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

Project ID: 730924

Deliverable number: D6.2

Deliverable title: Development of new equations from proxies to predict animal feed efficiency and its determinants in 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	18 May 2022		
V2	13 June 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 gold standard method(s) (GSM) 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 (CH_4) emissions.
Objectives	**The objective of this work was to develop new equations from 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. This concerns: - **the natural ^{15}N abundance in animal proteins (vs urea-N) and other plasma metabolites** for prediction of urinary nitrogen excretion (UNE) in dairy and beef cattle; **- the near-infrared spectra (NIRS) in faeces** for prediction of enteric CH4 emissions in cattle; **- the mid-infrared spectra (MIRS) in milk** for predicting rumen fermentation parameters (volatiles fatty acids, ammonia, pH) 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 (N urinary excretion, enteric CH4 emissions, rumen fermentation parameters) measured by GSM, and proxies (15N, NIRS, MIRS, plasma metabolites) 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).

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

	Collection of data and proxies were carried out from historical experiments and new experiments conducted during the SmartCow project. When proxies' data 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. First models were built to predict phenotypes across diets and between-individuals. Our choice in models was based on hypothesis of the literature and dataset availability acquired within SmartCow.
Results & implications	➔ **^{15}N abundance (vs urea-N) and other plasma metabolites for prediction of urinary N excretion (UNE) in dairy and beef cattle** The database included 513 individual records for N balance and laboratory data for $\Delta^{15}N_{animal-diet}$ in plasma or milk (n = 513) and urea concentration in plasma (n = 100), milk (n = 166) or both (n = 147). The natural ^{15}N abundance of animal proteins may predict changes in urinary N excretion (g/kg BW) across diets and experiments at the same level as urea-N in milk or in plasma do. However, this isotopic biomarker shows a slightly superiority compared to urea-N both in milk or plasma when urinary N excretion was expressed by unit of N intake. **Unlike what we have observed for the milk N use efficiency, the natural ^{15}N abundance of animal proteins failed to discriminate the between-animal variability in the urinary N excretion.** For plasma metabolites identification in beef cattle through a metabolic approach, the work was carried out with a dataset of 32 individual observations. Identification of 14 plasma metabolites with a moderate to high repeatability (r>0.4) have been identified as potential proxies to predict UNE in beef cattle. Among them, creatinine and serotonin were correlated at the individual level with total UNE in g/d but not when expressed in g/g N intake. Fourteen plasma metabolites were significantly impacted by the dietary CP level with Arg and Gly as the only ones repeatable. **Repeatable biomarkers of urinary N excretion identified could be used to discriminate dietary treatments impacting the N utilization and phenotyping individuals as high or low N polluters.** ➔ **Faecal NIRS for predicting enteric CH_4 emissions in cattle** The database included ~ 800 individual records for CH_4 emissions measured with 3 different methods (SF6, for SF6 tracer technique; RC for respiration chamber; GF for GreenFeed system) on beef and dairy cattle. Merging CH_4 emission datasets obtained from different measurement methods dit not impprove the potential of NIRS from faeces to predict enteric CH_4 emissions. **First promising results on the potential of faecal NIRS to predict enteric CH_4 emissions were obtained in beef cattle from a dataset of 346 CH_4 values measured with GF**. More CH_4 reference data and corresponding faecal NIRS are required to improve the robustness of this proxy and enlarge the range of application of the model (other breeds, diets, physiological status). ➔ **Milk MIRS for predicting rumen fermentation parameters in dairy cows**

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

	Potential milk MIRS for rumen diagnostics has been investigated using limited datasets (n~200) with a poor variability in terms of health status of the rumen to discriminate individuals. It was not possible to conclude if milk MIRS is relevant or not to predict rumen fermentation parameters. First results highlighted the need to measure repeatability of proxies across time on dedicated trials and according to common protocols and standardized procedures. This is a key step to be able to merge data sets and develop robust proxies while saving time and money. Another investigation of research might be to work on "cluster models" related to health status rather than to an exact value.

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

Table of contents

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

1 The use of ^{15}N, urea-N and other plasma metabolites to predict the urinary N excretion (UNE) in dairy and beef cattle

1.1 ^{15}N vs urea-N concentration in milk or plasma

The objective of this task was to explore the potential of the natural ^{15}N enrichment of animal proteins over the diet ($\Delta^{15}N_{animal\text{-}diet}$) as a proxy of urinary N excretion (UNE) in ruminants (Bernard et al., 2020) compared to more established biomarkers such as plasma urea N (PUN) or milk urea N (MUN) (Kohn et al., 2005).

1.1.1 Material & methods

A database was built using N balance data from 17 trials provided by INRAE, University of Reading, CRA-W, LUKE and SRUC and conducted in dairy cows (14 trials and 465 observations) and growing-fattening beef cattle (3 trials and 48 observations). The database included 513 individual records for N balance and laboratory data for $\Delta^{15}N_{animal\text{-}diet}$ in plasma or milk (n = 513) and urea concentration in plasma (n = 100), milk (n = 166) or both (n = 147). Measurement periods for the determination of total urinary N excretion ranged between 4 and 10 days.

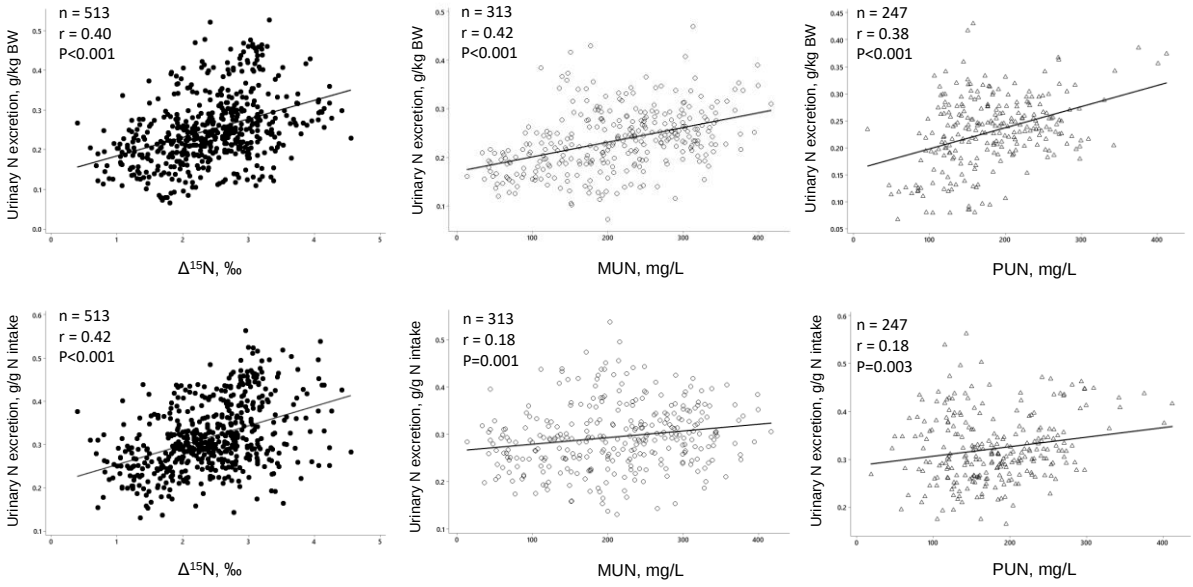


Figure 1.1. *Between-study relationships between urinary N excretion either expressed by kg of BW (top panels) or by unit of N intake (bottom panels) and the N isotopic fractionation ($\Delta^{15}N$animal-diet, ‰) or urea concentration (mg/L) in milk (MUN) or plasma (PUN) in fattening beef cattle.*

1.1.2 Results & discussion

Overall, $\Delta^{15}N_{animal\text{-}diet}$, MUN and PUN were all significantly correlated with UNE (between-study relationship, **Figure 1.1**) but in a different way depending on how UNE was expressed. When UNE was expressed in g/kg BW, both $\Delta^{15}N_{animal\text{-}diet}$

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

and MUN or PUN showed moderate and significant correlations (0.38≥r≥0.42; P<0.001). However, when UNE was expressed in relative terms (g of urinary N excreted by g of N intake) much lower, though significant, correlations were obtained with MUN/PUN (r=0.18) than with $\Delta^{15}N_{animal\text{-}diet}$ (r = 0.42). These results reinforce the idea that only $\Delta^{15}N_{animal\text{-}diet}$ seems to be biologically linked to the N partitioning between catabolism and anabolism (Cantalapiedra-Hijar et al., 2015) compared to urea in milk or blood which would rather reflect the absolute amounts of urinary N excretion.

spite an overall moderate correlation between both kind of biomarkers and UNE, they failed to discriminate individuals from the same contemporary group (same experiment, diet and measurement period) in terms of UNE since very low correlations were obtained when using $\Delta^{15}N_{animal\text{-}diet}$ or MUN to predict it at the individual level (r < 0.15; P<0.05). One exception that may deserve further exploration was the moderate correlation when using PUN rather than MUN to predict UNE at the individual level (0.27≤r≤0.33; P<0.001).

1.1.3 Conclusion

The natural ^{15}N abundance of animal proteins may predict changes in urinary N excretion (g/kg BW) across diets and experiments at the same level as milk or plasma urea do. However, this isotopic biomarker shows a slightly superiority compared to urea in milk or plasma when the urinary N excretion was expressed by unit of N intake. **Unlike what we have observed for the milk N use efficiency, the natural ^{15}N abundance of animal proteins failed to discriminate the between-animal variability in the urinary N excretion.**

1.1.4 Abstracts & peer-reviewed articles from this work

Nasrollahi, S. M., Nozière, P., Dewhurst, R. J., Chantelauze, C., Cheng, L., Froidmont, E., Martin, C., Cantalapiedra-Hijar, G. 2019. Natural 15N abundances in plasma and urea-N concentration in milk as biomarkers of urinary N excretion in dairy cows: a meta-analysis. 6th EAAP International Symposium on Energy and Protein Metabolism and Nutrition. 9-12.09.2019. Belo Horizonte, Brazil. Oral presentation

1.2 Identification of plasma metabolites in beef cattle through a metabolomics approach

The objective of this task was to identify repeatable plasma metabolites (i.e. metabolites whose plasma concentrations were highly correlated across time) of urinary N excretion (UNE) across diets but also between-individuals in beef cattle.

1.2.1 Material & methods

For this, plasma was sampled from animals involved in the N balance trial performed at INRAE (WP5; Bellagi et al., 2022) using two contrasted dietary CP levels (11.5% vs 17.0%CP). Sixteen Charolais young bulls were used and their UNE determined in metabolic stalls in two different measurement periods (first and second repetitions) with 1 month between them. Four consecutive measurement periods of N balance were conducted, each on 8 metabolism stalls and resulting in 32 observations. Blood was sampled before the morning meal on the last day of each of the N balance trial (d10), immediately centrifuged and the obtained plasma stored at -80°C. Plasma samples were analysed by LC-tandem mass spectrometry and colorimetric methods for quantifying 74 targeted metabolites. The effect of CP level and the interaction between CP level and the measurement period on metabolite plasma concentration was evaluated through repeated measures by mixed-effect models. For determining the repeatability and relationships at the individual level, all variables were first adjusted by the effects of CP level and measurement periods. Repeatability was then calculated as the regression coefficient of the relationship between plasma concentration obtained in first and second repetitions. Relationships at the individual level between UN and plasma metabolites were evaluated from adjusted values by mixed effect models integrating the interaction between plasma metabolite and CP level and the animal random effect. Significant effects or relationships were declared at P<0.05.

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

1.2.2 Results & discussion

The urinary N excretion was very contrasted across the two dietary CP levels, with 118% higher flux in high (98.8 g/d) vs low (45.3 g/d) CP level (P<0.001). Among the 72 metabolites measured in plasma, 14 showed moderate to high repeatability (r>0.40) with only 9 showing significant correlation between first and second repetition (**Table 1.1**). Highest repeatability was found for plasma concentration of 3-indole propionate (r = 0.86), 3-indole acetate (r = 0.70) and creatinine (r = 0.73), the two formers having a microbial origin while the latter having an endogenous metabolic origin. Fourteen metabolites were significantly impacted by the diet (P<0.05) regardless of the measurement periods (CP level × period; P>0.05). Phenylacetylglycine, p-cresol sulfate and urea were the plasma metabolites most significantly impacted by the dietary CP level (P<0.001), all of them already identified as urinary biomarkers of UN in dairy cows (Boudra et al., 2021). Phenylacetylglycine and p-cresol are microbial metabolites from phenylalanine and tyrosine degradation in the rumen, respectively. These three plasma metabolites were able to discriminate 100% of individuals in any of the two measurement periods according to their dietary treatment (and thus to their UN) and were related to each other through the protein degradation and deamination pathways in the rumen. Seven plasma metabolites showed significant correlations with UN (g/d) at the individual level (P<0.05) and regardless of dietary CP level, with homocysteine (r = -0.63), serotonin (r = 0.55) and creatinine (r = -0.51) showing the highest correlation (**Figure 1.2**).

Table 1.1. *Repeatable plasma metabolites in fattening beef cattle*

	Repeatability[#]
High repeatability[1] (0.7≤r≤0.9)	
Indolepropionic acid	0.86*
Indoleacetic acid	0.70*
Creatinine	0.70*
Moderate repeatability[1] (0.4≤r<0.7)	
5-Aminovaleric acid	0.63*
β-Aminobutyric acid	0.63*
Homoarginine	0.61*
Arginine	0.59*
Glutamine	0.52*
Glycine	0.51*
Citrulline	0.48
Serotonin	0.48
Tyrosine	0.47
p-Cresol sulfate	0.45
Betaine	0.44
L-Anserine	0.41

[#]Repeatability defined as the tendency of individuals to maintain over time their ranking based on plasma concentration of metabolites and calculated as the intra-class correlation coefficient (see material and methods)
[1]According to Martin and Bateson, 1986
* P< 0.05 between repetition 1 and 2

However, these 7 metabolites were not correlated at the individual level when UN was expressed by unit of N intake (g/g), in which case we found that the plasma concentration of 5 amino acids and of kynurenine were correlated (P<0.05). Among them, Ala, Thr and Val showed the highest correlations with the urinary N-to-N intake ratio (**Figure**

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

1.2). No common metabolites were found among those impacted by the dietary CP level and those correlated with urinary N excretion at the individual level expressed either as g/d or g/g of N intake.

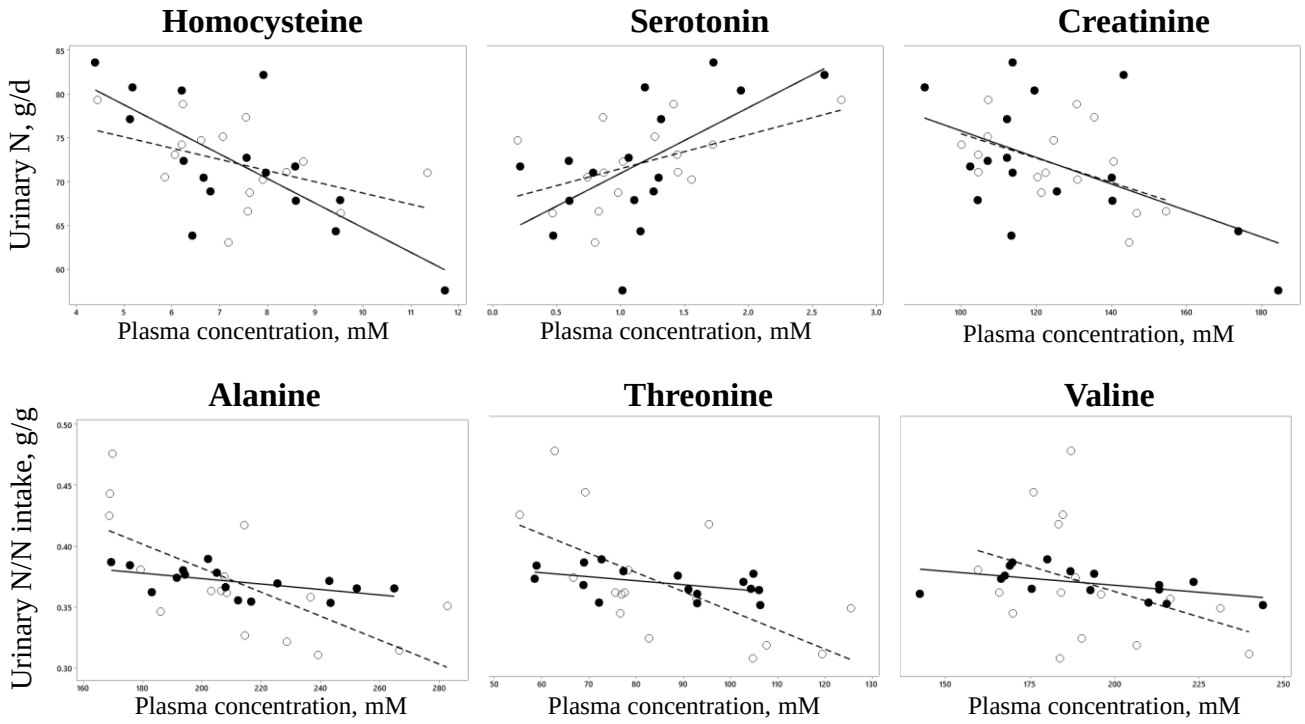


Figure 1.2. *Significant relationships (P<0.05) at the individual level between urinary N excretion (g/d, 3 panels on the top) or urinary N excretion by unit of N intake (g/g, 3 panels at the bottom) and plasma metabolite concentrations. Variables are adjusted by the measurement period (1 to 4) and diet (low vs high CP) effects so individual variability can be evaluated. Symbols: High CP diet (●) and Low CP diet (○).*

1.2.3 Conclusion

Fourteen plasma metabolites here evaluated showed a moderate to high repeatability (r>0.4). Among them, creatinine and serotonin were correlated at the individual level with total urinary N excretion in g/d, but not when expressed in g/g N intake, in which case no repeatable metabolites were found. Fourteen plasma metabolites were significantly impacted by the dietary CP level, with Arg and Gly as the only ones repeatable. Repeatable biomarkers of urinary N excretion here identified could be used to discriminate dietary treatments impacting the N utilization and phenotyping individuals as high or low N polluters

1.2.4 Abstracts & peer-reviewed articles from this work

Cantalapiedra-Hijar G., Bellagi R., Salis L., Baumont R., Noziere P. (2022). Plasma biomarkers of urinary N excretion in beef cattle: repeatability and relationships at the dietary and individual level. 7th EAAP International Symposium on Energy and Protein Metabolism and Nutrition. 12-15 September, Granada (Spain). Accepted

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

2 Faecal NIR spectra to estimate methane emissions

Proxies to estimate methane (CH_4) emissions exist (e.g. Milk MIR spectra as discussed in task 6.1) but as they are based on milk analyses, they are limited to lactating dairy cows. The objective of this task was to collect data to investigate the feasibility to develop a proxy to estimate individual CH_4 emissions if milk is not available, so more especially for young cattle, heifers, dry cows and beef cattle. Due to common metabolic origins at ruminal level during fermentation processes, faecal composition and more especially near-infrared spectra (NIR) of faeces are of interest for such proxy. Moreover, CH_4 emissions are closely related to the digestible OM (OMD) (Moss et al., 2000) while OMD can be quantified from faecal NIR spectra (Deacreuyenaere et al., 2009).

2.1 Material & methods

A crucial and challenging step was the phase of reference data collection. The historical data including CH_4 measurement and faecal NIR spectra and merged within SmartCow are coming from three partners as detailed below.

- **INRAE (France)** permitted the use of data from 7 different projects and 1 TNA project. Data were from dairy cows (Holstein) and beef cattle (Charolaise). Depending of the dataset animals received a diet based on grassland, hay, grass silage or corn silage (**n = 637**). Five of the French datasets presented weekly faecal NIR spectra while the three others were related to punctual spot sampling. Regarding CH_4 values, one data set is related to SF_6 tracer technique (SF_6), 3 to respiration chambers (RC) and 2 to GreenFeed (GF) systems only, while 1 is combining data acquired almost simultaneously with GF and SF_6, and another 1 with RC and SF6.

- **CRA-W (Belgium)** provided historical data from beef cattle (reformed Belgian blue cows, calf, heifers and suckling dual purpose Belgian blue cows) and from lactating dairy cows (Holstein) receiving diet based on grassland, grass silage or hay (**n = 93**); GreenFeed system was used to measure CH4 emissions and **"spot" faecal** sampling were taken.

- **Agroscope (Switzerland)** shared historical data from beef cattle (Simmental, Angus, Limousine) receiving a diet based on corn silage (**n = 56**). Methane values were obtained using GreenFeed system while faecal sampling was **"weekly"-type**.

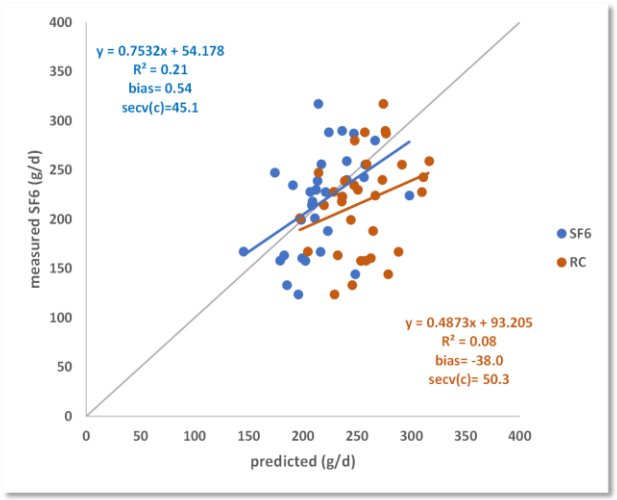
The amount of data that hase been brought together thanks to collaboration is noteworthy. The various type of diets, breeds, metabolic status and other animal variabilities is very interesting in reference values to develop a robust proxy and also permit to observe if one and unique proxy makes sense or if it is more relevant to develop a proxy per breed, per diet, etc. **However, the various reference methods used to measure CH_4, and the various protocols to collect faecal samples represent an issue.** Indeed, considering these datasets together will introduce cumulative bias and error in the model. One of the objectives of the SmartCow project was to highlight this issue on historical data and define recommended protocols for the future collection of data by all the research teams to facilitate the sharing of the data even if they are not directly collected with this purpose. In this case, with historical data, considering together data with identical protocols only is the most relevant way to proceed. However, this considerably reduce the amount of data included in a mode and multiply the number of models which are then highly dependent of the reference dataset considered and the reference technique used which is not the purpose. This is why calculation of corrective factors for CH_4 reference values obtained with different techniques has been investigated as some trials from INRAE included CH_4 values collected almost simultaneously with different techniques on the same cows. This step is detailed in the following section.

2.2 Results & discussion

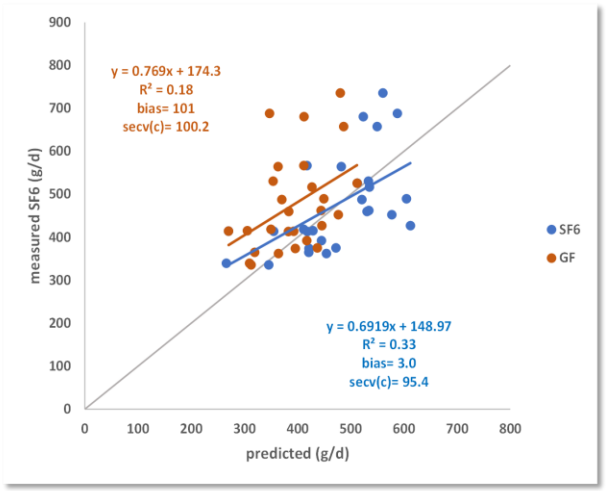
2.2.1 Correction of reference CH4 values as a function of the reference method considered

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

In this work based on 2 *in vivo* studies, we used SF6 measurements as the reference. In study 1, RC and SF6 measurements were performed simultaneously leading to 2 NIRS models based on 32 RC and SF6 methane values in non-lactating cows. In study 2, SF6 and GF values were collected simultaneously and 2 NIRS models were developed using 24 values from dairy cows. In both studies, bias was non-significant compared to the standard error of cross-validation corrected by the bias (SECVC), when values issued from models based on SF6 were compared to SF6 values (bias= 0.54 g/d and SECVC= 45.1 g/d for study 1 and bias= 3.00 g/d and SECVC= 95.4 g/d for study 2; **Figure 2.1**). However, if the cross-validation is done on RC or GF and compared to SF6 values, bias greatly increased (biais and SECVC= -38.0 g/d and 50.3 g/d in study 1 and 101.2 g/d and 100.4 g/d in study 2, respectively; **Figure 2.1**). These results suggest that SECVC values were similar for models built considering different methods within each study but there was a bias between methods. **This bias was more important for GF than for RC when compared to SF6 method. These results demonstrated that the combination of CH_4 emission datasets obtained from different measurement methods decrease the prediction potential of the models.**



CH4 measured SF_6
vs
CH4 predicted SF_6 or CH4 predicted RC values



CH4 measured SF_6
Vs
CH4 predicted SF_6 or CH4 predicted GF values

***Figure 2.1**: Comparison between CH_4 emissions measured with the SF6 method and the cross-validated CH_4 predicted values from faecal NIRS spectra with the SF6, RC or GF models.*

2.2.2 Development of first models to estimate enteric CH_4 from faecal NIR spectra

Considering that the previously detailed correction factors for CH_4 values collected with different reference techniques are very dependent of the reference data sets considered and that is risky for the moment to use them, as they are, only raw reference data have been used to develop first predictive models. Reference data collected from dairy cattle were considered separately to the ones collected on beef cattle.

Dairy cows

To optimize the standardisation of protocols between datasets (especially for CH_4 aspects), the number of data available for now and in the future (allowing a prospect of improvement) only datasets including data measured with GreenFeed system have been considered for the moment. At the end, 91 data were considered to develop the first model dedicated to dairy cows (**Table 2.1**). Data from a 4th dataset contain weekly faecal sample while other datasets include spot faecal sample.

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

Table 2.1: *Distribution of data considered for the proxy to estimate CH_4 from feaces NIR spectra for dairy cows*

Dataset	N	Ref. CH_4 (g/day, mean ± SD)	Based diet	Faecal sample Type
1	45	368 ± 68	Grassland or Corn silage	Spot
2	19	294 ± 59	Corn silage	Spot
4	27	395 ± 103	Corn silage	Weekly
Total	**91**	**361 ± 86**	**/**	**/**

Based on the 91 data, the first version of the prediction model dedicated to dairy cows presented a R^2 of calibration (R^2c) of 0.43 and a standard error of calibration (SEc) of 65 g of CH_4/day. Even if this R^2 is modest, this level of error (18% of the reference values mean) is encouraging regarding the feasibility to consider this proxy to estimate CH_4. **Figure 2.2** represents the measured values as a function of the predicted ones and the trend clearly appears. Moreover, the added values linked to the pooling of the datasets can be noted regarding the distribution of the CH_4 values.

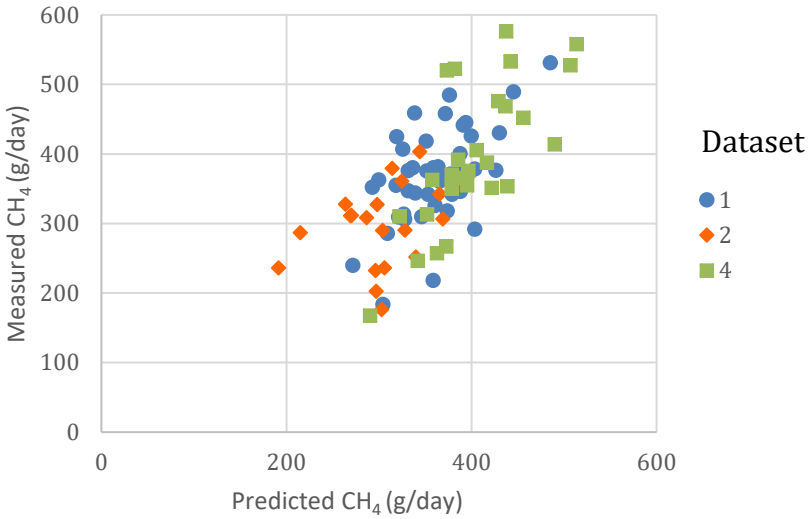


Figure 2.2*: Measured CH_4 values as a function of the predicted ones using the first version of the proxy to estimate CH_4 from faecal NIR spectra for dairy cows (n=91)*

However, considering a 4-groups cross-validation step, the R^2cv and SEcv observed were about 0.15 and 79 g CH_4/d respectively. The important decrease of the R^2 indicate no robustness of this model at this stage. That could be expected regarding the low amount of data considered. More data need to be collected to conclude about the relevance of this proxy dedicated to lactating dairy cows.

On another hand, considering faecal NIR spectra to estimate CH_4 is more relevant for cattle with no milk production. An interesting future prospect to estimate CH_4 for dairy cows would be the combination of milk MIR and faecal NIR spectral information to observe if the accuracy of prediction increase. However, in practice, this approach is much less easy than considering only milk MIR spectra only.

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

Beef cattle

To homogenize the reference data considered for the development of this first version of model dedicated to beef cattle, and regarding the amount of data available, data from Switzerland (based on a weekly-type of faecal sampling) were not considered in this first version. The 346 data considered to develop this model are detailed in **Table 2.2**.

Table 2.2: *Distribution of data considered for the proxy to estimate CH4 from feaces NIR spectra for beef cattle.*

Dataset	n	Breed	Ref. CH_4 (g/day, mean ± SD)	Based diet	Faecal sample type
CRA-W	83	Belgian blue (BB) and dual purpose BB	264 ± 47	Grass (fresh and silage) or hay	Spot
INRAE	263	Charolais	207 ± 32	Grass silage & hay	Spot
Total	**346**	**/**	**220 ± 42**	**/**	**/**

Even if both datasets were collected using spot faecal sampling, the spectral information are different depending of the dataset **(Figure 2.3)**. This is a good illustration of the interest to collate datasets to increase the covered information when a new model is in development. To be able to do this, sampling and analyses protocols need to be standardized.

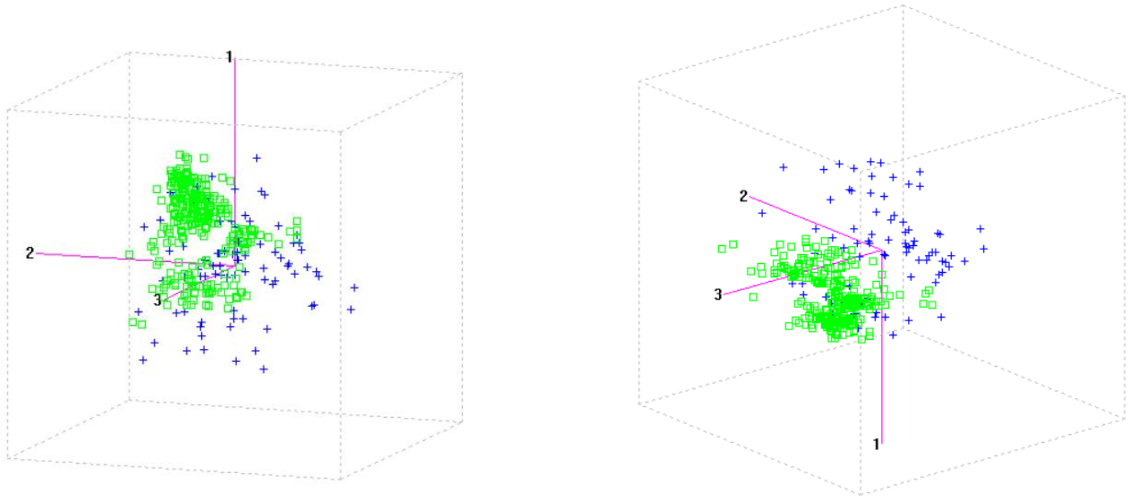


Figure 2.3: *Representations of variability of spectral information in funtion of the reference dataset (□ INRAE – Charolais, + CRA-W – Belgian blue and dual purpose belgiam blue) in function of the three first principal components related to milk spectral information.*

Based on the 346 data considered, the first version of the prediction model dedicated to beef cattle presented a R^2 of calibration (R^2c) of 0.62 and a standard error of calibration (SEc) of 26 g of CH_4/day. This error around 12% of the mean of reference data is very encouraging regarding the feasibility to consider this proxy to estimate CH_4. The **Figure 2.4** represents the measured values as a function of the predicted ones and the trend clearly appears. In addition to the added value linked to the pooling of the datasets regarding the variability of spectral information, the complementarity of the two reference datasets can be also be observed in **Figure 2.4** regarding the distribution of the CH_4 values. In this case, the range of CH_4 value cannot be attributed to a breed or a diet because the physiological status of the animals is also very different: data from INRAE are from heifers while data from CRA-W are mainly collected on reformed or suckling cows. A next step could be to consider the physiological status of animal in the model or in the unit (eg. CH_4 per kg of $BW^{0.75}$).

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

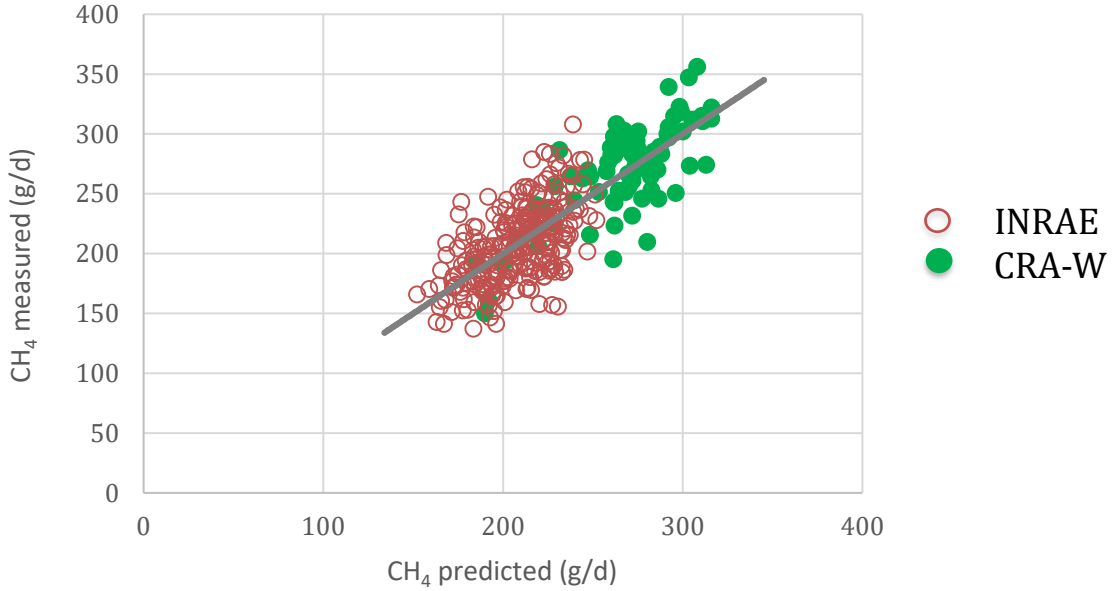


Figure 2.4: *Measured CH_4 values as a function of the predicted ones using the first version of the proxy to estimate CH_4 from faecal NIR spectra for beef cattle (n=346).*

Considering a 4-groups cross-validation step, the R^2cv and SEcv observed were about 0.55 and 29 g CH_4/d respectively. As the decreasing of the R^2 between calibration and cross validation is much less important than for the model dedicated to dairy cattle, this model dedicated to beef cattle is more robust. This can be expected as the amount and the variability (at zootechnical, spectral and reference values levels) or the reference data of this version is more important. **This first version of this proxy is very promising regarding the feasibility to estimate CH_4 from faeces NIR spectra. The integration of the other datasets has to be investigated and the collection of new reference data with identical protocols of sampling and analysis should be performed to include them in the model.**

2.3 Conclusion

Merging CH4 emission datasets obtained from different measurement methods (SF6, RC, GF) decreased the potential of NIRS from faeces to predict enteric CH4 emissions. First results on the potential of faecal NIRS to predict enteric CH_4 emissions in beef cattle are promising. More CH_4 reference data and corresponding faecal NIRS are required to confirm the trends observed on beef cattle, to improve the robustness of this proxy and enlarge the range of application of the model (other breeds, diets, physiological status). For dairy cows the combination of faecal NIR with milk MIR spectral information will be investigated to observe if the accuracy of prediction models increase. However, in practice, this approach is much less easy than considering only milk MIR spectra only.

2.4 Abstracts & peer-reviewed articles from this work

Andueza, D., Picard, F., Rochette, Y., Pourrat, J., Vanlierde, A., Dehareng, F., Morgavi, D., Martin, C. 2021. Influence of measurement method of methane on the performances of faecal NIRS models in cattle. In: 72nd Annual Meeting of the European Federation of Animal Science, Davos, Switzerland. 30.08.2021-03.09.2021 (poster)

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. In: 7th EAAP International Symposium on Energy and Protein Metabolism and Nutrition, Granada, Spain. September. 12-15.09.2022 (poster).

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

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, Namur, Belgium. 27-29.04.2022 (Oral presentation)

Vanlierde, A., Dehareng, F., Mertens, A., Mathot, M., Lefevre, A., Morel, I., Renand, G., Rochette, Y., Picard, F., Martin, C., Andueza, D., 2022. 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. 05-09.09-2022 (Oral presentation)

Vanlierde, A., et al. First models to estimate enteric CH4 emissions from faeces NIR spectra in cattle (in progress for Animals)

.

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

3 Milk MIR spectra as a new proxy for ruminal fermentation parameters

Due to metabolic processes, milk composition is influenced by ruminal fermentation processes. This is why it is relevant to test if it is feasible to quantify ruminal fermentation parameters of interest or a ruminal condition from milk mid-infrared (MIR) spectra. Indeed, milk MIR spectra reflect the chemical bounds in the milk and thus, indirectly, is a reflect of milk composition. Moreover, it is available in routine and at reasonable cost. Main volatile fatty acids (VFA) and ammonia (NH3) concentrations in the rumen, as well as the pH were the main parameters considered to be estimated from milk MIR spectra within SmartCow. Indeed, developing a proxy to estimate these parameters could permit to detect early stages of rumen disorders.

3.1 Material & methods

Reference data (n~200) from CRA-W, INRAE, FBN were coming from rumen juice sampling and related to a milk MIR spectrum collected the same day or at a close date. Most of the datasets shared by the partners thanks to SmartCow were historical dataset collected with specific protocols and purposes to each trial. Once again, we were facing to different data sets very difficult to merge due to these differences in sampling protocols. As detailed in **Table 3.1**, the delay between the rumen juice collection and the feeding moment (moment feeding line) varies between each dataset while the kinetic of the parameters of interest are deeply impacted by this delay. Consequently, merging datasets varying on this parameter is not an option. Moreover, the delay between rumen juice sampling is the same day for most data sets but is more important for the two others.

Table 3.1: *Description of reference data available including ruminal parameter(s) and corresponding milk MIR spectra*

	CRA-W	INRAE	INRAE	INRAE	FBN
Type	Historical	Historical	Historical	TNA	Historical
N animal	12	79	28	19	58
Delay between rumen juice sampling & feeding time (hours)	H0, +2, +4	H-1	H+3	H0, +1.5, +3.5, +5.5, +11.5, +21.5	H-1, +0.75
N data (animals x sampling time)	36	79	28	114	116
Delay between rumen juice sampling & milk spectra (days)	D-day	+4 days	+7-8 days	D-day, -1, +1	D-day, -1, +1
Spectral standardization	Ok	Ok	Ok	To be done	Ok

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

3.2 First results based on datasets available

3.2.1 pH estimation from milk MIR spectra

Regarding this parameter, the datasets available did not include enough variability to develop a model. Indeed, all pH data available vary between 7.8 and 6.0, which correspond to the same physiological status. There was no possibility to develop a calibration on these fixed values in parallel with milk spectral information. To distinguish animal with a different rumen status/health could be feasible but no data linked to cows presenting rumen health disorders were available so it was not possible to test it. To collect data from animals with a larger range of variation of pH values, including low rumen pH representative of health disorders, would be interesting.

3.2.2 Rumen VFA from milk MIR spectra

Only the results based on the dataset including the larger amount of data is detailed here: INRAE dataset including 79 reference data. Four data were considered as outliers. Based on this dataset, major VFAs (acetate, propionate, butyrate) were not related to milk MIR spectra. However, interesting trends were observed for isoButyrate and isoValerate. As detailed in **Figure 3.1**, more data need to be collected to make any conclusion about the feasibility or not to consider milk MIR spectra as a proxy to estimate rumen VFA profile. Moreover, as mentioned for pH, it would be much more relevant to try to predict abnormal proportions of VFA more than exact values of each of them but for this, data linked to animal presenting rumen disorders are needed and we did not have such data in SmartCow database.

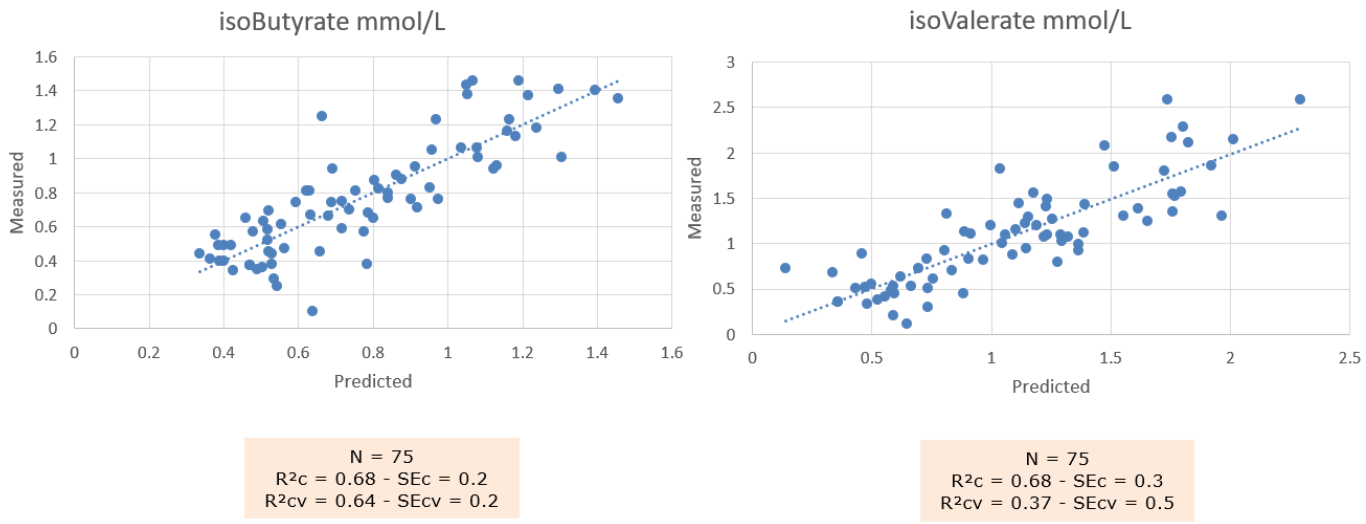


Figure 3.1. *Prediction of the rumen VFA concentration from corresponding milk MIR spectra in dairy cows from different data sets collected within SmartCow.*

3.2.3 Rumen ammonia from milk MIR spectra

Finally, regarding the estimation of rumen NH_3 level from milk MIR spectra, it was no feasible to merge dataset due difference of sampling protocols. Anyway, within each dataset available, interesting trend were observed between measured and predicted values as shown in **Figure 3.2**. This has to be confirmed with larger datasets collected with identical protocols. Moreover, as mentioned before, it would be very interesting to also have data related to sick cows to observe if it is possible to detect these animals based on milk MIR spectra.

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

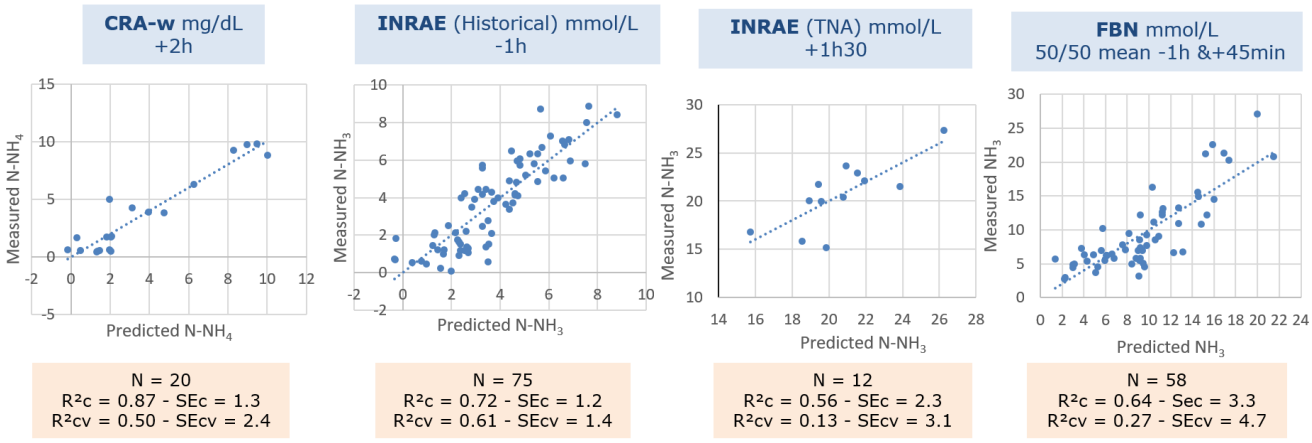


Figure 3.2. *Prediction of the rumen ammonia concentration (N-NH3) with corresponding milk MIR spectra in dairy cows from different data sets collected within SmartCow.*

3.3 Conclusion

It was not possible to conclude if milk MIR spectra are relevant or not to predict rumen fermentation parameters with the datasets available: low number of reference data with a poor variability in terms of health status of the rumen to discriminate individuals. First results highlighted the need to measure repeatability of proxies across time on dedicated trials and according to common protocols and standardized procedures. This is a key step to be able to merge data sets and develop robust proxies while saving time and money. Another investigation of research might be to work on "cluster models" related to health status rather than to an exact value. Finally, this approach is an interesting perspective regarding the strong limitation in the future to have fistulated animals to predict rumen status.

3.4 Abstracts & 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)

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