



ESTRUTURAÇÃO METODOLÓGICA DE UMA ABORDAGEM MIR-QUIMIOMÉTRICA PARA SEDIMENTOS DE CÓRREGOS EM ÁREAS DE PRESERVAÇÃO AMBIENTAL DO CERRADO

Joaquim Tomás da Conceição Santos¹; Fabiana Casarin²; Silvia Keli de Barros Alcanfor³; Wellington Bruno dos Santos Alves⁴

Agronomia – Universidade Católica de Brasília;

²Docente – Universidade Católica de Brasília;

³Troplia Ltda.;

⁴Pesquisador de Pós-doutorado – Universidade Católica de Brasília.

E-mail: joaquim.santos@a.ucb.br; fabiana.casarin@p.ucb.br; silviaalcanfor@gmail.com; dr.wellington.bio@gmail.com

RESUMO

Os sedimentos aquáticos registram alterações geoquímicas, hidrológicas e antrópicas, podendo complementar séries de monitoramento ambiental. Este trabalho correspondeu a uma etapa inicial, de aproximadamente três meses, destinada a estruturar uma abordagem baseada em espectroscopia no infravermelho médio (MIR) e quimiometria para sedimentos de seis córregos da APA Gama-Cabeça de Veado. O estudo concentrou-se na revisão metodológica, na organização das matrizes X e Y, na definição de pré-processamentos e nos critérios para PCA, PLS, PLS-DA e validação. Como resultado parcial, consolidou-se um protocolo analítico que utiliza a experiência da espectroscopia MIR aplicada a solos como referência metodológica, sem pressupor a transferência direta dos modelos para sedimentos. A experimentação teve início em agosto de 2026, incorporando o protocolo desenvolvido à continuidade experimental.

Palavras-chave: quimiometria. Espectroscopia MIR. Sedimentos.

INTRODUÇÃO

Os sedimentos aquáticos funcionam como arquivos ambientais, pois acumulam sinais erosivos, hidrológicos, geoquímicos e antrópicos por períodos superiores aos captados por medições pontuais da água. Na APA Gama-Cabeça de Veado, as séries históricas do PELD-AGCV tornam sua análise relevante para interpretar mudanças espaciais e temporais nos córregos Capetinga, Onça, Gama, Monjolo, Cabeça de Veado e Bonito (Fonseca & de Mendonça-Galvão, 2014; Marques et al., 2022).

A espectroscopia MIR produz assinaturas com numerosas medidas de absorbância ou refletância ao longo do número de onda. Em solos e sedimentos, essas respostas associam-se à matéria orgânica, água estrutural, argilominerais, carbonatos e silicatos, mas também são influenciadas pela granulometria, heterogeneidade, contato com o acessório ATR e preparo. Portanto, o ganho analítico depende da organização dos dados, do pré-processamento e da modelagem multivariada (Armenta & de la Guardia, 2014).

Estudos com solos relacionam carbono orgânico, textura e outros atributos ao perfil espectral por modelos multivariados, desde que haja amostras representativas, dados de referência confiáveis e validação adequada (Baumann et al., 2021; Hati et al., 2022). Contudo, solos e sedimentos diferem em mineralogia, matéria orgânica, granulometria, deposição e armazenamento. Assim, os estudos em solos orientam o método, mas não autorizam transferir coeficientes ou calibrações.

A PCA reduz a dimensionalidade e permite reconhecer tendências, amostras anômalas e relações entre scores e loadings. O PLS prevê variáveis quantitativas de referência, enquanto o PLS-



DA classifica categorias previamente definidas e exige validação da capacidade de generalização. Desse modo, a quimiometria transforma a pergunta ambiental em matriz de dados, modelo estatístico e critério de decisão (Souza & Poppi, 2012; Ferreira, 2026).

OBJETIVO

O objetivo deste estudo foi estruturar, na etapa inicial do projeto, uma estratégia MIR-quimiométrica para a futura caracterização de sedimentos da série histórica do PELD-AGCV, definindo a organização das matrizes espectrais e de referência, os pré-processamentos, a aplicação de PCA, PLS e PLS-DA e os critérios de validação necessários à continuidade experimental iniciada em agosto de 2026.

MÉTODO

A etapa metodológica, com duração de três meses, abrangeu revisão da literatura, estruturação dos dados e desenho do protocolo analítico. A experimentação foi iniciada em agosto de 2026. Cada espectro MIR constitui um vetor de primeira ordem. Ao empilhar n amostras nas linhas e p números de onda nas colunas, forma-se a matriz preditora $X (n \times p)$. Essa organização segue a lógica das aplicações de PCA e calibração multivariada (Souza & Poppi, 2012; Valderrama et al., 2009). A Tabela 1 apresenta uma representação conceitual, sem dados experimentais.

Tabela 1

Representação conceitual das matrizes X (MIR) e Y para sedimentos

Amostra	v_1	v_2	...	v_p	Y quantitativo	Y classe
Capetinga-2017	x_{11}	x_{12}	...	x_{1p}	pH, Corg, textura	classe A/B/C
Onça-2019	x_{21}	x_{22}	...	x_{2p}	pH, Corg, textura	classe A/B/C
...	...	...	...	...	...	...
Bonito-2025	x_{n1}	x_{n2}	...	x_{np}	pH, Corg, textura	classe A/B/C

Nota. v representa o número de onda; x_{ij} , a intensidade espectral da amostra i na variável j . A matriz Y contém somente variáveis de referência disponíveis e confiáveis. O exemplo é conceitual e não corresponde a dados experimentais. Fonte: elaboração própria (2026).

Antes da modelagem, devem-se comparar pré-processamentos que reduzam variações físicas ou instrumentais sem eliminar informação química. Na SNV, cada espectro é centralizado por sua média e dividido pelo desvio-padrão. A centralização por variável poderá preceder a PCA/PLS, e derivadas de Savitzky-Golay poderão reduzir a linha de base e realçar bandas sobrepostas.

$$\text{SNV: } x'_{ij} = (x_{ij} - \bar{x}_i) / s_i \text{ (normalização de cada espectro)} \quad (1)$$

Na PCA, a matriz pré-processada é decomposta em scores, loadings e resíduos. Os scores posicionam as amostras no espaço multivariado, e os loadings indicam as regiões espectrais relevantes, permitindo investigar agrupamentos, gradientes e anomalias sem impor classes prévias.

$$\text{PCA: } X = TP^T + E \text{ (} T = \text{scores; } P = \text{loadings; } E = \text{resíduos)} \quad (2)$$

Quando houver valores quantitativos de referência, como pH, carbono orgânico ou frações texturais, indica-se o PLS. O modelo busca combinações latentes de X com covariância em relação a Y .



A qualidade não se fundamenta apenas em R^2 , mas também na seleção do número de variáveis latentes e na avaliação dos erros de calibração, validação cruzada e predição (Valderrama et al., 2009).

$$\text{PLS: } \hat{y} = Xb + b_0 \text{ (predição de uma propriedade quantitativa)} \quad (3)$$

$$\text{RMSEP} = \sqrt{[\Sigma(y_i - \hat{y}_i)^2 / n_{\text{pred}}]} \text{ (erro de predição simplificado)} \quad (4)$$

Para classificação, o PLS-DA é aplicável somente quando as classes ambientais são definidas por critérios independentes do espectro. Em uma codificação de três classes, Y pode conter os vetores $[1,0,0]$, $[0,1,0]$ e $[0,0,1]$. Ferreira (2026) descreve a regressão discriminante pela relação matricial $Y = XB + E$. O desempenho deve ser avaliado com dados não utilizados no ajuste, por matriz de confusão, sensibilidade e especificidade.

$$\text{PLS-DA: } Y = XB + E \text{ (Y codificada por classes)} \quad (5)$$

$$\text{Sensibilidade} = \text{VP}/(\text{VP} + \text{FN}); \text{Especificidade} = \text{VN}/(\text{VN} + \text{FP}) \quad (6)$$

RESULTADOS

Como resultado, definiram-se o problema analítico, as ferramentas compatíveis com cada pergunta e as condições mínimas para que os futuros modelos sejam interpretados cientificamente.

Essa etapa é relevante porque, na espectroscopia MIR, a quantidade de variáveis é muito superior à de propriedades ambientais de interesse. Sem estrutura prévia, podem surgir separações visuais ou correlações aparentes que não se sustentam fora do conjunto usado para construir o modelo. A organização explícita das matrizes X e Y evita misturar variáveis preditoras, de referência e de classe. A definição antecipada da validação cruzada e do conjunto externo também reduz o risco de sobreajuste e vazamento de informações entre treinamento e teste.

A literatura de solos reforçou a viabilidade do desenho, mas recomendou cautela: assinaturas MIR podem estimar atributos do solo por modelos multivariados (Baumann et al., 2021; Hati et al., 2022), enquanto a heterogeneidade e a deposição ampliam a variabilidade dos sedimentos. A fase experimental deverá testar repetibilidade, granulometria, armazenamento e estabilidade do controle interno antes de atribuir diferenças espectrais a processos ambientais. Calibrações de solos exigirão construção e validação locais para uso em sedimentos.

A sequência metodológica consolidada foi: (1) controle e preparo das amostras; (2) aquisição MIR em replicata; (3) construção de X e organização de Y ; (4) comparação justificada de pré-processamentos; (5) PCA para exploração; (6) PLS apenas para respostas quantitativas de referência; (7) PLS-DA apenas para classes independentes e previamente definidas; e (8) validação compatível com a finalidade do modelo. A Tabela 2 resume a transformação da informação espectral em decisão analítica.

Tabela 2

Etapas do protocolo MIR-quimiométrico e interpretação esperada

Etapas	Equação/critério	Interpretação
Matriz MIR	$X(n \times p)$	Base para PCA/PLS.
Pré-processamento	$x' = (x - \bar{x})/s$	Reduz variações físicas/instrumentais.
PCA	$X = TP^T + E$	Explora scores, loadings e outliers.
PLS	$\hat{y} = Xb + b_0$	Prediz propriedades quantitativas.
Validação	RMSEP	Mede o erro de predição externa.
PLS-DA	$Y = XB + E$	Classifica categorias predefinidas.



Nota. As equações são apresentadas de forma didática e resumida; a implementação final deverá respeitar o algoritmo, o pré-processamento e a estratégia de validação selecionados para cada conjunto de dados. Fonte: elaboração própria (2026).

CONCLUSÃO

O estudo transformou a proposta de aplicação da quimiometria em protocolo analítico com perguntas, matrizes, métodos e critérios de validação definidos. A literatura sobre MIR em solos pode orientar o desenho experimental para sedimentos, mas não autoriza transferir calibrações entre matrizes. Esse amadurecimento metodológico foi incorporado à fase iniciada em agosto de 2026. Na continuidade, a prioridade será gerar espectros reproduzíveis, verificar a estabilidade das amostras e do sistema instrumental e, posteriormente, construir os modelos de PCA, PLS e PLS-DA. Assim, os modelos poderão ser avaliados não apenas pela separação visual ou pelo ajuste matemático, mas pela capacidade de produzir interpretações ambientais consistentes e generalizáveis para a série histórica do PELD-AGCV.

REFERÊNCIAS

- Armenta, S., & de la Guardia, M. (2014). Vibrational spectroscopy in soil and sediment analysis. *Trends in Environmental Analytical Chemistry*, 2, 43–52. <https://doi.org/10.1016/j.teac.2014.05.001>
- Baumann, P., Lee, J., Frossard, E., Schönholzer, L. P., Diby, L., Hgaza, V. K., Kiba, D. I., Sila, A., Shepherd, K., & Six, J. (2021). Estimation of soil properties with mid-infrared soil spectroscopy across yam production landscapes in West Africa. *SOIL*, 7(2), 717–731. <https://doi.org/10.5194/soil-7-717-2021>
- Ferreira, M. M. C. (2026). Quimiometria IV. Reconhecimento de padrões: métodos discriminativos e modelativos. Quando utilizá-los e qual deles. *Química Nova*, 49, e-20260072. <https://doi.org/10.21577/0100-4042.20260072>
- Fonseca, B. M., & de Mendonça-Galvão, L. (2014). Pristine aquatic systems in a Long Term Ecological Research (LTER) site of the Brazilian Cerrado. *Environmental Monitoring and Assessment*, 186(12), 8683–8695. <https://doi.org/10.1007/s10661-014-4035-8>
- Hati, K. M., Sinha, N. K., Mohanty, M., Jha, P., Londhe, S., Sila, A., Towett, E., Chaudhary, R. S., Jayaraman, S., Coumar, M. V., Thakur, J. K., Dey, P., Shepherd, K., Muchhala, P., Weulow, E., Singh, M., Dhyani, S. K., Biradar, C., Rizvi, J., Patra, A. K., & Chaudhari, S. K. (2022). Mid-infrared reflectance spectroscopy for estimation of soil properties of Alfisols from Eastern India. *Sustainability*, 14(9), 4883. <https://doi.org/10.3390/su14094883>
- Marques, N. C. S., Machado, R. B., Aguiar, L. M. S., Mendonça-Galvão, L., Tidon, R., Vieira, E. M., Marini-Filho, O. J., & Bustamante, M. (2022). Drivers of change in tropical protected areas: Long-term monitoring of a Brazilian biodiversity hotspot. *Perspectives in Ecology and Conservation*, 20(2), 69–78. <https://doi.org/10.1016/j.pecon.2022.02.001>
- Souza, A. M., & Poppi, R. J. (2012). Experimento didático de quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: Um tutorial, parte I. *Química Nova*, 35(1), 223–229.
- Valderrama, P., Braga, J. W. B., & Poppi, R. J. (2009). Estado da arte de figuras de mérito em calibração multivariada. *Química Nova*, 32(5), 1278–1287.

Financiamento: PIC/UCB – Universidade Católica de Brasília.