



European
Federation
of Animal
Science



LIÈGE université
Gembloux
Agro-Bio Tech

Sébastien Franceschini
Modeling and development
sfranceschini@uliege.be



Reconstructing representative daily milk FT-MIR spectra from partial milking robot samples using machine learning

J. Paquet, C. Sneessens, **S. Franceschini**, P.-E. Jacoby, E. Reding, J. Leblois, H. Soyeurt

EAAP AI4AS conference 2026



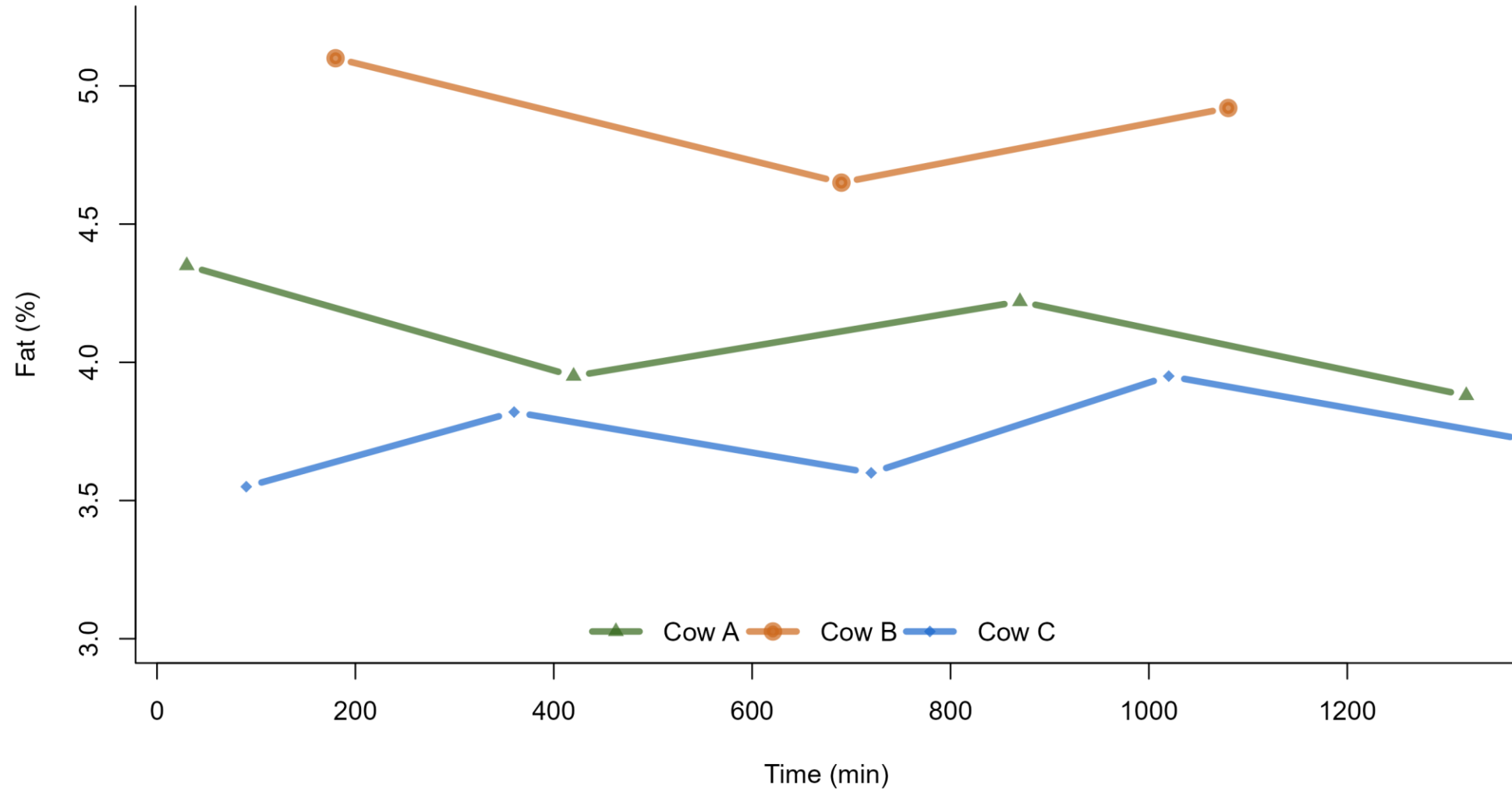
Automated Milking System (AMS)



- Increasingly common in modern dairy farming.
- One out of four dairy farms in Denmark and Sweden
- Continuous collection of data.

(Cogato et al., 2021; Hassoun et al., 2023; Zagidullin et al., 2023)

Evolution of milk fat over a day



Correction for Fat



THE GLOBAL STANDARD
FOR LIVESTOCK DATA

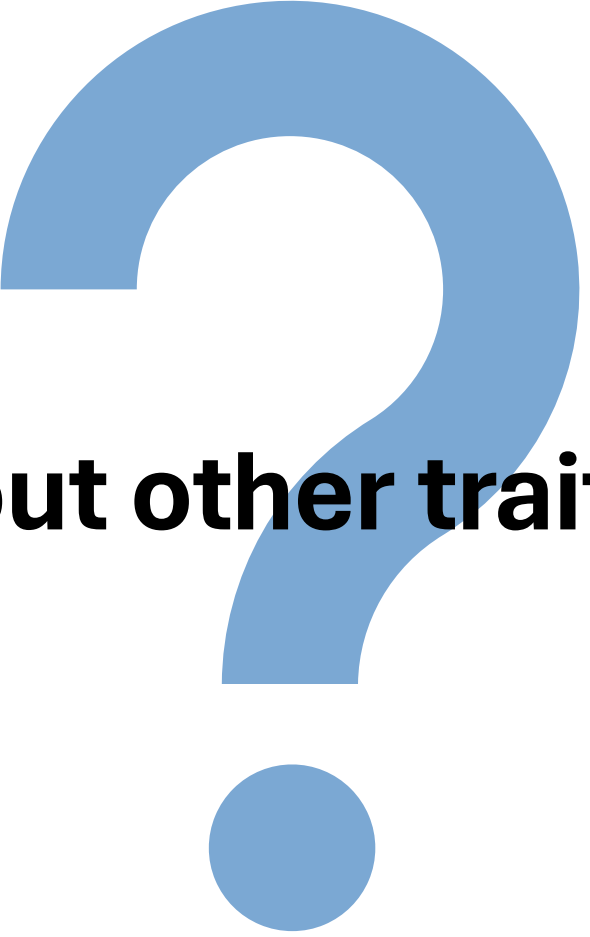
Procedure 1 of Section 2 of ICAR Guidelines - Computing 24-hour Yields

Methods to calculate 24-hour yields in Automatic Milking Systems

Computing 24-hour Yields

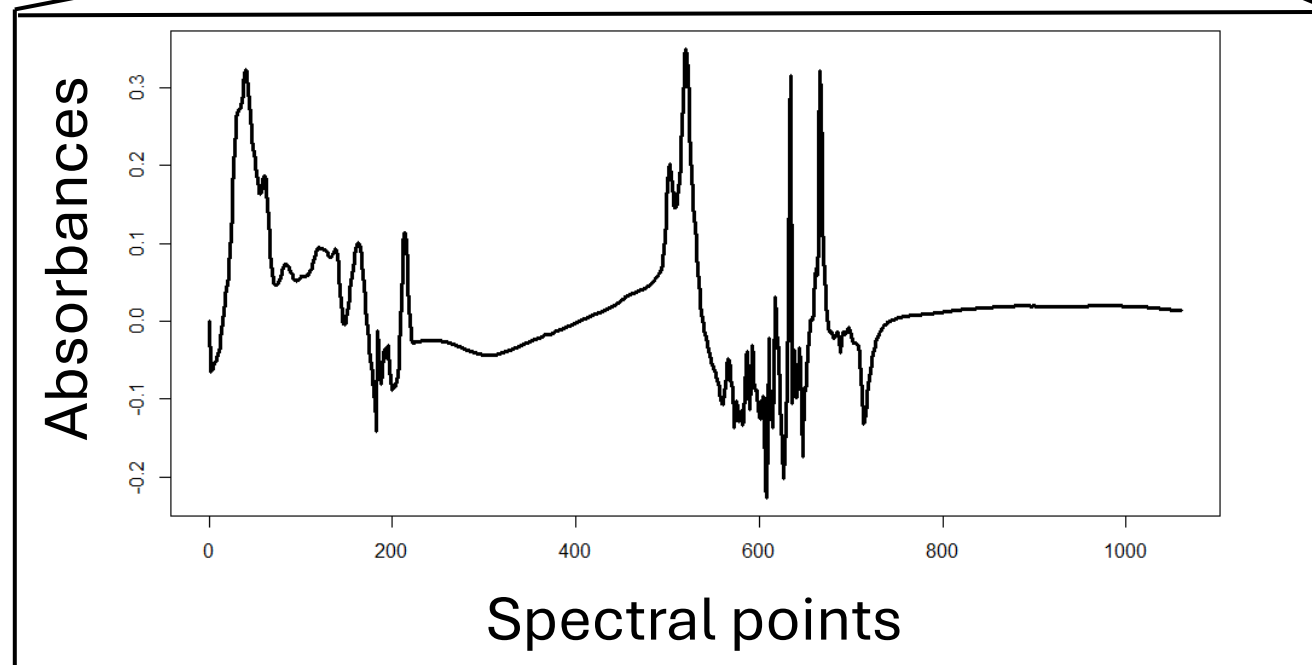
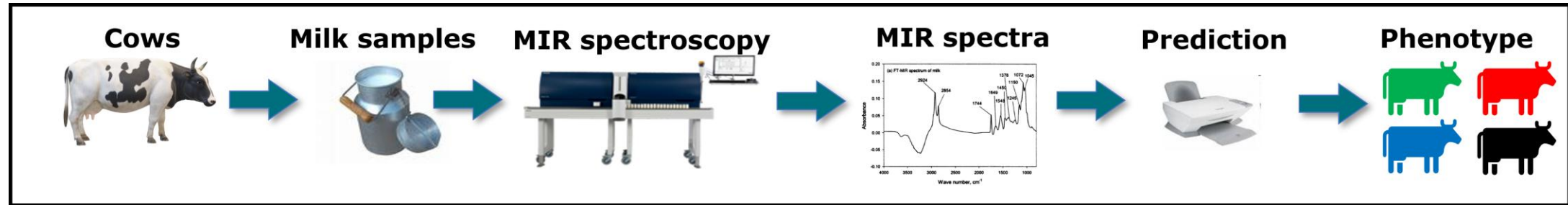
Version June, 2023

=> Peeters and Galesloot (2002)

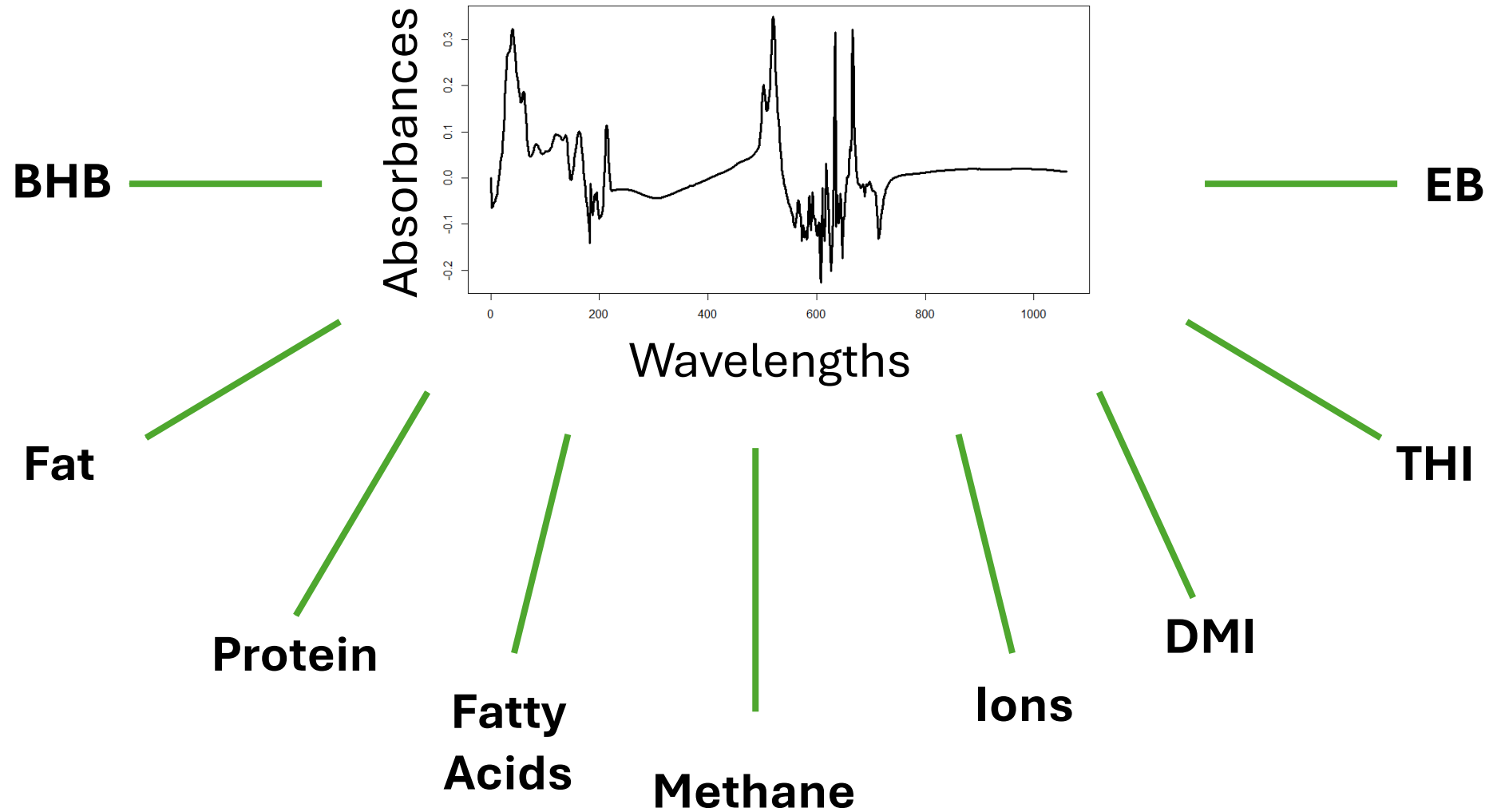


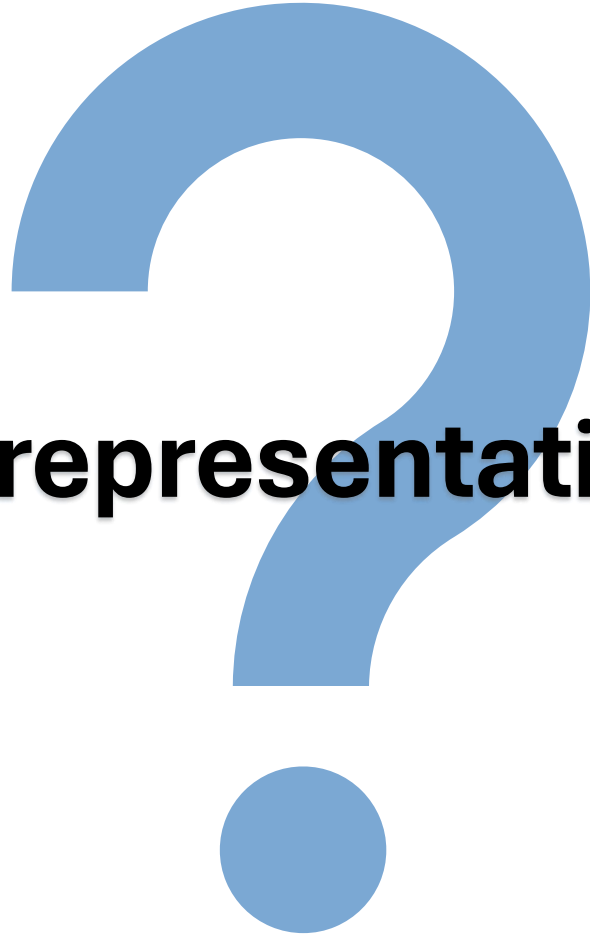
What about other traits than Fat?

What is Fourier Transform mid-infrared spectrometry ?



Predictions? Phenotypes ?





Can we create a representative daily spectrum?

Datasets

Data used to train the methods or find coefficients

Data Train: Fat and Prot

- 744 farms
- 2,748 Milk Recording Events
- Between 2009 and 2026
- 30,983 cows
- 140,257 milk samples

Data Train: Spectra

- 701 farms
- 1,985 Milk Recording Events
- Between 2014 and 2026
- 26,892 cows
- 102,464 milk samples

Data Validation

- 9 farms
- 9 Milk Recording Events
- Between 2025 and 2026
- 579 cows
- 4,866 milk samples

All samples were analyzed.

Methods for estimation of Fat 24h

1. PG = ICAR Certified: Peeters and Galesloot (2002)
2. WA = Weighted average
3. aNN = artificial neural network

Methods for estimation of Fat 24h

1. PG = Peeters and Galesloot (2002) with only one milking:

$$\%FAT_{24h} = a + b \times \%FAT + c \times \%PROT + d \times MI + e \times MY + f \times pMY$$

 6 coefficients (a, b, c, d, e et f)

	Milking Interval (minutes)				
	Class	]0;361[	[361;511[	[511;701[	[701;1441[
$\frac{\%FAT}{\%PROT}$	<1.1	1	2	3	4
	[1.1;1.25[	5	6	7	8
	[1.25;1.4[	9	10	11	12
	>1.4	13	14	15	16

16 classes of 6 coefficients
(a, b, c, d, e and f)

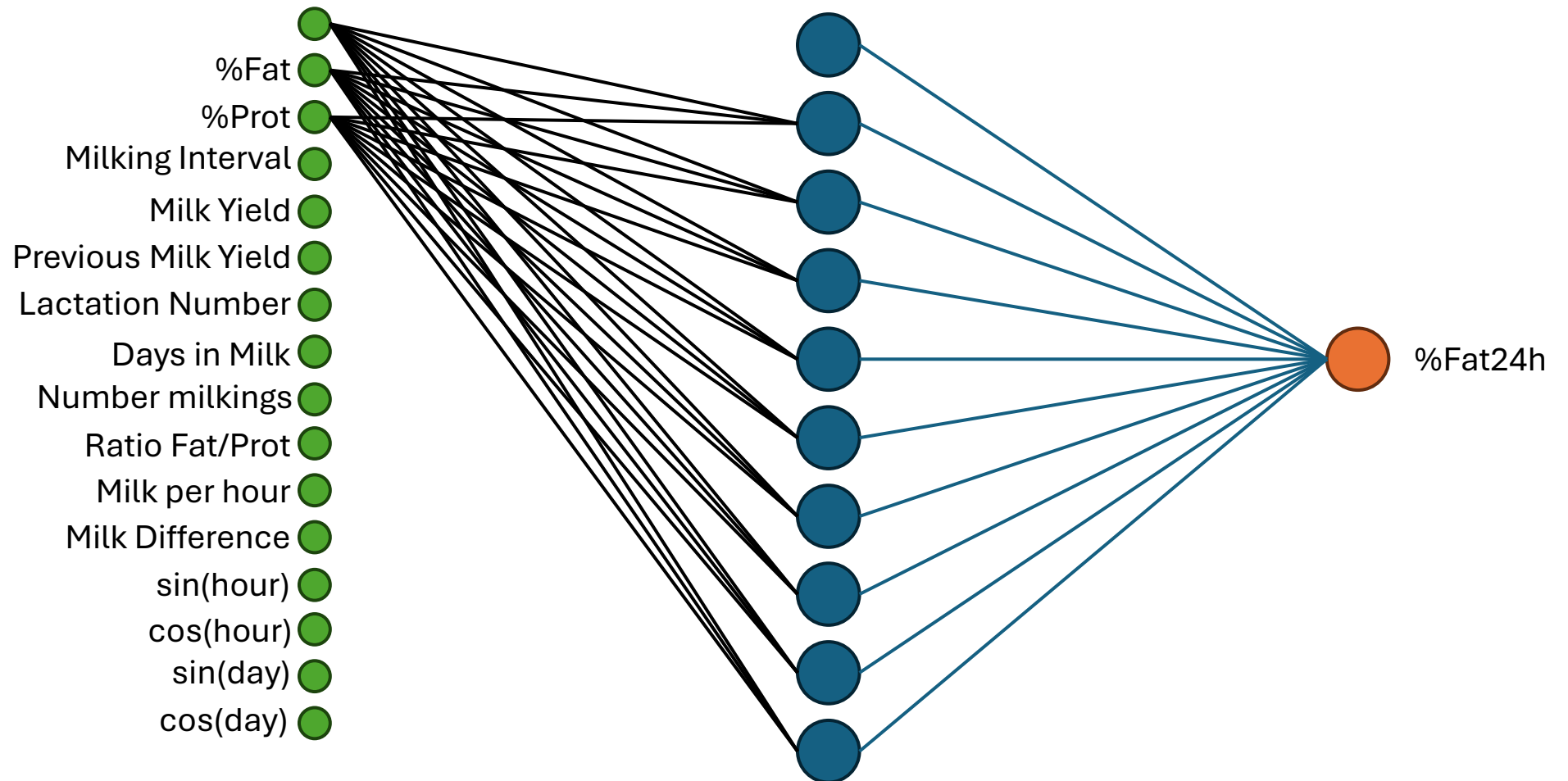
Methods for estimation of Fat 24h

2. **WA = Weighted average** with more than one milking:

$$\%FAT24h = \frac{\sum (\%FAT \times MY)}{\sum MY}$$

Methods for estimation of Fat 24h

3. aNN = artificial neural network with only one milking



RMSE of Fat

RMSE Fat24h (g/100g of milk)		
	one sample	two samples
No_Mod	0.35	/
PG	0.29	0.19*
WA	/	0.18
aNN	0.28	0.19*

- aNN is the best for one sample
- WA is the best for two samples**

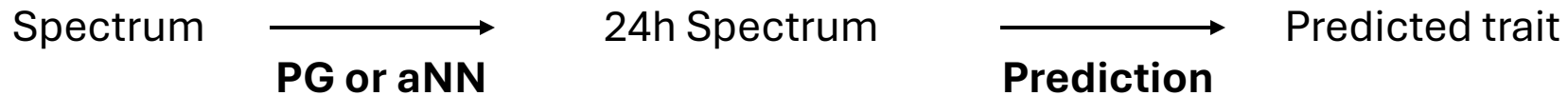
* Obtained by averaging the corrected values from two samples

** Particularly if we know more than 60% of total MY

Methods for estimation of Spectra 24h

Each wavelengths is processed independently:

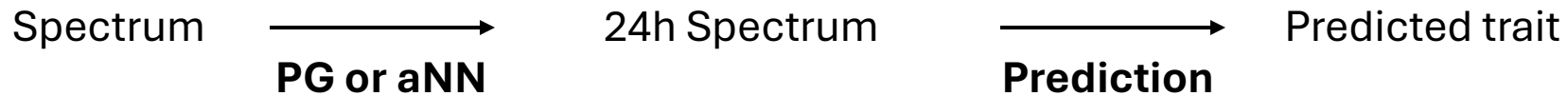
With one milk sample



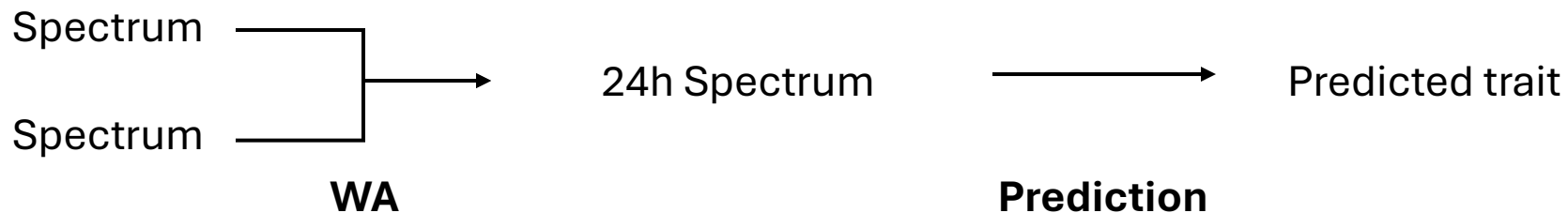
Methods for estimation of Spectra 24h

Each wavelengths is processed independently:

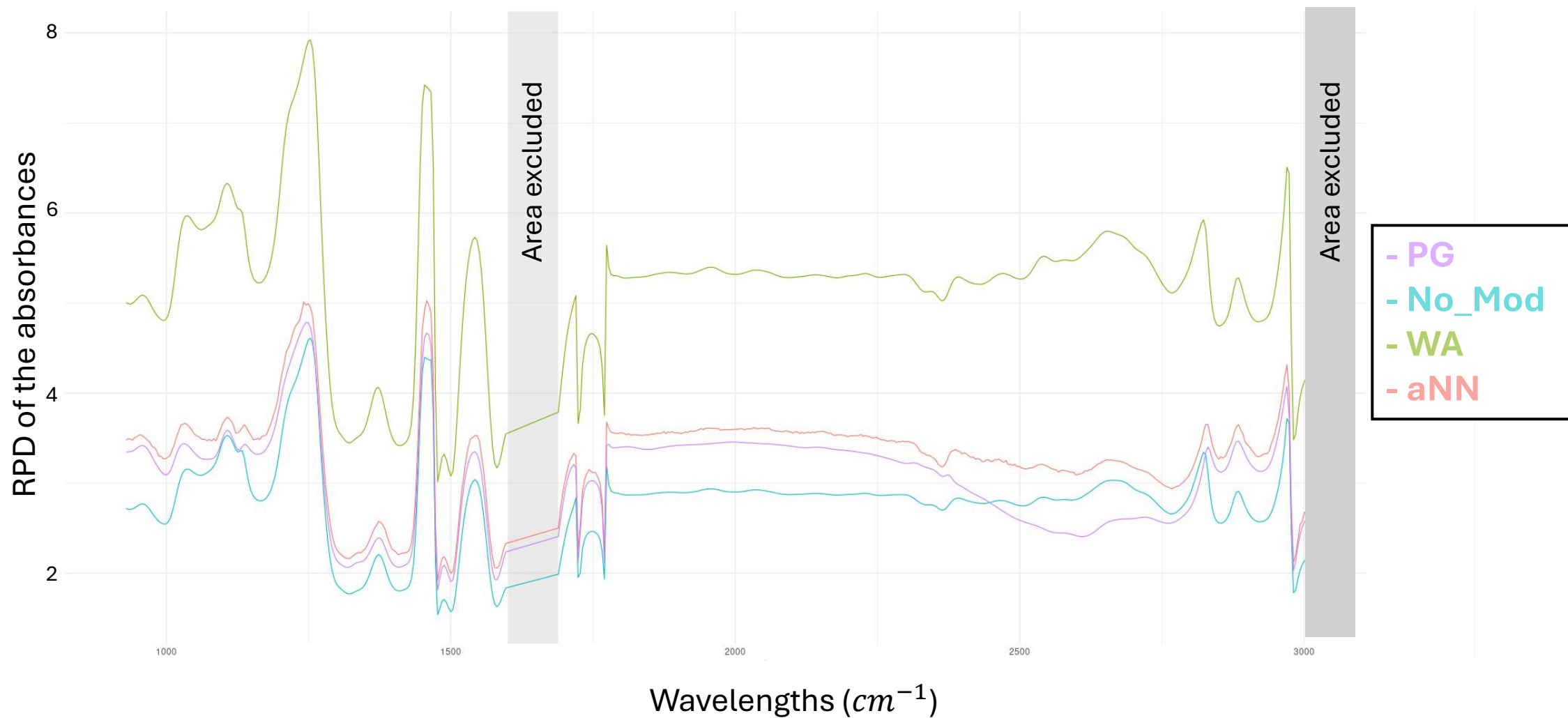
With one milk sample



With two milk samples



RPD of Spectra 24h



RMSE of Fat24h from Spectra24h

RMSE Fat24h (g/100g of milk)				
	Labo		Spectrum	
	one sample	two samples	one sample	two samples
No_Mod	0.35	/	0.35	/
PG	0.29	0.19*	0.30	/
WA	/	0.18	/	0.19
aNN	0.28	0.19*	0.29	/

RMSE of predictions from Spectra24h

Traits	Protein	Lactose	SFA	UFA	SCFA	MCFA	LCFA
Unit	g/100g	g/100g	g/100g	g/100g	g/100g	g/100g	g/100g
No_Mod	0.09	0.05	0.26	0.12	0.03	0.19	0.15
PG	0.09	0.06	0.22	0.10	0.03	0.17	0.13
WA*	0.06	0.03	0.13	0.04	0.02	0.10	0.08
aNN	0.09	0.05	0.25	0.13	0.03	0.25	0.18

* For 2 milk samples

Take Home Message

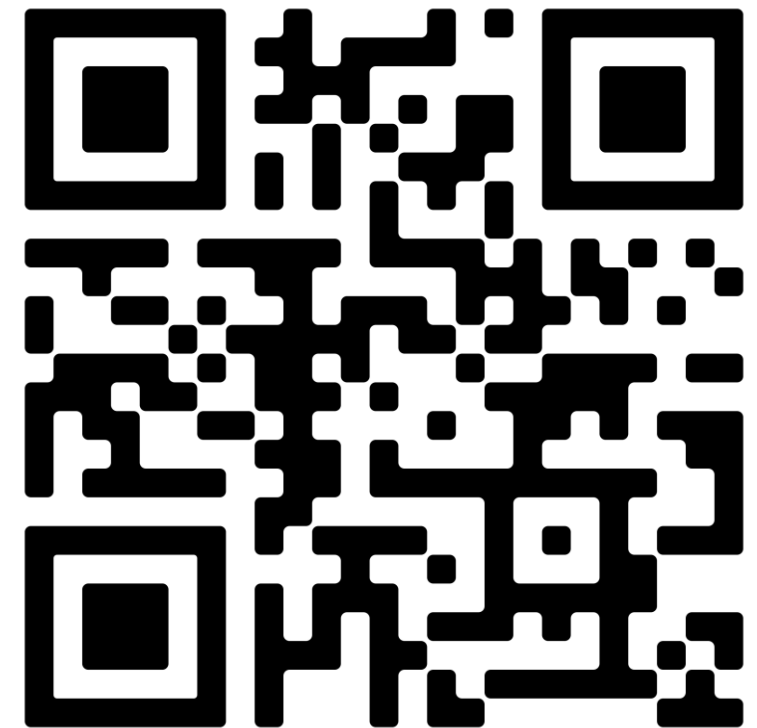
- 1. We can create a daily representative spectrum for AMS milk recording!**
- 2. Same RMSE for FAT24h corrected through PG or predicted from spectrum24h**
- 3. Weighted Average with two samples is always better**
- 4. aNN without 16 classes is promising to create spectra**
- 5. But aNN is underperforming for predictions!**

Thank you for attention



sfranceschini@uliege.be

Let us discuss!



Linked 

Methods for estimation of Fat and Prot 24h

3. LM = Linear regression

With one or more milking:

$$\%FAT24h = \frac{FAT_{known} + FAT_{pred}}{\sum MY}$$

Where $FAT_{pred} = MY + MY^2 + \%PROT + MI + pMI$

With MY = Milk Yield, MI = Milking Interval and pMI = previous Milking interval

RPD of Spectra 24h

- Lipids Area
- Proteins Area
- Carbohydrates Area

