Horizon 2020 Programme

INFRAIA-02-2017
Integrating Activities for Starting Communities

SmartCow: an integrated infrastructure for increased research capability and innovation in the European cattle sector

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

Project ID: 730924

Deliverable number: D6.1

Deliverable title: Improvement of existing equations from proxies to predict animal feed efficiency and its determinants across diets and individuals for dairy and beef cattle

EC version : V1

Due date of milestone	**31/01/2022 (M48)**
Actual submission date	**20/06/2022 (M53)**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n°730924

DOCUMENT INFO

1. Author(s)

Organisation name lead contractor	INRAE

Author	Organisation	e-mail
Cécile Martin	**INRAE**	cecile.martin@inrae.fr
Gonzalo Cantalapiedra	**INRAE**	gonzalo.cantalapiedra@inrae.fr
Donato Andueza	**INRAE**	donato.andueza@inrae.fr
Frédéric Dehareng	**CRAW**	dehareng@cra.wallonie.be
Amélie Vanlierde	**CRAW**	a.vanlierde@cra.wallonie.be

2. Revision history

Version	Date	Modified by	Comments
V1	16 May 2022		

3. Dissemination level

PU	Public	X
CO	Confidential , only for members of the consortium (including the Commission Services)	☐

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

EXECUTIVE SUMMARY

Background	In research, animal experimentation is becoming extremely difficult to justify for its high cost and constraints imposed to animals that can cause harm and animal welfare issues. It is important to reduce the use of animals on one hand, and to increase phenotyping capabilities of Research Infrastructure (RIs) on the other hand, in order to promote optimized use of experimental animals and efficient use of feed resources. Research activities in SmartCow **(https://www.smartcow.eu/)** aimed to increase phenotyping capabilities while implementing the 3R principles (refine, reduce and replace) in cattle nutrition and behaviour studies. In this context, the development and validation of non-invasive proxies to predict feed efficiency and its determinants were undertaken (Work package 6) with the goal of minimizing handling and constraints of experimental cattle in RIs. Proxies are defined as "predictors" measurable in different body matrices (milk, faeces, urine, blood, breath) easy to access and easier to implement than the reference methods used to measure the phenotypes of interest such as: - Feed efficiency, nitrogen partitioning and total tract digestibility - Rumen fermentation parameters (volatile fatty acids and ammonia concentrations, pH), and enteric methane (CH4) emissions.
Objectives	**The objective of this work was to validate proxies already identified for their potential (solid proxies) to predict feed efficiency and its determinants in cattle (dairy, beef) and to identify their range of applicability across diets and individuals when used alone or in combination.** We focused on proxies for their practical application at large scale in RIs and on farm. These concerns: - **the natural ^{15}N abundance in animal proteins (vs urea-N)** for prediction of feed efficiency (FE) in beef cattle (plasma) and milk nitrogen use efficiency (MNE) in lactating dairy cows; **- the near-infrared spectra (NIRS) in faeces** for prediction of total tract organic matter digestibility (OMD) in both beef and dairy cattle; **- the mid-infrared spectra (MIRS) in milk** for predicting enteric CH_4 emissions in lactating dairy cows.
Methods	Our strategy consisted in building a large and representative database of the European cattle production conditions, including both individual phenotypes (FE, MNE, OMD, CH_4) measured by reference methods and proxies (^{15}N, NIRS, MIRS) from different easily accessible body matrices (milk, faeces, blood) from beef and dairy cattle. This database building was possible thanks to a strong collaborative network among SmartCow partners but also with collaborators outside the consortium (TNA applicants including private companies and other research institutes like LUKE from Finland and Agroscope from Switzerland). Collection of data and proxies were carried out from historical experiments and new experiments conducted during the SmartCow project. When proxies' data

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

	were not available from historical or new experiments, samples were transferred to laboratories for analyses according to standardized sampling protocols (MS48) in order to improve power of the database for prediction. Models were tested to predict phenotypes across diets and between-individuals.
Results & implications	**^{15}N vs urea-N for prediction of Feed Efficiency in beef cattle** The database included 759 individual records for animal performance (growing heifers, steers, young bulls) and laboratory data for the difference in the natural abundance of ^{15}N between animal proteins and the diet ($\Delta^{15}N_{animal\text{-}diet}$) measured in plasma or muscle (n = 749) and plasma urea concentration (n = 659). Feed conversion efficiency (FCE; average daily gain/DM intake) and residual feed intake (RFI; observed minus predicted DM intake) criteria were calculated for a duration ranging between 56 and 259 days, depending on the trial. Better models of prediction were systematically obtained with ^{15}N compared to plasma urea, irrespective of using mean or individual values and regardless of the feed efficiency criterion (FCE, RFI). $\Delta^{15}N_{animal\text{-}diet}$ was significantly negatively correlated with FCE across diets and individuals. The proxy can discriminate significantly 11% of animals from the same contemporary group (same diet, place, and time) in terms of FCE), which prevents at this stage to propose ^{15}N as a robust phenotypical tool for assisting genetic selection. In addition, prediction models of FCE from ^{15}N would be dependent on the type of breed; the higher responses (slope) being obtained with late vs early maturing breeds. Finally, $\Delta^{15}N_{animal\text{-}diet}$ succeeded to discriminate the extreme animals within contemporary group in terms of FCE and RFI. These results highlight the potential of $\Delta^{15}N_{animal\text{-}diet}$ to form groups of animals in terms of feed efficiency (FCE or RFI) which could be useful in the context of precision feeding when information about intake and body weight gain is lacking. Combination of both candidate biomarkers did not improve feed efficiency prediction irrespective of the evaluated criteria. **^{15}N vs urea-N for prediction of Nitrogen Use Efficiency (MNE) in dairy cows** We used an extensive database built during the SmartCow project including 1,300 observations (animal x period) reporting milk N yield and N intake and thus enabling to calculate MNE in dairy cows. Partners also shared plasma (n = 696) or milk (n = 604) samples as well as representative diets (n = 74) in order to conduct isotopic analysis by elemental analyser – isotope ratio mass spectrometer (same laboratory) and calculate $\Delta^{15}N_{animal\text{-}diet}$ for each observation. Better models with ^{15}N compared to milk urea-N were reported to predict MNE in dairy cows. We confirmed the previously reported negative correlation between Δ^{15}N and MNE in lactating dairy cows across diets and individuals in mid and late lactation. However, the developed model with ^{15}N seems robust enough to differentiate only extreme cows in terms of MNE. In early lactation, the reported positive (rather than negative) association between Δ^{15}N and MNE might be explained by the considerable protein mobilization of body reserves at this physiological stage. Increases in repeatability of either MNE and Δ^{15}N improved the prediction fitness of the model to differentiate cows in terms of MNE when fed the same diet at the same time. This reinforces the need to identify best sampling protocols and to monitor the accuracy of measurements towards the identification and improvement of proxies to phenotype animals.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

Faecal NIRS for predicting OMD in cattle

The study used a total of 1,236 faeces samples and values of OMD (n = 496 measured by GSM and n = 740 measured using markers) from beef (females and males) cattle and lactating cows fed different diets. Another objective was to compare several modelling strategies in order to optimise the accuracy of the NIRS models. The accuracy of NIRS models (standard error of prediction of 6.4%) for predicting OMD is close to that of the GSM (prediction error of 5.2%). Hence, only a slightly higher number of animals is required when using faecal NIRS predictions than when using the GSM. Therefore, our results confirm that faecal NIRS is an interesting proxy to predict OMD in dairy and beef cattle. This proxy constitutes an alternative method to the GSM (using stalls constraining for animals) when taking into account the 3R principles in animal experimentation. This is an important point to consider for implementing this tool at large scale in practical conditions on farm. Next step will consist in including reference data not well represented in the database (beef males) to enlarge the domain of validity of the model and to improve its robustness. Local calibration methods were more accurate than classic partial least square regressions for predicting OMD measured by the gold standard method.

Milk MIRS for predicting enteric CH_4 emissions in dairy cows

A data set including 261 individuals CH_4 values (measured in respiration chambers) and the corresponding standardized milk MIRS from dairy cows were collected for an external validation of the existing predictive models built from CH4 reference data measured using both respiration chambers and SF6 tracer technique. External validation of the existing models demonstrates that milk MIRS proxy allows obtaining CH_4 predictions with an error of prediction (58 g/d) in line with the known error of the model (52 g/d). Precision of the model allows distinguish high and low CH_4 emitters from a herd fed a same diet. This is all the more possible since the milk MIRS is easily available on the farm because it is already routinely collected during the milk recording. Our data confirm that this high throughput approach offers the possibility to integrate CH_4 phenotype in dairy cows breeding programmes.

We observed also that it is not possible to merge phenotype data sets measured using different methods (ex: GreenFeed vs. respiration chambers and SF6 for CH4; GSM vs. markers for total tract digestibility) without increasing model prediction errors. This highlights the importance of common and standardized protocols for measurements and data recording for merging and enhancing all future data. This will help to enlarge the diversity of the reference databases and to update the models according to research recommendations.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

Table of contents

1 Proxies to predict feed efficiency (FCE, RFI) in beef cattle: urea-N vs 15N 7
1.1 Material & methods 7
1.2 Results & discussion 7
1.3 Conclusion 11
1.4 Abstracts & peer-reviewed articles from this work 12
2 Proxies to predict milk nitrogen use efficiency (MNE) in dairy cows: urea-N vs ^{15}N 13
2.1 Material & methods 13
2.2 Results & discussion 13
2.3 Conclusion 16
2.4 Abstracts & peer-reviewed articles from this work 16
3 Proxies to predict total tract digestibility in cattle: faecal NIR spectra 19
3.1 Material & methods 19
3.2 Results & Discussion 20
3.3 Conclusion 22
3.4 Abstracts & peer-reviewed articles from this work 22
4 Proxies to predict enteric methane emissions: milk MIR spectra 24
4.1 Material & methods 24
4.2 Results & discussion 24
4.3 Conclusion 25
4.4 Abstracts and peer-reviewed articles from this work 26

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

1 Proxies to predict feed efficiency (FCE, RFI) in beef cattle: urea-N vs 15N

The objective of this task was to confirm two candidate biomarkers of feed efficiency in growing cattle: the natural ^{15}N enrichment of animal proteins over the diet, also known as N isotopic discrimination or fractionation ($\Delta^{15}N_{animal\text{-}diet}$; Cantalapiedra-Hijar et al., 2015; Guarnido Lopez et al., 2021) and plasma urea concentration (Richardson et al., 2004) both of them previously reported to be related to N metabolism and use efficiency.

1.1 Material & methods

A database was built using performance data from 13 trials provided by 3 research centres (INRAE from France, SRUC from Scotland and Agroscope from Switzerland) conducted with growing heifers, steers and young bulls and testing 34 dietary treatments. The database included 759 individual records for animal performance and laboratory data for $\Delta^{15}N_{animal\text{-}diet}$ measured in plasma or muscle (n = 749) and plasma urea concentration (n = 659). Feed conversion efficiency (FCE; average daily gain/DM intake) and residual feed intake (RFI; observed minus predicted DM intake) criteria were calculated for a duration ranging between 56 and 259 days, depending on the trial. For FCE prediction, mixed models included the random effects of study, diet within-study and pen within-study (i.e. contemporary group) allowing these effects to be progressively excluded from the relationship and therefore assessing the ability of this biomarker to predict FCE variation both across diets (within-study relationship) and between-individuals (within-contemporary group relationship). **In addition, we tested if the type of breed (late vs early or intermediate maturing breeds) could have an effect on the FCE prediction from biomarkers.** For RFI prediction, simple linear regressions were tested with the contemporary group effect removed from biomarker values (likewise in the RFI model) before regression analysis.

1.2 Results & discussion

Better models were always obtained with $\Delta^{15}N_{animal\text{-}diet}$ compared to plasma urea, irrespective of using mean or individual values and regardless of the feed efficiency criterion. Overall, both biomarkers were negatively and significantly correlated with FCE across experiments (Figure 1A, 1B) or diets (Figure 1C, 1D) but only $\Delta^{15}N_{animal\text{-}diet}$ remained negatively associated with FCE on average at the individual level (Figure 1E) while plasma urea failed (Figure 1F).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

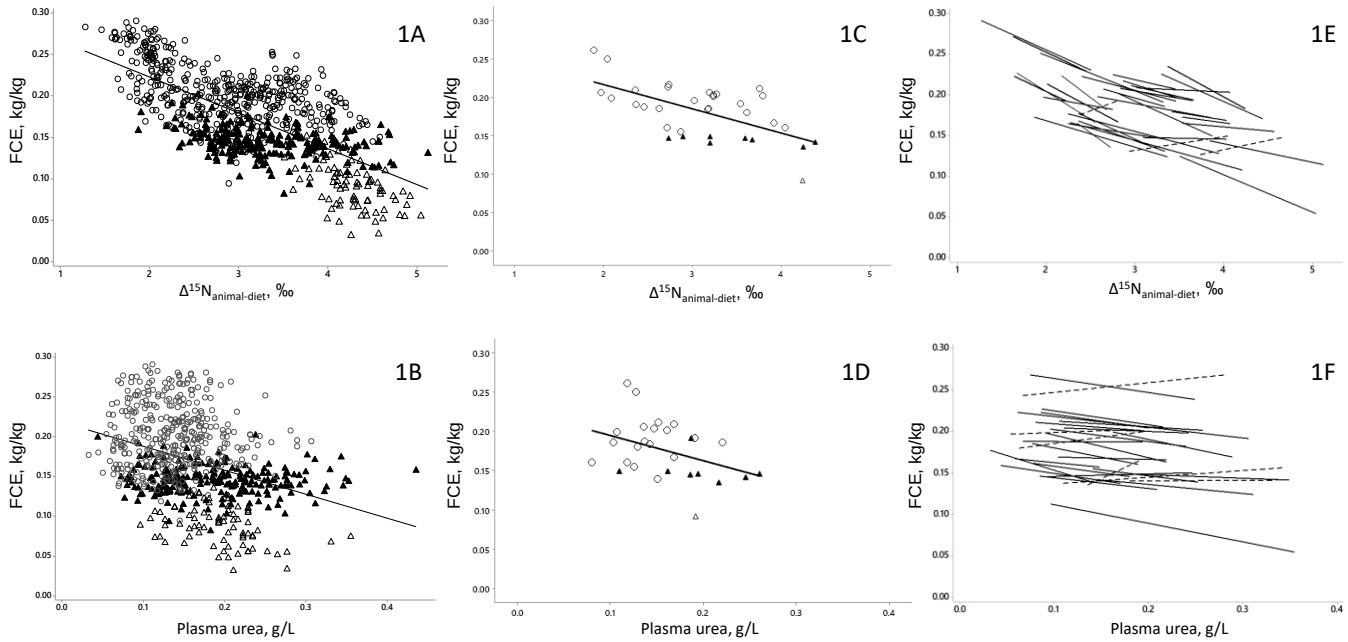


Figure 1.1. Relationship between feed conversion efficiency (FCE, kg/kg) and either N isotopic fractionation ($\Delta^{15}N_{animal\text{-}diet}$, ‰) or plasma urea concentration (g/L) in fattening beef cattle. On the left panels, simple linear regression across-studies using individual observations (1A: FCE = 0.31 – 0.043 × $\Delta^{15}N$ [n = 749; r = -0.69; *P*<0.001]; 1B: FCE = 0.22 – 0.30 × Plasma urea [n = 659; r = -0.36; *P*<0.001]) where open circles = young bulls, open triangles = beef heifers and closed triangles = beef steers. On the middle panels; simple linear regression using dietary treatment means (1C: FCE = 0.28 – 0.032 × $\Delta^{15}N$ [n = 34; r = -0.62; P<0.001]; 1D: FCE = 0.23 – 0.32 × Plasma urea [n = 28; r = -0.39; P=0.04]. On the right panels, simple linear regression within-treatment for $\Delta^{15}N_{animal\text{-}diet}$ (1E) and plasma urea (1F), where continuous line = negative relationship and dashed line = positive relationship. For $\Delta^{15}N_{animal\text{-}diet}$, 16 out of 32 within-treatment relationships were significant (*P*<0.05; 1E), whereas only 3 out of 28 were significant for plasma urea concentration (*P*<0.05; 1F).

When the between-contemporary group effect (pen within diet and study) was removed from both feed efficiency and biomarker values to explore the relationships at the individual level, significant correlations were obtained between efficiency traits (FCE or RFI) and either $\Delta^{15}N_{animal\text{-}diet}$ or plasma urea concentration (P<0.01) but with different fits depending on the trait and biomarker (Figure 2.1). On average, greater model fits were achieved for FCE than for RFI. From higher to lower model fits were the regressions of FCE on $\Delta^{15}N_{animal\text{-}diet}$ (r = -0.50; P<0.001, Figure 2A), RFI on $\Delta^{15}N_{animal\text{-}diet}$ (r = 0.23; P<0.001, Figure 2C), FCE on plasma urea (r = -0.15; P<0.001, Figure 2B) and RFI on plasma urea (r= 0.11; P=0.003, Figure 2D).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

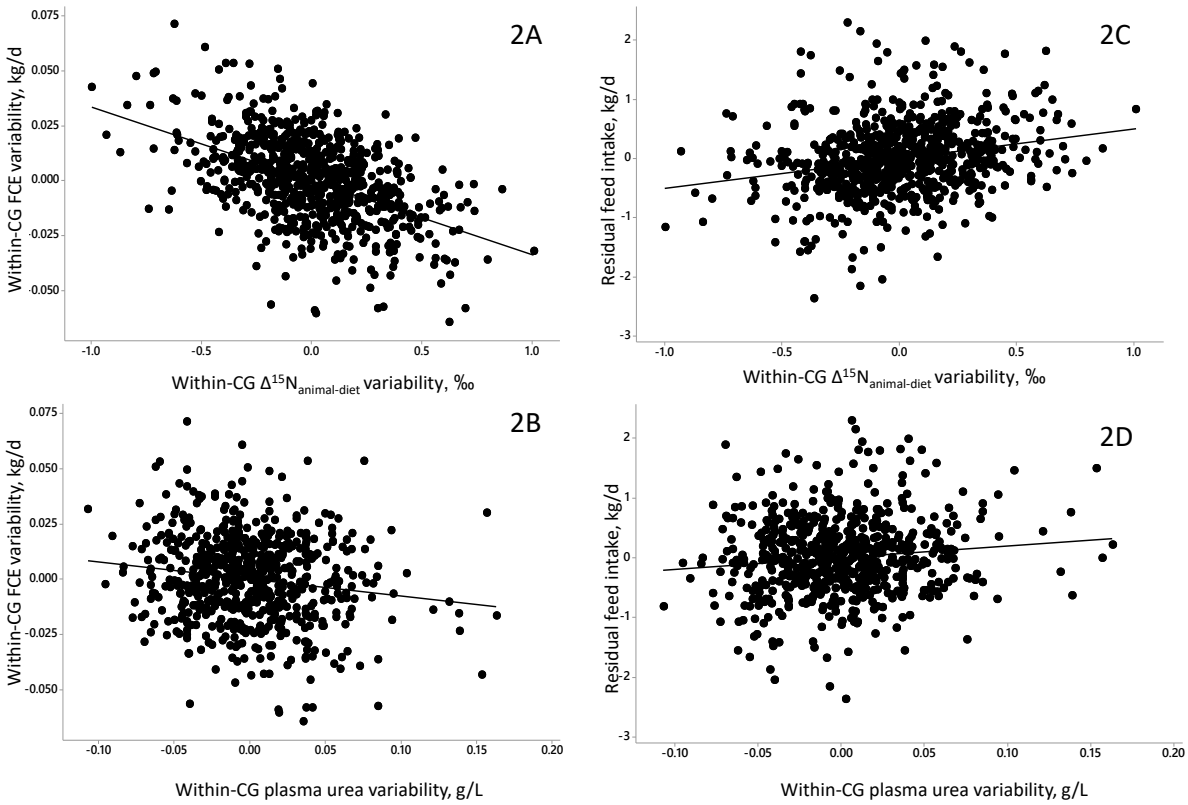


Figure 1.2. Relationships between feed efficiency traits (feed conversion efficiency; FCE [2A, 2C] and residual feed intake, RFI [2B, 2D]) and either N isotopic discrimination (Δ^{15}Nanimal-diet, ‰) or plasma urea concentration (g/L) in fattening beef cattle when variability across contemporary groups (CG) was previously removed (within-CG variability analysis). Equations (intercept not different from 0): (2A) Y = -0.034X (n = 749; r = -0.50; P<0.001); (2B): Y = -0.077X (n = 659; r = -0.15; P<0.001); (2C) Y = 0.50X (n = 748; r = 0.23; P<0.001); (2D): Y = 1.93X (n = 659; r = 0.11; P=0.003).

Prediction error (0.027 kg/kg) from mixed-effect models using mean FCE and $\Delta^{15}N_{animal\text{-}diet}$ values allows discrimination of 2 dietary treatments or production context in terms of FCE (95%CI) if they differ by more than 0.10 kg/kg (calculated as ± 1.96 × RMSEP from model 1 in Table 1), a difference observed among several studies from our database. Results from the mixed model analysis also highlighted that it is possible to significantly discriminate from $\Delta^{15}N_{animal\text{-}diet}$ or plasma urea concentration two animals randomly selected from the same contemporary group if they differ by at least 0.06 kg/kg and 0.08 kg/kg of FCE, respectively (calculated as ± 1.96 × RMSEP from model 4 and 9, respectively, in Table 1.1). This minimal detectable difference between two individuals from the same contemporary group was observed for about 11 and 6% of animals, for $\Delta^{15}N_{animal\text{-}diet}$ and plasma urea respectively, which prevents at this stage to propose them as robust phenotypical tools in support of breeding programmes. Interestingly, the models to predict at the individual level the animal FCE from $\Delta^{15}N_{animal\text{-}diet}$ seemed to be affected by the type of breed since a trend (P = 0.06) for higher responses (slope) were obtained with late vs early maturing breeds (Table 1.1). **This finding supports the close link existing between $\Delta^{15}N_{animal\text{-}diet}$ and protein deposition in growing ruminants and highlight the need to establish prediction models adapted to the type of breed.**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

Table 1.1. Simple linear and mixed-effect regression models of feed conversion efficiency on the N isotopic fractionation ($\Delta^{15}N_{animal\text{-}diet}$) or plasma urea concentration using either dietary treatment means or individual observations from different fattening cattle production systems

Item	Model	Intercept	Slope	RMSEP[1]	RSR global[2]	RSR condition[3]	R^2	AIC[4]	BIC[4]
$\Delta^{15}N_{animal\text{-}diet}$, ‰									
Treatment means, n = 34									
Simple linear model	1	0.283**± 0.023	-0.032** ± 0.007	0.027	0.768		0.39	-126	-122
Individual observations, n = 749									
Simple linear model	2	0.310**± 0.005	-0.043** ± 0.002	0.037	0.727		0.47	-2789	-2775
Mixed-effect models									
Study random effect	3	0.273**± 0.020	-0.032** ± 0.005	0.018	0.363	0.882	0.31	-3788	-3760
Treatment within study random effect[$]	4	0.277**± 0.018	-0.033** ± 0.004	0.016	0.325	0.846	0.33	-3826[$]	-3784[$]
Late vs Early maturing breeds[#]		+0.013±0.007[NS]	-0.005[†]± 0.002						
Contemporary group random effect	5	0.277**± 0.018	-0.033** ± 0.004	0.016	0.323	0.840	0.33	-3820	-3779
Plasma urea, g/L									
Treatment means, n = 28									
Simple linear model	6	0.224**± 0.025	-0.304* ± 0.150	0.033	0.912		0.15	-95.0	-91.2
Individual observations, n = 659									
Simple linear model	7	0.218**± 0.005	-0.301* ± 0.031	0.048	0.935		0.13	-2101	-2088
Mixed-effect models									
Study random effect	8	0.176**± 0.010	-0.019[NS] ± 0.028	0.021	0.406	0.982	0.00	-3114	-3087
Treatment within study random effect[$]	9	0.188**± 0.007	-0.066* ± 0.026	0.019	0.374	0.969	0.01	-3140[$]	-3112[$]
Late vs Early maturing breeds		+0.003±0.005[NS]	-0.019[NS] ± 0.002						
Contemporary group random effect	10	0.193**± 0.007	0.059* ± 0.026	0.019	0.369	0.956	0.01	-3073	-3046

[1]RMSEP = residual mean square error of prediction when comparing observed and predicted values (package chillR in R)
[2]RSR global = Ratio of the RMSEP (model prediction error) to the standard deviation of feed conversion efficiency observed in the whole dataset (across-study variability). The lower the better.
[3]RSR condition = Ratio of the RMSEP (model prediction error) to the standard deviation of feed conversion efficiency observed either within-study, or within-diet and study or within-CG. The lower the better
[4]AIC = Akaike information criterion (the lower the better); BIC = Bayesian information criterion (the lower the better)
[$]Best random structure model based on AIC/BIC criteria and the log-likelihood ratio test ($P < 0.05$).
[NS] Non significant (P>0.05) ; [**]$P \leq 0.001$; [*]$P \leq 0.05$; [†]$P \leq 0.10$.
[#]From a sub-dataset containing 312 observations from studies including both early or intermediate and late maturing breeds

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

The potential of both candidate biomarkers to form groups of animals based on their feed efficiency ranking was also evaluated. Although the two tested biomarkers were significantly correlated with feed efficiency traits at the individual level when using the whole dataset, their ability to discriminate the top 20% highest and lowest individuals within-contemporary group in terms of feed efficiency varied (Figure 1.3). **Indeed, both $\Delta^{15}N_{animal\text{-}diet}$ and plasma urea concentration succeeded to discriminate the most extreme animals within contemporary group in terms of FCE (Figure 3A and 3B; P<0.001 and P=0.002), but only $\Delta^{15}N_{animal\text{-}diet}$ was significantly different (P<0.001) for extreme RFI animals (Figure 3C). These results highlight the potential of $\Delta^{15}N_{animal\text{-}diet}$ to form groups of animals in terms of feed efficiency (FCE or RFI) which could be useful in the context of precision feeding.**

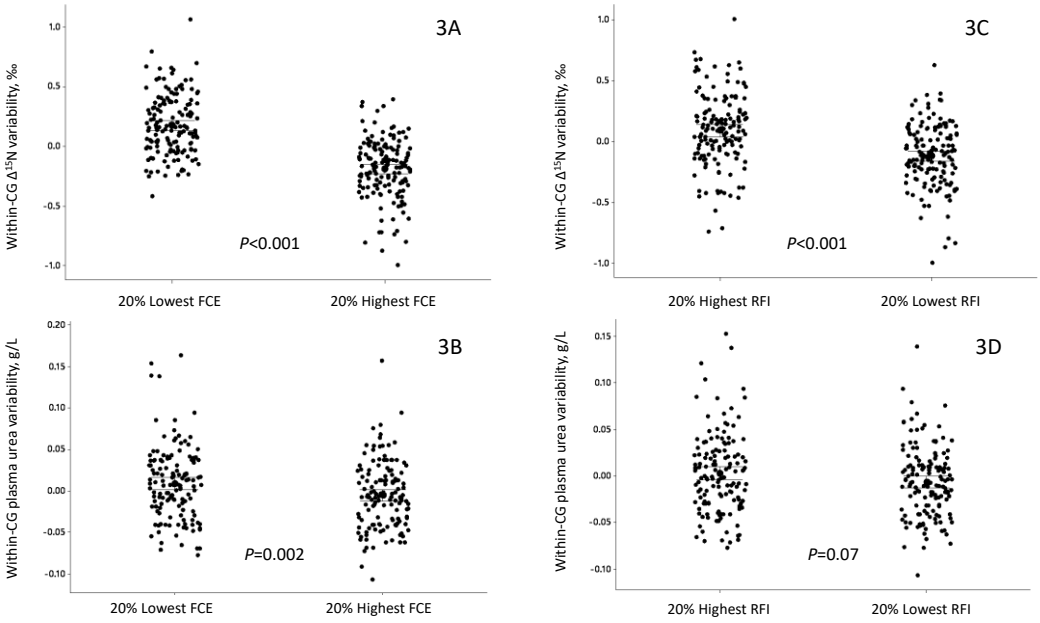


Figure 1.3. Within contemporary group values for N isotopic discrimination (Δ15N) and plasma urea concentration in the top 20% higher and lower efficient animals within contemporary group according to feed conversion efficiency (on the left; 3A and 3B) or residual feed intake (on the right; 3C and 3D).

Finally, no gain in feed efficiency prediction was observed when combining the two candidate biomarkers. However, for FCE, if the average daily gain is available on farm (best single predictor), its combination with $\Delta^{15}N_{animal\text{-}diet}$ strengthen the prediction at the individual level (R^2 = 0.51) compared to using only single predictors (R^2 = 0.25 and 0.39 for $\Delta^{15}N_{animal\text{-}diet}$ and ADG, respectively).

1.3 Conclusion

This work from task 6.1 demonstrated that $\Delta^{15}N_{animal\text{-}diet}$ could be proposed as a biomarker of feed efficiency to discriminate different production conditions or diets if they differ by at least 0.10 kg/kg of FCE. When the objective is to discriminate two animals from the same contemporary group, $\Delta^{15}N_{animal\text{-}diet}$ may succeed provided they differ by at least 0.06 kg/kg of FCE. This minimal detectable difference (0.06 kg/kg of FCE) across individuals represent a limitation to predict FCE at the individual level and calls into question its use as tool for breeding programmes. However, our results highlight that $\Delta^{15}N_{animal\text{-}diet}$ can significantly discriminate group of animals with contrasting FCE or RFI values (20% highest vs 20% lowest ranked

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

animals). **No gain in feed efficiency prediction was observed when combining both candidate biomarkers (^{15}N and plasma urea-N) irrespective of the evaluated criteria.** However, when average daily gain (ADG) data was combined with $\Delta^{15}N_{animal\text{-}diet}$ values the prediction of FCE at the individual level was strengthened compared to using only one of them, in which case ADG was the best single predictor. **Our findings confirm that $\Delta^{15}N_{animal\text{-}diet}$ is related to the between-animal variability of feed efficiency in growing cattle to a greater extent than plasma urea concentration. It may be useful to form groups of animals for precision feeding when information about intake and body weight gain is lacking. More studies are warranted however to evaluate the usefulness of these two biomarkers to assist the genetic selection and the gain to combine them with other promising biomarkers of feed efficiency.**

1.4 Abstracts & peer-reviewed articles from this work

Cantalapiedra-Hijar, G., Ortigues-Marty, I., Martin, C., Morel, I., Dewhurst, RJ., 2020. Invited speaker: Natural 15N abundance of animal proteins: a promising biomarker of feed efficiency in beef cattle. 71st Annual Meeting of European Federation of Animal Science. 1-4 December, Porto (Portugal).

Cantalapiedra-Hijar, G., Morel, I., Sepchat, B., Chantelauze, C., Miller, G.A., Duthie, C-A., Ortigues-Marty, I., Dewhurst, R.J., 2022. Identifying cattle with superior growth feed efficiency through their natural 15N abundance and plasma urea concentration: a meta-analysis. Peer Community Journal 2:e31. https://doi.org/10.24072/pcjournal.130

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

2 Proxies to predict milk nitrogen use efficiency (MNE) in dairy cows: urea-N vs 15N

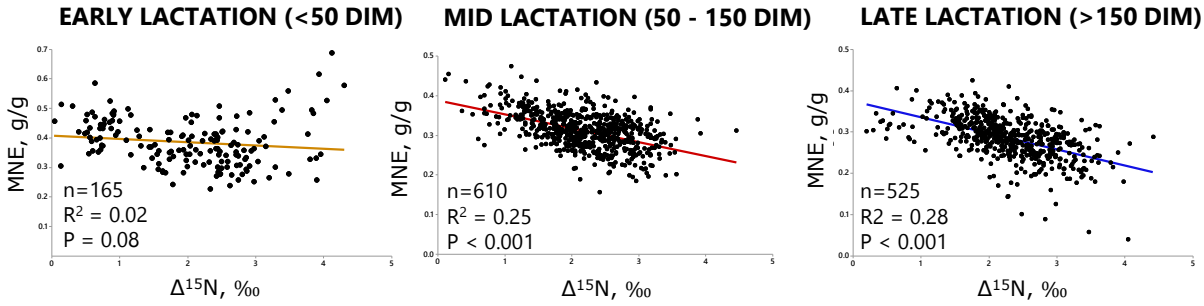
The aim of this task was to assess the ability of two biomarkers of the N status of ruminants, the natural ^{15}N enrichment of animal proteins over the diet also known as the nitrogen isotopic discrimination ($\Delta^{15}N$; Cantalapiedra-Hijar et al., 2018) and milk urea-N (**MUN**) concentration (Huhtanen et al., 2015), to predict diet and between-animal variations in the efficiency of N use for milk production (**MNE**) in dairy cows.

2.1 Material & methods

We used an extensive database built during the SmartCow project including 20 independent trials proposed by partners (INRAE, Arhus University, CRA-W, University of Reading and LUKE) with 1,300 observations (animal x period) reporting milk N yield and N intake and thus enabling to calculate MNE. Partners also shared plasma (n = 696) or milk (n = 604) samples as well as representative diets (n = 74 dietary treatments) in order to conduct isotopic analysis by EA-irms (same laboratory) and calculate $\Delta^{15}N$ for each observation. Data for MUN was available from 9 experiments and 703 observations and was analysed by the partner through different methods (mid-infrared spectroscopy, colorimetric or continuous flow analyser). Data were analysed through mixed-effect regression models considering the experiment, period and diet as random effects, and so allowing these effects to be progressively excluded from the relationship to assess the relationships both across diets (within-experiment relationship) and between-individuals (within diet, period and diet relationship). In addition, repeatability estimates were calculated for experiments where repeated measurements within the same animal were available to test the hypothesis of improved MNE predictions when measurements errors of biomarkers and MNE decreased.

2.2 Results & discussion

First of all, it was observed that the relationship between MNE and $\Delta^{15}N$ in early lactation (DIM < 50) was different compared to that observed in mid and late lactation (Figure 2.1).



This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

Figure 2.1. Relationships between the natural ^{15}N enrichment of animals proteins over the diet ($\Delta^{15}N$) and the efficiency of feed N utilization for milk protein yield (MNE) in lactation dairy cows in early (on the left pannel), mid (on the middle pannel) and late lactation.

The considerable protein mobilization in early lactation artificially increased both MNE and $\Delta^{15}N$ leading to a positive rather than negative relationship and this limited the implementation of this biomarker in early lactating cows. This phenomenon was confirmed in one specific TNA trial (Responsible L. Bahloul; Adisseo) included in our database and aiming to evaluate animal response to amino acid supplementation in transition dairy cows. In this TNA experiment it was observed that $\Delta^{15}N$ variations in early lactation mirrored on average the changes in body weight (Figure 2.2) and that body weight loss were associated with higher $\Delta^{15}N$ and MNE values (Correa-Luna et al., 2021).

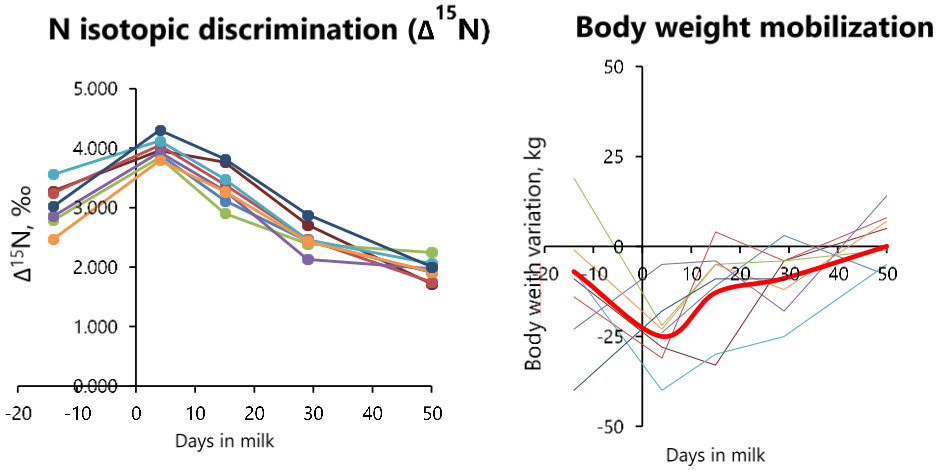


Figure 2.2. Association between N isotopic discrimination ($\Delta^{15}N$) and body weight mobilization in transition dairy cows.

In line with previous research (meta-analysis by Cantalapiedra-Hijar et al., 2018), we observed that, at least in mid and late lactation, $\Delta^{15}N$ was on average negatively and significantly correlated with MNE at the individual level (Table 2). In our conditions, $\Delta^{15}N$ was only able to discriminate significantly in terms of MNE two given cows within the same contemporary group (95%CI) if they differed by at least 0.112 g/g of MNE (calculated as ± 1.96 × RMSPE from model 5 ni Table 2.1). The error for predicting MNE from $\Delta^{15}N$ was considered still high (even more for milk urea) for discriminating individuals, and prevent at this stage proposing it as a robust phenotyping tool in support of breeding programmes. However, $\Delta^{15}N$ unlike milk urea was able to form groups of dairy cows showing contrasted MNE values (Figure 2.3).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

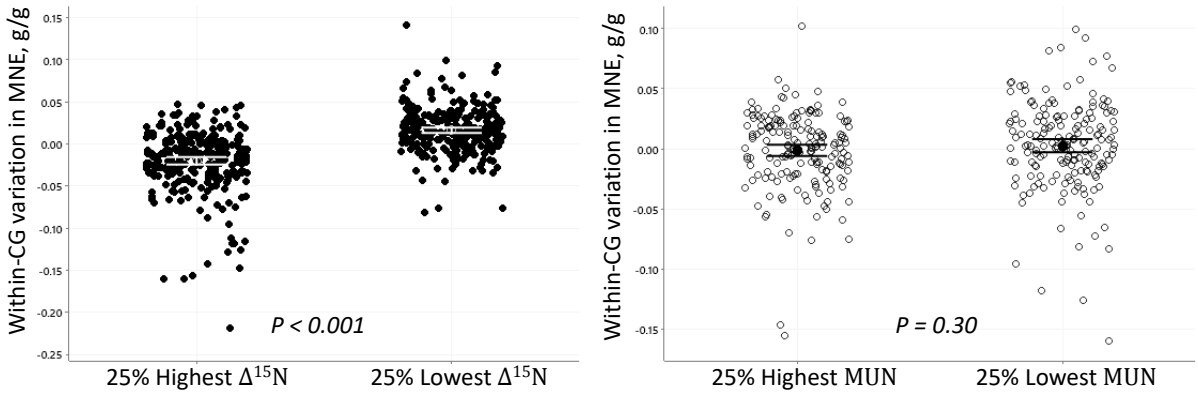


Figure 2.3. Within contemporary group (CG) variation in milk N use efficiency (MNE) when dairy cows were ranked as the 25% highest and lowest individuals in relation to $\Delta^{15}N$ (on the left) or MUN (on the right) values. Milk N use efficiency varied only when groups were formed from $\Delta^{15}N$ values ($P<0.001$).

Indeed, when the $\Delta^{15}N$ values from the 25%most and least efficient dairy cows in terms of NUE were compared, they significnatly differed ($P<0.001$), highlighting the potential of this isotopic biomarker to form groups of dairy cows based on their ability to transform the feed N into milk proteins for precision feeding.

Our meta-analysis also highlighted that the experiments showing the highest correlation between observed and predicted MNE from $\Delta^{15}N$ proxy (Y axis in Figure 2.4) were those where both traits (MNE phenotype and $\Delta^{15}N$ proxy) showed the highest repeatability values (X axis in Figure 2.Figure 2.4), suggesting an even higher potential of this promissing biomarker if measurements errors are diminished, a condition fulfilled in some experiements of our dataset (i.e. ID#7, ID#9 and ID#15 in Figure 2.4.a).

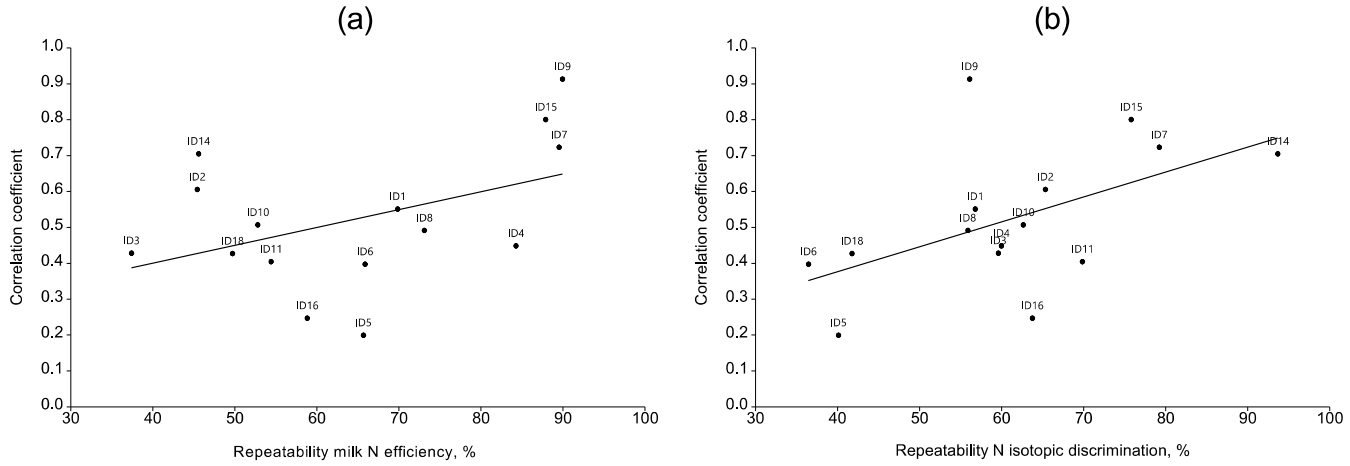


Figure 2.4. Relationship between correlation coefficient between observed vs. predicted MNE at the within-study level (Table 2) and repeatability of either (a) MNE ($R^2 = 0.49$; $P = 0.06$) or (b) $\Delta^{15}N$ ($R^2 = 0.54$; $P = 0.03$). Higher correlations between MNE and $\Delta^{15}N$ were obtained in those studies where the measurements errors of MNE and $\Delta^{15}N$ were lower (i.e. higher repeatability).

In the case of MUN, although it was negatively correlated with MNE, the responses were rather small and inconsistent at the individual level and this led only to marginal discrimination of dietary treatments in terms of MNE (Table 2.1; Figure 2.5). Although our MUN data came from different laboratories using different

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

methodologies the ability of MUN to discriminate conditions (diets and individuals) within-experiment could be evulated. Our results suggest that MUN can not be proposed as a biomarker of MNE at the individual level.

2.3 Conclusion

Results from this work confirmed the previously reported negative correlation between $\Delta^{15}N$ and MNE in lactating dairy cows regardless of experimental site, sampling period, and dietary treatment and thus demonstrated that this biomarker may reflect the between-animal variation in MNE of lactating dairy cows in mid and late lactation. The developed model seems however only robust enough to differentiate extreme cows in terms of MNE. The weak negative correlation between MUN and MNE within contemporary group confirms that MUN has a limited ability to differentiate between-animal variation of MNE. In early lactation both MNE and $\Delta^{15}N$ might be artificially increased because of the considerable protein mobilization of body reserves. **This was confirmed by observing a positive (rather than negative) association of $\Delta^{15}N$ along with MNE in early lactation.** Increases in repeatability of either MNE and $\Delta^{15}N$ improved the prediction fitness of the model to differentiate cows in terms of MNE when fed the same diet at the same time. **This reinforces the need to identify best sampling protocols and to monitor the accuracy of measurements towards the identification and improvement of proxies to phenotype animals.**

2.4 Abstracts & peer-reviewed articles from this work

Correa-Luna, M., Bahloul, L., Chantelauze, C., Larsen, M. Cantalapiedra-Hijar, G., 2021. Predictions of N use efficiency from natural 15N abundance in periparturient dairy cows are impaired by the protein mobilization. 2021 ADSA Annual Meeting. July 11-14 (Oral presentation).

Correa-Luna, M., Johansen, M., Nozière, P., Bayat, A.R., Compton, L.A., Reynolds, C.K., Froidmont, E., Eduard, N., Lund, P., Martin, C., Cantalapiedra-Hijar, G., 2021. Prediction of between-animal variation in nitrogen use efficiency from natural 15N abundance in animal protein: model evaluation in dairy cows. In: 72nd Annual Meeting of the European Federation of Animal Science. Davos, Switzerland. 30.08.2021-03.09.2021 (Oral presentation)

Correa-Luna, M., Johansen, M., Nozière, P., Chantelauze, C., Nasrollahi, S.M., Lund, P., Larsen, M., Bayat, A., Crompton, L., Reynolds, C., Froidmont, E., Edouard, N., Dewhurst, R., Bahloul, L., Martin, C., Cantalapiedra-Hijar, G., 2022. Nitrogen isotopic discrimination as a biomarker of between-cow variation in the efficiency of nitrogen utilization for milk production: A meta-analysis. Journal of Dairy Science 105 (6):5004-5023. https://doi.org/10.3168/jds.2021-21498

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

Table 2.1. Mixed-effect regression models of N use efficiency for milk protein yield (Y, g/g) on the N isotopic discrimination ($\Delta^{15}N$, X_1) and MUN (X_2) using either dietary treatment means or individual observations in mid and late lactating dairy cows

Item	Model no.	Intercept	Slope	AIC	RMSPE	R^2	CCC	RSR
Tier 1: dietary treatment means								
$\Delta^{15}N$, ‰ (n = 72)								
Experiment random effects§	1	0.378* ± 0.017	-0.037* ± 0.007	-289	0.034	0.28	0.478	0.822
MUN, mg/dl (n = 50)								
Experiment random effects§	2	0.3484* ± 0.0164	-0.0030* ± 0.0008	241	0.035	0.20	0.281	0.958
Tier 2: individual observations								
$\Delta^{15}N$, ‰ (n = 1,135)								
Experiment random effects	3	0.403* ± 0.014	-0.049* ± 0.007	-4,313	0.036	0.30	0.380	0.882
Experiment and period random effects	4	0.407* ± 0.013	-0.050* ± 0.007	-4,316	0.035	0.32	0.381	0.883
Contemporary group random effects$	5	0.417* ± 0.013	-0.056* ± 0.007	-4,498	0.028	0.36	0.400	0.851
MUN, mg/dl (n = 703)								
Experiment random effects$	6	0.3302* ± 0.0168	-0.0021** ± 0.0007	3,889	0.037	0.08	0.088	0.982
Experiment and period random effects	7	0.3231* ± 0.0137	-0.0020** ± 0.0006	3,929	0.037	0.08	0.077	0.985
Contemporary group random effects	8	0.2984* ± 0.0067	-0.0005** ± 0.0003	3,967	0.031	0.01	0.003	0.998

AIC = Akaike information criterion; RMSPE = square root of the mean square prediction error (× 100); R^2 = coefficient of determination calculated for equations according to the experimental factor nesting level included in each case; CCC = concordance correlation coefficient; RSR = square root of the mean square prediction error to standard deviation of observed values ratio.

§Models at the treatment means level were tested with random effects on the intercept.

†All models at the individual observations level were tested with random effects on the intercept, slope or both.

¥It refers to cows fed the same diet in the same experimental period within the same experiment.

$Best random structure model based on the AIC criterion.

*$P \leq 0.001$; **$P \leq 0.05$.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

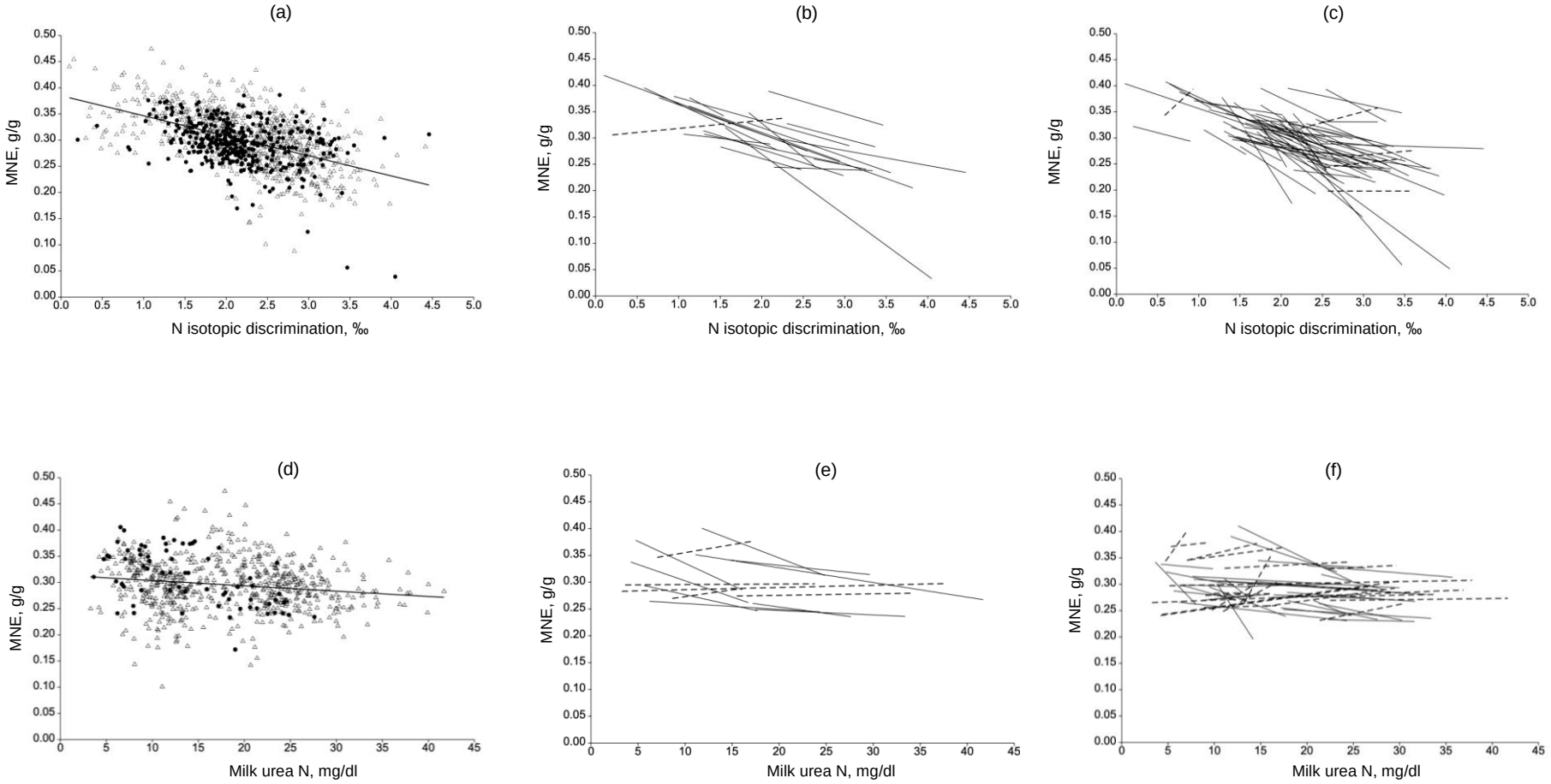


Figure 2.5. Relationship between milk N use efficiency (MNE) and either N isotopic fractionation ($\Delta^{15}N$) or milk urea N in lactating cows using individual values: (a, d) Simple linear regression analysis where open triangles represent multiparous cows and closed circles represent primiparous cows; (b, e) simple linear regression for each independent study (within-study regression) (c, f) simple linear regression analysis for each independent diet (; within-diet regression). In (b, e) and (c, f) solid lines represents negative slopes and dashed lines represents positive slopes. Correlations coefficients (and statistical significances) are presented in Table 2.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

3 Proxies to predict total tract digestibility in cattle: faecal NIR spectra

Near infrared spectroscopy (NIR) is a commonly used technology for the management of livestock systems. In particular, it is used by the industry and research for the estimation of the chemical composition and digestibility of feedstuffs. Its application on faeces is seen less often because of the difficulty in obtaining samples from the animals. However, an advantage of its use compared to feedstuff samples is the possibility of obtaining estimated values for individuals. Faeces are made of the undigested fraction of the diet but also of other components of non-dietary origin derived from the digestive processes of the diet such as undigested microbial and animal endogenous material. Chemical composition of faeces may reflect changes in diet digestibility (Demarquilly et al., 1995). Consequently, NIR spectra (NIRS) of faeces can provide information on the digestive use of the diet and particularly on the organic matter digestibility (OMD), which is closely related to the energy value of the diet. Near infrared spectroscopy of faeces has been successfully used for predicting OMD in tropical (Boval et al., 2004) and temperate (Decruyenaere et al., 2012; Jancewizcz et al., 2016) conditions. However, the two last models have been built with specific datasets limiting their domain of validity and their use in practice (lactating grazing dairy cow for Decruyenaere et al.; cattle for fattening for Jancewizcz et al.). The main objective of this work was to evaluate the ability of NIR spectroscopy performed on individual faeces of cattle to predict diet OMD by using a large dataset including both dairy and beef cattle fed a large variety of diets. Another objective was to compare several modelling strategies in order to optimise the accuracy of the NIRS models.

3.1 Material & methods

The study used 1,236 faeces samples from beef cows, lactating cows and young bulls fed different diets. Research centres from 6 countries (INRAE from France, CRA-W from Belgium, Reading University from UK, IRTA from Spain, Aarhus University from Denmark and Agroscope from Switzerland) have contributed by supplying faecal samples and organic matter digestibility (OMD) values. Organic matter digestibility have been measured *in vivo* using either gold standard method (GSM; n=496 samples; OMD_GSM) or estimated by markers (Cr_2O_3, AIA and TiO_2; n=740; OMD_M). Because the GSM is different from the marker method, the databases cannot be combined, and therefore 2 different models were developed: one from the OMD_GSM database and the other one from OMD_M data set. Some partners provided faecal samples mixed with urine but these samples were not used in this work.

Acquisition of faecal NIRS and alignment with in vivo reference data

For OMD_GSM, each experimental period comprised at least 8–day adaptation period to the diet and to the stalls of digestibility followed by 6 days of measurements. Diets were offered *ad libitum* (sometimes 0.95 times *ad libitum*). The animals had free access to water and vitamin–mineral blocks throughout the experimental period. During the measurement period, the total quantity of faeces excreted by each animal was collected daily and after weighing, they were subsampled. **Subsamples collected over six days were then combined to provide one sample of faeces per animal and per period.** Finally, the OMD is calculated as (OM ingested – OM excreted in faeces)/OM ingested as described in Mesgaran et al. (2020) in the Book of Method of SmartCow.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

For OMD_M, a marker was supplied to animals or an internal marker was analysed. **One spot faeces sample was obtained per animal**, and the marker content administered or present in the diet and in the faeces was analysed. Organic matter digestibility was calculated according to Mesgaran et al. (2020)
Faeces (and diet) samples were oven–dried at 60 °C for 72 h to measure DM (sometimes they were lyophilised), then ground through a 1 mm screen and stored in ambient laboratory conditions. Crude ash of diets and faeces samples were analysed (600°C for 6h) according to Association of Official Analytical Chemists (AOAC 942.05).

After sample homogenization, a dried sample of faeces (≈5 g) were placed in a 50mm diameter ring cup and scanned in reflectance mode at 2 nm intervals in the range of 400–2500 nm using a Foss NIRSystems model 6500 scanning visible–NIR spectrometer. Spectra and reference values were recorded using ISIScan software (Infrasoft International). Each spectrum was time-averaged from 32 scans. A scan (using the internal ceramic reference tile) was taken before and after each sample as a background reference. Reflectance values were converted into absorbance values using the formula: absorbance=log (1/reflectance). Faeces samples were scanned twice. If spectral differences were high (root mean square values higher than 500), samples were re-scanned and the average spectrum was calculated.

Calibration Model Development and Data Analysis
Calibration models for the prediction of OMD_GSM and OMD_M were developed using R software. The faeces samples and corresponding NIR spectra were divided using the Kennard–Stone algorithm into a calibration set and a validation set. Four calibration strategies were developed: (1) a single model with all the samples in the calibration set using the partial least squares regression method (GPLS) and (2) Three LOCAL approaches, where partial least square (PLS) models were developed for each prediction sample. Each model is based on a variable number of spectra from the total population selected based on their spectral similarity to the unknown sample. (2.1) Local weighted partial least square regression (LWPLSR) (Lesnoff et al., 2020). A weight was applied to each sample. This weight was calculated from the distance of Mahalanobis. (2.2) Local weighted partial least square regression performed using the k–nearest neighbours previously selected for each sample (kNN–LWPLSR). (2.3) Local weighted partial least square regression performed by averaging the prediction of the kNN–LWPLSR values obtained using models built from a different number of latent variables (kNN–LWPLSR_agg). The same validation set was used for the four models. Models performance for each model was assessed by the coefficient of determination of the external validation (R^2V) and the standard error of prediction (SEP). The SEP was decomposed into bias and SEP corrected by bias (SEP(c)). Bias and SEP(c) values associated to different models were compared using the methodology suggested by Fearn, (1996).

3.2 Results & Discussion

The mean, standard deviation, minimum and maximum values for the OMD of faeces for both calibration and validation sets are given in **Table 3.1**. Calibration and validation sets covered similar ranges for each component and gave similar mean and standard deviation values.

The values for the standard error of the GSM for *in vivo* measurements were calculated from a trial in which in vivo OMD was measured in 16 beef cows successively fed with 2 diets: a 100% permanent grassland hay and a diet based on 67% corn silage and 33% concentrate (De la Torre et al., 2019). For each diet, OMD measurement was repeated 2 times. For OMD_M, the standard error of the in vivo measurement was not possible to be calculated.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

Table 3.1: Descriptive statistics of calibration and validation samples for in vivo organic matter digestibility (OMD_GSM) and marker-estimated organic matter digestibility (OMD_M).

	Calibration					Validation					
	N	Mean	Min	Max	Sd	N	Mean	Min	Max	Sd	Se
OMD_GSM	380	71.88	59.73	88.40	4.17	96	70.37	62.96	77.50	3.75	1.33
OMD_M	553	73.44	55.11	87.89	4.38	143	73.78	61.24	83.72	3.92	-

N: number of samples; Min: minimum value; Max: maximum value; Sd: standard deviation; Se: standard error.

Faecal Near-Infrared Spectroscopy Predictions

The faecal NIR spectroscopy validation statistics for the prediction of OMD_GSM and OMD_M obtained by the four approaches are given in **Tables 3.2 and 3.3**, respectively.

For the prediction of OMD_GSM models were characterized by R^2 values comprised between 0.68 and 0.81 and SEP values ranging between 1.64 et 2.13%. When the SEP values were decomposed into the bias and SEP(c) the random error (SEP(c)) was higher than the systematic error (bias) for all models. No significant differences for bias were found between models whereas random errors associated to models on local PLS regression were lower than those associated to the GPLS regression.
Prediction errors associated to local models (around 1.65%) were close to that estimated for the Se_GSM (1.33%). This result suggests that 0.81 of the SEP error was explained by the Se_GSM whereas 0.19 of the SEP was attributed to the NIRS error (analytical repeatability + model prediction inaccuracy) but also to the variability of the faeces samples coming from a same animal/diet. From the SEP errors, the calculated minimum detectable difference for the OMD measured by the GSM (95%CI) between two animals was 5.20% whereas the calculated minimum detectable difference between two animals for OMD predicted by NIRS was 6.40%. That means that for OMD_GSM, with the obtained prediction error, we can theoretically detect significant differences (CI 95%) between 2 groups of animals differing in 3 points of OMD if each group is composed of 5 animals vs. 3 animals per group is required for detecting the same difference in OMD measured by the GSM. For detecting significant differences between two groups of animals differing 2 points of OMD, it would be necessary to predict the OMD_GSM of 11 animals per group vs. 7 animals per group is required for detecting the same difference in digestibility measured with the GSM.

For OMD_M, SEP and R^2 values were poorer than those obtained for OMD_GSM. The SEP values varied between 1.98 and 2.25% whereas R^2 values were comprised between 0.66 and 0.74. As for OMD_GSM, most of the SEP value was attributed to the random error. For SEP(c) kNN–LWPLSR_agg and LWPLSR models were characterized by lower bias and SEP(c) than GPLSR model.
For OMD_M the minimum detectable difference when the NIRS model is used is 7.8% (CI 95%): 7 animals per group would be necessary for detecting significant differences between 2 groups of animals differing in 3 points of OMD_M, whereas 16 animals per group would be necessary for detecting significant differences between 2 groups of animals differing in 2 points of OMD_M.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

Table 3.2. Performance of faecal NIRS models for predicting organic matter digestibility measured with the gold standard method (OMD_GSM).

	OMD_GSM				
	LV	SEP	R^2	Bias	SEP(c)
GPLSR	11	2.13	0.68	-0.25a	2.12a
LWPLSR	2	1.67	0.80	-0.09a	1.68b
kNN-LWPLSR	2	1.64	0.81	-0.14a	1.65b
kNN-LWPLSR_agg	1_20	1.64	0.81	0.1a	1.65b

LV: number of latent variables; SEP: standard error of prediction; R^2 coefficient of determination; SEP(c): standard error of prediction corrected by bias; GPLSR: classic partial least square regression; LWPLSR: Local weighted partial least square regression; kNN–LWPLSR: k–near neighbours-Local PLS regression; kNN–LWPLSR_agg: aggregated k–near neighbours-Local PLS regression

Table 3.3. Performance of faecal NIRS models for predicting organic matter digestibility estimated by markers (OMD_M).

	OMD_M				
	LV	SEP	R^2	Bias	SEP(c)
GPLSR	19	2.27	0.66	0.40a	2.25a
LWPLSR	12	2.00	0.74	0.23ab	1.99b
KNN–LWPLSR	12	2.06	0.72	0.23ab	2.05ab
KNN–LWPLSR_agg	2_20	1.98	0.74	0.17b	1.98b

LV: number of latent variables; SEP: standard error of prediction; R^2 coefficient of determination; SEP(c): standard error of prediction corrected by bias; GPLSR: classic partial least square regression; LWPLSR: Local weighted partial least square regression; kNN–LWPLSR: k–near neighbours-Local PLS regression; kNN–LWPLSR_agg: aggregated k–near neighbours-Local PLS regression

3.3 Conclusion

This work confirms that faecal NIRS is an interesting proxy to predict OMD in dairy and beef cattle and therefore constitutes an alternative method to the GSM (using stalls constraining for the animal), when taking into account the 3R principles in animal experimentation. The accuracy of NIRS for predicting OMD is close to that of the GSM. Hence, only a slightly higher number of animals is required when using faecal NIRS predictions than when using the GSM. This is an important point to consider for implementing this tool at large scale in research infrastructures. Next step will consist in including reference data not yet represented in the database (beef cows and males) to enlarge the domain of validity of the model. It also remains to test the robustness of the model when it is applied on spot faecal samples.

The classic GPLSR is commonly used to predict OMD in cattle. However, our results suggest that the Local methods were more appropriated than classic GPLSR for predicting OMD_GSM. Most of the Local methods, particularly LWPLSR_agg and LWPLSR, were also better than classic GPLSR model for predicting OMD_M. The Local methods improve the accuracy of prediction of OMD in cattle by 33% compared to that of conventional GPLSR.

3.4 Abstracts & peer-reviewed articles from this work

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

Andueza, D., Nozière, P., Herremans, S., De La Torre, A., Froidmont, E., Picard, F., Pourrat, J., Constant, I., Martin, C., Cantalapiedra-Hijar, G., 2019. Faecal-NIRS for predicting digestibility and intake in cattle: efficacy of two calibration strategies. Book of Abstracts of the 70th Annual Meeting of the European Federation of Animal Science, 475.

Andueza, D., Picard, F., Pourrat, J., De la Torre, A., Devant, M., Reynolds, C.K., Froidmont, E., Bernard, L., Martin, C., Nozière, P., Cantalapiedra-Hijar, G., 2020. Faecal-VIS/NIR spectroscopy as a tool to predict animal-to-animal variation in feed organic matter digestibility in cattle. Book of Abstracts of the 71th Annual Meeting of the European Federation of Animal Science, 508.

Andueza, D., Picard, F., Pourrat, J., Vanlierde, A., Nozière, P., Cantalapiedra, G., Morgavi, D., De la Torre, A., Dehareng, F., Martin, C., Renand, G., 2022. Phenotyping of enteric methane emissions and intake from near-infrared spectra of beef cattle faeces. The 7th EAAP International Symposium on Energy and Protein Metabolism and Nutrition (ISEP 2022), Granada, Spain, (Sep.12-15, 2022).

Picard et al., 2022. Comparison of locally partial least square strategies for regression on nutritive value NIR data. Chemometrics and intelligent laboratory systems (in progress).

Andueza et al., 2022. NIR faeces for predicting digestibility across diets & individuals in dairy and beef cattle (in progress).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

4 Proxies to predict enteric methane emissions: milk MIR spectra

A model to estimate individual daily emissions of methane (CH_4) from lactating dairy cows has already been developed (Vanlierde et al., 2021; https://doi.org/10.1002/jsfa.10969) and is based on 1,080 reference data from 7 research centres (CRA-W-Belgium, Teagasc-Ireland, INRAE-France, AFBI-United-Kingdom (Northern Ireland), FBN-Germany, Aarhus-Denmark and Agroscope-Switzerland). This proxy is based on CH_4 reference data collected with respiration chamber and the SF_6 tracer technique, connected with the corresponding standardized milk mid-infrared (MIR) spectra. Statistics of this existing predictive model are $R^2$of calibration (R^2c) and of cross-validation (R^2cv) about 0.68 and 0.64 respectively, and standard error of calibration (SEc) and of cross-validation (SEcv) about 58 and 61 g/d of CH_4, respectively. The purpose within SmartCow was to have access to international reference data (CH_4 values and the corresponding standardized milk MIR spectra) not yet included in the calibration dataset used to build the model to establish validation statistics.

4.1 Material & methods

Due to previous collaborations, the published model already includes an important part of historical data from the SmartCow partners. Moreover, collection of CH_4 measurement is particularly expensive and time consuming. This is why a moderate amount of reference data obtained in respiration chambers have been collected (n=261) and available for this validation step. In each case, 24 hours of CH_4 measurement were related to a daily milk MIR spectra representative of the two milking. Reference data of CH_4 measured with GreenFeed system (© C-lock) are also available but due to difference of methodology related to reference measurement technique; these data are considered separately for the moment and cannot be considered as a validation step for the existing model.

Four additional historical datasets from FBN (Germany) have been shared with a total of 233 combinations of CH_4 values and corresponding standardized milk MIR spectra collected from 75 different dairy cows (Holstein) with and mean ± SD about 392 ± 55 g of CH_4/d (min – max: 214 – 546 g/d) collected in respiration chambers.

During a TNA carried out at WUR with Kelly Nichols (The Netherlands) including CH_4 measurements (respiration chambers), additional milk samples have been collected by the local research team and sent to CRA-W (Belgium) for analyse and acquisition of standardized milk MIR spectra. Thanks to this, 28 reference data from 28 Holstein cows were available. They present a mean ± SD about 416 ± 35 g/d of CH_4 (min – max: 354 – 475 g/d).

Another trial leaded at UREAD (UK) in the framework of the SmartCow WP5 and including 4 Holstein cows (4 periods in a Latin square design) permitted the collection of additional CH_4 reference values and milk MIRS analysed in UK on a standardized spectrophotometer as required. However, spectral information is not available when this report is written. Data will be considered later and are not detailed here.

4.2 Results & discussion

As a first step, the milk spectral variability of the validation data set has been compared to the spectral variability present in the reference data set used to develop the model. No spectra with a Mahalanobis distance (GH) higher than 5 was observed meaning that the spectral information is well covered by the model. However, to perform the validation test, only GH ≥ 3 were considered leading to 238 references values.

Existing predictive model detailed in the material and methods paragraph has been applied on the 238 new milk MIR spectra. The CH_4 predictions in function of the measured values are shown in **Figure 4.1.** No aberrant value of predicted CH_4 is observed. A R^2 of prediction (R^2p) of 0.17 is observed with a standard error

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

of prediction (SEp) of 52 g/d of CH_4 and a RPDp of 2. This last parameter is the most relevant during a validation step. We can notice that this error is in line and event lower than the SEcv of the model that is a very good point. This leads to the conclusion that, as expected, if spectral variability is covered, this proxy permits to obtain CH_4 predictions from standardized milk MIR spectra with an error of prediction in line with the known error of the model. This is feasible to distinguish high and low CH_4 emitters and to follow animals, herds, populations in time especially because the great interest of milk MIR spectra is that it is easily available as it is already collected in routine during the milk recording.

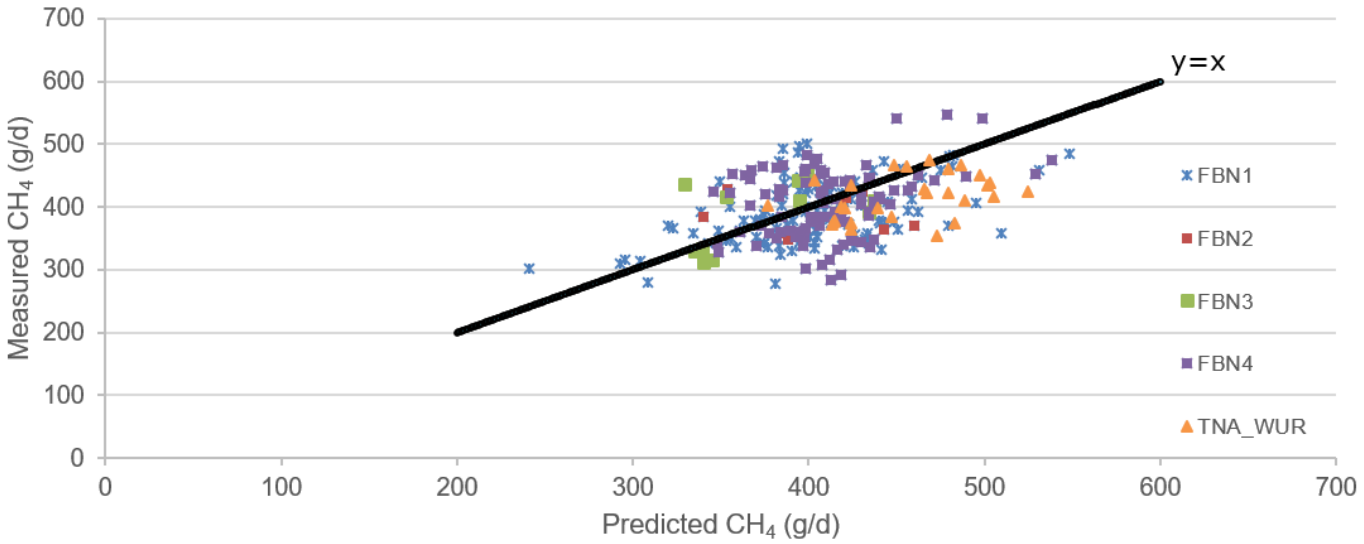


Figure 4.1: Measured CH_4 values in function of the predicted CH_4 values obtained from milk MIR spectra using the existing model during the validation step and considering the origin of the data.

As at this stage CH4 reference values collected with GreenFeed system cannot be considered for validation step of this existing equation based on reference values obtained with respiration chamber and SF_6 tracer technique. The best methodology to combine milk MIR spectra and CH_4 values obtained with GreenFeed have been investigated in Coppa et al. (2022). Various protocols for dataset management were tested to identify the best approach in using milk mid-infrared spectroscopy (MIR) to predict enteric methane (CH_4) emission data recorded with GreenFeed as a reference method. Increasing the duration of CH_4 measurement using GreenFeed and averaging spectra collected throughout the period optimized milk MIR predictive performance. Additional explanatory variables such as milk yield or fat and protein corrected milk improved performance. Specific models are required to reliably predict CH4 emissions from dairy cows fed CH4-mitigating diet without effect on milk composition.

4.3 Conclusion

Access to new reference data sets, including CH_4 values acquired in respiration chambers and corresponding milk MIR spectra, permitted to perform an external validation procedure about the existing predictive model. The performances obtained confirmed the expected range prediction accuracy. However, this model is evolving and acquire reference data including information not yet covered in the model (breed, diet…) to upgrade the model will be pursued.

For the moment, combining CH_4 reference data measured with the GreenFeed system with the existing model based on respiration chambers and SF6 tracer values is not relevant, without introducing noise in the model.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924

A specific model based on values collected with GreeeFeed system is in progress. Ideally, only one model of prediction (merging respiration chambers, SF6 tracer and GreenFeed values) would be the most relevant. This perspective is still investigated but a way to "correct" GreenFeed values to be exactly on the same foot than other reference values need to be found to avoid to introduce noise in the existing model, related to the difference of the measurement techniques.

4.4 Abstracts and peer-reviewed articles from this work

Vanlierde, A., Dehareng, F., Herremans, S., Nichols, K., Kuhla, B., Eugène, M., Martin, C., 2021. Milk mid-infrared spectra to estimate rumen fermentation parameters. In: 72nd Annual Meeting of the European Federation of Animal Science. Davos, Switzerland. 30.08.2021-03.09.2021 (Oral presentation)

Vanlierde, A., Martin, C., Picard, F., Rochette, Y., Dehareng, F., Andueza, D., 2022. Firsts results about potential use of fecal near infrared spectra to estimate daily methane emissions of dairy cows measured with GreenFeed system. In: Dair'Innov congress, 27-29 April 2022, Namur, Belgium. (Oral presentation)

Vanlierde A., Dehareng F., Mertens A., Mathot M., Lefevre A., Morel I., Renand G., Rochette Y., Picard F., Martin C. & Andueza D. Estimation of methane eructed by dairy and beef cattle using fecal near-infrared spectra. In: 73rd Annual Meeting of the European Federation of Animal Science. Porto, Portugal. (Oral presentation)

Coppa M., Vanlierde A., Bouchon M., Jurquet J., Musati M., Dehareng F., Martin C, 2022. Methodological approach for using milk mid-infrared spectra to predict enteric CH4 emission data from GreenFeed in cows. J3R, Paris

Coppa, M., Vanlierde, A., Bouchon, M., Jurquet, J., Musati, M., Dehareng, F., Martin, C., 2022. Methodological approach for using milk mid-infrared spectra to predict enteric CH4 emission data from GreenFeed in cows. Journal of Dairy Science (in revision).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°730924