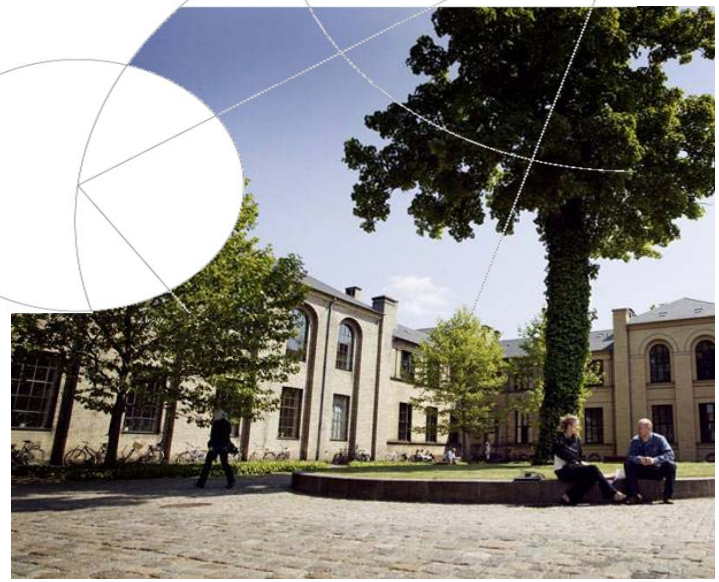
# Contaminant profiling: A set of next generation analytical tools to deal with contaminant complexity

Associate Professor

Jan H Christensen

Department of Basic
Sciences and Environment



# Risk Assessment

How can we improve current risk assessment procedures?

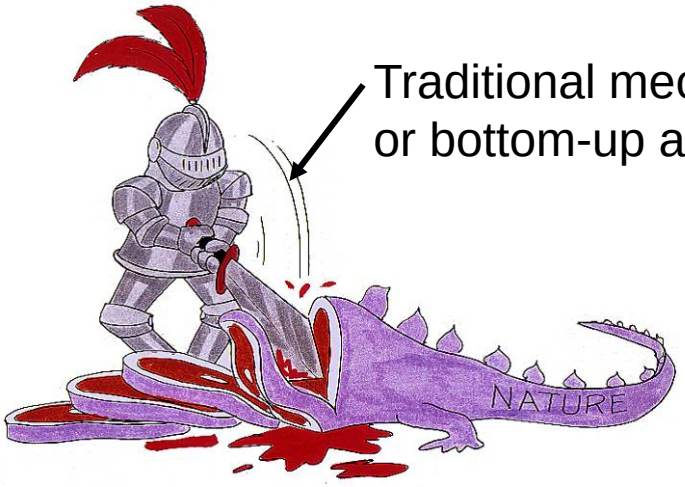
(1) Measure the entire contaminant complexity (‘the chemical fingerprint’)

(2) Investigate combination effects / Cocktail effects

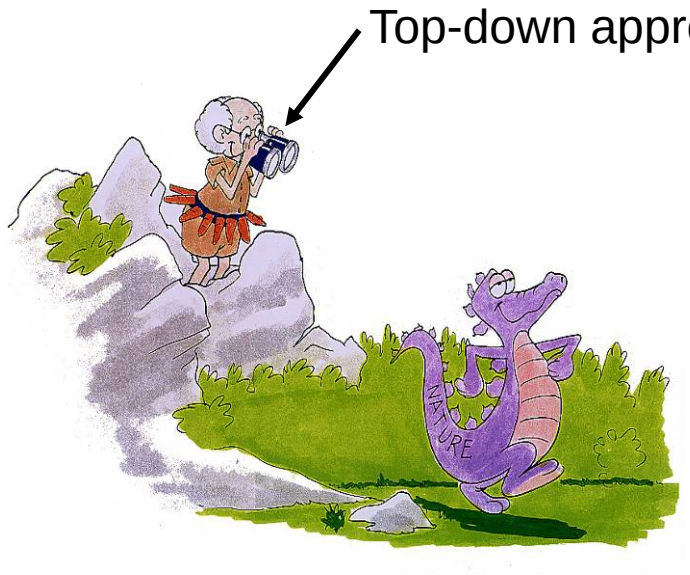
Comprehensive investigation of single compounds

Chemical analysis and effect studies

# A paradigm shift in (environmental) analytical chemistry?

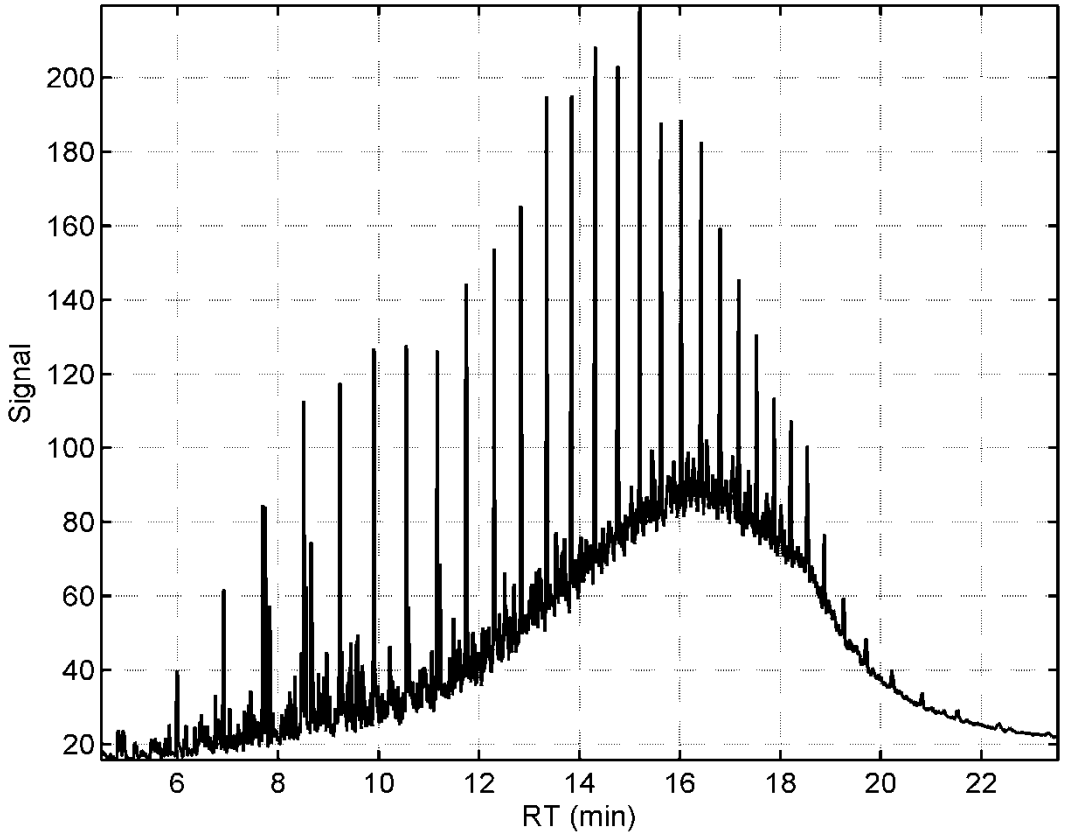



- Extensive sample preparation (to isolate)
- Chemical target analysis
- Fully quantitative analysis
- Extrapolation to complex systems

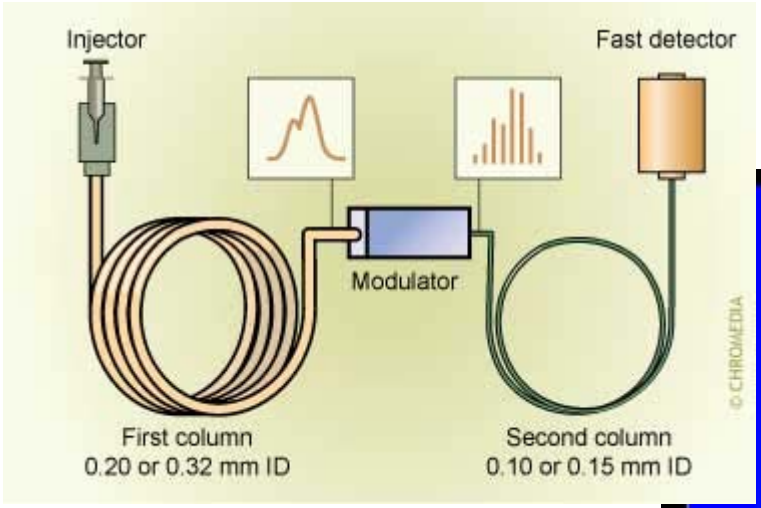


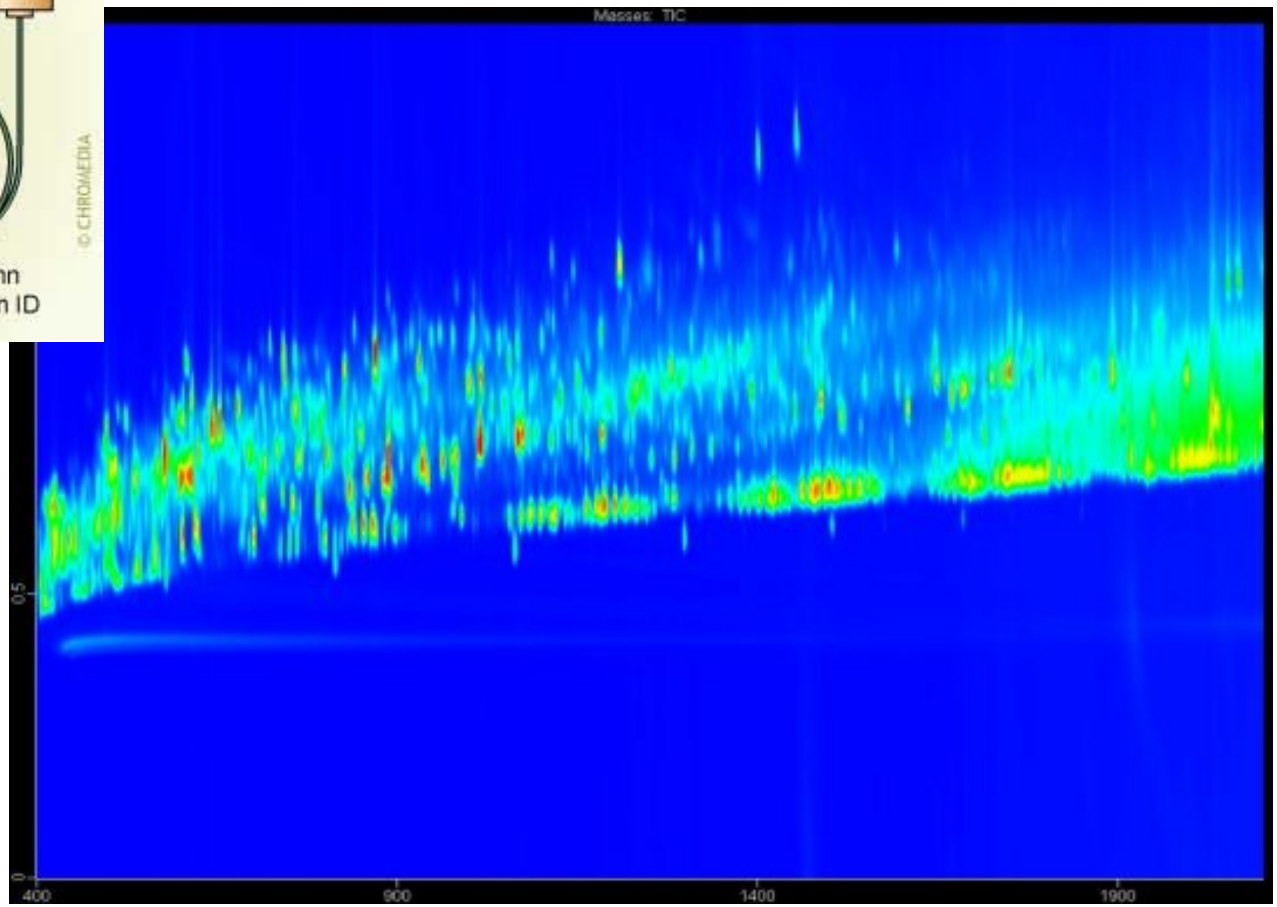

- Limited sample preparation
- Chemical profiling analysis
- (semi)quantitative analysis or patterns (fingerprints)
- Identification of relevant compounds

# Analytical data are multivariate



# …a trend that is increasing



# Human pattern recognition

Kevin Thomas

Barack Obama

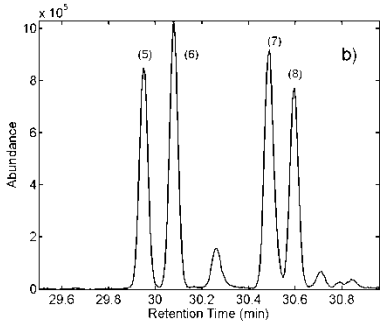
# Human pattern recognition

# Some mathematics

**+**

**= ”It is Elvis and not Berlusconi”**

**eyes + Brain**




**+**

**= ”It is oil pollution”**

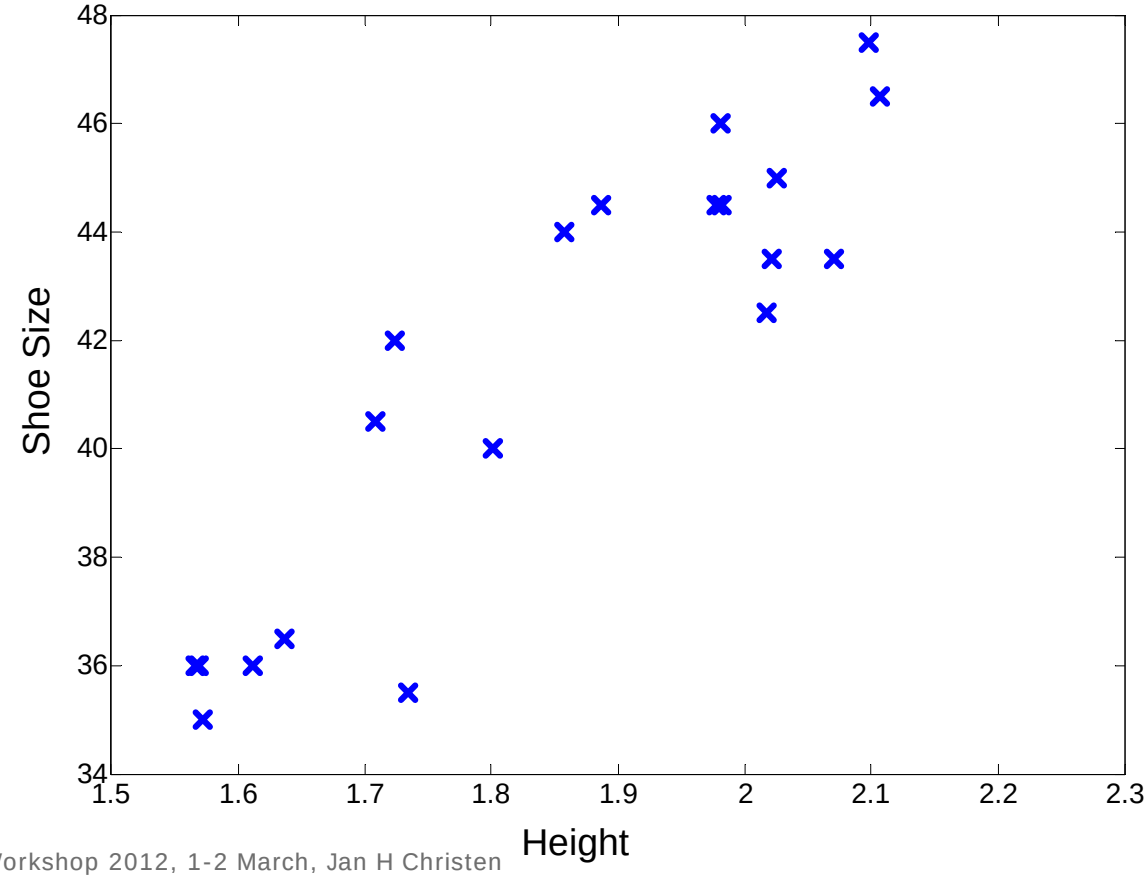
**Sensor + Computer with ’multivariate’ software**

# PCA – the most common multivariate method

Height and shoe size of 19 persons.

Each sample (person) can be described by two variables (height and shoe size)

| Height | Shoe size |
|---|---|
| 1.98 | 44.5 |
| 1.57 | 36.0 |
| 2.03 | 45.0 |
| 2.11 | 46.5 |
| 2.02 | 42.5 |
| 2.07 | 43.5 |
| 1.98 | 46.0 |
| 1.72 | 42.0 |
| 1.86 | 44.0 |
| 1.64 | 36.5 |
| 2.02 | 43.5 |
| 1.57 | 36.0 |
| 1.80 | 40.0 |
| 1.57 | 35.0 |
| 1.73 | 35.5 |
| 1.98 | 44.5 |
| 1.89 | 44.5 |
| 1.71 | 40.5 |
| 2.10 | 47.5 |

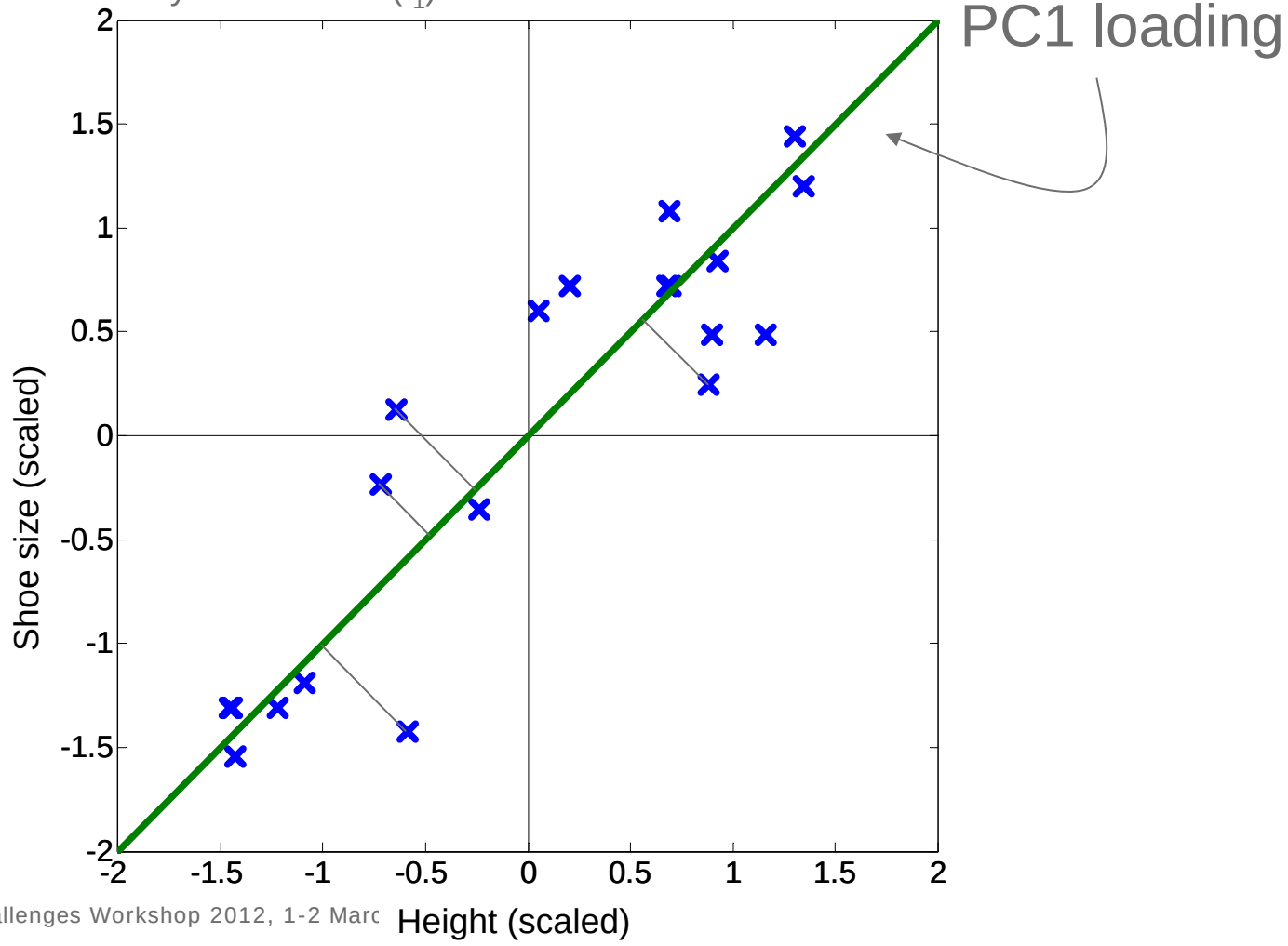


# PCA – projection (Scores)

The first score ($t_1$) is found by projection of the original points onto the 1. loading (PC1 loading).

Now each sample is represented by one number ($t_1$). That's a model!

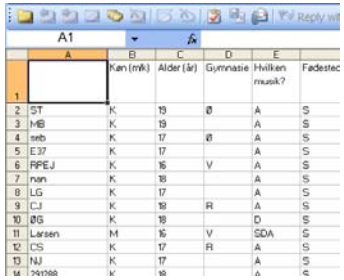
| $t_1$ |
|---|
| 0.9895 |
| -1.9458 |
| 1.2485 |
| 1.7989 |
| 0.7963 |
| 1.1601 |
| 1.2533 |
| -0.3647 |
| 0.4633 |
| -1.6113 |
| 0.9784 |
| -1.9535 |
| -0.4184 |
| -2.1039 |
| -1.4234 |
| 1.0027 |
| 0.6545 |
| -0.6732 |
| 1.9366 |

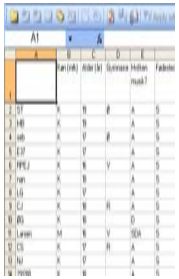


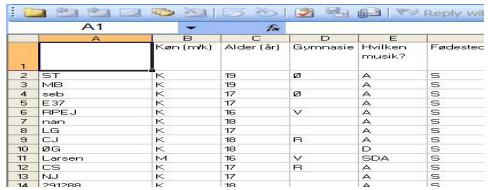
# Principal Component Analysis (PCA)

$$X = TP'$$

**One Excel sheet =** **Two new Excel sheets**



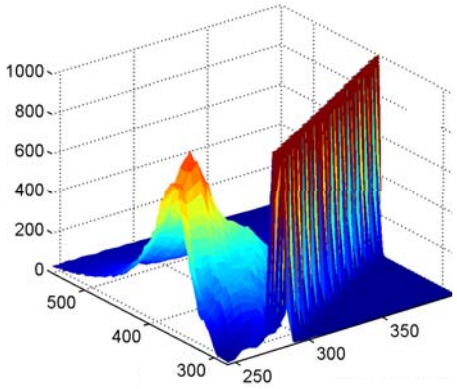


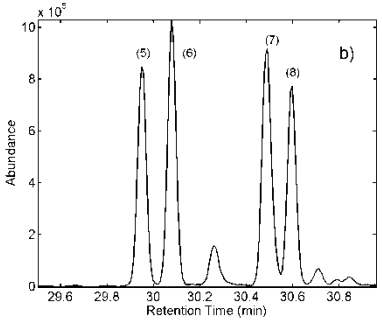


**Scores:**
**Info about samples**

**Loadings:**
**Info about measurements**

# Sensors: Analytical data




## 1. Spectroscopy

- Ultraviolet-visual
- Near-infrared, FT-Infrared, Raman
- Fluorescence (excitation or emission spectra)
- Fluorescence (time-resolved, excitation-emission landscapes)
- Nuclear Magnetic Resonance

## 2. Mass spectrometry

- Ionization (Desorption, spray and gas phase methods)
- Mass analyzers (nominal vs. accurate mass)

## 3. Chromatography

- Liquid chromatography (LC, HPLC, UPLC)
- Gas chromatography (GC)
- Electrophoresis (migration)
- Univariate detectors (ECD, TCD, sulphur, FID etc)

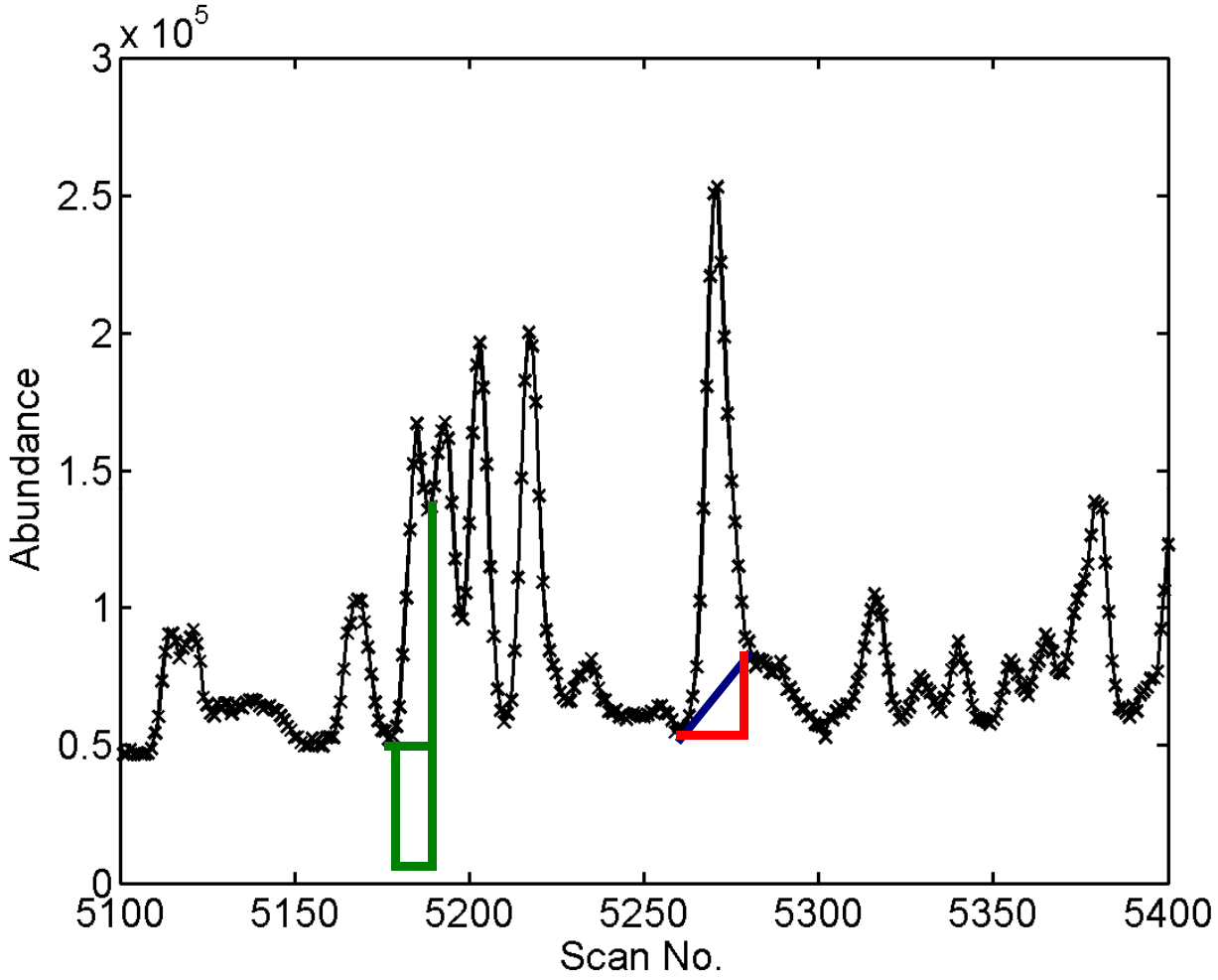
## 4. Hyphenated techniques

- Liquid chromatography – spectroscopy (fluorescence and UV/VIS)
- Gas or liquid chromatography with mass spectrometry detection (LC-MS, GC-MS)
- Two-dimensional chromatography (GC×GC, LC×LC)

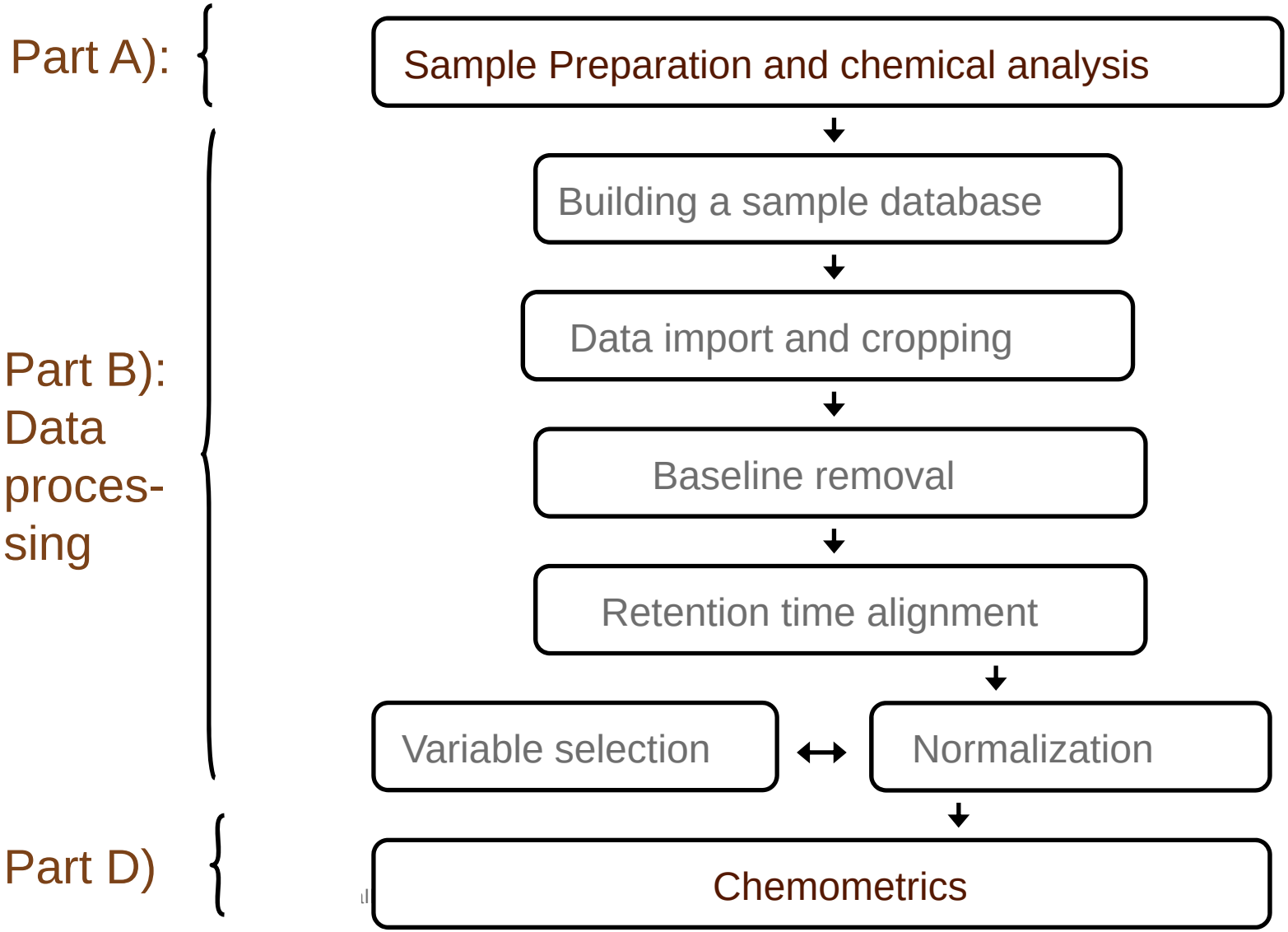
# Example 1: Oil hydrocarbon Fingerprinting

Christensen JH, Tomasi G and Hansen AB. Chemical Fingerprinting of Petroleum Biomarkers using Time Warping and PCA, *Environmental Science and Technology*, **2005**, 39 (1), 255-260.

# Conventional methods rely on peak integration

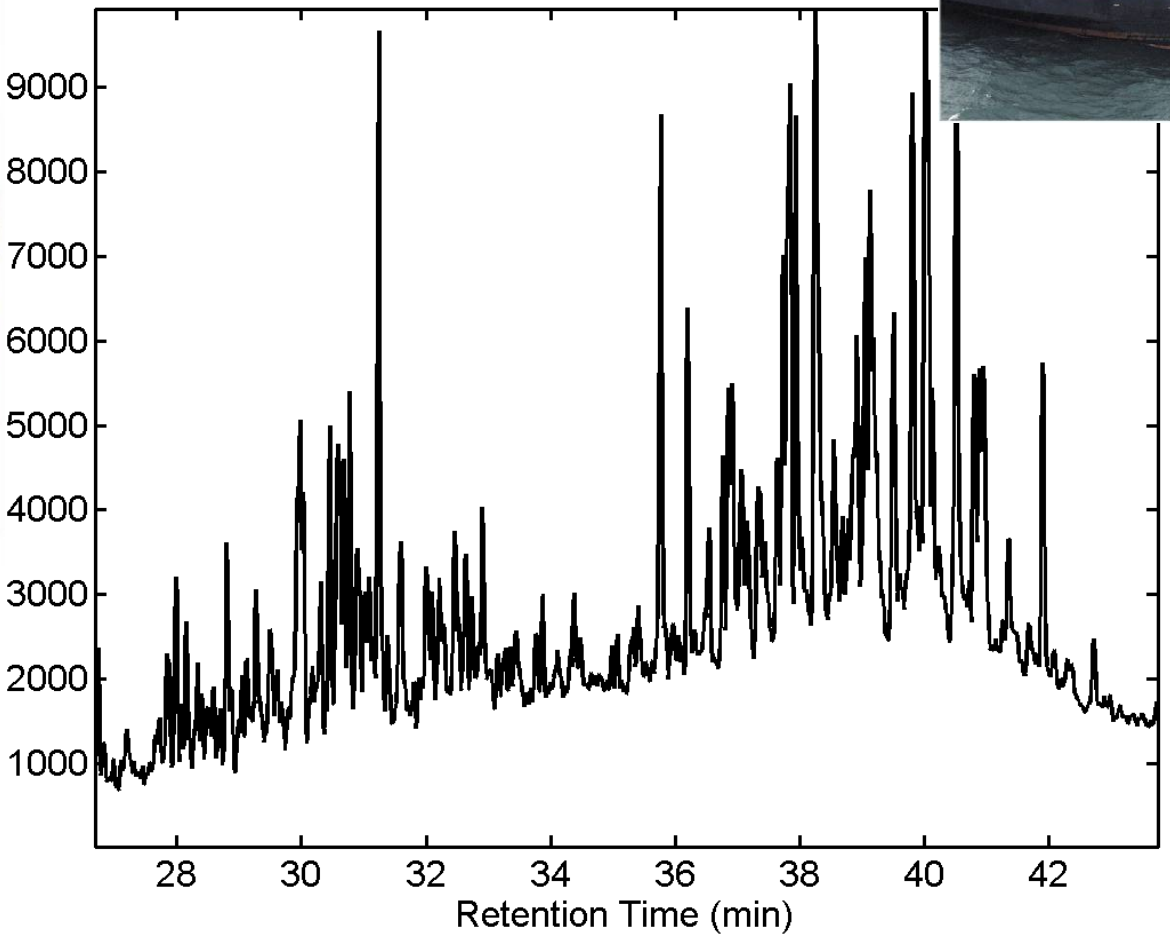


# The CHEMSIC method (How to avoid peak integration)



# PART B: Data pre-processing

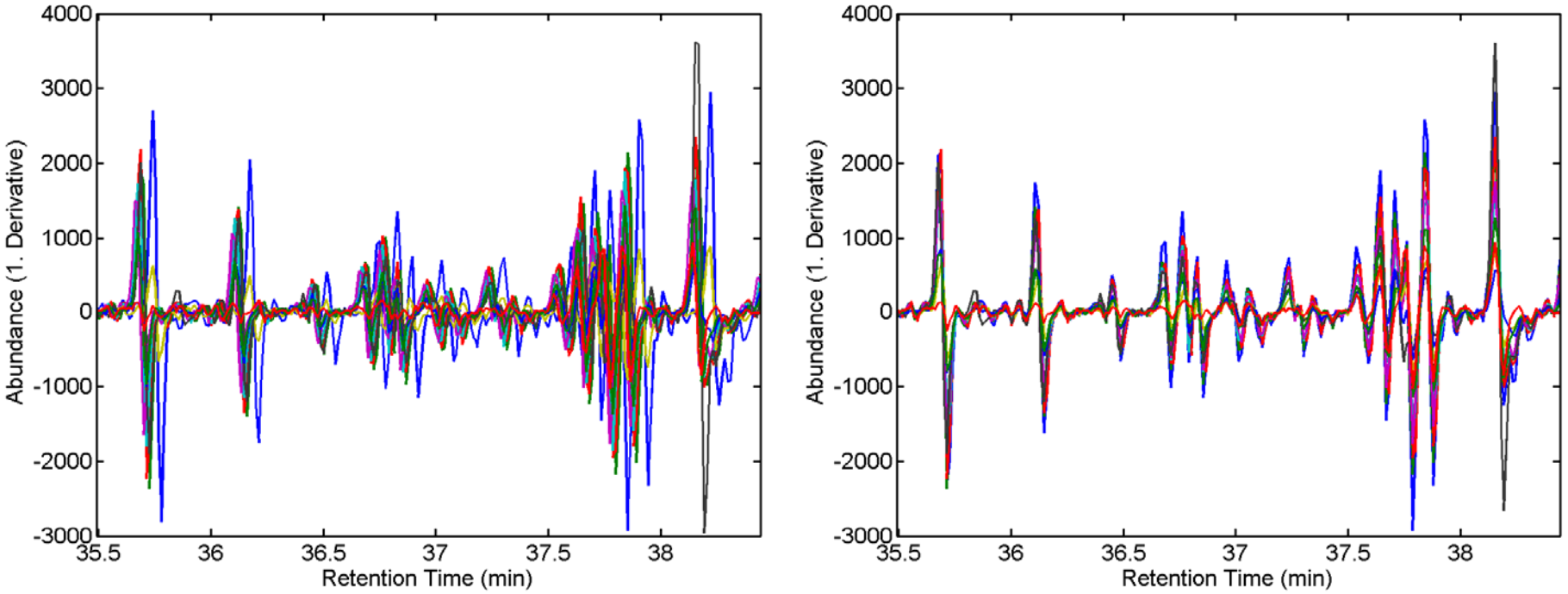
## Step I: Baseline removal



# PART B: Data pre-processing

## Step II: Retention Time Alignment

The effects of correlation optimized warping (COW) on a chromatographic section of m/z 217 SIC in 10 oil samples



# PART B: Data pre-processing

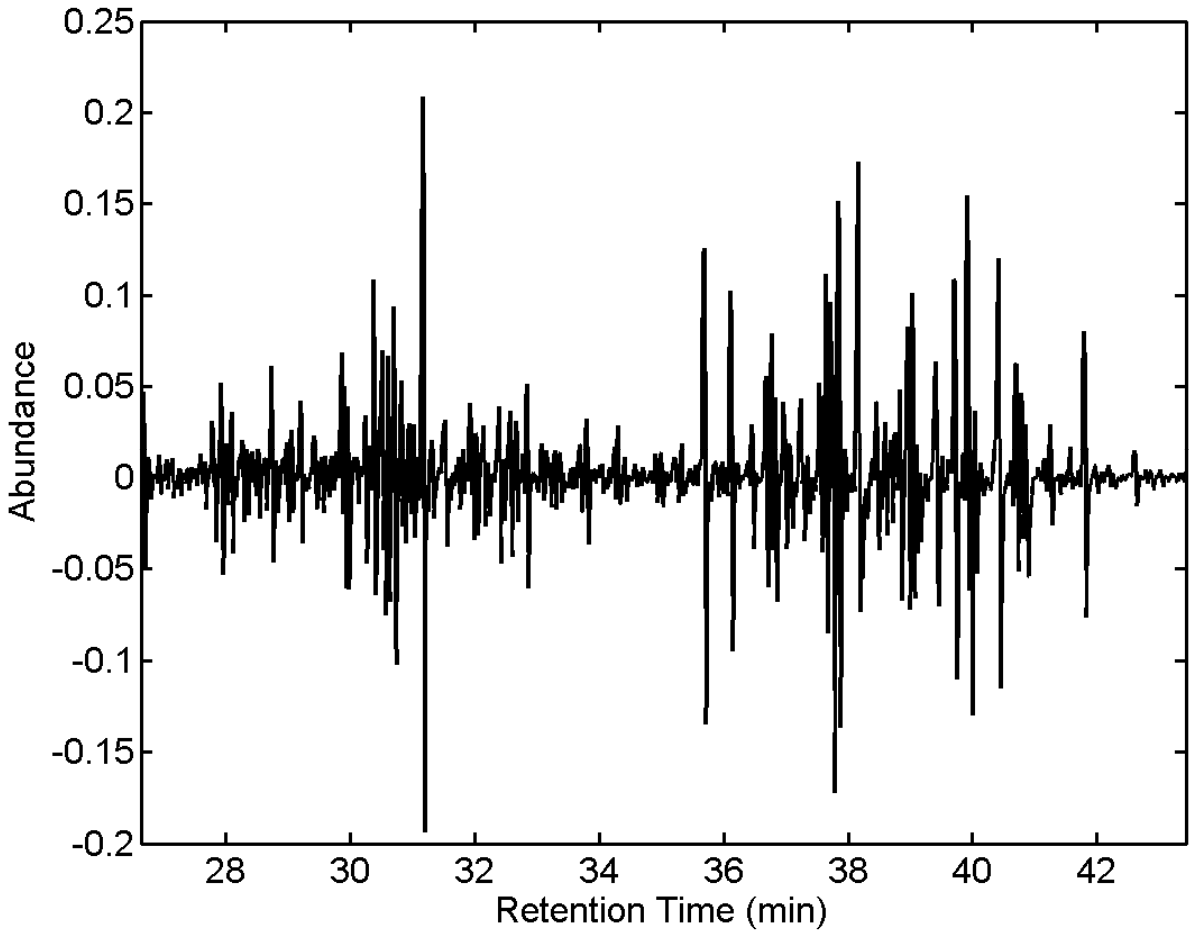
## Step III: Normalization

(I) Internal standards

(II) $$x_{nj}^{N} = \frac{x_{nj}}{\sqrt{\sum_{j=1}^{J} x_{nj}^{2}}}$$

(III) Normalization using selected data points

- Less affected by closure
- Less sensitive to noise

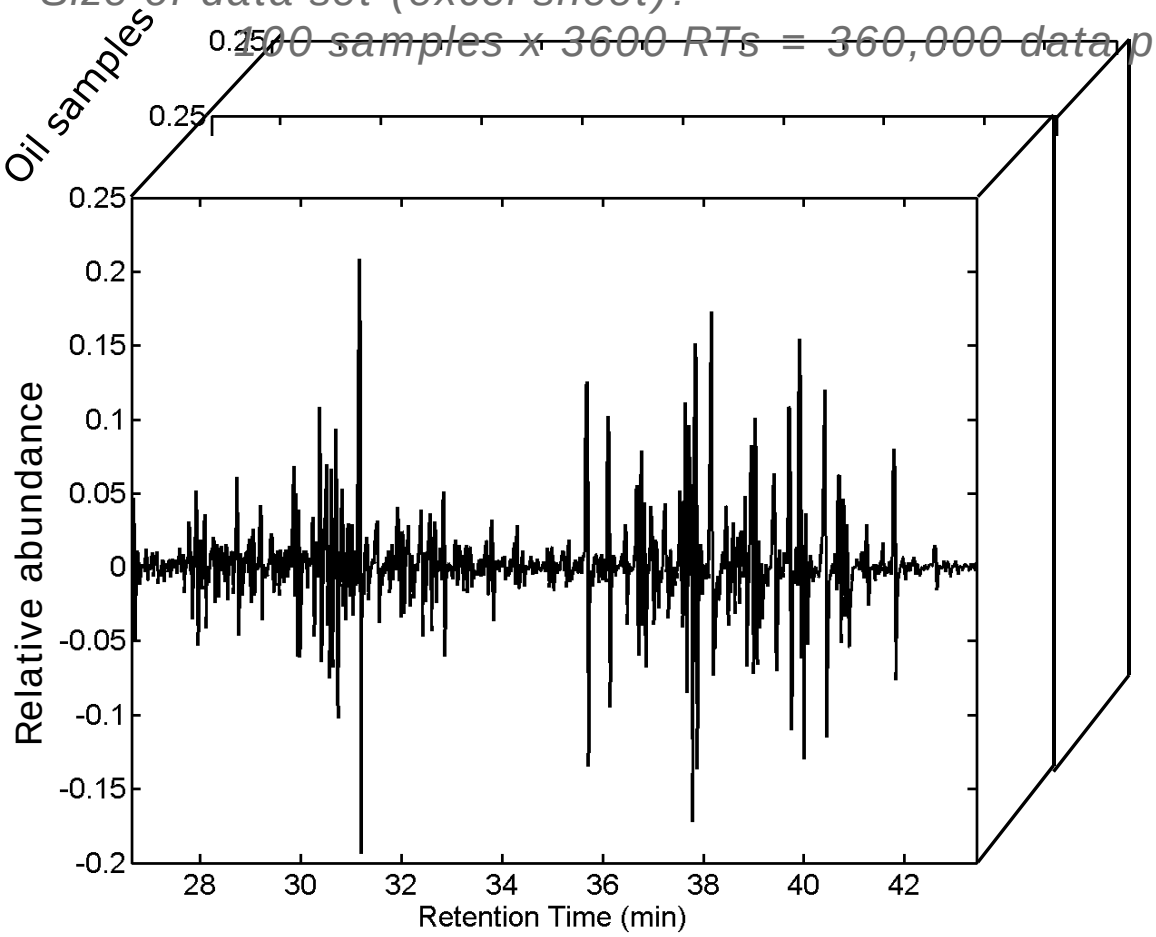


# PART C: Data structure

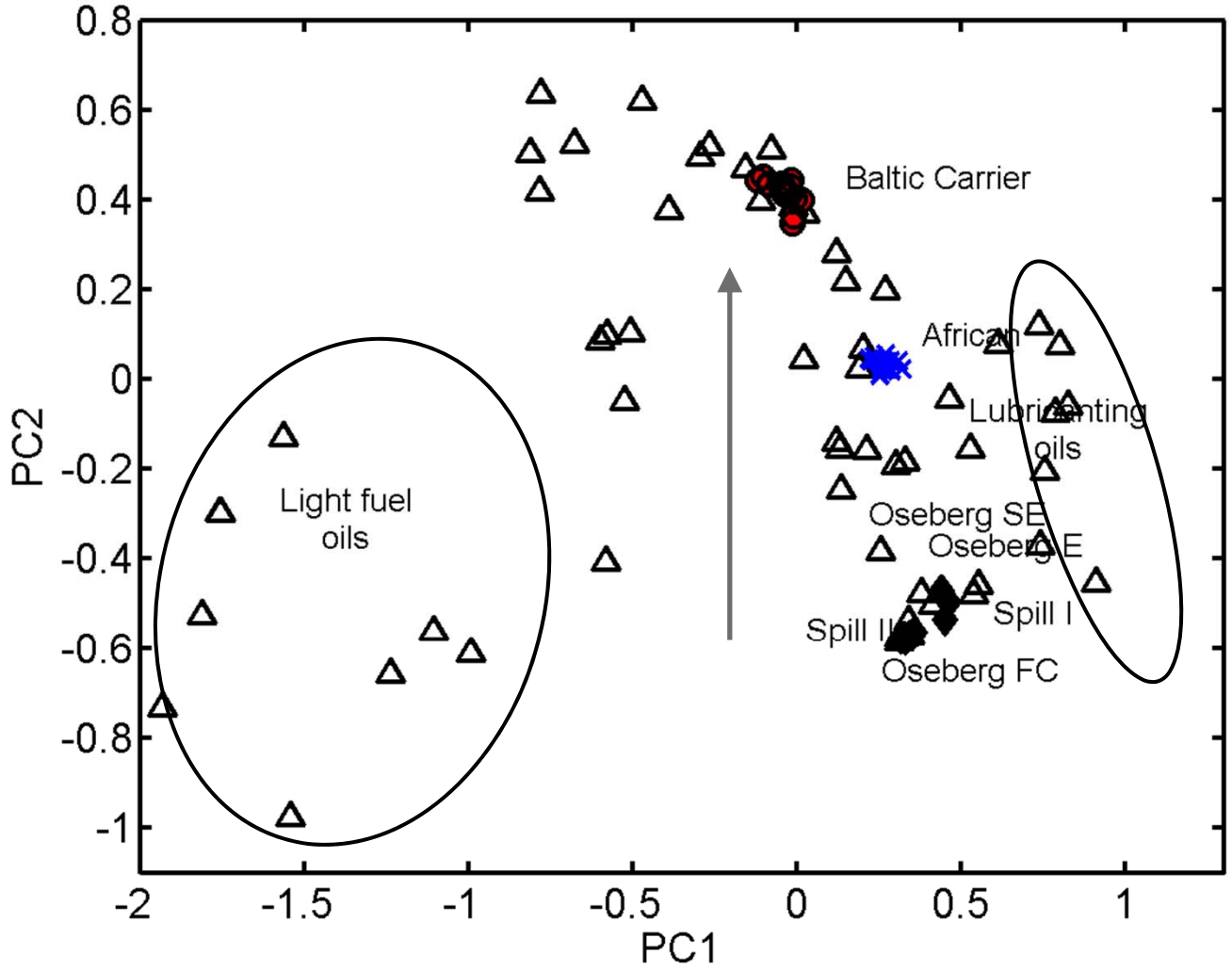
## Chemical profiles of 100 crude oils or refined produts

*Size of data set (excel sheet):*

*100 samples x 3600 RTs = 360,000 data points*



# PART C: Map of samples (the score plots)



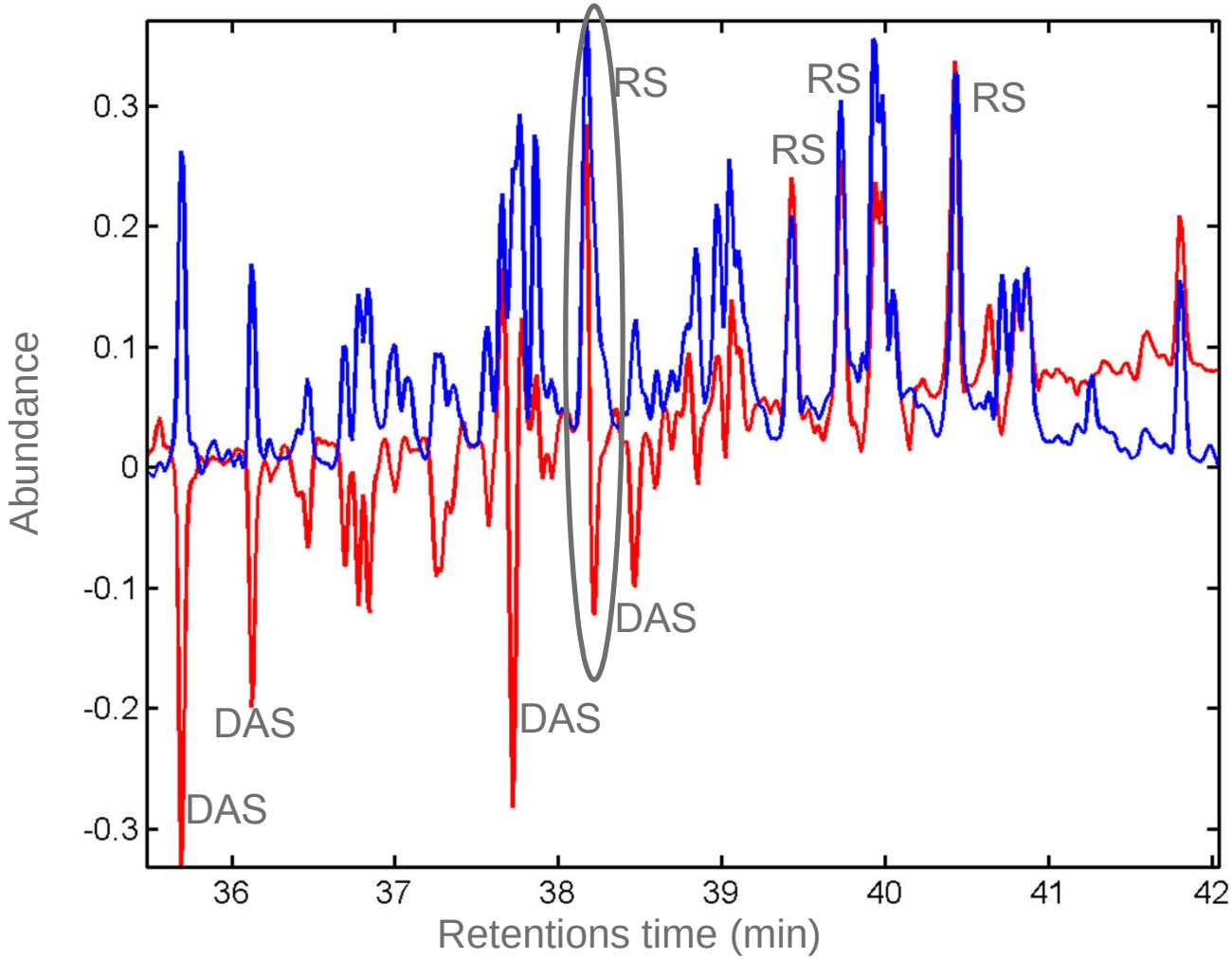
# PART C: Chemical interpretation (PC2 loadings)

...Why are the oils different?

Distinguish oils originating from source rocks with different clay content

High PC2 ⇔
High DAS/RS ⇔
High clay content

Low PC2 ⇔
Low DAS/RS ⇔
Low clay content and anoxic conditions



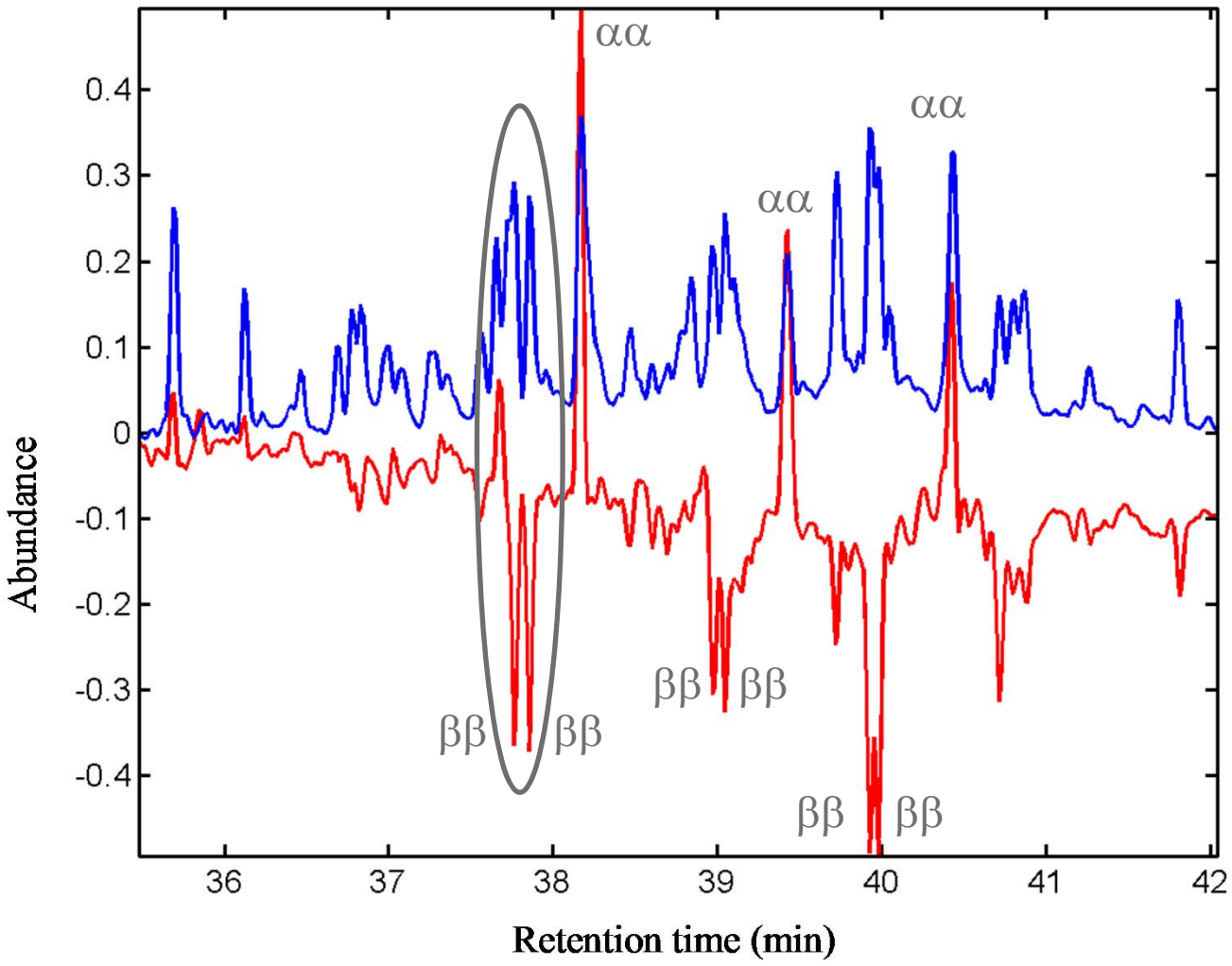
# PART C: Chemical interpretation (PC4 loadings)

## ...Why are the oils <u>also</u> different?

Distinguish oils originating from source rocks with different thermal maturity

High PC4 ⇔
High αα/ββ ratio of $C_{27}$-$C_{29}$ regular steranes ⇔
Low thermal maturity

Seems to be independent of source organic matter input

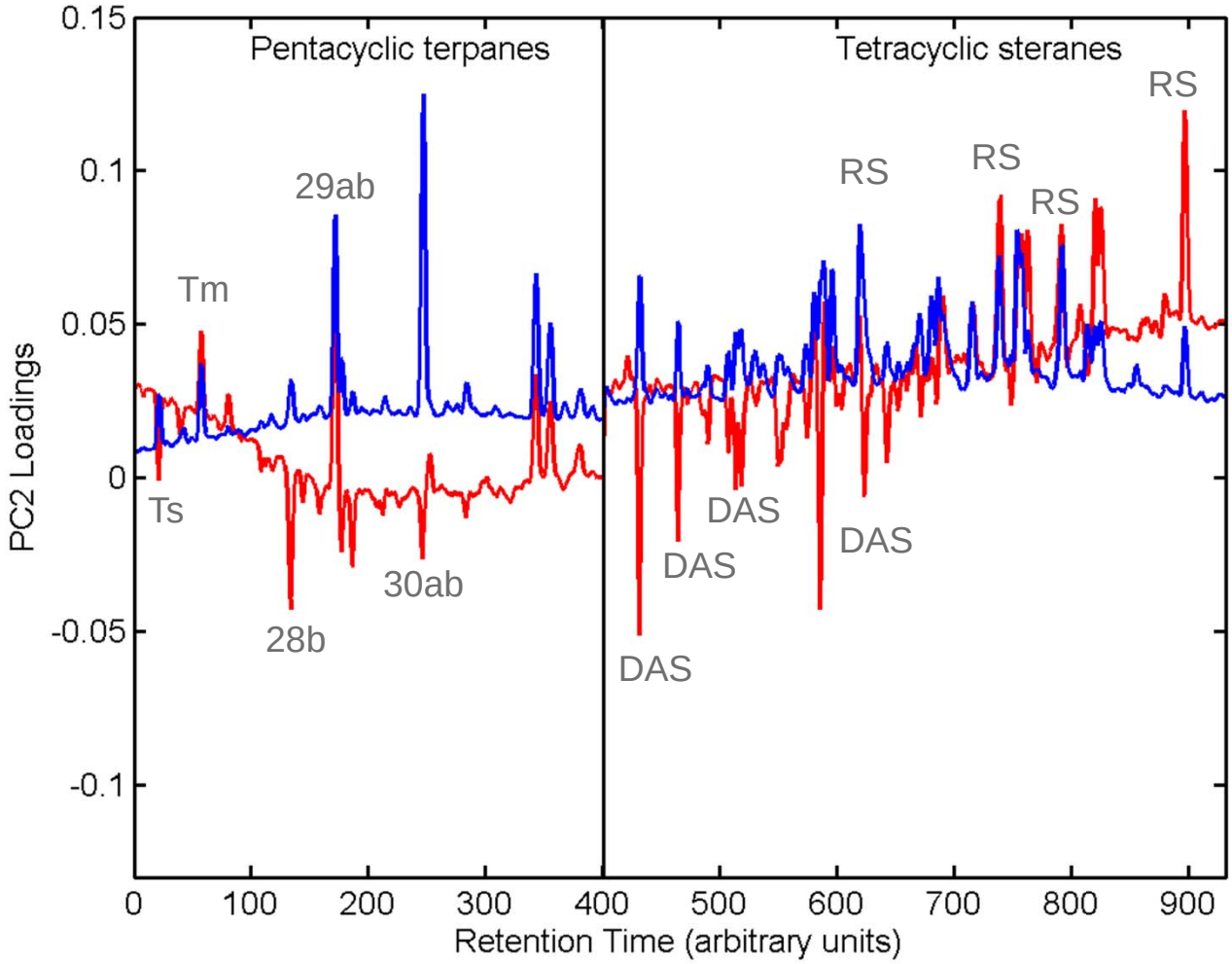


# Modifications to the CHEMSIC method






… adding more SICs (information)

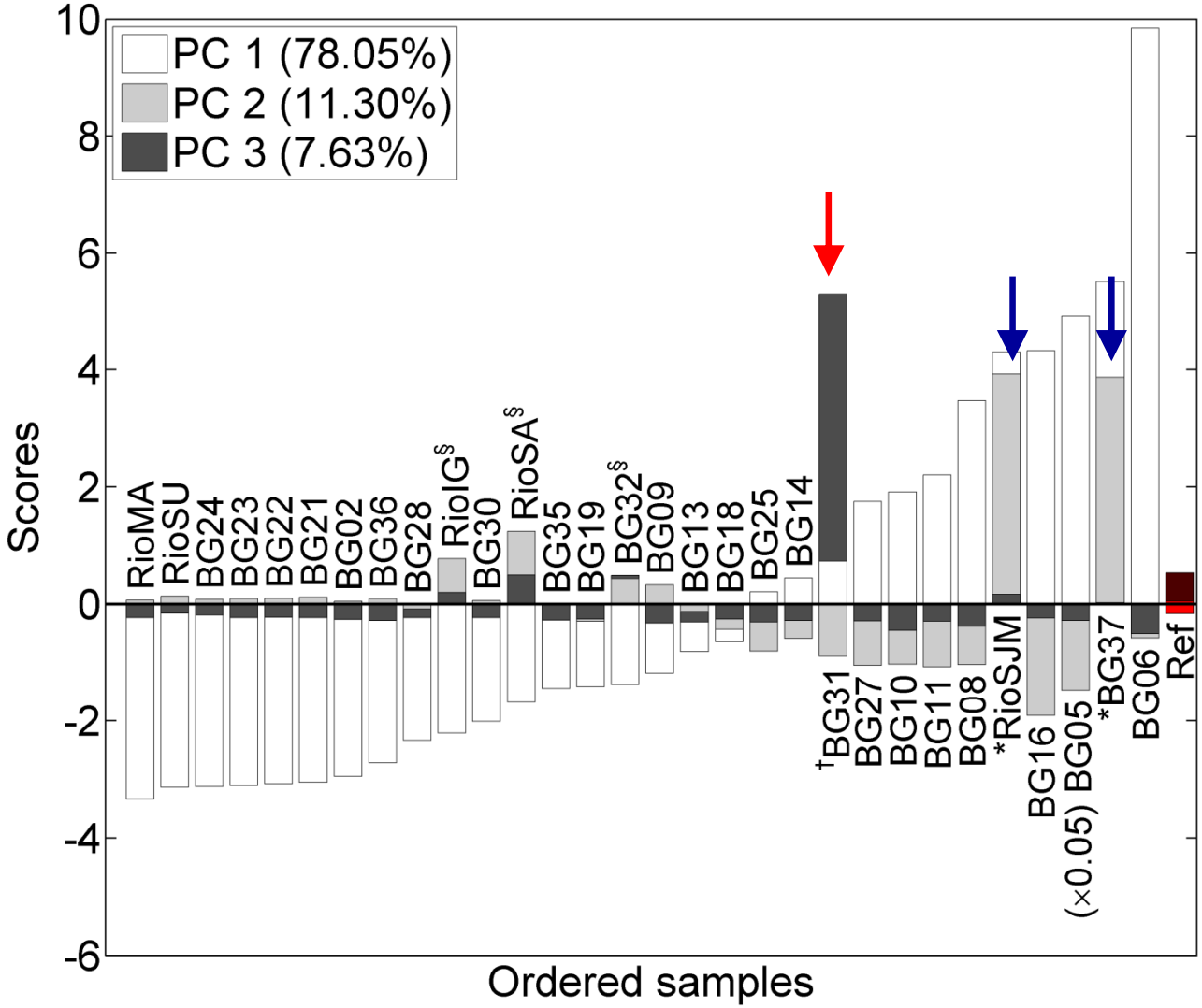


# Example 2: Does the CHEMSIC method also work under warmer climates?

January 2000:

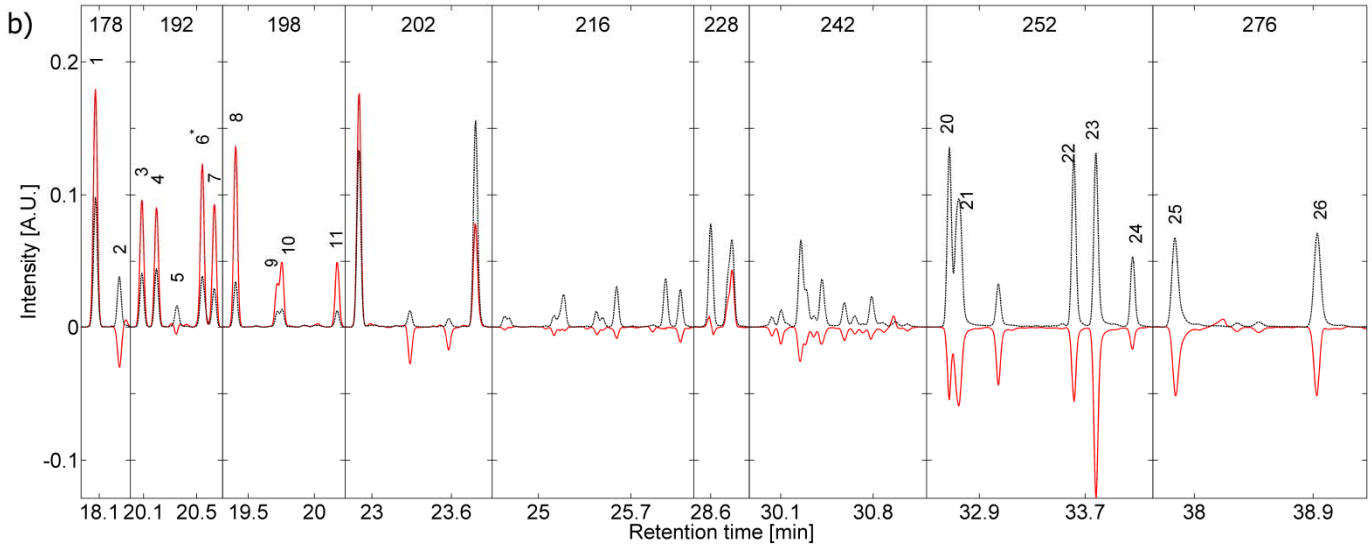
4 million litre of crude oil escaped from a burst underwater pipeline at the Petrobras refinery into and Guanabara Bay

# PART C: Map of samples (the score plots)

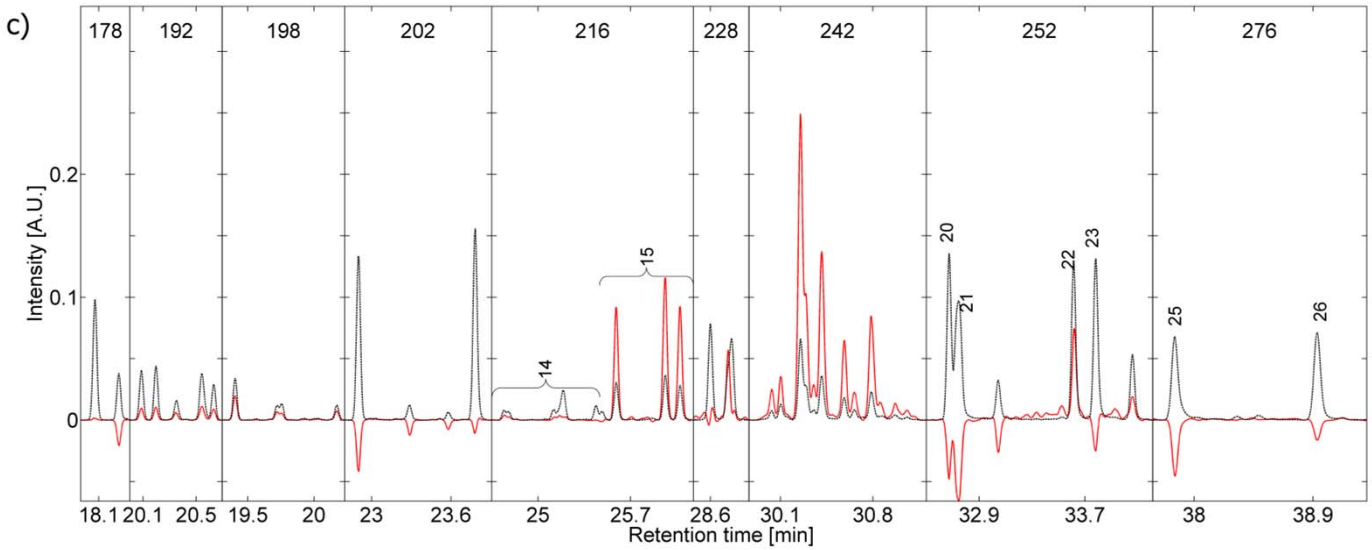


# PART C: Chemical interpretation (the loading plots)

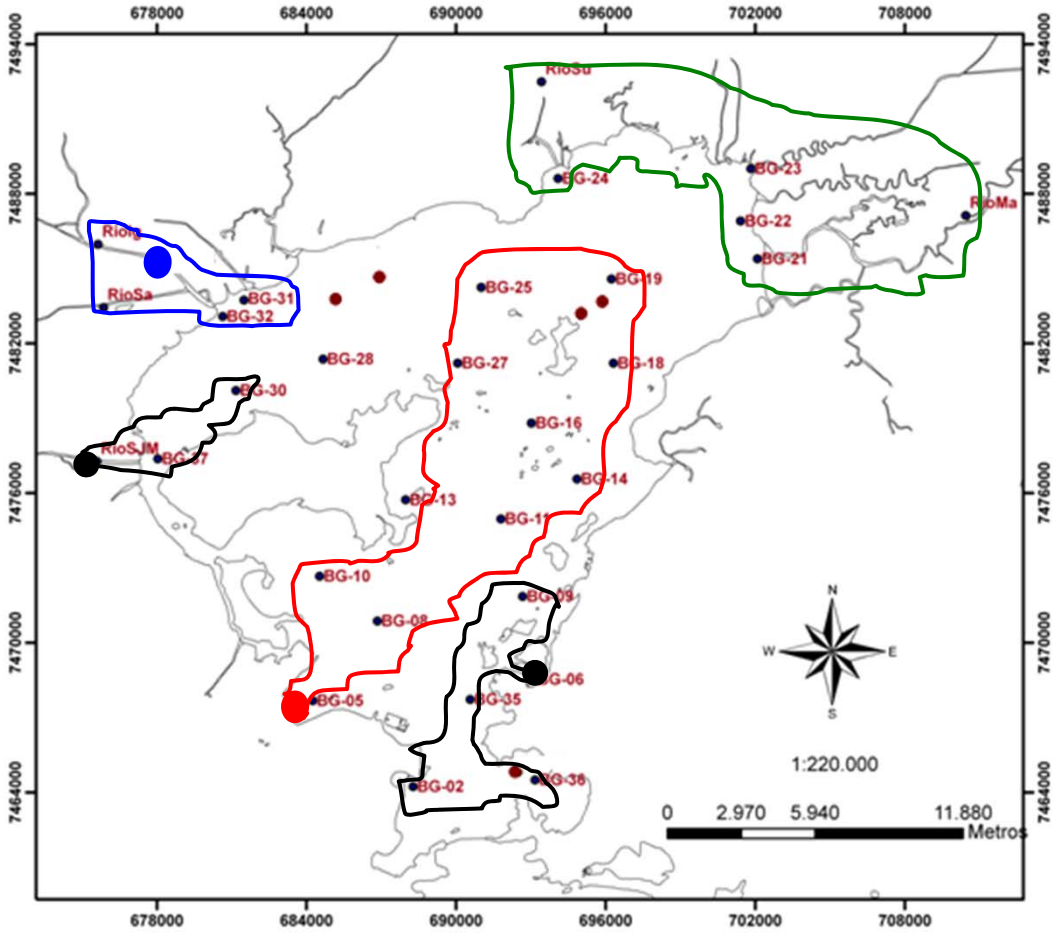
PC2




PC3

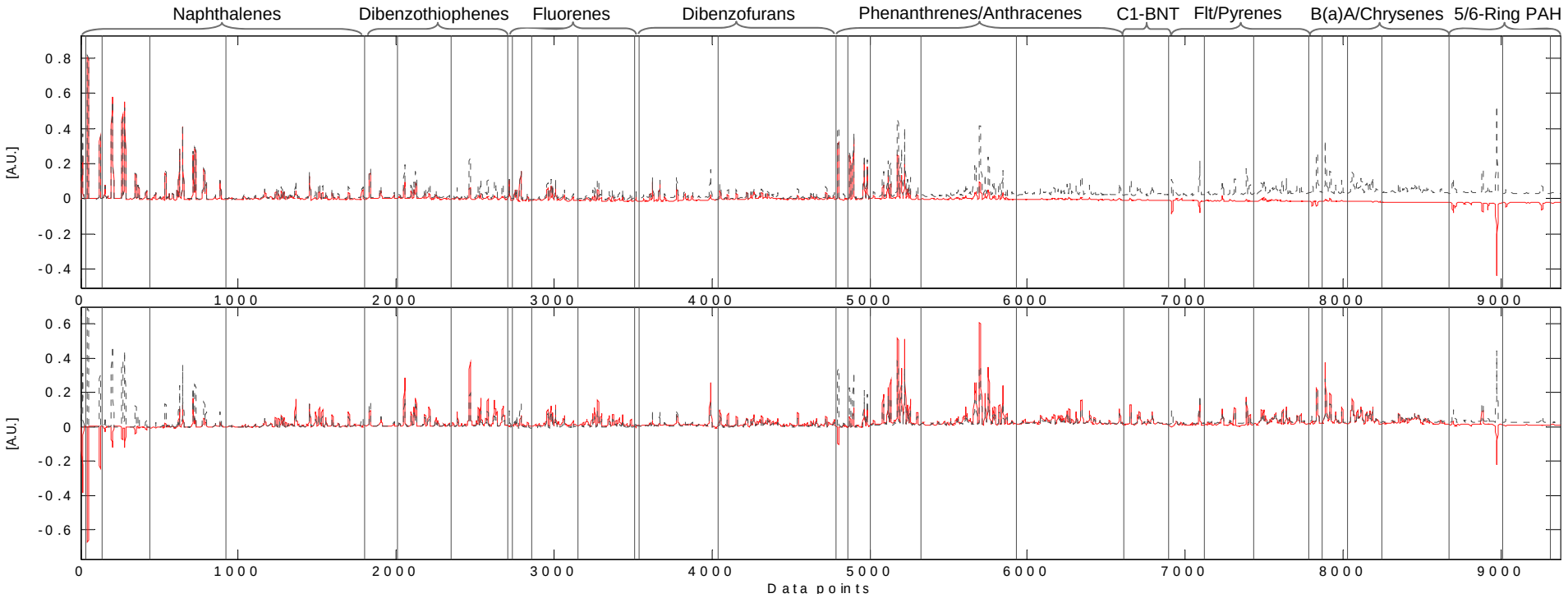


# PAH pollution sources, patterns and gradients



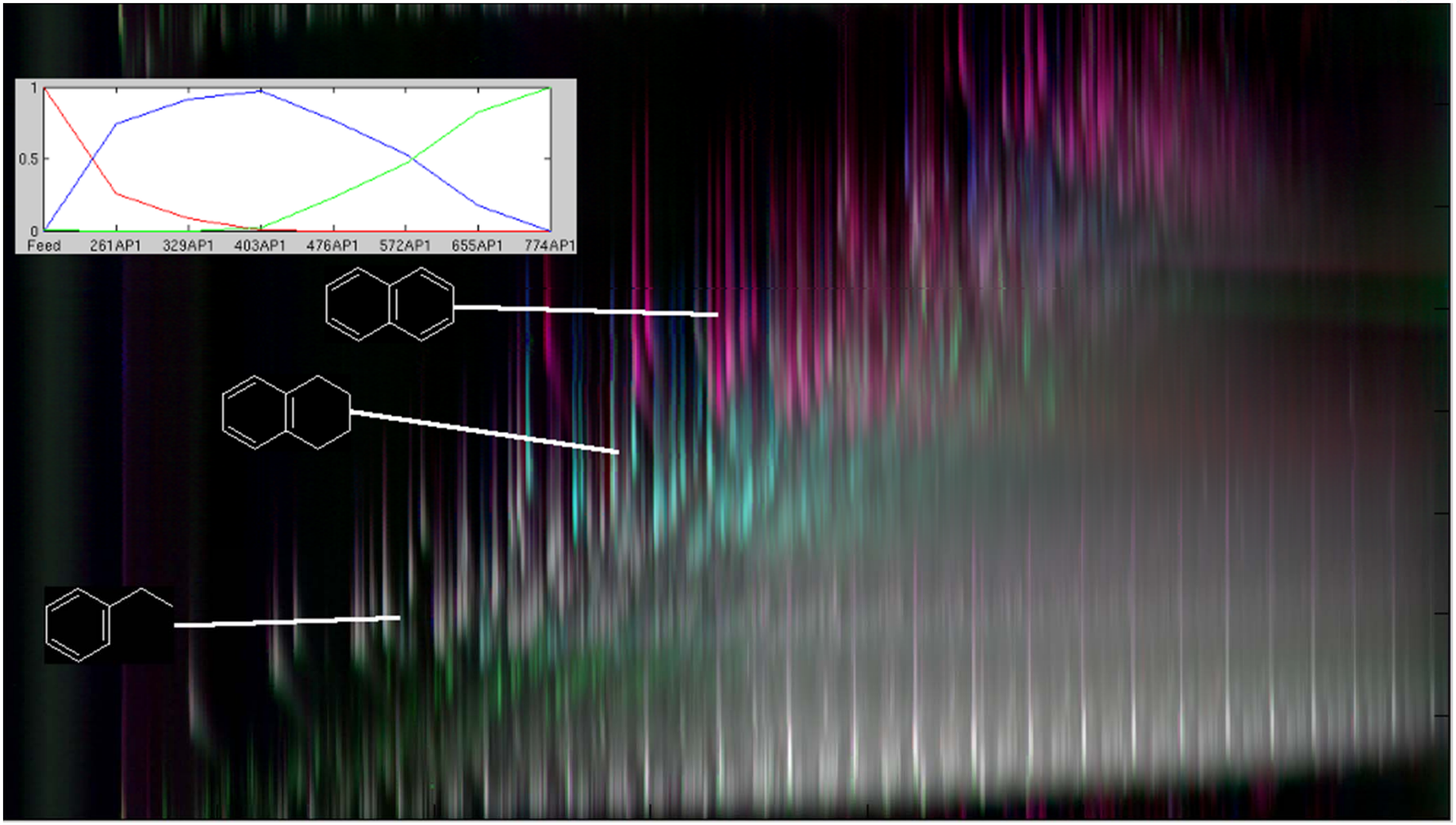
# Modifications to the CHEMSIC method

… and more SICs (information)



# ...and to GCxGC data

## Archetypal analysis



# Concluding remarks

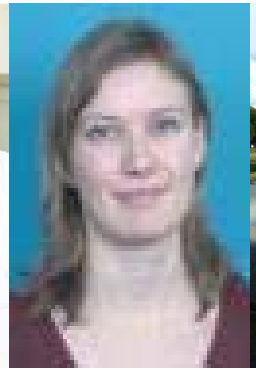
- **Contaminant profiling is <u>one way</u> to deal with contaminant complexity and it is <u>complementary</u> to conventional quantitative analysis of selected contaminants**

- **New tools are under development – and we are applying them to numerous different areas**
  - Source apportionment
  - Tracking bioremediation efficiency
  - Environmental metabolomics
  - Risk assessment of contaminant mixtures (correlation to toxicity)
  - …

# Acknowledgements

Post Doc
Giorgio Tomasi

Lab technician
Jette Petersen



Assist Prof
Nikoline Nielsen

Post Doc:
Raquel F Varela

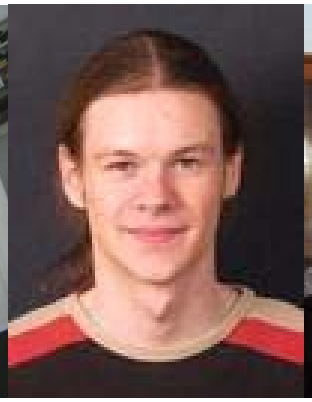
PhD. Stud:
Nichlas Weezel

PhD. Stud:
Linus Malmquist

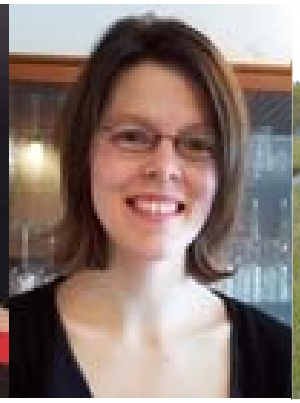
PhD. Stud:
Sarah Mohamed

PhD Stud:
Esther S Boll

PhD Stud
Søren Furbo



PhD Stud
Rune Græsbøl



PhD Stud
Majbrit Hansen

PhD Stud
Kuba
Modrzynski

PhD Stud
Peter Christensen