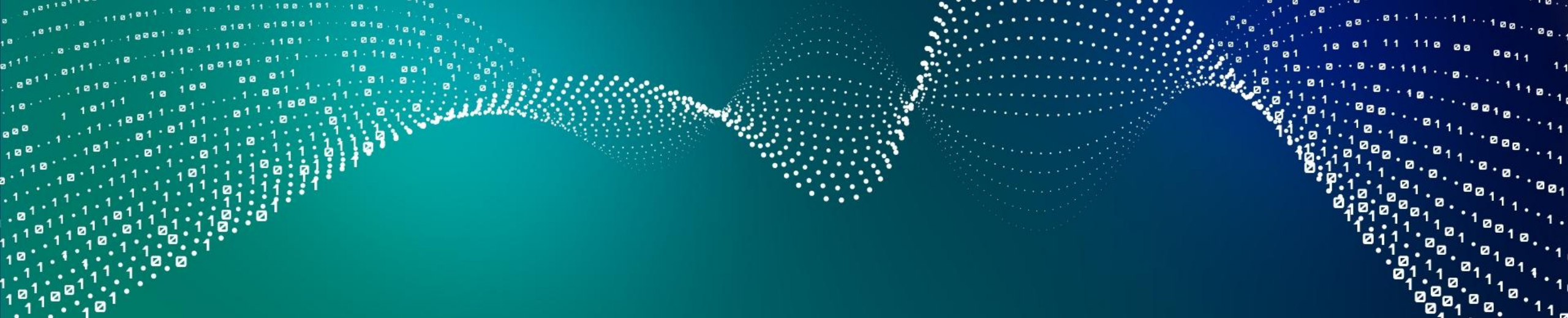




Ondalys

Headquarters :

4 rue Georges Besse
34 830 Clapiers
France

04 67 67 97 87

contact@ondalys.fr

www.ondalys.fr

L'Intelligence Artificielle au service de la spectroscopie

Dr. Sylvie ROUSSEL, CEO

Journée Utilisateurs
Spectroscopie
04 Juin 2026

 **Metrohm**
France SAS

Making sense of your data
Donnez du sens à vos données
 **ondalys**

Ondalys in short



- Founded in **2003**
- **Montpellier** (34), France



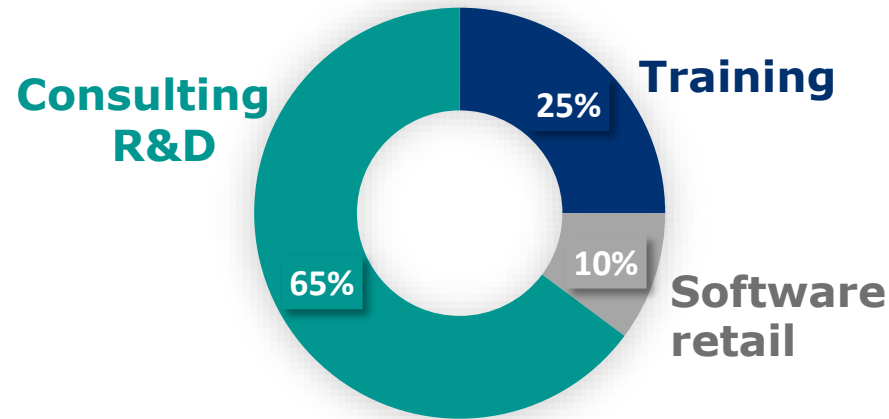
Multidisciplinary Data Scientist Team

- Machine Learning
- Instrumentation / Data
- Application sectors

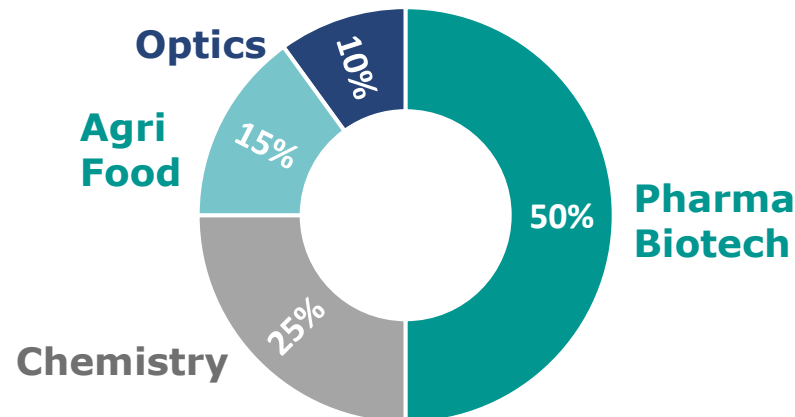
➔ Data analytics for customer satisfaction and quality services



Core business



Customers



Chemometrics Machine learning Predictive AI

- Data mining
- Spectroscopic calibrations
- Model robustness
- Sensor fusion
- Process monitoring (MSPC, BSPC)
- Design of Experiments (DoE / QbD)



Software partners



They trust us for consulting and training in Machine Learning



Pharma – Biotech



Chemicals – Petro.



Agriculture – Food



Optical sensors – R&D

sanofi

DEINOVE

MICHELIN

Bonduelle



Nestlé

bertin
technologies

SARTORIUS

GSK

GENETHON
CURE THROUGH INNOVATION

SYENSQO



TotalEnergies

MOËT & CHANDON
CHAMPAGNE

Perten
INSTRUMENTS
a PerkinElmer company

Metrohm

SERVIER

ORiL
INDUSTRIE

ARKEMA

ROYAL CANIN

MARS

BRUKER

BIO-RAD

ucb Pharma

ansm
Agence nationale de sécurité du médicament
et des produits de santé

Elkem

VALRHONA
Imaginons le meilleur du chocolat®

linea
SEMENCES

EADS

envea

Firmenich
for good, naturally

FERRING
PHARMACEUTICALS

ifp
Energies
nouvelles

Lidea
FRESH IDEAS FOR AGRICULTURE

euràlis
NOURRIR VOTRE CONFIANCE

cnrs

cirad

L'ORÉAL

Lonza

P&G

SOLVAY

BIP
Bureau national Interprofessionnel du Pruneau

IFV

ITERG
Expertise Corps Gras

RD
Institut de Recherche
pour le Développement
RANC

INRAE

Instrumental data



SPECTROSCOPY

- > Near Infrared (NIR)
- > Infrared (FTIR)
- > Raman
- > UV
- > Fluorescence
- > LIBS ...



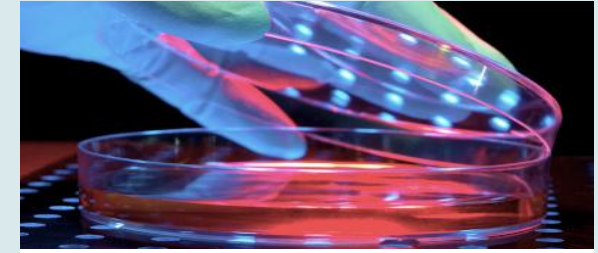
PROCESS PARAMETERS

- > In-line, on-line, at-line analyzers
- > Process parameters
- > Industrial IoT
- > In-line spectrometers ...



IMAGING

- > Hyperspectral
- > Multispectral
- > RGB
- > UV, visible, NIR, Raman, ...



PHYSICO-CHEMICAL

- > HPLC, GC
- > GC-MS
- > Physical measurements
- > NMR
- > ICP...



AGENDA

- What are the different layers of Artificial Intelligence?
- Data Analysis - Good Practices
- Machine Learning calibration models
- Calibration transfer
- Data fusion

The different layers of Artificial Intelligence (AI)

> Machine Learning

- › “Classical ML” – Chemometrics
- › “Predictive AI”
- › “Data Analytics”

> Deep Learning

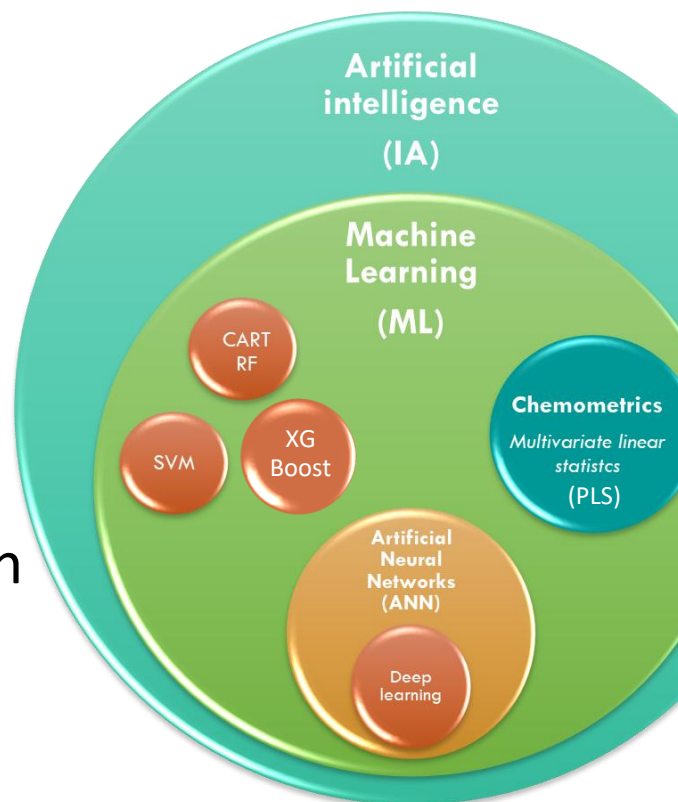
- › Usually on image, video, speech recognition

> Generative AI

- › Text and code generation

> Agentic AI

- › End-to-end Workflow



TYPES OF AI

Use cases
→ Automated customer service
→ End-to-end workflow execution

What it needs
→ Reliable underlying tools and APIs
→ Clear objectives and constraints
→ Human-in-the-loop escalation paths

AI Agents & Agentic AI
Plans, decides and takes actions.

Key risks
→ Unintended actions at scale
→ Compounding errors

Where value compounds and risk multiplies.
Constrain tightly, scale carefully.

Use cases
→ Writing content and code
→ Searching and summarising documents

What it needs
→ Strong prompt and UX design
→ Guardrails and human oversight

Generative AI
Creates content and summaries

Key risks
→ Hallucinations and confident errors
→ IP and data leakage

Can provide high leverage.
Treat it as an assistant, not an authority.

Use cases
→ Image and video analysis
→ Speech recognition

What it needs
→ Large labelled datasets
→ Significant compute

Deep Learning
Learns complex patterns automatically

Key risks
→ High cost & energy use
→ Low explainability
→ Overkill for simple problems

Powerful but expensive.
Use when rules and classical ML cannot do the job.

Use cases
→ Demand forecasting
→ Predicting customer churn

What it needs
→ Clean, representative historical data
→ Ongoing monitoring & retraining

Classical ML (Predictive AI)
Forecast outcomes from past data

Key risks
→ Bias baked into historical data
→ Model drift over time

Proven use cases with high ROI.
Less exciting than GenAI, but more reliable.

Claire Kitching,

[Linkedin post](#) on March 2026



AGENDA

- What are the different layers of Artificial Intelligence?
- **Data Analysis - Good Practices**
- Machine Learning calibration models
- Calibration transfer
- Data fusion

Good practices for spectroscopic data modeling

1. Data acquisition - Good Practices : the most important!

2. Data Analysis - Good Practices

➔ Take into account the factors of variability to measure representative samples

Environment

- Stray light
- Temperature
- Humidity
- Dust, vibrations

Spectral measurements (X)

- Measures (probe, resolution, signal/noise)
- Operators
- Different instruments

Samples

- Heterogeneity (ex: fruits)
- Preparation (ex: granulometry)
- Variety, location, etc
- Year, season, batch effect

Reference values (y)

- Large range to avoid extrapolation
- For all parameters
- Suitable distribution
- Sampling between X and y

Good practices for spectroscopic data modeling

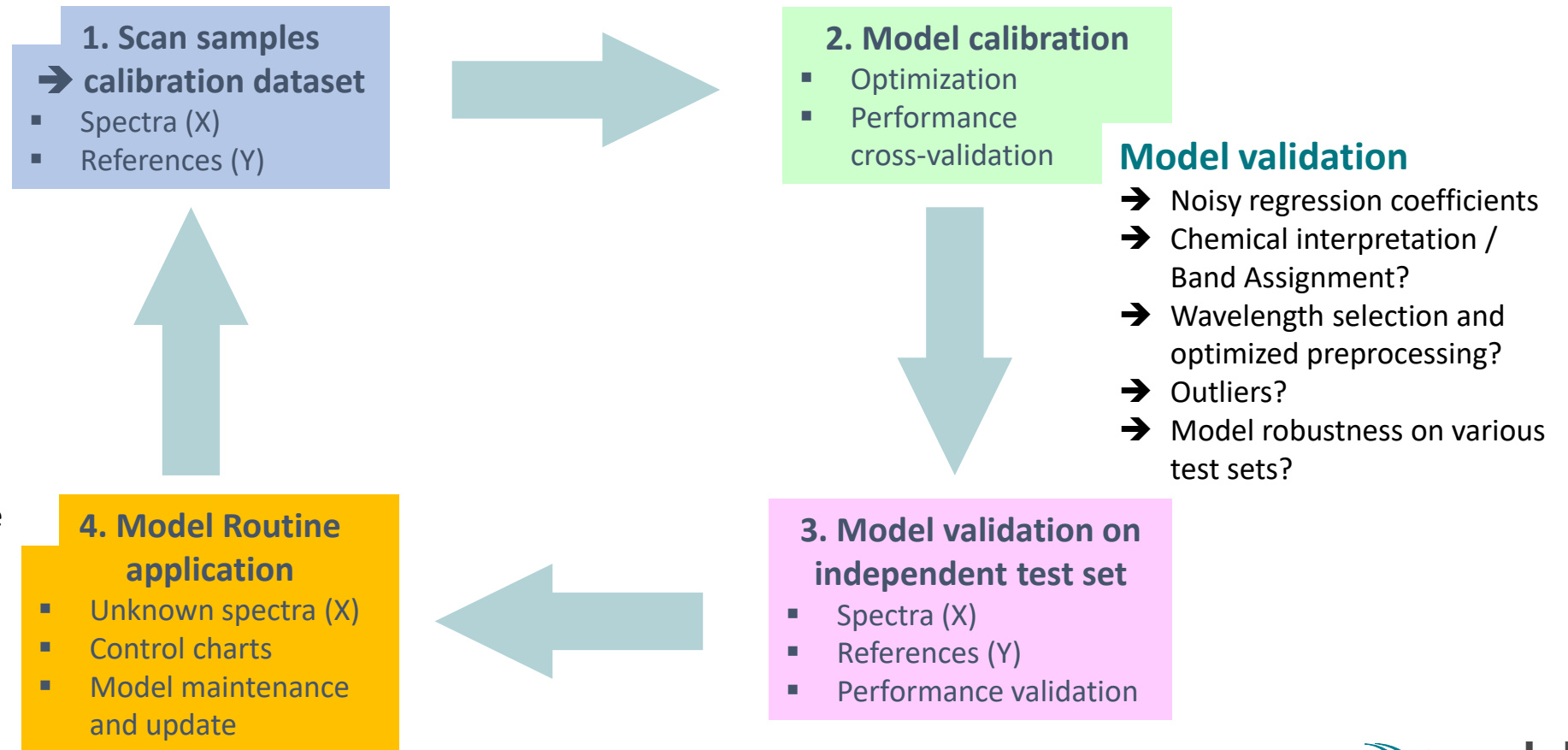
Exploratory analysis

- Outliers?
- Sources of variability?
- Preprocessing required?
- Reference Accuracy

Routine control

- Outlier detection
- Extrapolation?
 - > need for an update
- Drift?
 - > Need for spectrometer qualification or maintenance

Spectroscopic Data Analysis Workflow





AGENDA

- What is chemometrics / Machine Learning / Artificial Intelligence?
- Data Analysis - Good Practices
- **Machine Learning calibration models**
- Calibration transfer
- Data fusion

Machine Learning model comparison for meat composition prediction

Chemical composition of meat

- The most accurate predictions
- The most robust predictions for industrial application

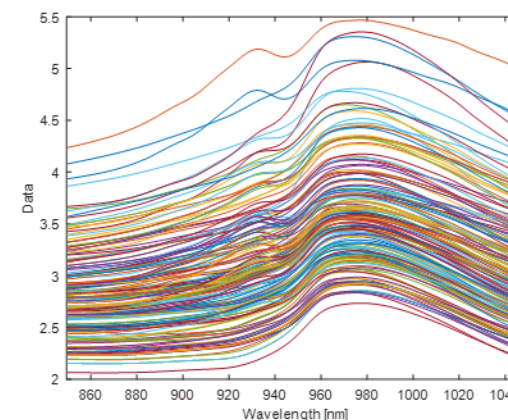


Ground meat data sets

- 172 calibration samples
- 43 validation samples
- Reference: humidity, protein and fat contents

Near Infrared (NIR) Spectroscopy

- Transmittance measurements
- Wavelength range: 850-1050nm



Machine Learning model comparison for meat composition prediction

Machine Learning Methods

- > **Linear model (PLS)**
 - > The most-widespread spectroscopic calibration method
- > **Local multivariate models (LWR)**
 - > Efficient for heterogeneous datasets
 - > Ex : different types of polymers
- > **Random Forest (RF)**
 - > Well suited for non-correlated variables
 - > Wavelength selection required
- > **Support Vector Machines (SVM)**
 - > Non linear model, but do not require large dataset
 - > Ex : non linear relationship between spectra and criteria to predict
 - > Ex : complex discrimination
- > **Artificial Neural Networks (ANN)**
 - > Non-Linear model, but require large dataset
 - > Ex : very large datasets such as commodity quality prediction

Fat content prediction based on Machine Learning Methods

ML Method	SEP (%)	Non linearity handling	Performance	Development Complexity	Risk of Overfitting
PLS	2.7	☹️	☹️	😊😊	😊😊
PLS on X^2	0.9	😊	😊	😊😊	😊
LWR	1.0	😊	😊	😊	😊
RF	1.0	😊	😊	😊	😊
SVR	0.5	😊	😊😊	☹️	☹️
ANN	0.4	😊	😊😊	☹️☹️	☹️☹️

➔ Ask us for the scientific study on ML method comparison

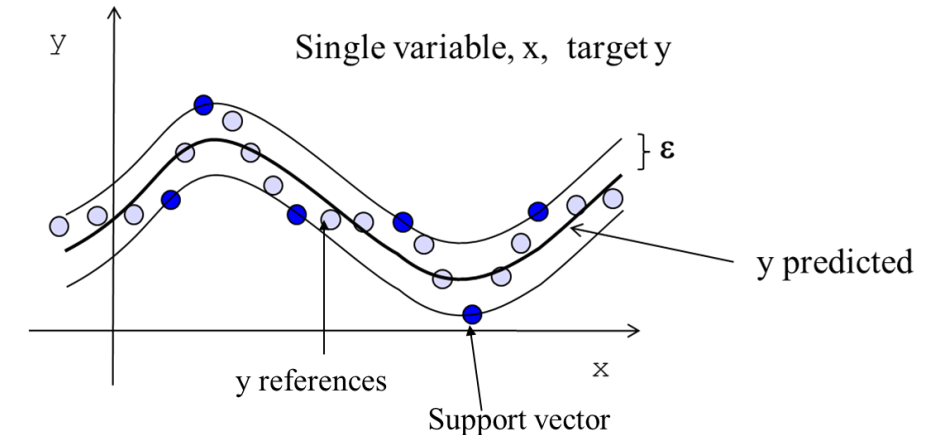


Focus on Support Vector Machines (SVM)



Supervised methods based on margins

- Only a few samples are used for the calculation of the final models
- Samples defining the margins = “**support vectors**”
- The algorithm tolerates small prediction errors and focuses on significant deviations

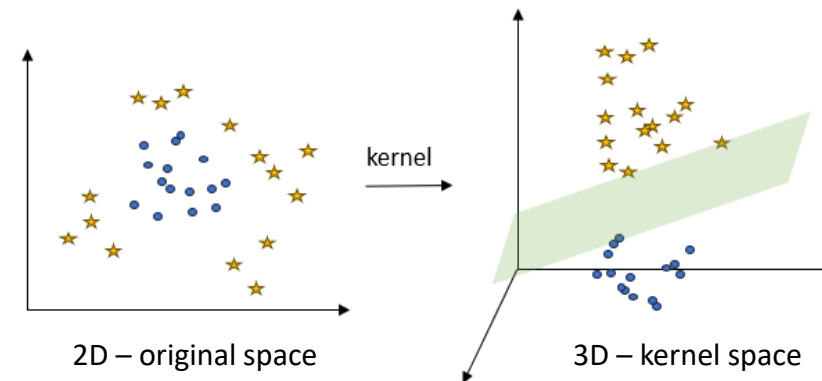


Source : Eigenvector Research Inc.



Versatile algorithm that does not necessarily need a lot of data and can perform well when non-linearities are present

- Non-linearities are managed using a **kernel transformation**
=> model training based on **neighbors**

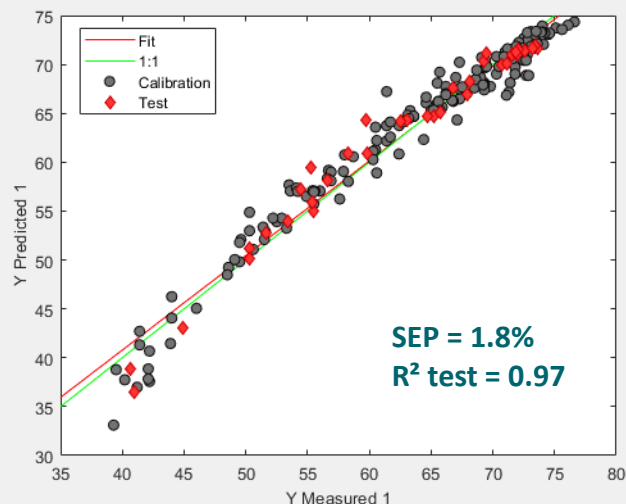


PLS vs SVM models when non linearities are present

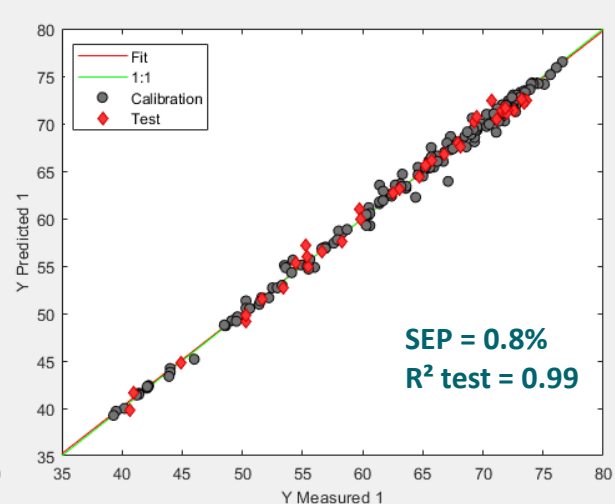
The relationships between spectra and the reference values show non-linearities

- PLS is not able to cope with large non-linearities
- SVM is able to model complex relationships (non-linearities, noise, clusters)

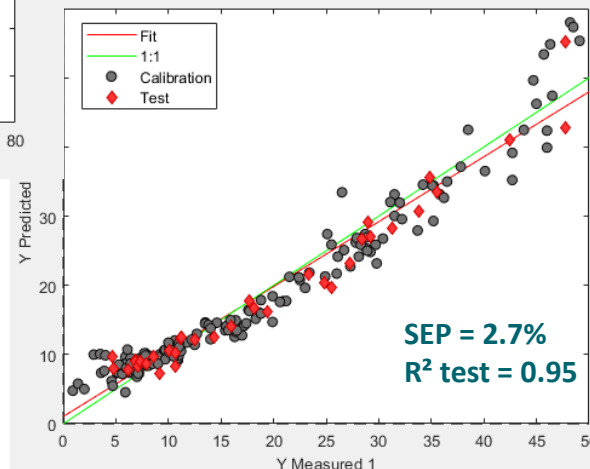
Moisture PLS



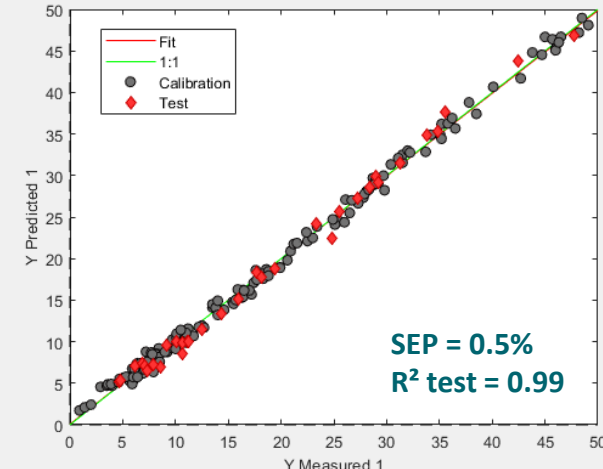
Moisture SVM-R



Fat content PLS



Fat SVM-R





AGENDA

- What are the different layers of Artificial Intelligence?
- Data Analysis - Good Practices
- Machine Learning calibration models
- **Calibration transfer**
- Data fusion

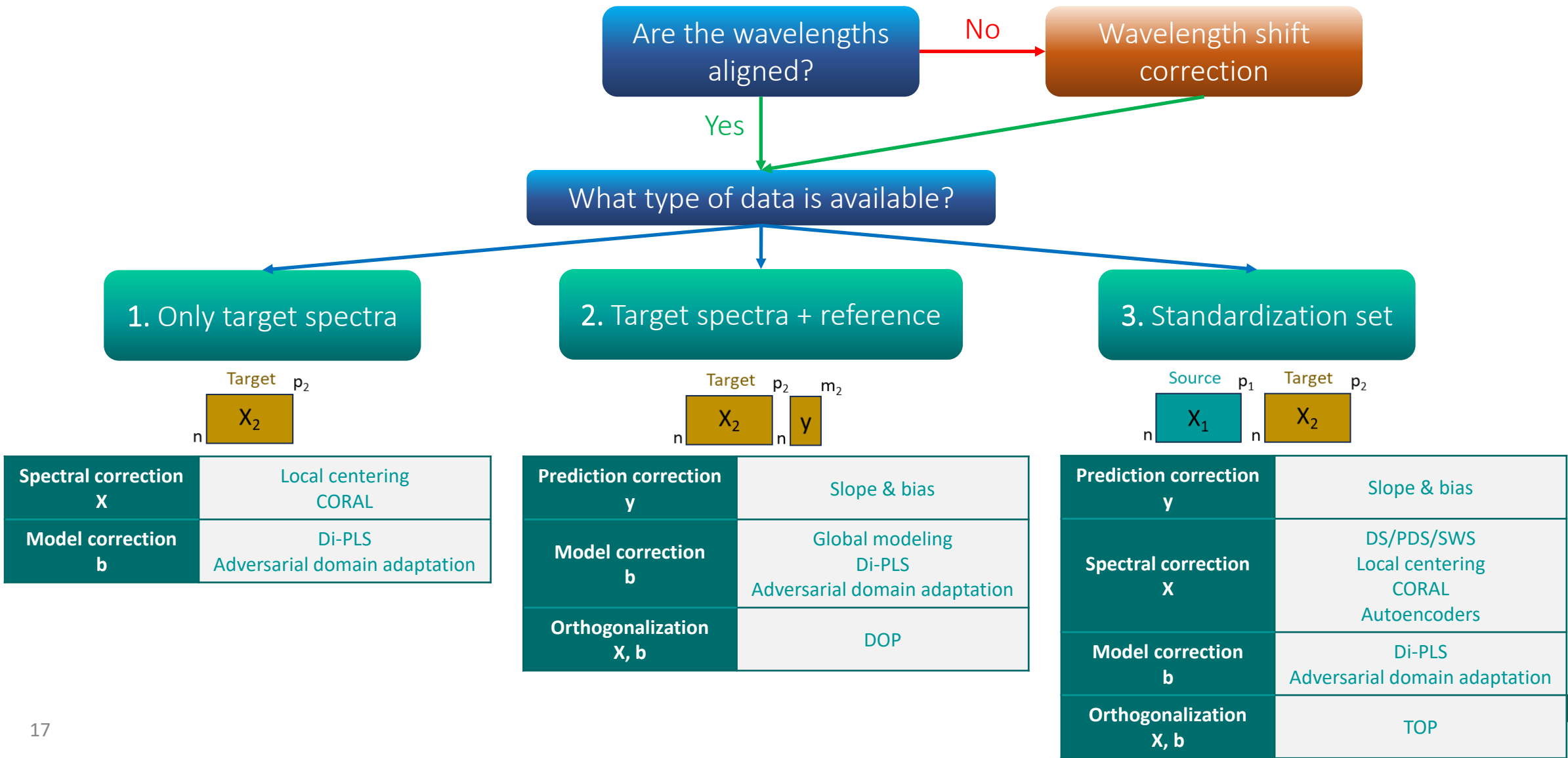
Calibration Transfer: A Robustness Issue

In what context should a calibration transfer be performed?

- A model running on a spectrometer to be used on another device
 - Same or different spectrometer brands
 - Different types: benchtop spectrometer versus handheld spectrometer
 - Different measurement conditions: laboratory measurement, field measurement, online measurement
 - Scale-up: transition from pilot to manufacturing
 - Transferring from spectrometer to hyperspectral imaging
- ➔ Spectroscopic models developed on one instrument often need to be adapted to the new conditions
- Differences in:
 - Instrumental response
 - Spectral resolution
 - Level of noise
 - Measurement conditions,...
- ➔ Calibration transfer is a common challenge in all spectroscopic applications



General principle of calibration transfer



Pour en savoir plus... Inscrivez-vous à notre webinaire!

WEBINAIRE le 11 Juin 2026 à 11h (CEST)

Transfert de calibrations spectrales :
Stratégies et bonnes pratiques!



Jordane POULAIN
Data Scientist
Expert in Machine Learning



Inscrivez-vous
au webinaire ici



Application example: Flash BMP® calibration transfer

Flash BMP®: Biochemical Methane Potential (BMP)

- Fast BMP evaluation for Methane production
- Metrohm Spectrometer DS2500
- Substrates (500+) :
 - > Green waste
 - > Agro-industrial
 - > Energy crops
 - > Manure
 - > Sludge



Improve waste stock management for higher methane yield with the NIRS DS2500 Analyzer



 **Metrohm**

 **ondalys**

Copyright

Application example: Flash BMP® calibration transfer

Flash BMP®: Biochemical Methane Potential (BMP)

- Inter-instrument transfer

- > Source instrument: Calibration developed on NIRFlex N-500 (Buchi)
- > Target instrument: Calibration transferred to a DS2500 (Metrohm)

- Standardization set available

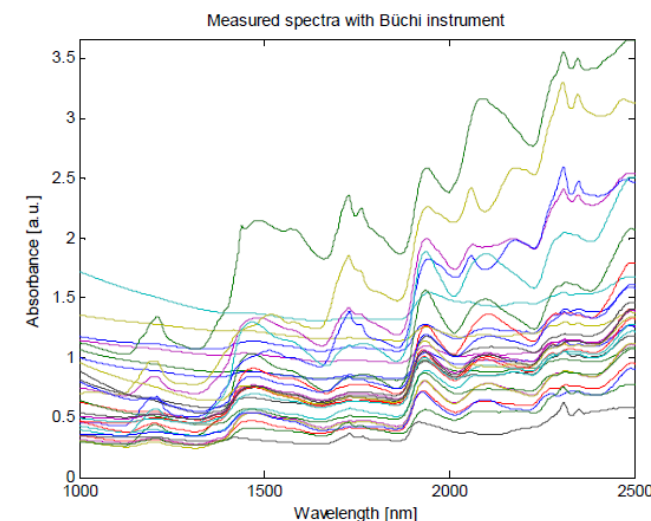
- > 20 samples for standardization
- > 10 samples for transfer validation

- Wavelength / Wavenumber alignment and Interpolation

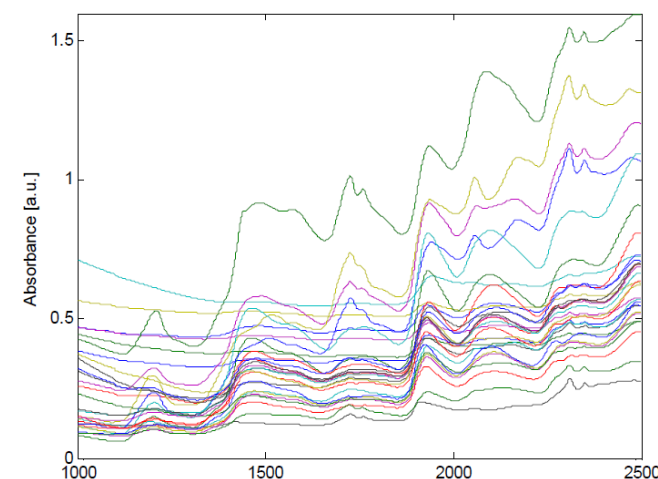
- > Buchi: 4 cm⁻¹
- > Metrohm : 0.5 nm

- Different transfer methods tested

- > Piecewise Direct Standardization (PDS)
 - > Most widespread method
 - > Less prone to overfit



NIRFlex N-500 (Buchi)

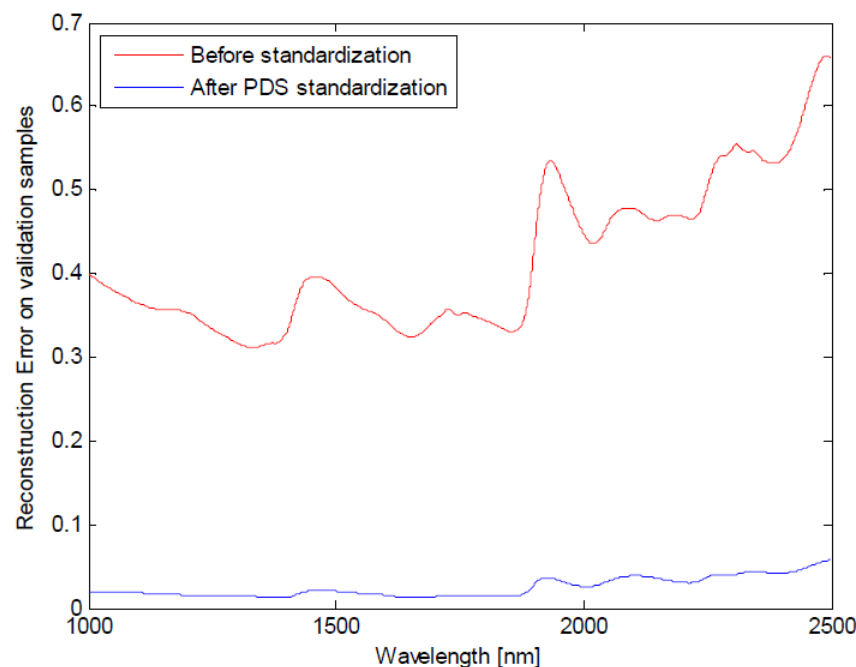


DS2500 (Metrohm)

Application example: Flash BMP® calibration transfer

Flash BMP®: Biochemical Methane Potential (BMP)

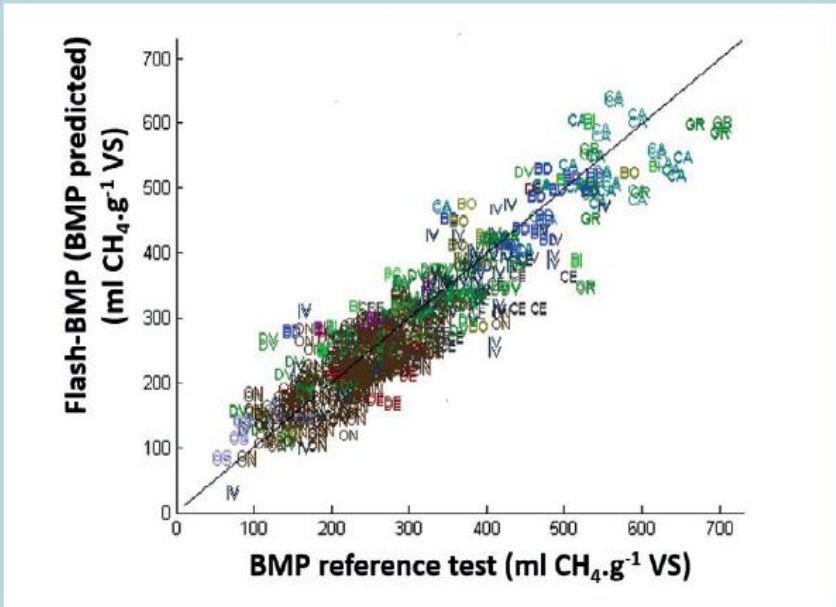
- Inter-instrument transfer
 - > Source instrument: Calibration developed on NIRFlex N-500 (Buchi)
 - > Target instrument: Calibration transferred to a DS2500 (Metrohm)
- Reduced difference between spectra after spectral standardization



Application example: Flash BMP® calibration transfer

Flash BMP®: Biochemical Methane Potential (BMP)

- **Inter-instrument transfer**
 - > Source instrument: Calibration developed on NIRFlex N-500 (Buchi)
 - > Target instrument: Calibration transferred to a DS2500 (Metrohm)
- **Efficient calibration transfert**
 - > Test on 10 samples for transfer validation
 $R^2_{\text{Buchi}}: 0,95$ vs $R^2_{\text{Metrohm}}: 0,97$
- **Metrohm FlashBMP® model performance**
 - > Similar to the source model
 - > Same number of PLS calibration factors



Package comes with the ready-to-use Flash BMP® calibration model, you can speed up your evaluation with effortless calibration development.



AGENDA

- What are the different layers of Artificial Intelligence?
- Data Analysis - Good Practices
- Machine Learning calibration models
- Calibration transfer
- Data fusion

Prediction models based on data fusion



Raw material quality control: Chocolate discrimination model



A Biancolillo, SPreys, et al. (2021). Multi-block classification of chocolate and cocoa samples into sensory poles, Food Chemistry 340 (2021) 127904.



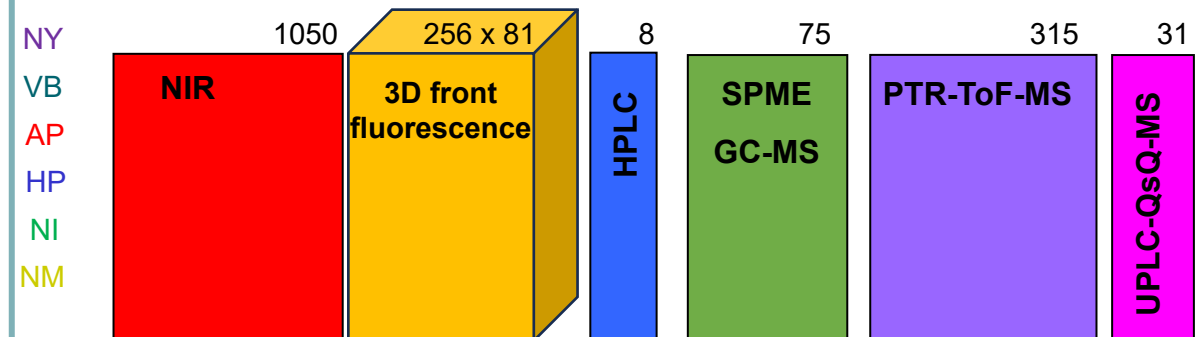
Cocoa beans / Chocolate samples

- > 4 key characteristics of chocolate flavor
- > 97 samples



Several blocks of data

- > NIR spectroscopy
- > 3D front fluorescence
- > Liquid and gas chromatography + Mass spectroscopy

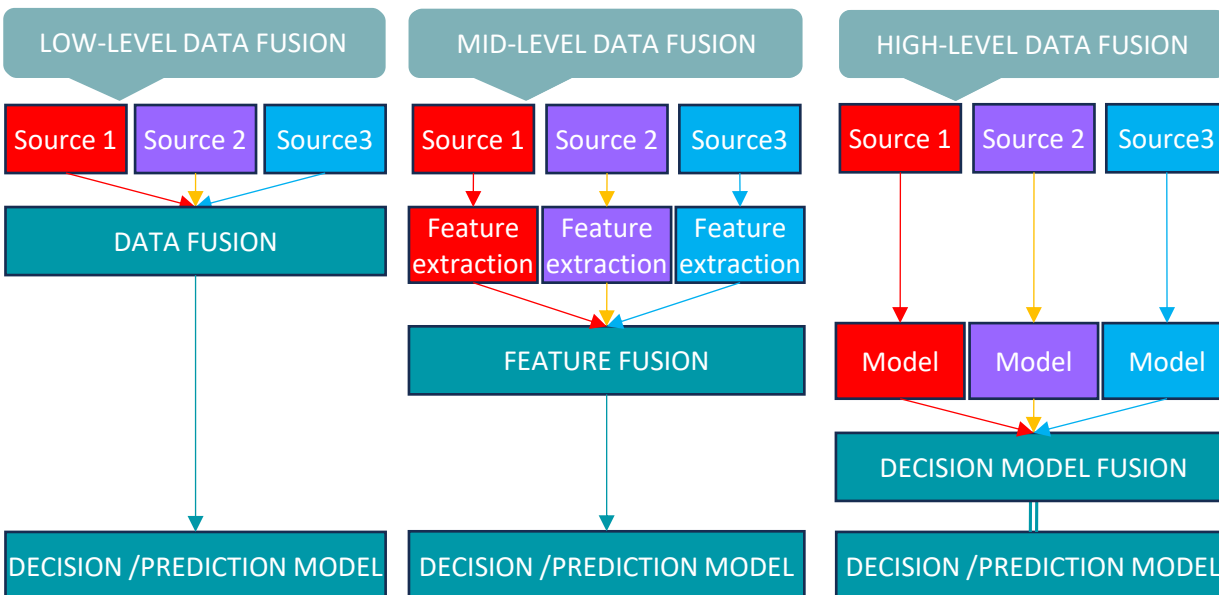


97

Prediction models based on data fusion

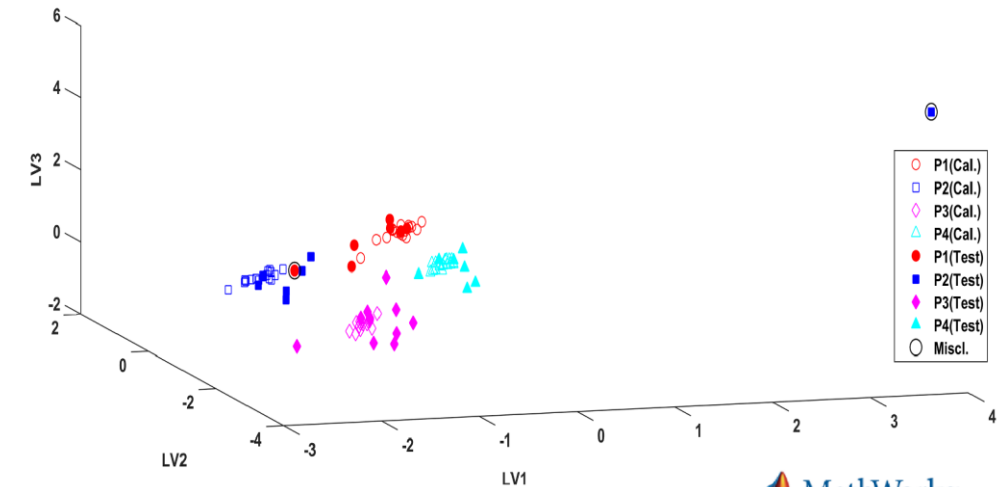


Raw material quality control: Chocolate discrimination model



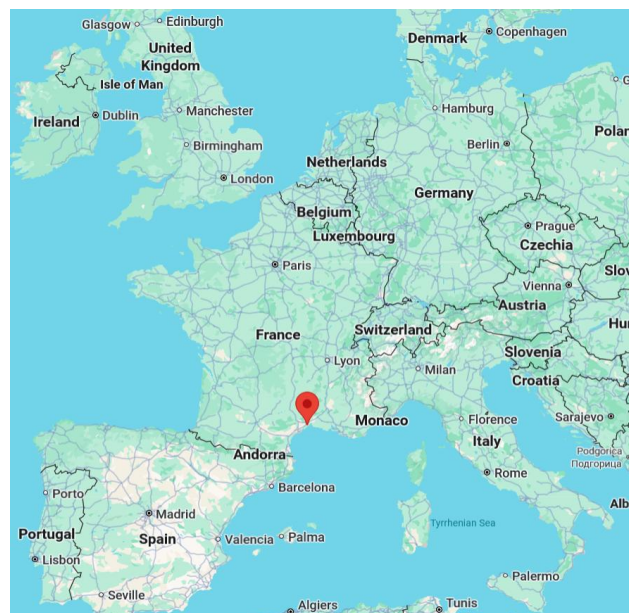
Multiblock (MBK) analysis

- Low-level fusion (SO-PLS)
- Understanding of block contributions
- Outlier detection
- Good results for class discrimination



A Biancolillo, SPReys, et al. (2021). Multi-block classification of chocolate and cocoa samples into sensory poles, Food Chemistry 340 (2021) 127904.

Thank you for your attention!



Ondalys

8 avenue de l'Europe
34 830 Clapiers
France

www.ondalys.fr

T: +33 4 67 67 97 87
contact@ondalys.fr



Dr. Sylvie ROUSSEL, CEO
sroussel@ondalys.fr

