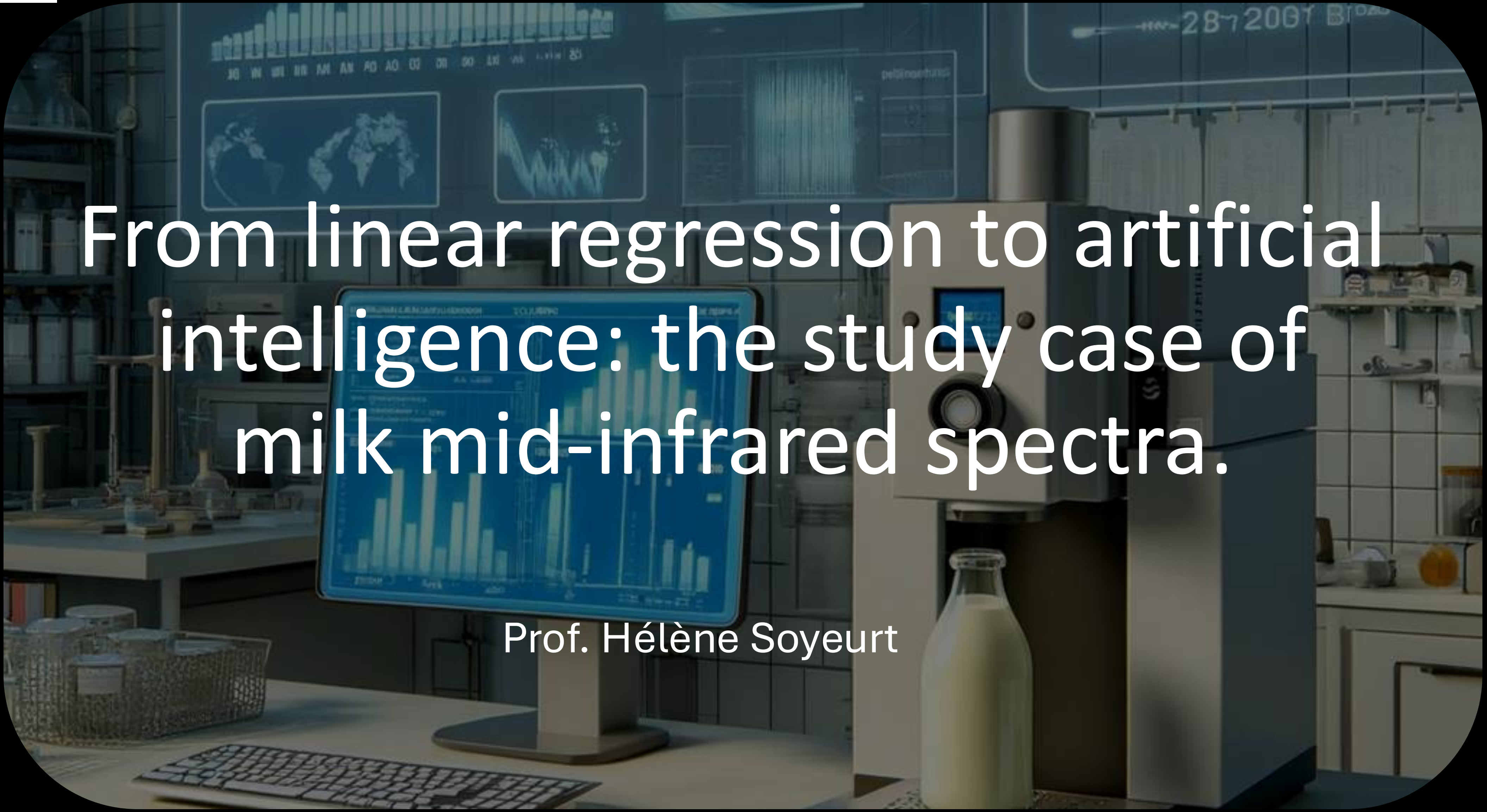
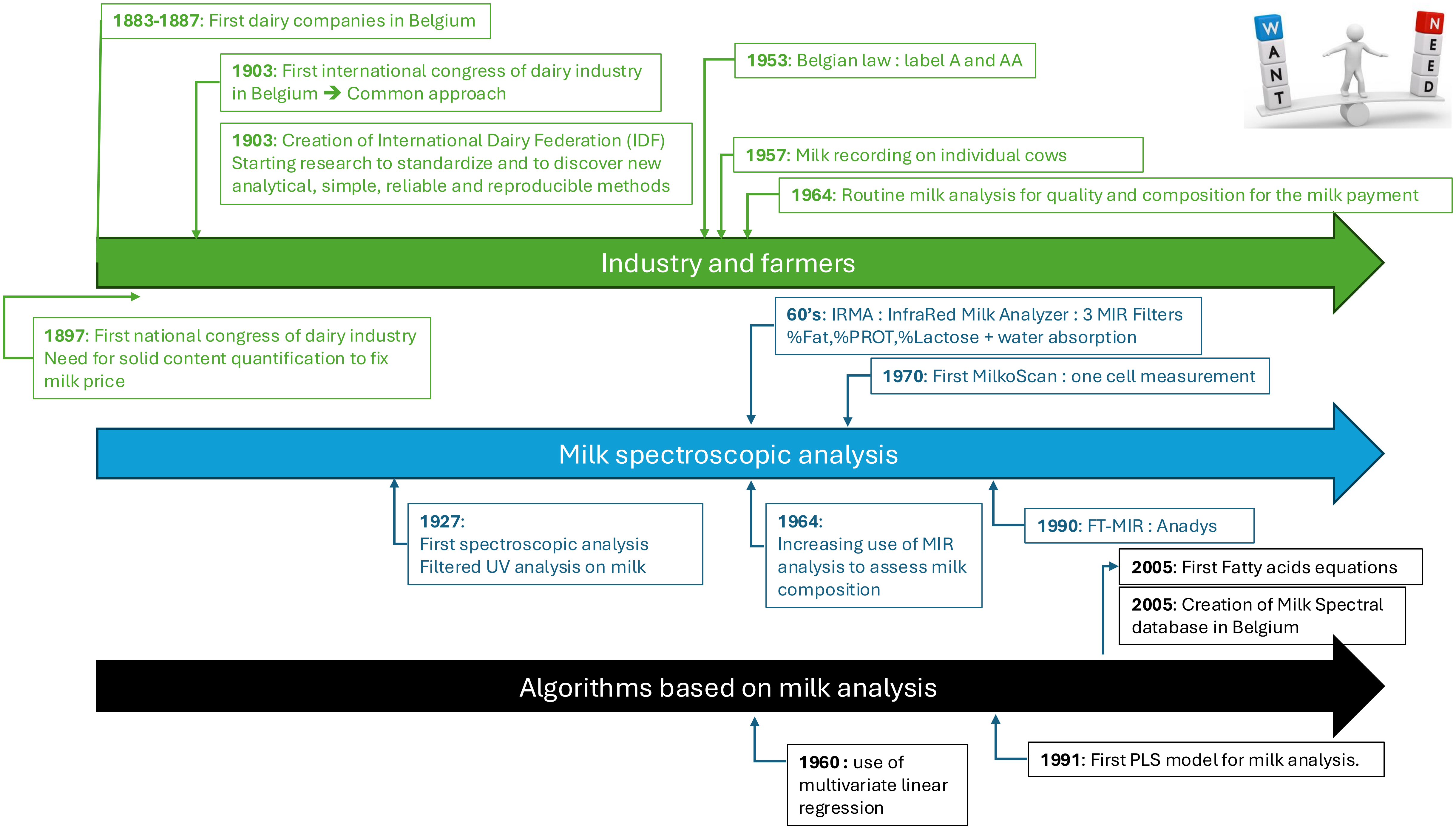
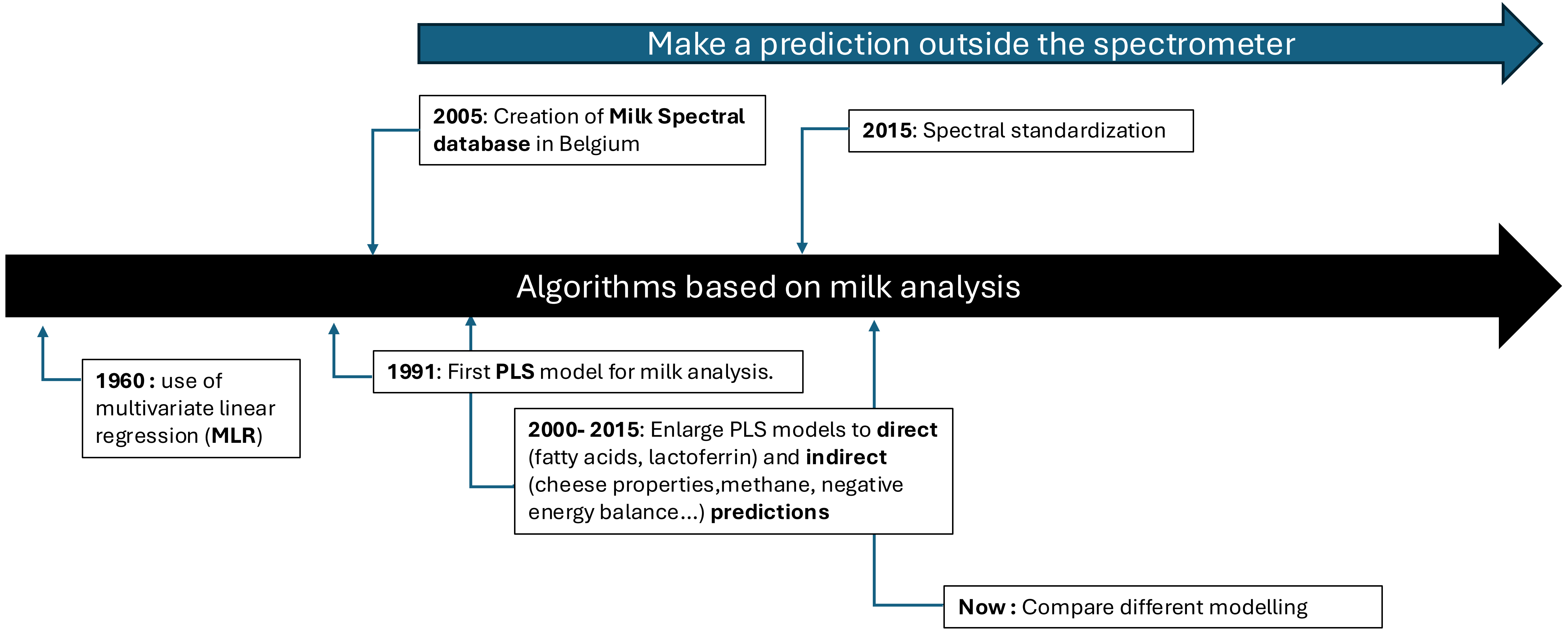


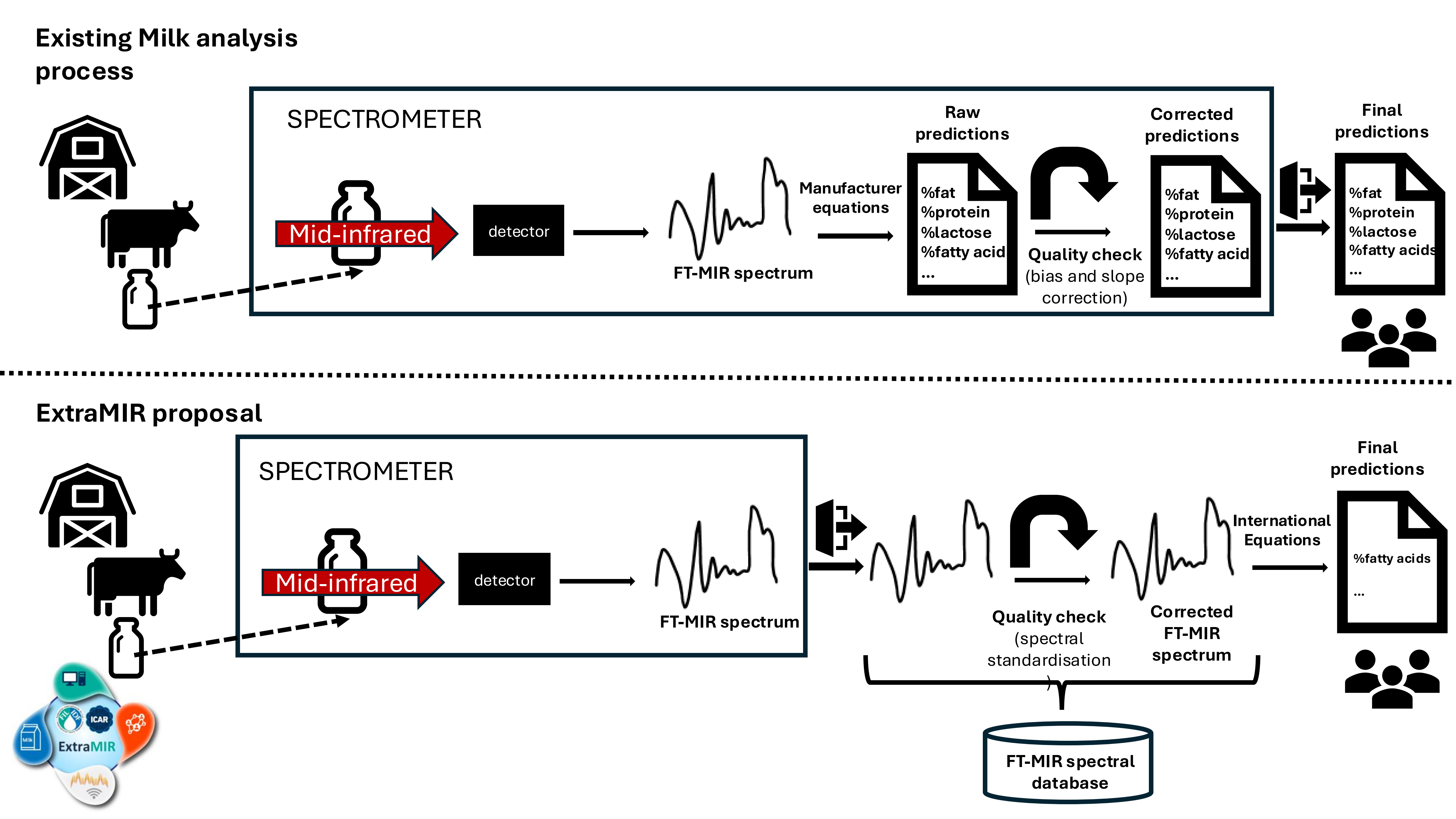
From linear regression to artificial intelligence: the study case of milk mid-infrared spectra.

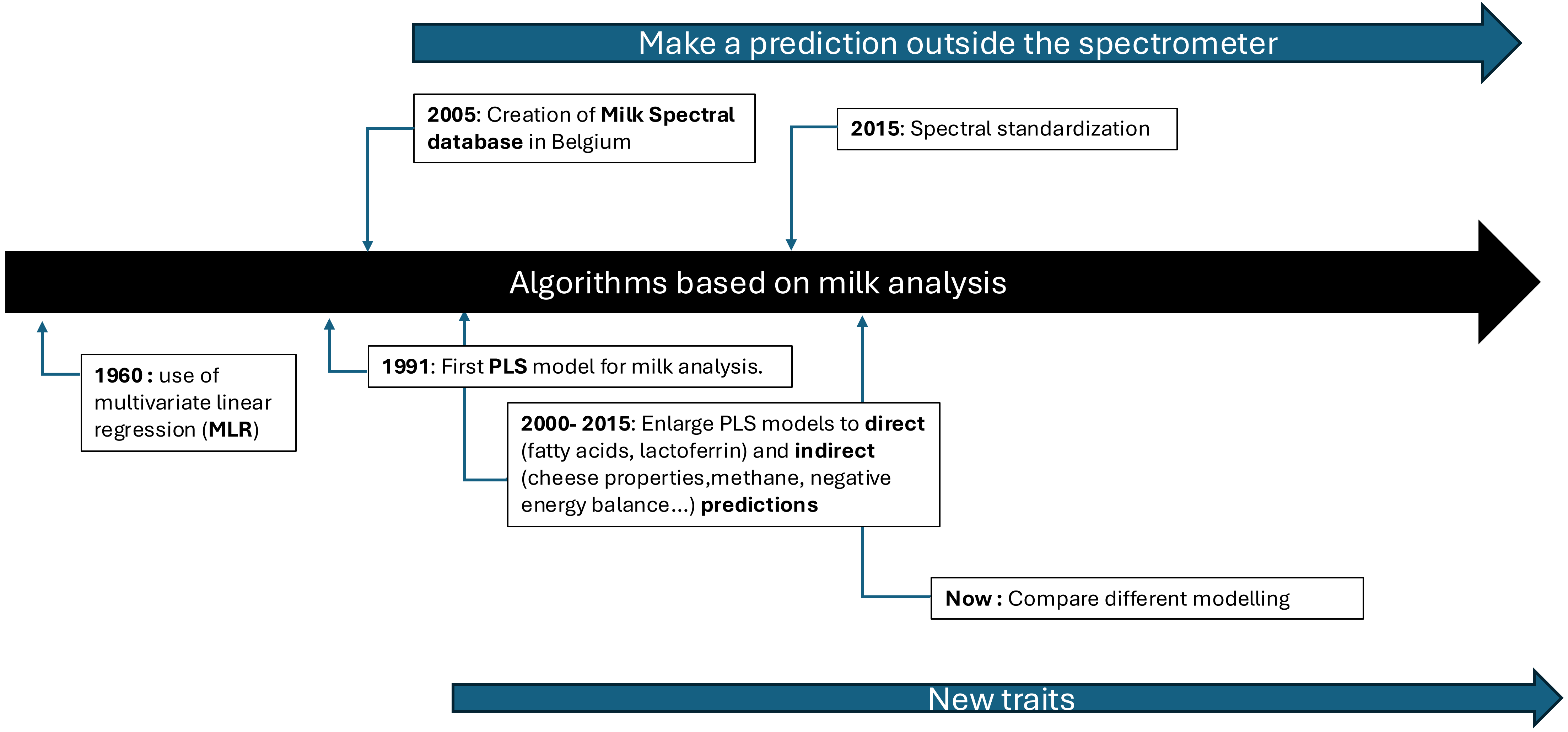
Prof. Hélène Soyeurt

46th Discovery Conference – Itasca – 6-9th May 2024


1883-1887: First dairy companies in Belgium
1903: First international congress of dairy industry in Belgium ➔ Common approach
1903: Creation of International Dairy Federation (IDF) Starting research to standardize and to discover new analytical, simple, reliable and reproducible methods
1953: Belgian law : label A and AA
1957: Milk recording on individual cows
1964: Routine milk analysis for quality and composition for the milk payment
WANT
NEED
Industry and farmers
1897: First national congress of dairy industry Need for solid content quantification to fix milk price
60's: IRMA : InfraRed Milk Analyzer : 3 MIR Filters %Fat,%PROT,%Lactose + water absorption
1970: First MilkoScan : one cell measurement
Milk spectroscopic analysis
1927: First spectroscopic analysis Filtered UV analysis on milk
1964: Increasing use of MIR analysis to assess milk composition
1990: FT-MIR : Anadys
2005: First Fatty acids equations
2005: Creation of Milk Spectral database in Belgium
Algorithms based on milk analysis
1960 : use of multivariate linear regression
1991: First PLS model for milk analysis.


Make a prediction outside the spectrometer
2005: Creation of Milk Spectral database in Belgium
2015: Spectral standardization
Algorithms based on milk analysis
1960 : use of multivariate linear regression (MLR)
1991: First PLS model for milk analysis.
2000- 2015: Enlarge PLS models to direct (fatty acids, lactoferrin) and indirect (cheese properties,methane, negative energy balance...) predictions
Now : Compare different modelling


Existing Milk analysis process
SPECTROMETER
Mid-infrared
detector
FT-MIR spectrum
Manufacturer equations
Raw predictions
%fat
%protein
%lactose
%fatty acid
...
Quality check
(bias and slope correction)
Corrected predictions
%fat
%protein
%lactose
%fatty acid
...
Final predictions
%fat
%protein
%lactose
%fatty acids
...
ExtraMIR proposal
SPECTROMETER
Mid-infrared
detector
FT-MIR spectrum
Quality check
(spectral standardisation)
Corrected FT-MIR spectrum
International Equations
Final predictions
%fatty acids
...
FT-MIR spectral database
ICAR
ExtraMIR


Make a prediction outside the spectrometer
2005: Creation of Milk Spectral database in Belgium
2015: Spectral standardization
Algorithms based on milk analysis
1960 : use of multivariate linear regression (MLR)
1991: First PLS model for milk analysis.
2000- 2015: Enlarge PLS models to direct (fatty acids, lactoferrin) and indirect (cheese properties,methane, negative energy balance…) predictions
Now : Compare different modelling
New traits

Open other possibilities

Nutritional quality: Fat, protein, lactose, fatty acids, Ca, lactoferrin,...

Technological properties: Cheese yield, yoghurt yield, butter yield, spreadability ...

Animal Health: Na, lactoferrin, Energy balance, body weight, dry matter intake, acetone, BHB, citrate ...

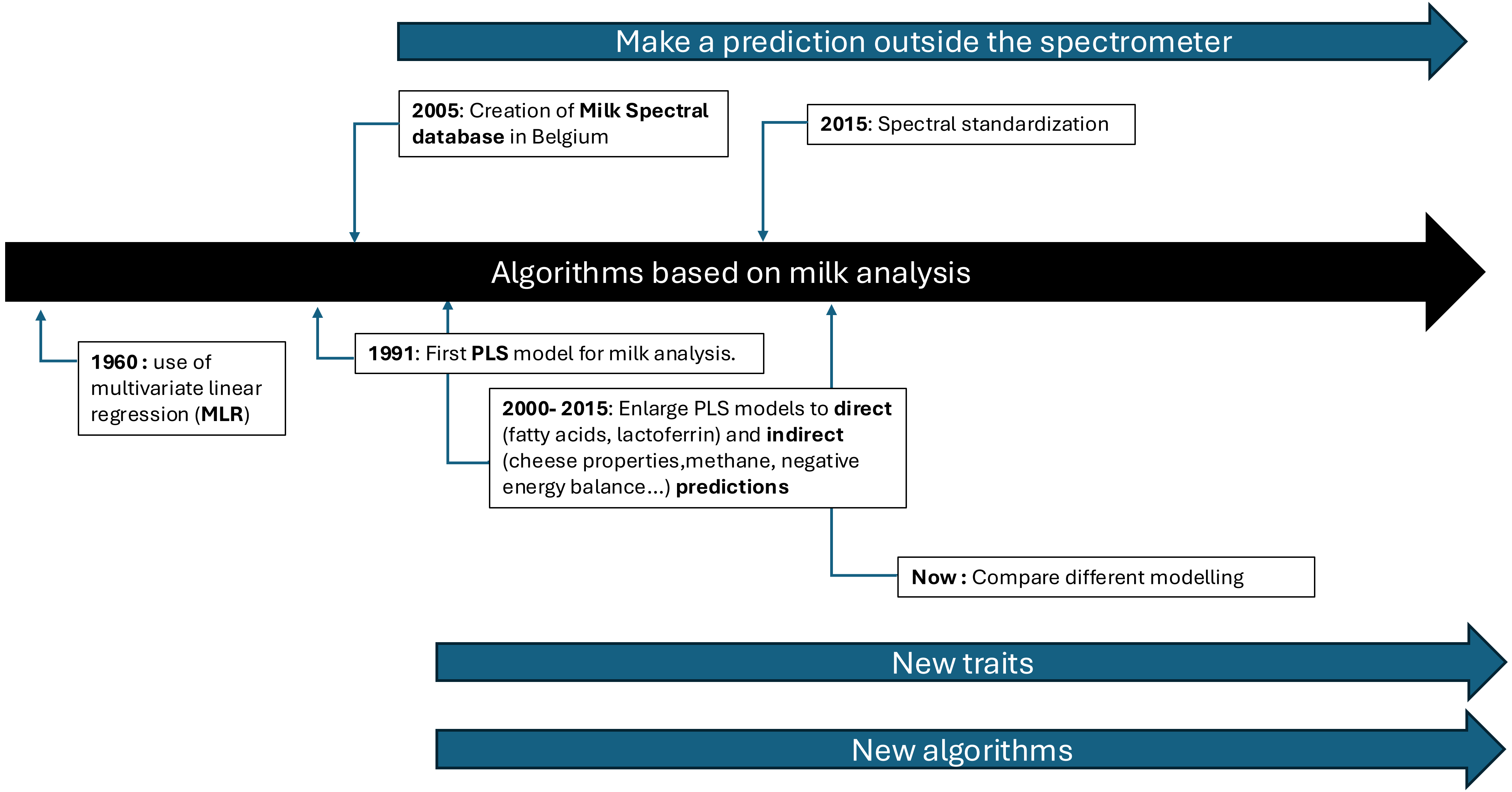
Environmental fingerprint: Methane, P, urea ...

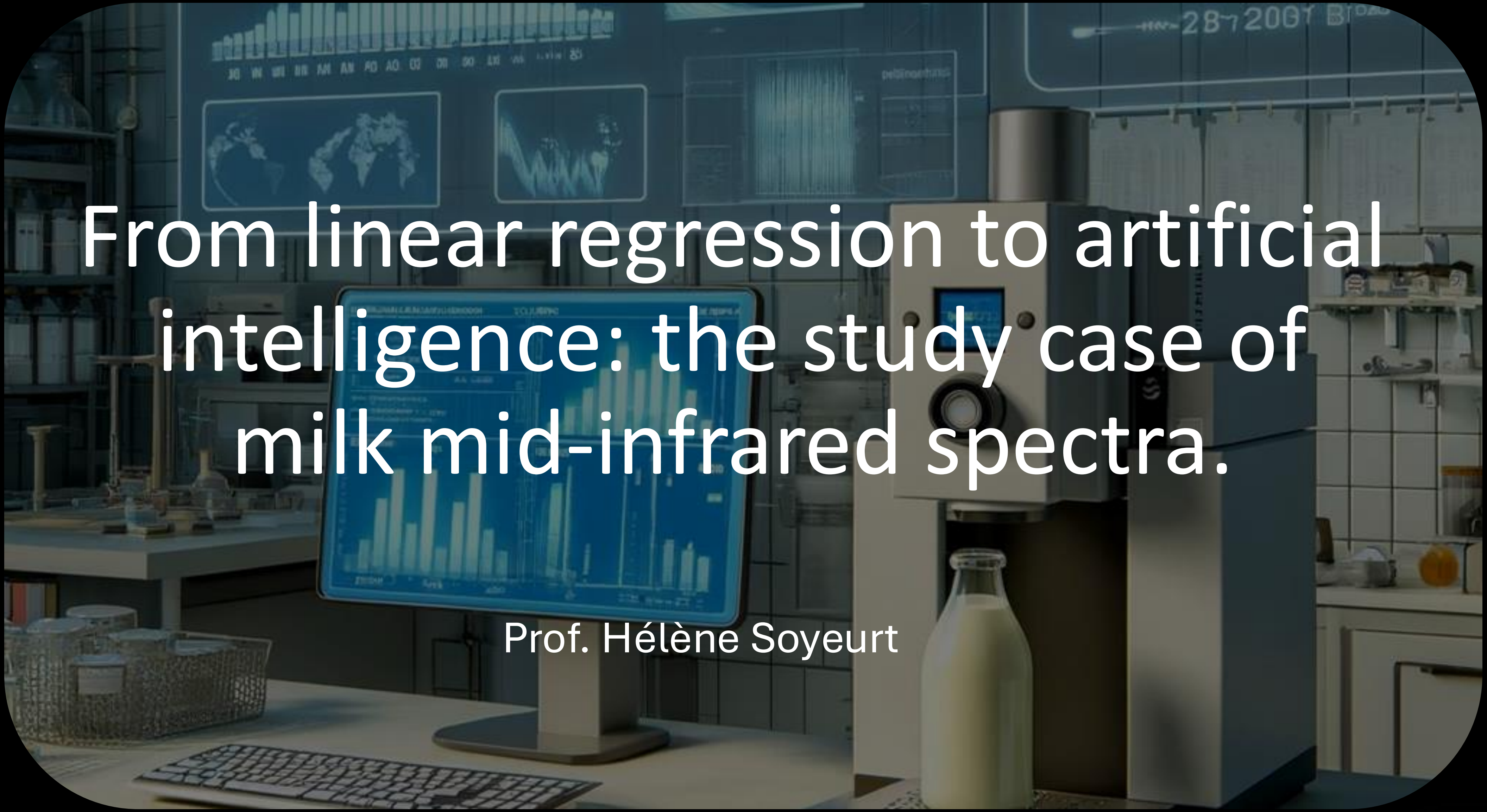
Sustainability: Consumption index, nitrogen efficiency ...

Abnormal milk samples, color ...

Soyeurt, 2023

Future Outlines





From linear regression to artificial intelligence: the study case of milk mid-infrared spectra.

Prof. Hélène Soyeurt

LIÈGE université Gembloux Agro-Bio Tech

60's: IRMA : InfraRed Milk Analyzer : 3 MIR Filters
%Fat,%PROT,%Lactose + water absorption

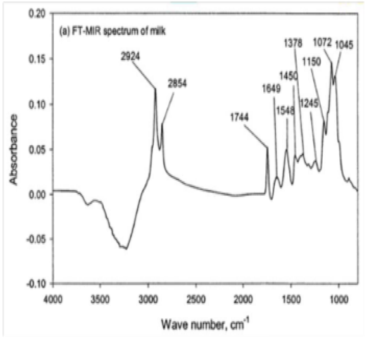
1970: First MilkoScan : one cell measurement

Milk spectroscopic analysis

1927:
First spectroscopic analysis
Filtered UV analysis on milk

1964:
Increasing use of MIR analysis to assess milk composition

1990: FT-MIR : Anadys

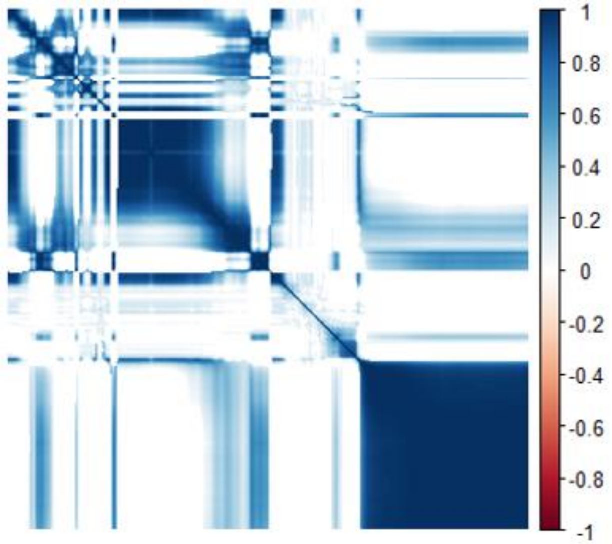


Higher dimensionality

Linear regression

$$Y = \mu + b_1 x_1 + b_2 x_2 + b_3 x_3$$

- Low computational resources
- Need a number of samples higher than the number of features in the regression
- Features cannot be highly correlated



60's: IRMA : InfraRed Milk Analyzer : 3 MIR Filters
%Fat,%PROT,%Lactose + water absorption

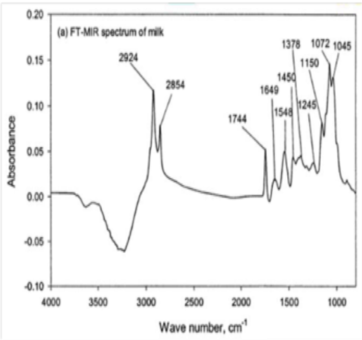
1970: First MilkoScan : one cell measurement

Milk spectroscopic analysis

1927:
First spectroscopic analysis
Filtered UV analysis on milk

1964:
Increasing use of MIR analysis to assess milk composition

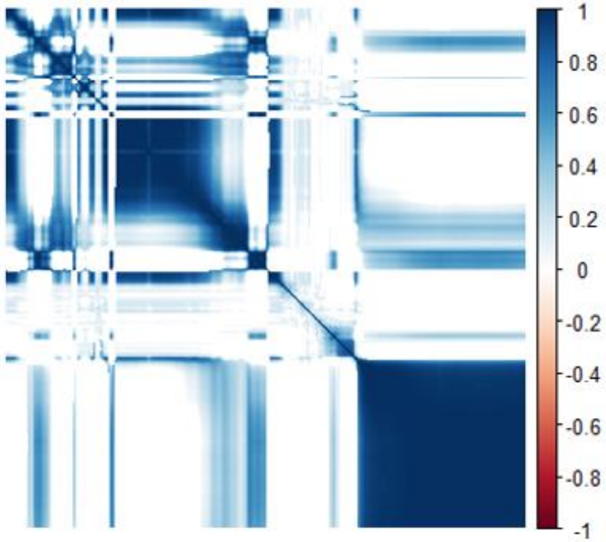
1990: FT-MIR : Anadys



Higher dimensionality

Linear regression
$Y = \mu + b_1 x_1 + b_2 x_2 + b_3 x_3$

- Low computational resources
- Need a number of samples higher than the number of features in the regression
- Features cannot be highly correlated



PLS regression
$Y = \mu + b_1 f_1 + b_{..} f_{..} + b_n x_n$

- Ok for colinearity
  - Create a new space where the correlations between features are minimized
- Ok if low number of samples
- Take information between spectral points

Conflict

New models with AI ?

Artificial intelligence

You

Give me a definition of artificial intelligence and machine learning

ChatGPT

Definition of Artificial Intelligence (AI)

Artificial Intelligence (AI) refers to the branch of computer science that is concerned with building smart machines capable of performing tasks that typically require human intelligence. AI integrates a variety of techniques and technologies, including computer science, cognitive science, linguistics, psychology, and more, to enable machines to sense, comprehend, act, and learn. AI systems are designed to handle tasks such as visual perception, speech recognition, decision-making, and language translation.

Definition of Machine Learning (ML)

Machine Learning (ML) is a subset of AI focused on the development of algorithms and statistical models that allow computers to perform specific tasks without explicit instructions, relying instead on patterns and inference. It is fundamentally about building programs with the ability to access data and use it to learn for themselves. ML enables systems to improve their performance on tasks over time without being directly programmed to do so, through exposure to new data.

Machine learning is an artificial intelligence

Arthur Lee Samuel

Machine Learning includes methods given the **ability to the computers to lean without being explicitly programmed** to achieve this objective.

Tom Mitchell introduces the concept of « Learning by experience »

« A computer program is said to learn from experience E with respect to some task (T) and some performance measure P, if its performance on T, as measured by P, improves with experience E. »

Tom Mitchell introduces the concept of « Learning by experience »

« A computer program is said to learn from experience E with respect to some task (T) and some performance measure P, if its performance on T, as measured by P, improves with experience E. »

When have we started to use it ?

➔ When we have started to estimate regression coefficients using interative procedure.

From linear regression to artificial intelligence

It is just for marketing !

But, the complexity of models changes

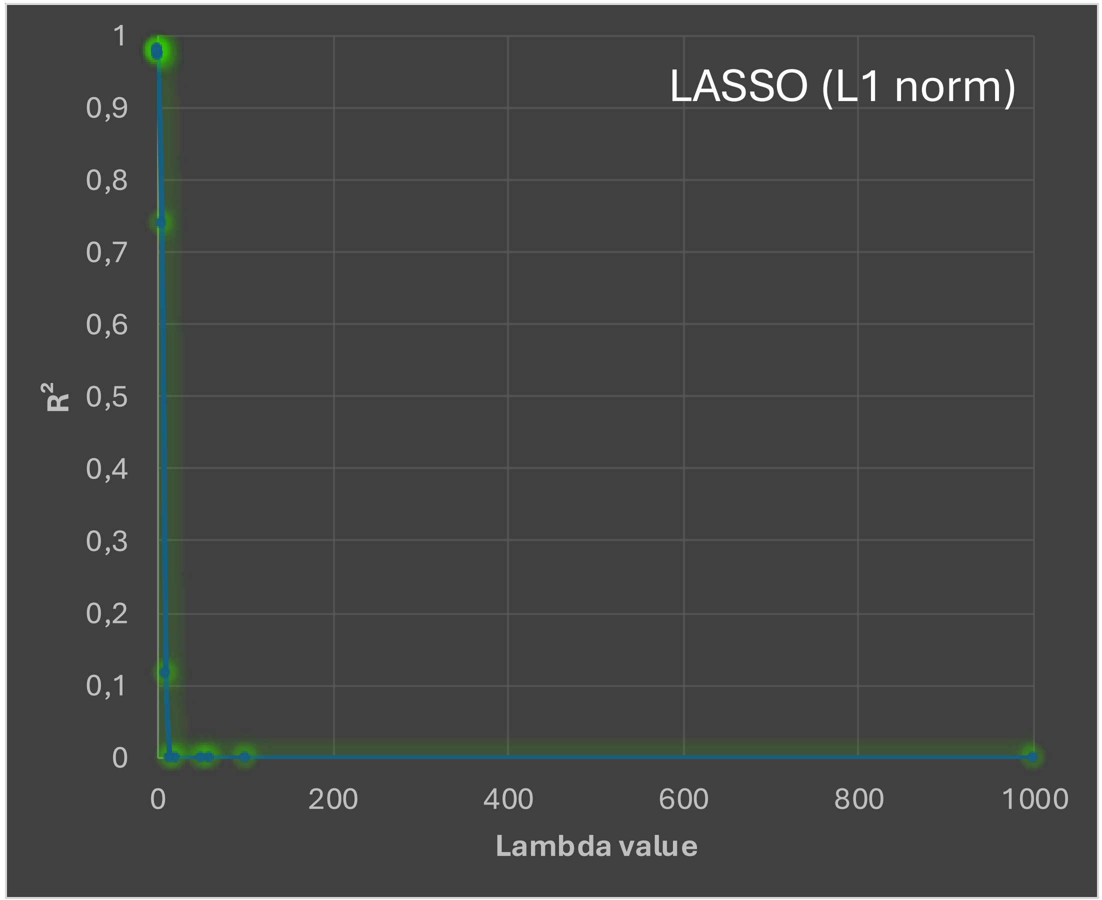
- Structure
- Number of parameters

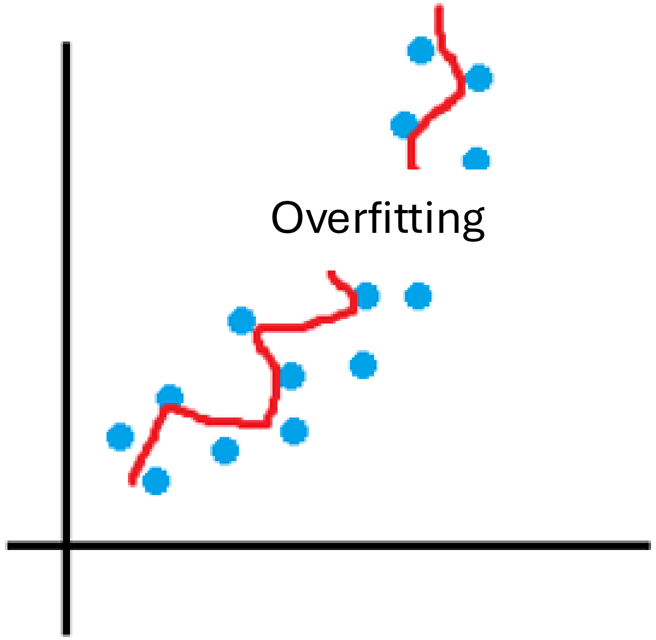
Models already tested in milk spectroscopic analysis

- Ordinary multivariate regression
- Penalized linear Regression (Ridge, Lasso, Elastic)
- Principal Component regression
- Partial Least Squares Regression
- Random Forest
- Support vector machine
- Neural Network – Multilayer perceptron
- Convolutional neural network

Linear until non-linear relationships

But the complexity of models changes





Higher features, higher train R^2

Impact of validation

- External validation:
  - Independent datasets
  - Take care to the hidden dependencies
- Internal validation:
  - Simple: Random splitting of the dataset
    - 2/3 calibration and 1/3 validation
  - Cross-validation
  - Bootstrapping …

→ **Following the validation, the results will be different**

But, the complexity of models changes

- Structure
- Number of parameters

Models already tested in milk spectroscopic analysis →

- Ordinary multivariate regression
- Penalized linear Regression (Ridge, Lasso, Elastic)
- Principal Component regression
- Partial Least Squares Regression
- Random Forest
- Support vector machine
- Neural Network – Multilayer perceptron
- Convolutional neural network

Linear until non-linear relationships →

Higher computational resources →

Journal of Dairy Science
Volume 103, Issue 12, December 2020, Pages 11585-11596

Research

A comparison of 4 different machine learning algorithms to predict lactoferrin content in bovine milk from mid-infrared spectra

H. Soyeurt [1], C. Grelet [2], S. McParland [3], M. Calmels [4], M. Coffey [5], A. Tedde [1], P. Delhez [1 6], F. Dehareng [2], N. Gengler [1]

Table 1. The 10-fold cross-validation and external-validation performances for predicting lactoferrin content in milk using 4 different machine learning algorithms[1]

		PLSR	PLS + Linear SVR	PLS + Polynomial SVR	PLS + ANN
Selection function		oneSE	oneSE	best	best = oneSE[2]
Calibration (n = 5,541)	Parameters	nLV[3] = 23	C[4] = 5	degree = 3; scale = 0.01; C = 1	size = 4; decay = 0.5
	R^2c	0.53	0.53	0.64	0.60
	RMSEc	140.94	144.32	125.89	130.59
Cross-validation	R^2cv	0.51	0.53	0.56	0.55
	R^2cv SD	0.03	0.03	0.03	0.03
	RMSEcv	144.31	144.60	138.40	139.01
	RMSEcv SD	5.77	5.61	8.08	5.05
	RPD	1.43	1.42	1.49	1.48
External validation (n = 836)	R^2v	0.61	0.63	0.62	0.60
	RMSEv	163.76	174.92	166.75	162.17

Soyeurt et al., JDS 2020

Example : **Lactoferrin**

Mean : 260 mg/L

$|\text{Difference}|_{\text{PLS vs ANN}}$: 1.59 mg/L → 0.61%

$|\text{Difference}|_{\text{PLS vs LSVM}}$: 11.16 mg/L → 4.29%

$|\text{Difference}|_{\text{PLS vs PSVM}}$: 2.99 mg/L → 1.15%

$|\text{Difference}|_{\text{ANN vs LSVM}}$: 12.75 mg/L → 4.90%

$|\text{Difference}|_{\text{ANN vs PSVM}}$: 4.58 mg/L → 1.76%

$|\text{Difference}|_{\text{PSVM vs LSVM}}$: 8.17 mg/L → 3.14%

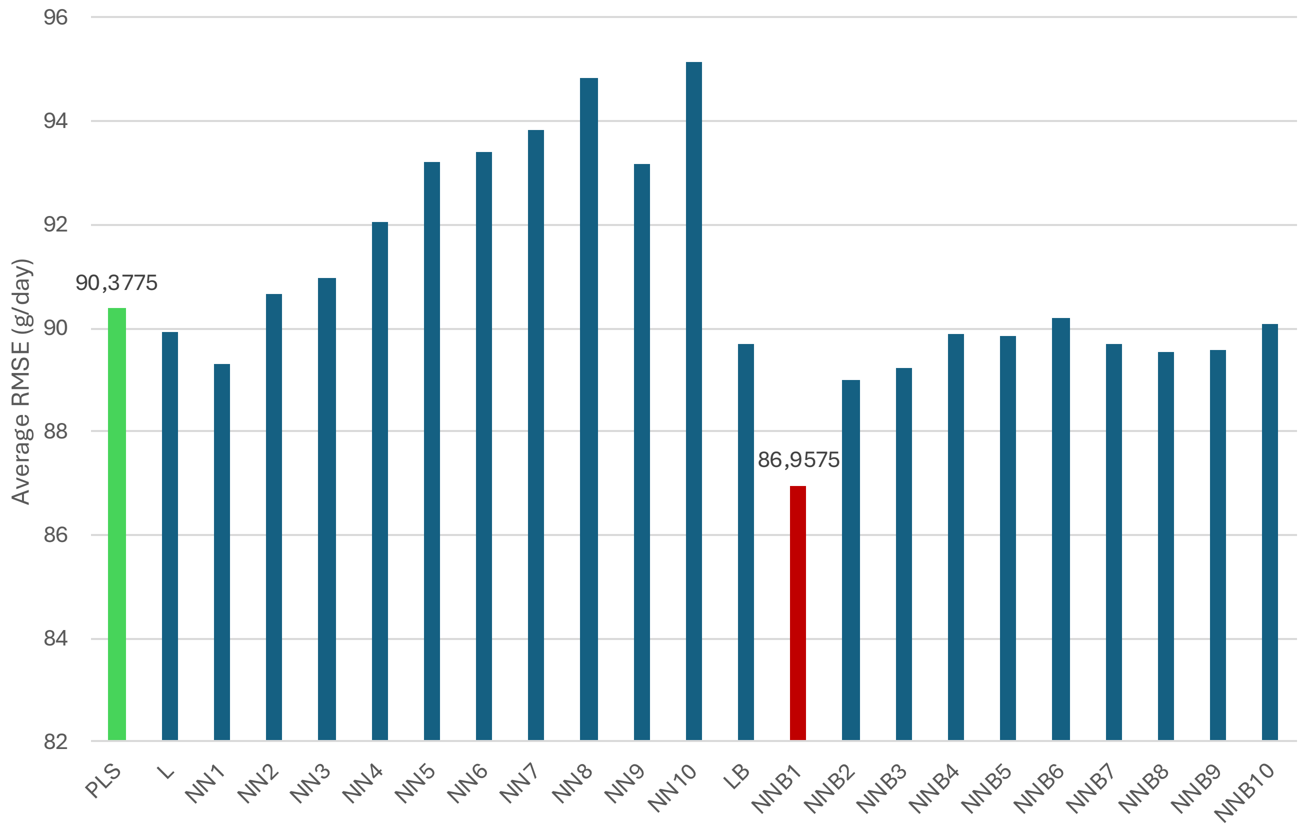
Small differences !

Journal of Dairy Science
Volume 105, Issue 10, October 2022, Pages 8272-8285

Research

Predicting methane emission in Canadian Holstein dairy cattle using milk mid-infrared reflectance spectroscopy and other commonly available predictors via artificial neural networks

Saeed Shadpour [1], Tatiane C.S. Chud [1], Dagnachew Hailemariam [2], Graham Plastow [2], Hinayah R. Oliveira [1], Paul Stothard [2], Jan Lassen [3], Filippo Miglior [1 4], Christine F. Baes [1], Dan Tulpan [1], Flavio S. Schenkel [1]



Example : **Methane**

Mean : +/- 400g/jour

|Difference| $_{\text{PLS vs NN}}$: 3g/day ➔ 0.75%

Journal of Dairy Science
Volume 107, Issue 2, February 2024, Pages 978-991

Research

Predicting methane emissions of individual grazing dairy cows from spectral analyses of their milk samples

S. McParland, M. Frizzarin, B. Lahart, M. Kennedy, L. Shalloo, M. Egan, K. Starsmore, D.P. Berry

Method	Predictor	Statistic											
		Bias	RMSE	r	b	RPIQ	Method		Bias	RMSE	r	b	RPIQ
	Previous							Previous					
	a.m.	-0.23 (8.97)	43.04 (2.36)	0.64 (0.03)	0.94 (0.11)	1.76 (0.16)		a.m.	0.12 (8.66)	42.50 (1.64)	0.64 (0.02)	0.90 (0.11)	1.76 (0.16)
	p.m.	0.03 (9.49)	43.73 (2.48)	0.63 (0.03)	0.93 (0.11)	1.74 (0.16)		p.m.	-0.11 (10.00)	43.15 (2.04)	0.64 (0.03)	0.87 (0.10)	1.73 (0.18)
	a.m. and p.m.	-0.01 (8.89)	42.15 (1.52)	0.65 (0.03)	0.94 (0.10)	1.77 (0.17)		a.m. and p.m.	-0.42 (8.26)	41.14 (1.86)	0.67 (0.02)	0.90 (0.09)	1.81 (0.16)
	Average a.m. and p.m.	-0.09 (8.81)	41.76 (1.98)	0.66 (0.03)	0.95 (0.11)	1.79 (0.18)		Average a.m. and p.m.	-0.42 (7.96)	41.45 (1.45)	0.66 (0.03)	0.91 (0.12)	1.80 (0.15)
	Following							Following					
	a.m.	-0.49 (9.34)	44.62 (1.46)	0.61 (0.02)	0.95 (0.09)	1.71 (0.11)		a.m.	-0.73 (8.93)	43.59 (1.42)	0.64 (0.02)	0.92 (0.11)	1.75 (0.11)
Partial least squares	p.m.	-0.06 (9.93)	45.23 (0.90)	0.60 (0.04)	0.93 (0.12)	1.69 (0.12)	**Neural networks**	p.m.	-0.36 (9.08)	44.36 (0.37)	0.62 (0.03)	0.89 (0.07)	1.72 (0.11)
	a.m. and p.m.	-0.20 (9.14)	44.14 (0.76)	0.62 (0.03)	0.94 (0.10)	1.73 (0.11)		a.m. and p.m.	-0.96 (8.84)	42.48 (1.02)	0.66 (0.03)	0.89 (0.08)	1.80 (0.13)
	Average a.m. and p.m.	-0.56 (9.43)	44.00 (1.40)	0.63 (0.03)	0.95 (0.10)	1.74 (0.11)		Average a.m. and p.m.	-0.46 (8.23)	42.33 (1.38)	0.66 (0.03)	0.92 (0.09)	1.81 (0.13)
	Flanking							Flanking					
	a.m.	-0.42 (9.08)	38.09 (1.87)	0.68 (0.03)	0.95 (0.11)	1.87 (0.17)		a.m.	0.16 (8.97)	38.23 (1.50)	0.68 (0.03)	0.92 (0.12)	1.86 (0.17)
	p.m.	-0.11 (9.71)	39.25 (1.57)	0.66 (0.03)	0.93 (0.12)	1.81 (0.15)		p.m.	-0.46 (9.27)	39.16 (0.94)	0.66 (0.02)	0.87 (0.07)	1.82 (0.14)
	a.m. and p.m.	-0.10 (8.83)	37.98 (1.36)	0.68 (0.03)	0.95 (0.11)	1.87 (0.16)		a.m. and p.m.	-0.77 (8.43)	37.17 (1.53)	0.70 (0.02)	0.91 (0.09)	1.91 (0.15)
	Average a.m. and p.m.	-0.25 (8.92)	37.56 (2.00)	0.69 (0.03)	0.96 (0.11)	1.90 (0.16)		Average a.m. and p.m.	-0.33 (7.95)	37.46 (4.01)	0.71 (0.03)	0.92 (0.11)	1.95 (0.16)

Example : **Methane**

Also low differences

Open Access Article

Identifying Health Status in Grazing Dairy Cows from Milk Mid-Infrared Spectroscopy by Using Machine Learning Methods

by Brenda Contla Hernández [1], Nicolas Lopez-Villalobos [2] and Matthieu Vignes [1,*]

[1] School of Fundamental Sciences, Massey University, Palmerston North 4442, New Zealand
[2] School of Agriculture and Environment, Massey University, Palmerston North 4442, New Zealand
[*] Author to whom correspondence should be addressed.

Animals **2021**, *11*(8), 2154; https://doi.org/10.3390/ani11082154

More nuanced in classification

Table 2. Performance of classification models obtained in 10 Monte Carlo cross-validation for classifying any health problem and healthy cows during lactation (early, mid and lactation) at two dairy farms during the 2016 production season [1].

Models [2]	Sensitivity	Specificity	Accuracy	PPV	NPV	AUC	MCC
PLS-DA	65.60 ± 5.97	79.59 ± 2.36	78.85 ± 2.23	15.25 ± 3.07	97.66 ± 0.5	72.59 ± 3.27	0.24 ± 0.04
RF	46.22 ± 8.62	79.26 ± 2.15	77.51 ± 1.75	10.94 ± 1.88	96.38 ± 0.73	62.74 ± 3.78	0.14 ± 0.04
SVM	66.39 ± 6.80	76.39 ± 2.92	75.84 ± 2.42	13.48 ± 1.62	97.61 ± 0.61	71.39 ± 2.37	0.22 ± 0.02
NN	61.74 ± 15.99	97.00 ± 2.85	95.16 ± 3.26	59.99 ± 26.20	97.87 ± 0.87	79.37 ± 9.16	0.58 ± 0.22
CNN	57.02 ± 12.70	92.5 ± 5.27	90.63 ± 4.98	33.82 ± 13.41	97.5 ± 0.75	74.76 ± 6.88	0.39 ± 0.13
ESA	57.15 ± 12.38	87.61 ± 6.19	86.02 ± 6.21	24.06 ± 13.07	97.36 ± 0.77	72.38 ± 8.48	0.31 ± 0.16
ESMJ	60.75 ± 5.98	83.57 ± 2.56	82.36 ± 2.27	17.18 ± 3.21	97.46 ± 0.55	72.16 ± 2.9	0.25 ± 0.04
ESWA	56.43 ± 14.56	85.13 ± 7.41	83.61 ± 7.36	21.33 ± 14.18	97.22 ± 0.97	70.78 ± 9.71	0.27 ± 0.17

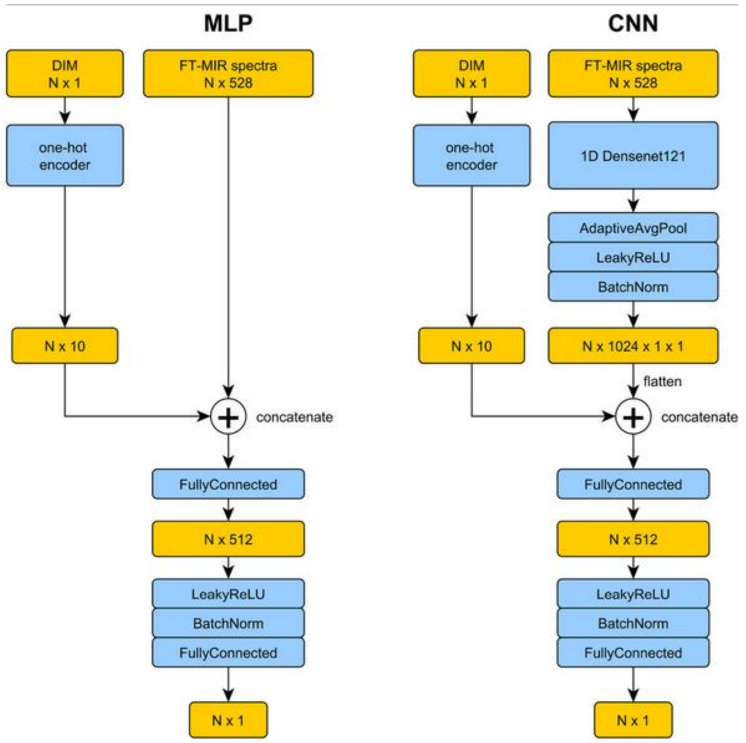
[1] These values correspond to the mean ± SD obtained by 10-fold Monte Carlo cross-validation for classifying any health problem (lameness, mastitis, reproductive disorder, etc.). From the cows' records, the positive cases were cows that had any illness (lameness, mastitis, reproductive disorder, etc.) and negative cases were cows who were healthy (no diagnosed disease); SD = Standard deviation; PPV = positive predicted value; NPV = negative predicted value; AUC = area under the receiver operating characteristic curve; MCC = Matthews correlation coefficient. [2] Models used to perform the classification: PLS-DA = partial least squares discriminant analysis, RF = random forest, SVM = support vector machine, NN = neural network, CNN = convolutional neural network, ESA = ensemble stacking average, ESMJ = ensemble stacking major voting and ESWA = ensemble stacking weighted average.

ELSEVIER
Journal of Dairy Science
Volume 105, Issue 4, April 2022, Pages 3615-3632

Research

Pregnancy status predicted using milk mid-infrared spectra from dairy cattle

K.M. Tiplady [1,2], M.-H. Trinh [1], S.R. Davis [1], R.G. Sherlock [1], R.J. Spelman [1], D.J. Garrick [2], B.L. Harris [1]



More nuanced in classification

Table 3. Model performance for multilayer perceptron (MLP) and convolutional neural network (CNN) approaches based on strategy 3 data[1]: accuracy (Acc), sensitivity (Sens), specificity (Spec), and area under the receiver operating characteristic curve (AUC) values within the training, herd-independent validation (VAL-Test) and pregnancy-associated glycoproteins validation (VAL-PAG) data sets

Deep learning approach[2] and model[3]	Training				Test validation (VAL-Test)				Glycoprotein-based validation (VAL-PAG)			
	Acc	Sens	Spec	AUC	Acc	Sens	Spec	AUC	Acc	Sens	Spec	AUC
MLP approach												
FT-MIR spectra	0.592	0.574	0.611	0.628	0.586	0.580	0.607	0.632	0.664	0.672	0.569	0.669
FT-MIR spectra + DIM	0.594	0.621	0.566	0.631	0.614	0.629	0.564	0.635	0.692	0.709	0.499	0.647
FT-MIR spectra (pre-adjusted for DIM)	0.559	0.554	0.564	0.583	0.562	0.567	0.547	0.581	0.554	0.547	0.636	0.636
CNN approach												
FT-MIR spectra	0.625	0.625	0.625	0.675	0.611	0.620	0.582	0.641	0.684	0.696	0.554	0.676
FT-MIR spectra + DIM	0.645	0.670	0.620	0.700	0.636	0.659	0.563	0.654	0.723	0.741	0.519	0.685
FT-MIR spectra (pre-adjusted for DIM)	0.982	0.975	0.988	0.998	0.668	0.790	0.273	0.551	0.759	0.805	0.266	0.564

Where is the highest error ?

Kind of model

Validation ?

The variability is bigger within the same model using different sets of features

ELSEVIER
Journal of Dairy Science
Volume 105, Issue 10, October 2022, Pages 8272-8285

Research

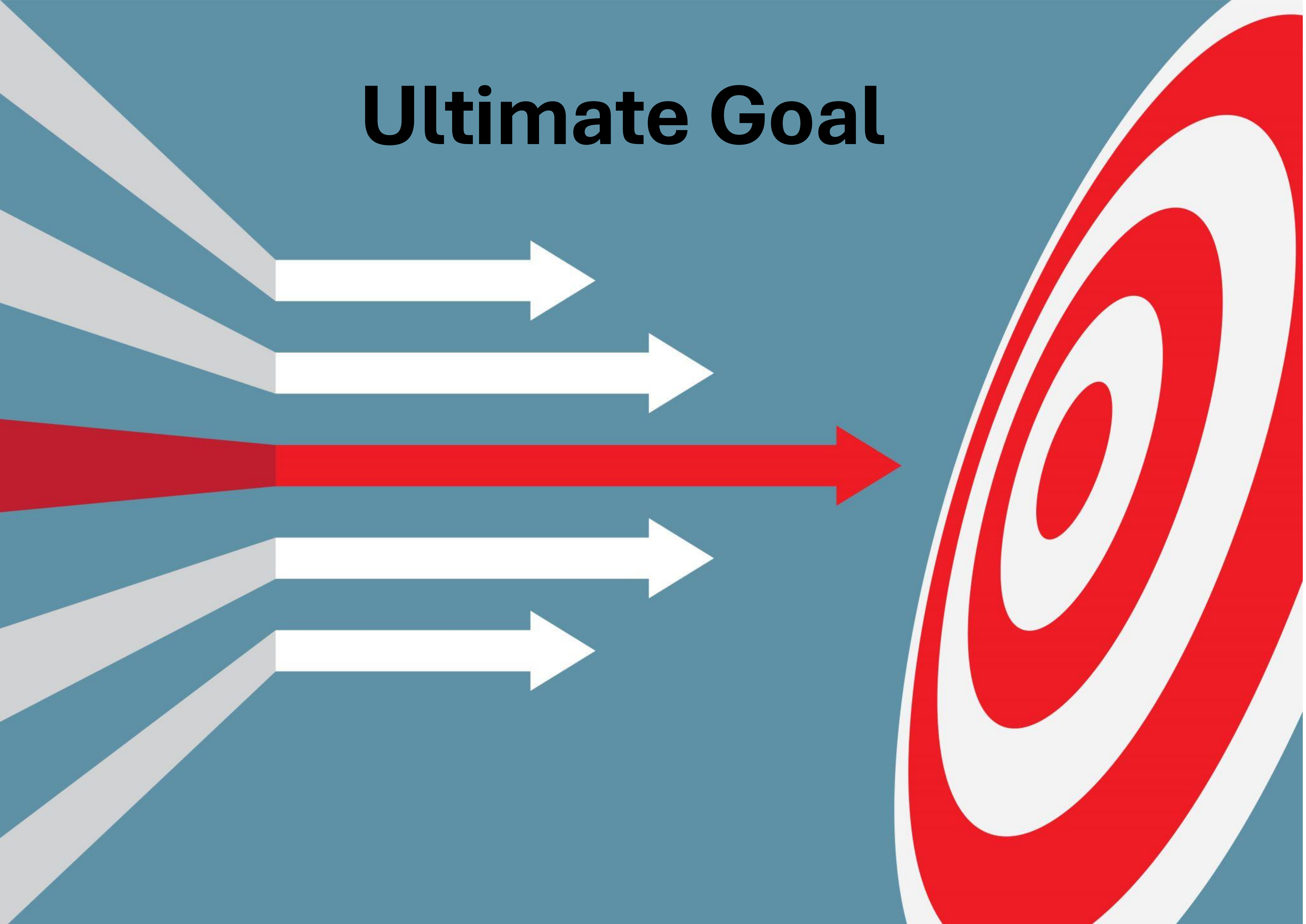
Predicting methane emission in Canadian Holstein dairy cattle using milk mid-infrared reflectance spectroscopy and other commonly available predictors via artificial neural networks

Saeed Shadpour [1], Tatiane C.S. Chud [1], Dagnachew Hailemariam [2], Graham Plastow [2], Hinayah R. Oliveira [1], Paul Stothard [2], Jan Lassen [3], Filippo Miglior [1 4], Christine F. Baes [1], Dan Tulpan [1], Flavio S. Schenkel [1]

Features	PLS	L	NN1	NN2	NN3	NN4	NN5	NN6	NN7	NN8	NN9	NN10	LB	NNB1	NNB2	NNB3	NNB4	NNB5	NNB6	NNB7	NNB8	NNB9	NNB10	SD
1	96.47	96.55	93.08	93.16	93.33	94.74	93.98	94.36	97.65	97.65	96.54	99.39	96.26	91.67	93.01	92.59	92.07	92.41	92.2	92.2	92.22	92.08	92.45	**2.26**
2	95.68	95.29	91.4	93.47	94.64	94.05	95.54	96.11	95.44	98.04	97.66	96.16	95.61	91.93	92.55	93.26	92.89	93.12	93.07	93.06	92.97	92.92	93.1	**1.78**
3	96.32	96.36	93.58	94.44	93.96	94.55	94.71	97.27	97.69	97.39	95.15	98.47	95.99	92.49	93.19	91.46	91.91	92.35	92.45	92.3	92.31	92.46	92.59	**2.13**
4	95.84	95.73	93.4	94.18	96.15	93.58	96.82	94.14	96.05	97.44	96.65	96.57	95.59	91.8	93.41	93.14	93.35	92.84	92.66	92.81	93.17	93.22	92.79	**1.67**
5	96.5	96.72	95.23	94.89	93.41	97.01	98.35	98.94	97.31	99.91	100.94	100.9	96.11	92.6	93.87	93.05	92.82	93.08	92.66	92.98	92.83	92.8	92.83	**2.84**
6	96.75	96.07	94.41	95.86	96.8	97.44	97.89	101.16	100.12	100.65	98.93	101.93	95.66	92.55	93.46	93.87	93.33	93.07	93.4	93.45	93.34	93.35	93.62	**2.94**
7	73.7	72.44	77.56	81.19	81.51	85.42	84.69	82.72	84.49	82.27	80.79	84.57	71.61	71.75	76.2	78.96	79.48	81.59	82.58	78.65	80.36	79.33	81.66	**4.15**
8	71.76	70.31	75.75	77.95	78.08	79.72	83.64	82.61	81.85	85.36	78.72	83.08	70.83	70.87	76.12	77.5	83.09	80.28	82.42	82.1	79.16	80.56	81.71	**4.44**
SD	**10.91**	**11.48**	**7.90**	**6.94**	**7.08**	**6.20**	**5.78**	**7.02**	**6.76**	**6.96**	**8.48**	**7.26**	**11.42**	**9.67**	**7.92**	**6.83**	**5.41**	**5.52**	**4.75**	**5.84**	**6.06**	**5.98**	**5.20**	

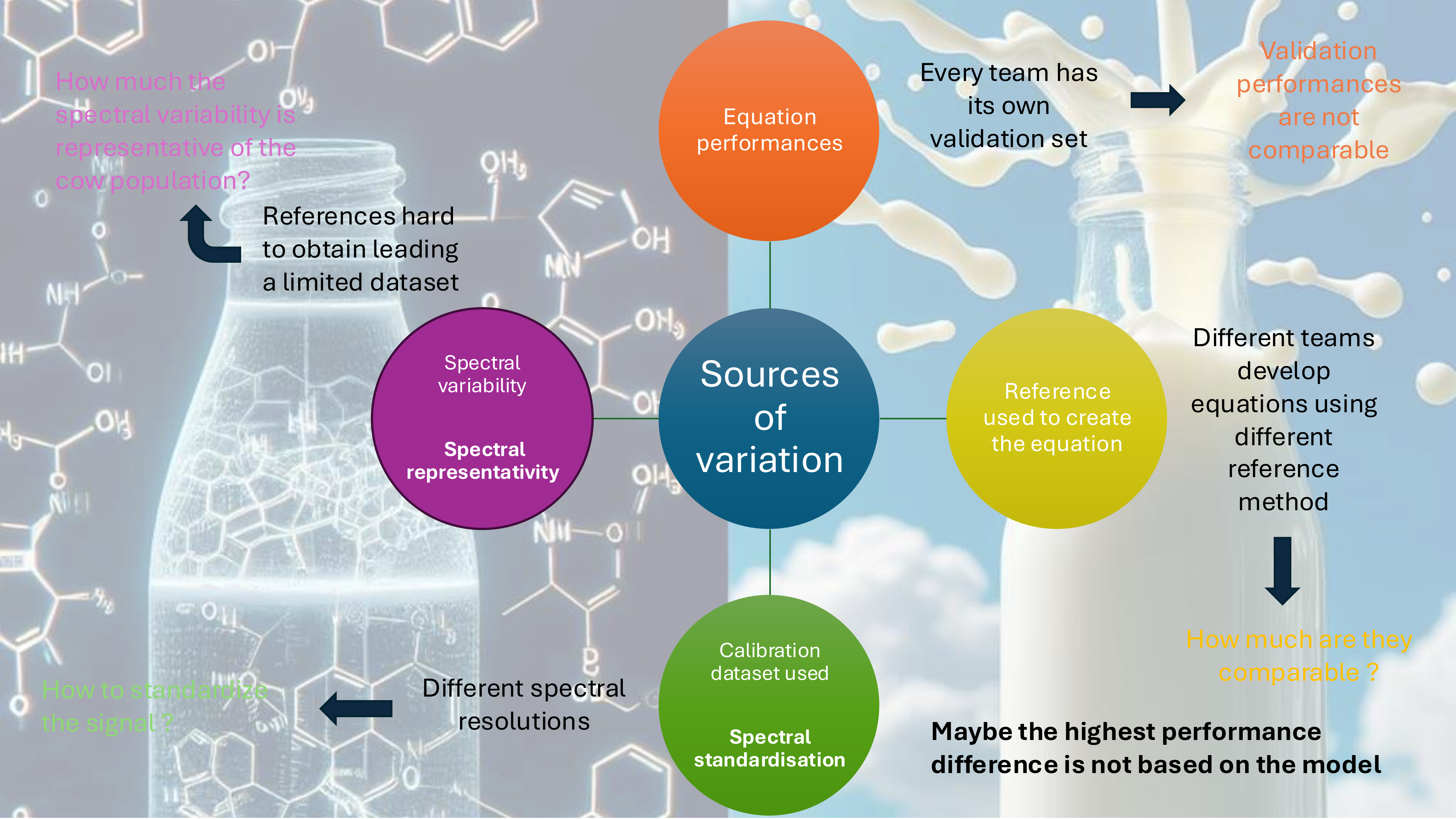
➔ **Not only focus on the model type**

Ultimate Goal



Having the same prediction for the same spectrum if you use a prediction model for the same trait

Not so simple …


Sources of variation
Equation performances
Every team has its own validation set
Validation performances are not comparable
Spectral variability
Spectral representativity
References hard to obtain leading a limited dataset
How much the spectral variability is representative of the cow population?
Reference used to create the equation
Different teams develop equations using different reference method
How much are they comparable ?
Calibration dataset used
Spectral standardisation
Different spectral resolutions
How to standardize the signal ?
Maybe the highest performance difference is not based on the model

ELSEVIER

Methods
Volume 186, February 2021, Pages 97-111

Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions

C. Grelet [a], P. Dardenne [a], H. Soyeurt [b], J.A. Fernandez [a], A. Vanlierde [a], F. Stevens [a], N. Gengler [b], F. Dehareng [a]

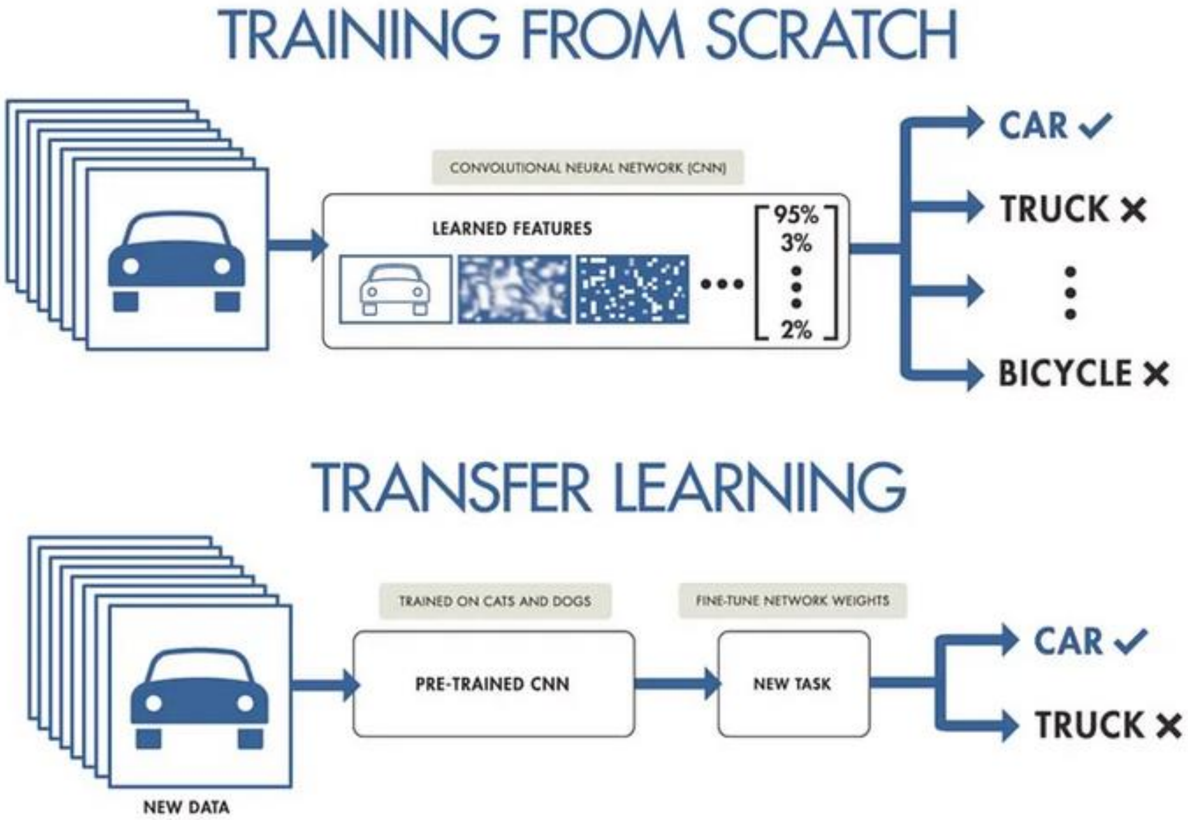
For regression mainly …

Highlights

- Robustness of models is essential to generate accurate phenotypes at a large scale.
- Variability in the reference and spectral data is a key element of robustness.
- In most cases, complexity of models is in contradiction with robustness.
- Models development should consider absorbance reliability of spectral regions.
- Quality assurance methods are necessary to evaluate models and predictions.

More complex models ...

Transfer learning is interesting



https://iamamadsiddiqui.medium.com/unleashing-the-power-of-transfer-learning-in-artificial-intelligence-e8a9deee62f3

Deeply used with image and video data source

More complex models ...

- Transfer learning is interesting:
    - Large number of data is available
    - Difficulty to share data
    - Save computation time

Take home message

All available models are developed from AI

More complex models did not seem to be significantly better for regression purposes

Maybe due to relatively low training set

However, in the future in the transfer learning is an interesting approach to build international models

Promotion of neural networks

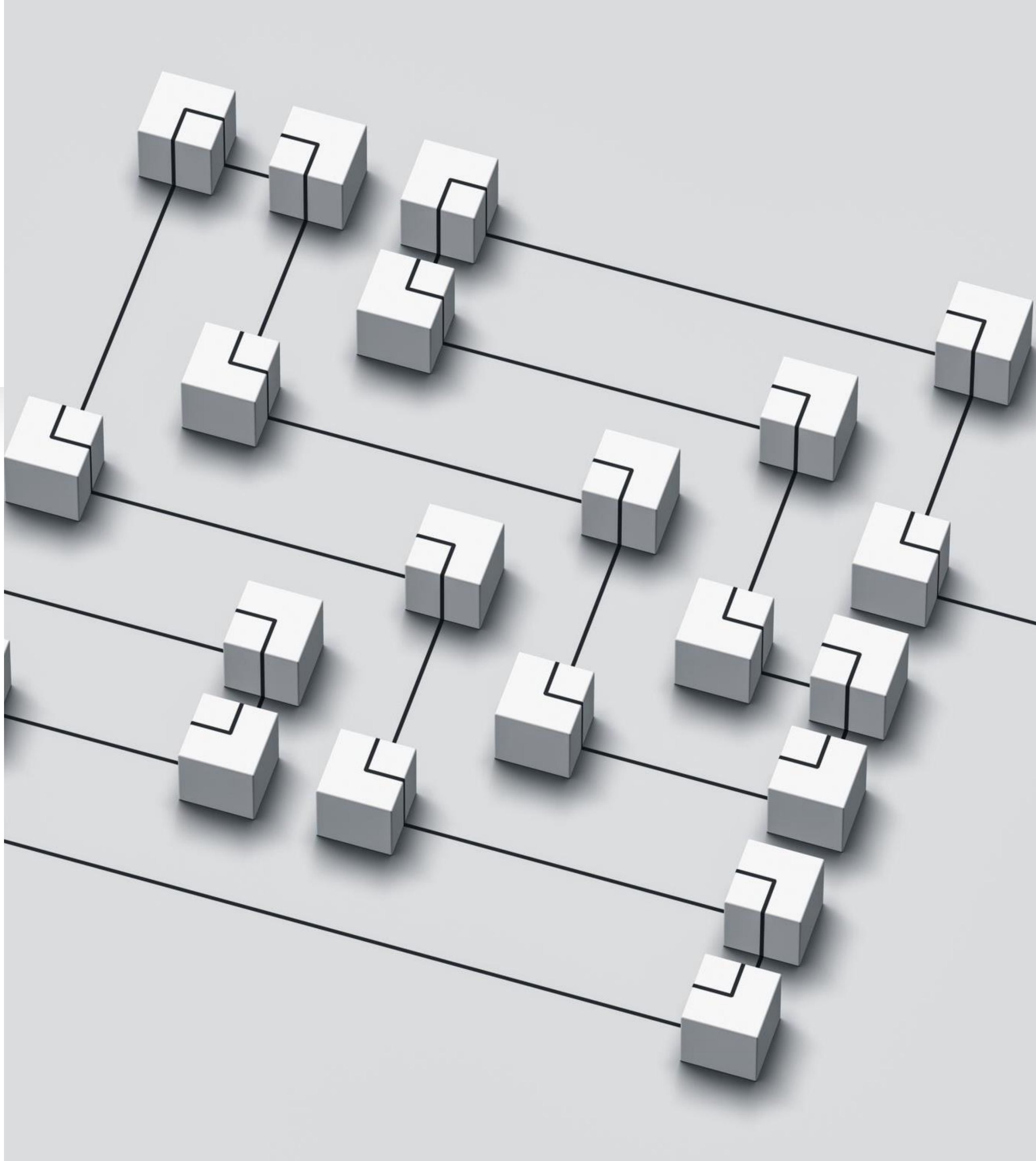
However, to achieve a good results we need to:

- Standardize the reference method used to estimate/quantify the trait of interest
- Standardize the spectral data

IDF/ICAR guidelines

New ideas

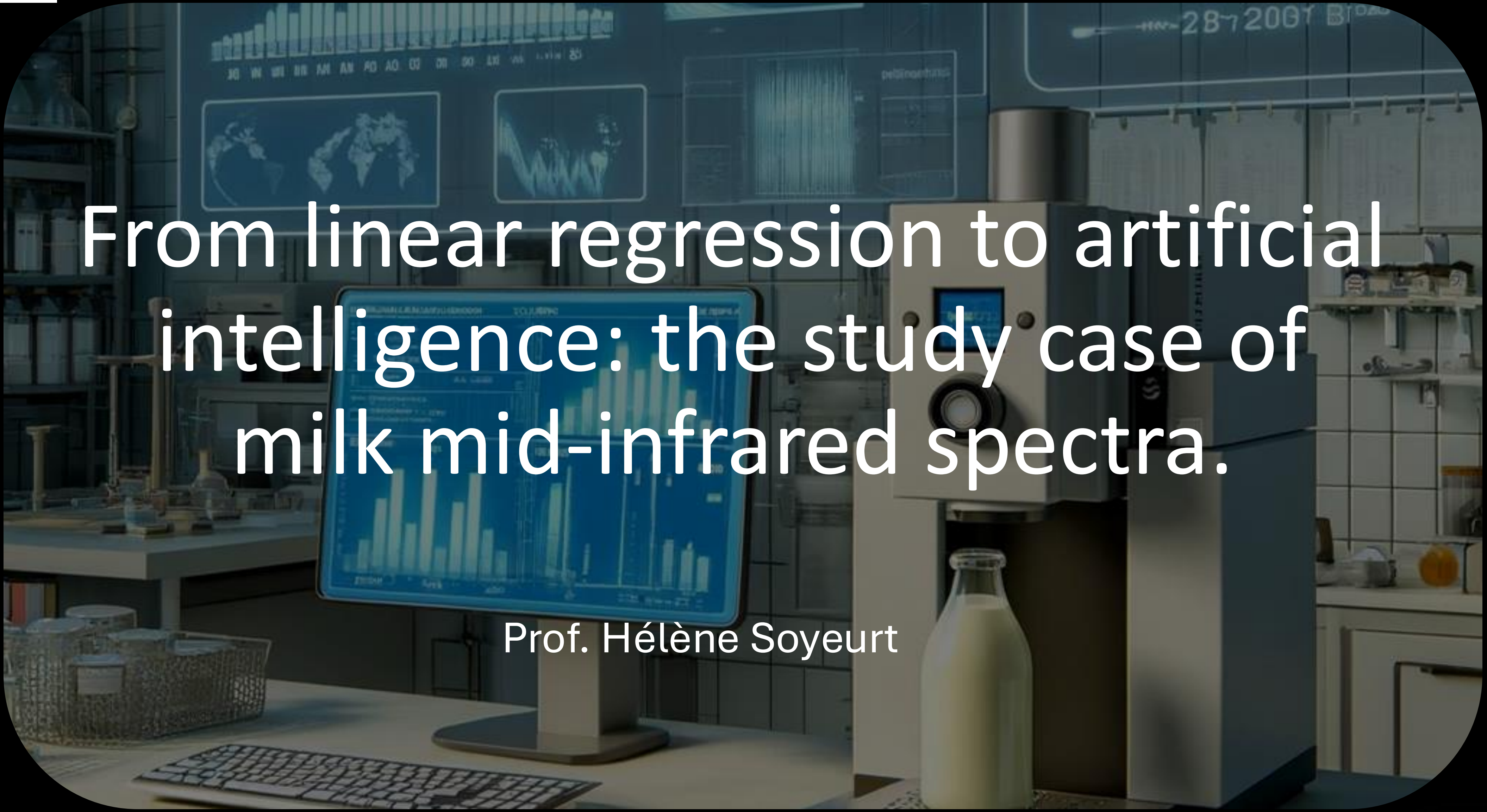
- Collect data routinely
  - Improvement of the prediction based on the repeated acquisition
- Collaborative learning
  - Improve the labelling of disease or useful information
- Consider the spectrum as an image
  - Use pre-trained neural network models



Environmental impact ?

Search the optimized situation. Not always needed to perform something complex and computationally demanding

From linear regression to artificial intelligence: the study case of milk mid-infrared spectra.



Prof. Hélène Soyeurt

46th Discovery Conference – Itasca – 6-9th May 2024