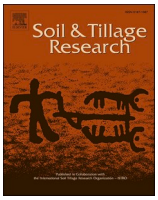
Soil & Tillage Research 216 (2022) 105250

Contents lists available at ScienceDirect

Soil & Tillage Research

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



Laser-Induced Breakdown Spectroscopy (LIBS) for tropical soil fertility analysis

Tiago R. Tavares [a,b,*], Abdul M. Mouazen [c], Lidiane C. Nunes [b], Felipe R. dos Santos [d], Fábio L. Melquiades [d], Thainara R. da Silva [e], Francisco J. Krug [b], José P. Molin [a]

[a] Laboratory of Precision Agriculture (LAP), Department of Biosystems Engineering, "Luiz de Queiroz" College of Agriculture (ESALQ), University of São Paulo (USP), 13418900 Piracicaba, São Paulo, Brazil
[b] Center for Nuclear Energy in Agriculture (CENA), University of São Paulo (USP), Piracicaba, São Paulo 13416000, Brazil
[c] Precision Soil and Crop Engineering Group (Precision SCoRing), Department of Environment, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, Blok B, 1st Floor, 9000 Gent, Belgium
[d] Laboratory of Applied Nuclear Physics (LFNA), Department of Physics, Londrina State University (UEL), 86057970 Londrina, Paraná, Brazil
[e] Laboratory of Agricultural Machinery and Precision Agriculture (LAMAP), Department of Biosystems Engineering, Faculty of Animal Science and Food Engineering (FZEA), University of São Paulo (USP), 13635900 Pirassununga, São Paulo, Brazil

ARTICLE INFO

Keywords:
Soil fertility testing
Proximal soil sensing
Hybrid laboratories
Precision agriculture
Matrix effect

ABSTRACT

The Laser Induced Breakdown Spectroscopy (LIBS) is a promising technique for soil fertility analysis in a rapid and environmentally friendly way. This application requires the selection of an optimal modelling procedure capable of handling the high spectral resolution of LIBS. This work aimed at comparing different modelling methods of LIBS data for the determination of key fertility attributes in Brazilian tropical soils. A benchtop LIBS system was used for the analysis of 102 soil samples, prepared in the form of pressed pellets. Models for the prediction of clay, organic matter, pH, cation exchange capacity, base saturation, and the extractable nutrients P, K, Ca, and Mg were developed using univariate linear regression (ULR), multiple linear regression (MLR) and partial least squares regression (PLS). The following input data for PLS were used: (i) the full spectra from 200 to 540 nm (38,880 variables), and (ii) variables selected by the interval successive projections algorithm (iSPA). The multivariate models achieved satisfactory predictions [residual prediction deviation (RPD) > 1.40] for eight out of nine fertility attributes. However, the best performances were obtained for the PLS with the variable ranges selected by the iSPA, which achieved satisfactory predictions (RPD ≥ 1.44) for seven out of the nine soil attributes studied. The MLR method obtained lower prediction performance than the iSPA-PLS using only 21 variables. The iSPA-PLS approach allowed a reduction from 3 to 160-fold in the total of variables compared to the full LIBS spectra, making it efficient and accurate modelling method that uses reduced number of variables. Although LIBS technique proved to be efficient for predicting fertility attributes in tropical soils, further research is encouraged in order to reduce the amount of sample preparation conducted in this study.

1. Introduction

Worldwide, sensor-based analytical methods have been seen to characterize key soil fertility attributes in a rapid way without producing chemical residues (Viscarra Rossel and Bouma, 2016; Molin and Tavares, 2019). Proximal soil sensing (PSS) methods allows to reduce the number of operations con-ducted in traditional analytical procedures (e.g. digestion and extraction), making the analysis faster, cost-effective and environmentally friendly. Different PSS techniques have been applied for determining soil fertility attributes, e.g., X-ray fluorescence (XRF), mid infrared and visible and near infrared (vis-NIR) diffuse reflectance spectroscopies, ion selective electrodes (ISE), are few to mention among others (Silva and Molin, 2018; Munnaf et al., 2019; Demattê et al., 2019; Tavares et al., 2020a). LIBS spectra allow a broad characterization of soil elementary constitution (e.g., N, P, K, Ca, Mg, Al, and Fe) (He et al., 2018; Nicolodelli et al., 2019), and satisfactory

* Correspondence to: Laboratory of Precision Agriculture, Department of Biosystems Engineering, "Luiz de Queiroz" College of Agriculture, University of São Paulo, 13418900 Piracicaba, São Paulo, Brazil.
E-mail addresses: tiagosrt@usp.br (T.R. Tavares), abdul.mouazen@ugent.be (A.M. Mouazen), lcnunes@cena.usp.br (L.C. Nunes), fe.chicoo@gmail.com (F.R. dos Santos), fmelquiades@uel.br (F.L. Melquiades), thainararebelo3@usp.br (T.R. da Silva), fjkrug@cena.usp.br (F.J. Krug), jpmolin@usp.br (J.P. Molin).

https://doi.org/10.1016/j.still.2021.105250
Received 8 May 2021; Received in revised form 11 October 2021; Accepted 1 November 2021
Available online 12 November 2021

determinations of the total contents of these elements in soil samples are commonly found in the literature (Hussain et al., 2007; Riebe et al., 2019; Erler et al., 2020; Xu et al., 2019a). LIBS spectra can also be used as a proxy to infer about soil fertility attributes [e.g., extractable (ex-) nutrients, pH, cation exchange capacity (CEC), organic matter (OM), pH, base saturation (V), and textural attributes] (Senesi, 2020). However, research regarding LIBS applications for soil fertility analysis are still incipient. In temperate soils, the results of the few studies carried out are not always satisfactory. For example, in German agricultural soils, Erler et al. (2020) have achieved good prediction performances for pH (R^2 between 0.91 and 0.95), reasonable predictions for humus content (R^2 between 0.54 and 0.66), and poor predictions for ex-P (R^2 between 0.22 and 0.35). Similarly, in Chinese agricultural soils, studies have shown poor prediction performance for ex-K and ex-P (Xu et al., 2019a), reasonable predictions for OM ($R^2 = 0.58$) (Xu et al., 2019b), and good prediction for CEC, OM, and pH ($R^2 \geq 0.81$) (Xu et al., 2019a). Xu et al. (2019c) reported satisfactory predictions for pH, OM, and total N (R^2 between 0.60 and 0.75), and poor performances ($R^2 < 0.35$) for ex-K and ex-P in Chinese soils. In Brazilian tropical soils, although good prediction performance ($R^2 \geq 0.85$) have been achieved for textural attributes (Villas-Boas et al., 2016) and pH (Ferreira et al., 2015), evaluation of the prediction accuracy for other key fertility attributes (e.g., CEC, V, OM, and extractable nutrients) has not been addressed in the literature yet.

The selection of an optimal modelling procedure for LIBS spectral data, capable of extracting useful but hidden information is crucial for accurate prediction of soil fertility attributes. Although there is no consensus regarding the best modelling strategy to adopt, recent studies suggested using multivariate techniques for the determination of soils elementary constitution when using LIBS spectra, in particular the partial least squares regression (PLS) (Takahashi and Thornton, 2017; Riebe et al., 2019). In addition, multivariate models are a simple and useful strategy to mitigate the matrix effect commonly present in soil samples (Takahashi and Thornton, 2017; Tavares et al., 2020b). Despite of the benefits, multivariate models using the entire spectrum require high computational capacity, especially due to the high spectral resolution of LIBS data that result in a large number of spectral variables (e.g., thousands data points) (Erler et al., 2020). In this regard, algorithms for the selection of most significant variables, such as deep learning or interval successive projection algorithm in partial least squares (iSPA-PLS) (Gomes et al., 2013; Riebe et al., 2019; Niu et al., 2021), can be a useful technique to reduce the amount of LIBS input variables and simplify data acquisition and processing procedures. To our best knowledge, no previous work reported the performance of iSPA-PLS for the prediction of key soil attributes in tropical soils, in comparison with full-spectra based PLS analyses.

This work aimed at evaluating different modelling approaches, namely, multiple linear regression (MLR) using selected emission lines of LIBS data, PLS using the full LIBS spectra from 200 to 540 nm (with 38,880 variables), and PLS using specific spectral regions selected by the iSPA algorithm for the prediction of key soil fertility attributes in Brazilian tropical soils. The best performing model was compared with univariate linear regression (ULR) in order to assess the influence of the matrix effect on the prediction of extractable nutrients (ex-P, ex-Ca, and ex-Mg).

2. Material and Methods

Soil samples and reference analysis

A set of 102 soil samples, belonging to the soil sample bank of the Precision Agriculture Laboratory (LAP) from Luiz de Queiroz College of Agriculture, University of São Paulo, were used in this study. These samples were collected from 0 to 20 cm depth in two Brazilian agricultural fields (designated as Field 1 and Field 2) and stored after being air-dried and sieved at 2 mm. The analysis results of the LAP's soil sample bank were used to choose samples with wide ranges of variability of key fertility attributes. Field 1 is of Lixisol (IUSS, 2014) or Ultisol (Soil Survey Staff, 2010), having a clayey texture, and Field 2 is of Ferralsol (IUSS, 2014) or Oxisol (Soil Survey Staff, 2010), with texture varying between a sandy loam and sand clay loam. Lixisols (Ultisol) and Ferralsols (Oxisol) are weathered soils, which are quite common and representative of Brazilian tropical areas (Santos, 2018).

After dataset selection, all samples were again sent to a commercial laboratory of soil fertility analyses for determining clay, organic matter (OM), cation exchange capacity (CEC), pH, base saturation (V), and extractable (ex-) nutrients (ex-P, ex-K, ex-Ca, and ex-Mg) contents, which were used as reference for the development of calibration models. Soil attributes were determined following the methods described by Van Raij et al. (2001), in which extractable nutrients were determined using ion exchange resin extraction, OM by oxidation with potassium dichromate solution, pH with calcium chloride solution, and clay by the Bouyoucos hydrometer method. The CEC was calculated by totaling the soil potential acidity (H + Al) plus the sum of bases (ex-Ca + ex-Mg + ex-K), and V was calculated by rationing the sum of the bases by the CEC. Finally, H + Al, used for CEC calculation, was extracted via the buffer solution method (SMP).

The pseudo total content (ptc) of macronutrients (P, K, Ca, and Mg) was deter-mined for comparison with its extractable contents. The determination of macronutrients ptc was performed following the United States Environmental Protection Agency (USEPA) Method 3051 A (Element, 2007). This microwave-assisted digestion method consists of a partial dissolution of the analytes in the samples with HNO_3 and

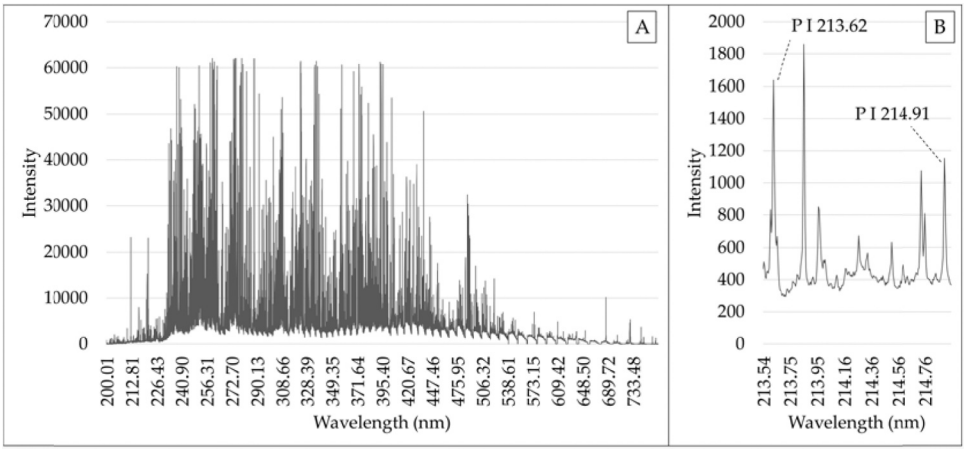


Fig. 1. Full Laser Induced Breakdown Spectroscopy (LIBS) spectra obtained from the pellet soil sample (A), and P emission intensities lines in the range between 213.50 and 215.00 nm (B), under the defined experimental conditions.

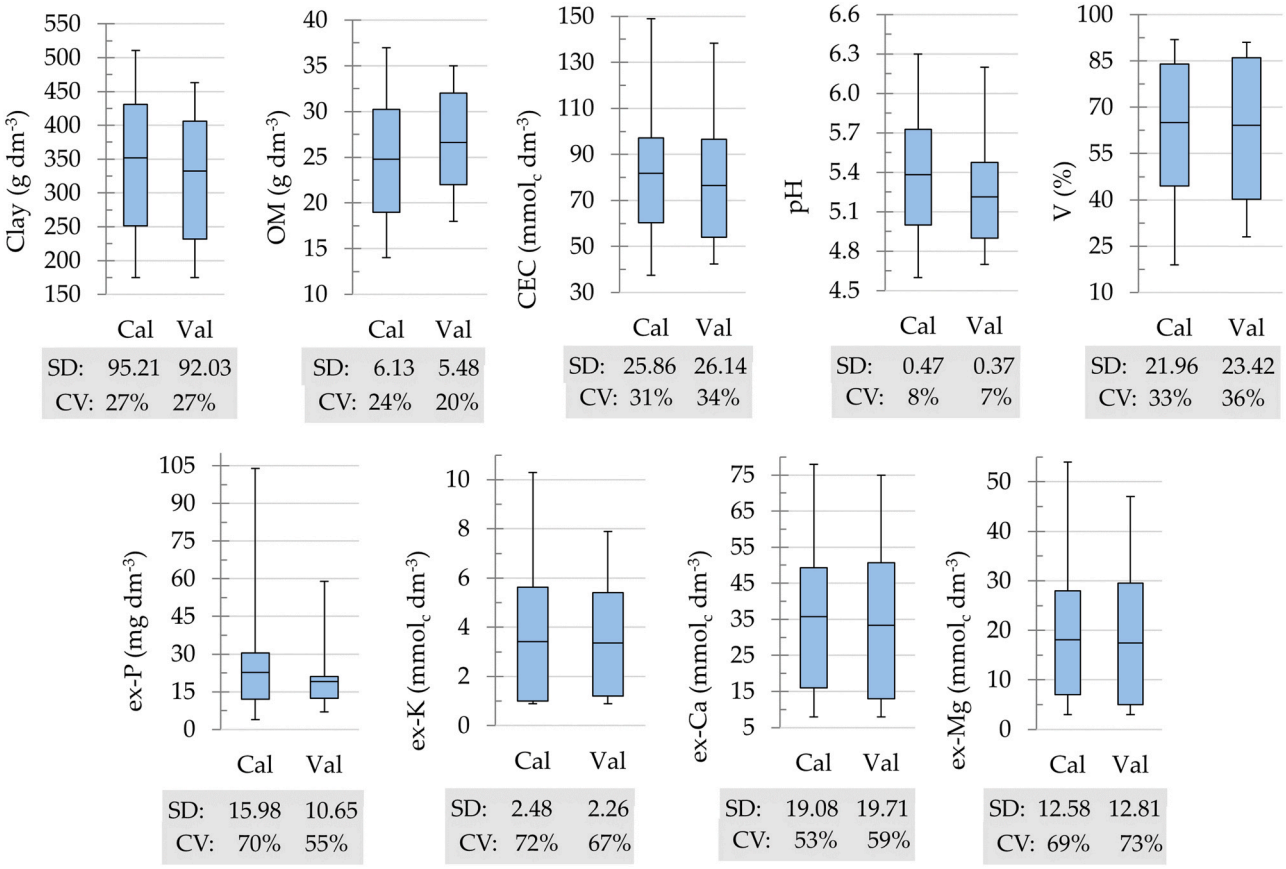


Fig. 2. Descriptive statistics of the studied soil fertility attributes provided for the calibration and validation set. The abbreviations CV, SD, OM, CEC, V, and the suffix ex- were used for coefficient of variation, standard deviation, organic matter, cation exchange capacity, base saturation, and extractable nutrients, respectively.

HCl, followed by the quantification via inductively coupled plasma optical emission spectrometry (ICP OES). In tropical soils, this method gives proportional recoveries compared to those based on total sample de-composition methods (Silva et al., 2014). The relations between ptc and extractable contents of macronutrients were evaluated by simple regression analysis, as well as by the ratio between both contents (ptc-ex ratio).

Test sample preparation for LIBS

The presentation of soil test samples in the form of pressed pellets is of fundamental importance for soil analysis by LIBS (Jantzi et al., 2016). The samples with 10% w w^{-1} of microcrystalline cellulose powder (Sigma-Aldrich, Merck, Darmstadt, Germany) were ground and homogenized using a planetary ball mill (PM 200 mill, Retsch, Haan, Germany), for 20 min (10 cycles of 2 min grinding), and pelletized (15 mm diameter) in a hydraulic press (SPEX 3624B X-Press, Metuchen, EUA), as described elsewhere by Tavares et al. (2019).

LIBS measurements

It was used a Q-switched Nd:YAG laser (Brilliant, Quantel, France) at 1064 nm, generating 5 ns pulses of up to 365 mJ, in a 6 mm diameter beam, at 10 Hz repetition rate. The laser was focused on the pellet surface by a convergent lens of 2.54 cm in diameter and 20 cm focal length (Newport, USA). Pellets were positioned into a plastic sample holder situated in a two axes manually-controlled translation stage movable in the plane orthogonal to the laser direction. In order to displace the atmospheric air from the sample surface, a laminar stream of argon (5.0 L min^{-1}) was continuously fed from the bottom of the sample holder as described elsewhere by Nunes et al. (2019).

The emission spectrum generated by the plasma was collected with a telescope composed by 50 mm and 80 mm focal length fused silica lenses (LLA Instruments GmbH, Germany) and coupled to the entrance slit of the spectrometer (model ESA 3000, LLA Instruments GmbH, Germany) with Echelle optics and ICCD detector. The collection angle with respect to the laser optical axis was 25°. The spectrum obtained presents range from 200 to 780 nm, with a resolution oscillating from 5 pm at 200 nm to 19 pm at 780 nm (Fig. 1). The same LIBS system used in this study was described elsewhere, e.g., by Nunes et al., (2019). ESAWIN software (LLA Instruments GmbH, Germany) and NIST Atomic Spectra Database (Ralchenko and Kramida, 2020) were used for identification of the analytes emission lines.

The instrumental conditions including (i) laser pulse energy, (ii) lens-to-sample distance (LTSD), (iii) number of accumulated laser pulses, (iv) delay time, and (v) integration time gate, were optimized in initial tests in order to obtain the maximum signal-to-noise ratio (SNR) of the emission lines of interest, in which the ICCD detector did not saturate (Nunes et al., 2009). The best results were obtained with 15 laser pulses, 2.0 μs delay time with a 7.0 μs integration time gate. The lens-to-sample distance and the pulse energy were adjusted at 19.5 cm and 65 mJ, respectively, leading to 250 J cm^{-2} at the sample surface.

In order to minimize drawbacks related to analytes micro heterogeneity, each pellet was analysed (interrogated) at 21 different sites. To avoid re-ablation at the edges of the neighboring craters, a distance of about 1 mm between them was maintained. The spectra acquired of each test sample were averaged in one representative spectrum for further analyses.

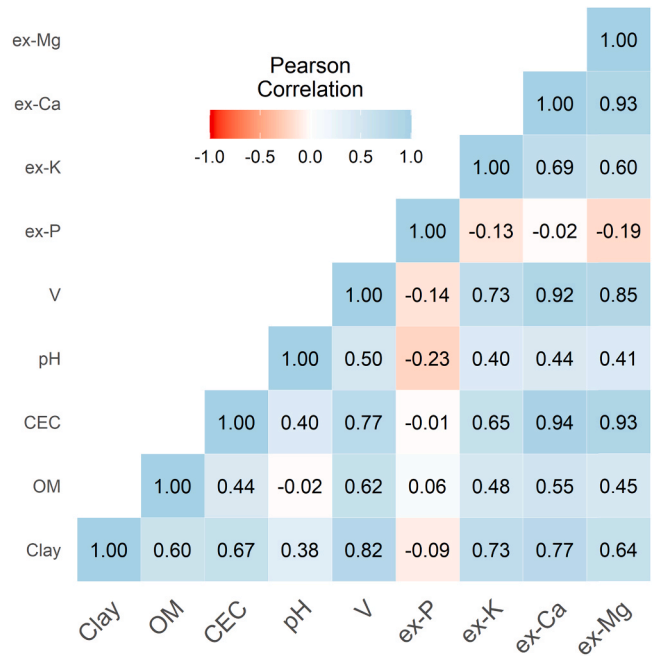


Fig. 3. Matrix of Pearson's correlation coefficient of the measured studied soil fertility attributes. The abbreviations OM, CEC, V, and the suffix ex- were used for organic matter, cation exchange capacity, base saturation, and extractable nutrients, respectively.

Modeling

For the development of the calibration models, 102 soil samples were separated using Kennard Stone algorithm (Kennard and Stone, 1969) into two sets of 68 and 34 samples, which were used for calibration and validation, respectively. Calibration models involving univariate linear regressions (ULR), multiple linear regressions (MLR), and partial least squares regressions (PLS) were built and validated.

Univariate models (ULR) were obtained from the regression of the peak area of the corresponding emission line (after background subtraction) against the reference soil extractable nutrient, by using the classical least squares regression model (Danzer and Currie, 1998). On the basis of spectral selectivity and sensitivity, P I 214.914 nm, Ca II 315.887 nm, and Mg I 277.983 nm emission lines were selected for the prediction of ex-P, ex-Ca, and ex-Mg, respectively (Nunes et al., 2010; Yu et al., 2016; Riebe et al., 2019). K was not determined since it is below the detection limit of the LIBS method, under the instrumental conditions used.

For the MLR models, 22 emission lines of elements commonly found in tropical soils (namely, C, Na, Mg, Al, Si, P, Ca, Ti, Mn, Fe, Cu, and Zn) were used as X-variables. Two well-defined emission lines for each of the above-mentioned element were select-ed. However, for C and Na only one clear emission line in the LIBS spectra was used. Only well-defined emission lines were chosen in order to avoid spectral interference from other elements. Following the mentioned criteria, the emission lines C I 247.856, Na II 330.135, Mg I 277.983, Mg I 285.213, Al I 308.2152, Al I 309.2709, Si I 212.412, Si I 390.552, P I 213.6182, P I 214.914, Ca II 315.887, Ca II 317.933, Ti II 353.541, Ti I 398.176, Mn II 294.920, Mn I 293.306, Fe I 305.908, Fe II 234.349, Cu I 324.754, Cu I 327.390, Zn II 206.200, and Zn I 213.857 nm were selected. No N emission lines were identified in our LIBS spectra.

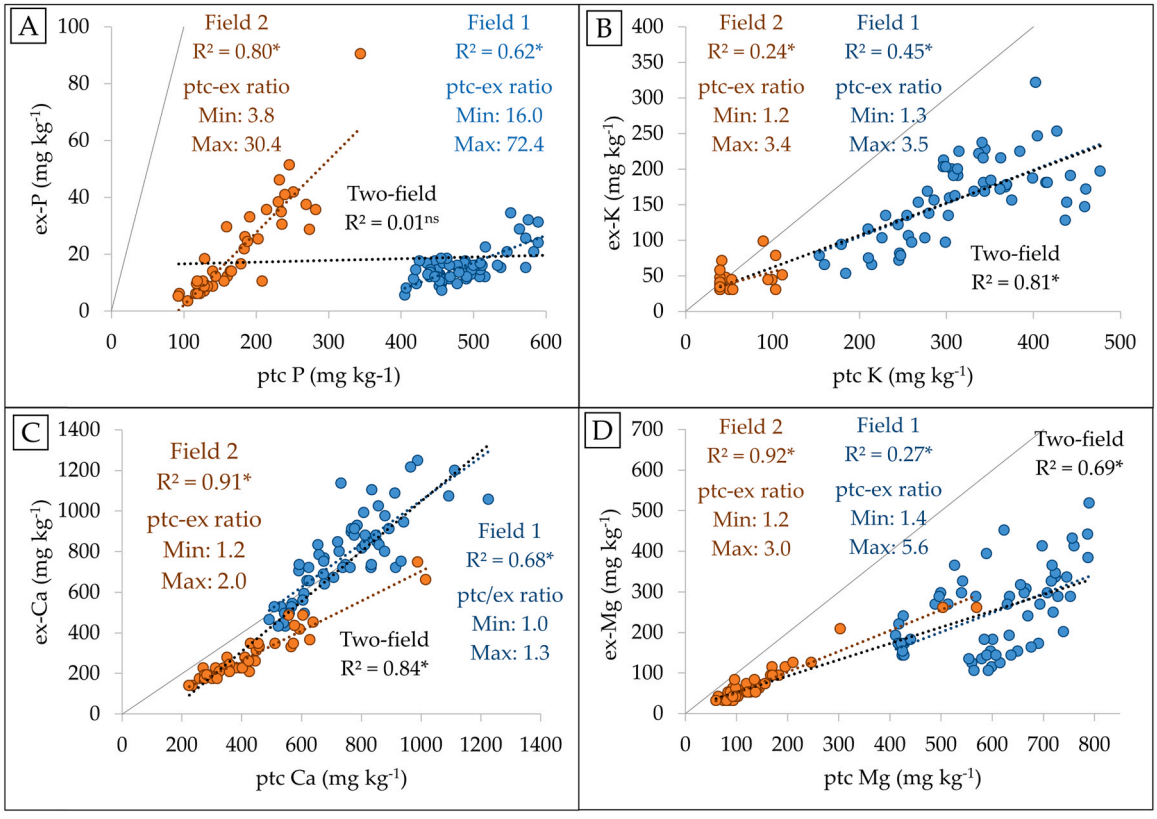


Fig. 4. Scatter plots of extractable and pseudo total content (ptc) of P (A), K (B), Ca (C), and Mg (D). The minimum and maximum ratio between ptc and extractable contents (ptc-ex ratio), as well as the coefficient of determination (R^2) of the regression analysis, were presented. Information in blue, orange, and black correspond, respectively, to Field 1, Field 2, and Two-field data. Significant and non-significant regression analysis, at the probability level of 0.05, were marked with "* " and "NS", respectively.

Table 1
Prediction results of the validation (n = 34) obtained from partial least squares regression (PLS) and multiple linear regression (MLR) using LIBS data. The PLS models were built using the LIBS full spectra (FS) and spectral regions selected by the interval successive projection algorithm (iSPA). The MLR models were calibrated with the 21 selected emission lines (EL). The number of latent variables (LV) used in each model is also presented.

Table 1. Prediction results of the validation (n = 34) obtained from partial least squares regression (PLS) and multiple linear regression (MLR) using LIBS data. The PLS models were built using the LIBS full spectra (FS) and spectral regions selected by the interval successive projection algorithm (iSPA). The MLR models were calibrated with the 21 selected emission lines (EL). The number of latent variables (LV) used in each model is also presented.

Regression approach	Variable selection	Clay	OM[1]	CEC[2]	pH	V[3]	ex-P[4]	ex-K[4]	ex-Ca[4]	ex-Mg[4]
		---------- R^2 ----------								
MLR	21 EL	0.80	0.64	0.88	0.06	0.86	0.59	0.58	0.91	0.89
PLS	FS	0.75	0.76	0.84	0.30	0.94	0.16	0.74	0.91	0.92
	iSPA	0.89	0.81	0.84	0.31	0.94	0.72	0.76	0.94	0.93
		---------- RMSE ----------								
MLR	21 EL	43.88	4.30	10.55	0.43	8.56	7.40	1.51	6.47	5.06
PLS	FS	47.16	2.89	10.75	0.34	5.74	10.69	1.14	6.19	4.16
	iSPA	27.68	2.56	10.69	0.33	5.49	5.84	1.10	4.85	3.51
		---------- RMSE% ----------								
MLR	21 EL	15.24	25.29	11.00	28.60	13.59	14.24	21.52	9.66	11.49
PLS	FS	16.38	18.52	14.95	23.02	10.62	21.47	16.42	9.48	9.45
	iSPA	12.38	15.05	11.14	21.95	8.71	11.23	15.75	7.23	7.97
		---------- RPD ----------								
MLR	21 EL	2.10	1.27	2.48	0.86	2.74	1.44	1.50	3.04	2.53
PLS	FS	1.95	1.89	2.43	1.08	4.07	1.00	1.98	3.19	3.08
	iSPA	3.06	2.14	2.45	1.12	4.27	1.82	2.05	4.07	3.65
		---------- LV ----------								
PLS	FS	4	6	4	3	5	6	2	5	6
	iSPA	5	6	6	4	3	5	4	5	4

[1] Organic matter; [2] cation exchange capacity; [3] base saturation; [4] extractable (ex-) nutrients. The coefficient of determination (R^2) and residual prediction deviation (RPD) values are presented on grayscale, highlighting the highest values. The root-mean-square error (RMSE) was given in g dm^{-3} for clay and OM; in $mmol_c$ dm^{-3} for CEC, ex-K, ex-Ca, and ex-Mg; in % for V; and, for ex-P, the RMSE was given in mg dm^{-3}.

[1] Organic matter; [2] cation exchange capacity; 3 base saturation; 4 extractable (ex-) nutrients. The coefficient of determination (R^2) and residual prediction deviation (RPD) values are presented on grayscale, highlighting the highest values. The root-mean-square error (RMSE) was given in g dm− 3 for clay and OM; in mmolc dm− 3 for CEC, ex-K, ex-Ca, and ex-Mg; in % for V; and, for ex-P, the RMSE was given in mg dm− 3.

For the PLS models, two calibration strategies have been performed: (i) using the full spectra within the region from 200 to 540 nm (designated in this study as FS-PLS), which covers most of the valuable and high-intensity emission lines of soil LIBS data (Yu et al., 2016), and (ii) using regions selected through the iSPA (Gomes et al., 2013). For the iSPA, the LIBS spectra from 200 to 540 nm were divided into 160 intervals of 243 variables. The iSPA uses projection operations iteratively to select intervals of variables with minimum collinearity to achieve good prediction ability (Gomes et al., 2013). The number of latent variables (LVs) used in PLS models was selected based on the lowest root mean square error value with 5-fold cross-validation (RMSECV).

The models quality was assessed on the validation set by means of the coefficient of determination (R^2), the root mean square error (RMSE), the relative error (RMSE%), and the residual prediction deviation (RPD). To calculate the relative error, the RMSE was divided by the range of the laboratory measured soil property. The RPD interpretations were carried out following the classes proposed by Chang et al. (2001), in which RPD values below 1.40 indicate bad models, values between 1.40 and 2.00 determine reasonable models, and values above 2.00 indicate good models.

3. Results

Laboratory measured soil properties.

The descriptive statistics of soil fertility attributes for the calibration and validation datasets are shown in Fig. 2. The Kennard Stone algorithm ensuring comparable range and standard deviation (SD) for the calibration and validation sets (Fig. 2) is fundamental for an effective

Table 2
Improvement of prediction performance, represented by the reduction of root-mean-square error (RMSE), achieved with the best predictive model compared to the performance reached with the 2nd and 3rd predictive approaches. The partial least squares regression (PLS) models were built with LIBS full spectra (FS) and spectral regions selected by the interval successive projection algorithm (iSPA). The multiple linear regression (MLR) models were calibrated with 21 selected emission lines (EL). The presented results were obtained by using the validation set (n = 34).

Table 2. Improvement of prediction performance, represented by the reduction of root-mean-square error (RMSE), achieved with the best predictive model compared to the performance reached with the 2nd and 3rd predictive approaches. The partial least squares regression (PLS) models were built with LIBS full spectra (FS) and spectral regions selected by the interval successive projection algorithm (iSPA). The multiple linear regression (MLR) models were calibrated with 21 selected emission lines (EL). The presented results were obtained by using the validation set (n = 34).

Regression approach	Variable selection	Clay	OM[1]	CEC[2]	pH	V[3]	ex-P[4]	ex-K[4]	ex-Ca[4]	ex-Mg[4]
		RMSE								
MLR	21 EL	43.88	4.30	**10.55**	0.43	8.56	7.40	1.51	6.47	5.06
PLS	FS	47.16	2.89	10.75	0.34	5.74	10.69	1.14	6.19	4.16
	iSPA	**30.03**	**2.56**	10.69	**0.33**	**5.49**	**5.84**	**1.10**	**4.85**	**3.51**
		RMSE improvement of the best approach								
compared to 2nd		-32%	-11%	-1%	-3%	-4%	-21%	-4%	-22%	-16%
compared to 3rd		-36%	-40%	-2%	-23%	-36%	-45%	-27%	-25%	-31%

[1] Organic matter; [2] cation exchange capacity; [3] base saturation; [4] extractable (ex-) nutrients. The RMSE values are presented on grayscale, highlighting the lowest ones. The lowest RMSE was also presented in bold. The RMSE was given in g dm^{-3} for clay and OM; in $mmol_c$ dm^{-3} for CEC, ex-K, ex-Ca, and ex-Mg; in % for V; and, for ex-P, the RMSE was given in mg dm^{-3}.

[1] Organic matter; [2] cation exchange capacity; 3 base saturation; 4 extractable (ex-) nutrients. The RMSE values are presented on grayscale, highlighting the lowest ones. The lowest RMSE was also presented in bold. The RMSE was given in g dm− 3 for clay and OM; in mmolc dm− 3 for CEC, ex-K, ex-Ca, and ex-Mg; in % for V; and, for ex-P, the RMSE was given in mg dm− 3.

evaluation of the models' predictive performance (Mourad et al., 2005; Mouazen et al., 2006). It is also important to mention that the selected samples have wide variability ranges of fertility attributes, represented by their high coefficients of variation (CV), which is always above 20%. The only exception is for pH, due to its logarithmic scale.

Pearson's correlation coefficients (r) between all fertility attributes presented in Fig. 3 aid to understand the covariations that explain indirect predictions (Tavares et al., 2020a), e.g., successful predictions of CEC, which has no emission lines in LIBS spectra. Soil CEC, V, ex-Ca, and ex-Mg are closely associated in our dataset, with strong positive correlations (r between 0.85 and 0.94). Ex-K possesses high positive correlations with clay (r = 0.73) and V (r = 0.73), and moderate positive correlations with CEC, ex-Ca, and ex-Mg (r between 0.60 and 0.69). Clay content showed strong positive correlations with V, ex-K, and ex-Ca (r between 0.73 and 0.82), and moderate positive correlations with OM, CEC, and ex-Mg (r between 0.60 and 0.67). OM had weak correlations with CEC, pH, ex-P, ex-K, and ex-Mg (r between −0.02 and 0.48), and moderate correlations with clay, V, and ex-Ca (r between 0.55 and 0.62). Ex-P and pH were the attributes that presented in general the weakest association with the other attributes, with r oscillating between − 0.23 and 0.06, and from − 0.23–0.50, respectively.

The scatter plots of Fig. 4 show the relationships between ptc and extractable contents of the macronutrients for the individual field and merged Two-field dataset (Field 1 + Field 2) sets. The relationship between ptc and extractable contents are characterized by the R^2 of the regression line, as well as by the ratio between ptc and extractable contents (ptc-ex ratio) observed. In the Two-field dataset, the scatter plots between ptc and extractable contents of K, Ca, and Mg showed a linear dispersion of points with a significantly high R^2 values of 0.81, 0.84, and 0.69, for K, Ca, and Mg, respectively. In contrast, a non-linear behavior with R^2 of 0.01 was observed for P. For the individual Field 1 and Field 2 sets it is possible to observe that all macronutrients show a linear behavior with significant R^2 values. The non-linearity of P observed for the Two-field set is explained by the contrasting ptc ratio presented by each individual field. In other words, P was the only nutrient that showed a contrasting ptc-ex ratio between both fields (oscillating from 16.0 to 72.4 for Field 1, and from 3.8 to 30.4 for Field 2; Figura 4 A), whereas for all other nutrients the ptc-ex ratio showed values below 5.6 (Fig. 4B, C, D). These linear relations observed for K, Ca and Mg suggest that their extractable fractions can be successfully assessed by LIBS using its elementary information (once their emission line is detected). On the other hand, if satisfactory predictions for the ex-P is obtained, this should be explained by the ability of the prediction model to comprehend the different relationships between ptc and extractable content of P observed for each individual field. Satisfactory predictions for ex-P could also be attributed to the high correlations P has with K, Ca, and Mg; however, this is not the case in this study due to the weak correlations of ex-P with the other attributes (Fig. 3).

Prediction performance of the different multivariate approaches.

The results of the fertility attributes prediction using MLR and PLS are presented in Table 1. Complementing these results, Table 2 shows the performance gain (in percentage of RMSE reduction) of the best performing method compared to the others. In general, the LIBS' prediction performance was satisfactory (RPD ≥ 1.40) for clay, OM, CEC, V, ex-P, ex-K, ex-Ca, and ex-Mg using at least one of the multivariate approaches. Only pH did not show satisfactory predictions (RPD ≤ 1.12). The best prediction performances were obtained for CEC, ex-Ca, and ex-Mg, which had RPD always greater than 2.0, regardless of the multivariate approach applied. Good prediction performances (RPD > 2.0) were also obtained for clay, OM and ex-K prediction, using at least one of

Table 3
Information about the intervals of variables selected by the interval successive projection algorithm (iSPA) used as input variables for partial least squares regression (PLS) analyses.

Attribute	Number of intervals selected	Total of variables used	Reduction*	Spectral ranges of the selected intervals (nm)
Clay	3	729	53 times	220.84 – 222.21, 246.77 – 248.29, and 251.37 – 252.94
OM[1]	1	243	160 times	295.08 – 296.92
CEC[2]	52	12,636	3 times	201.31 – 272.37, 287.89 – 298.73, and 414.15 – 416.63
pH	3	729	53 times	220.84 – 222.21, 237.82 – 239.27, and 391.78 – 394.22
V[3]	57	13,851	3 times	201.31 – 272.37, 279.18 – 300.60, and 396.59 – 399.10
ex-P[4]	1	243	160 times	214.14 – 215.44
ex-K[4]	1	243	160 times	375.28 – 377.53
ex-Ca[4]	1	243	160 times	389.37 – 391.77
ex-Mg[4]	1	243	160 times	284.36 – 286.14

[1] Organic matter; [2] cation exchange capacity; [3] base saturation; [4] extractable (ex-) nutrients. The variable importance of projection (VIP) scores of each interval of variables are shown in Figure A1 (Appendix Section). * Reduction in the number of variables used in the iSPA-PLS models when compared to the number of variables present in the range between 200 and 540 nm (n = 38,880 variables), which was used in the SF-PLS approach

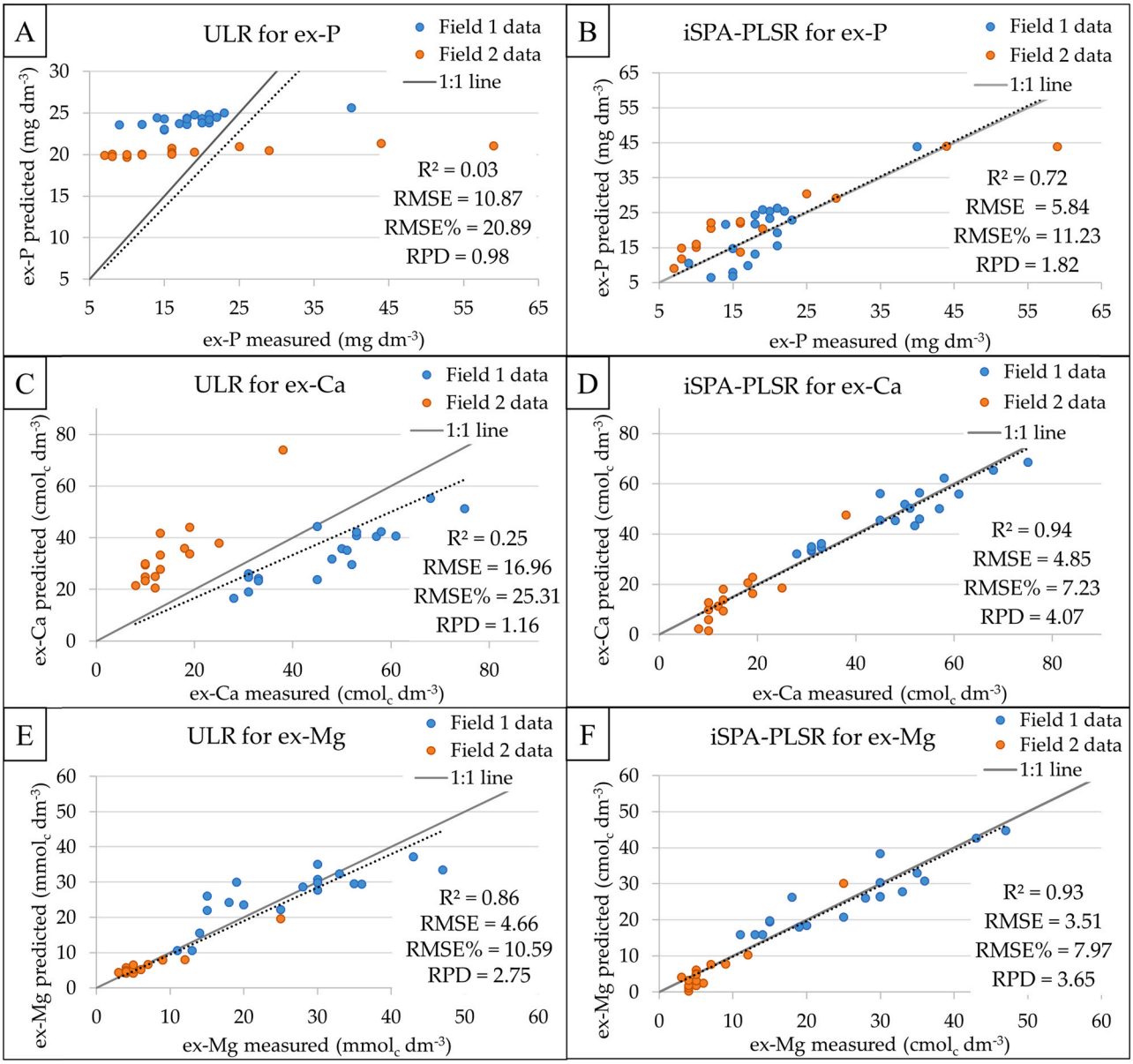


Fig. 5. Scatter plots of measured versus predicted extractable (ex-) P (A, B), Ca (C, D), and Mg (E, F). The A,C, and E plots were obtained using univariate linear regression (ULR), whereas the B,D, and F plots were obtained with partial least squares regression (PLS) after interval variables selection by the interval successive projection algorithm (iSPA). The figures concern the prediction performance of the models applied on the validation dataset (n = 34).

the evaluated multivariate methods.

The comparison of the different multivariate modelling approaches (Table 2) revealed that in general the best performing is the iSPA-PLS models, obtaining the lowest RMSE for all attributes. The only exception was for CEC, whose the lowest prediction error was obtained by using MLR. In any case, MLR reduced the error of CEC prediction by only 1%, compared to the error obtained with iSPA-PLS (Table 2). For the pre-diction of OM, pH, V, ex-K, ex-Ca and ex-Mg the sequence of methods giving the best to worst performance was iSPA-PLS > FS-PLS > MLR. For predicting clay and ex-P the sequence of best performing approaches was iSPA-PLS > MLR > FS-PLS, whereas for CEC, the sequence was MLR > iSPA-PLS > FS-PLS. Finally, it is worth noting that the gain in the prediction performance obtained for ex-P ranged from poor performance (RPD = 1.00 and RMSE = 10.69) with FS-PLS, to satisfactory performance with MLR and iSPA-PLS approaches (RPD of 1.44 and 1.82, respectively). This performance gain represented an error reduction of 21% (RMSE = 7.40) and 45% (RMSE = 5.84) for MLR and iSPA-PLS, respectively.

The spectral regions as well as the total number of variables used in the iSPA-PLS approach for predicting each soil attribute are presented in Table 3. Compared to the FS-PLS approach, iSPA-PLS allowed a 160-fold reduction in the total number of variables for OM, ex-P, ex-K, ex-K and ex-Mg prediction; a 53-fold reduction for clay and pH prediction; and about 3-fold reduction for CEC and V prediction. It is also note-worthy that the variables used as input lie between 201.31 and 416.63 nm, while no variables after 416.63 nm were used.

Matrix effect: comparison of univariate and multivariate approaches for extractable nutrients prediction.

Fig. 5 shows the scatter plots of measured and predicted values of ex-P, ex-Ca, and ex-Mg using ULR and the multivariate iSPA-PLS models. Univariate models for ex-P and ex-Ca (Fig. 5A,C) showed poor performance (RPD $\leq$ 1.16), which is attributed to matrix effect, originated from sample physicochemical properties. The multivariate predictive models (Fig. 5B,D,F) demonstrate less matrix-matching dependent, which showed markedly higher performances (RPD of 1.82, 4.07, and 6.65, for ex-P, ex-Ca, and ex-Mg, respectively) than the URL. For the multivariate models, points are closely distributed around the 1:1 line. The same trend is also seen for the other multivariate approaches (MLR and FS-PLS) in Figure A2 (Appendix Section). It is also noteworthy that the matrix effect for the prediction of ex-Mg was less significant than that observed for ex-P and ex-Ca, since ex-Mg provided very good prediction result (RPD = 2.75) with the ULR model, showing linear behavior with points closely spread around the 1:1 line (Fig. 5E). Nevertheless, the the iSPA-PLS model showed clearly superior performance (RPD = 3.65) and much closer points dispersion to the 1:1 line, when compared to the ULR.

4. Discussion

Predicting key soil fertility attributes using LIBS

The LIBS spectra obtained from the pelletized soil samples allowed satisfactory prediction results ($0.59 \leq R^2 \leq 0.94$) using multivariate models developed for eight out of nine fertility attributes, although pH was the only attribute that showed poor prediction performance ($R^2 \leq 0.31$). Soil fertility attributes are indirectly predicted via elemental analysis sensors due to relationship existing between such attributes and the elemental constitution of the soil samples. Textural attributes are often related to the content of Fe, Si, Al, Ti, among others, which are structural components of the different particle size fractions (Villas-Boas et al., 2016; Lima et al., 2019). The results obtained for clay prediction ($R^2 = 0.89$) were close to those obtained by Villas-Boas et al. (2016) ($R^2 = 0.83$), working with a LIBS sensor for the analysis of tropical soils. Ferreira et al. (2015) obtained successful models for pH ($R^2 = 0.87$) using Al, Ca, H, and O emission lines collected for tropical soils. However, the prediction results for pH were poor in our study. Satisfactory predictions of OM using LIBS were reported using C or N emission lines (Senesi and Senesi, 2016), or indirectly, through relationships with textural attributes (Lima et al., 2019). The performance of OM predictions obtained in this work oscillated according to the modelling approach used ($0.64 \leq R^2 \leq 0.81$), although this range of accuracy was similar to that reported in studies conducted on Chinese soils that showed R^2 oscillating between 0.58 and 0.83 (Xu et al., 2019a, 2019b, 2019c). Although satisfactory predictions of CEC and V using LIBS have not yet been reported in the literature, good predictive models ($R^2 \geq 0.75$) can also be achieved (as achieved in this work), once they present covariations with well-predicted fertility attributes (e.g., textural attributes, extractable nutrients, etc), as reported by Tavares et al. (2020a) and Silva et al. (2017), both using XRF sensors.

Regarding the extractable nutrients, the correlation between total and extractable contents is the central reason to explain successful predictions using the elemental concentration of a given nutrient. In this work, ptc and extractable contents of K, Ca and Mg showed significant correlations (Fig. 4) in both the individual and Two-field datasets. Thus, the ex-Ca and ex-Mg predictions were very successful, obtaining R^2 value higher than 0.88. On the other hand, ex-K predictions, although satisfactory, presented inferior prediction performance ($0.58 \leq R^2 \leq 0.74$), which is attributed to the absence of identifiable K emission lines in the LIBS spectrum, and the level of accuracy obtained may well be due to its covariation with other fertility attributes, e.g., clay (r = 0.73), ex-Ca (r = 0.69), and ex-Mg (r = 0.60) (Fig. 3). It is interesting to mention that due to the weathered nature of tropical soils, their composition are generally poor in primary minerals that are natural source of nutrients (Fontes, 2012). Thus, the relations between the extractable and total contents of nutrients, such as Ca and Mg, are largely controlled by agricultural management (e.g., fertiliser inputs) rather than natural processes (Andrade et al., 2020).

The soil chemistry regarding P in tropical soils is complex and characterized by strong association with Fe oxides (hematite and goethite) and Al (gibbsite). Due to this complexity, it is common to observe unbalanced relationships between its total content present in the soil and the amount of ex-P (Schäefer et al., 2008). This behavior was observed for P when merging the data from both fields (Two-field dataset). For this reason, we were not expecting satisfactory predictions for this nutrient via LIBS spectra using the Two-field dataset. On the other hand, in the individual field datasets, ex-P showed a significant correlation with its ptc (Fig. 4 A). This is what would explain our satisfactory predictions for ex-P when using both multivariate iSPA-PLS and MLR models, and not when using the univariate model. In other words, the multivariate models allowed to correct the different ratios between total and ex-tractable contents present in both individual fields, resulting in satisfactory models for ex-P (R^2 of 0.59 and 0.72 using MLR and iSPA-PLS, respectively).

Due to the indirect nature of predictions using LIBS sensors, it is important that further investigations address the temporal stability of the models' performance. Understanding the LIBS' performance for soil fertility prediction as a function of time/soil management will be fundamental to advance our knowledge regarding: (i) understanding and mitigating factors that influence sensor performance and (ii) developing spectral libraries of LIBS data for soil analysis.

Multivariate approaches, number of variables, and matrix effect mitigation

Using the iSPA algorithm, for the variables selection, combined with PLS showed to be the best method for predicting fertility attributes via LIBS sensor. iSPA-PLS provided accurate predictions (RPD $\geq$ 1.82) for eight out of the nine attributes (namely, clay, OM, CEC, V, ex-P, ex-K, ex-Ca, and ex-Mg), by employing a lower number of input variables than those used in the FS-PLS approach (Table 3). Although MLR generally underperformed the iSPA-PLS, its performance was satisfactory (RPD between 1.44 and 3.04) for the prediction of clay, CEC, V, ex-P, ex-K, ex-Ca, and ex-Mg by using 21 selected emission lines. Thus, MLR can also be

considered a simple alternative method for predicting fertility attributes when drastic reductions in the number of variables are required. Reducing the number of input variables of LIBS data is necessary because their spectra contain a large number of variables, most of them are irrelevant for predicting fertility attributes. High volumes of data can negatively affect the performance of model calibration, as well as creates an unnecessary computational burden (Erler et al., 2020).

Although iSPA-PLS proved to be the best approach for predicting fertility attributes and mitigating the matrix effect for P, Ca, and Mg, all the other multivariate models also mitigated the matrix effect in some degree and achieved more accurate predictions than ULR models (Figure A2). The matrix effect in soil samples are associated with differences in material matrices, such as hardness, chemical composition, and density, which affects the intensities of the elemental emission lines (Senesi, 2020). The mitigation of the matrix effect with multivariate models can be explained by the fact that chemical and physical matrix effects are shown as elements' peaks. Therefore, multivariate regressions can extract this information in the spectra without requiring for prior assumptions of the plasma condition, as it is commonly used e.g., in calibration-free LIBS (Takahashi and Thornton, 2017). The correction of the matrix effect via multivariate approaches has also been reported by Tavares et al. (2020b) in analysis of fertility attributes using XRF sensors, which is also a matrix dependent technique.

The challenge of soil sample preparation

It is important to highlight operational aspects of test sample preparation that are little discussed in scientific works aiming at using LIBS sensors as a practical method of soil analysis. Although some studies have evaluated the performance of LIBS sensors using non-pelletized soil samples (intact soil core) (Bricklemyer et al., 2013, 2018), most published studies used soil samples conformed in the form of pressed pellets (Ferreira et al., 2015; Villas-Boas et al., 2016; Yu et al., 2016; Riebe et al., 2019; Xu et al., 2019a, 2019b, 2019c; Erler et al., 2020). This is because the pelletizing is of paramount importance before the data acquisition, as it improves homogeneity, as well as allow congruent ablations by laser, maintaining the stoichiometric proportion of the craters formed at each pulse, which is crucial for the replicability of the calibrated models (Jantzi et al., 2016). Quantitative analyses of test samples in the form of loose powder would have drastically reduced performance compared to analyses on pelletized test samples. In addition, although soil pelletizing does not require a reagent and is apparently simple to perform, soil comminution and pressing requires specific equipment (e.g., mill and press), which adds costs and operational complexity to the method. Moreover, particle size reduction in soils containing different portions of quartz (common mineral in the sand fraction) require the use of mills with compatible hardness (e.g., tungsten carbide), which are relatively more expensive and present extra challenges such as: (i) possibility of contamination of the sample with the milled material (Tavares et al., 2019), (ii) higher wear of the equipment. Thus, the aforementioned aspects bring into question the practicality of LIBS analysis when using pelletized soil samples. In this sense, it is important that further research needs to develop new solutions that avoid LIBS deteriorated performance when applied in non-pelletized soil samples. This is a critical issue to be overcome for the expansion of LIBS as a practical soil analysis within the areas of precision agriculture and soil science.

5. Conclusion

LIBS proved to be efficient for predicting fertility attributes in tropical soils. The LIBS spectra obtained from the pelletized soil samples associated with the multivariate models achieved satisfactory predictions (RPD $>$ 1.40) for eight out of the nine key soil fertility attributes, with pH being the only exception that showed poor predictive performance (RPD $\leq$ 1.12). The best prediction performances were obtained for CEC, ex-Ca, and ex-Mg, which had RPD always greater than 2.0, regardless of the multivariate approach applied. Good performances (RPD $>$ 2.0) were also obtained for clay, OM and ex-K prediction, using at least one of the evaluated multivariate methods. Among the extractable nutrients, the prediction of ex-P proved to be the most challenging (best RPD reaching 1.82). Although LIBS shows P emission lines, the relationship between extractable and pseudo total content (ptc) of P is complex, varying considerably in different agricultural areas. Further studies involving the prediction of ex-P by combining LIBS with other direct analysis techniques are encouraged.

Comparing the different multivariate approaches, multiple linear regressions (MLR) in general showed slightly lower performance than iSPA-PLS, it allowed satisfactory predictions (RPD between 1.44 and 3.04) for clay, CEC, V, ex-P, ex-K, ex-Ca, and ex-Mg by employing only 21 variables. Anyhow, the best performing models were those obtained after the interval successive projection algorithm in combination with partial least squares regression (iSPA-PLS). The iSPA-PLS allowed reducing the amount of variables by 160 times for OM, ex-P, ex-K, ex-K and ex-Mg, by 53 times, for clay and pH and by approximately 3 times for CEC and V, compared with the amount of variables in the full LIBS spectra from 200 to 540 nm (38,880 variables). The iSPA-PLS is an efficient and accurate modelling method that uses reduced number of variables to predict fertility attributes via LIBS sensor.

Due to the challenges related to pelletizing of soil samples further research is encouraged to find solutions to allow LIBS applications in soil samples with a simpler sample preparation procedure. Overcoming this limitation would expand the possibilities of LIBS sensors application in precision agriculture and soil science.

Funding acknowledgements

T.R.T. was funded by São Paulo Research Foundation (FAPESP) [grant number 2017/21969-0 and 2020/16670-9]. Soil fertility tests were funded by National Council for Scientific and Technological Development (CNPq) – "Edital de Chamada Universal" [grant number 458180/2014–9]. L.C.N. was funded by CNPq [grant number 381261/2020–4]. LIBS facilities were funded by FAPESP (FAPESP 04/15965–2) [grant number 2015–19121–8]. Authors also acknowledge the financial support received from the Research Foundation - Flanders (FWO) for Odysseus I SiTeMan Project [grant number G0F9216N].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

See Appendix Figs. A1,A2.

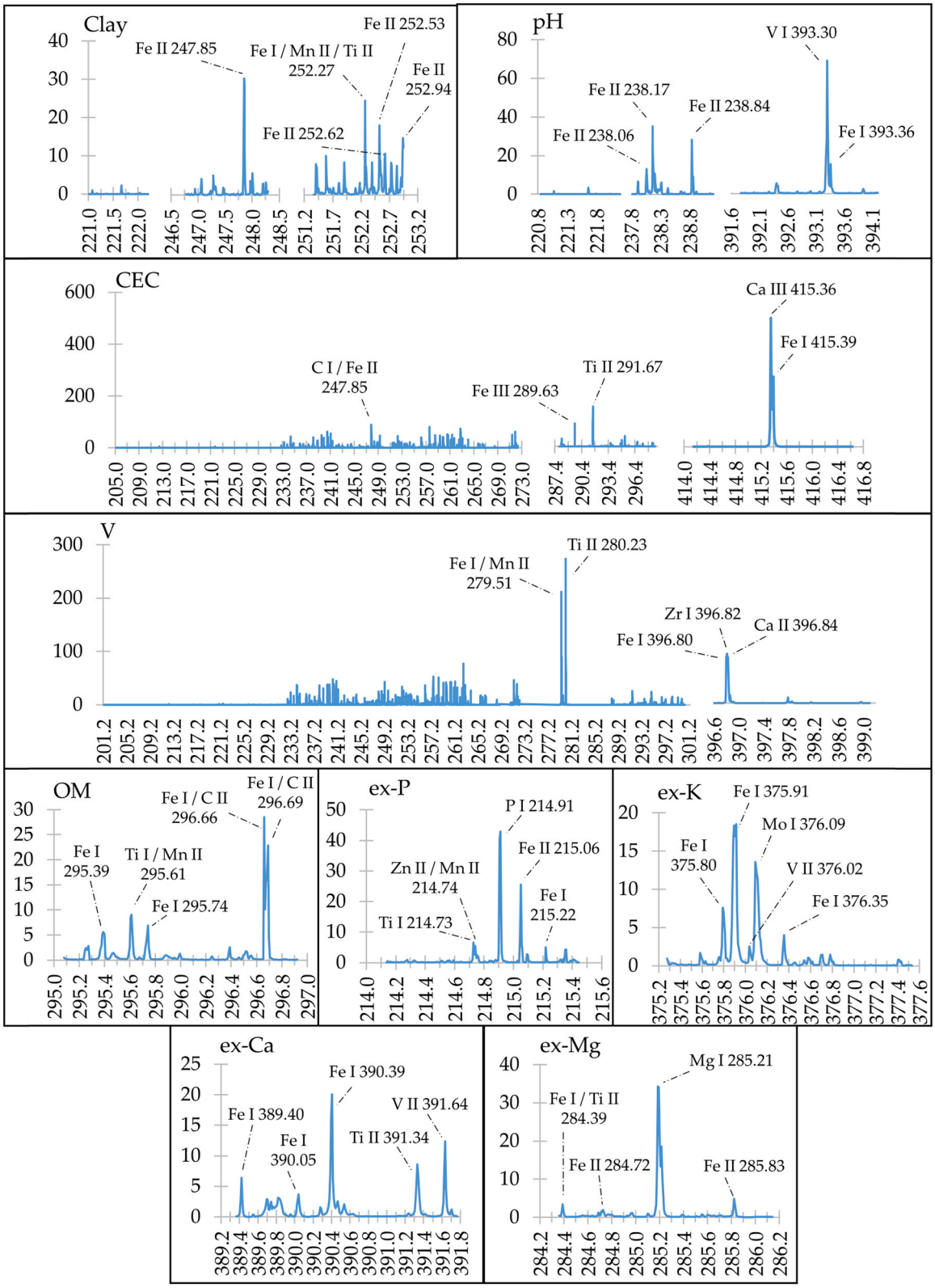


Fig. A1. Variable importance of projection (VIP) scores of the interval of variables selected by the interval successive projection algorithm (iSPA) for the prediction of the key soil fertility attributes. The X-axis of the graphs represent the wavelengths (expressed in nm) of the LIBS spectrum and the Y-axis represents the VIP scores. The top five highest contributing emission lines to the prediction of each soil attribute were also identified.

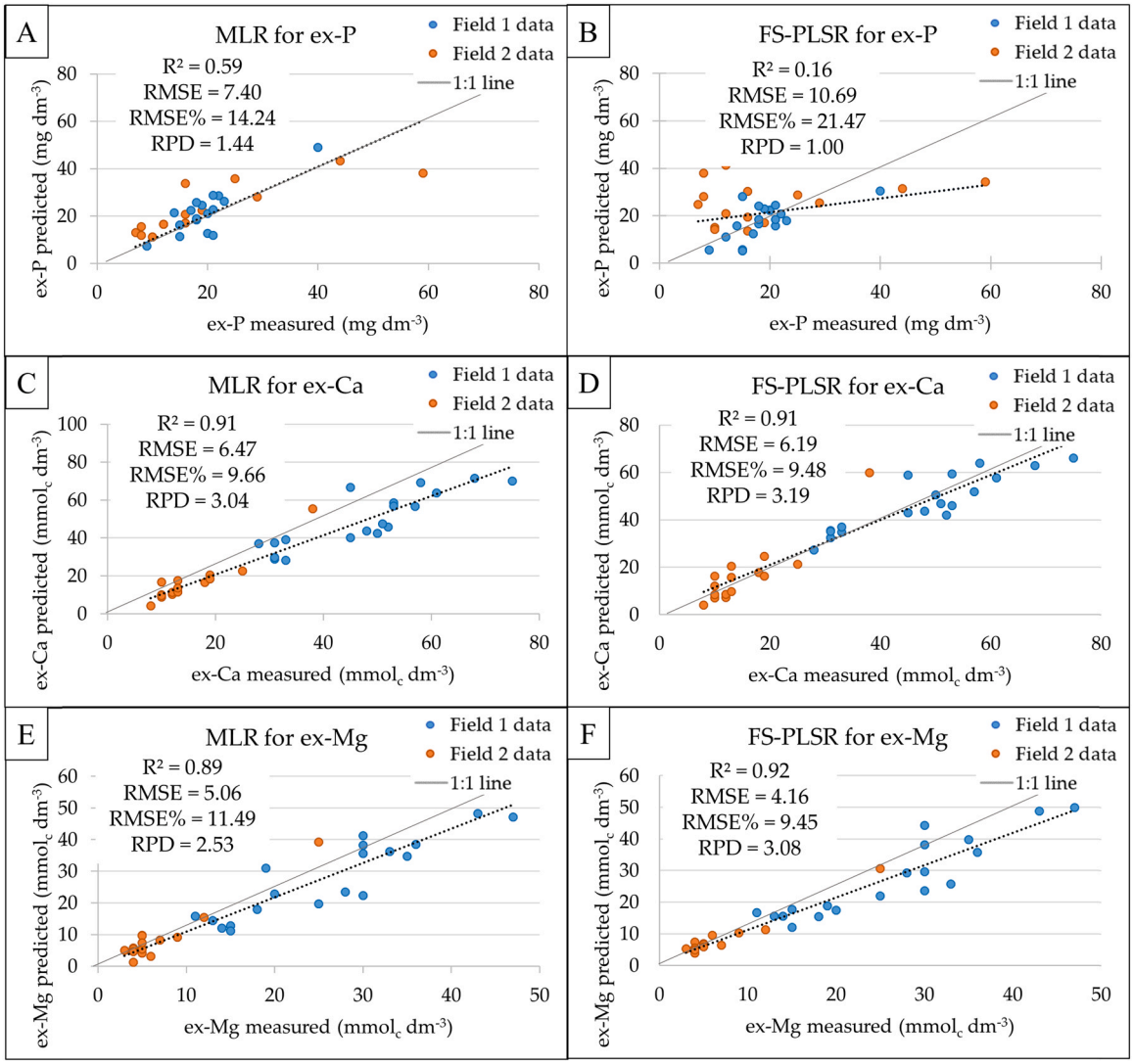


Fig. A2. Scatter plots of measured versus predicted extractable (ex-) P (A, B), Ca (C, D), and Mg (E, F). In the left side (A,C,E), the models were calibrated using multiple linear regression (MLR) calibrated using the 21 selected emission lines. In the right side (B,D,F), the models were calibrated with partial least square regression using the spectra within the region from 200 to 540 nm (FS-PLS). The figures of merit presented correspond to the prediction performance of the models applied on the validation dataset (n = 34).

References

Andrade, R., Faria, W.M., Silva, S.H.G., Chakraborty, S., Weindorf, D.C., Mesquita, L.F., Guilherme, L.R.G., Curi, N., 2020. Prediction of soil fertility via portable X-ray fluorescence (pXRF) spectrometry and soil texture in the Brazilian Coastal Plains. Geoderma 357, 113960. https://doi.org/10.1016/j.geoderma.2019.113960.

Bricklemyer, R.S., Brown, D.J., Turk, P.J., Clegg, S.M., 2013. Improved intact soil-core carbon determination applying regression shrinkage and variable selection techniques to complete spectrum laser-induced breakdown spectroscopy (LIBS). Appl. Spectrosc. 67 (10), 1185–1199. https://doi.org/10.1366/12-06983.

Bricklemyer, R.S., Brown, D.J., Turk, P.J., Clegg, S., 2018. Comparing vis–NIRS, LIBS, and combined vis–NIRS-LIBS for intact soil core soil carbon measurement. Soil Sci. Soc. Am. J. 82, 1482–1496. https://doi.org/10.2136/sssaj2017.09.0332.

Chang, C.W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. Soil Sci. Soc. Am. J. 65, 480–490. https://doi.org/10.2136/sssaj2001.652480x.

Danzer, K., Currie, L.A., 1998. Guidelines for calibration in analytical chemistry. Part I. Fundamentals and single component calibration (IUPAC Recommendations 1998). Pure Appl. Chem. 70 (4), 993–1014. https://doi.org/10.1351/pac199870040993.

Demattê, J.A.M., Dotto, A.C., Bedin, L.G., Sayão, V.M., Souza, A.B., 2019. Soil analytical quality control by traditional and spectroscopy techniques: Constructing the future of a hybrid laboratory for low environmental impact. Geoderma 337, 111–121. https://doi.org/10.1016/j.geoderma.2018.09.010.

Element, C.A.S., 2007. Method 3051A microwave assisted acid digestion of sediments, sludges, soils, and oils. Z. Für. Anal. Chem. 111, 362–366.

Erler, A., Riebe, D., Beitz, T., Löhmannsröben, H.G., Gebbers, R., 2020. Soil Nutrient Detection for Precision Agriculture Using Handheld Laser-Induced Breakdown Spectroscopy (LIBS) and Multivariate Regression Methods (PLS, Lasso and GPR). Sensors 20, 418. https://doi.org/10.3390/s20020418.

Ferreira, E.C., Neto, J.A.G., Milori, D.M., Ferreira, E.J., Anzano, J.M., 2015. Laser-induced breakdown spectroscopy: Extending its application to soil pH measurements. Spectrochim. Acta B. 110, 96–99. https://doi.org/10.1016/j.sab.2015.06.002.

Fontes, M.P.F., 2012. Intemperismo de rochas e minerais. In: Ker, J.C., Curi, N., Schaefer, C.E.G.R., Vidal-Torrado, P. (Eds.), Pedologia: Fundamentos. Sociedade Brasileira de Ciência do Solo, Viçosa, Brazil, pp. 181–205 (In Portuguese).

Gomes, A.A., Galvão, R.K.H., Araújo, M.C.U., Véras, G., Silva, E.C., 2013. The successive projections algorithm for interval selection in PLS. Microchem. J. 110, 202–208. https://doi.org/10.1016/j.microc.2013.03.015.

He, Y., Liu, X., Lv, Y., Liu, F., Peng, J., Shen, T., Zhao, Y., Tang, Y., Luo, S., 2018. Quantitative analysis of nutrient elements in soil using single and double-pulse laser-induced breakdown spectroscopy. Sensors 18, 1526. https://doi.org/10.3390/s18051526.

Hussain, T., Gondal, M.A., Yamani, Z.H., Baig, M.A., 2007. Measurement of nutrients in green house soil with laser induced breakdown spectroscopy. Environ. Monit. Assess. 124, 131–139. https://doi.org/10.1007/s10661-006-9213-x.

Jantzi, S.C., Motto-Ros, V., Trichard, F., Markushin, Y., Melikechi, N., Giacomo, A., 2016. Sample treatment and preparation for laser-induced breakdown spectroscopy. Spectrochim. Acta Part B 115, 52–63. https://doi.org/10.1016/j.sab.2015.11.002.

Kennard, R.W., Stone, L.A., 1969. Computer aided design of experiments. Technometrics 11, 137–148. https://doi.org/10.1080/00401706.1969.10490666.

Lima, T.M., Weindorf, D.C., Curi, N., Guilherme, L.R., Lana, R.M., Ribeiro, B.T., 2019. Elemental analysis of Cerrado agricultural soils via portable X-ray fluorescence spectrometry: Inferences for soil fertility assessment. Geoderma 353, 264–272. https://doi.org/10.1016/j.geoderma.2019.06.045.

Molin, J.P., Tavares, T.R., 2019. Sens. Syst. Mapp. Soil Fertil. Attrib.: Chall., Adv. Perspect. Braz. Trop. Soils Eng. Agric. 39, 126–147. https://doi.org/10.1590/1809-4430-Eng.Agric.v39nep126-147/2019.

Mouazen, A.M., De Baerdemaeker, J., Ramon, H., 2006. Effect of wavelength range on the measurement accuracy of some selected soil constituents using visual-near infrared spectroscopy. J. Infrared Spectrosc. 14 (3), 189–199. https://doi.org/10.1255/jnirs.614.

Mourad, M., Bertrand-Krajewski, J.L., Chebbo, G., 2005. Calibration and validation of multiple regression models for stormwater quality prediction: data partitioning, effect of dataset size and characteristics. Water Sci. Technol. 52, 45–52. https://doi.org/10.2166/wst.2005.0060.

Munnaf, M.A., Nawar, S., Mouazen, A.M., 2019. Estimation of secondary soil properties by fusion of laboratory and on-line measured Vis–NIR spectra. Remote Sens 11 (23), 2819. https://doi.org/10.3390/rs11232819.

Nicolodelli, G., Cabral, J., Menegatti, C.R., Marangoni, B., Senesi, G.S., 2019. Recent advances and future trends in LIBS applications to agricultural materials and their food derivatives: an overview of developments in the last decade (2010–2019). Part I. Soils and fertilizers. Trends Anal. Chem. 115, 70–82. https://doi.org/10.1016/j.trac.2019.03.032.

Niu, C., Cheng, X., Zhang, T., Wang, X., He, B., Zhang, W., Feng, Y., Bai, J., Li, H., 2021. Novel Method Based on Hollow Laser Trapping-LIBS-Machine Learning for Simultaneous Quantitative Analysis of Multiple Metal Elements in a Single Microsized Particle in Air. Anal. Chem. 93 (4), 2281–2290. https://doi.org/10.1021/acs.analchem.0c04155.

Nunes, L.C., Rocha, F.R., Krug, F.J., 2019. Slope ratio calibration for analysis of plant leaves by laser-induced breakdown spectroscopy. Journal of Analytical Atomic Spectrometry. https://doi.org/10.1039/C9JA00270G.

Nunes, L.C., Silva, G.A., Trevizan, L.C., Júnior, D.S., Poppi, R.J., Krug, F.J., 2009. Simultaneous optimization by neuro-genetic approach for analysis of plant materials by laser induced breakdown spectroscopy. Spectrochim. Acta Part B 64, 565–572. https://doi.org/10.1016/j.sab.2009.05.002.

Nunes, L.C., Braga, J.W.B., Trevizan, L.C., Souza, P.F., Carvalho, G.G.A., Júnior, D.S., Poppi, R.J., Krug, F.J., 2010. Optimization and validation of a LIBS method for the determination of macro and micronutrients in sugar cane leaves. J. Anal. . Spectrom. 25, 1453–1460. https://doi.org/10.1039/C003620J.

Ralchenko, Y., Kramida, A., 2020. Development of NIST Atomic Databases and Online Tools. Atoms 8, 56. https://doi.org/10.3390/atoms8030056.

Riebe, D., Erler, A., Brinkmann, P., Beitz, T., Löhmannsröben, H.G., Gebbers, R., 2019. Comparison of Calibration Approaches in Laser-Induced Breakdown Spectroscopy for Proximal Soil Sensing in Precision Agriculture. Sensors 19, 5244. https://doi.org/10.3390/s19235244.

Santos, Humberto Gonçalves, et al., 2018. Brazilian Soil Classification System, 5th. EMBRAPA, Brasília.

Schäefer, C.E.G.R., Fabris, J.D., Ker, J.C., 2008. Minerals in the clay fraction of Brazilian latosols (oxisols): a review. Clay Min. 43, 137–154. https://doi.org/10.1180/claymin.2008.043.1.11.

Senesi, G.S., Senesi, N., 2016. Laser-induced breakdown spectroscopy (LIBS) to measure quantitatively soil carbon with emphasis on soil organic carbon. A review. Anal. Chim. Acta 938, 7–17. https://doi.org/10.1016/j.aca.2016.07.039.

Senesi, G.S., 2020. A Survey of Recent Applications of Laser-induced Breakdown Spectroscopy (LIBS) to Soil Analysis. Int. J. Earth Environ. Sci 5, 172. https://doi.org/10.15344/2456-351X/2020/172.

Silva, Y.J.A.B., Nascimento, C.W.A., Biondi, C.M., 2014. Comparison of USEPA digestion methods to heavy metals in soil samples. Environ. Monit. Assess. 186, 47–53. https://doi.org/10.1007/s10661-013-3354-5.

Silva, S.H.G., Teixeira, A.F.D.S., Menezes, M.D.D., Guilherme, L.R.G., Moreira, F.M.D.S., Curi, N., 2017. Multiple linear regression and random forest to predict and map soil properties using data from portable X-ray fluorescence spectrometer (pXRF). Ciênc. Agrotec. 41 (6), 648–664. https://doi.org/10.1590/1413-70542017416010317.

Silva, F.C.D.S., Molin, J.P., 2018. On-the-go tropical soil sensing for pH determination using ion-selective electrodes. Pesqui. Agropecuária Bras. 53, 1189–1202. https://doi.org/10.1590/s0100-204×2018001100001.

Soil Survey Staff, 2010. Keys to soil taxonomy. United States Department of Agriculture (USDA), Soil Conservation Service: Washington, DC.

Takahashi, T., Thornton, B., 2017. Quantitative methods for compensation of matrix effects and self-absorption in Laser Induced Breakdown Spectroscopy signals of solids. Spectrochim. Acta B 138, 31–42. https://doi.org/10.1016/j.sab.2017.09.010.

Tavares, T.R., Nunes, L.C., Alves, E.E.N., Almeida, E., Maldaner, L.F., Krug, F.J., Carvalho, H.W.P., Molin, J.P., 2019. Simplifying sample preparation for soil fertility analysis by X-ray fluorescence spectrometry. Sensors 19, 5066. https://doi.org/10.3390/s19235066.

Tavares, T.R., Molin, J.P., Nunes, L.C., Alves, E.E.N., Melquiades, F.L., de Carvalho, H.W.P., Mouazen, A.M., 2020a. Effect of X-ray tube configuration on measurement of key soil fertility attributes with XRF. Remote Sens 12 (6), 963. https://doi.org/10.3390/rs12060963.

Tavares, T.R., Mouazen, A.M., Alves, E.E.N., dos Santos, F.R., Melquiades, F.L., Pereira de Carvalho, H.W., Molin, J.P., 2020b. Assessing soil key fertility attributes using a portable X-ray fluorescence: A simple method to overcome matrix effect. Agronomy 10 (6), 787. https://doi.org/10.3390/agronomy10060787.

Van Raij, B., Andrade, J.C., Cantarela, H., Quaggio, J.A., 2001. Análise química para avaliação de solos tropicais, IAC, Campinas, Brazil. (In Portuguese).

Viscarra Rossel, R.A., Bouma, J., 2016. Soil sensing: A new paradigm for agriculture. Agric. Syst. 148, 71–74. https://doi.org/10.1016/j.agsy.2016.07.001.

Villas-Boas, P.R., Romano, R.A., Franco, M.A.M., Ferreira, E.C., Ferreira, E.J., Crestana, S., Milori, D.M.B.P., 2016. Laser-induced breakdown spectroscopy to determine soil texture: a fast analytical technique. Geoderma 263, 195–202. https://doi.org/10.1016/j.geoderma.2015.09.018.

Xu, X., Du, C., Ma, F., Shen, Y., Zhou, J., 2019a. Fast and Simultaneous Determination of Soil Properties Using Laser-Induced Breakdown Spectroscopy (LIBS): A Case Study of Typical Farmland Soils in China. Soil Syst. 3, 66. https://doi.org/10.3390/soilsystems3040066.

Xu, X., Du, C., Ma, F., Shen, Y., Wu, K., Liang, D., Zhou, J., 2019b. Detection of soil organic matter from laser-induced breakdown spectroscopy (LIBS) and mid-infrared spectroscopy (FTIR-ATR) coupled with multivariate techniques. Geoderma 355, 113905. https://doi.org/10.1016/j.geoderma.2019.113905.

Xu, D., Zhao, R., Li, S., Chen, S., Jiang, Q., Zhou, L., Shi, Z., 2019c. Multi-sensor fusion for the determination of several soil properties in the Yangtze River Delta, China. Eur. J. Soil Sci. 70, 162–173. https://doi.org/10.1111/ejss.12729.

Yu, K.Q., Zhao, Y.R., Liu, F., He, Y., 2016. Laser-induced breakdown spectroscopy coupled with multivariate chemometrics for variety discrimination of soil. Sci. Rep. 6, 1–10. https://doi.org/10.1038/srep27574.