

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DO SOLO

**ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO
PARA ESTIMAR A CONCENTRAÇÃO DE ELEMENTOS EM SOLOS E
SEDIMENTOS EM UMA BACIA HIDROGRÁFICA RURAL**

**Gabriela Naibo
(Dissertação de Mestrado)**

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DO SOLO

**ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO
PARA ESTIMAR A CONCENTRAÇÃO DE ELEMENTOS EM SOLOS E
SEDIMENTOS EM UMA BACIA HIDROGRÁFICA RURAL**

GABRIELA NAIBO
Engenheira Florestal (Unoesc)

Dissertação apresentada como um
dos requisitos à obtenção do Grau
de Mestra em Ciência do Solo

Porto Alegre (RS) Brasil
Março de 2021

CIP - Catalogação na Publicação

Naibo, Gabriela
ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO PARA
ESTIMAR A CONCENTRAÇÃO DE ELEMENTOS EM SOLOS E
SEDIMENTOS EM UMA BÁCIA HIDROGRÁFICA RURAL / Gabriela
Naibo. -- 2021.
94 f.
Orientador: Tales Tiecher.

Coorientador: Jean Michel Moura-Bueno.

Dissertação (Mestrado) -- Universidade Federal do
Rio Grande do Sul, Faculdade de Agronomia, Programa de
Pós-Graduação em Ciência do Solo, Porto Alegre, BR-RS,
2021.

1. Espectroscopia de reflectância. 2. quimiometria.
3. machine learning. 4. pré-processamento espectral.
5. elementos químicos. I. Tiecher, Tales, orient. II.
Moura-Bueno, Jean Michel, coorient. III. Título.

GABRIELA NAIBO
Engenheira Florestal (UNOESC)

DISSERTAÇÃO

Submetida como parte dos requisitos
para obtenção do Grau de

MESTRE EM CIÊNCIA DO SOLO

Programa de Pós-Graduação em Ciência do Solo

Faculdade de Agronomia

Universidade Federal do Rio Grande do Sul

Porto Alegre (RS), Brasil

Aprovado em: **19/03/2021**
Pela Banca Examinadora

Homologado em: **26/08/2021**
Por



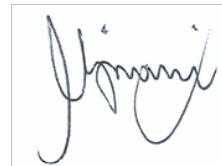
TALES TIECHER
Orientador-PPG Ciência do Solo



ALBERTO VASCONCELLOS INDA JUNIOR
Coordenador do
Programa de Pós-Graduação em
Ciência do Solo

CLÁUDIA ALESSANDRA PEIXOTO DE BARROS
UFRGS
(MConf UFRGS)

YURI JACQUES AGRA BEZERRA DA SILVA
UFPI
(MConf UFRGS)



LAURENT CANER
Université de Poitiers/França
(MConf UFRGS)

CARLOS ALBERTO BISSANI
Diretor da Faculdade
de Agronomia

" A tarefa não é ver o que ninguém ainda viu, mas pensar o que ninguém pensou sobre aquilo que todo mundo vê..."

Arthur Schopenhauer

*Dedico aos meus pais,
Rosani e Altair e
ao meu amado Everson,
amo vocês além da vida!*

AGRADECIMENTOS

À Deus e ao plano espiritual, pelo dom da vida, pelas bênçãos concedidas durante esta trajetória até a realização de mais um objetivo, da concretização de mais uma etapa, do sonho de ser Mestre, enfim, por tudo.

Aos meus pais, Altair Naibo e Rosani F. Dilberti Naibo, meus heróis e meus exemplos, agradeço a cada dia por estarem ao meu lado, nas dificuldades e comemorações, sempre e em tudo. Jamais esquecerei o esforço de vocês para realizar este sonho. Amo vocês eternamente!

Ao meu namorado/esposo Everson Kuhn Sbruzzi, por ter me apoiado em todos os momentos, com amor, dedicação, carinho, e principalmente paciência! Obrigada por tudo.

À minha família, em especial aos meus avós Lídio, Irma, Nadir e Itacir (*In memoriam*), minha tia Rosimara e primo Eduardo, por sempre se preocuparem comigo e me amarem sem medida, grata por tudo sempre!

Ao meu orientador, Prof. Dr. Tales Tiecher, por todos os ensinamentos, pela experiência acadêmica, pelo convívio, pela amizade e pela oportunidade de me orientar nesta jornada. Sua orientação é inspiradora e faz com que acreditemos que pode haver um mundo melhor através da educação. Grata eternamente professor!

Ao meu coorientador, Dr. Jean Michel Moura-Bueno por todos os ensinamentos repassados durante este período, sempre com sabedoria e paciência, sou eternamente grata a ti!

Aos meus amigos e parceiros de lida da espectroscopia, Rafael Ramon e Gustavo Pesini, pela troca de aprendizado dentro desta linha de pesquisa, obrigada por tudo e contem sempre comigo!!!

Aos amigos que fiz durante o mestrado e que devo eterna gratidão a estes, Ana Paula, Esraelda, Felipe, Luciana, Thamires e Prof. Claudia, obrigada por tudo que fizeram por mim, pela acolhida e paciência nos momentos difíceis.

Aos colegas do Programa de Pós-Graduação em Ciência do Solo da UFRGS e do Grupo de Pesquisa IRGEB, pela parceria e amizade de sempre.

Aos meus irmãos do coração que sempre estão comigo em todas as batalhas, Ana Carolina, Jonas, Rodrigo e Jaqueline obrigada por todo o apoio e coragem que depositam em mim.

Aos Professores do Programa de Pós-graduação em Ciência do Solo por todo o conhecimento repassado durante esta etapa da minha vida, sempre digo: sem professor não existe doutor, gratidão por tudo!

Ao CNPq pela concessão da bolsa de mestrado, a qual teve fundamental importância para conclusão desta etapa.

Aos que de alguma forma contribuíram para o desenvolvimento deste mestrado e que não foram mencionados, mas saibam que todos têm um lugar especial no meu coração, sou grata eternamente por tudo.

ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO PARA ESTIMAR A CONCENTRAÇÃO DE ELEMENTOS EM SOLOS E SEDIMENTOS EM UMA BACIA HIDROGRÁFICA RURAL¹

Autora: Gabriela Naibo
Orientador: Prof. Dr. Tales Tiecher
Coorientador: Dr. Jean Michel Moura-Bueno

RESUMO

A estimativa da concentração de elementos em solos e sedimentos é essencial no monitoramento da transferência de nutrientes e poluentes dos ambientes terrestres para os aquáticos. Contudo, o custo do monitoramento detalhado multielementar baseado em análises químicas convencionais para séries grandes de amostras em bacias hidrográficas pode ser impeditivo para muitos centros de pesquisa. Dessa forma, é preciso desenvolver técnicas de baixo custo para determinar a concentração de elementos em amostras de solo e sedimento. O objetivo desse estudo foi avaliar quais são as melhores combinações de métodos multivariados e técnicas de pré-processamento para aumentar a acurácia da estimativa da concentração de elementos em amostras de solo e sedimento utilizando o infravermelho próximo (NIR). Adicionalmente, foi avaliado o uso da espectroscopia na faixa do NIR para estimar a concentração de elementos presentes no sedimento usando modelos calibrados com amostras de solo de uma mesma bacia hidrográfica. O banco de dados de solo e sedimento utilizado é oriundo de amostras coletadas na bacia hidrográfica do Rio Guaporé, no Sul do Brasil. Um total de 316 amostras de solo e 196 de sedimento, foram secas, desagregadas e peneiradas em malha de 63 µm. O carbono orgânico (C) foi quantificado pela oxidação úmida e a concentração pseudo-total dos elementos (Al, Ba, Be, C, Ca, Co, Cr, Cu, Fe, K, La, Li, Mg, Mn, Na, Ni, P, Pb, Sr, Ti, V e Zn) foi quantificada por ICP-OES (*Inductively Coupled Plasma Optical Emission Spectrometry*) após digestão assistida por micro-ondas com HCl e HNO₃ (3:1). Os espectros NIR (1000-2500 nm) foram obtidos em todas as amostras de solo e sedimento. Os métodos multivariados de *machine learning* testados foram: *Partial Least Squares Regression* (PLSR) e *Support Vector Machine* (SVM), combinados com três pré-processamentos dos espectros: *Detrend* (DET), *Savitzky-Golay Derivate* (SGD) e *Standard Normal Variate* (SNV), comparados aos espectros brutos (RAW). Essas combinações foram aplicadas ao conjunto de amostras de solo e de sedimento isoladamente, e combinando ambos os conjuntos (solo + sedimento). O desempenho dos modelos e pré-processamentos foram avaliados pelo coeficiente de determinação (R²), raiz quadrada do erro médio (RMSE), e pela relação entre o desempenho e a distância interquartil (RPIQ). O modelo SVM resultou em melhores previsões comparado ao PLSR para solo, sedimento e solo+sedimento. O efeito do pré-processamento aumentou a acurácia das estimativas na seguinte ordem: RAW<SNV<DET=SGD. Os melhores modelos foram obtidos combinando o método de calibração multivariada SVM com o pré-processamento SGD. O ajuste dos modelos calibrados separadamente

¹Dissertação de Mestrado em Ciência do Solo. Programa de Pós-Graduação em Ciência do Solo, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul. Porto Alegre. (93p.) Março, 2021.

utilizando amostras de solo ($R^2 = 0,88$) e sedimento ($R^2 = 0,89$) foi superior comparado aos modelos calibrados combinando amostras de solo e sedimento ($R^2 = 0,85$). Na bacia hidrográfica estudada, a predição da concentração dos elementos presentes em sedimentos através de modelos calibrados com amostras de solo teve baixa acurácia, com R^2 médio de 0,11 (variando de 0,00 a 0,46). Portanto, é necessário calibrar modelos separadamente para estimar a concentração de elementos em solo e sedimento com maior acurácia. Os resultados demonstraram que é possível construir modelos espectroscópicos acurados para predizer a concentração de elementos tanto em amostras de solo quanto em amostras de sedimento para todos os elementos avaliados.

Palavras-chave: espectroscopia de reflectância, quimiometria, *machine learning*, pré-processamento espectral, sedimentos, elementos químicos.

NEAR INFRARED SPECTROSCOPY TO ESTIMATE THE CONCENTRATION OF ELEMENTS IN SOILS AND SEDIMENTS IN A RURAL CATCHMENT²

Authora: Gabriela Naibo
Adviser: Prof. Tales Tiecher
Coadviser: Dr. Jean Michel Moura-Bueno

ABSTRACT

Estimating the concentration of elements in soils and sediments is essential for monitoring the transfer of nutrients and pollutants from terrestrial to aquatic environments. However, the cost of detailed multi-element monitoring based on conventional chemical analysis for large series of samples can be cost prohibitive for many research center. Thus, it is necessary to develop low-cost techniques to estimate the concentration of elements in soil and sediment samples. The objective of this study was to evaluate which are the best combinations of multivariate methods and pre-processing techniques to increase the accuracy of estimating the concentration of elements in soil and sediment samples using near infrared (NIR). Additionally, the use of spectroscopy in the NIR range was evaluated to estimate the concentration of elements present in the sediment using calibrated models with soil samples from the same catchment. The database of soil and sediment used were collected in the Guaporé catchment, in southern Brazil. A total of 316 soil samples and 196 sediment samples were dried, disaggregated and sieved in a 63 µm mesh. Organic carbon was quantified by wet oxidation and the pseudo-total concentration of the elements (Al, Ba, Be, Ca, Co, Cr, Cu, Fe, K, La, Li, Mg, Mn, Na, Ni, P, Pb, Sr, Ti, V and Zn) were quantified by ICP-OES (Inductively Coupled Plasma Optical Emission Spectrometry) after microwave assisted digestion with HCl and HNO₃ (3:1). The NIR spectra (1000-2500 nm) were obtained in all soil and sediment samples. The multivariate methods of machine learning tested were: partial least squares regression (PLSR) and support vector machine (SVM), combined with three spectrum pre-processing: detrend (DET), Savitzky-Golay Derivate (SGD) and standard normal variate (SNV), compared to the raw spectra (RAW). These combinations were applied to the set of soil and sediment samples alone, and combining both sets (soil + sediment). The performance of the models and pre-processing techniques were evaluated by the coefficient of determination (R²), root mean square error (RMSE), and the relationship between performance and the interquartile distance (RPIQ). The SVM model resulted in better predictions compared to the PLSR for soil, sediment, and soil + sediment. The pre-processing effect increased the accuracy of the estimates in the following order: RAW<SNV<DET=SGD. The best models were obtained by combining the SVM multivariate calibration method with SGD pre-processing. The adjustment of the models calibrated separately with data of soil samples (R² = 0.88) and sediment (R² = 0.89) was superior compared to the models calibrated combining soil and sediment samples (R² = 0.85). In the studied catchment, the prediction of the concentration of the elements presents in sediments through models calibrated

²MSc dissertation in Soil Science. Graduate Program in Soil Science, Faculty of Agronomy, Universidade Federal do Rio Grande do Sul, Porto Alegre. (93p.) March, 2021.

with soil samples had low accuracy, with an average R^2 of 0.11 (ranging from 0.00 to 0.46). Therefore, it is necessary to calibrate models separately to estimate the concentration of elements in soil and sediment with greater accuracy. The results showed that it is possible to build accurate spectroscopic models to predict the concentration of elements in soil samples as well as in sediment samples for all the evaluated elements.

Key-words: reflectance spectroscopy, chemometry, machine learning, spectral pre-processing, sediments, chemical elements.

SUMÁRIO

1. INTRODUÇÃO	17
2. REVISÃO BIBLIOGRÁFICA	19
2.1 Espectroscopia de reflectância	19
2.2 Uso da espectroscopia de reflectância em solos.....	21
2.3 Uso da espectroscopia de reflectância em sedimentos.....	26
2.4 Pré-processamentos espectrais	27
2.4.1 <i>Detrend</i> (DET)	27
2.4.2 Derivada de <i>Savitzky-Golay</i> (SGD)	27
2.4.3 <i>Standard Normal Variate</i> (SNV)	29
2.5 Métodos de calibração multivariada	29
2.5.1 <i>Partial Least Squares Regression</i> (PLSR)	29
2.5.2 <i>Support Vector Machine</i> (SVM).....	30
3. HIPÓTESE	31
4. OBJETIVO	32
5. MATERIAL E MÉTODOS	33
5.1 Caracterização da bacia hidrográfica	33
5.2 Coleta das amostras de solo e sedimento.....	36
5.3 Análises químicas.....	38
5.4 Análises espectroscópicas	39
5.5 Pré-processamento espectral.....	39
5.6 Calibração e avaliação dos modelos	40
6. RESULTADOS E DISCUSSÃO	43
6.1 Análise do comportamento espectral no NIR	43
6.2 Estatística descritiva da concentração de elementos em amostras de solo e sedimento	46
6.3 Principais fatores afetando a acurácia dos modelos.....	48
6.3.1 Efeito dos modelos multivariados	49
6.3.2 Efeito das técnicas de pré-processamento.....	51
6.3.3 Efeito do elemento.....	52
6.3.4 Efeito do conjunto de dados e predição da concentração de elementos em amostras sedimentos com modelos espectrais calibrados com amostras solo de uma bacia hidrográfica	62

7. CONCLUSÕES	67
8. REFERÊNCIAS BIBLIOGRÁFICAS	68
9. ANEXOS	78
10. RESUMO BIOGRÁFICO.....	91

RELAÇÃO DE TABELAS

Tabela 1. Constituintes do solo, seus respectivos comprimentos de onda detectadas por espectroscopia na faixa de NIR e os grupos funcionais relacionados.....	22
Tabela 2. Seleção de trabalhos com uso da espectroscopia de reflectância para quantificar a concentração de elementos em solos no mundo.....	24
Tabela 3. Seleção de trabalhos com uso da espectroscopia de reflectância para quantificar a concentração de elementos em sedimentos no mundo.....	28
Tabela 4. Número de amostras de solo e de sedimento coletadas na bacia hidrográfica do Rio Guaporé, Sul do Brasil.....	39
Tabela 5. Estatística descritiva do conjunto de amostras de solo e sedimento coletadas na bacia do Rio Guaporé.....	48
Tabela 6. Modelos espectrais com melhor desempenho preditivo para os 22 elementos avaliados calibrados utilizando amostras de solo e amostras de sedimento.....	62
Tabela 7. Seleção dos melhores pré-processamentos para cada um dos 22 elementos utilizando amostras de solo e resultado da predição dos mesmos utilizando a assinatura espectroscópica de amostras de sedimento para a faixa espectral NIR e método multivariado SVM.....	65

RELAÇÃO DE FIGURAS

Figura 1. Ilustração do processo de espectroscopia de reflectância difusa para obtenção da assinatura espectral de solos e sedimentos.....	20
Figura 2. Localização e relevo da bacia hidrográfica do Rio Guaporé, Rio Grande do Sul, Sul do Brasil.....	35
Figura 3. Classes de solo da bacia hidrográfica do Rio Guaporé, Rio Grande do Sul, Sul do Brasil.....	35
Figura 4. Uso e ocupação do solo da bacia hidrográfica do Rio Guaporé, Rio Grande do Sul, Sul do Brasil.....	36
Figura 5. Distribuição dos pontos de coleta de amostras de solo e de sedimento da bacia hidrográfica do Rio Guaporé, Rio Grande do Sul, Sul do Brasil.....	37
Figura 6. Esquema da calibração dos modelos para amostras de solo, sedimento e solo + sedimento. a) conjunto de dados de solo; b) conjunto de dados de sedimento; c) conjunto de dados de solo + sedimento; d) melhores modelos calibrados com amostras de solo.....	43
Figura 7. Assinaturas espectrais médias do conjunto de amostras de solo e sedimento da bacia hidrográfica do Rio Guaporé para a faixa espectral NIR bruto sem nenhum pré-processamento (a) RAW e com os pré-processamentos, (b) <i>Detrend</i> (DET), (c) <i>Standard Normal Variate</i> (SNV), (d) <i>Savitsky-Golay derivative</i> (SGD).....	45
Figura 8. Árvore de regressão de inferência condicional utilizando os valores de R^2 para as variáveis de método multivariado, elemento e pré-processamento espectral.....	49
Figura 9. Árvore de regressão de inferência condicional utilizando os valores de RPIQ para as variáveis de método multivariado, elemento e pré-processamento espectral.....	50
Figura 10. Comparação do desempenho dos métodos de calibração multivariada indicado pelos valores de R^2 (a) e RPIQ (b). As barras representam valores médios e o desvio padrão de 264 modelos. PLSR: <i>Partial Least Squares Regression</i> . SVM: <i>Support Vector Machine</i>	51
Figura 11. Comparação do desempenho dos modelos calibrados com diferentes técnicas de pré-processamento indicado pelos valores de R^2 (a) e RPIQ (b). As	

barras representam valores médios e o desvio padrão de 132 modelos. RAW: espectros brutos, DET: *Detrend*, SNV: *Standard Normal Variate*, SGD: *Savitzky-Golay derivative*.....53

Figura 12. Resposta média dos quatro pré-processamentos e dos dois modelos analisados para predição de cada elemento de acordo com a faixa espectral NIR. a) R² e RPIQ para a faixa espectral NIR; b) R² e RPIQ para os métodos multivariados SVM e PLSR; c) R² e RPIQ para os pré-processamentos DET, SGD e SNV, mais os dados brutos sem pré-processamento.....54

Figura 13. Espectro NIR processado com SGD (a) e importância relativa das bandas na predição pelo modelo SVM-SGD na concentração dos elementos carbono (C), cromo (Cr), ferro (Fe), fósforo (P) e sódio (Na).....60

Figura 14. Comparação do desempenho dos modelos calibrados com conjuntos de dados de sedimento, solo e solo + sedimento indicado pelos valores de R² (a) e RPIQ (b). As barras representam valores médios e o desvio padrão de 176 modelos.....64

RELAÇÃO DE ABREVIATURAS

PLSR – *Partial Least Squares Regression*;

SVM – *Support Vector Machine*;

RMSE – Raiz quadrada do erro médio;

R^2 – Coeficiente de determinação;

RPIQ – Relação entre o desempenho interquartil;

nm – nanômetros;

SNV – *Standard Normal Variate*;

SGD – *Savitsky-Golay Derivative*;

DET – *Detrend*;

RAW – espectros brutos;

NIR – Espectroscopia de Infravermelho Próximo;

1. INTRODUÇÃO

A estimativa da concentração de elementos químicos em solos e sedimentos é imprescindível em estudos de monitoramento ambiental em bacias hidrográficas quando se busca compreender o processo de transporte de elementos do ambiente terrestre para os aquáticos. Contudo, um inventário detalhado em escala espacial e temporal da concentração de elementos em solos e sedimentos resulta em aumento dos custos de análises laboratoriais, necessidade de infraestrutura laboratorial mais robusta, devido à grande quantidade de amostra analisadas, além do custo de pessoas envolvidas, tempo e recursos financeiros. Esses fatores podem ser impeditivos em muitos centros de pesquisa no mundo. Além disso, muitas das análises para determinação da concentração de elementos em solos e sedimentos geram grande quantidade de resíduos químicos, que posteriormente devem ser tratados. Como alternativa, métodos baseados na espectroscopia de reflectância difusa têm sido utilizados para estimar a concentração de elementos químicos em solos e sedimentos nas últimas décadas (VISCARRA ROSSEL et al., 2006). Trata-se de um método que permite a estimativa de concentrações de vários elementos de forma rápida e não destrutiva, além de diminuir consideravelmente o consumo de reagentes químicos.

A espectroscopia de infravermelho próximo (NIR) acoplada com quimiometria, pode fornecer ferramentas exclusivas para o uso direto e simultâneo na quantificação de vários elementos químicos em amostras de solo (ARMENTA; GUARDIA, 2014; NOCITA et al., 2014; KHOSRAVI et al. 2018). Esses estudos são muito difundidos para estimar propriedades de solos, especialmente relacionados com o tamanho das partículas (BILGILI et al., 2010; SUMMERS et al., 2011; DEMATTÊ et al., 2019; SANTOS et al., 2019) e o teor

de carbono orgânico total (GOMEZ et al., 2008; GUERRERO et al., 2016; DOTTO et al., 2017; MOURA-BUENO et al., 2019; MOURA-BUENO et al., 2020). Contudo, essa técnica ainda é pouco explorada para elementos químicos e ainda não há consenso sobre qual a melhor combinação de método multivariado e pré-processamento espectral para a predição da concentração de elementos de interesse. Além disso, os estudos relacionados ao uso da espectroscopia para estimar as concentrações de elementos químicos em sedimentos (NELSON; SOMMERS, 1982; MOROS et al., 2009; JIANG et al., 2018) são incipientes comparados aos estudos utilizando solo (NOCITA et al., 2014; VISCARRA ROSSEL et al., 2016; KHOSRAVI et al., 2018; DEMATTÊ et al., 2019).

Adicionalmente aos custos envolvidos com as análises químicas de laboratório, a coleta de um conjunto representativo de amostras de sedimento em bacias hidrográficas é um trabalho oneroso e por vezes muito caro, especialmente em grandes bacias hidrográficas (COLLINS et al., 2010; ARMENTA; GUARDIA, 2014). Contudo, como os sedimentos são uma mistura das principais fontes de sedimento em uma bacia hidrográfica (MINELLA; MERTEN, 2011), espera-se que seja possível calibrar modelos a partir de amostras de solo para estimar a concentração de elementos químicos em sedimentos, desde que utilizado os pré-processamentos e métodos multivariados adequados no desenvolvimento dos modelos preditivos.

2. REVISÃO BIBLIOGRÁFICA

2.1 Espectroscopia de reflectância

A espectroscopia de reflectância foi desenvolvida no início dos anos 1970, com a finalidade de analisar a umidade de grãos (como soja, milho) na área agrônômica e posteriormente se expandiu para outras áreas devido à sua eficiência e robustez (ZORNOZA; GUERRERO, 2008). Na ciência do solo, a espectroscopia de reflectância é utilizada para análise das propriedades químicas e físicas do solo desde os anos 90 (DEMATTÊ et al., 2018); na indústria alimentícia, para análise de gordura; na indústria petroquímica, para análise dos componentes do petróleo; na indústria de polímeros, para identificar e quantificar compostos e ramificações de cadeias (CHAGAS, 2006).

A reflectância espectral [$\rho(\lambda)$] é função do comprimento de onda (λ), sendo definido matematicamente como a razão entre a energia do comprimento de onda (λ) refletida pelo objeto [$E_R(\lambda)$] e a energia do comprimento de onda (λ) incidente sobre o objeto [$E_I(\lambda)$] (LILLESAND; KIEFER, 1994). O princípio básico de funcionamento da espectroscopia de reflectância consiste na emissão de um feixe de radiação, com comprimento de onda conhecido, sobre uma amostra. Esse princípio é comumente conhecido como lei de Beer-Lambert-Bouguer por ser a união dos estudos de Pierre Bouguer em 1729, Johann Heinrich Lambert em 1760 e August Beer em 1852. A lei estabelece uma relação entre a absorção de energia de um material, por exemplo, o solo e suas propriedades, quando este é atravessado por uma radiação luminosa monocromática colimada (raios luminosos paralelos). Esta radiação irá interagir com as moléculas, deixando-as em um estado excitado (STONER; BAUMGARDNER, 1981) e, parte desta energia será absorvida e parte refletida (Figura 1a).

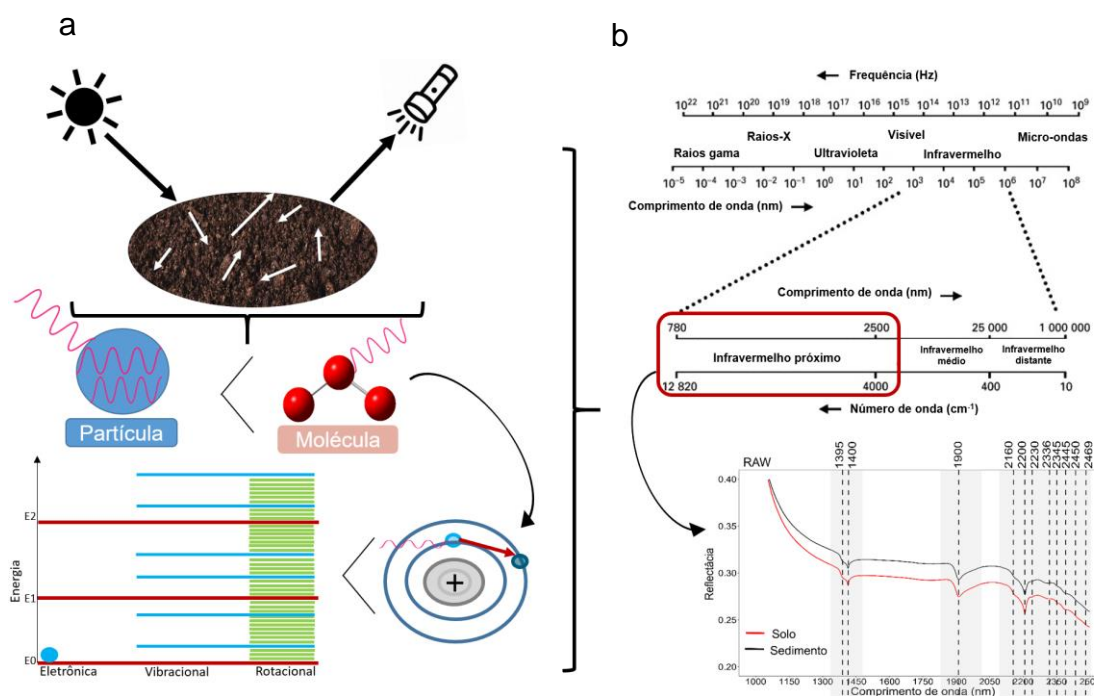


Figura 1. Ilustração do processo de espectroscopia de refletância difusa para obtenção da assinatura espectral de solos e sedimentos.

Fonte: A autora (2021).

A energia refletida é então quantificada por sensores para a obtenção do espectro (LILLESAND; KIEFER, 1994). As ondas eletromagnéticas são classificadas pelo seu comprimento de onda, frequência ou número de ondas por unidade de distância. O comprimento de onda é medido através da distância entre dois pontos equidistantes na onda e a frequência representa o número de ondas que passa por um ponto por unidade de tempo, medida em ciclos por segundo (Hertz) (NIEMZ, 2007). O espectro eletromagnético compreende todos os comprimentos de onda, das ondas menos energéticas, como as ondas de rádio, até as mais energéticas, como os raios gama (NIEMZ, 2007) (Figura 1b). Além de ser descrita na forma de ondas, a radiação eletromagnética ainda pode ser considerada uma forma de energia dinâmica, composta por partículas que carregam quantidades específicas de energia (BEN-DOR; IRONS; EPEMA, 1999). As partículas que carregam energia através das ondas eletromagnéticas são chamadas de fótons, sendo esta energia diretamente relacionada à frequência da radiação eletromagnética (NIEMZ, 2007). Sabe-se que quanto maior a frequência e menor o comprimento de onda e, conseqüentemente maior a energia contida na onda eletromagnética.

Dentre as faixas do espectro eletromagnético mais utilizadas e estudadas para estimar as propriedades dos solos, destaca-se (*Near Infrared Reflectance*) NIR, no intervalo dos comprimentos de onda compreendidos entre 780 a 2500 nm (BLANCO; VILLARROYA, 2002).

Alguns fatores como a intensidade da curva espectral; forma da curva; inclinação geral da curva espectral; aspecto ascendente, plano ou descendente; mudanças na inclinação da mesma curva; comparação de assinaturas espectrais entre horizontes do mesmo perfil, são imprescindíveis de serem analisados na assinatura espectral de um solo (DEMATTÊ et al., 2014).

2.2 Uso da espectroscopia de reflectância em solos

A espectroscopia de NIR combinado com alguns métodos de modelagem multivariada, vem sendo desenvolvida para quantificar as propriedades espectralmente ativas de amostras de solo (KHOSRAVI et al. 2018), de uma forma relativamente rápida e sem gerar resíduos.

Para ocorrer a interação entre o solo e a radiação NIR, é necessário que haja ligações covalentes e iônicas entre os átomos que compõem as moléculas do solo (NOCITA, 2014). Ainda, moléculas que absorvem energia na região do NIR, apresentam um momento dipolo elétrico associado a vibração ou rotação da molécula, o que serve como regra para selecionar os compostos que são infravermelho-ativos (STUART, 2004).

A energia absorvida pela radiação incidente é resultado das ligações presentes nas moléculas do solo, tipo de grupo funcional presente, moléculas adjacentes e presença de ligações de hidrogênio (ARMENTA; GUARDIA, 2014). A resposta espectral do solo para a radiação eletromagnética no comprimento de onda do NIR é resultado, principalmente, da interação da radiação com as moléculas H-O-H, metal-OH, O-H, C-H, N-H, S-H, O, Si-O, C-O, Fe-O, C-C, Al-O (SORIANO-DISLA et al., 2013; ARMENTA; GUARDIA, 2014). Alguns minerais e ligações associadas de acordo com a resposta espectral de diferentes constituintes do solo podem ser observadas na Tabela 1.

A espectroscopia NIR é um método direto de análise do solo, pois consegue medir diretamente as ligações de alguns constituintes do solo (Tabela 1). Posteriormente, é possível quantificar indiretamente outras propriedades do solo e sedimento, que possuem relação com as bandas de reflectância desses

constituintes organo-minerais. Porém, esta técnica necessita de métodos de referência de química úmida para calibração dos modelos utilizados para predição.

Tabela 3. Constituintes do solo, seus respectivos comprimentos de onda detectadas por espectroscopia na faixa de NIR e os grupos funcionais relacionados.

Constituintes do solo	Comprimento de onda (nm)	Grupo funcional
Água	1387 ^a	O-H
Caulinita, esmectita, mica, vermiculita hidroxientrecamada	1415 ^{b,c}	O-H
Matéria orgânica	1706 ^a	C-H
Matéria orgânica	1754 ^a	C-H
Água	1844 ^a	(H-O-H) + (O-H)
Caulinita	2160 ^{a,c}	Al-OH
Caulinita	2208 ^{a,c}	O-H
Esmectita	2206 ^{a,c}	O-H
Mica	2206 ^{a,c}	O-H
Esmectita	2230 ^{a,c}	Al-OH
Matéria orgânica (metil)	2316 ^b	C-H
Micas	2340 ^{b,c}	(O-H) + (Al-OH)
Matéria orgânica (carboidratos)	2381 ^a	C-O
Micas	2450 ^{b,c}	(O-H) + (Al-OH)

Viscarra Rossel; Behrens (2010)^a; Terra (2011)^b; Madejová et al. (2011)^c.

As análises quimiométricas se propõem a solucionar questões de origem química por meio de ferramentas que utilizam a matemática, a estatística e a computação (BEEBE; PELL; SEASHOLTZ, 1998; FERREIRA et al., 1999). As abordagens da quimiometria envolvem principalmente o planejamento e otimização de experimentos, pré-processamento de dados espectrais, reconhecimento de padrões, seleção de variáveis e amostras, calibração multivariada e transferência de calibração ou validação (BEEBE; PELL; SEASHOLTZ, 1998).

O espectro do solo, obtido pela espectroscopia NIR é utilizado para estabelecer um modelo, no qual as informações relevantes são concentradas em variáveis, com a finalidade de prever, com melhor correlação possível, as propriedades do solo (ZORNOZA; GUERRERO, 2008). Após a radiação ser

incidida na amostra de solo ou sedimento, esta radiação penetra na amostra, sofrendo múltiplas refrações e reflexões pelas partículas, resultando em uma radiação difusa, a qual é captada e pode prover informações sobre a composição organo-mineral do solo e sedimento.

Apesar das desvantagens do uso das técnicas analíticas de química úmida para a análise de solo e sedimento, como destruição da amostra, geração de resíduos químicos perigosos, ser relativamente demorada e necessitar de uma equipe qualificada para a realização do trabalho, estes ainda continuam sendo utilizados como métodos de referência. Assim, a crescente necessidade de se desenvolver métodos analíticos mais sustentáveis, que possibilitem a obtenção de dados em grande quantidade, com qualidade e que possam ser reproduzíveis se torna importante (NOCITA et al., 2014; CARVALHO, 2019).

Em contraste com as técnicas analíticas de química úmida usadas para a análise de solo e sedimento, a espectroscopia vibracional combinada com quimiometria pode fornecer ferramentas exclusivas para o uso direto e simultâneo da determinação de muitos parâmetros em amostras de solo baseados em medições de infravermelho (ARMENTA; GUARDIA, 2014). Estudos utilizando a espectroscopia de reflectância para quantificar as propriedades do solo foram desenvolvidos ao longo dos últimos anos a nível mundial (Tabela 2), avaliando uma quantidade expressiva de propriedades e elementos químicos presentes no solo. Estes estudos foram buscados na plataforma Web Of Science e filtrados pelos anos de 2010 até 2020.

Tabela 4. Seleção de trabalhos com uso da espectroscopia de reflectância para quantificar a concentração de elementos em solos no mundo.

Amostra	Elementos	Faixa espectral (nm)	Nº de amostras	Pré-processamentos	Mod. Multiv.	R ²	SEP (%)	RMSE	RPD	Referências
Solo	Carbono orgânico	350 - 2500	2471	Smoothing	PLSR	0,67	-	0,95	-	Rosin et al. (2021)
Solo	Carbono orgânico	350 - 2500	2471	Standard Normal Variate (SNV)	PLSR	0,70	-	0,91	-	Rosin et al. (2021)
Solo	Carbono orgânico	350 - 2500	2471	Savitzky Golay derivative	PLSR	0,72	-	0,87	-	Rosin et al. (2021)
Solo	Carbono orgânico	350 - 2500	280	Standard Normal Variate (SNV)	PLS	0,37	-	0,18	-	Seema et al. (2020)
Solo	Carbono orgânico	350 - 2500	280	Savitzky Golay derivative – SGD	PLS	0,30	-	0,17	-	Seema et al. (2020)
Solo	Carbono orgânico	350 - 2500	280	Savitzky Golay derivative – SGD	PLS	0,35	-	0,19	-	Seema et al. (2020)
Solo	Carbono orgânico	350 - 2500	2471	Raw Data	PLSR	0,67	-	0,91	-	Moura-Bueno et al. (2020)
Solo	Carbono orgânico	350 - 2500	2471	Savitzky-Golay Derivative (SGD)	PLSR	0,70	-	0,87	-	Moura-Bueno et al. (2020)
Solo	Carbono orgânico	350-1200	595	Savitzky-Golay Derivative (SGD)	PLSR	0,56	-	0,71	-	Dotto et al. (2019)
Solo	Carbono orgânico	350-1200	595	Raw Data	PLSR	0,75	-	0,56	-	Dotto et al. (2019)
Solo	Carbono orgânico	350-1200	595	Smoothing	PLSR	0,70	-	0,60	-	Dotto et al. (2019)
Solo	Carbono orgânico	350-1200	595	Normalization	PLSR	0,79	-	0,54	-	Dotto et al. (2019)
Solo	Carbono orgânico	350-1200	595	Detrend	PLSR	0,75	-	0,51	-	Dotto et al. (2019)
Solo	Carbono orgânico	350-1200	595	Standard Normal Variate (SNV)	PLSR	0,80	-	0,47	-	Dotto et al. (2019)
Solo	Carbono	350-2500	554	Savitzky Golay derivative	PLSR	0,86	-	0,17	2,68	Vasques, Grunwald, Sickman, (2008)
Solo	Carbono	350-1100	841	Smoothing	PLSR	0,72	-	0,56	-	Moura-Bueno et al. (2019)
Solo	Carbono	350-1100	841	Savitzky Golay derivative	PLSR	0,74	-	0,52	-	Moura-Bueno et al. (2019)
Solo	Carbono	350-1100	841	Standard Normal Variate (SNV)	PLSR	0,71	-	0,59	-	Moura-Bueno et al. (2019)
Solo	Carbono	350-1100	841	Detrend	PLSR	0,69	-	0,60	-	Moura-Bueno et al. (2019)
Solo	Carbono orgânico	-	-	1st derivada, normalização, "centering"	PLSR	0,60	-	-	-	Udelhoven et al. (2003)
Solo	Carbono orgânico	1100-2500	287	2ª derivada	PLSR	0,78	-	-	-	Martin et al. (2002)
Solo	Matéria orgânica	400-2500	125	MSC	PLS	0,93	-	-	-	He et al. (2005)
Solo	Matéria orgânica	1870 - 1038	150	1ª derivada	PLS	0,71	-	-	-	Sun et al. (2006)
Solo	Matéria orgânica	1000-2500	91	-	PLS	0,55-0,70	-	-	-	Ben-Dor; Banin, (1995)
Solo	Nitrogênio total	2240, 2090, 1901, 1827	150	1ª derivada	PLS	0,64	-	-	-	Sun et al. (2006)
Solo	Nitrogênio orgânico	1100-2500	287	2ª derivada	PLS	0,37	-	-	-	Martin et al. (2002)
Solo	Nitrogênio	400-2500	-	1st derivada, normalização, "centering"	PLSR	0,62	-	-	-	Udelhoven et al. (2003)
Solo	Arsênio (As)	350-2500	93	1st Derivative (FD) + 1st Derivative	PLSR	0,46	-	-	1,14	Cheng et al. (2019)

continuação >

continuação >

Amostra	Propriedade	Faixa espectral (nm)	Nº de amostras	Pré-processamentos	Mod. Multiv.	R ²	SEP (%)	RMSE	RPD	Referências
Solo	CaCO ₃	400-2500	52	Savitzsky Golay 1ª e 2ª derivada, BC, SNV,	PLSR	0,94	-	38,30	-	Gomes et al. (2008)
Solo	Cobre (Cu)	350-2500	33	MSC, CR	PLS	0,41	-	0,20	-	Ren et al. (2009)
Solo	Cobre (Cu)	350-2500	93	Raw Data	PLSR	0,92	-	-	1,54	Cheng et al. (2019)
Solo	Cromo (Cr)	350-2500	93	Raw Data	PLSR	0,99	-	-	1,01	Cheng et al. (2019)
Solo	Cádmio (Cd)	350-2500	93	Raw Data	PLSR	0,30	-	-	0,98	Cheng et al. (2019)
Solo	Chumbo (Pb)	350-2500	120	Suavização, MSC, MC, SNV, remoção do contínuo	PLSR	0,93	-	63,01	-	Khosravi et al. (2018)
Solo	Chumbo (Pb)	350-2500	93	1st Derivative com vetor de normalização (UVN)	PLSR	0,68	-	-	1,11	Cheng et al. (2019)
Solo	Ferro (Fe)	350-2500	33	1ª e 2ª derivada, BC, SNV, MSC, CR	PLS	0,78	-	0,08	-	Ren et al. (2009)
Solo	Ferro (Fe)	400-2500	-	1st derivada, normalização, "centering"	PLSR	0,84	-	-	-	Udelhoven et al. (2003)
Solo	Magnésio (Mg)	400-2500	-	1st derivada, normalização, "centering"	PLSR	0,91	-	-	-	Udelhoven et al. (2003)
Solo	Manganês (Mn)	400-2500	-	1st derivada, normalização, "centering"	PLSR	0,71	-	-	-	Udelhoven et al. (2003)
Solo	Potássio (K)	400-2500	-	1st derivada, normalização, "centering"	PLSR	0,85	-	-	-	Udelhoven et al. (2003)
Solo	Zinco (Zn)	-	-	Suavização, MSC, MC, SNV	PLSR	0,87	-	167,90	-	Khosravi et al. (2018)
Solo	Zinco (Zn)	350-2500	93	1st Derivative	PLSR	0,93	-	-	3,80	Cheng et al. (2019)
Solo	Umidade	1000-2500	91	-	PLS	0,55-0,70	-	-	-	Ben-Dor; Banin, (1995)
Solo	Argila	400-2500	52	Savitzsky Golay	PLS	0,85	-	31,20	-	Gomes et al. (2008)
Solo	Argila	1000-2500	91	-	PLS	0,55-0,70	-	-	-	Ben-Dor; Banin, (1995)
Solo	Área superficial específica	1000-2500	91	-	PLS	0,55-0,70	-	-	-	Ben-Dor; Banin, (1995)
Solo	Troca catiônica	1000-2500	91	-	PLS	0,550,70	-	-	-	Ben-Dor; Banin, (1995)

*nm: nanômetros; Mod. Mult.: Modelos multivariados; R²: Coeficiente de determinação; SEP: Coeficiente de correlação de Spearman; RMSE: Raiz do erro quadrático médio; RPD: Desvio padrão do erro.

2.3 Uso da espectroscopia de reflectância em sedimentos

O uso da espectroscopia para estimar a concentração de elementos tem aplicabilidade direta e de grande importância em trabalhos de monitoramento de qualidade de sedimentos em bacias hidrográficas. Contudo, existe uma grande lacuna de estudos avaliando o uso da espectroscopia para estimar a concentração de elementos em sedimentos.

Por meio da calibração e validação de modelos baseados nas respostas espectrais para estimar a concentração de elementos nos sedimentos, é possível estabelecer um programa de monitoramento mais detalhado. Uma vez que, o custo das análises é baixo e pode ser feita utilizando pequenas quantidades de material (VISCARRA ROSSEL et al., 2016; ARMENTTA; GUARDIA, 2014; CARVALHO et al., 2020). Esta é uma das grandes dificuldades, especialmente para amostras de sedimento em suspensão que possuem baixa concentração dependendo das características da bacia hidrográfica avaliada.

O processo erosivo é caracterizado pela desagregação, transporte, e deposição de partículas do solo podendo estas serem depositadas na rede de drenagem (MORGAN, 2005). Como resultado das diferentes fontes de erosão de uma bacia hidrográfica, a composição dos sedimentos representa a média da composição das principais fontes de sedimento (MINELLA; MERTEN, 2011; MIGUEL et al., 2011). Sendo possível que os modelos calibrados para estimar a concentração de elementos nas amostras de solo possam ser utilizados para obter a concentração de elementos nos sedimentos. Como a amostragem de sedimento é normalmente demorada e onerosa (CARVALHO et al., 2020), isso seria extremamente importante, pois seria possível construir uma biblioteca espectral em uma única campanha de amostragem de solo, e posteriormente, extrapolar esses modelos para estimar a concentração de elementos nos sedimentos. Além disso, esses modelos poderiam ser extrapolados para outras bacias hidrográficas semelhantes em termos de fontes de sedimentos, uso do solo, relevo, geomorfologia e clima. Contudo, não existem relatos na literatura atual que tenha avaliado essa aplicação da espectroscopia, nem utilizando a combinação de amostras de solo e sedimentos, bem como a utilização de diferentes pré-processamentos espectrais e modelos matemáticos.

Estudos utilizando a espectroscopia de reflectância para quantificar as propriedades do sedimento foram desenvolvidos ao longo dos últimos anos a

nível mundial (Tabela 3), avaliando uma quantidade expressiva de propriedades e elementos químicos presentes no sedimento. Estes estudos foram buscados na plataforma Web Of Science e filtrados pelos anos de 2010 até 2020.

2.4 Pré-processamentos espectrais

Na literatura é possível encontrar estudos que têm apontado que a grande confiabilidade e precisão dos modelos dependem diretamente do pré-processamento utilizado nos dados espectrais (VASQUES et al., 2008; DOTTO et al., 2017). Sabe-se que os pré-processamentos auxiliam na interpretação e extração de informações dos espectros de reflectância, já que estes são responsáveis por ressaltar feições espectrais relacionadas com as propriedades dos solos e de sedimentos (GHOLIZADEH et al., 2016). A seguir será detalhado os pré-processamentos que foram utilizados neste estudo, sendo eles: *Detrend* (DET), Derivada de *Savitsky-Golay* (SGD) e *Standard Normal Variate* (SNV).

2.4.1 *Detrend* (DET)

O pré-processamento *Detrend* serve para remover ruídos presentes nos dados espectrais, através da utilização da variação normal padrão seguida de um modelo de regressão polinomial de segundo grau, o qual é utilizado para remover variações de baixa frequência constantes presentes nos dados espectrais, realçando as variações de maior frequência (BARNES; DHANOA; LISTER, 1989).

2.4.2 Derivada de *Savitzky-Golay* (SGD)

O algoritmo de *Savitzky-Golay* (SAVITZKY; GOLAY, 1964) é geralmente utilizado para derivação espectral, composta de um polinômio de primeira ordem, com uma janela de busca de 11 nm, após testes preliminares e utilizando a primeira derivada. Esta calcula a taxa de mudança da reflectância em taxa de variação em relação ao comprimento de onda. Conseguindo, assim, remover as linhas de base e realçar feições de interesse no espectro eletromagnético.

Tabela 3. Seleção de trabalhos com uso da espectroscopia de reflectância para quantificar a concentração de elementos em sedimentos no mundo.

Amostra	Propriedade	Faixa espectral (nm)	Nº de amostras	Pré-processamentos	Mod. Multiv.	R ²	SEP (%)	RMSE	RPD	Referências
Sedimento	Carbono orgânico	2259 - 1217 -		First Derivative (FD)+Vector Normalization (VN)	PLS	0,88	-	2,20	2,10	Moros et al. (2010)
Sedimento	Carbono orgânico	700-1200	50	-	PLSR	0,92	-	-	-	Cadd et al. (2020)
Sedimento	Carbono orgânico	350-2500	103	Raw Data	PLSR	0,48	-	9,18	-	Jiang et al. (2018)
Sedimento	Nitrogênio	2259 - 1913 - 1217	79	First Derivative (FD)+Vector Normalization (VN)	PLSR	0,86	-	0,80	1,95	Moros et al. (2010)
Sedimento	Nitrogênio total	700-1200	50	-	PLSR	0,57	-	-	-	Cadd et al. (2020)
Sedimento	Arsênio (As)	833-2976	81	-	PLS	-	-	-	1,40	Nelson; Sommers, (1982)
Sedimento	Arsênio (As)	1566 - 1217	79	FD+Multiplicative Scatter Correction (MSC)	PLS	0,84	-	3,50	1,79	Moros et al. (2010)
Sedimento	Arsênio (As)	350-2500	103	Raw Data	PLS	0,01	-	224,31	-	Jiang et al. (2018)
Sedimento	Cromo (Cr)	1566 - 1217	79	First Derivative (FD)	PLS	0,86	-	29,00	1,94	Moros et al. (2010)
Sedimento	Cromo (Cr)	350-2500	103	Raw Data	PLS	0,03	-	164,75	-	Jiang et al. (2018)
Sedimento	Cádmio (Cd)	-	-	-	PLS	-	-	-	1,30	Nelson; Sommers, (1982)
Sedimento	Cádmio (Cd)	1566 - 1217	79	First Derivative (FD)+Vector Normalization (VN)	PLS	0,90	-	0,09	2,28	Moros et al. (2010)
Sedimento	Cádmio (Cd)	350-2500	103	Raw Data	PLSR	0,39	-	0,28	-	Jiang et al. (2018)
Sedimento	Chumbo (Pb)	1566 - 1217	79	Vector Normalization (VN)	PLS	0,74	-	7,40	1,42	Moros et al. (2010)
Sedimento	Chumbo (Pb)	350-2500	103	Raw Data	PLSR	0,32	-	1,78	-	Jiang et al. (2018)
Sedimento	Níquel (Ni)	350-2500	103	Raw Data	PLS	0,34	-	8,22	-	Jiang et al. (2018)
Sedimento	Merúrio (Hg)	350-2500	103	Raw Data	PLS	0,39	-	0,05	-	Jiang et al. (2018)
Sedimento	Antimônio (Sb)	1566 - 1217	79	First Derivative (FD)+Vector Normalization (VN)	PLSR	0,88	-	0,50	2,05	Moros et al. (2010)
Sedimento	Estanho (Sn)	1913-1217	79	First Derivative (FD)+Multiplicative Scatter Correction (MSC)	PLSR	0,85	-	2,40	1,92	Moros et al. (2010)

2.4.3 *Standard Normal Variate* (SNV)

O pré-processamento *Standard Normal Variate* é geralmente utilizado para remoção da interferência por variação multiplicativa causada pelo espalhamento da radiação e, também, pela interferência no espectro devido ao tamanho de partícula (BARNES; DHANOA; LISTER, 1989). Consiste basicamente na subtração da média e divisão pelo desvio padrão da reflectância de cada espectro (BARNES; DHANOA; LISTER, 1989).

2.5 Métodos de calibração multivariada

Os métodos de calibração multivariada referem-se a uma técnica utilizada para análise da quimiometria, a qual engloba o desenvolvimento de modelos matemáticos para a análise de dados complexos (CLEGG et al., 2009). Vários métodos são utilizados e tem como finalidade extrair a informação desejada de muitos dados espectrais melhorando cada vez mais a acurácia nas predições (GHOLIZADEH et al., 2016).

A estimativa dos componentes do solo através da espectroscopia depende diretamente da seleção de métodos adequados de calibração multivariada (VOHLAND et al., 2011). O desenvolvimento da tecnologia na última década, em relação à capacidade computacional, métodos estatísticos mais robustos e banco de dados com maior número de amostras, tem sido empregado na construção dos modelos de predição dos teores dos elementos nos solos (MOURA-BUENO, 2018). Algumas das técnicas de análise multivariada mais utilizadas na calibração multivariada são a *Partial Least Squares Regression* (PLSR) e *Support Vector Machine* (CHEN; MORRIS; MARTIN, 2007; PASQUINI, 2003; VISCARRA ROSSEL; BEHRENS, 2010).

2.5.1 *Partial Least Squares Regression* (PLSR)

O método foi criado inicialmente para um tratamento de cadeias de matrizes e aplicações em econometria por volta de 1975 (VARMUZA; FILMOSER, 2009). Trata-se de uma técnica de análise multivariada criada para o desenvolvimento de modelos matemáticos que apresentam relações complexas entre as variáveis resposta e preditora (WOLD; SJÖSTRÖM;

ERIKSSON, 2001). *Partial least squares regression* é uma técnica amplamente utilizada na análise multivariada, a qual suporta a alta colinearidade, ruídos e a análise de inúmeras variáveis (GRINAND et al., 2012; WOLD; SJÖSTRÖM; ERIKSSON, 2001), como no caso da análise de solo através da espectroscopia.

O método PLSR reduz grandes conjuntos de dados em um pequeno número de fatores ortogonais, não correlacionados, de modo a minimizar a soma dos quadrados dos erros entre os valores a serem preditos (VARMUZA; FILZMOSER, 2009). Utiliza fatores ortogonais, também chamados de variáveis latentes, para prever as variáveis resposta, sendo as variáveis latentes combinações lineares das variáveis preditoras (NOCITA et al., 2014).

2.5.2 *Support Vector Machine (SVM)*

A análise multivariada SVM baseia-se no método de aprendizagem estatística “*kernel*” (CORTES; VAPNIK, 1995), onde busca identificar uma função de interpolação. Ajusta-se os dados de treinamento para um tubo com raio definido, usando amostras limites, em que o melhor modelo é obtido usando uma função. O objetivo é minimizar simultaneamente o tamanho dos coeficientes e os erros de predição, em que conjuntos de dados com padrões não lineares podem ser representados mais facilmente (IVANCIUC, 2007). Utilizando esta técnica é possível derivar um plano espectral linear como uma função de decisão para problemas não lineares (VISCARRA ROSSEL; BEHRENS, 2010).

3. HIPÓTESE

A acurácia da estimativa da concentração de elementos químicos em amostras de solo e sedimento pode ser aprimorado quando utilizado técnicas de pré-processamento espectral, modelos multivariados não-paramétricos.

Modelos calibrados com amostras de solo e sedimento combinadas apresentam maior acurácia na estimativa da concentração de elementos químicos em amostras de solo e sedimento comparado aos modelos calibrados com conjunto de amostras isoladas.

Modelos calibrados com amostras de solo de uma bacia hidrográfica podem ser utilizados para estimar a concentração de elementos químicos em amostras de sedimento.

4. OBJETIVO

O objetivo desse estudo foi avaliar qual a melhor combinação de modelos multivariados e técnicas de pré-processamento espectral para aumentar a acurácia da estimativa da concentração de elementos em amostras de solo e sedimento. Adicionalmente, foi avaliado o uso da espectroscopia de reflectância para estimar a concentração de elementos presentes no sedimento, usando modelos calibrados com amostras de solo em uma bacia hidrográfica.

5. MATERIAL E MÉTODOS

5.1 Caracterização da bacia hidrográfica

A bacia hidrográfica do Rio Guaporé (2.032 km²) localiza-se no nordeste do Rio Grande do Sul, Brasil (Figura 2), compreendendo os municípios de Anta Gorda, Arvorezinha, Camargo, Casca, Guaporé, Gentil, Ibirapuitã, Ilópolis, Itapuca, Marau, Mato Castelhano, Montauri, Nova Alvorada, Santo Antônio do Palma, Serafina Corrêa, União da Serra e Vila Maria. O clima na região é classificado como Cfa – subtropical úmido segundo Köppen (ALVARES et al. 2013) com precipitação média anual de 1400 a 2000 mm. A área da bacia hidrográfica possui altitudes variando de 207 a 830 m de altitude. O material de origem é variável na região, predominando rochas vulcânicas extrusivas básicas (basalto) e ácidas (riodacito) (WILDNER et al., 2008). Os solos predominantes na região são Argissolo Vermelho, Argissolo Vermelho-Amarelo, Latossolo Vermelho, Luvisolo Crômico, Neossolo Regolítico e Nitossolo Bruno (STRECK et al., 2018) (Figura 3).

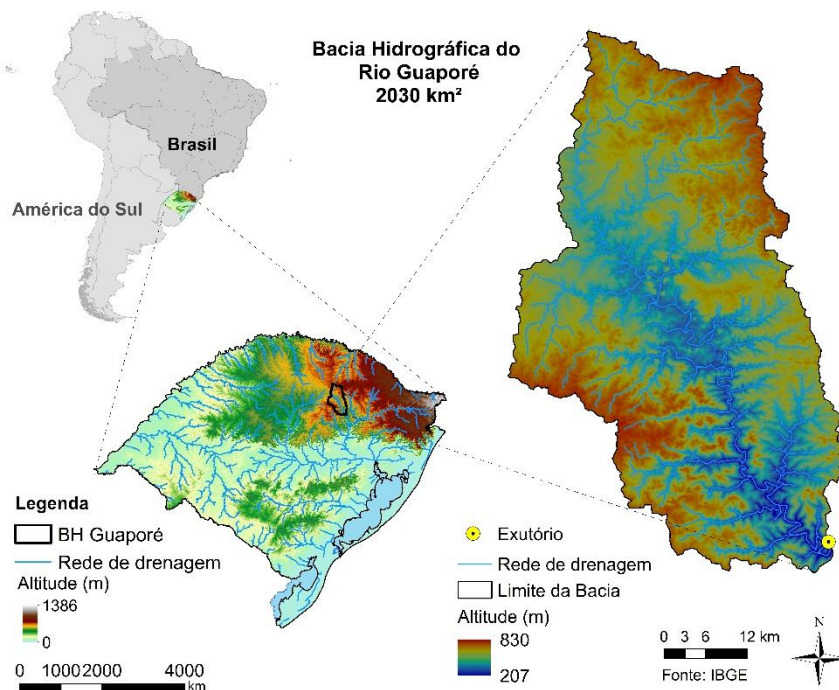


Figura 2. Localização e relevo da bacia hidrográfica do Rio Guaporé, Rio Grande do Sul, Sul do Brasil.

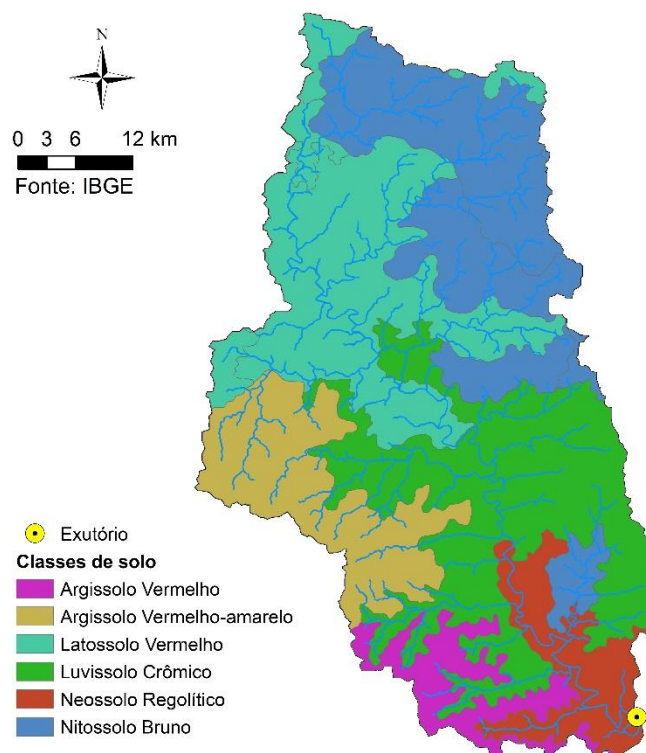


Figura 3. Classes de solo da bacia hidrográfica do Rio Guaporé, Rio Grande do Sul, Sul do Brasil.

O uso do solo é variado na bacia do Rio Guaporé (Figura 4), sendo 9,7 % para pastagem, 56,6 % para florestas e 32,5 % para lavoura (TIECHER et al.,

2015). No terço superior, onde o terreno é caracterizado por encostas suaves, há predominância de soja (*Glycine max*) e milho (*Zea mays*) cultivados no verão e trigo (*Triticum aestivum*) e outras plantas de cobertura durante o inverno em sistema de plantio direto (TIECHER et al, 2017). Nos outros dois terços da bacia (partes intermediárias e inferiores), os principais cultivos são o tabaco (*Nicotiana tabacum*), milho, erva-mate, reflorestamento com eucaliptos (*Eucalyptus spp.*) e pastagens perenes e anuais para gado leiteiro (TIECHER et al., 2015). Nessas áreas, o manejo do solo varia muito e o cultivo convencional e mínimo são as principais práticas de manejo do solo. Nas áreas de maior declive, principalmente nas margens dos rios, ocorrem grandes porções de áreas de mata nativa, bem como pequenas áreas urbanas (TIECHER et al., 2016). As características fisiográficas do terço intermediário e inferior da bacia determinam fragilidade ambiental, quando o solo é utilizado para agricultura sem levar em conta práticas conservacionistas, pois o relevo é mais acidentado e há predomínio de lavouras de tabaco com intenso revolvimento e preparo do solo.

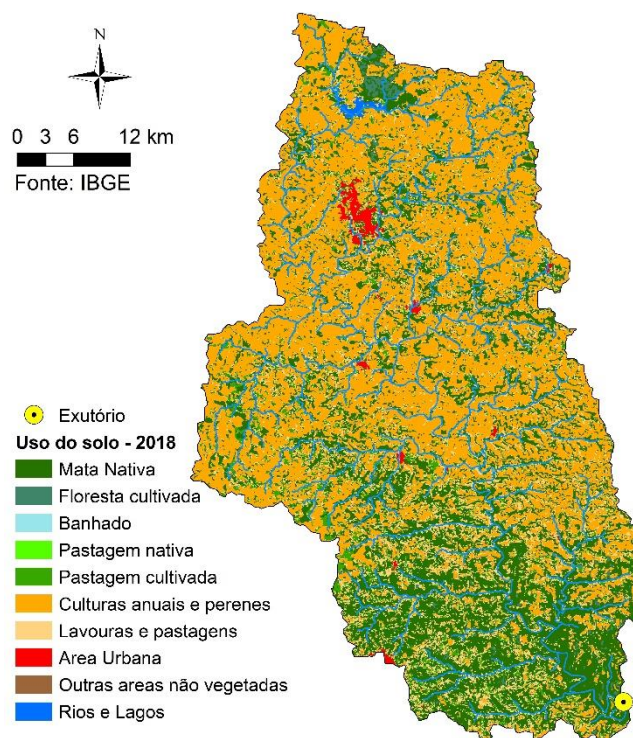


Figura 4. Uso e ocupação do solo da bacia hidrográfica do Rio Guaporé, Rio Grande do Sul, Sul do Brasil.

5.2 Coleta das amostras de solo e sedimento

As amostras de solo e sedimento foram coletadas durante o período de 2011 e 2014. As amostras de solo foram coletadas nas principais fontes potenciais de sedimento da bacia hidrográfica, sendo elas: lavoura anual, pastagens, estradas não-pavimentadas e as margens dos rios e riachos, na camada de (0-0,05 m). Para obter material representativo, cada amostra foi composta por 10 subamostras coletadas próximas do ponto de amostragem (Tabela 4). A distribuição espacial das amostras de solo e sedimento que foram coletadas na bacia hidrográfica do Rio Guaporé foi apresentada na Figura 5.

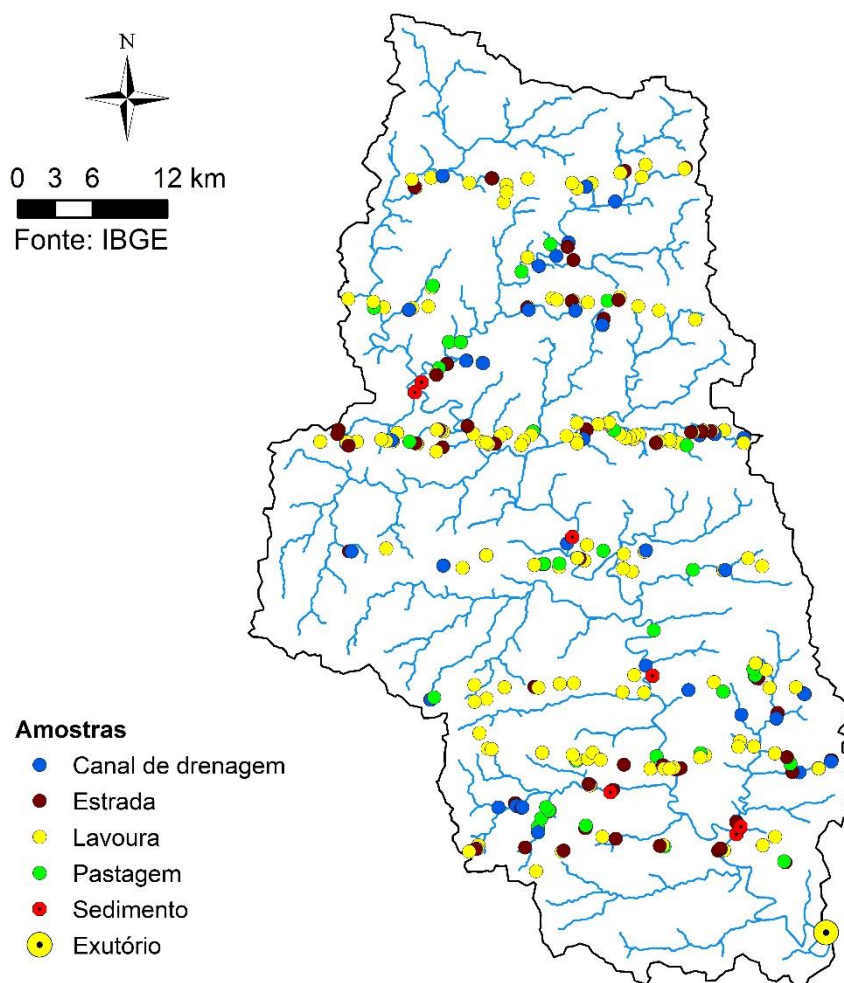


Figura 5. Distribuição dos pontos de coleta de amostras de solo e de sedimento da bacia hidrográfica do Rio Guaporé, Rio Grande do Sul, Sul do Brasil.

A amostragem de sedimentos foi realizada utilizando quatro diferentes métodos de obtenção das amostras, dentre estas, as abordagens I, III e IV estão

distribuídas na figura 5, nomeadas como sedimento, a abordagem II como exutório e, todas são demonstradas na sequência:

I) amostradores de sedimento em suspensão integrados no tempo do tipo Phillips et al. (2000). Em 10 pontos da bacia do Guaporé (Figura 5), dois amostradores integrados no tempo foram instalados para garantir que uma quantidade suficiente de sedimento fosse coletada para análises subsequentes. Este coletor possui um pequeno orifício que permite a passagem do fluxo de água, o qual baseia-se no princípio da sedimentação: o material em suspensão na água que passa pelo coletor ficará acumulado no fundo, integrando em uma única amostra o sedimento em suspensão transportado pela água por um determinado período. Os amostradores integrados no tempo coletam sedimentos finos em suspensão ($<63 \mu\text{m}$), (PHILLIPS; RUSSELL; WALLING, 2000).

II) amostras de sedimento em suspensão foram coletadas durante eventos pluviométricos, somente no exutório da bacia (Figura 5). Para este fim, um grande volume de água (50 a 200 litros) foi coletado manualmente em intervalos variados ao longo da ascensão e recessão do hidrograma para avaliar a variação intra-evento das contribuições da fonte de sedimento.

III) amostras de sedimento em suspensão coletados durante os eventos pluviométricos coletado por meio do amostrado de nível ascendente - ANA, amostradores que são uma adaptação do modelo US U-59 (CEW-EH-Y, 1995), os quais foram instalados no leito do rio.

IV) sedimento de fundo ou de leito fino: foram coletados com um amostrador de sucção de aço inoxidável que permitiu a coleta de sedimentos superficiais, sem perda do material fino na interface sedimento/água. As amostras foram compostas de 20 a 30 subamostras coletadas ao longo do canal do rio (Figura 5).

Tabela 4. Número de amostras de solo e de sedimento coletadas na bacia hidrográfica do Rio Guaporé, Sul do Brasil.

Material	Tipo de amostra	Número de amostras
Solo	Lavoura (32,5% da área total da bacia)**	159
	Canal de drenagem	46
	Estradas	58
	Pastagem (9,7% da área total da bacia)**	46
	Total	316
Sedimento	Amostrador integrado de SS*	50
	SS coletado manualmente durante eventos pluviométricos	12
	SS coletado durante eventos pluviométricos com ANA	26
	Sedimento de fundo/leito	62
	Total	196

*SS: Sedimento em suspensão; ANA: Amostradores de Nível Ascendente; ** Porcentagem referente ao uso e ocupação do solo da bacia estudada.

As amostras de solo e de sedimentos foram secas em estufa a 50 °C com circulação de ar forçado e levemente desagregadas usando gral de ágata. Estas amostras foram peneiradas a 63 μm , a fim de comparar as frações granulométricas de tamanho semelhante. Com o intuito de reduzir as incertezas relacionadas ao efeito da diluição de elementos e o efeito das concentrações das partículas de areia (MICHELAKI; HANCOCK, 2013) e que considera os agregados de argila suspensos (MINELLA; WALLING; MERTEN, 2008) utilizou-se a peneira de <63 μm . Devido a seletividade do tamanho de partícula no processo erosivo, é necessário usar frações granulométricas compatíveis para poder comparar amostras de solo e sedimento (<63 μm). Além disso, essa é a fração que concentra a maior parte dos elementos e, por isso, ambientalmente mais relevante para estudos relacionados a concentrações de elementos em bacias hidrográficas. Mais detalhes referentes a coleta das amostras de solo e sedimentos podem ser observados em Tiecher (2015).

5.3 Análises químicas

Nas amostras de solo e de sedimentos foram determinados por meio de métodos analíticos de química úmida uma gama de elementos químicos. O carbono orgânico total (C) foi estimado pela oxidação úmida com $\text{K}_2\text{Cr}_2\text{O}_7$ e

H₂SO₄ (WALKLEY; BLACK, 1934) e a concentração pseudo-total (<63 µm) dos elementos Al, Ba, Be, Ca, Co, Cr, Cu, Fe, K, La, Li, Mg, Mn, Na, Ni, P, Pb, Sr, Ti, V e Zn, foi estimada por ICP-OES (*Inductively Coupled Plasma Optical Emission Spectrometry*) após digestão assistida por micro-ondas por 9,5 min a 182 °C com HCl e HNO₃ concentrado na proporção de 3:1 (*aqua-regia*) (USEPA, 1998).

5.4 Análises espectroscópicas

Os espectros do NIR foram registrados da faixa de 10000–4000 cm⁻¹ (1000-2500 nm) utilizando um espectrômetro FTIR Nicolet 26700 (Waltham, Massachusetts, EUA) no modo de reflectância difusa com uma esfera de integração e um detector InGaAs intercooled com uma resolução de 2 cm⁻¹ e 100 leituras por espectro.

5.5 Pré-processamento espectral

Após a aquisição dos dados espectrais, os espectros brutos (RAW) foram submetidos a três técnicas de pré-processamento espectral: *Detrend (DET)* que remove os ruídos presentes nos dados espectrais através da utilização da *standard normal variate* seguida de um modelo de regressão polinomial de segunda ordem para padronizar os dados (BARNES; DHANOA; LISTER, 1989); *Standard Normal Variate (SNV)* utilizado para remover a dispersão de dados espectrais causada por ruídos e diferentes tamanhos de partículas e, consiste em suavizar o espectro a partir dos valores médios e desvio-padrão usando uma função de convolução aplicada em cada espectro individualmente; *Savitzky-golay derivative (SGD)* com a 1ª derivada usando um polinômio de primeira ordem, com janela de busca de 11 nm, após teste inicial para definir a melhor janela de busca. A definição dessas três técnicas foi baseada nas diferenças que existem no processo matemático de cada uma para transformar os dados espectrais. Todos os pré-processamentos foram realizados utilizando os pacotes *prospectrs* e *clusterSim* (STEVENS; RAMIREZ-LOPEZ, 2020; WALESIAK; DUDEK, 2020) implementados no R (R CORE TEAM, 2020).

5.6 Calibração e avaliação dos modelos

Modelos multivariados foram calibrados entre a concentração de C orgânico total e dos 21 elementos pseudo-totais (Al, Ba, Be, Ca, Co, Cr, Cu, Fe, K, La, Li, Mg, Mn, Na, Ni, P, Pb, Sr, Ti, V e Zn) e dados espectrais NIR, para amostras de solo (n=316), sedimento (n=196) e combinando solo + sedimento (n=512). Cabe salientar que para alguns elementos químicos o número de amostras é inferior devido a não obtenção dos dados químicos.

Na calibração dos modelos foi utilizado dois métodos multivariados, por se tratar de um modelo paramétrico e um não paramétrico, a saber: (i) *Partial Least Squares Regression* (PLSR) (R *pls* package (MEVIK et al., 2016)), técnica paramétrica amplamente utilizada na modelagem espectroscópica (DOTTO et al., 2018; ANGELOPOULOU et al., 2020) e (ii) *Support Vector Machine* (SVM) (R package *e1071* (MEYER et al., 2019)), técnica não-paramétrica.

Métodos com diferentes abordagens foram selecionados devido à ocorrência de correlações lineares e não lineares entre os componentes organominerais do solo e sedimento e as variáveis espectrais (VISCARRA ROSSEL; BEHRENS, 2010). O método PLSR reduz grandes conjuntos de variáveis preditoras a um pequeno número de fatores ortogonais não correlacionados para minimizar a soma dos quadrados dos erros das estimativas (VARMUZA; FILZMOSER, 2009). O método SVM foi utilizado com a função *kernel*, que separa os dados de calibração em hiperplanos e busca estabelecer correlações entre as variáveis dependentes e independentes quando estas apresentam comportamento não linear (IVANCIUC, 2007).

A calibração dos modelos com PLSR e SVM para as amostras de solo, sedimento e solo + sedimento combinando os três pré-processamentos espectrais mais os espectros brutos (RAW), totalizou 24 modelos preditivos para cada elemento, e um total de 528 modelos no estudo completo (Figura 6). Todos os procedimentos de modelagem foram implementados no *software* R (R CORE TEAM, 2020).

Todos os modelos foram calibrados por validação cruzada usando uma divisão aleatória em 10-fold (BRUS et al., 2011). No processo de calibração foi calculada a importância de cada banda no NIR na predição dos elementos por

meio da função *varImp* do pacote “caret” implementado no *software* R (R CORE TEAM, 2020).

O desempenho dos modelos foi avaliado pelo coeficiente de determinação (R^2) (Equação 1), relação entre o desempenho e a distância interquartil (RPIQ) (Equação 2) e o erro médio quadrático de predição (RMSE) (Equação 3) (BELLON-MAUREL et al., 2010). Para fins de interpretação, foi utilizado a classificação de qualidade dos modelos sugerida por Veum et al. (2015) e Chang et al. (2001), onde modelos com Categoria A possuem alto nível de confiabilidade: $R^2 \geq 0,75$ e $RPIQ \geq 3,0$; na sequência modelos com Categoria B possuem $R^2 \geq 0,63$ e $< 0,75$ e $RPIQ \geq 1,9$ e $< 3,0$, e modelos com Categoria C possuem $R^2 \geq 0,50$ e $< 0,63$ e $RPIQ \geq 1,5$ e $< 1,9$. Calculou-se também a incerteza (%) dos valores de RMSE, por meio dos valores das concentrações dos elementos obtidos por modelagem espectral, como pode ser visto na Equação 4.

$$R^2 = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (1)$$

$$RPIQ = \frac{(Q3 - Q1)}{RMSE} \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (3)$$

$$Incerteza (\%) = \frac{RMSE}{teor\ médio} * 100 \quad (4)$$

Posteriormente, os dados de R^2 e RPIQ foram submetidos a uma análise de árvore de regressão de inferência condicional, para destacar os fatores (método multivariado, pré-processamento, elemento, e conjunto de amostras) que mais influenciaram na acurácia dos modelos.

Os parâmetros de acurácia como R^2 , RMSE e RPIQ obtidos com os diferentes métodos de calibração, pré-processamentos espectrais, conjunto de dados (solo e sedimento) e elementos foram comparados pelo teste de média

de Tukey ao nível de significância de $p < 0,05$, utilizando o *software* R (R CORE TEAM, 2020).

Após a modelagem e interpretação dos resultados dos modelos calibrados com as amostras de solo, foram selecionados para cada elemento os modelos mais acurados entre as combinações de método multivariado e pré-processamento, a fim de realizar a estimativa dos elementos nas amostras de sedimento (Figura 6).

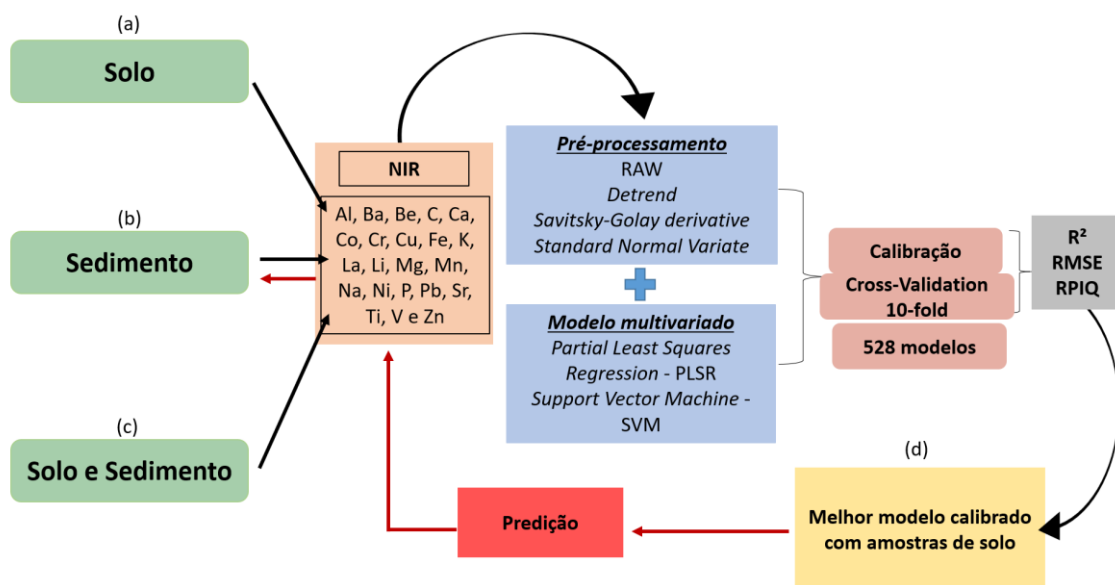


Figura 6. Esquema da calibração dos modelos para amostras de solo, sedimento e solo + sedimento. (a) conjunto de dados de solo; (b) conjunto de dados de sedimento; (c) conjunto de dados de solo e sedimento; (d) melhores modelos calibrados com amostras de solo.

6. RESULTADOS E DISCUSSÃO

6.1 Análise do comportamento espectral no NIR

As assinaturas espectrais são imprescindíveis para interpretar qualitativamente a amostra analisada. Os resultados das assinaturas para solo e sedimento são apresentados na Figura 7. As bandas presentes na assinatura espectral são alteradas de acordo com os teores de óxidos, matéria orgânica, teor de argila, minerais como quartzo, água (DALMOLIN et al., 2005) e vários outros constituintes presentes na amostra, resultando em maiores ou menores reflectâncias. A reflectância das amostras de sedimento é maior em comparação com as amostras de solo, principalmente pela composição do sedimento ser diferente do solo (Figura 7). Ao que tudo indica, as amostras de solo são mais ricas em quartzo e sabe-se que quanto maior a quantidade de quartzo presente em uma amostra maior será sua reflectância. Outro fator que deve ser levado em consideração é a quantidade de matéria orgânica presente no solo. Quanto maior o teor de matéria orgânica maior será a absorção ao longo de todo o espectro NIR, podendo apresentar maior redução do albedo ao longo de todo o espectro eletromagnético (DALMOLIN et al., 2005). Esses autores observaram a influência da matéria orgânica na forma e no albedo da curva espectral ao longo de todo o espectro Vis-NIR para amostras de Latossolo no Estado do RS. Os autores observaram que o aumento do teor de matéria orgânica no solo implica no escurecimento do mesmo, causando uma maior absorção de radiação eletromagnética e, conseqüentemente, uma reflectância menor.

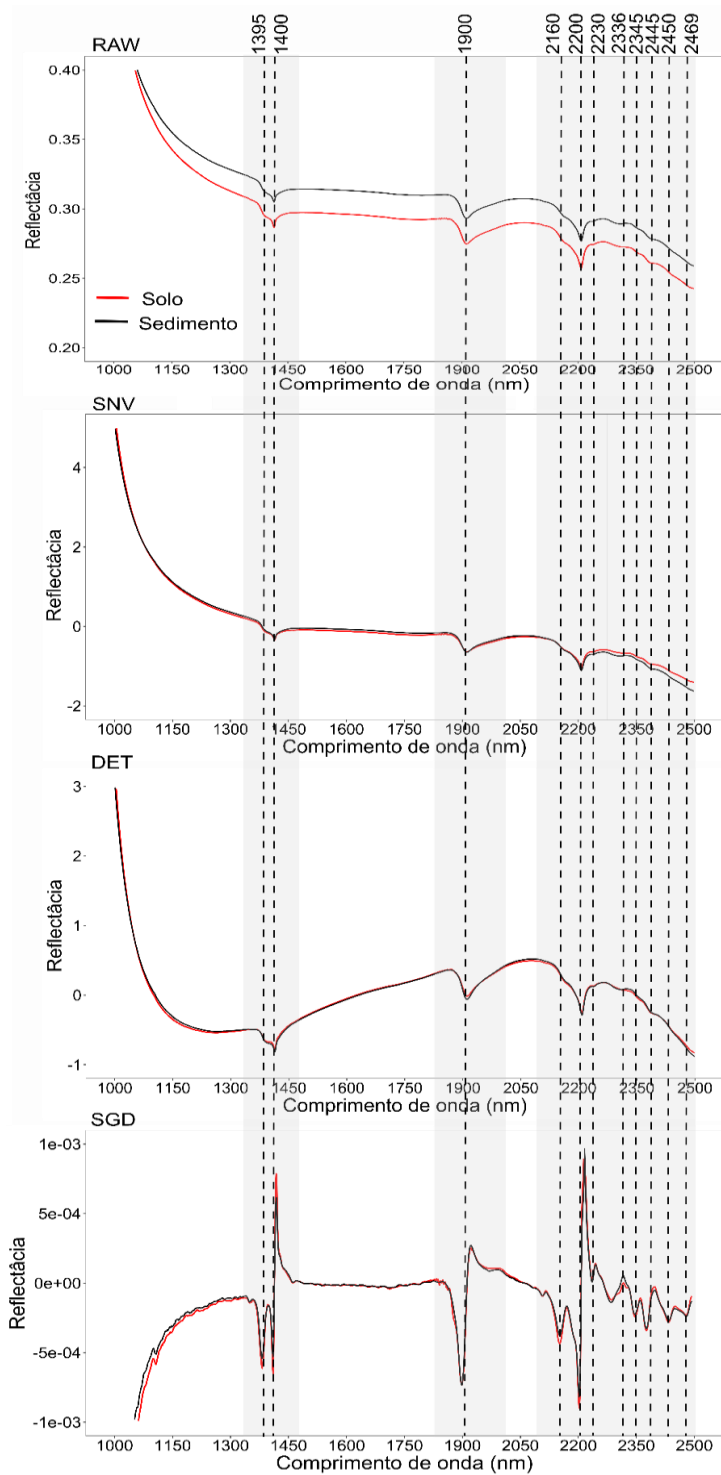


Figura 7. Assinaturas espectrais médias do conjunto de amostras de solo e sedimento da bacia hidrográfica do Rio Guaporé para a faixa espectral NIR bruto sem nenhum pré-processamento (a) RAW e com os pré-processamentos, (b) *Detrend* (DET), (c) *Standard Normal Variate* (SNV), (d) *Savitsky-Golay derivative* (SGD).

A curva espectral representa a média das amostras de solo e sedimento, onde são destacadas as principais bandas do espectro NIR. Vale salientar que alguns minerais presentes em solos e sedimentos são excitados pelas vibrações e podem responder em várias bandas do espectro eletromagnético, como pode ser o caso da caulinita e mica. A primeira banda de absorção é no comprimento de onda de 1395 nm, a qual corresponde a água, grupo funcional O-H (VISCARRA ROSSEL; BEHRENS 2010). Para o comprimento de onda de 1400 nm, onde observa-se um vale na curva espectral, trata-se de minerais como a caulinita, esmectita, micas e vermiculita. Em 1900 nm, a banda de absorção é atribuída à vibração molecular dos grupos OH da água adsorvida aos argilominerais (LINDBERG; SNYDER, 1972; HUNT, 1980).

Nos comprimentos de onda de 2200 nm, ocorrem bandas de absorção devido às vibrações moleculares dos grupos OH e Al-OH presente nos minerais (HUNT; SALISBURG, 1970; VISCARRA ROSSEL; BEHRENS, 2010; MADEJOVÁ et al., 2011), sendo típicas da presença de caulinita e/ou montmorilonita. A banda de absorção de 2336 nm está relacionada às vibrações dos grupos OH presentes na gibbsita (HUNT et al., 1971), e ao carbono (metil), grupo funcional C-H onde a vibração ocorre em 2316 nm, já o carbono (carboidratos), grupo funcional C-O, em 2381 nm (VISCARRA ROSSEL; BEHRENS, 2010; TERRA et al., 2011). A banda de absorção de 2340 nm, onde o grupo funcional é o (O-H) + (Al-OH), são atribuídas às micas (TERRA et al., 2011; MADEJOVÁ et al., 2011).

Após os dados espectrais serem submetidos aos pré-processamentos, observa-se que ocorreram alterações nas assinaturas espectrais. Os espectros SNV foram suavizados, resultando em redução do desvio-padrão. Isso porque a técnica SNV calcula os valores médios dos dados espectrais usando uma função de convolução (RINNAN et al., 2009). Por isso as assinaturas de solo e sedimento estão mais próximas no SNV comparado ao RAW. No entanto, a técnica SNV não realça as regiões de absorção como observado para os espectros processados com a técnica SGD.

A técnica SGD usa a primeira derivada para calcular a taxa de variação da refletância em relação aos comprimentos de onda. Isso resulta em transformações rigorosas nos espectros, como realçar feições de absorção e

remove o sinal de linha de base sem importância das amostras (RINNAN et al., 2009). Essas diferenças nas assinaturas espectrais podem ter efeito na calibração dos modelos de predição, refletindo na acurácia das predições (VASQUES et al., 2008; KNOX et al., 2015; DOTTO et al., 2018; MOURA-BUENO et al., 2019).

O DET normaliza os dados espectrais aplicando uma transformação de variável normal padrão e ajusta um modelo de regressão polinomial de segundo grau. Mas também não ocorre realce das regiões de absorção.

6.2 Estatística descritiva da concentração de elementos em amostras de solo e sedimento

A estatística descritiva dos elementos demonstra que existe uma grande amplitude na concentração dos 22 elementos químicos avaliados, tanto nas amostras de solo quanto nas amostras de sedimento na bacia hidrográfica do Rio Guaporé (Tabela 5). Para o elemento C, por exemplo, a concentração de C no solo variou de 1,98 a 47,04 g kg⁻¹, mostrando uma ampla variação das amostras (Tabela 5). A grande variabilidade nas concentrações de elementos químicos nas amostras de solo se deve às diferenças na geologia, tipos de solo (ALTHAUS et al., 2018) e, do efeito do uso do solo, que impacta diretamente nas concentrações dos elementos (TIECHER, 2015). Por consequência, a variabilidade na concentração dos elementos químicos nas amostras de sedimento reflete as principais fontes de sedimentos operantes em diferentes locais da bacia hidrográfica e diferentes épocas de amostragem de sedimento.

Tabela 5. Estatística descritiva do conjunto de amostras de solo e sedimento coletadas na bacia do Rio Guaporé.

Estatística	C	Al	Ca	Fe	K	Mg	Mn	Ti	Ba	Be	Co	Cr	Cu	La	Li	Na	Ni	P	Pb	Sr	V	Zn
Solo	(g kg⁻¹)								(mg kg⁻¹)													
Nº de amostras	316	316	314	316	316	316	316	316	316	316	316	316	316	316	316	304	316	316	316	316	315	316
Mínimo	1,98	22,32	0,05	21,34	0,17	1,31	0,28	1,22	60,80	0,12	0,36	6,30	14,70	6,70	6,62	0,04	2,07	84,46	5,28	7,01	24,11	6,24
Máximo	47,04	106,22	6,74	147,11	7,07	10,77	4,53	7,89	553,27	7,21	112,73	276,36	589,26	94,86	114,37	994,36	82,26	1482,79	60,87	127,27	774,54	33,86
Média	18,31	51,92	2,03	77,34	1,55	3,78	1,93	3,67	208,75	3,33	43,51	28,72	197,55	32,39	42,97	84,89	23,04	376,80	23,24	29,18	291,81	14,28
Mediana	17,73	48,80	1,91	83,03	1,42	3,39	1,92	3,55	204,06	3,34	45,04	23,22	213,15	29,46	38,91	54,08	21,37	348,56	22,29	26,52	311,33	13,41
Desvio Padrão	8,34	15,61	1,22	31,88	0,87	1,59	0,96	1,13	76,15	1,37	26,73	21,99	130,95	14,90	15,25	108,18	17,45	174,17	8,58	14,52	181,92	4,98
1st Qu	12,22	41,41	1,16	46,99	0,97	2,56	1,11	2,87	156,79	2,18	16,54	17,42	56,07	21,35	32,40	38,92	8,48	263,36	17,48	20,54	99,68	10,53
3rd Qd	24,13	57,66	2,62	101,72	1,98	4,60	2,66	4,42	243,07	4,34	64,83	32,18	291,69	40,94	49,37	87,13	29,94	456,75	28,38	34,06	424,91	16,68
Sedimento																						
Nº de amostras	50	196	196	196	196	196	196	196	196	196	167	196	196	196	196	69	151	196	196	196	196	196
Mínimo	6,90	13,88	1,22	18,27	0,07	1,67	0,63	0,28	32,93	0,01	9,49	7,89	31,58	10,79	9,86	0,23	3,05	126,38	5,00	13,61	48,48	8,80
Máximo	54,30	95,28	43,17	101,78	19,37	16,25	41,82	19,53	612,81	9,97	103,28	130,80	300,49	62,08	54,37	6107,86	38,80	1854,34	40,36	330,48	672,82	189,68
Média	29,13	37,14	44,63	53,35	2,23	4,65	1,90	7,02	319,61	4,19	45,35	33,19	154,83	32,83	35,11	658,44	20,88	440,82	19,03	52,90	284,92	32,77
Mediana	28,05	37,59	32,84	52,98	1,69	4,23	1,83	5,14	325,61	2,88	45,37	30,55	156,02	33,22	35,04	253,29	21,55	434,82	18,07	45,04	266,34	23,16
Desvio Padrão	11,02	8,29	4,95	15,35	2,35	1,84	0,63	4,94	101,12	3,15	18,84	16,15	55,56	9,07	6,44	1136,07	7,68	174,89	6,61	40,38	128,55	23,00
1st Qu	23,47	32,62	2,74	43,97	1,30	3,61	1,48	3,45	257,43	1,50	31,60	22,70	120,80	27,21	31,82	118,86	16,05	354,23	14,02	35,24	191,04	16,02
3rd Qd	33,30	41,73	4,12	62,19	2,16	5,15	2,26	11,61	389,79	7,68	56,35	39,14	186,63	38,80	37,78	551,95	25,25	510,51	23,40	55,12	367,30	50,68

* 1st Qd: primeiro interquartil; 3rd Qd: terceiro interquartil; C: carbono; Al: alumínio; Ba: bário; Be: berílio; Ca: cálcio; Co: cobalto; Cr: cromo; Cu: cobre; Fe: ferro; K: potássio; La: lantânio; Li: lítio; Mg: magnésio; Mn: manganês; Na: sódio; Ni: níquel; P: fósforo; Pb: chumbo; Sr: estrôncio; Ti: titânio; V: vanádio; Zn: zinco.

6.3 Principais fatores afetando a acurácia dos modelos

A partir dos resultados da análise de árvore de regressão de inferência condicional utilizando os valores de ajuste dos modelos (R^2 e RPIQ), foi possível verificar que os fatores mais importantes para definir a acurácia das predições da concentração de elementos foram: modelo multivariado > elemento = pré-processamento para R^2 (Figura 8) e RPIQ (Figura 9). Com base nos resultados obtidos nessas análises, a seguir são apresentados e discutidos detalhadamente sobre o efeito de cada um dos fatores na qualidade dos modelos.

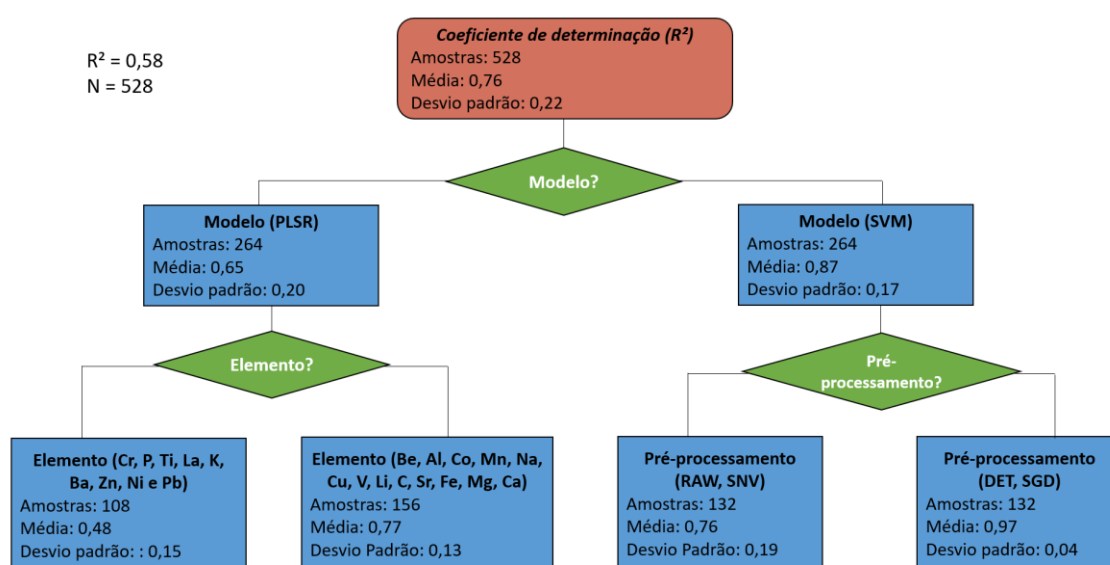


Figura 8. Árvore de regressão de inferência condicional utilizando os valores de R^2 para as variáveis de método multivariado, elemento e pré-processamento espectral.

* R^2 Coeficiente de determinação; PLSR: Regressão parcial de quadrados mínimos; SVM: modelo de aprendizagem de máquina (*Support Vector Machine*); Na: sódio; K: potássio; Cr: cromo; P: fósforo; Zn: zinco; Ti: titânio; Sr: estrôncio; Ca: cálcio; Ba: bário; La: lantânio; Ni: níquel; Pb: chumbo; Al: alumínio; Mn: manganês; C: carbono; Mg: magnésio; Be: berílio; Co: cobalto; Cu: cobre; Li: lítio; V: vanádio; Fe: ferro; SNV: Variação normal padrão (*Standart Normal Variate*), DET: *Detrend*; SGD: Derivada de *Savitsky-Golay*.

Importante ressaltar que para o método multivariado PLSR, o elemento tem mais efeito em comparação com as demais variáveis e, para o SVM, o pré-processamento tem mais efeito, tanto para R^2 e RPIQ. Isso pode ter ocorrido

porque o PLSR transforma as variáveis predictoras em variáveis ortogonais, o que minimiza o efeito dos pré-processamentos. Porém, isso não significa que vai reduzir a acurácia das predições, como ocorreu neste estudo para a maior parte dos elementos analisados.

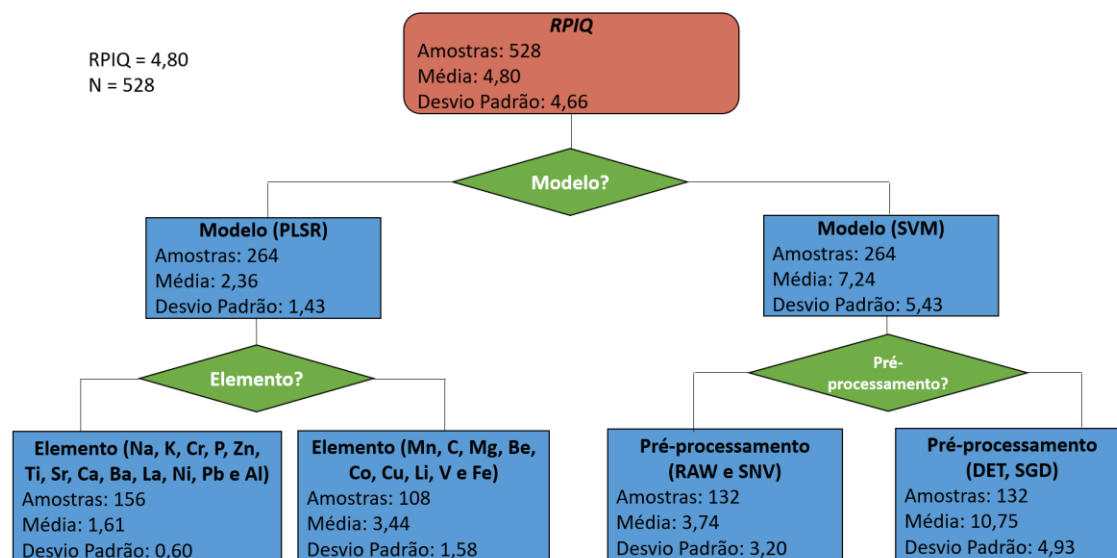


Figura 9. Árvore de regressão de inferência condicional utilizando os valores de RPIQ para as variáveis de método multivariado, elemento e pré-processamento espectral.

*RPIQ: Relação de desempenho interquartil; PLSR: Regressão parcial de quadrados mínimos; SVM: modelo de aprendizagem de máquina (*Support Vector Machine*); Na: sódio; K: potássio; Cr: cromo; P: fósforo; Zn: zinco; Ti: titânio; Sr: estrôncio; Ca: cálcio; Ba: bário; La: lantânio; Ni: níquel; Pb: chumbo; Al: alumínio; Mn: manganês; C: carbono; Mg: magnésio; Be: berílio; Co: cobalto; Cu: cobre; Li: lítio; V: vanádio; Fe: ferro; SNV: Variação normal padrão (*Standart Normal Variate*), DET: *Detrend*; SGD: Derivada de *Savitsky-Golay*.

6.3.1 Efeito dos modelos multivariados

O modelo multivariado SVM resultou em melhores predições que o PLSR em todos os casos avaliados, como indicado pelos valores médios de ajuste do modelo ($R^2 = 0,87$ SVM e $0,65$ para PLSR) (Figura 10), e pelos valores de RPIQ (2,30 PLSR e 7,25 SVM) (Figura 10). O melhor desempenho do método não-paramétricos SVM é devido ao uso da função de *kernel*, a qual separa os dados de calibração em hiperplanos e busca estabelecer correlações entre a variável dependente e independentes quando essas apresentam comportamento não linear (IVANCIUC, 2007). Essa abordagem tem mostrado vantagens na

modelagem de conjuntos de dados que apresentam relações não lineares comparado ao método PLSR, como observado por Tiecher et al. (2021) na estimativa da contribuição de fonte de sedimento. Seema et al. (2020) analisando amostras de solo, verificaram uma melhor predição do modelo SVM em comparação com o PLSR para o teor de carbono do solo.

Além disso, o método PLSR reduz o número de variáveis preditoras, transformando-as em fatores ortogonais. Nesse processo, pode ocorrer perda de variáveis explicativas importantes para a predição e, portanto, reduzir o desempenho preditivo dos modelos em comparação com o método SVM (Figura 10). Isso sugere que feições espectrais menos absorventes, mas ainda importantes, podem ser perdidas em alguns comprimentos de onda. Comportamento semelhante foi observado por Dotto et al. (2018) para predições de carbono orgânico total (COT) de solos da região sul do Brasil. Os autores encontraram o melhor desempenho para o método multivariado SVM e pré-processamento SGD ($R^2 = 0,80$ e RPIQ = 2,94) em relação do método multivariado PLSR e pré-processamento SGD ($R^2 = 0,67$ e RPIQ = 2,30).

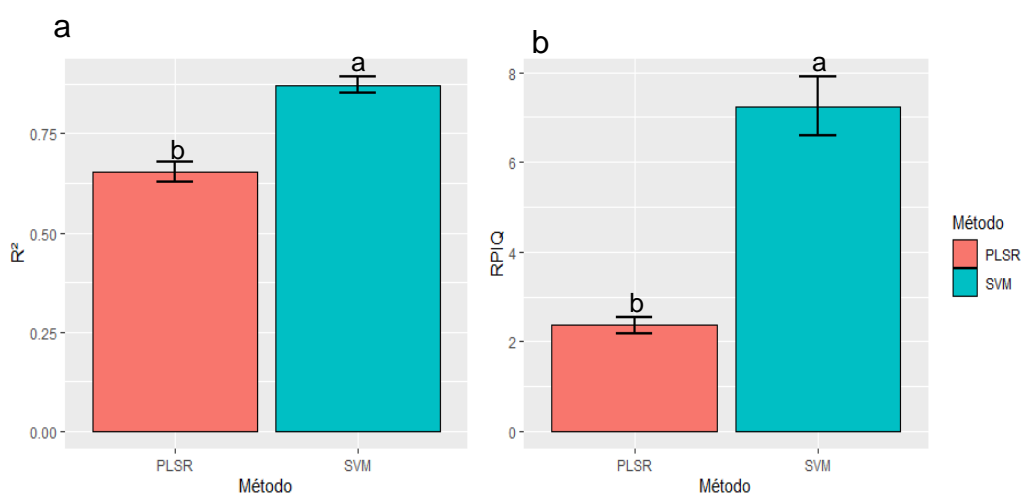


Figura 10. Comparação do desempenho dos métodos de calibração multivariada indicado pelos valores de R^2 (a) e RPIQ (b). As barras representam valores médios e o desvio padrão de 264 modelos. PLSR: *Partial Least Squares Regression*. SVM: *Support Vector Machine*.

*Médias seguidas pela mesma letra comparando os métodos de calibração multivariados não diferem estatisticamente pelo teste de Tukey a $p < 0,05$.

6.3.2 Efeito das técnicas de pré-processamento

Utilizando o método PLSR, não houve diferença entre os pré-processamentos testados, com valores de R^2 variando de 0,64 a 0,66 e RPIQ variando entre 2,23 e 2,40 (Figura 11). Por outro lado, a técnica de pré-processamento influenciou diretamente na predição dos dados quando comparado com o modelo SVM (Figura 11). Nesse caso, o desempenho dos modelos aumentou na seguinte ordem de pré-processamento utilizado: RAW<SNV<DET=SGD, como pode ser visto tanto pelo R^2 e RAW<SNV<DET<SGD pelo RPIQ (Figura 11). Além disso, é possível verificar que o desempenho dos modelos SVM sem pré-processamento (RAW) resultou em modelos com acurácia menor em comparação com os demais pré-processamentos e semelhante àqueles obtidos com PLSR. Isso indica que o método SVM é mais sensível a alterações causadas pelos pré-processamentos. Isso também foi observado por Khosravi et al. (2018), onde os autores analisaram amostras de solo e sedimento para os elementos Pb e Zn, utilizando vários pré-processamentos semelhantes aos deste estudo, combinados com PLSR e SVM como método de calibração multivariada. Os autores observaram resultados superiores para espectros com pré-processamentos (*Savitzky Golay* (SG), primeira derivada (FD), segunda derivada (SD) e SVM em comparação com o PLSR. Tiecher et al., (2021) também encontraram boas predições para a contribuição de fonte de sedimento para o pré-processamento SGD combinado com o modelo SVM.

Conforme visto no início da abordagem, o efeito do pré-processamento no realce das absorções do espectro NIR, teve maior efeito quando se utilizou o método multivariado SVM. Cabe ressaltar que os dados originais (RAW) e o pré-processamento SNV, o qual apenas suaviza o espectro, tiveram menor acurácia no SVM em comparação aos demais pré-processamento.

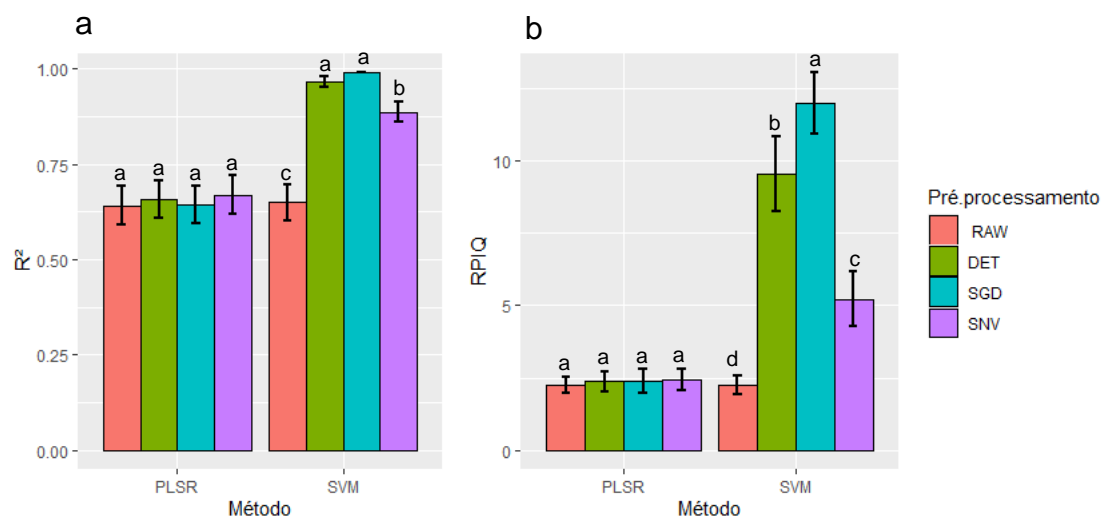


Figura 11. Comparação do desempenho dos modelos calibrados com diferentes técnicas de pré-processamento indicado pelos valores de R^2 (a) e RPIQ (b). As barras representam valores médios e o desvio padrão de 132 modelos. RAW: espectros brutos, DET: *Detrend*, SNV: *Standard Normal Variate*, SGD: *Savitzky-Golay derivative*.

*Médias seguidas pela mesma letra comparando os métodos de calibração multivariados não diferem estatisticamente pelo teste de Tukey a $p < 0,05$.

O desempenho dos modelos em relação à combinação de modelos e técnicas de pré-processamento espectral, foram relacionados com a particularidade da abordagem matemática utilizada. Embora seja bem conhecido que a combinação entre modelo e pré-processamento afeta a precisão da predição (VASQUES et al., 2008; KNOX et al., 2015; PINHEIRO et al., 2017; DOTTO et al., 2018; MOURA-BUENO et al., 2019), a melhor combinação entre eles permanece ser obscura e depende das características do solo e sedimento.

6.3.3 Efeito do elemento

Os elementos foram divididos em categorias conforme proposto por (VEUM et al., 2015; CHANG et al., 2001) sendo interpretadas pelo R^2 e RPIQ respectivamente e, os elementos que tiveram as maiores acurácias (Categoria A) nas predições quando observado o coeficiente de determinação (R^2), sendo: Al, Be, C, Ca, Co, Cu, Fe, Li, Mg, Mn, Na, Ni, Sr e V (Figura 12a), onde considera-se satisfatório $R^2 > 0,75$ (VEUM et al., 2015; CHANG et al., 2001). Estudos de O'Rourke et al. (2016) mostraram acurácia satisfatória na predição da

concentração de C, Al, Ba, Ca, Co, Cu, Fe, K, La, Mg, Na Sr, Ti e V, por meio de modelos espectrais Vis-NIR.

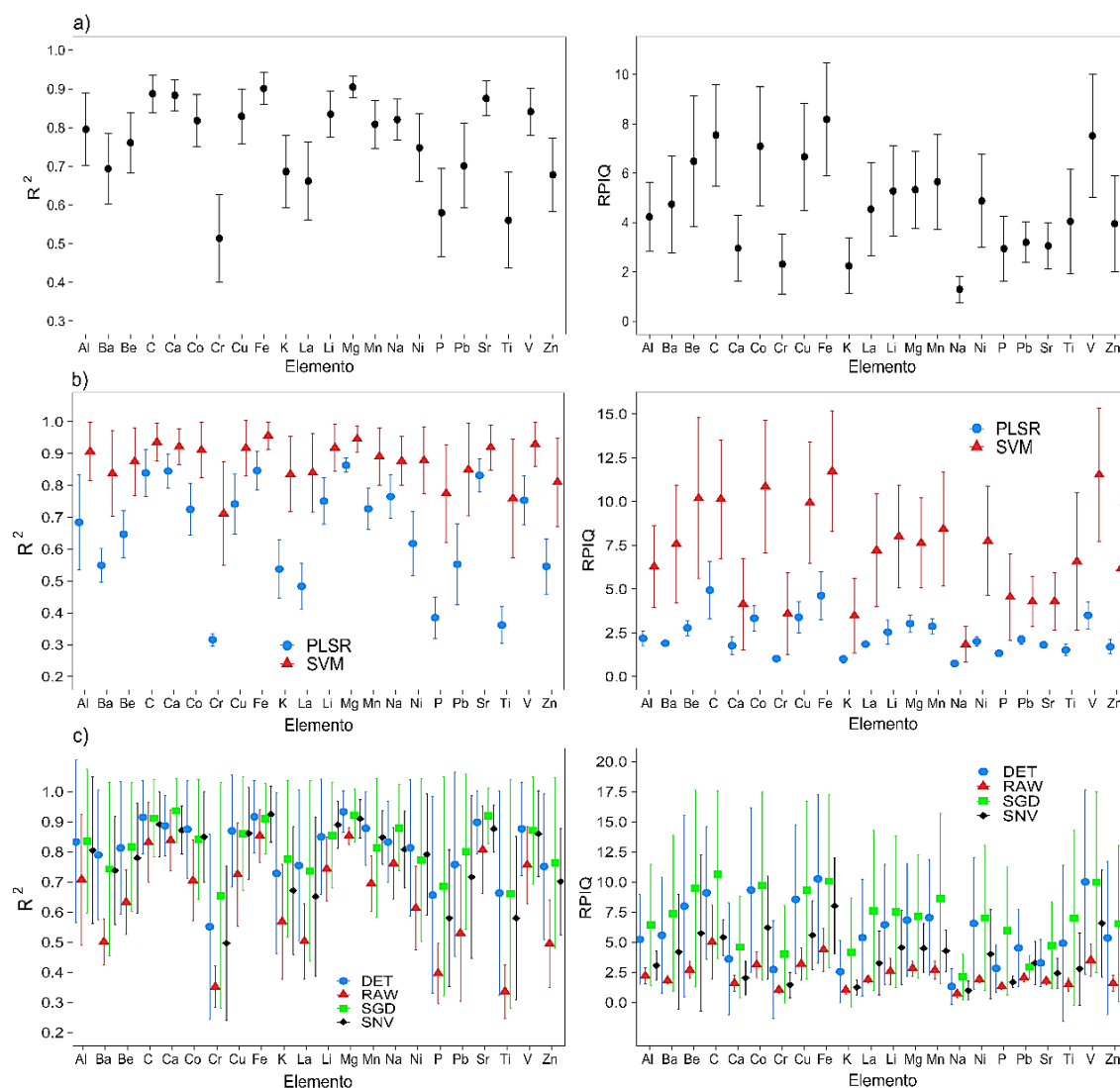


Figura 12. Resposta média dos quatro pré-processamentos e dos dois modelos analisados para predição de cada elemento de acordo com a faixa espectral NIR. a) R^2 e RPIQ para a faixa espectral NIR; b) R^2 e RPIQ para os métodos multivariados SVM e PLSR; c) R^2 e RPIQ para os pré-processamentos DET, SGD e SNV, mais os dados brutos sem pré-processamento.

*Al: alumínio; Ba: bário; Be: berílio; Ca: cálcio; Co: cobalto; Cr: cromo; Cu: cobre; Fe: ferro; K: potássio; La: lantânio; Li: lítio; Mg: magnésio; Mn: manganês; Na: sódio; Ni: níquel; P: fósforo; Pb; chumbo; Sr: estrôncio; Ti: titânio; V: vanádio; Zn: zinco.

A melhor acurácia observada para as predições de alguns desses elementos pode ser relacionada com a interação que se têm entre o elemento e a matriz do solo, como é o caso do C, Ca, Fe, Al, Mg e Mn. Esses elementos

estão diretamente relacionados com a composição da rocha matriz do solo, ou são componentes da matriz sólida orgânica do solo, ou ainda, pelo fato de serem cátions predominantes nos sítios de troca do solo (Ca, Mg, Na). No caso da composição organo-mineral do solo e sedimento, o teor de argila e matéria orgânica tem efeito direto nos espectros, conseqüentemente, nas predições, como observados em estudos realizados no Brasil (DALMOLIN et al., 2005; DEMATTÊ et al., 2019; MOURA-BUENO et al., 2020; TIECHER et al., 2021) e no mundo (MOROS et al., 2010; SORIANO-DISLA et al., 2013; VISCARRA ROSSEL et al., 2016; JIANG et al., 2018; REIDEL et al., 2018). Como a Capacidade de Troca Catiônica (CTC) é diretamente proporcional ao teor de argila e matéria orgânica, esses cátions (Ca, Mg, Na) tem uma correlação indireta com estas propriedades. Viscarra Rossel et al. (2016) alcançaram $R^2 = 0,73$ na predição da CTC para solos da biblioteca espectral mundial, e relacionaram esse resultado com a relação indireta de matéria orgânica e argila. Da mesma forma, Demattê et al. (2019) encontraram $R^2 = 0,68$ para solos de todo o Brasil e $R^2 = 0,64$ para solos do estado do Rio Grande do Sul.

Os valores de coeficiente de determinação precisam estar entre $R^2 > 0,63$ e $R^2 < 0,75$, o qual é considerado categoria B (VEUM et al., 2015; CHANG et al., 2001). Portanto, os elementos que estão dentro desta faixa ($R^2 > 0,63$ e $R^2 < 0,75$) são: Ba, K, La, Pb e Zn. O'Rourke et al. (2016), utilizando modelos espectrais na faixa do Vis-NIR e MIR encontraram acurácia moderadamente satisfatória para a concentração dos elementos: Ba, K, La e Zn.

Os elementos Cr, Ti e P alcançaram $R^2 = 0,51$, $R^2 = 0,56$ e $R^2 = 0,58$, respectivamente, sendo os elementos com acurácia inferior nas predições comparado aos demais elementos analisados (Figura 12a). Conforme a classificação sugerida por Veum et al. (2015) e Chang et al. (2001), esses valores são considerados de acurácia insatisfatória (categoria C - $R^2 > 0,50$ e $< 0,63$). A explicação para essa acurácia inferior na predição pode ser devido esses elementos estarem retidos em argilo-minerais do solo, onde relação com a reflectância espectral é indireta para muitos elementos, como por exemplo, Cr, Ti e P. Neste caso, pode ocorrer sobreposição de absorção na mesma banda espectral, sendo essa absorção relacionada a vários elementos. O mesmo comportamento ocorre para regiões de absorção relacionadas a grupos funcionais da matéria orgânica, em que grupos vibratórios C-H, N-H, O-H, C-O,

C-N, N-O, e C-C podem ou não estarem correlacionados com esses elementos. Ainda, existe a relação dos elementos com compostos espectralmente ativos na região do visível (Vis), tais como óxidos e oxi-hidróxidos de Al, Fe e Mn (O'ROURKE et al., 2016). No entanto, essa região do espectro não foi contemplada do presente estudo. Esse cenário resulta em dificuldades para os modelos estabelecerem relações entre a mudança na reflectância espectral de algumas bandas e a concentração de único elemento em específico. Ou seja, ocorre uma confusão, em que elementos que não possuam uma relação direta com a matéria orgânica e argila, ou que estão em menor concentração no solo e sedimentos, podem ser prejudicados na modelagem espectral.

Valores de RPIQ maiores que $\geq 3,0$, são considerados satisfatórios de acordo com Veum et al. (2015) e Chang et al. (2001). Os elementos que ficaram nesta faixa foram: Al, Ba, Be, C, Co, Cu, Fe, La, Li, Mg, Mn, Ni, Pb, Sr, Ti, e V. Na categoria B ($RPIQ \geq 1,9$ e $< 3,0$) foram: Ca, Cr, K e P. Como único elemento insatisfatório e considerado categoria C, temos o Na com $RPIQ = 1,31$ (Figura 12a).

Os elementos tiveram uma boa predição quando se observa os métodos multivariados utilizados sendo o modelo SVM superior ao PLSR para todos os elementos (Figura 12b). O método PLSR teve a menor acurácia nas predições da concentração dos elementos Cr, K, La, P, Pb e Ti, sendo que estes tiveram boas predições para o método SVM. Isso pode estar relacionado com o comportamento não linear entre a concentração do elemento e reflectância espectral. Portanto, a escolha do método multivariado para a predição dos elementos nos solos e nos sedimentos, interfere diretamente na acurácia dos mesmos.

Os dados com pré-processamentos tiveram melhores predições em comparação com o espectro bruto (RAW), demonstrando a importância de utilizar pré-processamentos para a predição dos elementos em solo e sedimento (Figura 12c).

Na Figura 13, é apresentada a importância relativa das bandas na predição dos elementos C, Cr, Fe, P e Na para o modelo que apresentou a melhor predição (SVM-SGD). Os elementos C e Fe foram escolhidos pois tiveram as melhores predições, Cr e P por terem tido as piores predições e Na por se tratar de um cátion trocável monovalente com maior potencial de

mobilidade do solo. É perceptível que no comprimento de onda de 1135 nm ocorreu diferenças entre a importância dessa banda para explicar a variação na concentração de cada elemento. Para o C, observa-se baixa importância (< 30%), seguido pelo P que teve uma importância intermediária (~ 40%) em comparação com os demais elementos. Os elementos Cr, Fe e Na apresentaram valores maiores de importância (> 80%). A relação do elemento Na nesta banda de absorção possivelmente seja por se tratar de um cátion trocável e monovalente, o qual possui alta solubilidade. Este comprimento de onda (1135 nm) corresponde às vibrações da molécula de água (VISCARRA ROSSEL; BEHRENS, 2010) e grupos aromáticos (C-H) (STENBERG et al., 2010). Neste sentido, a associação de cada elemento com a água e grupos funcionais gera um comportamento espectral diferente, refletindo a importância atribuída ao modelo para cada bandas/região espectral.

A faixa espectral de 1395–1400 nm teve pouca importância (~30%) para os elementos C, Cr e Na, intermediária (~ 50%) para P, e maior importância (~ 80%) para Fe, possivelmente por estar fortemente relacionado com os constituintes do solo como os minerais caulinita, esmectita, mica e vermiculita hidroxil-intercalada, associadas ao grupo funcional O-H (TERRA, 2011; MADEJOVÁ et al., 2011). As esmectitas exibem as conotações de água mais fortes em torno de 1410 nm, enquanto as caulinitas exibem os sobretons OH mais fortes próximos a 1380 e 1410 nm (BISHOP et al., 2008). Portanto, essas diferenças na importância das bandas na predição de cada elemento podem ser explicadas pela forma como os elementos estão associados aos grupos funcionais dos constituintes minerais do solo e sedimento.

Diferenças significativas na importância das bandas na predição também são observadas em 1730 e 1754 nm, para o elemento C (> 80%), onde ocorre a vibração de alguns grupos funcionais da matéria orgânica (VISCARRA ROSSEL; BEHRENS, 2010). Também na banda de absorção de 1730 nm encontram grupos funcionais C-H metil (STENBERG et al., 2010). Já para Cr, Fe e Na a importância é menor (< 40%) neste comprimento de onda (Figura 9), pois estes elementos não têm relação direta com o carbono presente nos solos e sedimentos. De 1800 a 1850 tem importância entre 70 e 80% para o Na, nota-se que indicando maior relação do Na com grupos funcionais de minerais de

argila que exibem feições de absorção nessas bandas (VISCARRA ROSSEL; BEHRENS, 2010).

Após 1900 nm, a importância das bandas na explicação da variação da concentração de Na no solo e sedimento é < 30%, indicando que este elemento não tem relação direta com a matéria orgânica e com os principais grupos funcionais desta região espectral. Nessa região ocorre bandas de absorção vibracionais derivadas do íon CO_3^{2-} , derivadas de carbonatos (CLARK, 1999). As fortes absorções perto de 1908 nm nos espectros dos grupos indicam a ocorrência de água ligada na entrecamada reticuladas (BISHOP et al., 1994). Os elementos C e P tiveram importâncias de (~ 90%) para a banda de absorção de 1900 e 1915, possivelmente por estarem intimamente ligados as vibrações de ácidos carboxílicos presentes na matéria orgânica dos solos e sedimentos (STENBERG et al., 2010). A esmectita também tem uma banda de absorção muito pronunciada próximo de 1900 nm devido à combinação de vibrações de água ligada na entrecamada como cátions hidratados e água adsorvida nas superfícies das partículas (WETTELIND et al., 2013). Bandas fracas de H-O-H são observadas perto 1920 nm na caulinita, embora a molécula de água não faça parte do mineral estrutural como no caso das esmectitas. Essa banda é provavelmente devido a uma pequena quantidade de moléculas de água.

A banda de absorção de 2033 nm corresponde aos grupos funcionais amida (STENBERG et al., 2010). Por este motivo, a importância da reflectância para o elemento C (~ 60%) ocorreu devido a correlação desde elemento com os grupos amida. Já em 2160 nm, Viscarra Rossel; Behrens (2010) e Madejová et al. (2011) encontraram caulinita nos grupos funcionais Al-OH. Em 2200 nm ocorre as vibrações moleculares relacionadas aos grupos Al-OH (FANG et al., 2018), podendo ser ocasionadas pela vermiculita, montmorilonita, illita e caulinita. Nessa banda, os elementos C, P e Fe tem maior importância (> 60%) em comparação aos demais (< 50%).

O comprimento de onda de 2200 nm também está relacionado às vibrações de CO_3^{2-} (FANG et al., 2018). E como pode ser observado na figura 12, as maiores importâncias dessa banda na predição são encontradas para o elemento C e P (> 90%), devido estarem diretamente ligados a matéria orgânica presentes nos solos e sedimentos. Os espectros de esmectitas e caulinitas, ambos têm fortes bandas de combinação de OH em 2200 nm (BISHOP et al.,

2008). Essas bandas de combinação NIR foram investigadas em termos de modos de alongamento e flexão OH em função de composição de cátions octaédricos variável (BISHOP et al., 2002a; BISHOP et al., 2002b; GATES, 2005). Para as bandas de absorção de 2230 nm ocorrem as vibrações de esmectita, grupos funcionais (Al-OH) (VISCARRA ROSSEL; BEHRENS, 2010; MADEJOVÁ et al., 2011).

Os carbonatos também têm várias bandas de absorção na região NIR, sendo a mais forte em torno de 2300 nm (CLARK et al., 1990). Na banda de absorção de 2336-2450 nm ocorre a vibração de micas com o grupo funcional (O-H) + (Al-OH) (TERRA, 2011; MADEJOVÁ et al., 2011). O elemento Fe teve altos valores de importância de reflectância (~ 80%) neste comprimento de onda, pois bandas entre 2400 - 2500nm remetem às combinações de alongamento e flexão OH e por estar associado a tetraédrica substituição de Al ou Fe por Si ou com ligação entre o cátion octaédrico e o cátion tetraédrico (BISHOP et al., 2002a; GATES, 2005). Na banda de absorção no comprimento de onda de 2469 nm ocorreu maior importância para o elemento Fe (> 90%), isso pode estar intimamente relacionado com o Al Fe-OH e bandas de combinação Fe₂-OH no nontronite e os espectros são observados em 2400 nm em espectros de biotita (BISHOP et al., 2008).

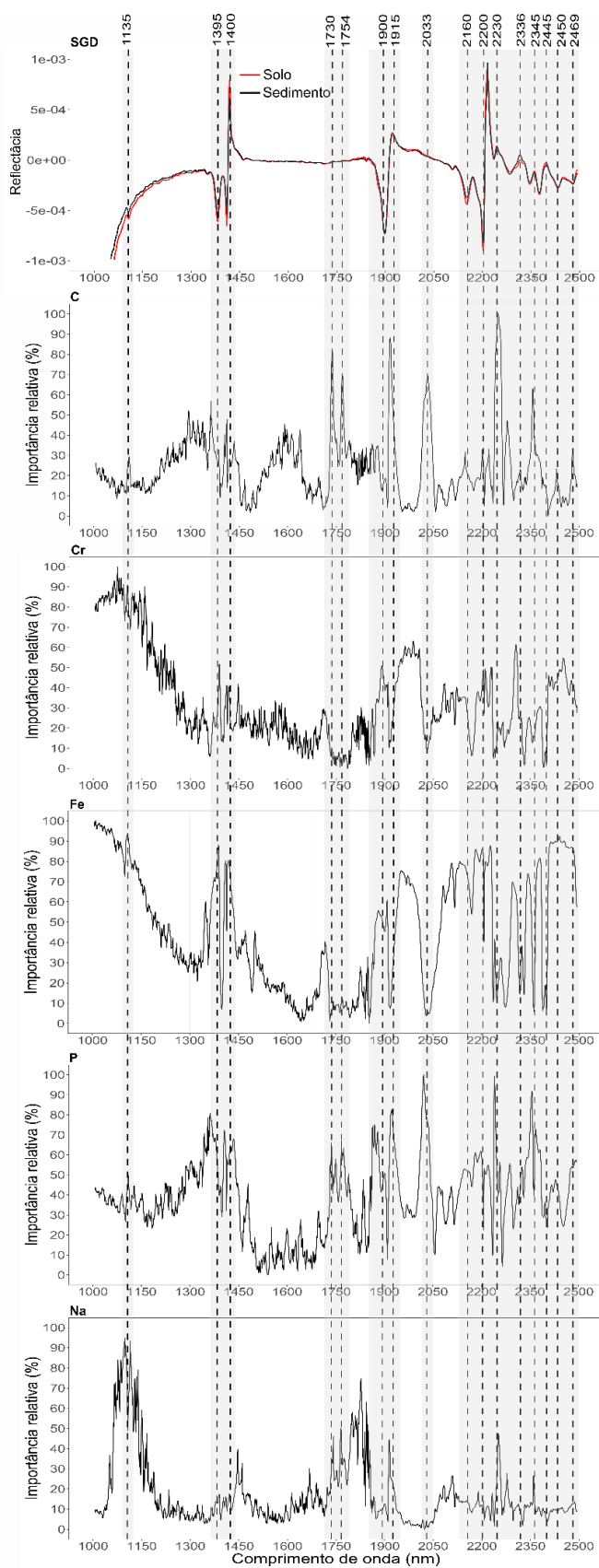


Figura 13. Espectro NIR processado com SGD (a) e importância relativa das bandas na predição pelo modelo SVM-SGD na concentração dos elementos carbono (C), cromo (Cr), ferro (Fe), fósforo (P) e sódio (Na).

Embora alguns modelos tenham resultado de confiabilidade intermediária segundo Veum et al. (2015) e Chang et al. (2001), para alguns elementos quando analisamos R^2 , RPIQ e RMSE, os resultados demonstram que utilizando a melhor combinação de modelo multivariado + pré-processamento, é possível fazer uma boa estimativa da concentração de todos os elementos estudados por meio da espectroscopia de reflectância tanto para amostras de solo quanto para amostras de sedimento (Tabela 6). Vale ressaltar que, por mais que os valores encontrados para RMSE serem altos, quando se observa a acurácia, os valores até são baixos com base no teor dos elementos. Logicamente que para isso, o usuário deve assumir o erro embutido nessa predição, que para muitos casos, foi extremamente baixo (Tabela 6).

Os maiores erros médios absolutos (RMSE) encontrados para solo foram para os elementos Mg e Ca, Ti (111 mg kg^{-1} , 149 mg kg^{-1} e 117 mg kg^{-1}), respectivamente (Tabela 6). Pode-se verificar que o erro embutido na predição para estes elementos é alto, porém se considerarmos o teor médio dos elementos é um valor aceitável. Contudo, deve ser levado em consideração na decisão da técnica espectroscópica a ser utilizada. Para o sedimento os elementos com maiores erros médios absolutos (RMSE) foram Mg (170 mg kg^{-1}), Ca (117 mg kg^{-1}) e o Na (97 mg kg^{-1}).

O elemento que teve a maior incerteza (%) nas predições tanto para solo quanto para sedimento foi o Na com 12,6% solo e 20% sedimento, respectivamente. Contudo, foi verificado que para muitos casos a incerteza na predição pode ser inferior a 5% para solo (Fe, Al, C, Mg, Ti, Mn, Ba, P, Li, Sr, La, Pb, Zn e Be) e para sedimento (Fe, Al, C, Mg, Ti, Mn, Ba, V, Cu, P, Co, Cr, Li, La, Ni, Pb, Ca e K) desde que usados os modelos e pré-processamentos adequados. Isso demonstra que para valores inferiores a 5%, a espectroscopia de reflectância tem grande potencial de uso em estudos de monitoramento da concentração de elementos em bacias hidrográficas, tanto para solo, quanto para sedimento.

Tabela 6. Modelos espectrais com melhor desempenho preditivo para os 22 elementos avaliados calibrados utilizando amostras de solo e amostras de sedimento.

Elemento	Unidade	Melhor Modelo Solo	RMSE	Teor médio	Incerteza (%)	Elemento	Unidade	Melhor Modelo Sedimento	RMSE	Teor médio	Incerteza (%)
Al	g kg ⁻¹	IVP - SVM - SGD	1,5	51,9	2,8	Al	g kg ⁻¹	IVP - SVM - SGD	0,8	37,0	2,2
Ba	mg kg ⁻¹	IVP - SVM - SGD	7,4	208,7	3,6	Ba	mg kg ⁻¹	IVP - SVM - SGD	9,7	325,0	3,0
Be	mg kg ⁻¹	IVP - SVM - SGD	0,1	3,3	3,9	Be	mg kg ⁻¹	IVP - SVM - DET	0,3	3,5	9,1
C	g kg ⁻¹	IVP - SVM - SGD	0,8	18,3	4,3	C	g kg ⁻¹	IVP - SVM - DET	1,3	29,1	4,6
Ca	mg kg ⁻¹	IVP - SVM - DET	116,9	2021,8	5,8	Ca	mg kg ⁻¹	IVP - SVM - SGD	447,7	4428,8	10,1
Co	mg kg ⁻¹	IVP - SVM - SGD	2,6	43,7	5,9	Co	mg kg ⁻¹	IVP - SVM - SNV	1,8	45,9	3,9
Cr	mg kg ⁻¹	IVP - SVM - SGD	2,1	28,7	7,5	Cr	mg kg ⁻¹	IVP - SVM - DET	1,5	33,5	4,6
Cu	mg kg ⁻¹	IVP - SVM - SGD	12,5	197,9	6,3	Cu	mg kg ⁻¹	IVP - SVM - DET	5,3	155,7	3,4
Fe	g kg ⁻¹	IVP - SVM - SGD	2,9	77,4	3,7	Fe	g kg ⁻¹	IVP - SVM - DET	1,4	53,2	2,7
K	mg kg ⁻¹	IVP - SVM - SGD	85,1	1549,8	5,5	K	mg kg ⁻¹	IVP - SVM - SGD	222,8	2209,8	10,1
La	mg kg ⁻¹	IVP - SVM - SGD	1,4	32,4	4,4	La	mg kg ⁻¹	IVP - SVM - DET	0,9	33,2	2,7
Li	mg kg ⁻¹	IVP - SVM - SGD	1,5	42,9	3,4	Li	mg kg ⁻¹	IVP - SVM - DET	1,7	35,1	4,8
Mg	mg kg ⁻¹	IVP - SVM - DET	149,3	3770,9	4,0	Mg	mg kg ⁻¹	IVP - SVM - DET	170,0	4642,9	3,7
Mn	mg kg ⁻¹	IVP - SVM - SGD	93,9	1934,9	4,9	Mn	mg kg ⁻¹	IVP - SVM - DET	60,0	1899,3	3,2
Na	mg kg ⁻¹	IVP - SVM - SGD	10,3	81,7	12,6	Na	mg kg ⁻¹	IVP - SVM - DET	97,9	489,0	20,0
Ni	mg kg ⁻¹	IVP - SVM - SGD	1,7	23,0	7,3	Ni	mg kg ⁻¹	IVP - SVM - SGD	0,7	21,5	3,4
P	mg kg ⁻¹	IVP - SVM - SGD	16,9	377,4	4,5	P	mg kg ⁻¹	IVP - SVM - SGD	16,9	441,9	3,8
Pb	mg kg ⁻¹	IVP - SVM - SGD	0,8	23,3	3,6	Pb	mg kg ⁻¹	IVP - SVM - SGD	0,6	19,2	3,4
Sr	mg kg ⁻¹	IVP - SVM - SGD	1,4	29,2	4,8	Sr	mg kg ⁻¹	IVP - SVM - SGD	3,7	53,0	7,1
Ti	mg kg ⁻¹	IVP - SVM - SGD	111,1	3667,8	3,0	Ti	mg kg ⁻¹	IVP - SVM - SGD	81,8	6951,7	1,2
V	mg kg ⁻¹	IVP - SVM - DET	16,9	291,6	5,8	V	mg kg ⁻¹	IVP - SVM - DET	12,5	287,4	4,3
Zn	mg kg ⁻¹	IVP - SVM - SGD	0,5	14,3	3,4	Zn	mg kg ⁻¹	IVP - SVM - DET	2,2	32,3	6,8

*Teor médio obtido pelo método de química úmida; Al: alumínio; Ba: bário; Be: berílio; Ca: cálcio; Co: cobalto; Cr: cromo; Cu: cobre; Fe: ferro; K: potássio; La: lantânio; Li: lítio; Mg: magnésio; Mn: manganês; Na: sódio; Ni: níquel; P: fósforo; Pb; chumbo; Sr: estrôncio; Ti: titânio; V: vanádio; Zn: zinco; RMSE: raiz quadrada do erro médio.

6.3.4 Efeito do conjunto de dados e predição da concentração de elementos em amostras sedimentos com modelos espectrais calibrados com amostras solo de uma bacia hidrográfica

No método PLSR, os valores de R^2 para o conjunto de dados foi: sedimento sendo inferior ao solo e igual ao solo + sedimento (Figura 14). Para o RPIQ o conjunto de dados que teve melhor desempenho aos demais foi o solo que teve valor de RPIQ = 2,87, sendo superior aos demais conjuntos de dados de solo + sedimento (RPIQ = 2,22) e sedimento (RPIQ = 1,81) (Figura 14). No método SVM, os três conjuntos de dados analisados não diferiram significativamente entre si, obtendo os valores de $R^2 = 0,88$ para sedimento, $R^2 = 0,87$ para solo e $R^2 = 0,85$ para solo + sedimento (Figura 14). Para RPIQ, o modelo SVM o conjunto com dados de solo + sedimento alcançou o menor desempenho (RPIQ = 6,03), sendo inferior aos demais e diferindo significativamente.

Nota-se que a combinação de conjunto de dados com amostras de solos e sedimentos não ganham em predição (Figura 14). Isso pode ser explicado por alguns fatores como: i) a composição organo-mineral ser diferente para solo e sedimento, já que o sedimento pode ser enriquecido por alguns elementos, como por exemplo, o carbono; ii) o comportamento espectral do espectro SGD para solo e sedimento, em geral, é semelhante, porém existem diferenças nas absorções das bandas 1400, 1915 e a partir de 2100 nm. As diferenças a partir de 2100 nm são ocasionadas por grupos funcionais da matéria orgânica, resultando em alterações distintas no espectro do NIR. Essas diferenças parecem pequenas, porém numericamente elas têm efeito na modelagem matemática. Diferenças nos espectros UV-Vis-NIR-MIR de sedimentos e fontes de sedimentos também foram observadas por Tiecher et al. (2021) em uma bacia hidrográfica no Sul do Brasil. Os autores atribuíram as diferenças no espectro NIR relacionadas ao teor e composição dos argilominerais e matéria orgânica.

As possíveis alternativas encontradas para resultar em uma melhor predição dos conjuntos amostrais de solo + sedimento, podem ser: i) separar as amostras de sedimento em função da técnica de amostragem; ii) separar amostras por uso do solo da região (estratificação da biblioteca espectral); iii)

comparar separadamente as amostras (solo x sedimento) do terço superior, médio e inferior; iv) estratificar diferentes momentos do hidrograma separadamente ou de forma conjunta com as estratégias de amostragem feitas; v) testar em bacias menores e mais homogêneas ou em escala de sub-bacias.

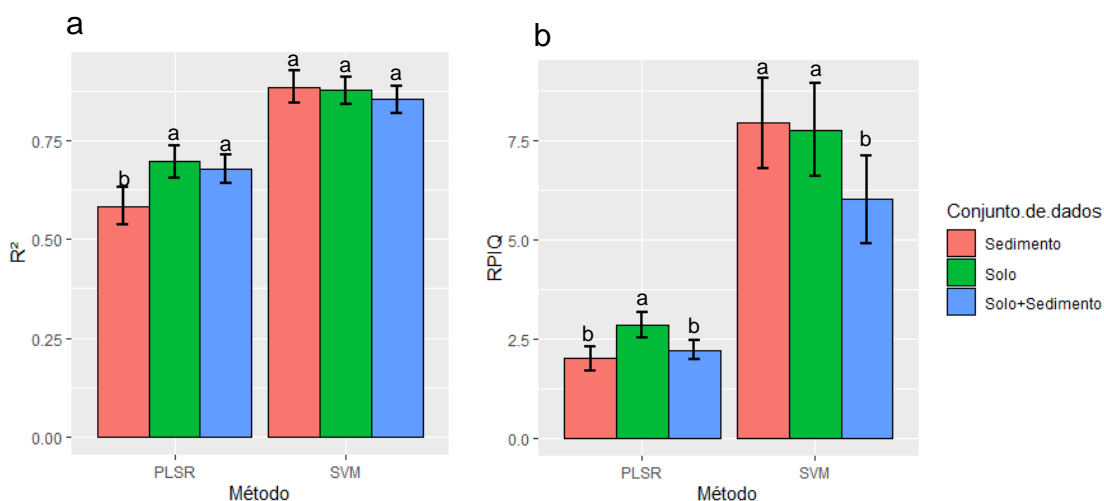


Figura 14. Comparação do desempenho dos modelos calibrados com conjuntos de dados de sedimento, solo e solo + sedimento indicado pelos valores de R^2 (a) e RPIQ (b). As barras representam valores médios e o desvio padrão de 176 modelos.

*Médias seguidas pela mesma letra comparando os métodos de calibração multivariados não diferem estatisticamente pelo teste de Tukey a $p < 0,05$.

Após usando os melhores modelos calibrados com amostras de solo, não foi possível prever com acurácia a concentração de nenhum dos 22 elementos avaliados em amostras de sedimento (Tabela 7). A melhor estimativa dentre os 22 elementos avaliados nessa abordagem foi obtida com o fósforo (P), e mesmo assim, com baixa acurácia para o R^2 , alto RMSE e baixo RPIQ ($R^2 = 0,46$, RMSE = 910,81 mg kg⁻¹ e RPIQ = 0,49 – Tabela 7). Esses resultados estão de acordo com os obtidos a partir dos modelos utilizando a combinação de amostras de solo e de sedimento (Figura 14), que produziram modelos com menor acurácia do que aqueles construídos utilizando os conjuntos isolados de solo e de sedimento. Embora se esperasse ser possível estimar o teor de elementos nas amostras de sedimento utilizando modelos calibrados com amostras de solo, pelo fato do sedimento ser uma mistura das principais fontes de sedimento

operante na bacia hidrográfica (TIECHER et al., 2021), os resultados acima evidenciam que para a bacia do Rio Guaporé existe diferença na matriz sólida desses materiais que impossibilitam essa abordagem. Uma das alternativas que podem ser utilizadas para melhorar a qualidade dos modelos em estudos futuros é a separação das amostras de sedimento por fonte. Além disso, outra alternativa seria separar amostras de sedimento obtidos a partir de diferentes estratégias de coletas, que podem impactar na distribuição de tamanho de partícula.

Tabela 7. Seleção dos melhores pré-processamentos para cada um dos 22 elementos utilizando amostras de solo e resultado da predição dos mesmos utilizando a assinatura espectroscópica de amostras de sedimento para a faixa espectral NIR e método multivariado SVM.

Elemento	Unidade	Pré-processamento	Calibração com amostras de solo			Predição com amostras de sedimento		
			R ²	RMSE	RPIQ	R ²	RMSE	RPIQ
Al	g kg ⁻¹	SGD	0,99	1473,54	11,03	0,13	25185,87	0,36
Ba	mg kg ⁻¹	SGD	0,99	7,43	11,62	0,03	132,25	1,00
Be	mg kg ⁻¹	SGD	0,99	0,13	16,42	0,07	3,85	1,73
C	g kg ⁻¹	SGD	0,99	0,79	15,01	0,06	22,03	0,33
Ca	mg kg ⁻¹	Detrend	0,99	116,99	12,63	0,03	5148,07	0,27
Co	mg kg ⁻¹	SGD	0,99	2,57	18,82	0,12	20,16	1,23
Cr	mg kg ⁻¹	SGD	0,99	2,16	6,82	0,00	37,02	0,44
Cu	mg kg ⁻¹	SGD	0,99	12,48	18,88	0,16	154,11	0,43
Fe	g kg ⁻¹	SGD	0,99	2900,43	18,87	0,10	45799,66	0,40
K	mg kg ⁻¹	SGD	0,99	85,13	11,86	0,41	1930,49	0,45
La	mg kg ⁻¹	SGD	0,99	1,44	13,65	0,00	19,97	0,58
Li	mg kg ⁻¹	SGD	0,99	1,46	11,64	0,09	40,18	0,80
Mg	mg kg ⁻¹	Detrend	0,99	149,32	13,71	0,15	1826,26	0,84
Mn	mg kg ⁻¹	SGD	0,99	93,86	16,48	0,02	939,58	0,84
Na	mg kg ⁻¹	SGD	0,99	10,31	4,68	0,03	1109,45	0,40
Ni	mg kg ⁻¹	SGD	0,99	1,67	12,84	0,00	22,80	0,40
P	mg kg ⁻¹	SGD	0,99	16,98	11,21	0,46	910,81	0,49
Pb	mg kg ⁻¹	SGD	0,99	0,83	13,21	0,00	9,97	0,94
Sr	mg kg ⁻¹	SGD	0,99	1,40	9,65	0,28	37,56	0,53
Ti	mg kg ⁻¹	SGD	0,99	111,05	13,96	0,02	5593,63	1,46
V	mg kg ⁻¹	Detrend	0,99	16,97	19,30	0,27	130,28	1,35
Zn	mg kg ⁻¹	SGD	0,99	0,48	12,72	0,02	27,23	1,27

Al: alumínio; Ba: bário; Be: berílio; Ca: cálcio; Co: cobalto; Cr: cromo; Cu: cobre; Fe: ferro; K: potássio; La: lantânio; Li: lítio; Mg: magnésio; Mn: manganês; Na: sódio; Ni: níquel; P: fósforo; Pb; chumbo; Sr: estrôncio; Ti: titânio; V: vanádio; Zn: zinco; SGD: Savitzky-Golay Derivative.

Os resultados obtidos no presente estudo reforçam que é imprescindível monitorar e coletar amostras tanto de solo quanto sedimento nas bacias hidrográficas, para posteriormente obter uma biblioteca espectral de amostras de solo e sedimento mais robusta, que proporcione a calibração de modelos para solo e para sedimento. Essa condição é similar à encontrada em alguns estudos de espectroscopia em solos usando diferentes escalas amostrais, onde se observa que previsões de carbono orgânico podem ser mais acuradas quando os modelos são calibrados com dados locais ou regionais (MOURA-BUENO et al., 2019; DEMATTÊ et al., 2019; MOURA-BUENO et al., 2020). Ao transferir os modelos desenvolvidos de um local (região) para outro pode ocorrer perda de acurácia nas previsões, devido às características intrínsecas do solo e sedimento, resultado da diversidade pedológica e geomorfológica (GRUNWALD et al., 2015).

Ou seja, os modelos espectrais devem ser regionais e considerar amostras locais e suas características, assim como em bacias hidrográficas, onde deve ser levado em consideração a relação das particularidades observadas na modelagem de sedimentos em bacias hidrográficas, e os modelos devem ser exclusivos para amostras de solo e para amostras de sedimento. Outro aspecto importante, é a relação com a variação dos metais em solos conforme a região geomorfológica do RS, conforme visto por Althaus et al. (2018). Os autores verificaram que os teores médios de metais como Cd, Co, Cr, Cu, Ni, Pb, V e Zn foram maiores nos solos derivados do material de origem basalto comparado a solos derivados de outras rochas (sedimentares e metamórficas) presentes no RS.

A bacia hidrográfica do Rio Guaporé possui grande extensão (2.032 km²) e, uma grande diversidade de tipos e uso do solo, com condições de relevo e variação geológica. É possível que o número de amostras e as principais fontes de sedimento selecionadas não tenham sido suficientes para caracterizar as áreas com maior contribuição para o sedimento que chega até a rede drenagem.

Além disso, o processo de desagregação e transporte das partículas de solo erodido até os locais de coleta na rede de drenagem podem ter alterado o material, seja pela seletividade de algumas partículas ou pelo enriquecimento/exaurimento de alguns elementos durante o processo erosivo.

Esses fatores podem ter interferido negativamente na estimativa da concentração de elementos em amostras de sedimento utilizando modelos calibrados com amostras de solo na bacia hidrográfica do Rio Guaporé. Dessa forma, estudos futuros devem ser realizados em bacias menores, mais homogêneas em termos geologia, e de tipos e usos de solo para avaliar a utilização dessa abordagem, que pode ser uma excelente opção para monitoramento de perdas de nutrientes em sedimentos.

7. CONCLUSÕES

A predição dos 22 elementos estudados, tanto em amostras de solo quanto em amostras de sedimento, foi mais acurada quando utilizando o modelo de calibração multivariada *Support Vector Machine* (SVM) combinado com o pré-processamento SGD. Neste sentido, a técnica espectroscopia de reflectância difusa (NIR) pode ser utilizada para prever as concentrações dos elementos em solos e sedimentos com acurácia, devendo levar em consideração o erro associado a técnica.

A predição da concentração dos elementos presentes em sedimentos através de modelos calibrados com amostras de solo teve baixa acurácia. Por este motivo, em estudos futuros deve-se levar em consideração bacias hidrográficas mais homogêneas em termos geomorfológicos, pedológicos e de uso do solo, dando ênfase para amostragem de solo mais representativa da bacia. Assim, a técnica espectroscopia de reflectância difusa (NIR) poderá proporcionar modelos mais acurados das concentrações de elementos em sedimentos.

8. REFERÊNCIAS

ALTHAUS, D. *et al.* Natural fertility and metals contents in soils of Rio Grande do Sul (Brasil). **Revista Brasileira de Ciência do Solo**, Viçosa, MG, v. 42, [art.] e0160418, [p. 1-15], 2018.

ALVARES, C. A. *et al.* Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, Stuttgart, v. 22, n. 6, p. 711-728, 2013.

ANGELOPOULOU, T. *et al.* Reflectance spectroscopy (Vis-NIR) for assessing soil heavy metals concentrations determined by two different analytical protocols, based on ISO 11466 and ISO 14869-1. **Water, Air & Soil Pollution**, Dordrecht, v. 228, [art.] 436, 2020.

ARMENTA, S.; GUARDIA, M. Vibrational spectroscopy in soil and sediment analysis. **Trends in Environmental Analytical Chemistry**, Amsterdam, v. 2, p. 43–52, 2014.

BARNES, R. J.; DHANOA, M. S.; LISTER, S. J. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. **Applied Spectroscopy**, Plainfield, v. 43, p. 772-777, 1989.

BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, M. B. **Chemometrics**: a practical guide. New York: John Wiley, 1998. 360 p.

BELLON-MAUREL, V. *et al.* Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. **Trends in Analytical Chemistry**, Amsterdam, v. 29, n. 9, p. 1073–1081, 2010.

BEN-DOR, E.; BANIN, A. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. **Soil Science Society of America Journal**, Madison, v. 59, n. 2, p. 364-372, 1995.

BEN-DOR, E.; IRONS, J. R.; EPEMA, G. F. Soil reflectance. *In*: RENCZ, A. N. (ed.). **Remote sensing for earth sciences**. New York: Wiley, 1999. p. 111-188.

BISHOP, J. L.; PIETERS, C. M.; EDWARDS, J. O. Infrared spectroscopic analyses on the nature of water in montmorillonite. **Clays and Clay Minerals**, Long Island City, v. 42, p. 701-715, 1994.

- BISHOP, J. L.; MURAD, E.; DYAR, M. D. The influence of octahedral and tetrahedral cation substitution on the structure of smectites and serpentines as observed through infrared spectroscopy. **Clay Minerals**, London, v. 37, p. 617-628, 2002a.
- BISHOP, J. L. *et al.* The influence of structural Fe, Al and Mg on the infrared OH bands in spectra of dioctahedral smectites. **Clay Minerals**, London, v. 37, p. 607-616, 2002b.
- BISHOP, J. L. *et al.* Reflectance and emission spectroscopy study of four groups of phyllosilicates: smectites, kaolinite-serpentines, chlorites and micas. **Clay Minerals**, London, v. 43, p. 35-54, 2008.
- BILGILI, V. A. *et al.* Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. **Journal of Arid Environments**, London, v. 74, n. 2, p. 229–238, 2010.
- BLANCO, M.; VILLARROYA, I. NIR spectroscopy: a rapid-response analytical tool. **Trends in Analytical Chemistry**, Amsterdam, v. 21, n. 4, p. 240–250, 2002.
- BRUS, D. J.; KEMPEN, B.; HEUVELINK, G. B. M. Sampling for validation of digital soil maps. **European Journal of Soil Science**, Oxford, v. 62, n. 3, p. 394-407, 2011.
- CADD, H. J. *et al.* The potential for rapid determination of charcoal from wetland sediments using infrared spectroscopy. **Palaeogeography, Palaeoclimatology, Palaeoecology**, Amsterdam, v. 542, [art.] 109562, [p. 1-13], 2020.
- CARVALHO, J. K. **Avaliação de métodos de pré-processamento e modelagem matemática aplicados à espectroscopia de infravermelho próximo para estimar teores de argila e matéria orgânica do solo.** 2019. 65 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência do Solo, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.
- CARVALHO, A. C. B. *et al.* Análise espectroscópica da matéria orgânica no sedimento superficial da Baía de Sepetiba, Rio de Janeiro, Brasil. **Química Nova**, São Paulo, v. 43, n. 5, p. 552-557, 2020.
- CHAGAS, I. P. **Desenvolvimento de um fotômetro portátil NIR para determinação do teor de água no álcool combustível e do teor de etanol na gasolina.** 2006. 151 f. Tese (Doutorado) – Instituto de Química, Universidade Estadual de Campinas, Campinas, 2006.
- CHANG, C. *et al.* Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. **Soil Science Society of America Journal**, Madison, v. 65, p. 480–490, 2001.

CHEN, T.; MORRIS, J.; MARTIN, E. Gaussian process regression for multivariate spectroscopic calibration. **Chemometrics and Intelligent Laboratory Systems**, Amsterdam, v. 87, n. 1, p. 85–97, 2007.

CHENG, H. *et al.* Estimating heavy metal concentrations in suburban soils with reflectance spectroscopy. **Geoderma**, Amsterdam, v. 336, p. 59-67, 2019.

CLARK, R. N. *et al.* High spectral resolution reflectance spectroscopy of minerals. **Journal of Geophysical Research**, Hoboken, v. 95, p. 12653-12680, 1990.

CLARK, R. N. Spectroscopy of rocks and minerals, and principles of spectroscopy. *In*: RENCZ, A. N. (ed.). **Remote sensing for earth sciences**. New York: Wiley, 1999. p. 3-52.

CLEGG, S. M. *et al.* Multivariate analysis of remote laser-induced breakdown spectroscopy spectra using partial least squares, principal component analysis, and related techniques. **Spectrochimica Acta - Part B. Atomic Spectroscopy**, Amsterdam, v. 64, n. 1, p. 79–88, 2009.

CORTES, C.; VAPNIK, V. Support-vector networks. **Machine Learning**, New York, v. 20, n. 3, p. 273–297, 1995.

COLLINS, A. L. *et al.* Assessing damaged road verges as a suspended sediment source in the Hampshire Avon catchment, Southern United Kingdom. **Hydrological Processes**, Chichester, v. 24, n. 9, p. 1106-1122, 2010.

DALMOLIN, R. S. D. *et al.* Relação entre os constituintes do solo e seu comportamento espectral. **Ciência Rural**, Santa Maria, v. 35, n. 2, p. 481-489, 2005.

DEMATTÊ, J. A. M. *et al.* Morphological interpretation of reflectance spectrum (MIRS) using libraries looking towards soil classification. **Scientia Agricola**, Piracicaba, v. 71, n. 6, p. 509-520, 2014.

DEMATTÊ, J. A. M. *et al.* Soil analytical quality control by traditional and spectroscopy techniques: constructing the future of a hybrid laboratory for low environmental impact. **Geoderma**, Amsterdam, v. 337, p. 111–121, 2018.

DEMATTÊ, J. A. M. *et al.* The Brazilian Soil Spectral Library (BSSL): a general view, application and challenges. **Geoderma**, Amsterdam, v. 354, [art.] 113793, [p. 1-21], 2019.

DOTTO, A. C. *et al.* Two preprocessing techniques to reduce model covariables in soil property predictions by Vis-NIR spectroscopy. **Soil and Tillage Research**, Amsterdam, v. 172, p. 59-68, 2017.

DOTTO, A. C. *et al.* A systematic study on the application of scatter-corrective and spectral- derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. **Geoderma**, Amsterdam, v. 314, p. 262-274, 2018.

DOTTO, A. C. *et al.* AlradSpectra: a quantification toll for soil properties using spectroscopic data in R. **Revista Brasileira de Ciência do Solo**, Viçosa, MG, v. 43, [art.] e0180263, [p. 1-15], 2019.

FANG, Q. *et al.* Visible and near-infrared reflectance spectroscopy for investigating soil mineralogy: a review. **Journal of Spectroscopy**, New York, v. 2018, [art.] ID 3168974, [p. 1-14], 2018.

FERREIRA, M. M. C. *et al.* Quimiometria I: calibracao multivariada, um tutorial. **Química Nova**, São Paulo, v. 22, n. 5, p. 724-731, 1999.

GATES, W. P. Infrared spectroscopy and the chemistry of dioctahedral smectites. *In*: KLOPROGGE, J. T. (ed.). **The application of vibrational spectroscopy to clay minerals and layered double hydroxides**. Aurora, Colorado: The Clay Minerals Society, 2005. p. 125-168.

GHOLIZADEH, A. *et al.* Comparing different data preprocessing methods for monitoring soil heavy metals based on soil spectral features. **Soil and Water Research**, Prague, v. 10, p. 218–227, 2016.

GOMEZ, C.; VISCARRA ROSSEL, R. A.; MCBRATNEY, A. B. Soil organic carbon prediction by hyperspectral remote sensing and field vis–NIR spectroscopy: an Australian case study. **Geoderma**, Amsterdam, v. 146, n. 3/4, p. 403–411, 2008.

GRINAND, C. *et al.* Prediction of soil organic and inorganic carbon contents at a national scale (France) using mid-infrared reflectance spectroscopy (MIRS). **European Journal of Soil Science**, Oxford, v. 63, p. 141–151, 2012.

GUERRERO, C. *et al.* Do we really need large spectral libraries for local scale SOC assessment with NIR spectroscopy? **Soil and Tillage Research**, Amsterdam, v. 155, p. 501–509, 2016.

HE, Y. *et al.* Measurement and analysis of soil nitrogen and organic matter content using near-infrared spectroscopy techniques. **Journal of Zhejiang University Science B**, Hangzhaou, v. 6, n. 11, p. 1081-1086, 2005.

HUNT, G. R.; SALISBURY, J. W. Visible and near-infrared spectra of minerals and rocks: I. Silicate minerals. **Modern Geology**, New York, v. 1, p. 283-300, 1970.

HUNT, G. R.; SALISBURY, J. W.; LENHOFF, C. J. Visible and near-infrared spectra of minerals and rocks: III Oxides and hydroxides. **Modern Geology**, New York, v. 2, p. 195-205, 1971.

HUNT, G. R. Electromagnetic radiation: the communication link in remote sensing. *In*: SIEGAL, B. S.; GILLESPIE, A. R. (ed.). **Remote sensing in geology**. New York: Wiley, 1980. p. 5-45.

IVANCIUC, O. Applications of support vector machines in chemistry. *In*: LIPKOWITZ, K.B.; CUNDARI, T. R. (ed.). **Reviews in computational chemistry**. Weinheim: Wiley, 2007. p. 291–400.

JIANG, Q. *et al.* Feasibility of using visible and near-infrared reflectance spectroscopy to monitor heavy metal contaminants in urban lake sediment. **Catena**, Cremlingen, v. 162, p. 72-79, 2018.

KHOSRAVI, V. *et al.* Monitoring soil lead and zinc contents via combination of spectroscopy with extreme learning machine and other data mining methods. **Geoderma**, Amsterdam, v. 318, p. 29-41, 2018.

KNOX, N. M. *et al.* Modelling soil carbon fractions with visible near-infrared (VNIR) and mid-infrared (MIR) spectroscopy. **Geoderma**, Amsterdam, v. 239/240, p. 229–239, 2015.

LILLESAND, T. M.; KIEFER, R. W.; CHIPMAN, J. **Remote sensing and image interpretation**. Hoboken: Wiley, 1994. 176 p.

LINDBERG, J. D.; SNYDER, D. G. Diffuse reflectance spectra of several clay minerals. **American Mineralogist**, Washington, DC, v. 57, p. 485-493, 1972.

MADEJOVÁ, J.; BALAN, E.; PETIT, S. Application of vibrational spectroscopy to the characterization of phyllosilicates and other industrial minerals. *In*: CHRISTIDIS, G. E. (ed.). **Advances in the characterization of industrial minerals**. London: European Mineralogical Union and the Mineralogical Society of Great Britain and Ireland, 2011. p. 171–226.

MARTIN, P. D. *et al.* Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. **Canadian Journal of Soil Science**, Ottawa, v. 82, p. 413-422, 2002.

MEVIK, B. H.; WEHRENS, R.; LILAND, K. H. **Partial least squares and principal component regression**. Vienna: R CRAN, 2016.

MEYER, D. *et al.* **R**: misc functions of the department of statistics, probability theory group (Formerly: E1071), TU Wien: R Packag. Version 1.7-3. Vienna: R CRAN, 2019.

MICHELAKI, K.; HANCOCK, R. G. V. Reassessment of elemental concentration data of sediments from the western delta of the Nile River. **Open Journal of Archaeometry**, Pavia, v. 1, [art.] e2, p. 6-9, 2013.

MIGUEL, P. *et al.* Solo e dinâmica de ocupação das terras em áreas do rebordo do planalto do Rio Grande do Sul. **Revista Brasileira de Agrociência**, Pelotas, v. 17, n. 4, p. 447-455, 2011.

MINELLA, J. P. G.; MERTEN, G. H. Monitoramento de bacias hidrográficas para identificar fontes de sedimentos em suspensão. **Ciência Rural**, Santa Maria, v. 41, n. 3, p. 424-432, 2011.

MINELLA, J. P. G.; WALLING, D. E.; MERTEN, G. H. Combining sediment source tracing techniques with traditional monitoring to assess the impact of improved land management on catchment sediment yields. **Journal of Hydrology**, Amsterdam, v. 348, n. 3/4, p. 546–563, 2008.

MOROS, J. *et al.* Estuarine sediment quality assessment by Fourier-Transform Infrared Spectroscopy. **Vibrational Spectroscopy**, Amsterdam, v. 53, p. 204-213, 2010.

MORGAN, R. P. C. **Soil erosion and conservation**. 3rd ed. London: Collins, 2005. 316 p.

MOROS, J. *et al.* Use of reflectance infrared spectroscopy for monitoring the metal content of the estuarine sediments of the Nerbioi-Ibaizabal River (Metropolitan Bilbao, Bay of Biscay, Basque Country). **Environmental Science & Technology**, Easton, v. 43, n. 24, p. 9314-9320, 2009.

MOURA-BUENO, J. M. **Uso de bibliotecas espectrais para a predição do carbono orgânico do solo**. 2018. 201 f. Tese (Doutorado em Ciência do Solo) – Programa de Pós-Graduação em Ciência do Solo, Universidade Federal de Santa Maria, Santa Maria, 2018.

MOURA-BUENO, J. M. *et al.* Stratification of a local VIS-NIR-SWIR spectral library by homogeneity criteria yields more accurate soil organic carbon predictions. **Geoderma**, Amsterdam, v. 337, p. 565-581, 2019.

MOURA-BUENO, J. M. *et al.* When does stratification of a subtropical soil spectral library improve predictions of soil organic carbon content? **Science of the Total Environment**, Amsterdam, v. 737, [art.] 139895, [p. 1-21], 2020.

NELSON, D. V.; SOMMERS, L. E. Total carbon, organic carbon, and organic matter. *In*: PAGE, A. L.; MILLER, R. H.; KEENEY, D. R. (ed.), **Methods of soil analysis. Part 2. Chemical and biological methods**. Madison: American Society of Agronomy and Soil Science of America, 1982. p. 539–579.

NIEMZ, M. H. Light and matter. *In*: HOUCK, M. M.; SIEGEL, J. A. **Fundamentals of forensic science**. Amsterdam: Academic Press, 2007. p. 94-119.

NOCITA, M. *et al.* Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. **Soil Biology and Biochemistry**, Amsterdam, v. 68, p. 337-347, 2014.

O'ROURKE, S. M. *et al.* An assessment of model averaging to improve predictive power of portable vis-NIR and XRF for the determination of agronomic soil properties. **Geoderma**, Amsterdam, v. 279, p. 31-44, 2016.

PASQUINI, C. Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society**, Campinas, v. 14, n. 2, p. 198–219, 2003.

PHILLIPS, J. M.; RUSSELL, M. A.; WALLING, D. E. Time-integrated sampling of fluvial suspended sediment: a simple methodology for small catchments. **Hydrological Processes**, Chichester, v. 14, n. 14, p. 2589–2602, 2000.

PINHEIRO, E. F. M. *et al.* Prediction of soil physical and chemical properties by Visible and Near-Infrared diffuse reflectance spectroscopy in the central Amazon. **Remote Sensing**, Basel, v. 9, n. 293, p. 1-22, 2017.

R CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2020. Disponível em: <http://www.R-project.org/>. Acesso em: 4 mar. 2021.

REN, H. Y. *et al.* Estimation of As and Cu contamination in agricultural soils around a mining area by reflectance spectroscopy: a case study. **Pedosphere**, Beijing, v. 19, n. 6, p. 719-726, 2009.

RIEDEL, F. *et al.* Prediction of soil parameters using the spectral range between 350 and 15,000 nm: a case study based on the Permanent Soil Monitoring Program in Saxony, Germany. **Geoderma**, Amsterdam, v. 315, p. 188-198, 2018.

RINNAN, A.; BERG, F. V.; ELGELSEN, S. B. Review of the most common pre-processing techniques for near-infrared spectra. **Trends in Analytical Chemistry**, Amsterdam, v. 28, p. 1201-1222, 2009.

ROSIN, N. A. *et al.* Diffuse reflectance spectroscopy for estimating soil organic carbon and make nitrogen recommendations. **Scientia Agricola**, Piracicaba, v. 78, n. 5, [art.] e20190246, [p. 1-13], 2021.

SANTOS, G. A.; PEREIRA, A. B.; KORNDÖRFER, G. H. Uso do sistema de análise por infravermelho próximo (NIR) para análises de matéria orgânica e fração argila em solos e teores foliares de silício e nitrogênio em cana-de-açúcar. **Bioscience Journal**, Uberlândia, v. 26, n. 1, p. 100-108, 2019.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. **Analytical Chemistry**, Amsterdam, v. 36, n. 8, p. 1627–1639, 1964.

SEEMA, K. *et al.* Application of VIS-NIR spectroscopy for estimation of soil organic carbon using different spectral preprocessing techniques and multivariate methods in the middle Indo-Gangetic plains of India. **Geoderma Regional**, Amsterdam, v. 23, [art.] e00349, [p. 1-10], 2020.

SORIANO-DISLA, J. M. *et al.* Prediction of the concentration of chemical elements extracted by aqua regia in agricultural and grazing European soils

using diffuse reflectance mid-infrared spectroscopy. **Applied Geochemistry**, Oxford, v. 39, p. 33-42, 2013.

STENBERG, B. *et al.* Visible and near infrared spectroscopy in soil science. *In*: DONALD, L. S. (ed.). **Advances in agronomy**. Burlington: Academic Press, 2010. p. 163-215.

STEVENS, A.; RAMIREZ-LOPEZ, L. An introduction to the prospectr package. [S.l.: s.n.], 2020. Disponível em: <https://cran.rproject.org/web/packages/prospectr/index.html>. Acesso em: 21 fev. 2020.

STRECK, E. V.; FLORES, C. A.; SCHENEIDER, P. **Solos do Rio Grande do Sul**. Porto Alegre: Emater/RS-Ascar, 2018. 252 p.

STUART, B. H. **Infrared spectroscopy: fundamentals and applications**. Hoboken: Wiley, 2004. 244 p.

STONER, E. R.; BAUMGARDNER, M. F. Characteristic variations in reflectance of surface soils. **Soil Science Society of America Journal**, Madison, v. 45, p. 1161-1165, 1981.

SUMMERS, D. *et al.* Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. **Ecological Indicators**, New York, v. 11, n. 1, p. 123–131, 2011.

SUN, J. Y. *et al.* Real-time analysis of soil moisture, soil organic matter, and soil total nitrogen with NIR spectra. **Spectroscopy and Spectral Analysis**, Beijing, v. 26, n. 3, p. 426-429, 2006.

TERRA, F. S. **Espectroscopia de reflectância do visível ao infravermelho médio aplicada aos estudos qualitativos e quantitativos de solos**. 2011. 375 f. Tese (Doutorado em Ciência do Solo) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2011.

TIECHER, T. *et al.* Tracing sediment sources in a subtropical rural catchment of southern Brazil by using geochemical tracers and near-infrared spectroscopy. **Soil and Tillage Research**, Amsterdam, v. 155, p. 478–491, 2016.

TIECHER, T. *et al.* Quantifying land use contributions to suspended sediment in a large cultivated catchment of Southern Brazil (Guaporé River, Rio Grande do Sul). **Agriculture, Ecosystems & Environment**, Amsterdam, v. 237, p. 95–108, 2017.

TIECHER, T. **Fingerprinting sediment sources in agricultural catchments in Southern Brazil**. 2015. 307 f. Tese (Doutorado em Ciência do Solo) – Programa de Pós-Graduação em Ciência do Solo, Universidade Federal de Santa Maria, Santa Maria, 2015.

TIECHER, T. *et al.* Improving the quantification of sediment source contributions using different mathematical models and spectral preprocessing techniques for individual or combined spectra of ultraviolet–visible, near- and middle-infrared spectroscopy. **Geoderma**, Amsterdam, v. 384, [art.] 114815, [p. 1-16], 2021.

UDELHOVEN, T.; EMMERLING, C.; JARMER, T. Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: a feasibility study. **Plant and Soil**, The Hague, v. 251, p. 319-329, 2003.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY - USEPA. **Test methods for evaluating solid waste, physical and chemical methods**. Washington, DC: USEPA, 1998.

VARMUZA, K.; FILZMOSER, P. **Introduction to multivariate statistical analysis in chemometrics**. New York: CRC Press, 2009. 313 p.

VASQUES, G. M.; GRUNWALD, S.; SICKMAN, J. O. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. **Geoderma**, Amsterdam, v. 146, p. 14-25, 2008.

VEUM, K. *et al.* Estimating a soil quality index with VNIR reflectance spectroscopy. **Soil Science Society of America Journal**, Madison, v. 2, p. 637-649, 2015.

VISCARRA ROSSEL, R. A. *et al.* A global spectral library to characterize the world's soil. **Earth-Science Reviews**, Amsterdam, v. 155, p. 198-230, 2016.

VISCARRA ROSSEL, R. A.; BEHRENS, T. Using data mining to model and interpret soil diffuse reflectance spectra. **Geoderma**, Amsterdam, v. 158, p. 46-54, 2010.

VISCARRA ROSSEL, R. A. *et al.* Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. **Geoderma**, Amsterdam, v. 131, p. 59-75, 2006.

VOHLAND, M. *et al.* Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy. **Geoderma**, Amsterdam, v. 166, n. 1, p. 198–205, 2011.

WALESIAK, M.; DUDEK, A. The choice of variable normalization method in cluster analysis. *In*: SOLIMAN, K. S. (ed.). **Education excellence and innovation management: a 2025 vision to sustain economic development during global challenges**. Sewilla: International Business Information Management Association, 2020. p. 325-340.

WALKLEY, A.; BLACK, I. A. An examination of the Degtjareff method for determining soil organic matter, and proposed modification of the chromic acid titration method. **Soil Science**, New Brunswick, v. 37, n. 1, p. 29-38, 1934.

WETTERLIND, J.; STENBERG, B.; VISCARRA ROSSEL, R. A. Soil analysis using visible and near infrared spectroscopy. *In*: MAATHUIS, F. J. M. (ed.). **Plant mineral nutrients: methods and protocols**. New York: Humana Press, 2013. p. 95-107.

WILDNER, W. *et al.* **Mapa geológico do Estado do Rio Grande do Sul**. Porto Alegre: Serviço Geológico do Brasil, 2008. 1 mapa, Escala 1:750.000.

WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. PLS-regression: a basic tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems**, Amsterdam, v. 58, n. 2, p. 109–130, 2001.

ZORNOZA, R.; GUERRERO, C. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. **Soil Biology & Biochemistry**, Amsterdam, v. 40, n. 7, p. 1923–1930, 2008.

9. ANEXOS

Tabela 8. Dados obtidos após a predição dos 22 elementos para dois métodos multivariado, três pré-processamentos e o espectro bruto, para os conjuntos de amostras de solo e sedimento isolados, e solo e sedimento combinado na faixa espectral NIR.

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	PLSR	RAW	C	Solo	0,81	3,63	3,28
NIR	PLSR	DET	C	Solo	0,84	3,29	3,62
NIR	PLSR	SNV	C	Solo	0,85	3,24	3,67
NIR	PLSR	SGD	C	Solo	0,78	3,91	3,04
NIR	PLSR	RAW	Al	Solo	0,86	5728,59	2,84
NIR	PLSR	DET	Al	Solo	0,87	5610,90	2,90
NIR	PLSR	SNV	Al	Solo	0,88	5385,26	3,02
NIR	PLSR	SGD	Al	Solo	0,86	5825,20	2,79
NIR	PLSR	RAW	Ba	Solo	0,40	58,82	1,47
NIR	PLSR	DET	Ba	Solo	0,59	48,62	1,77
NIR	PLSR	SNV	Ba	Solo	0,62	47,15	1,83
NIR	PLSR	SGD	Ba	Solo	0,45	56,25	1,53
NIR	PLSR	RAW	Be	Solo	0,76	0,67	3,24
NIR	PLSR	DET	Be	Solo	0,81	0,60	3,61
NIR	PLSR	SNV	Be	Solo	0,83	0,57	3,81
NIR	PLSR	SGD	Be	Solo	0,78	0,65	3,35
NIR	PLSR	RAW	Ca	Solo	0,74	616,73	2,36
NIR	PLSR	DET	Ca	Solo	0,81	529,38	2,75
NIR	PLSR	SNV	Ca	Solo	0,85	473,36	3,07
NIR	PLSR	SGD	Ca	Solo	0,74	626,17	2,32
NIR	PLSR	RAW	Co	Solo	0,81	11,53	4,19
NIR	PLSR	DET	Co	Solo	0,88	9,41	5,13
NIR	PLSR	SNV	Co	Solo	0,86	9,87	4,89
NIR	PLSR	SGD	Co	Solo	0,82	11,33	4,26
NIR	PLSR	RAW	Cr	Solo	0,30	18,33	0,81
NIR	PLSR	DET	Cr	Solo	0,30	18,36	0,80
NIR	PLSR	SNV	Cr	Solo	0,32	18,15	0,81

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	PLSR	SGD	Cr	Solo	0,32	18,13	0,81
NIR	PLSR	RAW	Cu	Solo	0,84	52,38	4,50
NIR	PLSR	DET	Cu	Solo	0,89	43,54	5,41
NIR	PLSR	SNV	Cu	Solo	0,89	42,81	5,50
NIR	PLSR	SGD	Cu	Solo	0,85	49,81	4,73
NIR	PLSR	RAW	Fe	Solo	0,92	9174,56	5,97
NIR	PLSR	DET	Fe	Solo	0,94	7833,53	6,99
NIR	PLSR	SNV	Fe	Solo	0,96	6104,17	8,97
NIR	PLSR	SGD	Fe	Solo	0,92	8796,48	6,22
NIR	PLSR	RAW	K	Solo	0,30	727,28	1,39
NIR	PLSR	DET	K	Solo	0,36	697,36	1,45
NIR	PLSR	SNV	K	Solo	0,33	712,14	1,42
NIR	PLSR	SGD	K	Solo	0,42	659,90	1,53
NIR	PLSR	RAW	La	Solo	0,59	9,57	2,05
NIR	PLSR	DET	La	Solo	0,63	9,07	2,16
NIR	PLSR	SNV	La	Solo	0,58	9,67	2,03
NIR	PLSR	SGD	La	Solo	0,60	9,41	2,08
NIR	PLSR	RAW	Li	Solo	0,77	7,23	2,35
NIR	PLSR	DET	Li	Solo	0,76	7,50	2,26
NIR	PLSR	SNV	Li	Solo	0,77	7,37	2,30
NIR	PLSR	SGD	Li	Solo	0,78	7,10	2,39
NIR	PLSR	RAW	Mg	Solo	0,88	550,14	3,72
NIR	PLSR	DET	Mg	Solo	0,91	475,80	4,30
NIR	PLSR	SNV	Mg	Solo	0,90	497,57	4,11
NIR	PLSR	SGD	Mg	Solo	0,85	622,83	3,29
NIR	PLSR	RAW	Mn	Solo	0,78	447,48	3,46
NIR	PLSR	DET	Mn	Solo	0,81	411,86	3,76
NIR	PLSR	SNV	Mn	Solo	0,82	402,46	3,84
NIR	PLSR	SGD	Mn	Solo	0,78	444,77	3,48
NIR	PLSR	RAW	Na	Solo	0,63	65,57	0,74
NIR	PLSR	DET	Na	Solo	0,62	66,45	0,73
NIR	PLSR	SNV	Na	Solo	0,63	65,93	0,73
NIR	PLSR	SGD	Na	Solo	0,68	61,51	0,78
NIR	PLSR	RAW	Ni	Solo	0,72	9,20	2,33
NIR	PLSR	DET	Ni	Solo	0,77	8,35	2,57
NIR	PLSR	SNV	Ni	Solo	0,81	7,61	2,82
NIR	PLSR	SGD	Ni	Solo	0,67	10,00	2,15
NIR	PLSR	RAW	P	Solo	0,48	125,27	1,52
NIR	PLSR	DET	P	Solo	0,49	123,91	1,54
NIR	PLSR	SNV	P	Solo	0,50	122,66	1,55
NIR	PLSR	SGD	P	Solo	0,56	115,72	1,65
NIR	PLSR	RAW	Pb	Solo	0,70	4,68	2,33

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	PLSR	DET	Pb	Solo	0,74	4,34	2,51
NIR	PLSR	SNV	Pb	Solo	0,72	4,55	2,40
NIR	PLSR	SGD	Pb	Solo	0,81	3,74	2,92
NIR	PLSR	RAW	Sr	Solo	0,66	8,47	1,60
NIR	PLSR	DET	Sr	Solo	0,75	7,27	1,86
NIR	PLSR	SNV	Sr	Solo	0,77	6,92	1,95
NIR	PLSR	SGD	Sr	Solo	0,77	7,00	1,93
NIR	PLSR	RAW	Ti	Solo	0,25	972,25	1,59
NIR	PLSR	DET	Ti	Solo	0,27	962,68	1,61
NIR	PLSR	SNV	Ti	Solo	0,26	968,72	1,60
NIR	PLSR	SGD	Ti	Solo	0,27	964,23	1,61
NIR	PLSR	RAW	V	Solo	0,87	65,96	4,93
NIR	PLSR	DET	V	Solo	0,88	63,20	5,15
NIR	PLSR	SNV	V	Solo	0,83	75,88	4,29
NIR	PLSR	SGD	V	Solo	0,89	60,56	5,37
NIR	PLSR	RAW	Zn	Solo	0,67	2,84	2,17
NIR	PLSR	DET	Zn	Solo	0,65	2,95	2,08
NIR	PLSR	SNV	Zn	Solo	0,77	2,36	2,60
NIR	PLSR	SGD	Zn	Solo	0,68	2,81	2,19
NIR	SVM	RAW	C	Solo	0,80	3,74	3,19
NIR	SVM	DET	C	Solo	0,99	0,79	14,99
NIR	SVM	SNV	C	Solo	0,94	2,03	5,88
NIR	SVM	SGD	C	Solo	0,99	0,79	15,01
NIR	SVM	RAW	Al	Solo	0,87	5571,60	2,92
NIR	SVM	DET	Al	Solo	0,99	1584,81	10,25
NIR	SVM	SNV	Al	Solo	0,96	3309,86	4,91
NIR	SVM	SGD	Al	Solo	0,99	1473,54	11,03
NIR	SVM	RAW	Ba	Solo	0,45	56,90	1,52
NIR	SVM	DET	Ba	Solo	0,97	14,70	5,87
NIR	SVM	SNV	Ba	Solo	0,81	35,58	2,43
NIR	SVM	SGD	Ba	Solo	0,99	7,43	11,62
NIR	SVM	RAW	Be	Solo	0,76	0,67	3,23
NIR	SVM	DET	Be	Solo	0,99	0,17	12,99
NIR	SVM	SNV	Be	Solo	0,90	0,44	4,87
NIR	SVM	SGD	Be	Solo	0,99	0,13	16,42
NIR	SVM	RAW	Ca	Solo	0,76	603,97	2,41
NIR	SVM	DET	Ca	Solo	0,99	116,47	12,48
NIR	SVM	SNV	Ca	Solo	0,92	351,24	4,14
NIR	SVM	SGD	Ca	Solo	0,99	118,05	12,31
NIR	SVM	RAW	Co	Solo	0,83	11,23	4,30
NIR	SVM	DET	Co	Solo	0,99	2,78	17,40
NIR	SVM	SNV	Co	Solo	0,94	6,44	7,50

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	SVM	SGD	Co	Solo	0,99	2,57	18,82
NIR	SVM	RAW	Cr	Solo	0,35	18,53	0,80
NIR	SVM	DET	Cr	Solo	0,69	2,33	1,20
NIR	SVM	SNV	Cr	Solo	0,56	15,18	0,97
NIR	SVM	SGD	Cr	Solo	0,99	2,16	6,82
NIR	SVM	RAW	Cu	Solo	0,86	51,02	4,62
NIR	SVM	DET	Cu	Solo	0,99	14,19	16,60
NIR	SVM	SNV	Cu	Solo	0,96	27,85	8,46
NIR	SVM	SGD	Cu	Solo	0,99	12,48	18,88
NIR	SVM	RAW	Fe	Solo	0,93	8458,56	6,47
NIR	SVM	DET	Fe	Solo	0,99	2957,33	18,51
NIR	SVM	SNV	Fe	Solo	0,98	4192,83	13,05
NIR	SVM	SGD	Fe	Solo	0,99	2900,43	18,87
NIR	SVM	RAW	K	Solo	0,39	696,62	1,45
NIR	SVM	DET	K	Solo	0,98	144,03	7,01
NIR	SVM	SNV	K	Solo	0,78	437,83	2,31
NIR	SVM	SGD	K	Solo	0,99	85,13	11,86
NIR	SVM	RAW	La	Solo	0,63	9,10	2,15
NIR	SVM	DET	La	Solo	0,98	2,32	8,45
NIR	SVM	SNV	La	Solo	0,84	6,10	3,21
NIR	SVM	SGD	La	Solo	0,99	1,44	13,65
NIR	SVM	RAW	Li	Solo	0,80	6,91	2,46
NIR	SVM	DET	Li	Solo	0,98	2,21	7,68
NIR	SVM	SNV	Li	Solo	0,92	4,42	3,84
NIR	SVM	SGD	Li	Solo	0,99	1,46	11,64
NIR	SVM	RAW	Mg	Solo	0,86	604,94	3,38
NIR	SVM	DET	Mg	Solo	0,99	149,32	13,71
NIR	SVM	SNV	Mg	Solo	0,96	305,77	6,69
NIR	SVM	SGD	Mg	Solo	0,99	150,38	13,61
NIR	SVM	RAW	Mn	Solo	0,80	434,61	3,56
NIR	SVM	DET	Mn	Solo	0,98	134,28	11,52
NIR	SVM	SNV	Mn	Solo	0,91	284,69	5,43
NIR	SVM	SGD	Mn	Solo	0,99	93,86	16,48
NIR	SVM	RAW	Na	Solo	0,62	69,53	0,69
NIR	SVM	DET	Na	Solo	0,86	40,75	1,18
NIR	SVM	SNV	Na	Solo	0,75	54,72	0,88
NIR	SVM	SGD	Na	Solo	0,99	10,38	4,64
NIR	SVM	RAW	Ni	Solo	0,70	10,16	2,11
NIR	SVM	DET	Ni	Solo	0,99	1,68	12,80
NIR	SVM	SNV	Ni	Solo	0,91	5,97	3,79
NIR	SVM	SGD	Ni	Solo	0,99	1,67	12,84
NIR	SVM	RAW	P	Solo	0,54	124,36	1,53

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	SVM	DET	P	Solo	0,97	33,68	5,65
NIR	SVM	SNV	P	Solo	0,81	80,61	2,36
NIR	SVM	SGD	P	Solo	0,99	16,98	11,21
NIR	SVM	RAW	Pb	Solo	0,73	4,47	2,44
NIR	SVM	DET	Pb	Solo	0,98	1,18	9,22
NIR	SVM	SNV	Pb	Solo	0,90	2,76	3,94
NIR	SVM	SGD	Pb	Solo	0,99	0,83	13,21
NIR	SVM	RAW	Sr	Solo	0,62	9,27	1,46
NIR	SVM	DET	Sr	Solo	0,98	2,43	5,57
NIR	SVM	SNV	Sr	Solo	0,87	5,45	2,48
NIR	SVM	SGD	Sr	Solo	0,99	1,40	9,65
NIR	SVM	RAW	Ti	Solo	0,26	980,93	1,58
NIR	SVM	DET	Ti	Solo	0,94	279,37	5,55
NIR	SVM	SNV	Ti	Solo	0,71	628,65	2,47
NIR	SVM	SGD	Ti	Solo	0,99	111,05	13,96
NIR	SVM	RAW	V	Solo	0,88	64,20	5,07
NIR	SVM	DET	V	Solo	0,99	17,03	19,10
NIR	SVM	SNV	V	Solo	0,97	30,93	10,52
NIR	SVM	SGD	V	Solo	0,99	17,28	18,82
NIR	SVM	RAW	Zn	Solo	0,64	3,05	2,02
NIR	SVM	DET	Zn	Solo	0,99	0,62	9,90
NIR	SVM	SNV	Zn	Solo	0,87	1,89	3,27
NIR	SVM	SGD	Zn	Solo	0,99	0,48	12,72
NIR	PLSR	RAW	C	Sedimento	0,98	1,69	5,81
NIR	PLSR	DET	C	Sedimento	0,98	1,44	6,82
NIR	PLSR	SNV	C	Sedimento	0,98	1,54	6,39
NIR	PLSR	SGD	C	Sedimento	0,99	0,82	11,99
NIR	PLSR	RAW	Al	Sedimento	0,41	6333,04	1,44
NIR	PLSR	DET	Al	Sedimento	0,33	6759,83	1,35
NIR	PLSR	SNV	Al	Sedimento	0,34	6710,45	1,36
NIR	PLSR	SGD	Al	Sedimento	0,41	6349,14	1,43
NIR	PLSR	RAW	Ba	Sedimento	0,50	71,35	1,85
NIR	PLSR	DET	Ba	Sedimento	0,57	66,45	1,99
NIR	PLSR	SNV	Ba	Sedimento	0,52	69,85	1,89
NIR	PLSR	SGD	Ba	Sedimento	0,46	74,35	1,78
NIR	PLSR	RAW	Be	Sedimento	0,55	2,10	2,94
NIR	PLSR	DET	Be	Sedimento	0,51	2,21	2,79
NIR	PLSR	SNV	Be	Sedimento	0,52	2,19	2,82
NIR	PLSR	SGD	Be	Sedimento	0,57	2,06	3,00
NIR	PLSR	RAW	Ca	Sedimento	0,96	1000,06	1,38
NIR	PLSR	DET	Ca	Sedimento	0,85	1913,11	0,72
NIR	PLSR	SNV	Ca	Sedimento	0,90	1595,14	0,87

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	PLSR	SGD	Ca	Sedimento	0,99	598,44	2,31
NIR	PLSR	RAW	Co	Sedimento	0,54	12,68	1,95
NIR	PLSR	DET	Co	Sedimento	0,60	11,82	2,09
NIR	PLSR	SNV	Co	Sedimento	0,59	11,98	2,07
NIR	PLSR	SGD	Co	Sedimento	0,52	12,96	1,91
NIR	PLSR	RAW	Cr	Sedimento	0,34	13,08	1,26
NIR	PLSR	DET	Cr	Sedimento	0,34	13,10	1,26
NIR	PLSR	SNV	Cr	Sedimento	0,34	13,07	1,26
NIR	PLSR	SGD	Cr	Sedimento	0,38	12,67	1,30
NIR	PLSR	RAW	Cu	Sedimento	0,51	38,61	1,71
NIR	PLSR	DET	Cu	Sedimento	0,54	37,42	1,76
NIR	PLSR	SNV	Cu	Sedimento	0,59	35,70	1,84
NIR	PLSR	SGD	Cu	Sedimento	0,53	38,05	1,73
NIR	PLSR	RAW	Fe	Sedimento	0,73	7989,72	2,28
NIR	PLSR	DET	Fe	Sedimento	0,70	8431,15	2,16
NIR	PLSR	SNV	Fe	Sedimento	0,76	7542,18	2,42
NIR	PLSR	SGD	Fe	Sedimento	0,71	8186,11	2,23
NIR	PLSR	RAW	K	Sedimento	0,72	1231,03	0,70
NIR	PLSR	DET	K	Sedimento	0,62	1443,51	0,60
NIR	PLSR	SNV	K	Sedimento	0,66	1372,73	0,63
NIR	PLSR	SGD	K	Sedimento	0,66	1360,25	0,64
NIR	PLSR	RAW	La	Sedimento	0,29	7,61	1,52
NIR	PLSR	DET	La	Sedimento	0,46	6,66	1,74
NIR	PLSR	SNV	La	Sedimento	0,31	7,51	1,54
NIR	PLSR	SGD	La	Sedimento	0,39	7,05	1,64
NIR	PLSR	RAW	Li	Sedimento	0,61	4,01	1,49
NIR	PLSR	DET	Li	Sedimento	0,53	4,37	1,36
NIR	PLSR	SNV	Li	Sedimento	0,87	4,21	1,42
NIR	PLSR	SGD	Li	Sedimento	0,58	4,14	1,44
NIR	PLSR	RAW	Mg	Sedimento	0,88	644,61	2,38
NIR	PLSR	DET	Mg	Sedimento	0,87	668,29	2,30
NIR	PLSR	SNV	Mg	Sedimento	0,86	695,02	2,21
NIR	PLSR	SGD	Mg	Sedimento	0,80	823,75	1,86
NIR	PLSR	RAW	Mn	Sedimento	0,68	355,22	2,21
NIR	PLSR	DET	Mn	Sedimento	0,80	278,89	2,81
NIR	PLSR	SNV	Mn	Sedimento	0,80	278,73	2,82
NIR	PLSR	SGD	Mn	Sedimento	0,46	456,07	1,72
NIR	PLSR	RAW	Na	Sedimento	0,90	357,73	1,21
NIR	PLSR	DET	Na	Sedimento	0,86	416,22	1,04
NIR	PLSR	SNV	Na	Sedimento	0,89	375,85	1,15
NIR	PLSR	SGD	Na	Sedimento	0,88	394,34	1,10
NIR	PLSR	RAW	Ni	Sedimento	0,38	6,03	1,53

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	PLSR	DET	Ni	Sedimento	0,46	5,62	1,64
NIR	PLSR	SNV	Ni	Sedimento	0,44	5,75	1,60
NIR	PLSR	SGD	Ni	Sedimento	0,38	6,04	1,52
NIR	PLSR	RAW	P	Sedimento	0,29	147,16	1,06
NIR	PLSR	DET	P	Sedimento	0,27	149,35	1,05
NIR	PLSR	SNV	P	Sedimento	0,28	148,15	1,05
NIR	PLSR	SGD	P	Sedimento	0,27	149,35	1,05
NIR	PLSR	RAW	Pb	Sedimento	0,29	5,55	1,69
NIR	PLSR	DET	Pb	Sedimento	0,25	5,72	1,64
NIR	PLSR	SNV	Pb	Sedimento	0,26	5,69	1,65
NIR	PLSR	SGD	Pb	Sedimento	0,40	5,12	1,83
NIR	PLSR	RAW	Sr	Sedimento	0,94	10,09	1,97
NIR	PLSR	DET	Sr	Sedimento	0,89	13,36	1,49
NIR	PLSR	SNV	Sr	Sedimento	0,91	12,35	1,61
NIR	PLSR	SGD	Sr	Sedimento	0,88	13,83	1,44
NIR	PLSR	RAW	Ti	Sedimento	0,36	3945,81	2,07
NIR	PLSR	DET	Ti	Sedimento	0,36	3944,21	2,07
NIR	PLSR	SNV	Ti	Sedimento	0,37	3907,31	2,09
NIR	PLSR	SGD	Ti	Sedimento	0,34	3997,70	2,04
NIR	PLSR	RAW	V	Sedimento	0,58	83,05	2,12
NIR	PLSR	DET	V	Sedimento	0,63	77,86	2,26
NIR	PLSR	SNV	V	Sedimento	0,63	77,95	2,26
NIR	PLSR	SGD	V	Sedimento	0,56	85,18	2,07
NIR	PLSR	RAW	Zn	Sedimento	0,32	18,94	1,83
NIR	PLSR	DET	Zn	Sedimento	0,45	17,04	2,03
NIR	PLSR	SNV	Zn	Sedimento	0,47	16,68	2,08
NIR	PLSR	SGD	Zn	Sedimento	0,35	18,52	1,87
NIR	SVM	RAW	C	Sedimento	0,98	1,41	6,97
NIR	SVM	DET	C	Sedimento	0,99	0,94	10,42
NIR	SVM	SNV	C	Sedimento	0,99	0,92	10,63
NIR	SVM	SGD	C	Sedimento	0,99	1,00	9,78
NIR	SVM	RAW	Al	Sedimento	0,48	6036,56	1,51
NIR	SVM	DET	Al	Sedimento	0,99	1138,46	7,99
NIR	SVM	SNV	Al	Sedimento	0,88	2961,11	3,07
NIR	SVM	SGD	Al	Sedimento	0,99	806,32	11,29
NIR	SVM	RAW	Ba	Sedimento	0,50	71,09	1,86
NIR	SVM	DET	Ba	Sedimento	0,99	9,89	13,39
NIR	SVM	SNV	Ba	Sedimento	0,99	9,86	13,42
NIR	SVM	SGD	Ba	Sedimento	0,99	9,68	13,68
NIR	SVM	RAW	Be	Sedimento	0,59	2,03	3,05
NIR	SVM	DET	Be	Sedimento	0,99	0,31	20,10
NIR	SVM	SNV	Be	Sedimento	0,99	0,34	18,17

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	SVM	SGD	Be	Sedimento	0,99	0,31	19,90
NIR	SVM	RAW	Ca	Sedimento	0,95	1241,00	1,11
NIR	SVM	DET	Ca	Sedimento	0,99	463,71	2,98
NIR	SVM	SNV	Ca	Sedimento	0,98	708,42	1,95
NIR	SVM	SGD	Ca	Sedimento	0,99	447,66	3,08
NIR	SVM	RAW	Co	Sedimento	0,55	12,75	1,94
NIR	SVM	DET	Co	Sedimento	0,99	1,82	13,59
NIR	SVM	SNV	Co	Sedimento	0,99	1,81	13,66
NIR	SVM	SGD	Co	Sedimento	0,99	1,84	13,45
NIR	SVM	RAW	Cr	Sedimento	0,48	11,86	1,39
NIR	SVM	DET	Cr	Sedimento	0,99	1,55	10,59
NIR	SVM	SNV	Cr	Sedimento	0,93	4,71	3,49
NIR	SVM	SGD	Cr	Sedimento	0,99	1,59	10,35
NIR	SVM	RAW	Cu	Sedimento	0,52	38,61	1,70
NIR	SVM	DET	Cu	Sedimento	0,99	5,30	12,43
NIR	SVM	SNV	Cu	Sedimento	0,98	7,51	8,76
NIR	SVM	SGD	Cu	Sedimento	0,99	5,38	12,24
NIR	SVM	RAW	Fe	Sedimento	0,77	7321,34	2,49
NIR	SVM	DET	Fe	Sedimento	0,99	1442,32	12,63
NIR	SVM	SNV	Fe	Sedimento	0,99	11825,98	9,98
NIR	SVM	SGD	Fe	Sedimento	0,99	1477,52	12,33
NIR	SVM	RAW	K	Sedimento	0,74	1246,23	0,70
NIR	SVM	DET	K	Sedimento	0,99	223,66	3,87
NIR	SVM	SNV	K	Sedimento	0,92	750,65	1,15
NIR	SVM	SGD	K	Sedimento	0,99	222,85	3,89
NIR	SVM	RAW	La	Sedimento	0,47	6,57	1,76
NIR	SVM	DET	La	Sedimento	0,99	0,88	13,18
NIR	SVM	SNV	La	Sedimento	0,98	1,40	8,29
NIR	SVM	SGD	La	Sedimento	0,99	0,89	13,03
NIR	SVM	RAW	Li	Sedimento	0,62	4,01	1,49
NIR	SVM	DET	Li	Sedimento	0,99	0,61	9,84
NIR	SVM	SNV	Li	Sedimento	0,99	0,61	9,71
NIR	SVM	SGD	Li	Sedimento	0,99	0,61	9,74
NIR	SVM	RAW	Mg	Sedimento	0,86	694,92	2,21
NIR	SVM	DET	Mg	Sedimento	0,99	170,00	9,03
NIR	SVM	SNV	Mg	Sedimento	0,99	218,87	7,01
NIR	SVM	SGD	Mg	Sedimento	0,99	171,73	8,94
NIR	SVM	RAW	Mn	Sedimento	0,56	417,69	1,88
NIR	SVM	DET	Mn	Sedimento	0,99	60,04	13,07
NIR	SVM	SNV	Mn	Sedimento	0,97	111,92	7,01
NIR	SVM	SGD	Mn	Sedimento	0,99	60,63	12,94
NIR	SVM	RAW	Na	Sedimento	0,84	465,54	0,93

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	SVM	DET	Na	Sedimento	0,99	104,35	4,15
NIR	SVM	SNV	Na	Sedimento	0,98	178,78	2,42
NIR	SVM	SGD	Na	Sedimento	0,99	106,66	4,06
NIR	SVM	RAW	Ni	Sedimento	0,56	5,13	1,79
NIR	SVM	DET	Ni	Sedimento	0,99	0,75	12,22
NIR	SVM	SNV	Ni	Sedimento	0,99	0,83	11,11
NIR	SVM	SGD	Ni	Sedimento	0,99	0,73	12,58
NIR	SVM	RAW	P	Sedimento	0,32	145,58	1,07
NIR	SVM	DET	P	Sedimento	0,97	34,58	4,52
NIR	SVM	SNV	P	Sedimento	0,79	86,19	1,81
NIR	SVM	SGD	P	Sedimento	0,99	16,96	9,22
NIR	SVM	RAW	Pb	Sedimento	0,23	5,85	1,60
NIR	SVM	DET	Pb	Sedimento	0,99	0,65	14,51
NIR	SVM	SNV	Pb	Sedimento	0,96	1,50	6,26
NIR	SVM	SGD	Pb	Sedimento	0,99	0,65	14,46
NIR	SVM	RAW	Sr	Sedimento	0,94	9,78	2,03
NIR	SVM	DET	Sr	Sedimento	0,99	3,80	5,23
NIR	SVM	SNV	Sr	Sedimento	0,99	4,18	4,75
NIR	SVM	SGD	Sr	Sedimento	0,99	3,74	5,31
NIR	SVM	RAW	Ti	Sedimento	0,36	4082,02	2,00
NIR	SVM	DET	Ti	Sedimento	0,99	477,21	17,10
NIR	SVM	SNV	Ti	Sedimento	0,97	958,93	8,51
NIR	SVM	SGD	Ti	Sedimento	0,99	81,84	16,93
NIR	SVM	RAW	V	Sedimento	0,64	77,64	2,27
NIR	SVM	DET	V	Sedimento	0,99	12,46	14,15
NIR	SVM	SNV	V	Sedimento	0,99	13,67	12,90
NIR	SVM	SGD	V	Sedimento	0,99	12,50	14,10
NIR	SVM	RAW	Zn	Sedimento	0,38	18,70	1,85
NIR	SVM	DET	Zn	Sedimento	0,99	2,21	15,68
NIR	SVM	SNV	Zn	Sedimento	0,88	8,49	4,08
NIR	SVM	SGD	Zn	Sedimento	0,99	2,25	15,43
NIR	PLSR	RAW	C	Solo+Sedimento	0,71	6,51	3,60
NIR	PLSR	DET	C	Solo+Sedimento	0,71	6,45	3,63
NIR	PLSR	SNV	C	Solo+Sedimento	0,71	6,43	3,64
NIR	PLSR	SGD	C	Solo+Sedimento	0,73	6,24	3,75
NIR	PLSR	RAW	Al	Solo+Sedimento	0,80	6788,69	2,15
NIR	PLSR	DET	Al	Solo+Sedimento	0,84	6068,91	2,41
NIR	PLSR	SNV	Al	Solo+Sedimento	0,84	5948,89	2,45
NIR	PLSR	SGD	Al	Solo+Sedimento	0,77	7225,80	2,02
NIR	PLSR	RAW	Ba	Solo+Sedimento	0,60	64,08	2,14
NIR	PLSR	DET	Ba	Solo+Sedimento	0,65	60,62	2,26
NIR	PLSR	SNV	Ba	Solo+Sedimento	0,65	60,49	2,27

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	PLSR	SGD	Ba	Solo+Sedimento	0,58	66,36	2,07
NIR	PLSR	RAW	Be	Solo+Sedimento	0,61	1,46	1,96
NIR	PLSR	DET	Be	Solo+Sedimento	0,61	1,46	1,95
NIR	PLSR	SNV	Be	Solo+Sedimento	0,63	1,43	2,00
NIR	PLSR	SGD	Be	Solo+Sedimento	0,58	1,52	1,87
NIR	PLSR	RAW	Ca	Solo+Sedimento	0,84	1366,91	1,33
NIR	PLSR	DET	Ca	Solo+Sedimento	0,76	1661,04	1,09
NIR	PLSR	SNV	Ca	Solo+Sedimento	0,78	1601,80	1,13
NIR	PLSR	SGD	Ca	Solo+Sedimento	0,92	958,93	1,89
NIR	PLSR	RAW	Co	Solo+Sedimento	0,74	12,16	3,17
NIR	PLSR	DET	Co	Solo+Sedimento	0,80	10,85	3,55
NIR	PLSR	SNV	Co	Solo+Sedimento	0,80	10,72	3,59
NIR	PLSR	SGD	Co	Solo+Sedimento	0,74	12,19	3,16
NIR	PLSR	RAW	Cr	Solo+Sedimento	0,31	16,63	1,01
NIR	PLSR	DET	Cr	Solo+Sedimento	0,27	17,09	0,98
NIR	PLSR	SNV	Cr	Solo+Sedimento	0,28	17,04	0,99
NIR	PLSR	SGD	Cr	Solo+Sedimento	0,29	16,84	1,00
NIR	PLSR	RAW	Cu	Solo+Sedimento	0,81	47,79	3,34
NIR	PLSR	DET	Cu	Solo+Sedimento	0,82	47,12	3,39
NIR	PLSR	SNV	Cu	Solo+Sedimento	0,82	46,34	3,45
NIR	PLSR	SGD	Cu	Solo+Sedimento	0,81	48,38	3,30
NIR	PLSR	RAW	Fe	Solo+Sedimento	0,88	10146,40	4,50
NIR	PLSR	DET	Fe	Solo+Sedimento	0,89	9748,07	4,69
NIR	PLSR	SNV	Fe	Solo+Sedimento	0,89	9490,92	4,82
NIR	PLSR	SGD	Fe	Solo+Sedimento	0,86	10756,96	4,25
NIR	PLSR	RAW	K	Solo+Sedimento	0,62	1002,85	0,94
NIR	PLSR	DET	K	Solo+Sedimento	0,56	1079,64	0,88
NIR	PLSR	SNV	K	Solo+Sedimento	0,59	1052,04	0,90
NIR	PLSR	SGD	K	Solo+Sedimento	0,61	1015,56	0,93
NIR	PLSR	RAW	La	Solo+Sedimento	0,50	9,12	1,89
NIR	PLSR	DET	La	Solo+Sedimento	0,54	8,82	1,95
NIR	PLSR	SNV	La	Solo+Sedimento	0,45	9,58	1,80
NIR	PLSR	SGD	La	Solo+Sedimento	0,46	9,48	1,81
NIR	PLSR	RAW	Li	Solo+Sedimento	0,83	9,00	3,82
NIR	PLSR	DET	Li	Solo+Sedimento	0,85	8,56	4,02
NIR	PLSR	SNV	Li	Solo+Sedimento	0,86	8,31	4,14
NIR	PLSR	SGD	Li	Solo+Sedimento	0,80	9,80	3,51
NIR	PLSR	RAW	Mg	Solo+Sedimento	0,82	734,43	2,63
NIR	PLSR	DET	Mg	Solo+Sedimento	0,85	683,33	2,83
NIR	PLSR	SNV	Mg	Solo+Sedimento	0,83	719,93	2,98
NIR	PLSR	SGD	Mg	Solo+Sedimento	0,91	519,09	3,72
NIR	PLSR	RAW	Mn	Solo+Sedimento	0,67	485,36	2,47

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	PLSR	DET	Mn	Solo+Sedimento	0,72	443,96	2,70
NIR	PLSR	SNV	Mn	Solo+Sedimento	0,73	440,50	2,72
NIR	PLSR	SGD	Mn	Solo+Sedimento	0,67	482,79	2,48
NIR	PLSR	RAW	Na	Solo+Sedimento	0,81	294,83	0,41
NIR	PLSR	DET	Na	Solo+Sedimento	0,76	336,91	0,36
NIR	PLSR	SNV	Na	Solo+Sedimento	0,77	324,85	0,37
NIR	PLSR	SGD	Na	Solo+Sedimento	0,75	341,95	0,35
NIR	PLSR	RAW	Ni	Solo+Sedimento	0,72	7,74	2,04
NIR	PLSR	DET	Ni	Solo+Sedimento	0,69	8,07	1,96
NIR	PLSR	SNV	Ni	Solo+Sedimento	0,75	7,28	2,17
NIR	PLSR	SGD	Ni	Solo+Sedimento	0,62	8,93	1,77
NIR	PLSR	RAW	P	Solo+Sedimento	0,36	141,87	1,36
NIR	PLSR	DET	P	Solo+Sedimento	0,39	138,19	1,40
NIR	PLSR	SNV	P	Solo+Sedimento	0,41	136,02	1,42
NIR	PLSR	SGD	P	Solo+Sedimento	0,32	146,32	1,32
NIR	PLSR	RAW	Pb	Solo+Sedimento	0,61	5,10	2,06
NIR	PLSR	DET	Pb	Solo+Sedimento	0,62	5,04	2,09
NIR	PLSR	SNV	Pb	Solo+Sedimento	0,61	5,06	2,08
NIR	PLSR	SGD	Pb	Solo+Sedimento	0,62	4,99	2,11
NIR	PLSR	RAW	Sr	Solo+Sedimento	0,87	10,61	2,08
NIR	PLSR	DET	Sr	Solo+Sedimento	0,81	12,93	1,71
NIR	PLSR	SNV	Sr	Solo+Sedimento	0,83	12,14	1,82
NIR	PLSR	SGD	Sr	Solo+Sedimento	0,90	9,62	2,29
NIR	PLSR	RAW	Ti	Solo+Sedimento	0,48	2579,92	0,90
NIR	PLSR	DET	Ti	Solo+Sedimento	0,50	2535,56	0,92
NIR	PLSR	SNV	Ti	Solo+Sedimento	0,49	2540,78	0,91
NIR	PLSR	SGD	Ti	Solo+Sedimento	0,39	2793,74	0,83
NIR	PLSR	RAW	V	Solo+Sedimento	0,78	77,46	3,25
NIR	PLSR	DET	V	Solo+Sedimento	0,78	76,78	3,28
NIR	PLSR	SNV	V	Solo+Sedimento	0,80	72,84	3,46
NIR	PLSR	SGD	V	Solo+Sedimento	0,81	70,95	3,55
NIR	PLSR	RAW	Zn	Solo+Sedimento	0,49	12,37	0,83
NIR	PLSR	DET	Zn	Solo+Sedimento	0,56	11,41	0,90
NIR	PLSR	SNV	Zn	Solo+Sedimento	0,56	11,40	0,90
NIR	PLSR	SGD	Zn	Solo+Sedimento	0,58	11,21	0,92
NIR	SVM	RAW	C	Solo+Sedimento	0,70	6,62	3,54
NIR	SVM	DET	C	Solo+Sedimento	0,98	1,55	15,11
NIR	SVM	SNV	C	Solo+Sedimento	0,89	3,98	5,89
NIR	SVM	SGD	C	Solo+Sedimento	0,99	1,16	20,26
NIR	SVM	RAW	Al	Solo+Sedimento	0,83	6288,56	2,32
NIR	SVM	DET	Al	Solo+Sedimento	0,98	2261,93	6,45
NIR	SVM	SNV	Al	Solo+Sedimento	0,93	4102,08	3,56

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	SVM	SGD	Al	Solo+Sedimento	0,99	1457,11	10,02
NIR	SVM	RAW	Ba	Solo+Sedimento	0,56	67,94	2,02
NIR	SVM	DET	Ba	Solo+Sedimento	0,97	16,80	8,16
NIR	SVM	SNV	Ba	Solo+Sedimento	0,84	41,81	3,28
NIR	SVM	SGD	Ba	Solo+Sedimento	0,99	9,86	13,61
NIR	SVM	RAW	Be	Solo+Sedimento	0,53	1,62	1,75
NIR	SVM	DET	Be	Solo+Sedimento	0,97	0,44	6,47
NIR	SVM	SNV	Be	Solo+Sedimento	0,81	1,03	2,76
NIR	SVM	SGD	Be	Solo+Sedimento	0,99	0,23	12,47
NIR	SVM	RAW	Ca	Solo+Sedimento	0,78	1876,93	0,97
NIR	SVM	DET	Ca	Solo+Sedimento	0,92	1121,71	1,62
NIR	SVM	SNV	Ca	Solo+Sedimento	0,80	1774,88	1,02
NIR	SVM	SGD	Ca	Solo+Sedimento	0,99	320,59	5,66
NIR	SVM	RAW	Co	Solo+Sedimento	0,76	11,84	3,25
NIR	SVM	DET	Co	Solo+Sedimento	0,99	2,71	14,23
NIR	SVM	SNV	Co	Solo+Sedimento	0,92	6,98	5,52
NIR	SVM	SGD	Co	Solo+Sedimento	0,99	2,33	16,51
NIR	SVM	RAW	Cr	Solo+Sedimento	0,33	16,65	1,01
NIR	SVM	DET	Cr	Solo+Sedimento	0,72	10,73	1,57
NIR	SVM	SNV	Cr	Solo+Sedimento	0,55	14,02	1,20
NIR	SVM	SGD	Cr	Solo+Sedimento	0,96	4,50	3,73
NIR	SVM	RAW	Cu	Solo+Sedimento	0,81	49,03	3,26
NIR	SVM	DET	Cu	Solo+Sedimento	0,99	13,62	11,73
NIR	SVM	SNV	Cu	Solo+Sedimento	0,93	29,14	5,48
NIR	SVM	SGD	Cu	Solo+Sedimento	0,99	10,66	14,99
NIR	SVM	RAW	Fe	Solo+Sedimento	0,89	10067,44	4,54
NIR	SVM	DET	Fe	Solo+Sedimento	0,99	2759,25	16,56
NIR	SVM	SNV	Fe	Solo+Sedimento	0,97	5174,23	8,83
NIR	SVM	SGD	Fe	Solo+Sedimento	0,99	2772,29	16,48
NIR	SVM	RAW	K	Solo+Sedimento	0,64	1031,57	0,92
NIR	SVM	DET	K	Solo+Sedimento	0,86	636,92	1,48
NIR	SVM	SNV	K	Solo+Sedimento	0,75	886,71	1,07
NIR	SVM	SGD	K	Solo+Sedimento	0,99	158,03	5,98
NIR	SVM	RAW	La	Solo+Sedimento	0,54	8,81	1,95
NIR	SVM	DET	La	Solo+Sedimento	0,93	3,58	4,80
NIR	SVM	SNV	La	Solo+Sedimento	0,75	6,57	2,62
NIR	SVM	SGD	La	Solo+Sedimento	0,99	1,27	13,54
NIR	SVM	RAW	Li	Solo+Sedimento	0,83	9,01	3,82
NIR	SVM	DET	Li	Solo+Sedimento	0,99	2,53	13,57
NIR	SVM	SNV	Li	Solo+Sedimento	0,93	5,84	5,89
NIR	SVM	SGD	Li	Solo+Sedimento	0,99	2,10	16,36
NIR	SVM	RAW	Mg	Solo+Sedimento	0,82	747,74	2,58

Continuação >

Espectro	Método	Pré-processamento	Elemento	Conjunto de dados	R ²	RMSE	RPIQ
NIR	SVM	DET	Mg	Solo+Sedimento	0,99	218,13	8,86
NIR	SVM	SNV	Mg	Solo+Sedimento	0,92	485,58	3,98
NIR	SVM	SGD	Mg	Solo+Sedimento	0,99	167,32	11,55
NIR	SVM	RAW	Mn	Solo+Sedimento	0,68	484,14	2,47
NIR	SVM	DET	Mn	Solo+Sedimento	0,97	143,10	8,37
NIR	SVM	SNV	Mn	Solo+Sedimento	0,86	312,25	3,83
NIR	SVM	SGD	Mn	Solo+Sedimento	0,99	82,13	14,58
NIR	SVM	RAW	Na	Solo+Sedimento	0,77	387,01	0,31
NIR	SVM	DET	Na	Solo+Sedimento	0,91	229,40	0,53
NIR	SVM	SNV	Na	Solo+Sedimento	0,83	325,83	0,37
NIR	SVM	SGD	Na	Solo+Sedimento	0,99	66,37	1,82
NIR	SVM	RAW	Ni	Solo+Sedimento	0,60	9,53	1,66
NIR	SVM	DET	Ni	Solo+Sedimento	0,98	1,93	8,17
NIR	SVM	SNV	Ni	Solo+Sedimento	0,85	6,04	2,61
NIR	SVM	SGD	Ni	Solo+Sedimento	0,99	1,41	11,20
NIR	SVM	RAW	P	Solo+Sedimento	0,39	140,48	1,38
NIR	SVM	DET	P	Solo+Sedimento	0,85	70,63	2,74
NIR	SVM	SNV	P	Solo+Sedimento	0,69	103,53	1,87
NIR	SVM	SGD	P	Solo+Sedimento	0,99	17,34	11,16
NIR	SVM	RAW	Pb	Solo+Sedimento	0,62	5,05	2,08
NIR	SVM	DET	Pb	Solo+Sedimento	0,97	1,47	7,17
NIR	SVM	SNV	Pb	Solo+Sedimento	0,85	3,21	3,28
NIR	SVM	SGD	Pb	Solo+Sedimento	0,99	0,80	13,14
NIR	SVM	RAW	Sr	Solo+Sedimento	0,81	14,25	1,55
NIR	SVM	DET	Sr	Solo+Sedimento	0,97	5,79	3,81
NIR	SVM	SNV	Sr	Solo+Sedimento	0,89	11,37	1,94
NIR	SVM	SGD	Sr	Solo+Sedimento	0,99	2,82	7,81
NIR	SVM	RAW	Ti	Solo+Sedimento	0,31	3078,37	0,76
NIR	SVM	DET	Ti	Solo+Sedimento	0,92	1037,66	2,24
NIR	SVM	SNV	Ti	Solo+Sedimento	0,68	2031,57	1,14
NIR	SVM	SGD	Ti	Solo+Sedimento	0,99	349,26	6,66
NIR	SVM	RAW	V	Solo+Sedimento	0,79	75,25	3,35
NIR	SVM	DET	V	Solo+Sedimento	0,99	15,64	16,10
NIR	SVM	SNV	V	Solo+Sedimento	0,94	42,20	5,97
NIR	SVM	SGD	V	Solo+Sedimento	0,99	15,77	15,96
NIR	SVM	RAW	Zn	Solo+Sedimento	0,47	13,79	0,75
NIR	SVM	DET	Zn	Solo+Sedimento	0,87	6,60	1,56
NIR	SVM	SNV	Zn	Solo+Sedimento	0,66	10,50	0,98
NIR	SVM	SGD	Zn	Solo+Sedimento	0,99	1,69	6,10

10. RESUMO BIOGRÁFICO

Gabriela Naibo, filha de Altair Naibo e Rosani Fatima Dilberti Naibo, nasceu em 22 de junho de 1996, em Ponte Serrada, Santa Catarina. Estudou na Escola Básica Municipal Antonio Paglia e Escola de Educação Básica Dom Vital, onde completou seus estudos de primeiro e segundo graus. Em 2014 ingressou na Universidade do Oeste de Santa Catarina - Unoesc, graduando-se como Engenheira Florestal em 2019. Em abril de 2019 iniciou seus estudos de Mestrado em Ciência do Solo no Programa de Pós-Graduação da Faculdade de Agronomia da Universidade Federal do Rio Grande do Sul. Atualmente, é Representante Titular do Programa de Pós-Graduação em Ciência do Solo da Universidade Federal do Rio Grande do Sul, membra da Sociedade Brasileira de Ciência do Solo (SBCS) e da União Europeia de Geociência (EGU).