UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO

CINTHIA MARIA CORDEIRO ATANÁZIO CRUZ SILVA

PREDIÇÃO DE ELEMENTOS TERRAS RARAS E METAIS PESADOS EM SOLOS DO NORDESTE BRASILEIRO POR ESPECTROSCOPIA DE REFLECTÂNCIA DIFUSA

RECIFE 2023

Cinthia Maria Cordeiro Atanázio Cruz Silva Engenheira Agrônoma

Predição de elementos terras raras e metais pesados em solos do Nordeste Brasileiro por espectroscopia de reflectância difusa

> Tese apresentada ao Programa de Pós-Graduação em Ciência do Solo, da Universidade Federal Rural de Pernambuco, como parte dos requisitos para obtenção do título de Doutora em Ciência do Solo.

> Orientador: Prof. Dr. Clístenes Williams Araújo do Nascimento

Coorientador: Prof. Dr. Ronny Sobreira Barbosa

Recife 2023

Dados Internacionais de Catalogação na Publicação Universidade Federal Rural de Pernambuco Sistema Integrado de Bibliotecas Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

 S586p Silva, Cinthia Maria Cordeiro Atanázio Cruz Predição de elementos terras raras e metais pesados em solos do nordeste brasileiro por espectroscopia de reflectância difusa / Cinthia Maria Cordeiro Atanázio Cruz Silva. - 2023.
95 f. : il.

> Orientador: Clistenes Williams Araujo do Nascimento. Coorientador: Ronny Sobreira Barbosa. Inclui referências e apêndice(s).

Tese (Doutorado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Ciência do Solo, Recife, 2023.

1. Infravermelho próximo. 2. Qualidade ambiental. 3. Elementos químicos. I. Nascimento, Clistenes Williams Araujo do, orient. II. Barbosa, Ronny Sobreira, coorient. III. Título

CDD 631.4

CINTHIA MARIA CORDEIRO ATANÁZIO CRUZ SILVA

Predição de elementos terras raras e metais pesados em solos do Nordeste Brasileiro por espectroscopia de reflectância difusa

Tese apresentada ao Programa de Pós-Graduação em Ciência do Solo, da Universidade Federal Rural de Pernambuco, como parte dos requisitos para obtenção do título de Doutora em Ciência do Solo.

Aprovada em 11 de agosto de 2023

Prof. Dr. Clístenes Williams Araújo do Nascimento Orientador Universidade Federal Rural de Pernambuco

BANCA EXAMINADORA

Prof. Dr. Edivan Rodrigues de Souza Universidade Federal Rural de Pernambuco

> Prof. Dr. Ronny Sobreira Barbosa Universidade Federal do Piauí

Dr. Rennan Cabral Nascimento Universidade Federal do Piauí

Profa. Dra. Maria Eugênia Ortiz Escobar Universidade Federal do Ceará

DEDICO

Com muito amor aos meus pais Onilde e Mário, meu alicerce. Ao meu amigo de todas as horas, meu maior professor e exemplo, meu amado esposo Yuri Jacques. Aos meus pedaços de céu, meus filhos Pedro e Cecília, vocês são meu tudo.

AGRADECIMENTOS

A Deus, por me conceder o dom da vida, pela saúde e por sempre renovar as minhas forças. O Teu amor é o que me conduz.

Aos meus pais, Mário e Onilde, por serem meu alicerce, pelas palavras que me impulsionam, pelos ensinamentos da fé e por nunca medirem esforços aos meus estudos. Vocês são essenciais em minha vida. Obrigada por cada batalha travada, cuja finalidade sempre foi nos manter de pé.

Aos meus avós, Mário e Maria Elizabeth (in memorian), com muito amor e saudade. Sem vocês, a concretização dos meus sonhos não seria possível. Vocês são o início de tudo.

Aos meus irmãos M^a Letícia e Ygor, pelas preocupações divididas, pelos conselhos, pela presteza sem dia e sem hora e pelo amor ofertado a todo momento. Ygor, obrigada por sempre cuidar da nossa família, sendo um exemplo de tamanha integridade. Leti, minha irmã, minha melhor amiga, você é sem dúvida a pessoa que tem o maior coração que possa existir. Obrigada por dividir os momentos desse doutorado do começo ao fim e em todos os momentos da minha vida.

A minha afilhada M^a Larissa, por toda admiração, amor e carinho. Que esta conquista te inspire a acreditar que não existe barreiras para os estudos. Persista nos teus objetivos que madrinha estará sempre ao seu lado. Ao meu sobrinho João, que seu sorriso e alegria continue a me contagiar.

Ao meu amor, meu esposo Yuri Jacques, por seu companheirismo incondicional, pelo seu exemplo de persistência e foco, por nunca soltar a minha mão, por deixar meus dias mais leves, por nunca colocar barreiras na concretização dos meus sonhos, sonhamos juntos. Obrigada por nossos dois milagres da vida, Pedro e Cecília, nossos filhos. A eles, obrigada por me permitir ser a mamãe de vocês, minha melhor versão. Vocês me ensinam os maiores sorrisos e formas de amar. Pedro, foi presente no mestrado. Cecília, presente no doutorado. Que sorte a minha!

Agradeço de forma carinhosa a vovó Zene, quem me estendeu a mão durante esses anos, por sua disponibilidade e amor. Pelo acolhimento além do coração.

Ao meu orientador, Clístenes Nascimento, pela confiança depositada e inspiração profissional. Por desempenhar um papel no meio acadêmico acreditando que a educação é a base de um futuro melhor.

Ao programa de pós-graduação em Ciência do Solo da UFRPE, pela responsabilidade na formação de recursos humanos de qualidade. Meus agradecimentos também aos funcionários terceirizados, técnicos, secretárias e ao excelente corpo docente. Um agradecimento especial ao professor Edivan e a professora Giselle, por agregar a humildade e empatia ao cotidiano profissional.

Ao doutor Olivier Evrard, que possibilitou meu doutorado sanduíche no CEA – França sendo meu supervisor, pela oportunidade de evoluir cientificamente, pela receptividade e convivência com pessoas e local de alto nível.

Ao professor Ronny Barbosa, que me acompanha desde o mestrado pela interação e cooperação que possibilitaram a construção desse trabalho.

Aos amigos em que a pós-graduação me proporcionou, Jaciane, Cintia, Marília, Luiz, Frank, pelas mãos que me foram estendidas, pelos grupos de estudos, pelas boas risadas e a todos os colegas que contribuíram para o meu crescimento pessoal e profissional de alguma forma.

Ao Grupo de Pesquisa de Química Ambiental de Solos e ao Grupo de pesquisa em Conservação do Solo e Qualidade Ambiental, em especial ao Angelo, pelas valiosas contribuições científicas, ao Rennan e Marcos.

À CAPES, pela concessão da bolsa de doutorado e pela oportunidade que me foi concedida a partir da aprovação do Edital CAPES-PRINT-UFRPE no Programa Institucional de Doutorado Sanduíche no Exterior (PDSE), que me permitiu um desenvolvimento pessoal e aperfeiçoamento profissional em um dos maiores desafios da minha vida.

Meu muito obrigada!

Não pense você que ao se tornar mãe uma mulher abandona todas as mulheres que já foi um dia. Bobagem. Ganha mais mulheres em si mesma. Com seus desejos aumentam sua audácia, sua garra, seus poderes.

Se já era impossível, cuidado: ela vira muitas.

Também não me venha imaginar mães como seres delicados e frágeis. Mães são fogo, ninguém segura. Se antes eram incapazes de matar um mosquito, adquirem uma fúria inédita. Montam guarda ao lado de suas crias, dispostas a enfrentar tudo o que zumbir perto delas:

pernilongos, lagartas, leões, gente.

Mães não têm tempo para o ensaio: estreiam a peça no susto.

(livro Mãe, Cris Guerra)

Predição de elementos terras raras e metais pesados em solos do Nordeste Brasileiro por espectroscopia de reflectância difusa

RESUMO

O solo é um sistema aberto formado por constituintes orgânicos e inorgânicos que podem ser avaliados por meio da espectroscopia de reflectância difusa. Dessa forma, o presente estudo teve como objetivo testar diferentes combinações de métodos de préprocessamentos e de calibração, para a predição de elementos químicos (Ba, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, Se, Sb, V, Zn, Sc, La, Ce, Pr, Nd, Sm, Eu, Y, Gd, Tb, Dy, Er, Ho, Yb, Lu), conjuntos de elementos (ETRLs, ETRPs, ETRLs/ETRPs, ETRs), atributos granulométricos, pH, capacidade de troca catiônica e carbono orgânico total usando espectros obtidos na região do infravermelho próximo (NIR) como atributos pedoindicadores, em amostras de solos representativos da diversidade geológica e pedológica da região nordeste do Brasil. As hipóteses desta pesquisa são que (i) a escolha do modelo matemático é mais importante do que a escolha dos pré-processamento para a predição de elementos químicos em solos derivados de diferentes materiais de origem e (ii) desde que sob contexto geológico, pedológico e climático semelhantes, modelos matemáticos de solos na faixa do NIR de outras áreas podem ser utilizados para a predição de elementos químicos em áreas onde existem apenas as assinaturas espectrais. Para esta pesquisa, solos dos estados de PE, PB e RN foram selecionados para representar a diversidade pedológica e geológica do Brasil, aumentando assim a aplicabilidade e alcance dos resultados, dado que as 13 ordens de solo do SiBCS foram coletadas na área de estudo. As medições na faixa do comprimento do NIR (1000-2500 nm) foram realizadas no FTIR/ NIR (Frontier/Perkin Elmer) com transformada de Fourier, onde foi acoplado o acessório de reflectância difusa (NIRA - Near Infrared Reflectance Accessory). Dentre os modelos testados, o Floresta Aleatória (RF) foi o que apresentou melhor desempenho. Além disso, observou-se também que a escolha do modelo foi mais relevante do que a escolha do pré-processamento. A melhor predição dos ETRs em comparação com os metais pesados pode ser observada pelos valores mais altos de desvio residual de previsão (RPD) e razão entre o desempenho e a distância interquartil (RPIQ) e pelos valores mais baixos dos erros associados. Recomenda-se em estudos futuros explorar mais a combinação de modelos com diferentes naturezas computacionais ao invés de testar várias técnicas de pré-processamento com modelos isolados. Os valores preditos para regiões geográficas em que existiam apenas as assinaturas espectrais mostraram dependência espacial moderada, exceto para Pr, Sm e Tb, que demonstraram forte dependência espacial. Isso reforça a qualidade dos mapas preditos para os ETRs, que são essenciais para identificar áreas suscetíveis a impactos ambientais - etapa importante no estabelecimento de políticas ambientais para proteção da saúde humana e do meio ambiente.

Palavras-chave: Infravermelho próximo. Qualidade ambiental. Elementos químicos.

Prediction of rare earth elements and heavy metals in soils from northeastern Brazil using diffuse reflectance spectroscopy

ABSTRACT

The soil is an open system formed by organic and inorganic constituents that can be evaluated by diffuse reflectance spectroscopy. Thus, the present study aimed at testing different combinations of preprocessing and calibration methods for the prediction of chemical elements (Ba, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, Se, Sb, V, Zn, Sc, La, Ce, Pr, Nd, Sm, Eu, Y, Gd, Tb, Dy, Er, Ho, Yb, Lu), group of elements (LREEs, HREEs, LREEs/HREEs, REEs), granulometric attributes, pH, cation exchange capacity and total organic carbon from the spectra obtained in the near-infrared region (NIR) as pedoindicator attributes. The soil samples were select to represent the geological and pedological diversity of the northeast Brazil. We hypothesized that (i) regarding the prediction of chemical elements in soils derived from different parent materials, the choice of the mathematical model is more important than the choice of pre-processes and (ii) under similar geological, pedological and climatic context, mathematical models of soils in the NIR range from other areas can be used for the prediction of chemical elements in areas where only spectral signatures exist. The states of PE, PB, and RN were select for this research because they represent the pedological and geological diversity of Brazil, which increases the applicability and scope of the results. All 13 soil orders in the SiBCS were collected in the study site. Measurements in the NIR length range (1000-2500 nm) were performed in a Fourier transform FTIR/ NIR spectometer (Frontier/Perkin Elmer), coupled with a near Infrared reflectance accessory (NIRA). The Random Forest (RF) model had the best performance. Furthermore, it was also found that the choice of the model was more relevant than the choice of pre-processing. The better prediction of REEs compared to heavy metals can be observed by the higher values of the predicted residual deviation (RPD) and the ratio between the performance and the interquartile distance (RPIQ) and the lower values of the associated errors. Future studies must explore the combination of models with different computational natures rather than testing various preprocessing techniques with single models. The predicted values for geographic regions where only spectral signatures existed showed moderate spatial dependence, except for Pr, Sm, and Tb, which showed strong spatial dependence. This reinforces the quality of the predicted maps for REEs, which are essential for identifying areas susceptible to environmental impacts - an important step in establishing environmental policies for the protection of human health and the environment.

Keywords: Near infrared. Environmental quality. Chemical elements.

LISTA DE FIGURAS

| Figura 1 – Espectro Eletromagnético com destaque para a faixa do NIR |
|---|
| Figura 2 – Etapas da predição de elementos químicos por meio do NIR |
| Figura 3 – Exemplos de comportamentos espectrais de perfis de solos |
| Figura 4 – Exemplo de assinaturas espectrais usando os dados originais (ausência de pré- processamento) |
| Figura 5 – Exemplo de assinaturas espectrais usando o pré-processamento DET 27 |
| Figura 6 – Exemplo de assinaturas espectrais usando o pré-processamento SNV 28 |
| Figura 7 – Exemplo de assinaturas espectrais usando o pré-processamento CR |
| Figura 8 – Exemplo de assinaturas espectrais usando o pré-processamento SG 30 |
| Figura 9 – Exemplo de assinaturas espectrais usando o pré-processamento MSC 31 |
| Figura 10 – Localização da área de estudo e dos pontos de amostragem |
| Figura 11 – Abertura das amostras em sistema fechado e transferência dos extratos para balões volumétricos certificados |
| Figura 12 – FTIR/ NIR Spectrometer (Frontier/PerkinElmer) com transformada de Fourier |
| Figura 13 – Giro da placa de Petri a cada 90° com o objetivo de obter melhor representatividade da superfície examinada |
| Figura 14 – Calibração do espectrômetro a cada 10 amostras de solo analisadas |
| Figura 15 – Etapas para a predição dos elementos químicos e dos atributos físicos e químicos do solo por meio do NIR |
| Figura 16 – Valores de RPIQ para as variáveis estudadas com o modelo PLSR. a) atributos, b) metais e c) ETRs |
| Figura 17 – Valores de RPIQ para as variáveis estudadas com o modelo RF. a) atributos, b) metais e c) ETRs |
| Figura 18 – Mapas dos valores observados (coluna da esquerda – solos do RN) e preditos (coluna da direita – solos da PB) de ETRs por meio do NIR |

LISTA DE TABELAS

| Tabela químicos. | 1 | _ | Potencial | do | NIR | para | a | predição | de | elementos 22 |
|-------------------------|------------|----------------|---------------|---------|----------|----------|--------|--------------|---------|------------------|
| Tabela 2 | - An | álise | descritiva da | s variá | íveis es | tudadas | (n = 1 | 210) | | |
| Tabela 3 - | – Me | lhore | s modelos ol | otidos | para ca | da um do | os atr | ibutos do so | lo anal | isados 45 |
| Tabela 4 - | – Me | lhore | s modelos ol | otidos | para os | ETRs | | | ••••• | 47 |
| Tabela 5 - | – Me | lhore | s modelos ol | otidos | para os | metais a | analis | sados | | |
| Tabela 6 ETRs | | Resu | mo dos res | sultado | os obti | dos atra | avés | do modelo |) PLSI | R para os 53 |
| Tabela 7 observado | – P sep | arâm redite | etros do ser | nivario | ograma | e mode | elos | ajustados a | partir | de valores 54 |

LISTA DE ABREVIATURAS E SIGLAS

Bias – Diferença sistemática entre os valores preditos e medidos

CCC – Coeficiente de correlação de concordância de Lin (*Lin's concordance correlation coefficient*)

CR – Remoção do contínuo (Continuum removal)

DET – Redução de tendência (*Detrend normalization*)

ETRs – Elementos terras raras

ETRLs – Elementos terras raras leves

ETRPs – Elementos terras raras pesados

ICP-OES – Espectroscopia de emissão óptica com plasma acoplado indutivamente (*Inductively coupled plasma optical emission spectrometry*)

MSC – Correção de dispersão multiplicativa (Multiplicative scatter correction)

NIR – Espectroscopia de infravermelho próximo (Near infrared spectroscopy)

PLSR – Regressão de mínimos quadrados parciais (Partial least squares regression)

 \mathbf{R}^2 – Coeficiente de Determinação

RAW – Dados originais (espectros originais)

RF – Floresta aleatória (Random Forest)

RMSE – Raiz do erro quadrático médio (Root mean square error)

RPD – Desvio residual de previsão (Ratio of prediction to deviation)

RPIQ – Razão entre o desempenho e a distância interquartil (*Ratio of performance to interquartile distance*)

SG – Savitzky-Golay

SNV – Variação normal padrão (Standard normal variate)

SVM – Máquina de Vetor de Suporte (Support Vector Machine)

TFSA – Terra fina seca ao ar

SUMÁRIO

| 1.INTRODUÇÃO |
|---|
| 1.1 Hipóteses |
| 1.2 Objetivos |
| 1.2.1 Objetivo Geral |
| 1.2.2 Objetivos Específicos |
| 2.REVISÃO BIBLIOGRÁFICA |
| 2.1. Relevância da área de estudo |
| 2.2. Espectroscopia no infravermelho próximo (NIR) em estudos de solo: histórico e conceito |
| 2.3 Potencial da espectroscopia no infravermelho próximo em solos 20 |
| 2.4. Pré-processamentos e modelagem de dados espectrais |
| 2.4.1. Pré-processamentos |
| 2.4.2. Modelos utilizados |
| 3. MATERIAL E MÉTODOS |
| 3.1 Áreas de estudo |
| 3.2. Amostragem e caracterização física e química das amostras |
| 3.3. Determinação dos elementos químicos e controle de qualidade das análises 35 |
| 3.4. Modelagem espectral |
| 3.5. Análise de dados |
| 3.6. Modelagem quimiométrica |
| 4.RESULTADOS E DISCUSSÃO |
| 4.1 Análise descritiva |
| 4.2 Modelagem |
| 4.3 Predição de ETRs em amostras contendo apenas as assinaturas espectrais |
| 5.CONCLUSÕES |
| REFERÊNCIAS BIBLIOGRÁFICAS |
| ANEXO A – Número de amostras de cada estado utilizada na calibração e validação dos modelos |
| ANEXO B - Modelos, pré-processamentos e parâmetros da predição dos elementos químicos e atributos do solo |
| ANEXO C - Melhores modelos obtidos para: a) Areia com RF-SG; b) Silte com RF-SG; |
| c) Argila com RF-RAW; d) COT com PLSR-CR77 |
| ANEXO D - Melhores modelos obtidos para: a) CTC com PLSR-SNV; b) pH com RF- SG |

| ANEXO E - Melhores modelos obtidos para: a) Ba com RF-SNV; b) Cd com PLSR-CR; c) Co com RF-DET; d) Cr com RF-DET |
|---|
| ANEXO F - Melhores modelos obtidos para: a) Cu com PLSR-CR; b) Fe com RF-DET; c) Mn com RF-DET; d) Ni com RF-DET |
| ANEXO G - Melhores modelos obtidos para: a) Pb com RF-CR; b) Sb com RF-RAW81 |
| ANEXO H - Melhores modelos obtidos para: a) Se com PLSR-RAW; b) V com RF-RAW; c) Zn com PLSR-DET |
| ANEXO I - Melhores modelos obtidos para: a) Ce com RF-MSC; b) Dy com RF-SG; c) Er com RF-SNV; d) Eu com RF-RAW |
| ANEXO J - Melhores modelos obtidos para: a) Gd com RF-DET; b) Ho com RF-RAW; c) La com PLSR-MSC; d) Lu com RF-CR |
| ANEXO K - Melhores modelos obtidos para: a) Nd com RF-RAW; b) Pr com PLSR-SG; c) Sc com PLSR-RAW; d) Sm com RF-RAW |
| ANEXO L - Melhores modelos obtidos para: a) Tb com RF-SG; b) Y com PLSR-RAW; c) Yb com RF-SG |
| ANEXO M - Melhores modelos obtidos para: a) ETRLs com RF-SNV; b) ETRPs com RF-SG; c) ETRLs/ETRPs com PLSR-RAW |
| ANEXO N - Semivariogramas dos teores observados (Rio Grande do Norte) |
| ANEXO O - Validação Cruzada (kriging) dos teores observados (Rio Grande do Norte) |
| ANEXO P - Semivariogramas dos teores preditos (Paraíba) |
| ANEXO Q - Validação Cruzada (kriging) dos teores preditos (Paraíba) |

1 INTRODUÇÃO

O solo é um sistema aberto formado por constituintes orgânicos e inorgânicos que podem ser estudados quantitativamente e/ou qualitativamente por meio da espectroscopia de reflectância. Dentre as faixas do espectro eletromagnético, o infravermelho próximo -NIR (750 a 2500 nm) tem sido mais utilizado pelo baixo custo das análises (exclui-se aqui o custo de aquisição do equipamento), rapidez na obtenção das assinaturas espectrais, e a não geração de resíduos. Essas assinaturas são resultantes da interação entre a energia eletromagnética e a amostra de solo, possibilitando a obtenção de informações de interesse baseado na transmissão, absorção ou reflexão de energia incidente (irradiância). As feições de absorção contidas nessas assinaturas espectrais (vales e picos) representam as ligações e grupos funcionais presentes nas moléculas do solo.

Dessa forma, o presente estudo teve como objetivo geral a predição de elementos químicos Bário (Ba), Cádmio (Cd), Cobalto (Co), Cromo (Cr), Cobre (Cu), Ferro (Fe), Manganês (Mn), Níquel (Ni), Chumbo (Pb), Selênio (Se), Antimônio (Sb), Vanádio (V), Zinco (Zn), Escândio (Sc), Lantânio (La), Cério (Ce), Praseodímio (Pr), Neodímio (Nd), Samário (Sm), Európio (Eu), Ítrio (Y), Gadolínio (Gd), Térbio (Tb), Disprósio (Dy), Érbio (Er), Hólmio (Ho), Itérbio (Yb), Lutécio (Lu), Somatório de elementos terras raras leves (ETRLs), Somatório de elementos terras raras pesados (ETRPs), Relação entre os somatório de elementos terras raras leves e pesados (ETRLs/ETRPs), Somatório de elementos terras raras (ETRs), teores de areia, silte, argila, pH, capacidade de troca catiônica (CTC), carbono orgânico total (COT) do solo usando o infravermelho próximo (NIR). Em comparação com técnicas mais tradicionais, em que a quantidade dos elementos químicos expressa em volume ou massa, o NIR fornece informações acerca das ligações químicas. Foram testados os modelos Regressão de Mínimos Quadrados Parciais (Partial Least Square Regression - PLSR) e Floresta Aleatória (Random Forest - RF), combinados com diferentes técnicas de pré-processamento em amostras de solos representativos, sobretudo, da diversidade geológica e pedológica da região Nordeste do Brasil (Pernambuco, Paraíba e Rio Grande do Norte).

Esses testes foram realizados com os dados originais e pré-processados, adotando os modelos de normalização Savitzky-Golay (SG), correção de dispersão multiplicativa (*Multiplicative Scatter Correction* – MSC), Variação Normal Padrão (*Standard Normal Variate* – SNV), Redução de tendência (*Detrend* – DET) e Remoção de contínuo (*Continuum Removal* – CR). Esses modelos de normalização foram usados para reduzir a relação sinal-ruído e filtrar as características espectrais irrelevantes para os modelos com abordagens distintas. As amostras de solo foram oriundas de áreas sob mínima influência antrópica. A partir desta abordagem inicial, foi possível criar modelos matemáticos para alguns elementos químicos (foco inicial são elementos terras raras – ETRs) e aplicá-los em áreas onde existem apenas as assinaturas espectrais. Trata-se de uma abordagem pioneira que requer cautela na utilização e pode ser o primeiro passo para trabalhos futuros.

1.1 Hipóteses

- A escolha do modelo matemático é mais importante do que a escolha dos préprocessamentos para a predição de elementos químicos em solos derivados de diferentes materiais de origem;
- Desde que sob contexto geológico, pedológico e climático semelhantes, modelos matemáticos de solos na faixa do NIR de outras áreas podem ser utilizados para a predição de elementos químicos em áreas onde existem apenas as assinaturas espectrais. Para tal, as assinaturas espectrais e as respectivas concentrações de elementos terras raras em solos do Rio Grande do Norte foram usadas para a calibração e validação de modelos usando apenas as assinaturas espectrais de amostras do estado da Paraíba. Espera-se gerar uma base de dados capaz de auxiliar no monitoramento de elementos químicos no solo.

1.2 Objetivos

1.2.1 Objetivo Geral

Predizer a concentração de elementos químicos e atributos físicos e químicos de solos representativos da região Nordeste do Brasil por meio da espectroscopia no infravermelho próximo e diferentes modelos matemáticos e técnicas de préprocessamento.

1.2.2 Objetivos Específicos

 Predizer a concentração dos elementos químicos Ba, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, Se, Sb, V, Zn, Sc, La, Ce, Pr, Nd, Sm, Eu, Y, Gd, Tb, Dy, Er, Ho, Yb, Lu, assim como o somatório de ETRLs, ETRPs, ETRs e a relação ETRLs/ETRPs usando os modelos PLSR e RF com os pré-processamentos SG, MSC, SNV, DET e CR, além dos dados originais, em amostras de solos sob diferentes contextos geológicos e pedológicos da região Nordeste, Brasil;

- Predizer os atributos físicos e químicos (areia, silte, argila, pH, CTC e COT) usando os modelos PLSR e RF com os pré-processamentos SG, MSC, SNV, DET e CR, além dos dados originais, em amostras de solos sob diferentes contextos geológicos e pedológicos da região nordeste, Brasil;
- Elaborar mapas da distribuição espacial de ETRs, em áreas onde esses elementos não foram quantificados, por meio de modelos criados em áreas com o contexto climático, geológico e pedológico semelhante.

2 REVISÃO BIBLIOGRÁFICA

2.1 Relevância da área de estudo

Para realização do trabalho foram escolhidos os estados de Pernambuco (PE), Paraíba (PB) e Rio Grande do Norte (RN), pois representam praticamente toda diversidade pedológica e geológica do Brasil, aumentando assim a aplicabilidade e alcance dos resultados. As 13 ordens de solo do Sistemas Brasileiro de Classificação de Solos (SiBCS) ocorrem na área de estudo (Neossolo, Vertissolo, Cambissolo, Chernossolo, Luvissolo, Nitossolo, Argissolo, Latossolo, Gleissolo, Organossolo, Espodossolo, Plintossolo e Planossolo), incluindo também grande parte das classes do segundo nível categórico. Tal diversidade pode estar vinculada aos inúmeros materiais de origens dos solos, tais como: sedimentos aluviais e coluviais, dunas, mangues, calcário, arenito, micaxistos, quartzitos, basaltos, granitos e gnaisses, assim como a precipitação pluvial média variando entre 500 mm e 2400 mm. A área total dos três estados compreende aproximadamente 207.000 km², superando aproximadamente 75% dos países da Europa em termos de tamanho. Diante da heterogeneidade do contexto climático, pedológico e geológico da área de estudo, é possível afirmar que os diversos processos pedogenéticos de formação dos solos estão representados.

Diante da dificuldade e necessidade da realização de estudos mais detalhados sobre a concentração de elementos químicos no solo, tanto do ponto de vista ambiental quanto do manejo do solo, entende-se como essencial a utilização de técnicas complementares na quantificação e predição de nutrientes e/ou contaminantes.

2.2 Espectroscopia no infravermelho próximo (NIR) em estudos de solo: histórico e conceito

A espectroscopia é a ciência que estuda a interação da radiação eletromagnética com a matéria. A lei de Beer-Lambert-Bouguer, que dá suporte a essa técnica, estabelece uma relação entre a absorção de energia de um material, por exemplo, o solo, quando este é atravessado por uma radiação eletromagnética. O espectro infravermelho próximo (NIR, do inglês "near-infrared") é uma faixa de comprimento de onda de luz entre 750-2500 nanômetros. O NIR é uma ferramenta importante em muitas aplicações de ciência, tecnologia e engenharia, porque muitas moléculas orgânicas e inorgânicas têm assinaturas espectrais únicas no NIR (VISCARRA ROSSEL et al., 2006). Isso significa que as moléculas podem ser identificadas e quantificadas com base em seus espectros NIR. A

história do NIR remonta há 50 anos, quando os primeiros espectrômetros comerciais foram introduzidos. Desde então, a tecnologia NIR tem evoluído constantemente, resultando em instrumentos mais precisos, confiáveis e fáceis de usar. Nos últimos anos, o NIR tornou-se cada vez mais importante em áreas como análise de alimentos, farmacêutica, agrícola, petroquímica e de materiais, graças à sua capacidade de análise rápida e não destrutiva (ZORNOZA; GUERRERO, 2008). Historicamente, a espectroscopia de reflectância foi desenvolvida na década de 70 (ZORNOZA; GUERRERO, 2008). Os estudos pioneiros focaram na umidade de grãos, a exemplo de milho e soja, e avançaram para a área da ciência do solo na década de 90 devido ao seu potencial preditivo.

A espectroscopia NIR engloba o espectro eletromagnético entre as regiões do visível e do infravermelho médio, que contém vibrações moleculares e pode ser usado para identificar diferentes tipos de componentes do solo. Esses componentes podem ser identificados porque absorvem comprimentos de onda específicos da luz infravermelha próxima. Essas informações podem então ser usadas para analisar várias características do solo por meio de um modelo estatístico e fornecer predições dessas características.

Dentre as faixas do espectro eletromagnético (**Figura 1**), o NIR (750 a 2500 nm) tem sido mais utilizado pelo menor custo dos equipamentos e maior rapidez na obtenção das assinaturas espectrais do solo. Essas assinaturas são resultantes da interação entre a energia eletromagnética e a amostra de solo, possibilitando a obtenção de informações de interesse baseado na transmissão, absorção ou reflexão de energia incidente (irradiância). As feições de absorção contidas nestas assinaturas espectrais (vales e picos) representam as ligações e grupos funcionais presentes nas moléculas do solo.



Figura 1 - Espectro Eletromagnético com destaque para a faixa do NIR

Fonte: <u>https://www.bruker.com/en/products-and-solutions/infrared-and-raman/ft-nir-spectrometers/what-is-ft-nir-spectroscopy.html</u>

Existem dois processos de interação da radiação eletromagnética com a superfície do solo, as interações macroscópicas e microscópicas. No primeiro caso, as interações dependem das características da superfície do material a ser analisado (i.e., granulometria, forma, tamanho, estrutura, esfericidade e compactação dos grãos), pois relacionam-se com as propriedades ópticas do material em questão (refração, difração e espalhamento); as interações microscópicas, por sua vez, resultam da interação da radiação incidente com os átomos e moléculas do solo em uma camada delgada (10 a 15 micrômetros). Essa interação é responsável pela identificação de feições de absorção específicas relacionadas à composição química e mineralógica do solo. A característica dessa interação talvez seja o aspecto mais importante da espectroscopia e depende diretamente do tipo de estrutura atômica/molecular interna dos constituintes (átomos), da força de ligação entre átomos e moléculas, impurezas iônicas contidas no material (elementos adsorvidos) e tamanho dos raios iônicos (TERRA et al., 2018). Um aspecto importante no espectro eletromagnético na faixa do Vis-NIR é que apresenta bandas largas, provenientes da sobreposição de absorções resultantes, sobretudo, de sobretons (transições de níveis de energia) e combinações de vibrações de ligações químicas O-H, N-H, C-H, C=O e Al-OH. Isso passa a ser um entrave em matrizes mais complexas, como o solo. Maiores informações sobre as dificuldades de interpretação no espectro NIR do solo podem ser obtidas em Viscarra Rossel e Behrens (2010).

Vale ressaltar que o NIR é uma ferramenta complementar, pois não substitui o uso inicial de técnicas convencionais. Em outras palavras, a utilização de dados espectrais não terá êxito se o controle de qualidade das análises de elementos químicos e outros atributos de interesse forem negligenciados, uma vez que os resultados iniciais serão utilizados para calibrar os modelos posteriormente (TERRA et al., 2018).

2.3 Potencial da espectroscopia no infravermelho próximo em solos

Um dos principais potenciais do uso da espectroscopia NIR para análise de solo é que ela pode ser usada em uma ampla variedade de tipos de solo. Também pode ser usado para analisar o solo remotamente e no local, reduzindo a necessidade de extenso trabalho de laboratório. Além disso, a espectroscopia NIR tem o potencial de auxiliar no manejo do solo, permitindo que os agricultores e outros administradores de terras melhorem o rendimento das culturas otimizando as condições do solo. A espectroscopia na faixa do NIR tem sido utilizada nas últimas décadas para a predição de elementos químicos em solos (O'ROURKE et al., 2016; WANG; ZHANG; PAN, 2017; MAIA et al., 2022) (Figura 2). Exemplos do potencial do NIR para a predição de carbono orgânico e elementos químicos podem ser encontrados na Tabela 1.





Fonte: autora.

Dentre as principais vantagens desta técnica, destacam-se a não geração de resíduos, rapidez, baixo custo em relação às análises convencionais e emprego em larga escala e com maior densidade amostral.

| Variável | Faixa espectral | N° | Pré-processamentos | Modelo | \mathbb{R}^2 | RMSE | RPD | RPIQ | Referências |
|----------|--------------------|----------|------------------------------|--------|----------------|---------|------|-------|------------------------------|
| | (nm) | amostras | | | | | | | |
| Cu | 350-2500 | 93 | Dados originais | PLSR | 0,92 | | 1,54 | | CHENG et al. (2019) |
| Cr | 350-2500 | 93 | Dados originais | PLSR | 0,99 | | 1,01 | | CHENG et al. (2019) |
| Cd | 350-2500 | 93 | Dados originais | PLSR | 0,30 | | 0,98 | | CHENG et al. (2019) |
| CO | 350-2500 | 2471 | Suavização | PLSR | 0,67 | 0,95 | | | ROSIN et al. (2021) |
| CO | 350-2500 | 2471 | SNV | PLSR | 0,70 | 0,91 | | | ROSIN et al. (2021) |
| CO | 350-2500 | 2471 | SG | PLSR | 0,72 | 0,87 | | | ROSIN et al. (2021) |
| СО | 350-1100 | 841 | Suavização | PLSR | 0,72 | 0,56 | | | MOURA-BUENO et al. (2019) |
| CO | 350-1100 | 841 | SG | PLSR | 0,74 | 0,52 | | | MOURA-BUENO et al. (2019) |
| CO | 350-1100 | 841 | SNV | PLSR | 0,71 | 0,59 | | | MOURA-BUENO et al. (2019) |
| CO | 350-1100 | 841 | DT | PLSR | 0,69 | 0,6 | | | (2019) |
| Pb | 350-2500 | 120 | Suavização, MSC, SNV e RC | PLSR | 0,93 | 63,01 | | | KHOSRAVI et al. (2018) |
| Al | 1000-2500 | 145 | SNV | PL SR | 0.76 | 11836 | 2 12 | 2 18 | MAIA et al. (2022) |
| Ce | 1000-2500 | 145 | DET | PLSR | 0.48 | 33.63 | 1 37 | 1 48 | MAIA et al. (2022) |
| Cr | 1000-2500 | 145 | CR | RF | 0.16 | 18.12 | 1 10 | 1 15 | MAIA et al. (2022) |
| Fe | 1000-2500 | 145 | DET | RF | 0.5 | 8743 | 1.36 | 1.81 | MAIA et al. (2022) |
| La | 1000-2500 | 145 | SG | RF | 0.6 | 14.06 | 1.46 | 1.32 | MAIA et al. (2022) |
| Mn | 1000-2500 | 145 | DET | RF | 0.53 | 136.58 | 1.45 | 1.82 | MAIA et al. (2022) |
| Pr | 1000-2500 | 145 | MSC | PLSR | 0,6 | 4,35 | 1,58 | 1,2 | MAIA et al. (2022) |
| Ni | 1000-2500 | 145 | Dados originais | PLSR | 0,36 | 5,19 | 1,25 | 1,19 | MAIA et al. (2022) |
| Th | 1000-2500 | 145 | MSC | PLSR | 0,72 | 9,84 | 1,89 | 1,69 | MAIA et al. (2022) |
| Cr | 1000-2500 | 316 | SG | SVM | 0,99 | 2,16 | | 6,82 | NAIBO et al. (2022) |
| Cu | 1000-2500 | 316 | SG | SVM | 0,99 | 12,48 | | 18,88 | NAIBO et al. (2022) |
| Fe | 1000-2500 | 316 | SG | SVM | 0,99 | 2900,43 | | 18,87 | NAIBO et al. (2022) |
| La | 1000-2500 | 316 | SG | SVM | 0,99 | 1,44 | | 13,65 | NAIBO et al. (2022) |
| Pb | 1000-2500 | 316 | SG | SVM | 0,99 | 0,83 | | 13,21 | NAIBO et al. (2022) |
| Zn | 1000-2500 | 316 | SG | SVM | 0,99 | 0,48 | | 12,72 | NAIBO et al. (2022) |
| Ba | 1000-2500 | 316 | SG | SVM | 0,99 | 7,43 | | 11,62 | NAIBO et al. (2022) |
| Zn | 350-2500 | 93 | 1st Derivada | PLSR | 0.93 | | 3.8 | | CHENG et al. (2019) |

Tabela 1 - Potencial do NIR para a predição de elementos químicos

RF Floresta Aleatória, PLSR Regressão de Mínimos Quadrados Parciais, RAW Dados Originais, MSC Correção de Dispersão Multiplicativa, SG Savitzky-Golay, SNV Variação Normal Padrão, DET Redução de tendência, CR Remoção de contínuo, R²Coeficiente de Determinação, RMSE Raiz do Erro Quadrático Médio, RPD Desvio Residual de Previsão, RPIQ Razão entre o desempenho e a distância interquartil.

Apesar da dificuldade de predição de elementos químicos com baixas concentrações no solo usando o infravermelho e da ausência de feições de absorção específicas para os elementos químicos, a relação entre as características espectrais e os

constituintes primários do solo (argila, matéria orgânica, óxidos de ferro) podem auxiliar na calibração e validação de modelos preditivos (STENBERG et al., 2010; WU et al., 2007; HORTA et al., 2015). Tais modelagens são possíveis devido a relações de primeira ou segunda ordem destes atributos com os espectros de solos.

A predição de elementos potencialmente tóxicos por meio de espectroscopia de reflectância difusa foi testada por Camargo et al. (2018) ao comparar diferentes superfícies geomórficas de Latossolos do Planalto Ocidental Paulista. Esta abordagem forneceu modelos bem calibrados de predição para Ba, Co, Cu, Mn e Ni. Cao et al. (2020) utilizaram um novo índice espectral baseado em espectro na faixa do Vis e NIR para melhorar o mapeamento de metais pesados no solo. Esse índice foi usado como covariável em cokrigagem ordinária, fornecendo maior acurácia em comparação com outros métodos comparativos. Esses autores forneceram uma contribuição importante para a predição de elementos químicos em baixas concentrações no solo. Maia et al. (2022) testaram o NIR para a predição de elementos potencialmente tóxicos em solos e sedimentos da bacia hidrográfica do rio Ipojuca que abrange o semiárido e litoral de Pernambuco. Foram utilizados os modelos PLSR e RF combinados com diferentes préprocessamentos. De forma geral, o modelo RF demonstrou maior acurácia e elementos com maiores concentrações no solo, a exemplo do Al e Ti, apresentaram as melhores predições. Estudando elementos terras raras em solos do maior depósito de fosfato de urânio do Brasil, Maia et al. (2020) também obtiveram as melhores predições no NIR para os elementos com maiores concentrações. Os autores destacaram que a sobreposição das feições de absorção, as baixas concentrações dos elementos químicos e as interferências espectrais recorrentes na determinação de ETRs por ICP-OES podem influenciar na acurácia da predição, como também destacado por Balaram (2019).

A predição de elementos químicos por espectroscopia na faixa do NIR também foi proposta por Naibo et al. (2022), ao estudarem solos e sedimentos de uma bacia hidrográfica rural. Foram investigados 22 elementos em 316 amostras de solos e 196 amostras de sedimentos. O modelo Máquina de Vetores de Suporte (SVM) combinado com o pré-processamento SG forneceu os melhores resultados em amostras de solo e sedimentos. Em outras palavras, os maiores valores de R² e RPIQ, assim como os menores valores para a raiz do erro quadrático médio (RMSE), foram oriundos da melhor combinação entre o modelo multivariado e pré-processamento. Khosravi et al. (2018) e Tiecher et al. (2021) reforçam que a performance do SVM é melhorada por meio da escolha mais adequada do pré-processamento. Vale ressaltar que em alguns casos os

dados originais podem fornecer os melhores resultados, a exemplo do que foi observado por Mendes et al. (2021) em amostras de solo para a predição de carbono e pH usando aprendizado de máquina.

Além do uso de dados de reflectância para quantificações de elementos químicos, publicações recentes têm demonstrado o potencial da abordagem na discriminação e classificação de perfis de solos (**Figura 3**) (TERRA et al., 2018). Neste caso, os espectros dos horizontes de cada perfil de solo são modelados conjuntamente com o auxílio de algoritmos de clusterização, como distâncias taxonômicas (métricas pedológicas), e/ou ferramentas estatísticas multivariadas, tais como análise por componentes principais ou canônica. Isto é possível, pois a pedogênese produz feições de absorção específicas nas assinaturas espectrais.

Figura 3 - Exemplos de comportamentos espectrais de perfis de solos. A) influência do conteúdo de areia na intensidade de reflectância; B) conteúdos similares de argila e mineralogia em profundidade



Fonte: Adaptado de Terra et al. (2018)

Além dos esforços voltados para o uso do NIR na predição de elementos químicos em solos, outras pesquisas focaram na predição de areia, silte e argila (MENDES et al., 2021), metais em fertilizantes (GUINDO et al., 2021) e salinidade do solo (LIU et al., 2015). Apesar de seu potencial, existem algumas limitações no uso da espectroscopia NIR na análise de solo. Isso inclui o domínio computacional para gerir e aplicar os modelos de predição, o que pode ser uma barreira para a consolidação de suas aplicações em diferentes segmentos da sociedade, custo de aquisição do equipamento, variações nas condições do solo, bem como outros fatores, como homogeneidade da amostra, umidade e temperatura do solo, que podem afetar a precisão dos resultados. Evidentemente que os dois últimos (i.e., umidade e temperatura) são entraves em medições direto no campo, uma vez que esses fatores são mais controlados em condição de laboratório.

2.4 Pré-processamentos e modelagem de dados espectrais

2.4.1 Pré-processamentos

Para interpretar adequadamente as assinaturas espectrais é importante identificar e realçar, na maioria dos casos, as principais feições de absorção (interações microscópicas relacionadas sobretudo ao tipo de estrutura atômica/molecular interna dos constituintes átomos) e variações nas intensidades de reflectância (interações macroscópicas relacionadas às características físicas e texturais do solo). Para tal, diversos préprocessamentos têm sido utilizados, pois influenciam diretamente na precisão e confiabilidade dos modelos. A escolha dos pré-processamentos depende da função destes e da natureza dos dados. Entretanto, ainda não existe consenso acerca dos melhores préprocessamentos e modelos multivariados. O conhecimento do solo pode ajudar consideravelmente no entendimento do espectro e vice-versa. A seguir será abordado de forma sucinta os pré-processamentos e modelos usados neste estudo. Vale ressaltar que resultados satisfatórios podem ser obtidos com os dados originais (Figura 4), descartando a necessidade do uso individual ou combinado de pré-processamentos (MAIA et al., 2020). Além disso, quando o equipamento não está devidamente calibrado e precisa de manutenção, costumam ocorrer muitos ruídos que se confundem com pequenas feições de absorção, principalmente no início e final dos espectros.



Figura 4 - Exemplo de assinaturas espectrais usando os dados originais (ausência de préprocessamento)

O pré-processamento DET remove ruídos dos dados espectrais por meio da normalização do espectro pela aplicação da variação normal padrão seguida de regressão polinomial de segundo grau (**Figura 5**). Esse pré-processamento realça as variações de maior frequência em detrimento das de baixa frequência que são removidas (BARNES; DHANOA; LISTER, 1989).



Figura 5 - Exemplo de assinaturas espectrais usando o pré-processamento DET

A SNV é uma importante técnica de pré-processamento na análise de dados de espectroscopia NIR, sendo usada para normalizar os espectros removendo os efeitos da variação da linha de base, que pode ser causada por fatores como diferenças na morfologia da amostra, tamanho da partícula, dispersão e absorção. Essa técnica transforma os espectros em novos valores que têm uma média de zero e um desvio padrão de um. A SNV também pode ajudar a reduzir os efeitos do ruído sistêmico, como desvios e mudanças de instrumentos (VISCARRA ROSSEL; BEHRENS, 2010) (**Figura 6**).

Ainda, a SNV é utilizada para remover as variações indesejadas no espectro decorrentes de ruídos e diferentes tamanhos de partículas por meio da subtração da média e divisão pelo desvio padrão da reflectância de cada espectro. A técnica de normalização SNV é útil na análise de dados espectrais, pois pode melhorar a exatidão e a precisão de modelos de análise multivariada, como a PLSR, reduzindo a variação devido ao ruído instrumental e aspectos físicos da amostra. A SNV também pode melhorar o desempenho dos algoritmos de aprendizado de máquina usados na análise de dados do NIR. No geral, a SNV aumenta a comparabilidade dos espectros, reduz as interferências causadas por variações indesejadas e aumenta a precisão da análise NIR.



Figura 6 - Exemplo de assinaturas espectrais usando o pré-processamento SNV

A remoção do espectro contínuo (CR) tem a finalidade de remover as feições contínuas do espectro e, por conseguinte, analisar e realçar os picos de absorção a partir de funções matemáticas (**Figura 7**). O espectro contínuo é identificado pela junção/ligação de pontos máximos do espectro; já a remoção deste ocorre pela divisão dos valores originais de reflectância pelos valores correspondentes da linha contínua (CLARK; ROUSH, 1984).



Figura 7 - Exemplo de assinaturas espectrais usando o pré-processamento CR

O pré-processamento SG é frequentemente aplicado a dados espectrais. Esse método matemático de suavização, baseado em uma regressão polinomial, consiste em definir uma janela de comprimento "n" e movê-la um ponto por vez ao longo de todo o sinal analítico (**Figura 8**). Essas janelas são identificadas por meio de testes preliminares. Em outras palavras, calcula-se a primeira derivada de um polinômio de primeira, permitindo entender a taxa de mudança da reflectância em relação ao comprimento de onda. Isso permite a remoção das linhas de base e realce das feições de absorção de interesse no espectro eletromagnético (TERRA et al., 2018).



Figura 8 - Exemplo de assinaturas espectrais usando o pré-processamento SG

A MSC é uma técnica usada em quimiometria e espectroscopia que consiste em corrigir o efeito de espalhamento, no qual a luz é esparsada em todas as direções (**Figura 9**). Essa dispersão ocorre quando a luz é desviada por partículas ou outros materiais na amostra, podendo levar à distorção nos dados espectrais, dificultando a identificação e quantificação mais exata dos componentes da amostra (KHOSRAVI et al., 2018).

A MSC realiza cálculos de um espectro de referência, no qual se refere a um espectro de amostra sob condições em que ocorre o mínimo de dispersão. Ao selecionar um espectro de referência, deve-se ter cuidado, pois ele deve ser representativo da amostra que está sendo medida. Esse espectro é utilizado para corrigir os espectros das demais amostras. O resultado da divisão do espectro medido pelo espectro de referência é multiplicado por um fator de correção, que é calculado usando um método de regressão. Dessa forma, garante-se que o espectro reparado corresponda ao espectro de referência.



Figura 9 - Exemplo de assinaturas espectrais usando o pré-processamento MSC

A técnica MSC é copiosamente utilizado em diversos campos por ser uma técnica bastante eficaz para corrigir a dispersão, aumentando a precisão e confiabilidade; sendo importante considerar que a MSC pode não ser eficaz nos casos em que o espalhamento não ocorre por fatores externos e sim pelas propriedades físicas.

2.4.2 Modelos utilizados

Para a realização desta pesquisa, dois métodos com diferentes abordagens, PLSR e RF foram selecionados para calibrar os modelos espectroscópicos. A PLSR talvez seja um dos métodos mais utilizados em função da interface mais fácil em termos da interpretação da relação entre as assinaturas espectrais e atributos do solo. A vantagem da PLSR é trabalhar com conjunto de dados com muitas variáveis independentes e altamente correlacionadas, ou seja, insensível à colinearidade que é algo muito recorrente em dados espectrais na faixa do infravermelho (REEVES; MCCARTY; MEISINGER, 1999; VISCARRA ROSSEL et al., 2006). Entretanto, métodos não lineares, como o modelo RF, têm se destacado por apresentarem melhores resultados em comparação com métodos lineares na predição de elementos químicos (TAN et al., 2019; MAIA et al., 2022).

A PLSR e o RF são métodos de regressão que possuem características distintas. A PLSR busca encontrar um conjunto de variáveis latentes que explicam a variância máxima nas variáveis preditoras e de resposta, permitindo criar um modelo linear para prever a variável de resposta. Já o RF utiliza árvores de decisão combinadas para gerar uma previsão final, sem assumir uma relação linear entre as variáveis (MAIA et al., 2022).

A PLSR é indicada quando há muitas variáveis preditoras com alta colinearidade, além de lidar bem com valores ausentes e apresentar menor risco de sobreajuste. No entanto, ele assume uma relação linear entre as variáveis e pode não ter bom desempenho em problemas não lineares. Já o RF funciona bem em situações não lineares, lidando com interações entre as variáveis e sendo robusto a outliers e valores ausentes. Contudo, ele pode apresentar sobreajuste e ser computacionalmente caro em grandes conjuntos de dados. A PLSR é comumente utilizada em quimiometria, bioinformática e ciências sociais para prever resultados em dados de alta dimensão, além de ser útil na seleção de características importantes para o modelo. O RF, por sua vez, é amplamente aplicado em diversas áreas, como ecologia, finanças e marketing, podendo lidar com relacionamentos complexos entre as variáveis e oferecer previsões precisas.

A escolha entre os dois métodos depende da natureza dos dados, dos recursos computacionais disponíveis e da pergunta de pesquisa em questão. A PLSR é mais adequada para problemas lineares com alta colinearidade, enquanto o RF é mais apropriado para problemas não lineares com interações complexas entre as variáveis. Vale destacar que o ganho de popularidade do RF decorre da alta acurácia, precisão e versatilidade em grandes conjuntos de dados, além da robustez em planilhas com dados faltantes, algo recorrente em estudos de cunho ambiental.

A avaliação da performance dos modelos pode ser obtida pela interpretação dos parâmetros R^2 , R^2 ajustado, diferença sistemática entre os valores preditos e medidos (bias), coeficiente de correlação de concordância de Lin (CCC), RMSE, desvio residual de predição (RPD) e razão entre o desempenho e a distância interquartil (RPIQ). Esses parâmetros devem ser avaliados de forma conjunta. O RPD e RPIQ permitem a comparação da performance dos modelos derivados de diferentes áreas. O RPIQ tem sido preferível atualmente como métrica de validade do modelo, em comparação com o RPD, uma vez que não assume que os dados apresentam distribuição normal (BELLON-MAUREL et al., 2010). Ainda vale ressaltar, que o RPIQ leva em consideração tanto o erro de predição e a variação de valores medidos (JOHNSON et al., 2019). De forma geral, quanto maior os valores de R^2 , R^2 ajustado, CCC, RPD e RPIQ e menores os valores

de bias e RMSE, melhor a capacidade preditiva dos modelos. As faixas de interpretação destes parâmetros diferem um pouco na literatura não sendo objetivo dessa revisão de literatura discutir essas peculiaridades. Mais informações nesse sentido serão fornecidas no material e métodos.

3 MATERIAL E MÉTODOS

3.1 Áreas de estudo

A área de coleta das amostras de solo compreende os estados do Rio Grande do Norte (RN), Pernambuco (PE) e Paraíba (PB) (Figura 10). As amostras das respectivas áreas de estudos foram oriundas de trabalhos de tese (COSTA, 2013; BIONDI, 2010; ALMEIDA JÚNIOR, 2014). O RN apresenta área de 52.796,79 km², precipitação pluvial média anual variando entre 500 mm para a região semiárida e 1500 mm para a faixa mais úmida localizada na porção leste do estado. As temperaturas médias anuais do ar ocorrem numa estreita faixa de 26 a 27 °C. O RN apresenta uma vasta variedade de formas de relevo, formadas em sedimentos da Bacia Potiguar e terrenos mais antigos do embasamento cristalino. O RN pode ser divido em dois grandes ambientes climáticos. A porção semiárida abrange grande parte do território do estado, apresentando precipitação pluvial média anual entre 500 mm e 750 mm. A faixa úmida, segundo ambiente mais dominante com precipitação média anual variando entre 750 mm e 1500 mm, encontrase na parte leste do estado. As regiões sub e semiúmida apresentam precipitação média anual variando de 800 a 1200 mm e 600 a 800 mm, respectivamente (SILVA et al., 2018). O estado tem grande variedade de solos, com destaque para os Luvissolos, Latossolos, Chernossolos, Planossolos, Gleissolos, Neossolos, Plintossolos e Espodossolos derivados de diversos materiais de origem (sedimentos aluviais e coluviais, dunas, mangues, calcário, arenito, micaxistos, quartzitos, basaltos, granitos e gnaisses).



Figura 10 - Localização da área de estudo e dos pontos de amostragem

A PB apresenta uma área de 56.438 km², distribuídos em 223 municípios de quatro mesorregiões (Mata Paraibana, Agreste Paraibano, Borborema e Sertão Paraibano). A precipitação pluvial anual média varia entre 400 e 800 mm na porção semiárida do estado (leste e o oeste paraibano); já na parte litorânea, os totais anuais podem superar os 1500 mm (ALMEIDA JÚNIOR et al., 2016). Destaca-se que a parte semiárida do estado é coberta pela vegetação caatinga e pode chegar a ter 11 meses de seca. Aproximadamente 80% do estado, exceto a faixa sedimentar litorânea, é constituída predominantemente por gnaisses e migmatitos (rochas do pré-cambriano: 3,45 até 542 milhões de anos), ocorrendo sempre associado à micaxisto e granitos. Os 20% remanescentes são formados por rochas do período Cretáceo (145 a 65 milhões de anos) e por coberturas sedimentares cenozoicas (65 milhões de anos até o recente) (CPRM, 2002). Os solos são predominantemente Neossolos, Luvissolos, Argissolos, Vertissolo, Espodossolo Latossolo e Planossolos (BRASIL, 1972; CAMPOS; QUEIROZ, 2006).

O estado de PE compreende uma área de 98.312 km² e estende-se desde a zona costeira até a zona semiárida. A precipitação pluvial média anual varia de 500 a 2.200 mm e a temperatura do ar varia de 24 °C a 28 °C para regiões semiáridas e costeiras, respectivamente. Na zona costeira, os solos desenvolvem-se frequentemente sobre sedimentos terciários ou rochas cristalinas, normalmente profundos e altamente intemperizados com relevo variando de ligeiramente a fortemente ondulado. Dentre as principais ordens de solos, destacam-se os Latossolos, Argissolos e Gleissolos (BIONDI,

2010). A vegetação é composta principalmente de floresta tropical perene de folhas largas. Na região semiárida, os solos são geralmente rasos devido à longa estação seca, que por sua vez diminui o intemperismo químico e o processo de lixiviação. Destacamse os Neossolos, Planossolos, Luvissolos e Cambissolos. A vegetação é predominantemente floresta caducifólia, conhecida localmente como caatinga hipóxerófila, e vegetação arbórea-arbustiva caducifólia densa, classificada localmente como caatinga hiperxerófila, algo semelhante a um cerrado espinheiro decidual da estação seca (NASCIMENTO et al., 2006).

3.2 Amostragem e caracterização física e química das amostras

Foram coletadas 210 amostras compostas de solo na profundidade de 0-0,20 m em áreas sob mínima influência antrópica. Cada amostra composta foi formada por 10 amostras simples. A granulometria foi determinada segundo Gee e Or (2002). A caracterização química foi realizada de acordo com metodologias sugeridas por Donagema et al. (2011): pH foi determinado em água (1:2,5); Os cátions Ca^{2+} , Mg^{2+} e Al^{3+} foram extraídos com KCl 1,0 mol L⁻¹ e dosados por titulometria. Potássio e Na⁺ foram extraídos com Mehlich-1 e dosados por fotometria de chama. A acidez potencial (H+Al) foi extraída com acetato de cálcio (0,5 mol L⁻¹) e dosada por titulometria. O carbono orgânico total (COT) pelo método de Walkley-Black modificado (SILVA et al., 1999). A partir dos resultados obtidos do complexo sortivo foram calculados os valores de soma de bases (SB) e capacidade de troca de cátions total (CTC). Os solos exibiram ampla variabilidade quanto às propriedades físicas e químicas. De forma geral, são levemente ácidos, com baixo teor de carbono orgânico (<1%) e, por conseguinte, apresentam baixa CTC. A concentração de areia e argila variou de 80 a 970 e 20 a 640 g kg⁻¹, respectivamente.

3.3 Determinação dos elementos químicos e controle de qualidade das análises

A extração dos elementos químicos (Ba, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, Se, Sb, V, Zn, Sc, La, Ce, Pr, Nd, Sm, Eu, Y, Gd, Tb, Dy, Er, Ho, Yb e Lu) foi realizada de acordo com o método proposto pela Agência de Proteção Ambiental dos Estados Unidos (USEPA, 1998), em sistema fechado (forno micro-ondas) (**Figura 11**), sendo colocado em tubo de teflon 1 g de solo com 9 mL de HNO₃ e 3 mL de HCl (ácidos de alta pureza - Merck PA). Os extratos foram vertidos em balões certificados de 25 mL (NBR ISO/IEC)
(**Figura 11**), com o volume dos balões completado com água ultra pura (Millipore Direct-Q System) e filtrados através de papel filtro lento (Macherey Nagel®). Todas as análises foram realizadas em duplicata.

Figura 11 - Abertura das amostras em sistema fechado e transferência dos extratos para balões volumétricos certificados



Amostras em branco e de certificação internacional - SRM 2709 San Joaquin Soil (NIST, 2002) - foram usadas para o controle de qualidade das análises. As concentrações dos elementos químicos foram determinadas por Espectroscopia de emissão óptica com plasma acoplado indutivamente (ICP-OES/Optima DV7000, Perkin Elmer). Em função da maior dificuldade, sobretudo para a leitura dos elementos terras raras, uma câmara ciclônica foi utilizada para aumentar a sensibilidade das leituras. As taxas de recuperação variaram entre 75 e 105%, demonstrando a confiabilidade dos dados.

3.4 Modelagem espectral

As medições na faixa do comprimento do NIR foram realizadas no FTIR/ NIR Spectrometer (Frontier/PerkinElmer) com transformada de Fourier, onde foi acoplado o acessório de reflectância difusa (NIRA – Near Infrared Reflectance Accessory) (**Figura 12**). A finalidade da transformada de Fourier é a retirada de ruídos de alta frequência.

Figura 12 - FTIR/ NIR Spectrometer (Frontier/PerkinElmer) com transformada de Fourier

As amostras de solo (terra fina seca ao ar – TFSA) foram adicionadas a placas de Petri com 6 cm de diâmetro e altura de 1 cm. O solo posicionado na placa de Petri foi encaminhado para secagem em estufa por 48h a 50 °C (TERRA; DEMATTÊ; ROSSEL, 2015). Os espectros foram adquiridos na faixa de 1000 – 2500 nm, resolução 2 nm e 32 varreduras (acumulações) por amostra de solo e intervalo de amostragem de 0,5 nm. Seguindo a metodologia de Terra et al. (2015) foram obtidas quatro medidas espectrais para a mesma amostra, girando a placa de Petri a cada 90° com o objetivo de obter melhor representatividade da superfície examinada (**Figura 13**). O spectralon branco com 100% de refletância (LabSphere, North Sutton, NH, EUA: L124-1634) foi utilizado para calibrar o espectrômetro a cada 10 amostras de solo analisadas (**Figura 14**). **Figura 13 -** Giro da placa de Petri a cada 90° com o objetivo de obter melhor representatividade da superfície examinada



Figura 14. Calibração do espectrômetro a cada 10 amostras de solo analisadas



O conjunto de dados espectral foi composto de observações na faixa espectral de 1000 – 2500 nm com uma janela de 0,5 nm. Posteriormente, esse conjunto de dados foi reamostrado para janela de 1 nm utilizando a função resample do pacote do R 'prospectr' (STEVENS; RAMIREZ-LOPEZ, 2020). A amostragem de hipercubo latino condicionada (cLHS - *conditioned Latin Hypercube Sampling*) foi utilizada para dividir o conjunto de dados em 80% para calibração do modelo e 20% para validação. O cLHS permite cobrir a variabilidade ótima das variáveis ambientais (MINASNY; MCBRATNEY, 2006), que nesse caso são as faixas espectrais. O pacote do R 'clhs' foi utilizado para realizar o cLHS.

Os dados de calibração e validação foram pré-processados usando métodos disponíveis no pacote "prospectr" do R: Savitzky-Golay 1ª derivada utilizando polinomial de 1ª ordem e janela de busca de 9-nm – SG (SAVITZKY; GOLAY, 1964), Multiplicative Scatter Correction (DHANOA et al., 1994), SNV (BARNES; DHANOA; LISTER, 1989), DET (BARNES; DHANOA; LISTER, 1989), CR (CLARK; ROUSH, 1984) e *MSC*. Os métodos de pré-processamento foram usados para reduzir a relação sinal-ruído e filtrar as características espectrais irrelevantes para a modelos com abordagens distintas (**Figura 15**).





Fonte: Autora

O algoritmo RF (BREIMAN, 2001) e PLSR foram selecionados para construir os modelos. O Random Forest é um método algorítmico de aprendizado por máquina que combina várias árvores de decisão para alcançar a classificação, que podem aumentar a precisão da previsão em comparação com modelos lineares, tais como a regressão PLSR

(KNOX et al., 2015; SANTANA; SOUZA; POPPI, 2018). Ambos modelos foram aplicados usando o pacote "caret" em R (KUHN, 2008) ao conjunto de dados de calibração com validação cruzada de 10 vezes por busca na grade. A performance dos modelos foi avaliada através da RMSE, coeficiente de determinação ajustado (R²adj), bias, RPIQ, CCC, para o conjunto de validação. A RMSE mede a precisão geral da predição do modelo, quanto mais próximo de 0, melhor a predição. O R² adj varia de 0 a 1, quanto mais próximo de 1 melhor o desempenho do modelo (COBLINSKI et al., 2020).

O bias é avaliado através da estimativa do erro das previsões (BELLON-MAUREL et al., 2010). O RPIQ foi utilizado para avaliar a qualidade dos modelos, baseando-se na distância interquartil para explicar a disseminação dos dados, conferindo maior acurácia das predições. Quanto maior o RPIQ melhor é o modelo (COBLINSKI et al., 2020), considerando-se como modelos excelentes (RPIQ > 2,5), modelos muito bons (2,5 > RPIQ > 2,0), bom modelo (2,0 > RPIQ > 1,7), regular (RPIQ 1,7 > RPIQ > 1,4) e modelo muito ruim (RPIQ < 1,4) (DOTTO et al., 2017). O CCC é um coeficiente de concordância proposto por Lin (1989) que verifica se as variáveis independentes de um modelo estão correlacionadas e concordantes entre si, ou seja, avalia a acurácia e precisão do modelo.

3.5 Análise de dados

Os dados foram analisados por meio de análise descritiva (mínimo, máximo, média, desvio padrão e correlação de Spearman). A estatística espacial foi usada para a predição dos ETRs em áreas onde existiam apenas as assinaturas espectrais. Foram usados modelos calibrados e validados em áreas com contexto climático, geológico e pedológico semelhante. Na análise da distribuição espacial dos atributos, a krigagem ordinária foi utilizada. Todas as variáveis apresentaram distribuição não normal para o teste de Kolmogorov–Smirnov (p < 0.05). Os semivariogramas isotrópicos e anisotrópicos foram avaliados. Os padrões de variabilidade espacial foram mantidos após a aplicação de diferentes direções nos semivariogramas (0, 45, 90 e 135°), por consequência, essa modelagem geoestatística foi realizada com base nos parâmetros dos semivariogramas isotrópicos.

Os padrões espaciais foram avaliados por meio dos parâmetros dos semivariogramas e da validação cruzada. Os modelos foram selecionados com base no grau de dependência espacial (*degree of spatial dependence -* DSD) e raiz quadrada média

padronizada próximo a 1. O DSD foi calculado seguindo Cambardella et al. (1994), seguindo a fórmula $C0/(C0+C1) \times 100$, onde C0 é o efeito pepita e C1 a contribuição. Um efeito pepita menor ou igual a 25% do patamar foi considerado forte, moderado entre 25 e 75% e fraco quando maior que 75%.

Os parâmetros da validação cruzada foram referentes a relação dos valores mensurados e estimados pelas diferentes abordagens geoestatísticas. As variáveis com fraca dependência espacial (DSD > 75%) ou raiz quadrada média padronizada com valores distantes de 1 não foram considerados. A análise geoestatística foi realizada com o GS+TM v.7 Build 17 software (Gamma Design Software, LLC, Plainwell, MI, USA) em conjunto com o Surfer v. 9.11.947 Copyright© 2023–2010 (Golden Software, LLC, Golden, CO, USA).

3.6 Modelagem quimiométrica

Para gerar os mapas de distribuição espacial dos ETRs, os dados foram separados em dois bancos de dados, 91 amostras do estado do Rio Grande do Norte foram selecionadas para calibração e validação cruzada do modelo e outras 90 amostras do estado da Paraíba foram utilizadas para predição (apenas assinaturas espectrais). Essa é a segunda modelagem usada no trabalho, uma vez que a primeira usou dados de solos de PE, PB e RN. Para calibrar os modelos quimiométricos, foi utilizado o modelo PLSR e o software ParleS[®] (VISCARRA ROSSEL, 2008). Os conjuntos de calibração foram utilizados para calibrar e validar a precisão da predição de cada modelo PLSR (YU et al., 2023). A SNV e suavização Savitzky-Golay combinando a primeira derivada foram adotados como pré-processamentos para eliminar os ruídos do sistema e distúrbios externos no espectro (GHOLIZADEH et al., 2015; SAVITZKY; GALOAY, 1964).

Para selecionar o número de variáveis preditivas, o modelo foi observado por meio de um gráfico de variância residual da validação cruzada *leave-one-out*, como também por critérios de informação de Akaike (AIC) para a predição de cada variável determinando o número de seus fatores (F) que devem ser mantidos no modelo (AKAIKE, 1998; EFRON; TIBSHIRANI, 1993). A precisão da validação cruzada é dada pelo RMSE (LEONE et al., 2012):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_{pred} - y_{ob})^2}$$
(1)

onde N é o tamanho da amostra, y_{pred} é o valor previsto e y_{ob} é o valor observado.

Além do RMSE, mais alguns parâmetros foram usados para avaliar a precisão da

validação cruzada: R^2 , R^2_{adj} , erro médio (ME), desvio padrão da distribuição de erro (SDE) e o RPD.

O ME foi usado para quantificar o viés (LEONE et al., 2012), segundo a equação abaixo:

$$ME = \frac{1}{N} \sum_{i=1}^{N} (y_{pred} - y_{ob})^2$$
(2)

Para o SDE

$$SDE = \frac{\sum_{i=1}^{N} (y_{pred} - y_{ob} - ME)^2}{N - 1}$$
(3)

Para o RPD

$$RPD = \frac{SD}{RMSE}$$
(4)

onde SD é o desvio padrão dos dados observados e RMSE é a raiz do erro quadrático médio do erro.

Quanto aos valores de RPD, diferentes níveis de interpretação foram abordados: RPD < 1.0 indica predições não confiáveis, inadequadas para análise; 1.0 < RPD < 1.4previsões ruins; 1.4 < RPD < 1.8 indica previsões aceitáveis, adequadas para avaliação e correlação; 1.8 < RPD < 2,0 indica boas previsões, adequadas para avaliação quantitativa; 2.0 < RPD < 2.5 indica previsões quantitativas muito boas; e RPD > 2.5 indica previsões excelentes (VISCARRA ROSSEL et al., 2006). Para que fosse selecionado o comprimento de onda que corresponda a variável na predição foi utilizado o índice de importância variável na projeção (VIP) (VISCARRA ROSSEL; FOUAD; WALTER, 2008), onde os picos mais altos no gráfico correspondem a melhores predições dos atributos.

4 RESULTADOS E DISCUSSÃO

4.1 Análise descritiva

A área de estudo é representada por solos sob diversos graus de desenvolvimento da região Nordeste, sob mínima influência antrópica. Esses solos apresentam, de forma geral, baixo teor de carbono orgânico. Os solos apresentam desde baixa CTC, sobretudo para os derivados de arenitos e demais rochas sedimentares, exceto calcário, a alta CTC (**Tabela 2**), com destaque para alguns solos da região semiárida oriundos de rochas metamórficas e ígneas. Os teores dos metais representam as concentrações naturais desses elementos nos solos.

| Propriedades do solo | Mínimo | Média | Máximo | Desvio Padrão |
|---|--------|----------|-----------|---------------|
| Areia (g kg ⁻¹) | 61,00 | 690,24 | 975,90 | 200,68 |
| Silte (g kg ⁻¹) | 2,40 | 132,73 | 503,20 | 105,89 |
| Argila (g kg ⁻¹) | 24,5 | 177,05 | 713,00 | 130,18 |
| pH | 3,93 | 5,85 | 8,23 | 0,86 |
| $CTC (cmol_c dm^{-3})$ | 3,68 | 12,61 | 68,23 | 10,08 |
| $COT (g kg^{-1})$ | 1,37 | 11,36 | 50,00 | 7,82 |
| Ba (mg kg ⁻¹) | 0,83 | 66,71 | 446,03 | 77,91 |
| $Cd (mg kg^{-1})$ | 0,01 | 0,16 | 1,90 | 0,29 |
| $Co (mg kg^{-1})$ | 0,01 | 8,49 | 51,86 | 8,41 |
| Cr (mg kg ⁻¹) | 0,36 | 28,00 | 266,08 | 32,24 |
| Cu (mg kg ⁻¹) | 0,10 | 10,45 | 81,98 | 11,00 |
| $Fe (mg kg^{-1})$ | 135,57 | 39653,34 | 274140,00 | 56199,98 |
| Mn (mg kg ⁻¹) | 0,55 | 278,25 | 2608,75 | 315,33 |
| Ni (mg kg ⁻¹) | 0,10 | 13,85 | 156,73 | 17,95 |
| Pb (mg kg ⁻¹) | 0,74 | 10,63 | 45,41 | 8,30 |
| Se (mg kg ⁻¹) | 0,02 | 0,30 | 1,70 | 0,24 |
| Sb (mg kg ⁻¹) | 0,02 | 0,37 | 4,26 | 0,55 |
| V (mg kg ⁻¹) | 0,77 | 22,39 | 183,94 | 23,84 |
| $Zn (mg kg^{-1})$ | 2,15 | 19,90 | 95,68 | 17,06 |
| Sc (mg kg ⁻¹) | 0,08 | 2,31 | 12,39 | 2,51 |
| La (mg kg ⁻¹) | 0,40 | 23,01 | 227,40 | 30,41 |
| Ce (mg kg ⁻¹) | 0,80 | 48,88 | 377,10 | 59,63 |
| $\Pr(\text{mg kg}^{-1})$ | 0,04 | 7,32 | 55,84 | 8,40 |
| Nd (mg kg ⁻¹) | 0,05 | 19,16 | 155,00 | 23,78 |
| $Sm (mg kg^{-1})$ | 0,08 | 3,66 | 33,00 | 4,43 |
| Eu (mg kg ⁻¹) | 0,01 | 0,67 | 6,84 | 0,92 |
| $Y (mg kg^{-1})$ | 0,16 | 4,45 | 19,70 | 4,80 |
| $Gd (mg kg^{-1})$ | 0,01 | 2,26 | 29,15 | 3,07 |
| Tb (mg kg ⁻¹) | 0,01 | 0,55 | 5,10 | 0,64 |
| $Dy (mg kg^{-1})$ | 0,01 | 2,40 | 51,30 | 4,91 |
| $\operatorname{Er}(\operatorname{mg} \operatorname{kg}^{-1})$ | 0,05 | 0,99 | 14,35 | 1,57 |
| Ho (mg kg ⁻¹) | 0,03 | 0,24 | 0,96 | 0,19 |
| Yb (mg kg ⁻¹) | 0,03 | 0,82 | 11,00 | 1,31 |
| Lu (mg kg ⁻¹) | 0,01 | 0,15 | 1,00 | 0,15 |
| ETRLs (mg kg ⁻¹) | 1,75 | 102,70 | 793,25 | 124,69 |
| ETRPs (mg kg ⁻¹) | 0,24 | 7,11 | 93,50 | 10,52 |
| ETRLs/ETRPs (mg kg ⁻¹) | 4,84 | 17,69 | 51,57 | 8,02 |
| ETRs (mg kg ⁻¹) | 1,99 | 109,81 | 812,85 | 133,10 |

Tabela 2 - Análise descritiva das variáveis estudadas (n = 210)

Os elementos terras raras (ETRs) apresentaram, de forma geral, uma distribuição relativamente mais uniforme e baixo desvio padrão, como exemplificado no somatório para os elementos terras raras leves (ETRLs) e pesados (ETRPs) (**Tabela 2**). Observa-se maior enriquecimento para os ETRLs, algo já relatado em estudos anteriores (SILVA et al., 2018; MAIA et al., 2022). Esses elementos possibilitam traçar processos pedogenéticos e erosivos (LAVEUF; CORNU, 2009; GUZMÁN et al., 2013; COLLINS et al., 2020), além de serem essenciais nas indústrias automotiva, metalúrgica, eletrônicos e aeroespacial (MIGASZEWSKI; GAŁUSZKA, 2015; WONDRACZEK et al., 2015; MORDOR INTELLIGENCE, 2022), em função da necessidade do desenvolvimento de tecnologias avançadas.

O conhecimento das propriedades do solo, sobretudo a composição química, mineralógica e a distribuição do tamanho de partículas, é fundamental para a predição e, por conseguinte, a compreensão do comportamento dos elementos químicos estudados nesta pesquisa. Esses elementos não apresentam feições de absorção específicas na região do NIR e, portanto, são estimados via intercorrelação com os atributos do solo que são espectralmente ativos na assinatura espectral (i.e., complexação pela matéria orgânica e associação com componentes mineralógicos nas diferentes frações inorgânicas do solo) (GHOLIZADEH et al., 2015).

A areia e argila se correlacionaram negativamente e positivamente com os elementos químicos. A CTC apresentou correlação positiva significativa (p < 0.01) com todos os elementos químicos estudados, exceto Selênio. Por outro lado e já discutido acima, o COT não apresentou correlação com os elementos químicos avaliados. Em ambientes tropicais, os compostos de carbono são rapidamente mineralizados, resultando em baixo teor de COT nos solos avaliados, como observado em solos em ambientes tropicais do Brasil (CAMARGO et al., 2018).

A elevada correlação positiva entre Cu, Zn, Pb, além de outros metais, reflete a semelhança em termos de geoquímica. Essa relação é bem conhecida e, portanto, esses elementos têm uma certa semelhança na geoquímica epigenética (LIU et al., 2020). Os ETRs apresentaram elevada correlação positiva entre eles (0.7 a 0.98; p < 0.01). Isso demonstra o potencial desses elementos para traçar processos pedogenéticos no solo (LAVEUF et al., 2008). Maiores detalhes sobre as associações geoquímicas entre os

ETRs em solos de Pernambuco e Rio Grande do Norte podem ser encontrados em Silva et al. (2016) e Silva et al. (2018). Vale destacar que a elevada correlação dos atributos do solo, sobretudo areia, argila e CTC, com os teores dos metais pesados e ETRs permitem que o primeiro grupo seja usado para a predição do segundo, abordagem empregada com sucesso em Camargo et al. (2018). A elevada correlação entre o Fe e os elementos químicos estudados (dados não apresentados), sugere que esse elemento e a susceptibilidade magnética podem ser uma alternativa em trabalhos futuros.

4.2 Modelagem

O modelo RF apresentou os melhores resultados para as variáveis analisadas. Os resultados da performance dos modelos e erros associados a predição encontram-se nas tabelas 3, 4 e 5. Para argila, Eu, Ho, Nd, Sc, Sm, Y, ETRLs/ETRPs, Sb, Se e V os dados espectrais originais, independente do modelo testado, forneceram os melhores resultados, com RPIQ variando de 0,85 a 2,25 e RMSE de 0,11 a 95,80. Isso ressalta a importância de testar os dados originais antes de aplicar os pré-processamentos.

Dados espectrais originais também forneceram bons resultados na predição de ETRs em solos do maior depósito de fosfato de urânio do Brasil (MAIA et al., 2020). De forma geral, os maiores valores de R² ajustado coincidem com os maiores valores de RPD (CTC, Eu, Yb, Co e Sb), mas ressalta-se que o R² avalia apenas a relação linear entre as variáveis (PASQUINI, 2018), ou seja, estes parâmetros de performance de modelos e erros/incertezas associadas devem ser interpretados de forma conjunta. Elevados valores de R² na calibração não representam unicamente a boa performance do modelo, o qual deve ser também avaliado por meio da RMSE, bias e CCC, que é determinante para avaliar se há erros sistemáticos na construção dos modelos de calibração, dessa forma funcionando como filtro para validação destes.

| Atributo | Modelo | PP | R ² adj | CCC | RMSE | bias | RPD | RPIQ |
|----------|--------|-----|--------------------|------|--------|--------|------|------|
| Areia | RF | SG | 0,48 | 0,53 | 165,17 | 4,15 | 1,35 | 1,70 |
| Silte | RF | SG | 0,35 | 0,49 | 93,17 | 0,71 | 1,27 | 1,43 |
| Argila | RF | RAW | 0,68 | 0,70 | 95,80 | -18,09 | 1,64 | 2,02 |
| COT | PLSR | CR | 0,11 | 0,24 | 7,28 | 0,05 | 1,09 | 1,04 |
| CTC | PLSR | SNV | 0,89 | 0,88 | 3,77 | -0,54 | 2,65 | 1,57 |
| pН | RF | SG | 0,48 | 0,64 | 0,54 | 0,10 | 1,39 | 1,43 |

 Tabela 3 - Melhores modelos obtidos para cada um dos atributos do solo analisados

PP Pré-processamento, RF Floresta Aleatória, PLSR Regressão de Mínimos Quadrados Parciais, RAW Dados Originais, MSC Correção de Dispersão Multiplicativa, SG Savitzky-Golay, SNV Variação Normal Padrão, DET Redução de tendência, CR Remoção de contínuo, CCC Coeficiente de Correlação de

Concordância de Lin, RMSE Raiz do Erro Quadrático Médio, RPD Desvio Residual de Previsão, RPIQ Razão entre o desempenho e a distância interquartil.

Dentre os atributos do solo (**Tabela 3**), a argila e CTC apresentaram os melhores resultados, com R² ajustado iguais a 0,68 e 0,89 e RPIQ de 2,02 e 1,57, respectivamente. Valores satisfatórios foram obtidos para areia, silte e pH. Entretanto, os piores resultados foram obtidos para o carbono orgânico. Como esta pesquisa estuda solos predominantemente inorgânicos, não se pretende explorar o efeito de grupos funcionais da matéria orgânica na predição de elementos químicos, algo que seria incoerente. Além disso, o carbono orgânico não apresentou correlação significativa positiva ou negativa com a maior parte dos elementos químicos estudados, reforçando ainda mais a limitação de buscar uma associação do carbono orgânico com a predição dos elementos químicos no presente estudo. Silva et al. (2018) também observaram ausência de correlação entre COT e ETRs em solos sob diferentes contextos geológicos, pedológicos e climático da região nordeste do Brasil.

Apesar dos elementos químicos não apresentarem feições espectrais específicas no NIR, estes podem ser preditos por meio de covariação com componentes espectralmente ativos como o carbono orgânico e argila (STENBERG et al., 2010; TODOROVA et al., 2014). A presença desses outros componentes e seus níveis de concentração podem afetar consideravelmente a assinatura espectral geral de uma amostra. Ao medir e analisar cuidadosamente a covariação desses componentes, é possível construir modelos preditivos precisos para a análise quantitativa de elementos químicos.

A melhor predição dos ETRs em comparação aos metais pesados pode ser observada pelos maiores valores de RPD, RPIQ e menores valores de RMSE, além dos valores mais próximos de 1 do CCC. Este último é muito útil para comparar duas medições da mesma variável. Apesar da falta de padronização na interpretação dos valores de CCC, é consenso que o mesmo pode ser avaliado da mesma forma que é feito para os coeficientes de correlação de Pearson ou Sperman e deve ser interpretado da forma mais rigorosa possível (MCBRIDE, 2005).

O modelo RF foi mais indicado para a predição de ETRs. Por outro lado, Wang, Zhang e Pan (2017) demonstraram que La, Pr, Nd, Sm e Σ ETRLs foram bem preditos em modelos PLSR. Os ETRLs representam mais de 80% do total. Apesar destes autores terem obtidos bons resultados, acredita-se que tanto o artigo supracitado quanto outros estudos que testaram unicamente o modelo PLSR teriam gerados resultados melhores com o RF (Tabela 4).

| Atributo | Modelo | PP | R^2_{adj} | CCC | RMSE | bias | RPD | RPIQ |
|-------------|--------|-----|-------------|------|-------|-------|------|------|
| Ce | RF | MSC | 0,45 | 0,45 | 31,26 | 7,70 | 1,27 | 1,55 |
| Dy | RF | SG | 0,28 | 0,44 | 0,65 | -0,07 | 1,22 | 1,85 |
| Er | RF | SNV | 0,71 | 0,78 | 0,36 | 0,09 | 1,84 | 3,36 |
| Eu | RF | RAW | 0,70 | 0,35 | 1,12 | -0,26 | 1,22 | 0,85 |
| Gd | RF | DET | 0,44 | 0,58 | 1,63 | -0,26 | 1,37 | 2,27 |
| Но | RF | RAW | 0,51 | 0,63 | 0,11 | 0,03 | 1,42 | 1,80 |
| La | PLSR | MSC | 0,51 | 0,55 | 16,50 | 2,30 | 1,40 | 0,89 |
| Lu | RF | CR | 0,34 | 0,51 | 0,08 | 0,01 | 1,29 | 1,50 |
| Nd | RF | RAW | 0,53 | 0,38 | 20,87 | -4,67 | 1,24 | 1,01 |
| Pr | PLSR | SG | 0,36 | 0,39 | 4,51 | 2,31 | 1,09 | 1,88 |
| Sc | PLSR | RAW | 0,31 | 0,55 | 1,50 | -0,29 | 1,35 | 2,25 |
| Sm | RF | RAW | 0,39 | 0,52 | 1,58 | 0,30 | 1,30 | 2,07 |
| Tb | RF | SG | 0,45 | 0,52 | 0,34 | 0,20 | 1,14 | 1,48 |
| Y | PLSR | RAW | 0,19 | 0,43 | 4,53 | -0,61 | 1,30 | 1,22 |
| Yb | RF | SG | 0,58 | 0,71 | 0,29 | 0,09 | 1,52 | 2,95 |
| ETRLs | RF | SNV | 0,34 | 0,38 | 79,35 | 14,30 | 1,21 | 1,18 |
| ETRPs | RF | SG | 0,22 | 0,38 | 4,16 | 0,25 | 1,18 | 1,90 |
| ETRLs/ETRPs | PLSR | RAW | 0,44 | 0,56 | 4,72 | 0,12 | 1,38 | 1,45 |

Tabela 4 - Melhores modelos obtidos para os ETRs

PP Pré-processamento, RF Floresta Aleatória, PLSR Regressão de Mínimos Quadrados Parciais, RAW Dados Originais, MSC Correção de Dispersão Multiplicativa, SG Savitzky-Golay, SNV Variação Normal Padrão, DET Redução de tendência, CR Remoção de contínuo, CCC Coeficiente de Correlação de Concordância de Lin, RMSE Raiz do Erro Quadrático Médio, RPD Desvio Residual de Previsão, RPIQ Razão entre o desempenho e a distância interquartil. ETRLs Elementos Terras Raras Leves, ETRPs Elementos Terras Raras Pesados. ETRLs/ETRPs Relação entre ETRs leves e pesados.

Como esperado, os maiores valores de RMSE foram obtidos para constituintes mais abundantes, a exemplo da areia, Fe, Mn e Ce (**Tabelas 4 e 5**). Entretanto, as melhores predições não foram obtidas de forma linear para os elementos mais abundantes (e.x., Fe e La), como observado por Maia et al. (2022), onde Al, Fe e Ti, elementos abundantes na crosta terrestre, resultaram em bons modelos de calibração, pois o espectro infravermelho expressa informações físico-químicas sobre a relação existente entre a reflectância e a concentração do elemento, descrita pela lei de Beer-Lambert. Isso pode influenciar na construção do modelo de calibração para concentrações de elementos com base em dados espectrais. Curiosamente, tanto elementos com baixas concentrações, como o Ho e Lu, e os supracitados Fe e La, apresentaram predições não satisfatórias no presente estudo. No primeiro caso, acredita-se por serem elementos com concentrações próximas do limite de detecção para o ICP-OES e com baixo desvio padrão, seja possivelmente inviável predizer esses elementos; para Fe e La, algumas possibilidades podem ser elencadas: 1) alta heterogeneidade dos solos estudados e 2) necessidade de incluir uma faixa espectral maior - emprego do visível (300-750 nm) e infravermelho na faixa do médio (2500-25000 nm) podem ajudar a fornecer informações mais detalhadas sobre a presença ou ausência de capacidade preditiva para esses elementos.

| Atributo | Modelo | PP | R ² adj | CCC | RMSE | bias | RPD | RPIQ |
|----------|--------|-----|--------------------|------|----------|----------|------|------|
| Ba | RF | SNV | 0,49 | 0,60 | 57,65 | 8,45 | 1,40 | 1,03 |
| Cd | PLSR | CR | 0,04 | 0,08 | 0,30 | -0,01 | 1,04 | 0,17 |
| Co | RF | DET | 0,63 | 0,68 | 6,18 | -1,19 | 1,57 | 1,70 |
| Cr | RF | DET | 0,48 | 0,63 | 17,23 | 4,96 | 1,37 | 1,16 |
| Cu | PLSR | CR | 0,24 | 0,40 | 12,39 | 0,59 | 1,17 | 0,68 |
| Fe | RF | DET | 0,45 | 0,54 | 36816,09 | -1941,20 | 1,34 | 0,78 |
| Mn | RF | DET | 0,35 | 0,53 | 198,35 | -1,95 | 1,27 | 1,42 |
| Ni | RF | DET | 0,38 | 0,50 | 8,42 | 3,53 | 1,19 | 1,44 |
| Pb | RF | CR | 0,42 | 0,43 | 5,78 | 0,71 | 1,26 | 1,32 |
| Sb | RF | RAW | 0,60 | 0,65 | 0,21 | 0,06 | 1,51 | 1,01 |
| Se | PLSR | RAW | 0,13 | 0,20 | 0,28 | -0,08 | 1,05 | 0,94 |
| V | RF | RAW | 0,44 | 0,52 | 16,39 | -2,71 | 1,35 | 1,59 |
| Zn | PLSR | DET | 0,40 | 0,53 | 9,63 | 0,80 | 1,31 | 1,59 |

 Tabela 5 - Melhores modelos obtidos para os metais analisados

PP Pré-processamento, RF Floresta Aleatória, PLSR Regressão de Mínimos Quadrados Parciais, RAW Dados Originais, SNV Variação Normal Padrão, DET Redução de tendência, CR Remoção de contínuo, CCC Coeficiente de Correlação de Concordância de Lin, RMSE Raiz do Erro Quadrático Médio, RPD Desvio Residual de Previsão, RPIQ Razão entre o desempenho e a distância interquartil.

Selênio e Cd não apresentaram predições satisfatórias (**Tabela 5**). Esses elementos não possuem características/respostas espectrais e são limitados em espectros vis-NIR. Outros autores observaram resultados semelhantes para Mn, Zn e As (WU et al., 2007; WANG et al., 2014; LIU et al., 2020). Na validação dos modelos utilizando o préprocessamento SG, a melhor predição foi observada para areia, silte, pH, Dy, Pr, Tb, Yb, ETRPs; CR para COT, Lu, Cd, Cu e Pb; SNV para CTC, Er, ETRLs e Ba; DET para Gd, Co, Cr, Fe, Mn, Ni, Zn; MSC Ce e La. Percebe-se que, independentemente do préprocessamento utilizado, a escolha do modelo foi mais relevante do que a escolha do préprocessamento. Essa tendência também foi observada por Naibo et al. (2022), que estudaram NIR para estimar a concentração de elementos químicos em solos e sedimentos em uma bacia hidrográfica rural. Portanto, recomenda-se para os mais diversos estudos, a utilização combinada de modelos com diferentes naturezas computacionais. Trabalhos que testam um único modelo, mesmo abrangendo o maior número de pré-processamentos possível, devem ser avaliados com cautela.

Os micronutrientes Fe, Co e Mn apresentaram as melhores predições para o modelo RF e pré-processamento DET, que remove as variações indesejadas no espectro. O micronutriente Zn apresentou a mesma tendência, exceto pelo modelo que foi o PLSR. Isso aumenta a aplicabilidade e confiabilidade dos resultados, uma vez que pode auxiliar no monitoramento e reposição de micronutrientes no solo em áreas a serem exploradas para agricultura.

Em grandes territórios com elevada diversidade geológica e pedológica e distintos padrões climáticos, como evidenciado neste estudo, a predição de elementos químicos pode se tornar um desafio ainda maior. Provavelmente, a densidade amostral do estudo, embora abrangendo as áreas com mínima influência antrópica, pode não ter sido suficiente para elementos químicos com altos valores de desvio padrão. Talvez por isso e pelo aspecto mais conservativo dos ETRs nos solos, estes elementos forneceram os melhores valores de RPD, RPIQ (**Figuras 16c e 17c**) e RMSE. Essa afirmativa pode ser confirmada pelos resultados obtidos por Camargo et al. (2018), Maia et al. (2020) e Maia et al. (2022) que avaliaram a predição de elementos químicos em uma área consideravelmente menor do que o presente estudo, 5 km², 1,54 km² e 3.430 km², respectivamente. Curiosamente, e de encontro ao que está posto acima, mesmo com um número de amostras de solo menor, como as 50 amostras compostas utilizadas em Maia et al. (2020) e as 44 amostras de solo utilizadas em Liu et al. (2020), a predição pode ser alta em áreas mais homogêneas em termos da concentração de metais.

CR → MSC DET → RAW 4.0 a) SG SNV . 3.5 3.0 2.5 RPIQ 2.0 1.5 1.0 0.5 0.0 -Ba Ċu Fe Mn Zn Ċd Ċr Ni Pb Sb Co Se v b) 4.0 -CR DET MSC RAW SG SNV + . 3.5 3.0 2.5 RPIQ 2.0 1.5 1.0 0.5 0.0 pH areia silte CÓT ctc argila c) 4.0 → MSC CR DET SG SNV 3.5 3.0 2.5 RPIQ 2.0 1.5 1.0 0.5 0.0 Ce ETRLS ETRPS ETRLS/ ETRPS Ďу Gd Ho La Lu Nd Ýb Ér Eu Pr Sc Sim Тb Ý

Figura 16 - Valores de RPIQ para as variáveis estudadas com o modelo PLSR. a) metais, b) atributos e c) ETRs

Figura 17 - Valores de RPIQ para as variáveis estudadas com o modelo RF. a) metais, b) atributos e c) ETRs



O RPIQ foi utilizado nesta pesquisa para avaliar a performance dos modelos. Esse parâmetro confere maior acurácia em comparação com o RPD, pois baseia-se na distância interquartil para explicar a distribuição dos dados (COBLINSKI et al., 2020). Segundo a classificação de Dotto et al. (2017), os modelos foram considerados bons (RPIQ > 1,5)

para areia, argila, CTC, Ce, Dy, Ho, Lu, Pr, Co, V, Zn; muito bons RPIQ > 2,0 para argila, Er, Gd, Sc, Sm, Yb.

4.3 Predição de ETRs em amostras contendo apenas as assinaturas espectrais

A partir dos bons resultados iniciais para os ETRs, esses elementos foram utilizados (i.e., elementos com maior capacidade preditiva) para criar e testar modelos nas áreas onde existem apenas as assinaturas espectrais. Foram escolhidos os ETRs pelo recente status de poluentes emergentes deste grupo de elementos e pelo crescimento dos setores agrícola e industrial que aumentou a demanda por elementos da série dos lantanídeos. Dessa forma, almeja-se responder a seguinte pergunta: Seria possível fazer a predição em áreas onde não foram realizadas as análises químicas usando apenas assinaturas espectrais? Para tal, foram usados modelos calibrados e validados com amostras que apresentam o mesmo contexto geológico e pedológico, conforme descrito no material e métodos, assim como padrão climático semelhante, uma vez que as duas regiões apresentam uma porção semiárida e outra mais úmida (precipitação pluvial variando de 500 a 1500 mm por ano).

Os valores de RPD foram considerados aceitáveis para Sm, Eu, Gd, Tb, Er, Yb e ETRPs (**Tabela 6**). Lantânio e Ce, ETRs mais abundantes e somam frequentemente mais de 80% do total destes elementos, devem ser interpretados com cautela. Os semivariogramas e gráficos de validação cruzada para os valores observados e preditos encontram-se no anexo.

| Variáveis | | Parâmetros da validação de teste $n = 91$ Parâmetros da validação cruzado | | | | | | | |
|-------------|---|---|--------|-------|------|---------------------|--------|--|--|
| | F | R ² adj | PP | RMSE | RPD | R ² ajus | RMSE | | |
| La | 8 | 0,33 | SNV-SG | 15,34 | 1,2 | 0,30 | 16,681 | | |
| Ce | 8 | 0,43 | SNV-SG | 26,47 | 1,31 | 0,38 | 28,929 | | |
| Pr | 8 | 0,39 | SNV-SG | 5,21 | 1,27 | 0,35 | 5,577 | | |
| Nd | 8 | 0,45 | SNV-SG | 10,18 | 1,34 | 0,37 | 11,248 | | |
| Sm | 9 | 0,50 | SNV-SG | 1,73 | 1,42 | 0,46 | 1,823 | | |
| Eu | 8 | 0,54 | SNV-SG | 1,48 | 1,48 | 0,48 | 0,334 | | |
| ETRLs | 8 | 0,42 | SNV-SG | 57,36 | 1,31 | 0,37 | 62,153 | | |
| Gd | 8 | 0,63 | SNV-SG | 1,30 | 1,65 | 0,55 | 1,44 | | |
| Tb | 9 | 0,60 | SNV-SG | 0,26 | 1,59 | 0,52 | 0,295 | | |
| Er | 8 | 0,58 | SNV-SG | 0,40 | 1,56 | 0,52 | 0,47 | | |
| Yb | 8 | 0,59 | SNV-SG | 0,34 | 1,57 | 0,56 | 0,359 | | |
| ETRPs | 8 | 0,63 | SNV-SG | 2,18 | 1,65 | 0,56 | 2,386 | | |
| ETRLs/ETRPs | 5 | 0,29 | SNV-SG | 7,11 | 1,18 | 0,20 | 8,021 | | |

Tabela 6 - Resumo dos resultados obtidos através do modelo PLSR para os ETRs

F, número de fatores PLSR usados no modelo; PP pré-processamento, SNV Variação normal padrão; SG Savitzky-Golay; coeficiente de determinação ajustado; RMSE Raiz do erro quadrático médio, RPD, desvio de previsão residual.

Maiores teores de ETRs, tanto para os valores observados e preditos, foram observados em solos derivados de rochas ígneas e metamórficas. Os valores observados para La, Ce, Pr, Sm, LREE, Tb e Er tiveram forte dependência espacial (**Tabela 7**). De forma geral, os valores preditos apresentaram dependência espacial moderada, exceto para Pr, Sm e Tb que apresentaram grau de dependência espacial variando de 0-25% (dependência espacial forte). Isso reforça a qualidade destes mapas (Figura 18) que são essenciais para identificar áreas susceptíveis aos impactos ambientais – etapa primordial no estabelecimento de futuras políticas ambientais que afetam a saúde humana e a proteção ambiental (SILVA et al., 2018), ajudando a identificar os hotspots e suas fontes (WANG; LIANG, 2016). Assim como observado por Pereira et al. (2019), La e Ce também apresentam valores mais baixos de alcance em função da alta aleatoriedade espacial foram desconsideradas. Segundo Gao et al. (2016), ETRs que não apresentaram dependência espacial podem estar associados à microvariabilidade dos resultados.

Os semivariogramas experimentais dos ETRs se ajustaram melhor ao modelo esférico ou exponencial (**Tabela 7**). A mesma tendência foi observada por Silva et al. (2018), estudando a distribuição de ETRs em solos da região do nordeste brasileiro e por Pereira et al. (2019) que estudaram ETRs em solos derivados de rochas sedimentares em área de fronteira agrícola do Piauí.

| Variáveis | Modelo | C_0^{a} | $C_0 + C_1^{b}$ | DSD ^c | a ^d | R ² | RSS | VC |
|-------------------------------|--------|-----------|-----------------|------------------|----------------|----------------|----------|-------|
| La | Exp. | 0,099 | 2,416 | 4,097 | 27300 | 0,594 | 0,356 | 0,434 |
| La _{pred} | Esf. | 0,769 | 1,728 | 44,50 | 104100 | 0,513 | 0,616 | 0,777 |
| Ce | Exp. | 0,27 | 5,611 | 4,81 | 25800 | 0,613 | 1,47 | 0,777 |
| Ce_{pred} | Esf. | 0,968 | 3,0363 | 31,88 | 818000 | 0,612 | 1,07 | 0,818 |
| Pr | Exp. | 0,049 | 1,139 | 4,30 | 25800 | 0,640 | 0,054 | 0,358 |
| Pr _{pred} | Exp. | 0,101 | 0,584 | 17,29 | 62700 | 0,414 | 0,0669 | 0,860 |
| Nd | Exp. | 0,843 | 2,22 | 37,97 | 43200 | 0,571 | 0,305 | 0,476 |
| Nd _{pred} | Esf. | 0,279 | 0,9027 | 30,95 | 101200 | 0,516 | 0,152 | 0,768 |
| Sm | Exp. | 0,051 | 0,561 | 9,09 | 32100 | 0,788 | 0,009 | 0,658 |
| Sm _{pred} | Esf. | 0,025 | 0,2186 | 11,44 | 105500 | 0,559 | 0,0146 | 0,550 |
| ETRLs | Exp. | 0,500 | 12,150 | 4,11 | 28500 | 0,709 | 5,92 | 0,445 |
| ETRLs _{pred} | Esf. | 0,039 | 0,0793 | 50,57 | 81200 | 0,517 | 0,000748 | 0,916 |
| Gd | Exp. | 0,062 | 0,552 | 11,23 | 30600 | 0,605 | 0,0208 | 0,264 |
| $\mathrm{Gd}_{\mathrm{pred}}$ | Exp. | 0,141 | 0,3592 | 39,42 | 150000 | 0,410 | 0,0232 | 0,608 |
| Tb | Exp. | 0,010 | 0,0793 | 13,11 | 27900 | 0,790 | 0,000297 | 0,372 |
| Tb _{pred} | Esf. | 0,007 | 0,0300 | 24,00 | 107300 | 0,639 | 0,00016 | 0,687 |
| Er | Exp. | 0,008 | 0,1076 | 7,43 | 22600 | 0,635 | 0,000404 | 0,360 |
| Er _{pred} | Exp. | 0,039 | 0,0977 | 39,98 | 186000 | 0,725 | 0,000619 | 0,786 |

 Tabela 7 - Parâmetros do semivariograma e modelos ajustados a partir de valores observados e preditos

Exp., exponencial; Esf., esférico; efeito pepita; b, patamar; c, grau de dependência espacial (DSD) é considerado forte DSD < 25%, moderado DSD \ge 25 a \le 75% e fraco DSD > 75%; d, alcance (m); RSS soma de quadrados residuais; VC, validação cruzada. Eu e Yb e ETRPs foram excluídos porque apresentaram efeito pepita puro para os valores preditos.

Em ambos os casos, na porção litorânea são observados os menores teores de ETRs, onde os solos são compostos por sedimentos arenosos inconsolidados do Grupo Barreiras. Nesses solos ocorre a predominância de quartzo e, por conseguinte, menores teores de ETRs.

Figura 18 - Mapas dos valores observados (coluna da esquerda – solos do RN) e preditos (coluna da direita – solos da PB) de ETRs por meio do NIR





Destaca-se que esse é o primeiro trabalho que gera mapas de distribuição de ETRs a partir apenas de assinaturas espectrais e amostras de áreas que apresentam condições climáticas, geológicas e pedológicas semelhantes. Ainda, não se recomenda a utilização da mesma abordagem em áreas sob influência agrícola e com contribuição considerável de eventos que possam alterar as concentrações naturais dos elementos químicos estudados, a exemplo da deposição atmosférica. A partir destes resultados iniciais, foram usados os elementos com a maior capacidade preditiva para criar e testar modelos para ETRs nas áreas dentro do estudo onde existem apenas as assinaturas espectrais.

5 CONCLUSÕES

O modelo RF apresentou os melhores resultados tanto para os atributos do solo quanto para os elementos químicos. Percebe-se que, independentemente do préprocessamento usado, a escolha do modelo foi mais relevante do que a escolha do préprocessamento. Portanto, recomenda-se, para os mais diversos estudos, a utilização combinada de modelos com diferentes naturezas computacionais. Trabalhos que testam um único modelo, mesmo abrangendo o maior número de pré-processamentos possível, devem ser interpretados com cautela. A melhor predição dos ETRs em comparação aos metais pesados pode ser observada pelos maiores valores de RPD, RPIQ e menores erros associados.

De forma geral, os valores preditos para regiões onde existiam apenas as assinaturas espectrais apresentaram dependência espacial moderada, exceto para Pr, Sm e Tb que apresentaram dependência espacial forte. Isso reforça a qualidade destes mapas que são essenciais para identificar áreas susceptíveis aos impactos ambientais – etapa primordial no estabelecimento de futuras políticas ambientais que afetam a saúde humana e a proteção ambiental, ajudando a identificar os hotspots e suas fontes.

REFERÊNCIAS BIBLIOGRÁFICAS

AKAIKE, H. Information Theory and an Extension of the Maximum Likelihood Principle. Em: KOTZ (Ed.). **Foundations and Basic Theory**. 1. ed. New York: Springer-Verlag, v. 1p. 199–213, 1998.

ALMEIDA JÚNIOR, A. B. **Teores naturais e valores de referência de qualidade para metais pesados em solos do Estado da Paraíba**. 2014. 93 f. Tese (Doutorado em Ciência do Solo) – Universidade Federal Rural de Pernambuco, Recife, 2014.

ALMEIDA JÚNIOR, A. B.; NASCIMENTO, C. W. A.; BIONDI, C. M.; SUZA, A. P.; BARROS, F. M. R. Background and Reference Values of Metals in Soil from Paraíba State, Brazil. **Revista Brasileira de Ciência do Solo**, v. 40, e0150122, 2016. https://doi.org/10.1590/18069657rbcs20150122

BALARAM, V. Rare earth elements: a review of applications, occurrence, exploration, analysis, recycling, and environmental impact. **Geoscience Frontiers**, v. 10, n. 4, p. 1285–1303, 2019. <u>https://doi.org/10.1016/j.gsf.2018.12.005</u>

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. https://doi.org/10.1366/0003702894202201

BELLON-MAUREL, V.; FERNANDEZ-AHUMADA, E.; PALAGOS, B.; ROGER, J.-M.; MCBRATNEY, A. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. **TrAC** -**Trends in Analytical Chemistry**, v. 29, p. 1073–1081. 2010. https://doi.org/10.1016/j.trac.2010.05.006

BRASIL. Mistério da Agricultura. Mapa Exploratório - Reconhecimento de solos do Estado da Paraíba, Rio de Janeiro, Sudene, 1972. (Esc. 1:500.000).

BREIMAN, L. Random Forests, Machine Learning, v. 45, n. 1, p. 5–32, 2001. https://doi.org/10.1023/A:1010933404324

CAMARGO, L. A.; MARQUES, J.; BARRÓN, V.; ALLEONI, L. R. F.; PEREIRA, G. T.; TEIXEIRA, D. B.; BAHIA, A. S. R. S. Predicting potentially toxic elements in tropical soils from iron oxides, magnetic susceptibility and diffuse reflectance spectra. **Catena**, v. 5, p. 503-515, 2018. <u>https://doi.org/10.1016/j.catena.2018.02.030</u>

CAMBARDELLA, C. A.; MOORMAN, T. B.; NOVAK, J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA, A. E. Field-Scale Variability of Soil Properties in Central Iowa Soils. **Soil Science Society of America Journal**, v. 58, n. 5, p. 1501–1511, 1994. <u>https://doi.org/10.2136/sssaj1994.03615995005800050033x</u>

CAMPOS, M. C. C.; QUEIROZ, S. B. Reclassificação dos perfis descritos no levantamento exploratório- Reconhecimento de solos do Estado da Paraíba. **Revista de Biologia e Ciência da Terra**, v. 6, n.1, p. 45-50, 2006.

CAO, J.; LI, C.; WU, Q.; QIAO, J. Improved Mapping of Soil Heavy Metals Using a Vis-NIR Spectroscopy Index in an Agricultural Area of Eastern China. **IEEE Access**, v. 8, p. 42584-42594, 2020. <u>https://doi.org/10.1109/ACCESS.2020.2976902</u>

BIONDI, C. M. **Teores naturais de metais pesados nos de referência do Estado de Pernambuco**. 2010. 70 f. Tese (Doutorado em Ciência do Solo) – Universidade Federal Rural de Pernambuco, Recife, 2010.

CHENG, H.; SHEN, R.; CHEN, Y.; WAN, Q.; SHI, T.; WANG, J.; WAN, Y.; HONG, Y.; LI, X. Estimating heavy metal concentrations in suburban soils with reflectance spectroscopy. **Geoderma**, Amsterdam, v. 336, p. 59-67, 2019. https://doi.org/10.1016/j.geoderma.2018.08.010

CLARK, R. N.; ROUSH, T. L. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. **Journal of Geophysical Research**, Washington, v. 89, p. 6329–6340, 1984. <u>https://doi.org/10.1029/JB089iB07p06329</u>

COLLINS, A. L.; BLACKWELL, M.; BOECKX, P.; CHIVERS, C. A.; EMELKO, M.; EVRARD, O.; FOSTER, I.; GELLIS, A.; GHOLAMI, H.; GRANGER, S.; HARRIS, P.; HOROWITZ, A. J.; LACEBY, J. P.; MARTINEZ-CARRERAS, N.; MINELLA, J.; MOL, L.; NOSRATI, K.; PULLEY, S.; SILINS, U.; SILVA, Y. J. A. B.; STONE, M.; TIECHER, T.; UPADHAYAY, H. R.; ZHANG, Y. Sediment source fingerprinting: benchmarking recent outputs, remaining challenges and emerging themes. Journal of Soils and Sediments, v. 20, p. 4160-93, 2020. <u>https://doi.org/10.1007/s11368-020-02755-4</u>

COBLINSKI, J. A.; GIASSON, É.; DEMATTÊ, J. A. M.; DOTTO, A. C.; COSTA, J. J. F.; VASAT, R. Prediction of soil texture classes through different wavelength regions of reflectance spectroscopy at various soil depths. **Catena**, v. 189, n. 104485, 2020. https://doi.org/10.1016/j.catena.2020.104485

COSTA, W. P. L. B. **Metais pesados em solos do Rio Grande do Norte: valores de referência de qualidade e relações geopedológicas**. 2013. 121 f. Tese (Doutorado em Ciência do Solo) – Universidade Federal Rural de Pernambuco, Recife, 2010.

CPRM- Serviço Geológico do Brasil. **Geologia e recursos minerais do Estado da Paraíba**. Recife, 2002, 234p.

DHANOA, M. S.; LISTER, S. J.; SANDERSON, R.; BARNES, R. J. The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra. **Journal of Near Infrared Spectroscopy**, v. 21, p. 43-47, 1994. <u>https://doi.org/10.1255/jnirs.30</u>

DONAGEMA, G. K.; CAMPOS, D. V. B.; CALDERANO, S. B.; TEIXEIRA, W. G.; VIANA, J. H. M. (Org.). Manual de métodos de análise de solo. 2.ed. Rio de Janeiro: Embrapa Solos, p. 225, 2011.

DOTTO, A. C.; DALMOLIN, R. S. D.; GRNWALD, S.; TEN CATEN, A.; PEREIRA FILHO, W. Two preprocessing techniques to reduce model covariables in soil property predictions by Vis-NIR spectroscopy. **Soil and Tillage Research**, v. 172, p. 59–68, 2017. <u>https://doi.org/10.1016/j.still.2017.05.008.</u>

EFRON, B.; TIBSHIRANI, R. J. An Introduction to the Bootstrap: Monographs on statistics and applied probability, v 57, 1993.

GAO, Z.; FU, W.; ZHANG, M.; ZHAO, K.; TUNNEY, H.; GUAN, Y. Potentially hazardous metals contamination in soilrice system and it's spatial variation in Shengzhou City, China. **Journal of Geochemical Exploration**, v. 167, p. 62–69, 2016. https://doi.org/10.1016/j.gexplo.2016.05.006

GEE, G. W.; OR, D. Particle Size Analysis. In: DANE, J. H. & Topp, G. C. **Methods of soil analysis**. 4. Ed. Physical methods - Madison, WI: Soil Science Society of America. p. 255-293, 2002.

GHOLIZADEH, A.; BORŮVKA, L.; SABERIOON, M. M.; KOZÁK, J.; VAŠÁT, R.; NĚMEČEK, K. Comparing different data preprocessing methods for monitoring soil heavy metals based on soil spectral features. **Soil and Water Research**, v. 10, n. 4, p. 218-227, 2015. <u>https://doi.org/10.17221/113/2015-SWR</u>

GUZMÁN, G.; QUINTON, J. N.; NEARING, M. A.; MABIT, L.; GÓMEZ, J. A. SEDIMENT tracers in water erosion studies: current approaches and challenges. **Journal of Soils and Sediments,** v. 13, p. 816-33, 2013. https://doi.org/10.1007/s11368-013-0659-5

GUINDO, M. L.; KABIR, M. H.; CHEN, R.; LIU, F. Potential of Vis-NIR to measure heavy metals in different varieties of organic-fertilizers using Boruta and deep belief network. **Ecotoxicology and Environmental Safety**, v. 228, p. 112996, 2021. https://doi.org/10.1016/j.ecoenv.2021.112996

HORTA, A.; MALONE, B.; STOCKMANN, U.; MINASNY, B.; BISHOP, T. F. A.; MCBRATNEY, A. B.; PALLASSER, R.; POZZA, L. Potential of integrated field spectroscopy and spatial analysis for enhanced assessment of soil contamination: a prospective review. **Geoderma**, v. 241, p. 180–209, 2015. https://doi.org/10.1016/j.geoderma.2014.11.024.

INTELLIGENCE. M. Rare earth elements market - growth, trends, covid-19 impact, and forecasts (2022 - 2027). Hyderabad, Telangana: Mordor Intelligence; 2022 [acesso 2 Nov 2022]. Disponível em: <u>https://www.mordorintelligence.com/industry-</u> <u>reports/rare-earth</u> <u>elementsmarket?gclid=Cj0KCQjwqoibBhDUARIsAH2OpWj0H_QRpqoA_yNWjMg6</u> mE3FU_zIgCVBsZwy8RfOjXi3K25bBTjr5RwaAhf-EALw_wcB

JOHNSON, J.; VANDAMME, E.; SENTHILKUMAR, K.; SILA, A.; SHEPHERD, K. D.; SAITO, K. Near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for assessing soil fertility in rice fields in sub-Saharan Africa, **Geoderma**, v. 354, p. 113840, 2019. <u>https://doi.org/10.1016/j.geoderma.2019.06.043</u>.

KHOSRAVI, V.; ARDEJANI, F. D.; YOUSEFI, S.; ARYAFAR, A. Monitoring soil lead and zinc contents via combination of spectroscopy with extreme learning machine and other data mining methods. **Geoderma**, v. 318, p. 29–41, 2018. https://doi.org/10.1016/j.geoderma.2017.12.025

KNOX, N. M.; GRUNWALD, S.; MCDOWELL, M. L.; BRULAND, G. L.; MYERS, D. B.; HARRIS, W. G. Modelling soil carbon fractions with visible near-infrared (VNIR) and mid-infrared (MIR) spectroscopy, **Geoderma**, v. 239, p. 229–239, 2015. https://doi.org/10.1016/j.geoderma.2014.10.019

KUHN, M. Building Predictive Models in R Using the caret Package. Journal of Statistical Software, v. 28, p. 1–26, 2008. <u>https://doi.org/10.18637/jss.v028.i05</u>

LAVEUF, C.; CORNU, S. A. Review on the potentiality of rare earth elements to trace pedogenetic processes. **Geoderma**, v. 154, p. 1-12, 2009. https://doi.org/10.1016/j.geoderma.2009.10.002

LAVEUF, C.; CORNU, S.; JUILLOT, F. Rare earth elements as tracers of pedogenetic processes. **Comptes Rendus Geosciences**, v. 340, p. 523-32, 2008. <u>https://doi.org/10.1016/j.crte.2008.07.001</u>

LEONE, A. P.; VISCARRA-ROSSEL, R. A.; AMENTA, P.; BUONDONNO, A. Prediction of Soil Properties with PLSR and vis-NIR Spectroscopy: Application to Mediterranean Soils from Southern Italy. **Current Analytical Chemistry**, v. 8, n. 2, p. 283–299, 1, 2012. <u>https://doi.org/10.2174/157341112800392571</u>

LIN, L. I. A concordance correlation coefficient to evaluate reproducibility. **Biometrics**, v. 45, n.1, p. 255-268, 1989.

LIU, Y.; PAN, X.; WANG, C.; LI, Y.; SHI, R. Predicting Soil Salinity with Vis–NIR Spectra after Removing the Effects of Soil Moisture Using External Parameter Orthogonalization. **PLoS ONE**, v. 10(10): e0140688, 2015. https://doi.org/10.1371/journal.pone.0140688

LIU, J., XIE, J., HAN, J.; WANG, H.; SUN, J.; LI, R.; LI, S. Visible and near-infrared spectroscopy with chemometrics are able to predict soil physical and chemical properties. **Journal of Soils and Sediments**, v. 20, p. 2749–2760, 2020. https://doi.org/10.1007/s11368-020-02623-1

MAIA, A. J.; NASCIMENTO, R. C.; SILVA, Y. J. A. B.; NASCIMENTO, C. W. A.; MENDES, W. S.; NETO, J. G. V.; FILHO, J. C. A.; TIECHER, T.; SILVA, Y. J. A. B. Near-infrared spectroscopy for prediction of potentially toxic elements in soil and sediments from a semiarid and coastal humid tropical transitional river basin. **Microchemical Journal**, v. 179, 107544, 2022. https://doi.org/10.1016/j.microc.2022.107544

MAIA, A. J.; DA SILVA, Y. J. A. B.; DO NASCIMENTO, C. W. A.; ESCOBAR, M. E. O.; CUNHA, C. S. M.; SILVA, Y. J. A. B.; NASCIMENTO, R. C.; PEREIRA, L. H. S. Near-infrared spectroscopy for the prediction of rare earth elements in soils from the largest uranium-phosphate deposit in Brazil using PLS, iPLS, and iSPA-PLS models. **Environ Monit Assess** v. 192, p. 675, 2020. <u>https://doi.org/10.1007/s10661-020-08642-2</u>

MCBRIDE, G. B. A proposal for strength-of-agreement criteria for Lin's concordance correlation, 2005. <u>https://www.medcalc.org/download/pdf/McBride2005.pdf</u>

MENDES, W. S.; BOECHAT, C. L.; GUALBERTO, A. V. S.; BARBOSA, R. S.; SILVA, Y. J. A. B.; SARAIVA, P.C.; SENA, A. F. S.; DUARTE, L. S. L. Soil spectral library of Piauí State using machine learning for laboratory analysis in Northeastern Brazil. **Revista Brasileira de Ciência do Solo**, 45, e0200115, 2021. <u>https://doi.org/10.36783/18069657rbcs20200115</u>.

MIGASZEWSKI, Z. M.; GAŁUSZKA, A. The characteristics, occurrence, and geochemical behavior of rare earth elements in the environment: A review. **Critical Reviews in Environmental Science and Technology**, v. 45, p. 429-71, 2015. https://doi.org/10.1080/10643389.2013.866622

MINASNY, B.; MCBRATNEY, A. B. A conditioned Latin hypercube method for sampling in the presence of ancillary information, **Computers and Geosciences**, v. 32, p. 1378–1388, 2006. <u>https://doi.org/10.1016/j.cageo.2005.12.009.</u>

MOURA-BUENO, J. M.; DALMOLIN, R. S. D.; CATEN, A.; DOTTO, A. C.; DEMATTÊ, J. A. M. 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. <u>https://doi.org/10.1016/j.geoderma.2018.10.015</u>

NAIBO, G.; RAMON, R.; PESINI, G.; MOURA-BUENO, J. M.; BARROS, C. A. P.; CANER, L.; SILVA, Y. J. A. B.; MINELLA, J. P. G.; SANTOS, D. R.; TIECHER, T. Near-infrared spectroscopy to estimate the chemical element concentration in soils and sediments in a rural catchment. **Catena**, v. 213, p. 106145, 2022. https://doi.org/10.1016/j.catena.2022.106145

NASCIMENTO, C. W. A.; OLIVEIRA, A. B.; RIBEIRO, M. R.; MELO, É. E. C. Distribution and availability of zinc and copper in benchmark soils of Brazil. **Communications in Soil Science and Plant Analysis,** v. 37, p. 109-25, 2006. https://doi.org/10.1080/00103620500403895

NIST — National Institute of Standards and Technology. Standard Reference Materials-SRM 2709, 2710 and 2711 Addendum Issue Date: 18 January 2002.

O'ROURKE, S. M.; MINASNY, B.; HOLDEN, N. M.; MCBRATNEY, A. B. Synergistic Use of Vis-NIR, MIR, and XRF spectroscopy for the determination of soil geochemistry. **Soil Science Society of America Journal**, v. 80, p. 888-899, 2016. https://doi.org/10.2136/sssaj2015.10.0361 PASQUINI, C. Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. **Analytica Chimica Acta**, v. 1026, p. 8-36, 2018. <u>https://doi.org/10.1016/j.aca.2018.04.004</u>

PEREIRA, B. A.; SILVA, Y. J. A. B.; NASCIMENTO, C. W. A.; SILVA, Y. J. A. B.; NASCIMENTO, R. C.; BOECHAT, C. L.; BARBOSA, R. S.; SINGH, V. P. Watershed scale assessment of rare earth elements in soils derived from sedimentary rocks. **Environmental Monitoring and Assessment**, v. 191, p. 514, 2019. https://doi.org/10.1007/s10661-019-7658-y

REEVES, J. B., III, MCCARTY, G. W., MEISINGER, J. J. Near infrared reflectance spectroscopy for the analysis of agricultural soils. **Journal of Near Infrared Spectroscopy.**, v. 7, p. 179–193, 1999. <u>https://doi.org/10.1255/jnirs.248</u>

ROSIN, N. A.; DALMOLIN, R. S. D.; HORST-HEINEN, T.; MOURA-BUENO, J. M.; SILVA-SANGOI, D. V.; SILVA, L. S. 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. <u>https://doi.org/10.1590/1678-992X-2019-0246</u>

SANTANA, F. B.; SOUZA, A. M.; POPPI, R. J. Visible and near infrared spectroscopy coupled to random forest to quantify some soil quality parameters. **Spectrochimica.** Acta Part A: Molecular and Biomolecular Spectroscopy, v. 191, p. 454–462, 2018. https://doi.org/10.1016/j.saa.2017.10.052

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, v.36, p. 1627–1639, 1964. <u>https://doi.org/10.1021/ac60214a047</u>

SILVA, A. C.; TORRADO, P. V.; ABREU JUNIOR, J. DE S. Métodos de quantificação da matéria orgânica do solo. **Revista da Universidade de Alfenas**, v.5, p. 21-26, 1999.

STEVENS, A.; RAMIREZ-LOPEZ, L. (2020). An introduction to the prospectr package. R package version 0.2.1. https://cran.rproject.org/web/packages/prospectr/citation.html (accessed 24 november 2021).

STENBERG, B.; VISCARRA ROSSEL, R.; MOUAZEN, A. M.; WETTERLIND, J. Visible and near infrared spectroscopy in soil science. **Advances in Agronomy**, v. 107, p. 163–215, 2010. <u>https://doi.org/10.1016/S0065-2113(10)07005-7</u>

SILVA, C. M. C. A. C.; BARBOSA, R. S.; NASCIMENTO, C. W. A.; SILVA, Y. J. A. B.; SILVA, Y. J. A. B. Geochemistry and spatial variability of rare earth elements in soils under different geological and climate patterns of the Brazilian Northeast. **Revista Brasileira de Ciência do Solo**, v. 42, e0170342, 2018. https://doi.org/10.1590/18069657rbcs20170342

SILVA, Y. J. A. B.; NASCIMENTO, C. W. A.; SILVA, Y. J. A. B.; BIONDI, C. M.; SILVA, C. M. C. A. C. Rare earth element concentrations in Brazilian benchmark soils.

Revista Brasileira de Ciência do Solo, v. 40, e0150413, 2016. https://doi.org/10.1590/18069657rbcs20150413

TAN, K.; MA, W.; WU, F.; DU, Q. Random forest–based estimation of heavy metal concentration in agricultural soils with hyperspectral sensor data. **Environmental Monitoring and Assessment**, v. 191 (7), p. 1–14, 2019. https://doi.org/10.1007/s10661-019-7510-4

TERRA, F. S.; DEMATTE, J. A. M.; VISCARRA ROSSEL, R. A. Proximal spectral sensing in pedological assessments: Vis-NIR spectra for soil classification based on weathering and pedogenesis. **Geoderma**, v. 318, p. 123–136, 2018. https://doi.org/10.1016/j.geoderma.2017.10.053

TERRA, F. S.; DEMATTÊ, J. A.; ROSSEL, R. A. V. Spectral libraries for quantitative analyses of tropical Brazilian soils: Comparing vis–NIR and mid-IR reflectance data. **Geoderma**, v. 255, p. 81-93, 2015. <u>https://doi.org/10.1016/j.geoderma.2015.04.017</u>

TODOROVA, M.; MOUAZEN, A. M.; LANGE, H.; ATANASSOVA, S. Potential of Near-Infrared Spectroscopy for Measurement of Heavy Metals in Soil as Affected by Calibration Set Size. **Water, Air, & Soil Pollution**, v. 225, n. 2036, 2014. https://doi.org/10.1007/s11270-014-2036-4

TIECHER, T.; MOURA-BUENO, J. M.; CANER, L.; MINELLA, J. P. G.; EVRARD, O.; RAMON, R.; NAIBO, G.; BARROS, C. A. P.; SILVA, Y. J. A. B.; AMORIM, F. F.; RHEINHEIMER, D.S. 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**, v. 384, p. 114815, 2021. https://doi.org/10.1016/j.geoderma.2020.114815.

USEPA - UNITED STATES ENVIRONMENTAL PROTECTION AGENCY. **Method 3051a** – Microwave assisted acid digestion of sediments, sludges, soils, and oils. 1998.

VISCARRA ROSSEL, R. A., WALVOORT, D. J. J., MCBRATNEY, A. B., JANIK, L. J.; SKJEMSTAD, J. O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. **Geoderma**, v. 131, p. 59–75, 2006. <u>https://doi.org/10.1016/j.geoderma.2005.03.007</u>

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. https://doi.org/10.1016/j.geoderma.2009.12.025

VISCARRA ROSSEL, R. A. ParLeS: Software for chemometric analysis of spectroscopic data. **Chemometrics and Intelligent Laboratory Systems**, v. 90, n. 1, 2008. <u>https://doi.org/10.1016/j.chemolab.2007.06.006</u>

VISCARRA ROSSEL, R. A.; FOUAD, Y.; WALTER, C. Using a digital camera to measure soil organic carbon and iron contents. **Biosystems Engineering**, v. 100, n. 2, p. 149–159, 2008. <u>https://doi.org/10.1016/j.biosystemseng.2008.02.007</u>

WANG, C.; ZHANG, T.; PAN, X. Potential of visible and near-infrared reflectance spectroscopy for the determination of rare earth elements in soil. **Geoderma**, v. 306, p. 120-126, 2017. <u>https://doi.org/10.1016/j.geoderma.2017.07.016</u>

WANG, J. J.; CUI, L.J.; GAO, W. X.; SHI, T. Z.; CHEN, Y. Y.; GAO, Y. Prediction of low heavy metal concentrations in agricultural soils using visible and near-infrared reflectance spectroscopy. **Geoderma**, v. 216, n. 4, p. 1–9, 2014. https://doi.org/10.1016/j.geoderma.2013.10.024

WANG, L.; LIANG, T. Anomalous abundance and redistribution patterns of rare earth elements in soils of a mining area in Inner Mongolia, China. **Environmental Science and Pollution Research**, v. 23(11), p. 11330–11338, 2016. https://doi.org/10.1007/s11356-016-6351-8

WONDRACZEK, L.; TYYSTJÄRVI, E.; MÉNDEZ-RAMOS, J.; MÜLLER, F. A.; ZHANG, Q. Shifting the sun: Solar spectral conversion and extrinsic sensitization in natural and artificial photosynthesis. **Advanced Science**, v. 2, p. 1500218, 2015. https://doi.org/10.1002/advs.201500218

WU, Y., CHEN, J., JUNFENG, J., GONG, P., LIAO, Q., TIAN, Q., MA, H. R. A mechanism study of reflectance spectroscopy for investigating heavy metals in soils. **Soil Science Society of American Journal**, v. 71, n. 3, p. 918–926, 2007. https://doi.org/10.2136/sssaj2006.0285

YU, B.; YAN, C.; YUAN, J.; DING, N.; CHEN, Z. Prediction of soil properties based on characteristic wavelengths with optimal spectral resolution by using Vis-NIR spectroscopy. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 293, p. 122452, 2023. <u>https://doi.org/10.1016/j.saa.2023.122452</u>

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. https://doi.org/10.1016/j.soilbio.2008.04.003

| Variável | PB | PE | RN | amostras | calibração | validação |
|-------------|----|----|----|----------|------------|-----------|
| argila | 89 | 35 | 91 | 215 | 172 | 43 |
| Ba | 87 | 35 | 82 | 204 | 163 | 41 |
| Cd | 72 | 30 | 79 | 181 | 144 | 37 |
| Ce | NA | 35 | 91 | 126 | 100 | 26 |
| Со | 76 | 34 | 91 | 201 | 160 | 41 |
| СОТ | 89 | 33 | 91 | 213 | 170 | 43 |
| Cr | 89 | 35 | 86 | 210 | 168 | 42 |
| CTC | 89 | NA | 91 | 180 | 144 | 36 |
| Cu | 85 | 34 | 85 | 204 | 163 | 41 |
| Dy | NA | 35 | 85 | 120 | 96 | 24 |
| Er | NA | 35 | 91 | 126 | 100 | 26 |
| Eu | NA | 34 | 91 | 125 | 100 | 25 |
| Fe | 89 | 35 | 91 | 215 | 172 | 43 |
| Gd | NA | 34 | 91 | 125 | 100 | 25 |
| Но | NA | 30 | 74 | 104 | 83 | 21 |
| ETRPs | NA | 35 | 91 | 126 | 100 | 26 |
| La | NA | 35 | 91 | 126 | 100 | 26 |
| ETRLs/ETRPs | NA | 35 | 91 | 126 | 100 | 26 |
| ETRLs | NA | 35 | 91 | 126 | 100 | 26 |
| Lu | NA | 22 | 80 | 102 | 81 | 21 |
| Mn | 89 | 34 | 91 | 214 | 171 | 43 |
| Nd | NA | 35 | 91 | 126 | 100 | 26 |
| Ni | 89 | 30 | 74 | 193 | 154 | 39 |
| Pb | 89 | 35 | 79 | 203 | 162 | 41 |
| pН | 89 | 35 | 91 | 215 | 172 | 43 |
| Pr | NA | 35 | 91 | 126 | 100 | 26 |
| Sb | 70 | NA | 59 | 129 | 103 | 26 |
| Sc | NA | 35 | NA | 35 | 28 | 7 |
| Se | 89 | NA | 91 | 180 | 144 | 36 |
| Sm | NA | 35 | 91 | 126 | 100 | 26 |
| Tb | NA | 35 | 91 | 126 | 100 | 26 |
| V | NA | NA | 91 | 91 | 72 | 19 |
| Y | NA | 35 | NA | 35 | 28 | 7 |
| Yb | NA | 34 | 91 | 125 | 100 | 25 |
| Zn | 89 | 35 | 91 | 215 | 172 | 43 |

ANEXO A – Número de amostras de cada estado utilizada na calibração e validação dos modelos

| Variável | Modelo | Preproc | R2 | conc. | MSE | | RMSE | bias | MSEc | RMSEc | RPD | RPIQ |
|----------|--------|---------|-------|-------|-----|----------|--------|--------|----------|--------|------|------|
| areia | RF | RAW | 0.39 | 0.51 | | 30170.64 | 173.70 | 29.89 | 29277.37 | 171.11 | 1.28 | 1.62 |
| areia | RF | SG | 0.48 | 0.53 | | 27282.39 | 165.17 | 4.15 | 27265.18 | 165.12 | 1.35 | 1.70 |
| areia | RF | CR | 0.40 | 0.47 | | 30362.08 | 174.25 | 13.40 | 30182.57 | 173.73 | 1.28 | 1.61 |
| areia | RF | DET | 0.24 | 0.37 | | 36393.92 | 190.77 | 14.13 | 36194.20 | 190.25 | 1.17 | 1.47 |
| areia | RF | SNV | 0.22 | 0.32 | | 37780.86 | 194.37 | 19.85 | 37386.96 | 193.36 | 1.15 | 1.45 |
| areia | RF | MSC | 0.33 | 0.28 | | 37070.03 | 192.54 | 0.67 | 37069.57 | 192.53 | 1.16 | 1.46 |
| argila | RF | RAW | 0.68 | 0.70 | | 9177.41 | 95.80 | -18.09 | 8850.21 | 94.08 | 1.64 | 2.02 |
| argila | RF | SG | 0.50 | 0.56 | | 12755.52 | 112.94 | -0.14 | 12755.50 | 112.94 | 1.39 | 1.72 |
| argila | RF | CR | 0.43 | 0.50 | | 14403.45 | 120.01 | -8.37 | 14333.44 | 119.72 | 1.31 | 1.62 |
| argila | RF | DET | 0.29 | 0.43 | | 16863.12 | 129.86 | -8.60 | 16789.13 | 129.57 | 1.21 | 1.49 |
| argila | RF | SNV | 0.17 | 0.30 | | 19459.89 | 139.50 | -7.25 | 19407.40 | 139.31 | 1.13 | 1.39 |
| argila | RF | MSC | -0.01 | -0.06 | | 27857.02 | 166.90 | 8.28 | 27788.41 | 166.70 | 0.94 | 1.16 |
| Ba | RF | RAW | 0.14 | 0.36 | | 5677.00 | 75.35 | 2.46 | 5670.93 | 75.31 | 1.07 | 0.79 |
| Ba | RF | SG | 0.45 | 0.49 | | 3754.58 | 61.27 | 2.46 | 3748.52 | 61.23 | 1.32 | 0.97 |
| Ba | RF | CR | 0.34 | 0.50 | | 4200.06 | 64.81 | 10.52 | 4089.40 | 63.95 | 1.24 | 0.91 |
| Ba | RF | DET | 0.47 | 0.56 | | 3440.83 | 58.66 | 2.75 | 3433.29 | 58.59 | 1.37 | 1.01 |
| Ba | RF | SNV | 0.49 | 0.60 | | 3323.89 | 57.65 | 8.45 | 3252.50 | 57.03 | 1.40 | 1.03 |
| Ba | RF | MSC | 0.34 | 0.30 | | 4849.66 | 69.64 | 12.86 | 4684.32 | 68.44 | 1.16 | 0.85 |
| Cd | RF | RAW | -0.01 | 0.13 | | 0.12 | 0.34 | 0.01 | 0.12 | 0.34 | 0.91 | 0.15 |
| Cd | RF | SG | -0.02 | 0.05 | | 0.11 | 0.33 | 0.03 | 0.11 | 0.33 | 0.94 | 0.15 |
| Cd | RF | CR | -0.02 | -0.08 | | 0.19 | 0.44 | 0.05 | 0.19 | 0.43 | 0.71 | 0.11 |
| Cd | RF | DET | -0.01 | -0.09 | | 0.12 | 0.35 | 0.02 | 0.12 | 0.35 | 0.89 | 0.14 |
| Cd | RF | SNV | -0.03 | 0.01 | | 0.11 | 0.33 | 0.02 | 0.11 | 0.33 | 0.95 | 0.15 |
| Cd | RF | MSC | -0.02 | 0.04 | | 0.10 | 0.32 | 0.06 | 0.10 | 0.31 | 0.97 | 0.16 |
| Ce | RF | RAW | 0.04 | 0.23 | | 1643.34 | 40.54 | 5.44 | 1613.72 | 40.17 | 0.98 | 1.19 |
| Ce | RF | SG | 0.18 | 0.33 | | 1223.56 | 34.98 | 5.54 | 1192.82 | 34.54 | 1.14 | 1.38 |
| Ce | RF | CR | 0.05 | 0.26 | | 1757.00 | 41.92 | 10.03 | 1656.33 | 40.70 | 0.95 | 1.15 |
| Ce | RF | DET | 0.16 | 0.31 | | 1259.14 | 35.48 | 5.80 | 1225.45 | 35.01 | 1.12 | 1.36 |
| Ce | RF | SNV | 0.11 | 0.26 | | 1304.01 | 36.11 | 2.97 | 1295.20 | 35.99 | 1.10 | 1.34 |
| Ce | RF | MSC | 0.45 | 0.45 | | 976.93 | 31.26 | 7.70 | 917.62 | 30.29 | 1.27 | 1.55 |
| Co | RF | RAW | 0.54 | 0.55 | | 49.08 | 7.01 | -1.16 | 47.75 | 6.91 | 1.38 | 1.50 |
| Co | RF | SG | 0.59 | 0.62 | | 43.04 | 6.56 | -1.19 | 41.62 | 6.45 | 1.47 | 1.60 |
| Co | RF | CR | 0.56 | 0.56 | | 48.20 | 6.94 | -1.16 | 46.86 | 6.85 | 1.39 | 1.51 |
| Co | RF | DET | 0.63 | 0.68 | | 38.15 | 6.18 | -1.19 | 36.74 | 6.06 | 1.57 | 1.70 |
| Co | RF | SNV | 0.29 | 0.44 | | 64.53 | 8.03 | -1.23 | 63.03 | 7.94 | 1.20 | 1.31 |
| Co | RF | MSC | 0.23 | 0.36 | | 69.70 | 8.35 | -0.88 | 68.94 | 8.30 | 1.16 | 1.26 |
| COT | RF | RAW | 0.00 | 0.13 | | 76.83 | 8.77 | 1.02 | 75.79 | 8.71 | 0.90 | 0.86 |
| COT | RF | SG | 0.13 | 0.30 | | 53.24 | 7.30 | 0.54 | 52.95 | 7.28 | 1.08 | 1.03 |
| COT | RF | CR | 0.07 | 0.21 | | 57.64 | 7.59 | 0.96 | 56.71 | 7.53 | 1.04 | 0.99 |
| COT | RF | DET | -0.02 | 0.05 | | 65.70 | 8.11 | 0.30 | 65.61 | 8.10 | 0.98 | 0.93 |
| COT | RF | SNV | 0.02 | 0.18 | | 72.96 | 8.54 | 1.75 | 69.89 | 8.36 | 0.93 | 0.88 |
| COT | RF | MSC | 0.01 | 0.14 | | 70.46 | 8.39 | 1.08 | 69.28 | 8.32 | 0.94 | 0.90 |

ANEXO B - Modelos, pré-processamentos e parâmetros da predição dos elementos químicos e atributos do solo

| Cr | RF | RAW | 0.30 | 0.54 | 478.30 | 21.87 | 4.70 | 456.22 | 21.36 | 1.08 | 0.91 |
|-----|----|-----|-------|-------|---------------|----------|----------|---------------|----------|------|------|
| Cr | RF | SG | 0.38 | 0.53 | 356.68 | 18.89 | 5.42 | 327.35 | 18.09 | 1.25 | 1.06 |
| Cr | RF | CR | 0.39 | 0.52 | 333.34 | 18.26 | 3.09 | 323.80 | 17.99 | 1.29 | 1.10 |
| Cr | RF | DET | 0.48 | 0.63 | 296.74 | 17.23 | 4.96 | 272.09 | 16.50 | 1.37 | 1.16 |
| Cr | RF | SNV | 0.26 | 0.44 | 408.85 | 20.22 | 4.33 | 390.11 | 19.75 | 1.16 | 0.99 |
| Cr | RF | MSC | 0.27 | 0.35 | 517.68 | 22.75 | 11.55 | 384.21 | 19.60 | 1.03 | 0.88 |
| CTC | RF | RAW | 0.84 | 0.78 | 25.84 | 5.08 | -0.04 | 25.84 | 5.08 | 1.97 | 1.16 |
| CTC | RF | SG | 0.92 | 0.83 | 19.59 | 4.43 | -0.74 | 19.04 | 4.36 | 2.26 | 1.34 |
| CTC | RF | CR | 0.82 | 0.77 | 27.02 | 5.20 | -0.28 | 26.94 | 5.19 | 1.92 | 1.14 |
| CTC | RF | DET | 0.76 | 0.76 | 29.80 | 5.46 | -0.54 | 29.50 | 5.43 | 1.83 | 1.08 |
| CTC | RF | SNV | 0.86 | 0.79 | 25.19 | 5.02 | -0.76 | 24.62 | 4.96 | 1.99 | 1.18 |
| CTC | RF | MSC | 0.83 | 0.69 | 35.85 | 5.99 | 1.74 | 32.83 | 5.73 | 1.67 | 0.99 |
| Cu | RF | RAW | 0.04 | 0.17 | 195.71 | 13.99 | 0.77 | 195.12 | 13.97 | 1.04 | 0.60 |
| Cu | RF | SG | 0.15 | 0.26 | 171.75 | 13.11 | 0.51 | 171.48 | 13.10 | 1.11 | 0.65 |
| Cu | RF | CR | 0.23 | 0.32 | 158.20 | 12.58 | 0.21 | 158.16 | 12.58 | 1.15 | 0.67 |
| Cu | RF | DET | 0.13 | 0.27 | 175.51 | 13.25 | 0.03 | 175.51 | 13.25 | 1.10 | 0.64 |
| Cu | RF | SNV | 0.04 | 0.13 | 193.59 | 13.91 | -0.60 | 193.23 | 13.90 | 1.04 | 0.61 |
| Cu | RF | MSC | 0.00 | 0.04 | 217.01 | 14.73 | 4.13 | 199.98 | 14.14 | 0.99 | 0.57 |
| Dy | RF | RAW | -0.01 | -0.14 | 0.86 | 0.93 | 0.04 | 0.86 | 0.93 | 0.86 | 1.30 |
| Dy | RF | SG | 0.28 | 0.44 | 0.42 | 0.65 | -0.07 | 0.42 | 0.65 | 1.22 | 1.85 |
| Dy | RF | CR | 0.18 | 0.37 | 0.49 | 0.70 | -0.10 | 0.48 | 0.69 | 1.13 | 1.72 |
| Dy | RF | DET | 0.29 | 0.45 | 0.44 | 0.66 | -0.16 | 0.41 | 0.64 | 1.20 | 1.82 |
| Dy | RF | SNV | 0.12 | 0.29 | 0.51 | 0.71 | -0.05 | 0.51 | 0.71 | 1.11 | 1.69 |
| Dy | RF | MSC | 0.17 | 0.38 | 0.49 | 0.70 | -0.03 | 0.49 | 0.70 | 1.13 | 1.72 |
| Er | RF | RAW | 0.61 | 0.61 | 0.19 | 0.44 | -0.07 | 0.19 | 0.43 | 1.50 | 2.74 |
| Er | RF | SG | 0.71 | 0.75 | 0.13 | 0.36 | 0.02 | 0.13 | 0.36 | 1.82 | 3.32 |
| Er | RF | CR | 0.71 | 0.76 | 0.13 | 0.36 | 0.00 | 0.13 | 0.36 | 1.84 | 3.36 |
| Er | RF | DET | 0.67 | 0.74 | 0.14 | 0.37 | 0.00 | 0.14 | 0.37 | 1.75 | 3.21 |
| Er | RF | SNV | 0.71 | 0.78 | 0.13 | 0.36 | 0.09 | 0.12 | 0.34 | 1.84 | 3.36 |
| Er | RF | MSC | 0.29 | 0.38 | 0.44 | 0.67 | 0.40 | 0.28 | 0.53 | 0.98 | 1.79 |
| Eu | RF | RAW | 0.70 | 0.35 | 1.24 | 1.12 | -0.26 | 1.18 | 1.09 | 1.22 | 0.85 |
| Eu | RF | SG | 0.18 | 0.20 | 1.55 | 1.25 | -0.26 | 1.49 | 1.22 | 1.09 | 0.76 |
| Eu | RF | CR | 0.15 | 0.19 | 1.59 | 1.26 | -0.27 | 1.51 | 1.23 | 1.08 | 0.75 |
| Eu | RF | DET | 0.19 | 0.18 | 1.58 | 1.26 | -0.27 | 1.50 | 1.23 | 1.09 | 0.76 |
| Eu | RF | SNV | 0.23 | 0.17 | 1.58 | 1.26 | -0.27 | 1.51 | 1.23 | 1.08 | 0.76 |
| Eu | RF | MSC | 0.16 | 0.17 | 1.60 | 1.27 | -0.27 | 1.53 | 1.24 | 1.08 | 0.75 |
| Fe | RF | RAW | 0.01 | 0.13 | 2532754540.00 | 50326.48 | 2887.50 | 2524416864.00 | 50243.58 | 0.98 | 0.57 |
| Fe | RF | SG | 0.12 | 0.25 | 2075848194.00 | 45561.48 | 3499.16 | 2063604089.00 | 45426.91 | 1.09 | 0.63 |
| Fe | RF | CR | 0.14 | 0.30 | 2034162857.00 | 45101.69 | -2268.00 | 2029019050.00 | 45044.63 | 1.10 | 0.64 |
| Fe | RF | DET | 0.45 | 0.54 | 1355424658.00 | 36816.09 | -1941.20 | 1351656393.00 | 36764.88 | 1.34 | 0.78 |
| Fe | RF | SNV | 0.25 | 0.36 | 1773031454.00 | 42107.38 | 2826.35 | 1765043211.00 | 42012.42 | 1.18 | 0.68 |
| Fe | RF | MSC | 0.04 | 0.12 | 4676519668.00 | 68385.08 | 47413.82 | 2428449248.00 | 49279.30 | 0.72 | 0.42 |
| Gd | RF | RAW | 0.02 | 0.20 | 5.19 | 2.28 | -0.29 | 5.11 | 2.26 | 0.98 | 1.62 |
| Gd | RF | SG | 0.26 | 0.37 | 3.70 | 1.92 | -0.52 | 3.43 | 1.85 | 1.16 | 1.92 |
| Gd | RF | CR | 0.43 | 0.51 | 2.92 | 1.71 | -0.44 | 2.72 | 1.65 | 1.31 | 2.17 |
| Gd | RF | DET | 0.44 | 0.58 | 2.66 | 1.63 | -0.26 | 2.59 | 1.61 | 1.37 | 2.27 |

| Gd | RF | SNV | 0.38 | 0.48 | 3.17 | 1.78 | -0.49 | 2.92 | 1.71 | 1.25 | 2.08 |
|-------------|----|-----|-------|------|-----------|--------|--------|----------|--------|------|------|
| Gd | RF | MSC | 0.36 | 0.46 | 3.21 | 1.79 | -0.44 | 3.01 | 1.74 | 1.24 | 2.07 |
| Ho | RF | RAW | 0.51 | 0.63 | 0.01 | 0.11 | 0.03 | 0.01 | 0.10 | 1.42 | 1.80 |
| Но | RF | SG | 0.30 | 0.48 | 0.02 | 0.14 | 0.06 | 0.01 | 0.12 | 1.11 | 1.40 |
| Но | RF | CR | 0.16 | 0.40 | 0.02 | 0.15 | 0.05 | 0.02 | 0.14 | 1.01 | 1.28 |
| Но | RF | DET | 0.16 | 0.39 | 0.02 | 0.15 | 0.04 | 0.02 | 0.14 | 1.03 | 1.31 |
| Но | RF | SNV | 0.18 | 0.39 | 0.02 | 0.14 | 0.04 | 0.02 | 0.13 | 1.07 | 1.35 |
| Но | RF | MSC | 0.17 | 0.19 | 0.03 | 0.18 | 0.12 | 0.02 | 0.13 | 0.85 | 1.08 |
| ETRPs | RF | RAW | 0.14 | 0.39 | 22.35 | 4.73 | 0.49 | 22.11 | 4.70 | 1.03 | 1.67 |
| ETRPs | RF | SG | 0.22 | 0.38 | 17.30 | 4.16 | 0.25 | 17.24 | 4.15 | 1.18 | 1.90 |
| ETRPs | RF | CR | 0.18 | 0.39 | 18.83 | 4.34 | 0.26 | 18.76 | 4.33 | 1.13 | 1.82 |
| ETRPs | RF | DET | 0.18 | 0.43 | 20.98 | 4.58 | 0.55 | 20.67 | 4.55 | 1.07 | 1.72 |
| ETRPs | RF | SNV | 0.05 | 0.25 | 22.84 | 4.78 | 0.37 | 22.71 | 4.77 | 1.02 | 1.65 |
| ETRPs | RF | MSC | 0.02 | 0.19 | 28.40 | 5.33 | 1.81 | 25.14 | 5.01 | 0.92 | 1.48 |
| La | RF | RAW | 0.25 | 0.33 | 383.08 | 19.57 | 0.73 | 382.55 | 19.56 | 1.18 | 0.75 |
| La | RF | SG | 0.24 | 0.30 | 391.51 | 19.79 | 0.15 | 391.49 | 19.79 | 1.17 | 0.75 |
| La | RF | CR | 0.18 | 0.26 | 410.34 | 20.26 | -0.19 | 410.30 | 20.26 | 1.14 | 0.73 |
| La | RF | DET | 0.15 | 0.21 | 430.64 | 20.75 | -0.08 | 430.63 | 20.75 | 1.11 | 0.71 |
| La | RF | SNV | 0.16 | 0.24 | 420.08 | 20.50 | 0.45 | 419.87 | 20.49 | 1.13 | 0.72 |
| La | RF | MSC | 0.46 | 0.48 | 305.92 | 17.49 | 2.27 | 300.78 | 17.34 | 1.32 | 0.84 |
| ETRLs/ETRPs | RF | RAW | 0.15 | 0.35 | 34.05 | 5.84 | -0.44 | 33.86 | 5.82 | 1.11 | 1.17 |
| ETRLs/ETRPs | RF | SG | 0.29 | 0.38 | 28.87 | 5.37 | -0.34 | 28.76 | 5.36 | 1.21 | 1.27 |
| ETRLs/ETRPs | RF | CR | 0.25 | 0.35 | 30.44 | 5.52 | 0.61 | 30.07 | 5.48 | 1.18 | 1.24 |
| ETRLs/ETRPs | RF | DET | 0.36 | 0.43 | 26.64 | 5.16 | -0.19 | 26.60 | 5.16 | 1.26 | 1.32 |
| ETRLs/ETRPs | RF | SNV | 0.26 | 0.43 | 28.70 | 5.36 | 0.00 | 28.70 | 5.36 | 1.21 | 1.28 |
| ETRLs/ETRPs | RF | MSC | 0.06 | 0.09 | 42.27 | 6.50 | -2.20 | 37.42 | 6.12 | 1.00 | 1.05 |
| ETRLSETRLS | RF | RAW | 0.04 | 0.25 | 10101.95 | 100.51 | 16.34 | 9834.99 | 99.17 | 0.96 | 0.93 |
| ETRLs | RF | SG | 0.29 | 0.38 | 6447.38 | 80.30 | 13.35 | 6269.18 | 79.18 | 1.20 | 1.16 |
| ETRLs | RF | CR | 0.10 | 0.29 | 8114.37 | 90.08 | 13.12 | 7942.16 | 89.12 | 1.07 | 1.04 |
| ETRLs | RF | DET | 0.22 | 0.31 | 6990.16 | 83.61 | 12.72 | 6828.26 | 82.63 | 1.15 | 1.12 |
| ETRLs | RF | SNV | 0.34 | 0.38 | 6295.96 | 79.35 | 14.30 | 6091.49 | 78.05 | 1.21 | 1.18 |
| ETRLs | RF | MSC | 0.09 | 0.24 | 9273.46 | 96.30 | 35.55 | 8009.82 | 89.50 | 1.00 | 0.97 |
| Lu | RF | RAW | -0.02 | 0.15 | 0.02 | 0.13 | 0.01 | 0.02 | 0.13 | 0.82 | 0.95 |
| Lu | RF | SG | 0.11 | 0.33 | 0.01 | 0.09 | 0.00 | 0.01 | 0.09 | 1.09 | 1.26 |
| Lu | RF | CR | 0.34 | 0.51 | 0.01 | 0.08 | 0.01 | 0.01 | 0.08 | 1.29 | 1.50 |
| Lu | RF | DET | 0.26 | 0.43 | 0.01 | 0.08 | 0.00 | 0.01 | 0.08 | 1.22 | 1.42 |
| Lu | RF | SNV | 0.16 | 0.34 | 0.01 | 0.09 | 0.00 | 0.01 | 0.09 | 1.14 | 1.33 |
| Lu | RF | MSC | 0.01 | 0.19 | 0.01 | 0.10 | 0.01 | 0.01 | 0.10 | 1.01 | 1.17 |
| Mn | RF | RAW | 0.03 | 0.18 | 65697.43 | 256.32 | 19.83 | 65304.17 | 255.55 | 0.99 | 1.10 |
| Mn | RF | SG | 0.33 | 0.51 | 40989.96 | 202.46 | -7.23 | 40937.69 | 202.33 | 1.25 | 1.39 |
| Mn | RF | CR | 0.34 | 0.50 | 40314.72 | 200.79 | -7.66 | 40256.04 | 200.64 | 1.26 | 1.41 |
| Mn | RF | DET | 0.35 | 0.53 | 39344.36 | 198.35 | -1.95 | 39340.56 | 198.34 | 1.27 | 1.42 |
| Mn | RF | SNV | 0.15 | 0.29 | 51799.63 | 227.60 | -12.40 | 51645.99 | 227.26 | 1.11 | 1.24 |
| Mn | RF | MSC | 0.11 | 0.23 | 132401.34 | 363.87 | 243.85 | 72937.56 | 270.07 | 0.69 | 0.78 |
| Nd | RF | RAW | 0.53 | 0.38 | 435.36 | 20.87 | -4.67 | 413.52 | 20.34 | 1.24 | 1.01 |
| Nd | RF | SG | 0.24 | 0.26 | 513.73 | 22.67 | -3.13 | 503.95 | 22.45 | 1.14 | 0.93 |

| 0 | | | | | | | | |
|--------|----|------|------|------|--------|-------|-------|--------|
| | | | | | | | | |
| Nd | RE | CR | 0.06 | 0.15 | 596 11 | 24.42 | -1 32 | 577 73 |
| Nd | DE | DET | 0.00 | 0.19 | 566 55 | 27.72 | 4.20 | 5/9 97 |
| Nd | | SNW | 0.14 | 0.10 | 562.00 | 23.80 | -4.20 | 542.25 |
| INU NI | КГ | SINV | 0.10 | 0.19 | 550.02 | 25.71 | -4.55 | 543.25 |
| Nd | RF | MSC | 0.17 | 0.17 | 550.93 | 23.47 | -1.56 | 548.50 |
| Ni | RF | RAW | 0.04 | 0.21 | 126.46 | 11.25 | 3.81 | 111.97 |
| Ni | RF | SG | 0.34 | 0.43 | 76.16 | 8.73 | 3.65 | 62.84 |
| Ni | RF | CR | 0.23 | 0.41 | 87.66 | 9.36 | 3.56 | 74.98 |
| Ni | RF | DET | 0.38 | 0.50 | 70.98 | 8.42 | 3.53 | 58.52 |
| Ni | RF | SNV | 0.24 | 0.33 | 96.34 | 9.82 | 4.95 | 71.84 |
| Ni | RF | MSC | 0.09 | 0.07 | 356.78 | 18.89 | 16.45 | 86.31 |
| Pb | RF | RAW | 0.11 | 0.25 | 47.61 | 6.90 | 1.59 | 45.08 |
| Pb | RF | SG | 0.35 | 0.45 | 34.01 | 5.83 | 0.62 | 33.63 |
| Pb | RF | CR | 0.42 | 0.43 | 33.36 | 5.78 | 0.71 | 32.86 |
| Pb | RF | DET | 0.21 | 0.28 | 40.69 | 6.38 | 0.54 | 40.40 |
| Pb | RF | SNV | 0.23 | 0.33 | 40.12 | 6.33 | 1.05 | 39.02 |
| Pb | RF | MSC | 0.25 | 0.27 | 49.46 | 7.03 | 3.26 | 38.86 |
| pН | RF | RAW | 0.19 | 0.38 | 0.45 | 0.67 | 0.06 | 0.44 |
| pН | RF | SG | 0.48 | 0.64 | 0.29 | 0.54 | 0.10 | 0.28 |
| pH | RF | CR | 0.44 | 0.60 | 0.31 | 0.56 | 0.11 | 0.30 |
| pH | RF | DET | 0.43 | 0.59 | 0.33 | 0.57 | 0.14 | 0.31 |
| pH | RF | SNV | 0.33 | 0.53 | 0.39 | 0.62 | 0.16 | 0.36 |
| pН | RF | MSC | 0.27 | 0.46 | 0.44 | 0.66 | 0.20 | 0.40 |
| Pr | RF | RAW | 0.06 | 0.25 | 54.88 | 7.41 | 3.08 | 45.40 |

24.04 1.06

23.43 1.08

1.09

1.10

0.89

1.14

1.07

1.19

1.02

0.53

1.05

1.24

1.26

1.14

1.14

1.03

23.31

23.42

10.58

7.93

8.66

7.65

8.48

9.29

6.71

5.80

5.73

6.36

6.25

6.23

0.86

0.89

0.89

0.90

1.08

1.39

1.30

1.44

1.24

0.64

1.11

1.31

1.32

1.20

1.21

1.09 1.15 1.43 1.38 1.35 1.23 1.16

1.14

1.87

1.44

1.51 1.45

1.47

1.01 0.85

0.81 0.77

0.99

0.28 2.17

1.68 1.58

1.45

1.36 1.47

0.94

0.88 0.91

0.87

0.92

0.93

| RF | RAW | 0.19 | 0.38 | 0.45 | 0.67 | 0.06 | 0.44 | 0.67 | 1.12 |
|----|-----|-------|-------|-------|------|-------|-------|------|------|
| RF | SG | 0.48 | 0.64 | 0.29 | 0.54 | 0.10 | 0.28 | 0.53 | 1.39 |
| RF | CR | 0.44 | 0.60 | 0.31 | 0.56 | 0.11 | 0.30 | 0.55 | 1.35 |
| RF | DET | 0.43 | 0.59 | 0.33 | 0.57 | 0.14 | 0.31 | 0.56 | 1.31 |
| RF | SNV | 0.33 | 0.53 | 0.39 | 0.62 | 0.16 | 0.36 | 0.60 | 1.20 |
| RF | MSC | 0.27 | 0.46 | 0.44 | 0.66 | 0.20 | 0.40 | 0.63 | 1.13 |
| RF | RAW | 0.06 | 0.25 | 54.88 | 7.41 | 3.08 | 45.40 | 6.74 | 0.66 |
| RF | SG | 0.36 | 0.48 | 20.44 | 4.52 | 2.46 | 14.41 | 3.80 | 1.09 |
| RF | CR | 0.21 | 0.38 | 34.64 | 5.89 | 3.40 | 23.11 | 4.81 | 0.84 |
| RF | DET | 0.21 | 0.35 | 31.61 | 5.62 | 3.40 | 20.07 | 4.48 | 0.88 |
| RF | SNV | 0.05 | 0.24 | 34.13 | 5.84 | 2.67 | 27.00 | 5.20 | 0.84 |
| RF | MSC | 0.02 | -0.08 | 33.13 | 5.76 | 2.51 | 26.82 | 5.18 | 0.86 |
| RF | RAW | 0.60 | 0.65 | 0.04 | 0.21 | 0.06 | 0.04 | 0.20 | 1.51 |
| RF | SG | 0.42 | 0.52 | 0.06 | 0.25 | 0.09 | 0.05 | 0.23 | 1.27 |
| RF | CR | 0.37 | 0.47 | 0.07 | 0.26 | 0.09 | 0.06 | 0.24 | 1.22 |
| RF | DET | 0.33 | 0.50 | 0.07 | 0.27 | 0.11 | 0.06 | 0.25 | 1.15 |
| RF | SNV | 0.60 | 0.64 | 0.05 | 0.21 | 0.07 | 0.04 | 0.20 | 1.48 |
| RF | MSC | 0.17 | 0.05 | 0.55 | 0.74 | 0.69 | 0.08 | 0.28 | 0.42 |
| RF | RAW | 0.21 | 0.44 | 2.41 | 1.55 | -0.33 | 2.30 | 1.52 | 1.30 |
| RF | SG | -0.20 | -0.01 | 4.03 | 2.01 | -0.57 | 3.71 | 1.93 | 1.00 |
| RF | CR | -0.09 | -0.11 | 4.52 | 2.13 | -0.56 | 4.20 | 2.05 | 0.95 |
| RF | DET | -0.04 | -0.20 | 5.42 | 2.33 | -0.63 | 5.02 | 2.24 | 0.87 |
| RF | SNV | 0.42 | -0.41 | 6.17 | 2.48 | -0.38 | 6.02 | 2.45 | 0.81 |
| RF | MSC | -0.02 | -0.23 | 5.23 | 2.29 | 0.06 | 5.22 | 2.29 | 0.88 |
| RF | RAW | 0.11 | 0.24 | 0.08 | 0.28 | -0.08 | 0.07 | 0.27 | 1.05 |
| RF | SG | 0.01 | 0.09 | 0.09 | 0.30 | -0.09 | 0.08 | 0.29 | 0.99 |
| RF | CR | 0.06 | 0.12 | 0.08 | 0.29 | -0.08 | 0.08 | 0.28 | 1.02 |

0.09

0.08

0.08

0.30

0.29

0.29

-0.09

-0.07

-0.03

0.08

0.08

0.08

0.29

0.28

0.28

0.97

1.03

1.03

Pr

Pr

Pr

Pr Pr

Sb

Sb Sb

Sb

Sb Sb

 \mathbf{Sc} Sc

Sc

 \mathbf{Sc} Sc

Sc

Se Se

Se

Se

Se

Se

RF

RF

RF

DET

SNV

MSC

-0.01

0.06

0.02

0.05

0.18

0.08

| silte | RF | RAW | 0.08 | 0.21 | 12351.57 | 111.14 | -0.19 | 12351.54 | 111.14 | 1.06 | 1.20 |
|-------|----|-----|-------|-------|----------|--------|-------|----------|--------|------|------|
| silte | RF | SG | 0.35 | 0.49 | 8681.47 | 93.17 | 0.71 | 8680.96 | 93.17 | 1.27 | 1.43 |
| silte | RF | CR | 0.19 | 0.31 | 10735.02 | 103.61 | 1.08 | 10733.85 | 103.60 | 1.14 | 1.29 |
| silte | RF | DET | 0.21 | 0.35 | 10485.13 | 102.40 | -0.81 | 10484.47 | 102.39 | 1.15 | 1.30 |
| silte | RF | SNV | 0.27 | 0.36 | 9969.22 | 99.85 | -8.47 | 9897.53 | 99.49 | 1.18 | 1.34 |
| silte | RF | MSC | 0.20 | 0.32 | 10698.44 | 103.43 | -7.51 | 10642.06 | 103.16 | 1.14 | 1.29 |
| Sm | RF | RAW | 0.39 | 0.52 | 2.50 | 1.58 | 0.30 | 2.41 | 1.55 | 1.30 | 2.07 |
| Sm | RF | SG | 0.25 | 0.41 | 3.28 | 1.81 | 0.56 | 2.96 | 1.72 | 1.14 | 1.81 |
| Sm | RF | CR | 0.38 | 0.48 | 3.20 | 1.79 | 0.86 | 2.46 | 1.57 | 1.15 | 1.83 |
| Sm | RF | DET | 0.13 | 0.32 | 3.83 | 1.96 | 0.49 | 3.59 | 1.89 | 1.05 | 1.67 |
| Sm | RF | SNV | -0.01 | 0.16 | 4.98 | 2.23 | 0.42 | 4.81 | 2.19 | 0.92 | 1.46 |
| Sm | RF | MSC | 0.00 | 0.17 | 5.03 | 2.24 | 0.57 | 4.70 | 2.17 | 0.92 | 1.46 |
| Tb | RF | RAW | 0.27 | 0.48 | 0.14 | 0.38 | 0.15 | 0.12 | 0.34 | 1.04 | 1.34 |
| Tb | RF | SG | 0.45 | 0.52 | 0.12 | 0.34 | 0.20 | 0.08 | 0.28 | 1.14 | 1.48 |
| Tb | RF | CR | 0.30 | 0.49 | 0.13 | 0.37 | 0.16 | 0.11 | 0.33 | 1.07 | 1.38 |
| Tb | RF | DET | 0.13 | 0.35 | 0.18 | 0.42 | 0.15 | 0.15 | 0.39 | 0.93 | 1.20 |
| Tb | RF | SNV | 0.17 | 0.36 | 0.14 | 0.38 | 0.14 | 0.12 | 0.35 | 1.04 | 1.34 |
| Tb | RF | MSC | -0.01 | 0.11 | 0.24 | 0.49 | 0.25 | 0.17 | 0.42 | 0.80 | 1.04 |
| Th | RF | RAW | -0.08 | 0.01 | 55.93 | 7.48 | -2.36 | 50.37 | 7.10 | 0.85 | 0.93 |
| Th | RF | SG | 0.42 | 0.58 | 24.19 | 4.92 | -1.93 | 20.47 | 4.52 | 1.29 | 1.42 |
| Th | RF | CR | -0.04 | 0.16 | 44.85 | 6.70 | -1.40 | 42.88 | 6.55 | 0.95 | 1.04 |
| Th | RF | DET | 0.08 | 0.28 | 34.91 | 5.91 | -1.29 | 33.26 | 5.77 | 1.08 | 1.18 |
| Th | RF | SNV | -0.06 | 0.11 | 53.91 | 7.34 | -2.11 | 49.45 | 7.03 | 0.87 | 0.95 |
| Th | RF | MSC | -0.02 | 0.17 | 39.35 | 6.27 | 0.30 | 39.26 | 6.27 | 1.02 | 1.11 |
| V | RF | RAW | 0.44 | 0.52 | 268.76 | 16.39 | -2.71 | 261.40 | 16.17 | 1.35 | 1.59 |
| V | RF | SG | 0.06 | 0.26 | 431.77 | 20.78 | -2.83 | 423.78 | 20.59 | 1.07 | 1.26 |
| V | RF | CR | 0.28 | 0.43 | 329.26 | 18.15 | -3.16 | 319.27 | 17.87 | 1.22 | 1.44 |
| V | RF | DET | 0.24 | 0.45 | 342.26 | 18.50 | -1.42 | 340.26 | 18.45 | 1.20 | 1.41 |
| V | RF | SNV | 0.08 | 0.33 | 477.22 | 21.85 | -0.15 | 477.20 | 21.84 | 1.02 | 1.20 |
| V | RF | MSC | -0.02 | 0.14 | 494.04 | 22.23 | 0.87 | 493.28 | 22.21 | 1.00 | 1.18 |
| Y | RF | RAW | -0.11 | -0.12 | 46.50 | 6.82 | -2.79 | 38.72 | 6.22 | 0.86 | 0.81 |
| Y | RF | SG | 0.10 | -0.16 | 41.90 | 6.47 | -2.05 | 37.72 | 6.14 | 0.91 | 0.86 |
| Y | RF | CR | -0.19 | -0.04 | 34.92 | 5.91 | -0.81 | 34.26 | 5.85 | 1.00 | 0.94 |
| Y | RF | DET | 0.09 | -0.19 | 45.42 | 6.74 | -2.34 | 39.96 | 6.32 | 0.87 | 0.82 |
| Y | RF | SNV | 0.27 | -0.38 | 54.40 | 7.38 | -1.43 | 52.35 | 7.23 | 0.80 | 0.75 |
| Y | RF | MSC | -0.13 | -0.11 | 40.49 | 6.36 | 1.80 | 37.24 | 6.10 | 0.93 | 0.87 |
| Yb | RF | RAW | 0.18 | 0.42 | 0.25 | 0.50 | 0.16 | 0.23 | 0.48 | 0.89 | 1.73 |
| Yb | RF | SG | 0.58 | 0.71 | 0.09 | 0.29 | 0.09 | 0.08 | 0.28 | 1.52 | 2.95 |
| Yb | RF | CR | 0.42 | 0.62 | 0.12 | 0.35 | 0.08 | 0.12 | 0.34 | 1.27 | 2.47 |
| Yb | RF | DET | 0.23 | 0.45 | 0.15 | 0.38 | 0.01 | 0.15 | 0.38 | 1.17 | 2.27 |
| Yb | RF | SNV | 0.36 | 0.57 | 0.13 | 0.37 | 0.09 | 0.13 | 0.36 | 1.22 | 2.37 |
| Yb | RF | MSC | 0.25 | 0.48 | 0.17 | 0.41 | 0.11 | 0.16 | 0.40 | 1.09 | 2.12 |
| Zn | RF | RAW | 0.20 | 0.41 | 131.12 | 11.45 | 1.90 | 127.52 | 11.29 | 1.10 | 1.34 |
| Zn | RF | SG | 0.18 | 0.39 | 151.77 | 12.32 | 3.63 | 138.57 | 11.77 | 1.02 | 1.25 |
| Zn | RF | CR | 0.36 | 0.55 | 102.31 | 10.11 | 2.07 | 98.01 | 9.90 | 1.25 | 1.52 |
| Zn | RF | DET | 0.37 | 0.54 | 98.42 | 9.92 | 1.53 | 96.06 | 9.80 | 1.27 | 1.55 |
| Zn | RF | SNV | 0.38 | 0.53 | 95.36 | 9.77 | 0.79 | 94.73 | 9.73 | 1.29 | 1.57 |
|--------|------|-----|-------|-------|----------|--------|--------|----------|--------|------|------|
| Zn | RF | MSC | 0.23 | 0.24 | 135.46 | 11.64 | 3.45 | 123.60 | 11.12 | 1.08 | 1.32 |
| areia | PLSR | RAW | 0.33 | 0.50 | 31993.82 | 178.87 | 18.08 | 31667.02 | 177.95 | 1.25 | 1.57 |
| areia | PLSR | SG | 0.33 | 0.42 | 32867.21 | 181.29 | 2.50 | 32860.94 | 181.28 | 1.23 | 1.55 |
| areia | PLSR | CR | 0.28 | 0.51 | 36582.46 | 191.27 | -4.64 | 36560.90 | 191.21 | 1.16 | 1.47 |
| areia | PLSR | DET | 0.25 | 0.39 | 35916.89 | 189.52 | 21.00 | 35476.09 | 188.35 | 1.18 | 1.48 |
| areia | PLSR | SNV | 0.30 | 0.48 | 33621.95 | 183.36 | 23.03 | 33091.61 | 181.91 | 1.21 | 1.53 |
| areia | PLSR | MSC | 0.10 | 0.14 | 45655.90 | 213.67 | 47.23 | 43425.06 | 208.39 | 1.04 | 1.32 |
| argila | PLSR | RAW | 0.55 | 0.64 | 11247.58 | 106.05 | -11.74 | 11109.82 | 105.40 | 1.48 | 1.83 |
| argila | PLSR | SG | 0.37 | 0.44 | 15777.77 | 125.61 | -2.90 | 15769.35 | 125.58 | 1.25 | 1.54 |
| argila | PLSR | CR | 0.38 | 0.59 | 14850.40 | 121.86 | -1.70 | 14847.51 | 121.85 | 1.29 | 1.59 |
| argila | PLSR | DET | 0.31 | 0.40 | 16901.91 | 130.01 | -15.46 | 16662.78 | 129.08 | 1.21 | 1.49 |
| argila | PLSR | SNV | 0.29 | 0.41 | 17129.37 | 130.88 | -14.62 | 16915.65 | 130.06 | 1.20 | 1.48 |
| argila | PLSR | MSC | 0.09 | 0.19 | 22575.84 | 150.25 | -33.53 | 21451.33 | 146.46 | 1.04 | 1.29 |
| Ba | PLSR | RAW | 0.18 | 0.37 | 5106.89 | 71.46 | 1.61 | 5104.30 | 71.44 | 1.13 | 0.83 |
| Ba | PLSR | SG | 0.30 | 0.43 | 4394.89 | 66.29 | -3.80 | 4380.45 | 66.18 | 1.22 | 0.89 |
| Ba | PLSR | CR | 0.41 | 0.63 | 4081.26 | 63.88 | 7.32 | 4027.68 | 63.46 | 1.26 | 0.93 |
| Ba | PLSR | DET | 0.31 | 0.54 | 4726.33 | 68.75 | 8.43 | 4655.19 | 68.23 | 1.17 | 0.86 |
| Ba | PLSR | SNV | 0.34 | 0.58 | 5003.92 | 70.74 | 8.89 | 4924.91 | 70.18 | 1.14 | 0.84 |
| Ba | PLSR | MSC | 0.23 | 0.14 | 5545.97 | 74.47 | -4.31 | 5527.40 | 74.35 | 1.08 | 0.80 |
| Cd | PLSR | RAW | 0.03 | 0.17 | 0.09 | 0.30 | 0.01 | 0.09 | 0.30 | 1.02 | 0.16 |
| Cd | PLSR | SG | -0.02 | 0.03 | 0.10 | 0.31 | 0.01 | 0.10 | 0.31 | 0.99 | 0.16 |
| Cd | PLSR | CR | 0.04 | 0.08 | 0.09 | 0.30 | -0.01 | 0.09 | 0.30 | 1.04 | 0.17 |
| Cd | PLSR | DET | -0.02 | 0.04 | 0.10 | 0.32 | 0.00 | 0.10 | 0.32 | 0.98 | 0.16 |
| Cd | PLSR | SNV | -0.03 | -0.03 | 0.15 | 0.39 | 0.04 | 0.15 | 0.39 | 0.80 | 0.13 |
| Cd | PLSR | MSC | -0.01 | 0.06 | 0.09 | 0.31 | -0.01 | 0.09 | 0.31 | 1.02 | 0.16 |
| Ce | PLSR | RAW | -0.02 | 0.07 | 1503.57 | 38.78 | -0.39 | 1503.42 | 38.77 | 1.03 | 1.25 |
| Ce | PLSR | SG | 0.17 | 0.26 | 1231.50 | 35.09 | 1.29 | 1229.83 | 35.07 | 1.13 | 1.38 |
| Ce | PLSR | CR | 0.11 | 0.23 | 1301.39 | 36.07 | 2.11 | 1296.93 | 36.01 | 1.10 | 1.34 |
| Ce | PLSR | DET | 0.00 | 0.18 | 1796.09 | 42.38 | 2.72 | 1788.69 | 42.29 | 0.94 | 1.14 |
| Ce | PLSR | SNV | -0.03 | 0.06 | 1620.39 | 40.25 | 2.07 | 1616.12 | 40.20 | 0.99 | 1.20 |
| Ce | PLSR | MSC | -0.03 | 0.04 | 1572.81 | 39.66 | 0.73 | 1572.28 | 39.65 | 1.00 | 1.22 |
| Co | PLSR | RAW | 0.20 | 0.30 | 76.27 | 8.73 | -2.11 | 71.82 | 8.47 | 1.11 | 1.20 |
| Co | PLSR | SG | 0.23 | 0.36 | 71.68 | 8.47 | -1.84 | 68.27 | 8.26 | 1.14 | 1.24 |
| Co | PLSR | CR | 0.60 | 0.67 | 40.07 | 6.33 | -1.60 | 37.51 | 6.12 | 1.53 | 1.66 |
| Co | PLSR | DET | 0.55 | 0.62 | 47.06 | 6.86 | -2.17 | 42.33 | 6.51 | 1.41 | 1.53 |
| Co | PLSR | SNV | 0.48 | 0.56 | 51.87 | 7.20 | -1.93 | 48.16 | 6.94 | 1.34 | 1.46 |
| Co | PLSR | MSC | 0.59 | 0.75 | 45.06 | 6.71 | -0.92 | 44.22 | 6.65 | 1.44 | 1.57 |
| COT | PLSR | RAW | 0.06 | 0.24 | 59.11 | 7.69 | -0.12 | 59.10 | 7.69 | 1.03 | 0.98 |
| COT | PLSR | SG | 0.10 | 0.24 | 55.08 | 7.42 | 0.82 | 54.42 | 7.38 | 1.07 | 1.02 |
| COT | PLSR | CR | 0.11 | 0.24 | 53.04 | 7.28 | 0.05 | 53.04 | 7.28 | 1.09 | 1.04 |
| СОТ | PLSR | DET | -0.02 | 0.04 | 64.06 | 8.00 | 0.33 | 63.95 | 8.00 | 0.99 | 0.94 |
| COT | PLSR | SNV | -0.02 | -0.02 | 65.63 | 8.10 | 0.22 | 65.59 | 8.10 | 0.98 | 0.93 |
| COT | PLSR | MSC | -0.02 | -0.02 | 64.37 | 8.02 | 0.09 | 64.37 | 8.02 | 0.99 | 0.94 |
| Cr | PLSR | RAW | 0.16 | 0.40 | 525.10 | 22.91 | 2.49 | 518.90 | 22.78 | 1.03 | 0.87 |
| Cr | PLSR | SG | 0.26 | 0.47 | 501.52 | 22.39 | 8.00 | 437.46 | 20.92 | 1.05 | 0.89 |

| C. | | CD | 0.20 | 0.50 | 450.00 | 21.29 | C 11 | 415 57 | 20.20 | 1 1 1 | 0.04 |
|-----|------|-----------|-------|------|----------------|-----------|--------------|----------------|-----------|-------|------|
| Cr | PLSK | UK DET | 0.30 | 0.58 | 452.88 | 21.28 | 0.11 5.22 | 415.57 | 10.39 | 1.11 | 1.05 |
| Cr | PLSK | DEI | 0.37 | 0.55 | 567.72 | 19.15 | 5.32 | 520.78 | 18.57 | 1.23 | 1.05 |
| Cr. | PLSK | SINV | 0.04 | 0.19 | 266.51 | 25.65 | 3.29 | 339.78 | 19.40 | 1.22 | 1.04 |
| CTC | PLSK | MSC | 0.58 | 0.59 | 300.51 | 19.14 | 4.95 | 341.98 | 18.49 | 1.23 | 1.04 |
| CTC | PLSK | KAW | 0.60 | 0.65 | 45.22 | 0.72 | -1.20 | 43.78 | 0.02 | 1.49 | 0.88 |
| CTC | PLSK | SU | 0.50 | 0.59 | 48.02 | 0.97 | -0.98 | 47.00 | 6.90 | 1.45 | 0.85 |
| CTC | PLSK | UK DET | 0.01 | 0.57 | 48.25 | 0.95 | -0.64 | 47.84 | 0.92 | 1.44 | 0.85 |
| CTC | PLSK | DEI | 0.84 | 0.87 | 16.59 | 4.07 | -0.45 | 10.39 | 4.05 | 2.45 | 1.45 |
| CIC | PLSK | SNV | 0.89 | 0.88 | 14.19 | 3.// | -0.54 | 13.90 | 3./3 | 2.65 | 1.57 |
| cic | PLSK | MSC | 0.84 | 0.67 | 180.93 | 13.45 | 2.24 | 175.93 | 13.26 | 0.74 | 0.44 |
| Cu | PLSR | RAW | 0.14 | 0.24 | 173.45 | 13.17 | 0.04 | 173.45 | 13.17 | 1.10 | 0.64 |
| Cu | PLSR | SG | 0.14 | 0.28 | 173.51 | 13.17 | 0.59 | 173.17 | 13.16 | 1.10 | 0.64 |
| Cu | PLSR | CR | 0.24 | 0.40 | 153.57 | 12.39 | 0.59 | 153.21 | 12.38 | 1.17 | 0.68 |
| Cu | PLSR | DET | 0.21 | 0.27 | 164.12 | 12.81 | -0.72 | 163.59 | 12.79 | 1.13 | 0.66 |
| Cu | PLSR | SNV | 0.19 | 0.28 | 164.50 | 12.83 | -0.68 | 164.03 | 12.81 | 1.13 | 0.66 |
| Cu | PLSR | MSC | 0.07 | 0.19 | 188.55 | 13.73 | 0.09 | 188.54 | 13.73 | 1.06 | 0.62 |
| Dy | PLSR | RAW | -0.04 | 0.06 | 0.71 | 0.84 | 0.04 | 0.71 | 0.84 | 0.94 | 1.43 |
| Dy | PLSR | SG | 0.14 | 0.30 | 0.50 | 0.71 | 0.00 | 0.50 | 0.71 | 1.12 | 1.70 |
| Dy | PLSR | CR | 0.09 | 0.25 | 0.53 | 0.73 | 0.00 | 0.53 | 0.73 | 1.09 | 1.65 |
| Dy | PLSR | DET | 0.07 | 0.22 | 0.55 | 0.74 | -0.04 | 0.54 | 0.74 | 1.07 | 1.63 |
| Dy | PLSR | SNV | -0.04 | 0.02 | 0.66 | 0.81 | 0.02 | 0.66 | 0.81 | 0.98 | 1.49 |
| Dy | PLSR | MSC | -0.04 | 0.02 | 0.63 | 0.79 | 0.01 | 0.63 | 0.79 | 1.00 | 1.52 |
| Er | PLSR | RAW | 0.36 | 0.58 | 0.28 | 0.52 | -0.09 | 0.27 | 0.52 | 1.25 | 2.28 |
| Er | PLSR | SG | 0.37 | 0.59 | 0.28 | 0.53 | 0.06 | 0.27 | 0.52 | 1.24 | 2.27 |
| Er | PLSR | CR | 0.36 | 0.58 | 0.26 | 0.51 | 0.00 | 0.26 | 0.51 | 1.28 | 2.34 |
| Er | PLSR | DET | 0.30 | 0.54 | 0.30 | 0.55 | -0.03 | 0.30 | 0.55 | 1.19 | 2.17 |
| Er | PLSR | SNV | 0.28 | 0.52 | 0.46 | 0.68 | 0.02 | 0.46 | 0.68 | 0.96 | 1.76 |
| Er | PLSR | MSC | 0.24 | 0.49 | 0.32 | 0.57 | -0.01 | 0.32 | 0.57 | 1.15 | 2.10 |
| Eu | PLSR | RAW | 0.09 | 0.15 | 1.68 | 1.30 | -0.31 | 1.58 | 1.26 | 1.05 | 0.73 |
| Eu | PLSR | SG | 0.06 | 0.11 | 1.74 | 1.32 | -0.33 | 1.63 | 1.28 | 1.03 | 0.72 |
| Eu | PLSR | CR | 0.04 | 0.12 | 1.75 | 1.32 | -0.33 | 1.65 | 1.28 | 1.03 | 0.72 |
| Eu | PLSR | DET | 0.02 | 0.07 | 1.79 | 1.34 | -0.32 | 1.69 | 1.30 | 1.02 | 0.71 |
| Eu | PLSR | SNV | 0.21 | 0.16 | 1.62 | 1.27 | -0.30 | 1.53 | 1.24 | 1.07 | 0.75 |
| Eu | PLSR | MSC | 0.62 | 0.26 | 1.40 | 1.18 | -0.29 | 1.32 | 1.15 | 1.15 | 0.80 |
| Fe | PLSR | RAW | -0.01 | 0.07 | 2438649571.00 | 49382.68 | 176.51 | 2438618417.00 | 49382.37 | 1.00 | 0.58 |
| Fe | PLSR | SG | -0.02 | 0.04 | 2473514503.00 | 49734.44 | 1745.70 | 2470467043.00 | 49703.79 | 0.99 | 0.58 |
| Fe | PLSR | CR | 0.05 | 0.16 | 2293577817.00 | 47891.31 | -7849.10 | 2231969397.00 | 47243.72 | 1.03 | 0.60 |
| Fe | PLSR | DET | -0.01 | 0.06 | 2382614520.00 | 48812.03 | 1362.93 | 2380756949.00 | 48793.00 | 1.01 | 0.59 |
| Fe | PLSR | SNV | 0.00 | 0.06 | 2334061436.00 | 48312.13 | 1086.08 | 2332881855.00 | 48299.92 | 1.02 | 0.59 |
| Fe | PLSR | MSC | 0.00 | 0.04 | 95232353196.00 | 308597.40 | -59944.28 | 91639036744.00 | 302719.40 | 0.16 | 0.09 |
| Gd | PLSR | RAW | 0.15 | 0.38 | 4.53 | 2.13 | -0.41 | 4.36 | 2.09 | 1.05 | 1.74 |
| Gd | PLSR | SG | 0.25 | 0.45 | 3.56 | 1.89 | -0.31 | 3.46 | 1.86 | 1.18 | 1.96 |
| Gd | PLSR | CR | 0.18 | 0.37 | 3.82 | 1.95 | -0.27 | 3.75 | 1.94 | 1.14 | 1.89 |
| Gd | PLSR | DET | 0.19 | 0.33 | 3.82 | 1.96 | -0.32 | 3.72 | 1.93 | 1.14 | 1.89 |
| Gd | PLSR | SNV | 0.09 | 0.21 | 4.36 | 2.09 | -0.42 | 4.18 | 2.05 | 1.07 | 1.77 |
| Gd | PLSR | MSC | 0.03 | 0.09 | 4.63 | 2.15 | -0.42 | 4.46 | 2.11 | 1.04 | 1.72 |

| 4 | | | | | | | |
|-------------|-------|-----|-------|-------|--------|-------|-------|
| | | | | | | | |
| | DI GD | DAW | 0.01 | 0.41 | 0.02 | 0.14 | 0.05 |
| Но | PLSR | RAW | 0.21 | 0.41 | 0.02 | 0.14 | 0.05 |
| Но | PLSR | SG | 0.23 | 0.47 | 0.02 | 0.14 | 0.04 |
| Но | PLSR | CR | 0.11 | 0.31 | 0.02 | 0.15 | 0.05 |
| Но | PLSR | DET | 0.08 | 0.23 | 0.02 | 0.14 | 0.04 |
| Но | PLSR | SNV | 0.09 | 0.21 | 0.02 | 0.15 | 0.06 |
| Но | PLSR | MSC | 0.09 | 0.17 | 0.02 | 0.15 | 0.05 |
| ETRPs | PLSR | RAW | 0.19 | 0.42 | 18.79 | 4.33 | -0.09 |
| ETRPs | PLSR | SG | 0.06 | 0.23 | 22.28 | 4.72 | 0.72 |
| ETRPs | PLSR | CR | 0.14 | 0.28 | 19.03 | 4.36 | 0.03 |
| ETRPs | PLSR | DET | 0.06 | 0.23 | 21.30 | 4.62 | 0.40 |
| ETRPs | PLSR | SNV | 0.08 | 0.22 | 20.71 | 4.55 | -0.50 |
| ETRPs | PLSR | MSC | 0.07 | 0.08 | 21.44 | 4.63 | 0.19 |
| La | PLSR | RAW | 0.03 | 0.18 | 494.68 | 22.24 | -1.24 |
| La | PLSR | SG | -0.02 | 0.09 | 519.74 | 22.80 | -0.20 |
| La | PLSR | CR | -0.03 | -0.10 | 767.12 | 27.70 | -3.00 |
| La | PLSR | DET | -0.04 | 0.00 | 529.28 | 23.01 | -0.49 |
| La | PLSR | SNV | 0.05 | -0.15 | 645.57 | 25.41 | -1.03 |
| La | PLSR | MSC | 0.51 | 0.55 | 272.30 | 16.50 | 2.30 |
| ETRLs/ETRPs | PLSR | RAW | 0.44 | 0.56 | 22.24 | 4.72 | 0.12 |
| | DLOD | | 0.00 | 0.44 | 21.55 | F 2 | 0.07 |

| ETRPs | PLSR | MSC | 0.07 | 0.08 | 21.44 | 4.63 | 0.19 | 21.40 | 4.63 | 1.06 | 1.70 |
|-------------|------|-----|-------|-------|----------|--------|--------|----------|--------|------|------|
| La | PLSR | RAW | 0.03 | 0.18 | 494.68 | 22.24 | -1.24 | 493.13 | 22.21 | 1.04 | 0.66 |
| La | PLSR | SG | -0.02 | 0.09 | 519.74 | 22.80 | -0.20 | 519.70 | 22.80 | 1.01 | 0.65 |
| La | PLSR | CR | -0.03 | -0.10 | 767.12 | 27.70 | -3.00 | 758.11 | 27.53 | 0.83 | 0.53 |
| La | PLSR | DET | -0.04 | 0.00 | 529.28 | 23.01 | -0.49 | 529.04 | 23.00 | 1.00 | 0.64 |
| La | PLSR | SNV | 0.05 | -0.15 | 645.57 | 25.41 | -1.03 | 644.51 | 25.39 | 0.91 | 0.58 |
| La | PLSR | MSC | 0.51 | 0.55 | 272.30 | 16.50 | 2.30 | 267.00 | 16.34 | 1.40 | 0.89 |
| ETRLs/ETRPs | PLSR | RAW | 0.44 | 0.56 | 22.24 | 4.72 | 0.12 | 22.23 | 4.71 | 1.38 | 1.45 |
| ETRLs/ETRPs | PLSR | SG | 0.23 | 0.44 | 31.67 | 5.63 | -0.95 | 30.76 | 5.55 | 1.16 | 1.21 |
| ETRLs/ETRPs | PLSR | CR | 0.21 | 0.43 | 31.93 | 5.65 | -0.59 | 31.58 | 5.62 | 1.15 | 1.21 |
| ETRLs/ETRPs | PLSR | DET | 0.23 | 0.40 | 30.14 | 5.49 | -0.45 | 29.93 | 5.47 | 1.18 | 1.25 |
| ETRLs/ETRPs | PLSR | SNV | 0.16 | 0.38 | 34.43 | 5.87 | -0.57 | 34.11 | 5.84 | 1.11 | 1.16 |
| ETRLs/ETRPs | PLSR | MSC | 0.15 | 0.31 | 33.05 | 5.75 | -0.13 | 33.03 | 5.75 | 1.13 | 1.19 |
| ETRLs | PLSR | RAW | 0.03 | 0.20 | 8874.21 | 94.20 | 6.44 | 8832.75 | 93.98 | 1.02 | 0.99 |
| ETRLs | PLSR | SG | 0.25 | 0.26 | 7064.92 | 84.05 | 8.13 | 6998.90 | 83.66 | 1.15 | 1.11 |
| ETRLs | PLSR | CR | 0.10 | 0.16 | 7976.50 | 89.31 | 11.67 | 7840.31 | 88.55 | 1.08 | 1.05 |
| ETRLs | PLSR | DET | -0.01 | 0.15 | 10534.67 | 102.64 | 13.25 | 10359.08 | 101.78 | 0.94 | 0.91 |
| ETRLs | PLSR | SNV | -0.04 | 0.03 | 9571.17 | 97.83 | 13.77 | 9381.65 | 96.86 | 0.98 | 0.95 |
| ETRLs | PLSR | MSC | -0.03 | 0.03 | 9118.45 | 95.49 | 10.96 | 8998.29 | 94.86 | 1.01 | 0.98 |
| Lu | PLSR | RAW | 0.10 | 0.33 | 0.01 | 0.10 | 0.01 | 0.01 | 0.10 | 1.07 | 1.24 |
| Lu | PLSR | SG | 0.07 | 0.25 | 0.01 | 0.10 | 0.02 | 0.01 | 0.10 | 1.04 | 1.21 |
| Lu | PLSR | CR | 0.32 | 0.52 | 0.01 | 0.08 | 0.00 | 0.01 | 0.08 | 1.27 | 1.48 |
| Lu | PLSR | DET | 0.26 | 0.32 | 0.01 | 0.09 | 0.03 | 0.01 | 0.09 | 1.15 | 1.33 |
| Lu | PLSR | SNV | 0.13 | 0.28 | 0.01 | 0.09 | 0.02 | 0.01 | 0.09 | 1.11 | 1.29 |
| Lu | PLSR | MSC | -0.05 | 0.04 | 0.01 | 0.10 | 0.02 | 0.01 | 0.10 | 0.99 | 1.15 |
| Mn | PLSR | RAW | 0.02 | 0.16 | 65750.52 | 256.42 | -1.07 | 65749.37 | 256.42 | 0.99 | 1.10 |
| Mn | PLSR | SG | 0.00 | 0.15 | 73582.43 | 271.26 | -1.20 | 73580.99 | 271.26 | 0.93 | 1.04 |
| Mn | PLSR | CR | 0.14 | 0.29 | 52701.79 | 229.57 | -9.56 | 52610.43 | 229.37 | 1.10 | 1.23 |
| Mn | PLSR | DET | 0.05 | 0.19 | 59235.99 | 243.38 | -12.81 | 59072.02 | 243.05 | 1.04 | 1.16 |
| Mn | PLSR | SNV | 0.09 | 0.21 | 56157.12 | 236.97 | -17.29 | 55858.20 | 236.34 | 1.07 | 1.19 |
| Mn | PLSR | MSC | 0.07 | 0.28 | 75471.17 | 274.72 | -59.75 | 71901.15 | 268.14 | 0.92 | 1.03 |
| Nd | PLSR | RAW | 0.05 | 0.16 | 607.82 | 24.65 | -4.54 | 587.18 | 24.23 | 1.05 | 0.86 |
| Nd | PLSR | SG | 0.05 | 0.16 | 606.36 | 24.62 | -4.95 | 581.84 | 24.12 | 1.05 | 0.86 |
| Nd | PLSR | CR | -0.03 | -0.07 | 810.28 | 28.47 | -6.74 | 764.90 | 27.66 | 0.91 | 0.74 |
| Nd | PLSR | DET | 0.33 | -0.49 | 1442.64 | 37.98 | -7.61 | 1384.74 | 37.21 | 0.68 | 0.56 |

0.02

0.02

0.02

0.02

0.02

0.02

18.78

21.76

19.03

21.15

20.47

0.13 1.06

0.13 1.08

0.14 1.02

0.14 1.06

0.14 1.00

0.14 1.02

4.33 1.13

4.66 1.04

4.52 1.07

1.12

1.06

4.36

4.60

1.34

1.37

1.29

1.34

1.27

1.29

1.82

1.67

1.81

1.71

1.73

| Nd | PLSR | SNV | -0.04 | 0.02 | 666.28 | 25.81 | -4.77 | 643.49 | 25.37 | 1.00 | 0.82 |
|-------|------|-----|-------|-------|----------|--------|-------|----------|--------|------|------|
| Nd | PLSR | MSC | -0.02 | 0.04 | 644.48 | 25.39 | -4.20 | 626.82 | 25.04 | 1.02 | 0.83 |
| Ni | PLSR | RAW | 0.07 | 0.20 | 117.64 | 10.85 | 4.92 | 93.46 | 9.67 | 0.92 | 1.12 |
| Ni | PLSR | SG | 0.09 | 0.24 | 120.72 | 10.99 | 5.38 | 91.77 | 9.58 | 0.91 | 1.11 |
| Ni | PLSR | CR | 0.17 | 0.36 | 96.48 | 9.82 | 3.55 | 83.91 | 9.16 | 1.02 | 1.24 |
| Ni | PLSR | DET | 0.26 | 0.45 | 95.54 | 9.77 | 4.29 | 77.10 | 8.78 | 1.02 | 1.24 |
| Ni | PLSR | SNV | 0.09 | 0.24 | 106.36 | 10.31 | 4.03 | 90.12 | 9.49 | 0.97 | 1.18 |
| Ni | PLSR | MSC | 0.09 | 0.24 | 111.44 | 10.56 | 4.55 | 90.72 | 9.52 | 0.95 | 1.15 |
| Pb | PLSR | RAW | 0.16 | 0.31 | 42.56 | 6.52 | 0.80 | 41.91 | 6.47 | 1.11 | 1.17 |
| Pb | PLSR | SG | 0.16 | 0.31 | 42.24 | 6.50 | 0.62 | 41.85 | 6.47 | 1.12 | 1.18 |
| Pb | PLSR | CR | 0.14 | 0.34 | 43.87 | 6.62 | 0.06 | 43.87 | 6.62 | 1.09 | 1.15 |
| Pb | PLSR | DET | 0.25 | 0.13 | 46.09 | 6.79 | 0.94 | 45.21 | 6.72 | 1.07 | 1.13 |
| Pb | PLSR | SNV | 0.13 | 0.15 | 46.25 | 6.80 | 1.08 | 45.08 | 6.71 | 1.07 | 1.12 |
| Pb | PLSR | MSC | 0.13 | 0.13 | 46.96 | 6.85 | 1.05 | 45.86 | 6.77 | 1.06 | 1.11 |
| pH | PLSR | RAW | 0.11 | 0.29 | 0.51 | 0.71 | 0.10 | 0.50 | 0.71 | 1.05 | 1.08 |
| pH | PLSR | SG | 0.19 | 0.39 | 0.46 | 0.68 | 0.09 | 0.45 | 0.67 | 1.11 | 1.14 |
| pH | PLSR | CR | 0.31 | 0.53 | 0.42 | 0.65 | 0.13 | 0.40 | 0.63 | 1.16 | 1.19 |
| pН | PLSR | DET | 0.29 | 0.51 | 0.42 | 0.65 | 0.15 | 0.40 | 0.63 | 1.15 | 1.18 |
| pН | PLSR | SNV | 0.16 | 0.38 | 0.51 | 0.71 | 0.14 | 0.49 | 0.70 | 1.05 | 1.08 |
| pН | PLSR | MSC | 0.15 | 0.39 | 0.58 | 0.76 | 0.16 | 0.55 | 0.74 | 0.99 | 1.01 |
| Pr | PLSR | RAW | 0.17 | 0.33 | 27.28 | 5.22 | 2.76 | 19.64 | 4.43 | 0.94 | 1.62 |
| Pr | PLSR | SG | 0.36 | 0.39 | 20.32 | 4.51 | 2.31 | 14.96 | 3.87 | 1.09 | 1.88 |
| Pr | PLSR | CR | 0.12 | 0.23 | 26.59 | 5.16 | 2.59 | 19.86 | 4.46 | 0.96 | 1.64 |
| Pr | PLSR | DET | 0.00 | 0.17 | 53.67 | 7.33 | 3.09 | 44.10 | 6.64 | 0.67 | 1.16 |
| Pr | PLSR | SNV | -0.04 | -0.02 | 36.27 | 6.02 | 2.13 | 31.75 | 5.63 | 0.82 | 1.41 |
| Pr | PLSR | MSC | -0.04 | 0.02 | 36.61 | 6.05 | 2.42 | 30.74 | 5.54 | 0.81 | 1.40 |
| Sb | PLSR | RAW | 0.09 | 0.31 | 0.09 | 0.31 | 0.06 | 0.09 | 0.30 | 1.02 | 0.68 |
| Sb | PLSR | SG | 0.14 | 0.21 | 0.09 | 0.29 | 0.09 | 0.08 | 0.28 | 1.07 | 0.71 |
| Sb | PLSR | CR | 0.10 | 0.18 | 0.09 | 0.31 | 0.11 | 0.08 | 0.29 | 1.03 | 0.69 |
| Sb | PLSR | DET | 0.29 | 0.47 | 0.07 | 0.27 | 0.10 | 0.07 | 0.26 | 1.15 | 0.77 |
| Sb | PLSR | SNV | 0.02 | 0.16 | 0.11 | 0.33 | 0.12 | 0.09 | 0.31 | 0.95 | 0.63 |
| Sb | PLSR | MSC | 0.07 | 0.22 | 0.10 | 0.32 | 0.13 | 0.09 | 0.30 | 0.98 | 0.66 |
| Sc | PLSR | RAW | 0.31 | 0.55 | 2.24 | 1.50 | -0.29 | 2.15 | 1.47 | 1.35 | 2.25 |
| Sc | PLSR | SG | -0.15 | 0.15 | 4.08 | 2.02 | -0.52 | 3.81 | 1.95 | 1.00 | 1.67 |
| Sc | PLSR | CR | -0.19 | 0.03 | 4.31 | 2.08 | -0.88 | 3.53 | 1.88 | 0.97 | 1.62 |
| Sc | PLSR | DET | 0.13 | 0.14 | 3.45 | 1.86 | -0.72 | 2.93 | 1.71 | 1.09 | 1.81 |
| Sc | PLSR | SNV | -0.07 | 0.06 | 4.38 | 2.09 | -1.06 | 3.24 | 1.80 | 0.96 | 1.61 |
| Sc | PLSR | MSC | 0.01 | 0.07 | 4.06 | 2.01 | -0.93 | 3.19 | 1.79 | 1.00 | 1.67 |
| Se | PLSR | RAW | 0.13 | 0.20 | 0.08 | 0.28 | -0.08 | 0.07 | 0.27 | 1.05 | 0.94 |
| Se | PLSR | SG | -0.01 | 0.05 | 0.09 | 0.30 | -0.07 | 0.09 | 0.29 | 0.98 | 0.88 |
| Se | PLSR | CR | 0.00 | 0.09 | 0.09 | 0.29 | -0.05 | 0.08 | 0.29 | 1.01 | 0.91 |
| Se | PLSR | DET | 0.03 | 0.11 | 0.08 | 0.29 | -0.06 | 0.08 | 0.28 | 1.02 | 0.91 |
| Se | PLSR | SNV | 0.00 | 0.07 | 0.09 | 0.30 | -0.08 | 0.08 | 0.29 | 0.99 | 0.89 |
| Se | PLSR | MSC | 0.00 | 0.08 | 0.09 | 0.30 | -0.07 | 0.08 | 0.29 | 0.99 | 0.89 |
| silte | PLSR | RAW | 0.04 | 0.18 | 13062.11 | 114.29 | -6.92 | 13014.15 | 114.08 | 1.03 | 1.17 |
| silte | PLSR | SG | 0.14 | 0.29 | 11445.01 | 106.98 | 0.81 | 11444.35 | 106.98 | 1.10 | 1.25 |

| silte | PLSR | CR | 0.12 | 0.32 | 12104.07 | 110.02 | 5.55 | 12073.22 | 109.88 | 1.07 | 1.21 |
|-------|------|-----|-------|-------|----------|--------|--------|----------|--------|------|------|
| silte | PLSR | DET | 0.18 | 0.34 | 10945.68 | 104.62 | -4.51 | 10925.34 | 104.52 | 1.13 | 1.28 |
| silte | PLSR | SNV | 0.15 | 0.31 | 11373.19 | 106.65 | -11.07 | 11250.61 | 106.07 | 1.11 | 1.25 |
| silte | PLSR | MSC | 0.01 | 0.11 | 13751.38 | 117.27 | -16.15 | 13490.71 | 116.15 | 1.01 | 1.14 |
| Sm | PLSR | RAW | 0.30 | 0.49 | 3.21 | 1.79 | 0.63 | 2.81 | 1.68 | 1.15 | 1.83 |
| Sm | PLSR | SG | 0.12 | 0.27 | 3.99 | 2.00 | 0.68 | 3.53 | 1.88 | 1.03 | 1.64 |
| Sm | PLSR | CR | 0.15 | 0.28 | 3.67 | 1.92 | 0.55 | 3.36 | 1.83 | 1.08 | 1.71 |
| Sm | PLSR | DET | -0.03 | 0.06 | 4.78 | 2.19 | 0.50 | 4.53 | 2.13 | 0.94 | 1.50 |
| Sm | PLSR | SNV | -0.04 | -0.02 | 5.14 | 2.27 | 0.19 | 5.11 | 2.26 | 0.91 | 1.44 |
| Sm | PLSR | MSC | -0.03 | -0.04 | 4.46 | 2.11 | 0.30 | 4.37 | 2.09 | 0.98 | 1.55 |
| Tb | PLSR | RAW | 0.20 | 0.45 | 0.15 | 0.39 | 0.10 | 0.14 | 0.38 | 0.99 | 1.29 |
| Tb | PLSR | SG | 0.16 | 0.40 | 0.17 | 0.41 | 0.14 | 0.15 | 0.39 | 0.95 | 1.23 |
| Tb | PLSR | CR | 0.24 | 0.41 | 0.33 | 0.57 | 0.20 | 0.29 | 0.54 | 0.68 | 0.88 |
| Tb | PLSR | DET | 0.25 | 0.48 | 0.16 | 0.40 | 0.14 | 0.14 | 0.38 | 0.97 | 1.26 |
| Tb | PLSR | SNV | 0.17 | 0.39 | 0.24 | 0.49 | 0.19 | 0.20 | 0.45 | 0.80 | 1.03 |
| Tb | PLSR | MSC | 0.17 | 0.36 | 0.15 | 0.38 | 0.15 | 0.12 | 0.35 | 1.02 | 1.32 |
| Th | PLSR | RAW | 0.02 | 0.16 | 40.68 | 6.38 | -2.46 | 34.60 | 5.88 | 1.00 | 1.09 |
| Th | PLSR | SG | -0.01 | 0.20 | 50.49 | 7.11 | -2.56 | 43.93 | 6.63 | 0.90 | 0.98 |
| Th | PLSR | CR | 0.04 | 0.29 | 48.45 | 6.96 | -2.09 | 44.10 | 6.64 | 0.91 | 1.00 |
| Th | PLSR | DET | 0.11 | 0.35 | 39.31 | 6.27 | -2.05 | 35.11 | 5.93 | 1.02 | 1.11 |
| Th | PLSR | SNV | 0.08 | 0.29 | 39.61 | 6.29 | -2.24 | 34.59 | 5.88 | 1.01 | 1.11 |
| Th | PLSR | MSC | 0.08 | 0.29 | 38.87 | 6.23 | -2.23 | 33.89 | 5.82 | 1.02 | 1.12 |
| V | PLSR | RAW | 0.16 | 0.38 | 404.49 | 20.11 | -4.33 | 385.73 | 19.64 | 1.10 | 1.30 |
| V | PLSR | SG | 0.00 | 0.14 | 476.44 | 21.83 | -4.71 | 454.24 | 21.31 | 1.02 | 1.20 |
| V | PLSR | CR | -0.02 | 0.15 | 521.45 | 22.84 | -3.91 | 506.19 | 22.50 | 0.97 | 1.14 |
| V | PLSR | DET | 0.24 | 0.46 | 344.91 | 18.57 | -1.42 | 342.90 | 18.52 | 1.19 | 1.41 |
| V | PLSR | SNV | -0.02 | 0.14 | 480.01 | 21.91 | -1.66 | 477.26 | 21.85 | 1.01 | 1.19 |
| V | PLSR | MSC | 0.08 | 0.27 | 413.18 | 20.33 | -2.10 | 408.76 | 20.22 | 1.09 | 1.29 |
| Y | PLSR | RAW | 0.19 | 0.43 | 20.48 | 4.53 | -0.61 | 20.11 | 4.48 | 1.30 | 1.22 |
| Y | PLSR | SG | -0.20 | 0.01 | 38.55 | 6.21 | -2.20 | 33.72 | 5.81 | 0.95 | 0.89 |
| Y | PLSR | CR | -0.07 | -0.17 | 43.55 | 6.60 | -1.54 | 41.17 | 6.42 | 0.89 | 0.84 |
| Y | PLSR | DET | 0.34 | -0.31 | 57.08 | 7.56 | -2.91 | 48.60 | 6.97 | 0.78 | 0.73 |
| Y | PLSR | SNV | 0.08 | -0.23 | 50.16 | 7.08 | -2.41 | 44.35 | 6.66 | 0.83 | 0.78 |
| Y | PLSR | MSC | -0.20 | -0.01 | 35.18 | 5.93 | -2.21 | 30.31 | 5.51 | 0.99 | 0.93 |
| Yb | PLSR | RAW | 0.44 | 0.65 | 0.12 | 0.35 | 0.04 | 0.12 | 0.35 | 1.27 | 2.47 |
| Yb | PLSR | SG | 0.22 | 0.42 | 0.15 | 0.39 | 0.05 | 0.15 | 0.38 | 1.16 | 2.25 |
| Yb | PLSR | CR | 0.44 | 0.61 | 0.11 | 0.34 | 0.09 | 0.10 | 0.32 | 1.34 | 2.60 |
| Yb | PLSR | DET | 0.20 | 0.45 | 0.25 | 0.50 | 0.08 | 0.25 | 0.50 | 0.89 | 1.73 |
| Yb | PLSR | SNV | 0.18 | 0.43 | 0.27 | 0.52 | 0.12 | 0.26 | 0.51 | 0.86 | 1.66 |
| Yb | PLSR | MSC | 0.13 | 0.37 | 0.19 | 0.43 | 0.07 | 0.18 | 0.43 | 1.03 | 2.01 |
| Zn | PLSR | RAW | 0.32 | 0.52 | 105.23 | 10.26 | 0.61 | 104.85 | 10.24 | 1.23 | 1.50 |
| Zn | PLSR | SG | 0.34 | 0.54 | 103.26 | 10.16 | 0.84 | 102.55 | 10.13 | 1.24 | 1.51 |
| Zn | PLSR | CR | 0.29 | 0.54 | 138.49 | 11.77 | 1.30 | 136.79 | 11.70 | 1.07 | 1.30 |
| Zn | PLSR | DET | 0.40 | 0.53 | 92.77 | 9.63 | 0.80 | 92.13 | 9.60 | 1.31 | 1.59 |
| Zn | PLSR | SNV | 0.35 | 0.51 | 99.81 | 9.99 | 0.89 | 99.01 | 9.95 | 1.26 | 1.54 |
| Zn | PLSR | MSC | 0.19 | 0.36 | 123.46 | 11.11 | 0.48 | 123.23 | 11.10 | 1.13 | 1.38 |



ANEXO C - Melhores modelos obtidos para: a) Areia com RF-SG; b) Silte com RF-SG; c) Argila com RF-RAW; d) COT com PLSR-CR



ANEXO D - Melhores modelos obtidos para: a) CTC com PLSR-SNV; b) pH com RF-SG



ANEXO E - Melhores modelos obtidos para: a) Ba com RF-SNV; b) Cd com PLSR-CR; c) Co com RF-DET; d) Cr com RF-DET



ANEXO F - Melhores modelos obtidos para: a) Cu com PLSR-CR; b) Fe com RF-DET; c) Mn com RF-DET; d) Ni com RF-DET



ANEXO G - Melhores modelos obtidos para: a) Pb com RF-CR; b) Sb com RF-RAW



ANEXO H - Melhores modelos obtidos para: a) Se com PLSR-RAW; b) V com RF-RAW; c) Zn com PLSR-DET



ANEXO I - Melhores modelos obtidos para: a) Ce com RF-MSC; b) Dy com RF-SG; c) Er com RF-SNV; d) Eu com RF-RAW



ANEXO J - Melhores modelos obtidos para: a) Gd com RF-DET; b) Ho com RF-RAW; c) La com PLSR-MSC; d) Lu com RF-CR



ANEXO K - Melhores modelos obtidos para: a) Nd com RF-RAW; b) Pr com PLSR-SG; c) Sc com PLSR-RAW; d) Sm com RF-RAW



ANEXO L - Melhores modelos obtidos para: a) Tb com RF-SG; b) Y com PLSR-RAW; c) Yb com RF-SG



ANEXO M - Melhores modelos obtidos para: a) ETRLs com RF-SNV; b) ETRPs com RF-SG; c) ETRLs/ETRPs com PLSR-RAW



Exponential model (Co = 0.099; Co + C = 2.416; Ao = 9100.00; Exponential model (Co = 0.270; Co + C = 5.611; Ao = 8600 RSS = 1.47) RSS = 0.356)



Pr: Isotropic Variogram

88



Nd: Isotropic Variogram



Exponential model (Co = 0.049; Co + C = 1.139; Ao = 8600.00; Exponential model (Co = 0.843; Co + C = 2.220; Ao = 14400; Co = RSS = 0.0542)



RSS = 0.305)





Exponential model (Co = 0.0510; Co + C = 0.5610; Ao = 10700 Exponential model (Co = 0.00570; Co + C = 0.08160; Ao = 990 RSS = 9.958E-03)





RSS = 5.92)

RSS = 3.305E-04)

Gd: Isotropic Variogram



Exponential model (Co = 0.50; Co + C = 12.15; Ao = 9500.00; r Exponential model (Co = 0.0620; Co + C = 0.5520; Ao = 10200 RSS = 0.0208)



Er: Isotropic Variogram



Exponential model (Co = 0.01040; Co + C = 0.07930; Ao = 930 Exponential model (Co = 0.0088; Co + C = 0.1076; Ao = 7600.C RSS = 1.297E-04) RSS = 4.043E-04)







Exponential model (Co = 0.0091; Co + C = 0.1102; Ao = 9100.(Spherical model (Co = 0.00010; Co + C = 0.06420; Ao = 49200 RSS = 6.208E-04) RSS = 2.275E-03)



ANEXO O - Validação Cruzada (kriging) dos teores observados (Rio Grande do Norte)

Regression coefficient = 0.445 (SE = 0.187, r2 = 0.062, y intercept = 44.5, SE Prediction = 53.818)



Regression coefficient = 0.372 (SE = 0.227, r2 =0.031, y intercept = 0.35, SE Prediction = 0.377)



93.29

Regression coefficient = 0.360 (SE = 0.250, r2 =0.024, y intercept = 0.42, SE Prediction = 0.491)

0.58

Estimated Er

1.12

1.65

0.05





Regression coefficient = 0.597 (SE = 0.235 , r2 =0.069, y intercept = 0.26, SE Prediction = 0.448)

Regression coefficient = 0.643 (SE = 0.240 , r2 =0.188, y intercept = 0.29, SE Prediction = 0.385)



ANEXO P - Semivariogramas dos teores preditos (Paraíba)

Exponential model (Co = 0.1416; Co + C = 0.3592; Ao = 50 Spherical model (Co = 0.00729; Co + C = 0.03006; Ao = 10730 RSS = 0.0232) RSS = 1.628E-04)

Er: Isotropic Variogram



Exponential model (Co = 0.03908; Co + C = 0.09770; Ao = 620 RSS = 6.194E-04)



ANEXO Q - Validação Cruzada (kriging) dos teores preditos (Paraíba)