J. Dairy Sci. 105:5124–5140
https://doi.org/10.3168/jds.2021-20158

Integrating heterogeneous across-country data for proxy-based random forest prediction of enteric methane in dairy cattle

Enyew Negussie,[1]* Oscar González-Recio,[2] Mara Battagin,[3] Ali-Reza Bayat,[4] Tommy Boland,[5] Yvette de Haas,[6] Aser Garcia-Rodriguez,[7] Philip C. Garnsworthy,[8] Nicolas Gengler,[9] Michael Kreuzer,[10] Björn Kuhla,[11] Jan Lassen,[12] Nico Peiren,[13] Marcin Pszczola,[14] Angela Schwarm,[15] Hélène Soyeurt,[9] Amélie Vanlierde,[16] Tianhai Yan,[17] and Filippo Biscarini[18]

[1]Animal Genomics and Breeding, Natural Resources Institute Finland (Luke), 31600 Jokioinen, Finland
[2]Department of Animal Breeding, Instituto Nacional de Investigacion y Tecnologia Agraria y Alimentaria (INIA-CSIC), 28040 Madrid, Spain
[3]Italian Brown Cattle Breeders' Association, Verona, Italy
[4]Animal Nutrition, Natural Resources Institute Finland (Luke), 31600 Jokioinen, Finland
[5]Agriculture and Food Science Centre, School of Agriculture and Food Science, University College Dublin, Belfield, Belfield, Dublin 4, Ireland
[6]Animal Breeding and Genomics, Wageningen University and Research, 6700 AH Wageningen, the Netherlands
[7]Department of Animal Production, NEIKER—Basque Institute for Agricultural Research and Development, 01192 Arkaute, Spain
[8]School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD, United Kingdom
[9]TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, University of Liège, 5030 Gembloux, Belgium
[10]ETH Zurich, Institute of Agricultural Sciences, Universitaetstrasse 2, 8092 Zurich, Switzerland
[11]Research Institute for Farm Animal Biology (FBN), Institute of Nutritional Physiology "Oskar Kellner," Wilhelm-Stahl-Allee 2, 18196 Dummerstorf, Germany
[12]VikingGenetics, Ebeltoftvej 16, 8960 Randers, Denmark
[13]Institute for Agricultural and Fisheries Research (ILVO), Merelbeke, Belgium
[14]Department of Genetics and Animal Breeding, Poznan University of Life Sciences, Wołynska 33, 60-637 Poznan, Poland
[15]Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, PO Box 5003, 1432 Ås, Norway
[16]Productions in Agriculture Department, Walloon Agricultural Research Centre (CRA-W), BEL-5030 Gembloux, Belgium
[17]Livestock Production Science Branch, Agri-Food and Biosciences Institute, Hillsborough, Co. Down BT26 6DR, United Kingdom
[18]National Research Council, Institute of Agricultural Biology and Biotechnology (CNR-IBBA), Via Bassini 15, 20133 Milan, Italy

ABSTRACT

Direct measurements of methane (CH_4) from individual animals are difficult and expensive. Predictions based on proxies for CH_4 are a viable alternative. Most prediction models are based on multiple linear regressions (MLR) and predictor variables that are not routinely available in commercial farms, such as dry matter intake (DMI) and diet composition. The use of machine learning (ML) algorithms to predict CH_4 emissions from across-country heterogeneous data sets has not been reported. The objectives were to compare performances of ML ensemble algorithm random forest (RF) and MLR models in predicting CH_4 emissions from proxies in dairy cows, and assess effects of imputing missing data points on prediction accuracy. Data on CH_4 emissions and proxies for CH_4 from 20 herds were provided by 10 countries. The integrated data set contained 43,519 records from 3,483 cows, with 18.7% missing data points imputed using k-nearest neighbor imputation. Three data sets were created, 3k (no missing records), 21k (missing DMI imputed from milk, fat, protein, body weight), and 41k (missing DMI, milk fat, and protein records imputed). These data sets were used to test scenarios (with or without DMI, imputed vs. nonimputed DMI, milk fat, and protein), and prediction models (RF vs. MLR). Model predictive ability was evaluated within and between herds through 10-fold cross-validation. Prediction accuracy was measured as correlation between observed and predicted CH_4, root mean squared error (RMSE) and mean normalized discounted cumulative gain (NDCG). Inclusion of DMI in the model improved within and between-herd prediction accuracy to 0.77 (RMSE = 23.3%) and 0.58 (RMSE = 31.9%) in RF and to 0.50 (RMSE = 0.327) and 0.13 (RMSE = 42.71) in MLR, respectively than when DMI was not included in the predictive model. When missing DMI records were imputed, within and between-herd accuracy increased to 0.84 (RMSE = 18.5%) and 0.63 (RMSE = 29.9%), respectively. In all scenarios, RF models out-performed MLR models. Results suggest routinely measured variables from dairy farms can be used in developing globally robust prediction models for CH_4 if coupled with state-of-the-art techniques for imputation and advanced ML algorithms for predictive modeling.

Key words: enteric methane, machine learning, prediction models, proxies for methane

Received January 15, 2021.
Accepted February 9, 2022.
*Corresponding author: enyew.negussie@luke.fi

INTRODUCTION

Food production and agriculture face major challenges under climate change, in terms of expected negative effect on productivity as well as implementation of sectoral actions to limit greenhouse gas (**GHG**) emissions. Sustainable farming, livestock husbandry, fisheries, and forestry can help countries identify opportunities for reducing emissions while addressing their food security, resilience, and rural development goals (FAO, 2016).

Agricultural activities contribute 10 to 14% of global anthropogenic GHG emissions (Jantke et al., 2020). Livestock production systems account for 40% of CH_4 emissions in agriculture (FAO, 2018), where the largest part originates from CH_4 that is produced and released from the rumen. There is also an indirect contribution through, for example, feed-production activities, deforestation, and manure (Cassandro, 2020). Agriculture, particularly livestock, is increasingly being recognized as both a contributor to the process and a potential victim of it (Cassandro et al., 2013; Cassandro, 2020). However, in the agriculture sector, livestock production has a great potential for reducing GHG emissions and a tremendous ability to contribute to climate change mitigation and adaptation (Cassandro 2020). This is in part because, of the many available options for mitigation of GHG, mitigation of CH_4 is particularly efficient given its relatively short half-life and therefore any mitigation effort is expected to result in quick returns. The growing demand for meat and milk, which is predicted to double by 2050 (Rojas-Downing et al., 2017) calls for an accurate inventory of CH_4 emissions for setting up effective and sustainable mitigation strategies.

Direct measurement of CH_4 emissions from individual animals using respiration chambers provides reliable information which can be used for national inventories, assessment of dietary mitigation strategies, genetic selection, and calculation of energy loss through exhaled CH_4 (Appuhamy et al., 2016). However, this approach is not suitable for large-scale assessment and is expensive and labor intensive (Kebreab et al., 2006; Moraes et al., 2014; Negussie et al., 2017b). There have been several efforts in recent decades to develop low-cost and portable methods for direct measurement of CH_4 emissions in animals (Negussie et al., 2017a,b; Zhao et al., 2020). Although such handheld and portable applications for direct measurement have the potential for high throughput, they are generally based on CH_4 concentration as opposed to flux assays and are in some cases considered to be less accurate than respiration chambers (Garnsworthy et al., 2019). Instead, the use of combinations of proxies for CH_4 has been suggested as a valid alternative to direct measurement of CH_4. Proxies for CH_4 are traits that are directly or indirectly related to CH_4 and that can easily be measured on a large scale and at low-cost (Negussie et al., 2017a). Some proxies (e.g., milk yield, milk composition, lactation stage) are easily and readily available from routine national recording schemes in many countries and thus their use may be promising in developing robust prediction models for CH_4. In a comprehensive review (Negussie et al., 2017a) highlighted that use of combinations of readily available proxies for CH_4 could increase accuracy of CH_4 predictions by 15 to 35%. This is mainly because, different proxies describe independent sources of variation in CH_4 emissions and one proxy can correct for shortcomings of others.

Several equations have been developed for proxy-based prediction of CH_4 emission in dairy cattle using primarily some powerful yet expensive proxies such as feed intake and diet composition (Ellis et al., 2010; Hristov et al., 2013; Nielsen et al., 2013; Ramin and Huhtanen, 2013; Storlien et al., 2014; Appuhamy et al., 2016; Charmley et al., 2016). A comprehensive analysis of data consisting of these proxies was reported recently by Niu et al. (2018). They used multiple linear regression (**MLR**) models for prediction of enteric CH_4 emissions based on traits such as energy intake, diet composition, and milk yield as predictor variables (Niu et al., 2018). Unfortunately, large-scale availability of such data containing energy intake or DMI and diet composition is limited. They are especially difficult and expensive to record from commercial farms. When available, they are mainly resourced from relatively small numbers of animals or from a single herd, that limits their applicability to other regions or production systems. Furthermore, most of these prediction models used so far are based on conventional statistical methods fitting MLR models. Such models cannot approximate potentially nonlinear relationships between proxies and emissions unless resorting to generalized additive model extensions and to modeling nonlinear relationships explicitly. Therefore, the use of low-cost and routinely recorded traits (e.g., milk yield, milk composition, age, lactation stage) as predictor variables can be a practical option. For these reasons, a more comprehensive database needs to be collated to develop enteric CH_4 emission prediction models at both global and regional scales (Niu et al., 2018) applying more flexible state-of-the-art statistical and analytical tools.

The smart farming revolution is a global trend based on key innovative technologies, such as Internet of things, cloud computing, big data, and machine learning (**ML**), which are reshaping modern agriculture (Wolfert et al., 2017). A vast array of sensors and phenotyping platforms for farm applications are now generating an enormous and continuous stream of data. Several of the above-mentioned proxies are already being generated

from such applications and more are likely to follow. These data are high throughput, relatively low-cost and can be used for development of robust models for accurate prediction of CH_4 emissions. Effective and efficient utilization of information contained in such large heterogeneous data sets requires advanced and versatile statistical tools, such as ML algorithms. In predictive modeling ML provides an excellent solution to identify hidden trends in heterogeneous and noisy data sets, and to accommodate nonlinear relationships between variables (Zhang and Ma, 2012; Al-Jarrah et al., 2015). Use of ML methods for proxy-based predictions of CH_4 emissions from combined across-country heterogeneous data sets has not yet been reported. Importantly, their predictive performance in comparison with conventional statistical methods, involving routinely recorded proxies for CH_4 remains unexplored. The main objectives of the current study were (1) to combine heterogeneous across-country data on routinely measured proxies for CH_4 into an integrated data set; (2) to apply a ML ensemble algorithm random forest (**RF**) to the proxy-based predictions of CH_4 and compare its performance with that of MLR models; (3) to explore the possibility of imputing missing data points and compare accuracy of CH_4 prediction from imputed and nonimputed data sets, because combining data from heterogeneous across-country sources is bound to generate a proportion of missing records.

MATERIALS AND METHODS

Data

Data on enteric CH_4 production and proxies for CH_4 that are routinely collected from dairy farms were provided by 13 research centers from 10 European partner countries (Belgium, Denmark, Finland, Germany, Ireland, the Netherlands, Poland, Spain, Switzerland and UK) of the METHAGENE consortium (EU-COST Action FA1302) on large-scale methane measurements on individual ruminants for genetic evaluations (www.methagene.eu). The data sets were from 20 herds covering a diverse geographical and production-systems mix. Individual cow records from different breeds, parity (from 1–3+), age and stage of lactation (DIM from 1–349) were combined into a large integrated data set. Variables included in the combined data set were herd, breed, parity, DIM, BW, DMI, CH_4 production, CH_4 measurement method, milk yield, milk fat, and milk protein. The breeds included were mainly Holstein Friesians, Nordic Red, Brown Swiss, Norwegian Red, and Nordic crosses. The majority (90%) of the herds kept Holstein Friesian breed and only 2 herds kept breeds other than Holstein Friesian. In all herds, BW was measured and provided as weight in kilograms along with measurement of CH_4. Because of cost and associated technical difficulties, in most commercial dairy herds recordings of feed intake and other diet-related information are not routinely conducted. As a result, for this study extensive diet composition and feeding management information were not provided except for DMI, which few of the participating herds were able to provide. Methane emissions were measured with 5 measurement techniques: cattle respiration chambers, the SF_6 tracer technique, and sniffers (F10, Gasmet, and NG Guardian), and description of the methods as well as some measurement details are provided in Garnsworthy et al. (2019). In total, the combined data set included 47,129 repeated records from 3,886 cows, belonging to 5 dairy breeds (Table 1). Detailed description of the data by participating herds is provided in Supplemental Table S1 (https://doi.org/10.7910/DVN/BINDG9, Negussie, 2022).

Data Integration

Individual data sets collated from the 20 herds were standardized by making sure that all variables were expressed in the same units such as milk yield in kilograms, protein and fat as percentages, and enteric CH_4 as grams per day. When expressed in liters per day, enteric CH_4 production was converted to grams per day using the following conversion equation:

Table 1. Descriptive statistics of CH_4 and the main proxy variables included in the integrated data set

Variable	No. of observations	Mean	SD	Minimum	Maximum
CH_4, g/d	43,519	372.5	133.2	100	983
Milk yield, kg/d	43,507	31.3	9.2	9.0	89.5
Milk fat, %	23,783	3.84	0.64	3.23	9.91
Milk protein, %	23,783	3.63	0.55	3.00	6.94
DMI, kg/d	3,427	20.2	3.3	7.4	39.3
BW, kg	43,392	585	114.5	478	955
DIM	43,519	156	87.5	1	349
Parity	43,472	1.99	1.6	1	14

$$CH_4\ (L/d) \times 0.668 = CH_4\ (g/d), \quad [1]$$

where 0.668 is the density of CH_4 at normal temperature and pressure (20°C and 1 atm) (Engineering ToolBox, 2003, 2004). Weekly averages of CH_4 production were calculated and used in data analyses. Categorical variables for breed and CH_4 measurement method were standardized by making sure that all categories were labeled consistently across data sets. All date variables were standardized to a common DD-MM-YYYY format. After data standardization, data from the 20 individual herds were combined into a large integrated data set. The combined data set was then filtered by retaining only records with DIM ≤350 and CH_4 ≤1,000 g/d. Outliers for enteric CH_4 production (g/d; outside 3.5 SD, within herd) were also excluded. The final filtered data set contained 43,519 records from 3,483 cows.

Imputation of Missing Proxies

Integrated and filtered data from the combined 20 individual data sets contained 18.7% missing data points for DMI and milk fat and protein. A nonparametric *k*-nearest neighbor imputation approach (Troyanskaya et al., 2001) was used to impute missing proxy data to obtain a larger and complete data set to compare the predictive performance of algorithms on imputed and nonimputed data sets. The imputation process first involved calculation of Gower distances (Gower, 1971) between records as

$$S_{ij} = \left(\sum_{k=0}^{n} S_{ijz}.\delta_{ijz}\right) \Big/ \sum_{z=1}^{v} \delta_{ijz}, \quad [2]$$

where S_{ij} is the similarity between samples i and j, $\delta_{ijz} \in \{0, 1\}$ is an indicator variable that specifies whether a comparison between samples i and j over variable z is possible ($\delta_{ijz} = 1$) or not ($\delta_{ijz} = 0$); S_{ijz} is the similarity coefficient at any given variable z. Similarities were calculated differently depending on whether variable z was binary, categorical, or quantitative: (1) association tables; (2) Hamming distances; (3) $1 - |x_i - x_j|/$ (range). Similarity coefficients were then averaged over the v variables for which comparisons were possible, to give a general similarity coefficient S_{ijz} between samples i and j. Based on the matrix of Gower similarities, *k*-nearest neighbors for any given record were selected, and missing data points were imputed based on the average values (quantitative variables) or majority-vote (if binary or categorical variables) of the k neighbors. In the present study, a value of $k = 4$ was chosen for the imputation. In a few cases, the neighborhood did not have enough information to make imputation possible and missing values were left unimputed and consequently removed. This left 40,532 records in the final imputed data set.

Data Analyses

After data integration and imputation, 3 data sets were generated for predictive modeling: (1) 3k, containing 3,337 records with no missing data on any variables (no imputation); (2) 21k, containing 21,215 records and only the missing 84% of DMI records were imputed; and (3) 41k, containing 40,532 records [i.e., all filtered records where all missing data points (81% of DMI and 48% of milk fat and protein records) were imputed]. All data sets comprised the 11 proxies used in this study to model CH_4 production: breed (5 classes), country (10 classes), herd (20 classes), DIM, parity (3 classes), methane measurement technique (5 classes), DMI (kg/d; imputed in the 21k and 41k data sets), BW (kg), milk yield (kg/d), milk fat (%; imputed in the 41k data sets) and milk protein (%; imputed in the 41k data sets). The 3 data sets were all used to predict CH_4 emissions with both RF and MLR predictive models, either including or not including DMI in the model. Figure 1 provides a visual representation of the adopted data analysis.

Prediction Models for CH_4 Emissions from Proxies

Random Forest Predictive Model. Random forest is a supervised ML algorithm used for predictive modeling in both classification and regression problems (Breiman, 2001). Random forest generates and aggregates a "forest" of trees built on resampled subsets of the data. In the current study, RF was used to predict individual CH_4 emissions from imputed and nonimputed data sets of proxies for CH_4. As a semiparametric ML ensemble algorithm, RF is robust to overfitting and able to capture complex interactions in data, thereby handling efficiently problems associated with complex and heterogeneous data structures (González-Recio and Forni, 2011). The RF works by building large number of decision trees on bootstrapped samples of the data and random subsets of the predictor variables. Predictions from each single decision tree are then averaged to get the final prediction for CH_4 emissions:

$$\hat{f}(\mathbf{x}_i) = \frac{1}{B}\sum_{b=1}^{B} \hat{f}_b(\mathbf{x}_i), \quad [3]$$

where $\mathbf{x}_i$ is the vector of predictors for record i (i.e., CH_4 proxies), and $\hat{f}_b(\mathbf{x}_i)$ is the corresponding predicted CH_4 emission from the regression tree built on the bth bootstrapped sample of the data. Each tree heuristically minimizes the loss function (residual sum of

squares) through top-down greedy recursive binary splitting. The number of trees in the forest (B), the number of variables or proxies (m) randomly used in each node and the maximum size of each node for further partition were tuned based on the lowest generalized error on the out-of-bag samples. Convergence of the generalization error was assessed from 1,000 trees onward and lowest generalized error was obtained with $m = 5$ and minimum node [size = 5]. Equation 3 was used to build RF models to accommodate the different scenarios with and without DMI and imputation of missing DMI as well as other missing proxy variables as described above.

Multiple Linear Regression Model. To provide a benchmark against which performance of RF models can be compared, MLR models were run for all data sets and scenarios considered.

The basic MLR model for CH_4 as a function of proxies was:

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i, \quad [4]$$

where y_i is methane production of individual i, β_0 is the intercept of the model, $\beta_1, \ldots, \beta_m$ are the regression coefficients of the proxy variables included in the model, x_{ij} is the value for proxy j in animal i (for categorical proxies, x_{ij} are indicator variables), and ε_i is the residual term. The variables in [1; p] depend on the scenario considered (data set, imputation). The QR factorization was used to solve the MLR models. Although residual variance may vary between subsets of the records, the most relevant in this study being methane measurement techniques, for the MLR model homoscedasticity was assumed. Equation 4 was applied to the 3 data sets and all envisaged scenarios, the same as with RF models.

Accuracy of Prediction

Predictive ability of RF and MLR models was evaluated within and between herds in the 3 data sets (3k, 21k, and 41k). In each data set, inclusion or not inclusion of DMI (measured or imputed) was tested. The effect of imputing missing proxies was evaluated indirectly by comparing the 3 data sets: no imputation (3k), imputation of DMI (21k), imputation of DMI and milk fat and protein (41k). Predictive ability of the models was estimated through 10-fold cross-validation

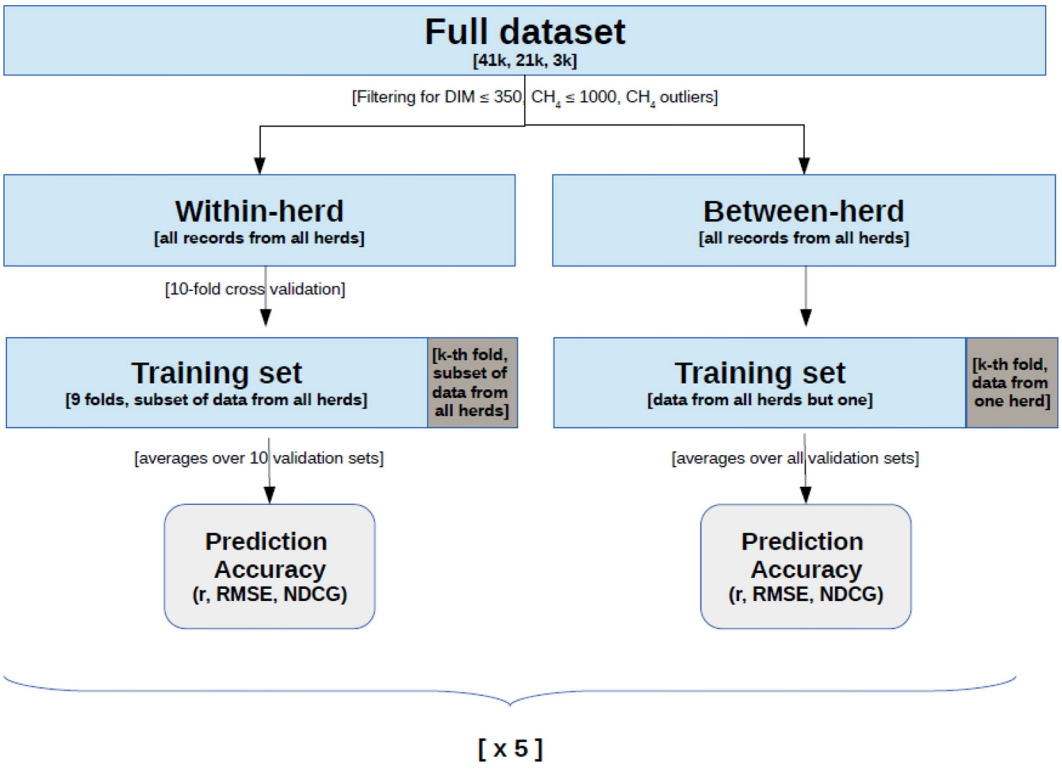

Figure 1. Cross-validation scheme used for estimation of the predictive ability from random forest and multiple linear regression models under within- and between-herd prediction scenarios. Each model (data set size, inclusion of DMI, within- or between-herd scenario) was replicated 5 times. r = Pearson correlation between observed and predicted CH_4 values; RMSE = root mean squared error; NCDG = normalized discounted cumulative gain.

replicated 5 times. The data were split into 10 partitions: 9 were used to train the model and one was used to test the model, until all 10 partitions were used once as test set. Multiple records from the same cows were always assigned to the same fold. In the within-herd cross-validation, records from all herds were used both in the training and test sets: CH_4 records from a given cow were predicted from data of herd mates plus cows from other herds and countries. In the between-herd cross-validation, all records from one herd "k" (each herd in turn) were set aside as test set, and the remaining $k-1$ herds were used to train the model (Figure 1).

Three accuracy metrics were used to measure predictive ability of the models: (1) Pearson correlation between observed and predicted CH_4 values, (2) root mean squared error (**RMSE**), as percentage of the mean of the observed response variable, and (3) normalized discounted cumulative gain (**NDCG**). Normalized discounted cumulative gain is a ranking metric developed in information theory (Järvelin and Kekäläinen, 2002), which has been applied to evaluation of genomic selection models (Blondel et al., 2015). The NDCG metric evaluates the top individuals in the ranking, which are supposed to be the most relevant when comparing models. The top 20% emitters were considered here and their CH_4 outputs were ranked based on observed and predicted ranks; NDCG was calculated as:

$$NDCG\ (y,\hat{y}) = \left(\frac{\sum_{i=1}^{k} \left(y\left[\pi(\hat{y})\right]_i \cdot d(i) \right)}{\sum_{i=1}^{k} \left(y\left[\pi(y)\right]_i \cdot d(i) \right)} \right), \qquad [5]$$

where $y\left[\pi(\hat{y})\right]$ and $y\left[\pi(y)\right]$ are the top k observed CH_4 emission records according to either their predicted or observed ranking; $d(i) = \log_2(i+1)$ is the weight by which ranked values are discounted with ($i \in [1, k]$); k is the number of individuals included in the top 20%. The NDCG values range between 0 and 1, with values close to 1 indicating better performance of the model to correctly predict the most relevant individuals (e.g., identifying those cows that emit more CH_4 than other cows).

Variable Importance

In predictive modeling, important proxies drive the outcome of the model and have a significant effect on accuracy of prediction. In RF, the relative importance of proxies included in the predictive models was automatically retrieved. The relative importance of proxies was estimated by running the out-of-bag samples through the RF trees after randomly permuting the values at each proxy variable and comparing the resulting predictive accuracy (or loss function) with that obtained from the original data (nonpermuted). The relative importance of proxies was then scaled to be in the [0, 100] range, which provided insights into the predictive and biological roles played by the proxies in prediction of CH_4 emissions from cows. Two measures of variable importance were used: percent increase in mean squared error and increase in node purity.

Software and Computing Environment

All data handling, processing, and analysis were performed using the R environment for statistical computing (https://r-project.org). Specifically, the R package VIM (Kowarik and Templ, 2016) was used for imputation of missing proxy data and RF (Liaw and Wiener, 2002) was used for RF predictions. For the MLR models we used the lm() function from the stats R package, which uses QR factorization to solve the model (https://r-project.org). Plots were generated using the ggplot2 R package (Wickham, 2009).

RESULTS

Integrated Across-Country Data

Average enteric CH_4 emission across the 20 herds in 10 European countries was 372.5 g/d (±133.2; SD) and ranged from 280 to 543 g/d (Table 1). Figure 2 summarizes CH_4 emissions by herd and shows that there is marked variability between herds, with most (two-thirds) measurements lying between 250 and 500 g/d. The effect of country of origin on methane emissions was tested in a multiple regression model. In a naive model, the effect of country of origin was significant ($P < 0.01$). When accounting for co-dependencies between records (cows nested within herds, nested within countries) the herd effect absorbs a large portion of the variation from the country effect, which is no longer significant ($P > 0.01$). This indicates that CH_4 emission levels vary between countries mainly because of between-herd variations.

Regarding the distribution of methane, we have checked visually the per-country histograms and numerically the descriptive statistics and the mean and median for CH_4 correspond very well. The formal Shapiro-Wilk test tells us however that in all countries except 3 there are some deviations from normality. This may be linked to different sample sizes and different CH_4 variability among countries. Additionally, these deviations are limited (in the sense that distributions do not appear to be binary, or bimodal, or strongly skewed) and linear models are known to be rather robust to deviations from normality and distributional

assumptions (Schielzeth et al., 2020; Knief and Forstmeier, 2021). General summary statistics for the different proxy variables in the integrated data set are given in Table 1.

Prediction Accuracy of CH_4 from RF Versus MLR Models.

Within-Herd CH_4 Prediction Accuracy. Table 2 shows average Pearson correlations, RMSE, and NDCG for all predictive models and scenarios from both RF and MLR. Figures 3, 4, and 5 show RF results for each of the 5 replicates per model and scenario. In the 3k data set, when measured DMI was included in the RF model, prediction accuracy measured as Pearson correlation between observed and predicted CH_4 $r(\hat{y}, y)$ increased from 0.52 to 0.77, RMSE was reduced from 31.3 to 23.3, and NDCG increased from 0.75 to 0.89 (Table 2). In the 21k data set, when missing DMI records were imputed and included in the prediction model, prediction accuracy $r(\hat{y}, y)$ increased from 0.80 to 0.84, RMSE declined from 20.0 to 18.5, and NDCG increased from 0.91 to 0.92. In the 41k data set, when all missing variables including DMI were imputed and included in the prediction model, prediction accuracy $r(\hat{y}, y)$ increased slightly from 0.81 to 0.82, RMSE decreased from 20.6 to 20.0, and NDCG remained the same. When moving from 3k to 21k and 41k data sets, predictions become progressively less variable.

Within-herd prediction accuracy from RF models varied with CH_4 measurement method. In general, CH_4 measurements from chambers tended to be predicted more accurately and more robustly (lower between-replicate variability) than CH_4 records from sniffers and SF_6. This was especially true when DMI was not included in the model, although with DMI included in the model sniffers gave comparable results to chambers. The SF_6 almost always gave the least accurate predictions and was very variable across replicates (Figures 3 to 5).

Marked differences were found between predictive models (RF vs. MLR) in within-herd prediction accuracy. Prediction accuracies from the MLR model were consistently lower than RF across data sets and scenarios. Across the different data sets and scenarios, when prediction was made using MLR instead of RF, $r(\hat{y}, y)$ declined from 0.77 to 0.50, RMSE increased from 23.3 to 32.7, and NDCG decreased from 0.89 to 0.73 for the 3k data set. For the 21k data set, within-herd $r(\hat{y}, y)$ declined from 0.84 to 0.75, RMSE increased from 18.5

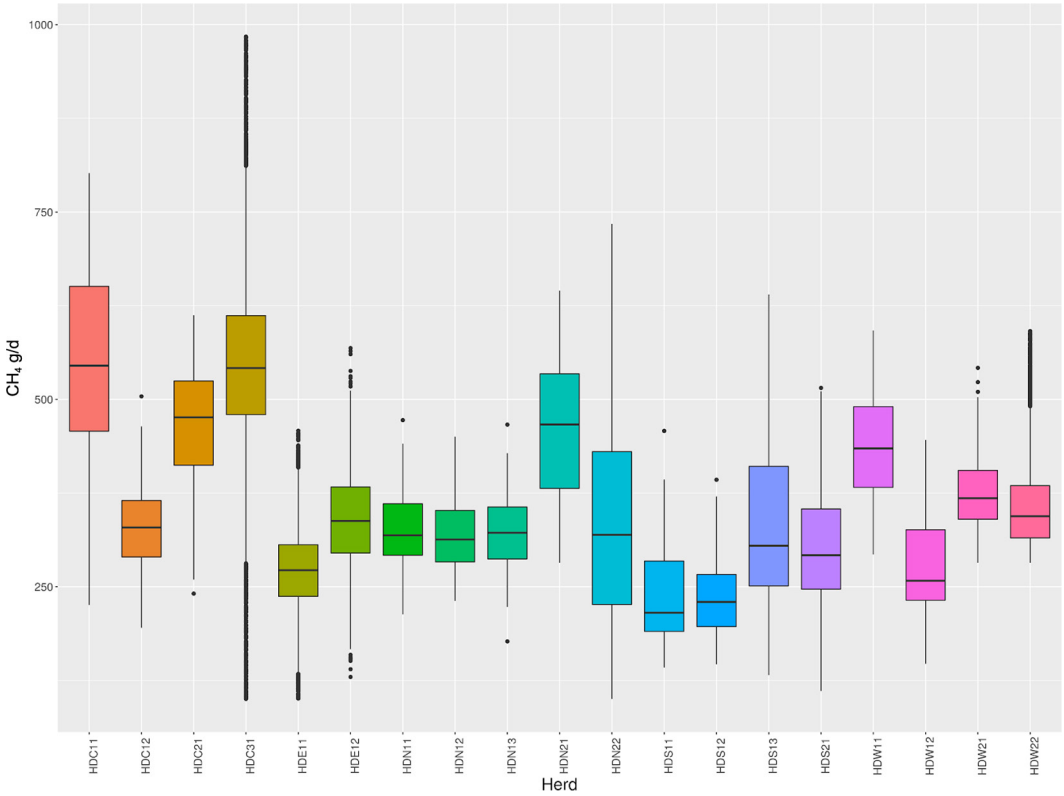

Figure 2. Distribution of mean methane production (g/d) across herds in the combined data set. Boxes correspond to the interquartile range (IQR); whiskers are 1.5 times the IQR on both sides; dots describe data points which fall outside the ±1.5 × IQR boundaries around the box.

Table 2. Within- and between-herd predictive ability of random forest (RF) and multiple linear regression (MLR) models with different data sets: 3k, 21k, 41k, measured by 3 accuracy metrics: Pearson correlation, RMSE, and NDCG from 10-fold cross-validation that is averaged over 5 replicates

Prediction model	Data attributes			No. of cross-validation replicates	Within-herd			Between-herd		
	Size	Proxies imputed	DMI included		$r(\hat{y},y)$[1]	RMSE[2]	NDCG[3]	$r(\hat{y},y)$	RMSE	NDCG
RF	3k	No	No	5	0.52	31.3	0.75	0.32	35.4	0.73
RF	3k	No	Yes	5	0.77	23.3	0.89	0.58	31.9	0.82
RF	21k	No	No	5	0.80	20.3	0.91	0.33	33.1	0.65
RF	21k	Yes	Yes	5	0.84	18.5	0.92	0.63	29.9	0.83
RF	41k	Yes	No	5	0.81	20.6	0.86	0.20	34.8	0.55
RF	41k	Yes	Yes	5	0.82	20.0	0.86	0.39	33.6	0.65
MLR	3k	No	No	5	0.28	36.6	0.70	0.12	45.6	0.56
MLR	3k	No	Yes	5	0.50	32.7	0.73	0.13	42.7	0.56
MLR	21k	No	No	5	0.71	22.4	0.89	0.07	38.4	0.70
MLR	21k	Yes	Yes	5	0.75	22.7	0.90	0.14	32.7	0.87
MLR	41k	Yes	No	5	0.77	21.6	0.86	0.19	41.4	0.81
MLR	41k	Yes	Yes	5	0.79	21.4	0.86	0.21	39.6	0.83

[1]Pearson correlations (r) between observed and predicted CH_4 production.

[2]Root mean squared error (expressed as percentage of the mean CH_4 production g/d).

[3]Mean normalized discounted cumulative gain.

to 22.7, and NDCG decreased from 0.92 to 0.90. Similarly, for 41k data set, within-herd $r(\hat{y},y)$ declined from 0.82 to 0.79, RMSE increased from 20.0 to 21.4, and NDCG was the same at 0.86 for both MLR and RF predictive models.

Between-Herd CH_4 Prediction Accuracy. For the between-herd scenario, again a general pattern toward more accurate and especially less variable predictions with increasing data size was observed, although somewhat less clear than in the within-herd scenario. In the 3k data set, inclusion of measured DMI in the prediction model increased between-herd $r(\hat{y},y)$ to 0.13 and 0.58, reduced RMSE to 42.7 and 31.9, and increased NDCG to 0.56 and 0.82 in MLR and RF models, respectively (Table 2). In the 21k data set, when missing DMI imputed, between-herd $r(\hat{y},y)$ increased to 0.14 and 0.63, RMSE decreased to 32.7 and 29.9, and NDCG decreased to 0.87 and 0.83 in MLR and RF models, respectively. In 41k data set, when missing proxy variables were imputed, between-herd $r(\hat{y},y)$ increased to 0.21 and 0.39, RMSE decreased to 39.6 and 33.6, and NDCG decreased to 0.83 to 0.65 in MLR and RF models, respectively.

Between-herd CH_4 prediction accuracies among herds using different CH_4 measurement methods are shown in Figures 3 to 5. In general, whether or not DMI was included in predictive models, between-herd prediction accuracies for chamber CH_4 measurement methods were higher than sniffer CH_4 measurement methods. For instance, when measured DMI was included or missing DMI was imputed and included in the prediction model, between-herd $r(\hat{y},y)$ for chamber measurement methods was 20% higher than for sniffer measurement methods. Similarly, RMSE for chamber CH_4 measurement methods was 30 to 50% lower than sniffer measurement methods. However, when NDCG metric was used, only a small difference was observed in between-herd prediction accuracy between chamber and sniffer measurement methods.

The RF and MLR predictive models had varied between-herd prediction accuracy over the 3 different data sets (Table 2). For instance, in the 3k data set, between-herd $r(\hat{y},y)$ declined from 0.58 to 0.13, RMSE increased from 31.9 to 42.7, and NDCG decreased from 0.82 to 0.56 when MLR was used instead of RF model. Similarly, for the 21k data set, between-herd $r(\hat{y},y)$ declined from 0.63 to 0.14, RMSE increased from 29.9 to 38.4, and NDCG increased from 0.83 to 0.87. For the larger 41k data set, in which all missing proxy variables were imputed, between-herd $r(\hat{y},y)$ declined from 0.39 to 0.21, RMSE increased from 33.6 to 39.6, and NDCG increased from 0.65 to 0.83. In general, the differences between the models in $r(\hat{y},y)$, RMSE and NDCG in Table 2 were all significant at ($P < 0.0001$) except for the 41k data with imputation and DMI in the model, where the difference was not significant ($P = 0.28$).

Variable Importance

When building decision trees, RF computes how much each variable is contributing to the prediction, which is a measure of variable importance. Two measures of variable importance were used in the present study: percent increase in mean square error and increase in node purity, which are presented in Figure 6. When DMI was not included in the prediction model,

the most relevant proxies for prediction were DIM, milk yield, BW, milk fat, and milk protein. When DMI was included in the prediction model DMI ranked first in variable importance, followed by milk yield, DIM, milk fat, and BW. In general, breed, parity and CH_4 measurement method ranked at the bottom of variable importance.

DISCUSSION

Accurate inventories of GHG emission are essential to reflect a country's national emissions from livestock production systems. Productivity and associated emissions intensity of livestock farming differ widely around the world and the potential for change is large (Niu et al., 2018). Understanding national, regional, and global variations in GHG emissions are therefore essential for concerted global actions to mitigate emissions (Hristov et al., 2018). Particularly at a time when there are uncertainties in proportion of increase in CH_4 emissions solely attributable to livestock sources (Hristov et al., 2018), accurate estimation of CH_4 emissions across different national borders and production systems is needed. Average CH_4 production (g/d) from ruminants varies with diet, animal populations (e.g., species and breeds), production system, production level, and DIM (Bell et al., 2014; de Haas et al., 2011; Garnsworthy et al., 2012; Negussie et al., 2017b). Different estimates of average CH_4 production of dairy cows were reported from different production systems (Waghorn et al., 2008; O'Neill et al., 2011; Hellwing et al., 2013; Deighton et al., 2014; Negussie et al., 2017b; Niu et al., 2018). Using data compiled over 8 experiments and covering 30 diets (Hellwing et al., 2013) reported an average CH_4 production of 412 g/d for Danish lactating cows. Negussie et al. (2017b) working on Nordic red cows estimated an average CH_4 production of 396 g/d whereas (Bayat et al., 2017) using chambers reported a range from 335 to 492 g/d in an experiment designed to test different dietary treatments. Bell et al. (2014) reported an average CH_4 production of 418 g/d with a range between 220 and 480 g/d in 1964 lactating Holstein Friesian cows across 21 UK herds. In Australia (Williams et al., 2013; Deighton et al., 2014; Moate et al., 2014) CH_4 emission ranged from 369 to 458 g/d for cows fed harvested pasture grass. From other pasture grass based dairy productions systems, somewhat different values have been reported. For instance, O'Neill et al. (2011) reported CH_4 emissions of 251 g/d in cows fed harvested perennial ryegrass. In New Zealand, Waghorn et al. (2008) reported CH_4 emissions ranging from 273 to 352 g/d in cows fed harvested pasture

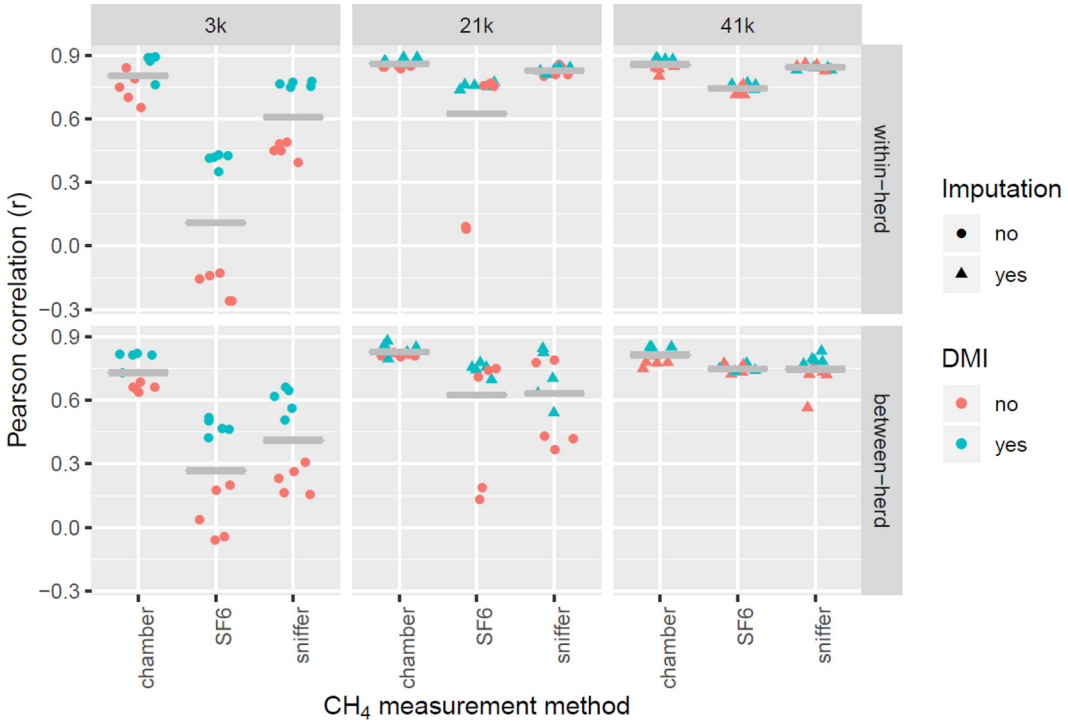

Figure 3. Within-herd (top) and between-herd (bottom) prediction accuracies in terms of Pearson correlations $r(\hat{y}, y)$ between observed and predicted CH_4 emissions (g/d) by CH_4 measurement method and prediction scenarios within 3k, 21k, and 41k data sets. Colors indicate inclusion (blue) or not (red) of DMI in the predictive model; shapes indicate imputation (triangles) or no imputation (dots) of missing proxies in the data set. Data points represent the 5 replicates of each predictive model. The gray horizontal bars are average prediction accuracies per scenario.

grass. This selection of results demonstrates the large between-country variability. In our combined across-country data, estimated overall mean CH_4 production was 372 g/d, which is in the middle of the ranges described above. Our estimate also corresponds to recent estimates from a combined regional data set reported in (Niu et al., 2018). Using a global data set collated from the United States, European Union, Australia, and New Zealand, the authors reported mean CH_4 production of 345 g/d per cow for European Union, 354 g/d per cow for the United States, and 347 g/d per cow for the combined intercontinental data set.

Accuracy of Proxy-Based Prediction of CH_4 Using RF Versus MLR Models

In a comprehensive review, Negussie et al. (2017a) concluded that whenever direct animal measurements are difficult and expensive to procure, use of combinations of proxies for CH_4 in empirical prediction equations has great potential. Empirical models have long been used involving different predictor variables to predict CH_4 emissions from cows, some as early as the 1930s (Kriss, 1931). There are several examples of such quantitative approaches to predict CH_4 production in cattle using mainly dietary and animal factors as proxies (Kebreab et al., 2008; Ellis et al., 2010; Ramin and Huhtanen, 2013; Appuhamy et al., 2016; Niu et al., 2018; Benaouda et al., 2019). Appuhamy et al. (2016) listed 40 such models that were developed in North America, Europe, Australia, and New Zealand. They suggested that comprehensive CH_4 emission models need examining and testing against CH_4 emission measurements from dairy cows in different regions of the world, an idea that was recently implemented in Benaouda et al. (2019). So far, although many prediction models have been reported, a closer look at most empirical models indicates that there are still a range of limitations that may preclude their practical applicability. These limitations include the following: (1) Most prediction models are not based on individual cow observations but on treatment means from different studies. Depending on sample size and other factors including measurement methods, this can be associated with different degrees of uncertainty (e.g., SD) (Appuhamy et al., 2016). (2) In most cases, data sets used as an input were from a single herd, a specific diet or from only few laboratories (Blaxter and Clapperton, 1965; Yan et al., 2000; Jentsch et al., 2007; Ellis et al., 2010). (3) Most prediction models were based on measure-

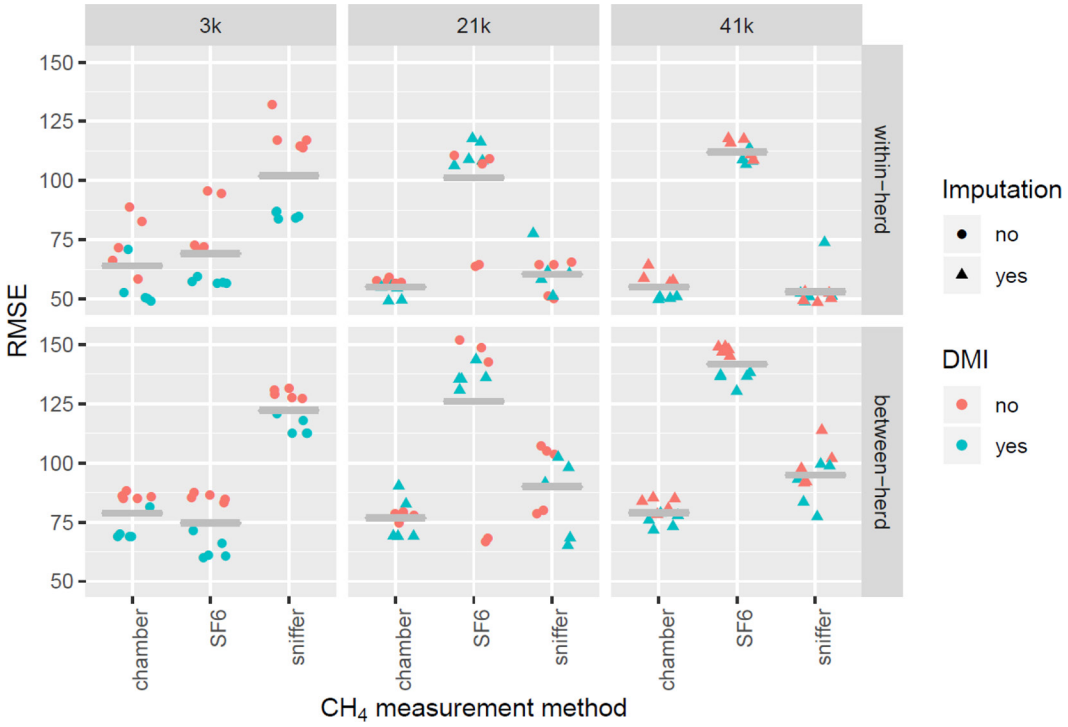

Figure 4. Within-herd (top) and between-herd (bottom) prediction accuracies in terms of root mean squared error (RMSE) by CH_4 measurement method and prediction scenarios within 3k, 21k, and 41k data sets. Colors indicate inclusion (blue) or no inclusion (red) of DMI in the predictive model; shapes indicate imputation (triangles) or no imputation (dots) of missing proxies in the data sets. Data points represent the 5 replicates of each predictive model. The gray horizontal bars are average prediction accuracies per scenario.

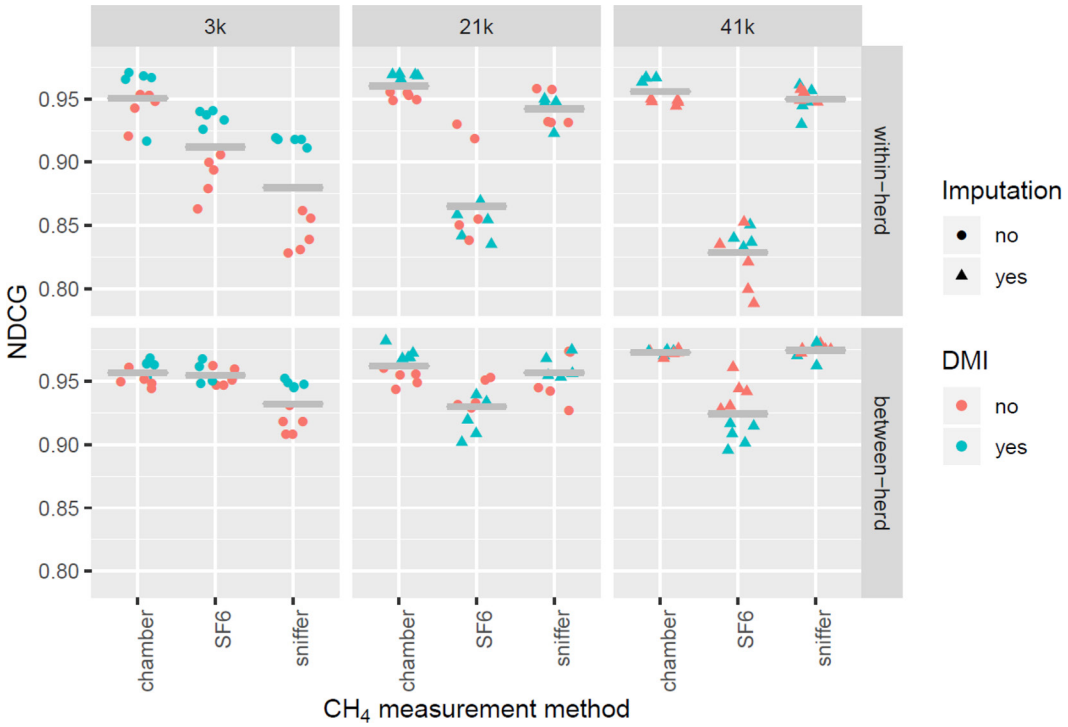

Figure 5. Within-herd (top) and between-herd (bottom) prediction accuracies in terms of mean normalized discounted cumulative gain (NDCG) by CH_4 measurement method and prediction scenarios within 3k, 21k, and 41k data sets. Colors indicate inclusion (blue) or no inclusion (red) of DMI in the predictive model; shapes indicate imputation (triangles) or no imputation (dots) of missing proxies in the data sets. Data points represent the 5 replicates of each predictive model. The gray horizontal bars are average prediction accuracies per scenario.

ments from relatively small numbers of animals, which may limit their broad applicability. (4) Most prediction models used simple or MLR analyses without appropriate modeling of the fixed and random components (Ramin and Huhtanen, 2013). In many instances, possible nonlinear relationships in the data were not taken into consideration, leading to biased estimates of parameters (St-Pierre, 2001). (5) As enteric CH_4 emissions are strongly related to feed intake, almost all models included a measure of intake such as DMI, intake of gross energy or metabolizable energy, or fiber intake, as prime predictor variables. However, these variables are currently not readily available under commercial conditions and none of them is routinely measured on individual animal's on-farm. Thus, there is a need to develop robust prediction models that do not rely completely on feed intake measures or estimates (Hristov et al., 2018). Nevertheless, models without these variables (such as DMI or dietary composition) could be less accurate and thus during model development, it is essential to consider the trade-offs between cost, practicability and prediction accuracy (Appuhamy et al., 2016). (6) Finally, the advent of the smart farming revolution, new phenotyping platforms, such as sensors, on-line recording and imaging tools, have started to generate an enormous amount of data on proxies for CH_4 from heterogeneous sources. Compilation, analysis and utilization of such large sets of information require the latest, robust and versatile statistical tools, which are now common in the ML approach. However, none of the CH_4 prediction models reported so far in dairy cattle has attempted to use ML algorithms; and our study represents the first such effort in this direction. Machine learning algorithms have great potential for identifying hidden trends in unstructured heterogeneous data sets and offer predictive modeling that can accommodate nonlinear relationships among variables. Combining across-country heterogeneous data along with the application of ML algorithms should be a logical step toward developing robust and globally relevant CH_4 prediction models (Negussie et al., 2019). The robustness of ML methods is partly due to their extraordinary ability to input and use information from heterogeneous sources. Consequently, predictions from ML model RF are reliable and robust and are applicable under diverse production and environmental conditions. Furthermore, the approach outlined in the current study offers analytical tools to support future attempts to build globally representative large-scale GHG emission databases, on the basis of which, ac-

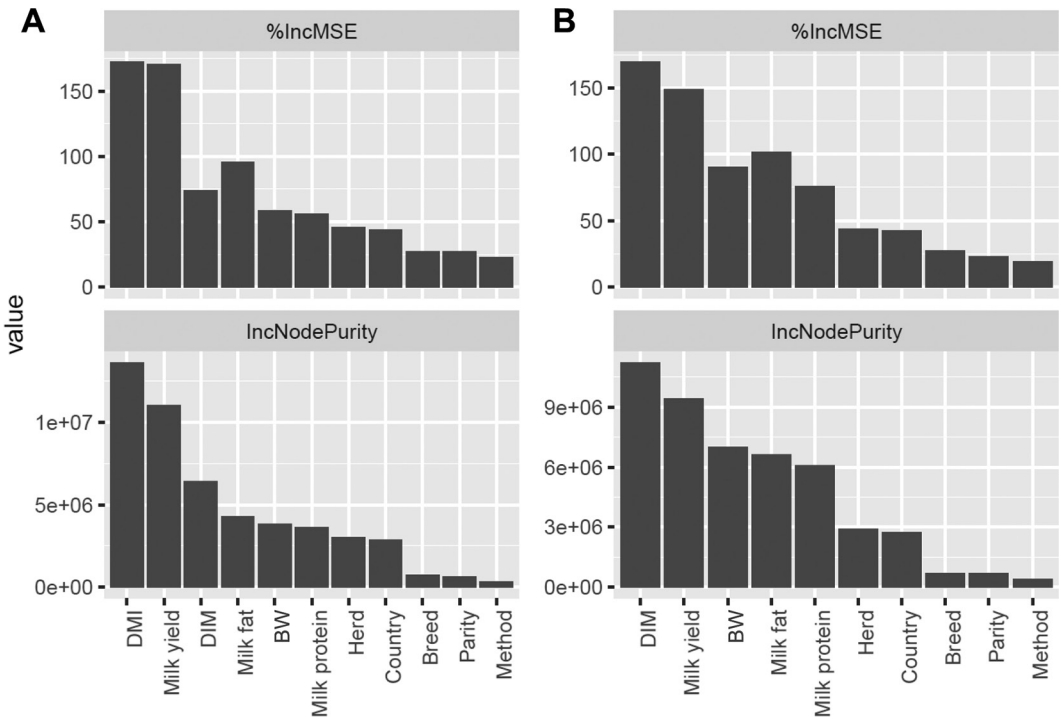

Figure 6. Variable importance in terms of percent mean square error reduction (%IncMSE) and increase in node purity (IncNodePurity) from the random forest model (A) without including DMI and (B) with inclusion of DMI in the prediction model.

curate regional and intercontinental inventories as well as concerted global mitigation strategies could be developed.

Within-Herd and Between-Herd Methane Prediction Accuracy

Both in within- and between-herd predictions, inclusion of DMI, either measured or imputed, increased the predictive ability of RF and MLR models. As expected for decreasing marginal increments, the largest prediction accuracy increase was observed for the 3k data set (from RF models: 17–48% for within-herd predictions and 1–83% for between-herd predictions, depending on the metrics used). In modeling the stage of lactation, quadratic terms or high order polynomials can be used to account for possible nonlinear relationships. In Equation 4, an additive model with linear terms only was chosen. However, we have also tested MLR model with a quadratic DIM term, but no significant changes in the results (negligible or no improvements in predictive accuracy) was observed. The modeling of lactation stage for CH_4 prediction could be an area of future further investigation.

Between-herd predictions are more challenging than within-herd predictions. For between-herd predictions, RF showed a much better performance than MLR, with accuracy increasing by as much as 350% for $r(\hat{y}, y)$ with 21k data. This enhanced performance substantiates the robustness of RF predictions and the ability of RF models to make effective use of information coming from different herds with heterogeneous management and farm routines. Overall, across the 3 data sets and accuracy metrics, between-herd prediction accuracies were lower than within-herd prediction accuracies which is in line with the results reported in Wang and Bovenhuis (2019). This is probably because observations of an individual cow will predict the CH_4 output of its herd mates with higher accuracy than predicting CH_4 output of animals in other herds in the combined across-country data. Furthermore, diet composition and other factors that influence CH_4 output, vary less within herds than between herds. Evaluating random cross-validation and block cross-validation (with farms as blocks) which corresponds to within- and between-herd cross-validation in our study, Wang and Bovenhuis (2019) reported that random cross-validation could result in an over optimistic view on the ability of milk IR spectra to predict CH_4 emission and leads to misleading conclusions. Roberts et al. (2017) explained that when validation data are randomly selected for cross-validation from the entire spatial domain, training and validation data from nearby locations will be dependent (spatial autocorrelation). Consequently, if the objective

is to project outside the spatial structure of the training data, error estimates from random cross-validations will be overly optimistic. To address this, they suggested that blocks can be designed across the spatial structure itself (i.e., in contiguous geographic space). This effectively forces testing on more spatially distant records, thus decreasing optimism in error estimates which underscores the power and practicality of the between-herd or block cross-validation as implemented in the current study. In addition to cross-validation with nonrandom blocks, carefully chosen modeling objectives can offer more reliable error estimates (Roberts et al., 2017).

When comparing methods used for CH_4 measurement, records coming from respiration chambers consistently displayed the most accurate and least variable predictions, across data sets, scenarios and accuracy metrics. Possible reasons for this include within-day variation in CH_4 emissions, which may not be accounted for in spot-sample sniffer techniques, and the influence of herd and environmental variability on sniffer measurements. In addition, sniffer measurements are influenced by cow activity, feeding behavior, and relationships between cow herd mates, which are excluded when cows are placed in chambers (Garnsworthy et al., 2019). As a result, sniffer data were not as robust as chamber data in predicting CH_4 emission in other herds. Nevertheless, when within-herd prediction accuracies were compared, sniffer data were as accurate as chamber data in prediction of CH_4 as they are mostly tailored to specific herd environments. On the other hand, for the SF_6 method, when DMI was added or missing proxies were imputed, with-herd prediction accuracies were close to estimates from the chamber herds. However, estimates from SF_6 were in general highly variable owing to the small number of observations available for the method.

Variable Importance and Effect of Imputation of Missing Data Points

In predictive statistics, it is fundamentally important to have an accurate model; however, it may also be desirable to have a model that is easy to interpret, and where variable features that contribute most to predictive ability can be identified. In the present study, relative contributions of proxy variables were provided by RF models. When DMI was not included in the model, the proxies that contributed most to prediction accuracy were DIM, milk yield, BW, milk fat, and milk protein. On the other hand, when included in the model, DMI was identified as the most important variable by all the metrics used to measure variable importance. On the contrary, breed, parity and CH_4 measurement method were the variables that contributed least to prediction accuracy.

Adding DMI to the prediction models is expected to increase accuracy of prediction of CH_4 because of the clear biological relationship between DMI and CH_4 production. For instance, in dairy cows that consume more feed, more CH_4 is produced due to the greater availability of substrate for microbial fermentation (Hristov et al., 2018). Conversely, increased intake may potentially increase passage rate and shorten digesta retention time in the rumen, thus decreasing rumen fermentation and organic matter digestibility, which ultimately decreases CH_4 per unit of feed (Boadi et al., 2004). Dry matter intake and ME intake are the variables most used for prediction of CH_4 emission (Johnson and Johnson, 1995; Mills, et al., 2003; Ellis et al., 2007). Consequently, prediction equations including such energy intake variables showed low root mean square prediction error (**RMSPE**) and are therefore important in prediction of enteric CH_4 emission (Sobrinho et al., 2019). Ellis et al. (2007) also confirmed that use of DMI in prediction equations for CH_4 emission in cattle resulted in lower RMSPE. Sobrinho et al. (2019) working on Nellore cattle reported that equations that included intakes of DM, total carbohydrate, ME, cellulose and nonfiber carbohydrates were the most accurate for the prediction of enteric CH_4 emission. In our study, however, DMI was measured in few herds, and data on other dietary variables were not available. Therefore, imputation of missing DMI data points and their inclusion in prediction models had a marked positive effect on prediction accuracies, especially on between-herd prediction accuracies. This was particularly true for herds using the sniffer method, the majority of which did not have measured DMI records. This is consistent with reports in the literature (Appuhamy et al., 2016; Bayat et al., 2017). Appuhamy et al. (2016) evaluated 40 prediction equations using data that included measured or estimated DMI and some feed quality attributes. They reported that models using estimated DMI predicted enteric CH_4 emissions as accurately as the measured DMI, provided DMI could be estimated with reasonable accuracy. They also reported that enteric CH_4 emissions from dairy cows can be predicted successfully (RMSPE = 12.7%) without DMI, but more accurately (RMSPE = 7.7%) with estimated DMI. Similarly, in our study, using heterogeneous across-country data RMSPE for RF models ranged from 23.3 to 31.3% when DMI was not in the prediction model and ranged from 18.5 to 20.3% when imputed DMI was included in the prediction model. This in general indicates that by imputing missing DMI data points it is possible to achieve satisfactory predic-

tion of CH_4 emissions provided the right statistical and imputation methods are used.

Niu et al. (2018) evaluated the potential contribution of predictor variables by adding sequentially each of the predictor variables during a model development process. They observed that accuracy of prediction of CH_4 production was improved in models that included DMI, diet composition, milk production and composition, and BW. In particular, complex models that used all available variable information consistently improved prediction performance compared with simpler models. However, models using only milk yield or diet composition were the least accurate. When DMI was removed from the model to predict CH_4 production, ECM was selected instead due to its high correlation with DMI, but model predictive ability was reduced. This is consistent with results of the current study where, when available, DMI was ranked first, followed by milk yield and milk compositional variables. Under the scenario when DMI was omitted from training of the regression trees, RF ranked milk yield and milk compositional traits at the top. Breed, parity and methods used for the measurement of CH_4 ranked lowest, indicating their relatively low importance, which could also be due in part to their correlations with highly related variables.

Benefits of Heterogeneous Across-Country Data and Potentials of Machine Learning in Predictive Modeling

In recent years, marked progress has been made in developing empirical CH_4 prediction models at national, regional, and global levels (Ellis et al., 2010; Appuhamy et al., 2016; Niu et al., 2018). Despite this progress, much remains to be done. Especially, collating diverse and heterogeneous intercontinental data into one usable set, data harmonization, standardization, model validation, and correction for heterogeneity of variances in across-country data are all areas of interest. The statistical methods most used in developing CH_4 predictive models to date have been questioned because of their limitations (Hristov et al., 2018), such as not including random effects of animals or studies. Furthermore, models based on MLR assume predominantly linear relationships among the features of the target variables, although nonlinear relationships in the data set may be equally likely. To the best of our knowledge, the current study is one of the first attempts to predict CH_4 by applying ML on combined data on low-cost routinely recorded proxies for CH_4 from individual animals and from diverse international sources.

In our heterogeneous data set, because most herds had no measured DMI records, most of the missing DMI data points were imputed from routinely recorded proxy variables. The finding that imputed DMI records improved prediction accuracies clearly shows the potential that CH_4 output could be predicted with reasonable accuracy from routinely available variables and estimated DMI, provided that DMI is imputed accurately. This opens a great opportunity to include many herds in large global or regional databases for intensive data analysis, such as across-country genetic evaluations when direct measurements of DMI are not available. In general, addition of many routinely recorded predictor variables into the predictive model can contribute to increased prediction accuracy. Therefore, our results emphasize the great value of using proxy variables that are recorded routinely on-farm and imputing missing DMI observations to generate a reasonably accurate prediction of CH_4 when direct measurements of CH_4 from individual animals are difficult or expensive to obtain on a large scale.

Because predictive ability of models is likely to be enhanced with increasing model complexity (Moraes et al., 2014), during model development the trade-off between availability of variable inputs on-farm and prediction accuracy must be carefully considered. In this regard, a review by Negussie et al. (2017a) provided an extensive list of potential proxies with critical evaluation of their attributes. The present study substantiated the practical use of such proxies in improving the accuracy of prediction of CH_4. Therefore, from now onward, relaxing the threshold on predictor variables to include low-cost and routinely available proxies will open possibilities to collate more and diverse information on GHG emissions from as many regions and production systems as possible. More and diverse information, when properly compiled and analyzed, will have a significant implication in improving the accuracy of predictions. However, Hristov et al. (2018) highlighted that one of current challenges is to make more data available and the next frontier should be the collation of GHG information from as wide and diverse livestock populations as possible to develop robust models that are globally applicable. At the same time, efforts should also be directed toward sharpening analytical tools by which important missing data points could be accurately imputed from routinely measured proxies.

Future Considerations

Predictive modeling has a great significance and value particularly for traits that are difficult and expensive to record routinely under commercial farm conditions. This includes traits such as DMI and CH_4 emission. The accuracy of predictive models can be influenced by the type and the completeness of the data used. In most situations especially with integrated data from diverse

sources missing variables are quite common. The imputation of missing variables form other correlated predictor variables is a common solution to increase sample size and hence accuracy as shown in this study. However, imputation techniques can introduce uncertainty in the data with trade-offs between increase in sample size and accuracy on one hand and possible increase in uncertainty on the other. Our results showed the benefits of imputation in terms of predictive performance across the different scenarios. Nevertheless, future studies on the measurement of accuracy of imputations are warranted. Furthermore, heterogeneity of residual variances is one important area. In multilevel models, residual variance may vary between subsets of the data, such as between methane measurement techniques, and this was not accounted for in our MLR model in Equation 4, where homoscedasticity was assumed. In classical statistical modeling, notably in mixed linear models for animal genetics, heterogeneity of variance is known to bias the estimates of model solutions (e.g., Visscher and Hill, 1992). However, Hill (1984) on the contrary has shown that ignoring heterogeneity of variances decreases the efficiency of genetic evaluation procedures and consequently the response to selection. Generally, less obvious is the effect of heterogeneous variance on the performance of predictive models in ML, where most of the methods used (including RF) do not rely on the homoscedasticity assumption. In a simulation study, W. Ruth and T. Loughin (Simon Fraser University, Burnaby, BC, Canada; unpublished data) showed that heterogeneity of variance negatively affects the performance of single regression trees. However, less clear is how this could affect the performance of ensemble methods such as RF, which are based on a large number of trees that could counteract the negative effects of heterogeneous variance by reducing the variance of predictions through averaging over trees. Assessing the effect of heterogeneous variance on the predictive ability of ML models for prediction of CH_4 emission is an interesting topic that will be taken up in our follow-up studies.

In conclusion, the present study describes a novel way forward for developing accurate and robust CH_4 prediction models. These models will help in designing effective and sustainable GHG mitigation strategies as well as aiding national, regional and global GHG inventories. The broad applicability of such models requires collation of input data from wide and heterogeneous sources. It helps to overcome the difficulty of procuring predictor variables related to intake and diet composition on-farm, and the need for versatile statistical tools for compiling and analysis of unstructured, heterogeneous across-country data. In this way, low-cost and routinely measured proxy variables can be used to provide a reasonably accurate prediction of CH_4 when coupled with imputation of missing DMI data points. As a predictive model, the use of the ML ensemble algorithm RF consistently gave more accurate predictions than conventional multiple regression models. This provides a great potential for building a globally representative large-scale CH_4 emission database on the basis of which an accurate regional and intercontinental inventory as well as a concerted global mitigation strategy could be developed. Results from this study lay strong foundations for our next thorough comparison of various state-of-the-art ML methods for prediction of dairy-cow CH_4 emissions in a much larger integrated global data set.

ACKNOWLEDGMENTS

This paper is the result of the concerted effort of all participants and support from the networks of COST Action FA1302 “METHAGENE: Large-scale methane measurements on individual ruminants for genetic evaluations.” The authors thank all individuals and groups who have directly or indirectly contributed to this work; special thanks are due to the technical and financial support from the COST Action FA1302 of the European Union. In addition, all financial and technical support from all participating countries and research centers involved in this work is greatly acknowledged. The authors have not stated any conflicts of interest.

REFERENCES

Al-Jarrah, O. Y., P. D. Yoo, S. Muhaidat, G. K. Karagiannidis, and K. Taha. 2015. Efficient machine learning for big data: A review. Big Data Research. 2:87–93. http://dx.doi.org/10.1016/j.bdr.2015.04.001.

Appuhamy, J. A. D. R. N., J. France, and E. Kebreab. 2016. Models for predicting enteric methane emissions from dairy cows in North America, Europe, and Australia and New Zealand. Glob. Chang Biol. 22:3039–3056. https://doi.org/10.1111/gcb.13339.

Bayat, A. R., L. Ventto, P. Kairenius, T. Stefański, H. Leskinen, I. Tapio, E. Negussie, J. Vilkki, and K. J. Shingfield. 2017. Dietary forage to concentrate ratio and sunflower oil supplement alter rumen fermentation, ruminal methane emissions, and nutrient utilization in lactating cows. Transl. Anim. Sci. 1:277–286. https://doi.org/10.2527/tas2017.0032.

Bell, M. J., N. Saunders, R. H. Wilcox, E. M. Homer, J. R. Goodman, J. Craigon, and P. C. Garnsworthy. 2014. Methane emissions among individual dairy cows during milking quantified by eructation peaks or ratio with carbon dioxide. J. Dairy Sci. 97:6536–6546. https://doi.org/10.3168/jds.2013-7889.

Benaouda, M., C. Martin, X. Li, E. Kebreab, A. N. Hristov, Z. Yu, D. R. Yáñez-Ruiz, C. K. Reynolds, L. A. Crompton, J. Dijkstra, A. Bannink, A. Schwarm, M. Kreuzer, M. McGee, P. Lund, A. L. F. Hellwing, M. R. Weisbjerg, P. J. Moate, A. R. Bayat, K. J. Shingfield, N. Peiren, and M. Eugène. 2019. Evaluation of the performance of existing mathematical models predicting enteric methane emissions from ruminants: Animal categories and dietary mitigation strategies. Anim. Feed Sci. Technol. 255:114207. https://doi.org/10.1016/j.anifeedsci.2019.114207.

Blaxter, K. L., and J. L. Clapperton. 1965. Prediction of the amount of methane produced by ruminants. Br. J. Nutr. 19:511–522. https://doi.org/10.1079/BJN19650046.

Blondel, M., A. Onogi, H. Iwata, and N. Ueda. 2015. A ranking approach to genomic selection. PLoS One 10:e0128570. https://doi.org/10.1371/journal.pone.0128570.

Boadi, D. A., K. M. Wittenberg, S. L. Scott, D. Burton, K. Buckley, J. A. Small, and K. H. Ominski. 2004. Effect of low and high forage diet on enteric and manure pack greenhouse gas emissions from a feedlot. Can. J. Anim. Sci. 84:445–453. https://doi.org/10.4141/A03-079.

Breiman, L. 2001. Random forests. Mach. Learn. 45:5–32. https://doi.org/10.1023/A:1010933404324.

Cassandro, M. 2020. Animal breeding and climate change, mitigation and adaptation. J. Anim. Breed. Genet. 137:121–122. https://doi.org/10.1111/jbg.12469.

Cassandro, M., M. Marcello, and B. Stefanon. 2013. Genetic aspects of enteric methane emission in livestock ruminants. Ital. J. Anim. Sci. 12:450–458. https://doi.org/10.4081/ijas.2013.e73.

Charmley, E., S. R. O. Williams, P. J. Moate, R. S. Hegarty, R. M. Herd, V. H. Oddy, P. Reyenga, K. M. Staunton, A. Anderson, and M. C. Hannah. 2016. A universal equation to predict methane production of forage-fed cattle in Australia. Anim. Prod. Sci. 56:169–180. https://doi.org/10.1071/AN15365.

de Haas, Y., J. J. Windig, M. P. L. Calus, J. Dijkstra, M. de Haan, A. Bannink, and R. F. Veerkamp. 2011. Genetic parameters for predicted methane production and the potential for reducing enteric emissions through genomic selection. J. Dairy Sci. 94:6122–6134. https://doi.org/10.3168/jds.2011-4439.

Deighton, M. H., S. R. O. Williams, M. C. Hannah, R. J. Eckard, T. M. Boland, W. J. Wales, and P. J. Moate. 2014. A modified sulphur hexafluoride tracer technique enables accurate determination of enteric methane emissions from ruminants. Anim. Feed Sci. Technol. 197:47–63. https://doi.org/10.1016/j.anifeedsci.2014.08.003.

Ellis, J. L., A. Bannink, J. France, E. Kebreab, and J. Dijkstra. 2010. Evaluation of enteric methane prediction equations for dairy cows used in whole farm models. Glob. Chang. Biol. 16:3246–3256. https://doi.org/10.1111/j.1365-2486.2010.02188.x.

Ellis, J. L., E. Kebreab, N. E. Odongo, B. W. McBride, E. K. Okine, and J. France. 2007. Prediction of methane production from dairy and beef cattle. J. Dairy Sci. 90:3456–3466. https://doi.org/10.3168/jds.2006-675.

Engineering ToolBox. 2003. Gases - densities. Accessed Jun. 11, 2018. https://www.Engineeringtoolbox.Com/Gas-Density-d_158.Html.

Engineering ToolBox. 2004. STP – standard temperature and pressure and NTP – normal temperature and pressure. Accessed Jun. 11, 2018. https://www.engineeringtoolbox.com/stp-standard-ntp-normal-air-d_772.html.

FAO (Food and Agriculture Organization of the United Nations). 2016. Greenhouse Gas Emissions from Agriculture, Forestry and Other Land Use. FAO. http://www.fao.org/3/a-i6340e.pdf.

FAO (Food and Agriculture Organization of the United Nations). 2018. Enteric fermentation. FAOSTAT. Accessed June 10, 2018. http://www.fao.org/faostat/en/#data/ge.

Garnsworthy, P. C., J. Craigon, J. H. Hernandez-Medrano, and N. Saunders. 2012. Variation among individual dairy cows in methane measurements made on farm during milking. J. Dairy Sci. 95:3181–3189. https://doi.org/10.3168/jds.2011-4606.

Garnsworthy, P. C., G. F. Difford, M. J. Bell, A. R. Bayat, P. Huhtanen, B. Kuhla, J. Lassen, N. Peiren, M. Pszczola, D. Sorg, M. H. P. W. Visker, and T. Yan. 2019. Comparison of methods to measure methane for use in genetic evaluation of dairy cattle. Animals (Basel) 9:837. https://doi.org/10.3390/ani9100837.

González-Recio, O., and S. Forni. 2011. Genome-wide prediction of discrete traits using Bayesian regressions and machine learning. Genet. Sel. Evol. 43:7. https://doi.org/10.1186/1297-9686-43-7.

Gower, J. C. 1971. A general coefficient of similarity and some of its properties. Biometrics 27:857–874. https://doi.org/10.2307/2528823.

Hellwing, A. L. F., P. Lund, J. Madsen, and M. R. Weisbjerg. 2013. Comparison of enteric methane production predicted from the CH_4/CO_2 ratio and measured in respiration chambers. Adv. Anim. Biosci. 4:557.

Hill, W. G. 1984. On selection among groups with heterogenous variance. Anim. Prod. 39:473–477.

Hristov, A. N., E. Kebreab, M. Niu, J. Oh, A. Bannink, A. R. Bayat, T. M. Boland, A. F. Brito, D. P. Casper, L. A. Crompton, J. Dijkstra, M. Eugène, P. C. Garnsworthy, N. Haque, A. L. F. Hellwing, P. Huhtanen, M. Kreuzer, B. Kuhla, P. Lund, J. Madsen, C. Martin, P. J. Moate, S. Muetzel, C. Muñoz, N. Peiren, J. M. Powell, C. K. Reynolds, A. Schwarm, K. J. Shingfield, T. M. Storlien, M. R. Weisbjerg, D. R. Yáñez-Ruiz, and Z. Yu. 2018. Symposium review: Uncertainties in enteric methane inventories, measurement techniques, and prediction models. J. Dairy Sci. 101:6655–6674. https://doi.org/10.3168/jds.2017-13536.

Hristov, A. N., J. Oh, J. L. Firkins, J. Dijkstra, E. Kebreab, G. Waghorn, H. P. S. Makkar, A. T. Adesogan, W. Yang, C. Lee, P. J. Gerber, B. Henderson, and J. M. Tricarico. 2013. Special topics: Mitigation of methane and nitrous oxide emissions from animal operations: I. A review of enteric methane mitigation options. J. Anim. Sci. 91:5045–5069. https://doi.org/10.2527/jas.2013-6583.

Jantke, K., M. J. Hartmann, L. Rasche, B. Blanz, and U. A. Schneider. 2020. Agricultural greenhouse gas emissions: Knowledge and positions of German farmers. Land (Basel) 9:130. https://doi.org/10.3390/land9050130.

Järvelin, K., and J. Kekäläinen. 2002. Cumulated gain-based evaluation of IR techniques. ACM Trans. Inf. Syst. 20:422–446. https://doi.org/10.1145/582415.582418.

Jentsch, W., M. Schweigel, F. Weissbach, H. Scholze, W. Pitroff, and M. Derno. 2007. Methane production in cattle calculated by the nutrient composition of the diet. Arch. Anim. Nutr. 61:10–19. https://doi.org/10.1080/17450390601106580.

Johnson, K. A., and D. E. Johnson. 1995. Methane emissions from cattle. J. Anim. Sci. 73:2483–2492. https://doi.org/10.2527/1995.7382483x.

Kebreab, E., K. Clark, C. Wagner-Riddle, and J. France. 2006. Methane and nitrous oxide emissions from Canadian animal agriculture: A review. Can. J. Anim. Sci. 86:135–157. https://doi.org/10.4141/A05-010.

Kebreab, E., K. A. Johnson, S. L. Archibeque, D. Pape, and T. Wirth. 2008. Model for estimating enteric methane emissions from United States dairy and feedlot cattle. J. Anim. Sci. 86:2738–2748. https://doi.org/10.2527/jas.2008-0960.

Knief, U., and W. Forstmeier. 2021. Violating the normality assumption may be the lesser of two evils. Behav. Res. Methods 53:2576–2590. https://doi.org/10.1101/498931.

Kowarik, A., and M. Templ. 2016. Imputation with the *R* package VIM. J. Stat. Softw. 74:1–16. https://doi.org/10.18637/jss.v074.i07.

Kriss, M. 1931. A comparison of feeding standards for dairy cows, with especial reference to energy requirements. J. Nutr. 4:141–161. https://doi.org/10.1093/jn/4.1.141.

Liaw, A., and M. Wiener. 2002. Classification and regression by randomforest. R News 2:18–22. http://CRAN.R-project.org/doc/Rnews/.

Mills, J. A. N., E. Kebreab, C. M. Yates, L. A. Crompton, S. B. Cammell, M. S. Dhanoa, M. S. R. E. Agnew, and J. France. 2003. Alternative approaches to predicting methane emissions from dairy cows. J. Anim. Sci. 81:3141–3150. https://doi.org/10.2527/2003.81123141x.

Moate, P. J., S. R. O. Williams, C. Grainger, M. C. Hannah, E. N. Ponnampalam, and R. J. Eckard. 2011. Influence of cold-pressed canola, brewers grains and hominy meal as dietary supplements suitable for reducing enteric methane emissions from lactating dairy cows. Anim. Feed Sci. Technol. 166–167:254–264. https://doi.org/10.1016/j.anifeedsci.2011.04.069.

Moraes, L. E., A. B. Strathe, J. G. Fadel, D. P. Casper, and E. Kebreab. 2014. Prediction of enteric methane emissions from cattle. Glob. Chang. Biol. 20:2140–2148. https://doi.org/10.1111/gcb.12471.

Negussie, E. 2022. Supplemental Table S1. Harvard Dataverse, V1. https://doi.org/10.7910/DVN/BINDG9

Negussie, E., Y. de Haas, F. Dehareng, R. J. Dewhurst, J. Dijkstra, N. Gengler, D. P. Morgavi, H. Soyeurt, S. van Gastelen, T. Yan, and F. Biscarini. 2017a. Invited review: Large-scale indirect measurements for enteric methane emissions in dairy cattle: A review of proxies and their potential for use in management and breeding decisions. J. Dairy Sci. 100:2433–2453. https://doi.org/10.3168/jds.2016-12030.

Negussie, E., O. González-Recio, Y. de Haas, N. Gengler, H. Soyeurt, N. Peiren, M. Pszczola, P. Garnsworthy, M. Battagin, A. R. Bayat, J. Lassen, T. Yan, T. Boland, B. Kuhla, T. Strabel, A. Schwarm, A. Vanlierde, and F. Biscarini. 2019. Machine learning ensemble algorithms in predictive analytics of dairy cattle methane emission using imputed versus non-imputed datasets. Page 40 in Proceedings of 7th GGAA (Greenhouse Gas and Animal Agriculture) Conference, Iguassu Falls, Brazil. Embrapa Southeast Livestock.

Negussie, E., J. Lehtinen, P. Mäntysaari, A. R. Bayat, A.-E. Liinamo, E. A. Mäntysaari, and M. H. Lidauer. 2017b. Non-invasive individual methane measurement in dairy cows. Animal 11:890–899. https://doi.org/10.1017/S1751731116002718.

Nielsen, N. I., H. Volden, M. Åkerlind, M. Brask, A. L. F. Hellwing, T. Storlien, and J. Bertilsson. 2013. A prediction equation for enteric methane emission from dairy cows for use in NorFor. Acta Agric. Scand. A Anim. Sci. 63:126–130. https://doi.org/10.1080/09064702.2013.851275.

Niu, M., E. Kebreab, A. N. Hristov, J. Oh, C. Arndt, A. Bannink, A. R. Bayat, A. F. Brito, T. Boland, D. Casper, L. A. Crompton, J. Dijkstra, M. A. Eugène, P. C. Garnsworthy, M. N. Haque, A. L. F. Hellwing, P. Huhtanen, M. Kreuzer, B. Kuhla, P. Lund, J. Madsen, C. Martin, S. C. McClelland, M. McGee, P. J. Moate, S. Muetzel, C. Muñoz, P. O'Kiely, N. Peiren, C. K. Reynolds, A. Schwarm, K. J. Shingfield, T. M. Storlien, M. R. Weisbjerg, D. R. Yáñez-Ruiz, and Z. Yu. 2018. Prediction of enteric methane production, yield, and intensity in dairy cattle using an intercontinental database. Glob. Chang. Biol. 24:3368–3389. https://doi.org/10.1111/gcb.14094.

O'Neill, B. F., M. H. Deighton, B. M. O'Loughlin, F. J. Mulligan, T. M. Boland, M. O'Donovan, and E. Lewis. 2011. Effects of a perennial ryegrass diet or total mixed ration diet offered to spring-calving Holstein-Friesian dairy cows on methane emissions, dry matter intake, and milk production. J. Dairy Sci. 94:1941–1951. https://doi.org/10.3168/jds.2010-3361.

Ramin, M., and P. Huhtanen. 2013. Development of equations for predicting methane emissions from ruminants. J. Dairy Sci. 96:2476–2493. https://doi.org/10.3168/jds.2012-6095.

Roberts, D. R., V. Bahn, S. Ciuti, M. S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J. J. Lahoz-Monfort, B. Schröder, W. Thuiller, D. I. Warton, B. A. Wintle, F. Hartig, and C. F. Dormann. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40:913–929. https://doi.org/10.1111/ecog.02881.

Rojas-Downing, M. M., A. P. Nejadhashemi, T. Harrigan, and S. A. Woznicki. 2017. Climate change and livestock: Impacts, adaptation, and mitigation. Clim. Risk Manage. 16:145–163. https://doi.org/10.1016/j.crm.2017.02.001.

Schielzeth, H., N. J. Dingemanse, S. Nakagawa, D. F. Westneat, H. Allegue, C. Teplitsky, D. Réale, N. A. Dochtermann, L. Z. Garamszegi, and Y. G. Araya-Ajoy. 2020. Robustness of linear mixed-effects models to violations of distributional assumptions. Methods Ecol. Evol. 11:1141–1152. https://doi.org/10.1111/2041-210X.13434.

Sobrinho, T. L. P., R. H. Branco, E. Magnani, A. Berndt, R. C. Canesin, and M. E. Z. Mercadante. 2018. Development and evaluation of prediction equations for methane emission from Nellore cattle. Acta Sci. Anim. Sci. 41:e42559. https://doi.org/10.4025/actascianimsci.v41i1.42559.

St-Pierre, N. R. 2001. Invited review: Integrating quantitative findings from multiple studies using mixed model methodology. J. Dairy Sci. 84:741–755. https://doi.org/10.3168/jds.S0022-0302(01)74530-4.

Storlien, T. M., H. Volden, T. Almøy, K. A. Beauchemin, T. A. McAllister, and O. M. Harstad. 2014. Prediction of enteric methane production from dairy cows. Acta Agric. Scand. A Anim. Sci. 64:98–109. https://doi.org/10.1080/09064702.2014.959553.

Troyanskaya, O., M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein, and R. B. Altman. 2001. Missing value estimation methods for DNA microarrays. Bioinformatics 17:520–525. https://doi.org/10.1093/bioinformatics/17.6.520.

Visscher, P. M., and W. G. Hill. 1992. Heterogeneity of variance and dairy-cattle breeding. Anim. Sci. 55:321–329. https://doi.org/10.1017/S0003356100021012.

Waghorn, G. C., H. Clark, V. Taufa, and A. Cavanagh. 2008. Monensin controlled-release capsules for methane mitigation in pasture-fed dairy cows. Aust. J. Exp. Agric. 48:65–68. https://doi.org/10.1071/EA07299.

Wang, Q., and H. Bovenhuis. 2019. Validation strategy can result in an overoptimistic view of the ability of milk infrared spectra to predict methane emission of dairy cattle. J. Dairy Sci. 102:6288–6295. https://doi.org/10.3168/jds.2018-15684.

Wickham, H. 2009. Ggplot2: Elegant Graphics for Data Analysis. 2nd ed. Springer Nature.

Williams, S. R. O., T. Clarke, M. C. Hannah, L. C. Marett, P. J. Moate, M. J. Auldist, and W. J. Wales. 2013. Energy partitioning in herbage-fed dairy cows offered supplementary grain during an extended lactation. J. Dairy Sci. 96:484–494. https://doi.org/10.3168/jds.2012-5787.

Wolfert, S., L. Ge, C. Verdouw, and M.-J. Bogaardt. 2017. Big data in smart farming–A review. Agric. Syst. 153:69–80. https://doi.org/10.1016/j.agsy.2017.01.023.

Yan, T., R. E. Agnew, F. J. Gordon, and M. J. Porter. 2000. The prediction of methane energy output in dairy and beef cattle offered grass silage-based diets. Livest. Prod. Sci. 64:253–263. https://doi.org/10.1016/S0301-6226(99)00145-1.

Zhang, C., and Y. Ma. 2012. Random forest. Pages 157–175 in Ensemble Machine Learning: Methods and Applications. A. Cutler, D. R. Cutler, and J. R. Stevens, ed. Springer. https://doi.org/10.1007/978-1-4419-9326-7_5.

Zhao, Y., X. Nan, L. Yang, S. Zheng, L. Jiang, and B. Xiong. 2020. A review of enteric methane emission measurement techniques in ruminants. Animals (Basel) 10:1004. https://doi.org/10.3390/ani10061004.

ORCIDS

Enyew Negussie https://orcid.org/0000-0003-4892-9938
Oscar González-Recio https://orcid.org/0000-0002-9106-4063
Mara Battagin https://orcid.org/0000-0001-7309-6793
Ali-Reza Bayat https://orcid.org/0000-0002-4894-0662
Tommy Boland https://orcid.org/0000-0002-7433-130X
Yvette de Haas https://orcid.org/0000-0002-4331-4101
Aser Garcia-Rodriguez https://orcid.org/0000-0001-5519-6766
Philip C. Garnsworthy https://orcid.org/0000-0001-5131-3398
Nicolas Gengler https://orcid.org/0000-0002-5981-5509
Michael Kreuzer https://orcid.org/0000-0002-9978-1171
Björn Kuhla https://orcid.org/0000-0002-2032-5502
Jan Lassen https://orcid.org/0000-0002-1338-8644
Nico Peiren https://orcid.org/0000-0001-5500-1607
Marcin Pszczola https://orcid.org/0000-0003-2833-5083
Angela Schwarm https://orcid.org/0000-0002-5750-2111
Hélène Soyeurt https://orcid.org/0000-0001-9883-9047
Amélie Vanlierde https://orcid.org/0000-0002-4619-1936
Tianhai Yan https://orcid.org/0000-0002-1994-5202
Filippo Biscarini https://orcid.org/0000-0002-3901-2354