



## Good practices in LIBS analysis: Review and advices

Josette El Haddad, Lionel Canioni, Bruno Bousquet

### ► To cite this version:

Josette El Haddad, Lionel Canioni, Bruno Bousquet. Good practices in LIBS analysis: Review and advices. Spectrochimica Acta Part B: Atomic Spectroscopy, Elsevier, 2014, 101, pp.171-182. 10.1016/j.sab.2014.08.039 . hal-01094344

**HAL Id: hal-01094344**

**<https://hal.archives-ouvertes.fr/hal-01094344>**

Submitted on 12 Dec 2014

**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.



Distributed under a Creative Commons Attribution - NonCommercial| 4.0 International License

# Good practices in LIBS analysis: Review and advices

J. El Haddad, L. Canioni, B. Bousquet\*

Univ. Bordeaux, LOMA, CNRS UMR 5798, F-33400 Talence, France

## ABSTRACT

### Keywords:

Laser-induced breakdown spectroscopy (LIBS)  
Classification  
Concentration measurement  
Assessment  
Good practices

This paper presents a review on the analytical results obtained by laser-induced breakdown spectroscopy (LIBS). In the first part, results on identification and classification of samples are presented including the risk of misclassification, and in the second part, results on concentration measurement based on calibration are accompanied with significant figures of merit including the concept of accuracy. Both univariate and multivariate approaches are discussed with special emphasis on the methodology, the way of presenting the results and the assessment of the methods. Finally, good practices are proposed for both classification and concentration measurement.

## Contents

1.	Introduction . . . . .	171
2.	Identification and classification . . . . .	172
3.	Concentration measurement . . . . .	174
3.1.	Univariate analysis . . . . .	174
3.2.	Multivariate analysis . . . . .	177
3.2.1.	Subsets of data and validation . . . . .	177
3.2.2.	Figures of merit . . . . .	178
3.3.	Evaluation and limits . . . . .	178
3.3.1.	Y-randomization . . . . .	178
3.3.2.	Applicability domain . . . . .	178
3.3.3.	Influence of the predictors . . . . .	179
4.	Conclusion . . . . .	179
	References . . . . .	180

## 1. Introduction

Laser-induced breakdown spectroscopy (LIBS) has been investigated since the invention of the laser in 1960 and during the first years there was an impressive activity to render it mature and reliable [1]. LIBS has been applied to many fields with specific adaptations and improvements concerning the related instrumentation. Thus, in addition to standard laboratory conditions, LIBS analyses have been carried out at long distance in order to analyze potentially hazardous materials [2] and this technique has even been employed also to analyze the surface

of Mars [3]. In an opposite approach, LIBS has been exploited at the micro scale in order to provide chemical mapping of a given sample [4,5] and for high spatial resolution [6]. Moreover, many portable or transportable LIBS instruments have been developed in order to open the way to on-site measurements [7–9]. LIBS has thus been applied to e.g. medical science [10], geomaterials [11], explosives [12], recycling [13], forensics [14] and agriculture [15]. It is difficult to quote an exhaustive list of LIBS applications, however the most significant ones can be found in the specialized literature [16,17] and recent review papers [18,19]. It should be noticed that LIBS allows analyzing solids, liquids, aerosols and gases in various experimental conditions. Consequently, LIBS is potentially one of the most promising methods for elemental analysis.

\* Corresponding author. Tel.: +33 540002870; fax: +33 540006970.  
E-mail address: [bruno.bousquet@u-bordeaux.fr](mailto:bruno.bousquet@u-bordeaux.fr) (B. Bousquet).

Nowadays, LIBS is considered to be a well-established analytical technique when considering instrumentation and experimental conditions, despite of the large variety of experimental setups and conditions. Moreover, several handheld LIBS systems are already commercialized for on-site measurements. However, the treatment of the LIBS spectra is still often subject to endless discussions and finally the actual ability of the technique is not sufficiently assessed. Based on this observation, we propose in this review paper to give a general outline about LIBS data processing including the methods commonly applied, the way the results are usually presented and also how the reliability of the LIBS analysis is assessed. The first part of this review is dedicated to qualitative LIBS, including identification of species (ions, atoms and molecules) as well as classification and sorting. We present how the results have been obtained and presented, and how the reliability has been determined. Finally, the risk of misclassification is also discussed. The second part concerns concentration measurement by LIBS based on calibration [20]. Indeed, self-calibrated or calibration-free methods [21] have been discussed elsewhere since they are based on physical models involving specific physical parameters playing a role in the accuracy. Thus, the predictive ability of LIBS analysis is emphasized through the selection of significant figures of merit. And finally, a selection of good practices is proposed for both classification and concentration measurement by LIBS.

## 2. Identification and classification

In the frame of decision-making, one may want to classify a series of samples against the presence/absence of a given element or against the concentration of this element relative to a threshold value. In principle, such classification could be achieved from a single variable or predictor  $x$  selected as the most relevant for the decision. In LIBS, the predictor is usually the intensity or the peak area of the relevant atomic or ionic line [22], or even the one of the relevant molecular lines [23]. Moreover, in order to overcome some experimental effects, it is common to prefer the ratio of two spectral lines [24] as a kind of normalization.

In LIBS, the first question concerns the identification of each spectral line. Actually, it may be not necessary to identify each single line separately since elements such as iron or titanium for example display hundreds of lines. Thus, when iron is present, it is easy to verify its presence but it is not helpful to strictly identify each individual line. The real question to be discussed concerns the ability to detect by LIBS any chemical element of the periodic table. The most popular atomic database is the one of the NIST [25] but it should be noticed that some complementary data could be found in the Kurucz database [26]. In addition, some research groups and companies have built their own LIBS database from LIBS experiments. As an example, the company Applied Photonics [27] edited a LIBS database in the form of the periodic table based on selected LIBS papers in which each element has been actually detected. Whatever the database, element identification requires a method to assign the relevant spectral lines to the chemical elements. Ideally, this step should be automated. In this context, Amato et al. [28] presented an algorithm inspired by text retrieval for unassisted element identification from LIBS spectra. In addition, molecular bands have also been observed in LIBS spectra. The most commonly reported ones are the  $C_2$  and CN bands [29–31].

When no line is detected in the LIBS spectrum for a given element, it doesn't necessarily mean that the element is absent from the sample's composition. It might be present at a concentration lower than the limit of detection of the instrument. This limit is very dependent on the experimental parameters and consequently one should never conclude that an element is absent, but rather that its concentration is below the limit of detection. In addition, spectral interferences sometimes observed in the LIBS spectra could drive to wrong conclusions [32]. To overcome this problem, it is highly recommended to take advantage of the redundancy by checking the coexistence of a series of different lines related to the same element. Finally, the spectral

range available for analysis is determined by the spectrometer. As a consequence, interesting spectral lines might fall outside of the available spectral range and the related element not detected. Thus, the spectral range is also an important parameter for successful element identification.

It has been reported that classification of unknown samples in two classes regarding their LIBS spectra was possible through the identification of a significant element. For example, in-situ LIBS analysis of white pigments used in painting [33] has been achieved via the detection of the spectral lines of lead (Pb I) at 357.27, 363.96, 368.35, 373.99, and 405.78 nm and titanium (Ti I) at 429.87–430.59, 445.74, 453.32–453.60, and 498.17–503.99 nm. The presence of lead allowed classifying original paintings while titanium was related to the retouched parts. It should be noted that in this work, the authors didn't provide any figures of merit. The detection of different lines of lead (Pb I) was simply correlated to the first type of white pigment based on lead carbonate  $2PbCO_3 \cdot Pb(OH)_2$ . Similarly, the detection of different lines of titanium (Ti I) was simply correlated to the second type of white pigment based on titanium dioxide  $TiO_2$ .

Another example concerns the classification of treated and untreated wood samples by LIBS [34]. Thanks to the analysis of the emission lines of As I at 228.8 nm and Cr II at 267.7 nm, it was possible to proof the presence of chromated copper arsenate (CCA), which is a characteristic of treated wood. Indeed, for untreated wood, there was no peak at these two wavelengths. Thus, it was straightforward to separate the samples into two classes. In addition, the authors analyzed average values of intensity for the two spectral lines of interest and presented error bars corresponding to 3 times the standard deviation over 600 measurements, i.e. the 600 points of each single sample image, in order to overcome possible issues with heterogeneity of the samples. However, it should be noticed that no figures of merit were given to assess the real ability of classification of the method.

Moreover, classification of samples by LIBS has been performed by the use of multiple LIBS data simultaneously. For example, the ratios Ca (396.847 nm)/K (766.491 nm) and Cu (327.39 nm)/K (766.491 nm) were compared for both normal and malignant tissues [35]. Malignant tissues were identified since the concentration of potassium was kept unchanged while the concentration of calcium was higher and that of copper was lower in malignant tissues. But in this study, again, no figures of merit were given and this paper was mainly presented as a demonstration. Moreover, in the context of biomedical applications, the ratios H (656.7 nm)/C (247.9 nm) and Ca (310.9 nm)/C (247.9 nm) have been compared in order to identify organic and inorganic compounds in kidney stones [36]. This approach allowed discriminating hydroxyapatite from other stones. By extension, other ratios were calculated in order to identify different types of stones. This work based on the calculation of selected ratios was also presented as a demonstration and no figures of merit were given. At the opposite, in the field of forensics, it is common to provide statistical parameters such as the type I and type II errors, which are related to false positive and false negative, respectively. For example Naes et al. [37] used these two types of errors to discriminate glass fragments.

However, the authors decided to implement methods of chemometrics such as principal component analysis (PCA) and partial least square-discriminant analysis (PLS-DA) in order to enhance the confidence of classification. They presented their results of classification by a confusion matrix (see Table 1). Unfortunately, they didn't compare the ability of multivariate analysis to the one of the univariate approach. In

**Table 1**  
Confusion matrix in the simplest case of classification into two classes.

		Reference results	
		Positive	Negative
LIBS results	Positive	TP	FP
	Negative	FN	TN

addition, Kongbonga et al. [38] demonstrated that it was possible to discriminate the palm oil from other oils through the value of the ratio  $C_2$  (516.6 nm)/ $C_1$  (247.8 nm). The results of this study were interpreted by the analysis of variance (ANOVA).

Multivariate analyses have been already widely applied to LIBS data for classification purposes [39,40]. In this case, one should consider a matrix  $X$  of predictors instead of a single value  $x$ . First of all, the  $X$ -matrix is usually described through PCA calculation [41]. Once the principal components are calculated, a second step consisting in measuring the distances between the points is necessary to conclude about the presence of classes. This approach represents with PLS-DA [42, 43] the most popular multivariate method for classification. PCA is not supervised and provides results about similarities and differences between a series of samples without knowing anything about them. This method based on the calculation of the covariance of the  $X$ -matrix of predictors allows for projecting most of the information from the original dataset into a compressed subspace of a few principal components. In LIBS, PCA has been demonstrated to be very efficient to detect outliers, i.e. samples which display spectra that are very different from all the others and have been successfully applied to the classification of LIBS data [44]. PLS-DA is similar to PCA but is supervised. It relates the predictors to numbers arbitrarily set for classification, each number being associated with a class. This technique being supervised, it needs a learning step prior to its application to unknown samples. PLS-DA has also been applied to the classification of LIBS data [45,46]. In addition, one should also notice other multivariate methods of classification such as independent component analysis (ICA) [47], support vector machine (SVM) [48,49], artificial neural networks (ANN) [50], hierarchical cluster analysis (HCA) [50], soft independent modeling of class analogy (SIMCA) [51], and the method of the  $k$ -nearest neighbors (KNN) [52].

The common advantage of all the chemometric techniques is to simultaneously take advantage of different features within the LIBS spectra. As an example, in the framework of forensics, different biological tissues such as chicken brain, liver, kidney, spleen, lung and skeletal muscle were classified after LIBS analysis [50]. The authors exploited three chemometric methods, namely HCA, PLS-DA and ANN. Then, they calculated both the rates of correct and incorrect classification. This evaluation was done via samples analyzed by LIBS during an experimental run different from the calibration. The conclusion was not easy because each method presented specific advantages: ANN demonstrated a lower rate of incorrect classification for muscle and spleen since PLS demonstrated best results for the brain, liver and kidney.

More generally, the methodology implemented for classification purpose should be detailed. Indeed, a learning step is always required first and then the classification model can be evaluated during a second step. These two steps can be run from a single dataset by implementing the methods of leave-one-out (LOO) or leave-many-out (LMO) [49,54, 55]. These methods consist in excluding one sample or more from the dataset, then building the model, and finally checking the ability of the model to correctly classify the rejected sample(s). Finally, each sample is iteratively excluded from the calibration set and exploited at least once for validation. These methods are known as cross-validation methods [54]. They are implemented by default in many algorithms of chemometric methods. However, a more robust approach than cross-validation consists in preparing two separate subsets from the original dataset: the calibration set dedicated to the construction of the model and the validation set dedicated to its evaluation. This approach is known as external validation. It has been exploited in forensics [50] and revealed some advantages to avoid any risk of over-fitting and allows for checking if the calibration set or training set was correctly chosen to describe all the variability of the samples.

However, in order to implement either internal or external validation, one needs to calculate some figures of merit. For classification purposes, one may want to know if the samples are correctly classified or not. This assessment can only be done by comparing the results of the LIBS data processing and the ones provided by another method,

which is considered as the reference method. Consequently, the latter results are considered as “true” values. Finally, any test of classification should be assessed via an appropriate question, e.g. of the type: is this patient healthy? Is this sample polluted by lead? Is this sample made of recycled material? When the answer to the question is yes, as expected, the result is categorized as true positive (TP). Symmetrically, when the answer is no, as expected, the result is categorized as true negative (TN). At the opposite, when the answer is different from the expected one, one should consider false positive (FP) as well as false negative (FN) results. Table 1 displays the so-called confusion matrix [56] including the four figures of merit given before.

By extension, the confusion matrix can be implemented in order to assess the classification when more than two classes are present. In this case, the diagonal of the confusion matrix displays the results of correct classification, i.e. the values of TP related to each class and often normalized in order to present percentage values while the out-of-diagonal values indicate the cases of incorrect classification. Thus, in the case of 5 classes for example, the confusion matrix becomes a  $5 \times 5$  matrix [36]. Some examples of confusion matrices are given in Refs. [36,55,57]. It should be noticed that a confusion matrix designed for more than two classes doesn't display the values of TN, FP and FN as described in Table 1 but necessarily contains more detailed information about the distribution of the misclassified samples into the other classes.

In order to reduce the risk of misclassification, some authors introduced an additional threshold value. Thus, once the predicted value is found to be close to the target value corresponding to a given class according the tolerance given by the threshold value, the sample under study is attributed to this class. In addition, it is sometimes preferred to obtain unclassified instead of misclassified samples. This can be done by simply adding a class meaning “all that is not belonging to any of the classes explicitly defined” [44,49]. More generally, while hard classification, which consists in classifying each sample in a single class, is the most commonly applied, soft classification, which consists in calculating the probability of membership of each sample to each class, also provides very interesting results in LIBS. As an example, soft classification has been successfully applied to the LIBS analysis of soil samples [53]. Thus, the soil samples were located inside a ternary diagram after the calculation of factors related to the probability of belonging to each of the three major types of soils, namely silicates, calcareous and ores. This approach of soft classification was very efficient prior to concentration measurements.

In Table 2, we present a compilation of the most common figures of merit that have been exploited for classification purposes based on LIBS data processing.

**Table 2**  
Figures of merit exploited for classification by LIBS.

Figure of merit	Calculation	Reference
Rate of correct classification	Percentage of samples correctly classified for a given class	[46,54]
Rate of wrong classification	Percentage of samples incorrectly classified for a given class	[46,54]
Rate of no classification	Percentage of samples not classified in any class	[46,54]
Overall accuracy (%)	Number of samples correctly classified (all classes)/total number of samples	[55,56]
Sensitivity	$TP/(TP + FN)$ for a given class	[46,56]
Specificity	$TN/(TN + FP)$ for a given class	[56]
Negative predictive value	$TN/(TN + FN)$ for a given class	[56]
Positive predictive value	$TP/(TP + FP)$ for a given class	[56]
Robustness	After suppression of a given class, the rate of wrong classification of the samples in the other classes	[41,46]

It should be noticed that the rate of no classification is given by the samples for which the result of the LIBS data processing is out of the intervals delimited by the threshold values for all the classes under study. The overall accuracy is by definition less descriptive than the complete confusion matrix but it allows partially assessing the methods of classification. Sensitivity is given by the rate of correctly classified samples for each class, while specificity determines how the model is able to predict that a sample does not belong to a specific class. Robustness can be calculated by: i) eliminating all the samples of a given class and ii) testing afterward if the model of classification is really able to conclude that all these leftover samples are not incorrectly affected to any class. Of course, a model is considered as robust when the rate of misclassification is low.

In the field of plastic recycling [58,59], the results of classification of polymer samples after LIBS analysis have been presented through several figures of merit given in Table 2. The authors presented two relevant statistical parameters, namely the Mahalanobis distance (or M-distance) and the spectral residual, i.e. the square of the difference between the original and the simulated spectra. For a series of samples belonging to the same class, the M-distance is expected to be small as well as the spectral residual values and the results presented in [58,59] were in good agreement with this statement. The authors defined the best threshold for the M-distance and the result was evaluated through the so-called receiver-operating characteristic often called the ROC curve, and displaying the sensitivity versus the quantity  $(1 - \text{specificity})$ . They obtained a sensitivity of 90% and specificity above 76% for the four classes.

Classification after LIBS analysis has also been successfully demonstrated in security applications. Different explosive residues (RDX, TNT, etc.) have been classified by PLS-DA [60], and the results were presented through the sensitivity that the authors defined as the true positive rate TPR and the false positive rate which is given by  $\text{FPR} = (1 - \text{specificity})$ . The best model obviously corresponds to  $\text{TPR} = 100\%$  and  $\text{FPR} = 0\%$ . After optimal selection of the input data for PLS-DA, the authors obtained  $\text{TPR} = 99.5\%$  and  $\text{FPR} = 0.16\%$  for different models.

The evaluation of any model of classification should be carried out by analyzing unknown samples. Different methods of validation have been implemented, namely cross-validation methods such as i) the leave-one-out procedure (LOO) [41,49], ii) the k-fold method (or LMO) [40], and iii) the leave-one-subject-out procedure (LOSO) [61] and also external validation based on the use of an independent dataset for testing the model [50,60]. It should be emphasized that in the case of LOSO, no replicate of any given sample could be found in both the learning and the validation sets. As a consequence, LOSO is expected to be more robust by construction than LOO and LMO. Actually, Remus et al. [61] demonstrated that LOSO was better than the k-fold method for generalizing the classification to unknown samples. As a conclusion, external validation should be systematically preferred in order to obtain robust models for classification.

Finally, the choice of the predictors is of major importance. Cisewski et al. [48] classified spore samples and highlighted the importance of removing outliers prior to any calculation. This recommendation had also been stated by other studies [50]. Nowadays, the coupling between different multivariate approaches has become quite common in order to build robust models. Thus, PCA could be used to compress the original dataset into a reduced number of factors before introducing them into the model of classification. As an example, PCA was used prior to HCA [50] and PCA was also applied to compress the LIBS data related to different classes of proteins, prior to their treatment by SVM [62]. One can conclude that the results of classification strongly depend on the choice of the predictors. And in this context, Corsi et al. [63] proposed a very original approach for analyzing ancient copper artifacts. Indeed, they first calculated the concentrations of several elements (Ag, Pb, As, Fe, Sb) via calibration-free LIBS and exploited these values of

concentrations as input data of a PCA model. The scores plot clearly revealed the separation between two classes of samples.

It should be also pointed out that results of classification are sensitive to data preprocessing. As an example, Sahoo et al. [64] studied the interest of outlier removal before classification applied to LIBS spectra thanks to different approaches, namely PCA, dendograms, nearest neighbors and distance matrix. The average classification accuracy was thus increased for high energy materials (HEM). Dimensionality reduction was also successfully applied as an alternative to outlier removal for both HEM and non-HEM classification.

Concluding on this first section dedicated to the identification and classification by LIBS, it should be noted that the corresponding scientific papers have become more and more descriptive, most of them presenting now significant figures of merit required for objective assessment. Nevertheless, few papers still present classification results in an insufficient way. Indeed, the very important point of validation is often insufficiently described and thus it can be difficult to generalize the reported results. Consequently, we highly recommend for the future work on classification by LIBS to adopt in generally a series of good practices based on: i) clever selection of predictors, ii) separation of the original dataset into independent subsets, iii) calculation of the classification ability of the model by external validation, and iv) presentation of the relevant figures of merit.

### 3. Concentration measurement

In the case of concentration measurement, one should build a regression model establishing the best relationship between the predictors, namely the LIBS data ( $X$ ), and the concentration values ( $Y$ ) of the analyte. The simplest case of regression model is univariate, which means that only one predictor  $x$  per sample is exploited instead of the  $X$ -matrix and only the concentration values of one single analyte are predicted, and referenced as  $y$ . In this case, both  $x$  and  $y$  are vectors of dimension  $N$ ,  $N$  being the number of samples exploited for calibration. The corresponding graphical display is the so-called calibration curve, which consists in plotting the measured signals against the analyte's concentrations. In LIBS, the signal corresponds usually to the intensity or the peak area of the most relevant line, namely exempt of spectral interferences and self-absorption effects and with a good signal-to-noise ratio [65]. To overcome experimental effects, the ratio between two lines is usually preferred. In addition, normalization by an internal standard is widely applied in LIBS in the case of a series of samples characterized by a single matrix [66,67].

#### 3.1. Univariate analysis

In order to get a first understanding of the relationship between the physical parameters, namely the peak intensity or the peak area on one hand, given by  $x$ , and the analyte's concentration on the other hand, given by  $y$ , it is common to calculate the correlation factor also called the Pearson coefficient, assuming a linear relationship between  $x$  and  $y$ , which is given by [68]:

$$R = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (1)$$

where  $x_i$  is the value of the LIBS signal for the sample  $i$ , and  $\bar{x}$  the average value of  $x_i$  over the  $N$  samples. Similarly,  $\bar{y}$  is the reference value of the analyte's concentration for sample  $i$ , and  $\bar{y}$  the average value of  $y_i$  over  $N$  samples. The Pearson coefficient can vary between  $-1$  and  $+1$ . Values close to  $0$  are synonym of poor correlation while values close to  $+1$  reveal strong correlation and values close to  $-1$  reveal strong anti-correlation.



Most of the LIBS papers report  $R^2$ , the value of the square of the Pearson coefficient. Indeed, the  $R^2$  factor provides fast information about the correlation of the data and consequently a fast first knowledge about the prediction ability of the model since poor correlation necessarily implies poor predictive ability. As a first example, in the case of quantitative LIBS analysis of palladium, a very good correlation characterized by  $R^2 = 0.99$  has been reported [69]. In addition, the values of concentrations were equally distributed along the whole calibration range, indicating that the value of  $R^2$  was really significant. High values of  $R^2$  have also been reported by other authors [20,70]. At the opposite, very poor correlation characterized by  $R^2 = 0.01$  has been reported for aluminum in the frame of LIBS analysis of soil samples [71].

However, A. Golbraikh et al. [72] demonstrated that a model with a value of  $R^2$  close to 1 may indeed have a poor accuracy for prediction. Therefore, the authors introduced a new factor, namely  $R^2_{reg}$ , in order to calculate the correlation between the values deduced from the regression (on the fitting curve) and from the reference values, and they recommend having simultaneously  $R^2$  and  $R^2_{reg}$  close to 1. In fact,  $R^2$  could be close to 1 in the case of low degrees of freedom, or variable multi-collinearity [54].

Another figure of merit,  $Q^2$ , allows evaluating the ability of a model to predict values of concentrations close to the reference value [54].  $Q^2$  is given by:

$$Q^2 = 1 - \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2} \quad (2)$$

where  $\hat{y}_i$  are the reference values of concentration,  $y_i$  the predicted ones, and  $\bar{y}$  the average value of  $\hat{y}_i$  over  $N$  samples in the dataset. The values of  $R^2$  and  $Q^2$  can be separately calculated for the calibration, the validation and the test datasets. Ideally,  $R^2$  and  $Q^2$  values should be equal to 1, however some authors considered that a model was acceptable once  $Q^2 > 0.5$  and  $R^2 > 0.6$  [54,72,73]. Moreover, the  $Q^2$  value could be high for the calibration dataset and for cross-validation, but the related model could have poor predictive ability in the case of external validation, especially in case of non-linear behavior [72]. Consequently  $R^2$  and  $Q^2$  are definitely not sufficient to assess the predictive ability of a model. It should be noticed that the  $Q^2$  factor has been rarely applied in the case of univariate calibration.

Finally, it is mandatory to estimate the accuracy of a quantitative method [74]. Accuracy simultaneously contains trueness, which is related to systematic error and measured by the value of bias, and precision, which is related to random error and measured by the value of standard deviation. In most of the LIBS analyses, the accepted “true” values are given by ICP-AES or ICP-MS analyses. While trueness is simply given by the difference between the value of concentration retrieved after LIBS analysis and the “true” value, precision [75] includes instrument and independent repeatability as well as intermediate precision and reproducibility [74]. Precision is described by the standard deviation (SD) or by the relative standard deviation (RSD in %), which is a very common figure of merit in the framework of LIBS analysis [20,39]. These two factors are given by:

$$SD = \sqrt{\frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N}} \quad (3)$$

and:

$$RSD(\%) = \frac{SD}{\bar{y}} \times 100 \quad (4)$$

where  $\bar{y}$  and SD represent the mean value and the standard deviation of the  $y_i$  predicted concentrations of the  $N$  samples.

Furthermore, once the quantitative model is built, it provides predicted values of concentrations. Thus, it becomes possible to

compare the predicted values with the “true” ones provided by the reference method. Consequently, it is possible to define new figures of merit based on the error of prediction [20]. The root mean square error (RMSE) is thus given by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}} \quad (5)$$

where  $\hat{y}_i$  are the reference values of concentration,  $y_i$  the predicted ones, and  $N$  the number of samples in the dataset. It should be noticed that the unit of RMSE is identical to the unit of the concentrations  $y_i$ . Thus, if the concentrations  $y_i$  are given in part per million (ppm), RMSE will be also obtained in ppm. In addition, the quantity  $(\hat{y}_i - y_i)^2$  is expected to be higher for the highest concentrations. This means that the high values of concentration have a strong influence on the RMSE value.

Another way to evaluate the error between the predicted and the actual values of concentration consists in calculating the mean relative error [8,39] defined as:

$$RE(\%) = 100 \frac{1}{N} \sum_{i=1}^N \frac{|\hat{y}_i - y_i|}{\hat{y}_i} \quad (6)$$

In this case, the percentage error is calculated for each value of concentration. So both low concentrations and high concentrations have the same influence on the result. Consequently, RE (%) appears to be an adequate factor to estimate the ability of a quantitative model to correctly predict the analyte's concentrations. For real-life applications, this factor can be easily and rapidly calculated. For example, on-site LIBS quantitative analysis of polluted soils has been achieved with  $RE < 20\%$ , which was considered as a satisfying result for on-site analysis [8]. Other authors calculated the relative error for each sample, namely the quantity  $\frac{|\hat{y}_i - y_i|}{\hat{y}_i} \times 100$  and finally identified the maximum of these values over all the dataset (MRE) [76]:

$$MRE(\%) = \text{Max} \left( \frac{|\hat{y}_i - y_i|}{\hat{y}_i} \times 100 \right) \quad (7)$$

Li et al. [76] reported  $MRE = 13.61\%$  with their method of standardization in the case of LIBS analysis of carbon in coal. They concluded that their original method of normalization provided a significant enhancement of the predictive ability. Another way to evaluate the predictive ability of a quantitative model consists in calculating its limits of confidence [77–79]. These limits are detailed in Ref. [68]. Briefly, let us consider a linear regression between the values of concentration  $x$  and the LIBS data  $y$ , described by the equation  $y = bx + a$ . It should be noticed that the notation might be confusing since  $x$  was previously defined as the LIBS signal and  $y$  the value of concentration. At the opposite, in the latter mathematical expression,  $x$  represents the analyte's concentration and  $y$  the LIBS signal. So the reader is invited to pay attention to this alert.

The slope ( $b$ ) of the resulting straight line is given by [68]:

$$b = \frac{\sum_{i=1}^N (x_i - x_m)(y_i - y_m)}{\sum_{i=1}^N (x_i - x_m)^2} \quad (8)$$

And the intercept ( $a$ ) is given by:

$$a = y_m - bx_m \quad (9)$$

where  $y_m$  is the mean value of the LIBS data,  $x_m$  the mean value of the concentrations,  $N$  the number of samples and  $x_i$  the reference value of the concentration of the sample  $i$ . Then, the residual standard deviation

between the measured values of the LIBS signal and the values deduced from regression is given by [68]:

$$S = \sqrt{\frac{\sum_{i=1}^N (y_i - y_c)^2}{N-2}} \quad (10)$$

where  $y_i$  is the mean value of the replicates of the measured LIBS signals recorded for a given value of concentration  $x_i$ ;  $y_c$  is the value deduced from regression at the value of concentration  $x_i$  and considering the degree of freedom being  $N - 2$ . Based on this description, Mermet [68] explained that an uncertainty always exists when determining the slope and the intercept of the regression line. Indeed, for a given value of concentration  $x_o$ , the corresponding mean value of the LIBS signal  $y_o$  can be determined by the regression law within a confidence limit of  $\pm t \cdot s_o$  giving:

$$y_o = a + bx_o \pm t \cdot s_o \quad (11)$$

where  $t$  is the Student's coefficient  $t(1 - \alpha / 2)$  at a given risk  $\alpha$  or at a confidence percentage of  $[100(1 - \alpha)\%]$  and  $N - 2$  the degree of freedom. The value of  $s_o$  is given by:

$$s_o = S \cdot \sqrt{\frac{1}{N} + \frac{(x_o - x_m)^2}{\sum_{i=1}^N (x_i - x_m)^2}} \quad (12)$$

Finally, following the detailed calculation proposed by Mermet [68], and applying relevant approximations, it becomes possible to calculate the uncertainty range of concentration. Thus, for any unknown sample, one may measure in a first time the LIBS signal  $y_u$  and deduce two values of concentration given by:

$$x_1 = x_u - \frac{t \cdot S}{b} \left( \frac{1}{p} + \frac{1}{N} \right)^{1/2} \quad (13)$$

$$x_2 = x_u + \frac{t \cdot S}{b} \left( \frac{1}{p} + \frac{1}{N} \right)^{1/2} \quad (14)$$

where  $p$  represents the number of replicates.

Graphically, these limits are represented by hyperbolas surrounding the regression line [78]. This type of advanced analysis about accuracy allows for understanding that the predictive ability of any quantitative model is always better in the middle of the range of concentrations. At the end, two relevant parameters, namely the limit of detection and the limit of quantification can be calculated in order to describe the lower limits of a quantitative model [74]. In the case of a linear regression described by equation  $y = bx + a$ , where  $y$  describes the values of the LIBS data and  $x$  the values of concentrations, the definition of the limit of detection the most frequently used is [70,74,80]:

$$\text{LOD} = \frac{3\sigma_a}{b} \quad (15)$$

And the limit of quantification is generally defined by [74]:

$$\text{LOQ} = \frac{10\sigma_a}{b} \quad (16)$$

It should be noticed that the LOQ is rarely presented in LIBS. Moreover, the definitions of LOQ and LOD are not unique [74] and the discussion regarding these two factors is not specific to LIBS but concerns all techniques of analytical chemistry. Finally, it should be mentioned that the experimental response, i.e. the LIBS signal versus the concentration of the analyte might be nonlinear, especially for relatively high values of concentrations. As a consequence, the linear regression could be misfit and a quadratic regression might be

**Table 3**

Figures of merit exploited for univariate quantitative LIBS.

Abbreviation	Full description	References
$R^2$	Correlation coefficient	[62,66]
$Q^2$	Prediction ability	[66]
RSD (%)	Relative standard deviation	[19,36]
RMSE	Root mean square error	[19,70]
RE (%)	Relative error	[19,36]
LOD	Limit of detection	[68]
LOQ	Limit of quantification	[68]
	Uncertainty (limits of confidence)	[62]

advantageous. In this latter case, some figures of merit such as  $R^2$ ,  $S$ , and the confident limits have to be revisited [74]. Table 3 gives a compilation of the figures of merit commonly exploited in the case of univariate quantitative LIBS and examples of related papers.

After this general introduction about the relevant figures of merit that should be used for presenting results in the frame of concentration measurement by LIBS, we propose to observe how the results of LIBS experiments have been reported through a selection of recent papers, all based on univariate analysis. It should be mentioned that the only ambition of this short list of articles is to show some examples of practices in data processing and assessment, in order to finally propose few recommendations for better valorization of the results.

In a study dedicated to the analysis of arsenic [80], the experimental parameters were optimized in order to get the best signal-to-noise ratio and the As I line at 228.812 nm was normalized by the Al I line at 235.1256 nm since aluminum was considered as an internal standard. Then, a calibration curve was built but in a reduced range of concentrations (here 0–50 ppm) in order to avoid any risk nonlinear behavior due to possible saturation effects. And finally, figures of merit such as the coefficient  $R^2$ , the intercept and the slope of the calibration curve were calculated. Moreover, the limit of detection was also calculated but it is interesting to notice that the method for calculating it was not presented. We can conclude that the authors correctly describe the origin of the input data and the results extracted from their linear model of calibration. However, the predictive ability and the robustness of the model are insufficiently discussed and the definition of LOD is missing.

In a work focused on fluoride detection in commercial toothpaste [81], a linear calibration curve was built from the fluorine line at 731.102 nm and the authors calculated a limit of detection different from the one given in Table 3, with a factor of 2 instead of 3. This new definition of LOD clearly highlights the necessity for each author to carefully define the statistical indicators they calculate. Indeed, only rigorous description of each figure of merit will allow to successfully run further interlaboratory comparisons. In this article, again, after careful description of the way to get the best input data, the LOD is calculated but the predictive ability and the robustness of the model are not sufficiently discussed.

In another study dedicated to the quantification of palladium in water by LIBS, the peak intensity of the 340.46 nm Pd I line was analyzed [69]. In addition, due to very specific experimental conditions, the hit frequency defined as the ratio of the number of recorded hits to the total number of laser shots, was calculated as a relevant indicator. Finally, the coefficient of  $R^2$  was calculated and a limit of particle-size detection. This paper illustrates that the figures of merit reported in Table 3 may be not sufficient in some specific cases for which the hit probability of the particle-sample is very low.

As a last example, Xiu et al. [70] analyzed different elements contained in thin oil layers covering a pure aluminum substrate. They presented different analytical figures of merit:  $R^2$ , RSD, slope and intercept of the calibration curve, SD, and LOD for 10 elements: Fe, Mg, Sn, Si, Cu, Ag, Ti, Ni, Cr, and Pb. It should be mentioned that in

addition to the usual LOD, this article also provides interesting discussions about precision of the LIBS measurement. Thus, the height of oil on the surface of the aluminum substrate and the lens-to-sample distance were accurately taken into account. In addition, the experimental parameters were optimized in order to get the best signal-to-noise ratio and the spectral lines carefully selected in order to avoid spectral interferences and self-absorption. As preprocessing, continuous background was removed and the areas of the spectral lines were calculated after fitting by Lorentzian profiles. In addition, the experimental fluctuations were evaluated thanks to the analysis of 8 replicates per sample and advantage of normalization by internal standard was also assessed. Based on this very careful methodology, the authors presented a calibration curve based on the average values and error bars corresponding to the RSD values, calculated over the 8 replicates. The concentration range for the calibration model was limited to 20–400 µg/g, in order to eliminate any risk of nonlinear behavior. Then, the linear regression coefficient  $R^2$ , the intercept and the slope of the calibration curve were presented. Finally, this article provides most of the relevant information one could expect in the frame of concentration measurement by LIBS and is a good guidance paper for presenting LIBS results.

As a conclusion, the results presented in the LIBS articles are most of the time partially discussed and often insufficiently assessed. Thus, in order to finally propose good practices, we first present the multivariate approach, which brings nowadays important added-value to the LIBS.

### 3.2. Multivariate analysis

A very important issue in LIBS concerns the well-known matrix effects. Indeed, for real-life analyses, a wide diversity of matrices can be encountered and consequently the results of concentration measurement can be drastically affected. This drawback could make LIBS analyses unacceptable for on-line or on-site measurements. This has been evidenced in the frame of copper smelting industry where Pb, Cu, Al and Ni should be quantified on-line in lead brass alloy samples [82]. Strong matrix effects prevented to apply simple calibration curves and consequently, a multivariate approach was applied.

Generally speaking, the most common chemometric technique applied to concentration measurement by LIBS is partial least square regression (PLS) [83–86]. This method has been implemented either to calculate the concentrations of a single element (PLS-1 algorithm) or to simultaneously calculate the concentrations of more than one element (PLS-2 algorithm). PLS is a linear method since it is based on linear algebra calculations. Among linear methods, LIBS analyses have also been achieved by multi-linear regression (MLR) and principal component regression (PCR) [87]. One should also notice the methods called least-absolute-shrinkage-and-selection-operator (LASSO) [84, 88] and sparse multivariate-regression-with-covariance-estimation (MRCE) [88]. In 2014, a new method has been applied to LIBS data, namely the wavelet transform-hybrid model [89], and was considered as being more efficient than PLS. In addition, few authors also applied nonlinear methods such as artificial neural networks (ANN) to take into account the possible nonlinear relationship between the X and Y values [39,71,90]. For instance, in the case of copper smelting reported above [82], the PLS-2 method based on the simultaneous regression of the four elements was applied. Other authors reported comparisons between the calibration curve on one hand and multivariate analysis on the other hand. Among them, Andrade et al. [85] analyzed brass samples by LIBS and X-ray fluorescence (XRF) and applied both the calibration curves and the PLS regression for each single element of the following list: Zn (250.199 nm), Fe (238.204 nm), Sn (242.948 nm), Pb (405.781 nm) and Cu (330.795 nm). But it should be pointed out that multivariate quantitative analyses present a high risk of overfitting. As a consequence, it is very important to properly establish and then to properly evaluate the quantitative models.

#### 3.2.1. Subsets of data and validation

Any regression model should be evaluated through a relevant validation method as already commented in the case of classification. This can be done by internal validation or cross-validation methods such as LOO (leave-one-out) or LMO (leave-many-out), namely by excluding one or more samples from the calibration set by a series of consecutive permutations in order to use it or them *a posteriori* in order to test the model. The permutations are realized in a way to exclude each sample at least once from the calibration set [54]. But external validation can also be performed by preparing a dataset fully independent from the calibration one and dedicated to the evaluation. Following this strategy, Anderson et al. [91] built two sets of data, a training set and a test set, in order to evaluate the ability of PLS. In the case of ANN and genetic algorithm methods, the use of an external dataset has been highly recommended for testing the models and insuring both fast learning and optimization of the models [92]. In order to fulfill this requirement, the initial dataset should be split into three parts: the calibration set, the validation set and the test set. The calibration set is the dataset that is used to elaborate the model itself, namely to find the best fit between the predictors and the output values. The validation set is used to compare the different models by checking the risk of over fitting. Thanks to the validation dataset, it becomes possible to determine which model – built from the calibration set – is the best. Finally, the test set is required to post-evaluate the model. The samples to be included into each dataset should be considered on a case-to-case basis, i.e. not only their number but also their distribution over the whole range of concentrations. Thus, we should emphasize that each of the three datasets should not be built randomly but rather cleverly selected in order to include a larger range of concentrations of the analyte into each dataset. Finally, the data included in the test set should be totally independent, namely with no replicates of samples exploited in the other datasets. This methodology theoretically provides the best compromise to get both good learning and correct evaluation of the model. It has been already recommended by other authors in the case of quantitative nonlinear models like ANN, pointing out that the only satisfying method to validate an ANN model was the training-validation-testing approach [93]. Indeed, in the case of ANN, one should understand that the calibration set is used to train the ANN models, and the validation set to determine the best ANN model as a function of adjustable parameters, namely the number of layers and nodes, the learning speed and momentum and the number of iterations. Finally the test set is exploited to post-evaluate the best model. More precisely, the prediction of independent datasets is the only way to correctly assess the risk of overfitting during the training process. As an example, in the case of soil analysis, a series of ANN models were built and an independent validation set was used in order to select the best one, which was finally applied to a third independent dataset in order to get the performance of the model [53]. However, it should be mentioned that internal cross-validation is still commonly applied, especially in the case of linear models. It consists in mixing the calibration and validation sets and, in this case, the risk of overfitting can only be assessed afterward through the prediction of concentrations for samples of an independent dataset [94]. This methodology is completely general and depends neither on the nature of data nor on the type of the quantitative model. We suggest that all the future work about quantitative LIBS adopts this approach.

Finally, even when the right methodology is applied, one should pay attention to the number of data contained into each subset before extracting general results. As an example, ANN was reported to present better predictive ability than the univariate regression model [95], but this result was established with only one sample for the external validation. In this case, the number of samples appears to be too low for a complete generalization of the results. When considering in more detail the work already mentioned about copper smelting industry [82], and based on the use of the PLS-2 method, it appears that the



evaluation of the model was based on the use of replicates of samples exploited for the calibration. This is not the best way of evaluating the method and independent samples should be preferred in order to properly evaluate the ability of the model. In the work of Andrade et al. [85] who analyzed brass samples, the PLS method was evaluated through the LOO cross-validation method. For each analyte, they plotted the predicted values of concentration versus the reference ones, and they found a slope of 0.75 instead of 1 in the case of lead. This result was explained by the fact that the values of concentration were close to the lower range limit of concentrations. It should be emphasized that, even if the question of building subsets of data and validating the models is essential in the case of multivariate approaches, this protocol should also be applied to the case of univariate analysis and thus constitutes a common step that should be applied to all the cases.

### 3.2.2. Figures of merit

In the frame of multivariate analysis, the predictors are multiple and consequently it is impossible to calculate the Pearson coefficient of a predictor  $x$  and the concentration values  $y$ . In this case, we should directly exploit the correlation coefficient  $R^2$  relating the predicted and the reference values of concentration as defined earlier. For instance, Martin et al. [96] presented the factor  $R^2$  between the concentration values of total inorganic carbon (TIC) in soils predicted by PLS and the concentration values considered as reference values and calculated as the difference between the total amount of carbon present in untreated soils and the total amount of organic carbon (TOC). Both the values of TOC and total carbon were measured by ICP-MS. They reported  $R^2 = 0.9445$  for the calibration set and  $R^2 = 0.8713$  for the validation set.

The mathematical expressions of the other factors given in Table 3, namely  $Q^2$ , SD, RSD (%), RMSE, and RE (%) are kept unchanged while exploiting a multivariate approach. These figures of merit allowed Doucet et al. [87] comparing MLR, PCR and PLS with special highlight on the preprocessing step. The comparison was performed through the predictive residual error sum of squares (PRESS), the root mean square error of calibration (RMSEC), the root mean square error of prediction (RMSEP) and the square of the Pearson coefficient ( $R^2$ ). In addition, Barbieri Gonzaga et al. [97] compared multivariate and univariate approaches by the evaluation of relative errors (%), and Li et al. [98] calculated  $R^2$ , RSD, RMSEP and the maximum relative error (MRE) in order to evaluate the best preprocessing prior to PLS. Based on the regression law  $\hat{y} = by + a$ , the mathematical expression of the limits of detection (LOD) and quantification (LOQ) are also kept unchanged. However, they don't correspond to the same physical meaning that in the case of univariate analysis [39,99–101]. Consequently, it is absolutely necessary to build all the regression models on the same basis, namely giving the predicted vs. reference values of concentration prior to any comparison.

### 3.3. Evaluation and limits

#### 3.3.1. Y-randomization

One should take into account that in some cases, quantitative models could fit the data reasonably well but only by chance and without the existence of true correlation. To avoid this risk, the so-called Y-randomization method has been applied [102] since it is considered to be the most powerful validation procedure [103]. Indeed, the cross-validation and external validation methods only assess the predictive ability of a model but not the statistical significance of the estimated results, which can only be given by the Y-randomization procedure. For LIBS analysis, the Y vector is composed of an ordered list of concentration values of the analyte. If this list is randomly reorganized, there is no reason for the model to be able to predict the new list of concentrations with comparable success. This approach is actually an elegant way to evaluate the quantitative model and to demonstrate the existence of a true correlation between the predictors  $X$  and the output values  $Y$ .

Consequently, Y-randomization should be systematically applied to quantitative LIBS analysis. In practice, for a given dataset ( $X$ ) introduced into a quantitative model, the output values ( $Y$ ) are randomly permuted and the figures of merit  $R^2$ ,  $Q^2$  and RMSE are calculated after each permutation. The permutation procedure should be repeated many times in order to get a statistical meaning. Rucker et al. recommended to repeat the permutation procedure at least 25 times [102]. Obviously, the values of  $R^2$  and  $Q^2$  should be higher in the case of the best-optimized model than the ones calculated after Y-randomization, while the values of RMSE should be lower. If this is the case, one can conclude that the best model is statistically significant. To guarantee the result of the Y-randomization method, some authors introduced quantitative criteria, namely  $R^2 < 0.3$  and  $Q^2 < 0.05$  [73]. The Y-randomization method allows for concluding that the quantitative model converges on the basis of real physical parameters, and not by chance. These requirements should allow avoiding not only over-fitting but also any chance of physically meaningless good agreement between the data to finally obtain a robust and efficient model. The Y-randomization procedure has been successfully applied in the frame of LIBS analysis of soil samples based on ANN models [94]. In this study, the authors obtained, after random permutation of the concentration values during the training step, very low values of  $R^2$  et  $Q^2$  and very high values of errors of prediction.

#### 3.3.2. Applicability domain

Another important point to consider in the frame of quantitative analysis is the applicability domain of the model. Indeed, the upper and lower limits of applicability of the model should be determined [54]. Considering that any quantitative model based on calibration exploits a series of known samples, one may want to conclude that both the upper and the lower limits of applicability should be deduced from the extreme values of the calibration set. This is actually the case for the upper limit since the model could deviate from its initial trend and consequently, it would become impossible to predict any value of concentration higher than the maximum value of concentration of the calibration set. However, for the lower limit, it is common to extrapolate the model to values lower than the minimum value of the calibration set. This extrapolation is possible because the model is always considered as linear for the lowest values of concentration. This extrapolation allowed defining the limit of detection (LOD) and the limit of quantification (LOQ), according to the definitions given above. Finally, the applicability domain should also be considered regarding the matrix of the sample. This means that the matrix of any unknown sample should be similar to the matrices considered during the calibration step of the model otherwise the quantitative model might be not adapted for the quantification of the unknown sample.

More generally, it should be mentioned that the applicability domain is a question of major interest in the frame of the quantitative models, whatever the application fields [104]. Technically, the applicability domain can be determined by the calculation of threshold values well established in a guidance document [105]. As an example of direct application of this guide, in the frame of gas chromatography applied to the analysis of essential oils, Qin et al. [94] defined the applicability domain by leverage and thus guaranteed that their study met the OECD principle #3 of the guidance document. Basically, a sample related to a value above a given threshold was considered as outlier and consequently out of the applicability domain. Moreover, in another study dedicated to the exploitation of NIR and fluorescence spectra by quantitative models based on genetic algorithm coupled to PLS, the authors discussed the question of applicability domain through the calculation of average prediction error [106]. Finally, it should be noticed that applicability domain refers not only to the limits of the concentration range but also to the nature of unknown samples compared to the samples belonging to the calibration set, namely to the matrix effect.

### 3.3.3. Influence of the predictors

Similarly to the classification purpose, concentration measurement should be based on the optimal choice of the predictors and on the optimal preparation of the raw data. As a consequence one should select one or few spectral lines in the LIBS spectra as the future predictors, namely either raw or preprocessed data. Moreover, one should split the original dataset into three different subsets: the calibration set, the validation set and the test set as already mentioned. Both the choice of predictors and data subsets can have huge consequence on the predicting ability of the quantitative model. The general rule is that one should select from the original dataset accurate, precise, and consistent experimental data [54,72,73]. As a consequence, one may want to optimize the signal-to-noise ratio or signal-to-background ratio of the LIBS experiment and may prefer selecting specific atomic or ionic (or even molecular) lines with “good” properties, i.e. with a minimum risk of spectral interference and self-absorption [67]. Moreover, the samples should ideally be homogeneous. This is typically the case for metal alloy or liquid samples but not for soil samples clearly considered as heterogeneous at the scale of the laser spot size. In this latter case, one should introduce into the model a mean value resulting from a series of repetitions of LIBS experiments at different locations on the sample surface. In addition, the sample should be ideally plane and with similar physical properties from one point to the other. Finally, statistics over the values of the predictors for the series of samples or for a series of repeated measurements on the same sample may reveal outliers, namely samples related to LIBS data statistically very different from the others. Rejecting outliers at the early stage of the processing allows minimizing the influence of abnormal data and consequently optimizing the results of the model [107], since outliers have a negative effect on the quantitative ability of the models [74]. As an example, Cisewski et al. [48], who classified samples as spores or not spores, removed outliers during the first step of preprocessing. This procedure of removing outliers was also applied by Yueh et al. [50] in the frame of tissue classification. In addition, R.C. Wiens et al. [86] also exploited some preprocessing prior to PLS analysis of LIBS spectra provided by the ChemCam instrument involved in the analysis of Mars geology. The authors thus removed outliers by using independent component analysis (ICA) and calculated after each outlier removal the new error value (RMSE) in order to find an optimum. Indeed, since a large number of data is required for good calibration and good ability to generalize the predictions, removing input data could potentially decrease the accuracy of the model.

Moreover, for quantitative purpose, all the samples should ideally come from the same matrix since huge matrix effects have been reported in LIBS [108–110] with dramatic consequences on the predictive ability of the quantitative model. Nevertheless, when it is not possible to classify the samples prior to quantitative analysis, one should calibrate the model from data revealing all the diversity of the matrices that could be encountered, otherwise the model could be unable to predict a part of the samples. In addition, as already mentioned, the selection of the calibration samples should not be random since an optimized selection allows taking into account all the diversity of the original dataset while a random selection could be somehow restrictive. It is well established that this approach of optimized instead of random selection of the input data provides better statistical predictive ability [54]. To illustrate this point, it has been demonstrated, in the frame of multivariate analysis of soil samples that introducing in the quantitative model spectral data from the matrix in addition to those related to the analyte was very efficient for rising up the prediction ability of the model [71]. Regarding the classification of salt samples by PCA, the authors demonstrated that selecting reduced spectral windows and excluding the lines related to potassium were advantageous. Moreover, they also found that including the lines related to aluminum allowed recognizing the method of production and that normalizing by a CN band as a preprocessing step was very helpful for classification [57]. Other types of preprocessing have been applied to the LIBS data; they

consist in smoothing the spectra, subtracting the background and fitting the spectral lines with the Lorentzian function [36]. The selection of the most relevant data has also been achieved by the method of wavelet compression [48] prior to their introduction into a SVM model. Moreover, in the frame of soil analysis, each individual LIBS spectrum was first pre-treated by the method known as Standard Normal Variate (SNV). Then the average spectrum for each sample was calculated, the data were compressed by wavelet compression and finally three techniques of variable selection were applied, namely genetic, successive projections, and stepwise algorithms [55]. Furthermore, principal component analysis (PCA) has also been used to compress the initial data before their introduction into a model of cluster analysis [50], and PCA was also used to compress the LIBS data of different types of proteins before their introduction into a SVM model [62].

The selection of the relevant data was also achieved by variable importance in projection (VIP), which estimates the importance of variables in the model. Then, the variables selected by VIP were introduced into PLS-DA and SVM models and VIP was recognized to decrease the time of training of the models [43]. The authors thus classified rocks through the VIP variables, which were the wavelengths of the elements representing the rock matrix, i.e. Al, Si, Ca, Mg, K, and O. The k-fold cross validation method was finally applied in order to evaluate the models. Finally, in the frame of classification of explosive residues, it was also demonstrated that introducing the VIP variables into a PLS-DA model was very advantageous [60]. PCA has also been exploited in order to class samples prior to univariate quantitative analysis [111]. Moreover, Anderson et al. [91] applied different methods of classification before quantitative analysis achieved by PLS.

Finally, the selection of predictors should be clearly documented, especially in the case of complex datasets requiring the use of multivariate analysis. This may allow useful verification of the results and inter laboratory comparisons. As an example, ANN appears to be a very efficient method but very sensitive to the input data. However, this essential information about the input data is often not provided or insufficiently detailed, preventing any complete evaluation. As an example, Mukhono et al. [90] presented results about quantitative analysis by PLS and ANN of trace elements such as As, Cr, Cu, Pb, and Ti present in soils and rocks. This study was conducted in the frame of high background radiation areas (HBRA), namely areas characterized by high background natural radioactivity compared to the recommended dose limit. In this article, despite of good description of the models, the discussion about the selection of input data was missing as well as error values for ANN. Finally, the predicted values of concentration were given for only one soil and one rock, which is clearly not sufficient to fully assess the models.

## 4. Conclusion

We presented a review about classification and concentration measurement after LIBS analysis based on a selection of papers, which was obviously not exhaustive. We concluded that the results were often insufficiently assessed by the relevant figures of merit. In the frame of classification, we recommended to systematically present the confusion matrix in order to properly evaluate the ability of the models. And since preprocessing was found to be highly influent on the results of the classification models, we also recommended that this important step should be much more detailed in the scientific papers. This additional information could allow performing in the future accurate comparison between the different methods. The predictors should also be carefully selected and the outliers should be removed, since the classification ability of any model is potentially strongly influenced by the input data.

Similarly, in the case of quantitative analysis, we recommended to calculate a series of statistical indicators in order to properly evaluate the models. As a general advice, we emphasized that the evaluation of the models should be based on the exploitation of three independent subsets

of data and on external validation. Finally, we pointed out that any quantitative model should be evaluated through the test of Y-randomization in order to verify that the predictive ability of the model is really based on the correlation between the data and not obtained by chance. The applicability domain should also be systematically discussed, considering not only the range of concentrations but also the nature of the samples, namely the matrix effects. We finally pointed out that the selection of input data including preprocessing and outlier removal should be considered in detail since it has a very high impact on the results of LIBS experiments. And obviously, because the input data must have the best quality, it is highly recommended to optimize the experimental parameters and to control them along the LIBS measurements, as a very first but essential step to perform good LIBS analyses.

## References

- [1] M. Baudelet, B.W. Smith, The first years of laser-induced breakdown spectroscopy, *J. Anal. At. Spectrom.* 28 (2013) 624–629.
- [2] S. Palanco, C. López-Moreno, J.J. Laserna, Design, construction and assessment of a field-deployable laser-induced breakdown spectrometer for remote elemental sensing, *Spectrochim. Acta B* 61 (2006) 88–95.
- [3] S. Maurice, R. Wiens, M. Saccoccio, B. Barraclough, O. Gasnault, O. Forni, N. Mangold, D. Baratoux, S. Bender, G. Berger, The ChemCam instrument suite on the Mars Science Laboratory (MSL) rover: science objectives and mast unit description, *Space Sci. Rev.* 170 (2012) 95–166.
- [4] D. Menut, P. Fichet, J.-L. Lacour, A. Rivoallan, P. Mauchien, Micro-laser-induced breakdown spectroscopy technique: a powerful method for performing quantitative surface mapping on conductive and nonconductive samples, *Appl. Opt.* 42 (2003) 6063–6071.
- [5] C. Fabre, B. Lathuilière, Relationships between growth-bands and paleoenvironmental proxies Sr/Ca and Mg/Ca in hypercalcified sponge: a micro-laser induced breakdown spectroscopy approach, *Spectrochim. Acta B* 62 (2007) 1537–1545.
- [6] Y. Lu, V. Zorba, X. Mao, R. Zheng, R.E. Russo, UV fs–ns double-pulse laser induced breakdown spectroscopy for high spatial resolution chemical analysis, *J. Anal. At. Spectrom.* 28 (2013) 743–748.
- [7] S. Guirado, F. Fortes, V. Lazic, J. Laserna, Chemical analysis of archeological materials in submarine environments using laser-induced breakdown spectroscopy. On-site trials in the Mediterranean Sea, *Spectrochim. Acta B* 74 (2012) 137–143.
- [8] J. El Haddad, M. Villot-Kadri, A. Ismaël, G. Gallou, K. Michel, D. Bruyère, V. Laperche, L. Canioni, B. Bousquet, Artificial neural network for on-site quantitative analysis of soils using laser induced breakdown spectroscopy, *Spectrochim. Acta B* 79–80 (2012) 51–57.
- [9] A. Ismaël, B. Bousquet, K. Michel-Le Pierrès, G. Travaillé, L. Canioni, S. Roy, In situ semi-quantitative analysis of polluted soils by laser-induced breakdown spectroscopy (LIBS), *Appl. Spectrosc.* 65 (2011) 467–473.
- [10] S. Rehse, H. Salimnia, A. Miziolek, Laser-induced breakdown spectroscopy (LIBS): an overview of recent progress and future potential for biomedical applications, *J. Med. Eng. Technol.* 36 (2012) 77–89.
- [11] F.C. De Lucia Jr, J.L. Gottfried, Rapid analysis of energetic and geo-materials using LIBS, *Mater. Today* 14 (2011) 274–281.
- [12] J.L. Gottfried, F.C. De Lucia Jr, C.A. Munson, A.W. Miziolek, Laser-induced breakdown spectroscopy for detection of explosives residues: a review of recent advances, challenges, and future prospects, *Anal. Bioanal. Chem.* 395 (2009) 283–300.
- [13] H. Xia, M. Bakker, Reliable classification of moving waste materials with LIBS in concrete recycling, *Talanta* 120 (2014) 239–247.
- [14] E. Rodríguez-Celis, I. Gornushkin, U. Heitmann, J. Almirall, B. Smith, J. Winefordner, N. Omenetto, Laser induced breakdown spectroscopy as a tool for discrimination of glass for forensic applications, *Anal. Bioanal. Chem.* 391 (2008) 1961–1968.
- [15] R.A. Multari, D.A. Cremers, T. Scott, P. Kendrick, Detection of pesticides and dioxins in tissue fats and rendering oils using laser-induced breakdown spectroscopy (LIBS), *J. Agric. Food Chem.* 61 (2013) 2348–2357.
- [16] R. Noll, *Laser-Induced Breakdown Spectroscopy: Fundamentals and Applications*, Springer, 2012.
- [17] D.A. Cremers, L.J. Radziemski, *Handbook of Laser-Induced Breakdown Spectroscopy*, Wiley, 2013.
- [18] D.W. Hahn, N. Omenetto, Laser-induced breakdown spectroscopy (LIBS), part II: review of instrumental and methodological approaches to material analysis and applications to different fields, *Appl. Spectrosc.* 66 (2012) 347–419.
- [19] R. Gaudioso, M. Dell'Aglio, O.D. Pascale, G.S. Senesi, A.D. Giacomo, Laser induced breakdown spectroscopy for elemental analysis in environmental, cultural heritage and space applications: a review of methods and results, *Sensors* 10 (2010) 7434–7468.
- [20] L. Huang, M. Yao, Y. Xu, M. Liu, Determination of Cr in water solution by laser-induced breakdown spectroscopy with different univariate calibration models, *Appl. Phys. B* 111 (2013) 45–51.
- [21] S. Pandhija, A.K. Rai, Calibration free LIBS approach for quantitative measurement of constituents in environmental samples, *Emerg. Trends Laser Spectrosc. Appl.* (2010) 237.
- [22] A.W. Andrzej, V. Palleschi, I. Schechter, *Laser-Induced Breakdown Spectroscopy (LIBS): Fundamentals and Applications*, Cambridge University Press, 2006.
- [23] J.H. Scholten, J.M. Teule, V. Zafirooulos, R.M.A. Heeren, Controlled laser cleaning of painted artworks using accurate beam manipulation and on-line LIBS-detection, *J. Cult. Herit.* 1 (2000) S215–S220.
- [24] A. Stankova, N. Gilon, L. Dutruv, V. Kanicky, A simple LIBS method for fast quantitative analysis of fly ashes, *Fuel* 89 (2010) 3468–3474.
- [25] NIST basic atomic spectroscopic data, National Institute of Standards and Technology, 2014. (<http://physics.nist.gov/PhysRefData/Handbook/periodictable.htm>, Last access date 12/02/2014).
- [26] P.L. Smith, C. Heise, J.R. Esmond, R.L. Kurucz, Atomic spectral line database, Built From Atomic Data Files From R.L. Kurucz' CD-ROM, 23, 2014, (<http://www.pmp.uni-hannover.de/cgi-bin/ssi/test/kurucz/sekur.html>, last access date 24/02/2014).
- [27] Analytical capabilities of LIBS, Applied Photonics, 2014, ([http://www.appliedphotonics.co.uk/Libs/capabilities\\_libs.htm](http://www.appliedphotonics.co.uk/Libs/capabilities_libs.htm), last access date 24/02/2014).
- [28] G. Amato, G. Cristoforetti, S. Legnaioli, G. Lorenzetti, V. Palleschi, F. Sorrentino, E. Tognoni, Progress towards an unassisted element identification from laser induced breakdown spectra with automatic ranking techniques inspired by text retrieval, *Spectrochim. Acta B* 65 (2010) 664–670.
- [29] S. Harilal, C. Bindhu, R.C. Issac, V. Nampoori, C. Vallabhan, Electron density and temperature measurements in a laser produced carbon plasma, *J. Appl. Phys.* 82 (1997) 2140–2146.
- [30] C. Vivien, J. Hermann, C. Boulmer-Leborgne, Plasma study in laser ablation process for deposition, ALT'97 International Conference on Laser Surface Processing, International Society for Optics and Photonics, 1998, pp. 359–364.
- [31] M. Baudelet, L. Guyon, J. Yu, J.-P. Wolf, T. Amodeo, E. Fréjafon, P. Laloi, Spectral signature of native CN bonds for bacterium detection and identification using femtosecond laser-induced breakdown spectroscopy, *Appl. Phys. Lett.* 88 (2006) 063901.
- [32] W. Hübner, G. Ankerhold, Elemental misinterpretation in automated analysis of LIBS spectra, *Anal. Bioanal. Chem.* 400 (2011) 3273–3278.
- [33] D. Anglos, S. Couris, C. Fotakis, Laser diagnostics of painted artworks: laser-induced breakdown spectroscopy in pigment identification, *Appl. Spectrosc.* 51 (1997) 1025–1030.
- [34] Y. Aono, K. Ando, N. Hattori, Rapid identification of CCA-treated wood using laser-induced breakdown spectroscopy, *J. Wood Sci.* 58 (2012) 363–368.
- [35] A. Kumar, F.-Y. Yueh, J.P. Singh, S. Burgess, Characterization of malignant tissue cells by laser-induced breakdown spectroscopy, *Appl. Opt.* 43 (2004) 5399–5403.
- [36] B.G. Oztoprak, J. Gonzalez, J. Yoo, T. Gulecen, N. Mutlu, R.E. Russo, O. Gundogdu, A. Demir, Analysis and classification of heterogeneous kidney stones using laser-induced breakdown spectroscopy (LIBS), *Appl. Spectrosc.* 66 (2012) 1353–1361.
- [37] B.E. Naes, S. Umpierrez, S. Ryland, C. Barnett, J.R. Almirall, A comparison of laser ablation inductively coupled plasma mass spectrometry, micro X-ray fluorescence spectroscopy, and laser induced breakdown spectroscopy for the discrimination of automotive glass, *Spectrochim. Acta B* 63 (2008) 1145–1150.
- [38] Y.G. Mbesse Kongbonga, H. Ghalila, M.B. Onana, Z. Ben Lakhdar, Classification of vegetable oils based on their concentration of saturated fatty acids using laser induced breakdown spectroscopy (LIBS), *Food Chem.* 147 (2014) 327–331.
- [39] J.B. Sirven, B. Bousquet, L. Canioni, L. Sarger, S. Tellier, M. Potin-Gautier, I.L. Hecho, Qualitative and quantitative investigation of chromium-polluted soils by laser-induced breakdown spectroscopy combined with neural networks analysis, *Anal. Bioanal. Chem.* 385 (2006) 256–262.
- [40] M.Z. Martin, N. Labbé, T.G. Rials, S.D. Wullschlegel, Analysis of preservative-treated wood by multivariate analysis of laser-induced breakdown spectroscopy spectra, *Spectrochim. Acta B* 60 (2005) 1179–1185.
- [41] J.R. Cordeiro, M. Martinez IV, R.W. Li, A.P. Cardoso, L.C. Nunes, F.J. Krug, T.R. Paixão, C.S. Nomura, J. Gruber, Identification of four wood species by an electronic nose and by LIBS, *Int. J. Electrochem.* 2012 (2012).
- [42] R.A. Multari, D.A. Cremers, J.M. Dupre, J.E. Gustafson, The use of laser-induced breakdown spectroscopy for distinguishing between bacterial pathogen species and strains, *Appl. Spectrosc.* 64 (2010) 750–759.
- [43] X. Zhu, T. Xu, Q. Lin, L. Liang, G. Niu, H. Lai, M. Xu, X. Wang, H. Li, Y. Duan, Advanced statistical analysis of laser-induced breakdown spectroscopy data to discriminate sedimentary rocks based on Czerny–Turner and Echelle spectrometers, *Spectrochim. Acta B* (2014) 8–13.
- [44] J.-B. Sirven, B. Salle, P. Mauchien, J.-L. Lacour, S. Maurice, G. Manhes, Feasibility study of rock identification at the surface of Mars by remote laser-induced breakdown spectroscopy and three chemometric methods, *J. Anal. At. Spectrom.* 22 (2007) 1471–1480.
- [45] J.F.C. De Lucia, J.L. Gottfried, C.A. Munson, A.W. Miziolek, Multivariate analysis of standoff laser-induced breakdown spectroscopy spectra for classification of explosive-containing residues, *Appl. Opt.* 47 (2008) G112–G121.
- [46] J.L. Gottfried, R.S. Harmon, F.C. De Lucia Jr, A.W. Miziolek, Multivariate analysis of laser-induced breakdown spectroscopy chemical signatures for geomaterial classification, *Spectrochim. Acta B* 64 (2009) 1009–1019.
- [47] O. Forni, S. Maurice, O. Gasnault, R.C. Wiens, A.S. Cousin, S.M. Clegg, J.-B. Sirven, J.R. m. Lasue, Independent component analysis classification of laser induced breakdown spectroscopy spectra, *Spectrochim. Acta B* (2013) 31–41.
- [48] J. Cisewski, E. Snyder, J. Hannig, L. Oudejans, Support vector machine classification of suspect powders using laser-induced breakdown spectroscopy (LIBS) spectral data, *J. Chemom.* 26 (2012) 143–149.
- [49] N.C. Dingari, I. Barman, A.K. Myakalwar, S.P. Tewari, M. Kumar Gundawar, Incorporation of support vector machines in the LIBS toolbox for sensitive and robust classification amidst unexpected sample and system variability, *Anal. Chem.* 84 (2012) 2686–2694.
- [50] F.-Y. Yueh, H. Zheng, J.P. Singh, S. Burgess, Preliminary evaluation of laser-induced breakdown spectroscopy for tissue classification, *Spectrochim. Acta B* 64 (2009) 1059–1067.



- [51] A.K. Myakalwar, S. Sreedhar, I. Barman, N.C. Dingari, S. Venugopal Rao, P. Prem Kiran, S.P. Tewari, G. Manoj Kumar, Laser-induced breakdown spectroscopy-based investigation and classification of pharmaceutical tablets using multivariate chemometric analysis, *Talanta* 87 (2011) 53–59.
- [52] Q. Godoi, F.O. Leme, L.C. Trevizan, E.R. Pereira Filho, I.A. Rufini, D. Santos Jr., F.J. Krug, Laser-induced breakdown spectroscopy and chemometrics for classification of toys relying on toxic elements, *Spectrochim. Acta B* 66 (2011) 138–143.
- [53] J. El Haddad, D. Bruyère, A. Ismaël, G. Gallou, V. Laperche, K. Michel, L. Canioni, B. Bousquet, Application of a series of artificial neural networks to on-site quantitative analysis of lead into real soil samples by laser induced breakdown spectroscopy, *Spectrochim. Acta B* 97 (2014) 57–64.
- [54] A. Tropsha, P. Gramatica, V.K. Gombar, The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, *QSAR Comb. Sci.* 22 (2003) 69–77.
- [55] M.R.J.C. Pontes, J. Cortez, R.K.H. Galvão, C. Pasquini, M.R.C.S.U. Araújo, R.M. Coelho, M.R.K. Chiba, M.N.F. de Abreu, B.T.E.K. Madari, Classification of Brazilian soils by using LIBS and variable selection in the wavelet domain, *Anal. Chim. Acta* 642 (2009) 12–18.
- [56] E. Benfenati, Quantitative Structure–Activity Relationships (QSAR) for Pesticide Regulatory Purposes, Chapter 6, Elsevier Science, 2011.
- [57] M.M. Tan, S. Cui, J. Yoo, S.-H. Han, K.-S. Ham, S.-H. Nam, Y. Lee, Feasibility of laser-induced breakdown spectroscopy (LIBS) for classification of sea salts, *Appl. Spectrosc.* 66 (2012) 262–271.
- [58] V. Unnikrishnan, K. Choudhary, S.D. Kulkarni, R. Nayak, V. Kartha, C. Santhosh, Analytical predictive capabilities of laser induced breakdown spectroscopy (LIBS) with principal component analysis (PCA) for plastic classification, *RSC Adv.* 3 (2013) 25872–25880.
- [59] M. Boueri, V. Motto-Ros, W. Lei, Q. Ma, L. Zheng, H. Zeng, J. Yu, Identification of polymer materials using laser-induced breakdown spectroscopy combined with artificial neural networks, *Appl. Spectrosc.* 65 (2011) 307–314.
- [60] F.C. De Lucia, J.L. Gottfried, Classification of explosive residues on organic substrates using laser induced breakdown spectroscopy, *Appl. Opt.* 51 (2012) B83–B92.
- [61] J. Remus, K.S. Dunsin, Robust validation of pattern classification methods for laser-induced breakdown spectroscopy, *Appl. Opt.* 51 (2012) B49–B56.
- [62] T. Vance, N. Reljin, A. Lazarevic, D. Pokrajac, V. Kecman, N. Melikechi, A. Marciano, Y. Markushin, S. McDaniel, Classification of LIBS protein spectra using support vector machines and adaptive local hyperplanes, *Neural Networks (IJCNN)*, the 2010 International Joint Conference on, IEEE, 2010, pp. 1–7.
- [63] M. Corsi, G. Cristoforetti, M. Giuffrida, M. Hidalgo, S. Legnaioli, L. Masotti, V. Palleschi, A. Salvetti, E. Tognoni, C. Vallebona, Archaeometric analysis of ancient copper artefacts by laser-induced breakdown spectroscopy technique, *Microchim. Acta* 152 (2005) 105–111.
- [64] T.K. Sahoo, S. Sahoo, Sensitivity of outlier(s) removal on laser induced breakdown spectroscopy (LIBS) spectral data set classification, *Int. J. Adv. Res. Comput. Commun. Eng.* 3 (2014).
- [65] A. Sarkar, D. Alamelu, S.K. Aggarwal, Gallium quantification in solution by LIBS in the presence of bulk uranium, *Opt. Laser Technol.* 44 (2012) 30–34.
- [66] M.A. Ismail, H. Imam, A. Elhassan, W.T. Younis, M.A. Harith, LIBS limit of detection and plasma parameters of some elements in two different metallic matrices, *J. Anal. At. Spectrom.* 19 (2004) 489–494.
- [67] W.T.Y. Mohamed, Improved LIBS limit of detection of Be, Mg, Si, Mn, Fe and Cu in aluminum alloy samples using a portable Echelle spectrometer with ICCD camera, *Opt. Laser Technol.* 40 (2008) 30–38.
- [68] J.-M. Mermet, Calibration in atomic spectrometry: a tutorial review dealing with quality criteria, weighting procedures and possible curvatures, *Spectrochim. Acta B* 65 (2010) 509–523.
- [69] S.C. Snyder, W.G. Wickun, J.M. Mode, B.D. Gurney, F.G. Michels, The detection of palladium particles in proton exchange membrane fuel-cell water by laser-induced breakdown spectroscopy (LIBS), *Appl. Spectrosc.* 65 (2011) 642–647.
- [70] J. Xiu, V. Motto-Ros, G. Panczer, R. Zheng, J. Yu, Feasibility of wear metal analysis in oils with parts per million and sub-parts per million sensitivities using laser-induced breakdown spectroscopy of thin oil layer on metallic target, *Spectrochim. Acta B* 91 (2014) 24–30.
- [71] J. El Haddad, M. Villot-Kadri, A. Ismaël, G. Gallou, K. Michel, D. Bruyère, V. Laperche, L. Canioni, B. Bousquet, Artificial neural network for on-site quantitative analysis of soils using laser induced breakdown spectroscopy, *Spectrochim. Acta B* 79–80 (2013) 51–57.
- [72] A. Golbraikh, A. Tropsha, Beware of q<sup>2</sup>! *J. Mol. Graph. Model.* 20 (2002) 269–276.
- [73] L. Eriksson, J. Jaworska, A.P. Worth, M.T. Cronin, R.M. McDowell, P. Gramatica, Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs, *Environ. Health Perspect.* 111 (2003) 1361–1375.
- [74] J.-M. Mermet, Limit of quantitation in atomic spectrometry: an unambiguous concept? *Spectrochim. Acta B* 63 (2008) 166–182.
- [75] A. Menditto, M. Patriarca, B. Magnusson, Understanding the meaning of accuracy, trueness and precision, *Accred. Qual. Assur.* 12 (2007) 45–47.
- [76] X. Li, Z. Wang, Y. Fu, Z. Li, J. Liu, W. Ni, The application of spectrum standardization method for carbon analysis in coal using laser-induced breakdown spectroscopy, *arXiv2014*. (preprint arXiv:1402.2060).
- [77] F.C. De Lucia, J.L. Gottfried, Characterization of a series of nitrogen-rich molecules using laser induced breakdown spectroscopy, *Propellants Explos. Pyrotech.* 35 (2010) 268–277.
- [78] J. Jasik, J. Heitz, J.D. Pedarnig, P. Veis, Vacuum ultraviolet laser-induced breakdown spectroscopy analysis of polymers, *Spectrochim. Acta B* 64 (2009) 1128–1134.
- [79] G.G.A. de Carvalho, L.C. Nunes, P.F. de Souza, F.J. Krug, T.C. Alegre, D. Santos Jr., Evaluation of laser induced breakdown spectrometry for the determination of macro and micronutrients in pharmaceutical tablets, *J. Anal. At. Spectrom.* 25 (2010) 803–809.
- [80] A. Haider, M. Hedayet Ullah, Z. Khan, F. Kabir, K. Abedin, Detection of trace amount of arsenic in groundwater by laser-induced breakdown spectroscopy and adsorption, *Opt. Laser Technol.* 56 (2014) 299–303.
- [81] M. Gondal, Y. Maganda, M. Dastageer, F. Al Adel, A. Naqvi, T. Qahtan, Detection of the level of fluoride in the commercially available toothpaste using laser induced breakdown spectroscopy with the marker atomic transition line of neutral fluorine at 731.1 nm, *Opt. Laser Technol.* 57 (2014) 32–38.
- [82] Z.-b. Cong, L.-x. Sun, Y. Xin, Y. Li, L.-f. Qi, Comparison of calibration curve method and partial least square method in the laser induced breakdown spectroscopy quantitative analysis, *J. Comput. Commun.* 1 (2013) 14.
- [83] M.M. Tripathi, K.K. Srinivasan, S.R. Krishnan, F.-Y. Yueh, J.P. Singh, A comparison of multivariate LIBS and chemiluminescence-based local equivalence ratio measurements in premixed atmospheric methane–air flames, *Fuel* 106 (2012) 318–326.
- [84] M.D. Dyar, M.L. Carmosino, E.A. Breves, M.V. Ozanne, S.M. Clegg, R.C. Wiens, Comparison of partial least squares and lasso regression techniques as applied to laser-induced breakdown spectroscopy of geological samples, *Spectrochim. Acta B* 70 (2012) 51–67.
- [85] J.M. Andrade, G. Cristoforetti, S. Legnaioli, G. Lorenzetti, V. Palleschi, A.A. Shaltout, Classical univariate calibration and partial least squares for quantitative analysis of brass samples by laser-induced breakdown spectroscopy, *Spectrochim. Acta B* 65 (2010) 658–663.
- [86] R.C. Wiens, S. Maurice, J. Lasue, O. Forni, R.B. Anderson, S. Clegg, S. Bender, D. Blaney, B.L. Barraclough, A. Cousin, L. Deflores, D. Delapp, M.D. Dyar, C. Fabre, O. Gasnault, N. Lanza, J. Mazoyer, N. Melikechi, P.Y. Meslin, H. Newsom, A. Ollila, R. Perez, R.L. Tokar, D. Vaniman, Pre-flight calibration and initial data processing for the ChemCam laser-induced breakdown spectroscopy instrument on the Mars Science Laboratory rover, *Spectrochim. Acta B* 82 (2013) 1–27.
- [87] F.R. Doucet, T.F. Belliveau, J.-L. Fortier, J. Hubert, Use of chemometrics and laser-induced breakdown spectroscopy for quantitative analysis of major and minor elements in aluminium alloys, *Appl. Spectrosc.* 61 (2007) 327–332.
- [88] R.S. Bricklemeyer, D.J. Brown, P.J. Turk, S.M. Clegg, Improved intact soil–core carbon determination applying regression shrinkage and variable selection techniques to complete spectrum laser-induced breakdown spectroscopy (LIBS), *Appl. Spectrosc.* 67 (2013) 1185–1199.
- [89] T. Yuan, Z. Wang, Z. Li, W. Ni, J. Liu, A partial least squares and wavelet-transform hybrid model to analyze carbon content in coal using laser-induced breakdown spectroscopy, *Anal. Chim. Acta* 807 (2014) 29–35.
- [90] P.M. Mukhono, K.H. Angeyo, A. Dehayem-Kamadjeu, K.A. Kaduki, Laser induced breakdown spectroscopy and characterization of environmental matrices utilizing multivariate chemometrics, *Spectrochim. Acta B* 87 (2013) 81–85.
- [91] R.B. Anderson, J.F. Bell III, R.C. Wiens, R.V. Morris, S.M. Clegg, Clustering and training set selection methods for improving the accuracy of quantitative laser induced breakdown spectroscopy, *Spectrochim. Acta B* 70 (2012) 24–32.
- [92] P. Gramatica, P. Pilutti, E. Papa, Validated QSAR prediction of OH tropospheric degradation of VOCs: splitting into training–test sets and consensus modeling, *J. Chem. Inf. Comput. Sci.* 44 (2004) 1794–1802.
- [93] B.J. Taylor, M.A. Darrah, C.D. Moats, Verification and validation of neural networks: a sampling of research in progress, *Proceedings of Society of Photo-optical Instrumentation Engineers*, 2003, pp. 8–16.
- [94] L.-T. Qin, S.-S. Liu, F. Chen, Q.-F. Xiao, Q.-S. Wu, Chemometric model for predicting retention indices of constituents of essential oils, *Chemosphere* 90 (2013) 300–305.
- [95] F. Rezaei, P. Karimi, S. Tavassoli, Effect of self-absorption correction on LIBS measurements by calibration curve and artificial neural network, *Appl. Phys. B* (2013) 1–10.
- [96] M.Z. Martin, M.A. Mayes, K.R. Heal, D.J. Brice, S.D. Wulschleger, Investigation of laser-induced breakdown spectroscopy and multivariate analysis for differentiating inorganic and organic C in a variety of soils, *Spectrochim. Acta B* (2013) 100–107.
- [97] F. Barbieri Gonzaga, C. Pasquini, A compact and low cost laser induced breakdown spectroscopic system: application for simultaneous determination of chromium and nickel in steel using multivariate calibration, *Spectrochim. Acta B* 69 (2012) 20–24.
- [98] X. Li, Z. Wang, S.-L. Lui, Y. Fu, Z. Li, J. Liu, W. Ni, A partial least squares based spectrum normalization method for uncertainty reduction for laser-induced breakdown spectroscopy measurements, *Spectrochim. Acta B* 88 (2013) 180–185.
- [99] A.C. Olivieri, N.K.M. Faber, Standard error of prediction in parallel factor analysis of three-way data, *Chemom. Intell. Lab. Syst. 70* (2004) 75–82.
- [100] P. Valderrama, J.W.B. Braga, R.J. Poppi, Variable selection, outlier detection, and figures of merit estimation in a partial least-squares regression multivariate calibration model. A case study for the determination of quality parameters in the alcohol industry by near-infrared spectroscopy, *J. Agric. Food Chem.* 55 (2007) 8331–8338.
- [101] J.W.B. Braga, L.C. Trevizan, L.C. Nunes, I.A. Rufini, D. Santos Jr., F.J. Krug, Comparison of univariate and multivariate calibration for the determination of micronutrients in pellets of plant materials by laser induced breakdown spectrometry, *Spectrochim. Acta B* 65 (2010) 66–74.
- [102] C. Rucker, G. Rucker, M. Meringer, y-Randomization and its variants in QSPR/QSAR, *J. Chem. Inf. Model.* 47 (2007) 2345–2357.
- [103] H. Kubinyi, Comparative molecular field analysis (CoMFA), *Handbook of Chemoinformatics*, Wiley-VCH Verlag GmbH, 2008, pp. 1555–1574.



- [104] S. Dimitrov, G. Dimitrova, T. Pavlov, N. Dimitrova, G. Patlewicz, J. Niemela, O. Mekenyan, A stepwise approach for defining the applicability domain of SAR and QSAR models, *J. Chem. Inf. Model.* 45 (2005) 839–849.
- [105] OECD, Guidance document on the validation of (quantitative) structure–activity relationships [(Q)SAR] models, Environment Directorate, Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology, Paris, 2007.
- [106] M.S. Escobar, H. Kaneko, K. Funatsu, Flour concentration prediction using GAPLS and GAWLS focused on data sampling issues and applicability domain, *Chemom. Intell. Lab. Syst.* 137 (2014) 33–46.
- [107] P. Sobron, A. Wang, F. Sobron, Extraction of compositional and hydration information of sulfates from laser-induced plasma spectra recorded under Mars atmospheric conditions — implications for ChemCam investigations on curiosity rover, *Spectrochim. Acta B* 68 (2012) 1–16.
- [108] E. Cerrai, R. Trucco, On the matrix effect in laser sampled spectrochemical analysis, *Energia Nucleare* 15 (1968) 581–587.
- [109] K.W. Marich, P.W. Carr, W.J. Treytl, D. Glick, Effect of matrix material on laser-induced elemental spectral emission, *Anal. Chem.* 42 (1970) 1775–1779.
- [110] W.T.Y. Mohamed, A. Askar, Study of the matrix effect on the plasma characterization of heavy elements in soil sediments using LIBS with a portable Echelle spectrometer, *Prog. Phys.* 1 (2007) 46–52.
- [111] S.-J. Choi, K.-J. Lee, J. Yoh, Quantitative laser-induced breakdown spectroscopy of standard reference materials of various categories, *Appl. Phys. B* (2013) 1–10.