Brazilian Journal of Experimental Design, Data Analysis and Inferential Statistics
Federal University of Rio de Janeiro
BJEDIS, Rio de Janeiro, BJEDIS 2 YEARS, v. 3 (1), 2024
DOI: https://doi.org/10.55747/bjedis.v1i1.57177
ISSN: 2763-6925

Technical Notes | Received 20/11/2023. Revised 25/11/2023. Accepted 30/11/2023. Published 12/01/2024.

Laser-Induced Breakdown Spectroscopy Optimization Using Response Surface Methodology

Rodrigo Rossi de Araújo[1], Débora Marcondes Bastos Pereira Milori[2], Milene Corso Mitsuyuki[2], Célia Regina Montes[1], Adolpho José Melfi[1]

UNIVERSIDADE DE SÃO PAULO – USP
EMBRAPA - BRASIL
rodrigorossi300@gmail.com , debora.milori@embrapa.com, milene.corso@embrapa.br , crmlauar@usp.br, ajmelfi@usp.br

Abstract:
Laser-induced breakdown spectroscopy (LIBS) is a modern analytical technique that is capable of fast, multi-elemental, low-cost and environmental friendly analysis, which does not require complex sample preparation. Albeit its potential, LIBS analysis still presents many limitations in terms of sensitivity and reproducibility, especially when analyzing complex matrices such as of sediments and soil samples. In order to reduce these matrix related effects, it is highly recommended that the system temporal parameters, which are responsible for controlling plasma evolution and signal collection, are optimized beforehand. In this work, we proposed the design of experiments (DOE) tool – specifically, the response surface methodology (RSM) – as an approach to optimize LIBS's most important parameters (delay-time, interpulse delay, gate width and accumulated pulse). Signal-to-noise ratio (SNR) of the emission lines for Zn, Cd, Mg, Al, Ni, Cu, Ca, Cr, Sr, Fe were the response variable assessed during the procedure**.** The results showed that the RSM was an effective optimization tool for LIBS parameters and the final condition improved SNR ratios by up to a 48 ratio, when comparing to the not-optimal conditions.

Keywords: LIBS; DP-LIBS, Response surface methodology; Experimental design; Factorial designs; Optimization; Soil.

Adherence to the BJEDIS' scope: This essay is related to BJEDIS' scope as it presents experimental design and factorial design as applicable methodologies in the optimization of LIBS system for soil analysis.

1. INTRODUCTION

Laser-Induced Breakdown Spectroscopy (LIBS) is a type of optical emission technique that uses a high-energy laser beam focused by a set of optical apparatus on the surface of a sample, to ablate some of its mass and generate a small size, luminous, plasma in the process [1-2]. The atoms and molecules present in the sample absorb the energy from this laser source and, as part of the sample vaporizes and the plasma forms and expands, all of these species are excited onto higher energy levels [2-3]. When these high energy levels atoms and ions return to their respective lower energy levels, and the plasma begins cooling, a photon is emitted, which is then collected by an optical fiber cable and processed by a spectrometer, thus resulting in continuous emission spectra [4]. Since each atom has unique energy levels distribution, sample's elemental composition can be easily assessed by separating and identifying each of their respective wavelengths in the spectra, and the intensity of these emission lines can be directly correlated to the quantity of the element in the sample [1-2].

LIBS is a very recent technique that has been drawing more attention in the past two to three decades because of its distinct advantages over the other more commonly-used, or classical, techniques, such as inductively coupled plasma atomic emission spectroscopy (ICP-AES), atomic absorption spectroscopy (AAS) or even X-ray fluorescence (XRF) [1, 4]. The main advantages that can be noted are: (i) the analysis is faster and relatively simpler to perform; (ii) there are no, or very few, requirements for sample preparation before analysis; (iii) it's a multi-element, simultaneous, analysis; and there are plenty of emission wavelengths and information available in online databases; (iv) little damage is done to the sample; (v) only a few milligrams of sample are necessary for each analysis; (vi) LIBS has the possibility for in-situ, real-time, analysis [5-6].

Although this list of advantages come forward as very attractive, LIBS still presents many challenges and limitations for their users, and some of its more blatant flaws come off, in better display, during quantitative analysis. Matrix effects are a very common by-product of the unavoidable chaotic nature of LIBS analysis, and occur by fluctuations in both the intensity of the emitted laser itself and the non-linear interaction between the laser and the sample, which can undermine its sensitivity and precision [7-8]. In the past years, a number of studies [7-10] have been carried out trying to find alternatives to overcome this problem, and much information and advancements have already been attained. One of the few ways that matrix effects can be reduced, before any analysis is performed, is by having a firm grasp over the time-control parameters that the instrument provides. By understanding and changing these parameters accordingly, some of the matrix-effects can be diminished substantially, because they can dictate the manner that the plasma is generated and collected [9 -11]. When trying to optimize which conditions are the most effective for the analysis, many approaches are possible, but overall, the design of experiments (DOE) is usually a very effective way to try to optimize experiments [11], such as the Response Surface Methodology (RSM), which we hoped to prove in this study.

1.1 LIBS Theory Summary

The evolution/duration of the plasma in a standard LIBS analysis is very short (about 10^{-9} seconds), but it is a very complex phenomenon and it can be divided, temporally, into separate stages [2]. Initially, the energy of the laser is absorbed by the sample's surface particles, which then promotes its ablation, vaporizing and heating (up to 20.000 K), at the focal point of the incidence of the beam [12]. This initial stage is very rapid (just 10^{-13} seconds) and both the heat and the electric field gradient generated helps the breaking of molecular bonds, ionization of elements and the formation of a vapor cloud consisted of free electrons and charged species [13]. This high-pressure vapor cloud expands quickly, creating a shockwave during the process and promoting the appearance of small size condensed particles (which increase scattering and absorption of laser radiation and, therefore, the temperature of the plasma) [2, 14]. The next stages of the plasma evolution will vary a lot depending on the matrix of the sample, the potency and wavelength of the laser and the interaction between the laser and the sample. Nevertheless, two main radiation-absorbing effects can occur at this stage: the inverse Bremsstrahlung or multiphoton ionization [13-14]. The first is the continuous absorption of laser photons by free electrons in the plasma, increasing their kinetic energy and, by collision, inducing other atoms present into ionization [15]. This condition is more favored for plasmas with higher pressures and electronic densities [16]. On the other hand, multiphoton ionization occurs with lower pressures and electronic densities, and it takes place when atoms absorbs an amount of energy greater than that required for their ionization and, thus, causing the loss of one or more electrons [16]. By the end of the laser's pulse, the plasma is still expanding and cooling, and starts loosing energy due to the emission of radiation when the excited species return to their lower energy states [2, 13-14].

In the beginning, the initial emissions are comprised largely of continuous radiation, or background radiation, that happens due to the inverse Brehmsstrahlung and multiphoton ionization effects [14]. These emissions are the

result of free transitions of charged particles and accelerated electrons in the plasma, and can last for a few hundreds of nanoseconds. The atomic and ionic transitions normally only start after about one microsecond after the laser pulse, when the continuous radiation begins to drop significantly [2, 17]. Consequently, by understanding the behavior of the plasma evolution until extinction, we can manually adjust LIBS temporal parameters to better control the radiation collection and, thus, improve the overall quality of the final spectra.

The standard LIBS system parameters are: i) delay-time; ii) gate width; iii) interpulse-delay (for double-pulse LIBS, or DP-LIBS, systems); iv) accumulated pulses; v) energy and wavelength of the lasers; vi) order of the laser pulses and geometry configuration (DP-LIBS only) [18]. The delay-time corresponds to the time interval given between the start of the laser pulse and the signal acquisition by the spectrometer, and by altering the delay-time it is possible to minimize the collection of continuous radiation emissions [19]. The gate width is the time duration which the spectrometers remains "opened" for signal integration; the accumulated pulses is the number of laser pulses/discharges released into the sample; and, finally, the interpulse delay corresponds to the time interval between the first laser pulse and the second laser pulse. The optimization of theses parameters is imperative to reduce background noise and improve signal intensities in the LIBS spectra [19].

1.2 Response Surface Methodology

DOE can be a powerful tool when assessing the effects of a group of factors over a response variable in experimental processes, because they are capable of providing the most information of a process from a reduced number of experiments [20]. In the scope of the DOE, the RSM is a statistical and mathematical procedure that aims to optimize and improve experimental processes through the application of two basic principles: the generation of models and displacement towards the optimal region. The methodology has come a long way since it was initially proposed by Box and Wilson, in 1951, and it has been used in a wide variety of setups since then, especially in the chemometrics' field [21]. Simply put, the idea of the RSM is to have linear or quadratic functions being fit into the experimental observations, from factorial designs, and assessing the goodness of this fit to that particular dataset [20-22]. Then, the displacement is followed accordingly to the results and towards the path of steepest ascent; or, in other words, towards the most desired response for the response variable [22]. These two steps can be repeated a number of times until the optimal region is reached and/or the model is considered satisfactory. The standard RSM procedure is: screening design, path of steepest ascent and three/multi-levels design [23], which are described in better detail ahead.

1.2.1 Screening Design

The factorial designs are a matrix of b^k experiments combinations, where b and k correspond to the levels and the number of factors respectively [24]. For example, in the simplest model possible, there could be a study about the influence of two factors in two different levels, or 2^2, that would result in 4 distinct experiments. Knowing which independent variables are the most relevant/significant for the experiment and establishing the range (levels) which these variables are going to be studied, are two very important steps prior to applying the RSM. They can be determined by a screening design, literature research or simply by experience of the researcher with the process [21].

The levels are codified in a dimensionless scale that represent their location in the experimental domain, with (+1) being the upper limit levels and (-1) the lower limit levels. These levels can be transformed back into the traditional scale by using the formula:

$$x_i = \frac{z_i - \overline{z_i}}{\Delta z_i} \tag{1}$$

Where x_i is the codified value of factor i; z_i is the original value of factor i; $\overline{z_i}$ is the real value in the central point; and Δz_i is the distance between the real value in the central point and the real value of one of the limits level.

The simplest model possible that can be used in the RSM (for a 2^k factorial design) is a linear model, by applying a linear regression on the experimental data. The model, in this case, is mathematically described as in:

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon \tag{2}$$

Where k is the number of factors; β_0 is the constant term of the model; β_i are the coefficients of the linear model; x_i are the factors; and ε is the residual term, or the error associated to that model.

A linear model can be sufficient for a screening design, even though it will not be able to provide enough information about curvature or second-order effects [21-22]. Nevertheless, the main objectives for a screening design

should be the evaluation/selection of variables, and establishing the direction in which better responses are attained in the experimental domain [23]. In this step, the Pareto's chart can be a very helpful tool for the researcher in both scenarios. The Pareto's chart displays graphically, in order of importance – or significance – the problems of the experimental system. In other words, it arranges the factors by their "urgency", or relevancy, to the experimental process, in the range that they were studied [22]. It also discriminates which variables were indeed significant for the experiment (for $p \leq 0,05$), and displays, quantitatively, values for each of these effects. For positive values, the chart suggests that an increase in the modular value of the factor could potentially increase, as well, the modular value of the response variable, and the opposite for the negative value effects [22].

1.2.2 Path of the Steepest Ascent (PSA)

After establishing the significant factors through the screening design, the resulting model will indicate a direction as to progress, in order to reach the optimal region. The next step in the RSM is called the path of steepest ascent (PSA), which can be roughly described as "walking" tangentially on the surface, or curvature, of the optimal region, following the direction of increasing responses, until you begin "descending" on the other side [21-23]. With this procedure, it is possible to establish, with a certain degree of confidence, the approximate location of the optimal region [23]. To determine the design for this stage of the RSM, we take the factor from the previous design with the highest modular coefficient (most significant) as reference for coordination, and calculate, proportionally, the values for the remaining factors (Eq. 3). Where b_i and b_j are the coefficients of the adjusted model for the reference factor and the calculated factor, respectively; Δx_i is the codified step of the reference value; and Δx_j is the value, proportional to the step, of the factor being calculated.

$$\Delta x_j = \frac{b_j}{b_i}\Delta x_i \tag{3}$$

1.2.3 Central Composite Design (CCD)

When the optimal region is determined by one of the experiments in the PSA, a new factorial design is performed around this local maximum, with a higher number of levels (>2). This way, there can be a higher number of degrees of freedom and more information about the curvature and second-order effects can be determined from the model [25], which is described mathematically in Eq. 4:

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{1 \leq i \leq j}^{k} \beta_{ij} x_i x_j + \varepsilon \tag{4}$$

There are a number of popular three or more levels factorial designs that fulfil this premise e.g. Box-Behnken design, Doehlert design and central composite designs [26]. For this study, the central composite design (CCD) will be the one assessed for the RSM. The CCD can be divided in three different portions, which are comprised of their own set of experiments: i) a full factorial or fractional factorial design; ii) a new axial, or star, design, where the experimental points are rotated into an α distance from its center; iii) central points repetitions.

As stated before, with this higher-level design, it is possible to estimate the curvature of the response surface, as well as the lack of fit for the model and the residual error through the central point repetitions [26].

The maximum/minimum point, or the critical point, of the model will correspond to the optimal condition of the study, and can be obtained by the first derivate of the quadratic function describing the model [21-22]. For example, for a CCD evaluating two factors, from the final model equation (Eq. 5), the critical point coordinate can be calculated by solving the first grade system formed by Eq. 6 and 7, and finding the values for x_1 and x_2.

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 \tag{5}$$

$$\frac{\partial y}{\partial x_1} = \beta_1 + 2\beta_{11} x_1 + 2\beta_{12} x_2 = 0 \tag{6}$$

$$\frac{\partial y}{\partial x_2} = \beta_2 + 2\beta_{22} x_2 + 2\beta_{12} x_1 = 0 \tag{7}$$

1.2.4 Analysis of Variance (ANOVA)

The ANOVA is a statistical tool that is often used to evaluate how well the final model is adjusted to the experimental data and if it is satisfactory describing the experimental domain [27]. Fundamentally, it is an algebraic breakdown (Eq. 8) of the sum of the variances of the model responses and the observed global media $(y_i - \bar{y})^2$ into two portions:

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y} - \bar{y})^2 + \sum (y_i - \hat{y})^2 \quad (8)$$

$$SS_T = SS_R + SS_r \quad (9)$$

The first portion is the variance contained in the regression model itself, which corresponds to the sum of the differences between the predicted responses from the model and the observed global media responses $(\hat{y}_i - \overline{y})$. The second portion is the residual variance, which corresponds to the sum of the differences between the observed responses and the predicted responses from the model $(y_i - \hat{y}_i)$. Eq. 8 can also be expressed in terms of SS_T, SS_R, SS_r (Eq. 9), or the "sums of the square" for each term. In a well-fitted model, SS_R should amount to a much larger value in comparison to that of SS_r. The SS_r can be further split into two other individual terms, corresponding to the pure error SS_{pe} (associated with the repetitions) and the lack of fit of the model SS_{lof} (Eq. 10).

$$SS_r = SS_{pe} + SS_{lof} \quad (10)$$

Each of these *SS* components can be also expressed in terms of "media of the square" (*MS),* when divided by their respective degrees of freedom.

With the information provided by the ANOVA table, it is possible to verify the significance of the regression by applying an F test (Fisher distribution) (Eq. 11). In this test, the calculated value of F must be higher than the critical value of F, taking into consideration the respective degrees of freedom from the regression and the residual respectively, for the model to be significant [28].

$$F_{p-1,n-p} \approx \frac{MS_R}{MS_r} \quad (11)$$

The F* test for the lack of fit can also be applied (Eq. 12). It implies that, if the model is well fitted to the experimental data, MS_{lof} should be equivalent to the random errors only Since MS_{pe} is also a parameter that is associated to the random errors, the F* test evaluates if there exists a significant difference between the two medias. Just as for the F test, the lack of fit test also compares the calculated value of F* to its critical counterpart. [29].

$$F^*_{m-p,n-m} \approx \frac{MS_{lof}}{MS_{pe}} \quad (11)$$

2. MATERIALS AND METHOD

2.1 Sample

The soil sample chosen for this study was collected from Pantanal's Region (Nhecolândia, Mato-Grosso do Sul) in Brazil, and went through standard sample preparation procedure, which included freeze-drying, crushing and grounding in a mortar, and filtering through a 100-mesh sieve for waste removal and particle homogenization. The homogenized soil sample was then used to produce pellets for LIBS analyses. Pellets were produced using about 500 mg of soil, mixed with 35% (w/w) of KBr to help pellet cohesion and agglutination. The mixture was then pressed under 5 tons for 30 seconds.

2.2 DP-LIBS Setup

The DP-LIBS spectra were obtained using a system manually assembled on an optical board in the Optics and Photonics Laboratory in Embrapa Instrumentation (São Carlos, São Paulo). The system is comprised of two Nd:YAG lasers operating at distinct wavelengths: one at 1064 nm, reaching the sample at a 90º angle, and the other at 532 nm, reaching the sample at a 45º angle. The system can be operated both as SP mode, for each of them, or as DP mode, in a crossed-beam geometry. The first laser is a Q-switched Ultra (Quantel) with 50 mJ maximum energy, width of 8 ns and repetition rate of 20 Hz and the second laser is a Q-switched Brilliant (Quantel) with 180 mJ of maximum energy, width of 4 ns and repetition rate of 10 Hz, coupled with a second harmonic generation module. A pulse delay generator (Quantum Composers, model 9618) is used for temporal control of both lasers in DP mode and for the data acquisition system. The detection is performed by an echelle spectrometer (Aryelle Butterfly 400, Lasertechnik Berlin) equipped with an intensified charge-coupled device (ICCD) with 1024 x 1024 pixels, that operates in two different spectral bands (175-330 and 275-750 nm, with a resolution of 13-24 and 29-80 pm, respectively). The emitted radiation is collected by

a fiber optic cable paired with two fused silica converging lenses, both with a diameter of 25,4 mm and with focal distances of 50 and 70 mm, respectively.

Samples are positioned inside an ablation-chamber onto a small support that can be moved by the X-Y axis with a precision of 1 µm, the chamber has 3 optical-access windows made of fused silica. The 1064 nm laser beam is directed into the sample by two mirrors (1064 nm) with a diameter of 37,8 and 25,4 mm, respectively, and focused on the sample by a bk-7 converging lens, with 100 mm focal distance, 25,4 mm diameter and anti-reflective coating. The 532 nm laser beam is directed into the sample by a rectangular prism and a mirror (532 nm) with a 25,4 mm diameter, and is focused in the chamber by another bk-7 converging lens, with the same 100 mm focal distance and 25,4 mm diameter, but without anti-reflective coating.

2.3 Emission Lines And Spectra Analysis

In this study, several elements and their respective emission lines were evaluated (Table 1).

Table **1.** Spectroscopic atomic emission lines evaluated.

Element	Wavelenght (nm)
Zn I	213.85
Cd I	228.80
Fe II	259.94
Cu I	327.39
Ni I	352.45
Ca II	393.38
Al I	396.15
Sr II	421.55
Mg I	518.36
Cr I	520.60

Elements were chosen with the objective of finding an optimal system configuration that would benefit all the various emission lines at the same time. They were carefully picked as to present minimal adjacent spectra interferences, high accuracy (C+ or above) and high transition probability (Einstein's coefficient of 10^7 or above) (Table 1). The relative intensity of each emission line was also taken into consideration because emission lines can be easily saturated by macroelements such as Fe, Ca and K. On the other hand, microelements such as Zn, Ni, Cr, Cu, depending on their concentration in the sample, can present the opposite behavior. Therefore, as a manner to avoid these potential problems, less intense emission lines were chosen for the more abundant elements, and more intense lines were chosen for the less abundant elements. All the information was assessed in the NIST atomic spectra database [30].

The LIBS spectra was processed using LIBS Spectra Analyzer (LIBSSA) software, developed in the Python programming language. The software allows users to open and view LIBS spectra, exclude outliers (by the scalar product between the average and each measured spectrum), perform baseline correction and isolate peaks for area and height calculations [31].

According to previous studies (11), around 40–45 spectra per sample are required for a satisfactory reproducibility in LIBS analyses, without compromising too much the rapidness of the technique. Therefore, in this study, 40 spectra were collected per sample and all the measurements were made in triplicates in order to improve the precision of the results and reduce furthermore some of the matrix effects.

2.4 Experimental Design

Four factors were chosen as independent variables to be optimized in this study: (i) delay-time, (ii) interpulse delay, (iii) gate width (integration time) and (iv) accumulated pulses. The energy of both lasers pulses were set at 60 mJ since previous studies [11] showed that it had little influence on the overall response and sample variance. As for the response, or dependent variable, signal-to-noise ratios (SNR) were evaluated. They were calculated for each element (Table 1) by dividing the maximum peak height ($I_{máx}$), or intensity, of each emission line by the standard deviation of the noise (σ_{noise}) that is adjacent to that respective peak (Eq. 12).

Since one of the goals in this study was to propose an optimization procedure that would work for all the elements at once, and simultaneously in a single step, the proposed response variable was a multiplication of all the SNR of each emission line (Eq. 13), where *n* is the number of emission lines evaluated. By multiplying all the SNR,

we take into account the effects of each individual emission line and guarantee that the optimization is adequate for each of the elements, in a sort of “common-ground” scenario [22].

$$SNR = \frac{I_{máx}}{\sigma_{noise}} \tag{12}$$

$$y = \prod_{i}^{n} SNR_i \tag{13}$$

The levels, or limits, of each factors were asserted accordingly to previous studies [11], and the initial screening factorial design was comprised of four factors variating on two specified upper and lower limits (Table 2). In other words, the first factorial design was a full 2^4 factorial design with triplicates, with 48 experiments in total.

Table **2**. Screening factorial design matrix.

Factor	Limits	
	(-1)	(+1)
Delay-time (ns)	200	1500
Interpulse delay (ns)	500	3000
Integration time (µs)	5	20
Accumulated pulses	3	7

The remaining factorial designs are explained furthermore along the article, as we progress through the necessary steps of the RSM. All the statistical DOE analyses, as well as the figures and graphs, were processed and generated by STATISTICA12® software.

3. RESULTS AND DISCUSSION

3.1 Screening Factorial Design

The results of the initial factorial design showed that only three factors (in order of magnitude: gate width, accumulated pulses and delay-time) were significant ($p \leq 0{,}05$) in the range that they were studied, as can be observed in Pareto’s chart (Figure 1). Since interpulse delay was not a significant factor in the range of the first design study, it was removed from further experimental designs and its value was fixed at the lowest value of 500 ns. Because their effects were positive, an increase in the module of these variables would also ensue an overall increase in the results (SNR) attained. In addition, the initial linear response surface models (Figure 2) corroborates with this idea. Clearly, for all the models (which the other two factors are fixed in its center value), we can observe that the upper limit (+1) of the experimental design were the ones with the best overall response.

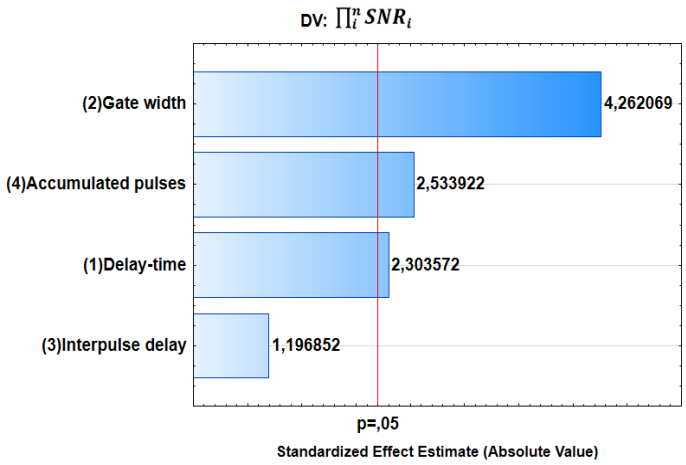

Figure **1.** Pareto’s chart for the screening design.

The error of this first part of the experiment was roughly estimated by the relative standard deviation (RSD) from the response variable triplicates. The total error calculated for this screening design was of RSD = 14,01%, which is acceptable considering LIBS analyses are prone to plasma and signal fluctuations, matrix effects, heterogeneity of the pellets, etc. The function of the linear model for this part of the RSM is shown in Eq. 14:

$$\hat{y} = 8{,}02E9 + 6{,}48E9x_1 + 1{,}19E10x_2 + 3{,}36E9x_3 + 7{,}13E9x_4 \quad (14)$$

3.2 Path of Steepest Ascent (PSA)

Following the steps of the RSM theory, and accordingly to the method description in subsection 1.3, a new design matrix was created with increasingly proportional values for each independent variable, using the most significant factor (gate width) from the previous design as reference (Table 3).

Table **3.** PSA design matrix for the 3 significant factors.

Step	Δx_1 (gate width)	Δx_2 (delay-time)	Δx_3 (accumulated pulses)	Gate width (μs)	Delay-time (ns)	Accumulated Pulses
Center	0	0	0	12,5	850	5
Center + Δ	1	0,54	0,59	20	1200	6
Center + 2Δ	2	1,08	1,19	28	1550	7
Center + 3Δ	3	1,62	1,78	35	1900	9
Center + 4Δ	4	2,16	2,38	43	2250	10
Center + 5Δ	5	2,70	2,97	50	2600	10

By plotting the results of this matrix, it becomes visually easier to grasp which step had the best results among the others (Figure 3). In this case, step 3 (Center + 2Δ) was the one with the steepest ascension and the subsequent

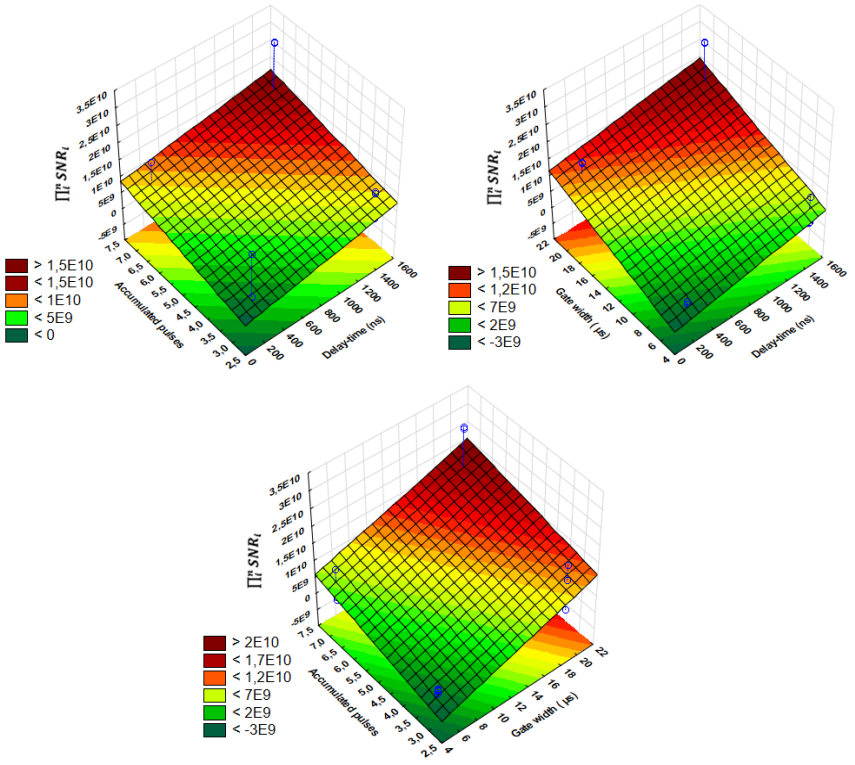

Figure **2.** Response surfaces of the screening design, for the significant factors. Two factors are plotted at a time, and the third factors is set in its central point.

steps/experiments resulted in an overall declining of the SNR. It is safe to pinpoint then, that step 3 is located very closely to the optimal region.

The decline of the SNR observed throughout the third step ahead is due to the correlation between the temporal characteristics of plasma formation, the emission lines of each element and the parameters controlling the LIBS system. As explained earlier, in section 1.2, some emission lines are influenced differently depending on the value of these factors. However, in general, high values of delay-time mean that the spectra are the product of emissions close to a cooler plasma, as it is cooling until extinction. Obviously, in these extreme conditions, the emission collected diminishes notably for most elements, as the plasma grows cooler through time. The accumulated pulses can also disrupt the quality of the signal, even though the higher number of pulses could assist, in theory, the ablation of a bigger portion of the sample, and thus improving the electronic density of the plasma, and providing more energy/heat. The counterpart, however, is that it can also increase matrix effects, dispersion and even cause unwanted auto-absorption effects, which affects the signal being obtained. Therefore, it is important to find a perfect balance for these parameters, as to maximize the overall SNR.

3.3 Central Composite Design (CCD)

After determing that step 3 (28 µs gate width, 1550 ns delay-time, 7 accumulated pulses) was the condition with the best S/N ratio results, a new design – this time with 3 levels and a central point – was performed (Table 4).

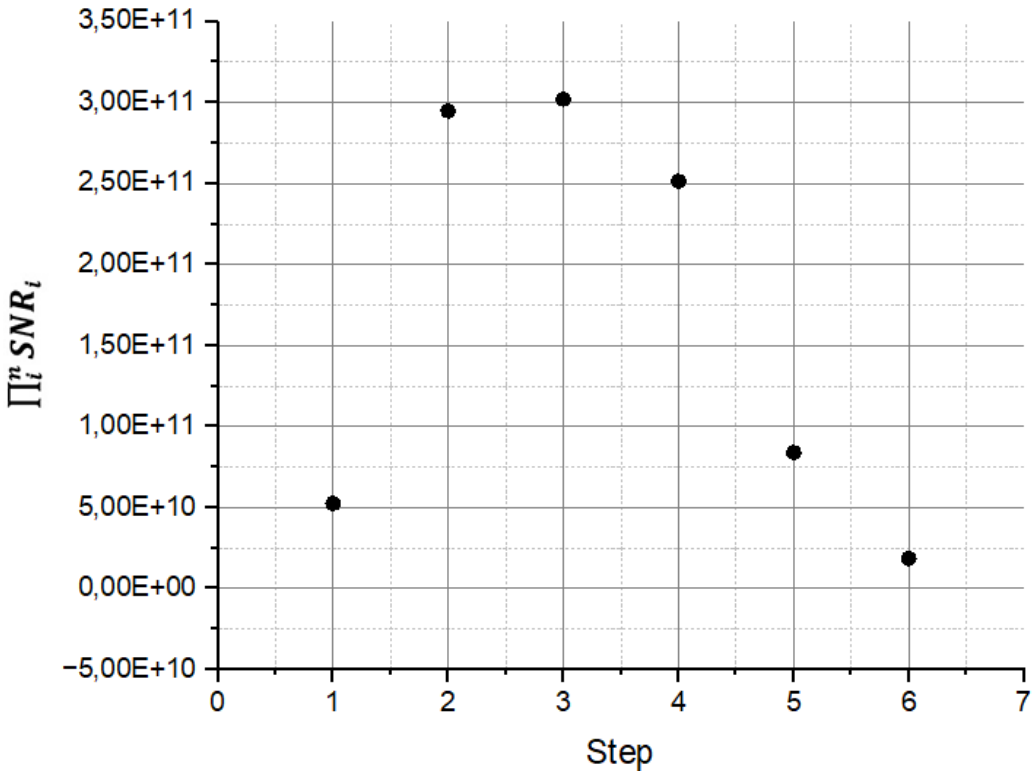

Table **4.** CCD matrix with lower, upper and central points. The (-1,682) and (+1,682) limits correspond to the axial design rotational angle α for 3 factors, calculated by $(2^k)^{0,25}$, with k being the number of factors.

Factors	Limits				
	Lower		Center	Upper	
	(-1,682)	(-1)	(0)	(+1)	(+1,682)
Gate width (μs)	15	20	28	35	40
Delay-time (ns)	450	900	1550	2200	2650
Accumulated Pulses	4	5	7	9	10

In this stage, 20 different experiments were carried out: 8 being the factorial points (-1, +1), 6 being the axial points (-1,682, +1,682) and 6 central point (0) repetitions. Since all the upper and lower limits experiments were triplicates, a total of 48 experiments were performed.

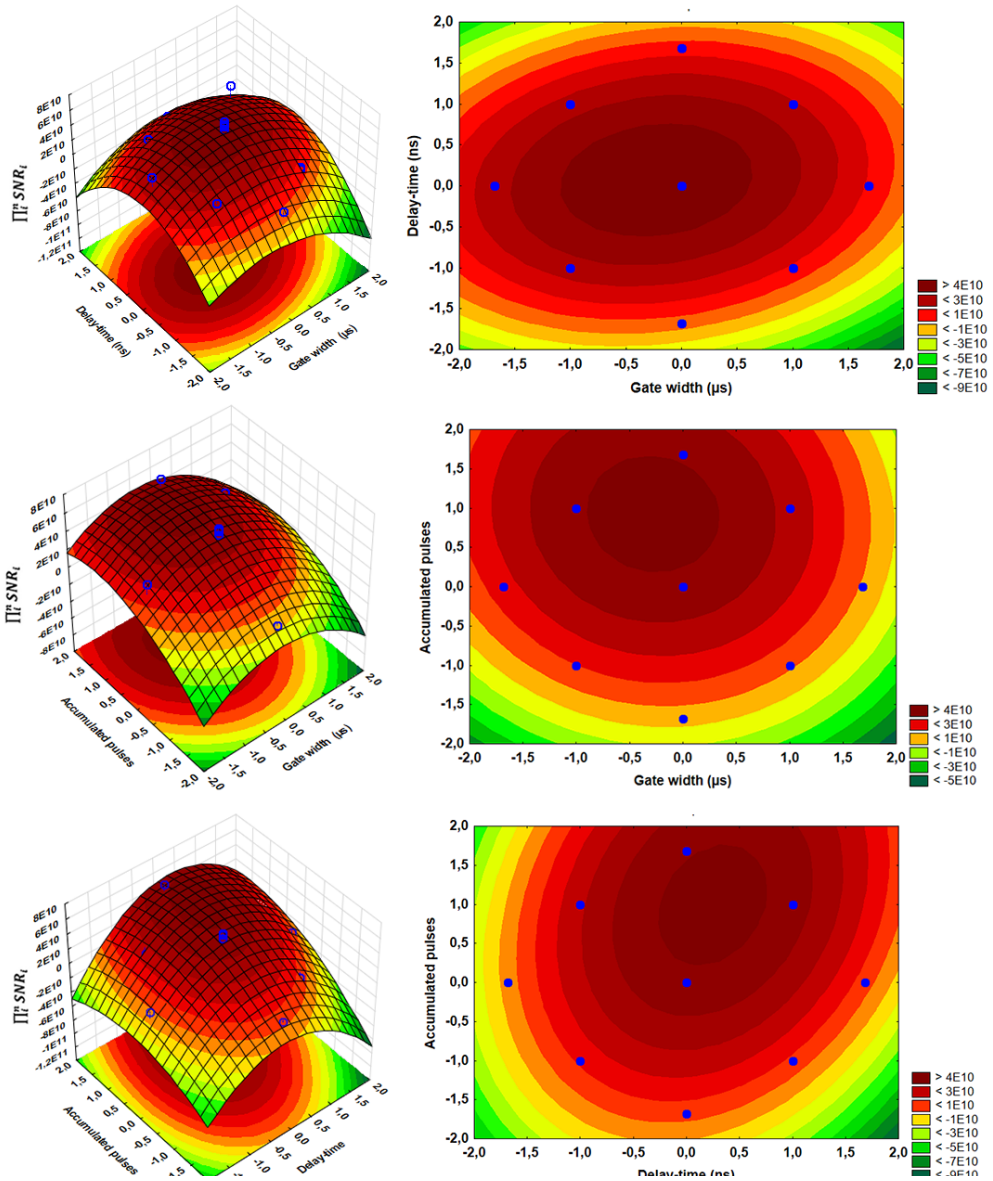

Figure **4.** Response surfaces and contour surfaces of the CCD design. Factors are plotted two at a time while the third factor is set at its central point.

The 3-D shaped response and the contour surfaces (Figure 4) for each pair of factors are the outcome of the CCD design results from the final model attained (Eq. 13), where x_1, x_2, x_3 are the corresponding factors (gate width, delay-time and accumulated pulses, respectively).

$$\hat{y} = 4{,}73E10 - 1{,}00E10x_1 - 2{,}11E10x_1^2 + 6{,}21E9x_2 - 3{,}36E10x_2^2 + 2{,}68E10x_3 - 1{,}48E10{x_3}^2 + 7{,}16E9x_1x_2 - 2{,}03E9x_1x_3 + 1{,}05E10x_2x_3 \quad (13)$$

The dark red region in the models (Figure 4), which is reminiscent of a "plateau", is our optimal region, where experiments will yield better responses. Noticeably, our CCD matrix – and the limits we proposed for each factor – were very near the optimal region, as we can observe the central point of the matrix being next to the maximum curvature of the model. To help mathematically evaluate the quality of the final model, an analysis of variance (ANOVA) was performed (Table 5).

Table **5.** ANOVA table of the model. SS = Sum of Squares, df = degrees of freedom, MS = Mean Square.

Source	SS	df	MS
Regression (R)	8,82E+21	9	9,80E+20
Residual (r)	9,74E+20	10	9,74E+19
Lack of fit	6,67E+20	5	1,33E+20
Pure error	3,07E+20	5	6,14E+19
Total	9,79E+21	19	
R^2	0,90		
R^2_{max}	0,97		

Using the information provided by the ANOVA table, the calculated F for the model resulted in: MS_R/MS_r = 10,06 which is higher than the critical $F_{(9,10)}$ = 2,35. This result indicates that the proposed model is significant, and that there exists a correlation between the regression variables and our response variable. The F* test for the lack of fit can also be performed, in which $F^* = MS_{LoF}/MS_{PE}$ = 2,17. When compared to the critical $F^*_{(5,5)}$ = 3,45 it is safe to assume that there is no lack of fit in the final model and the observed model is capable of explaining most of the observed results/variance. The coefficient of determination (R^2 = 0,90) can also be deemed satisfactory, even though its value seems underwhelming at first. However, as specified in section 1.2, LIBS analysis is very susceptible to fluctuations, matrix effects and other errors, which affects reproducibility. This fact becomes more apparent throughout the steps of the RSM in this study, since the RSD from the experiments were in the range of 10-15%. Considering the RSM was performed using a soil sample, which has a very complex matrix, and that many elemental lines were assessed at the same time, the relative low value of R^2 is still inside the range of the variance being observed.

Finally, after evaluating the final model, the critical values for each variable (the maximum point in the response surface) were estimated (Table 6) and an experiment was performed in the optimal conditions (Table 7). The result shows that, after optimizing LIBS parameters, the group of emission lines studied had 482 times better SNR (overall, not individually) than when comparing to the non-optimized system from the start of the study.

Table **6.** Best stipulated values for each variable, by the CCD experiment.

Factors	Critical values
Gate width (µs)	26
Delay-time (ns)	1600
Accumulated pulses	9
Interpulse delay (ns)	500

Table **7.** Comparison between the S/N ratios before and after system optimization with RSM.

	$\prod_i^n SNR_i$	Ratio
Optimized condition	1,24E+11	48,2
Not-optimized condition	2,57E+09	

CONCLUSION

In this work, we were able to show that the use of design of experiments, and the response surface methodology, could be a powerful tool when trying to simultaneously optimize a group of close-related parameters, such as the ones presented in the LIBS system. Although the interpulse-delay was not significant in the data range evaluated, the remaining three parameters (delay-time, accumulated pulses, gate width) were constantly refined after every stage of the RSM. In the end, the best overall setting reached was of 500 ns interpulse-delay, 1600 ns delay-time, 9 accumulated pulses and 26 µs gated width. The final model was significant, did not present any lack of fit (as evaluated by both F and F* tests) and the coefficient of determination obtained (R^2 = 0,90) was acceptable and inside of the expected LIBS error and imprecision. An experiment was performed at this optimal condition and the result showed an increase of 48,2 times in the S/N ratios for all the assessed emission lines (for the group of elements). This heavily suggests that the methodology applied was, indeed, successful in establishing the best overall condition for analyses, while performing the least number of experiments as possible.

REFERENCES

[1] Shah SKH, Iqbal J, Ahmad, P, *et al*. Laser induced breakdown spectroscopy methods and applications: A comprehensive review. Radiat Phys Chem 2020, 170.

[2] Cremers DA, Radziemski LJ. Handbook of Laser-Induced Breakdown Spectroscopy. New York City, New York: John Wiley & Sons, 2013.

[3] Hussain T, Gondal MA. Laser induced breakdown spectroscopy (LIBS) as a rapid tool for material analysis. J Phys Conf 2013, 439.

[4] Guo LB, Zhang D, Sun LX, *et al*. Development in the application of laser-induced breakdown spectroscopy in recent years: A review. Front Phys 2021, 16(2).

[5] Elnasharty IY, Doucett FR, Gravel JFY, Bouchard P, Sabsabi M. Double-pulse LIBS combining short and long nanosecond pulses in the microjoule range. J Anal At Spectrom 2014, 29(9): 660-6.

[6] Jantzi SC, Motto-Ros V, Trichard F, *et al*. Sample treatment and preparation for laser-induced breakdown spectroscopy. Spectrochim Acta B At Spectrosc 2016, 115.

[7] Lepore KH, Fasset CI, Breves EA, *et al*. Darby. Matrix Effects in Quantitative Analysis of Laser-Induced Breakdown Spectroscopy (LIBS) of Rock Powders Doped with Cr, Mn, Ni, Zn, and Co. Appl Spectrosc 2017, 71(4): 600-26.

[8] Segnini A, Xavier AAP, Otaviani-Junior PL, *et al*. Physical and Chemical Matrix Effects in Soil Carbon Quantification Using Laser-Induced Breakdown Spectroscopy. Am J Anal Chem 2014, 5(11): 722-9.

[9] Sun L, Yu H. Correction of self-absorption effect in calibration-free laser-induced breakdown spectroscopy by an internal reference method. Talanta 2009, 79(2): 388–95.

[10] Haider Z, Munajat Y, Kamarulzaman R. Review: Laser-induced breakdown spectroscopy (LIBS) a promising technique, its limitations and a proposed method. J Teknol (Sci Eng) 2012, 57.

[11] Morais CP, Nicolodelli G, Mitsuyuki MC., Mounier S, Milori DMBP. Optimization of laser-induced breakdown spectroscopy parameters from the design of experiments for multi-element qualitative analysis in river sediment. Spectrochim Acta B Atom Spectrosc 2021, 177.

[12] Junior DS, Tarelho LVG, Krug FJ, Neto LM, Milori DMBP. Espectrometria de emissão ópica com plasma induzido por laser (LIBS) – Fundamentos, aplicações e perspectivas. Analytica 2006, 24: 71-81**.**

[13] Pasquini C, Cortez J, Silva LMC, Gonzaga FB. Laser Induced Breakdown Spectroscopy. J Braz Chem Soc 2007, 18(3): 463-512.

[14] Ranulfi AC. LIBS como ferramenta diagnóstica em plantas: um estudo nutricional de folhas de soja em busca pelos efeitos da infestação por Aphelenchoides besseyi. PhD Thesis, University of São Paulo: São Carlos, February, 2019.

[15] Costa VC, Augusto AS, Castro JP, *et al*. Laser-induced breakdown spectroscopy (LIBS): History, fundamentals, applications and potentialities. Quim Nova 2019, 42(5): 527–45.

[16] Moll M, Schlanges M, Bornath TH, Krainov VP. Inverse bremsstrahlung heating beyond the first Born approximation for dense plasmas in laser fields. New J Phys 2012, 114(6).

[17] Borduchi LCL, Milori DMBP, Villas-Boas PR. Study of the effects of detection times in laser-induced breakdown spectroscopy and missed variation of plasma parameters with gate width. Spectrochimi Acta B Atomic Spectroscopy 2022, 191.

[18] Nicolodelli G, Senesi GS, Romano RA, Perazzoli ILO, Milori DMBP. Signal enhancement in collinear double-pulse laser-induced breakdown spectroscopy applied to different soils. Spectrochim Acta B Atom Spectrosc 2015, 111: 23–9.

[19] Markiewicz-Keszycka M, Cama-Moncunill X, Casado-Gavalda MP, *et al*. Laser-induced breakdown spectroscopy (LIBS) for food analysis: A review. Trends Food Sci Technol 2017, 65: 80–93.

[20] Ghelich R, Jahannama MR, Abdizadeh H, Torknik FS, Vaezi MR. Central composite design (CCD)-Response surface methodology (RSM) of effective electrospinning parameters on PVP-B-Hf hybrid nanofibrous composites for synthesis of HfB 2 -based composite nanofibers. Compos B Eng 2019, 166: 527–41.

[21] Bezerra MA, Santelli RE, Oliveira EP, Villar LS. and ESCALEIRA, Luciane A. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. Talanta 2008, 76(5): 965–77.

[22] Barros Neto B, Scarminio IS, Bruns RE. Como fazer experimentos: aplicações na ciência e na indústria. 3ª ed. Campinas, São Paulo: Ed. UNICAMP, 2007.

[23] Lee DH, Kim SH, Byun JH. A method of steepest ascent for multiresponse surface optimization using a desirability function method. Qual Reliab Eng Int 2020, 36(6): 1931–48.

[24] Breig SJM, Luti KJK. Response surface methodology: A review on its applications and challenges in microbial cultures. Mat Today Proc 2021, 42: 2277–84.

[25] Said KAM, Amin MAM. Overview on the Response Surface Methodology (RSM) in Extraction Processes. J Appl Sci Process Eng 2015, 2(1).

[26] Dejaegher B, Heyden YV. Experimental designs and their recent advances in set-up, data interpretation, and analytical applications. J Pharm Biomed Anal 2011, 56(2): 141–58.

[27] Stahle L, Wold S. Analysis of Variance (ANOVA). Chemometr Intell Lab Syst 1989, 6: 259-72.

[28] Dostanić J, Lončarević D, Rožić L, *et al*. Photocatalytic degradation of azo pyridone dye: Optimization using response surface methodology. Desal Water Treat 2013, 51: 2802–12.

[29] Wellek S. Testing for goodness rather than lack of fit of continuous probability distributions. PLoS ONE 2021, 16(9).

[30] National Institute of Standards and Technology (NIST). NIST Atomic Spectra Database Lines Form [online]. Avaiable at: http://physics.nist.gov/PhysRefData/ASD/lines_form.html. Acessed at: 08/10/2022.

[31] Stenio K, Villas-Boas PR, Milori DMBP. Desenvolvimento de software para análise de espectros LIBS. In: 6th Biotechnology Simposium of UFSCar (VI SBU), 2018, São Carlos.