Prediction of soil organic matter and clay contents by near-infrared spectroscopy - NIRS

Bruno Pedro Lazzaretti[1] Leandro Souza da Silva[2] (iD) **Gerson Laerson Drescher[2]*** (iD)
André Carnieletto Dotto[3] (iD) **Darines Britzke[2] José Laerte Nörnberg[4]**

[1]Programa de Pós-graduação em Ciência do Solo, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brasil.
[2]Departamento de Solos, Centro de Ciências Rurais (CCR), Universidade Federal de Santa Maria (UFSM), 97105-900, Santa Maria, RS, Brasil.
E-mail: gersondrescher@gmail.com. *Corresponding author.
[3]School of Molecular and Life Sciences, Curtin University, Bentley, WA, Australia.
[4]Departamento de Tecnologia e Ciência dos Alimentos, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brasil.

ABSTRACT: *Among the soil constituents, special attention is given to soil organic matter (SOM) and clay contents, since, among other aspects, they are key factors to nutrient retention and soil aggregates formation, which directly affect the crop production potential. The methods commonly used for the quantification of these constituents have some disadvantages, such as the use of chemical reactants and waste generation. An alternative to these methods is the near-infrared spectroscopy (NIRS) technique. The aim of this research is to evaluate models for SOM and clay quantification in soil samples using spectral data by NIRS. A set (n = 400) of soil samples previously analyzed by traditional methods were used to generate a NIRS calibration curve. The clay content was determined by the hydrometer method while SOM content was determined by sulfochromic solution. For calibration, we used the original spectra (absorbance) and spectral pretreatment (Savitzky-Golay smoothing derivative) in the following models: multiple linear regression (MLR), partial last squares regression (PLSR), support vector machine (SVM) and Gaussian process regression (GPR). The curve validation was performed with the SVM model (best performance in the calibration based on R² and RMSE) in two ways: with 40 random samples from the calibration set and another set with 200 new unknown samples. The soil clay content affects the predictive ability of the calibration curve to estimate SOM content by NIRS. Validation curves showed poorer performance (lower R² and higher RMSE) when generated from unknown samples, where the model tends to overestimate the lower levels and to underestimate the higher levels of clay and SOM. Despite the potential of NIRS technique to predict these attributes, further calibration studies are still needed to use this technique in soil analysis laboratories.*
Key words: *calibration, validation, mathematical models, spectral pretreatment.*

Predição dos teores de matéria orgânica e argila do solo por espectroscopia no infravermelho próximo – NIRS

RESUMO: *Dentre os constituintes do solo, especial atenção é voltada aos teores de argila e de matéria orgânica do solo (MOS), pois, entre outros aspectos, são determinantes para retenção de nutrientes e a formação de agregados no solo, os quais afetam diretamente o potencial produtivo das culturas. Os métodos mais comumente utilizados para quantificação destes constituintes apresentam algumas desvantagens, como o uso de reagentes químicos e a geração de resíduos. Uma alternativa a estes métodos é o uso da espectroscopia no infravermelho próximo (near infrared spectroscopy – NIRS). O objetivo deste trabalho é avaliar modelos de quantificação dos teores de argila e de MOS em amostras de solo utilizando dados espectrais por meio da técnica NIRS. Foram utilizadas 400 amostras de solos com amplitude nos teores de MOS e argila para geração de uma curva de calibração. A argila foi determinada pelo método do densímetro e a MOS por meio da solução sulfocrômica. Para calibração, utilizou-se os espectros originais (absorbância) e com pré-tratamento espectral (Savitzky-Golay derivative) das 400 amostras nos seguintes modelos: multiple linear regression (MLR), partial last squares regression (PLSR), support vector machine (SVM) e Gaussian process regression (GPR). A validação da curva foi realizada com o modelo que apresentou melhor desempenho na calibração (SVM) de duas maneiras: com 40 amostras aleatórias oriundas daquelas utilizadas na calibração e com outras 200 novas amostras desconhecidas. O teor de argila das amostras de solo afeta a capacidade preditiva da curva de calibração da estimativa do teor de MOS pelo NIRS. A validação das curvas apresentou pior desempenho (menor R² e maior RMSE) quando feita a partir de amostras desconhecidas, cujo modelo tende a superestimar os teores mais baixos e subestimar os teores mais elevados de argila e MOS com a curva gerada. Apesar do potencial de predição destes atributos via NIRS, outros estudos de calibração ainda são necessários para que esta técnica possa ser utilizada em laboratórios de análises de solos.*
Palavras-chave: *calibração, validação, modelos matemáticos, pré-tratamento espectral.*

INTRODUCTION

The soil organic matter (SOM) and clay contents are key factors in the soil nutrient dynamics and crop production. As a result, these attributes are used in soil fertility evaluation and fertilizer recommendation systems, such as the clay content for phosphorus availability interpretation by

Mehlich-1 method and the SOM content as an index for nitrogen fertilizer recommendation (CQFS-RS/SC, 2016). There are different procedures for this soil constituent's quantification, and the hydrometer method is commonly used for clay content analysis (TEIXEIRA et al., 2017), while SOM is generally analyzed by wet digestion method and quantified by colorimetric titration (TEDESCO et al., 1995).

Despite the effectiveness of these methods, there are some disadvantages regarding its use in routine soil analysis laboratories. The commonly used methods for clay and SOM determination use hazardous chemical reagents (such as sodium hydroxide for clay content quantification and sulfochromic solution for SOM analysis) and are relatively laborious and time-consuming to execute. Therefore, the scientific research has sought for alternative methods that can be used without such disadvantages, but with accuracy in these attributes' estimation. Among the potential methods for predicting clay and SOM contents in soil analysis laboratories, is the near-infrared spectroscopy (NIRS). This technique brings several advantages, such as speed in the analysis execution, easier sample preparation, and requires only a small soil sample to perform readings (FERRARESI et al., 2012). Furthermore, the NIRS is not harmful to the environment, since it uses no chemical extractors (VANDRAME et al., 2015) and the technique is easy to perform.

The NIRS technique presents several advantages and has potential use for soil analysis, although an effective calibration with traditional methods used for SOM and clay quantification is necessary (VAN VUUREN et al., 2006). The relationship between the soil spectrum from NIRS readings and its chemical/physical attributes is complex (STEVENS et al., 2013); therefore, the database used in the method calibration must display a wide SOM range, which allows the setting of a robust model and enables assertive and accurate analysis of soil samples. Hence, studies regarding the development of new strategies for the calibration and subsequent validation of NIRS should be performed, enabling the use of this technique for soil testing laboratories with high predictive capacity. We are looking for a model to predict clay and SOM contents with NIRS and hypothesized that soil clay content affects SOM predictive capacity by NIRS and that a better calibration curve is adjusted using different soil clay classes.

The aim of this study was to evaluate models for soil clay and SOM quantification using spectral data by NIRS technique. Therefore, a calibration curve for soil clay and SOM quantification was developed; the influence of soil clay content on the calibration curve and the SOM predictive capacity was evaluated; and the calibration curves were validated with unknown soil samples.

MATERIALS AND METHODS

This research was carried out with samples from the Soil Analysis Laboratory (SAL) of the Federal University of Santa Maria (UFSM). A set of 400 soil samples was selected based on the four clay classes defined by Liming and Fertilization Manual for the Rio Grande do Sul and Santa Catarina States (CQFS-RS/SC, 2016). Thus, the set of samples was composed of 100 samples with clay content > 60% (C1); 100 samples with clay content between 41 and 60% (C2); 100 samples with clay content between 21 and 40% (C3), and 100 samples with clay content ≤ 20% (C4).

The clay and SOM contents were quantified according to the analytical procedures described by TEDESCO et al. (1995). The samples spectroscopic analysis was performed in a spectrophotometer NIRSystem 5000 (Foss NIRSystem® Inc., Silver Spring, MD, USA) coupled to a computer with Vision® software. Readings were taken in the reflectance mode and the data were stored as absorbance (log 1/reflectance). For NIRS readings, the soil samples (oven-dried, grounded, and sieved to 2.0 mm openings) were placed in the machine cell and subjected to analysis in duplicate, with readings at every 2 nm and wavelengths range from 1100 to 2500 nm. Thereafter, spectra were imported and analyzed using R (R CORE TEAM, 2019) program and the graphical interface AlradSpectra (DOTTO et. al., 2019).

As the original soil spectra were provided in absorbance, the set of spectra was subjected to Savitzky-Golay derivative (SGD) treatment using 5 smoothing points, with polynomial order and derivative order, both of first degree. The pretreatment was performed to remove reference signals that have no importance to the measures, including variations from equipment instability, sample heterogeneity, scan failures, noise occurrence, among other influences. The mathematical description of this model is presented in the following equation:

$$x_j = \frac{1}{N} \sum_{-m}^{m} c_h \, x_{j+m} \qquad (1)$$

where x_j is the new value, N is the coefficient of normalization, m is neighboring values of j, and C_n is pre-computed coefficients that depend on the choice of polynomial and derivative order.

For the models' calibration, the 400 spectra, with and without SGD spectral pretreatment, were

submitted to four algorithms: multiple linear regression (MLR), partial last squares regression (PLSR), support vector machine (SVM), and Gaussian process regression (GPR). The calibration curves evaluation was performed by comparing the determination coefficients (R^2) and the root mean square error (RMSE). Adjustments of each model parameters were kept as in AlradSpectra interface default.

The validation was performed with the model that presented the best calibration in two ways. First, we used a set of 40 samples (10 for each soil clay class, randomly selected) obtained from the set samples used for model calibration. Subsequently, the validation was performed with a set of 200 new unknown samples provided by SAL-UFSM (50 for each soil clay class). The last validation test showed the model's ability to predict the clay and SOM contents in unknown samples. This

evaluation was performed considering the Pearson correlation coefficient value between the clay and SOM values quantified by SAL-UFSM and NIRS predicted values, as well as considering the RMSE value obtained for each attribute. The significance level ($p < 0.05$) of the correlation coefficient was evaluated by the Student's t test.

RESULTS AND DISCUSSION

Calibration models for clay and soil organic matter

The original spectra of NIRS (absorbance) and the pretreated SGD spectra of the set of 400 soil samples were associated with clay and SOM contents by various mathematical calibration models (Table 1). In general, the best results for both clay and SOM contents were obtained when calibration was carried out from

Table 1 - Determination coefficients (R^2) and root mean square error (RMSE) values obtained in the models' calibration for clay and soil organic matter (SOM) contents.

Models		Clay	SOM	SOM - C1	SOM - C2	SOM - C3	SOM - C4
		---without spectra pretreatment---					
MLR	R^2	0.86	0.67	0.87	0.81	0.69	0.84
	RMSE %	8.13	0.68	0.31	0.49	0.57	0.40
PLSR	R^2	0.91	0.76	0.98	0.87	0.68	0.86
	RMSE %	6.38	0.58	0.12	0.40	0.58	0.38
Linear SVM	R^2	0.85	0.68	0.65	0.79	0.66	0.76
	RMSE %	8.40	0.67	0.52	0.54	0.60	0.50
SVM Radial	R^2	0.86	0.68	0.76	0.71	0.83	0.90
	RMSE %	8.06	0.67	0.43	0.64	0.43	0.32
GPR Linear	R^2	0.79	0.60	0.67	0.55	0.61	0.71
	RMSE %	9.91	0.75	0.50	0.78	0.66	0.55
GPR Radial	R^2	0.80	0.64	0.42	0.60	0.79	0.67
	RMSE %	9.65	0.74	0.73	0.76	0.62	0.62
		---with spectra pretreatment (Savitzky-Golay smoothing derivative)---------------------------------------					
MLR	R^2	0.81	0.67	0.82	0.82	0.69	0.84
	RMSE %	9.49	0.68	0.37	0.48	0.57	0.40
PLSR	R^2	0.93	0.77	0.99	0.84	0.66	0.92
	RMSE %	5.86	0.57	0.09	0.46	0.60	0.29
Linear SVM	R^2	0.95	0.91	0.99	0.99	0.98	0.99
	RMSE %	4.71	0.36	0.08	0.11	0.14	0.10
SVM Radial	R^2	0.99	0.99	1.00	0.99	0.99	0.99
	RMSE %	2.01	0.12	0.08	0.14	0.10	0.09
GPR Linear	R^2	0.94	0.90	0.99	0.97	0.96	0.98
	RMSE %	5.08	0.38	0.09	0.20	0.21	0.16
GPR Radial	R^2	0.93	0.86	0.91	0.86	0.91	0.91
	RMSE %	6.18	0.50	0.37	0.56	0.48	0.42

MLR: multiple linear regression, PLSR: partial last squares regression, SVM: support vector machine and GPR: Gaussian process regression; SOM: soil organic matter, SOM - C1: samples with clay content > 60%, SOM - C2 samples with clay content between 60 and 40%, SOM - C3: samples with clay content between 40 and 21%, SOM - C4: samples with clay content ≤ 20%.

pre-processed spectra since the variations arising from equipment instability, sample heterogeneity, failures in the scan or due to noise occurrence are corrected/smoothed with the SGD procedure.

For clay content, the SVM model with pretreated spectra showed the best calibration fit with $R^2 = 0.99$ and RMSE = 2.01%. SILVA et al. (2017), working with 84 samples and using the PLSR mathematical model, found $R^2 = 0.84$ while the RMSE value was higher, up to 8.63%. The divergence between RMSE values reported for different models can be ascribed to the sample set size, intrinsic characteristics of each sample, as well as the mathematical parameters of the model.

For SOM content, SVM model with pretreated spectra also showed the best result, with $R^2 = 0.99$ and RMSE = 0.12%. Similar results were reported with organic carbon calibration using the SVM model by SÁ et al. (2010), obtaining $R^2 = 0.92$ and RMSE = 0.14%. This behavior indicated that the SVM model has a good predictive capacity and can be used to estimate the clay and SOM contents by NIRS in soil samples.

In order to verify the clay content influence on the mathematical models' calibration for SOM prediction, we evaluated the performance of each model within each texture class (Table 1). The best results were also obtained with the SVM model using pretreated spectra ($R^2 = 1.00$, 0.99, 0.99, and 0.99 and RMSE = 0.08%, 0.14%, 0.10%, and 0.09% for C1, C2, C3, and C4, respectively). Previous studies indicated that soil particle size can exert an influence on soil spectral response (SOUSA JR. et al., 2008). FELIX et al. (2016) observed that samples sieved to 2.0 mm openings had worse calibration ($R^2 = 0.57$ and standard error = 3.63) compared to samples with particle size below 0.2 mm ($R^2 = 0.86$ and error standard = 2.09). Thus, the calibration model for SOM prediction may be performed using samples from different textural classes, since the particle size uniformity of the sample is maintained.

According to the R^2 and RMSE data in the present study, the SVM model showed the best performance in the calibration process, both for clay and SOM contents, as well for SOM content within each texture class. Thus, SVM model was selected and then used as a standard model in the validation step.

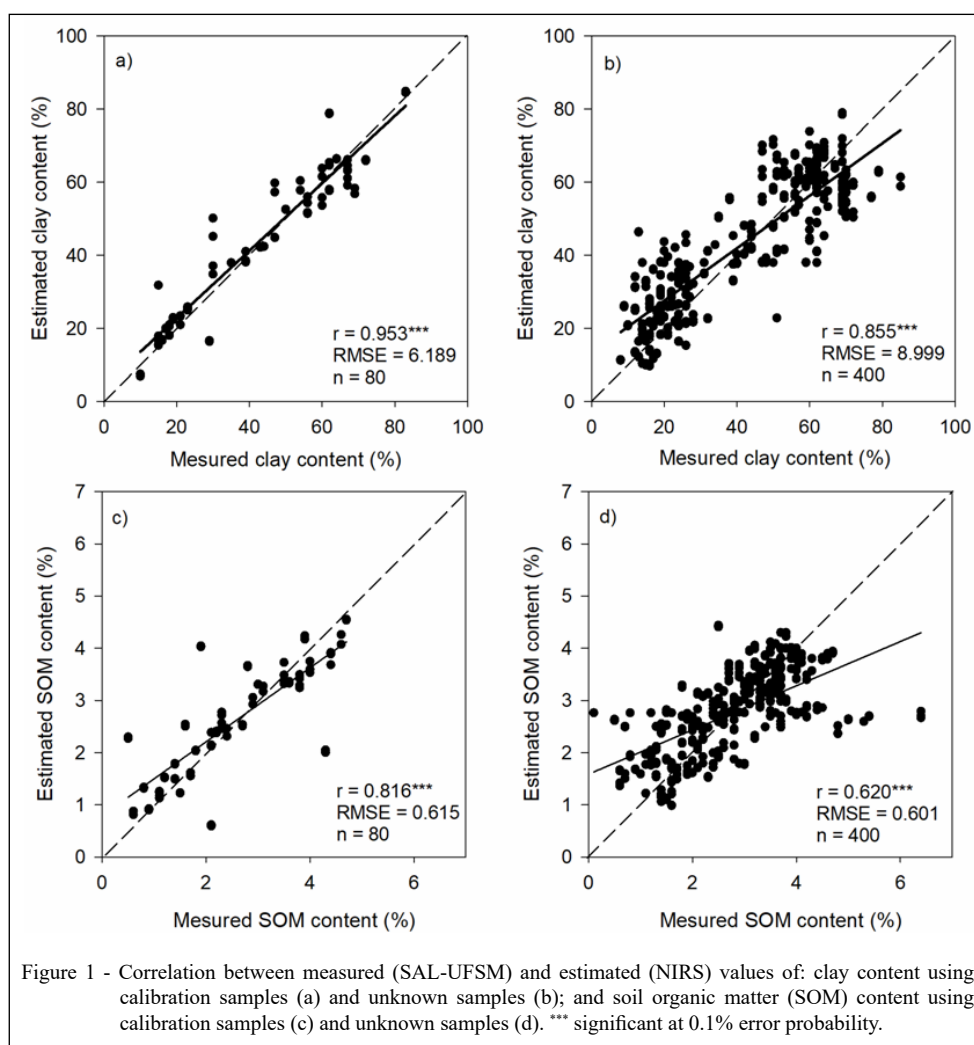*Validation of NIRS calibration model to estimate clay content*

There was a high correlation (r = 0.95) between clay content values estimated by NIRS and those determined by the hydrometer method in random samples (n = 80) from the same set used to generate the calibration curve (Figure 1a). The slope of the relationship between these two variables was also very close to the ideal, which demonstrated the ability of this predictive model to quantify the clay content in those samples. This behavior is important and functional for soil surveys. In such case, only a part of the soil survey samples can be analyzed in the laboratory with the standard methods and used to generate a specific calibration curve with NIRS, while the remained samples can be estimated by NIRS, saving time and resources.

However, despite the model's high correlation, some samples showed an expressive deviation between measured and estimated values. This deviation, measured by the RMSE, indicated an average error of ± 6.19%. Thus, some samples may exhibit a larger discrepancy between measured and estimated values, compromising the results of the analysis.

When the clay estimative was made for a set (n = 400) of unknown samples (i.e., samples that were not used in the calibration process) the correlation coefficient was lower (r = 0.85) (Figure 1b), but statistically high. However, there is a different correlation slope compared to the ideal slope, overestimating lower clay contents and underestimating higher clay contents. Therefore, when considering some samples deviation, the values may change the soil texture class for some samples. For instance, there are samples with 20% clay content measured by the hydrometer method that had the clay content expressively overestimated by NIRS, ranging from 23.0% up to 43.7%. With such prediction error, a sample with soil texture classified as C4 could be classified as C3 or even C2. Likewise, a soil sample with 70% clay content measured by the hydrometer method (i.e., C1) could be classified as C2 due to the prediction error, which can directly affect the phosphorus fertilizer recommendation for crops (CQFS-RS/SC, 2016).

Similar positive results using SVM model for predicting clay content were reported by JACONI et al. (2019). These authors concluded that clay content had better prediction than silt and sand contents, confirming a greater influence of clay particles in the spectral properties of the soil samples in the near-infrared range. Furthermore, similar behavior was observed by ARAÚJO (2013) ($R^2 = 0.81$ and RMSE = 11.00%) using SVM model for predicting clay content in soil samples from Goiás, Minas Gerais, Mato Grosso do Sul and São Paulo States, and SILVA et al. (2017) for Roraima state soils ($R^2 = 0.84$ and RMSE = 8.63%) using the PLSR model. This behavior highlights the ability of NIRS technique for predicting clay content in different soil types while presenting a potential use for routine laboratories.

Figure 1 - Correlation between measured (SAL-UFSM) and estimated (NIRS) values of: clay content using calibration samples (a) and unknown samples (b); and soil organic matter (SOM) content using calibration samples (c) and unknown samples (d). [***] significant at 0.1% error probability.

Validation of NIRS calibration model to estimate the SOM levels

There was a high correlation (r = 0.82) between the SOM values determined by the routine colorimetric method and the values estimated by NIRS when using a set (n = 80) from the same soil samples used to generate the calibration curve (Figure 1c). The slope of the relationship between these variables is not ideal, indicating that the model slightly overestimates lower SOM contents and somewhat underestimates higher SOM contents.

When the SOM estimation was performed for a set (n = 400) of unknown samples (i.e., that were not used in the calibration process) there was a lower correlation coefficient (r = 0.62) (Figure 1d). In this case, the slope is also driven away from the ideal, overestimating the lowest SOM levels and underestimating the higher SOM levels, which can be identi-

fied by the value's greater dispersion. Although the RMSE obtained for the unknown samples is 0.60%, the SOM differences for some samples is high. For instance, a sample with 1.5% SOM content measured with sulfochromic solution has an estimated amount of SOM ranging from 1.1% up to 2.8%, while at samples with 4.0% SOM content the NIRS estimated SOM content ranged from 2.7% up to 4.0%. This discrepancy between the measured and estimated SOM values may have a direct impact on nitrogen fertilizer management for grasses if the predicted value changes SOM class since nitrogen recommendation is based on SOM content (CQFS-RS/SC, 2016).
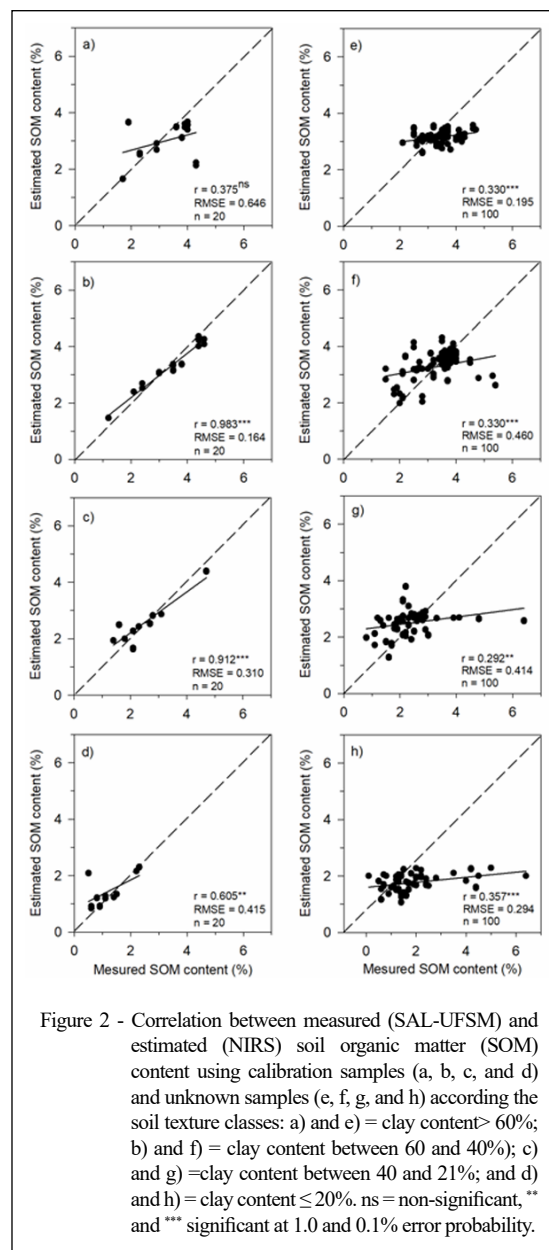
The light absorption is directly related to molecular vibration and rotation of organic functional groups in the near-infrared region (GUO et al., 2019) and several bands are important in the SOM prediction process (HONG et al., 2019). Thus, the performance of

the prediction model depends on the soil characteristics from the samples used to create the calibration curve. In our study, the range of SOM content in the unknown samples was greater than the SOM content in the samples used in the models' curve calibration (Figure 1d), reducing its accuracy. Moreover, a lack of similarity between the soil matrix of the samples used to calibrate the model's curve and the unknown samples may compromise the spectral models' accuracy (MOURA-BUENO et al., 2019) and; therefore, the SOM prediction. This is an additional difficulty regarding routine soil analysis laboratories, which cannot identify

and/or guarantee the origin of the new samples delivered to be analyzed. In this case, a specific strategy should be developed to solve this problem and avoid misleading.

Model validation for SOM levels by soil texture class

In order to optimize the predictive capacity of the model, we tested the model performance for SOM prediction for each soil texture class (i.e., C1, C2, C3, and C4). This prediction presented better results when using a set (n = 20) with the same samples used in the model's curve calibration (Figures 2a, 2b, 2c, and 2d). The correlation coefficient values

Figure 2 - Correlation between measured (SAL-UFSM) and estimated (NIRS) soil organic matter (SOM) content using calibration samples (a, b, c, and d) and unknown samples (e, f, g, and h) according the soil texture classes: a) and e) = clay content> 60%; b) and f) = clay content between 60 and 40%); c) and g) =clay content between 40 and 21%; and d) and h) = clay content ≤ 20%. ns = non-significant, ** and *** significant at 1.0 and 0.1% error probability.

were 0.37, 0.98, 0.91, and 0.60, while the RMSE was 0.65%, 0.16%, 0.31%, and 0.41%, for C1, C2, C3, and C4; respectively, representing reasonable results, except for C1. However, with unknown samples (n = 100) (Figures 2e, 2f, 2g, and 2h), the model's prediction had no satisfactory results for either the classes. The correlation values were expressively lower and ranged between 0.29 and 0.36, whereas the RMSE values ranged from 0.19% up to 0.46%.

In general, the model tends to overestimate the lower SOM content and underestimate the higher SOM content within the same texture class. Such behavior may be related to the size of the set of samples and the soil samples' mineralogy. Although the texture class has sorted the soil samples, mineralogical changes within the same texture class may significantly influence the results. Furthermore, the range of SOM content and SOM constitution can also have influence since different chemical groups behave differently when subjected to near-infrared radiation (DALMOLIN et al., 2005).

Considering the high correlation between the values obtained by the laboratory quantification and those estimated by the model, it can be stated that the NIRS analysis technique has great potential to predict clay and SOM contents. However, considering that the model presented some errors for both clay and SOM estimation, further studies should be performed adopting different strategies (e.g., using a higher number of samples to generate the calibration curve or using samples with greater similarity in the soil matrix for the curve calibration and analysis), improving the model prediction performance. Moreover, the purpose of the soil analysis should be considered as well. Analyzes carried out for soil fertility diagnosis, for example, do not require an exact result for a fertilizer recommendation and/or allow some deviations with no greater implications. Conversely, analyzes frameworks for legal purposes or penalties require highly accurate results compared to the reference methods. Thus, it is still necessary to improve the NIRS technique for routine soil analysis, contributing to lower waste production while reducing the risk of laboratory contamination due to chemical reagents handling.

CONCLUSION

The SVM model showed the best performance in the calibration process for both clay and SOM contents, and for SOM content calibrated within each soil texture class. However, the soil clay content affects the predictive capacity of the NIRS calibration curve to estimate the SOM content.

The model's curve validation had poorer performance (lower R^2 and higher RMSE) when created from unknown samples than from a set of soil samples used in the calibration. The model tends to overestimate lower levels of clay and SOM while underestimating the highest clay and SOM contents.

Despite the potential use of the NIRS technique to predict clay and SOM contents in soil analysis laboratories, other calibration studies are still needed in order to improve the model's prediction performance.

ACKNOWLEDGMENTS

DECLARATION OF CONFLICT OF INTERESTS

The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

AUTHORS' CONTRIBUTIONS

BPL, LSS, and GLD designed the experiment. BPL, DB, and JLN performed the experiment and analysis in the laboratory. BPL and ACD performed statistical analyses of the data. BPL, LSS, and GLD prepared the manuscript.

REFERENCES

ARAÚJO, S.R. **Reflectance spectroscopy vis-NIR and mid-IR applied for soil studies**. 2013. 182p. Tese (Doutorado em Ciências). Universidade de São Paulo, Escola Superior de Agricultura "Luiz de Queiroz".

COMISSÃO DE QUÍMICA E FERTILIDADE DO SOLO – RS/SC. **Manual de calagem e adubação para os Estados do Rio Grande do Sul e de Santa Catarina**. Sociedade Brasileira de Ciência do Solo – Núcleo Regional Sul. – [s. I.], 2016. 376p.

DALMOLIN, R.S.D. et al. Relationship between the soil constituents and its spectral behavior. **Ciência Rural**, v.35, n.2, p.481-489, 2005. Available from: <http://dx.doi.org/10.1590/S0103-84782005000200042>. Accessed: Mar. 18, 2019.

DOTTO, A.C. et al. AlradSpectra: a quantification tool for soil properties using spectroscopic data in R. **Revista Brasileira de Ciência do Solo**, v.43, e0180263, 2019. Available from: <http://dx.doi.org/10.1590/18069657rbcs20180263>. Accessed: Sep. 25, 2019.

FELIX, J.C. et al. Phosphorus, carbon, and nitrogen prediction in basaltic soils by NIR spectroscopy. **Pesquisa Agropecuária Brasileira**, v.51, n.9, p.1405-1416, 2016. Available from: <http://

dx.doi.org/10.1590/s0100-204x2016000900039>. Accessed: Apr. 19, 2019.

FERRARESI, T.M. et al. Infrared spectroscopy in determination of soil texture. **Revista Brasileira de Ciência do Solo**, v.36, p.1769-1777, 2012. Available from: <http://dx.doi.org/10.1590/S0100-06832012000600010>. Accessed: Mar. 11, 2019.

GUO, L. et al. Prediction of soil organic carbon stock by laboratory spectral data and airborne hyperespectral images. **Geoderma**, v.337, p.32-41, 2019. Available from: <https://doi.org/10.1016/j.geoderma.2018.09.003>. Accessed: Jun. 11, 2019.

HONG, Y. et al. Combination of fractional order derivative and memory-based learning algorithm to improve the estimation accuracy of soil organic matter by visible and near-infrared spectroscopy. **Catena**, v.174, p.104-116, 2019. Available from: <https://doi.org/10.1016/j.catena.2018.10.051>. Accessed: Jun. 11, 2019.

JACONI, A. et al. Near infrared spectroscopy as an easy and precise method to estimate soil texture. **Geoderma**, v.337, p.906-913, 2019. Available from: <https://doi.org/10.1016/j.geoderma.2018.10.038>. Accessed: Jun. 13, 2019.

SOUSA JR., J.G.A. et al. Spectral response of soils in the landscape based on terrestrial and orbital data acquisition levels. **Revista Brasileira de Ciência do Solo**, v.32, p.727-738, 2008. Available from: <http://dx.doi.org/10.1590/S0100-06832008000200027>. Accessed: Jun. 13, 2019.

MOURA-BUENO, J.M. et al. Stratification of a local VIS-NIR-SWIR spectral library by homogeneity criteria yields more accurate soil organic carbon predictions. **Geoderma**, v.337, p.565-581. 2019. Available from: <https://doi.org/10.1016/j.geoderma.2018.10.015>. Accessed: Jun. 19, 2019.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing. 2019. Available from: <https://www.r-project.org>. Accessed: Jan. 10, 2019.

SÁ, S.O. et al. Application of LS-SVM-NIR spectroscopy for carbon and nitrogen prediction in soils under sugarcane. In:19th WORLD CONGRESS OF SOIL SCIENCE, SOIL SOLUTIONS FOR A CHANGING WORLD. **Conference paper**, Brisbane, Australia, p.17-20, 2010.

SILVA, E.E. et al. Prediction of chemical and physical properties of Roraima soils by near infrared spectroscopy. **Biota Amazônia**, v.7, n.2, p.31-35, 2017. Available from: <https://periodicos.unifap.br/index.php/biota/article/view/2959>. Accessed: Jun. 26, 2019.

STEVENS, A. et al. Prediction of soil organic carbon at the European scale by visible and near infrared reflectance spectroscopy. **Plos One**, v.8, n.6, e66409, 2013. Available from: <https://doi.org/10.1371/journal.pone.0066409>. Accessed: Jun. 26, 2019.

TEDESCO, M.J. et al. **Análises de solo, plantas e outros materiais**. Porto Alegre: Departamento de Solos, UFRGS. 1995. 174p. (Boletim técnico n. 5).

TEIXEIRA, P.C. et al. **Manual de métodos de análise de solo**. 3.ed. Brasília, DF. Embrapa, 2017. 574p.

VANDRAME, P.R.S. et al. **Predição da granulometria e mineralogia de latossolos do cerrado utilizando espectroscopia de infravermelho próximo**. Planaltina: Embrapa. 2015. 25p. (Boletim de Pesquisa e Desenvolvimento n.326).

VAN VUUREN, J.A.J. et al. Potential use of near infrared reflectance monitoring in precision agriculture. **Communications in Soil Science and Plant Analysis**, v.37, p.2171-2184, 2006. Available from: <https://doi.org/10.1080/00103620600817382>. Accessed: Jun. 14, 2019.