WITH THE SUPPORT OF

www.optimir.eu

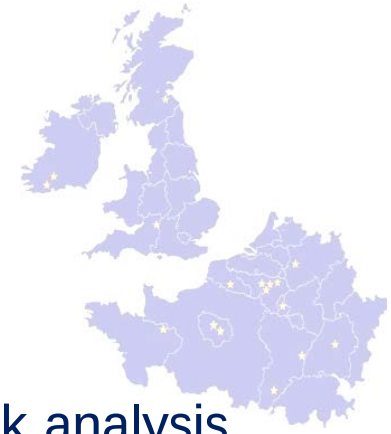
Potential use of mid-infrared milk spectrum in pregnancy diagnosis of dairy cows

A. Lainé[1], A. Goubau[1], L. M. Dale[1], H. Bel Mabrouk[1], H. Hammami[1,2], N. Gengler[1]

[1]University of Liège, Gembloux Agro-Bio Tech, Gembloux, Belgium
[2]National Fund for Scientific Research, Bruxelles, Belgium

64th EAAP Annual Meeting, Nantes, France, August 26th – 30th 2013

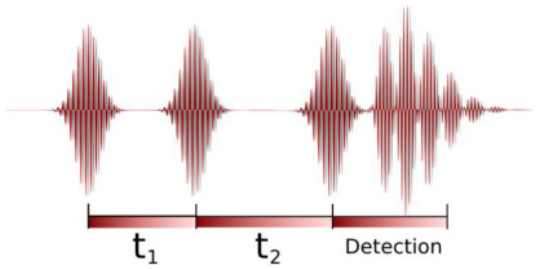
Context



- OptiMIR project:
  - 17 European partners → Common database
  - Milk recording organizations, research centers, milk analysis laboratory

"New tools for a more sustainable dairy sector"

- Based on mid-infrared spectral information from milk



www.optimir.eu

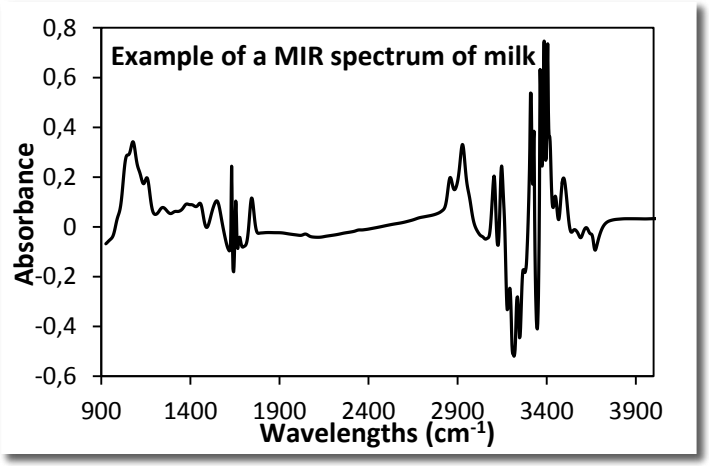
- Fertility
- Feeding
- Health
- Rejection of pollutants
- Milk quality

Context

Milk recording

Mid-Infrared Spectroscopy (MIR)

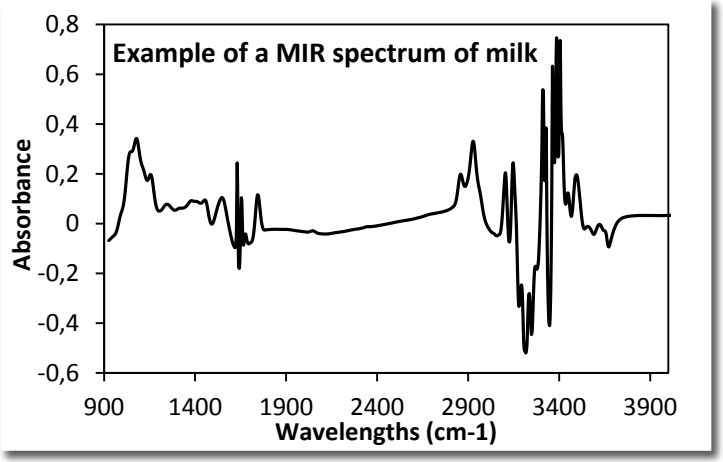
Spectral database (>5,000,000 spectra)



Major components: Fat, protein, lactose, urea

Fine milk composition: Fatty acids, minerals, lactoferrin, ...

Objectives



Direct response from the MIR spectrum about the animal status

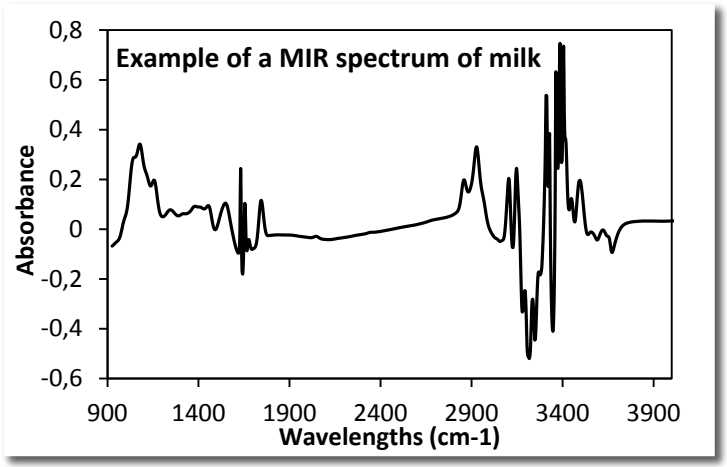
- Identification of a spectrum coming from a pregnant cow or an open cow → Pregnancy Diagnosis
  - Important cost for the dairy sector
  - Milk recording organizations

Does the observed MIR spectrum belong to a pregnant cow or not ?

Principles

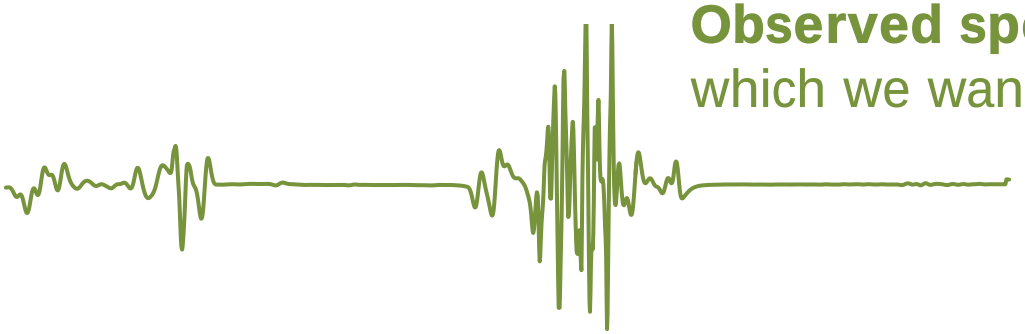
- Many factors influence the shape of the milk MIR spectra:
  - Days in Milk, Parity, Breed, Farm management, ...

➔ *How to observe differences in spectra due to the pregnancy ?*



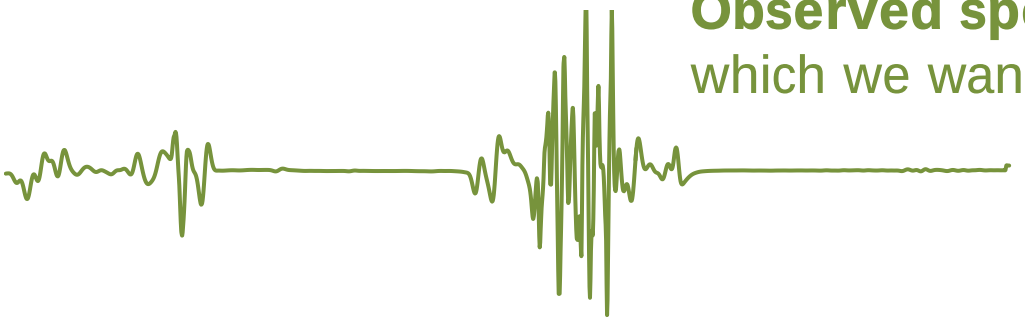
- Literature examples :
  - Sloth et al. 2003: Adjustment of milk parameters on a subset of healthy samples applied on a whole dataset (healthy and not) to assess udder health from milk samples
  - Staib et al. 2001: Diagnosis of rheumatoid arthritis with discriminant analysis on human blood IR spectra

Principles



Observed spectrum = Milk sample on which we want to test the pregnancy

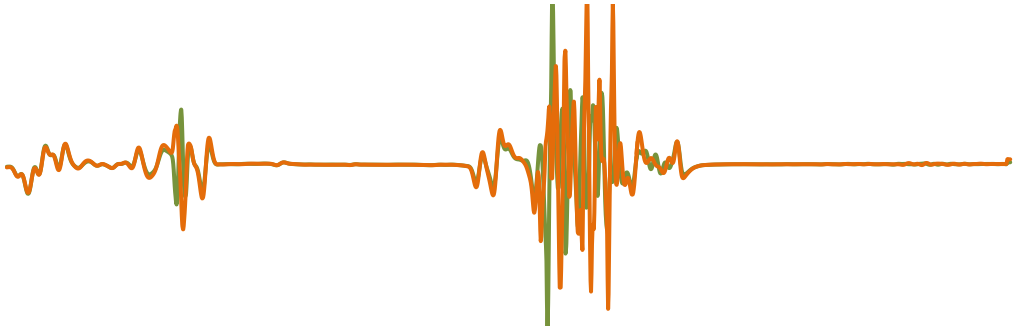
Principles



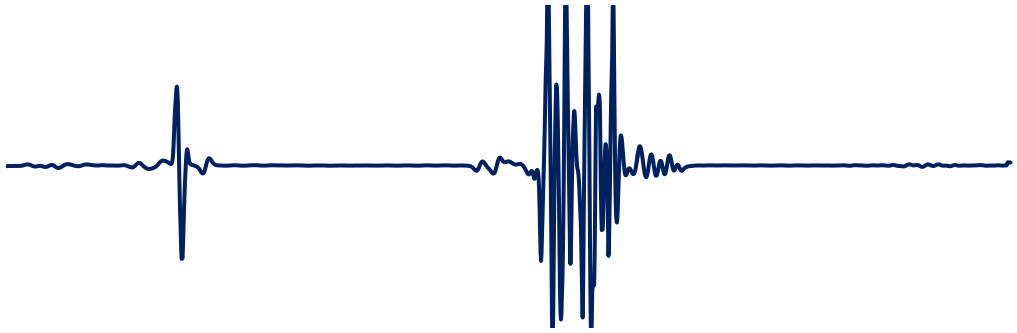
Observed spectrum = Milk sample on which we want to test the pregnancy

Expected spectrum = Expected spectrum for the same day in milk if the animal was not pregnant

Principles



Residual spectrum =
Observed spectrum – Expected spectrum



Reproductive status
Unaccounted factors
Errors

Residual spectra are used to perform discrimination between two groups of classification (Pregnant cow and non-pregnant cow)

Data

- Dataset from Walloon Region of Belgium
  - 388,951 observations = spectra
  - TD from January 2010 to December 2012 → 3 years
  - Only lactations that started within the period
  - At least 1 observation per animal x lactation for which the cow is open
- Pre-processing of MIR spectra
  - First derivative: Set all spectra to a common baseline
  - Informative area: Avoid noises and non-useful area
- Modeling an expected spectrum which is based on history of the animal
  - Animal, parity, breed, days in milk, ...
  - Modeling based on a subset of non-pregnant data: 197,109 spectra

Data

- Discriminant analysis
  - Groups of classification: Pregnant and Non-pregnant
  - Predictors: Residual spectral points
- Training set = construction of the discriminant equation
  - TD from January 2010 to December 2011 → 2 years
  - Maximum 120 days after the insemination
  - 217,148 observations (36.6% pregnant & 63.4% non-pregnant)
- Validation = test data for applying the discriminant equation
  - TD from January 2012 to December 2012 → 1 year
  - TD from lactations that started after 1st January 2012
  - Maximum 120 days after the insemination
  - 51,109 observations (15.0% pregnant & 85.0% non-pregnant)

Results

Specificity = 96.8% (NP observations correctly classified)

Sensibility = 77.0% (P observations correctly classified)

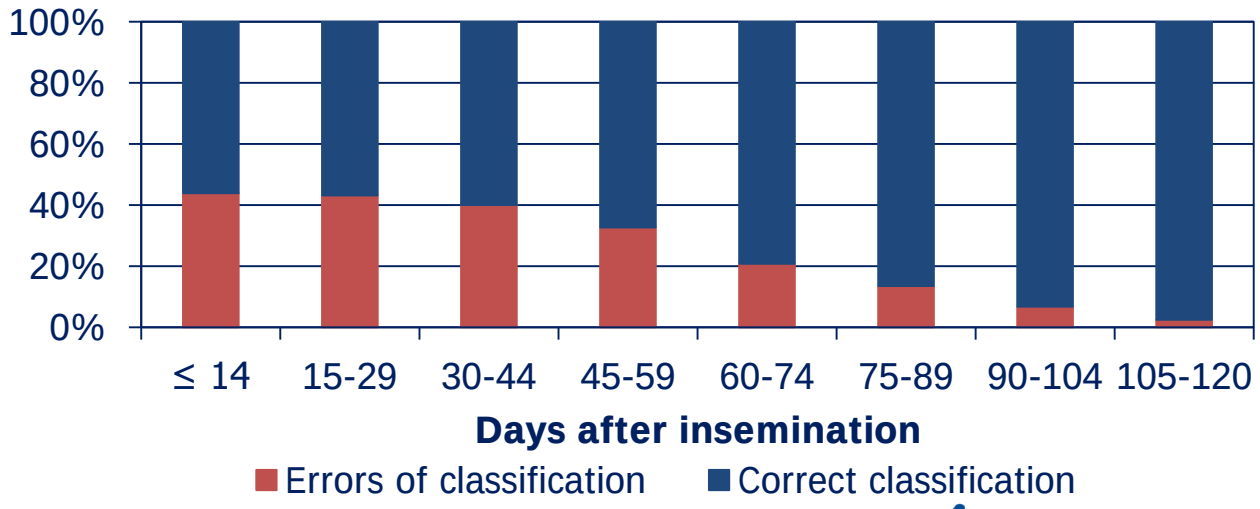
- Average error: 13.1%

Results

Specificity = 96.8% (NP observations correctly classified)

Sensibility = 77.0% (P observations correctly classified)

- Average error: 13.1%
- Classification from validation dataset
  - Distribution by classes of 15 days after insemination
  - Decrease of errors



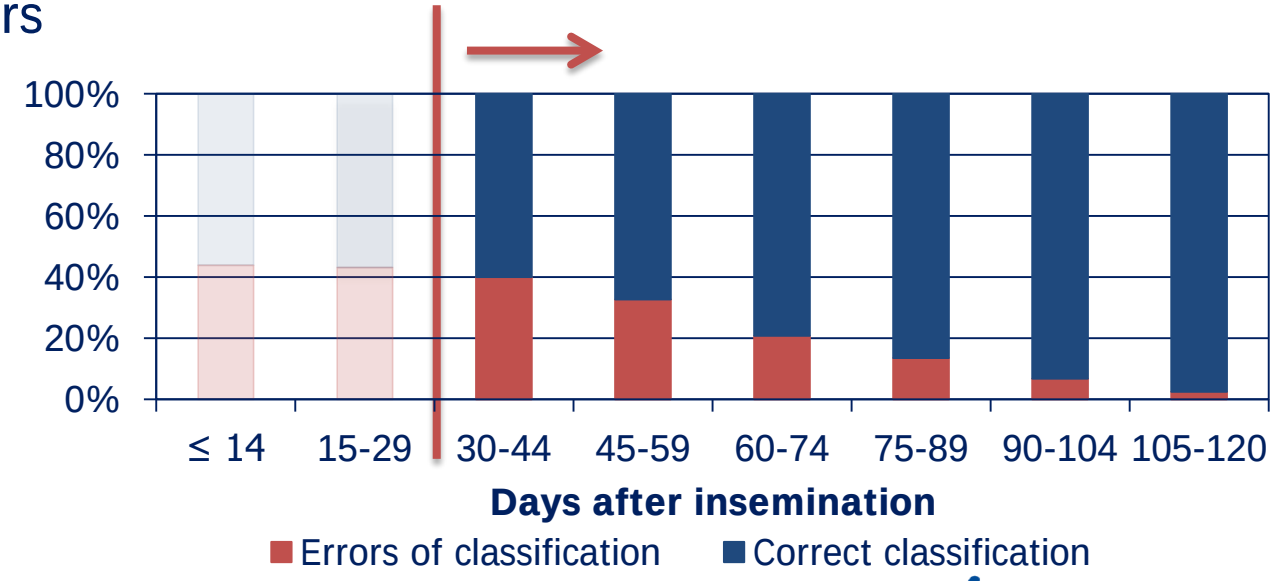
Results

Specificity = 96.8% (NP observations correctly classified)

Sensibility = 77.0% (P observations correctly classified)

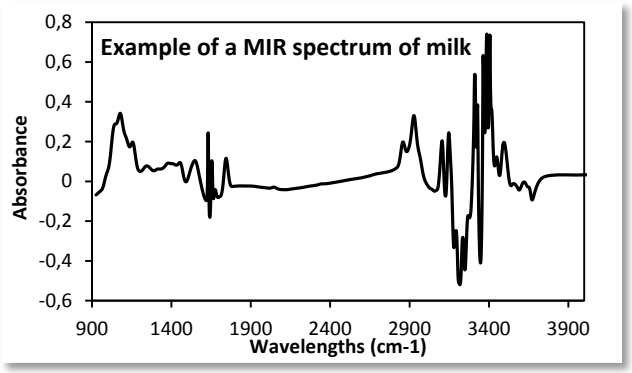
- Average error: 13.1%
- Classification from validation dataset
  - Distribution by classes of 15 days after insemination
  - Decrease of errors

What is the good answer for dairy producers?



Conclusion

- Direct use of the MIR spectra
  - Cheap
  - Easily transferable
  - MIR spectra already obtained in routine
  - Need an adjustement for other factors



- “ Diagnosis ”
  - Information about a cow status
  - Usefull as a warning for dairy producers

Next steps

- Still under development
  - How are the errors distributed among data?
  - Other options than discriminant function?
  - ...
- Optimisation and validation in the field
  - Test in pilot farms in the Walloon Region
- Development of the tool
  - Milk recording organizations involved in OptiMIR
  - Opportunity to use the same approach for mastitis detection or other metabolic disorders

WITH THE SUPPORT OF

www.optimir.eu

Acknowledgments

Service Public of Wallonia SPW – DGO3

European Commission (ERDF) through project Interreg IVb OptiMIR

Author's contact e-mail:

aurelie.laine@ulg.ac.be

Results - Cross-validation (Leave-one-out)

→ Residual spectra

Specificity = 95.0% (NP observations correctly classified)

Sensibility = 66.6% (P observations correctly classified)

- Average error: 19.2%

→ Raw spectra (no adjustments)

Specificity = 42.8 % (NP observations correctly classified)

Sensibility = 95.4 % (P observations correctly classified)

- Average error: 30.1 %

Results - External validation

→ Residual spectra

Specificity = 96.8% (NP observations correctly classified)

Sensibility = 77.0% (P observations correctly classified)

- Average error: 13.1%

→ Raw spectra (no adjustments)

Specificity = 99.1% (NP observations correctly classified)

Sensibility = 0.3% (P observations correctly classified)

- Average error: 50.3%

Principles

Expected spectrum = Expected spectrum for the same day in milk if the animal was not pregnant

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\varepsilon}$$

$\boldsymbol{y}$ = Vector of observations (spectral points)
$\boldsymbol{\beta}$ = Fixed effects
$\boldsymbol{\gamma}$ = Random effects
$\boldsymbol{\varepsilon}$ = Residual errors
$\boldsymbol{X}$ and $\boldsymbol{Z}$ = Incidence matrix

Mixed model on a subset of non pregnant data !

Solutions applied on the whole dataset to obtain all the expected spectrum

Principles

- Modeling an expected spectra which is based on history of the animal
  - Fixed effects: *parity*, *breed*, *month of TD*
  - Regression coefficients: *DIM* and *DIM²*
  - Random effects: *Cow x lactation*
  - Random regression coefficients: *DIM x cow x lactation* and *DIM² x cow x lactation*
- Subset of non-pregnant data: 197,109 spectra