We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

4,400 Open access books available
117,000 International authors and editors
130M Downloads

154 Countries delivered to
TOP 1% Our authors are among the most cited scientists
12.2% Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter 5

Calibration Methods of Laser-Induced Breakdown Spectroscopy

Hongbo Fu, Junwei Jia, Huadong Wang, Zhibo Ni and Fengzhong Dong

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.72888

Abstract

Laser-induced breakdown spectroscopy (LIBS) has gained great attention over the past two decades due to its many advantages, such as needless sample preparation, capability of remote measurement and fast multielement simultaneous analysis. However, because of its inherent uncertainty features of plasma, it is still a big challenge for LIBS community worldwide to realize high sensitivity and accurate quantitative analysis. Currently, many chemometric analytical methods have been applied to LIBS calibration analysis, including univariate regression, multivariate regression, principal component regression (PCR), partial least squares regression (PLSR) and so on. In addition, appropriate sample and spectral pretreatment can effectively improve the analytical performance (i.e., limit of detection (LOD), accuracy and repeatability) of LIBS. In this chapter, we briefly summarize the progress of these calibration methods and their applications on LIBS and provide our recommendations.

Keywords: laser-induced breakdown spectroscopy, sample and spectral pretreatment, calibration methods, chemometrics, calibration-free laser-induced breakdown spectroscopy

1. Introduction

Laser-induced breakdown spectroscopy (LIBS), also sometimes called laser-induced plasma spectroscopy (LIPS), has developed rapidly as an analytical technique over the past two decades. LIBS is a kind of atomic emission spectroscopy, which uses a high-energy pulsed laser as the excitation source. The laser is focused on the sample surface, thereby evaporating and atomizing the sample and generating a plasma. The light emitted by the plasma is detected by a spectrometer. One can obtain sample composition and concentration information by analyzing the plasma
emission spectra. LIBS most commonly used experimental instruments such as mainly lasers, spectrometers, detectors and computers, as shown in Figure 1.

The laser which is the most widely used in LIBS is the Nd:YAG solid-state lasers operated in the Q-switch mode. Typically, it is operated at the fundamental wavelength of 1064 nm, pulse energy is 30–100 mJ, pulse width is 5–15 ns and repetition rate is 1–10 Hz. In addition, the researchers tested the effects of lasers of different types and parameters on the LIBS. Trautner et al. [1] investigated polyethylene (PE) and a rubber material from tire production by employing 157 nm F2 laser and 532 nm Nd:YAG laser ablation in nitrogen and argon gas background or in air. The effects of laser wavelength on depth resolution of thin-film solar cell are investigated by Choi et al. [2] using an ultraviolet (λ = 266 nm) and a visible (λ = 532 nm) nanosecond Nd:YAG lasers. Labutin et al. [3] summarize nearly two decades of studies on femtosecond laser-induced breakdown spectrometry (fs-LIBS). Picosecond pulse train and nanosecond pulse were compared for laser ablation and LIBS measurements by Lednev et al. [4].

Spectrometers disperse the emitted radiation of the laser-induced plasma to get a spectrum in terms of intensity as a function of the wavelength. The dominant spectrometer types used for LIBS are multichannel fiber spectrometer and echelle spectrometer coupled with an intensified CCD. The echelle spectrometer offers a wide spectral range, a high spectral resolution, and the possibility of time-resolved. The plasma parameters (plasma temperature and electron density) are constantly changing with respect to the delay time, and an echelle spectrometer with time resolution is needed when calculating these parameters. However, time-resolved broadband spectrometers are expensive and strongly dependent on external circumstances. The multichannel fiber spectrometer is robust and reliable for the use in mobile and portable LIBS instruments, providing an accredited spectral resolution, but their integration time is typically much longer than the plasma lifetime.

In recent years, with the rapid development of lasers, spectrometers and detectors and the urgent demand of in situ and online analysis, LIBS has developed rapidly. Compared with many other types of elemental analysis techniques, LIBS has obvious advantages:

Figure 1. A schematic of a general apparatus for laser-induced breakdown spectroscopy illustrating the principal components.
1. Simple equipment: few instruments, low cost and easy integration.

2. Noncontact analysis: LIBS uses pulsed laser as the excitation source, which makes it noncontact analysis, especially in the dangerous environment or space exploration field, has broad application prospects.

3. No sample preparation: LIBS directly focused pulsed laser bombardment of the sample without processing the sample.

4. Various samples: samples can be gas, aerosols, liquids and solids.

5. Nondestructive analysis: The laser converges to the surface of the sample, and only a small amount of the sample is excited. It can be considered as nondestructive or near nondestructive.

6. Three-dimensional analysis: LIBS can collect laser at different positions on the sample surface or repeat measurements at the same location to analyze the sample surface and its different depths of the sample composition and content.

7. Total element analysis: The laser energy can simultaneously excite all the elements in the sample, so all elements in the sample can be analyzed simultaneously.

8. Remote analysis: The long-distance analysis of the LIBS can be achieved by remotely transmitting the laser energy and collecting the plasma emission spectrum through the fiber.

9. Online analysis: LIBS is a very fast technology that provides analytical results in seconds, making it particularly suitable for rapid analysis or online industrial monitoring.

Figure 2. The number of articles published in the Web of Science search by laser-induced breakdown spectroscopy (LIBS) in 1963–2016.
Because of many advantages of LIBS, it has been applied to a number of analytical domains, for example, various alloys [5–7], slags [8, 9], soil [10], rocks [11, 12] and isotopes [13]. We searched for all scientific papers from 1963 to 2016 on Web of Science with laser-induced breakdown spectroscopy (LIBS). The statistical results are presented in Figure 2. It can be seen that LIBS has developed rapidly since 1990.

2. Pretreatment of samples and spectra

2.1. Sample pretreatment

One of the most widely cited advantages of LIBS is that it does not require sample preparation, but this may also be the biggest limitation for improving its consistancy. In general, LIBS performance may be enhanced using two main approaches: pretreatment of samples and spectra. Many homogeneous solid samples require no sample preparation, for example, glass, alloy and plastic. For powder samples (e.g., cement [14], soil [10] and coal [15]) which can press the cake directly, it must be consistent with the standard sample preparation process used for calibration during the pressing process. Comparing with solid samples, the direct analysis of liquid samples by using LIBS has many disadvantages: splash, less excitation and fluctuation of liquid level. The simplest way to change a liquid sample into a solid sample is to freeze it [16, 17]. Sobral et al. [18] investigated the detection sensitivity of Cu, Mg, Pb, Hg, Cd, Cr and Fe traces in water and ice samples under the same experimental conditions by using LIBS. Another effective way that can be used for liquid analysis in a solid matrix configuration consists of using an absorbent substrate, for example, plant fiber spunlace nonwoven [19], absorbent paper [20], thin wood sample [21, 22] and membrane-based filter paper [23]. Now LIBS analysis of aerosols is mainly of two categories: direct analysis and enrichment. However, the detection limit and the statistical results of the direct analysis are still relatively poor. On the other hand, the substrate-based collection does not provide as instantaneous information and does allow one to achieve lower detection limits by increasing the sample flow rates and sampling times.

2.2. Correction and removal of continuum background

The detected plasma emission spectrum at a given wavelength in a spectrum is the sum of the analyte signal and the continuum background. The analyte signal is often overwhelmed by the continuum background, which interferes the true intensity of signal and compromises spectral clarity and hence reduces the accuracy of quantitative analysis. Zou et al. [24] developed a modified algorithm of background removal based on wavelet transform for spectrum correction and applied to low-alloy steel samples. This method can effectively improve the quality of the signals and the accuracy of the regression model. Sun et al. [25] presented a method that can automatically estimate and correct varying continuum background emission. Simulations and experiments were made to successfully prove the efficiency of the method. The proposed method scarcely needs people’s intervention and can automatically and flexibly estimate varying continuum backgrounds over a very wide spectrum range. Another way to deduct a continuous background is to add a polarizer to the collected light path. Penczak et al.’s [26]
research results show that the continuous background of the Al plasma emission spectra induced by 800 nm femtosecond pulse laser is strongly polarized. The use of a polarizer can effectively filter the continuous spectrum, thus improving the signal-to-noise ratio and the signal-to-back ratio of the characteristic spectrum.

2.3. Spectral normalization

In order to increase the stability of the signal, the analyte signal intensity can be normalized using a parameter representative of the actual plasma conditions. In general, there are three main standardized methods [27]: (1) normalization by using the intensity of an internal standard line; (2) normalization by using a reference signal; and (3) compensation for the plasma conditions. Castro et al. [28] used 12 different types of data normalization to reduce the interference matrix and to improve the calibration models. Their findings show that the application of normalization modes was useful to compensate for the differences among sample matrices. Models without normalization presented two- to fivefold higher errors. Karki et al. [29] studied the analytical performance of six different spectrum normalization techniques, namely internal normalization, normalization with total light, normalization with background along with their three-point smoothing methods for quantification of Cr, Mn and Ni in stainless steel. The final results show the superiority of internal normalization technique over normalization with total light and normalization with background techniques irrespective of whether it is Cr, Ni or Mn analysis. Wang et al. presented three spectrum standardized methods in order to improve the reproducibility of LIBS measurements which are named the spectrum standardization approach [30], the sampled spectrum standardization approach [31] and the multivariate spectrum standardization method, respectively. In spectral standardization, a particular example is the use of acoustic signals [32, 33] or laser-induced plasma image [34].

2.4. Automatic identification of emission lines

LIBS can excite all the elements of the spectrum of the sample, so reliable and fast identification of emission lines in laser ablation of multicomponent samples is crucial. Labutin et al. [35] applied an algorithm to automatically identify emission lines in LIBS. The algorithm is implemented by three parts: simulation of the set of spectra corresponding to different temperature and electron density, searching the best correlated pair of a model spectrum and an experimental one, and attributing the peaks with certain lines. Ukwatta et al. [36] consider the problem of element detection as a multilabel classification problem, using support vector machines (SVMs) and artificial neural networks (ANNs) for multielement classification. The proposed algorithm is evaluated by using the LIBS image obtained from the experiment. The accuracy of the machine learning method to identify the elements correctly can reach 99%. Mateo et al. [37] developed the software package SALIPS, which can quickly and semiautomatically identify the spectrum peak and give the element composition of the analytical sample. The software package simulates the spectrum by using the relative intensity of the atomic line in the NIST database. In order to facilitate visual comparison, it can present both the simulated and experimental spectra on the same plot.
3. Calibration methods of LIBS

A number of calibration methods have been applied to various research fields and physical state samples of LIBS quantitative analysis. We cannot involve all published research articles, and only a few of the most commonly used calibration methods are reviewed.

3.1. Univariate analysis

The fitting area intensity \( I \) corresponding to the transition between lower level \( E_l \) and upper level \( E_u \) of an atomic species \( \alpha \) can be expressed as:

\[
I = FN'^\alpha I \frac{g_u}{U'(T)} \exp \left( -\frac{E_u}{k_B T} \right)
\]  

(1)

where \( F \) is the experimental parameter, \( N'^\alpha \) is the atomic number density, \( A_{ul} \) is the transition probability, \( g_u \) is the upper-level degeneracy and \( U'(T) \) is the partition function at the temperature \( T \). For the same sample, if the temperature and density of each laser-induced plasma are constant, then \( I \) is proportional to the elemental concentration \( C \). If there are a series of samples with different \( C \), one can establish a calibration line between spectral intensity and element concentration.

\[
C_i = b_0 + b_1 I_i + e_i
\]

(2)

where \( b_0 \) and \( b_1 \) are model parameters, \( e_i \) is the random error and \( i \) is the number of samples. The parameter \( \hat{C}_i \) is the estimate value of \( C_i \), namely

\[
\hat{C}_i = b_0 + b_1 I_i
\]

(3)

and

\[
e_i = \hat{C}_i - C_i
\]

(4)

In the regression analysis, the best estimate of \( b_0 \) and \( b_1 \) is obtained based on a set of \( I \) and \( C \), which makes the \( \hat{C} \) and \( C \) to the nearest degree. For example, Bhatt et al. [38] choose Ce II 413.38, 418.65, and 439.16 nm to establish a univariate linear calibration curve, as shown in Figure 3.

Correlation coefficient \( R \), also called the Pearson coefficient, is often used to denote the correlation between \( I \) and \( C \), which is defined as:

\[
R = \frac{\sum_{i=1}^{n} (I_i - \bar{I})(C_i - \bar{C})}{\sqrt{\sum_{i=1}^{n} (I_i - \bar{I})^2 \sum_{i=1}^{n} (C_i - \bar{C})^2}}
\]

(5)

Correlation coefficient \( |R| \leq 1 \), and it is closer to 1, indicating the better relevance. Most of the LIBS papers report \( R^2 \) which 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. However, it should be noticed that a model with a value of $R^2$ close to 1 may indeed have a poor accuracy for prediction [39].

Precision is described by the standard deviation (SD), the relative standard deviation (RSD in %) and the root-mean-square error (RMSE), which can be expressed as:

$$SD = \sqrt{\frac{\sum_{i=1}^{n} (C_i - \bar{C})^2}{n}}$$  \hspace{1cm} (6)

$$RSD(\%) = \frac{SD}{\bar{C}} \times 100$$  \hspace{1cm} (7)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{C}_i - C_i)^2}{n}}$$  \hspace{1cm} (8)

In order to describe the lower limits of a quantitative model, the limit of detection (LOD) can be calculated by the following equation:

$$LOD = 3 \frac{\sigma \bar{C}}{T}$$  \hspace{1cm} (9)

where $\sigma$ is the standard deviation of the background and $C/I$ is the reciprocal of the slope of the calibration curve. The calculated values of LOD for different elements are presented in Table 1.

3.2. Multivariate analysis

3.2.1. Multiple linear regression

For LIBS, the line of an element is not one. If there are $m$ variables and $n$ samples, then
\[ C_i = b_0 + b_1 I_{i1} + \ldots + b_m I_{im} + e_i \quad (i = 1, 2, \ldots, n), \quad (10) \]

\[ Q = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (C_i - \bar{C}_i)^2 \quad (11) \]

where \( I \) is the intensity of different spectral lines from the same element, \( e_i \) is residual, \( Q \) is the sum of squares of the residuals and \( \bar{C}_i \) is the estimate value of \( C_i \). One can get \( b_0, b_1, \ldots, b_m \) when \( Q \) value achieves the minimum value.

For example, Chen et al. [40] used the multiple linear regression method to quantitatively analyze chromium in potatoes. The characteristic line of Cr can be considered that the concentration of Cr (\( C_i \)) has a relationship with the intensity of Cr and/or the other corresponding elements. They normalize the quantitative analysis of Cr by considering the influence of the Ca matrix. Four independent variables (\( I_{Cr}, I_{Cr}^{sum}, I_{Ca}, \) and \( I_{Ca}^{sum} \)) are used to test the performance of different linear regression methods, where \( I_{Cr} \) is the intensity of Cr I 425.43 nm, \( I_{Cr}^{sum} \) is the sum of three Cr lines (Cr I 425.43, Cr I 427.48 and Cr I 428.97 nm), \( I_{Ca} \) is the intensity of Ca I 431.86 nm and \( I_{Ca}^{sum} \) is the sum of five Ca lines (Ca I 422.67, Ca I 428.30, Ca I 430.25, Ca I 430.77, and Ca I 431.86 nm). Different combinations of the four independent variables were selected for unary, binary, ternary and quaternary linear regression analyses. The results of quantitative analysis of Cr element by the linear regression method with different variables are indicated in Table 2.

### 3.2.2. Principal component regression

In the LIBS quantitative analysis, the calculated concentration is affected by the lines of objective elements and other elements. In the study of empirical questions, in order to analyze the problem comprehensively and systematically, one must consider many spectral lines of many elements. Because each line reflects the information of the element concentration in varying degrees, and the lines have a certain correlation with each other, the information reflected in the calculation overlaps to some extent. In the study of multivariate problems by statistical methods, too many variables will increase the amount of calculation and increase the complexity of the problem. It is hoped that the variables involved in the process of quantitative analysis are less and the amount of information is more. Principal component analysis (PCA) is adapted to this requirement and is an ideal tool for solving such problems.

### Table 1. Limit of detection (LOD) estimated for different elements.

<table>
<thead>
<tr>
<th>Element</th>
<th>LOD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ce</td>
<td>0.098</td>
</tr>
<tr>
<td>Eu</td>
<td>0.052</td>
</tr>
<tr>
<td>Gd</td>
<td>0.077</td>
</tr>
<tr>
<td>Nd</td>
<td>0.047</td>
</tr>
<tr>
<td>Sm</td>
<td>0.25</td>
</tr>
<tr>
<td>Y</td>
<td>0.036</td>
</tr>
</tbody>
</table>

In the table, LOD is the limit of detection for different elements.
Principal component analysis can transform a set of variables that may have correlation into a set of linearly uncorrelated variables by orthogonal transformation. The variable after the conversion is called the principal component. Principal component regression (PCR) is a regression analysis method for analyzing multiple regression and is based on PCA. In general, predicting the concentration by PCR can be divided into three steps: first, the PCA is performed on the data (spectral) matrix of the original independent variables, and the appropriate number of principal components is selected by finding the eigenvalue, eigenvector, variance contribution rate and cumulative contribution rate. Second, the selected principal component is analyzed by the ordinary least-squares method. Finally, the strongest possible correlations between the orthogonal PC scores and elemental composition are established.

When selecting principal components by PCR, only the independent variables are taken into account, and the dependent variables are ignored. It can reduce the dimension of variables and address the problem of multiple collinearity but cannot distinguish noise when there is a lot of noise in the independent variable (signal) and lose some information of the original variables, so a better regression model will not be obtained.

Death et al. [41] applied PCR to determine the elemental composition of a series of run-of-mine (ROM) iron ore samples. LIBS spectral data were recorded in three separate spectral regions (250 nm, 400 nm and 750 nm) to measure major, minor and trace components of the iron ore sample pellets. Background stripping, normalization and spectral cleaning were applied to minimize RSD of the LIBS data. PCR analysis was used to produce calibration models of Fe, Al, Si, Mn, K and P. Independent LIBS measurement data are used to verify these calibration models. The model \( R^2 \) for Fe, Al, Si and K is 0.99, 0.98, 0.99 and 0.84, respectively. As an example, PCR calibration model of Fe is shown in Figure 4 [41].

### Table 2. Quantitative results of Cr by different linear regression methods.

<table>
<thead>
<tr>
<th>Calibrate method</th>
<th>Input variables</th>
<th>( R^2 )</th>
<th>Predicted value (μg/g)</th>
<th>Relative error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unary</td>
<td>( I_{Cr} )</td>
<td>0.007</td>
<td>S1:138.208</td>
<td>7.674</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S2:140.308</td>
<td>7.771</td>
</tr>
<tr>
<td>Binary</td>
<td>( I_{Cr} + \sum_{i} I_{Cr} )</td>
<td>0.887</td>
<td>S1:143.649</td>
<td>11.913</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S2:132.129</td>
<td>13.147</td>
</tr>
<tr>
<td>Ternary</td>
<td>( I_{Cr} + \sum_{i} I_{Cr} + I_{Ca} )</td>
<td>0.890</td>
<td>S1:142.413</td>
<td>10.950</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S2:132.687</td>
<td>12.780</td>
</tr>
<tr>
<td>Quaternary</td>
<td>( I_{Cr} + \sum_{i} I_{Cr} + I_{Ca} + I_{P} )</td>
<td>0.987</td>
<td>S1:133.659</td>
<td>4.130</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S2:159.965</td>
<td>5.151</td>
</tr>
</tbody>
</table>

3.2.3. Partial least squares regression

The main purpose of PCR is to extract relevant information hidden in the spectral line and then used to predict the concentration. This approach allows one to use only those independent variables, and the noise will be eliminated, so as to improve the quality of the predictive model. However, PCR still has some defects, and some useful variables whose correlation is very small are easily missed when the principal component is selected. If we choose for
each component, it is too difficult. Partial least squares regression (PLSR) is a new multivari-ate statistical data analysis method. It mainly studies the regression modeling of multiple dependent variables to multiple independent variables. PLSR is more effective, especially when the variables are highly linearly correlated. In addition, PLSR solves the problem that the number of samples is less than the number of variables. Partial least squares (PLS) is the advantage of three analytical methods, which are PCA, canonical correlation analysis and

![Figure 4](image)

PCR calibration model determined for iron using the 250-nm LIBS data.

<table>
<thead>
<tr>
<th>Region</th>
<th>8 mJ</th>
<th>3 mJ</th>
<th>8 mJ</th>
<th>3 mJ</th>
<th>8 mJ</th>
<th>3 mJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated spectra</td>
<td>WBC</td>
<td>3.47</td>
<td>3.27</td>
<td>2.74 (2)</td>
<td>2.72 (2)</td>
<td>2.63 (2)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3.67</td>
<td>3.76</td>
<td>3.36</td>
<td>2.92</td>
<td>2.85</td>
</tr>
<tr>
<td></td>
<td>BC</td>
<td>1.97</td>
<td>2.27</td>
<td>2.56</td>
<td>2.27</td>
<td>2.58 (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.67</td>
<td>2.78</td>
<td>2.09 (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.22</td>
<td>2.54</td>
<td>2.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.36</td>
<td>3.57</td>
<td>3.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.46</td>
<td>3.57</td>
<td>3.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.32</td>
<td>2.9</td>
<td>4.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.36</td>
<td>2.56</td>
<td>3.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.11</td>
<td>3.46</td>
<td>3.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.07</td>
<td>3.46</td>
<td>3.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.32</td>
<td>2.9</td>
<td>4.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.36</td>
<td>3.57</td>
<td>3.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.02</td>
<td>2.73 (2)</td>
<td>3.96 (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.22</td>
<td>3.09</td>
<td>3.73 (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.78</td>
<td>2.78</td>
<td>2.78</td>
<td></td>
</tr>
<tr>
<td>Time-resolved spectra</td>
<td>WBC</td>
<td>3.07</td>
<td>3.13</td>
<td>2.23</td>
<td>2.47 (2)</td>
<td>3.73 (2)</td>
</tr>
<tr>
<td></td>
<td>BC</td>
<td>3.07</td>
<td>3.13</td>
<td>2.23</td>
<td>2.47 (2)</td>
<td>3.73 (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.32</td>
<td>2.9</td>
<td>4.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.36</td>
<td>3.57</td>
<td>3.31</td>
<td></td>
</tr>
<tr>
<td>Values in brackets correspond to number of factors if different from three. WBC = without background correction, BC = with background correction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Standard error of calibration (SEP) and prediction (SEV, in italics) estimated during the determination of silver for autoscaled data.
multiple linear regression analysis. Both PLS and PCA try to extract the maximum information reflecting the data variation, but PCA only considers one independent variable matrix, while PLS has a response matrix, so it has predictive function. PLS avoids potential problems such as nonnormal distribution of data, factor indeterminacy and unidentifiable models. PLS has two types (PLS-1 and PLS-2), and PLS-1 corresponds to the case where there is only one dependent variable. PLS-2 corresponds to the case where there are several dependent variables. Although FLSR is more complex than PCR, and the tendency of overfitting is stronger, better results can be obtained by using PLS to analyze low-precision data or high-complexity systems.

The input variable of PLS can be characteristic spectral lines [42], partial spectral region [43] or full spectrum [44]. Amador-Hernandez et al. [45] used PLS-1 to quantify gold and silver gold and silver in Au-Ag-Cu alloys. The influence of spectral region (266–340 nm, 266–269/326–340 nm and 269–313 nm), laser energy (3 mJ, 8 mJ), background correction and integration time on the quantitative analysis of PLS was studied, respectively (Table 3).

3.2.4. Artificial neural network

To overcome the poor precision of the calibration curve methods and the limitations of nonlinear problems, scholars have proposed the use of statistical methods for the quantitative analysis by LIBS. Artificial neural networks (ANNs) are computing systems inspired by the biological neural networks that constitute animal brains. Such systems learn (progressively improve performance) to do tasks by considering examples, generally without task-specific programming. The following outstanding advantages of artificial neural networks have attracted great attention in recent years: (1) it can fully approximate any complex nonlinear relationship; (2) all the quantitative or qualitative information is stored in the neurons in the network, so it has strong robustness and fault tolerance; (3) ANNs adopts the parallel distribution processing method, so that it can perform a large number of operations quickly; (4) ANNs can learn and adapt to unknown or uncertain system; and (5) it can handle both quantitative and qualitative information at the same time. An artificial neuron network (ANN) is a computational model based on the structure and functions of biological neural networks. Artificial neural networks usually contain input layer, hidden layer (competitive layer) and output layer, as shown in Figure 5.

Figure 5. Schematic of the three-layer artificial neural network [46].
For example, El Haddad et al. [46] used an artificial neuron network to analyze the heavy metals in soil and predict the concentration of element. They used average relative error of calibration REC (%) and the average relative error of prediction REP (%) to evaluate the predictive quality of the ANN models. REC and REP were preferred to RMSE because they provide percentage instead of absolute values. Before using the artificial neural network, it is necessary to optimize the parameters such as the number of neurons and training times.

\[
REC(\%) = \frac{100}{N_c} \sum_{i=1}^{N_c} \frac{|c_i - \tilde{c}_i|}{c_i}
\]
\[
REP(\%) = \frac{100}{N_p} \sum_{i=1}^{N_p} \frac{|c_i - \tilde{c}_i|}{c_i}
\]

where \(N_c\) is the number of samples in the calibration and \(N_p\) is the number of samples in the prediction sets, respectively (Table 4).

### 3.2.5. Support vector regression

The neural network structure design depends on the designer’s experience and prior knowledge, while support vector machine (SVM) is based on statistical theory, which has a strict theoretical and mathematical basis. The neural network learning algorithm lacks quantitative analysis and complete theoretical support, and it also needs a lot of samples to learn. SVM is often used to pattern recognition, classification and regression analysis of small samples, nonlinear and high-dimensional data and can achieve very good results. SVM is based on the principle of structural risk minimization, which can ensure that the learning machine has a good generalization ability. SVM for regression prediction is called support vector regression (SVR). SVR also can guarantee the global optimality of the algorithm and avoid the local minimum problem that the neural network cannot solve. Therefore, when there are a small number of samples, it is better to use SVR instead of neural network. It is important to note that the selection of optimized significant penalty parameter \(C\) and the kernel parameter of RBF-\(\delta\) is more sensitive for the SVR model.

For example, Gu et al. [47] used three segmental spectra of 393–397 nm, 422–423 nm and 425–427 nm as the input variables of SVR model to predict the content of Cr in soil samples. They got better predictive results that \(R^2 = 0.999\) and the absolute relative error is 2.61% and the slope of the calibration curve is closer to 1, as shown in Figure 6.

<table>
<thead>
<tr>
<th>Output</th>
<th>Input element</th>
<th>REC (%)</th>
<th>REP (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al</td>
<td>Al, Ca, Ba, Fe, Ti</td>
<td>18.7 ± 0.8</td>
<td>19.3 ± 2.1</td>
</tr>
<tr>
<td>Ca</td>
<td>Ca, Ba, Fe, Ti</td>
<td>9.4 ± 0.4</td>
<td>15.2 ± 0.8</td>
</tr>
<tr>
<td>Fe</td>
<td>Fe, Ba, Ca, Ti</td>
<td>15.5 ± 0.6</td>
<td>16.8 ± 0.9</td>
</tr>
</tbody>
</table>

Table 4. Average relative errors of calibration (REC) and prediction (REP).
3.3. Calibration-free laser-induced breakdown spectroscopy

LIBS offers a strong potential for analysis in situ and in real time, not requiring complex sample preparation. This allows it to be applied quickly and extensively to qualitative analysis, but quantitative analysis is very difficult. Even with a given experimental configuration, the laser-induced breakdown spectrum is not only dependent on the concentration of the analyte but also dependent on the composition of the matrix and their polymerization state. Matrix effects play an important role in quantitative analysis of LIBS. In order to overcome the matrix
effect, Ciucci et al. [48] proposed the calibration-free laser-induced breakdown spectroscopy (CF-LIBS) approach which takes the matrix into account as a part of the analytical problem. In local thermodynamic equilibrium (LTE), excited levels are populated according to the Boltzmann distribution and ionization states are populated according to the Saha-Boltzmann equilibrium equation. Each spectral line is represented as a point in a Boltzmann plane where the slope and intercept correspond to the plasma temperature and the concentration of the corresponding element, respectively (Figure 7).

Ciucci et al. [48] proposed CF-LIBS for the first time and used it for quantitative analysis of the composition of metallic alloys and quantitative determination of the composition of the atmosphere. CF-LIBS has been applied to many samples, such as aluminum alloys, steel and iron alloys, precious alloys for jewelry, copper alloys, archeological copper artifacts, glasses, pigments on roman frescoes and on parchments, soils and rocks, meteorites, coral skeletons and human hair. However, the accuracy of CF-LIBS is still not high.

4. The comparison of calibration methods

LIBS is an analytical technique that can inspire all the elements in the sample. Univariate analysis uses only partial spectral information and suffers from the strong effects of plasma instability. More importantly, strong matrix effects prevented to apply simple calibration curves. There is no doubt that the multivariate analysis is superior to univariate analysis. This has been proven by numerous researchers, for instance, the analysis of rocks [49], rare earth elements [38], glass [50], cerium oxide [51], alloy steel [52], liquid steel [53], soil [54, 55], soybean oil [56], PZT (Lead Zirconate Titanate) ceramics [57], Pb in navel orange [58], Marcellus Shale [59], tailing cores [60], geologically diverse samples [49], steel melt [61], slurry [62], iron ore [63] and pellets of plant materials [64].

Many multivariate analysis methods have been applied to the quantitative analysis of LIBS, especially chemometric. Generally, the most common chemometric technique applied to concentration measurement by LIBS is PLS. It has been applied to many fields of analysis, such as soil [55, 65, 66], steel [67–69], glass [50], rock [70], iron ore [63] and coal [71, 72]. The rest of the analysis methods are PCR [50, 73, 74], LASSO [75, 76], kNN [77], ANN [78–81], SVM [82, 83] and so on. PLS has been implemented either to calculate the concentrations of a single element (PLS-1) or to simultaneously calculate the concentrations of more than one element (PLS-2). In addition to PLS, other linear (MLR [84, 85], PCR [50, 73] and LASSO [75–77]) and nonlinear regression methods (ANN [86, 87]) have been applied to LIBS quantitative analysis. In order to discern the most effective models for interpreting chemical abundances from LIBS spectra of geological samples, Boucher et al. [77] had studied nine kinds of linear and nonlinear regression methods. The advantages and disadvantages of various methods are introduced, as shown in Table 5. The final results show that nonlinear methods tend to overfit the data and predict less accurately, while the linear methods proved to be more generalizable with better predictive performance. The performance of different models for different oxides is different. At present, multivariate analysis, especially
PLS, is our best choice. But it should be pointed out that multivariate quantitative analyses present a high risk of overfitting.

CF-LIBS can overcome the influence of matrix effect, but the poor analysis accuracy has been the fatal shortcoming of CF-LIBS. This is mainly due to the fact that the laser-induced plasma is a very complex object and its realistic description is not attainable with simple mathematical models [88]. A number of researchers have made some modifications to the CF-LIBS algorithm, such as self-absorption [89, 90]. In recent years, several research groups [91–94] began using standard samples to improve the accuracy of the nonstandard analysis. Cavalcanti et al. [92] presented and used one-point-calibration CF-LIBS to analyze a set of copper-based samples. The results show that the new method achieves similar or even higher accuracy than the calibration curve.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant nos. 11075184 and 61505223) and the Knowledge Innovation Program of the Chinese Academy of Sciences (Grant no. Y03RC21124).
Author details

Hongbo Fu1,2, Junwei Jia1,3, Huadong Wang1,3, Zhibo Ni1 and Fengzhong Dong1,2,3*

*Address all correspondence to: fzdong@aiofm.ac.cn

1 Anhui Provincial Key Laboratory of Photonic Devices and Materials, Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, China
2 Key Laboratory of Environmental Optics and Technology, Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, China
3 University of Science and Technology of China, Hefei, China

References


(SVM) and partial least square (PLS) methods. Journal of Analytical Atomic Spectrometry. 2015;30:368-374. DOI: 10.1039/c4ja00421c


[60] Kuhn K, Meima JA, Rammlmair D, et al. Chemical mapping of mine waste drill cores with laser-induced breakdown spectroscopy (LIBS) and energy dispersive X-ray fluorescence


