

Accepted Manuscript

Multivariate calibration in Laser-Induced Breakdown Spectroscopy quantitative analysis: The dangers of a ‘black box’ approach and how to avoid them

A. Safi, B. Campanella, E. Grifoni, S. Legnaioli, G. Lorenzetti, S. Pagnotta, F. Poggialini, L. Ripoll-Seguer, M. Hidalgo, V. Palleschi



PII: S0584-8547(17)30354-3
DOI: [doi:10.1016/j.sab.2018.03.007](https://doi.org/10.1016/j.sab.2018.03.007)
Reference: SAB 5395

To appear in: *Spectrochimica Acta Part B: Atomic Spectroscopy*

Received date: 5 August 2017
Revised date: 4 March 2018
Accepted date: 14 March 2018

Please cite this article as: A. Safi, B. Campanella, E. Grifoni, S. Legnaioli, G. Lorenzetti, S. Pagnotta, F. Poggialini, L. Ripoll-Seguer, M. Hidalgo, V. Palleschi, Multivariate calibration in Laser-Induced Breakdown Spectroscopy quantitative analysis: The dangers of a ‘black box’ approach and how to avoid them. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Sab(2017), doi:[10.1016/j.sab.2018.03.007](https://doi.org/10.1016/j.sab.2018.03.007)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Multivariate calibration in Laser-Induced Breakdown Spectroscopy quantitative analysis: the dangers of a 'black box' approach and how to avoid them

A.Safi¹, B. Campanella¹, E. Grifoni¹, S. Legnaioli¹, G. Lorenzetti¹, S. Pagnotta¹,
F. Poggialini¹, L. Ripoll-Seguer², M. Hidalgo², and V. Palleschi¹

¹Applied and Laser Spectroscopy Laboratory, Institute of Chemistry of Organometallic Compounds, Research Area of National Research Council, Via G. Moruzzi, 1 – 56124 Pisa, Italy.

²Department of Analytical Chemistry and Food Science, University of Alicante, Apdo. 99, Alicante E-03080, Spain.

Abstract

The introduction of multivariate calibration curve approach in Laser-Induced Breakdown Spectroscopy (LIBS) quantitative analysis has led to a general improvement of the LIBS analytical performances, since a multivariate approach allows to exploit the redundancy of elemental information that are typically present in a LIBS spectrum. Software packages implementing multivariate methods are available in the most diffused commercial and open source analytical programs; in most of the cases, the multivariate algorithms are robust against noise and operate in unsupervised mode. The reverse of the coin of the availability and ease of use of such packages is the (perceived) difficulty in assessing the reliability of the results obtained which often leads to the consideration of the multivariate algorithms as 'black boxes' whose inner mechanism is supposed to remain hidden to the user. In this paper, we will discuss the dangers of a 'black box' approach in LIBS multivariate analysis, and will discuss how to overcome them using the chemical-physical knowledge that is at the base of any LIBS quantitative analysis.

Keywords: LIBS; Cast Iron; Calibration Curves, Partial Least Squares Analysis; Artificial Neural Networks

Corresponding author: Vincenzo Palleschi, e-mail: vincenzo.palleschi@cnr.it

1. Introduction

The procedures for quantitative analysis of materials by Laser-Induced Breakdown Spectroscopy (LIBS) technique are traditionally based, as in many other analytical applications, on the use of calibration curves [1]. With this term, we intend, in a broad sense, the determination of the unique correspondence between input data (derived from the LIBS spectra) and output results (the concentration of given element or elements in the sample). The dependence between the output(s) and input(s) can be obtained using a single input (in general, the intensity of a characteristic emission line of a given element) or exploiting several inputs (for example, the intensity of the LIBS spectrum at different wavelengths). In all the cases, for the building of the calibration curve, the outputs must be known quantities; for that purpose, reference materials of known composition (either certified or characterized by means of other analytical techniques) are typically used.

One of the specific problems related to LIBS analysis is the limited number of reference samples available. The presence of a strong matrix effect [2] in LIBS and the use, in most applications, of solid samples, makes less applicable the strategies used in other wet chemistry applications for increasing the number of reference samples, which in LIBS applications rarely exceeds a set of ten. It is also difficult the a priori choice of the set of standards, when the composition of the sample is not known, as typically occurs in analytical applications. Besides the matrix effect, self-absorption [3-5] of the LIBS emission lines often makes the calibration curves non-linear; moreover, the lack of reliable blanks might produce errors in the quantification of low-concentration elements and give unrealistic limits of detection (LODs) [6]. The guidelines for building univariate calibration curves in analytical chemistry were published by IUPAC in 1998 [7], although very few LIBS applications, if any, have complied with them in the last two decades.

A calibration approach (univariate or multivariate) starts from the hypothesis that the output(s) are linked to the input(s) by a relation through an analytical relation:

$$\mathbf{O} = f(\mathbf{I}) \quad (1)$$

where \mathbf{O} and \mathbf{I} are vectors of dimension m and n , respectively, where m represents the number of outputs (dependent variables) and n represents the number of inputs (predictors or independent variables). The function f linking the outputs with the inputs can be any multidimensional, possibly nonlinear function. The univariate case, corresponding to $m=n=1$, implies a relation between a single input and a single output (for example, in a LIBS spectrum, the link between the intensity of a given spectral line and the concentration of the corresponding element). The input, in this case, can be the value of the LIBS spectrum at a given wavelength (peak of the line) or the line integral or a predefined combination of spectral points (for example, the sum of the integral intensities of two or more lines from the same element). Whatever the choice for the input, this single value extracted from the LIBS spectrum is related to a single output (for example, the concentration of a given element) through the function $f(I)$. In this case, building a univariate calibration curve means the empirical determination of the function $f(I)$ through a best fitting procedure on a set of data obtained from samples whose output (composition) is already known and the input is obtained from a measurement done in the exact same conditions of the measurements on the unknown samples whose composition has to be determined. If the relation between inputs and outputs is linear, the best fitting function $f(I)$ can be easily determined through a linear fit of the data in an output vs. input plane. In fact, it is customary for univariate calibration to invert the role of the variables, considering the outputs (concentrations) as the independent variable and putting the values of the inputs (LIBS signal) on the y axis of the calibration curve plot. This apparent incongruence is due to the fact that the indetermination on the outputs is, in general, much lower than the experimental indetermination on the inputs. Since it is much easier to determine analytically the parameters of a linear fitting, using the minimum χ -square method, when the errors are on the dependent variable, the calibration curve is traditionally determined from the best fit of the inputs vs. outputs plot, and then the relation is inverted for the determination of the unknown outputs (concentrations) from the measured inputs (LIBS signal). It should be clear, however, that the calibration curve (or surface, in the multivariate case) is, by definition, the relation which links the experimentally determined inputs to the outputs, as in eq. 1.

This relation, as already said, is empirically determined from the analysis of a set of samples of known composition. The outputs would be, in the case of LIBS quantitative analysis, the concentration of one or more elements in the samples, which will be correlated to the LIBS signal coming from one or more spectral points in the corresponding LIBS spectrum.

In the univariate case, the quality of the calibration curve can be visually assessed and, in case of a linear calibration curve, also easily quantified through a series of parameters (R^2 , indetermination on the slope and intercept of the calibration line, etc.).

What is often neglected, in LIBS quantitative analysis, is the fact that having a good calibration curve does not automatically imply that the measurement of an unknown sample would be similarly good. This is true for both univariate and multivariate calibration approaches, although the problems become more difficult to spot in multivariate analysis. We will see in the following that the Partial Least Square (PLS) method, one of the most used multivariate analytical methods in LIBS, is designed for optimizing the factors of merit (linearity, precision, etc.) of the calibration surface on the calibration set. This optimization process is very specific of the calibration set used and does not imply in any way that the same linearity and precision would be obtained in the analysis of unknown samples. On the other hand, it is true that in a real LIBS analysis the only possible checks on the reliability of the calibration curve or surface are the ones performed using the known samples from the calibration set. Therefore, any verification other than the one implying the usual factors of merit of the calibration curve seems to be precluded. However, we should keep in mind that any calibration approach is based on a physical-chemical model of the relation between inputs and outputs. Our proposal is to exploit this idea to perform a reliability check of the calibration curves obtained through a multivariate approach exactly as we usually do (more or less unconsciously) in the analysis of a univariate calibration curve.

In other words, if our physical-chemical model implies a linear relation between input and outputs (for example, in the analysis of trace elements in a given matrix), from the saturation at high concentrations of a univariate calibration curve we can immediately spot a breach of the model, probably due to the presence of a non-negligible self-absorption of the emission lines considered. We will show in the following that a similar check can be (and must be, in our opinion) performed on the calibration surfaces obtained from multivariate methods.

2. Materials and Experimental setup

We will present, as an example of the application of different calibration approaches, the analysis of cast iron samples that was performed in the framework of an inter-laboratory test organized by the BAM Institute in Berlin, Germany. The laboratories who participated to the test were provided with seven reference samples, of known composition. Three additional samples (U1, U2, U3), of unknown composition, were given to the participants. The elements to quantify on the unknown samples were copper, nickel and chromium. The samples were provided in the form of powder or small chunks. For the analysis, the cast iron powders or chunks were put on double sticky tape.

Given the heterogeneous nature of the samples, we ruled out the possibility of using a classical LIBS setup with a relatively long focal focusing lens, in favor of a micro-LIBS setup recently realized in our laboratory [8]. The system is based on the Modì double-pulse instrument [9]; the laser pulses are sent through the Modì articulated arm into a Zeiss Axio microscope, and then focused onto the sample. The typical diameter of the craters produced on the sample is around 20 μm , using a 5+5 mJ energy of the laser pulses (about 20 ns FWHM). The irradiance on the sample was thus about $8 \times 10^{10} \text{ W/cm}^2$ per pulse. We used a delay between the pulses of 1 μs .

The LIBS signal was collected by an optical fiber, placed at 45 degrees with respect to the laser direction, at a distance of about 1 cm from the sample; a ball lens in front of the fiber guarantees the optimal collection of the LIBS signal from the whole plasma. The μ -Modì instrument uses an Avantes double spectrometer (AvaSpec-2048-2) which covers the spectral region from 190 to 900 nm (0.1 nm resolution from 190 to 450 nm, 0.3 nm resolution from 450 to 900 nm). We acquired the LIBS spectra 250 ns after the second laser pulse. The minimum acquisition time of the spectrometer is about 2 ms, much longer than the plasma lifetime; the spectra were therefore acquired in time-integrated mode.

For each sample, we acquired 10 LIBS spectra in 12 different points (120 spectra in total). The focus of the system was optimized at each point, and the spectra were saved singularly. A typical LIBS spectrum is shown in figure 1.

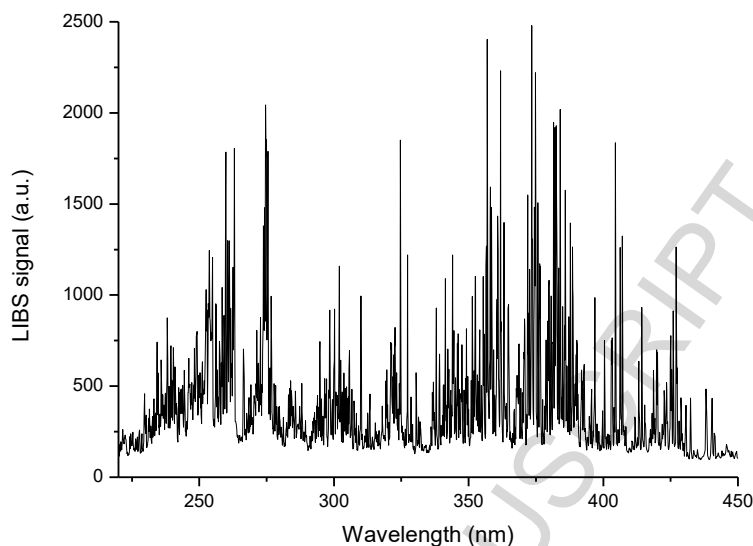


Figure 1 – A typical LIBS spectrum (detail) of the cast iron samples.

In spite of the care used for optimizing the laser beam focus at each point, the heterogeneity of the samples produced large variations between the LIBS spectra acquired on the same sample. For this reason, we decided to consider for the calibration only the spectra that passed a selection based on the total spectral intensity. For each sample, only the spectra having a total integrated intensity within 67% of the maximum would be retained. The spectra were then further normalized by their integrated intensity. After this selection, about 2/3 of the LIBS spectra were discarded (from the initial 840 spectra – 120 x 7 samples – only 265 were used for the analysis).

In the following, we will concentrate our discussion on the build and optimization of the calibration curve(s) for the determination of copper concentration. The reason of this choice is due to the large variability of copper concentration in the reference samples (from 0.0256 to 6.29 percent in weight, see Table I), which will give us the possibility of discussing about the non-linearity produced in the calibration curve by the self-absorption effect.

Table I – Cu content of the reference materials (the concentrations are in percent in weight), uncertainty at 95% confidence.

	Cu (w%)
C1	6.29 ± 0.02
C2	0.10 ± 0.01
C3	0.0256 ± 0.0004
C4	0.584 ± 0.013
C5	0.0303 ± 0.0006
C6	0.2445 ± 0.0025
C7	1.231 ± 0.008

Moreover, the presence in the LIBS spectra of two relatively close Cu I lines, of similar intensity (324.75 and 327.40 nm) would be useful in the discussion about multivariate calibration. Most of the arguments discussed in the following in the case of copper are, however, valid for the build of the calibration curves of other minor elements (Ni, Cr, Mn,...) in cast iron, too.

3. Building of the calibration curves

3.1 Univariate analysis

We will start our discussion from a classical univariate calibration curve approach. For building the univariate calibration curve of a minor element, it is customary in LIBS working on averaged spectra [10]. Although, from a statistical point of view, calculating the concentration of an element from the average LIBS spectrum is not the same as calculating the average concentration obtained from the single spectra, when the number of spectra is high, the choice of working with averaged spectra might be justified. The advantage in terms of elaboration time is evident, since in our case we should analyze only seven averaged spectra (one per sample), instead of 265 (one per shot). Moreover, the pre-selection operated on the spectra guarantees a good representativity of the average spectrum if compared with the single realizations.

In figure 2 we show the portion of the LIBS spectrum which comprises the two intense Cu I lines considered for the calibration.

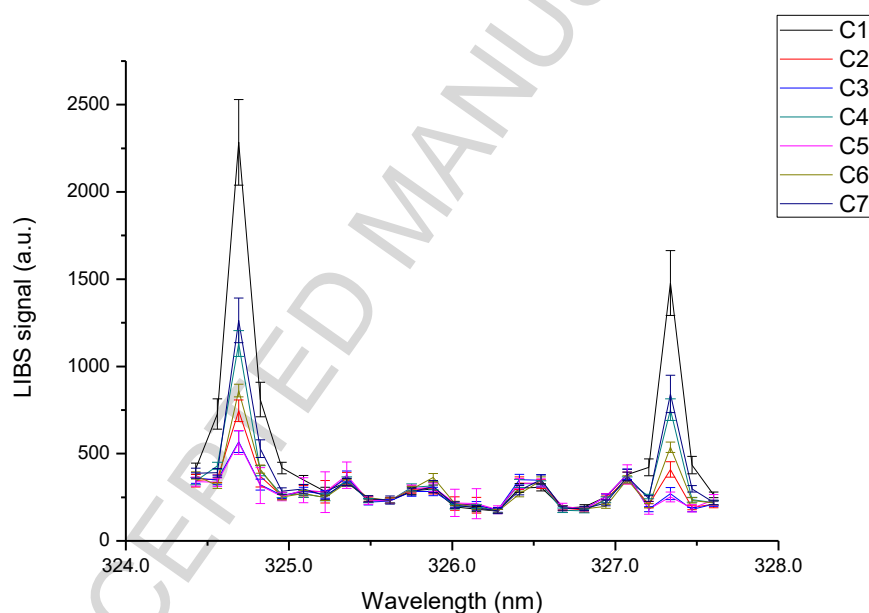


Figure 2 – The portion of the LIBS spectrum considered for the analysis of copper. The seven superimposed spectra corresponds to the reference samples as in the legend of the figure.

For each spectral point and each sample, in figure 2 we plotted the standard deviation of the LIBS signal, in the form of an error bar. One can easily see that the portion of the spectrum in between the two lines does not show any evident difference from sample to sample. Therefore, in the framework of a univariate analysis it would be reasonable to build a calibration curve for copper using either the peak value of one of the two Cu lines (or, for better results, the sum of the peaks of the two lines), after subtracting a constant background corresponding to the continuum under the curve or, alternatively, using as input the integral intensity obtained through a fit of one or both the Cu lines with a pseudo-Voigt profile [11]. In all the cases, a single input would be associated to a single output, and the univariate calibration curve would be calculated in the output vs. input plane.

The two calibration curves obtained using as input the sum of the peak values and the sum of the integrals of the two Cu lines are plotted in figure 3. Note that we used the conventional representation of the

calibration curve, with the outputs as independent variables, instead of the more general form with the LIBS intensities on the x-axis.

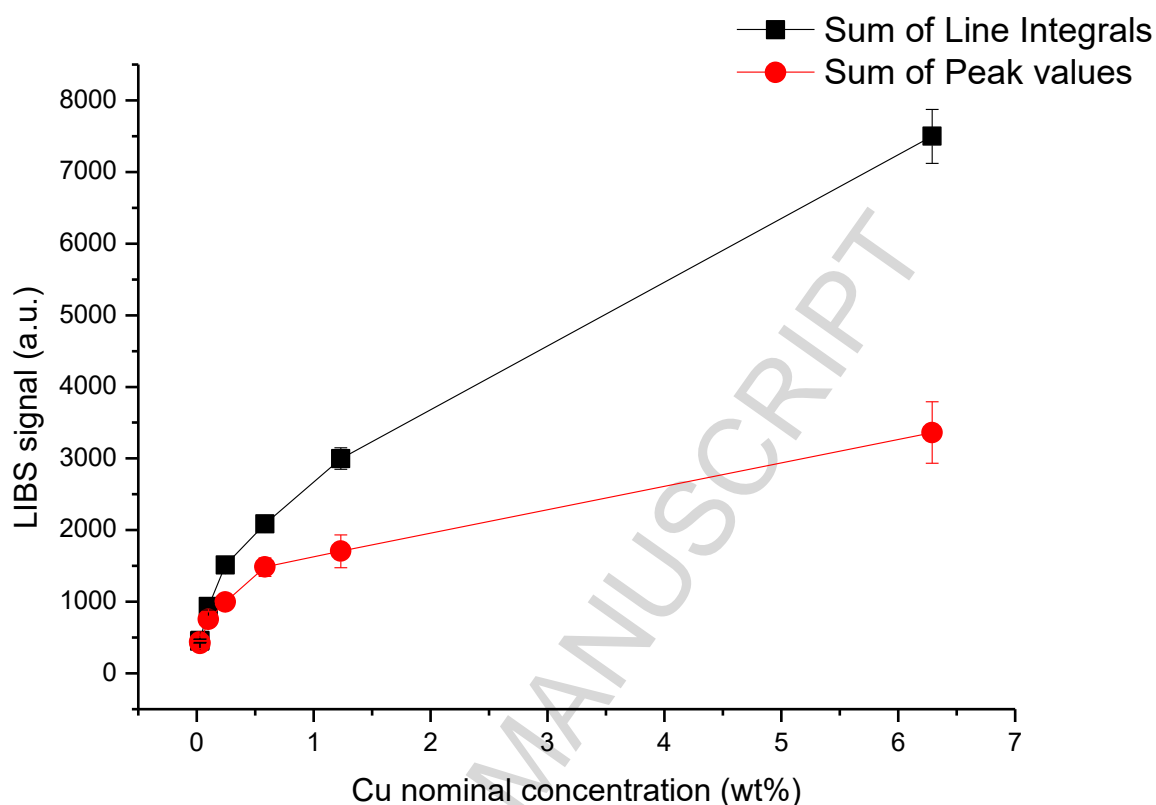


Figure 3 – Univariate calibration curves for Cu in cast iron. The black curve corresponds to the sum of the integral intensities, the red one to the sum of the peak intensities of the two Cu lines.

The effect of self-absorption is larger for the peak intensity than the line integral intensity. In fact, while the effect of self-absorption acts linearly on the peak intensity, the integral intensity is only affected as the square root of the self-absorption coefficient [12]. It is thus preferable, from an analytical point of view, to use the curve based on the integral value of the lines. Although the use of moderately self-absorbed lines has been recently suggested for improving the precision of LIBS analysis [13-15], these methods are based on a completely different approach with respect to the one used for building calibration curves, as recently discussed in ref. [16], and will not be considered for the present analysis.

We note from figure 3 that the linear part of the calibration curve (in the following, we will consider only the one calculated from the integral of the two Cu I lines) is quite limited. As the Cu concentration increase over 0.1 wt %, the curve starts to deviate from the linear behavior (see the detail of figure 3 reported in figure 4).

Another effect which is clearly visible in figure 4 is the fact that the intercept of the calibration curve at zero Cu concentration is not zero. This is probably due to the presence of some spectral interference, possibly related to the presence of weak iron lines underlying the Cu lines.

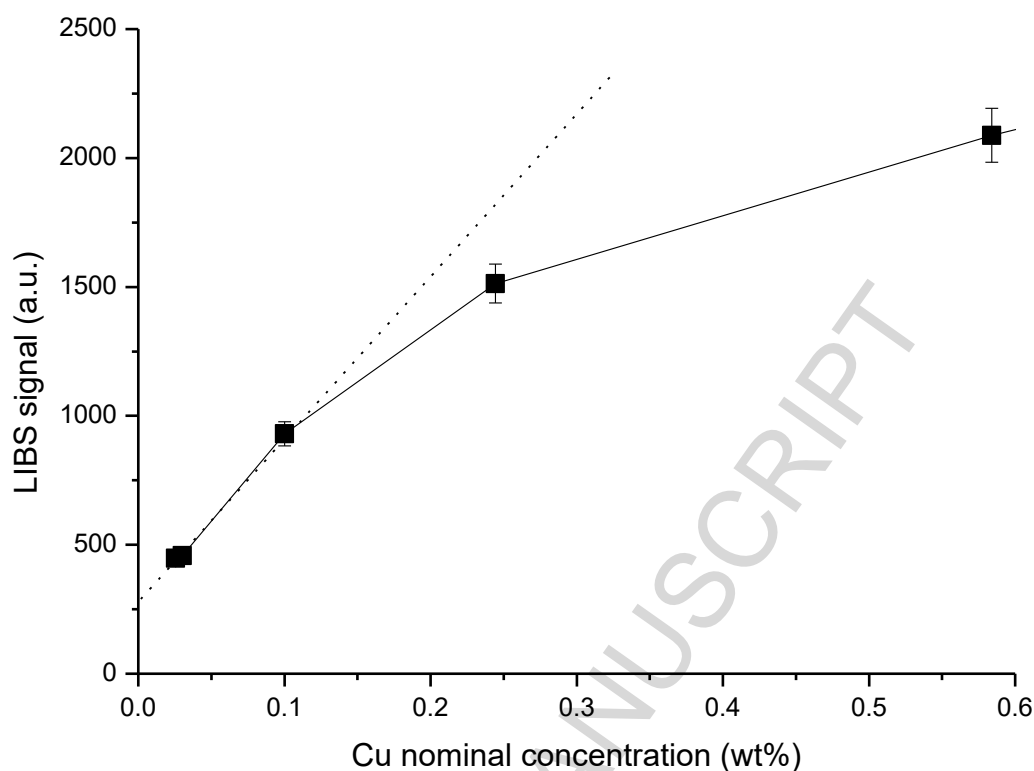


Figure 4 – Univariate calibration curve for Cu in cast iron. The LIBS intensity was calculated as the sum of the integral intensities of the two Cu lines.

The non-linearity of the univariate calibration curve does not prevent its use for the quantitative determination of the Cu content in the unknown samples; however, the lack of linearity and the indetermination on the values of the intercept do not allow the use of the standard formulas for the Limit of Detection and indetermination of the results that are derived under the assumptions of linearity and fixed intercept. From an analytical point of view, it must be noted that the error in the determination of the intercept of the calibration curve, as pointed out by Winefordner and Long [17] several years ago, must be considered for the determination of the limits of detection (LODs) along with the classical 3σ criterion.

3.2 Multivariate analysis – Partial Least Squares Analysis

Using the peaks or the integral value of the Cu lines for the determination of the univariate calibration curve implies the use of only a part of the information that can be extracted from the LIBS spectra. For example, in the calculation of the line integral intensity, all the spectral points contribute in the same way to the final value of the integral. However, we know from the transport equation of radiation that the peak of the emission curve is more affected by self-absorption than the tails of the curve (this is the reason why the self-absorption effect is somewhat mitigated using the integral of the line, instead of the peak). It seems thus logic that building a calibration curve using a linear combination of the LIBS spectral components could reduce the non-linearity of the calibration curve by increasing the importance of the spectral components more correlated with the element concentration, and proportionally reducing the influence of the components more affected by the self-absorption effect.

Partial Least Square Analysis (PLS) is a multivariate technique which produces a vector of weights which, scalarly multiplied with the spectral components, would maximize the linear correlation between the

nominal and predicted concentrations of the element of interest. As discussed in ref. [16], many commercial or open source statistical packages provides simple routines for the calculation of the weights that link the spectral components of the LIBS spectrum to the corresponding element concentration, in the form:

$$C = a + \sum_{i=1}^N w_i S(\lambda_i) \quad (2)$$

where $S(\lambda_i)$ are the N spectral components of the LIBS spectrum (we can call them independent variables, predictors or inputs, depending on the kind of multivariate approach used) and the concentration C is the dependent variable (output or target).

The PLS algorithm is very fast and runs automatically, without any intervention of the operator. For this reason, it is possible to apply the calculation to all the single spectra selected (265); this would give us a better idea of the indetermination involved in this kind of calculations.

The relation between nominal and calculated Cu concentration, obtained using all the spectral components in the interval in figure 2, is shown in figure 5. Each point corresponds to a different spectrum.

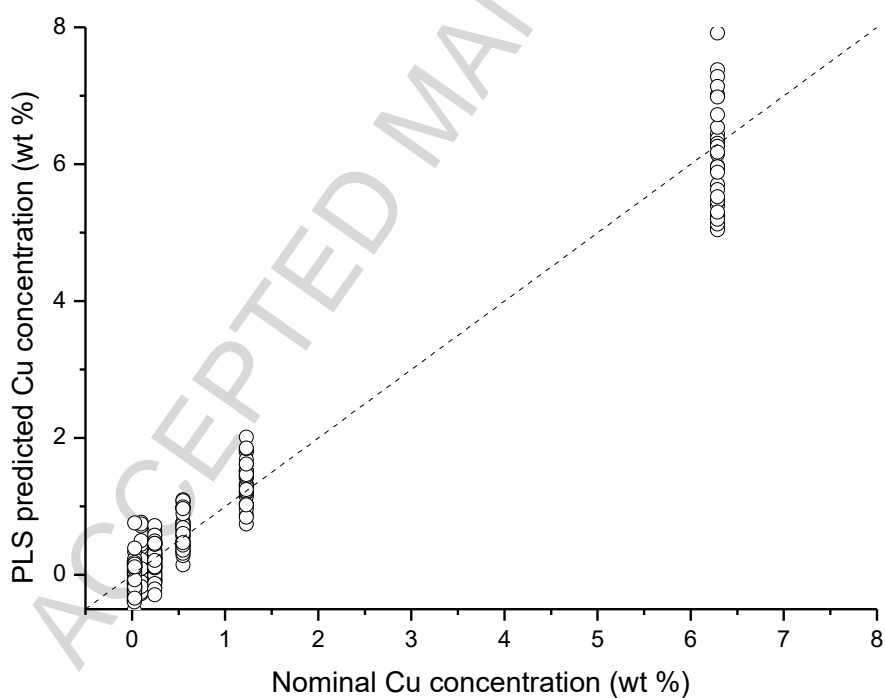


Figure 5 – Relation between nominal and PLS predicted Cu concentration for the seven samples considered. The dashed line corresponds to a perfect correlation between nominal and predicted concentrations.

In analyzing the results of figure 5, first we must note that the curve represented is not a calibration curve. The linearity and almost perfect correlation between nominal and predicted Cu concentrations does not mean that the method is necessarily more accurate than the univariate calibration curve. The risk always present in multivariate analysis is the possibility of overfitting the data, which results in a very good capability of the method to reproduce the data used for the calibration and in a corresponding very poor capability of predicting the concentration of unknown samples, not used for building the (multivariate) calibration curve. An *a posteriori* check of the goodness of the relation found between output and inputs

could be performed using the leave-one-out method [18]. With this method, in turn each sample is removed from the calculation of the multivariate calibration, and is treated as an unknown sample whose concentration has to be determined by the PLS method. However, since the samples with the lowest and highest Cu concentration cannot be used as unknowns (the basic rule on the use of uni- or multivariate calibration curve is that the concentration of the unknown samples must be comprised between the highest and the lowest concentration of the same element in the calibration samples), in our case only five samples over seven can be used for the leave-one-out test.

In figure 6 are shown the results of the test for the five samples comprised between the maximum and minimum of Cu concentration.

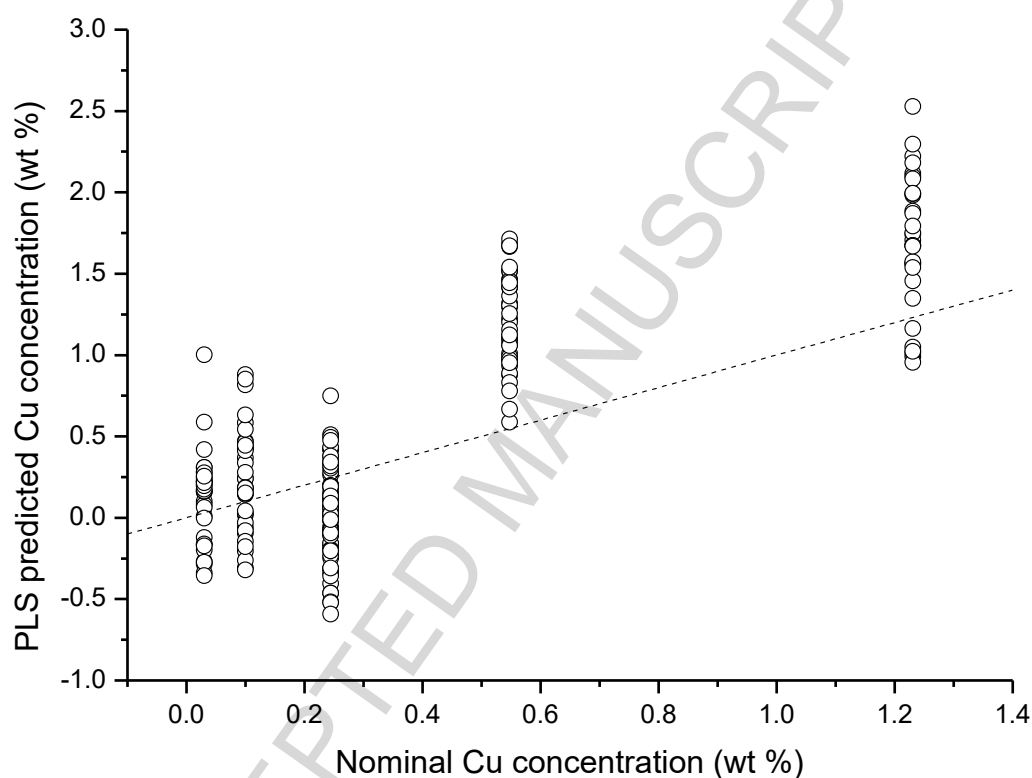


Figure 6 – Relation between nominal and PLS predicted Cu concentration evaluated after the leave-one-out procedure. The dashed line corresponds to a perfect correlation between nominal and predicted concentrations.

As we can see, after the leave-one-out procedure, the agreement between nominal and predicted concentrations appears (realistically) worse with respect to what one could have been deduced from the analysis of figure 5.

From eq. 2 we know that the Cu concentration is given by the scalar product of the LIBS spectral components times the vector of weights (plus a constant). For giving the same importance to all the spectral points, independently on their intensity, the PLS algorithm performs a scaling of the spectral points is usually performed, which are centered with respect to their average (over all the samples) and normalized by their standard deviation. The unwanted side effect of this procedure is the possible occurrence of casual correlation between a given input and the output, which could associate, as we observed, a high weight to points that, from a chemical-physical point of view, should not be correlated at all with the output. On the other hand, the weights associated to spectral points at low intensity would not affect too much the scalar product in eq. 2; a way for checking the reliability of the PLS results is thus plotting the normalized w_i (weight times the standard deviation of the corresponding spectral component) at the different wavelengths, as shown in figure 7. We observe from the figure that the only spectral points

with non-negligible normalized weights are the one corresponding to the two main Cu emission lines. Therefore, only this spectral information is important for the determination of the Cu concentration, as expected.

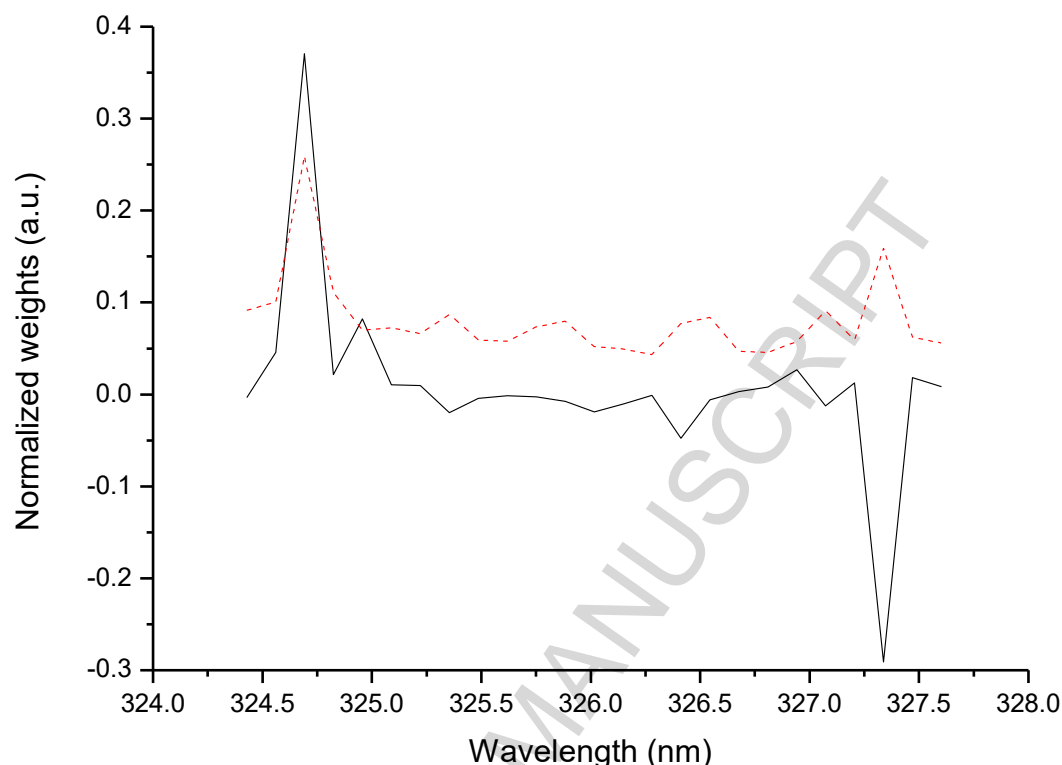


Figure 7 – Normalized weights obtained by PLS analysis (black). The red dashed curve superimposed corresponds to the LIBS spectrum (averaged over all the 265 spectra used for the calibration) in the region considered for the calculation.

On the other hand, we can see from figure 7 that the peak intensity of the Cu line at 327.40 nm is actually anti-correlated with the Cu concentration (negative normalized weight) in the calibration curve given by the PLS algorithm. This clearly makes no sense from a chemical-physical point of view. The main problem associated to the PLS results shown above is the attempt made by the algorithm to find a linear relation between nominal and predicted concentrations, in the presence of a strong non-linearity produced by the self-absorption effect. The PLS algorithm finds the best compromise between these two contrasting requirements by subtracting the intensity of the Cu line at 327.40 nm from the one at 324.75 nm. The ‘trick’ works relatively well on the seven samples used for the calibration, but it might fail dramatically on an unknown sample.

For obtaining a more realistic description of the linear part of the curve, corresponding to Cu concentrations lower than 0.5%, we should thus remove from the PLS calibration set the samples corresponding to high Cu concentration.

In figure 8 we show the equivalent of figure 7, obtained removing from the calibration set the samples with Cu at higher concentrations (samples C1 and C7).

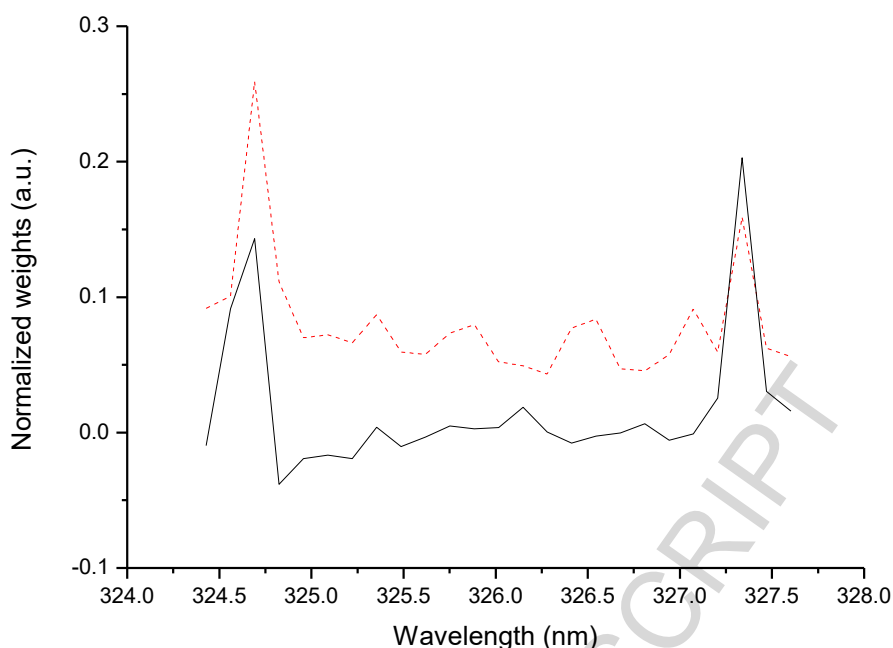


Figure 8 – Normalized weights obtained by PLS analysis (black) after removing the high Cu concentration samples from the calibration set. The red dashed curve superimposed corresponds to the LIBS spectrum (averaged over all the 265 spectra used for the calibration) in the region considered for the calculation.

We can observe that the vector of the normalized weights w_i now shows a behavior that could be expected, with both the Cu lines peak intensity positively correlated with the Cu concentration. However, the effect of the Cu line at 327.40 nm is over-estimated, with respect to the one of the most intense Cu line at 324.75 nm.

The leave-one-out check of the predictions of PLS approach obtained with the high Cu concentration sample removed from the dataset is shown in figure 9.

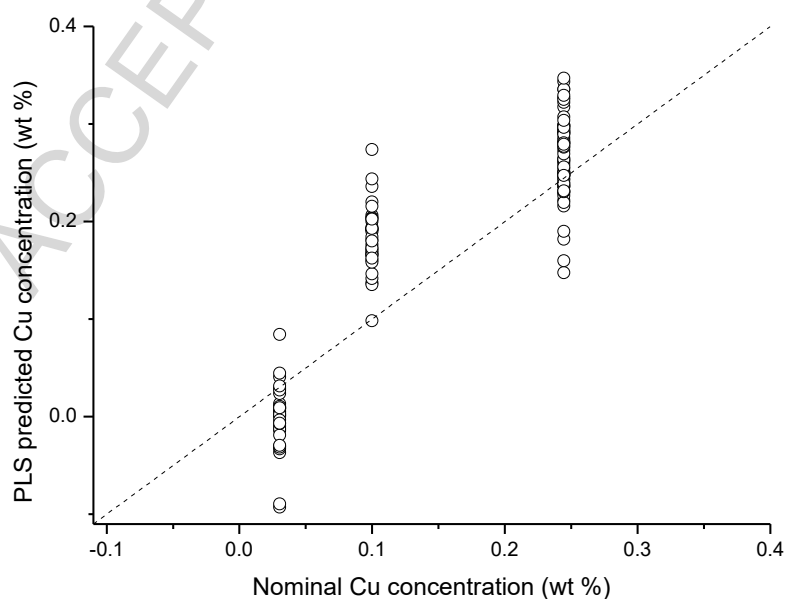


Figure 9 – Relation between nominal and PLS predicted Cu concentration evaluated after the leave-one-out procedure. The dashed line corresponds to a perfect correlation between nominal and predicted concentrations.

Figure 9 should be compared with the corresponding figure 6. We observe that the predictive capability of the PLS method at low Cu concentration improves after removing from the calibration dataset the samples which are more affected by the self-absorption effect. However, the overall correspondence between nominal and predicted Cu concentration is still rather poor and limited to low Cu concentrations, where the effect of self-absorption is almost negligible.

3.3 Multivariate analysis – Artificial Neural Networks

Another multivariate calibration approach that has become popular in LIBS quantitative analysis is based on the use of Artificial Neural Networks (ANN) [19-23]. The ANN approach is an extremely powerful algorithm; however, as for many other statistical approaches, the structure of the network has to be carefully considered for avoiding issues that might compromise the predictive capabilities of the method. One of the most common errors in the application of ANN to LIBS quantitative analysis is the evaluation of the network quality based on the correlation between predicted and nominal concentrations on the calibration dataset. As we already observed in the case of PLS, the correlation between predicted and nominal concentration can be extremely good (see figure 5), while the actual prediction capabilities of the method can be much poorer (figure 6). In fact, when the number of reference samples is limited, as in our case, the risk of overfitting the data is quite high. As in PLS analysis, the ANN method connects the inputs (independent variables, in our case the spectral components of the LIBS spectrum in the selected region of interest) to the outputs (in our case, the concentration of Cu in the samples). However, while in PLS the connection between inputs and outputs is supposed to be linear (eq. 1), the ANN approach in principle takes into account the possibility of a non-linear relationship between independent and dependent variables.

The introduction of a possible non-linear dependence between inputs and outputs is an obvious improvement with respect to the PLS approach; however, the intrinsic complexity of the ANN method often makes difficult a check of the chemical-physical consistence of the results obtained.

Most commercial and open source statistical packages provide algorithms for ANN analysis. One of the most diffused system is the ANN package included in Matlab®, which provides to the user a step-by-step guided procedure that allows to train and apply a Neural Network to any set of data, without the need of a specific experience and knowledge of the method. In most cases, however, the default settings proposed are not adequate for obtaining meaningful results.

Let us consider the architecture of a layered feedforward ANN, as implemented in Matlab®. The default settings for connecting N independent inputs to M outputs use a single hidden layer with 10 neurons; the outputs are linked to the inputs as follows:

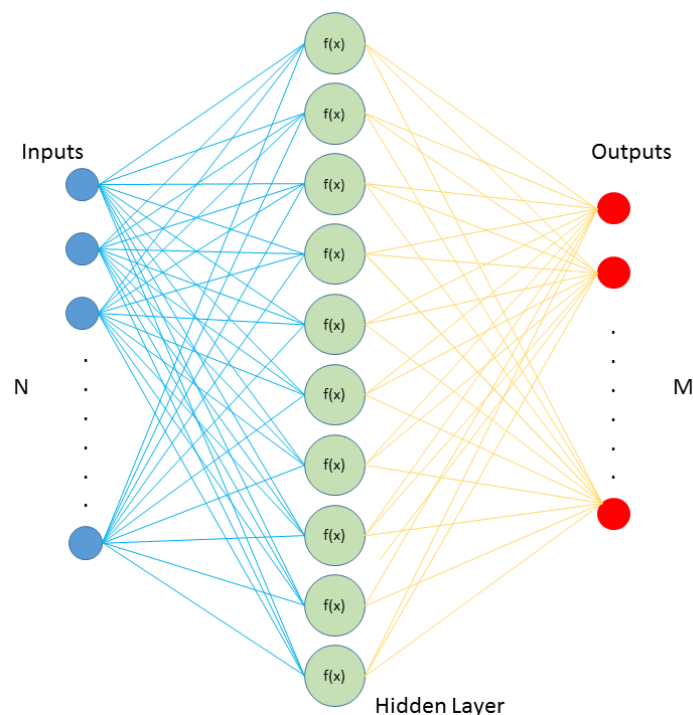


Figure 10 – Structure of the default ANN used in Matlab®.

The link between inputs and outputs is realized as follows (in the example, we will specialize the discussion to our experimental situation, where the inputs are the N spectral components of the LIBS spectra $S(\lambda_i)$ in the region considered and the output ($M=1$) is the corresponding Cu concentration C . First, the inputs are connected to all the hidden layer neurons in a linear way:

$$X_j = b_j + \sum_{i=1}^N w_{ij} S(\lambda_i) \quad (3)$$

The hidden layer neurons modify the X_j values by applying a function:

$$Y_j = f(X_j) \quad (4)$$

(by default, Matlab® uses a sigmoidal function $f(x) = \frac{2}{(1+\exp(-2x))} - 1$).

Then, the Y_j values are linearly combined to give the output (in our case, the copper concentration):

$$C = c + \sum_{j=1}^{10} W_j Y_j \quad (5)$$

The training of the ANN is the equivalent of the building of a univariate calibration curve or the determination of the weights in PLS. The default algorithm used in Matlab® determines the weights and the biases by minimizing the distance between the predicted and nominal outputs, using a Levenberg–Marquardt algorithm [24].

The Matlab® default parameters are intended to cover and reproduce very complex situation, where a high non-linear dependence of the outputs from the inputs is expected. However, in our case we expect to have a much simpler and regular connection between the LIBS spectral components and the Cu concentration

(ideally, in the absence of self-absorption, we would expect a perfect linear dependence between inputs and outputs). It is thus clear that using the Matlab® defaults for the number of neurons in the hidden layer would only increase the risk of overfitting. In fact, the proper optimization of the network would imply a series of tests varying of the number of hidden neurons [25]. This procedure is time-consuming, and totally empirical.

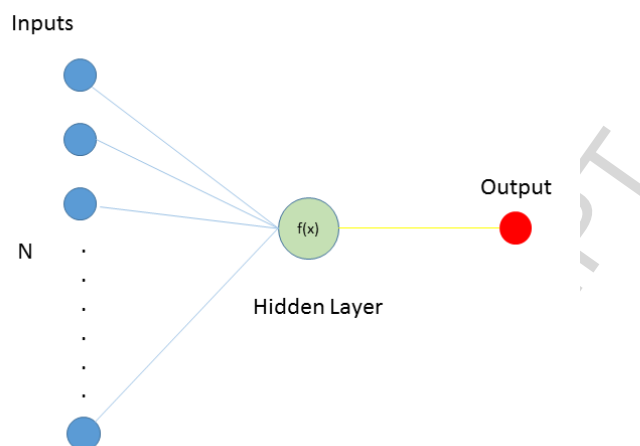


Figure 11 – The simplified ANN used for the determination of Cu concentration in cast iron.

On the other hand, one could consider that a drastic reduction of the number of hidden neurons from 10 to just one (figure 11) would result in a relation between the inputs and the output in the form of:

$$\begin{aligned} X &= a + \sum_{i=1}^N w_i S(\lambda_i) \\ C &= c + W f(X) \end{aligned} \quad (6)$$

which would become identical to eq. (1) of PLS when the function $f(x)$ is linear.

Therefore, with the introduction of a single non-linear term we expect to be able to reproduce the chemical-physical behavior of our system without introducing in the model any additional free parameter that would negatively affect the performances of the network.

As seen in figure 12, the correlation between nominal and predicted Cu concentration, on the calibration dataset, is very good and better than the one obtained using the PLS method (figure 6). However, we have already seen that the good correlation obtained on the calibration dataset does not necessarily imply a corresponding good predictivity of the method when unknown samples are analyzed.

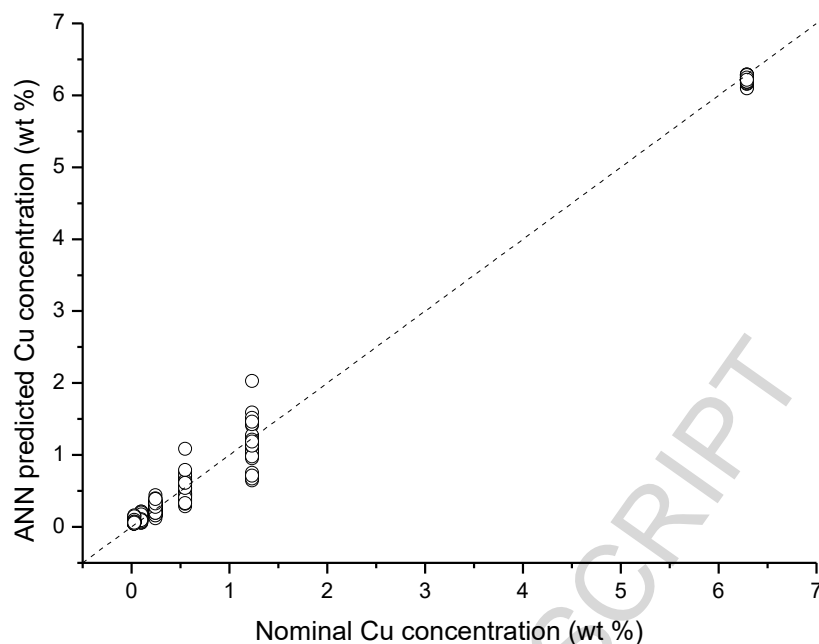


Figure 12 – Correlation between nominal and predicted concentration of Cu using the ANN method. The dashed line corresponds to a perfect correlation between nominal and predicted concentrations.

Figure 13 shows the results of a leave-one-out validation of the ANN results. We see that in this case the performances of the ANN remain acceptable, with a linear correlation between predicted and nominal Cu concentration that extends up to a Cu concentration higher than 1%, although with a relatively large variability from spectrum to spectrum.

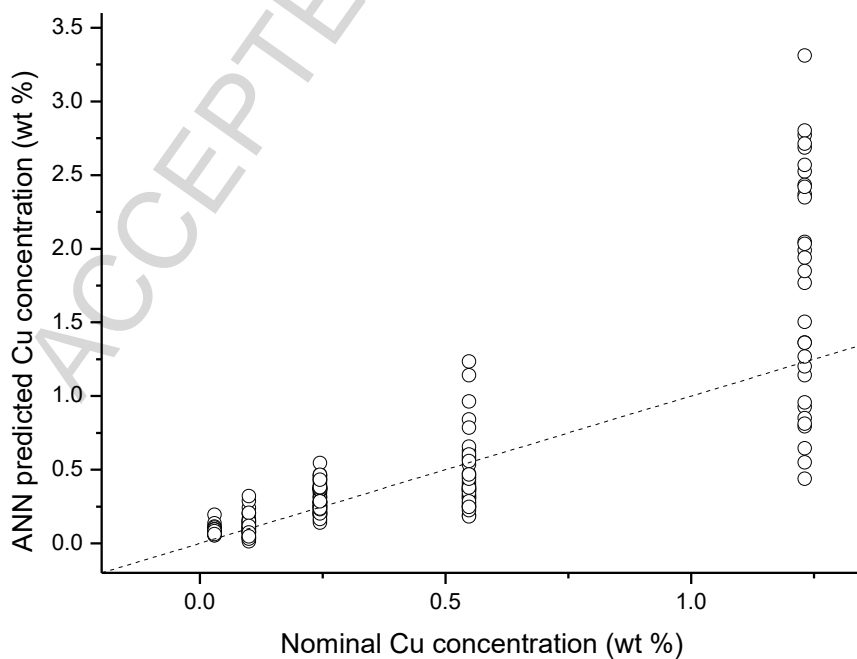


Figure 13– Correlation between nominal and predicted concentration of Cu after the leave-one-out procedure. The dashed line corresponds to a perfect correlation between nominal and predicted concentrations.

As in the case of PLS, we can obtain from the ANN the normalized weights w_i that link the inputs to the hidden neuron (eq. 6). These weights are shown in figure 14.

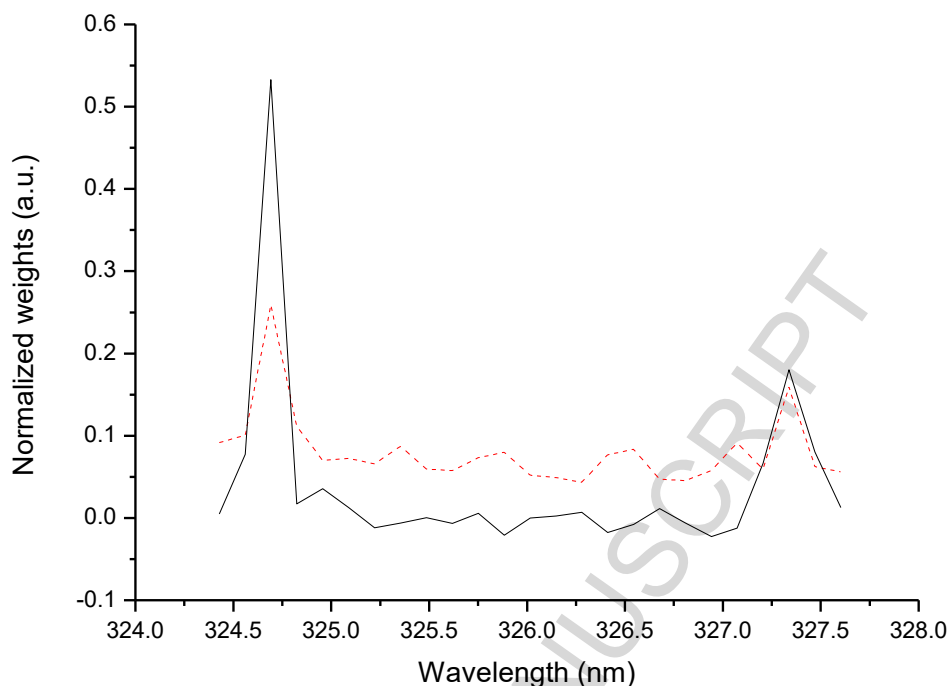


Figure 14 –Normalized weights obtained by the Artificial Neural Network method before the application of the non-linear transfer function. The red dashed curve superimposed corresponds to the LIBS spectrum (averaged over all the 265 spectra used for the calibration) in the region considered for the calculation.

In this case, contrarily to what we have observed in the application of the PLS method, the peak of the main Cu I line is associated to the higher weight. In fact, the effect of self-absorption is mainly corrected by the neuron in the hidden layer. For the samples with Cu at low concentration, the value of X would remain in the linear zone of the transfer function X ; at higher concentrations (higher X values), on the other hand, the non-linearity of the transfer function would compensate the saturation effect due to self-absorption. The ANN method thus provides a calibration curve that naturally changes from linear to non-linear with the increase of the copper concentration.

4. Determination of the best calibration approach

In this work, we have discussed different approaches (univariate calibration, Partial Least Squares Analysis and Artificial Neural Network) that can be used for the determination of the concentration of a minor element (in our case, copper) in solid (cast iron) samples. We have shown that each method requires an optimization of the analytical strategy and a careful choice of the samples used for calibration. In particular the multivariate approaches, based on linear correlations as the Partial Least Square analysis or allowing non-linear relations as in Artificial Neural Networks, require a special attention to the validation of the results obtained. At the end of the discussion, one could wonder which approach, among the ones described, provides the best analytical results on the unknown samples distributed for the inter-laboratory test. The discussion on the results of the test is not in the scope of the present work (a full report will be published by the colleagues at BAM who organized the test and collected the results). For the purposes of our discussion, however, the disclosure of the information about samples' composition which followed the end of the test give us an occasion to introduce another important issue related to the advantages and drawbacks of using a multivariate calibration approach instead of the usual univariate calibration curve method. A naïve approach to the problem would result in the consideration that an Artificial Neural Network (which would take into account all the spectral information and would naturally manage the non-

linearity of the calibration surface) would perform better than PLS (which, on the contrary, does not cope well with the non-linearity of the calibration surface); the same consideration will place the univariate calibration approach at the bottom of the rank. However, the performances of a calibration approach (univariate or multivariate) depends dramatically on how the calibration set is chosen with respect to the unknown samples to analyse. In other words, a simple univariate calibration approach can perform well (even better than a multivariate approach) if the input parameter is chosen on the basis of a sound chemical-physical model of the system under analysis.

In our case, the three unknown samples showed LIBS signals corresponding to the almost linear zone of the univariate calibration curve (see figure 4 and figure 15)

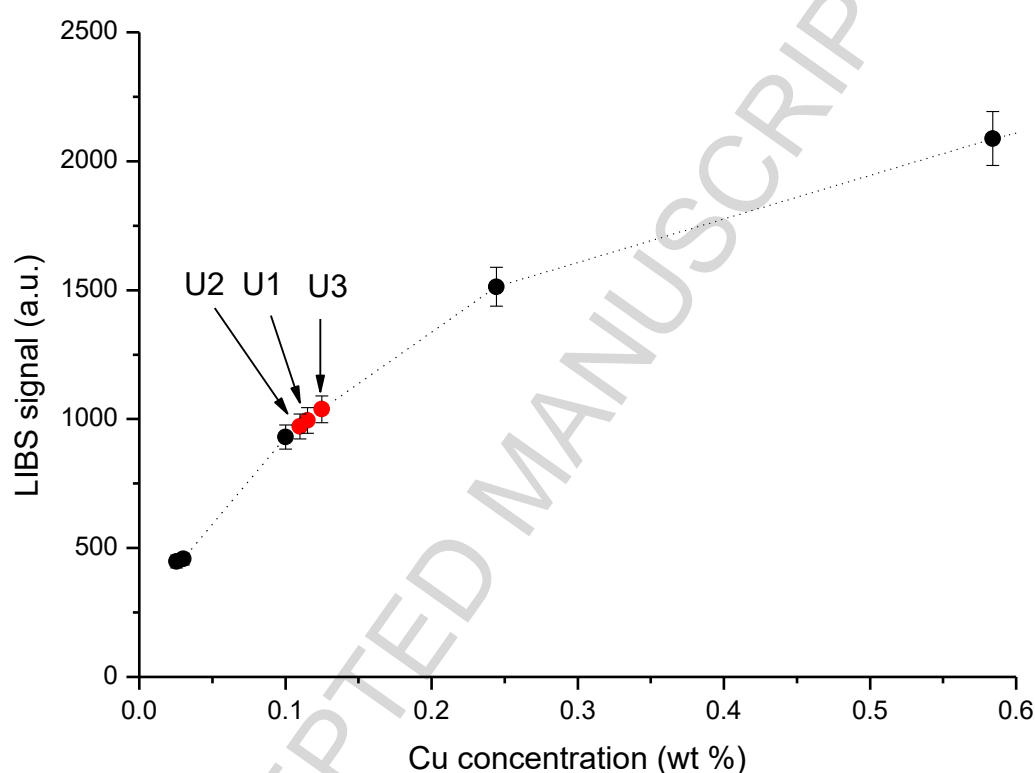


Figure 15 – Measurement of the Cu concentration in the unknown samples using the univariate calibration curve shown in figure 5

In Table II, the results obtained using the multivariate PLS and ANN approach are compared with the ones obtained using the univariate calibration. The nominal Cu concentrations, disclosed by the colleagues at BAM at the end of the test, are also reported in the table.

Table II – Cu concentration in the unknown samples, determined using univariate calibration, PLS and ANN, compared with the nominal concentration (certified by BAM, Berlin, Germany). Recovery, bias and z-scores are also reported for each method and sample.

	Univariate calibration curve				PLS				ANN				Nominal concentration
	Cu (wt %)	R (%)	b (wt %)	$ \zeta $	Cu (wt %)	R (%)	b (wt %)	$ \zeta $	Cu (wt %)	R (%)	b (wt %)	$ \zeta $	Cu (wt %)
U1	0.115 ± 0.045	90 ± 35	-0.01 ± 0.05	0.28	0.207 ± 0.027	162 ± 21	0.08 ± 0.03	2.93	0.118 ± 0.033	92 ± 26	-0.01 ± 0.03	0.29	0.1276 ± 0.0019
U2	0.110 ± 0.045	73 ± 30	-0.04 ± 0.05	0.89	0.240 ± 0.038	160 ± 26	0.09 ± 0.04	2.36	0.121 ± 0.047	81 ± 31	-0.03 ± 0.05	0.61	0.150 ± 0.004
U3	0.125 ± 0.045	83 ± 30	-0.03 ± 0.05	0.55	0.217 ± 0.021	145 ± 15	0.07 ± 0.02	3.13	0.133 ± 0.030	89 ± 20	-0.02 ± 0.03	0.56	0.150 ± 0.004

From a metrological point of view, the univariate approach is substantially equivalent to ANN in terms of recovery, calculated as:

$$R = \frac{x_{exp}}{x_{ref}} \times 100 \quad (7)$$

with uncertainty

$$u_R = R \times \sqrt{\left(\frac{u_{x_{exp}}}{x_{exp}}\right)^2 + \left(\frac{u_{x_{ref}}}{x_{ref}}\right)^2} \quad (8)$$

and bias, calculated as

$$b = x_{exp} - x_{ref} \quad (9)$$

with uncertainty

$$u_b = \sqrt{(u_{x_{exp}})^2 + (u_{x_{ref}})^2} \quad (10)$$

It is evident from the data that the univariate approach gives a satisfactory estimate of the Cu concentration in the three samples, which is only marginally improved by the use of a multivariate approach based on ANN. As we already observed in the discussion of the PLS approach, on the other hand, we see that the method systematically overestimates the Cu concentration, even when the samples with Cu at higher concentrations (samples C1 and C7) are removed from the calibration set. The leave-one-out check (figure 9) was indeed showing that the Cu concentration in the calibration samples around 0.1 wt% was overestimated by about a factor of two.

The z-score, a parameter which gives a measure of the goodness of the analysis, is defined as

$$\zeta = \frac{x_{exp} - x_{ref}}{\sqrt{(u_{x_{exp}})^2 + (u_{x_{ref}})^2}} = \frac{b}{u_b} \quad (11)$$

$|\zeta|$ is less than 1 for all the samples using the univariate and ANN approach, while in the PLS approach it is greater than 2 for samples U1 and U2 and greater than 3 for sample U3.

In the analysis of Proficiency Testing Schemes is customary to consider z-scores $|\zeta| \leq 2$ as 'satisfactory', while $2 < |\zeta| < 3$ are defined as 'questionable' and $|\zeta| > 3$ are considered 'unsatisfactory' [26].

5. Conclusion

The multivariate approaches, based on linear correlations as the Partial Least Square analysis or allowing non-linear relations as in Artificial Neural Networks, require a particular attention to the validation of the results obtained. In most of the situations, the structure of the statistical model used is simple enough to allow a direct check of the chemical-physical reliability of the model. Generally speaking, any increase in the complexity of the analytical model, from univariate to linear or non-linear multivariate, should be justified only in view of the complexity of the chemical-physical system under study. In the problem discussed in this work, a simple univariate calibration strategy is adequate for the determination of Cu concentrations around 0.1-0.2 wt %. At higher concentrations, the non-linear effects due to self-absorption would call for the use of more complex multivariate methods; a simplified ANN, with a single neuron in the hidden layer

and a single output, corresponding to the Cu concentration, would be probably able to provide a reliable evaluation of the copper concentration in the cast iron unknown samples.

Acknowledgement

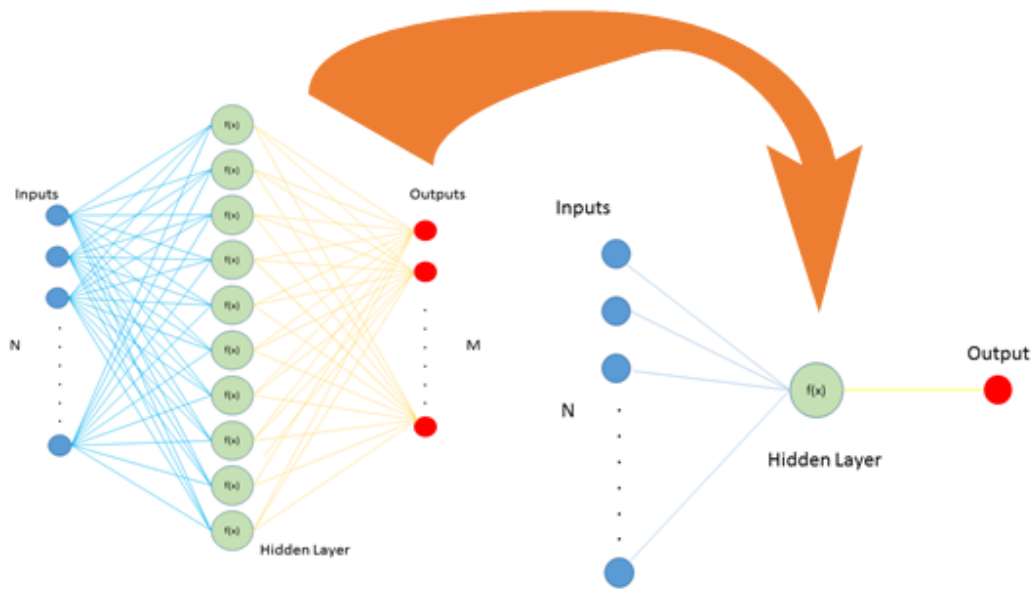
The authors would like to thank Dr. Igor Gornuskin and Dr. Wolfram Bremser from BAM, Berlin, Germany, for providing them the cast iron certified samples.

ACCEPTED MANUSCRIPT

References

- [1] Hahn, D.W., Omenetto, N., Laser-induced breakdown spectroscopy (LIBS), part II: Review of instrumental and methodological approaches to material analysis and applications to different fields (2012) *Appl. Spectrosc.* 66 (4), pp. 347-419.
- [2] Hahn, D.W., Omenetto, N., Laser-induced breakdown spectroscopy (LIBS), part I: Review of basic diagnostics and plasmaparticle interactions: Still-challenging issues within the analytical plasma community (2010) *Appl. Spectrosc.* 64 (12), pp. 335A-366A.
- [3] Aragón, C., Bengoechea, J., Aguilera, J.A., Influence of the optical depth on spectral line emission from laser-induced plasmas, *Spectrochim. Acta B* 56 (6), pp. 619-628,
- [4] Bulajic, D., Corsi, M., Cristoforetti, G., Legnaioli, S., Palleschi, V., Salvetti, A., Tognoni, E., A procedure for correcting self-absorption in calibration free-laser induced breakdown spectroscopy (2002) *Spectrochim. Acta B* 57 (2), pp. 339-353.
- [5] Gornushkin, I.B., Anzano, J.M., King, L.A., Smith, B.W., Omenetto, N., Winefordner, J.D., Curve of growth methodology applied to laser-induced plasma emission spectroscopy (1999) *Spectrochim. Acta B* 54, pp. 491-503.
- [6] Nomenclature, symbols, units and their usage in spectrochemical analysis II (1978) *Spectrochim. Acta B* 33B, pp. 242-245.
- [7] K. Danzer, K., Currie L.A., Guidelines for calibration in analytical chemistry. Part I. Fundamentals and single component calibration (IUPAC Recommendations 1998) (1998) *Pure & App. Chem* 70 (4), pp. 993-1014.
- [8] Schiavo, C., Menichetti, L., Grifoni, E., Legnaioli, S., Lorenzetti, G., Poggialini, F., Pagnotta, S., Palleschi, V., High-resolution three-dimensional compositional imaging by double-pulse laser-induced breakdown spectroscopy (2016) *Journal of Instrumentation* 11, C08002.
- [9] Bertolini, A., Carelli, G., Francesconi, F., Francesconi, M., Marchesini, L., Marsili, P., Sorrentino, F., Cristoforetti, G., Legnaioli, S., Palleschi, V., Pardini, L., Salvetti, A., Modì: A new mobile instrument for in situ double-pulse LIBS analysis (2006) *Analytical and Bioanalytical Chemistry* 385 (2), pp. 240-247.
- [10] Tognoni, E., Palleschi, V., Corsi, M., Cristoforetti, G., Quantitative micro-analysis by laser-induced breakdown spectroscopy: A review of the experimental approaches (2002) *Spectrochim. Acta B* 57 (7), pp. 1115-1130.
- [11] Olivero, J.J., Longbothum, R.L., Empirical fits to the Voigt line width: A brief review (1977) *Journal of Quantitative Spectroscopy and Radiative Transfer*, 17 (2), pp. 233-236.
- [12] El Sherbini, A.M., El Sherbini, Th.M., Hegazy, H., Cristoforetti, G., Legnaioli, S., Palleschi, V., Pardini, L., Salvetti, A., Tognoni, E., Evaluation of self-absorption coefficients of aluminum emission lines in laser-induced breakdown spectroscopy measurements (2005) *Spectrochim. Acta B* 60 (12), pp. 1573-1579.
- [13] Cristoforetti, G., Tognoni, E., Calculation of elemental columnar density from self-absorbed lines in laser-induced breakdown spectroscopy: A resource for quantitative analysis (2013) *Spectrochim. Acta B* 79-80, pp. 63-71.

- [14] Aragón, C., Aguilera, J.A., CSigma graphs: A new approach for plasma characterization in laser-induced breakdown spectroscopy (2014) *Journal of Quantitative Spectroscopy and Radiative Transfer* 149, pp. 90-102.
- [15] Aragón, C., Aguilera, J.A., Quantitative analysis by laser-induced breakdown spectroscopy based on generalized curves of growth (2015) *Spectrochim. Acta B* 110, pp. 124-133.
- [16] Grifoni, E., Legnaioli, S., Lorenzetti, G., Pagnotta, S., Poggialini, F., Palleschi, V., From Calibration-Free to Fundamental Parameters Analysis: a comparison of three recently proposed approaches (2016) *Spectrochim. Acta B* 124, pp. 40-46
- [17] Winefordner, J. D., Long, G. L., Limit of Detection, a closer look at the IUPAC definition (1983) *Analytical Chemistry* 55 (7), pp. 712A-724A.
- [18] Andrade, J.M., Cristoforetti, G., Legnaioli, S., Lorenzetti, G., Palleschi, V., Shaltout, A.A., Classical univariate calibration and partial least squares for quantitative analysis of brass samples by laser-induced breakdown spectroscopy (2010) *Spectrochim. Acta B* 65 (8), pp. 658-663.
- [19] Boueri, M., Motto-Ros, V., Lei, W.-Q., Ma, Q.-L., Zheng, L.-J., Zeng, H.-P., Yu, J., Identification of polymer materials using laser-induced breakdown spectroscopy combined with artificial neural networks (2011) *Appl. Spectrosc.* 65(3), pp. 307-314.
- [20] Caceres, J.O., Moncayo, S., Rosales, J.D., De Villena, F.J.M., Alvira, F.C., Bilmes, G.M., Application of laser-induced breakdown spectroscopy (LIBS) and neural networks to olive oils analysis (2013) *Appl. Spectrosc.* 67(9), pp. 1064-1072.
- [21] El Haddad, J., Bruyère, D., Ismaël, A., Gallou, G., Laperche, V., Michel, K., Canioni, L., Bousquet, B., Application of a series of artificial neural networks to on-site quantitative analysis of lead into real soil samples by laser induced breakdown spectroscopy (2014) *Spectrochim. Acta B* 97(1), pp. 57-64.
- [22] Vítková, G., Novotný, K., Prokeš, L., Hrdlička, A., Kaiser, J., Novotný, J., Malina, R., Prochazka, D., Fast identification of biominerals by means of stand-off laser induced breakdown spectroscopy using linear discriminant analysis and artificial neural networks (2012) *Spectrochim. Acta B* 73, pp. 1-6.
- [23] D'Andrea, E., Pagnotta, S., Grifoni, E., Legnaioli, S., Lorenzetti, G., Palleschi, V., Lazzerini, B., A hybrid calibration-free/artificial neural networks approach to the quantitative analysis of LIBS spectra (2015) *Appl.Phys. B* 118 (3), pp. 353-360.
- [24] Haykin, S.S., *Neural Networks. A Comprehensive Foundation* (2nd ed.) Prentice-Hall, New Jersey, U.S.A. (1999).
- [25] D'Andrea, E., Pagnotta, S., Grifoni, E., Lorenzetti, G., Legnaioli, S., Palleschi, V., Lazzerini, B., An artificial neural network approach to laser-induced breakdown spectroscopy quantitative analysis (2014) *Spectrochim. Acta B* 99, pp. 52-58.
- [26] "Selection, Use and Interpretation of Proficiency Testing (PT) Schemes by Laboratories", Eurachem webpage: <https://eurachem.org/index.php/publications/guides/usingpt>



Graphical abstract

ACCEPTED MANU

Highlights

- The build and optimization of calibration curves for minor elements in cast iron is discussed
- The example of the analysis of copper traces in cast iron is presented
- Univariate and Multivariate methods are used
- The dangers of using multivariate methods as 'black boxes' are discussed
- Methods are described for assessing the reliability of multivariate approaches
- A simple univariate approach may give results equivalent to complex multivariate approaches

ACCEPTED MANUSCRIPT