



XXXIII Congresso Brasileiro de Ciência do Solo

Solos nos biomas brasileiros: sustentabilidade e mudanças climáticas
31 de julho à 05 de agosto - Center Convention - Uberlândia/Minas Gerais

ANÁLISE EXPLORATÓRIA MULTIVARIADA EMPREGANDO ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO PARA CARACTERIZAÇÃO DE FERTILIZANTES ORGANOMINERAIS

André Marcelo de Souza⁽¹⁾; David Vilas Boas de Campos⁽²⁾; Ronei Jesus Poppi⁽³⁾; Vinicius de Melo Benites⁽²⁾; José Carlos Polidoro⁽²⁾

⁽¹⁾ Analista; Embrapa Solos, Rua Jardim Botânico, 1.024 - Jardim Botânico Rio de Janeiro, RJ - CEP 22460-000, Andre.souza@cnpes.embrapa.br;

⁽²⁾ Pesquisador; Embrapa Solos, Rua Jardim Botânico, 1.024 - Jardim Botânico Rio de Janeiro, RJ - CEP 22460-000;

⁽³⁾ Pesquisador; Instituto de Química, Universidade Estadual de Campinas, Caixa Postal 6154, Campinas, SP, CEP 13083-970.

Resumo – Cinco formulações diferentes de fertilizantes organominerais foram caracterizadas por Espectroscopia de Refletância no Infravermelho Próximo (Near-infrared Reflectance Spectroscopy, NIRS) e Análise por Componentes Principais (*Principal Component Analysis*, PCA). Diferentes métodos de pré-processamento do conjunto de espectros foram avaliados: primeira ou segunda derivada (com alisamento para eliminação de ruído) utilizando o algoritmo de Savitzky-Golay e correção de espelhamento multiplicativo (*Multiplicative Scatter Correction*, MSC). O método da segunda derivada com alisamento foi considerado o pré-processamento mais adequado para o conjunto de dados devido à eliminação na inclinação de linha de base e ruído. A PCA possibilitou a visualização de agrupamentos distintos entre as diferentes formulações evidenciando as diferenças e semelhanças entre elas com base nos perfis espectrais.

Palavras-Chave: NIRS; PCA;

INTRODUÇÃO

A Embrapa Solos vem desenvolvendo tecnologias de produção de novos fertilizantes, destacando-se os fertilizantes organominerais, que comparados aos fertilizantes minerais, apresentam uma solubilização mais gradativa no decorrer do período de desenvolvimento da cultura, aumentando assim a eficiência agrônoma (Kiehl, 2008). Diversos métodos espectroscópicos tem sido utilizados na caracterização de fertilizantes organominerais.

A espectroscopia de refletância no infravermelho próximo (*Near infrared Spectroscopy*, NIRS), associada à calibração multivariada, tem sido empregada por diversos autores para a caracterização química e física do solo (Francioso *et al.*, 2000; Mouazen *et al.*, 2006), caracterização de matéria orgânica (Moron and Cozzolino, 2004; Albrecht *et al.*, 2008) e caracterização de plantas de interesse agrícola (Gishen and Cozzolino, 2007).

A região espectral do infravermelho próximo estende-se de 10.000 a 4.000 cm⁻¹. As bandas de absorção nessa região são harmônicas ou combinações de bandas de estiramento fundamentais que ocorrem na

região de 3.000 a 1.700 cm⁻¹. As vibrações envolvidas são normalmente de ligações C-H, N-H e O-H (Skoog *et al.*, 2002).

Dentre as principais vantagens da espectroscopia NIR pode-se citar a velocidade de análise, a simplicidade na preparação de amostras (ou nenhum preparo), a facilidade de produzir replicatas de medidas, e ausência de reagentes químicos na obtenção dos espectros.

A análise de Componentes Principais (*Principal Component Analyses*, PCA) é um dos métodos mais importantes utilizados na Quimiometria e é a base para diversos métodos de reconhecimento de padrões, classificação e calibração multivariada. Normalmente, a PCA é utilizada com o objetivo de visualizar a estrutura dos dados, encontrar similaridades entre amostras, detectar amostras anômalas (*outliers*) e reduzir a dimensionalidade do conjunto de dados (Wold *et al.*, 1987). O reconhecimento de padrões através da PCA é largamente empregado em Quimiometria tendo cada vez mais novos usuários do meio acadêmico (alunos de graduação, pós-graduação e pesquisadores) e da indústria interessados na sua utilização.

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 (PC), permitindo a visualização da natureza multivariada dos dados em poucas dimensões. No espaço original, as amostras são pontos localizados em um espaço n- dimensional, sendo n igual ao número de variáveis. Com a redução de dimensionalidade proporcionada pela PCA, as amostras passam a ser pontos localizados em espaços de dimensões reduzidas definidos pelas PCs, por exemplo, bi ou tri dimensionais. Matematicamente, na PCA, a matriz **X** é decomposta em um produto de duas matrizes, denominadas *scores* (**T**) e *loadings* (**P**), mais uma matriz de erros (**E**), como mostrado na Equação 1:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad \text{Equação 1}$$

Os *scores* representam as coordenadas das amostras no sistema de eixos formados pelas Componentes Principais. Cada PC é formada pela combinação linear das variáveis originais e os coeficientes da combinação são denominados *loadings*. Os *loadings* são os cossenos dos ângulos entre as

variáveis originais e a Componente Principal, representando, portanto, o quanto cada variável original contribui para uma determinada PC. A Primeira Componente Principal (PC1) é traçada no sentido da maior variação no conjunto de dados; a segunda (PC2) é traçada ortogonalmente à primeira, com o intuito de descrever a maior porcentagem da variação não explicada pela PC1 e assim por diante. Enquanto os *scores* representam as relações de similaridade entre as amostras, a avaliação dos *loadings* permite entender quais variáveis mais contribuem para os agrupamentos observados no gráfico dos *scores* (Wold *et al.*, 1987). Diversos algoritmos estão disponíveis para a realização da PCA e quatro deles aparecem frequentemente na literatura: Nonlinear Iterative Partial Least Squares (NIPALS), Singular Value Decomposition (SVD), os quais utilizam a matriz de dados X , POWER e *Eigenvalue Decomposition* (EVD) que trabalham com a matriz de produto cruzado $X^T X$.

A etapa de pré-processamento dos dados é fundamental para o sucesso da análise multivariada. Os principais objetivos da aplicação das técnicas de pré-processamento são: eliminar informações não relevantes do ponto de vista químico e tornar a matriz de dados melhor condicionada para a análise, possibilitando a subsequente análise exploratória do conjunto de dados com eficiência. Existe uma vasta literatura disponível a respeito dos diversos métodos de processamento de dados em espectroscopia (Fearn *et al.*, 2009). Centrar os dados na média (*Mean Center*), derivar, alisar, aplicar correção de espalhamento multiplicativo (*Multiplicative Scatter Correction*, MSC) são alguns dos métodos mais aplicados.

Centrar os dados na média (Equação 2) consiste em calcular a média das intensidades para cada comprimento de onda e subtrair de cada uma das intensidades o valor médio. Desta maneira, cada variável passará a ter média zero, ou seja, as coordenadas são movidas para o centro dos dados, permitindo que diferenças nas intensidades relativas das variáveis sejam mais fáceis de perceber.

$$X_{ij} = X_{ij} - \bar{X}_{ij} \quad \text{Equação 2}$$

Deslocamento e inclinação de linha de base podem ser corrigidos por derivação dos espectros. Os métodos de alisamento são utilizados para reduzir matematicamente o ruído, aumentando com isto a relação sinal/ruído. Nestes métodos, é selecionada uma janela, a qual contém certo número de variáveis. Os pontos na janela são então utilizados para determinar o valor no ponto central da janela e assim, o tamanho da janela influencia diretamente o resultado do alisamento. No método de Savitzky-Golay, um polinômio de ordem baixa é ajustado aos pontos da janela e utilizado para recalculer o ponto central.

A correção de espalhamento multiplicativo é um método de transformação utilizado para compensar os efeitos aditivos e/ou multiplicativos em dados espectrais. Este método remove a influência de efeitos físicos nos espectros, tais como o tamanho de partícula,

a rugosidade e opacidade, os quais não trazem informações químicas sobre as amostras e introduz variações espectrais como o deslocamento da linha de base. Para isto fazer a correção, o método MSC assume que cada espectro é determinado pelas características químicas da amostra somadas às características físicas indesejadas. A Equação 3 descreve o princípio de funcionamento do MSC:

$$x_{ik}(\text{transformado}) = \frac{[x_{ik}(\text{original}) - a_i]}{b_i} \quad \text{Equação 3}$$

onde: x_{ik} (original) e x_{ik} (transformado) são os valores de absorvância antes e depois correção com o MSC em k comprimentos de onda; a_i e b_i são constantes estimadas a partir de uma regressão em mínimos quadrados de um espectro individual x_{ik} contra um espectro médio \bar{x} do conjunto de calibração em todos os comprimentos de onda ou em um subconjunto, seguindo a Equação 4:

$$x_{ik} = a_i + b_i \bar{x} + e_{ik} \quad \text{Equação 4}$$

onde e_{ik} corresponde a todos os outros efeitos nos espectros que não foram modelados (Fearn *et al.*, 2009).

Este estudo preliminar teve como objetivo caracterizar cinco diferentes formulações de fertilizantes organominerais utilizando espectroscopia de refletância de infravermelho próximo e Análise por Componentes Principais.

MATERIAL E MÉTODOS

Amostras

Cinco formulações diferentes de fertilizantes organominerais (Tabela 1) foram preparadas no Laboratório de Tecnologia de fertilizantes na Embrapa Solos.

Tabela 1. Formulações dos fertilizantes.

Organomineral	Fórmula	% MO
ORG	4-20-8	30
SUINO	3-20-0	25
FARTER	2-10-0	44
SS	2-10-0	44
TERMO	2-10-0	41

Espectroscopia no infravermelho próximo

Um total de vinte espectros de NIRS forma adquiridos para cada uma das cinco formulações de fertilizantes. Os espectros foram obtidos em um espectrômetro Perkin Elmer *Spectrum 100N* equipado com acessório de reflectância *SPIN* do Laboratório de Quimiometria em Química Analítica (LQQA) do Instituto de Química da Universidade Estadual de Campinas, com resolução espectral de 4 cm^{-1} , 16 varreduras, na faixa de 4000 a 10000 cm^{-1} . Os espectros foram obtidos em porcentagem de refletância (% R) e transformados para $\log 1/R$, que é proporcional à concentração. Posteriormente, os dados foram exportados para formato ASC e processados em Matlab 7.0 (Mathworks Co).

Análise por Componentes Principais

Diferentes métodos de pré-processamento de espectros NIR foram estudados: 1) primeira e 2) segunda derivada

seguidas de alisamento utilizando o algoritmo de Savitzky-Golay com 45 pontos e polinômio de segunda ordem, e correção de espalhamento multiplicativo (Multiplicative Scatter Correction, MSC) foram estudados. A PCA foi aplicada utilizando o algoritmo SVD após os dados terem sido centrados na média.

RESULTADOS E DISCUSSÃO

A Figura 1 apresenta um gráfico contendo os cem espectros NIR sem nenhum pré-processamento.

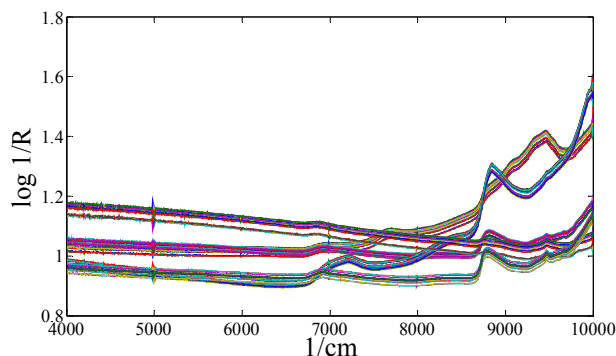


Figura 1. Conjunto de espectros NIR sem pré-processamento das amostras de fertilizantes.

Foi selecionada uma faixa espectral de interesse com o objetivo de eliminar informações irrelevantes do ponto de vista químico.

A Figura 2 apresenta o conjunto de dados pré-processados pela primeira derivada.

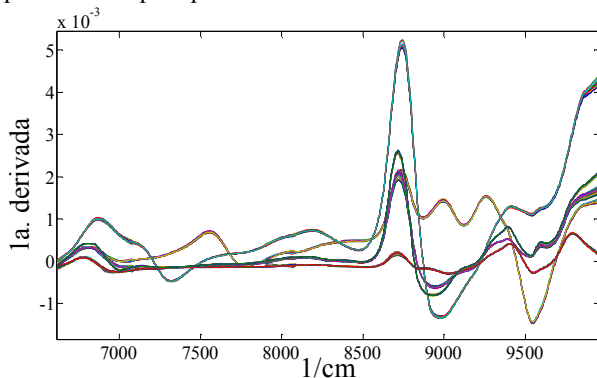


Figura 2. Primeira derivada do conjunto de dados.

Foi possível observar que os deslocamentos de linha de base e ruído foram minimizados pela aplicação desta técnica de pré-processamento. A Figura 3 apresenta o conjunto de dados pré-processados pela segunda derivada.

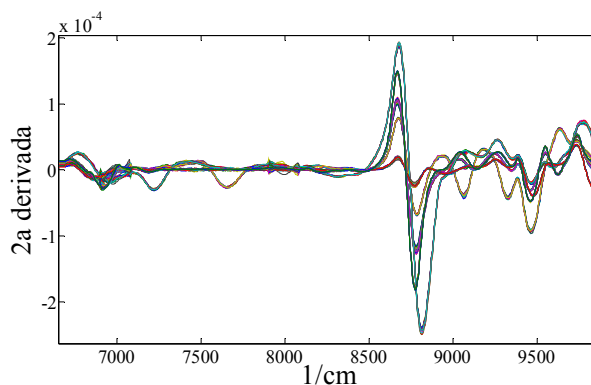


Figura 3. Segunda derivada do conjunto de dados.

Foi possível observar que inclinação de linha de base remanescente foi eliminada pela aplicação da segunda derivada. A Figura 4 apresenta o conjunto de dados pré-processados por MSC.

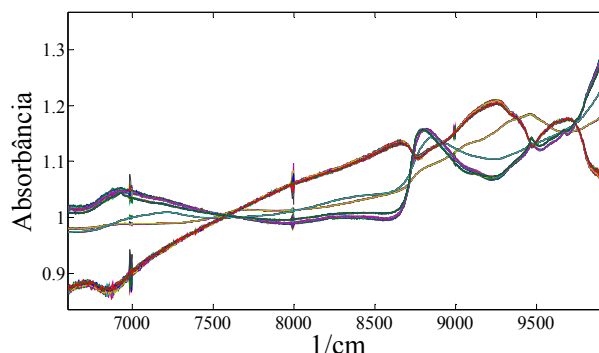


Figura 4. Espectros após correção por MSC.

A escolha do pré-processamento adequado para o conjunto dados foi realizada da seguinte forma: o pré-processamento envolvendo a primeira derivada foi descartado por não remover os efeitos de inclinação de linha de base. Portanto, o pré-processamento utilizado para a construção dos modelos de PCA foi a segunda derivada com alisamento, pois os espectros apresentaram pouco ou nenhum deslocamento de linha de base e o ruído instrumental foi minimizado. Portanto, após a aplicação da segunda derivada, os dados foram centrados na média e o modelo de PCA foi construído. O gráfico dos *scores* da PC1 vs PC2 é mostrado na Figura 5. Neste gráfico foi possível observar a formação de agrupamentos distintos entre os diferentes fertilizantes. A PC1 separa as amostras dos fertilizantes ORG e TERMO (localizados na porção negativa) das amostras de fertilizantes SUINO e SS (localizados na porção positiva) enquanto a PC2 separa as amostras FARTER, TERMO e SS (localizadas na porção positiva) das amostras ORG e SUINO (localizadas na porção negativa).

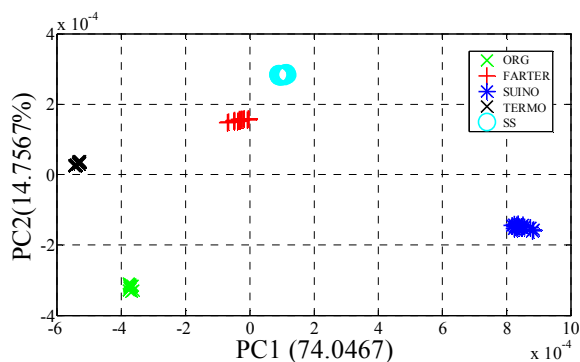


Figura 5. Gráfico dos scores da PC1 vs PC2.

O gráfico dos *loadings* mostrado na Figura 6 apresenta as variáveis de 8677 e 8817 cm^{-1} como sendo as mais importantes para a PC1.

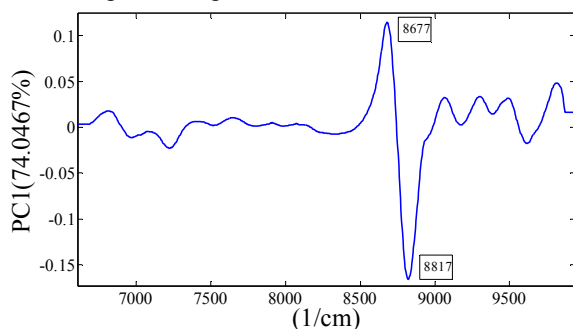


Figura 6. Gráfico dos *loadings* na PC1.

O gráfico dos *loadings* na PC2 (Figura 7) apresenta as variáveis 8753 e 9637 cm^{-1} como sendo as mais importantes para este PC.

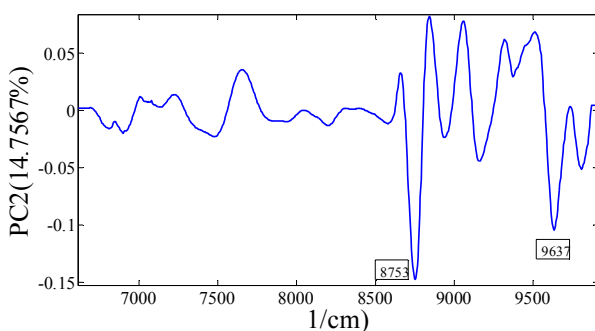


Figura 7. Gráfico dos *loadings* na PC2.

CONCLUSÕES

1. O método de processamento mais indicado para o conjunto de dados é o da segunda derivada seguido de alisamento.
2. A técnica espectroscópica NIR é adequada para um estudo de caracterização de fertilizantes devido a sua simples aplicação.
3. Através da PCA é possível verificar a existência de diferenças espectrais entre as formulações e possível verificar quais variáveis são as mais importantes. Portanto, esta é uma ferramenta muito importante para controle de matéria prima e do processo de produção de fertilizantes.

AGRADECIMENTOS

Os autores agradecem ao Etelvino Henrique Novothy e Marcelo Francisco Costa Saldanha (Embrapa Solos), a Felícia Miranda de Jesus e Rojane Chapeta Peixoto Santos (LASP, Embrapa Solos) ao Guilherme Post Sabin e Márcia Cristina Breikreitz (LQQA, IQ-Unicamp) pelo constante apoio, e os órgãos de fomento a pesquisa CAPES, CNPQ, FAPERJ e FAPESP pelo suporte financeiro.

REFERÊNCIAS

- ALBRECHT, R., JOFFRE, R., GROS, R., LE PETIT, J., TERROM, G., PERISSOL, C. Efficiency of near-infrared reflectance spectroscopy to assess and predict the stage of transformation of organic matter in the composting process. *Bioresour. Technol.*, 448-455, 2008.
- FEARN, T., RICCIOLI, C., GARRIDO-VARO, A., GUERRERO-GINEL, J.E. On the geometry of SNV and MSC. *Chemom. Intell. Lab. Syst.*, 22-26, 2009.
- FRANCOSO, O., CIAVATTA, C., SANCHEZ-CORTES, S., TUGNOLI, V., SITTI, L., GESSA, C. Spectroscopic characterization of soil organic matter in long-term amendment trials. *Soil Sci.*, 495-504, 2000.
- GISHEN, M., COZZOLINO, D. Feasibility study on the potential of visible and near infrared reflectance spectroscopy to measure alpaca fibre characteristics. *Animal*, 899-904, 2007.
- KIEHL, E. J. Fertilizantes Organominerais. 2.ed. Piracicaba, Degaspari, 2008. 160p.
- MORON, A., COZZOLINO, D. Determination of potentially mineralizable nitrogen and nitrogen in particulate organic matter fractions in soil by visible and near-infrared reflectance spectroscopy. *Journal of Agricultural Science*, 335-343, 2004.
- MOUAZEN, A.M., KAROUI, R., DE BAERDEMAEKER, J., RAMON, H. Characterization of soil water content using measured visible and near infrared spectra. *Soil Sci. Soc. Am. J.*, 1295-1302, 2006.
- WOLD, S., ESBENSEN, K., GELADI, P. Principal component analysis. *Chemom. Intell. Lab. Syst.*, 37-52, 1987.