



Determinação de matéria orgânica no solo empregando espectroscopia no infravermelho próximo e máquina de vetores de suporte

Sandro K. Otani*, Felipe B. de Santana, André M. de Souza, Ronei J. Poppi.

Resumo

Neste trabalho foi realizada a determinação de matéria orgânica em solos por espectroscopia no infravermelho próximo empregando máquina de vetores de suporte para tratamento dos dados. Para o desenvolvimento do modelo foi utilizada uma biblioteca com cerca de 26 mil amostras de solos, resultando em erros quadráticos médios da ordem de 4,3g/dm³, para matéria orgânica na faixa de 0 a 47 g/dm³. Os resultados mostraram que é possível desenvolver uma metodologia alternativa de menor custo, mais rápida que o método tradicional e ambientalmente sustentável.

Palavras-chave:

Solo, Espectroscopia no infravermelho próximo, Aprendizado de máquina.

Introdução

A análise de matéria orgânica do solo (MOS) tem grande importância para a agricultura, visto que é um parâmetro de fertilidade utilizado para avaliar a viabilidade de futuros plantios, além de contribuir para o manejo adequado do solo.

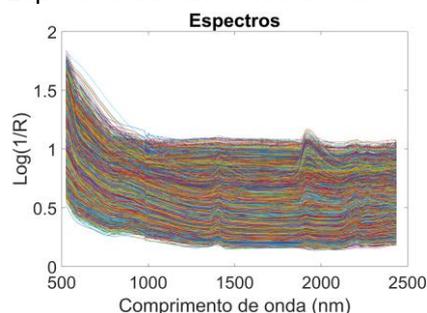
O procedimento padrão utilizado nesse tipo de análise é o método Walkley-Black que faz uso de uma solução de H₂SO₄ e K₂Cr₂O₇ na digestão da amostra e em seguida é realizada uma retrotitulação com FeSO₄ do Cr₂O₇²⁻ residual, processo que gera resíduos tóxicos e agressivos ao meio ambiente. Devido a esses fatos, a espectroscopia no infravermelho próximo (NIR) tem se destacado como método alternativo rápido, de baixo custo, não destrutivo e conseqüentemente sem geração de resíduos. Porém os dados obtidos são muito complexos e heterogêneos, requerendo o uso de métodos multivariados para extrair informação útil. No entanto os métodos multivariados convencionais não apresentam bons resultados, devido à enorme quantidade de dados. Neste sentido, o presente trabalho tem por objetivo o emprego do método de máquina de vetores de suporte para determinar os teores de matéria orgânica utilizando uma biblioteca espectral composta por cerca de 26 mil amostras de solo.

Resultados e Discussão

Nesse estudo foi utilizado como função kernel a Função de Base Radial (RBF) e otimizados no algoritmo o parâmetro (C) que pondera os erros dados por (ξ) na criação do modelo e o (γ) que é um parâmetro que regula a largura da função kernel, ou seja, que controla a complexidade do modelo. Para isto foi utilizada a otimização bayesiana que consiste em minimizar a raiz dos erros quadráticos médios da validação cruzada (RMSECV) gerando combinações de diferentes valores para os parâmetros (C) e (γ).

Os espectros das 25.926 amostras que constituem a biblioteca espectral estão ilustrados na Figura 1. Para minimizar a variação da linha base e acentuar as variações espectrais, os espectros foram pré-processados empregando a primeira derivada com suavização Savitzky-Golay. Dessa biblioteca espectral 10 mil amostras foram utilizadas para construir o modelo e o restante das amostras foram utilizadas na etapa de validação.

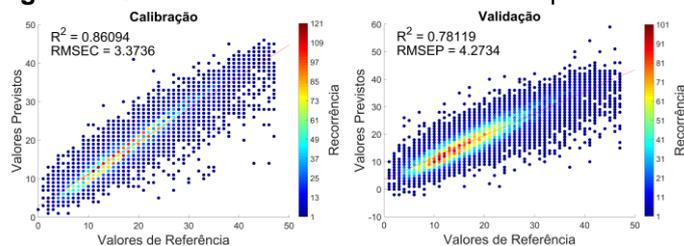
Figura 1. Espectros das amostras de solo



Na Figura 2 estão representados os gráficos dos valores de referência versus os valores previstos pelo modelo. Através da recorrência dos valores previstos observa-se que a dispersão dos pontos é mínima.

As proximidades entre os valores das raízes dos erros quadráticos médios da calibração (RMSEC=3,4 g/dm³) e da validação (RMSEP=4,3 g/dm³) e entre os coeficientes de determinação (R², 0,86 e 0,78), mostram uma ótima habilidade de previsão do modelo desenvolvido.

Figura 2. Gráficos de valores de referência vs previsto



Conclusões

Concluiu-se que a espectroscopia no infravermelho próximo aliada a máquina de vetores de suporte foi eficiente em lidar com o grande número de amostras e sua alta complexidade, indicando que pode ser uma metodologia alternativa para análise de matéria orgânica no solo.

Agradecimentos

CNPq (processo nº 102243/2019-0); IQ-Unicamp.

¹ Stenberg, B.; Viscarra Rossel, R. A.; Mouazen, M.; Wetterlind, J. Visible and near infrared spectroscopy in soil Science. Adv. Agron., 107 (2010), pp. 163-215