Spectrochimica Acta Part B 64 (2009) 565–572

Contents lists available at ScienceDirect

Spectrochimica Acta Part B

journal homepage: www.elsevier.com/locate/sab

Simultaneous optimization by neuro-genetic approach for analysis of plant materials by laser induced breakdown spectroscopy ☆

Lidiane Cristina Nunes [a,b], Gilmare Antônia da Silva [c], Lilian Cristina Trevizan [b], Dario Santos Júnior [d], Ronei Jesus Poppi [e], Francisco José Krug [b,*]

[a] Departamento de Química, Universidade Federal de São Carlos, Rodovia Washington Luís, km 235, 13565-905, São Carlos, SP, Brazil
[b] Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Av. Centenário 303, 13416-000, Piracicaba, SP, Brazil
[c] Departamento de Quimica, Universidade Federal de Ouro Preto, Morro do Cruzeiro S/N, 35400-000, Ouro Preto, MG, Brazil
[d] Universidade Federal de São Paulo—UNIFESP, Rua Prof. Artur Riedel 275, 09972-270, Diadema, SP, Brazil
[e] Instituto de Química, Universidade Estadual de Campinas, Caixa Postal 6154, 13084-971, Campinas, SP, Brazil

ARTICLE INFO

Article history:
Received 31 October 2008
Accepted 6 May 2009
Available online 13 May 2009

Keywords:
Laser induced breakdown spectroscopy (LIBS)
Plant analysis
Bayesian regularized neural network
Genetic algorithm
Simultaneous optimization

ABSTRACT

A simultaneous optimization strategy based on a neuro-genetic approach is proposed for selection of laser induced breakdown spectroscopy operational conditions for the simultaneous determination of macro-nutrients (Ca, Mg and P), micro-nutrients (B, Cu, Fe, Mn and Zn), Al and Si in plant samples. A laser induced breakdown spectroscopy system equipped with a 10 Hz Q-switched Nd:YAG laser (12 ns, 532 nm, 140 mJ) and an Echelle spectrometer with intensified coupled-charge device was used. Integration time gate, delay time, amplification gain and number of pulses were optimized. Pellets of spinach leaves (NIST 1570a) were employed as laboratory samples. In order to find a model that could correlate laser induced breakdown spectroscopy operational conditions with compromised high peak areas of all elements simultaneously, a Bayesian Regularized Artificial Neural Network approach was employed. Subsequently, a genetic algorithm was applied to find optimal conditions for the neural network model, in an approach called neuro-genetic. A single laser induced breakdown spectroscopy working condition that maximizes peak areas of all elements simultaneously, was obtained with the following optimized parameters: 9.0 μs integration time gate, 1.1 μs delay time, 225 (a.u.) amplification gain and 30 accumulated laser pulses. The proposed approach is a useful and a suitable tool for the optimization process of such a complex analytical problem.

1. Introduction

The determination of macro- and micro-nutrients in plant materials is frequently recommended for evaluation of the nutritional status of crops of economic interest. Essential elements play a decisive role in plant nutrition and can affect crop yields when they are not present in appropriate concentrations.

Laser induced breakdown spectroscopy (LIBS) is a useful technique for analysis of solid samples of environmental and geological interest. One of the major advantages of LIBS is that only minimal or no sample preparation is required. Rapid multielemental analysis can be carried out in situ; qualitative and quantitative determination of elements can be carried out in various test environments [1,2].

There is a lack of information for LIBS analysis of agricultural samples. Sun et al. [1] evaluated LIBS for the determination of some micro-nutrients (Fe, Cu, Mn, and Zn) as well as macro-nutrients (P, Ca, and Mg) in powdered leaf samples applied directly on a double-sided tape on a glass slide. Most recently a LIBS system furnished with a Nd:YAG at 1064 nm was evaluated in our laboratory for the determination of macro and micro-nutrients in plant materials [3,4] using pellets of certified reference materials (CRM) for calibration. Results of the analysis of different plant samples carried out by LIBS compared reasonably with those obtained by inductively coupled plasma optical emission spectrometry (ICP OES) after wet acid digestion. An fs-LIBS was also evaluated for the determination of Ca, Cu, Fe, K, Mg, Na and P in animal tissues [5]; however, the optimal experimental parameters for LIBS implementation have not yet been found, and consequently, there is a large need for methodical experimental optimization.

Some variables that can influence LIBS measurements are laser related (i.e. wavelength, energy, pulse duration and shot-to-shot power fluctuation), focusing spot size, ambient conditions, physical properties of the sample, amplification detector gain, ambient atmosphere, pressure and the detection window (delay time, integration time gate) [6]. In fact, the analytical performance of LIBS methods depends strongly on the choice of experimental conditions, which have to be studied in each particular application. Optimization of LIBS

☆ This paper was presented at the 10th Rio Symposium on Atomic Spectrometry, held in Salvador-Bahia, Brazil, 7–12 September 2008, and is published in the special issue of Spectrochimica Acta Part B, dedicated to that conference.

* Corresponding author. Fax: +55 19 34294610.
E-mail address: fjkrug@cena.usp.br (F.J. Krug).

0584-8547/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sab.2009.05.002

methods is habitually performed by univariated procedures [7,8–13] and little has been reported on multivariate procedures [14–16]. Conventional optimization through univariate methodology requires a great amount of experiments and is time consuming; moreover, interactions between variables are not evaluated. Another aspect is the fact that applying univariate procedures, it will be rarely possible to reach really optimized values. Multivariate techniques are alternative, economical and effective methods that allow the simultaneous optimization of more than one variable.

These multivariate statistical methods most used in chemistry can be conveniently classified according to how one decides which experiments are to be executed. All methods require the user to supply minimum and maximum values for each factor that defines the experimental domain to be investigated during the optimization procedure. Optimization procedures are frequently performed by experimental designs. The most commonly used designs to determine response surfaces are the full and fractional factorial and the more complex central composite [17], Box–Behnken [18], Doehlert [19,20] and mixture designs [21].

Doehlert design was applied in the optimization of the factors affecting LIBS performance for analysis of samples on planet Mars [14]. Experimental designs arranged according to Doehlert matrixes are useful to generate the necessary data, demanding fewer experiments and the quantity of the levels related to each factor can be selected in order to obtain more information about significant or problematic factors [22,23]. If several responses are being investigated, other approaches are extremely useful, especially those that allow the simultaneous modeling. In multi-response procedures, artificial neural network (ANN) approaches are very interesting alternatives because they provide one model for all responses, macro- and micronutrients and Al and Si, and that this model cannot be written in closed algebraic form. The main chemical applications of ANN, the theory and the type of learning scheme for their use have been reviewed by Zupan and Gasteiger [24].

2. Multivariate approaches

Artificial neural network is a computational modeling tool that consists of groups of highly interconnected processing elements called neurons. The neurons are arranged in one input layer (independent variables), one output layer (dependent variables), and several hidden layers that associate the inputs with outputs. A neuron collects a series of input signals and transforms them into the output signal via a transfer function [25].

During the network general training process, weights (w) and bias (b) for each connection output value are obtained as close as possible to the actual outputs, minimizing the error function given by Eq. (1):

$$E = \sum_{m=1}^{p} \sum_{k=1}^{n} \left(y_{m,k} - t_{m,k}\right)^2 \tag{1}$$

where t is the target value; y is the predicted value (net output); n is the number of experiments in training set and p is the number of dependent variables or number of responses to be optimized. The correction is continued until the error function converges to a minimum. Improvement of generalization is generally obtained by the use of large data sets, where all possible data examples are presented to the network; however, in some situations, the acquisition of large data sets is difficult or even impossible [26,27].

A method for improving generalization without using large data sets is the regularization. In this method, a modification of the error function (Eq. (1)) is accomplished by adding a term that consists of the mean of the sum of squares of the network weights and biases (Eq. (2)):

$$F_{\text{reg}} = \gamma E + (1 - \gamma) msw \tag{2}$$

where γ is the performance ratio and msw is given by the relation expressed in Eq. (3):

$$msw = \frac{1}{n} \sum_{j=1}^{n} w_j^2 \tag{3}$$

Weights and bias can be assumed as random variables with specified distributions. In this case, Bayesian regularized artificial neural network (BRANN) is used. This model considers all possible values of network parameters weighted by the probability of each set of weights. Bayesian methods are complementary to neural networks as they overcome the tendency of an over-flexible network to discover nonexistent, or overly complex, data models. Other drawbacks include long training times, possibility of over-trained sets and loss of an explicit function to model the system can be cincumvented [26–28].

ANN have been used for response surface modelling [29,30] and they provided several benefits as compared with conventional optimization techniques. They do not require the determination of a model expressible in a close mathematical form, and they can result in better generalisations, as they are able to model complex relationships. However the neural networks can only model the system and another algorithm is necessary to optimize the established model. Genetic algorithm (GA) can be an interesting way in this task. General information on genetic algorithms can be found in reference [31].

The GA routine utilized in this work demands a function that will be optimized by the algorithm. It was found that the loss-minimization function [32] (Eq. (4)), calculated from the sum of the weighted relative deviations, is the most suitable alternative:

$$\varphi = \sum_{i=1}^{m} w_i \left[\frac{Q(x_1, x_2, \cdots, x_n) - Q^*}{Q^*}\right]^2 \tag{4}$$

where φ is the loss-minimization function; w_i is the weight factor with $i = 1,2,\ldots,m$ and m is the number of objectives, or in this case, the number of elements. The term $Q(x_1,x_2,\ldots,x_n)$ corresponds to the objective function and n indicates the number of independent variables. The objective function is the output values of the trained BRANN. Q^* corresponds to the individual optimum value of the objective function, defined as the maximum peak area possible for each element studied. In this case, the GA algorithm will try to minimize the loss function φ.

GA can define the ideal conditions (maximum or minimum) from the model developed by the neural network, generating a hybrid approach named neuro-genetic [26]. In analytical chemistry, the neuro-genetic approach has been successfully applied for optimization of operational conditions in analytical chromatographic procedures [27]. In LIBS neural network have been applied to analysis of metals and painted metals [33] and of soils [34]. However, the aim of these works was for LIBS classification purposes. To our best knowledge, there is no work in the literature dealing with the use of ANN for the optimization of LIBS operation conditions.

According to all the facts described, the purpose of this work was to develop a strategy based on neuro-genetic approach to simultaneously optimize LIBS operational conditions for the determination of macro-nutrients (Ca, Mg and P), micro-nutrients (B, Cu, Fe, Mn and Zn), Al and Si in plant materials by using a commercially available LIBS system equipped with a Nd:YAG at 532 nm. Effects of integration time gate, delay time, amplification gain and number of accumulated laser pulses were evaluated through Doehlert design for definition of significant variables. A model considering all peak areas simultaneously was determined by neural network and the genetic algorithm was used to find the LIBS operation conditions where all peaks express compromised maximum area values.