FROM QUALITATIVE TO QUANTITATIVE LIBS USING UNIVARIATE ANALYSIS

Bruno Bousquet  
*CELIA, Univ. Bordeaux, France*  
bruno.bousquet@u-bordeaux.fr

Vincent Motto-Ros  
*ILM, Univ. Lyon, France*  
v Vincent.motto-ros@univ-lyon1.fr
Example of LIBS spectrum

From the wavelength, I can say if a given element has been detected!

From the line intensity, I can give the concentration of any detected element!

THE END!
How to conduct a reliable, robust and high-quality LIBS quantitative analysis?

Supposed to be mastered:
- **Samples**
  - Homogeneous
  - Adequate range of concentrations
  - Matrix matched
- **Protocol**
  - Sample preparation
  - Number of shots
  - Repeatable procedures
- **Acquisition**
  - Stable and reproducible instrumentation
  - Adequate detection limits

Raw Spectra

This lecture:
- **Calibration**
  - Calibration model
  - Figures of merit ($R^2$, LoD, LoQ, SNR, etc.)
- **Extraction**
  - Line selection
  - Method of extraction
  - Justification
- **Pre-processing**
  - Baseline
  - Noise
  - Normalization (?)
  - Line identification

Selection of the best sequence

Routine Quantification
1) Preprocessing
2) Line identification
3) Signal extraction
4) Towards quantification
5) Conclusion
Preprocessing: Why do we need to “prepare” the raw LIBS data?

Example of LIBS spectrum

- Noise?
- Dark, background signals?
- Calibration? Normalization?
- Saturation?
- Line broadening?

Potential error of line identification and bias for signal extraction

Error of quantification

Take the time to deeply observe the dataset!
LIBS spectrum = hundreds or thousands of variables/wavelengths
Exploration tool: principal component analysis – PCA (see 2\textsuperscript{nd} workshop by J. El Haddad)

Preprocessing: reveal and correct wavelength calibration issue

- Raw data: 5 classes instead of 3
- Corrected data: 3 classes as expected

Abnormal features related to wavelength drift

Normal features (here Ca\textsuperscript{+})
**How to theoretically describe the line profile?**

<table>
<thead>
<tr>
<th>Thermal Doppler Broadening</th>
<th>Stark Broadening</th>
</tr>
</thead>
<tbody>
<tr>
<td>The optical frequency of a photon emitted from an atom within the plasma is &quot;red&quot;- or &quot;blue&quot;-shifted by the Doppler effect depending on the velocity of the atom relative to the observer.</td>
<td>Due to the interaction of the emitting atom with fast moving electrons and the slowly moving ions in plasma*.</td>
</tr>
<tr>
<td>[ \gamma_{Doppler}(T) = 3.58 \times 10^{-7} \cdot \nu_0 \cdot \sqrt{\frac{T}{M}} ]</td>
<td>[ \Delta \lambda_{Stark} \propto Ne ] (electron density)</td>
</tr>
<tr>
<td>( \gamma_{Doppler} ): half width at half maximum, ( \nu_0 ) line frequency, ( M ) (g/mol), and T (K)</td>
<td>* “Spectral Line Broadening by Plasmas”, Hans Griem, Academic Press (1974)</td>
</tr>
</tbody>
</table>

\[ g_D(\nu - \nu_0) = \frac{\sqrt{\ln 2}}{\Delta \nu_D \sqrt{\pi}} e^{-\left(\frac{\nu - \nu_0}{\Delta \nu_D}\right)^2 \ln 2} \]

\[ g_L(\nu - \nu_0) = \frac{1}{\pi} \left( \frac{\Delta \nu_L}{(\nu_0 - \nu)^2 + \Delta \nu_L^2} \right) \]
How does the instrument broadening influence the line width?

**Instrumental broadening**

The instrumental broadening is generally modeled by a Gaussian profile to be convoluted with the calculated Voigt profile.

The measured linewidth $\Delta \nu_{meas}$ depends on the intrinsic width of the line $\Delta \nu_{line}$ and the instrumental width $\Delta \nu_{inst}$:

$$\Delta \nu_{meas} = \sqrt{\Delta \nu_{line}^2 + \Delta \nu_{inst}^2}$$

If $\Delta \nu_{inst} \ll \Delta \nu_{line}$:  $\Delta \nu_{meas} \approx \Delta \nu_{line}$

High spectral resolution is preferred!

**Resolution Power**

The resolution power of a spectrometer is driven by:
- the width of the entrance slit
- the diffraction grating

By injecting the light from a low-pressure spectral lamp (typ. Hg-Ar) into the spectrometer, one can plot for a given grating and slit width:

$$\frac{\lambda}{\Delta \lambda_{1/2}} = 9010 \pm 150$$

$\Delta \lambda_{1/2} \approx 0.7 \text{ Å}$

By plotting the instrumental half width (nm) versus wavelength (nm), one can see that the resolution is not constant over the spectrum.
**Preprocessing: Line Broadening**

What is a good approximation of the line profile?

<table>
<thead>
<tr>
<th>The Voigt profile*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convolution of the Gaussian and Lorentz profiles</td>
</tr>
<tr>
<td>( g_V(\nu) = g_D(\nu) \otimes g_L(\nu) )</td>
</tr>
<tr>
<td>( = \int_{-\infty}^{+\infty} g_L(\nu - \nu').g_D(\nu').d\nu' )</td>
</tr>
</tbody>
</table>

![Graph showing Gaussian, Lorentzian, and Voigt profiles](image)


**N.B.**

The Voigt profile does not take into account the change in collision induced speeds.

**Residue**

Distortion of the fitted lines - Dicke effect

---

The Voigt profile is a good approximation to describe the width and the intensity of a line with an accuracy of the order of the %. 

---

PREPROCESSING: LINE BROADENING

Example: Molybdate powder

Self-reversed

Spike

Self-absorbed

Bias on peak intensity or area
In most of the cases, the spectral noise can be considered as Gaussian and follows a normal distribution.
**Preprocessing: Baseline**

- **Dark current**: constant offset // can be simply subtracted
- **Background**: nonlinear variations with plasma conditions
  - constant offset only for constant plasma conditions // can’t be simply subtracted
- **Overlap**: related to spectral resolution and sample composition

**Baseline calculation on a case-to-case basis**
Why should we normalize the LIBS spectra?

The LIBS spectra can exhibit significant variations due to changes in experimental conditions.

- Same sample composition
- Same LIBS instrument
- Changing moisture conditions

Example:

Normalization can be required, especially for out of lab operation.
PREPROCESSING: NORMALIZATION

How to normalize the LIBS spectra?

Option 1: Looking for potential correlations of other signals with the LIBS signal

- Ablated mass, crater volume/depth
- Acoustic signal

Microphone onboard the Mars 2020 rover // Supercam


Option 2: Looking for strategies of spectral-based normalization // LIBS signal only

TOP 4 of the spectral based methods in LIBS

Background

\[ I_{k}^{\text{norm}} = \frac{I_k}{I_{bkgd}} \]

Internal standard

\[ I_{k}^{\text{norm}} = \frac{I_k}{I_{\text{Int.Std}}} \]

Total area

\[ I_{k}^{\text{norm}} = \frac{I_k}{\sum_{j=1}^{n} I_j} \]

Standard Normal Variate - SNV

\[ I_{k}^{\text{snv}} = \frac{I_k - I_{\text{mean}}}{s} \]

J. Guezenoc et al., Spectrochimica Acta Part B, Volume 160, October 2019, 105688
OUTLINE

FROM QUALITATIVE TO QUANTITATIVE LIBS USING UNIVARIATE ANALYSIS

1) Preprocessing
2) Line identification
3) Signal extraction
4) Towards quantification
5) Conclusion
### Line Identification

Example of line parameters extracted from NIST for Al I

<table>
<thead>
<tr>
<th>Observed Wavelength Air (nm)</th>
<th>Ritz Wavelength Air (nm)</th>
<th>Rel. Int. (s⁻¹)</th>
<th>Acc.</th>
<th>$E_i$ (cm⁻¹)</th>
<th>$E_k$ (cm⁻¹)</th>
<th>Lower Level Conf., Term, J</th>
<th>Upper Level Conf., Term, J</th>
<th>Type</th>
<th>TP Ref.</th>
<th>Line Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>308.21529</td>
<td>308.21510</td>
<td>24g</td>
<td>B+</td>
<td>0.000</td>
<td>32 435.453</td>
<td>$3s^23p^2 \ ^3P^o \ ^1/2$</td>
<td>$3s^23d \ ^5D \ ^3/2$</td>
<td>c40, T8653</td>
<td>L4737</td>
<td></td>
</tr>
<tr>
<td>308.702</td>
<td>308.6989</td>
<td>5</td>
<td>44</td>
<td>168.847</td>
<td>76 553.47</td>
<td>$3s^23d \ ^5D \ ^3/2$</td>
<td>$3s3p (3P^o)^4 d \ ^2D^o \ ^5/2$</td>
<td>L3514</td>
<td></td>
<td></td>
</tr>
<tr>
<td>309.982</td>
<td>309.9786</td>
<td>1</td>
<td>44</td>
<td>166.398</td>
<td>76 521.7</td>
<td>$3s^23d \ ^5D \ ^3/2$</td>
<td>$3s3p (3P^o)^4 d \ ^2D^o \ ^3/2$</td>
<td>L3514</td>
<td></td>
<td></td>
</tr>
<tr>
<td>309.27099</td>
<td>309.27084</td>
<td>26g</td>
<td>B+</td>
<td>112.061</td>
<td>32 436.796</td>
<td>$3s^23p \ ^3P^o \ ^1/2$</td>
<td>$3s^23d \ ^5D \ ^5/2$</td>
<td>c40, T7070</td>
<td>L4787</td>
<td></td>
</tr>
<tr>
<td>309.20386</td>
<td>309.20369</td>
<td>20g</td>
<td>B</td>
<td>112.061</td>
<td>32 435.453</td>
<td>$3s^23p \ ^3P^o \ ^1/2$</td>
<td>$3s^23d \ ^5D \ ^5/2$</td>
<td>c40, T7070</td>
<td>L4787</td>
<td></td>
</tr>
<tr>
<td>320.339</td>
<td>320.3394</td>
<td>4</td>
<td>45</td>
<td>345.594</td>
<td>76 553.47</td>
<td>$3s^23d \ ^5D \ ^1/2$</td>
<td>$3s3p (3P^o)^4 d \ ^2D^o \ ^5/2$</td>
<td>L3514</td>
<td></td>
<td></td>
</tr>
<tr>
<td>324.159</td>
<td>324.1605</td>
<td>3</td>
<td>42</td>
<td>237.783</td>
<td>73 077.8</td>
<td>$3s^23d \ ^5D \ ^1/2$</td>
<td>$3s3p (3P^o)^4 d \ ^2D^o \ ^7/2$</td>
<td>L3514</td>
<td></td>
<td></td>
</tr>
<tr>
<td>325.160</td>
<td>325.1607</td>
<td>2</td>
<td>42</td>
<td>233.742</td>
<td>72 978.9</td>
<td>$3s^23d \ ^5D \ ^1/2$</td>
<td>$3s3p (3P^o)^4 d \ ^2D^o \ ^5/2$</td>
<td>L3514</td>
<td></td>
<td></td>
</tr>
<tr>
<td>328.220</td>
<td>328.2139</td>
<td>3</td>
<td>46</td>
<td>094.312</td>
<td>76 553.47</td>
<td>$3s^27d \ ^5P \ ^1/2$</td>
<td>$3s3p (3P^o)^3 d \ ^2F^o \ ^7/2$</td>
<td>L3514</td>
<td></td>
<td></td>
</tr>
<tr>
<td>341.9347</td>
<td>341.9346</td>
<td>6g</td>
<td>0.000</td>
<td>29 066.96</td>
<td>$3s^23p \ ^3P^o \ ^1/2$</td>
<td>$3s3p^2 \ ^3P \ ^3/2$</td>
<td>L4737</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>341.3610</td>
<td>341.3610</td>
<td>9g</td>
<td>112.061</td>
<td>29 142.78</td>
<td>$3s^23p \ ^3P^o \ ^1/2$</td>
<td>$3s3p^2 \ ^3P \ ^5/2$</td>
<td>L4737</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>341.4965</td>
<td>341.4963</td>
<td>6g</td>
<td>0.000</td>
<td>29 020.41</td>
<td>$3s^23p \ ^3P^o \ ^1/2$</td>
<td>$3s3p^2 \ ^3P \ ^1/2$</td>
<td>L4737</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>345.2657</td>
<td>345.2658</td>
<td>2g</td>
<td>112.061</td>
<td>29 066.96</td>
<td>$3s^23p \ ^3P^o \ ^1/2$</td>
<td>$3s3p^2 \ ^3P \ ^3/2$</td>
<td>L4737</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>345.8216</td>
<td>345.8145</td>
<td>6g*</td>
<td>44</td>
<td>168.847</td>
<td>73 077.8</td>
<td>$3s^23p \ ^3P^o \ ^1/2$</td>
<td>$3s3p (3P^o)^3 d \ ^2F^o \ ^7/2$</td>
<td>L4737</td>
<td></td>
<td></td>
</tr>
<tr>
<td>345.8216</td>
<td>345.8218</td>
<td>6g*</td>
<td>112.061</td>
<td>29 020.41</td>
<td>$3s^23p \ ^3P^o \ ^1/2$</td>
<td>$3s3p^2 \ ^3P \ ^3/2$</td>
<td>L4737</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

https://physics.nist.gov/PhysRefData/ASD/lines_form.html
Example of homemade software:

- Kurucz database
- Saha-Boltzmann equations

Input parameters:

- # elements
- elemental fraction
- plasma parameters (T, Ne)
- spectral resolution
1) Preprocessing
2) Line identification
3) **Signal extraction**
4) Towards quantification
5) Conclusion
The “quest” of all Spectroscopists is to extract the line surface (S) with the best accuracy! BUT...

**Intensity Extraction: Principle**

- **Emission measurement with a $\Delta y$ accuracy**

- **Line profile**

- **Fitting procedure**

- **Line parameters**
  - Position: $v_c \pm f_1(\Delta x, \Delta y)$
  - Width: $\Delta v \pm f_2(\Delta x, \Delta y)$
  - Surface: $S \pm f_3(\Delta x, \Delta y)$

- **Frequency measurement Accuracy $\Delta x$**
If the spectrum structure is not too complex ("simple" baseline, no line interference, etc.), it is possible to determine the intensity of the transitions by a "basic" Peak detection.

For a fixed linewidth, the peak maximum \( I_{\text{max}} \) is proportional to the line surface \( S \)

\[
I_{\text{max}} = \frac{0.47 S}{\Delta \nu_D}
\]

**Gaussian profile**

\[
I_{\text{max}} = \frac{S}{\pi \Delta \nu_L}
\]

**Lorentzian Profile**

Temperature (K)

Delay (µs)

TCu (peak maximum)

TCu (peak area)
Subtracting the baseline can be tricky

Fused bead sample

Polynomial Baseline fit
Single Baseline point

\[ R^2 = 0.9998 \]

\[ R^2 = 0.997 \]

\[ R^2 = 0.9998 \]

### INTENSITY EXTRACTION: A SHORT RÉSUMÉ

<table>
<thead>
<tr>
<th>Peak detection</th>
<th>Line fitting</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td><strong>The Fastest and simplest method</strong></td>
<td><strong>The more accurate method</strong></td>
</tr>
<tr>
<td>[ V = I_{\text{max}} - I_{\text{bg}} ]</td>
<td>[ V = I_{\text{Voigt}} ]</td>
</tr>
<tr>
<td>[ I_{\text{bg}} = \frac{I_1 + I_2}{2} ]</td>
<td>[ \text{Fit Voigt + baseline} ]</td>
</tr>
<tr>
<td><strong>BUT...</strong></td>
<td><strong>BUT...</strong></td>
</tr>
<tr>
<td>- No need for human supervision</td>
<td>- Time consuming</td>
</tr>
<tr>
<td>- The peak detection and the base line model are dissociated</td>
<td>- Need for human supervision</td>
</tr>
<tr>
<td>- Required constant linewidth</td>
<td>- Possible divergence of the algorithm</td>
</tr>
<tr>
<td>- Should be adapted case by case</td>
<td>- Need a real expertise</td>
</tr>
</tbody>
</table>
EXAMPLE: PRACTICAL CASE WITH LIBS IMAGING

Specific constraints:
• Large amount of data to process (no eye supervision, fast algorithms, etc.)
• Single shot spectra (noise, weak signals, etc.)

The proposed idea: test different extraction algorithm from a synthetic dataset*
EXAMPLE: PRACTICAL CASE WITH LIBS IMAGING
**EXAMPLE: PRACTICAL CASE WITH LIBS IMAGING**

\[
I_S = \sum_{i=1}^{n_S} I_i - n_S I_{BG}
\]

\[
u_S^2 = (n_S \times \sigma)^2 \left( \frac{1}{n_S} + \frac{1}{n_{BG}} \right)
\]
EXAMPLE: PRACTICAL CASE WITH LIBS IMAGING
BUT REMEMBER THAN LINE FITTING IS ALWAYS BETTER
OUTLINE

FROM QUALITATIVE TO QUANTITATIVE LIBS USING UNIVARIATE ANALYSIS

1) Preprocessing
2) Line identification
3) Signal extraction
4) Towards quantification
5) Conclusion
TOWARDS QUANTIFICATION – SYNOPTIC OF THE LIBS ANALYSIS

- Calibration set
- Validation set
- LIBS data
- Model
- Reference concentrations
- Predicted concentrations
- Errors of calibration
- Errors of prediction
- Reference concentrations
- Model uncertainties
- Unknown samples
- LIBS data
- Predicted concentrations + errors + uncertainties
Towards quantification – Basic definitions - I

Decreasing error

Decreasing uncertainty

Modified from https://www.researchgate.net/
**Towards Quantification – Basic Definitions**

**Accuracy**
- The average value obtained from a large series of repetitions is **close** to the reference value.
- Data obtained from independent measurements under stipulated conditions are **very similar**.

**Uncertainty**
- The average value obtained from a large series of repetitions is **far from the reference value**.
- Data obtained from independent measurements under stipulated conditions are **significantly spread**.

**Decreasing error**
- Accuracy means trueness and precision.

---

Modified from https://www.researchgate.net/
**Towards Quantification – Basic Definitions - II**

**Step-by-step assessment**

**Repeatability**

The measurement is repeated but only very few changes are allowed
- Same analyst
- Same instrument
- Same conditions (short period of time)

**Within-laboratory reproducibility (intermediate precision)**

Within the same laboratory, any relevant influence factor is allowed to vary

**Reproducibility**

Takes into account the precision resulting from different laboratories

---

Towards Quantification - Interlaboratory Comparisons

**LIBS 2008 conference in Berlin, Germany**
- First international interlab comparison organized by the BAM
  Dedicated session at the conference

**LIBS 2016 conference in Chamonix, France**
- Second international interlab comparison organized by the BAM
  *Second LIBS inter-laboratory comparison : analysis of cast iron samples. W. Bremer et al.*

- Interlab comparison within the French network
  *Round robin tests on laser ablation and quantitative analysis by LIBS. J.-B. Sirven et al.*

**BUT**
Repeatability and intermediate precision must be evaluated **before** starting an interlab comparison

**Reproducibility**
- Trueness
- Precision

**No publication**
Towards quantification - Interlaboratory comparisons

There are too many paths from the sample to the analyte’s predicted concentration

Experiment

Many experimental conditions!

- Laser wavelength
- Pulse energy
- Fluence
- Focusing
- Nb of laser shots
- Spectral resolution
- Delay, gate
- Wavelength calib
- Noise
- Moisture
- Gas
- Etc.

Data processing

Many strategies!

- Denoising
- Baseline subtraction
- Filtering
- Line selection
- Fitting
- Intensity/area
- Calibration model

All the interlab comparisons consisted in analyzing replicate samples

In 2015, we have proposed a new strategy:
Share the spectra instead of samples
to better understand the influence of the strategies of data processing
TOWARDS QUANTIFICATION — NEW INTERLABORATORY COMPARISON

Contest

- 8 participants
- Start: August 2015
- All the participant received by email:
  - The average spectra of the 8 samples and a brief description of their origin
  - The reference concentration values of the 6 samples from the calibration set
  - The template to send back the details about the regression model and the predicted concentration values of the 2 unknown samples
  - The schedule of the contest

Objective

Retrieve the concentrations of Na, Si, Ca, Al et Cr for the two unknown samples $U_1$ and $U_2$

V. Motto-Ros et al., Critical aspects of data analysis for quantification in LIBS, SAB B 140 (2018).
Towards Quantification – New Interlaboratory Comparison

**Samples**

- \( \text{Li}_2\text{B}_4\text{O}_7 \) glass prepared by Cetim (prev. CRITT Matériaux Alsace)
  - 6 for calibration (values of concentration provided)
  - 2 for prediction (unknown)

**Experiment**

- **Laser**: 266nm-5ns-15mJ
- **Echelle spectrometer + ICCD / Resolution 5000 / 220-850 nm**
  - **Delay**: 5µs / **gate**: 500 ns

- 10x10 matrix at the surface of the sample with 10 laser shots by point

- 1 sample = 1 average spectrum (from 1000 spectra)
First campaign

The participants were free to select each step from A to Z

Is the selection of the emission line the most critical factor?

- Different emission lines
- Different methods of baseline subtraction
- Different methods of signal extraction
- Different types of calibration models
Second campaign

The participants were asked to provide the predicted values of concentration of aluminum from each of the two Al I lines at 308.22 nm and 309.27 nm.

- Different lines
- Different methods of baseline subtraction
- Different methods of signal extraction
- Different types of calibration models

The very high variability in the results is not only due to the selection of the emission line!
Towards quantification – Evaluation tools

Exp. data = n points

\[ y_i = \text{LIBS} \]

\[ x_i = \text{Concentration} \]
TOWARDS QUANTIFICATION – EVALUATION TOOLS

Exp. data = n points

\[ y_i = \text{LIBS} \]

\[ x_i = \text{Concentration} \]
Towards quantification — Evaluation tools

Linear regression model

\[ y = b_0 + b_1 \cdot x + e \]

- \( y \): Observed value
- \( x \): Independent variable
- \( b_0 \): Intercept
- \( b_1 \): Slope
- \( e \): Error

**Coefficient of determination**

\[ r^2 = \frac{[\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})]^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2} \]

**Warning:** \( r^2 \) can’t be the only indicator to evaluate a model !!!

Mermet, Jean-Michel SAB 65(7), 509-523 (2010).
Towards quantification — Evaluation tools

Limits

\[ y = b_0 + b_1 \cdot x + e \]

\( \hat{y}_i \)  Concentration value of sample i predicted by the model

\[
\begin{align*}
\text{Standard deviation of the residue, } e &\quad S_e = \left[ \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{n-2} \right]^{1/2} \\
\text{Standard deviation of } b_0 &\quad S_{b_0} = S_e \left[ \frac{1}{n} + \frac{x^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \right]^{1/2} \\
\text{Standard deviation of } b_1 &\quad S_{b_1} = \frac{s_e}{\left[ \sum_{i=1}^{n}(x_i - \bar{x})^2 \right]^{1/2}}
\end{align*}
\]

\[ y = \text{LIBS} \]

\[ x = \text{Concentration} \]

\[ \text{Exp. data = } n \text{ points} \]

Limit of detection \( \text{LOD} = \frac{3S_{b_0}}{b_1} \)

Limit of quantification \( \text{LOQ} = \frac{10S_{b_0}}{b_1} \)

Below the LOD
the analyte is not detected

Between the LOD and the LOQ
the analyte is detected but it is not possible to rigorously predict its concentration // Qualitative

Above the LOQ
The concentration can be rigorously predicted // Quantitative
Limit of Quantification: *should depend on the model!!!*

Limit of detection: ???

\[ LoD = \frac{3\sigma}{I_{max} - I_{BG}} \]

**Confidence Level**
(Student’s coefficient \( t \))

**Gaussian noise: Normal distribution**
\[ f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{(x - \bar{x})^2}{2\sigma^2} \right] \]

\( \sigma \): stdv
\( \mu \): mean

**N.B.:** Many methods exist for the estimation of the LoD, it is important to always mention how it was evaluated
Towards quantification – Evaluation tools

Confidence limit and LOQ

\[ \hat{y}_0(x_0) = \hat{\beta}_0 + \hat{\beta}_1 x_0 \pm t \cdot s_{\hat{y}_0} \]

- \( \hat{y}_0 \) is given within a confidence interval
- \( t \) is the Student’s coefficient calculated for \( n-2 \) degrees of freedom at a given risk of error usually set to 5%

Confidence limit

LOQ

\[ \text{LOQ} = \frac{2t \cdot s_e}{b_1} \left[ \frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2} \right]^{1/2} \]
Uncertainties

\[ y = b_0 + b_1 \cdot x + e \]

Standard deviation of the predicted response for a concentration value \( x_0 \)

\[ s_{\hat{y}_0} = s_e \left[ \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \right]^{1/2} \]

Standard deviation of the residue, \( e \)
TOWARDS QUANTIFICATION – EXAMPLE OF REAL QUANTITATIVE MODEL

Chromium (Cr I 357.87 nm)

$R^2 = 0.99995$
LOQ = 2.6 ppm

Confidence limit
Let's consider an unknown sample and its corresponding LIBS signal $\bar{y}_k$ resulting from the average of $p$ measurements

$$\bar{y}_k = \sum_{j=1}^{p} y_{kj}$$

regression model

$$x_k = \frac{\bar{y}_k - b_0}{b_1}$$

The standard deviation of $x_k$ is:

$$s_{x_k} = \frac{s_e}{b_1} \left[ \frac{1}{p} + \frac{1}{n} + \frac{(\bar{y}_k - \bar{y})^2}{b_1^2 \sum_{i=1}^{n} (x_i - \bar{x})^2} \right]^{1/2}$$

The value $x_k$ is given within the limits:

$$x_1 = x_k - t \cdot s_{x_k}$$
$$x_2 = x_k + t \cdot s_{x_k}$$

$t$ is the Student's coefficient calculated for $n-2$ degrees of freedom at a given risk of error usually set to 5%
Towards quantification – New interlaboratory comparison

Numerical parametric study - quantification of aluminum

2 emission lines – 6 methods of signal extraction – 2 types of regression: linear/quadratic

In this case, only the quadratic models allow to quantify the sample $U_2$. The best model has been selected \textit{a posteriori} on the basis of analytical performance.
Towards quantification – New interlaboratory comparison

Influence of the model - An example with Al I at 308 nm

- Linear model
- Confidence hyperbols
- Experiment

\[ R^2 = 0.99779 \]
\[ U_{\text{Xm}} = 7\% \]
\[ \text{LOQ} = 0.0089\% \]

- Quadratic model
- Confidence hyperbols
- Experiment

\[ R^2 = 0.99996 \]
\[ U_{\text{Xm}} = 2.7\% \]
\[ \text{LOQ} = 0.0019\% \]
Summary

*The interlab comparison focused on data processing has been published*

Critical aspects of data analysis for quantification in laser-induced breakdown spectroscopy


Demonstrated that the results of the LIBS analyses are strongly influenced by data processing

Standardization of LIBS data processing may allow to:
- Optimize preprocessing, normalization
- Optimize instrumentation and measurement
- Start new interlab comparison to assess the reproducibility
A model is robust if it efficiently predicts the concentrations of samples not used to build it.
A model is robust if it efficiently predicts the concentrations of samples not used to build it. For a model to be robust, the root-mean-square error (RMSE) between the predicted concentrations and the reference concentrations should be similar for both calibration and validation samples.

\[ \text{RMSE} = \sqrt{\frac{\sum_{p=1}^{N} (c_p - \hat{c}_p)^2}{N}} \]

where:
- \( N \) : Number of samples
- \( c_p \): Reference concentration
- \( \hat{c}_p \): Predicted concentration of sample \( p \)

RMSE for calibration samples (\( RMSEC \)) and validation samples (\( RMSEP \)) are calculated similarly:

\[ RMSEC = \sqrt{\frac{\sum_{p=1}^{N_c} (c_p - \hat{c}_p)^2}{N_c}} \]
\[ RMSEP = \sqrt{\frac{\sum_{p=1}^{N_v} (c_p - \hat{c}_p)^2}{N_v}} \]

The model is considered good if \( RMSEC \approx RMSEP \). Otherwise, the selection of the calibration/validation sets might be changed if \( RMSEC \ll RMSEP \).
OUTLINE

FROM QUALITATIVE TO QUANTITATIVE LIBS USING UNIVARIATE ANALYSIS

1) Preprocessing
2) Line identification
3) Signal extraction
4) Towards quantification

5) Conclusion
CONCLUSION

Data processing is a critical step to conduct quantitative LIBS analysis!

Now mastered

Allows to go back to pre-processing and measurement
CONCLUSION

• **Repeatability, intermediate precision and reproducibility** can now be evaluated.

• **Opportunity for the LIBS community to make a new step towards standardization via an inter-laboratory contest**, with «good» samples, and standardized data processing

• **Limit**: complex and/or un-prepared samples (heterogeneous, no matrix matched, etc.) induce additional complexity.
  → **Multivariate data analysis (Josette El Haddad, W2)**