervised machine

learning in cattle. Theriogenology. 2019;123:90-99. https://doi.org/10.1016/j.theriogenology.2018.09.035

21. Islam MM, Li Y, Zhang C, Mehmood K, Rehman MU, Nabi F, et al. Metabolic modeling elucidates the transactions in the rumen microbiome and the shifts. Front Microbiol. 2019;10:2412. https://doi.org/10.3389/fmicb.2019.02412
22. Davoudkhani M, Haghdoost NS, Koike S, Kobayashi Y. Towards the integration of microbial genomic information into mechanistic models of the rumen microbiome a theoretical study. J Anim Sci. 2022;100:178-179. https://doi.org/10.1093/jas/skac178.318
23. Danielsson R, Dicksved J, Sun L, Gonda H, Muller B, Schnurer A, et al. Methane production in dairy cows correlates with rumen methanogenic and bacterial community structure. Front Microbiol. 2017;8:226. https://doi.org/10.3389/fmicb.2017.00226
24. Hristov AN, Oh J, Giallongo F, Frederick T, Harper MT, Weeks HL, et al. The use of an automated system (GreenFeed) to monitor enteric methane and carbon dioxide emissions from ruminant animals. J Vis Exp. 2015;103:52904. https://doi.org/10.3791/52904

Tables and Figures

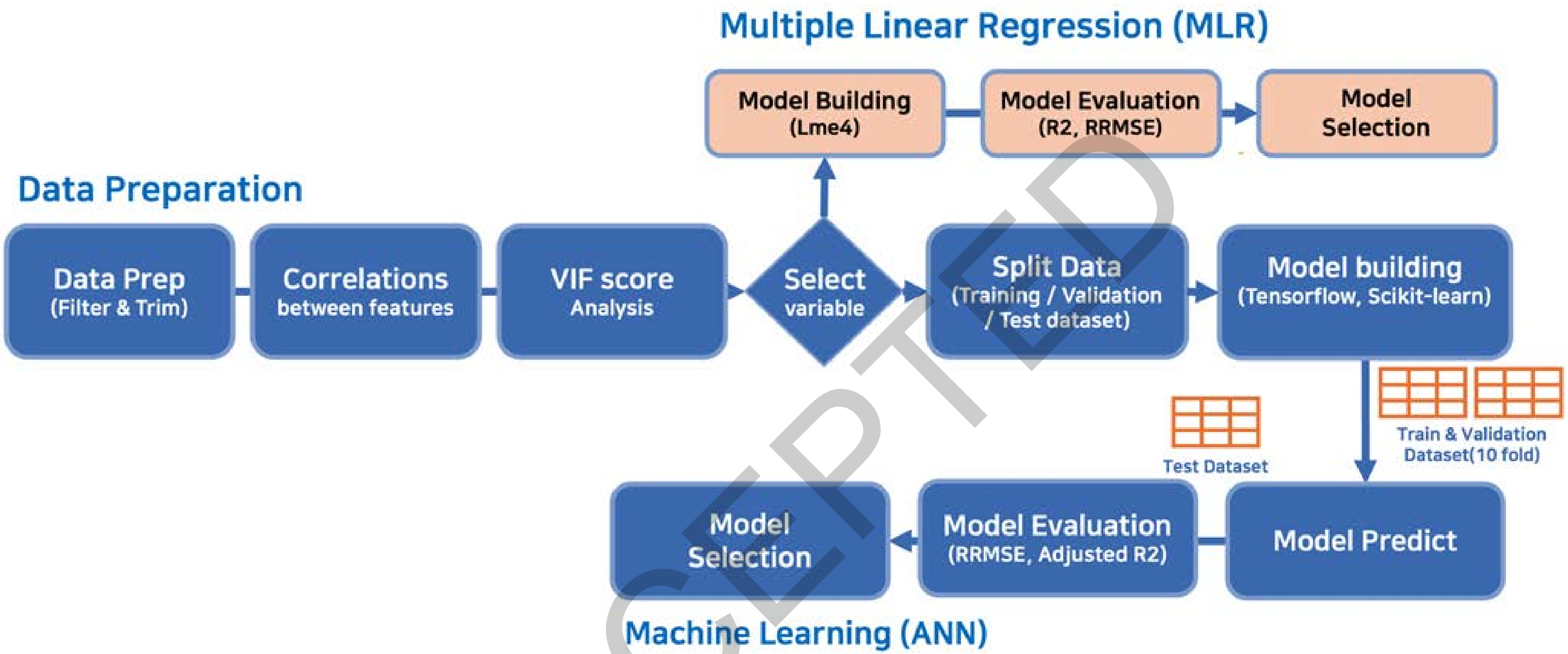

Figure 1. Schematic of analysis workflow for the development of MLR and ML models.

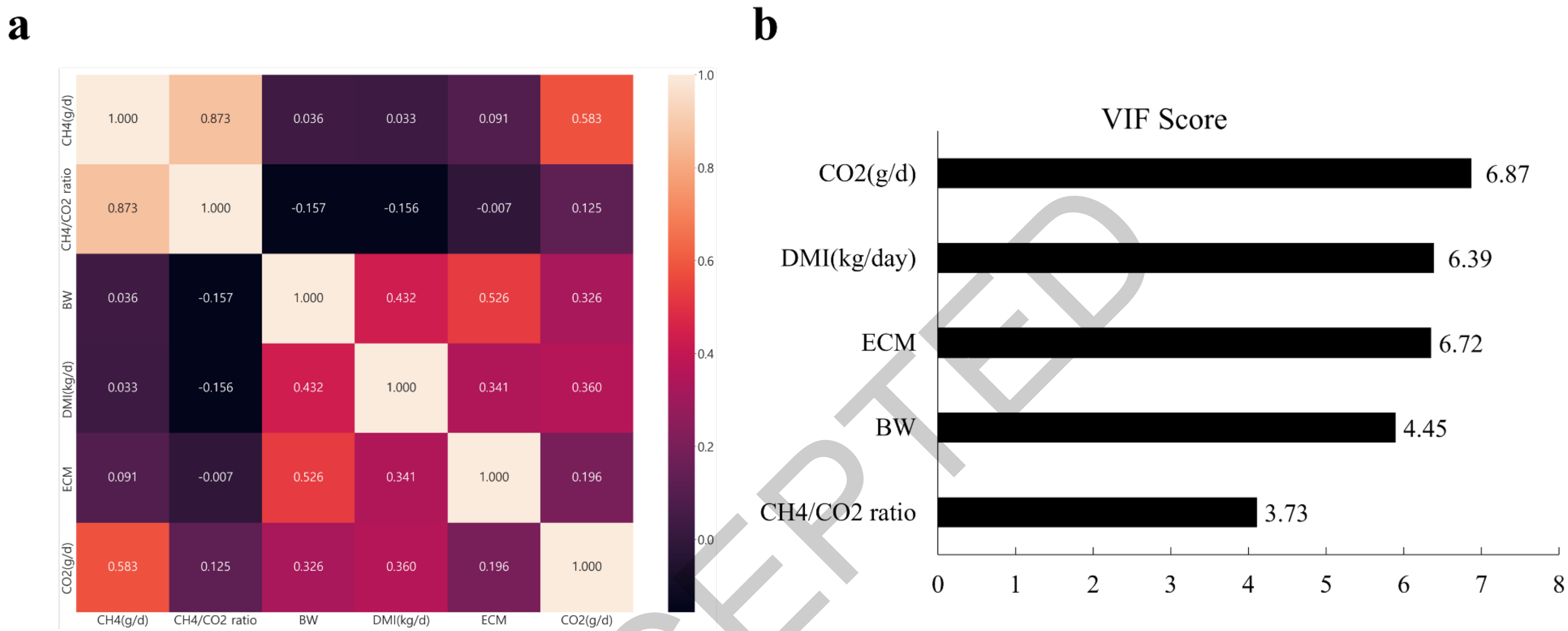

Figure 2. Correlation coefficient between features and variance inflation factors (VIF) of four features used in the study. (a) Correlation coefficient between features. (b) Variance inflation factors (VIF) of features. A VIF score was calculated for each feature and those with high values were removed. The threshold score for the VIF analysis was 10 and features with a VIF score be-low this threshold were selected. The VIF score was computed by VIF function in python.

Table 1. Descriptive statistics of five features used in this study.

	CH4 (g/d)	CH_4/CO_2*	BW (kg)*	DMI (kg/day)*	ECM*	CO_2(g/d)
Data no.	352	346	352	352	352	346
Mean	343.484	0.023	633.496	25.732	22.981	14834.839
SD	94.618	0.005	49.660	2.402	5.045	1977.571
Min	106.452	0.010	557.000	17.602	6.611	6483.103
Max	723.400	0.040	754.000	30.591	38.039	21799.246

* CH_4/CO_2, CH_4/CO_2 ratio; BW, body weight; DMI, dry matter intake; ECM, energy-corrected milk (kg/d), CO_2(g/d)

Table 2. Regression equations for ruminal methane emissions in Holstein cow

Equations for Holstein cow in Korea		Adjusted Marginal R^2	Adjusted Conditional R^2	RMSE	CCC	Random Intercept Variance	Residual Variance	ICC
CH4 (g/day) =								
$-275.301 + 0.353BW + 0.115ECM + 16{,}990.884CH_4/CO_2$	(1)	0.787	0.811	38.738	0.900	200.175	1568.956	0.113
$-364.919 + 0.266BW – 0.630ECM + 6.172DMI + 17{,}187.275CH_4/CO_2$	(2)	0.797	0.819	37.688	0.906	183.694	1490.404	0.110
$-334.788 + 0.025BW +0.038ECM – 1.094DMI +0.023CO_2 + 15283.633CH_4/CO_2$	(3)	0.991	0.991	2.489	0.996	1.606	75.909	0.021

BW, body weight; ECM, energy-corrected milk (kg/day); DMI, dry matter intake (kg/day); CH4/CO_2, CH_4/CO_2 ratio; CO_2, CO_2(g/d); RMSE, root mean square error (n=346; values in parentheses indicate standard error).

Table 3. Prediction accuracy and optimized hyperparameters of ANN models using different feature sets for predicting enteric methane emissions in dairy cattle.

Model	Feature	Alpha	Learning rate	Layer (node no.)	Batch size	Adjusted R^2	RMSE*	CCC*
Model 1	BW, ECM, CH4/CO2	0.01	0.001	3-16-1	8	0.770	43.485	0.876
Model 2		0.001	0.001	3-32-1	8	0.769	43.506	0.877
Model 3	BW, ECM, CH4/CO2, DMI	0.01	0.001	4-16-1	8	0.776	42.833	0.877
Model 4		0.001	0.001	4-32-1	8	0.768	43.555	0.876
Model 5	BW, ECM, CH4/CO2, DMI, CO2	0.01	0.001	5-16-1	8	0.999	2.658	1.000
Model 6		0.001	0.001	5-32-1	8	0.999	2.775	1.000

* RMSE = root mean square error, CCC = Concordance Correlation Coefficient.

Table 4. Comparative performance of scikit-learn and ANN models using three- and four-feature sets for predicting enteric methane emissions in dairy cattle..

Feature	ML Model*	Adjusted R^2	RMSE	CCC*
BW, ECM, CH_4/CO_2	Scikit-learn	0.767	43.723	0.872
	ANN(Model1)	0.770	43.485	0.876
BW, ECM, DMI, CH_4/CO_2	Scikit-learn	0.780	42.425	0.880
	ANN(Model3)	0.776	42.833	0.877
BW, ECM, DMI, CO_2, CH_4/CO_2	Scikit-learn	0.990	9.261	0.995
	ANN(Model5)	0.999	2.658	1.000

ML: machine learning; ANN: artificial neural network; RMSE: root mean square error; CCC = Concordance Correlation Coefficient.

Table 5. Comparison of the predicted performance between MLR and ML models for predicting methane emissions based on three and four features.

Equations	Model*	Adjusted R^2	RMSE*	CCC*
CH_4 = BW + ECM + CH_4/CO_2	LMM	0.811	38.738	0.900
	Scikit-learn	0.767	43.723	0.872
	ANN(Model1)	0.770	43.485	0.876
CH_4 = BW + ECM + DMI + CH_4/CO_2	LMM	0.820	37.688	0.906
	Scikit-learn	0.780	42.425	0.880
	ANN(Model3)	0.776	42.833	0.877
CH_4 = BW + ECM + DMI + CO_2 + CH_4/CO_2	LMM	0.991	2.489	0.996
	Scikit-learn	0.990	9.261	0.995
	ANN(Model5)	0.999	2.658	1.000

MLR, multiple linear regression; ANN, artificial neural network; RMSE, root mean square error; CCC=Concordance Correlation Coefficient.

--- Added clarification regarding statistical model structure ---

The dataset consisted of repeated measurements collected from the same animals. To account for this hierarchical structure, methane emissions were also analyzed using a linear mixed-effects

model implemented with the lme4 package in R. In this model, animal identity was included as a random effect to account for within-animal correlation among repeated observations.
Mixed-effects model equation:

$$CH4\_ij = \beta0 + \beta1(BW\_ij) + \beta2(DMI\_ij) + \beta3(ECM\_ij) + \beta4(CH4/CO2\_ij) + u\_i + \varepsilon\_ij$$

where i represents the individual cow and j represents repeated observations.

Machine learning validation strategy:
Machine-learning models were trained using a random train–test split (80:20) combined with k-fold cross-validation. Because the dataset included repeated measurements from the same animals, observations from the same cow could appear in both training and testing sets. Therefore, the machine-learning results should be interpreted primarily as exploratory comparisons of modeling approaches.

Artificial neural network implementation:
Artificial neural networks were implemented using the Keras deep-learning framework in TensorFlow. Hyperparameter tuning was performed using KerasTuner, which evaluated architectures with varying numbers of hidden layers, nodes per layer, epochs, and batch sizes.

Overfitting considerations:
Because the dataset consisted of only 44 independent animals, complex neural network architectures could lead to overfitting. Therefore, simpler architectures (e.g., a single hidden layer) were ultimately favored.

Final revisions incorporated according to Tables 2–5

Abstract/Results update: Linear mixed-effects models demonstrated moderate predictive performance (adjusted conditional $R^2 = 0.811–0.991$; CCC = 0.900–0.996). Adding DMI slightly improved prediction accuracy, whereas inclusion of CO_2(g/d) substantially improved model performance. Among machine-learning models, ANN models achieved adjusted R^2 values up to 0.999 and CCC values up to 1.000.

Results update: For the three-feature set (BW, ECM, and CH_4/CO_2 ratio), LMM, scikit-learn, and ANN models achieved adjusted R^2 values of 0.811, 0.767, and 0.770, respectively. After including DMI, adjusted R^2 values became 0.820, 0.780, and 0.776, indicating modest performance improvements.

Additional feature effect: The inclusion of CO_2(g/d) as an additional predictor markedly improved predictive performance. The five-feature models achieved adjusted R^2 values of 0.991 (LMM), 0.990 (scikit-learn), and 0.999 (ANN), with CCC values of 0.996, 0.995, and 1.000, respectively.

Discussion update: DMI likely improved predictive performance because feed intake reflects substrate availability for ruminal fermentation and methane formation. The substantial improvement observed after incorporating CO_2(g/d) is biologically plausible because CH_4 and CO_2 are simultaneously produced during ruminal fermentation and therefore share strong physiological relationships.