SOIL PHOSPHORUS TEST USING A LOW-COST SPECTROPHOTOMETER AND MACHINE LEARNING

Grégory O. Mayrink[1*], Daniel M. de Queiroz[1], Andre L. de F. Coelho[1], Domingos S. M. Valente[1]

[1*]Corresponding author. Universidade Federal de Viçosa/Viçosa - MG, Brasil.
E-mail: gregorymairinky@hotmail.com | ORCID ID: https://orcid.org/0000-0001-9845-7233

KEYWORDS

precision agriculture, proximal spectroscopy sensors, embedded systems, ion exchange resins, machine learning by multiple linear regression.

ABSTRACT

Phosphorus concentration is one of the main attributes determined in laboratory analyses of soil samples collected in the field. The objective is to develop a soil phosphorus test using a low-cost spectrophotometer and a machine learning technique. For reflectance measurements, a low-cost system consisting of a Sparkfun AS7625x spectrophotometer and an Arduino Uno is used. Ion exchange resins under standard saturated solutions and modified conditions are used to extract phosphorus ions from the soil samples. Reflectance and phosphorus concentrations determined by the reference method are used in the training and testing of a machine learning. A modification procedure of the ion-exchange resin saturation solution allows the establishment of a strong correlation between the reflectance in 18 spectral bands and P concentration of the soil samples. The obtained model uses five reflectance of the modified resins at wavelengths of 410, 460, 560, 705, and 645 nm to predict the phosphorus concentration. This model presents an R^2t accuracy of 0.97 in the training stage with an R^2v of 0.96, RMSEv of 9.05, and ratio of prediction to deviation) of 3.81 in the test step.

INTRODUCTION

Phosphorus (P) availability in tropical soils is one of the main factors that influence crop yield. After being absorbed from the soil solution by the roots, this chemical element participates in basic physiological processes in the plant, such as photosynthesis and respiration. Determining the soil P content is critical for prescribing the correct dose of phosphate fertilizer to be applied to the soil (Prezotti et al., 2013).

Conventional laboratory analyses used to assess available P and other physical and chemical soil attributes are generally time consuming and expensive. The total cost and time spent characterizing soil fertility depends on the size of the crop field and adopted agricultural management system (Mayrink et al., 2019). In precision agriculture (PA), a higher number of soil samples per unit of area is essential compared to conventional agriculture. To reduce labor usage and costs, farmers generally select inadequate sampling densities. This decision can lead to errors in the generation of fertilizer prescription maps, thus reducing the profitability of crops (Resende & Coelho, 2017).

Proximal sensing is a promising technique for assessing soil fertility in crops, especially in those that adopt PA (Kopačková et al., 2017). This technique is based on reflectance spectroscopy and uses sensors to measure the electromagnetic energy reflected by the surface in different spectra (Frei, 2019). Owing to the speed of data collection and lower cost, this technique stands out for enabling a significant increase in the sample density. In addition, spectroscopy enables the analysis of different soil attributes by using a single sensor (Rossel et al., 2006).

Research has shown that analysis of the spectral signature of the soil enables the estimation of texture, water and organic matter content, cation exchange capacity, and soil chemical attributes, such as total carbon, pH, and concentrations of nitrate, calcium, potassium, magnesium,

and phosphorus (Jiang et al., 2017; Soriano-Disla et al., 2017; Mayrink et al., 2019; Nisha & Prasad, 2020; Pätzold et al., 2020). However, it is necessary to use appropriate spectral signature analysis techniques to estimate soil attributes from spectroscopic data (Pinheiro et al., 2017). Among these techniques, machine-learning algorithms have great potential (Molin & Tavares 2019).

Several spectroscopy-based proximal sensors have been developed and marketed for rapid determination of soil nutrient availability (Rossel et al., 2011; Molin & Tavares 2019; Adamchuk & Rossel, 2010; Guerrero et al., 2021). ClorofiLOG, GreenSeeker, Crop Circle, Crop Meter, and Crop Spec are examples of commercial spectroscopy-based sensors used to assess nitrogen stress in plants. The Optic Mapper sensor used to evaluate the texture and organic matter in the soil is pulled by a vehicle. The data collected by these sensors can be influenced by various soil and plant attributes, making it difficult to accurately determine a specific attribute (Rossel et al. 2011).

Solid-phase extraction (SPE) is a promising low-cost alternative for increasing the accuracy of determining the concentration of a specific soil chemical attribute. Using SPE and ion exchange resins as pre-concentration matrices, Mayrink et al. (2019) obtained satisfactory results with R^2 greater than 0.90 and RPD (ratio of prediction to deviation) greater than 2 for prediction models of P, Ca, Mg, and K concentrations in soil samples. As resins are sensitive to the same forms of P assimilated by plants (inorganic), the concentration of inorganic P has a high correlation with the amount of the nutrient extracted by the plant (Gonçalves et al., 2012; Oliveira et al., 2015). In addition to being commercial and consolidated products for this application, ion-exchange resins eliminate the interference of other physical attributes of the soil and concentrate the chemical elements that are diluted in the soil. Although Mayrink et al. (2019) demonstrated the potential of using SPE to determine soil chemical attributes, this determination depended on a spectrophotometer, a high-cost laboratory equipment, to measure the reflectance of the samples at different wavelengths.

To increase sensitivity and selectivity, chromophore reagents can be used to modify the ion-exchange resins used in SPE. These reagents provide a greater molar absorptivity of the resin in specific electromagnetic spectra (Silva et al., 2016). Resin modifications can occur by substitution mechanisms inside the matrix between the ions of interest and the reagent used, as well as by the formation of complexes (ternary complexes) between the matrix, interfering ions, and reagents. The use of these reagents can facilitate the determination of the concentration of chemical attributes of the soil using the reflectance of soil samples from a low-cost spectrophotometer. Therefore, the objective of the present work is to develop a method for predicting P concentration in soil samples using a low-cost spectrophotometer sensor using standard and modified ion exchange resins with chromophore reagents and a machine learning algorithm.

MATERIAL AND METHODS

The development of a method for predicting P concentration in soil samples comprised of: i) preparing soil samples; ii) extracting P from the soil samples using a standard ion exchange resin test method and a modified method by addition of a chromophore reagent during extraction by the ion exchange resin; ii) measuring the reflectance of the samples of standard and modified resins after extraction using a low-cost spectrophotometer sensor; iii) determining the P concentration of samples using the reference method in soil laboratory analysis; and iv) training and testing a machine learning algorithm to predict P concentrations using the resin reflectance, measured using a low-cost spectrophotometer, and the soil P values obtained by the reference method of laboratory soil analysis.

Soil sample preparation

To develop and test the P concentration prediction method, 200 samples from different soil types in the state of Minas Gerais, Brazil were used. The samples were provided by the Soil Fertility Routine Laboratory of the Federal University of Vicosa, (Vicosa, MG, Brazil). Using the attributes pH, cation exchange capacity (CTC), organic matter (OM), and remaining phosphorus (P-rem), determined in the laboratory using the respective reference methods (Silva et al., 2009) The samples were divided into four classes of soil buffering capacity: very low (VL), low (L), medium (M), and high (H).

The soil samples were air-dried, macerated, sieved through 2 mm holes, and homogenized. Each soil sample was divided into three fractions. Two fractions were used in the P extraction procedure with standard and modified ion-exchange resins for further reflectance measurements using a low-cost spectrophotometer sensor. The third fraction was used to determine P concentration using the reference method for laboratory soil analysis.

Phosphorus extraction from soil samples with standard and modified ion exchange resin

For soil sample P ion extraction, PUROLITE A400 anionic resin (Purolite Corporation, Philadelphia, USA) was used under two saturation conditions. In the first condition (referred to as standard resin), the resin was prepared according to the methodology proposed by Van Raij & Quaggio (2001). In the second condition (referred to as modified resin), the resins were prepared by modifying the saturation step using a chromophore reagent.

The resin preparation process consists of two basic steps: pre-conditioning and preparation for use. The preconditioning step promoted the expansion of the active sites of the resins. In this step, the resins were first passed through a sieve with a mesh opening of 0.5 mm. They were then mixed in a saturated electrolyte solution (with approximately 5 g of KH_2PO_4, 4 g of $CaCl_2.2H_2O$, and 2 g of $MgSO_4.7H_2O$) and left in contact for one week with occasional stirring. After the contact period, the resins were washed with deionized water in a batch system.

As the name suggests, the preparation for the use step consists of saturating the resins with a solution containing standard ions, leaving them ready for use. In this step, using a batch system, a volume of resin sufficient for use in the experiment was mixed with 5 volumes of a 1 mol L^{-1} $NaHCO_3$ solution at pH 8.5, left in contact for 1 h, with occasional stirring. Subsequently, the resins were transferred to an elution column, through which 20 volumes of 1 mol L^{-1} $NaHCO_3$ solution were passed. Finally, 20 volumes of water were slowly passed through to remove the excess ions.

In the second condition, the modification involved replacing sodium bicarbonate solution ($NaHCO_3$) with a solution of chromophore reagents. The solution was prepared using ammonium molybdate, bismuth subcarbonate, sulfuric acid, and ascorbic acid (vitamin C). The proportion of each reagent used was defined according to the chemical analysis methods manual for soil fertility assessment (Silva et al., 1998). The ion-exchange resin was mixed with the modified saturation solution (prepared 2 h in advance) in the proportion of one volume of resin to five volumes of solution (Fig. 1a). The excess saturated solution contained in the resin was removed using deionized water. Saturated resins under standard and modified conditions were placed in individual permeable polyester bags. In each bag, 2.5 cm³ of resin was inserted (Fig. 1b).

For each of the 200 soil samples, the P ion extraction process consisted of adding each permeable polyester bag filled with 2.5 cm³ of saturated resin into a container with 2.5 cm³ of soil sample in suspension with 25 mL of deionized water. The container was stirred in a circular-horizontal motion at a stirring speed of 220 rpm for 16 h at room temperature (25 °C). After the extraction process, each resin bag was removed from contact with the soil solution. The bag was opened and the resins were dried by centrifugation, and placed in cylindrical containers to form tablets with a diameter of 2.5 cm (Fig. 1c).


(a)   (b)   (c)
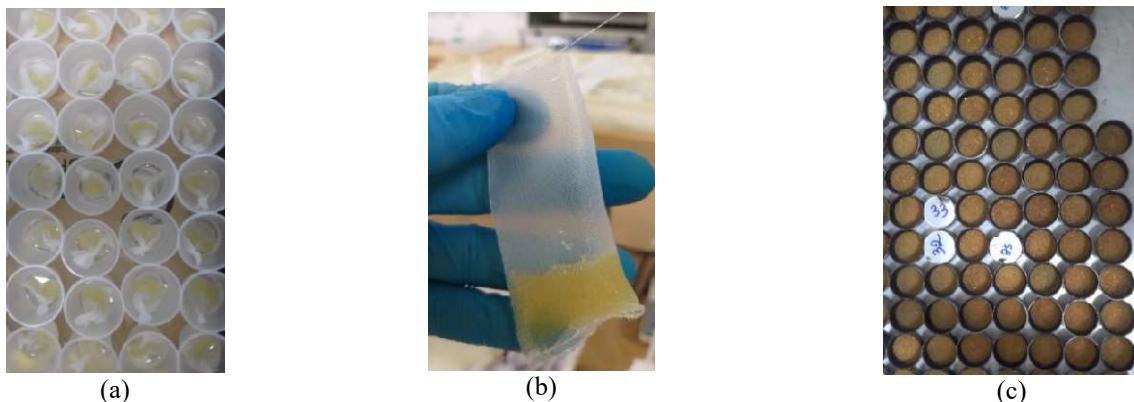
FIGURE 1. Procedure for saturation of the ion exchange resin and phosphorus extraction: (a) Saturation of the resin with standard or modified solutions, (b) Insertion of the saturated resin in permeable polyester bags, and (c) Formation of cylindrical tablets with the resin after extraction of the phosphorus ions from soil samples.

Measuring the reflectance using a low-cost spectrophotometer sensor

To measure the reflectance of the resin tablets, a SparkFun AS7265x spectrophotometer (SparkFun Electronics, USA) was used. AS7265x, a low-cost, open-hardware sensor, measures the reflectance in the 410–940 nm electromagnetic spectrum in 18 spectral bands. The three capture elements AS72651, AS72652, and AS72653 of the sensor sensitize to the ultraviolet, visible, and infrared spectra and measured reflectance in the bands 410, 435, 460, 485, 510, 535, 560, 585, 610, 645, 680, 705, 730, 760, 810, 860, 900, and 940 nm. The sensor contains three individual light-emitting diode (LED) elements that emit white (color temperature of 5700 K), ultraviolet (405 nm), and infrared (875 nm) light. A computer program, written in C++ language and executed using the Arduino Uno microcontroller board (Arduino, Italy), was developed and used to activate the LEDs, capture the reflectance measurements on the three capture elements, and save them in a file for further analysis. A white barium sulfate ($BaSO_4$) plate was used as the reference standard for AS7265x sensor calibration. This plate was assumed to have unity reflectance (1.0) in the 18 spectral bands.

Determining the phosphorus content in the soil sample using the reference method

The reference method for determining the P concentration in 200 soil samples was proposed by Van Raij & Quaggio (2001). In this method, P from a soil sample is extracted using ion exchange resins. To facilitate the understanding of the text, the concentration of P determined using the reference method is referred to as Presin. The Presin was determined by measuring the absorbance of the supernatant liquid in an acrylic cuvette with an optical path of 0.5 mm. This was done using a bench spectrophotometer (Azzota, SM100 UV-VIS), measuring the absorbance at a wavelength of 725 nm, corresponding to the visible spectral region.

Machine learning algorithm to predict soil P using the reflectance of standard and modified resins

For each of the two conditions, standard and modified resins, a dataset containing the reflectance of the 200 tablets in the 18 bands and the respective Presin values was generated. First, the Pearson correlations between the reflectances, and between the reflectance and Presin value were determined. To predict the P concentration of the soil samples using reflectance (in the text, the predicted P concentration is denoted as $\hat{P}$), machine learning models based on multiple linear regression were developed and evaluated.

To create, train, and test the models, a computer program was developed in Python 3.9, using the statsmodels package, version 0.12.2, (Seabold & Perktold, 2010). The scikit-learn package, version 0.24.1, (Pedregosa et al., 2011) was used to divide the dataset into training and testing data and to obtain metrics for evaluating the accuracy of the models. Using the k-fold procedure, the dataset was divided into 10 parts to use 90% for training and 10% for testing. An iterative process was implemented to combine all 10 possible combinations of data pieces for training and testing. The metrics used to evaluate the accuracy of the models were the coefficient of determination of the training step (R^2t), coefficient of determination of the testing step

(R^2v), square root of the squared error for the testing step (RMSEv, Equation 1), and the ratio of prediction to deviation in the testing step (RPD, Equation. 2).

$$RMSE_v = \sqrt{\frac{\sum_{i=1}^{n_p}\left(Presin_i - \widehat{P}_i\right)^2}{n_p}} \qquad (1)$$

Where:

RMSE$_v$ is the root mean square error of the model in the testing step;

Presin$_i$ is the sample P concentration of the test set measured using the reference method;

$\widehat{P}_i$ is the P concentration of a sample of the test set estimated using the machine learning model, and

n_p is the number of samples in the test set.

$$RPD = \frac{SD}{RMSE_v} \qquad (2)$$

Where:

RPD is the ratio of prediction to deviation in the testing step, and

SD is the standard deviation of the Presin$_i$ values.

Two strategies for obtaining machine-learning models were evaluated. In the first strategy, the models were created using an iterative process that tested models with one to twelve independent variables (reflectance in each spectral band) to calculate $\widehat{P}$. Initially, models with only one variable were tested. The variable that resulted in the highest coefficient of determination for the training stage was selected as a part of the model. At each iteration, a new variable was added to the model by selecting the variable that maximized the coefficient of determination of the training stage. The criterion for selecting the number of variables used in the model stabilized the metrics used to assess accuracy. Once the variables that resulted in the best model performance were defined, a new model was generated to determine $\widehat{P}$ using 80% of the data for training and 20% for testing.
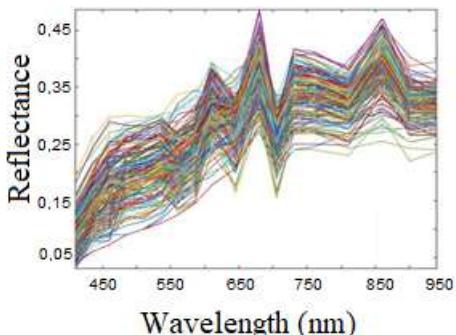
In the second strategy, the model was obtained using n reflectances that presented the highest Pearson correlation with the Presin, where n denotes the number of variables used in the first strategy. The model was trained and validated using 80% of the data for training and 20% for determining $\widehat{P}$ using 80% data for training and 20% for testing. All procedures for creating, testing, and validating the models for the two strategies were performed for the dataset obtained using standard and modified resins.

RESULTS AND DISCUSSION

When comparing the reflectance of the resin tablets in the 18 spectral bands under the two saturation conditions (standard, Fig. 2a, and modified, Fig. 2b), similar behavior was identified at certain wavelengths. This similar behavior occurred in the 810, 860, 900, and 940 nm bands, which correspond to the near-infrared region. This similarity is partly due to the physical and chemical constituents of the resin used, such as the polystyrene matrix and quaternary ammonium functional groups.
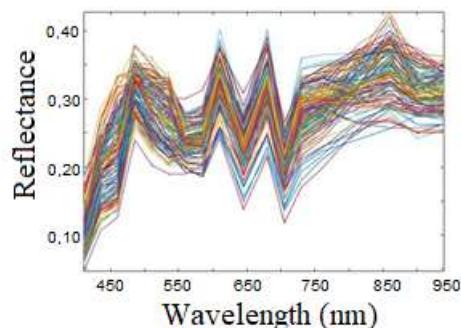

FIGURE 2. Reflectance spectra of resin tablets after P extraction: a) Saturated using standard solution proposed by Van Raij & Quaggio (2001); and b) Saturated using modified solution (chromophore reagent). Each line represents the reflectance of a tablet in the 18 spectral bands.

The differences in the behavior shown in Fig. 2 are mainly due to the solutions used to saturate the resin. For the saturated resins, a sodium bicarbonate solution was used according to the standard methodology. For the modified methodology, a solution of chromophore reagents (ammonium molybdate, bismuth subcarbonate, sulfuric acid, and ascorbic acid) is used for saturation. When present in the resin matrix, ammonium molybdate reacts with the P ions extracted from the soil, forming a phosphomolybdic complex. In acidic media, this complex promotes blue coloration in the resins (Bortolon & Gianello, 2010; Mumbach et al., 2018) proportional to the concentration of P ions. Thus, the bluish color acquired by the modified resins after the extraction of P from the soil was responsible for the change in reflectance in the spectral bands at 485, 510, and 730 nm. The last band is strongly correlated with the presence of the phosphomolybdic complex (Nagul et al., 2015).

**Correlation analysis among the reflectance values**

The use of chromophore reagent solution altered Pearson's correlations among the reflectances at 18 wavelengths. It also affected the correlation between the reflectance at each wavelength and Presin value (obtained using soil laboratory analysis for determination of soil P value). In the correlation matrix for the reflectance measured in the standard resin, there was a negligible correlation (in the range of 0–0.3) between the reflectance and Presin values (Fig. 3a). For the correlations between reflectances, it was observed that the behavior of strong

correlations (in the range of 0.7–0.9) for reflectances in closer spectra and weaker or negligible correlations for reflectances in spectra that were more different from each other. In the correlation matrix for the reflectance measured using the modified resin, there was a strong correlation (greater than 0.9) between the reflectance and Presin values (Fig. 3a). Strong or very strong correlations were observed among the reflectances in the 18 spectra.


(a)

(b)
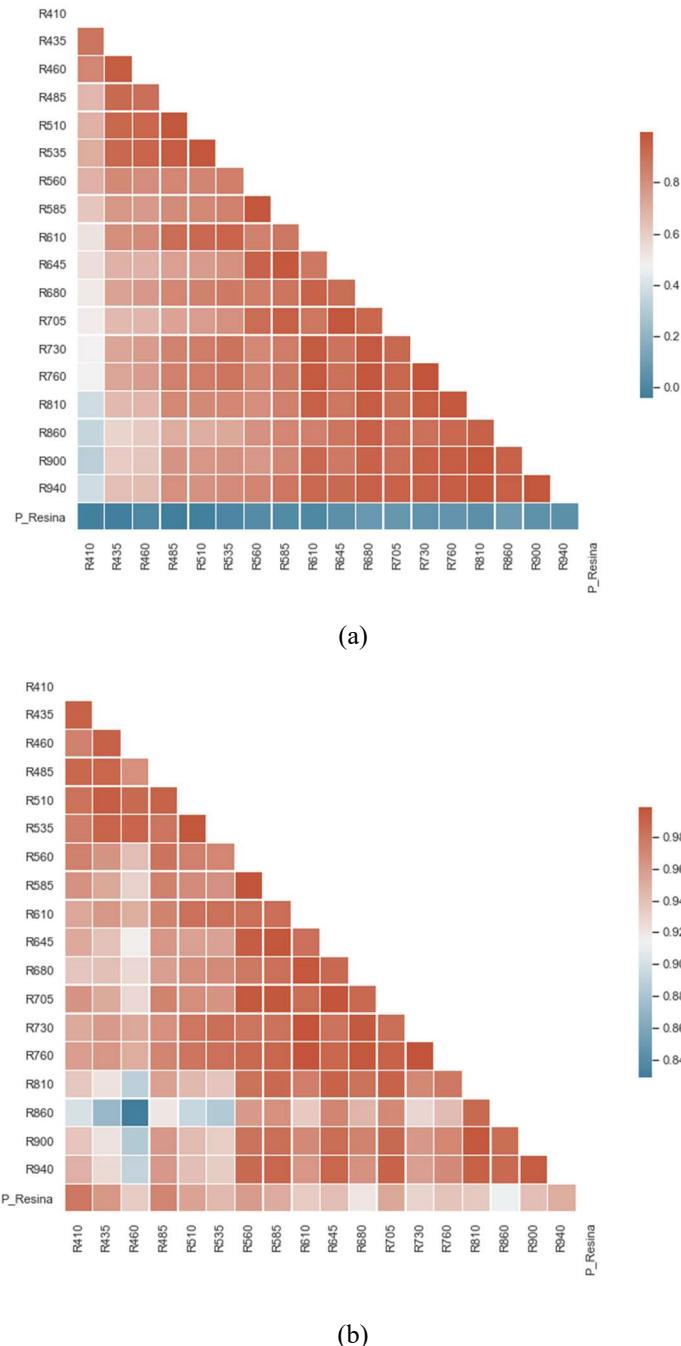
FIGURE 3. Pearson's correlation matrix between the reflectance in different wavelengths with Presin values (identified in the plot as P_resina) for (a) Saturated resin for standard solution and (b) Saturated resin for modified solution.

Model for predicting soil P using resin reflectance

For the first model generation strategy, the soil P prediction model using reflectances from modified resin tablets obtained from a solution of chromophore reagents (Fig. 4) showed superior accuracy compared to the model generated with reflectances from standard resin tablets (Fig. 5). The model developed using standard resin reflectance did not allow the prediction of phosphorus concentration in soil samples because of the negligible correlation observed between Presin values and reflectance. Based on the correlation matrix analysis (Fig. 3a and b), it can be stated that the use of the chromophore reagent solution was responsible for establishing the correlation between Presin and the reflectance of the resins.

Grégory O. Mayrink, Daniel M. de Queiroz, Andre L. de F. Coelho, et al.

(a)

(b)

(c)
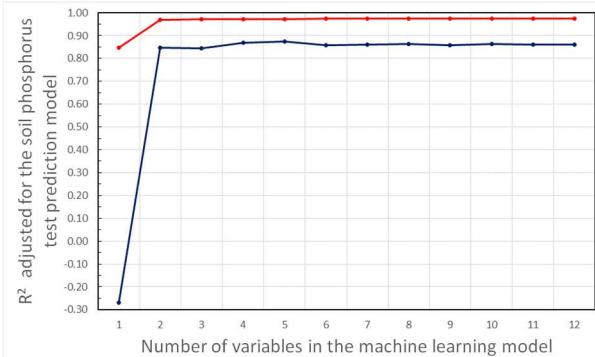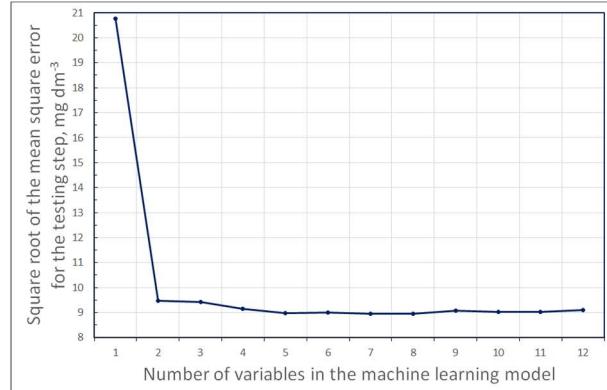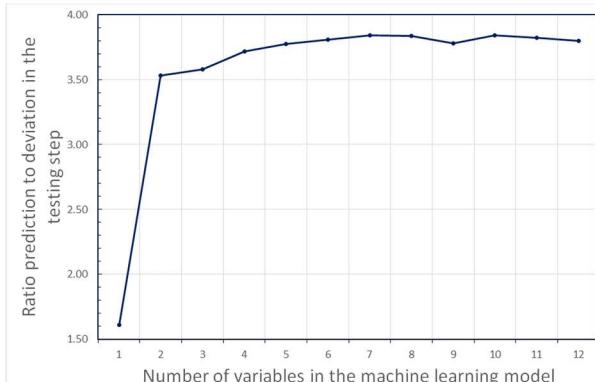
Number of variables in the prediction model	Variables in the machine learning model using modified solution (with a chromophore reagent)											
1	R410											
2	R410	R460										
3	R410	R460	R560									
4	R410	R460	R560	R705								
5	R410	R460	R560	R705	R645							
6	R410	R460	R560	R705	R645	R940						
7	R410	R460	R560	R705	R645	R940	R860					
8	R410	R460	R560	R705	R645	R940	R860	R485				
9	R410	R460	R560	R705	R645	R940	R860	R485	R510			
10	R410	R460	R560	R705	R645	R940	R860	R485	R510	R760		
11	R410	R460	R560	R705	R645	R940	R860	R485	R510	R760	R610	
12	R410	R460	R560	R705	R645	R940	R860	R485	R510	R760	R610	R585

(d)
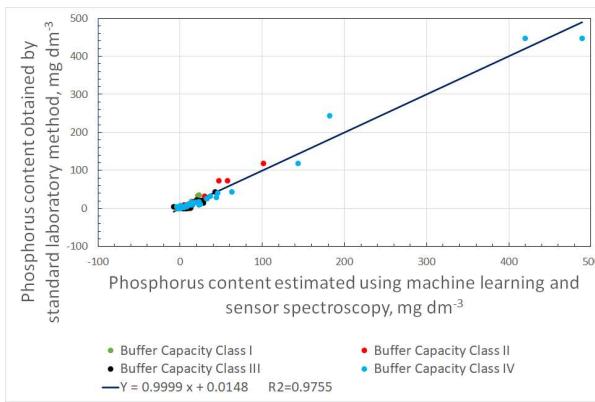
(e)
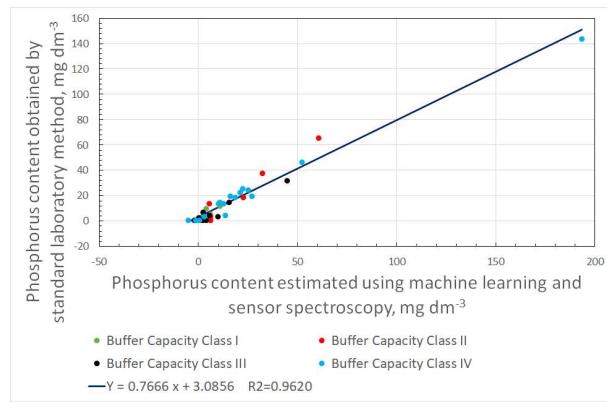
(f)

FIGURE 4. Metrics of the phosphorus prediction model using modified resin reflectance (a) R^2 of the model adjusted for the training and testing steps; (b) Square root of the mean square error of the model in the testing step; (c) Ratio prediction to deviation for the testing step; (d) Variables that produced the best fit of the machine learning model as a function of the number of variables in the model, in red is the highlighted number of ideal variables to predict the phosphorus concentration; (e) P concentration predicted by the model, in the training step, with the optimal number of variables compared to the P concentration obtained using the reference method; (f) P concentration predicted by the model, in the testing step, with the optimal number of variables compared to the P concentration obtained using the reference method.

(a)

(b)

(c)
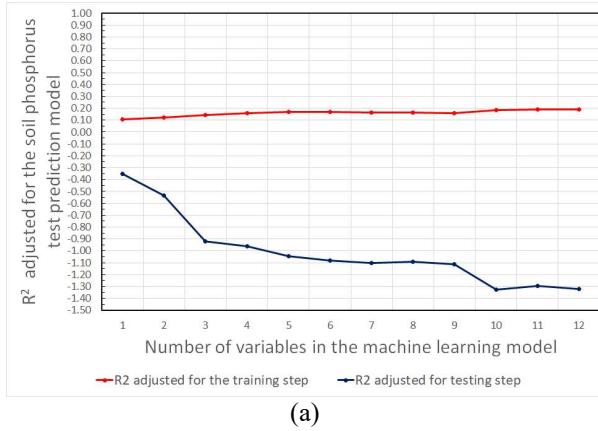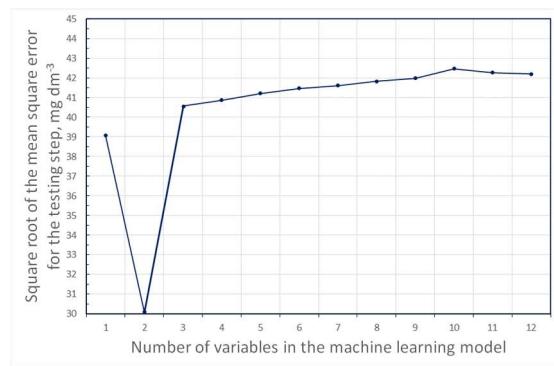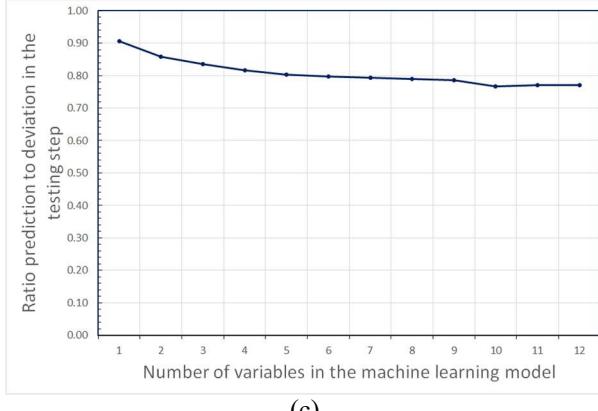
Number of variables in the prediction model	Variables in the machine learning model for standard solution											
1	R860											
2	R860	R485										
3	R860	R485	R680									
4	R860	R485	R680	R610								
5	R860	R485	R680	R610	R810							
6	R860	R485	R680	R610	R810	R900						
7	R860	R485	R680	R610	R810	R900	R435					
8	R860	R485	R680	R610	R810	R900	R435	R460				
9	R860	R485	R680	R610	R810	R900	R435	R460	R705			
10	R860	R485	R680	R610	R810	R900	R435	R460	R705	R645		
11	R860	R485	R680	R610	R810	R900	R435	R460	R705	R645	R730	
12	R860	R485	R680	R610	R810	R900	R435	R460	R705	R645	R730	R410

(d)
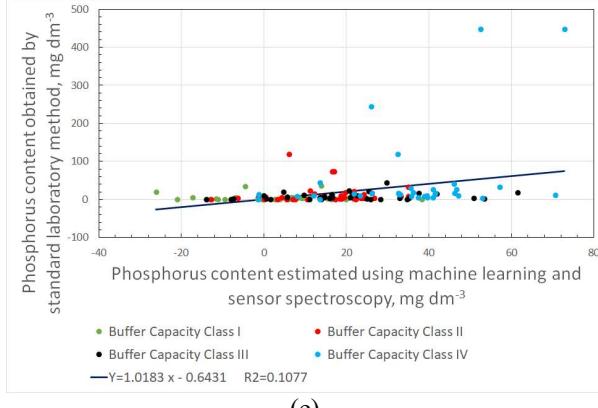
(e)
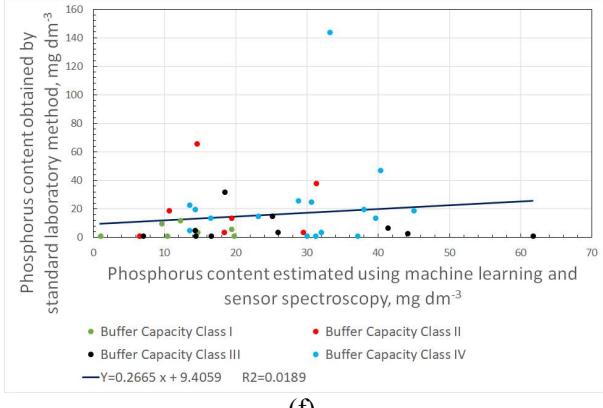
(f)

FIGURE 5. Metrics of the phosphorus prediction model using reflectance of the standard resins (a) R^2 of the model adjusted for the training and testing steps; (b) Square root of the mean square error of the model for the testing step; (c) Ratio prediction to deviation for the testing step; (d) Variables that produced the best fit of the machine learning model as a function of the number of variables in the model, in red is the highlighted ideal number of variables to predict the phosphorus concentration (e) P concentration predicted by the model, in the training step, with an optimal number of variables compared to the P concentration obtained using the reference method; (f) P concentration predicted by the model, in the testing step, with an optimal number of variables compared with P concentration obtained using the reference method

For the modified resin reflectance, an increase in the model accuracy was observed by increasing the number of model variables to five (Fig. 4). Models developed using more than five variables, that is, reflectance at five wavelengths, did not exhibit an increase in accuracy. The prediction of the P concentration can be obtained using resins saturated with a modified solution and reflectances in the bands 410 nm (in the violet band), 460 nm (in the blue band), 560 nm (in the green band), 705 nm (in the near-infrared band), and 645 nm (in the red band). The model used for testing (Fig. 5f) is given by [eq. (3)].

$$\hat{P} = 391{,}70\,R410 - 118{,}51\,R460 - 88{,}87\,R560 + 518{,}21\,R705 - 380{,}50\,R645 \tag{3}$$

Where:

$\hat{P}$ is the predicted phosphorus concentration, in mg dm^{-3}, and

R410, R460, R560, R705, and R645 are the reflectances of saturated resins in the modified solution at wavelengths of 410, 460, 560, 705, and 645 nm, respectively.

A comparison of the values of the ratio of prediction to deviation in the testing step (RPD, Fig. 5c) with the recommended values of this metric proposed by Williams (2014) was made. According to the metric, an RPD value above 3.5 for soil analysis, as obtained in models with two or more variables, is considered to have a good fit. This RPD value is similar to that obtained by Kweon & Maxton (2013) for soil organic matter determination based on reflectance at wavelengths of 660 and 940 nm. In contrast, the RPD values for the model generated with the reflectance of the standard resins (Fig. 5c), in which the RPD value was less than one for all fitted models, is considered a poor fit, and the recommendation is not to use this model in predictions.

In the second strategy for machine learning model generation, P prediction models were obtained using five reflectances from the standard and modified resins, which showed the highest correlation with the Presin values. The models generated using the first strategy performed slightly better than those generated using the second strategy (Table 1). That is, combining variables in an iterative process to maximize the accuracy of the prediction allows models with a higher performance to be obtained.

TABLE 1. Comparison of performance metrics for the models generated in the two strategies, for standard and modified resin. R^2t: coefficient of determination in the testing step. R^2v: coefficient of determination in the testing step. RMSEv: root mean square error of the model in the testing step. RPD: ratio of prediction to deviation in the testing step.

Parameters	Strategy 1		Strategy 2	
	Standard Resina	Modified Resin	Standard Resin	Modified Resin
Selected wavelength (nm)	860; 485; 680; 610; 810	410; 460; 560; 705; 645	860; 680; 705; 760; 900	410; 435; 485; 510; 705
R^2t	0.11	0.97	0.10	0.85
R^2v	0.01	0.96	-0.54	0.83
RMSEv	41.10	9.05	40.02	9.75
RPD	0.81	3.81	0.87	3.44

One of the drawbacks of machine learning based on multiple linear regression was that $\hat{P}$ negative values were found under certain conditions (Figs. 5e and 5f). However, it was possible to verify that for the modified resins, when the model predicted a negative value $\hat{P}$, this value was closer to zero and occurred when the Presin value (obtained using the soil laboratory analysis method) was closer to zero. As these situations occurred when the Presin value in the samples was low, negative values could be replaced by zero. It is important to note that according to Embrapa (2015), the concentration of adequate P is higher than 20, 15 and 8 mg dm^{-3}, for soils of sandy, medium, and clayey texture classes, respectively. Phosphorus concentrations below 4 mg dm^{-3} are considered low for all soil textural classes and, therefore, should receive a higher phosphorus dosage. An alternative would be to use other machine learning models; however, the simplicity of the linear models and good fit obtained were the reasons for not trying other alternative machine learning algorithms.

Although the soil samples were classified according to their buffer capacity class, the model for calculating $\hat{P}$ developed was generalized to all four soil buffer capacity classes. As future work, novel studies with a greater number of soil samples are suggested, making it possible to develop, test, and validate specific models for each soil buffer capacity class. In the proposed P concentration estimation procedure, the time required for the extraction of P ions using resins was 16 h. This time was selected because it was the extraction time adopted in the reference method for P determination. However, for the proposed procedure to be fast, further research is needed to evaluate the influence of extraction time on the accuracy of the P prediction models.

CONCLUSIONS

A low-cost and open-design embedded system for proximal spectroscopy was developed using an ATMEGA328P microcontroller and a SparkFun AS7265x spectrophotometer sensor. Using machine learning models, the reflectances measured with the embedded system allowed soil phosphorus tests to be performed. The procedure of modifying the saturation solution of the ion exchange resin, replacing the standard solution with a chromophore solution, allowed the establishment of a strong correlation between the reflectance in 18 spectral bands and the phosphorus concentration of the soil samples. The machine learning model, based on multiple linear regression, showed greater accuracy ($R^2t = 0.97$, $R^2v = 0.96$, RMSEv = 9.05, RPD = 3.81) using five reflectances of the modified resins (410, 460, 560, 705, and 645 nm). In addition, the interactive process strategy was obtained by selecting variables that maximize the accuracy of the model.

ACKNOWLEDGMENTS

REFERENCES

Adamchuk VI, Rossel RA (2010) Development of on-the-go proximal soil sensor systems. Proximal soil sensing 15-28.

Bortolon L, Gianello C (2010) Fósforo extraído pela solução de Mehlich-1 determinado por colorimetria e ICP em solos do sul do Brasil. Revista Brasileira de Ciência do Solo 34:263-268.

Embrapa (2015) Guia prático para interpretação de resultados de análises de solo. Aracaju, Embrapa Tabuleiros Costeiros.

Frei RW (2019) Espectroscopia de refletância difusa resolução de problemas ambientais. Imprensa CRC.

Gonçalves GK, Bortolon L, Meurer EJ, Gonçalves DRN, Sousa ROD, Fagundes SM (2012) Extratores de fósforo para o arroz irrigado em solos adubados com fosfato natural reativo. Revista de ciências agroveterinárias 11(3):196-204.

Guerrero A, Neve S de, Mouazen AM (2021) Current sensor technologies for in situ and on-line measurement of soil nitrogen for variable rate fertilization: A review. Advances in Agronomy 168:1-38.

Jiang Q, Li Q, Wang X, Wu Y, Yang X, Liu F (2017) Estimation of soil organic carbon and total nitrogen in different soil layers using VNIR spectroscopy: Effects of spiking on model applicability. Geoderma 293:54-63.

Kopačková V, Ben-Dor E, Carmon N, Notesco G (2017) Modelling diverse soil attributes with visible to longwave infrared spectroscopy using PLSR employed by an automatic modelling engine. Remote Sensing 9(2):134.

Kweon G, Maxton C (2013) Soil organic matter sensing with an on-the-go optical sensor. Biosystems engineering 115(1):66-81.

Mayrink GO, Valente DSM, Queiroz DM, Pinto FAC, Teofilo RF (2019) Determination of chemical soil properties using diffuse reflectance and ion-exchange resins. Precision Agriculture 20(3):541-561.

Molin JP, Tavares TR (2019) Sensor systems for mapping soil fertility attributes: Challenges, advances, and perspectives in brazilian tropical soils. Engenharia Agrícola 39:126-147.

Mumbach GL, Oliveira DAD, Warmling MI, Gatiboni LC (2018) Quantificação de fósforo por Mehlich 1, Mehlich 3 e Resina Trocadora de Ânions em solos com diferentes teores de argila. Revista Ceres 65:546-554.

Nagul EA, McKelvie ID, Worsfold P, Kolev SD (2015) The molybdenum blue reaction for the determination of orthophosphate revisited: Opening the black box. Analytica chimica acta 890:60-82.

Nisha S, Prasad S (2020) Absorption spectrometric macronutrients review of soil health during taro crop production. Applied Spectroscopy Reviews 55(5):378-392.

Oliveira CMB de, Gatiboni LC, Ernani PR, Boitt G, Brunetto G (2015) Capacidade de predição da disponibilidade de fósforo em solo com aplicação de fosfato solúvel e natural. Científica 43(4):413-419.

Pätzold S, Leenen M, Frizen P, Heggemann T, Wagner P, Rodionov A (2020) Predicting plant available phosphorus using infrared spectroscopy with consideration for future mobile sensing applications in precision farming. Precision Agriculture 21(4):737-761.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Duchesnay MPE (2011) Scikit-learn: Machine learning in Python. the Journal of machine Learning research 12:2825-2830.

Pinheiro ÉF, Ceddia MB, Clingensmith CM, Grunwald S, Vasques GM (2017) Prediction of soil physical and chemical properties by visible and near-infrared diffuse reflectance spectroscopy in the central Amazon. Remote Sensing 9(4):293.

Prezotti LC, Oliveira JA, Gomes JA, Dadalto GG (2013) Manual de recomendação de calagem e adubação para o Estado do Espírito Santo: 5ª aproximação. Viçosa, UFV.

Resende AV de, Coelho AM (2017) Amostragem para mapeamento e manejo da fertilidade do solo na abordagem de agricultura de precisão. Sete Lagoas, Embrapa Milho e Sorgo.

Rossel RV, Adamchuk VI, Sudduth KA, McKenzie NJ, Lobsey C (2011) Proximal soil sensing: An effective approach for soil measurements in space and time. Advances in agronomy 113:243-291.

Rossel RV, Walvoort DJJ, McBratney AB, Janik LJ, Skjemstad JO (2006) Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131(1-2):59-75.

Seabold S, Perktold J (2010) Statsmodels: econometric and statistical modeling with python. In Python in Science Conference. Austin, Proceedings…

Silva FC da, Abreu MF de, Perez DV, Eira PA da, Abreu CA de, Raij BV, Gianello C, Coelho AM, Quaggio JA, Tedesco MJ, Silva CA, Barreto WO (2009) Métodos de análises químicas para avaliação da fertilidade do solo. Rio de Janeiro, Embrapa Solos.

Silva FC da, Eira PA da, Barreto WDO, Pérez DV, Silva CA (1998) Manual de métodos de análises químicas para avaliação da fertilidade do solo. Rio de Janeiro, Embrapa Solos.

Silva JPRD, Cruz GFB, Cassella RJ, Pacheco WF (2016) Desenvolvimento de um sistema de análise por injeção em fluxo para pré concentração de cd sobre resina de xad-4 modificada com fenilfluorona e determinação por espectrometria de absorção atômica. Química Nova 39:561-566.

Soriano-Disla JM, Janik LJ, Allen DJ, McLaughlin MJ (2017) Evaluation of the performance of portable visible-infrared instruments for the prediction of soil properties. Biosystems Engineering 161:24-36.

Van Raij B, Quaggio JA (2001) Determinação de fósforo, cálcio, magnésio e potássio extraídos com resina trocadora de íons. In: van Raij B, Andrade JC, Cantarella H, Quaggio JA, et al. Ed. Análise química para avaliação da fertilidade de solos tropicais. Campinas, IAC, p189-199.

Williams P (2014) The RPD statistic: a tutorial note. NIR news 25(1):22-26.