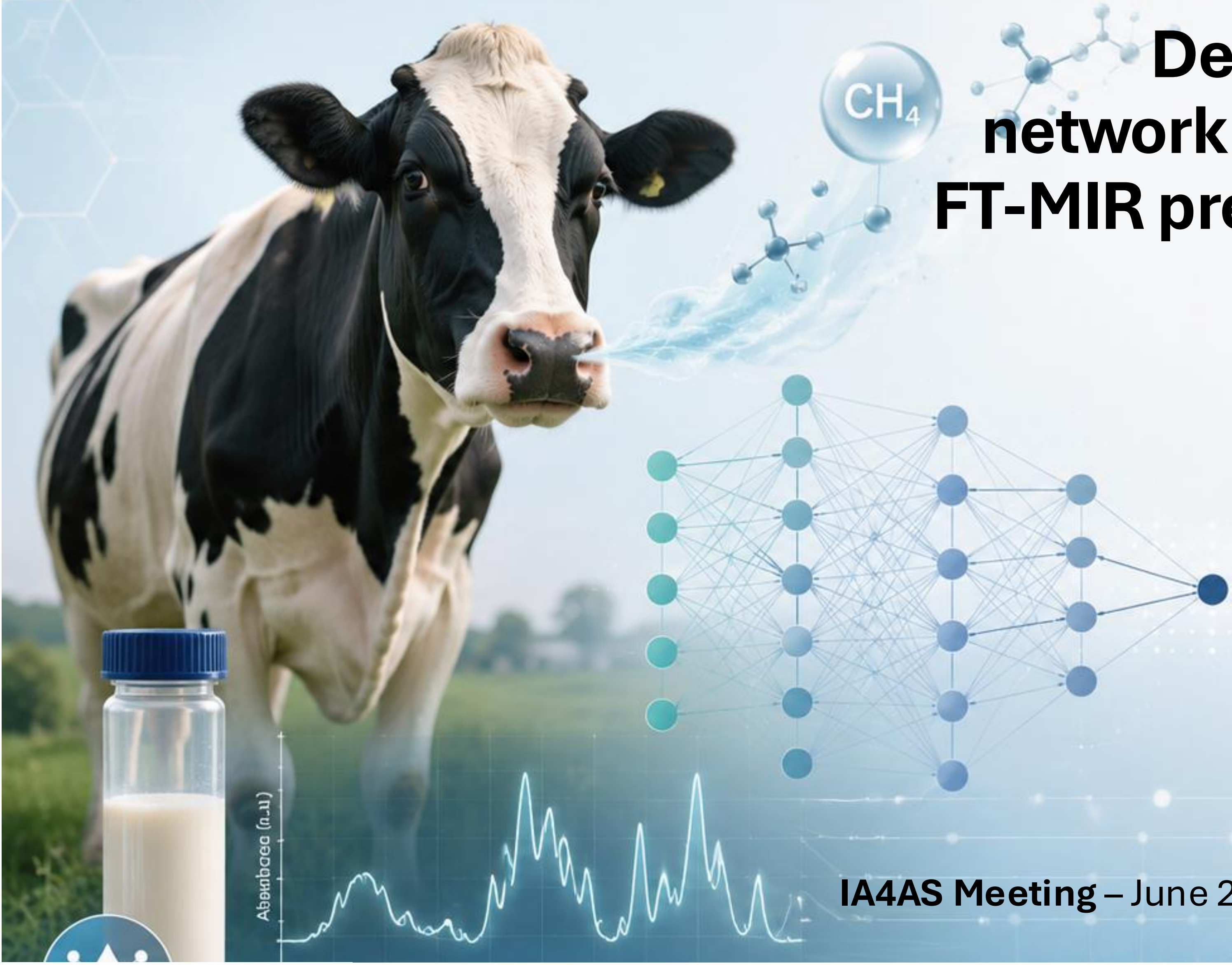
Development of robust neural network architectures for federated FT-MIR prediction of enteric methane emissions in dairy cows

H. Soyeurt, F. Dehareng, N. Gengler, S. McParland, M. Kreuzer, P. Lund, C. Martin, B. Kuhla, A. Vanlierde



IA4AS Meeting – June 2026 - Ghent

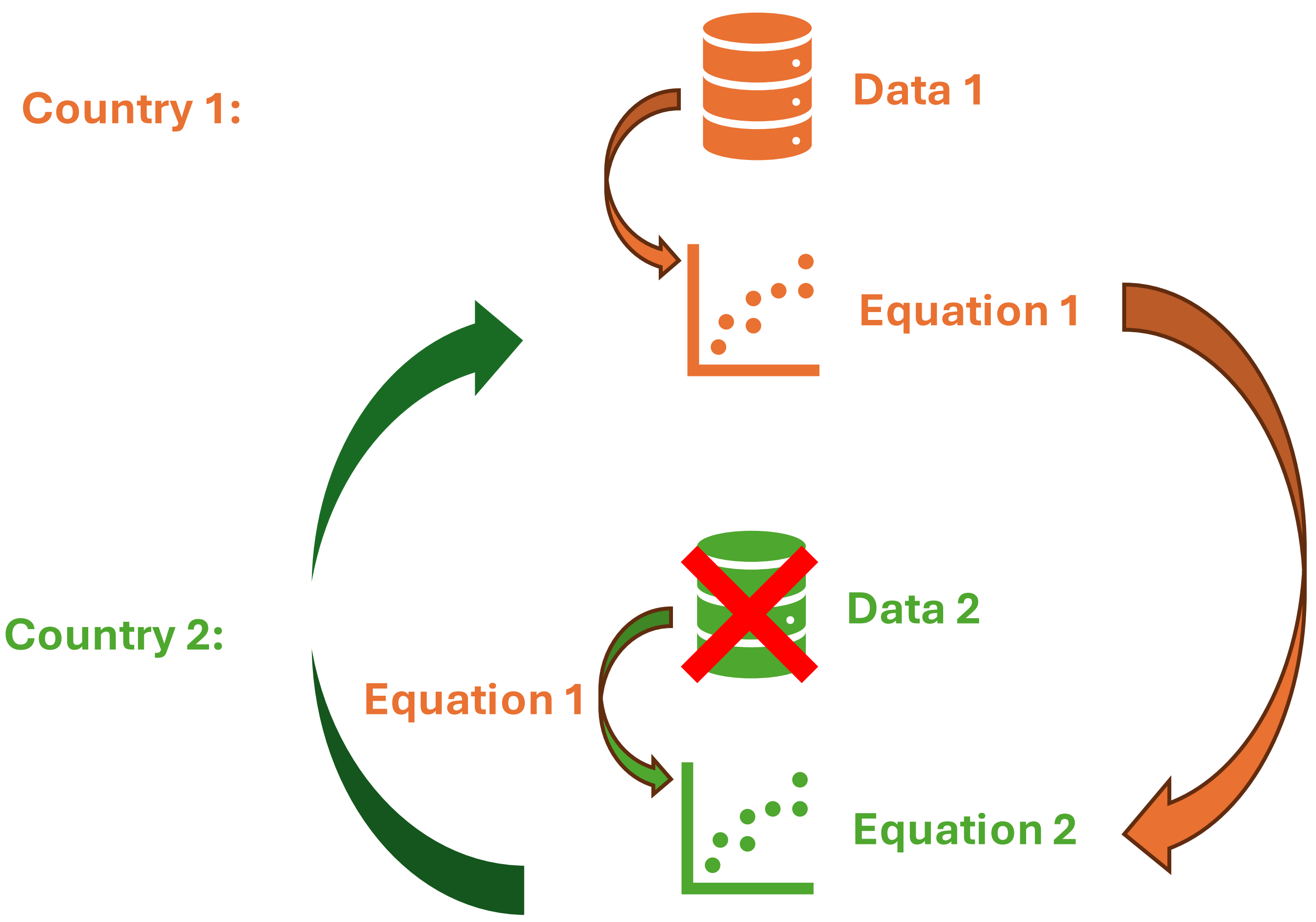
LIÈGE université Gembloux Agro-Bio Tech
Wallonie recherche CRA-W
teagasc Agriculture and Food Development Authority
ETH zürich
AARHUS UNIVERSITY
INRAE
fbn Research Institute for Farm Animal Biology

Why neural network ?

- Currently, Partial Least Squares (PLS) Regression
  - High correlation between some spectral points ➔ decrease the dimensionality
  - Known to be robust with low size dataset

- Why neural network ?
  - Considering potential non-linear relationships
  - Transfer learning can facilitate the improvement of models without exchanging data

Transfer learning



This is possible by using neural networks !

Model comparison

- **PLS**
  - **867 features** = 289 spectral points regressed until the second order Legendre polynomials
- Multi-layer neural perceptron (**ANN**)
  - **867 features**
- 1-D Convolutional neural network (**cANN**)
  - 1-D spectral image : **867 pixels**
- 2-D Convolutional neural network (**cANN2**)
  - 2-D spectral image : **289 * 3 pixels**, one row for each Legendre Polynomial

1	2	3	…	867

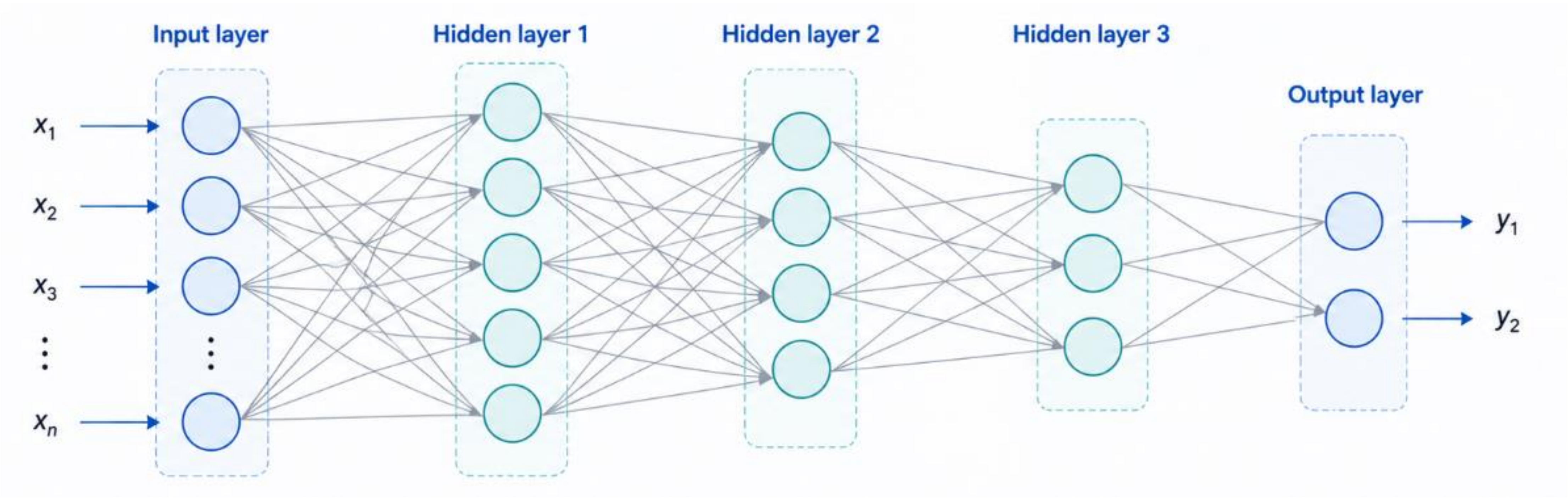
1	2	3	…	289	Raw spectrum
1	2	3	…	289	1st Legendre regressed spectrum
1	2	3	…	289	2nd Legendre regressed spectrum

Define the best structure for PLS

HYPERPARAMETER –
NUMBER OF PLS FACTORS

COW-INDEPENDENT **5-FOLDS** CROSS-VALIDATION

Multi-layer perceptron (ANN) - Theory



For us, only **one output** as it is a regression

Minimum hyperparameters:

- **Number of hidden layers**
- Number of **nodes** in each hidden layers

Define the best structure for ANN

- **Max. 10 hyperparameters** :
  - **Batch size** : 32 , 64 samples
  - **Learning rate**: 1e-4, 1e-2
  - **Weight decay** : 1e-8, 1e-2
  - **Number of hidden layers** : 1 to 5
  - **Number of nodes in each hidden layer** : 16 to 128 (step = 16)
  - **Drop-out importance** : 0 to 0.5 (step = 0.1)

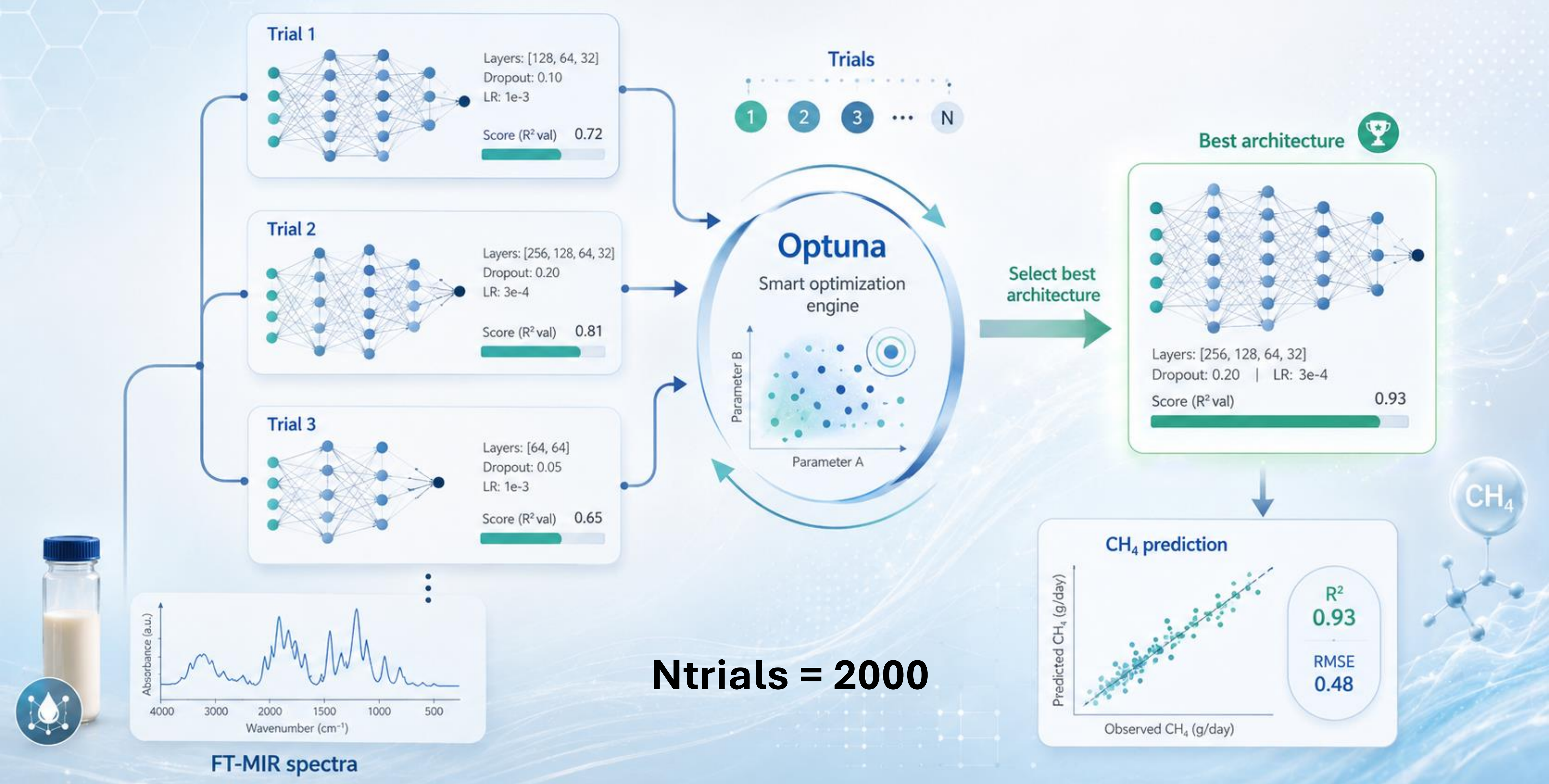
- A lot of possibilities to try :

$$N_{models} = n_1 * n_2 * \ldots * n_k$$

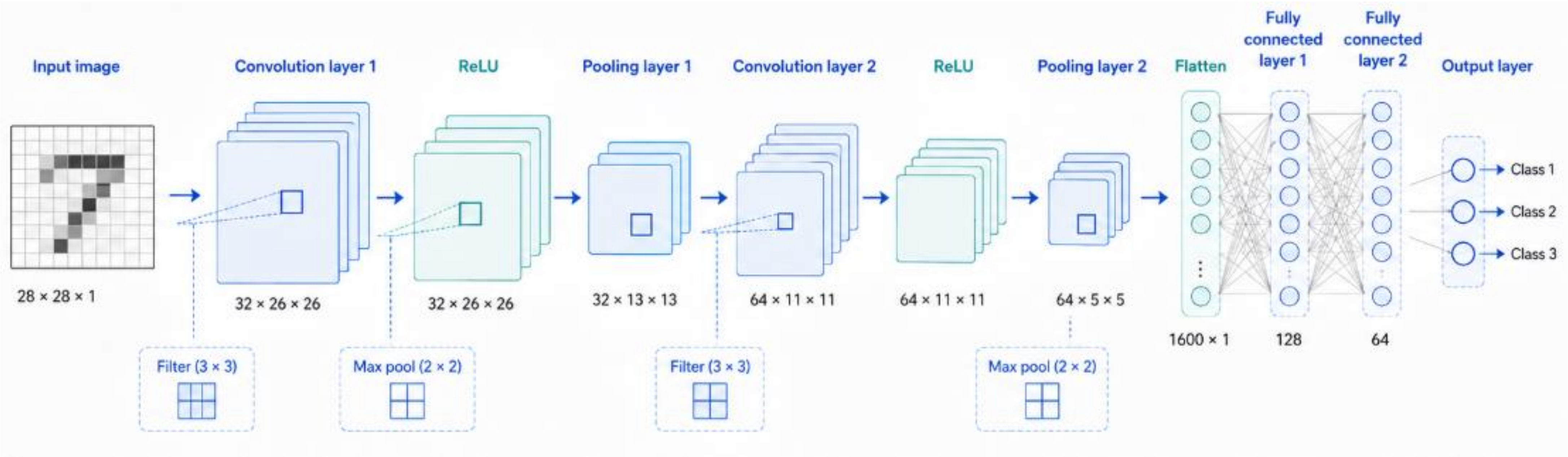
$n_1 = number\ of\ possibilities\ for\ hyperparameter\ 1$
$n_2 = number\ of\ possibilities\ for\ hyperparameter\ 1$
$n_k = number\ of\ possibilities\ for\ hyperparameter\ k$

373,160 models

Too many possibilities


Trial 1
Layers: [128, 64, 32]
Dropout: 0.10
LR: 1e-3
Score (R² val) 0.72
Trial 2
Layers: [256, 128, 64, 32]
Dropout: 0.20
LR: 3e-4
Score (R² val) 0.81
Trial 3
Layers: [64, 64]
Dropout: 0.05
LR: 1e-3
Score (R² val) 0.65
Trials
1 2 3 ... N
Optuna
Smart optimization engine
Parameter B
Parameter A
Select best architecture
Best architecture
Layers: [256, 128, 64, 32]
Dropout: 0.20 | LR: 3e-4
Score (R² val) 0.93
CH₄ prediction
Predicted CH₄ (g/day)
Observed CH₄ (g/day)
R² 0.93
RMSE 0.48
CH₄
Absorbance (a.u.)
4000 3000 2000 1500 1000 500
Wavenumber (cm⁻¹)
FT-MIR spectra
Ntrials = 2000

Convolutional neural network



Minimum Hyperparameters:

- **Number of filters**
- Number of **nodes** in the hidden layer

For us, only **one output** as it is a regression. We tried only **one fully connected hidden** layer.

Define the structure for **cANN**

Hyperparameters :

- **Batch size** : 8, 16, 32, 64 samples
- **Learning rate**: 1e-4, 1e-2
- **Number of filters in the conv layer**: 16, 32, 64
- **Drop-out rate (ndrop)**: 0.3, 0.5, 0.75
- **Number of neurons in the hidden later (nout)** : 10, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550

- Number of possibilities : 864 models
- Optuna selection – Ntrials : 200 models

Hyperparameters :

- **Batch size** : 8, 16, 32, 64 samples
- **Learning rate**: 1e-4, 1e-2
- **Number of filters in the conv layer**: 16, 32, 64
- **Drop-out rate (ndrop):** 0.3, 0.5, 0.75
- **Number of neurons in the hidden later (nout)** : 10, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550
- **Weight decay** : 1e-8,1e-2

Define the structure for **cANN2**

- Number of possibilities : 1,728 models
- Optuna selection – Ntrials : 200 models

Dataset used in a previous study

- N=**1,089 records**
- Respiration chambers and SF_6
- 7 countries
- Different breeds
- Different diets

Journal of the
Science of Food and
Agriculture

SCI
where science
meets business

Research Article

Improving robustness and accuracy of predicted daily methane emissions of dairy cows using milk mid-infrared spectra

Amélie Vanlierde, Frédéric Dehareng, Nicolas Gengler, Eric Froidmont, Sinead McParland, Michael Kreuzer, Matthew Bell, Peter Lund, Cécile Martin, Björn Kuhla, Hélène Soyeurt

First published: 22 November 2020 | https://doi.org/10.1002/jsfa.10969 | VIEW METRICS

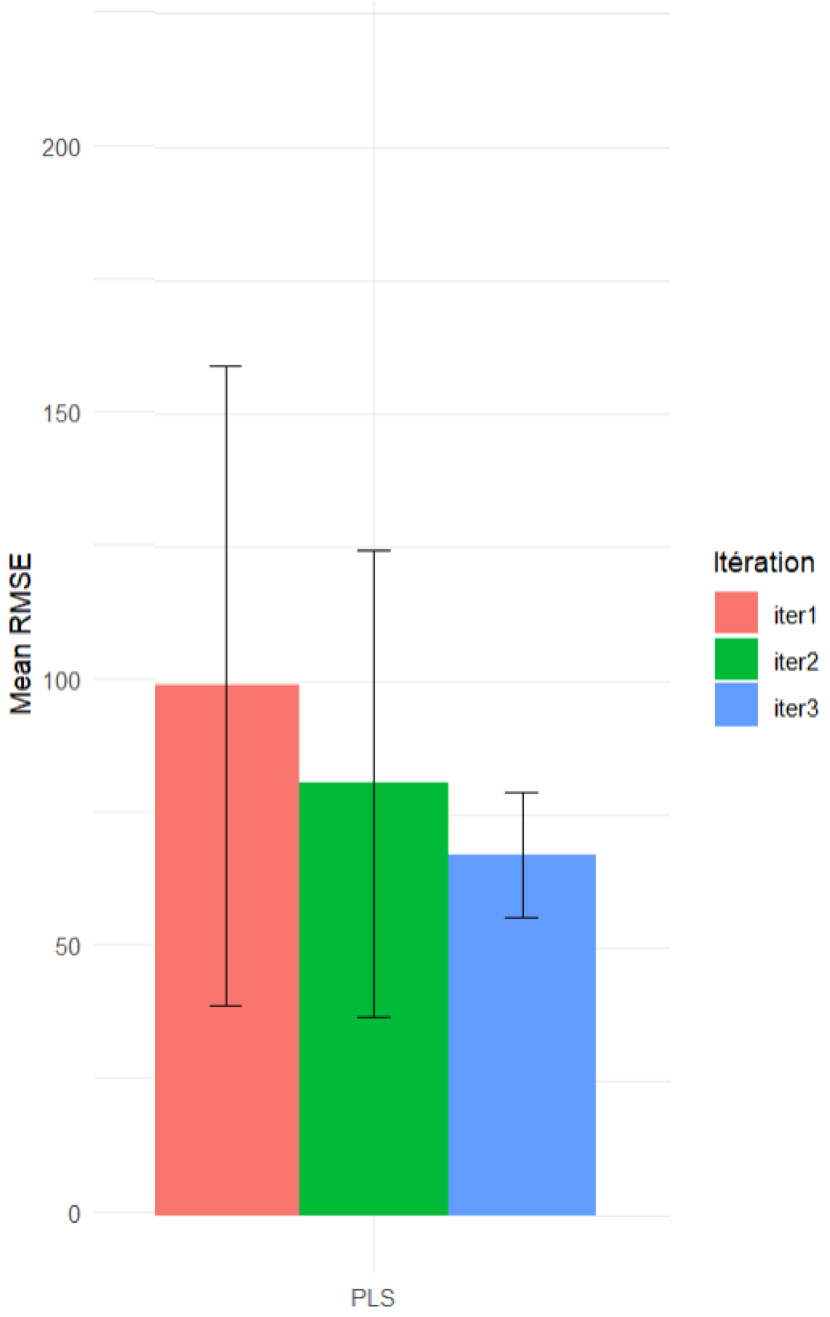
Data splitting

- Separation of dataset per country
  - 7 datasets

- Split between training and test sets for each dataset
  - 20 % for test set
  - 80 % for training set

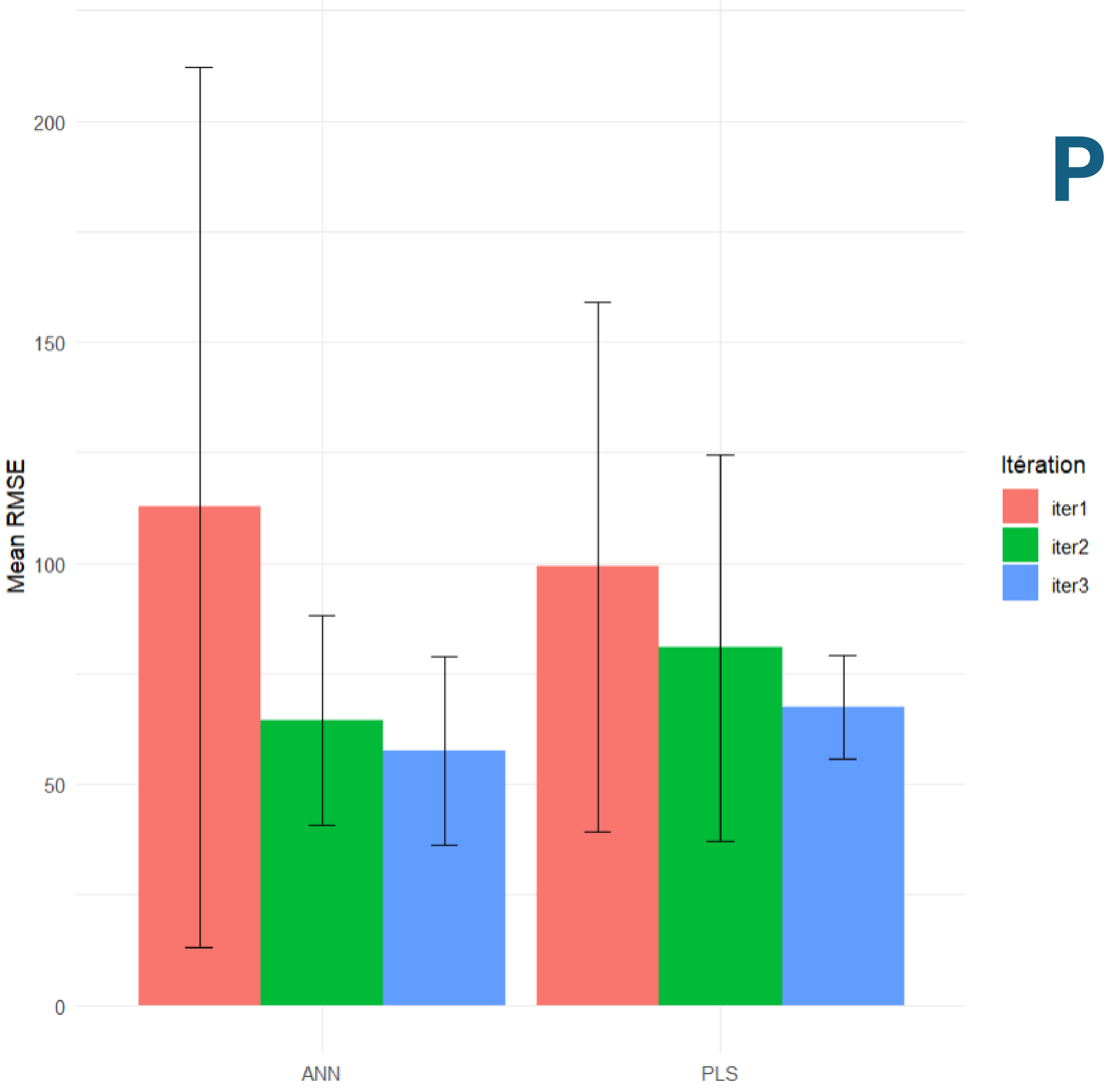
- Hyper-parameters optimization
  - Cow-independent 5-folds cross-validation

- See the impact of increasing the datasets
  - 3 iterations : [countries 1 & 2], [iter1 + countries 3,4 & 5], [iter2 + countries 6 & 7]



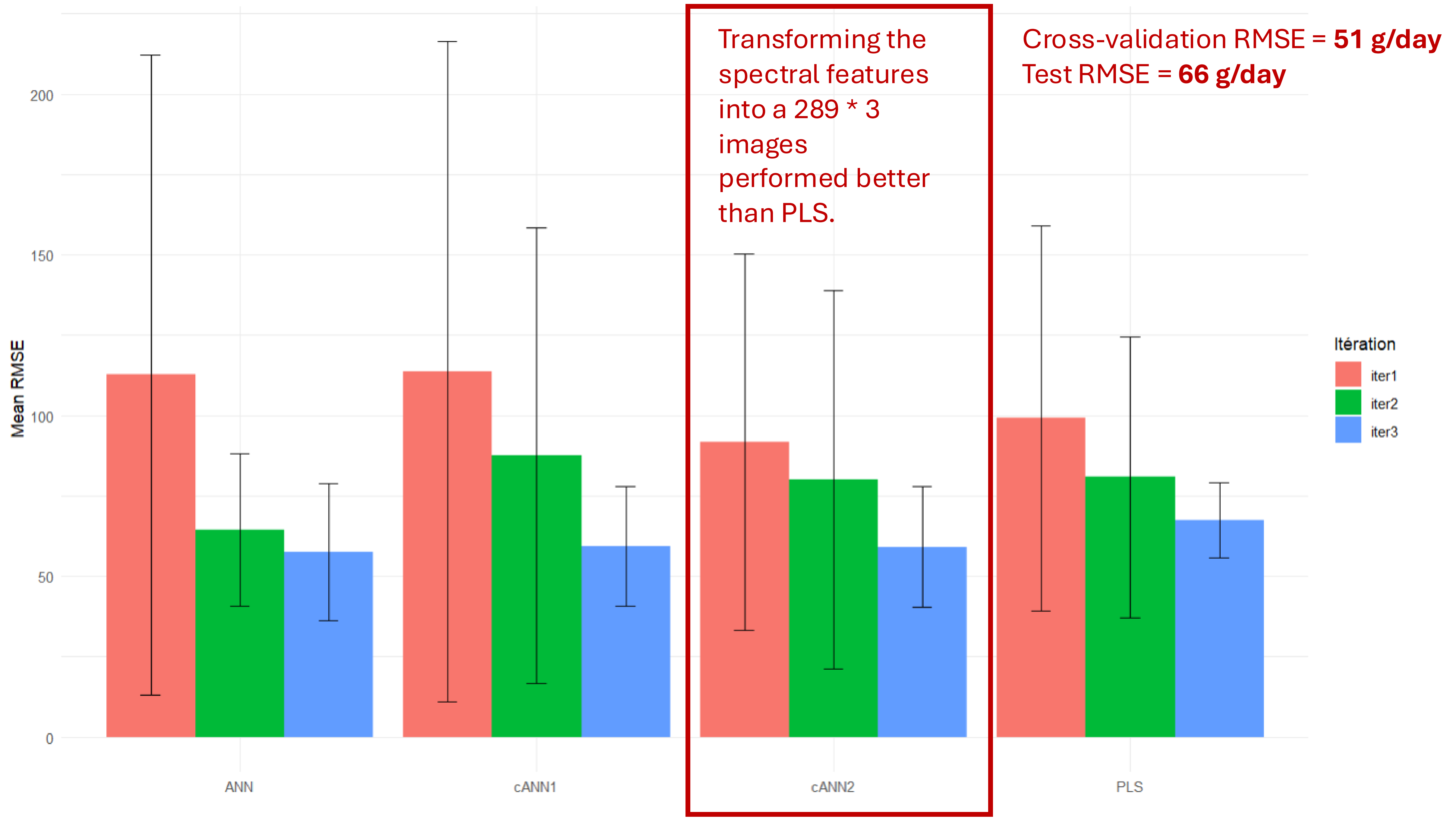
PLS model

- Different training set:
  - Iteration 1 : 410 samples
  - Iteration 2 : 611 samples
  - Iteration 3 : 890 samples
- Same dataset for each iteration (N= 199)
- Test RMSE ↓ when the training dataset increases as well as the SD



PLS vs. ANN model

- Same dataset for each iteration (N= 199)
- Test RMSE ↓ when the training dataset increases as well as the SD
- The performance is better when the training dataset is sufficient.


Transforming the spectral features into a 289 * 3 images performed better than PLS.
Cross-validation RMSE = 51 g/day
Test RMSE = 66 g/day
Mean RMSE
200
150
100
50
0
ANN
cANN1
cANN2
PLS
Itération
iter1
iter2
iter3

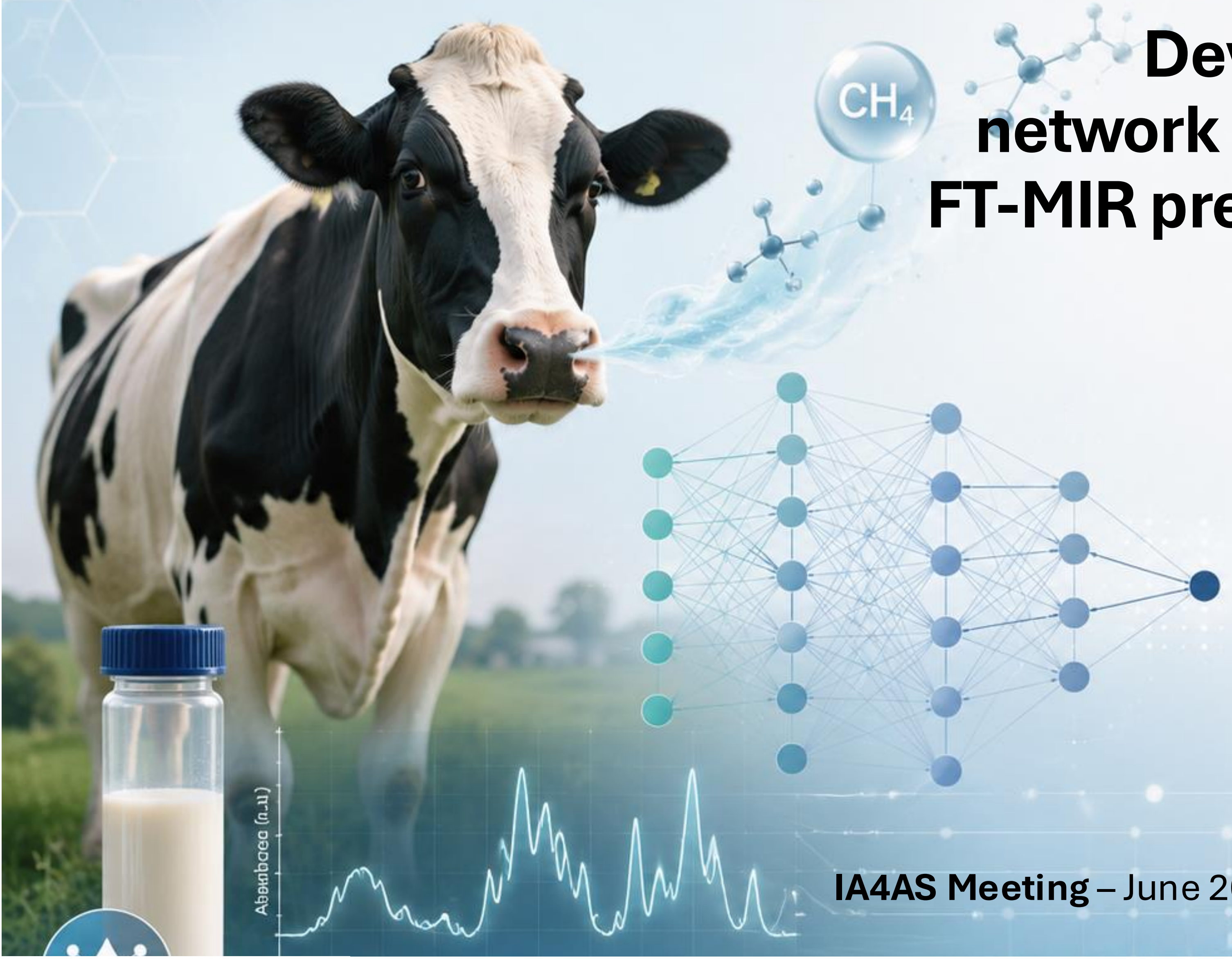
Conclusions

2D-convolutional neural network (cANN2) can improve the performances of PLS

- Cow-independent 5-folds cross-validation for cANN2 gives a prediction error of **51 g of CH_4/day**.
- The test error for cANN2 was of **66 g of CH_4/day**.
- The test RMSE SD was greater than PLS ➔ check the robustness
- When the structure of cANN2 will be fixed, transfer learning approaches could be tested.

Development of robust neural network architectures for federated FT-MIR prediction of enteric methane emissions in dairy cows

H. Soyeurt, F. Dehareng, N. Gengler, S. McParland, M. Kreuzer, P. Lund, C. Martin, B. Kuhla, A. Vanlierde



IA4AS Meeting – June 2026 - Ghent

LIÈGE université Gembloux Agro-Bio Tech
Wallonie recherche CRA-W
teagasc Agriculture and Food Development Authority
ETH zürich
AARHUS UNIVERSITY
INRAE
fbn Research Institute for Farm Animal Biology