Prediction of enteric methane production, yield and intensity of beef cattle using an intercontinental database

van Lingen, H. J., Niu, M., Kebreab, E., Valadares Filho, S. C., Rooke, J. A., Duthie, C. A., Schwarm, A., Kreuzer, M., Hynd, P. I., Caetano, M., Eugène, M., Martin, C., McGee, M., O'Kiely, P., Hünerberg, M., McAllister, T. A., Berchielli, T. T., Messana, J.D, Peiren, N., Chaves, A. V., Charmley, E., Cole, N. A., Hales, K. E., Lee, S. S., Berndt, A., Reynolds, Chris, Crompton, Les, Bayat, A. R., Yáñez-Ruiz, D. R., Yu, Z., Bannink, A., Dijkstra, J., Casper, D. P. and Hristov, A. N. (2019) Prediction of enteric methane production, yield and intensity of beef cattle using an intercontinental database. Agriculture Ecosystems & Environment, 283. 106575. ISSN 0167-8809 doi: https://doi.org/10.1016/j.agee.2019.106575 Available at http://centaur.reading.ac.uk/84405/

To link to this article DOI: http://dx.doi.org/10.1016/j.agee.2019.106575

Publisher: Elsevier

Title

Prediction of enteric methane production, yield and intensity of beef cattle using an intercontinental database

Running Head

Predicting enteric methane from beef cattle

Author List

Henk J. van Lingen[a,*], Mutian Niu[a,b], Ermias Kebreab[a], Sebastião C. Valadares Filho[c], John A. Rooke[d], Carol-Anne Duthie[d], Angela Schwarm[e,1], Michael Kreuzer[e], Phil I. Hynd[f], Mariana Caetano[f], Maguy Eugène[g], Cécile Martin[g], Mark McGee[h], Padraig O'Kiely[h], Martin Hünerberg[i,j], Tim A. McAllister[i], Telma T. Berchielli[k], Juliana D. Messana[k], Nico Peiren[l], Alex V. Chaves[m], Ed Charmley[n], N. Andy Cole[o], Kristin E. Hales[p], Sang-Suk Lee[q], Alexandre Berndt[r], Christopher K. Reynolds[s], Les A. Crompton[s], Ali-Reza Bayat[t], David R. Yáñez-Ruiz[u], Zhongtang Yu[v], André Bannink[w], Jan Dijkstra[x], David P. Casper[y], Alexander N. Hristov[z].

[a]Department of Animal Science, University of California, Davis, CA 95616, USA; [b]Farmer's Business Network Inc., San Carlos, CA 94070, USA; [c]Animal Science Department, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil; [d]SRUC, West Mains Road, Edinburgh EH9 3JG, UK; [e]ETH Zurich, Institute of Agricultural Sciences, 8092 Zürich, Switzerland; [f]Department of Animal and Veterinary Bioscience, The University of Adelaide, Roseworthy Campus, Roseworthy, SA 5371, Australia; [g]INRA, UMR Herbivores, VetAgro Sup, Université Clermont Auvergne, F-63122 Saint-Genès-Champanelle, France; [h]Teagasc, Grange, Dunsany, Co. Meath, Ireland; [i]Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, AB, Canada; [j]Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, AB, Canada; [k]Animal Science Department, São Paulo State University, UNESP, Jaboticabal, SP 14884-900, Brazil; [l]Flanders Research Institute for Agriculture, Fisheries and Food, Animal Sciences Unit, Scheldeweg 68, Melle 9090, Belgium; [m]The University of Sydney, Faculty of Science, School of Life and Environmental Sciences, Sydney, NSW, Australia; [n]CSIRO Agriculture and Food, Private Mail Bag, PO Aitkenvale, Qld 4814, Australia; [o]USDA-ARS, Bushland, TX 79012; [p]USDA-ARS, Clay Center, NE 68933; [q]Department of Animal Science and Technology, Sunchon National University, Suncheon, South Korea; [r]Research and Development, EMBRAPA Southeast Livestock, Rod Washington Luiz, km 234, PO Box 339, 13560-970 Sao Carlos, SP, Brazil; [s]School of Agriculture, Policy and Development, University of Reading, Reading, UK; [t]Milk Production, Production Systems, Natural Resources Institute Finland (Luke), Jokioinen, Finland; [u]Estación Experimental del Zaidin (CSIC), Granada, Spain; [v]Department of Animal Sciences, The Ohio State University, Columbus, OH, USA; [w]Wageningen Livestock Research, Wageningen University & Research, Wageningen, The Netherlands; [x]Animal Nutrition Group, Wageningen University & Research, Wageningen, The Netherlands; [y]Furst McNess Company, Freeport, IL, 61032, USA; [z]Department of Animal Science, The Pennsylvania State University, University Park 16802.

[*] Corresponding Author.
E-mail address: hjvanlingen@ucdavis.edu
[1] Present address: Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, 1433 Ås, Norway

Abstract

Enteric methane (CH_4) production attributable to beef cattle contributes to global greenhouse gas emissions. Reliably estimating this contribution requires extensive CH_4 emission data from beef cattle under different management conditions worldwide. The objectives were to: 1) predict CH_4 production (g d^{-1} $animal^{-1}$), yield [g (kg dry matter intake; DMI)$^{-1}$] and intensity [g (kg average daily gain)$^{-1}$] using an intercontinental database (data from Europe, North America, Brazil, Australia and South Korea); 2) assess the impact of geographic region, and of higher- and lower-forage diets. Linear models were developed by incrementally adding covariates. A *K*-fold cross-validation indicated that a CH_4 production equation using only DMI that was fitted to all available data had a root mean square prediction error (RMSPE; % of observed mean) of 31.2%. Subsets containing data with ≥ 25% and ≤ 18% dietary forage contents had an RMSPE of 30.8 and 34.2%, with the all-data CH_4 production equation, whereas these errors decreased to 29.3 and 28.4%, respectively, when using CH_4 prediction equations fitted to these subsets. The RMSPE of the ≥ 25% forage subset further decreased to 24.7% when using multiple regression. Europe- and North America-specific subsets predicted by the best performing ≥ 25% forage multiple regression equation had RMSPE of 24.5 and 20.4%, whereas these errors were 24.5 and 20.0% with region-specific equations, respectively. The developed equations had less RMSPE than extant equations evaluated for all data (22.5 *vs*. 23.2%), for higher-forage (21.2 *vs*. 23.1%), but not for the lower-forage subsets (28.4 *vs*. 27.9%). Splitting the dataset by forage content did not improve CH_4 yield or intensity predictions. Predicting beef cattle CH_4 production using energy conversion factors, as applied by the Intergovernmental Panel on Climate Change, indicated that adequate forage content-based and region-specific energy conversion factors improve prediction accuracy and are preferred in national or global inventories.

Keywords: empirical modeling, geographical region, forage content, dietary variables, methane emission

1. Introduction

The livestock sector emits about 7.1 gigatonnes of CO_2 equivalents of greenhouse gases per year, which represented approximately 14.5% of total global anthropogenic greenhouse gas emissions in 2005 (Gerber *et al.*, 2013). Cattle emitted 4.6 gigatonnes CO_2 equivalents, of which 2.5 gigatonnes originated from beef and 2.1 gigatonnes from dairy cattle, whereas small ruminants and buffalos emitted 0.47 and 0.62 gigatonnes CO_2 equivalents, respectively. Methane from enteric fermentation contributed about 45% of the combined CO_2 equivalents emissions from the two cattle types. World-wide beef cattle systems produced 35 million tonnes of meat, whereas dairy cattle systems produced 27 million tonnes. Meat protein greenhouse gas emission intensity from beef cattle, and combined meat and milk protein intensity from dairy cattle vary from about 200 to 1100, and 50 to 350 kg CO_2 equivalents per kg edible protein, respectively, depending on the region of the world (Opio *et al.*, 2013). Based on expected farming and consumer lifestyle practices and the predicted world population growth, compared to 1995, global enteric CH_4 emissions are predicted to increase by 70% by 2055 (Popp *et al.*, 2010). To offset this increase and to deal with the highly variable and typically greater CH_4 emission intensity of beef cattle systems, accurate prediction of beef cattle CH_4 emissions across regions are urgently required.

Various beef cattle CH_4 prediction equations, for which a variety of diet and animal characteristics were used as covariates, based on treatments means (*e.g.*, Ellis *et al.*, 2009; Escobar-Bahamondes *et al.*, 2017a) or individual animal data (Ellis *et al.*, 2007; Moraes *et al.*, 2014) have been published. Although the use of individual animal data as applied in the latter two studies contributes to more explained variation of CH_4 production due to dry matter intake (DMI) differences at the animal level, all previously mentioned studies only comprised data from specific geographical locations. In contrast to these equations, which may be appropriate for cattle systems under similar regional conditions, the widely used Intergovernmental Panel on Climate Change (IPCC) methodology recommends a generic CH_4 energy conversion factor (Y_m) without any adjustment for different geographical locations (IPCC, 2014). The Y_m quantifies enteric CH_4 emission as a fraction of the gross energy intake and discriminates between diets with forage contents of ≤ 10 and $> 10\%$ DM, with Y_m being 3.0% and 6.5% of the gross energy intake, respectively. However, more complex equations accounting for dietary nutrient composition and individual animal characteristics in addition to total feed intake may perform better than those that ignore these covariates for various cattle categories (Ellis *et al*., 2007, 2009; Moraes *et al*., 2014; Santiago-Suarez *et al*., 2016). Therefore, more complex beef cattle CH_4 prediction equations that draw from databases with a broad range of diets and geographic conditions may more accurately predict global CH_4 emissions. Publications of inventories that investigated cattle enteric CH_4 emissions in certain countries or regions (*e.g*., Basarab *et al*., 2005; Kebreab *et al*., 2008; Bannink *et al*., 2011; Castelan-Ortega *et al*., 2014; Charmley *et al*., 2016) compared to an intercontinental evaluation (*e.g*., Niu *et al*., 2018) confirm the utility of the latter approach.

The objectives of the current study were: 1) to collate an intercontinental database of enteric CH_4 production of individual animal records of beef cattle; 2) to determine the key variables for predicting beef cattle enteric CH_4 production (g d^{-1} $animal^{-1}$), yield [g $(kg\ DMI)^{-1}$] and intensity [g $(kg\ average\ daily\ body\ weight\ gain)^{-1}$] and their respective relationships; 3) to develop and cross-validate intercontinental and region-specific models, and models for lower- and higher-forage diets.

2. Materials and Methods

2.1 Database

The ‘GLOBAL NETWORK’ project is an international collaborative initiative of animal scientists (http://animalscience.psu.edu/fnn; accessed May 16, 2017). All animal scientists with an interest in greenhouse gas research and with access to CH_4 measurements from beef cattle were invited to collaborate and contribute data to this collaborative CH_4 mitigation data analysis. The resultant beef cattle CH_4 database that was developed from this initiative contains 2015 individual beef cattle records from 52 studies conducted from 1969 to 2015 by research entities from Europe (n = 869 from 18 studies), North America (n = 649 from 14 studies), Brazil (n = 313 from 12 studies), Australia (n = 174 from 7 studies) and South Korea (n = 10 from 1 study). The European studies were conducted in the UK (n = 313 from 7 studies), Switzerland (n = 96 from 1 study), Belgium (n = 72 from 4 studies), Ireland (n = 147 from 2 studies) and France (n = 241 from 4 studies). Eleven North American studies were from the United States (n = 492), and 3 were conducted in Canada (n = 157). The database includes records of enteric CH_4 production along with corresponding DMI, dietary gross energy, crude protein, ether extract (EE), neutral detergent fiber (NDF), starch, ash and forage contents, average daily body weight gain (ADG) and body weight (BW). The database comprised a broad variety of beef cattle that included growing and finishing steers, bulls and heifers, pregnant heifers, and pregnant, non-pregnant, dry and lactating beef cows. Various pure beef breeds and crossbreeds were included, *viz*., Aberdeen Angus, Blonde d’Aquitaine, Belgian Blue, Brahman, Brown Swiss × Limousin, Charolais, Hanwoo, Holstein × Zebu, Hereford × Angus, Luing and Nellore.

The original studies in the database (complete data bibliography is provided in Supplementary information) investigated the impact of diet composition on enteric CH_4 production or cattle metabolism. However, some studies tested the effect of a specific feed additive, nutrient or the use of hormone supplementation, and the data from these treatments were excluded. The excluded treatments included rapeseed cake and nitrate (Troy *et al*., 2015), limestone (Zanetti *et al*., 2017), Acacia tannins, maca, garlic and lupine seeds (Staerfl *et al*., 2012), monensin (Caetano *et al*., 2016, 2018), organosulfur compounds (garlic extracts) (Peiren *et al*., unpublished) and essential oils (Castro Montoya *et al*., 2015), lipids (Duthie *et al*., 2015), dried corn distillers grains (Hünerberg *et al*., 2013ab), linseed oil and protected fat (Fiorentini *et al*., 2014), soybean oil and protected fat (Silva *et al*., 2018), glycerin (Lage *et al*., 2016), whole soybeans (Rossi *et al*., 2017), monensin (Hales *et al*., 2012, 2013, 2014 2015, unpublished; Berndt *et al*., unpublished), diethylstilbestrol (Rumsey *et al*., 1981) and growth hormone-releasing factor (Lapierre *et al*., 1992). After removal of data associated with the aforementioned treatments, 1413 individual records were retained.

Records with missing CH_4 or DMI values were removed from the database; records from respiration chambers in which two animals were housed simultaneously were combined by averaging the CH_4 and DMI and all other variables regarding the two animals; records from repeated measurements within the same experimental period were averaged over the individual measurements recorded. In total, 1366 individual animal records were subsequently retained. In addition, records from growing cattle with negative ADG, and a study for which DMI varied from 9.0 to 32.5 kg d^{-1} (Rooke *et al*., 2015, unpublished) were discarded from the dataset, leaving 1257 records retained. Finally, studies were screened on the basis of mean CH_4 yield after which two studies, for which the control treatments contained 60 and 82% forage had unrealistically low CH_4 yields of 10.3 and 11.3 g (kg DMI)$^{-1}$ (San Vito *et al*., 2016; De Carvalho *et al*., 2016), respectively, were considered outliers and removed from the dataset. This resulted in the retention of 1248 records.

2.2 Model development

Production, yield and intensity of CH_4 were predicted by fitting mixed-effects models according to:

$$y_{ij} = \beta_0 + \beta_1 x_{ij1} + \beta_2 x_{ij2} + \cdots + \beta_k x_{ijk} + s_i + \varepsilon_{ij},$$

where y_{ij} denotes the j^{th} response variable of CH_4 production (g d^{-1} $animal^{-1}$), CH_4 yield [g (kg $DMI)^{-1}$] or CH_4 intensity [g (kg $ADG)^{-1}$] from the i^{th} study; β_0 denotes the fixed effect of intercept; x_{ij1} to x_{ijk} denote the fixed effects of predictor variables and β_1 to β_k are the corresponding slopes; s_i and ϵ_{ij} denote the random effect of study and residual error, respectively, distributed as $s_i \sim N(0, \sigma_s^2)$, $\epsilon_{ij} \sim N(0, y_{ij}\sigma_e^2)$ for CH_4 production, and $\epsilon_{ij} \sim N(0, \sigma_e^2)$ for CH_4 yield and intensity; σ_s^2 is the between-study variance, σ_e^2 is the residual variance, and $y_{ij}\sigma_e^2$ is the residual error variance being proportional to the dependent variable.

In order to provide equations that depend on various predictor variables, eight categories of CH_4 production models were developed, of which four used a fixed and another four a selected combination of covariates: DMI only (DMI_C), DMI and dietary NDF content (DMI+NDF_C), DMI and dietary starch content (DMI+STA_C), DMI and dietary EE content (DMI+EE_C); a selection of DMI and the dietary NDF, starch, forage, EE, crude protein and ash contents (Diet_C), the Diet_C covariates plus BW (Animal_C), the Animal_C covariates except DMI (Animal_no_DMI_C), and DMI, the dietary NDF and crude protein contents, and BW (Global_C). Global_C was exclusively associated with covariates that had few or no missing data points. In addition to these eight categories, CH_4 production was predicted using Y_m only. The mixed-effects model to estimate Y_m of this GLOBAL NETWORK Tier 2 equation only included y_{ij}, β_0, s_i, and ϵ_{ij} of the previously shown model, with $\epsilon_{ij} \sim N(0, \sigma_e^2)$. According to the CH_4 production models, six categories of CH_4 yield prediction models were developed: dietary NDF content only (NDF_C), dietary starch content only (STA_C), dietary EE content only (EE_C); a selection of the dietary NDF, starch, forage, EE, crude protein and ash contents (Diet_no_DMI_C), the Animal_no_DMI_C covariates, and dietary NDF, forage and crude protein contents and BW (Global_no_DMI_C). Finally, eight categories of CH_4 intensity prediction models were developed: DMI_C, DMI+NDF_C, DMI+STA_C, DMI+EE_C, Diet_C, Animal_C, Animal_no_DMI_C and Global_C.

Covariates that play a key role in predicting CH_4 production were selected for Diet_C, Diet_no_DMI_C, Animal_C, Animal_no_DMI_C, Global_C and Global_no_DMI_C using a multistep selection approach. Model selection started with all potential covariates associated with the particular model category. Subsequently, one or more next selection steps were performed if not all records without missing values for the selected covariates were used in the previous step. A backward selection approach was applied throughout the different steps, *i.e.*, only covariates selected in a previous step could be selected for the next step. The model selection procedure stopped when the selected covariates were the same as the ones selected in the previous step. With this procedure, a model equation was constructed based on records that contained no missing values for the final selection of covariates and data sufficiency was maximized for the development of model equations throughout the different categories.

The Bayesian information criterion (BIC; *e.g.*, James *et al.*, 2014) was computed for all fitted models. The BIC is a well-known quantitative approach to model selection that favors more parsimonious models over more complex models by penalizing the number of parameters included in the model. Models with the smallest BIC were selected, as a smaller BIC indicates a better tradeoff between the goodness of fit and the number of model parameters. In addition, the presence of multicollinearity of fitted models was examined based on the variance inflation factor. The largest variance inflation factor among all predictor variables was considered as an indicator of multicollinearity (Kutner *et al*., 2005). The identified predictor variables were removed from the model one at a time using a stringent variance inflation factor cutoff value of 3 (Zuur *et al*., 2010). All models were fitted using the *lme* function (Pinheiro and Bates, 2000) of R language and environment for statistical computing (R Core Team 2017; version 3.5.2).

2.3 Data handling

The entire database contained a wide variety of dietary forage contents (57.6 ± 29.8% DM; average ± SD), ranging from 8 to 100% DM. The database was split into a higher-forage subset containing the records with ≥ 25% forage, and a lower-forage subset containing all data with ≤ 18% forage. No studies tested forage contents between 18 and 25%. Because of the small coefficient of variation, dietary starch could not be selected for the lower-forage Diet_C, Animal_C, and Animal_no_DMI_C equations. To explore the geographical impact of CH_4 production, all European, North American and Brazilian higher-forage data were also used as separate subsets. Because of the scarcity of data from Australia and South Korea, no specific equations for the latter two regions were developed. Data from growing and finishing cattle for which ADG was measured were selected for a growing cattle subset, which enabled the development of CH_4 intensity [g (kg ADG)$^{-1}$] equations. Other outliers were identified using the interquartile range method (Zwillinger and Kokoska, 2000) based on all dependent and independent variables as in Niu *et al*. (2018). A factor of 1.5 for extremes was used in constructing boundaries to identify outliers for dependent variables and a factor of 2.5 for independent variables. Outliers were identified only for the complete database. The CH_4 intensity [g (kg ADG)$^{-1}$] values were log transformed to stabilize normality before outlier identification. After removal of records with interquartile range identified outliers in the CH_4 production and variables, 1021 records from 114 dietary treatments and 39 studies were retained. Of these records 882 were from 104 treatments and 38 studies in the higher-forage subset, 139 from 10 treatments and 8 studies from the UK, Ireland, France, Canada and Brazil in the low-forage subset, 307 from 28 treatments and 15 studies in the European higher-forage subset, 394 from 36 treatments and 10 studies the North American higher-forage subset, 104 from 17 treatments and 7 studies the Brazilian higher-forage subset, 72 from 22 treatments and 5 studies from Australia, and 5 from 1 treatment and 1 study from South Korea.

The cleaned dataset used for analysis comprised measurements of enteric CH_4 emission that were obtained from respiration chambers (n = 676), the GreenFeed system (n = 87), and the sulfur hexafluoride (SF_6) tracer technique (n = 258). Animals were either kept in confinement or on pasture (n = 991 *vs*. 30, respectively). Types of forage frequently used in higher-forage diets included fresh alfalfa, sugarcane, sugarcane bagasse, corn silage, barley straw, whole-crop barley silage, whole-crop wheat silage, grass herbage, elephant grass, grass silage, grass seed hay, grass hay wrapping, timothy and natural grassland hay. Types of forage frequently used in lower-forage diets were barley straw, wheat straw, whole-crop wheat silage, corn silage and whole-crop barley silage. Concentrate ingredients in higher-forage and lower-forage diets included dried distillers grains, barley, canola meal, soybean meal, soybean hulls, crude glycerin, corn grain, cereal by-products, dehydrated alfalfa, dehydrated beet pulp, citrus pulp, wheat distillers grains, whole grain oats and minerals.

2.4 Cross-validation and model evaluation

The predictive accuracy of the developed CH_4 prediction models was evaluated using a leave-one-out cross-validation (*e.g.*, James *et al*., 2014), in which all individual studies were consecutively taken as the testing set for model evaluation, while all remaining studies were taken as the training set for model fitting. Currently, most national enteric CH_4 inventories are based on energy conversion factors recommended by the IPCC (2006), which were evaluated, *i.e.*, not cross-validated. The IPCC models and the developed models throughout all categories were, if applicable, evaluated on the various (sub)sets using a combination of model evaluation metrics. Furthermore, equations from Yan *et al*. (2000, 2009) based on data from Northern Ireland, Ellis *et al*. (2007) based on data from North America, Ellis *et al*. (2009) based on data from Canada, Patra (2017) based on data from Brazil, India, Australia and Zimbabwe, Escobar-Bahamondes *et al*. (2017a) based on data from North America, Europe, Australia, Japan and New Zealand, Charmley *et al*. (2016) based on data from Australia, and the Mitscherlich equation from Mills *et al*. (2003) based on data from the UK were evaluated given that the covariates used in these published equations were available in the present database. Of these previously published extant equations, the equation that performed the best using our data and the single regression equation that only depended on DMI and performed the best using our data were reported in the present study. Data from studies included in the present database used for the development of these extant equations were excluded from evaluations of those extant equations to ensure independent evaluation.

First, the mean square prediction error (MSPE) was calculated according to Bibby and Toutenburg (1977) as:

$$\text{MSPE} = \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{n},$$

where O_i and P_i denote the observed and predicted value of the response variable for the i^{th} observation, respectively, and n denotes the number of observations. The square root of the mean square prediction error (RMSPE) was used to assess overall model prediction error. In the present study, RMSPE was expressed as a proportion of observed CH_4 production, yield or intensity means. The MSPE was decomposed into mean bias (MB), slope bias (SB) and random bias to identify systematic biases, of which the MB and SB were calculated as follows:

$$\mathrm{MB} = (\bar{O} - \bar{P})^2,$$

$$\mathrm{SB} = (s_p - r s_o)^2,$$

where $\bar{O}$ and $\bar{P}$ denote the predicted and observed means, s_p denotes the standard deviation of predicted values, s_o denotes the standard deviation of observed values, and r denotes the Pearson correlation coefficient. Second, the ratio of RMSPE and s_o, namely RMSPE-observations standard deviation ratio (RSR), which accounts for the specific variability of the data used for evaluation (Moriasi *et al*., 2007), was used to compare the performance of models based on data from different (sub)sets. Smaller values of RSR indicate less variation in the prediction error compared to the standard deviation of the observations, with RSR = 1 indicating the RMSPE variance is equal to observed data variance. If RSR > 1, $\bar{O}$ is a better predictor than P_i. Third, the concordance correlation coefficient (CCC; Lin, 1989), which quantifies both accuracy and precision based on the bias correction factor (C_b) and r by comparing the best-fit line and observations to the identity line ($y = x$), respectively, was calculated. The CCC is given as:

$$\mathrm{CCC} = r \cdot C_b,$$

The closer the CCC of a model to 1, the better the model performance.

Different forage proportion cutoff values with increments of 5% from 15 to 50% were tested to evaluate the effect of the cutoff for splitting the database into higher-forage and lower-forage subsets on equation performance. Cutoff values of 0, 15, 20, 25, 30, 35, 40, 45, and 50% forage DM were used for evaluation. Per cutoff value, an RSR weighted to the number of observations for the DMI_C equation was calculated for the higher-forage and lower-forage CH_4 production equations, after which the optimal cutoff value could be determined.

3. Results

The inclusion criterion for dietary treatment had different effects on the variables means, *viz*., DMI (8.13 *vs*. 8.06 kg d^{-1}; cleaned *vs*. uncleaned averages, respectively), and NDF (35.0 *vs*. 35.0% of DM), starch (34.0 *vs*. 30.5% of DM), EE (3.02 *vs*. 3.52% of DM), ash (6.29 *vs*. 7.26% of DM), and forage (51.0 *vs*. 58.1% of DM) content of the diet, BW (478 *vs*. 487 kg), CH_4 production (161 *vs*. 164 g d^{-1} $animal^{-1}$), CH_4 yield [20.0 *vs*. 20.4 g (kg DMI)$^{-1}$], CH_4 intensity [145 *vs*. 207 g (kg ADG)$^{-1}$] and Y_m (6.0 *vs*. 6.0 % of the gross energy intake). Summary statistics for the (sub)sets of the present cleaned database that included intake, dietary nutrient composition, BW, ADG and CH_4 variables are presented in Tables 1 and S1.

3.1 Methane production equations

The DMI_C all-data CH_4 production (g d^{-1} $animal^{-1}$) equation indicated a positive relationship of DMI with CH_4 production (Eq. 1; Table 2; regression coefficient ± 2·SE gives a rough estimate of the 95% confidence interval boundaries that correspond to a *P*-value of 0.05, all *P*-values < 0.05 were not reported). The DMI+NDF_C, DMI+STA_C and DMI+EE_C equations had positive, negative and negative regression coefficients for dietary NDF, starch and EE in relation to CH_4 production, respectively (Eqns. 2-4). The RSR, which is the most appropriate statistic for evaluating equations based on different numbers of observations, for the DMI_C, DMI+NDF_C, DMI+STA_C and DMI+EE_C equations indicated similar predictive performance, whereas the CCC indicated the DMI+NDF_C equation performed better than the DMI_C and DMI+EE_C equations (0.63 *vs*. 0.60 and 0.61, respectively). Dietary forage content and DMI were selected for the Diet_C and Animal_C equations (Eqns. 5-6), with BW also selected for the Animal_C equation. Dietary forage and ash and BW were selected for the Animal_no_DMI_C equation (Eq. 7). The Animal_C was the best performing all-data equation developed in the present analysis, with RSR and CCC of 0.61 and 0.76, respectively. Across the developed all-data equations, slope bias ranged from 1.01-12.7%, which was consistently associated with under-prediction at the high end and over-prediction at the low end of production (Fig. 1). Overall, models with a higher number of covariates tended to have less slope bias and had less between-study variance (σ_s^2 not shown).

The RSR of the all-data DMI_C CH_4 production equation was 0.71 (Table 3). Splitting the database into higher-forage and lower-forage subsets at cutoffs of 15 to 50% resulted in very similar weighted average RSR values of 0.68 to 0.69. The cutoff of 20% that was applied resulted in an RSR of 0.94 for the lower-forage subset at this cutoff value, whereas the cutoff values from 25 to 50% had all lower RSR values for the lower-forage subset. This might suggest that the lower-forage subset is a better predictor at a higher cutoff. However, the prediction of the data associated with ≤ 20% forage did not improve at cutoff values > 20% (results not shown), indicating that data with > 20% forage decreased the RSR of the lower-forage subset, but not the data associated with ≤ 20% forage. Based on these differences in performance and the fact that diets containing ≤ 20% forage are commonly fed to cattle in intense feedlot production systems, the data were split at 20% forage throughout the present study, which made all lower-forage data contain ≤ 18% forage and the higher-forage ≥ 25% forage.

The higher-forage CH_4 production equations overlapped with the all-data equations, where DMI and dietary NDF, starch and EE in the DMI_C, DMI+NDF_C, DMI+STA_C and DMI+EE_C equations showed regression coefficients with the same sign (Eqns. 12-15; Table 4). Moreover, similar covariates were selected for the Diet_C, Animal_C and Animal_no_DMI_C equations as for the all-data equations, although the Animal_no_DMI_C equation did not contain dietary ash (Eqns. 16-18). The higher-forage equations predicted the higher-forage subset better than the all-data equations, with mean RSR of 0.62 *vs.* 0.66 and CCC of 0.70 *vs.* 0.68, respectively, for the DMI_C, DMI+NDF_C, DMI+STA_C, DMI+EE_C, Diet_C, Animal_C and Animal_no_DMI_C equations. The developed higher-forage equations under-predicted CH_4 production at the high end and over-predicted it at the low end of production, with the multiple regression equations having less slope bias than the DMI_C equation (Fig. 2). In line with the all-data equations, models with a higher number of covariates had less between-study variance.

In accordance with the all-data and the higher-forage equations, DMI was positively related to CH_4 production in the lower-forage DMI_C equation (Eq. 20; Table 5). The DMI+NDF_C, DMI+STA_C and DMI+EE_C equations indicated no significant relationships between the corresponding dietary NDF, starch and EE contents with CH_4 production (Eqns. 21-23; *P*-values of 0.14, 0.10 and 0.57, respectively). The lower-forage DMI_C equation predicted the lower-forage subset better than the all-data equations based on RSR, whereas the highest CCC of 0.35 for the lower-forage subset were obtained from the all-data DMI+STA_C and Animal_C equations (Eqns. 3, 6; Table 2). Systematic bias, that is the sum of mean and slope bias, was less than 5.75% for these developed lower-forage equations (Table 5), except for the DMI+STA_C equation that had 3.70 and 20.18% mean and slope bias, respectively. The minor slope bias of the lower-forage DMI_C equation (≤ 0.03%) was due to under-prediction of CH_4 production at the high end and over-prediction at the low end (Fig. 3).

In contrast to the higher-forage equations, dietary NDF and starch contents in the European higher-forage DMI+NDF_C and DMI+STA_C equations were not related to CH_4 production (Eqns. 29-30, Table 6; *P*-values of 0.20 and 0.69, respectively). Furthermore, DMI, dietary NDF and EE were selected for the Diet_C equation (Eq. 32) with BW also being selected for the Animal_C equation (Eq. 33), whereas DMI and BW, and BW were selected for the Global_C and Animal_no_DMI_C equations, respectively (Eqns. 34-35). The North American higher-forage equations were largely in line with the higher-forage equations. However, the Animal_no_DMI_C equation also contained dietary ash (Eq. 44; Table 7) as obtained for the all-data equation, and the Global_C equation also contained dietary crude protein (Eq. 45), The European higher-forage and North American higher-forage equations under-predicted CH_4 at the high end and over-predicted it at the low end of production, except for the European higher-forage DMI+EE_C equation, which under-predicted CH_4 at the low end and over-predicted at the high end (Figs. 4-5). Dietary NDF and EE contents in the Brazilian higher-forage DMI+NDF_C and DMI+EE_C equations were not significantly related to CH_4 production (Eqns. 49-50, Table S2; P-values of 0.28 and 0.05, respectively), the Diet_C equation contained DMI and dietary ash (Eq. 51), whereas the Animal_no_DMI_C equation contained dietary forage (Eq. 52). Slope bias varied from 9.05 to 18.9% for the developed Brazilian higher-forage equations, except for the Animal_no_DMI_C equation for which 32.9% slope bias was obtained. Equations under-predicted CH_4 production at the low end and over-predicted at the high end, whereas the Animal_no_DMI_C equation showed a negative observed *vs*. predicted correlation (Fig. S1). Compared to the higher-forage equations, the European higher-forage, North American higher-forage and Brazilian higher-forage data were more adequately predicted by the European higher-forage (mean RSR of 0.80 *vs.* 0.85, mean CCC of 0.50 *vs.* 0.48; respectively; Tables 4, 6), North American higher-forage (mean RSR of 0.53 *vs.* 0.57, mean CCC of 0.80 *vs.* 0.77; respectively; Tables 4, 7) and Brazilian higher-forage (mean RSR of 1.13 *vs.* 1.35, respectively; Tables 4, S2), although mean CCC indicated Brazilian higher-forage data was more adequately predicted using the higher-forage than the Brazilian higher-forage equations (0.17 *vs*. 0.11, respectively; Tables 4, S2).

The IPCC (2006) Tier 2 higher-forage equation had an RSR of 0.68 and a CCC of 0.75 when evaluated using all data (Eq. 9; Table 2). Predicting the higher-forage subset with this equation resulted in RSR and CCC of 0.53 and 0.84, respectively (Eq. 9; Table 4). Despite this high accuracy of prediction of the Tier 2 approach, increased variance appeared along the unity line of the predicted *vs*. observed plots (Figs. 1-2). The IPCC Tier 2 (2006) lower-forage equation had an RSR of 1.38, a CCC of 0.17 and 59.6% mean bias for the lower-forage subset (Eq. 25; Table 5). The GLOBAL NETWORK Tier 2 equations with Y_m of 6.1% and 6.3% (Eqns. 8, 19; Tables 2, 4) performed slightly better than the IPCC Tier 2 (2006) equation for the all-data and higher-forage (sub)sets, respectively [note that the IPCC equations were validated, the GLOBAL NETWORK equations were cross-validated], whereas the lower-forage GLOBAL NETWORK Tier 2 equation with Y_m of 4.5% resulted in RSR of 0.90, a CCC of 0.43 and 0.47% of mean bias (Eq. 24) performed obviously better than the lower-forage IPCC Tier 2 equation. Although the IPCC currently uses a 10% forage cutoff, a Y_m of 4.5% is still more accurate than a Y_m of 3.0% for the present data, with RSR being 0.98 and 1.51, and CCC being 0.40 and 0.16 for the GLOBAL NETWORK and IPCC Tier 2 lower-forage equations, respectively (Eqns. 24-25). The European higher-forage and North American higher-forage subsets were associated with RSR of 0.66 and 0.48, and CCC of 0.71 and 0.88 for the IPCC Tier 2 (2006) equation, respectively (Eq. 9; Tables 6-7), whereas RSR of 1.81 and CCC of 0.21 were obtained for the Brazilian higher-forage subset (Eq. 9; Table S2). Compared to the latter equation, the GLOBAL NETWORK Tier 2 equations with Y_m of 6.6 and 6.3% performed similarly based on RSR and CCC for the European higher-forage and North American higher-forage subset (Eqns. 36, 46; Tables 6-7), whereas less mean bias was obtained with 1.89 *vs*. 3.54% and 2.51 *vs*. 8.70%, respectively. The Brazilian higher-forage subset was better predicted when using the GLOBAL NETWORK Tier 2 approach resulted in a Y_m of 5.5%, an RSR of 1.29, and a CCC of 0.28 (Eq. 53; Table S2).

Equations developed by Ellis *et al*. (2009), Charmley *et al*. (2016) and Escobar-Bahamondes *et al*. (2017a) were among the best performing extant equations and outperformed the Yan *et al*. (2000, 2009), Mills *et al*. (2003), Ellis *et al*. (2007) and Patra (2017) equations for all (sub)sets. The best performing equation of Charmley *et al*. (2016) performed better than the all-data DMI_C equation (Eqns. 1, 10; Table 2). The all-forage equation of Escobar-Bahamondes *et al*. (2017a) appeared to perform most accurately among all of the equations (Eq. 11). However, only 646 data points were available for independent evaluation. Based on RSR, it did not outperform the Animal_C equation for these 646 data points. For the higher-forage subset, the best Charmley *et al*. (2016) and the Escobar-Bahamondes *et al*. (2017a) equations performed the best based on CCC (Eqns. 20, 11; Table 4), but not on RSR. The Ellis *et al*. (2009) equation that also depended on the NDF:starch ratio (Eq. 26; Table 5) performed the best for the lower-forage data with RSR of 0.89 and CCC of 0.41. For the European higher-forage subset, the best Charmley *et al*. (2016) and the Escobar-Bahamondes *et al*. (2017a) equations (Eqns. 37, 11; Table 6) did not perform better than the Animal_C equation when just considering RSR and CCC values, although the Animal_C equation was evaluated using fewer data points. For the North American higher-forage subset, the best performing Charmley *et al*. (2016) equation (Eq. 47; Table 7) performed similarly to the Global_C equation based on RSR, whereas the Charmley *et al*. (2016) equation performed even slightly better based on CCC. Despite the accuracy of the various equations of Charmley *et al*. (2016) and in contrast to the Animal_C equations, the predicted *vs*. observed plots showed increasing variation along the unity line for all-data in particular (Fig. 1). However, the best-performing equations that were developed, which was the Animal_C equation for most subsets, did not show increasing variation along the unity line. This indicates that the best performing equations that were developed explain variation that is not captured by the Charmley *et al*. (2016) equations. These higher precisions obtained from the best performing equations is also indicated by the correlation coefficients of predicted *vs*. observed values on which the CCC is calculated (result not shown).

3.2 Methane yield equations

Positive, negative and negative slope regression coefficients were obtained for the NDF_C, STA_C and EE_C all-data CH_4 yield [g (kg DMI)$^{-1}$] equations (Eqns. 54-56, Table S3), respectively, which aligned with the all-data CH_4 production equations. The Diet_no_DMI_C and Global_no_DMI_C equations selected dietary forage (Eqns. 57-58), whereas dietary EE and ash were also selected for the Diet_no_DMI_C equation, and dietary crude protein for the Global_no_DMI_C equation. The NDF_C, STA_C, EE_C, Diet_no_DMI_C and Global_no_DMI_C equations had RSR values of 0.98, 1.06, 1.01, 0.97 and 0.96, respectively. The NDF_C, STA_C and EE_C higher-forage CH_4 yield equations indicated positive, negative and negative relationships to CH_4 yield, respectively (Eqns. 59-61, Table S4), whereas only dietary forage content was selected for the Diet_no_DMI_C equation (Eq. 62). The higher-forage CH_4 yield was associated with RSR of 1.03 to 1.21 (Table S3) when predicted by the all-data equations, whereas the higher-forage equations predicted CH_4 yield of this subset with RSR values from 0.98 to 1.04 (Table S4). The higher-forage equations reproduced the observed variation in CH_4 yield less adequately than the all-data equations, with even a negative observed *vs*. predicted relationship for the higher-forage STA_C and EE_C equations (Figs. S2-S3).

3.3 Methane intensity equations

In contrast to the CH_4 production equations, the DMI regression coefficients in the all-data DMI_C and DMI+EE_C CH_4 intensity equations [g (kg ADG)$^{-1}$] contained zero in their confidence intervals (*P*-values of 0.14 and 0.22, respectively), whereas the DMI+NDF_C and

DMI+STA_C equations had a positive regression coefficient for DMI (Eqns. 63-66; Table S5). In line with the CH_4 production equations, dietary NDF, starch and EE contents in the DMI+NDF_C, DMI+STA_C and DMI+EE_C equations had positive, negative and negative relationships with CH_4 intensity, respectively. Dietary forage content was selected for the Diet_C, Animal_no_DMI_C and Global_C equations (Eqns. 67-69), with DMI also being selected for the Diet_C equation and BW also being selected for the Global_C equation. The Diet_C, Animal_no_DMI_C and Global_C equations had RSR values of 0.99, 1.00 and 0.96, respectively, and appeared to predict the variation in CH_4 intensity most adequately (Fig. S4), whereas the other all-data CH_4 intensity equations had RSR greater than 1 and appeared to predict the variation in CH_4 intensity less adequately.

The higher-forage DMI_C, DMI+STA_C and DMI+EE_C equations did not indicate that DMI was related to CH_4 intensity (Eqns. 70, 72-73, Table S6; *P*-values of 0.06, 0.52 and 0.93, respectively). Dietary NDF was positively related to CH_4 intensity (Eq. 71), whereas dietary starch and EE contents were not related to CH_4 intensity (Eqns. 72-73; $P = 0.32$). Dietary ash content was selected for the Diet_C equation (Eq. 74), whereas BW were selected for the Animal_C and Global_C equations (Eqns. 75-76), with dietary NDF also being selected for the Global_C equation. All higher-forage CH_4 intensity equations had RSR $\geq$ 1.03. Furthermore, as also obtained for the higher-forage CH_4 yield equations, the higher-forage CH_4 intensity equations did not reproduce the observed variation in CH_4 intensity of the higher-forage subset more adequately than the all-data CH_4 intensity equations (Figs. S4-S5), which was also indicated by the RSR and CCC values.

4. Discussion

Global applicability is an important attribute of prediction equations of beef cattle enteric CH_4 emission. Various beef cattle systems that are applied world-wide may fit in our analysis. For more details about these beef cattle fattening systems, we refer to *e.g.*, De Vries *et al.* (2015), Gerssen-Gondelach *et al*. (2017) and Drouillard (2018). Our database, in which data (1021 individual records) from a variety of geographical regions across the world is represented, therefore, contributes to the overall robustness and global applicability of our all-data and higher-forage equations in particular. Hence, CH_4 production of beef cattle will be accurately predicted for data samples that represent a wider set of conditions throughout the world, which is a unique feature of the present equations. Several CH_4 prediction equations for beef cattle have been published previously, but they were developed from relatively small databases and only for one specific geographic region, such as Yan *et al*. (2009) using 108 individual animal records from 5 studies from Northern Ireland, Ellis *et al*. (2007) using 83 treatment means from 14 studies from North America, Ellis *et al*. (2009) using 872 individual animal records from 12 studies from Alberta (Canada), and Moraes *et al*. (2014) using individual records from 414 heifers and 458 steers housed at one research station in the United States.

In the present study, we collated a wide-ranging database that included a large number of studies from Europe, North America, Brazil, Australia and South Korea, which represented diverse global beef production systems. Studies from tropical areas were, however, not predominant in the present analysis, for which we refer to Charmley *et al.* (2016) who included studies from tropical Australia, and Patra (2017) who included studies from India, Zimbabwe, Australia and Brazil. Furthermore, Escobar-Bahamondes *et al*. (2017a) had a database comprising 148 treatment means from 38 studies with diets containing > 40% forage, and a database comprising 43 treatment means from 17 studies with diets containing < 20% forage. Therefore, their analysis for lower-forage diets, in particular, included more data from more studies than ours, but their cutoff values for lower and higher forage were based on differences in microbiome composition rather than the prediction error used in the present analysis. Furthermore, their analysis did not explore intercontinental variation in beef cattle CH_4 emissions and did not have the benefit of using individual animal records. Other unique strengths of the present study are the development of CH_4 yield and intensity equations, whereas beef cattle studies are commonly limited to only total CH_4 production, and the inclusion of dietary forage content as a covariate of the three CH_4 emission metrics.

Our database includes data obtained with different CH_4 (*viz.*, respiration chambers, GreenFeed system, SF_6) and DMI (*viz.*, weighing and estimating using marker techniques) measurement methods. The different CH_4 measurement techniques have their strengths and weaknesses (Hammond *et al*., 2016; Hristov *et al*., 2018), whereas directly weighing the amount of feed offered and refusals and their dry matter content is regarded as more accurate than the ytterbium and *n*-alkane markers used for some studies in the present database, which may over- or underestimate DMI (Pérez-Ramírez *et al*., 2012). However, the development of a DMI_C CH_4 production equation specific for respiration chamber, GreenFeed system and SF_6 subsets did not consistently improve the RSR and CCC of the corresponding subsets compared with the all-data DMI_C equation (results not shown). Furthermore, including CH_4 measurement method as a covariate in the statistical model of an all-data DMI_C equation did not improve the model fit. Similar results were obtained for measurement method of DMI. Therefore, CH_4 and DMI measurement methods did not have a major effect on the performance of the equations developed in the present analysis. However, the relatively high and low accuracies with which the region-specific subsets could be predicted may be related to the CH_4 measurement methods, because the percentage of use of respiration chambers in the European higher-forage, North American higher-forage and Brazilian higher-forage subsets differed substantially (48, 95 and 0%, respectively). Finally, statistically accounting for cattle breed or cattle type (*e.g.*, steers, heifers, cows) was considered, but did not or not consistently improve the prediction of CH_4 production throughout the subsets.

Non-linear CH_4 prediction equations such as the Mitscherlich equation were previously found to outperform linear equations in some studies (*e.g.*, Mills *et al.*, 2003; Patra, 2017). However, for the present database, fitting non-linear equations, *viz.*, Monomolecular, Exponential, Mitscherlich and Power forms, did not result in improved prediction of CH_4 production compared to the linear DMI_C equations (result not shown). The latter result is in line with the non-linear Mills *et al*. (2003) and Patra (2017) equations that did not outperform the linear Charmley *et al.* (2016) equations. This suggests that a multiple linear regression approach, as used for the development of our Animal_C and Global_C equations rather than non-linear approaches, improves the precision and accuracy of prediction of CH_4 production. The utility of ADG and digestibility of EE, NDF, nitrogen, gross energy, DM and organic matter for predicting CH_4 was also evaluated, but these covariates did not result in better prediction of CH_4 production than achieved by the various equations that are presented.

The linear regression equations of Charmley *et al.* (2016) that depended on DMI and outperformed our DMI_C equations were fitted using models that included more terms than just DMI, which resulted in nearly unbiased predictions of CH_4. Furthermore, the data Charmley *et al.* (2016) used were only from certain regions in Australia and may have been relatively homogeneous. The equations developed using these data may then result in accurate prediction of CH_4 production based on only DMI. Therefore, the prediction bias for our various Animal_C and

Global_C equations and some potential overestimation of between-study variance that remained may vanish by the inclusion of even more covariates in the statistical model. The negligible bias obtained for the European higher-forage Animal_C equation, for which dietary crude protein and EE contents were available for all individual animal records used for fitting this model, and suggests that multiple regression equations are associated with less bias. Despite prediction biases of mixed-effects models being associated with the inclusion of the random study effect (see also White *et al.*, 2017), which applies to models with fewer covariates in particular, omission of the random study effect will affect the inference made on the covariates and may result in type II errors (St-Pierre, 2001). Therefore, for achieving unbiased predictions, mixed-effects models are ideally applied to datasets without missing values throughout the different covariates. Such datasets will result in greater variation of the dependent variable explained by multiple fixed-effects terms and less overestimation of the random study effect.

4.1 Key predictor variables

Dry matter intake was the most important predictor of enteric CH_4 production as it was significantly and positively related to CH_4 production for all-data and the higher-forage, lower-forage, European higher-forage and North American higher-forage subsets. A positive relationship between DMI and CH_4 production is in agreement with previous dairy and beef cattle studies (*e.g.*, Ellis *et al.*, 2007; Hristov *et al.*, 2013a; Richmond *et al.*, 2015; Bell *et al.*, 2016; Charmley *et al.*, 2016; Niu *et al.*, 2018) and this is because more CH_4 is produced when more substrate is available for microbial fermentation and in turn methanogenesis. In addition, all Diet_C and Animal_C models based on these five (sub)sets selected DMI for the prediction of CH_4 production, and the Animal_no_DMI_C equations did not perform as well as the Animal_C equations, indicating the importance of DMI relative to other covariates.

The positive relationship between the all-data CH_4 production and dietary NDF content also aligns with previous results (*e.g.*, Ellis *et al*., 2007; Yan *et al*., 2009; Niu *et al*., 2018). The coefficients of variation were 43.8, 45.4 and 30.3% for CH_4 production, and 32.0, 30.6 and 8.6% for dietary NDF content for the all-data and the higher-forage and lower-forage subsets, respectively. This decrease in variation is in line with the disappearance of this positive relationship for the all-data and higher-forage *vs.* the lower-forage equations. Therefore, developing subsets with limited variation in forage percentage seems to have masked the positive relationship between CH_4 production and dietary NDF content. Furthermore, dietary nutrient contents change at the expense of other nutrients. Dietary NDF content may increase at the expense of more rapidly fermentable carbohydrates, which is positively associated with CH_4 production (Hatew *et al*., 2015). The latter hypothesis aligns with a model with DMI and dietary NDF and starch fitted to all data having regression coefficients that were positive, positive and not different from zero for DMI and dietary NDF and starch, respectively (result not shown). However, the lignin fraction of NDF being undegradable indicates that increased dietary NDF may not result in more CH_4 production in case of high lignin contents. Warner *et al*. (2016) observed lower CH_4 production but higher CH_4 yield per unit of digestible organic matter for dairy cattle fed grass silage of high lignin and NDF content, compared with grass silage of low lignin and NDF content. The observation of Na *et al*. (2017) who found different CH_4 yields per unit of DMI for deer and goats, but not per unit of digestible DMI may also support this hypothesis.

Dietary starch content is negatively related to CH_4 production as it typically increases propionate production in the rumen, yielding less H_2 for the reduction of CO_2 to CH_4 (Martin *et al*., 2010; Grainger and Beauchemin, 2011). The effect of dietary starch on CH_4 production appeared to be less pronounced for higher-forage diets in dairy cows (Van Gastelen *et al*., 2015), which may explain why no relationship between dietary starch content and CH_4 production was found for the European higher-forage subset, which had the highest forage content of all subsets. Furthermore, it was suggested that a critical dietary content of starch is required to decrease CH_4 production (Martin *et al*., 2010; Van Gastelen *et al*., 2015), possibly more than approximately 20% of DM, and that slight differences in intakes of starch, and other major carbohydrates (*e.g*., hemicellulose, cellulose and lignin) cannot explain the difference in CH_4 emissions of cattle (Moe and Tyrrell, 1979; Moate *et al*., 2018). This may also explain why no relationship between dietary starch content and CH_4 production was obtained based on the European higher-forage subset. The lack of a relationship between CH_4 production and dietary starch content for the lower-forage subset may be related to the small variation in starch content (coefficient of variation is 13.1%).

The positive relationship that was obtained between CH_4 production and dietary forage aligns with previously published studies (*e.g*., Yan *et al*., 2000; Hristov *et al*., 2013) stating that either increased forage or decreased concentrate proportion in the diet yielded more CH_4. Johnson and Johnson (1995) referred to cattle fed more than 90% concentrate producing only half of the CH_4 produced by cattle fed more common concentrate proportions, and Aguerre *et al*., 2011 observed a linear increase in CH_4 yield upon increasing dietary forage content from 47 to 68%. Nevertheless, a modeling study by Sauvant and Giger-Reverdin (2009) predicted that a decrease in CH_4 yield is only observed for dietary forage contents less than 65%. Despite the latter prediction, the frequent appearance of dietary forage in the equations developed in the present study indicates dietary forage content is a decent predictor of CH_4 emission, possibly more robust than dietary NDF content that was less frequently selected for the developed equations.

Dietary lipid content is commonly negatively related to CH_4 production (Grainger and Beauchemin, 2011). Lipids may inhibit cellulolytic bacteria, protozoal and archaeal activity, decrease NDF digestibility, and supply non-fermentable energy to the rumen, outcomes that can decrease CH_4 production (Maia *et al*., 2007; Beauchemin *et al*., 2008; Guyader *et al*., 2014). Long-chain saturated fatty acids may have a minimal inhibitive effect on archaeal activity and CH_4 production, whereas fatty acids such as C12:0 and C18:3 were found to be relatively potent reducers (Machmüller and Kreuzer, 1999; Patra, 2013). Therefore, the actual decrease in CH_4 production obtained from lipids may depend on their fatty acid composition, although this is not confirmed by all *in vivo* studies (*e.g*., Grainger and Beauchemin, 2011). More importantly, the removal of data associated with dietary lipid and oil supplements excluded data with higher dietary EE contents, which more potently decrease CH_4 production (Patra, 2013), may explain why DMI+EE_C equations did not perform better than the DMI_C equations, despite the significant relationships that were obtained for CH_4 production and yield with dietary EE content.

Dietary crude protein content being positively associated with the all-data CH_4 yield in the present analysis aligns with the observation that dietary nitrogen content is positively related to fiber digestibility (Dijkstra *et al*., 1996). However, decreased CH_4 production may only be observed from cattle fed a diet that is deficient in rumen degradable protein (Sutter *et al*., 2017). In the present study, we did not observe any relationship between crude protein content and NDF digestibility, $r = 0.04$. However, we did observe a correlation between crude protein content and organic matter and dry matter digestibility, $r = 0.42$ and $r = 0.37$, respectively. This is possibly due to higher starch degradability, which could not be verified because of the lack of starch degradability data. Van Lingen *et al*. (2018) applying a multivariate regression approach found that the methodological issues such as the structure of random-effects (co)variance matrices and the combination of fixed-effects variables affect the statistical inference regarding the relationship between dietary crude protein and CH_4 production or yield. Therefore, also based on dietary crude protein selected for only one equation, the latter relationship may not be commonly strong as well as it may not generally exist. Dietary crude protein may actually be associated with lesser CH_4 production when sufficient rumen degradable protein is fed so as not to limit fermentation in the rumen due to N shortage (Dijkstra *et al*., 2011), and may be considered a less robust predictor of CH_4 production than dietary NDF and starch.

A positive relationship between BW and CH_4 production observed in various equations in the present analysis aligns with previous cattle research (Yan *et al*., 2009; Moraes *et al*., 2014; Escobar-Bahamondes *et al*., 2017a). Demment and Van Soest (1985) and Smith and Baldwin (1974) observed rumen volume and weight proportional to BW of animals. Consequently, smaller animals ingest less feed and emit less CH_4 (Hristov *et al*., 2013b). In addition, empirical modeling (Sauvant and Nozière, 2016) and mechanistic model simulations (Huhtanen *et al*., 2015, 2016) indicated the DMI/BW ratio to be an important factor for CH_4 yield. At similar DMI, smaller cattle tend to produce less CH_4 as the passage rate from the rumen to the intestine may be higher due to a greater DMI/BW ratio. This has been confirmed in sheep for which animals yielding less CH_4 had smaller rumen size (Goopy *et al*., 2014). Therefore, BW influences DMI, and DMI and rumen volume determine the passage rate of ruminal digesta, which affects feed digestibility, rumen fermentation conditions, and ultimately CH_4 production and yield.

4.2 Best performing equations

Various equations and model categories for predicting beef cattle CH_4 emission have been applied on various subsets in the present study. The Y_m models have only one parameter and are the simplest models, the DMI_C models are still fairly simple, whereas the Animal_C models are potentially the most complex. The Animal_C model commonly performed best among all models and outperformed the GLOBAL NETWORK Tier 2 equations, except for the European higher-forage subset. The DMI appeared to be the major predictor of enteric CH_4 production in beef cattle, but may not always be available for individual animals on commercial farms, which points to the value of the Animal_no_DMI_C models. Using dietary forage content and BW as a covariate commonly improved the prediction of CH_4 compared with a DMI_C equation. Therefore, the on-farm availability of all previously mentioned variables is recommended. This availability also enables the evaluation of the effect of dietary nutrient composition on CH_4 production. Moreover, the DMI+STA_C equation (Eq. 3) appeared to perform well, and might also be used for the prediction of beef cattle CH_4 production, although this equation was based on fewer observations. Nevertheless, dietary starch content was never selected for the Diet_C, Animal_C, Animal_no_DMI_C and Global_C equations, whereas NDF content was, indicating that DMI+STA_C equations may be slightly less robust than Diet_C and Animal_C equations.

If dietary forage content is known to be > 25%, we recommend the use of the higher-forage equations, because the RSR and CCC of these equations are lower and higher, respectively, compared to the higher-forage subset evaluation of the all-data equations. Based on their predictive performance, the higher-forage Animal_C and the Escobar-Bahamondes *et al*. (2017a) equations (Eqns. 17, 11; Table 4) are specifically recommended. Despite its lower precision, the Charmley *et al*. (2016) equation (Eq. 20) will still give an accurate estimate of CH_4 production if only DMI is available. If dietary forage percentage is ≤ 18%, we recommend the Ellis *et al*. (2009) equations (Eqns. 26-27; Table 5). In addition, we recommend the lower-forage DMI_C or all-data DMI+NDF_C, DMI+STA_C and Animal_C equations that performed relatively well (Eqns. 20, 2-3, 6; Tables 2, 5). If dietary forage content is between 18 and 25%, we suggest an all-data equation that includes dietary forage, or dietary NDF or starch, because of the forage content that is commonly related to the latter two carbohydrate fractions. The European higher-forage and

North American higher-forage equations performed somewhat better on RSR and CCC than the higher-forage equations for the European higher-forage and North American higher-forage subsets, and less systematic bias was obtained for the region-specific equations. Therefore, we most strongly recommend the Diet_C, Animal_C and Charmley *et al*. (2016) equations (Eq. 32-33, 37; Table 6). For North American higher-forage data, the Global_C and Charmley *et al*. (2016) equations are recommended in particular (Eqns. 45, 47; Table 7).

Models that assumed a fixed Y_m, such as the IPCC and GLOBAL NETWORK Tier 2 equations, performed nearly as good as the developed more complex best performing equations in most cases. The Tier 2 equations may, therefore, have a high potential for predicting beef cattle CH_4 production as well, in particular for higher-forage diets, although the higher variance along the unity lines of the predicted *vs*. observed plots indicates a lack of precision. Moreover, the substantial mean bias that was obtained for the lower-forage subset in particular emphasizes the importance of an accurate estimate of Y_m. In cases where dietary forage contents are not close to the means of the present data (sub)sets, we do not recommend the use of the Y_m equations considered in the present study, but an equation that contains dietary forage, NDF or starch. A Y_m of 4.5% that was obtained for lower-forage diets may be fairly accurate given a Y_m of 5.2% that was reported for 42 treatments means with $\leq$ 17% forage (Escobar-Bahamondes *et al*., 2017b), and a Y_m of 3.8% for 34 treatments means with $\leq$ 18% forage (Escobar-Bahamondes *et al*., 2017a). Both of these studies reported 9.5% forage on average for studies collected from multiple continents. These Y_m values are all higher than the 3.0%, which the IPCC uses for $\leq$ 10% forage diets. The $\leq$ 10% forage records in the present analysis, which also had a Y_m of 4.5% suggests that the Y_m value for lower-forage diets used by the IPCC needs to be reconsidered. However, practices such as feeding steam-flaked corn (Hales *et al*., 2012) and dietary supplementation with monensin

(Appuhamy *et al*., 2013) may require alternative prediction as these diets may have a Y_m value of 3.0%. This also applies to fat supplemented diets (Grainger and Beauchemin, 2011; Patra, 2013).

For CH_4 yield predictions, the all-data NDF_C, Diet_no_DMI_C and Global_no_DMI_C CH_4 yield equations (Eqns. 54, 57-58) had RSR values < 1 and are suitable for use if dietary forage content is unknown. The Global_no_DMI_C equations may also be used if dietary forage content is known to be ≤ 18%. The higher-forage Diet_no_DMI_C CH_4 yield equation is the only equation to consider for forage contents > 25% (Eq. 62; Table S4), based on RSR values > 1 for the other higher-forage CH_4 yield equations. For forage contents between 18 and 25% we recommend an all-data equation with RSR < 1 and the highest CCC value when evaluated with all data, which is the Global_no_DMI_C yield equation (Eq. 58). Given that all CH_4 intensity equations were associated with an RSR value > 1 for the higher- and lower-forage subsets, we recommend the observed average values of 108 and 161 [g (kg ADG)$^{-1}$] for dietary forage contents of ≤ 18% and ≥ 25%, respectively. For dietary contents between 18 and 25% or if forage content is unknown we recommend the all-data Global_C equation (Eq. 69; Table S5).

5. Conclusion

Our analysis is based on the large GLOBAL NETWORK dataset comprising data from several continents and a wide variety of forage contents. As observed previously, DMI is the key factor for predicting beef cattle enteric CH_4 production. Non-linear models with DMI as the only independent variable did not outperform their counterpart linear models. However, linear models depending on DMI and dietary forage content or these two covariates plus BW commonly had an improved predictive ability. Separate equations for lower-forage (≤ 18%) and higher-forage (≥ 25%) data also improved predictive ability. Model evaluation specific to European higher-forage, North American higher-forage and Brazilian higher-forage diets compared with that of intercontinental higher-forage diet models suggests that overall enteric CH_4 production is more accurately predicted by region-specific models, although in many cases the best intercontinental and region-specific models may perform similarly. The equations developed in the present study commonly had higher precision and less prediction error with similar accuracy compared to the extant equations that were evaluated. Evaluation of CH_4 emission conversion factors indicated that region-specific and in particular dietary forage content-based Y_m values are required for adequately predicting beef cattle CH_4 production in national or global inventories.

Acknowledgments

Authors gratefully acknowledge project funding from the USDA National Institute of Food and Agriculture Federal Appropriations under Project PEN 04539 and Accession number 1000803; the Joint Programming Initiative on Agriculture, Food Security and Climate Change (FACCE-JPI)'s 'GLOBAL NETWORK' project and the 'Feeding and Nutrition Network' (http://animalscience.psu.edu/fnn) of the Livestock Research Group within the Global Research Alliance for Agricultural Greenhouse Gases (www.globalresearchalliance.org); the Sesnon Endowed Chair program (UC Davis); the Swiss Federal Office of Agriculture, Berne, Switzerland; AHDB Beef and Lamb, the Scottish Government, Defra and the devolved administrations through the UK Agricultural Greenhouse Gas Inventory Research Platform; French National Research Agency through the FACCE-JPI program (ANR-13-JFAC-0003-01); the Cooperative Research Program for Agriculture Science, (Project No. PJ013448012018), RDA, Republic of Korea; the Australian Government Department of Agriculture, Fisheries and Forestry (Carbon Farming Futures Action on the Ground program; AOTGR2-0400); the financial support of the Reducing Emissions from Livestock Research Program, the National Livestock Methane Program, Meat and Livestock Australia, CSIRO and Ridley AgriProducts Pty, Ltd; the Institute of Science and Technology in Animal Science (INCTCA 465377/2014-9), the Department of Agriculture, Food and the Marine (DAFM), Ireland (AGRI-I project); European Commission through SMEthane (FP7-SME-262270); Beef Cattle Research Council of the Canadian Cattlemen's Association; the Cofund for Monitoring & Mitigation of Greenhouse Gases from Agri- and Silvi-culture (FACCE ERA-GAS)'s project Capturing Effects of Diet on Emissions from Ruminant Systems and the Dutch Ministry of Agriculture, Nature and Food Quality (AF-EU-18010 & BO-4400159-01).

References

Aguerre, M.J., Wattiaux, M.A., Powell, J.M., Broderick, G.A., Arndt, C., 2011. Effect of forage-to-concentrate ratio in dairy cow diets on emission of methane, carbon dioxide, and ammonia, lactation performance, and manure excretion. J. Dairy Sci. 94, 3081-3093. https://doi:10.3168/jds.2010-4011

Appuhamy, J.A.D.R.N., Strathe, A.B., Jayasundara, S., Wagner-Riddle, C., France, J., Dijkstra, J., Kebreab E., 2013. Anti-methanogenic effects of monensin in dairy and beef cattle: A meta-analysis. J. Dairy Sci. 96, 5161-5173. https://doi.org/10.3168/jds.2012-5923

Bannink, A., Schijndel, M.W. van, Dijkstra, J., 2011. A model of enteric fermentation in dairy cows to estimate methane emission for the Dutch National Inventory Report using the IPCC Tier 3 approach. Anim. Feed Sci. Technol. 166-167, 603-618. https://doi.org/10.1016/j.anifeedsci.2011.04.043

Basarab, J.A., Okine, E.K., Baron, V.S., Marx, T., Ramsey, P., Ziegler, K., Lyle, K., 2005. Methane emissions from enteric fermentation in Alberta’s beef cattle population. Can. J. Anim. Sci. 85, 501-512. https://doi.org/10.4141/A04-069

Beauchemin, K.A., Kreuzer, M., O’Mara, F.O., McAllister, T.A., 2008. Nutritional management for enteric methane abatement: a review. Austr. J. Exp. Agric. 48, 21-27. https://doi.org/10.1071/EA07199

Bell, M., Eckard, R., Moate, P.J., Yan, T., 2016. Modelling the effect of diet composition on enteric methane emissions across sheep, beef cattle and dairy cows. Animals, 6, 54. https://doi.org/10.3390/ani6090054

Benchaar, C., Hassanat, F., Martineau, R., Gervais, R., 2015. Linseed oil supplementation to dairy cows fed diets based on red clover silage or corn silage: Effects on methane production, rumen fermentation, nutrient digestibility, N balance, and milk production. J. Dairy Sci. 98, 7993-8008. https://doi.org/10.3168/jds.2015-9398

Bibby, J., Toutenburg, T., 1977. Prediction and Improved Estimation in Linear Models. John 785 Wiley Sons, Chichester.

Castelan-Ortega, O.A., Ku-Vera, J.C., Estrada-Flores, J.G., 2014. Modeling methane emissions and methane inventories for cattle production systems in Mexico. Atmosfera 27, 185-191. https://doi.org/10.1016/S0187-6236(14)71109-9

Charmley, E., Williams, S.R.O., Moate, P.J., Hegarty, R.S., Herd, R.M., Oddy, V.H., Reyenga, P., Staunton, K.M., Anderson, A., Hannah, M.C. 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

Demment, M.W., Van Soest, P.J., 1985. A nutritional explanation for body-size patterns of ruminant and non-ruminant herbivores. American Naturalist 125, 641-672. https://doi.org/10.1086/284369

Dijkstra, J., France, J., Assis, A.G., Neal, H.D.St.C., Campos, O.F., Aroeira, L.J.M., 1996. Simulation of digestion in cattle fed sugarcane: prediction of nutrient supply for milk production with locally available supplements. J. Agric. Sci. 127, 247-260. https://doi.org/10.1017/S0021859600078023

Dijkstra, J., Oenema, O., Bannink, A., 2011. Dietary strategies to reducing N excretion from cattle: Implications for methane emissions. Curr. Opin. Environ. Sustain. 3, 414-422. https://doi.org/10.1016/j.cosust.2011.07.008

Drouillard, J.S., 2018. Current situation and future trends for beef production in the United States of America - A review. Asian-Australas. J. Anim. Sci. 31, 1007-1016. https://doi.org/10.5713/ajas.18.0428

Ellis, J.L., Kebreab, E., Odongo, N.E., McBride, B.W., Okine, E.K., France, J., 2007. Prediction of methane production from dairy and beef cattle. J. Dairy Sci. 90, 3456-3467. https://doi.org/10.3168/jds.2006-675

Ellis, J.L., Kebreab, E., Odongo, N.E., Beauchemin, K.A., McGinn, S., Nkrumah, J.D., Moore, S.S., Christopherson, R., Murdoch, G.K., McBride, B.W., Okine, E.K., France, J., 2009. Modeling methane production from beef cattle using linear and nonlinear approaches. J. Anim. Sci. 87, 1334-1345. https://doi.org/10.2527/jas.2007-0725

Escobar-Bahamondes, P., Oba, M., Beauchemin, K.A., 2017a. Universally applicable methane prediction equations for beef cattle fed high- or low-forage diets. Can. J. Anim. Sci. 97, 83-94. https://doi.org/10.1139/cjas-2016-0042

Escobar-Bahamondes, P., Oba, M., Beauchemin, K.A., 2017b. An evaluation of the accuracy and precision of methane prediction equations for beef cattle fed high-forage and high-grain diets. Animal 11, 68-77. https://doi.org/10.1017/S175173111600121X

Gastelen, S. van, Antunes-Fernandes, S.C., Hettinga, K.A., Klop, G., Alferink, S.J.J., Hendriks, W.H., Dijkstra, J., 2015. Enteric methane production, rumen volatile fatty acid concentrations, and milk fatty acid composition in lactating Holstein-Friesian cows fed grass silage- or corn silage-based diets. J. Dairy Sci. 98, 1915-1927. https://doi.org/10.3168/jds.2014-8552

Gerber, P.J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., Falcucci, A., Tempio, G, 2013. *Tackling climate change through livestock – A global assessment of emissions and mitigation opportunities*. Food and Agriculture Organization of the United Nations (FAO), Rome.

Gerssen-Gondelach, S.J., Lauwerijssen, R.B.G., Havlík, P., Herrero, M., Valin, H., Faaij, A.P.C., Wicke, B., 2017. Intensification pathways for beef and dairy cattle production systems: Impacts on GHG emissions, land occupation and land use change. Agric. Ecosyst. Environ. 240, 135-147. https://doi.org/10.1016/j.agee.2017.02.012

Goopy, J., Donaldson, A., Hegarty, R., Vercoe, P., Haynes, F., Barnett, M., Oddy, V., 2014. Low-methane yield sheep have smaller rumens and shorter rumen retention time. Brit. J. Nutr. 111, 578-585. https://doi.org/10.1017/S0007114513002936

Grainger, C., Beauchemin, K.A., 2011. Can enteric methane emissions from ruminants be lowered without lowering their production? Anim. Feed Sci. Technol. 166-167, 308-320. https://doi.org/10.1016/j.anifeedsci.2011.04.021

Guyader, J., Eugène, M., Nozière, P., Morgavi, D., Doreau, M., Martin, C., 2014. Influence of rumen protozoa on methane emission in ruminants: A meta-analysis approach. Animal 8, 1816-1825. https://doi.org/10.1017/S1751731114001852

Hammond, K.J., Crompton, L.A., Bannink, A., Dijkstra, J., Yáñez-Ruiz, D.R., O'Kiely, P., Kebreab, E., Eugène, M.A., Yu, Z., Shingfield, K.J., Schwarm, A., Hristov, A.N., Reynolds, C.K., 2016. Review of current *in vivo* measurement techniques for quantifying enteric methane emission from ruminants. Anim. Feed Sci. Technol. 219, 13-30. https://doi.org/10.1016/j.anifeedsci.2016.05.018

Hales, K.E., Cole, N.A., MacDonald, J.C., 2012. Effects of corn processing method and dietary inclusion of wet distillers grains with solubles on energy metabolism, carbon-nitrogen balance, and methane emissions of cattle. J. Anim. Sci. 90, 3174-3185. https://doi.org/10.2527/jas.2011-4441

Hatew, B., Cone, J.W., Pellikaan, W.F., Podesta, S.C., Bannink, A., Hendriks, W.H., Dijkstra, J., 2015. Relationship between in vitro and in vivo methane production measured simultaneously with different dietary starch sources and starch levels in dairy cattle. Anim. Feed Sci. Technol. 202, 20-31. https://doi.org/10.1016/j.anifeedsci.2015.01.012

Hristov, A.N., Oh, J., Firkins, J., Dijkstra, J., Kebreab, E., Waghorn, G., Makker, M.P.S., Adesogan, A.T., Yang, W., Lee, C., Gerber, P.J., Henderson, B., Tricarico, J.M., 2013a. 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

Hristov, A.N., Ott, J., Tricarico, J.M., Rotz, A., Waghorn, G., Adesogan, A.T., Dijkstra, J., Montes, F., Oh, J., Kebreab, E., Oosting, S.J., Gerber, P.J., Henderson, B., Makker, H.P.S., Firkins, J., 2013b. SPECIAL TOPICS - Mitigation of methane and nitrous oxide emissions from animal operations: III. A review of animal management mitigation options. J. Anim. Sci. 91, 5095-5113. https://doi.org/10.2527/jas.2013-6585

Hristov, A.N., Kebreab, E., Niu, M., Oh, J., Bannink, A., Bayat, A.R., Boland, T., Brito, A.F., Casper, D., Crompton, L.A., Dijkstra, J., Eugène, M.A., Garnsworthy, Ph.C., Haque, M.N., Hellwing, A.L.F., Huhtanen, P., Kreuzer, M., Kuhla, B., Lund, P., Madsen, J., Martin, C., Moate, P.J., Muetzel, S., Muñoz, C., Peiren, N., Powell, J.M., Reynolds, C.K., Schwarm, A., Shingfield, K.J., Storlien, T.M., Weisbjerg, M.R., Yáñez-Ruiz, D.R., Yu, Z., 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

Huhtanen, P., Ramin, M., Udén, P., 2015. Nordic dairy cow model Karoline in predicting methane emissions: 1. Model description and sensitivity analysis. Livest. Sci. 178, 71-80. https://doi.org/10.1016/j.livsci.2015.05.009

Huhtanen, P., Ramin, M., Cabezas-Garcia, E.H., 2016. Effects of ruminal digesta retention time on methane emissions: a modelling approach. Anim. Prod. Sci. 56, 501-506. https://doi.org/10.1071/AN15507

IPCC (2006. *2006 IPCC Guidelines for National Greenhouse Gas Inventories*. Intergovernmental Panel on Climate Change, IGES, Kanagawa, Japan.

IPCC (2014. *2013 Revised Supplementary Methods and Good Practice Guidance Arising from the Kyoto Protocol*, Hiraishi, T., Krug, T., Tanabe, K., Srivastava, N., Baasansuren, J., Fukuda, M., Troxler, T.G. (eds) IPCC, Switzerland.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2014. *An introduction to statistical learning: With applications in R*. Springer New York, US.

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

Kebreab, E., Johnson, J.A., Archibeque, S.L., Pape, D., Wirth, T., 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

Kutner, M.H., Nachtsheim, C., Neter, J., Li, W., 2005. *Applied linear statistical models*. Boston: McGraw-Hill Irwin. New York, US.

Lin, L.I., 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics, 45, 255-268. https://doi.org/10.2307/2532051

Lingen H.J., van, Fadel, J.G., Bannink, A., Dijkstra, J., Tricarico, J., Pacheco, D., Casper, D.P., Kebreab, E., 2018. Multi-criteria evaluation of dairy cattle feed resources and animal characteristics for nutritive and environmental impacts. Animal 12, s310-s320. https://doi.org/10.1017/S1751731118001313

Machmüller, A., Kreuzer, M., 1999. Methane suppression by coconut oil and associated effects on nutrient and energy balance in sheep. Can. J. Anim. Sci. 79, 65-72. https://doi.org/10.4141/A98-079

Maia, M.R.G., Chaudhary, L.C., Figueres, L., Wallace, R.J., 2007. Metabolism of polyunsaturated fatty acids and their toxicity to the microflora of the rumen. Anton. Leeuw. 91, 303-314. https://doi.org/10.1007/s10482-006-9118-2

Martin, C., Morgavi, D.P., Doreau, M., 2010. Methane mitigation in ruminants: from microbe to the farm scale. Animal 4, 351-365. https://doi.org/10.1017/S1751731109990620

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

Moate, P.J., Williams, S.R.O., Deighton, M.H., Hannah, M.C., Ribaux, B.E., Morris, G.L., Jacobs, J.L., Hill, J., Wales, W.J., 2019. Effects of feeding wheat or corn and of rumen fistulation on milk production and methane emissions of dairy cows. Anim. Prod. Sci. 59, 891-905. https://doi.org/10.1071/AN17433

Moe, P.W., Tyrrell, H.F., 1979. Methane production in dairy cows. J. Dairy Sci. 62, 1583-1586. https://doi.org/10.3168/jds.S0022-0302(79)83465-7

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

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. T. ASABE 50, 885-900. https://doi.org/10.13031/2013.23153

Na, Y., Li, D.H., Lee, S.R., 2017. Effects of dietary forage-to-concentrate ratio on nutrient digestibility and enteric methane production in growing goats (*Capra hircus hircus*) and Sika deer (*Cervus nippon hortulorum*). Asian-Australas. J. Anim. Sci. 7, 967-972. https://doi.org/10.5713/ajas.16.0954

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

Opio, C., Gerber, P., Mottet, A., Falcucci, A., Tempio, G., MacLeod, M., Vellinga, T., Henderson, B., Steinfeld, H., 2013. Greenhouse gas emissions from ruminant supply chains – A global life cycle assessment. Food and Agriculture Organization of the United Nations (FAO), Rome.

Patra, A.K., 2013. The effect of dietary fats on methane emissions, and its other effects on digestibility, rumen fermentation and lactation performance in cattle: A meta-analysis. Livest. Prod. Sci. 155, 244-254. https://doi.org/10.1016/j.livsci.2013.05.023

Patra, A.K., 2017. Prediction of enteric methane emission from cattle using linear and non-linear statistical models in tropical production systems. Mitig. Adapt. Strateg. Glob. Change 22, 629-650. https://doi.org/10.1007/s11027-015-9691-7

Pérez-Ramírez, E., Peyraud, J.L., Delagarde, R., 2012. *N*-alkanes *v*. ytterbium/faecal index as two methods for estimating herbage intake of dairy cows fed on diets differing in the herbage:maize silage ratio and feeding level. Animal 6, 232-244. https://doi.org/10.1017/S1751731111001480

Pinheiro, J.C., Bates, D.M., 2000. *Mixed-effects models in S and S-PLUS*. Springer Verlag, New York, Inc.

Popp, A., Lotze-Campen, H., Bodirsky, B., 2010. Food consumption, diets shifts and associated non-CO_2 greenhouse gases from agricultural production. Glob. Environ. Change 20, 451-462. https://doi.org/10.1016/j.gloenvcha.2010.02.001

R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Richmond, A.S., Wylie, A.R.G., Laidlaw, A.S., Lively, F.O., 2015. Methane emissions from beef cattle grazing on semi-natural upland and improved lowland grasslands. Animal 9, 130-137. https://doi.org/10.1017/S1751731114002067

Santiago-Suarez, B., Moraes, L.E., Appuhamy, J.A.D.R.N., Pellikaan, W.F., Casper, D.P., Tricarico, J., Kebreab, E., 2016. Prediction and evaluation of enteric methane emissions from lactating dairy cows using different levels of covariate information. Anim. Prod. Sci. 56, 557-564. https://doi.org/10.1071/AN15496

Sauvant, D., Giger-Reverdin, S., 2009. Modélisation des interactions digestives et de la production de méthane chez les ruminants. (In French.) INRA Prod. Anim. 22, 375-384.

Sauvant, D., Nozière, P., 2016. Quantification of the main digestive processes in ruminants: the equations involved in the renewed energy and protein feed evaluation systems. Animal 10, 755-770. https://doi.org/10.1017/S1751731115002670

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

Sutter, F., Schwarm, A., Kreuzer, M., 2017. Development of nitrogen and methane losses in the first eight weeks of lactation in Holstein cows subjected to deficiency of utilisable crude protein under restrictive feeding conditions. Arch. Anim. Nutr. 71, 1-20. https://doi.org/10.1080/1745039X.2016.1258880

Vries, M. de, Middelaar, C.E. van, Boer, I.J.M. de, 2015. Comparing environmental impacts of beef production systems: A review of life cycle assessments. Livest. Sci. 178, 279-288. https://doi.org/10.1016/j.livsci.2015.06.020

Warner, D., Hatew, B., Podesta, S.C., Klop, G., van Gastelen, S., van Laar, H., Dijkstra, J., Bannink, A., 2016. Effects of nitrogen fertilisation rate and maturity of grass silage on methane emission by lactating dairy cows. Animal 10, 34-43. https://doi.org/10.1017/S1751731115001640

White, R.R., Roman-Garcia, Y., Firkins, J.L., VanDeHaar, M.J., Armentano, L.E., Weiss, W.P., McGill, T., Garnett, R, Hanigan, M.D., 2017. Evaluation of the National Research Council (2001) dairy model and derivation of new prediction equations. 1. Digestibility of fiber, fat, protein, and non-fiber carbohydrate. J. Dairy Sci. 100, 3591-3610. https://doi.org/10.3168/jds.2015-10800

Yan, T., Agnew, R.E., Gordon, F.J., Porter, M.G., 2000. 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

Yan, T., Porter, M.G., Mayne, C.S., 2009. Prediction of methane emission from beef cattle using data measured in indirect open-circuit respiration calorimeters, Animal 3, 1455-1462. https://doi.org/10.1017/S175173110900473X

Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1, 3-14. https://doi.org/10.1111/j.2041-210X.2009.00001.x

Zwillinger, D., Kokoska, S., 2000. *CRC Standard Probability and Statistics Tables and Formulae*, CRC Press, Boca Raton, US.

Table 1. Variable summary statistics for all data, higher-forage (data associated with a forage content ≥ 25%), lower-forage (data associated with a forage content ≤ 18%), European higher-forage (EUR-HF) and North American higher-forage (NrAm-HF) entries of the GLOBAL NETWORK beef cattle database.

	All data (*n* = 1021)				Higher-forage (*n* = 882)				Lower-forage (*n* = 139)				EUR-HF (*n* = 307)				NrAm-HF (*n* = 394)			
Item*	Mean	Min	Max	SD	Mean	Min	Max	SD	Mean	Min	Max	SD	Mean	Min	Max	SD	Mean	Min	Max	SD
DMI (kg d^{-1})	8.13	2.26	17.5	2.82	7.81	2.26	17.5	2.86	10.1	4.77	15.6	1.89	9.46	3.17	15.7	2.12	6.16	2.26	14.1	2.51
GEI (MJ d^{-1})	150	42.6	317	53.7	144	42.6	317	53.1	191	88.2	300	37.1	177	57.8	299	42.0	114	42.6	254	45.4
Diet composition (% of DM)																				
CP	14.6	6.19	22.5	2.56	14.6	6.19	21.3	2.60	14.6	11.4	22.5	2.35	14.3	7.80	19.2	1.84	15.6	10.0	21.3	2.52
EE	3.02	0.372	7.02	1.20	2.87	0.372	7.02	1.05	3.90	0.377	5.63	1.57	3.65	0.372	5.80	1.40	2.58	0.669	5.50	0.825
Ash	6.29	3.22	13.7	2.11	6.52	3.22	13.7	2.13	4.85	3.50	8.00	1.24	6.06	3.40	11.4	1.80	6.51	3.22	13.7	2.29
NDF	35.0	17.2	73.9	11.2	36.6	17.2	73.9	11.2	24.7	19.8	33.3	3.12	37.5	26.1	68.4	7.25	32.9	17.5	67.8	9.76
ADF	19.3	6.92	50.8	8.13	20.6	7.50	50.8	8.06	11.6	6.92	14.5	1.99	21.7	14.0	40.3	5.15	17.5	7.50	36.5	7.28
STA	34.0	2.50	64.1	13.6	32.2	2.50	64.1	14.3	42.0	32.0	56.9	4.52	25.3	2.50	40.3	10.3	41.2	16.8	64.1	12.6
For	51.0	8.0	100	27.7	57.7	25.0	100	24.0	9.8	8.0	18.1	3.18	64.6	31.0	100	16.9	47.9	25.0	100	23.6
ADG (kg d^{-1})	1.25	0.060	3.38	0.431	1.19	0.060	3.38	0.438	1.46	0.552	2.22	0.330	1.22	0.088	1.99	0.348	NA	NA	NA	NA
BW (kg)	478	133	791	148	454	133	791	144	625	376	734	76.5	571	133	791	128	391	196	699	116
Methane emissions																				
CH_4 (g d^{-1})	161	37.0	372	70.5	162	37.0	372	73.5	153	45.1	310	46.4	215	40.9	372	71.8	125	37.0	313	57.5
CH_4/DMI (g kg^{-1})	20.0	6.29	35.1	5.05	20.7	6.29	35.1	4.75	15.2	7.50	30.9	4.29	22.5	6.64	35.1	5.19	20.3	6.29	33.3	4.21
CH_4/ADG (g kg^{-1}) ♪	4.98	3.31	6.68	0.522	5.08	3.31	6.68	0.528	4.68	3.84	5.71	0.362	5.20	3.31	6.68	0.517	NA	NA	NA	NA
Y_m (% of GEI)§	6.0	1.9	10.4	1.5	6.3	1.9	10.4	1.4	4.5	2.3	8.7	1.2	6.7	2.0	10.3	1.5	6.2	1.9	10.4	1.3

*DM = dry matter, DMI = dry matter intake, GEI = gross energy intake, CP = dietary crude protein, EE = dietary ether extract, NDF = dietary neutral detergent fiber, ADF = dietary acid detergent fiber, STA = dietary starch, For = dietary forage, ADG = average daily body weight gain, BW = body weight.

‡Min = minimum, Max = maximum, SD = standard deviation.

♪ ln transformed values.

§Methane conversion factor (%): energy of CH_4 as a proportion of GEI; the specific energy of CH_4 is 55.65 MJ kg^{-1}.

Table 2. All-data CH$_4$ emission (g d^{-1} animal^{-1}) prediction equations for various categories and model performance across the data (sub)sets based on root mean square prediction error (RMSPE; % of mean), RMSPE-observations-standard-deviation-ratio (RSR), mean and slope bias (MB and SB; % of mean square prediction error), and concordance correlation coefficient (CCC).

Model development				Model performance						
Eq.	Category§	Prediction equation*	*n*†	(Sub)set‡	*p*ǂ	RMSPE, %	RSR	MB, %	SB, %	CCC
[1]	DMI_C	54.2 (7.6) + 12.6 (0.6) × DMI	1021	All-data	991	31.2	0.71	0.69	12.55	0.60
				Higher-forage	852	30.8	0.67	3.86	27.26	0.64
				Lower-forage	139	33.9	1.12	34.96	0.59	0.27
[2]	DMI+NDF_C	–16.4 (9.0) + 12.1 (0.6) × DMI + 2.10 (0.16) × NDF	1021	All-data	991	31.4	0.71	0.92	1.99	0.63
				Higher-forage	852	31.8	0.69	1.82	2.81	0.65
				Lower-forage	139	28.7	0.94	3.45	2.10	0.32
[3]	DMI+STA_C	126 (11) + 11.5 (0.9) × DMI – 1.75 (0.16) × STA	704	All-data	704	28.9	0.71	6.09	1.01	0.65
				Higher-forage	575	28.7	0.70	13.87	1.80	0.68
				Lower-forage	129	30.1	0.96	15.77	0.64	0.35
[4]	DMI+EE_C	83.0 (9.8) + 11.9 (0.6) × DMI – 7.31 (1.69) × EE	754	All-data	754	29.4	0.71	1.35	8.83	0.61
				Higher-forage	644	29.2	0.67	0.08	19.59	0.64
				Lower-forage	110	30.3	1.21	37.59	4.20	0.25
[5]	Diet_C	–0.767 (7.493) + 12.0 (0.5) × DMI + 1.12 (0.06) × For	1021	All-data	991	29.5	0.67	2.39	1.24	0.70
				Higher-forage	852	29.5	0.64	1.30	2.12	0.72
				Lower-forage	139	29.8	0.98	17.57	0.30	0.32
[6]	Animal_C, Global_C		1003	All-data	991	26.9	0.61	2.20	1.37	0.76
				All-data♪,♯	646	22.5	0.52	3.05	0.41	0.84

		–28.3 (8.3) + 10.3 (0.6) × DMI + 1.12 (0.06) × For + 0.0885 (0.0150) × BW		Higher-forage	852	26.8	0.58	2.00	1.83	0.78
				Lower-forage	139	27.8	0.91	3.72	0.65	0.35
[7]	Animal_no_DMI_C	6.03 (10.40) + 1.25 (0.07) × For – 2.29 (0.77) × Ash + 0.212 (0.015) × BW	992	All-data	991	30.9	0.70	0.55	1.98	0.65
				Higher-forage	852	30.9	0.67	0.23	3.04	0.68
				Lower-forage	139	31.3	1.03	5.86	3.26	0.11
[8]	GLOBAL NETWORK Tier 2	[0.061 (0.001) × GEI] / 0.05565	1021	All-data	991	28.5	0.64	0.59	0.27	0.75
				All-data♪,♫	991	28.3	0.64	0.23	0.87	0.76
[9]	IPCC Tier 2 (2006) ¶	(0.065 × GEI) / 0.05565	-	All-data♪	991	29.9	0.68	7.97	3.04	0.75
[10]	Charmley *et al*. (2016)	–6.10 + 20.6 × DMI		All-data♪,‖	939	28.9	0.66	0.00	1.15	0.74
[11]	Escobar-Bahamondes *et al*. (2017a)	–35.0 + 0.08 × BW + 1.2 × For – 69.8 × EEI^3 + 3.14 × GEI		All-data♪,♯	646	23.2	0.54	11.66	2.12	0.85

§ Category acronyms (*e.g*., DMI_C) are explained in the 'Model development' subsection of the 'Methods and Materials' section.

* Equations are presented with regression coefficient standard errors in parenthesis; DMI = dry matter intake (kg d^{-1}), NDF = dietary neutral detergent fiber (% of DM), STA = dietary starch (% of DM), EE = dietary ether extract (% of DM), Ash = dietary ash (% of DM), For = dietary forage (% of DM), BW = body weight (kg), GEI = gross energy intake (MJ d^{-1}), EEI = ether extract intake (kg d^{-1}).

†*n* = number of observations used to fit model equations

‡All-data = all data collected for analysis, Higher-forage = data associated with a forage content ≥ 25%, Lower-forage = data associated with a forage content ≤ 18%.

ǂ*p* = numbers of observations used for model evaluation.

¶IPCC = Intergovernmental Panel on Climate Change.

♪Performance was evaluated, not cross-validated.

♫No independent evaluation.

‖The 991 data points minus data from Tomkins *et al*. (2011) and Kennedy and Charmley (2012) to ensure independent evaluation.

♯The 991 data points minus data from Pinares-Patiño *et al*. (2003), Chaves *et al*. (2006), McGeough *et al*. (2010ab), Doreau *et al*. (2011), Staerfl *et al*. (2012), Hünerberg *et al*. (2013ab) and Troy *et al*. (2015) to ensure independent evaluation.

Table 3. Root mean square prediction error-standard-deviation-ratio (RSR) of the DMI_C CH_4 production (g d^{-1}) equations based on lower-forage (≤ forage content cutoff) and higher-forage (> forage content cutoff) subsets, their number of observations (*n*), and the average RSR weighted to the number of higher- and lower-forage observations (All) for various diet forage content cutoff values to split the entire dataset into lower-forage and higher-forage subsets.

(Sub)set	Forage content cutoff (% of DM)								
	0	15	20	25	30	35	40	45	50
Higher-forage	NA	0.65	0.64	0.65	0.68	0.68	0.69	0.70	0.62
n	NA	902	882	783	672	664	602	579	474
Lower-forage	NA	0.95	0.94	0.78	0.69	0.70	0.68	0.68	0.73
n	NA	119	139	238	349	357	419	442	547
All	0.71	0.68	0.68	0.68	0.69	0.69	0.69	0.69	0.68

Table 4. Higher-forage CH_4 emission (g d^{-1} animal^{-1}) prediction equations for various categories and model performance across the data subsets based on root mean square prediction error (RMSPE; % of mean), RMSPE-observations-standard-deviation-ratio (RSR), mean and slope bias (MB and SB; % of mean square prediction error), and concordance correlation coefficient (CCC).

Model development				Model performance						
Eq.	Category§	Prediction equation*	*n*†	(Sub)set‡	*p*ǂ	RMSPE, %	RSR	MB, %	SB, %	CCC
[12]	DMI_C	52.8 (7.5) + 13.8 (0.6) × DMI	882	Higher-forage	852	29.3	0.64	0.28	22.78	0.68
				EUR-HF	307	29.3	0.88	26.42	15.91	0.43
				NrAm-HF	394	27.3	0.59	16.37	27.26	0.75
				BRZ-HF	75	26.1	1.40	31.70	22.96	0.23
[13]	DMI+NDF_C	23.8 (9.1) + 13.5 (0.6) × DMI + 0.844 (0.165) × NDF	882	Higher-forage	852	29.1	0.64	0.20	17.83	0.69
				EUR-HF	307	29.0	0.87	27.08	14.11	0.45
				NrAm-HF	394	25.2	0.55	13.46	21.70	0.80
				BRZ-HF	75	24.9	1.34	24.24	26.05	0.25
[14]	DMI+STA_C	83.4 (11.4) + 13.6 (0.8) × DMI – 0.594 (0.161) × STA	575	Higher-forage	575	26.6	0.65	1.42	16.37	0.68
				EUR-HF	273	26.0	0.93	21.41	4.37	0.35
				NrAm-HF	269	24.7	0.56	16.08	20.02	0.78
				BRZ-HF	14	29.2	2.47	77.05	9.49	-0.04
[15]	DMI+EE_C	66.4 (9.5) + 13.3 (0.6) × DMI – 3.69 (1.56) × EE	644	Higher-forage	644	27.8	0.64	1.32	15.45	0.69
				EUR-HF	122	26.1	1.00	43.22	6.45	0.39
				NrAm-HF	394	28.4	0.61	18.29	28.79	0.73
				BRZ-HF	104	24.0	1.33	30.48	19.06	0.23
[16]	Diet_C		882	Higher-forage	852	27.9	0.61	0.49	15.52	0.72

		23.4 (8.1) + 13.2 (0.5) × DMI + 0.571 (0.080) × For		EUR-HF	307	27.9	0.84	25.38	12.04	0.49
				NrAm-HF	394	23.1	0.50	8.40	17.91	0.83
				BRZ-HF	75	22.6	1.21	14.48	25.25	0.28
[17]	Animal_C, Global_C	–6.41 (8.31) + 11.3 (0.6) × DMI + 0.557 (0.077) × For + 0.0996 (0.0142) × BW	864	Higher-forage	852	24.6	0.54	0.80	14.74	0.80
				Higher-forage♯	567	21.2	0.47	0.11	10.43	0.86
				EUR-HF	307	24.5	0.73	18.25	11.11	0.61
				NrAm-HF	394	20.3	0.44	3.82	11.92	0.88
				BRZ-HF	75	21.2	1.14	1.36	27.11	0.22
[18]	Animal_no_DMI_C	17.9 (10.4) + 0.732 (0.091) × For + 0.226 (0.015) × BW	864	Higher-forage	852	30.8	0.67	0.01	13.23	0.65
				EUR-HF	307	26.3	0.79	14.80	11.89	0.52
				NrAm-HF	394	33.1	0.72	18.90	5.82	0.65
				BRZ-HF	75	27.0	1.45	31.74	25.94	-0.14
[19]	GLOBAL NETWORK Tier 2	[0.063 (0.002) × GEI] / 0.05565	882	Higher-forage	852	24.6	0.54	2.12	1.37	0.82
				Higher-forage♪,♫	852	24.0	0.52	0.14	0.62	0.83
[9]	IPCC Tier 2 (2006) ¶	(0.065 × GEI) / 0.05565	-	Higher-forage♪	852	24.1	0.53	1.19	0.06	0.84
[20]	Charmley *et al*. (2016)	21.0 × DMI		Higher-forage ♪,‖	829	25.4	0.57	0.10	0.05	0.81
[11]	Escobar-Bahamondes *et al*. (2017a)	–35.0 + 0.08 × BW + 1.2 × For – 69.8 × EEI^3 + 3.14 × GEI		Higher-forage♪,♯	567	23.1	0.51	9.61	1.56	0.86

§ Category acronyms (*e.g.*, DMI_C) are explained in the ‘Model development’ subsection of the ‘Methods and Materials’ section.

* Equations are presented with regression coefficient standard errors in parenthesis; DMI = dry matter intake (kg d^{-1}), NDF = dietary neutral detergent fiber (% of DM), STA = dietary starch (% of DM), EE = dietary ether extract (% of DM), For = dietary forage (% of DM), BW = body weight (kg), GEI = gross energy intake (MJ d^{-1}), EEI = ether extract intake (kg d^{-1}).

†*n* = number of observations used to fit model equations.

‡Higher-forage = data associated with a forage content ≥ 25%, EUR-HF = European data associated with a forage content ≥ 25%, NrAm-HF = North American data associated with a forage content ≥ 25%, BRZ-HF = Brazilian data associated with a forage content ≥ 25%.

ǂp = numbers of observations used for model evaluation.

¶IPCC = Intergovernmental Panel on Climate Change.

♪Performance was evaluated, not cross-validated.

♫No independent evaluation.

‖The 852 data points minus data from Tomkins *et al*. (2011) and Kennedy and Charmley (2012) to ensure independent evaluation.

♯The 852 data points minus data from Pinares-Patiño *et al*. (2003), Chaves *et al*. (2006), McGeough *et al*. (2010ab), Doreau *et al*. (2011), Staerfl *et al*. (2012), Hünerberg *et al*. (2013ab) and Troy *et al.* (2015) to ensure independent evaluation.

Table 5. Lower-forage CH_4 emission (g d^{-1} $animal^{-1}$) prediction equations and model performance using the lower-forage subset based on root mean square prediction error (RMSPE; % of mean), RMSPE-observations-standard-deviation-ratio (RSR), mean and slope bias (MB and SB; % of mean square prediction error), and concordance correlation coefficient (CCC).

Model development				Model performance						
Eq.	Category§	Prediction equation*	*n*†	(Sub)set‡	*p*ǂ	RMSPE, %	RSR	MB, %	SB, %	CCC
[20]	DMI_C, Diet_C, Animal_C, Global_C	46.6 (19.4) + 9.54 (1.80) × DMI	139	Lower-forage	139	28.4	0.94	4.81	0.03	0.26
[21]	DMI+NDF_C	112 (47) + 9.46 (1.79) × DMI – 2.58 (1.72) × NDF	139	Lower-forage	139	29.3	0.96	4.74	1.34	0.25
[22]	DMI+STA_C	42.0 (38.9) + 9.85 (1.88) × DMI + 0.0331 (0.7546) × STA	129	Lower-forage	129	34.5	1.11	3.70	20.18	0.23
[23]	DMI+EE_C	57.0 (18.1) + 8.84 (1.74) × DMI – 1.17 (2.03) × EE	110	Lower-forage	110	24.1	0.96	4.54	1.22	0.26
[24]	GLOBAL NETWORK Tier 2	[0.045 (0.002) × GEI] / 0.05565	139	Lower-forage	139	27.9	0.92	3.13	3.02	0.39
				Lower-forage♪,♫	139	27.3	0.90	0.47	3.99	0.43
				Lower-forage♪,‖	101	25.2	0.98	2.27	12.91	0.41
[25]	IPCC Tier 2 (2006) Lower-forage¶	(0.030 × GEI) / 0.05565	-	Lower-forage♪	139	42.1	1.38	59.60	0.08	0.17
				Lower-forage♪,‖	101	39.0	1.51	64.08	0.35	0.16
[26]	Ellis *et al*. (2009); Eq. N	48.2 + 14.1 × DMI – 20.5 × (STA/NDF)		Lower-forage♪	129	27.8	0.89	0.26	2.04	0.41
[27]	Ellis *et al*. (2009); Eq. A	41.2 + 12.0 × DMI		Lower-forage♪	139	27.9	0.92	6.19	0.15	0.34

§Category acronyms (*e.g.*, DMI_C) are explained in the ‘Model development’ subsection of the ‘Methods and Materials’ section; no Animal_no_DMI_C equation available.

*Equations are presented with regression coefficient standard errors in parenthesis; DMI = dry matter intake (kg d^{-1}), NDF = dietary neutral detergent fiber (% of DM), STA = dietary starch (% of DM), EE = dietary ether extract (% of DM), GEI = gross energy intake (MJ d^{-1}).

†n = number of observations used to fit model equations.

‡Lower-forage = data associated with a forage content ≤ 18%.

ǂp = numbers of observations used for model evaluation.

¶IPCC = Intergovernmental Panel on Climate Change.

♪Performance was evaluated, not cross-validated.

♫No independent evaluation.

‖A subset containing ≤ 10% forage records only was used (as recommended by the IPCC, 2006)

Table 6. European higher-forage CH_4 emission (g d^{-1} $animal^{-1}$) prediction equations for various categories and model performance based on root mean square prediction error (RMSPE; % of mean), RMSPE-observations-standard-deviation-ratio (RSR), mean and slope bias (MB and SB; % of mean square prediction error), and concordance correlation coefficient (CCC).

Model development				Model performance						
Eq.	Category§	Prediction equation*	*n*†	(Sub)set‡	*p*ǂ	RMSPE, %	RSR	MB, %	SB, %	CCC
[28]	DMI_C	60.5 (16.4) + 15.0 (1.4) × DMI	307	EUR-HF	307	26.3	0.79	4.86	15.69	0.48
[29]	DMI+NDF_C	38.1 (23.3) + 14.9 (1.4) × DMI + 0.598 (0.470) × NDF	307	EUR-HF	307	25.9	0.77	5.21	13.15	0.51
[30]	DMI+STA_C	92.4 (21.7) + 11.7 (2.0) × DMI + 0.113 (0.285) × STA	273	EUR-HF	273	25.6	0.92	9.92	2.20	0.30
[31]	DMI+EE_C	133 (34) + 14.5 (2.0) × DMI – 18.4 (6.6) × EE	122	EUR-HF	122	23.7	0.91	1.86	13.87	0.54
[32]	Diet_C	–20.9 (43.6) + 14.3 (2.0) × DMI + 4.04 (1.06) × NDF – 15.4 (3.8) × EE	122	EUR-HF	122	18.4	0.70	5.57	0.97	0.70
[33]	Animal_C	–102 (40.5) + 11.6 (2.1) × DMI + 3.74 (0.79) × NDF – 11.1 (3.0) × EE + 0.164 (0.054) × BW	122	EUR-HF	122	16.7	0.64	3.49	0.00	0.75
				EUR_HF♯	109	15.5	0.58	1.31	0.13	0.79
[34]	Animal_no_DMI_C	34.1 (18.7) + 0.287 (0.028) × BW	307	EUR-HF	307	27.0	0.81	10.37	3.73	0.50
[35]	Global_C	24.3 (17.7) + 9.37 (2.06) × DMI + 0.153 (0.040) × BW	307	EUR-HF	307	24.5	0.73	8.88	11.91	0.58
[36]	GLOBAL NETWORK Tier 2	[0.066 (0.003) × GEI] / 0.05565	307	EUR-HF	307	22.9	0.69	7.03	2.04	0.68
				EUR-HF♪,♬	307	21.8	0.65	1.89	1.34	0.71

[9]	IPCC Tier 2, 2006¶	(0.065 × GEI) / 0.05565	-	EUR-HF♪	307	22.0	0.66	3.54	1.61	0.71
[37]	Charmley *et al*. (2016)	–15.3 + 24.7 × DMI		EUR-HF♪	307	21.9	0.66	0.61	0.16	0.72
					122	19.7	0.75	2.09	3.18	0.66
[11]	Escobar-Bahamondes *et al*. (2017a)	–35.0 + 0.08 × BW + 1.2 × For – 69.8 × EEI^3 + 3.14 × GEI		EUR-HF♪,♯	109	16.0	0.60	0.61	0.37	0.77

---

§Category acronyms (*e.g.*, DMI_C) are explained in the ‘Model development’ subsection of the ‘Methods and Materials’ section.

*Equations are presented with regression coefficient standard errors in parenthesis; DMI = dry matter intake (kg d^{-1}), NDF = dietary neutral detergent fiber (% of DM), STA = dietary starch (% of DM), EE = dietary ether extract (% of DM), BW = body weight (kg), GEI = gross energy intake (MJ d^{-1}), EEI = ether extract intake (kg d^{-1}).

†*n* = number of observations used to fit model equations.

‡EUR-HF = European data associated with a forage content ≥ 25%.

ǂ*p* = numbers of observations used for model evaluation.

¶IPCC = Intergovernmental Panel on Climate Change.

♪Performance was evaluated, not cross-validated.

♫No independent evaluation.

♯ The 307 data points minus data from Pinares-Patiño *et al*. (2003), McGeough *et al*. (2010ab), Doreau *et al*. (2011), Staerfl *et al*. (2012) and Troy *et al.* (2015) to ensure independent evaluation.

Table 7. North American higher-forage CH_4 emission (g d^{-1} animal^{-1}) prediction equations for various categories and model performance based on root mean square prediction error (RMSPE; % of mean), RMSPE-observations-standard-deviation-ratio (RSR), mean and slope bias (MB and SB; % of mean square prediction error), and concordance correlation coefficient (CCC).

Model development				Model performance						
Eq.	Category§	Prediction equation*	*n*†	(Sub)set‡	*p*‡	RMSPE, %	RSR	MB, %	SB, %	CCC
[38]	DMI_C	33.9 (7.7) + 14.7 (0.6) × DMI	394	NrAm-HF	394	25.3	0.55	0.00	26.45	0.78
[39]	DMI+NDF_C	1.58 (8.30) + 14.2 (0.6) × DMI + 1.05 (0.16) × NDF	394	NrAm-HF	394	23.8	0.52	0.02	16.29	0.82
[40]	DMI+STA_C	89.7 (10.7) + 14.2 (0.8) × DMI – 1.17 (0.17) × STA	269	NrAm-HF	269	21.4	0.49	1.90	9.73	0.84
[41]	DMI+EE_C	43.7 (8.8) + 14.7 (0.6) × DMI – 3.72 (1.52) × EE	394	NrAm-HF	394	25.0	0.54	0.01	25.47	0.78
[42]	Diet_C	7.41 (7.12) + 14.1 (0.6) × DMI + 0.632 (0.069) × For	394	NrAm-HF	394	22.2	0.48	0.01	11.10	0.85
[43]	Animal_C,	–15.1 (7.5) + 12.7 (0.6) × DMI + 0.644 (0.066) × For + 0.0779 (0.0134) × BW	394	NrAm-HF	394	20.1	0.43	0.02	6.12	0.88
[44]	Animal_no_DMI_C	14.0 (12.3) + 0.965 (0.104) × For + 0.207 (0.018) × BW – 3.02 (0.95) × Ash	394	NrAm-HF	394	32.2	0.70	0.32	5.45	0.63
[45]	Global_C	–38.8 (10.9) + 12.7 (0.6) × DMI + 0.605 (0.066) × For + 1.61 (0.56) × CP + 0.0779 (0.0133) × BW	394	NrAm-HF	394	20.0	0.43	0.04	3.03	0.89
[46]	GLOBAL NETWORK Tier 2	[0.063 (0.003) × GEI] / 0.05565	394	NrAm-HF	394	21.9	0.48	0.81	0.03	0.87
				NrAm-HF♪,♫	394	21.4	0.46	2.51	0.02	0.88
[9]	IPCC Tier 2, 2006¶	(0.065 × GEI) / 0.05565	-	NrAm-HF♪	394	22.2	0.48	8.70	0.50	0.88
[47]	Charmley *et al.* (2016)	20.5 × DMI		NrAm-HF♪	394	20.0	0.43	0.45	0.02	0.90

§Category acronyms (*e.g.*, DMI_C) are explained in the 'Model development' subsection of the 'Methods and Materials' section.

*Equations are presented with regression coefficient standard errors in parenthesis; DMI = dry matter intake (kg d^{-1}), NDF = dietary neutral detergent fiber (% of DM), STA = dietary starch (% of DM), EE = dietary ether extract (% of DM), For = dietary forage (% of DM), GEI = gross energy intake (MJ d^{-1}), BW = body weight (kg).

†n = number of observations used to fit model equations.

‡NrAm-HF = North American data associated with a forage content ≥ 25%.

ǂp = numbers of observations used for model evaluation.

¶IPCC = Intergovernmental Panel on Climate Change.

♪Performance was evaluated, not cross-validated.

♫No independent evaluation.

Figure 1. Observed *vs*. predicted plots for all-data methane emission (g d^{-1} $animal^{-1}$) prediction equations for the different categories, *viz*., dry matter intake (DMI_C), dry matter intake and neutral detergent fiber (DMI+NDF_C), dry matter intake and starch (DMI+STA_C), dry matter intake and ether extract (DMI+EE_C), diet (Diet_C), animal (Animal_C), animal without DMI (Animal_no_DMI_C), GLOBAL NETWORK Tier 2, IPCC Tier 2 (2006), and the extant Charmley *et al*. (2016) and Escobar-Bahamondes *et al*. (2017) equations. The gray and black solid lines represent the fitted regression line for the relationship between observed and predicted values, and the identity line ($y = x$), respectively.

Figure 2. Observed *vs*. predicted plots for higher-forage methane emission (g d^{-1} $animal^{-1}$) prediction equations for the different categories, *viz*., dry matter intake (DMI_C), dry matter intake and neutral detergent fiber (DMI+NDF_C), dry matter intake and starch (DMI+STA_C), dry matter intake and ether extract (DMI+EE_C), diet (Diet_C), animal (Animal_C), animal without DMI (Animal_no_DMI_C), GLOBAL NETWORK Tier 2, IPCC Tier 2 (2006), and the extant Charmley *et al*. (2016) and Escobar-Bahamondes *et al*. (2017) equations. The gray and black solid lines represent the fitted regression line for the relationship between observed and predicted values, and the identity line ($y = x$), respectively.

Figure 3. Observed *vs*. predicted plots for lower-forage methane emission (g d^{-1} $animal^{-1}$) prediction equations for the different categories, *viz*., dry matter intake (DMI_C), dry matter intake and neutral detergent fiber (DMI+NDF_C), dry matter intake and starch (DMI+STA_C), dry matter intake and ether extract (DMI+EE_C), GLOBAL NETWORK Tier 2, IPCC Tier 2 (2006), and the extant Ellis *et al*. (2009) equations. The gray and black solid lines represent the

fitted regression line for the relationship between observed and predicted values, and the identity line ($y = x$), respectively.

Figure 4. Observed *vs*. predicted plots for European higher-forage methane emission (g d^{-1} animal^{-1}) prediction equations for the different categories, *viz.*, dry matter intake (DMI_C), dry matter intake and neutral detergent fiber (DMI+NDF_C), dry matter intake and starch (DMI+STA_C), dry matter intake and ether extract (DMI+EE_C), dietary (Diet_C), animal (Animal_C), animal without DMI (Animal_no_DMI_C), global (Global_C), GLOBAL NETWORK Tier 2, and IPCC Tier 2 (2006), and the extant Charmley *et al*. (2016) and Escobar-Bahamondes *et al*. (2017) equation**s**. The gray and black solid lines represent the fitted regression line for the relationship between observed and predicted values, and the identity line ($y = x$), respectively.

Figure 5. Observed *vs*. predicted plots for North American higher-forage methane emission (g d^{-1} animal^{-1}) prediction equations for the different categories, *viz.*, dry matter intake (DMI_C), dry matter intake and neutral detergent fiber (DMI+NDF_C), dry matter intake and starch (DMI+STA_C), dry matter intake and ether extract (DMI+EE_C), dietary (Diet_C), animal (Animal_C), animal without DMI (Animal_no_DMI_C), global (Global_C), GLOBAL NETWORK Tier 2 (2006), IPCC Tier 2 (2006), and the extant Charmley *et al*. (2016) equation. The gray and black solid lines represent the fitted regression line for the relationship between observed and predicted values, and the identity line ($y = x$), respectively.