

# Improving the understanding and optimization of LIBS analysis using chemometrics

J.-B Sirven, E. Vors, K. Tchepidjian, M. El Rakwe

▶ **To cite this version:**

J.-B Sirven, E. Vors, K. Tchepidjian, M. El Rakwe. Improving the understanding and optimization of LIBS analysis using chemometrics. LIBS 2014 - 8ème conférence internationale sur la spectroscopie de plasma, Sep 2014, Beijing, China. cea-02874853

**HAL Id: cea-02874853**

**https://cea.hal.science/cea-02874853**

Submitted on 19 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DE LA RECHERCHE À L'INDUSTRIE

www.cea.fr

# IMPROVING THE UNDERSTANDING AND OPTIMIZATION OF LIBS ANALYSIS USING CHEMOMETRICS

**J.-B. Sirven, E. Vors, K. Tchepidjian, M. El Rakwe**

*CEA, DEN, DPC, F-91191 Gif-sur-Yvette, France*

*jean-baptiste.sirven@cea.fr*

LIBS' 2014, 8-12 September 2014, Beijing, China

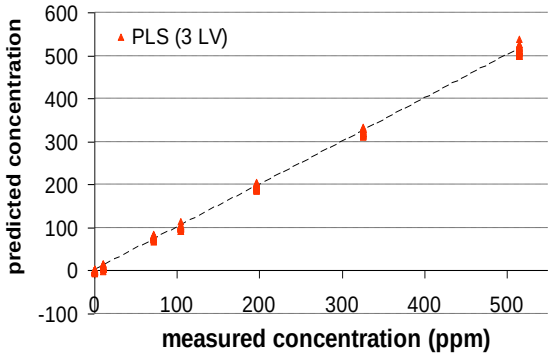
# CONTEXT

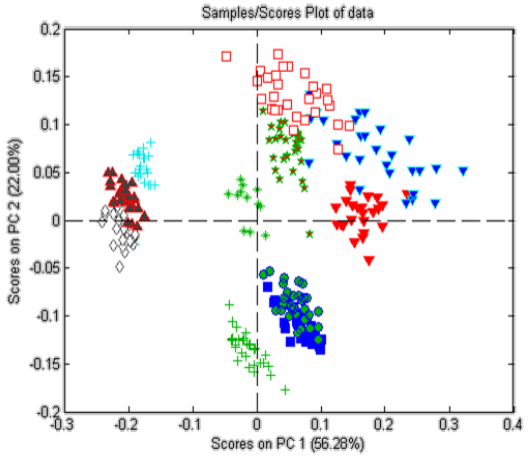
LIBS is **fast** and with Echelle spectrometers **high dimension spectra** can be acquired

→ MUCH data can be easily collected in a short time



**Chemometrics** has been extensively applied to LIBS data for ~ 20 years, to perform samples identification (PCA, SIMCA, PLS-DA, ANN…) or quantitative measurements (PCR, PLS, ANN…)



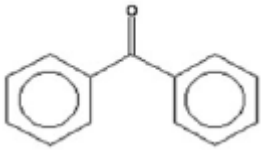




LIBS spectra do not contain only chemical but also **instrumental** and **physical** information.

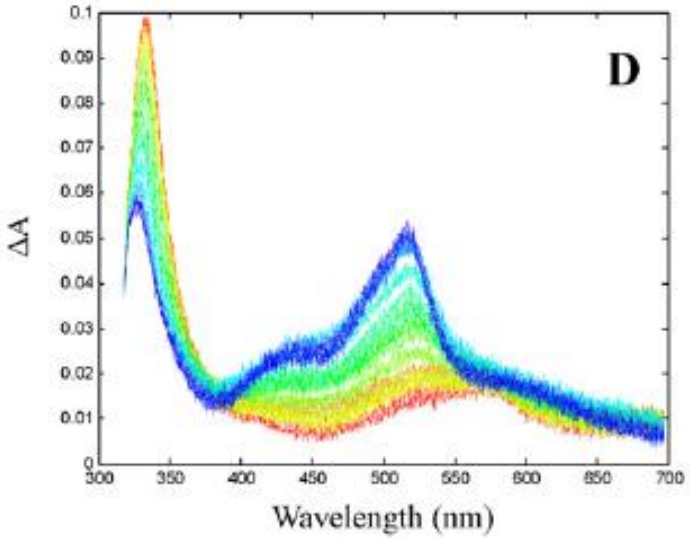
# AN INSPIRING EXAMPLE

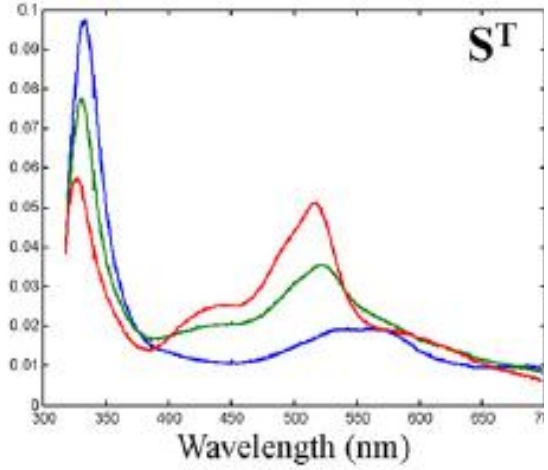
C. Ruckebusch et al. / Journal of Photochemistry and Photobiology C: Photochemistry Reviews *13 (2012) 1– 27*

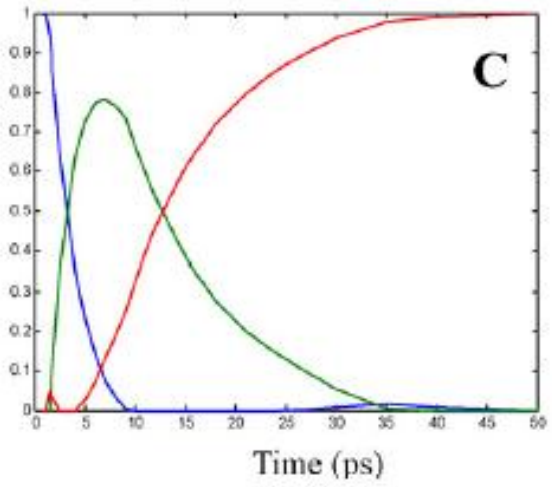
Femtosecond time-resolved absorption spectroscopy of benzophenone



Identification of an **intermediate state** (IS) of benzophenone and determination of its **spectrokinetic properties using chemometrics** (EFA, MCR-ALS, HS-MCR)




$S_1(n,\pi^*) \longrightarrow IS \longrightarrow T_1(n,\pi^*)$

# OTHER EXAMPLES



*Anal. Chem.* **2005,** *77,* 556–563

## Use of In-Line Near-Infrared Spectroscopy in Combination with Chemometrics for Improved Understanding of Pharmaceutical Processes

**Jukka Rantanen,†,‡ Håkan Wikström,‡ Rebecca Turner,‡ and Lynne S. Taylor*,‡**

*Viikki Drug Discovery Technology Center (DDTC), Pharmaceutical Technology Division, FIN-00014 University of Helsinki, Helsinki, Finland, and Department of Industrial and Physical Pharmacy, School of Pharmacy, Purdue University, West Lafayette, Indiana 47907-1336*

*Applied Spectroscopy Reviews*, 49:462–491, 2014
Copyright © Taylor & Francis Group, LLC
ISSN: 0570-4928 print / 1520-569X online
DOI: 10.1080/05704928.2013.855637

Taylor & Francis
Taylor & Francis Group

## A Review of Online Real-Time Process Analyses of Melt-State Polymer Using the Near-Infrared Spectroscopy and Chemometrics

MASAHIRO WATARI

Department of Chemistry and Research Center for Near-Infrared Spectroscopy, School of Science and Technology, Kwansei-Gakuin University, Sanda, Japan
Industry Marketing Department, Industrial Automation Marketing Headquarters, Yokogawa Electric Corporation, Musashino, Japan

ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect

Chemometrics and Intelligent Laboratory Systems 87 (2007) 173 – 184

Chemometrics and intelligent laboratory systems
www.elsevier.com/locate/chemolab

## Linking chemical knowledge and genetic algorithms using two populations and focused multimodal search

M.P. Gómez-Carracedo [a], M. Gestal [b], J. Dorado [b], J.M. Andrade [a,*]

[a] *Department Analytical Chemistry, University of A Coruña, Campus da Zapateira s/n, E-15071, A Coruña, Spain*
[b] *Department Information Technologies and Communication, University of A Coruña, Spain*

e at www.sciencedirect.com
cienceDirect

nt Laboratory Systems 91 (2008) 151 – 163

Chemometrics and intelligent laboratory systems
www.elsevier.com/locate/chemolab

## Acoustic chemometrics on liquid flow: Shift in the frequency spectra and its relationship to the physical properties of the liquid and the pipe

Andriy Kupyna *, Reidar Barfod Schüller, Elling-Olav Rukke, Tomas Isaksson

*Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, NO-1432, Aas, Norway*

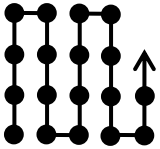
# OUTLINE

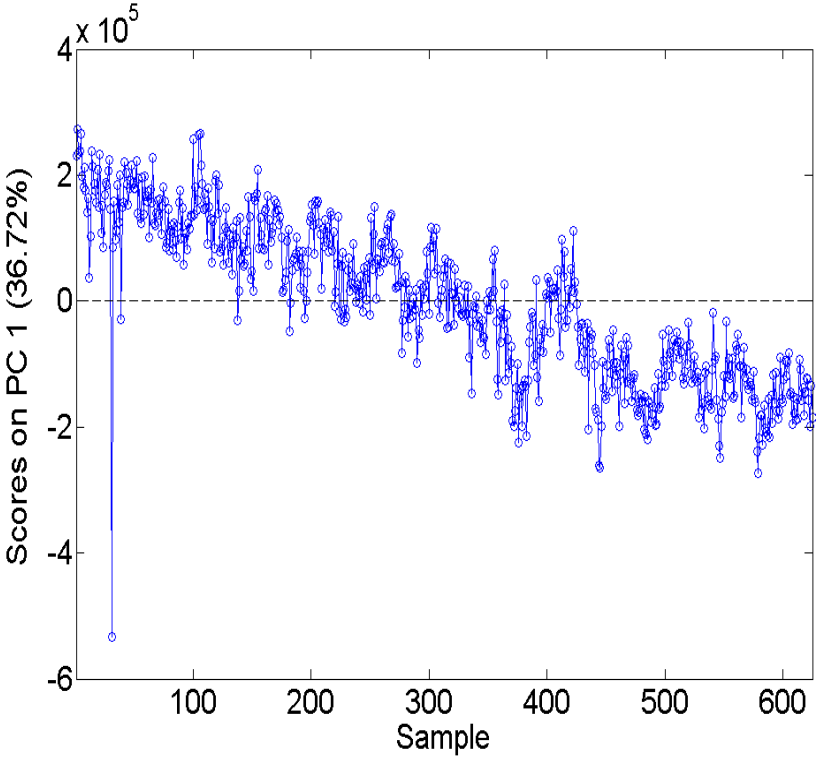
**Can we apply chemometric techniques to LIBS data in order to:**

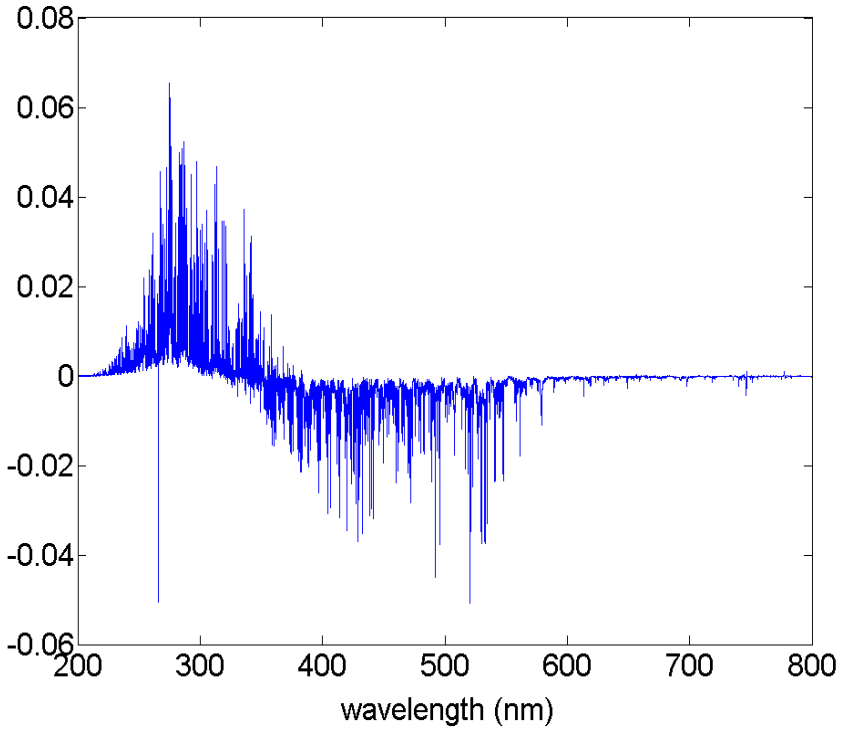
1. **Diagnose the instrumentation?**
2. **Optimize the analytical process?**
3. **Better understand the LIBS physics?...**

# DIAGNOSING THE INSTRUMENTATION: SAMPLE FLATNESS, DRIFT

**Data**: 625 spectra of steel (repeatability) acquired with an Echelle spectrometer. Matrix 25x25 on the sample.





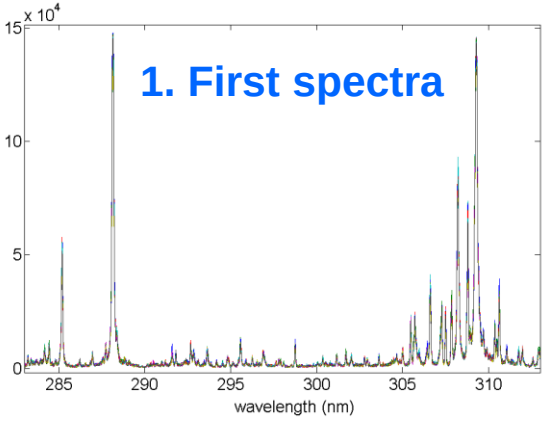



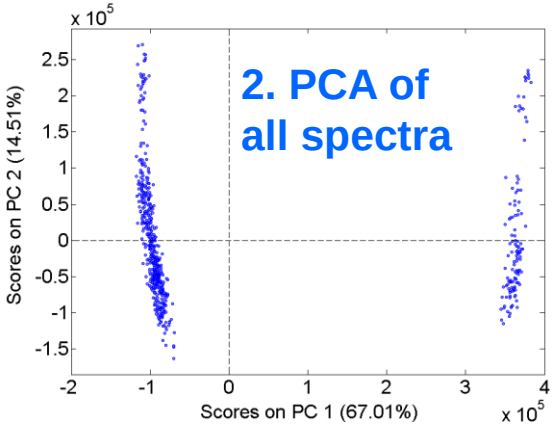

**1st PC:**

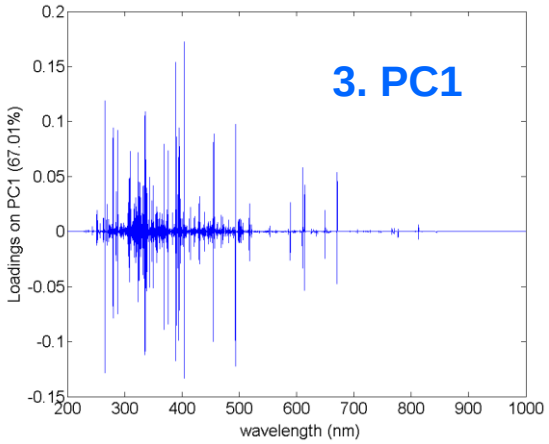
- **Scores**: slow drift + temporal periodicity (sample flatness)
- **Loadings**: different behaviour UV / VIS → Laser energy drift? Plasma temperature drift?

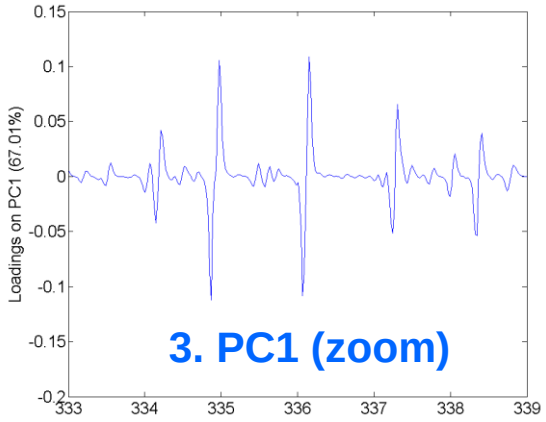
# DIAGNOSING THE INSTRUMENTATION : THERMAL DRIFT

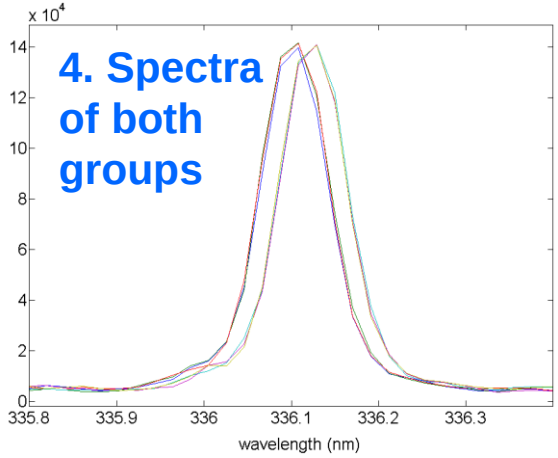
**Data**: 600 spectra of a glass sample (repeatability) acquired with an Echelle spectrometer.

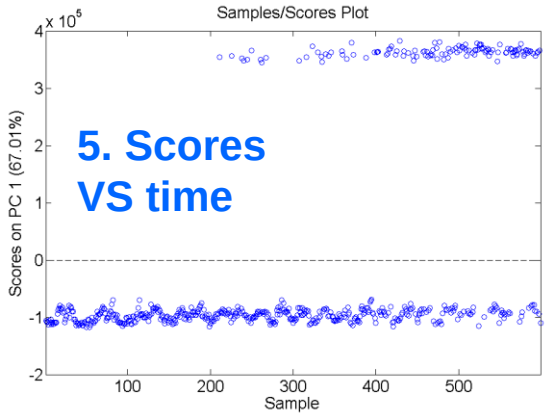



➔ **spectral shift**

➔ **slow thermal drift (~ 1 hour)**

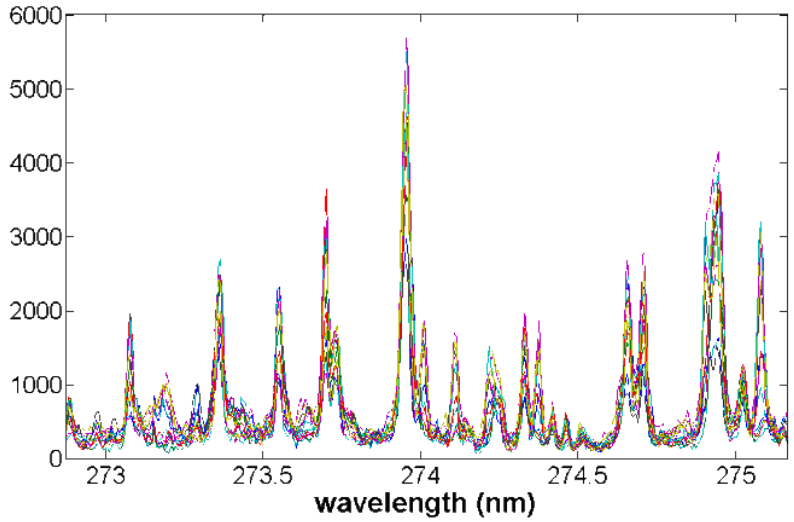
# OPTIMIZING A SIMCA MODEL: ROBUSTNESS OF PREDICTIONS

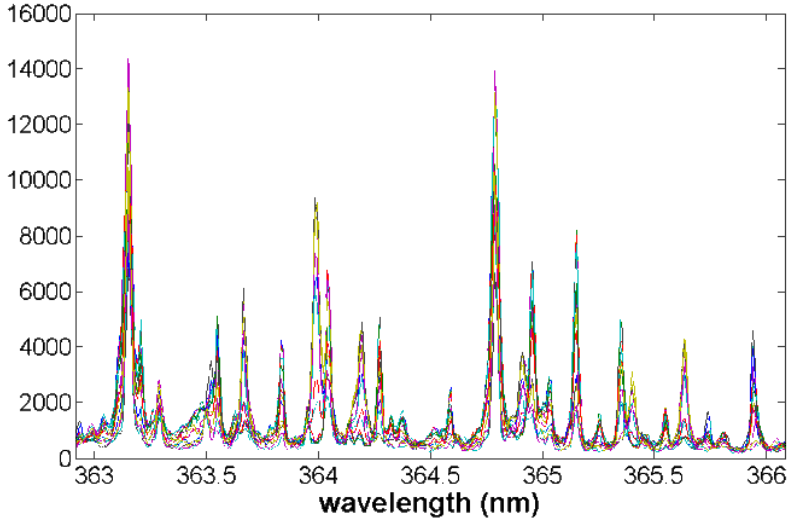
**Data**: 13 steel samples, 25 spectra each acquired with an Echelle spectrometer

**Objective**: robust samples identification using a SIMCA model

**2011 data**
80% spectra → model calibration
20% spectra → model test
(reference performances)

**2012 data (same samples)**
100% spectra → test of model robustness



# OPTIMIZING A SIMCA MODEL: ROBUSTNESS OF PREDICTIONS

**2011 test set:**

3.1% error rate

**2011 test set – confusion matrix**

| Predicted class \ Real class | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 1 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 |

**2012 data – confusion matrix**

| Predicted class \ Real class | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 47 | 32 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 3 | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 45 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 34 | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 47 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 29 | 0 | 1 | 0 | 0 |
| 10 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 3 | 20 | 50 | 14 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 35 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 41 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 |

**2012 data:**

16% error rate

➔ poor robustness

# OPTIMIZING A SIMCA MODEL: ROBUSTNESS OF PREDICTIONS

## Optimization using a full factorial DOE:

| Factors | Levels |
|---|---|
| Input data | 4 |
| Preprocessing | 6 |
| Distance to the model | 3 |
| Classification threshold | 6 |

**2011 test set – confusion matrix**

| Predicted class \ Real class | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 |

**2012 data – confusion matrix**

| Predicted class \ Real class | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 |

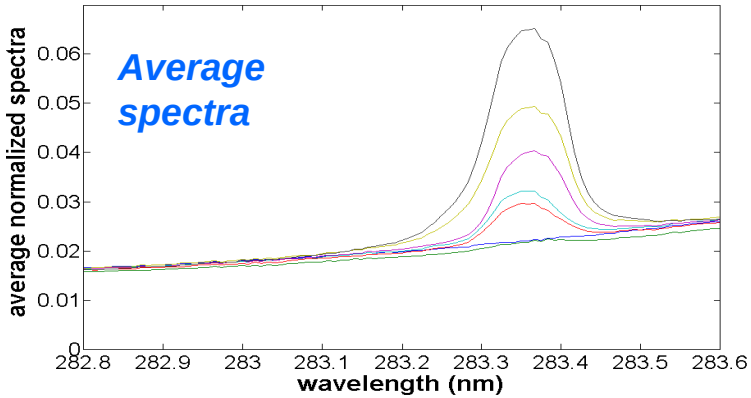
## Optimal parameters:

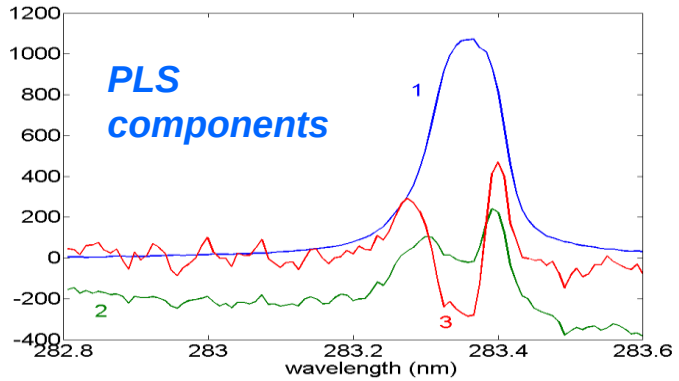
- Selection of lines using PCA components and an intensity threshold
- Preprocessing: normalization by the sum of emission lines + mean centering
- Distance criterion: residuals
- Any classification threshold

# PHYSICALLY INTERPRETING A MODEL: SELF-ABSORPTION

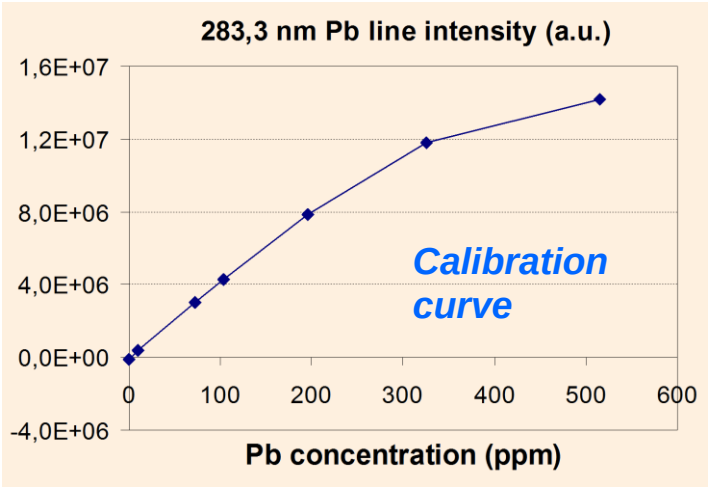
**Data:** calibration of lead concentration in liquid sodium by the standard addition method.

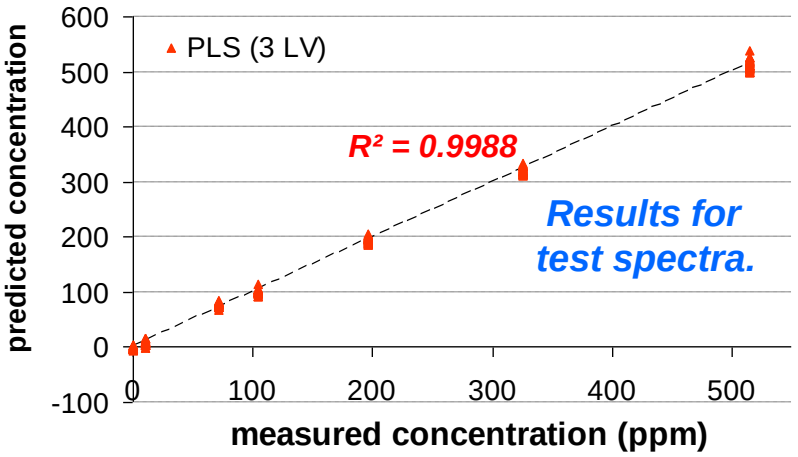
100 spectra per concentration, 67 for calibration, 33 for test. Only the **Pb line at 283.3 nm** is used in order to directly compare PLS and the calibration curve.



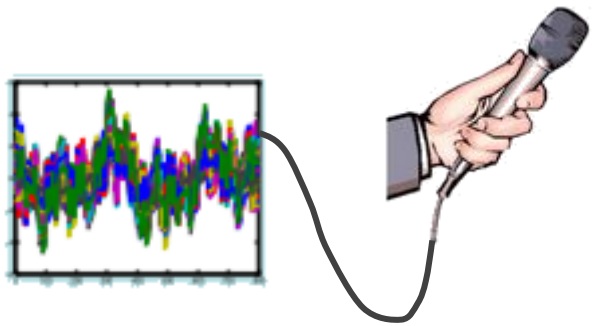

**Components 2 and 3 show that non linearity due to self-absorption is taken into account**



# PRINCIPLE OF ICA

**Independent Components Analysis (Blind Source Separation or Cocktail party):** *how to retrieve independent conversations from a mixture*



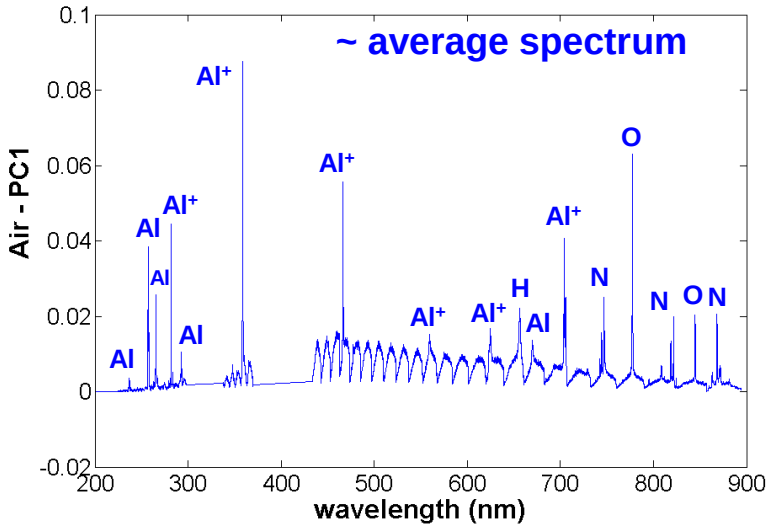


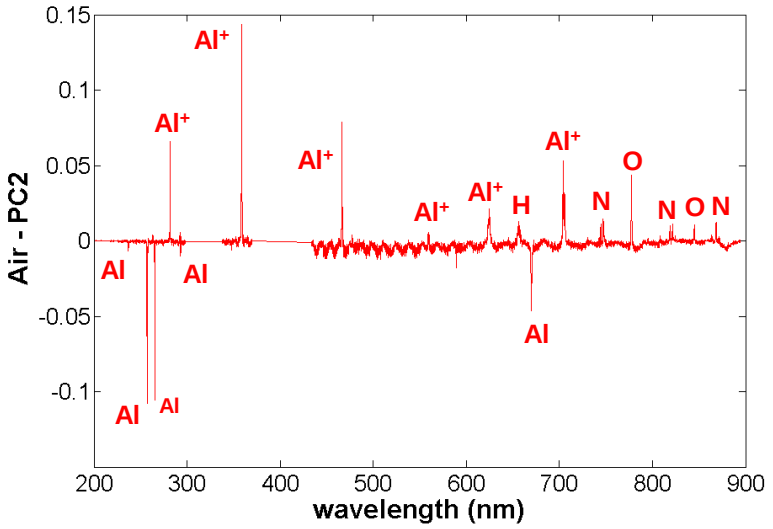
Independent Components ⇔ pure elements spectra *(? In LIBS ?)*

Experimental spectra = linear combination of components (like in PCA)

# ICA VS PCA: LIBS OF A PURE AL SAMPLE IN AIR

**Data**: pure Al sample spectra in air, gate delay 0.2 to 15 µs, 50 acc. laser shots (from R. Saad's PhD work)

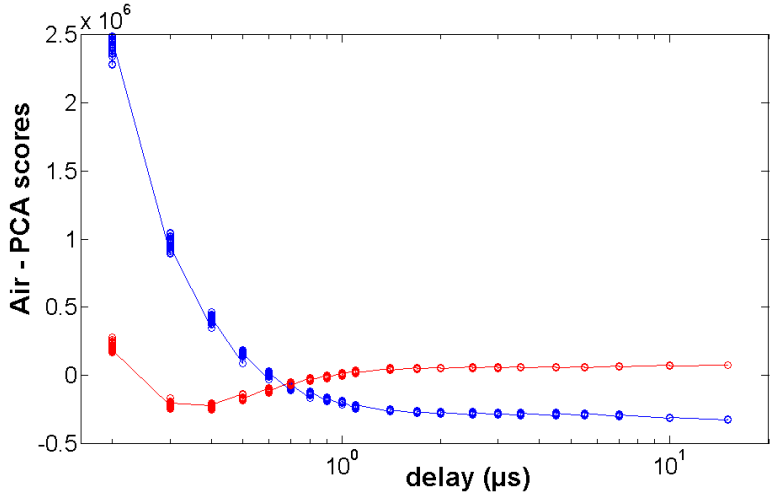



**PC1 scores** monotonic decrease (OK)

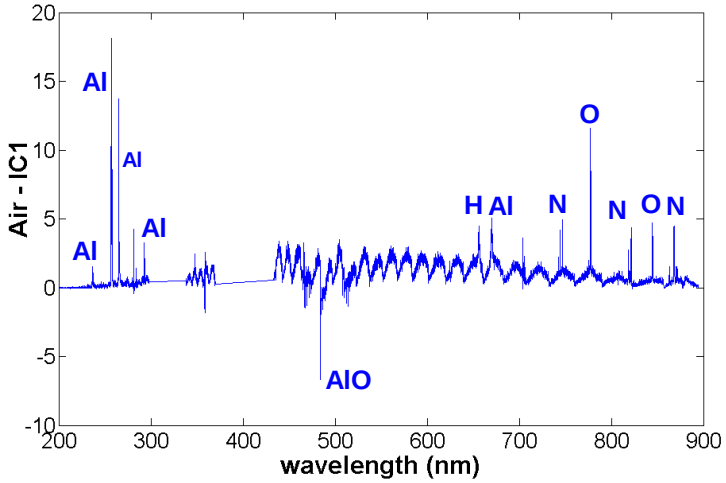
**PC2 scores** are not monotonic and can be negative

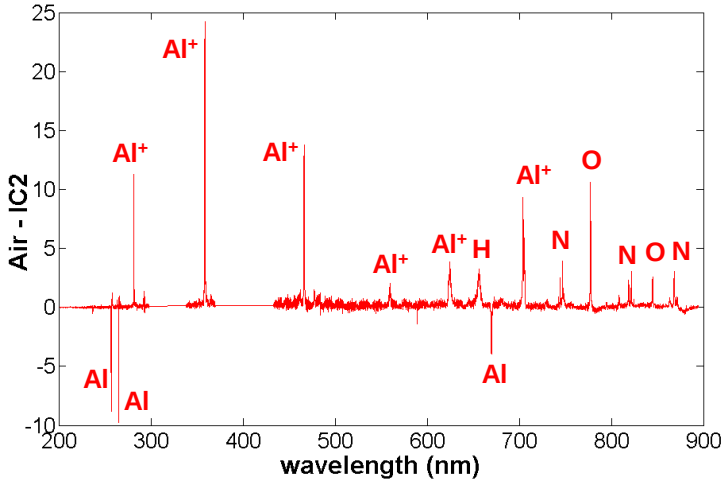
**Physical interpretation?**



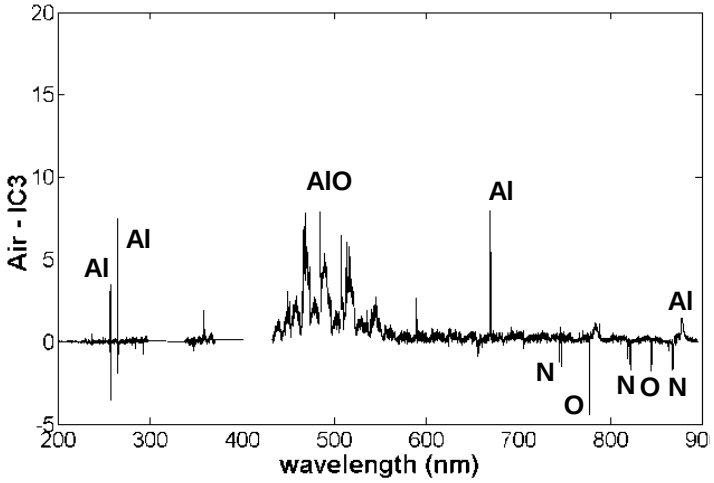
# ICA VS PCA: LIBS OF A PURE AL SAMPLE IN AIR

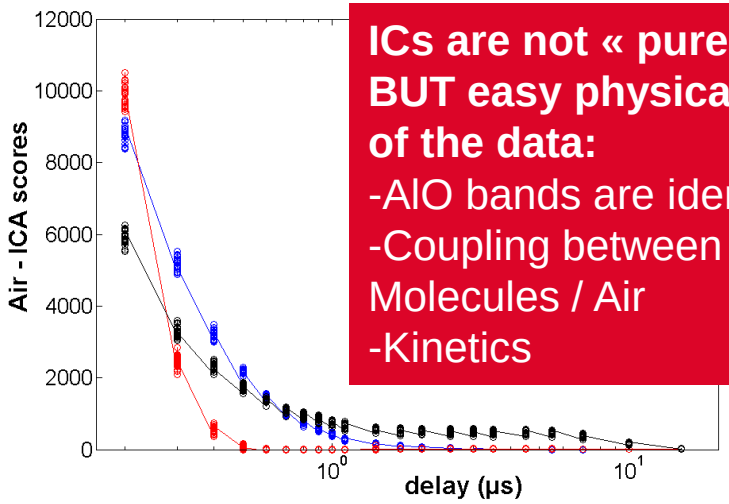
**Data**: pure Al sample spectra in air, gate delay 0.2 to 15 µs, 50 acc. laser shots (from R. Saad's PhD work)




*(IC2 ~ PC2)*




**ICs are not « pure components » BUT easy physical interpretation of the data:**

-AlO bands are identified (≠ PCA)

-Coupling between Ions / Neutrals / Molecules / Air

-Kinetics

# SUMMARY

**Chemometric techniques** are very powerful to efficiently extract relevant information from complex, highly multi-dimensional experimental data

They proved efficient for « classical » applications of LIBS analysis:

- Samples identification
- Quantification

But they can also be used for **other purposes**:

- Instrumentation diagnosis
- System / method optimization
- Description / understanding of physical phenomena

**Independent Components Analysis** is a very promising technique for this last point, e.g. for the treatment of time-resolved and / or molecular LIBS data

**Physical interpretation** of chemometric models is always possible (and desirable)

**Special thanks to:**

**D. N. Rutledge** (AgroParisTech, France) for codes and advice on ICA

**R. Saad** (CEA) for experimental data on time-resolved Al plasmas

# Thank you for your attention!

jean-baptiste.sirven@cea.fr

Commissariat à l'énergie atomique et aux énergies alternatives
Centre de Saclay | 91191 Gif-sur-Yvette Cedex
DEN/DANS/DPC/SEARS/LANIE – Bât. 467 – PC 56
T. +33 (0)1 69 08 21 00 | F. +33 (0)1 69 08 78 84

Etablissement public à caractère industriel et commercial
RCS Paris B 775 685 019

Direction de l'énergie nucléaire
Direction déléguée aux activités nucléaires de Saclay
Département de physico-chimie
Service d'études analytiques et de réactivité des surfaces
Laboratoire de développement analytique nucléaire, isotopique et élémentaire

# ABSTRACT

Laser-Induced Breakdown Spectroscopy is relatively simple to implement experimentally, but its underlying physics is very complex. This is especially due to two main factors. First, sampling and excitation occur in the same process, driven by the laser-matter interaction, and therefore are strongly coupled. Secondly, laser ablation, plasma formation and emission are transient by nature, and plasma features vary very rapidly with time.

As a consequence, this observation has stimulated the use of more empirical approaches based on chemometrics, which have been more and more widespread in LIBS for a dozen years. The power of multivariate methods for LIBS data treatment was indeed demonstrated in many papers. However, in most cases, those approaches aim at retrieving some chemical information on the sample, either qualitative (presence of some elements, sample nature) or quantitative (concentration measurements).

But LIBS spectra contain more information than the chemical one that could also be exploited by chemometric methods. Thus, multivariate models can be used or interpreted in order to diagnose possible instrumental issues, to optimize a LIBS system or analysis for a given application, or to improve our physical understanding of LIBS data. The talk will illustrate the use of chemometric methods to address these objectives.