



Determinação de carbono orgânico do solo por espectroscopia NIR e calibração multivariada⁽¹⁾.

André Marcelo de Souza⁽²⁾; Maurício Rizzato Coelho⁽³⁾; Ademir Fontana⁽³⁾; Camila Bressan⁽⁴⁾; Ronei Jesus Poppi⁽⁵⁾.

⁽¹⁾ Trabalho executado com recursos do Projeto NIRSOLOS (MP3 - 03.12.03.010.00.00) - Embrapa.

⁽²⁾ Analista; Empresa Brasileira de Pesquisa Agropecuária (Embrapa Solos); Rio de Janeiro, RJ; Rua Jardim Botânico, 1024, Jardim Botânico, Rio Janeiro/RJ, CEP 22460-000 (andremarcelo.souza@embrapa.br); ⁽³⁾ Pesquisador; Empresa Brasileira de Pesquisa Agropecuária (Embrapa Solos); Rio de Janeiro, RJ; Rua Jardim Botânico, 1024, Jardim Botânico, Rio Janeiro/RJ, CEP 22460-000; ⁽⁴⁾ Estudante; Empresa Brasileira de Pesquisa Agropecuária (Embrapa Solos); Rio de Janeiro, RJ; Rua Jardim Botânico, 1024, Jardim Botânico, Rio Janeiro/RJ, CEP 22460-000; ⁽⁵⁾ Professor, Instituto de Química, Universidade Estadual de Campinas, Caixa Postal 6154, Campinas, SP, CEP 13083-970.

RESUMO: As pesquisas atuais apontam que a espectroscopia no infravermelho próximo (NIR) é a técnica alternativa mais promissora para a determinação de carbono orgânico do solo (SOC) em substituição total ou parcial aos métodos tradicionais de via úmida. Considerado este fato, foi desenvolvido e validado um método para a determinação de SOC por espectroscopia NIR, visando seu emprego como método de rotina em laboratórios de análise de solos do Brasil. Para este fim, foram construídos modelos de calibração multivariada a partir de um número expressivo de amostras de solos (1490 amostras, 2980 espectros) que englobam a variabilidade de solos brasileiros. Estes modelos foram validados através da submissão dos valores previstos das concentrações de matéria orgânica do solo (SOM) ao Programa de Análise de Qualidade de Laboratórios de Fertilidade (PAQLF).

O sucesso da validação através do ensaio interlaboratorial não deixa dúvida sobre a robustez do modelo e, conseqüentemente sobre o potencial de aplicação da espectroscopia NIR como método de rotina para análise de solos no Brasil. Essa assertiva está suportada pela alta capacidade preditiva que os modelos de calibração demonstraram quando submetidos ao ensaio PAQLF de 2013 (em nível de simulação). Os resultados alcançados foram suficientes para classificar o laboratório hipotético dentre aqueles que receberiam o selo de qualidade A no Programa, igualando e superando (em alguns modelos), em termos de classificação geral, o desempenho alcançado pelo método de via úmida (método de dicromato).

Termos de indexação: Quimiometria, PLS, estoque de carbono.

INTRODUÇÃO

As pesquisas atuais apontam que a espectroscopia NIR é a técnica alternativa mais

promissora para a análise de solos nos laboratórios de rotina em substituição total ou parcial aos métodos tradicionais, sobretudo na determinação de SOC pelo método de via úmida. O crescente interesse pela sua aplicação em análise de solos pode ser justificado pelas inúmeras e marcantes vantagens que esta apresenta em relação às análises convencionais: (i) ser uma técnica não destrutiva; (ii) livre de resíduos indesejáveis e, portanto, impactos ambientais; (iii) ser de baixo custo; ser uma técnica (iv) rápida e inovadora, requerendo pouco manuseio das amostras quando combinada à Quimiometria. Entretanto, sua principal limitação está na dependência da Quimiometria para tratamento dos dados espectrais. Devido a este e demais fatores, até o presente momento, a técnica não é amplamente utilizada nos laboratórios de análise de solos, destacando-se à dificuldade de se construir modelos de calibração robustos (Madari et al., 2006; Stenberg et al., 2010).

São muitos os métodos quimiométricos e suas aplicações dependem da natureza do problema que se deseja resolver ou do tipo de informação que se pretende obter. Dentre os métodos quimiométricos mais utilizados em espectroscopia NIR estão a análise por componentes principais (*Principal Component Analysis*, PCA) e a regressão por quadrados mínimos parciais (*Partial Least Squares*, PLS). A PCA é um método que permite a redução da dimensionalidade através da representação do conjunto de dados em um novo sistema de eixos, denominados Componentes Principais (*Principal Component - PC*), permitindo a visualização da natureza multivariada dos dados em poucas dimensões (Wold, 2002; Souza and Poppi, 2012). Por sua vez, no PLS a variação da resposta espectral está relacionada aos atributos (ou propriedades) de interesse através de uma regressão multivariada. O PLS utiliza a PCA para a redução da dimensionalidade do conjunto de dados para posterior correlação entre os espectros (matriz X) e as propriedades de interesse (vetor y). A propriedade de interesse muitas vezes é a



concentração de um analito, porém não limitada à esta, podendo até mesmo ser propriedades físico-químicas, tais como densidade e viscosidade, as quais são relacionadas a composição da amostra (Wold et al., 2001; Souza et al., 2013).

O PLS foi empregado na criação dos modelos de calibração multivariada empregados nesse estudo que, em termos gerais, objetiva desenvolver e validar um método para a determinação de carbono orgânico do solo, por espectroscopia NIR e calibração multivariada, para ser empregado como método de rotina em laboratórios de análise de solos do Brasil. Como objetivos específicos destacam-se: (1) construir uma biblioteca espectral empregando amostras representativas de solos brasileiros; e (2) validar o modelo de calibração multivariada submetendo os valores previstos de matéria orgânica do solo (*Soil Organic Matter*, SOM) ao Programa de Análise de Qualidade de Laboratórios de Fertilidade (PAQLF) (Quaggio, 1987).

MATERIAL E MÉTODO

Determinação do carbono orgânico do solo

O carbono orgânico é determinado pelo método volumétrico do dicromato de potássio ($K_2Cr_2O_7$ $0,2 \text{ mol L}^{-1}$) descrito no Manual de Método de Análise de Solos da Embrapa Solos (Embrapa, 2011). No referido método, o excesso de dicromato é titulado com sulfato ferroso amoniacal e os resultados são expressos em g kg^{-1} . Para o cálculo do teor de matéria orgânica do solo (SOM), necessário para a etapa de validação empregando ensaio interlaboratorial, multiplica-se o valor de SOC por 1,724 (presumindo que a matéria orgânica do solo contenha 58% de carbono).

Preparo das amostras de solo

As amostras de solos foram secas em estufa a 40°C durante 48 horas, destorroadas com um martelo de borracha e passadas em peneira $< 2 \text{ mm}$, obtendo-se a fração conhecida como terra fina seca ao ar (TFSA) (Embrapa, 2011).

Aquisição dos espectros NIR

Os espectros foram adquiridos em duplicata, na região de 4000 a 10000 cm^{-1} , com resolução de 4 cm^{-1} , 64 varreduras, no modo de porcentagem de refletância (% R). O método de pré-processamento espectral empregado foi o da primeira derivada de Savitzky-Golay, com polinômio de segunda ordem e janela de 15 pontos. As análises foram realizadas em espectrofotômetro NIR PerkinElmer®, equipado

com acessório de reflectância difusa (NIRA). O espectrofotômetro e a esfera de integração são equipados com detectores de DTGS/AsInGa, respectivamente.

Modelos de calibração multivariada

A construção dos modelos de calibração envolveu 1.490 amostras de solos de perfis completos (2.980 espectros) representativas do Brasil, selecionadas a partir da coleção nacional de amostras de terra (Soloteca) da Embrapa Solos. O critério de seleção das amostras considerou sua representatividade ambiental dos solos brasileiros de tal forma que foram selecionadas todas as classes de solos em nível de subordem taxonômica (segundo nível categórico), de acordo com o Sistema Brasileiro de Classificação de Solos (SiBCS) (Santos et al., 2013). Foram 11 os modelos de calibração construídos empiricamente com base nos valores da variável y (concentração de SOC, g.kg^{-1}). O método descrito por Kennard e Stone (1969) foi empregado na seleção dos subconjuntos de dados para a construção dos modelos de calibração multivariada (2/3 para calibração e 1/3 para validação). Os modelos foram construídos empregando-se os seguintes intervalos de concentração de SOC: (M0) $0,1 < y < 36,4$; (M1) $5 < y < 10$; (M2) $10 < y < 15$; (M3) $15 < y < 20$; (M4) $20 < y < 25$; (M5) $25 < y < 40$; (M6) $25 < y < 70$; (M7) $10 < y < 50$ (gkg^{-1}); (M8) $5 < y < 70$ (gkg^{-1}); (M9) $5 < y < 55$ (gkg^{-1}); (M10) $3 < y < 60$ (gkg^{-1}).

Ensaio interlaboratorial

O Programa de Análise de Qualidade de Laboratórios de Fertilidade (PAQLF) originalmente classifica os laboratórios baseando nos resultados dos oito atributos necessários para avaliação da fertilidade do solo. Entretanto, sendo o foco deste trabalho avaliar apenas a matéria orgânica do solo (SOM), obtida através análise do carbono orgânico do solo (SOM), os cálculos foram adaptados para avaliação apenas desse atributo. Conforme mencionado anteriormente, a percentagem de MOS é calculada multiplicando o resultado do SOC por 1,724 (Embrapa, 2011). Nessa etapa de validação dos modelos de calibração foram empregadas amostras do ensaio interlaboratorial de 2013 (12 amostras: 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226).

As amostras do PAQLF foram analisadas por espectroscopia NIR. Posteriormente esses espectros foram utilizados nos modelos de calibração multivariada para previsão das



concentrações de SOM nessas amostras. De posse dos valores de SOM previstos pelos modelos, dos resultados de todos os laboratórios participantes do PAQLF de 2013, cada modelo de calibração foi avaliado como se fosse um novo laboratório participante do ensaio, via simulação. Para fins de comparação, as amostras do programa foram analisadas pelo método de referência (oxidação por dicromato de potássio) e tiveram seus resultados avaliados pelo programa de maneira similar (via simulação).

RESULTADOS E DISCUSSÃO

A Tabela 1 apresenta um resumo dos resultados obtidos e as figuras de mérito avaliadas neste trabalho.

A fim de estimar qualitativamente a representatividade das amostras do programa interlaboratorial de 2013, realizou-se a PCA dos conjuntos de calibração e validação (Figura não mostrada). O gráfico dos escores (não mostrado) das duas primeiras componentes principais foi aquele que melhor demonstrou a relação entre os conjuntos de calibração, validação e PAQLF, portanto, as demais combinações não serão comentadas.

Alguns modelos de calibração foram robustos o suficiente para gerar resultados capazes de classificar o laboratório na categoria B do referido programa, cujos laboratórios desse nível apresentam boa qualidade de seus procedimentos analíticos. Portanto, os resultados alcançados pelos modelos de calibração habilitam a espectroscopia NIR a substituir o método de referência. Além disso, foi demonstrado que a espectroscopia NIR pode fornecer resultados ainda melhores se houver seleção de modelos locais (mais apropriados) para a previsão de uma amostra específica e com a inserção de novas amostras representativas das faixas de concentração de interesse (resultados não foram mostrados nesse resumo).

CONCLUSÕES

Foram desenvolvidos e validados através de conjunto externo de dados, modelos de calibração para determinação de carbono orgânico do solo (ou de matéria orgânica no solo) por espectroscopia NIR, a partir de um conjunto de dados altamente representativo dos tipos de solos encontrados em território brasileiro.

A validação através do ensaio interlaboratorial foi o principal resultado alcançado nesse trabalho e não deixa dúvida sobre a robustez do modelo e, conseqüentemente sobre o potencial de aplicação

da espectroscopia NIR como método de rotina. Essa assertiva está suportada pela alta capacidade preditiva que os modelos de calibração demonstraram quando submetidos ao ensaio interlaboratorial (em nível de simulação).

AGRADECIMENTOS

Aos estagiários da Embrapa Solos que se dedicaram com afinco a execução de etapas essenciais do Projeto.

Ao Instituto de Química da Unicamp.

REFERÊNCIAS

- Embrapa, 2011. Manual de Métodos de Análise de Solos, 2nd ed. Rio de Janeiro.
- Madari, B.E., Reeves, J.B., Machado, P.L.O.A., Guimarães, C.M., Torres, E., McCarty, G.W., 2006. Mid- and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two Ferralsols. *Geoderma* 136, 245–259.
- Quaggio, J.A., 1987. Programa de laboratórios de análise de solo do Estado de São Paulo. *Bol. Inf. da Soc. Bras. Cienc. do Solo* 12, 60–70.
- Souza, A. de, Breikreitz, M., Filgueiras, P., Poppi, R., 2013. Experimento didático de quimiometria para calibração multivariada na determinação de paracetamol em comprimidos comerciais utilizando espectroscopia no infravermelho próximo: um tutorial, parte II. *Quim. Nov.* 36, 1057–1065.
- Souza, A. de, Poppi, R., 2012. Experimento didático de quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: um. *Quim. Nova* 35, 223–229.
- Stenberg, B., Rossel, R.A.V., Mouazen, A.M., Wetterlind, J., 2010. Visible and Near Infrared Spectroscopy in Soil Science, in: *Advances in Agronomy*. Elsevier Inc, pp. 163–215.
- Wold, S., 2002. Principal component analysis. *Chemom. Intell. Lab. Syst.* 2, 37–52.
- Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* 58, 109–130.

Tabela 1 – Figuras de mérito dos modelos de calibração multivariada construídos a partir da biblioteca de espectros de amostras de solo representativas do território brasileiro.

Figuras de mérito		M0	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Exatidão	RMSEC	2.7	0.8	0.8	0.8	1.0	2.0	5.0	2.4	2.6	2.5	2.3
	RMSEP	2.5	0.9	0.9	1.0	0.9	2.7	6.2	2.5	2.3	2.3	2.1
Coeficiente de variação (%)	CV (%)	36.4	12.6	7.3	5.5	4.0	9.2	15.6	16.1	21.4	21.5	23.6
RPD	RPD	2.1	1.1	1.1	1.1	1.0	0.9	1.6	1.4	1.9	1.9	2.2
Precisão	(%)	0.5	0.4	0.3	0.3	0.1	0.2	2.1	1.1	0.6	0.6	0.5
Sensibilidade	SEN (g kg ⁻¹)	4.1E-05	8.2E-05	1.5E-04	2.5E-04	7.2E-04	3.0E-04	3.4E-05	4.0E-05	4.7E-05	5.0E-05	5.1E-05
Sensibilidade analítica	SEN(A) (g kg ⁻¹)	1.5	3.4	4.2	5.1	11.8	3.9	0.7	1.5	1.7	1.8	1.8
Inverso da sensibilidade analítica	1/SEN(A) 1/(g kg ⁻¹)	0.7	0.3	0.2	0.2	0.1	0.3	1.5	0.6	0.6	0.5	0.6
Limite de detecção	LOD (g kg ⁻¹)	2.1	0.9	0.7	0.6	0.3	0.8	4.4	1.9	1.8	1.6	1.7
Limite de quantificação	LOQ (g kg ⁻¹)	6.8	2.9	2.4	2.0	0.8	2.6	14.5	6.5	5.9	5.4	5.6
Ajuste	Coeficiente de determinação	0.83	0.65	0.61	0.69	0.67	0.70	0.86	0.75	0.79	0.80	0.83
Intervalo de confiança	Máximo	5.5	1.7	1.8	1.8	2.3	5.1	12.7	5.2	5.6	5.1	4.7
	Mínimo	5.3	1.5	1.6	1.6	2.0	4.2	10.3	4.8	5.1	4.9	4.5
Detecção de <i>outliers</i>		M0	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Variáveis latentes	vl	20	20	13	9	6	5	9	17	17	17	18
Número de amostras	cal	1986	484	260	138	81	36	54	528	1026	1026	1368
Número de amostras	val	994	243	130	70	41	18	28	264	513	505	684
Remoção de <i>outliers</i> 1	cal	1772	405	217	116	70	32	47	463	908	898	1181
Remoção de <i>outliers</i> 1	val	936	210	106	49	26	2	16	246	481	470	632
Remoção de <i>outliers</i> 2	cal	1544	351	185	104	57	32	47	396	780	758	992
Remoção de <i>outliers</i> 2	val	871	188	73	32	19	2	16	222	441	430	575