

Utilização do sensoriamento remoto e da técnica quimiométrica PLSR na predição de atributos mineralógicos da região Noeroeste do Estado do Paraná

Franciele Romagnoli ⁽¹⁾
Marcos Rafael Nanni ⁽¹⁾
Anderson Antônio da Silva ⁽¹⁾
Carlos Antonio da Silva Junior ⁽¹⁾
Aline de Carvalho Gasparotto ⁽¹⁾
Everson Cezar ⁽¹⁾
Rafael Moreno Campos ⁽¹⁾
Mônica Sacioto ⁽¹⁾

⁽¹⁾ Universidade Estadual de Maringá (UEM), Avenida Colombo, 5790 – Maringá – Brasil - CEP: 87020900, (44) 3011-1359, franciele_romagnoli@hotmail.com, marcos.nanni@gmail.com, alinegaspa@hotmail.com, carlos-junior89@gmail.com, eversoncezar@yahoo.com.br Anderson_agrouem@hotmail.com, monicasacioto@gmail.com e rafaelmorenocampos@gmail.com.

ABSTRACT: This study aimed to evaluate the chemometric technique Partial Least Squares Regression (PLSR) performed by software in prediction of mineralogical soil properties (SiO_2 , Al_2O_3 , Fe_2O_3 and TiO_2) . Was used for both data soil reflectance (Vis_NIR) using a spectroradiometer laboratory . Chosen for the study areas are located in the northwest of Paraná , in lithosequence located in the transition zone between basalt and sandstone . Soil samples were collected in order to characterize the horizon A (0 - 0.20 m) and horizon B (0.8 - 1.0 m) and these were subjected to mineralogical analysis at the State University of Maringá . The spectral measurements were performed on a spectroradiometer in the range 350-2500 nm , enabling the prediction of soil attributes through its spectral response. Was possible to determine a prediction model to determine the soil through their spectral response using the multivariate technique of partial least squares regression (PLSR) through Parles program. Both in obtaining prediction models (calibration) and validation of these models , the performance of these tested by R^2 , RMSE and RPD were similar . The prediction model (calibration) and validation for TiO_2 did not perform well .

keywords : Vis - NIR spectroscopy , multivariate modeling, chemometrics and soils

1 Introdução

Uma vez realizado os modelos de calibração para predição dos atributos, pode-se analisar muitas amostras em poucos minutos sem a necessidade de recorrer aos métodos de análises convencionais. Mas, para que esses modelos de estimativa de atributos do solo, utilizando dados espectrais, possam ser utilizados para realizar predições em conjunto com os métodos convencionais, ou que venham a substituí-los, pesquisas deverão ser continuadas para melhorar a precisão e confiabilidade desses modelos de calibração.

A faixa espectral do Visível até o infravermelho médio (VIS-NIR-MIR) tem sido muito utilizada para quantificar os atributos do solo. Segundo Stenberg et al., (2010) a faixa espectral MIR (2500-25000 nm) tem provado a dar melhores predições quando

comparado com VIS-NIR (400-2500), mas é menos viável devido aos custos, portabilidade e preparação das amostras. Assim, a maioria das pesquisas de espectroscopia em solos se restringe a região VIS-NIR.

Na tentativa de correlacionar atributos do solo com sua resposta espectral, muitas técnicas multivariadas tem sido utilizadas. Segundo Stenberg et al., (2010), os métodos de calibração mais utilizados na estimativa de atributos é a Regressão Linear Múltipla com a seleção Stepwise, PCR e PLSR. A PCR e a PLSR são técnicas capazes de lidar com grande número de variáveis preditas altamente colineares. Todos esses métodos são lineares, porque os modelos resultantes descrevem uma relação linear entre as variáveis respostas e as concentrações dos componentes das amostras (dados espectrais) (Morgano, 2005).

Duas etapas são importantes na calibração multivariada: a obtenção do modelo (calibração) e verificação de sua validade, para analisar se o modelo é ou não realístico. A obtenção de um modelo de calibração significa desenvolver expressões matemáticas que possam descrever o comportamento da variável a ser estimada (y). Ao obter modelos de predição, segundo Stenberg et al. (2010) a seleção das amostras para obtenção dos modelos de predição, e amostras para a validação são de extrema importância.

Diante disto, o objetivo deste trabalho, foi de estimar atributos mineralógicos do solo por meio de seus dados espectrais e modelagens PLSR e verificar seus desempenhos no ajuste e validação dos modelos de predição.

2. Materiais e Métodos

O trabalho foi realizado na região Noeroeste do estado do Paraná, em uma zona transicional entre as formações basálticas e o arenito, apresentando certa porção de material arenítico inconsolidado, cobrindo o leito basáltico. Esta associação de materiais geológicos com variáveis características e diferentes produtos advindos do processo intempérico possibilitou a formação de diversas classes de solos, devido a variação de composição dos minerais presentes no solo. Segundo a carta de solos do Estado do Paraná Iapar (2008), as classes de solos que ocorrem nessa região são: Argissolos, Latossolos e Neossolos.

Para coleta dos solos e georreferenciamento dos pontos, As coletas com trado foram realizadas nas profundidades de 0 – 0,2 m, para caracterização de horizontes superficiais, e de 0,8 – 1,0 m para horizontes subsuperficiais. Foram coletadas um total de 156 amostras decorrentes dos 78 pontos de coleta.

As amostras foram encaminhadas ao laboratório de análises de solos da Universidade Estadual de maringá, onde os elementos mineralógicos, Fe_2O_3 , Al_2O_3 e TiO_2 foram determinados por espectrofotometria de absorção atômica no extrato do ataque sulfúrico após filtragem lenta, como recomendado pela (Embrapa, 1997). O SiO_2 foi determinado gravimetricamente após calcinação a 550 °C do material residual retido no papel de filtro (Embrapa, 1997).

As leituras espectrais foram realizadas no laboratório de Geoprocessamento e Sensoriamento Remoto Aplicado ao Meio Ambiente, do Departamento de Agronomia da Universidade Estadual de Maringá. O equipamento utilizado foi um espectrorradiômetro FieldSpec 3, com intervalo espectral de 350 – 2500 nm (Vis-NIR) e resolução espectral de 3nm até 700 e de 30 nm de 700 a 2500nm, presente em ambiente controlado de umidade, luminosidade e temperatura. O sensor ótico presente na ponta da fibra foi colocado em posição vertical a 8 cm de distância da amostra, sendo medida a luz refletida numa área aproximada de 2 cm² no centro da amostra. A fonte de

iluminação utilizada foi uma lâmpada halógena de 650 W, com feixe não colimado para o plano visado, sendo posicionada a 35 cm da amostra e com um ângulo zenital de 30° (Figura 1).



Figura 1. Disposição e geometria do equipamento FieldSpec para obtenção dos dados espectrais das amostras.

Para ajustar um modelo de calibração e posterior validação para predição dos atributos do solo, SiO_2 , Al_2O_3 , Fe_2O_3 , TiO_2 e C.O., foram utilizados os Softwares, ParLes 3.1, proposto por Viscarra Rossel, (2008). Para tanto, foi utilizada a técnica PLSR, introduzida por Word (1975), por meio da validação cruzada. A PLSR é uma técnica de análise multivariada, muito utilizada em análises de predição por meio dos dados espectrais, no qual as variáveis preditoras são altamente colineares (Belmonte, 2006). Durante a etapa de calibração, a modelagem PLS utiliza as informações da matriz de dados X e matriz de concentração Y, obtendo-se novas variáveis denominadas, variáveis latentes (fatores PLS).

Para a fase de calibração dos modelos de predição por meio do PLSR e da regressão Stepwise foram selecionadas aleatoriamente 2/3 das curvas espectrais junto com seus respectivos valores de atributos determinados em laboratório para ajuste do modelo restando 1/3 das amostras para posterior validação. Proporções semelhantes de divisão entre amostras de calibração e validação foram utilizadas por Reeves et al. (1999), McCarty et al (2002), Shepherd e Walsh (2002) e Islam et al. (2003).

De acordo com Viscarra Rossel et al., (2007), ao ajustar um modelo utilizando-se PLSR, pretende-se encontrar o menor número possível de fatores PLS necessários para explicar a maior parte da variação entre as variáveis. Durante a validação cruzada (cross validation), n-1 amostras foram utilizadas para a calibração do modelo de predição.

De acordo com Morgano (2005) o tratamento dos dados é aplicado antes da fase de calibração para melhorar a eficiência do método, produzindo modelos de regressão com melhor capacidade de predição. Foram realizados diferentes pré-processamentos, que resultasse no menor RMSE, apresentados na Tabela 1.

Tabela 1. Transformações, pré processamentos e pré tratamento dos dados

Atributos	Pré processamento	Pré tratamento
SiO ₂ (g kg ⁻¹)	MSC ¹	Mean centre ²
Al ₂ O ₃ (g kg ⁻¹)	MSC	Mean centre
Fe ₂ O ₃ (g kg ⁻¹)	MSC	Mean centre
TiO ₂ (g kg ⁻¹)	MSC	Mean centre

¹ Correção multiplicativa de sinal, ² Dados centrados na média

Para assegurar a capacidade preditiva de um modelo é necessário realizar um processo de validação do mesmo, aplicando ao modelo, novas amostras (que não fizeram parte da etapa de calibração).

Neste trabalho, 52 amostras foram utilizadas para validação, correspondendo a 1/3 das amostras totais. A seleção das amostras para calibração e validação dos modelos de predição é de fundamental importância. De acordo com Stenberg et al., (2010), a quantidade e a variabilidade das amostras selecionadas para a calibração, deve ser aquela que represente bem a área de estudo, e as amostras utilizadas na fase de validação dos modelos, não podem ser aquelas utilizadas nas calibrações do mesmo.

3. Resultados e Discussão

Os resultados na análise estatística esta apresentada na Tabela 2, a seguir. Os atributos SiO₂, Al₂O₃ e Fe₂O₃ apresentaram, para o modelo de predição, R² e RPD de 0,73 e 1,92; 0,74 e 1,96 e 0,73 e 1,93 respectivamente, sendo os modelos considerados como modelos aceitáveis, segundo (Dunn et al., 2002; Chang et al., 2001 e Sayes et al., 2005).

Tabela 2. Fatores PLS, R². RMSE e RPD dos modelos para os atributos avaliados

Atributos	Fatores PLS	R ²	RMSE	RPD
SiO ₂ (g kg ⁻¹)	6	0,73	71,17	1,92
Al ₂ O ₃ (g kg ⁻¹)	5	0,74	22,77	1,96
Fe ₂ O ₃ (g kg ⁻¹)	6	0,75	29,69	1,93
TiO ₂ (g kg ⁻¹)	8	0,59	10,02	1,34

O modelos do TiO₂ (R² = 0,59 e RPD = 1,51) indicou apenas a possibilidade de discriminação de altas e baixas concentrações e modelagem não confiável para predição, respectivamente (Sayes et al., 2005).

Depois de obtido os modelos para predições dos atributos do solo, foram realizadas suas validações, a fim de testar a sua capacidade de previsão. Essa validação foi realizada utilizando as amostras que não entraram na fase de construção do modelo. Para isto, foram utilizadas 52 amostras correspondendo, portanto, a um terço das amostras totais.

A validação dos modelos para predição dos atributos realizados pelo método PLSR, estão apresentados na Tabela 3.

Tabela 3. R², RMSE e RPD resultantes da validação dos modelos de predição

Atributos	R ²	RMSE	RPD
SiO ₂ (g kg ⁻¹)	0,74	55	1,96
Al ₂ O ₃ (g kg ⁻¹)	0,74	22,77	2,10
Fe ₂ O ₃ (g kg ⁻¹)	0,79	21,17	1,45
TiO ₂ (g kg ⁻¹)	0,54	9,15	1,67

Melhor desempenho na fase de validação foi obtido para o atributo Fe₂O₃ (R² = 0,79 e RPD = 2,10). Na fase de obtenção dos modelos (calibração), esse atributo também tiveram apresetou bom desempenho (Tabela 2).

Summers et al. (2006) trabalhando na Austrália obtiveram, para os modelos da argila e Fe₂O₃, R² de 0,66 e 0,61 e RPD de 2,0 e 1,7 respectivamente. Os modelos para o SiO₂ e o Al₂O₃ apresentaram, para fase de validação, R² 0,74 e RPD 1,9 e 1,96 e a validação do TiO₂, R² de 0,54 e RPD de 1,45.

Para os atributos SiO₂, Fe₂O₃ e TiO₂, , melhores modelos de predição, foram encontrados por Nanni e Demattê (2006) que, , encontraram R² de 0,92, 0,95 e 0,95 respectivamente.

Genú e Demattê (2010), utilizando dados espectrais e orbitais e dados topográficos - MDT (modelo digital do terreno), no município de Rafard – SP, obtiveram para SiO₂, Fe₂O₃ e TiO₂, R² de 0,12, 0,55 e 0,41 respectivamente.

4. Conclusão

- Foi possível determinar modelos de predição para os atributos do solo por meio de sua resposta espectral utilizando a técnica multivariada regressão por mínimos quadrados parciais (PLSR) por meio do programa ParLes.

- Tanto na fase de obtenção dos modelos de predição (calibração), quanto na fase de validação desses modelos, o desempenho destes, testados pelo R², RMSE e RPD foram similares.

Referências bibliográficas

Belmonte, R.Z. Evaluación de la calidad ambiental en suelos de la provincia de Alicante: desarrollo y aplicación de diferentes técnicas. **Tese** (Doutorado). Elche. 2006.

Chang, C.; Laird, D. A.; Mausbach, M. J.; Hurburg Junior, C. R. Near infrared reflectance spectroscopy – principal components regression analyses of soil properties. **Soil Science Society of American Journal**, v.25, n.2, p. 480-490, 2001.

Dunn, B.W.; Beecher, H. G.; Batten, G. D.; Ciavarella, S. The potential of near-infrared reflectance spectroscopy for soil analysis — a case study from the Riverine Plain of south-eastern Australia. **Australian Journal of Experimental Agriculture**, v.42, p.607-614, 2002.

Embrapa. **Manual de métodos de análises de solo**. 2.ed. Rio de Janeiro: Ministério da Agricultura e do Abastecimento, 1997

Genú, A. M.; Demattê, J.A.M. Informações espectroeletrônicas e topográficas na determinação de teores de componentes do solo. **Bragantia**.v.69, n.1, p. 157-164, 2010

Instituto Agronômico do Paraná (IAPAR). **Carta de solos do Estado do Paraná**. 2008.

Islam, K., Singh, B., McBratney, A.B., Simultaneous estimation of various soil properties by ultra-violet, visible and near-infrared reflectance spectroscopy. **Australian Journal of Soil Research**, v. 41, 1101– 1114, 2003.

McCarty, G.W., Reeves III, J.B., Reeves, V.B., Follett, R.F., Kimble, J.M., Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurements. **Soil Science Society of America Journal**, v.66, 640–646, 2002.

Morgano, M. A. Aplicação do métodos Quimiométrico em análise de alimentos. **Tese** (Doutorado). Universidade de Campinas. Campinas. 2005.

Nanni, M.R.; Demattê, J.A.M. Spectral reflectance methodology in comparison to traditional soil analysis. **Soil Science Society of America Journal**, v. 70, p. 393-407, 2006.

Reeves III, J.B., McCarty, G.W., Meisinger, J.J., Near infrared reflectance spectroscopy for the analysis of agricultural soils. **Journal of Near Infrared Spectroscopy**, v. 7, 179– 193. 1999.

Sayes, W.; Mouazen, A.M.; Ramon, H. Potencial for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy. **Biosystems Engineering**, v.91, p.393-402, 2005.

Shepherd, K.D., Walsh, M.G., Development of reflectance spectral libraries for characterization of soil properties. **Soil Science Society of America Journal** v. 66, 988– 998, 2002.

Stenberg, B.; Viscarra Rossel, R.A.; Mouazen, A.M.; Wetterlind. Visible and near infrared spectroscopy in soil science. **Advances in Agronomy**, v.107, p.163-215, 2010.

Viscarra Rossel, R. A. ParLeS: Software for chemometric analysis of spectroscopic data. **Chemometrics Intelligent Laboratory Systems**, v.90, p.72-83, 2008.

Viscarra Rossel, R.A. Robust modelling of soil diffuse reflectance spectra by —bagging — partial least squares regression. **Australian Centre for Precision Agriculture**. p.39-47, 2007.