

UFG
IQ

UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE QUÍMICA

S
I
A
R
O
M

A
R
I
E
V
I
L
O

E
D

O
T
S
U
G
U
A

O
R
D
E
P

DETERMINAÇÃO DE PARÂMETROS
DE FERTILIDADE DO SOLO POR
MEIO DA ANÁLISE
MULTIVARIADA DE IMAGENS E DE
ESPECTROS DE
INFRAVERMELHO

PEDRO AUGUSTO DE OLIVEIRA MORAIS

ORIENTADOR: PROF. DR. ANSELMO ELCANA DE OLIVEIRA

TESE DE DOUTORADO

2021

GOIÂNIA- 2021



UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE QUÍMICA

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

Dissertação Tese

2. Nome completo do autor

Pedro Augusto de Oliveira Moraes

3. Título do trabalho

Determinação de parâmetros de fertilidade do solo por meio da Análise Multivariada de Imagens e de espectros de infravermelho

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

- a) consulta ao(a) autor(a) e ao(a) orientador(a);
 - b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.
- O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Anselmo Elcana De Oliveira, Professor do Magistério Superior**, em 08/07/2021, às 16:44, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **PEDRO AUGUSTO DE OLIVEIRA MORAIS, Discente**, em 08/07/2021, às 18:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2198394** e o código CRC **519AD9C6**.

PEDRO AUGUSTO DE OLIVEIRA MORAIS

Determinação de parâmetros de fertilidade do solo
por meio da *Análise Multivariada de Imagens e de
espectros de infravermelho*

Tese apresentada ao Instituto de Química da
Universidade Federal de Goiás para obtenção
do título de Doutor em Química pelo
Programa de Pós-graduação em Química.

Área de Concentração: Química

Orientador: Prof. Dr. Anselmo Elcana de Oli-
veira

Goiânia

2020

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Morais, Pedro Augusto de Oliveira

Determinação de parâmetros de fertilidade do solo por meio da Análise Multivariada de Imagens e de espectros de infravermelho [manuscrito] / Pedro Augusto de Oliveira Moraes. - 2021. CXXXIX, 139 f.: il.

Orientador: Prof. Dr. Anselmo Elcana de Oliveira.

Tese (Doutorado) - Universidade Federal de Goiás, Instituto de Química (IQ), Programa de Pós-Graduação em Química, Goiânia, 2021.

Bibliografia. Apêndice.

Inclui siglas, fotografias, abreviaturas, símbolos, gráfico, tabelas, lista de figuras, lista de tabelas.

1. Textura do solo. 2. Carbono orgânico do Solo. 3. Óxidos. 4. MIA. 5. MIR. I. Oliveira, Anselmo Elcana de, orient. II. Título.



UNIVERSIDADE FEDERAL DE GOIÁS

INSTITUTO DE QUÍMICA

ATA DE DEFESA DE TESE

Ata nº **113** da sessão da Defesa de Doutorado de **Pedro Augusto de Oliveira Moraes**, que confere o título de **Doutor em Química**, na área de concentração em **Química**.

Aos **16 (dezesesseis) dias do mês de abril de 2020 (dois mil e vinte)**, a partir das **14h00m**, via **videoconferência de todos os membros**, realizou-se a sessão pública da Defesa de Doutorado intitulada "**Determinação de parâmetros de fertilidade do solo por meio da Análise Multivariada de Imagens e de espectros de infravermelho**". Os trabalhos foram instalados pelo Orientador, **Prof. Dr. Anselmo Elcana de Oliveira**, com a participação dos demais membros da Banca Examinadora: **Drª. Beata Emoki Madari (EMBRAPA)**, **Prof. Dr. Nelson Roberto Antoniosi Filho (UFG)**, **Prof. Dr. Clarimar José Coelho (PUC-GO)** e **Prof. Dr. Aparecido Ribeiro de Souza (UFG)**. Durante a arguição os membros da banca não fizeram sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Defesa de Doutorado, tendo sido o candidato **aprovado** pelos seus membros. Proclamados os resultados pelo Prof. Dr. Anselmo Elcana de Oliveira, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos 16 (dezesesseis) dias do mês de abril de 2020 (dois mil e vinte).

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Nelson Roberto Antoniosi Filho, Professor do Magistério Superior**, em 16/04/2020, às 18:13, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Aparecido Ribeiro De Souza, Professor do Magistério Superior**, em 16/04/2020, às 18:14, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Anselmo Elcana De Oliveira, Professor do Magistério Superior**, em 16/04/2020, às 18:15, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Beata Eموke Madari, Usuário Externo**, em 16/04/2020, às 18:31, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

Documento assinado eletronicamente por **CLARIMAR JOSÉ COELHO, Usuário Externo**, em 22/04/2020, às 14:39, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

*À minha família, amigos, e por mais estranho que pareça a mim mesmo, que nunca pensei
que chegaria até aqui.*

Agradecimentos

Ao Dr. Diego Mendes de Souza do Instituto de Criminalística da Polícia Civil - DF por viabilizar a realização deste estudo, assim como pela colaboração em cada etapa deste trabalho.

À pesquisadora Dra. Beata Eموke Madari do Centro Nacional de Pesquisa de Arroz e Feijão (CNPAP/EMBRAPA) pelo incentivo e colaboração.

Ao professor Dr. Anselmo Elcana de Oliveira pela orientação e paciência ao longo do curso.

Ao professor Dr. Wendell Karlos Tomazelli Coltro pelo incentivo e colaboração.

Aos membros do laboratório de Análise Agroambiental (LAA/CNPAP), Silvio, Adilson, Roberto Carlos, Ivã, Demétrio, Camila, Selma e Tatiana por possibilitarem um ambiente de trabalho dinâmico e agradável.

Aos integrantes do Laboratório de Química Teórica e Computacional (LQTC/UFG), Gabriel e Marcelo pelas discussões e descontrações.

Ao instituto de Química (IQ/UFG) pelo suporte.

Aos colegas de curso Lucas da Costa, Roger e Wanderson pelo companherismo no decorrer do curso.

À Gabriela Matos pelos conselhos, pelo afeto, incentivo e apoio.

À minha família, por sempre me depositar confiança, até quando eu duvidava de mim mesmo.

À Natália Bueno pela amizade e confiança.

À Veronica Lilja pela amizade e apoio, mesmo estando longe.

“Eis uma força motriz mais poderosa que a eletricidade, o vapor e a energia atômica: a

vontade.”

(Albert Einstein)

Resumo

Morais, Pedro Augusto de Oliveira. **Determinação de parâmetros de fertilidade do solo por meio da Análise Multivariada de Imagens e de espectros de infravermelho**. 2020. 139 f. Tese (Doutorado em Química) – Instituto de Química, Universidade Federal de Goiás, Goiânia, 2020.

A análise do solo é uma ferramenta importante na avaliação e monitoramento do impacto ambiental da atividade agropecuária e ao mesmo tempo viabiliza o planejamento do uso racional de insumos, contribuindo assim à redução e melhor sustentabilidade ambiental e econômico da produção. Conseqüentemente, há crescente demanda pelos serviços dos laboratórios de análise de solo. Entretanto, algumas das metodologias empregadas, além de gerar volumes consideráveis de resíduos, possuem alto custo de implementação. Logo, alternativas ambientalmente sustentáveis e mais baratas devem ser desenvolvidas. Nesse sentido, este estudo propõe a utilização de análise multivariada de imagens (MIA) digitais e da espectroscopia na região do infravermelho médio (MIR) para estimativa do carbono orgânico e da textura do solo, assim como da determinação dos teores de óxidos de ferro, alumínio e silício extraídos do solo. Para isso, 177 amostras de solo de diferentes regiões do país foram consideradas para as análises pelos métodos de referência em laboratório, para obtenção de imagens digitais no sistema de cor RGB (*Red, Green, Blue*) em formato Tiff e para aquisição dos espectros MIR. A correlação entre imagens digitais, e espectros de MIR, e os parâmetros de fertilidade foi realizada por meio dos métodos: Regressão por Mínimos Quadrados Parciais (PLS), Regressão Linear Múltipla associada com o Algoritmo das Projeções Sucessivas (SPA-MLR) e Máquina de Vetor Suporte por Mínimos Quadrados (LS-SVM). Os melhores modelos de predição dos parâmetros de fertilidade obtiveram correlações superiores a 90% e valores de desvio de previsão residual (RPD) superiores a 3. A utilização desses métodos na rotina laboratorial pode levar a um aumento significativo da produtividade, redução dos custos de análise e minimização do impacto ambiental, pois além de não gerarem resíduo, não utilizam reagentes químicos. O acesso a essas análises de forma rápida e de baixo custo permitiria um melhor acompanhamento dos solos para todos os produtores rurais, sejam eles grandes, médios ou, até mesmo, pequeno, o que resultaria, conseqüentemente, em um aumento da produtividade no campo.

Palavras-chaves: Textura do solo, Carbono orgânico do Solo, óxidos, MIA, MIR, PLS, SPA-MLR, LS-SVM.

Abstract

Morais, Pedro Augusto de Oliveira. **Determination of soil fertility parameters by Multivariate Image Analysis and infrared spectra**. 2020. 139 p. Thesis (Doctorate of Chemistry) – Institute of Chemistry, Federal University of Goiás, Goiânia, 2020.

Soil analysis is an important tool when monitoring the environmental impact of agricultural activity. It also allows for the rational planning inputs contributing to a better environmental sustainability and economic production. Consequently, there is a growing demand for the services of soil analysis laboratories. However, methodologies currently employed in the field not only generate a considerable amount of waste, but also have a high set up cost. Therefore, cheaper and environmentally sustainable alternatives should be developed. In this sense, this study proposes the use of soil digital images and mid-infrared spectroscopy (MIR) to estimate soil organic carbon (SOC), predict and classify soil texture, as well estimate iron, aluminium, and silicon oxides contents. For this purpose, 177 samples from different regions of the country were analyzed by standard methods. Soil digital images were acquired using RGB (Red, Green, Blue) in Tiff format. The correlation between digital images, MIR spectrum, and soil fertility parameters was obtained using Partial Least Squares Regression (PLS), Multiple Linear Regression algorithm associated with the Successive Projections (SPA-MLR), and Least Squares Support Vector Machines (LS-SVM). The best models present correlations higher than 90% and Residual Prediction Deviation (RPD) values greater than 3.0. The use of these methods in test soil analysis would allow a significant increase in productivity, reduction of the cost of analysis, and minimization of environmental impact. The proposed analyses do not produce waste and do not employ chemicals. As a result, farmers can benefit from the proposed methods taken into account that the analyses are quick and inexpensive and might lead to an increase in productivity in the field.

Keywords: Soil texture, Soil Organic Carbon, oxides, MIA, MIR, PLS, SPA-MLR, LS-SVM.

Lista de figuras

Figura 1 – Diagrama triangular utilizado na classificação textural do solo.	33
Figura 2 – Representação das componentes iluminância (I) e reflectância (R).	36
Figura 3 – Espectro eletromagnético.	37
Figura 4 – Sistema RGB de cores.	38
Figura 5 – Modelo de cores HSI.	39
Figura 6 – Exemplo de histograma de uma imagem dispendo de dez níveis de cores.	40
Figura 7 – Esquemática da etapa de pré-processamento centrar na média.	45
Figura 8 – Representação da MLR.	47
Figura 9 – Representação da decomposição da matriz \mathbf{X} e do vetor \mathbf{y} pelo método PLS	48
Figura 10 – Fluxograma da análise de solo por meio de imagens digitais.	54
Figura 11 – Fluxograma da análise de solo por meio de imagens digitais.	55
Figura 12 – Scanner de mesa utilizado para a aquisição das imagens de solo.	61
Figura 13 – Imagens digitais para seis amostras de solo (teores de COS entre 0,590 a 47,590 $g\ kg^{-1}$)	61
Figura 14 – Seleção de uma região de interesse da primeira imagem de amostra de solo obtida.	62
Figura 15 – Passos realizados para obtenção dos histogramas de ocorrência de cor de amostras de solo.	63
Figura 16 – Distribuição dos teores de carbono orgânico e granulometria do solo das 177 amostras.	65
Figura 17 – Classificação textural de todas as amostras	66
Figura 18 – Etapas utilizadas na quantificação do COS e da textura do solo pelo método de referência e pela MIA.	67
Figura 19 – Matriz de dados referentes aos histogramas de ocorrência dos modelos RGB, HSI e escala cinza de todas as imagens de solo.	68
Figura 20 – Etapas utilizadas para estimação dos teores de COS, areia e argila em amostras de solo por MIA.	68
Figura 21 – Concentrações de COS de referência versus previstas pelo modelo MIA- PLS-RHC. Calibração (*) e validação (▲).	69

Figura 22 – Concentrações de COS de referência versus previstas pelo modelo MIA-SPA-MLR-RC. Calibração (*) e validação (▲).	70
Figura 23 – Concentrações de COS de referência versus previstas pelo modelo MIA-LS-SVM-RHC. Calibração (*) e validação (▲).	71
Figura 24 – Teores de argila de referência versus os teores de argila previstos pelo modelo PLS. Calibração (*) e validação (▲).	76
Figura 25 – Teores de argila de referência versus os teores de argila previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).	77
Figura 26 – Teores de argila de referência versus os teores de argila previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).	78
Figura 27 – Teores de areia de referência versus os teores de areia previstos pelo modelo PLS. Calibração (*) e validação (▲).	81
Figura 28 – Teores de areia de referência versus os teores de areia previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).	81
Figura 29 – Teores de areia de referência versus os teores de areia previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).	82
Figura 30 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo PLS. Calibração (*) e validação (▲).	89
Figura 31 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).	90
Figura 32 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).	90
Figura 33 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo PLS. Calibração (*) e validação (▲).	92
Figura 34 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).	92
Figura 35 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).	92
Figura 36 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo PLS. Calibração (*) e validação (▲).	93
Figura 37 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).	94

Figura 38 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).	94
Figura 39 – Etapas utilizadas na quantificação do COS e da textura do solo pelo método de referência e por MIR.	95
Figura 40 – Matriz de dados referentes aos espectros de infravermelho médio de todas as amostras de solo.	95
Figura 41 – Etapas utilizadas para estimação do COS e da textura em amostras de solo por MIR.	96
Figura 42 – Concentrações de COS de referência versus previstas pelo modelo PLS. Calibração (*) e validação (▲).	97
Figura 43 – Concentrações de COS de referência versus previstas pelo modelo SPA-MLR. Calibração (*) e validação (▲).	97
Figura 44 – Concentrações de COS de referência versus previstas pelo modelo LS-SVM. Calibração (*) e validação (▲).	97
Figura 45 – Teores de argila de referência versus previstas pelo modelo PLS. Calibração (*) e validação (▲).	99
Figura 46 – Teores de argila de referência versus previstas pelo modelo SPA-MLR. Calibração (*) e validação (▲).	99
Figura 47 – Teores de argila de referência versus previstas pelo modelo LSSVM. Calibração (*) e validação (▲).	100
Figura 48 – Teores de areia de referência versus previstas pelo modelo PLS. Calibração (*) e validação (▲).	101
Figura 49 – Teores de areia de referência versus previstas pelo modelo SPA-MLR. Calibração (*) e validação (▲).	102
Figura 50 – Teores de areia de referência versus previstas pelo modelo LSSVM. Calibração (*) e validação (▲).	102
Figura 51 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo PLS. Calibração (*) e validação (▲).	104
Figura 52 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).	105
Figura 53 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).	105

Figura 54 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo PLS. Calibração (*) e validação (▲).	106
Figura 55 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).	107
Figura 56 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).	107
Figura 57 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo PLS. Calibração (*) e validação (▲).	108
Figura 58 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).	108
Figura 59 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).	109

Lista de tabelas

Tabela 1 – Faixas espectrais do infravermelho.	41
Tabela 2 – Condições de preparo das soluções padrão utilizadas na curva de calibração.	57
Tabela 3 – Bandas de absorção referentes aos óxidos de silício, ferro e alumínio na espectroscopia de infravermelho médio (MIR) (SAIKIA; PARTHASARATHY et al., 2010; IBRAHIM; HAMEED; JALBOUT, 2008; MADEJOVÁ, 2003; DJOMGOUE; NJOPWOUO et al., 2013).	59
Tabela 4 – Bandas de absorção referentes ao carbono orgânico na espectroscopia de infravermelho médio (MIR) (STEVENSON, 1994; BAES; BLOOM, 1989; NIEMEYER; CHEN; BOLLAG, 1992; MADARI et al., 2006).	60
Tabela 5 – Atributos físicos e químicos das 177 amostras de solo.	65
Tabela 6 – Proporções de amostras de solo para cada grupo textural.	66
Tabela 7 – Parâmetros estatísticos da calibração e validação obtidos utilizando o método PLS.	69
Tabela 8 – Parâmetros estatísticos da calibração e validação obtidos utilizando o método SPA-MLR.	70
Tabela 9 – Parâmetros estatísticos da calibração e validação obtidos utilizando o método LS-SVM.	70
Tabela 10 – Resultados obtidos na detecção da não linearidade dos dados por meio dos testes: <i>Runs</i> e Durbin-Watson	73
Tabela 11 – Parâmetros estatísticos da calibração e validação na quantificação de argila utilizando o método PLS.	75
Tabela 12 – Parâmetros estatísticos da calibração e validação na quantificação de argila utilizando o método SPA-MLR.	76
Tabela 13 – Parâmetros estatísticos da calibração e validação na quantificação de argila utilizando o método LS-SVM.	77
Tabela 14 – Resultados obtidos na detecção da não linearidade dos dados por meio dos testes: <i>Runs</i> e Durbin-Watson	79
Tabela 15 – Parâmetros estatísticos da calibração e validação na quantificação de areia utilizando o método PLS.	80
Tabela 16 – Parâmetros estatísticos da calibração e validação na quantificação de areia utilizando o método SPA-MLR.	80

Tabela 17 – Parâmetros estatísticos da calibração e validação na quantificação de areia utilizando o método LS-SVM.	82
Tabela 18 – Resultados obtidos na detecção da não linearidade dos dados por meio dos testes: <i>Runs</i> e Durbin-Watson	84
Tabela 19 – Taxas de acerto na classificação textural por meio de imagens.	87
Tabela 20 – Propriedades estatísticas das 52 amostras de solo.	88
Tabela 21 – Resultados estatísticos obtidos na estimação de Fe_2O_3 em solo por meio dos métodos: SPA-MLR, PLS e LS-SVM.	89
Tabela 22 – Resultados estatísticos obtidos na estimação de SiO_2 em solo por meio dos métodos: SPA-MLR, PLS e LS-SVM.	91
Tabela 23 – Resultados estatísticos obtidos na predição de Al_2O_3 por meio dos métodos SPA-MLR, PLSR e LS-SVM.	93
Tabela 24 – Parâmetros estatísticos da calibração e validação na quantificação de COS utilizando os métodos SPA-MLR, PLS e LS-SVM.	96
Tabela 25 – Parâmetros estatísticos da calibração e validação na quantificação dos teores de argila utilizando os métodos SPA-MLR, PLS e LS-SVM.	99
Tabela 26 – Parâmetros estatísticos da calibração e validação na quantificação dos teores de areia utilizando os métodos SPA-MLR, PLS e LS-SVM.	101
Tabela 27 – Resultados estatísticos obtidos na estimação de Fe_2O_3 em solo por meio dos métodos SPA-MLR, PLS e LS-SVM.	104
Tabela 28 – Resultados estatísticos obtidos na estimação de SiO_2 em solo por meio dos métodos SPA-MLR, PLS e LS-SVM.	106
Tabela 29 – Resultados estatísticos obtidos na estimação de Al_2O_3 em solo por meio dos métodos SPA-MLR, PLS e LS-SVM.	108
Tabela 30 – Comparação de resultados na estimação de COS e da textura do solo por meio de imagens digitais e MIR.	113

Lista de abreviaturas e siglas

CCD	<i>Charge Coupled Device</i> , ou Dispostivo de carga acoplada
CMY	<i>Cyan, Magenta, Yellow</i> , ou Ciano, Magenta e Amarelo
CONAB	Companhia Nacional de Abastecimento
COS	Carbono Orgânico do Solo
Ctot	Carbono Total
CV	<i>Cross Validation</i> , ou Validação Cruzada
DPI	<i>Dots per Inch</i> , ou Pontos Por Polegada
DRIFT	<i>Diffuse Reflectance Infrared Fourier Transform Spectroscopy</i> , ou reflectância difusa no infravermelho médio com transformada de Fourier
FAO	<i>Food and Agriculture Organization of the United Nations</i> , ou Organização das Nações Unidas para Agricultura e Alimentação
FIR	<i>Far Infrared</i> , ou Espectroscopia na região do infravermelho distante
HSI	<i>Hue, Saturation, Intensity</i> , ou Matiz, Saturação e Intensidade
ICP-OES	<i>Inductively Coupled Plasma Optical Emission Spectrometry</i> , ou Espectrofotometria de emissão ótica com plasma indutivamente acoplado
IR	<i>Infrared</i> ou Infravermelho
KS	Algoritmo Kennard-Stone
LAA	Laboratório de Análise Agroambiental
LS-SVM	<i>Least Squares Support Vector Machines</i> , ou Máquinas de Vetor Suporte por mínimos quadrados
LV	Variável Latente
MAPA	Ministério de Agricultura, Pecuária e Abastecimento
MIA	<i>Multivariate Image Analysis</i> , ou Análise Multivariada de Imagens

MIR	<i>Mid-infrared Spectroscopy</i> , ou Espectroscopia na região do infravermelho médio
MLR	<i>Multiple Linear Regression</i> , ou Regressão Linear Múltipla
NIRS	<i>Near Infrared Spectroscopy</i> , ou Espectroscopia na região do infravermelho próximo
NT	Nitrogênio Total
OCDE	Organização para a Cooperação e Desenvolvimento da Europa
PCR	<i>Principal Components Regression</i> , ou Regressão por Componentes principais
PIB	Produto Interno Bruto
PLS	<i>Partial Least Square</i> , ou Mínimos Quadrados Parciais
RI	<i>Redness Index</i> , ou índice de vermelhidão
RGB	<i>Red, Green, Blue</i> ou Vermelho, Verde e Azul
RMSE	<i>Root Mean Square Error</i> , ou Raiz Quadrada do Erro Médio Quadrático
RMSEC	<i>Root Mean Square Error of Calibration</i> , ou Raiz Quadrada do Erro Médio Quadrático de Calibração
RMSECV	<i>Root Mean Square Error of Cross-Validation</i> , ou Raiz Quadrada do Erro Médio Quadrático da Validação Cruzada
RMSEP	<i>Root Mean Square Error of Prediction</i> , ou Raiz Quadrada do Erro Médio Quadrático de Previsão
RMSEV	<i>Root Mean Square Error of Validation</i> , ou Raiz Quadrada do Erro Médio Quadrático de Validação
ROI	<i>Region of Interest</i> , ou Região de Interesse
RPD	<i>Residual Prediction Deviation</i> , ou Desvio Residual de Previsão
RPM	Rotação por minuto

SEP	<i>Standard Error of Prediction</i> , ou Erro Padrão de Revisão
SPA	<i>Successive Projection Algorithm</i> , ou Algoritmo das Projeções Sucessivas
SPXY	<i>Sample set Partitioning based on joint x-y distances</i>
SVM	<i>Support Vector Machines</i> , ou Máquinas de Vetor Suporte
TIFF	<i>Tagged Image File Format</i>
Vis	<i>Visible Spectroscopy</i> , ou Espectroscopia na região do visível
YIQ	<i>Luminance, In-phase, Quadrature</i>

Sumário

1	Introdução	24
2	Objetivos	26
2.1	Objetivos específicos	26
3	Revisão Bibliográfica	27
4	Metodologia	31
4.1	Carbono orgânico do solo	31
4.1.1	Quantificação do carbono orgânico do solo	31
4.2	A textura do solo	32
4.2.1	Quantificação da textura do solo	32
4.3	Óxidos de alumínio, silício e ferro em solo	34
4.3.1	Quantificação dos óxidos de alumínio, silício e ferro em solo	34
4.4	Imagens Digitais	35
4.4.1	Cor	36
4.4.2	Modelos de cores	37
4.4.2.1	Modelo RGB	38
4.4.2.2	Modelo HSI	39
4.5	Análise Multivariada de imagens	40
4.5.1	Histogramas de ocorrência de cor	40
4.6	Espectroscopia no infravermelho	41
4.6.1	Absorção de radiação na região do infravermelho	42
4.6.2	Aplicações da espectroscopia no infravermelho	42
4.7	Calibração Multivariada	43
4.7.1	Organização dos dados	43
4.7.2	Pré-processamento de dados	44
4.7.2.1	Centrar na Média	44
4.7.3	Algoritmo Kennard-Stone	45
4.7.4	Algoritmo SPXY	46

4.7.5	Regressão Linear Múltipla	46
4.7.6	PLS	47
4.7.7	Seleção de LVs	48
4.7.8	LS-SVM	49
4.7.9	Seleção de variáveis	50
4.7.9.1	SPA	50
4.7.10	Avaliação dos modelos de calibração multivariada por meio de métodos estatísticos	51
4.7.10.1	Linearidade	52
4.8	Análise de solo por meio de MIA	53
4.9	Análise de solo por meio de espectros de MIR	55
5	Parte experimental	56
5.1	Amostragem	56
5.1.1	Coleta e tratamento das amostras	56
5.2	Metodologia para quantificação do carbono orgânico do solo	56
5.3	Metodologia para quantificação da textura do solo	57
5.4	Metodologia para quantificação de fósforo, sódio e potássio disponíveis	58
5.5	Metodologia para quantificação de micronutrientes disponíveis	58
5.6	Metodologia para determinação de ferro, alumínio e silício em extrato sulfúrico	58
5.7	Aquisição dos espectros de infravermelho	59
5.8	Aquisição das imagens digitais das amostras de solo	60
5.9	Análise de solo por meio de MIA	61
5.10	Análise de solo por meio de espectros de MIR	62
5.11	Calibração e Validação	64
5.12	Manipulação e tratamento dos dados	64
6	Resultados e Discussões	65

6.1	Descrição estatística das propriedades das amostras de solo	65
6.2	Análise de solo por meio dos métodos de referência . .	66
6.2.1	Análise de solo por meio da MIA	67
6.2.1.1	Quantificação do COS	68
6.2.1.2	Quantificação da textura do solo	75
6.2.1.2.1	Quantificação dos teores de argila	75
6.2.1.2.2	Quantificação dos teores de areia	80
6.2.1.3	Classificação textural por meio de imagens digitais	87
6.2.1.4	Determinação de fósforo, sódio, potássio e micronutrientes disponíveis	88
6.2.1.5	Determinação dos óxidos de ferro, silício e alumínio em solo . . .	88
6.2.1.5.1	Determinação de Fe_2O_3 em amostras de solo	89
6.2.1.5.2	Determinação de SiO_2 em amostras de solo	91
6.2.1.5.3	Determinação de Al_2O_3 em amostras de solo	93
6.2.2	Análise de solo por meio de espectros MIR	94
6.2.2.1	Quantificação do COS	96
6.2.2.2	Quantificação da textura do solo	98
6.2.2.2.1	Quantificação dos teores de argila	98
6.2.2.2.2	Quantificação dos teores de areia	101
6.2.2.3	Classificação textural por meio de espectros de MIR	103
6.2.2.4	Determinação dos óxidos de ferro, silício e alumínio em solo . . .	103
6.2.2.4.1	Determinação de Fe_2O_3 em amostras de solo	104
6.2.2.4.2	Determinação de SiO_2 em amostras de solo	106
6.2.2.4.3	Determinação de Al_2O_3 em amostras de solo	107
6.2.3	Análise de solo por meio de MIA e MIR em relação aos métodos de referência	109
6.2.3.1	Operacionalidade das metodologias de análise de solo	110
6.2.3.2	Segurança dos procedimentos de análise de solo	111
6.2.3.3	Impacto ambiental dos procedimentos de análise de solo	111
6.2.3.4	Custo de implementação dos métodos de análise de solo	112
6.2.4	Análise de solo por meio de MIA e MIR: comparação de resultados	113

7	Conclusões	116
	Referências ¹	117

¹ De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

1 Introdução

A atividade agrícola é um dos setores que mais impulsiona a economia brasileira (EMBRAPA, 2018). Mesmo no período de recessão econômica, a agricultura juntamente com a pecuária foram os únicos ramos que tiveram alta na geração de empregos no país (CNA, 2017). Atualmente, esses dois setores respondem por um terço dos empregos gerados no país (MAPA, 2019). Além disso, o agronegócio já representa quase um quinto do Produto Interno Bruto (PIB) nacional (Agência Brasil, 2019).

Esses resultados são reflexos das constantes altas da produção agrícola nacional. De acordo com a Companhia Nacional de Abastecimento (CONAB), são estimados para a safra de 2019/2020 uma produção de grãos de 246,4 milhões de toneladas dispostos em uma área plantada de 64,1 milhões de hectares (CONAB, 2019a). Isto é, um aumento de quase 2% da produção agrícola e de 1,4% área plantada em relação a safra 2018/2019 (CONAB, 2019b).

Diante desse cenário de desenvolvimento da agricultura, o Brasil deve, de acordo com a Organização para a Cooperação e Desenvolvimento da Europa (OCDE) e a Organização das Nações Unidas para Agricultura e Alimentação (FAO), assumir a primeira posição na produção de produtos agrícolas a partir de 2024 (Estado de Minas, 2015; Globo Rural, 2015). Outra projeção apresentada pelo Ministério de Agricultura, Pecuária e Abastecimento (MAPA), estima para a safra 2026/27 uma produção nacional de grãos superior a 288 milhões de toneladas (MAPA, 2017). Isto é, um aumento de cerca de 15% em relação a produção atual. Nessa safra, o país contará com uma área de plantio estimada em mais de 70,8 milhões de hectares, ou seja, um aumento de quase 10% em relação a área cultivada atual.

Embora a expansão da fronteira agrícola contribua com o aumento da produção, ela pode gerar impactos ambientais como a remoção da vegetação nativa (BEUCHLE et al., 2015), o uso intensivo da água, sobretudo na irrigação (FAO, 1998), e a degradação do solo (FAO, 2015). Nessa direção, o crescimento da área plantada deve ser rigorosamente controlado. Além disso, práticas de agricultura sustentável devem ser incentivadas como, por exemplo, o uso racional da água e insumos e o manejo sustentável do solo (SCOPEL et al., 2013). Nessa perspectiva, solos degradados também devem ser recuperados para que possam tornar-se produtivos novamente. Para tanto, os mesmos necessitam ser avaliados e, em seguida, medidas de correção do solo devem ser aplicadas (WADT, 2003).

A avaliação e monitoramento do solo podem ser realizados por meio das análises físicas, químicas e mineralógicas do solo. Essas análises são ferramentas importantes na determinação do nível adequado das medidas corretivas, assim como na caracterização do solo que possibilita o seu uso e manejo de forma sustentável (SANTOS et al., 2010). Mesmo possibilitando informações valiosas sobre o solo, principalmente de interesse agrícola, alguns métodos de análises de solo apresentam várias desvantagens, como a baixa produtividade, o alto custo de implementação e também a utilização de reagentes químicos, muitas vezes de alta toxicidade. Portanto, o desenvolvimento de metodologias alternativas mais rápidas, de baixo custo e mais limpas, comparadas aos métodos atuais considerados de referência ou padrão, deve ser incentivado.

Diante disso, o objetivo deste trabalho é propor uma metodologia mais rápida, barata e de menor impacto ambiental para quantificação dos parâmetros: carbono orgânico do solo (COS), textura do solo, óxidos de alumínio, silício e ferro extraídos do solo empregando a análise multivariada de imagens (MIA - *Multivariate Image Analysis*) e de espectros de infravermelho médio (MIR - *Mid Infrared*).

2 Objetivos

A proposta deste trabalho envolve o desenvolvimento de novas metodologia para quantificação dos parâmetros de fertilidade do solo: carbono orgânico, textura do solo e dos óxidos de ferro, silício e alumínio, por meio da técnica já consolidada na literatura, a análise multivariada de imagens (MIA), assim como por meio de espectros de infravermelho, com o intuito de reduzir custos, minimizar os impactos ambientais e maximizar a eficiência das análises de solos.

2.1 Objetivos específicos

- Determinação do carbono orgânico do solo;
- Classificação textura do solo;
- Determinação dos óxidos de alumínio, silício e ferro presente no solo;
- Obtenção das imagens digitais e dos espectros de infravermelho de cada amostra de solo;
- Desenvolvimento de método alternativo para determinação do carbono orgânico, textura do solo e dos óxidos de alumínio, silício e ferro presente no solo por meio de imagens digitais obtidas por um *scanner* de mesa e de espectros de infravermelho;
- Comparar as metodologias baseadas em imagens digitais e espectros de infravermelho em relação aos métodos de referência quanto a três requisitos: *a*) operacionalidade, *b*) segurança em laboratório, *c*) impacto ambiental e *d*) custos de implementação.

3 Revisão Bibliográfica

As imagens são ilustrações gráficas de seres inanimados, seres humanos ou qualquer contexto ou cenário real. Elas podem ser obtidas por diversos meios, como câmeras fotográficas e *scanners*. As imagens desempenham um papel fundamental na sociedade, principalmente na comunicação (PRATS-MONTALBÁN; JUAN; FERRER, 2011). No campo científico, a partir do desenvolvimento da tecnologia computacional, as imagens se apresentaram como valiosos recursos da ciência (GELADI et al., 1992). Já na década de 60, as imagens já eram utilizadas no campo da medicina na análise de cromossomos (LEDLEY, 1964). As imagens já foram utilizadas para outros interesses médicos, como a detecção de aneurismas do coração (BOTVINICK et al., 1976) e a detecção de tumores (BALLARD; SKLANSKY, 1973).

A análise de imagens também já foi aplicada no ramo da biologia, na contagem de unidades formadoras de colônias (ALVES, 2006), na classificação de bactérias (ALMEIDA et al., 2014), em medições da dimensão do girassol (SUNOJ et al., 2018) e na avaliação de colite induzida em camundongos (KOZLOWSKI et al., 2013). Em análises de grãos, na determinação de umidade em café (PINTO et al., 2003), na detecção de danos físicos em soja (FLOR et al., 2004) e na avaliação da forma da semente de trigo (WILLIAMS; MUNKVOLD; SORRELLS, 2013). Em análise de tecidos vegetais, na detecção de pragas em folhas de bananeira (SILVA, 2008), na detecção de deformações em cascas de frutas cítricas (LÓPEZ-GARCÍA et al., 2010), na avaliação da qualidade de plumas de algodão (GONÇALVES et al., 2016), na detecção de várias doenças em plantas (BARBEDO; KOENIGKAN; SANTOS, 2016) e na classificação de mamona (VILAR et al., 2015).

Em análise clínicas, como no diagnóstico da dengue (OLIVEIRA et al., 2016), no monitoramento de diabetes e hiperlipidemia (LI et al., 2019), na detecção de *doping* em urina (LUIZ et al., 2019) e na quantificação da glicose, proteína e colesterol (SOUZA; ALVES; COLTRO, 2012). Em meteorologia, na detecção de céu claro (PAWAR et al., 2019). Em engenharia civil, na avaliação de sal úmido em paredes (VÁZQUEZ et al., 2013) e na avaliação da qualidade de argamassas (DZAYE et al., 2019). Em geografia, no sensoriamento remoto e em geoprocessamento (FIGUEIREDO, 2005).

As imagens digitais também já foram aplicadas na indústria, por exemplo, na inspeção de peças (GOLDSTEIN; NAGLER, 1988), no monitoramento de processos e no controle de qualidade (BHARATI; MACGREGOR, 1998; DUCHESNE; LIU; MACGREGOR,

2012; VIDAL et al., 2011). Em arqueologia na otimização de registros fotográficos antigos (MENESES, 2013), em física, no aprimoramento da técnica de microscopia (CRUZ, 2007). No campo da odontologia, as imagens já foram utilizadas na avaliação de epidemiologia dentária (HOGAN et al., 2018).

Em química, as imagens digitais também já foram utilizadas em várias aplicações, como na análise de bebidas, por exemplo, na determinação quantitativa de *Lactobacillus* em leite (BORIN et al., 2007), na determinação de íons de ferro em vinhos brancos (NETO et al., 2019), na quantificação de etanol em bebidas (BENEDETTI et al., 2015), na determinação de cobre em cachaça (PESSOA et al., 2017), na predição do pH de águas potáveis (DAMASCENO et al., 2015), no monitoramento de cloro em água (SUMRIDDETKHAJORN; CHAITAVON; INTARAVANNE, 2014), na estimativa da dureza da água potável (DAMASCENO et al., 2016), na quantificação de sódio e cálcio em leite em pó (LYRA et al., 2014) e na classificação de refrigerantes (GODINHO et al., 2008), chás não processados (DINIZ et al., 2012) e cervejas (SILVA; GODINHO; OLIVEIRA, 2011).

Além de bebidas, as imagens também já foram aplicadas na análise de alimentos (CHEVALLIER et al., 2006; BOTELHO, 2014; FOCA et al., 2011; PATACA, 2006; YU; MACGREGOR, 2003; PACIORNIK et al., 2006; BOTELHO; DANTAS; SENA, 2017; CATELANI et al., 2019), em análises ambientais (CHOODUM; SRIPROM; WONGNIRAMAIAKUL, 2019; FILHO; PETRUCI; CARDOSO, 2015; FIRDAUS et al., 2014; BARROS et al., 2017; FREITAS et al., 2016), em determinações farmacêuticas (GOMES et al., 2008), na avaliação de nanofluidos (LEMES; RABELO; OLIVEIRA, 2017), na determinação de titânio em plástico (LOPEZ-MOLINERO et al., 2010). Em análises forenses, na análise de cocaína (MARCELO et al., 2016), na estimativa da idade de manchas de sangue (THANAKIATKRAI; YAODAM; KITPIPIT, 2013), na quantificação de metanfetamina (CHOODUM et al., 2014), na semi-quantificação de drogas opiáceas (CHOODUM; DAEID, 2011). Além disso, as imagens também já foram utilizadas na análise de óleos lubrificantes (RIBEIRO et al., 2019), na caracterização de membranas de filtração (SUN et al., 2011), na análise de combustíveis (SOARES et al., 2019; SOARES; LIMA; ROCHA, 2017), na caracterização da corrosão do ferro (IGOE; PARISI, 2016) e na detecção e identificação de bactérias em água potável (AL-QADIRI et al., 2006).

Em análise de solos, a utilização de imagens como sinal analítico também já foi reportada na literatura. Oliveira et al. (2013) analisaram quantitativamente Cr (VI) em matrizes de solo por meio de imagens digitais adquiridas por meio de um *scanner* comercial. Choodum et al. (2013) quantificaram teores de trinitrotolueno por meio de imagens

adquiridas por um *smartphone*. [Viscarra-Rossel, Fouad e Walter \(2008\)](#) determinaram alguns parâmetros de fertilidade do solo, como os teores extraíveis de ferro por intermédio de imagens digitais de forma simples e rápida. De forma semelhante, [Chung et al. \(2012\)](#) quantificaram as partículas físicas do solo, argila, silte e areia por meio de histogramas de cor, obtidos de imagens digitais de amostras de solo. [Breul e Gouvers \(2006\)](#) determinaram a granulometria, assim como a mineralogia e a compactação do solo por intermédio de imagens, obtidas pela ferramenta de endoscopia. [Levin, Ben-Dor e Singer \(2005\)](#) avaliaram os teores de óxido de ferro livre em solo utilizando imagens digitais. [Morais et al. \(2019\)](#) determinaram os teores de argila e areia por meio de imagens digitais obtidas por um microscópio estéreo.

Além de imagens digitais, outras metodologias para análise de solo que também apresentam rapidez, baixo custo e com menor impacto ambiental já foram reportadas. [Brunet et al. \(2007\)](#) propuseram um método rápido e de baixo custo para quantificação do carbono total (C_{tot}) e nitrogênio total do solo (N_{tot}) por meio da espectroscopia de reflectância no infravermelho próximo (NIRS). De forma similar, [Madari et al. \(2005\)](#) determinaram os teores de carbono orgânico (COS) e carbono total por meio de NIRS e da reflectância no infravermelho médio (DRIFTS).

[Moron e Cozzolino \(2003\)](#) quantificaram, de modo rápido e de baixo custo, as concentrações dos metais ferro (Fe), cobre (Cu), manganês (Mn) e zinco (Zn) em solo por meio de NIRS. [Vohland et al. \(2014\)](#) determinaram os teores de argila e areia em amostras de solo utilizando NIRS e DRIFT. [Viscarra-Rossel et al. \(2006b\)](#) quantificaram os parâmetros do solo: pH, COS, textura e condutividade elétrica por meio de NIRS, da espectroscopia de reflectância no infravermelho médio (MIRS) e da espectroscopia na região do visível (Vis). De forma similar, [Shepherd e Walsh \(2002\)](#) determinaram várias propriedades do solo, dentre elas, os teores de argila, silte e areia, COS e as concentrações extraíveis de fósforo (P), por meio da regressão multivariada desses parâmetros com a reflectância do solo, mensurada por um espectrômetro portátil.

Nessa mesma linha, [Cécillon et al. \(2009\)](#) avaliaram a qualidade do solo por meio da quantificação de parâmetros físicos, químicos e biológicos do solo utilizando a NIRS. [Stenberg et al. \(2010\)](#) apresentam uma revisão sobre a quantificação de vários atributos do solo, em especial a matéria orgânica, argilominerais, textura, nutrientes, água, pH e metais pesados por meio da espectroscopia na região do visível e do infravermelho próximo (Vis-NIR). De forma similar, [Soriano-Disla et al. \(2014\)](#) apresentam uma revisão sobre a

aplicação de Vis, NIRS e MIRS em análises quantitativas do solo, principalmente para estimar a granulometria, COS, Ntot, pH, teor de água assim como algumas propriedades biológicas do solo. [Nocita et al. \(2015\)](#) descrevem o estado da arte da aplicação da espectroscopia Vis-NIR-MIR em análises de solos como alternativas aos métodos laboratoriais. [Cozzolino \(2015\)](#) apresenta uma revisão bibliográfica da utilização de NIR na análise, classificação e monitoramento de solos de alguns países da América do Sul: Brasil, Colômbia e Uruguai.

Este trabalho apresenta novas metodologias para determinação dos parâmetros de fertilidade bastante requisitados aos laboratórios de solo: carbono orgânico e textura do solo, assim como os óxidos de ferro, silício e alumínio extraídos do solo por meio da análise multivariada de imagens (MIA) e de espectros de infravermelho médio (MIR). Histogramas de ocorrência de cor foram obtidos das imagens digitais, adquiridas utilizando um *scanner* comercial. Espectros de infravermelho médio também foram obtidos para cada amostra de solo. As correlações dos histogramas de ocorrência de cor, assim como dos espectros de infravermelho e os parâmetros de fertilidade mencionados, foram realizadas por meio de três métodos de calibração multivariada, sendo PLS e SPA-MLR, de caracteres lineares e o LS-SVM de caráter não linear.

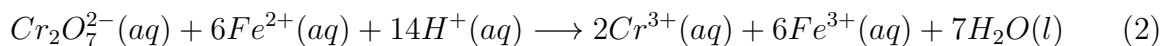
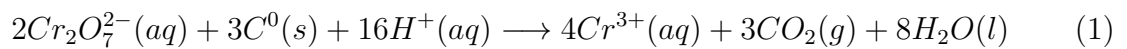
4 Metodologia

4.1 Carbono orgânico do solo

O carbono orgânico do solo (COS) corresponde a todos os compostos orgânicos no solo, incluindo tecidos vegetal e animal em decomposição, assim como os microorganismos vivos ou não (NOVAIS et al., 2007). O COS é um parâmetro de fertilidade de fundamental importância, pois ele influencia várias propriedades do solo, como a sua estruturação (OADES, 1993; BRONICK; LAL, 2005), a cor (VISCARRA-ROSSEL et al., 2006a; BAUMGARDNER et al., 1969) e a retenção de água (LADO; PAZ; BEN-HUR, 2004). Diante da relevância desse parâmetro, sua análise é largamente solicitado aos laboratórios de análise de fertilidade de solos.

4.1.1 Quantificação do carbono orgânico do solo

A quantificação do carbono orgânico do solo é realizada geralmente por duas formas: (i) análise elementar, que consiste na combustão da amostra de solo, seguida da determinação do carbono orgânico a partir da diferença entre os teores de carbono total e carbono inorgânico (MA, 2001) e (ii) com o emprego de métodos úmidos, que consistem na oxidação do carbono orgânico pelo dicromato de potássio ($Cr_2O_7^{2-}$) em meio ácido, Eq. (1), seguido da determinação do excesso do oxidante por meio da titulação com sulfato de Ferro ($FeSO_3$), Eq. (2).



Embora os métodos úmidos apresentem inconvenientes, como a baixa operacionalidade e a alta geração de resíduos, eles são mais utilizados que a análise elementar, pois são mais simples e apresentam um menor custo de implementação. Os métodos úmidos mais utilizados são Walkley-Black (WALKLEY; BLACK, 1934), Mebius (MEBIUS, 1960) e o método da Embrapa (SANTOS et al., 2010). Esses métodos se diferenciam pela presença, ou não, de aquecimento externo. As metodologias de Mebius e da Embrapa utilizam aquecimento externo com o intuito de otimizar a oxidação do carbono orgânico. Já no método Walkley-Black, em vez do aquecimento aplica-se um fator de correção de 1,33 ao

teor calculado de carbono orgânico com a finalidade de compensar o aquecimento (PANSU; GAUTHEYROU, 2003).

Esses três métodos baseiam-se na titrimetria. Logo ao final do ensaio, um volume considerável de resíduo é gerado e deve ser tratado antes do descarte. Além disso, essas metodologias são de baixa operacionalidade, dessa forma é possível analisar somente um número limitado de 100 amostras, em média, a cada dia. Diante disso, Souza et al. (2016) desenvolveram uma adaptação para cada um desses três métodos de modo a otimizar os ensaios e minimizar os resíduos gerados por meio da leitura espectrofotométrica em vez da titulação. A adaptação desses métodos além de ser uma alternativa mais econômica, possibilitou uma redução de pelo menos 90% do volume de resíduo gerado.

4.2 A textura do solo

A textura do solo também denominada como granulometria, representa a distribuição das partículas físicas do solo de tamanhos inferior a 2mm. Essas partículas são classificadas em três grupos de acordo com seus tamanhos, areia (0,05 a 2 mm), silte (0,002 a 0,05 mm) e argila (menor que 0,002 mm) (GEE; BAUDER; KLUTE, 1986). Desse modo, as amostras de solo podem ser classificadas de acordo com as proporções de cada partícula. Geralmente, essa classificação é realizada utilizando o diagrama textural apresentado pela Figura 1. Tal sistema de classificação foi adotado pelo Departamento de Agricultura dos Estados Unidos (DAVIS; BENNETT et al., 1927), assim como pela Sociedade Brasileira de Ciência do Solo (SOUZA; CELLIGOI, 2012).

A textura do solo é um parâmetro de fertilidade que afeta principalmente a retenção de água (SALTER; WILLIAMS, 1965), o teor de matéria orgânica (PLANTE et al., 2006), os estoques de nutrientes (SILVER et al., 2000) e a cor (VISCARRA-ROSSEL et al., 2006a; KONEN; BURRAS; SANDOR, 2003a).

4.2.1 Quantificação da textura do solo

A quantificação da textura do solo é realizada geralmente por duas metodologias, (*i*) método da pipeta (ROBINSON, 1922) e método do hidrômetro (BOUYOCOS, 1926). Ambos métodos consistem inicialmente na dispersão física e química da amostra de solo em meio aquoso, seguido da separação e quantificação dos três grupos texturais. A dispersão química



Figura 1 – Diagrama triangular utilizado na classificação textural do solo.

consiste na adição de reagentes com o intuito de eliminar qualquer tipo de agregantes de argila, especialmente a matéria orgânica (MAURI et al., 2011). A dispersão física consiste na desagregação das partículas por meio de choques mecânicos, utilizando, por exemplo, mesa agitadora (DOURADO; SILVA; MARINHO, 2012) ou equipamento de ultrassom (GENRICH; BREMNER, 1972).

Após a dispersão, realizam-se a separação da areia por peneiramento, e das frações de argila e silte por meio da sedimentação. A determinação da argila para o método do hidrômetro é realizada por meio da medida da densidade, já para o método da pipeta, por meio da determinação da massa de argila contida em uma alíquota coletada da superfície (PANSU; GAUTHEYROU, 2003).

As duas metodologias citadas, embora amplamente utilizadas principalmente no Brasil devido ao baixo custo de implementação, apresentam algumas desvantagens como, por exemplo, a geração de resíduos e a baixa produtividade (CORÁ et al., 2009).

4.3 Óxidos de alumínio, silício e ferro em solo

O silício é um dos elementos mais abundantes presentes no solo (KOVIDA, 1973), entretanto, mesmo em alta quantidade esse elemento não é essencial para o desenvolvimento das plantas. Por outro lado, o silício pode apresentar diversos efeitos benéficos em várias espécies de plantas (MARSCHNER, 1995), especialmente as gramíneas, como o arroz (MA et al., 2006) e a cana de açúcar (MATICHENKOV; CALVERT, 2002). O silício está presente no solo geralmente na forma de óxido, SiO_2 . Esse óxido corresponde a pelo menos 50% da massa total do solo (MA, 2005).

Depois do oxigênio e do silício, o alumínio é o elemento mais abundante no solo (WILLARD, 1979). No solo, esse elemento está presente na forma de óxidos e aluminossilicatos (MA, 2005). Esses compostos, embora presentes em altas concentrações, não são nocivos para as plantas. Por outro lado, em solos ácidos ($pH < 5$) o ânion Al^{3+} , um componente danoso ao desenvolvimento das plantas é formado em soluções do solo (KOCHIAN, 1995).

Os óxidos de ferro são componentes facilmente encontrados em solos. Esses óxidos correspondem a um importante parâmetro de fertilidade e por isso suas determinações são bastantes requeridas (SCHWERTMANN; TAYLOR, 1989). Os teores dos óxidos de ferro além de auxiliarem na classificação do solo (CAMARGO; KLAMT; KAUFFMAN, 1987), também influenciam outros parâmetros, especialmente a cor do solo (SCHWERTMANN, 1993; TORRENT et al., 1983; BARRÓN; TORRENT, 1986). Os teores de óxidos de ferro, assim como de alumínio também influenciam a estruturação e algumas reações iônicas no solo (MELO et al., 2001).

Por meio da quantificação dos teores dos óxidos de silício, alumínio e ferro, é possível avaliar o grau de intemperização dos solos, assim como a mineralogia das partículas de argila (UEHARA, 1988; TEIXEIRA et al., 2017).

4.3.1 Quantificação dos óxidos de alumínio, silício e ferro em solo

Anteriormente a quantificação dos óxidos de alumínio e ferro em solo, realiza-se o ataque sulfúrico, que consiste na extração desses componentes utilizando ácido sulfúrico concentrado. Após a extração e filtração, os teores de óxidos de ferro e alumínio são determinados utilizando o extrato sulfúrico. Já o teor de silício é determinado utilizando o resíduo sólido retido na etapa de filtração. Para tanto, os óxidos de silício são solubilizados

utilizando uma solução de hidróxido de sódio. Posteriormente a extração e filtração, determina-se os teores de óxido de silício na solução filtrada.

Os óxidos de alumínio presentes no extrato sulfúrico são normalmente determinados por meio de titulação (MCLEAN, 1965), colorimetria (WOLF, 1982), assim como por meio de métodos espectrofotométricos (OTOMO, 1963). Os teores de óxidos de ferro são geralmente quantificados por meio da espectrofotometria de absorção molecular e atômica (FREITAS et al., 2015). Já a determinação dos teores de óxido de silício no solo é realizada, de forma geral, utilizando espectrofotômetro de absorção molecular (MORRISON; WILSON, 1963), assim como por meio de espectrofotômetro de absorção atômica (TEIXEIRA et al., 2017).

4.4 Imagens Digitais

Uma imagem pode ser descrita como uma função matemática $f(x, y)$, aonde cada ponto dessa função corresponde a seu respectivo nível de cinza (GELADI; GRAHN, 1996). A função bidimensional $f(x, y)$ é uma combinação de dois parâmetros: (a) *iluminância*, $i(x, y)$, que corresponde a intensidade de luz incidente, e (b) *reflectância*, $r(x, y)$, que corresponde à intensidade de luz refletida. Essa função está representada pela Eq. (3).

$$f(x, y) = i(x, y).r(x, y) \quad (3)$$

Dessa forma, a componente iluminância é dependente da forma ou tipo de iluminação e a reflectância depende das propriedades dos elementos ou objetos presentes no cenário, conforme ilustrado na Figura 2 (GONZALEZ; WOODS, 2010).

A obtenção de uma imagem consiste na transformação de um cenário real de três dimensões em uma imagem de duas dimensões. Essa etapa é realizada geralmente utilizando um equipamento de CCD (*Charge Coupled Device*), que gera uma imagem analógica a partir do sinal elétrico produzido em função da intensidade de luz captada no sensor. Esse dispositivo está presente na maioria das câmeras digitais, assim como em outros equipamentos de captura de imagens (GONZALEZ; WOODS, 2010).

Posteriormente à aquisição da imagem, é possível construir uma imagem digital por meio de dois processos: (a) amostragem, que digitaliza as coordenadas x e y da função $f(x, y)$, e (b) quantização, que digitaliza os valores da função $f(x, y)$ (GONZALEZ; WOODS, 2010). Em suma, essas duas etapas transformam a imagem analógica em uma matriz bidimensional. Dessa forma, cada elemento da matriz, denominado *pixel*, corresponderá

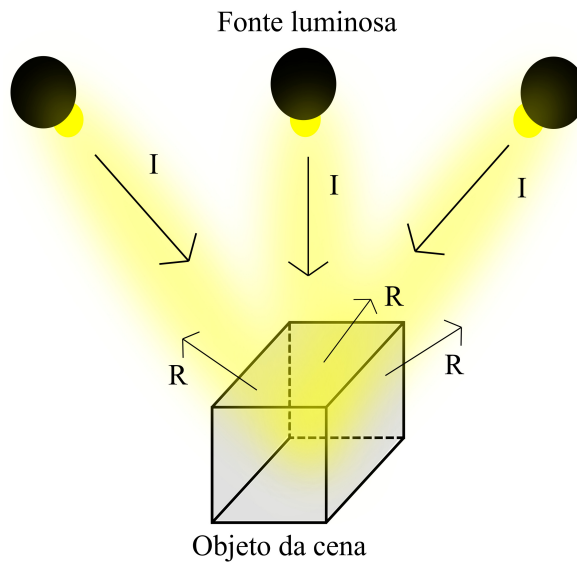


Figura 2 – Representação das componentes iluminância (I) e reflectância (R).

à um valor inteiro de 0 a $2^n - 1$ referente à intensidade de nível cinza naquele ponto (BHABATOSH; MAJUMDER, 2011). Após essas duas etapas, a imagem digital pode ser armazenada e manuseada em meios computacionais.

De maneira análoga às imagens de duas dimensões (monocromáticas), as imagens coloridas (tridimensionais) também são descritas matematicamente por uma função tridimensional $f(x, y, z)$. As imagens coloridas também podem ser digitalizadas, convertendo-se em imagens digitais, por meio dos processos de amostragem e quantização. Após a digitalização, em vez de uma matriz bidimensional, como no caso de imagens monocromáticas, obtêm-se três matrizes sendo uma para a imagem monocromática vermelha, uma para imagem verde e outra para imagem azul. Assim, o produto dessa combinação é uma imagem RGB (*Red Green Blue*) (GELADI; GRAHN, 1996).

4.4.1 Cor

A luz visível é uma forma de radiação eletromagnética perceptível ao ser humano. A faixa visível compreende seis cores: violeta, azul, verde, amarelo, laranja e vermelho. Essa faixa visível se inicia no comprimento de onda de 435,8 nm referente à cor violeta e finaliza no comprimento de onda de 700 nm referente à cor vermelha, Figura 3 (LU, 1997).

O reconhecimento da cor de um objeto pelo ser humano é resultado da interação da luz refletida pelo objeto com o olho humano. Nessa direção, se um objeto apresenta

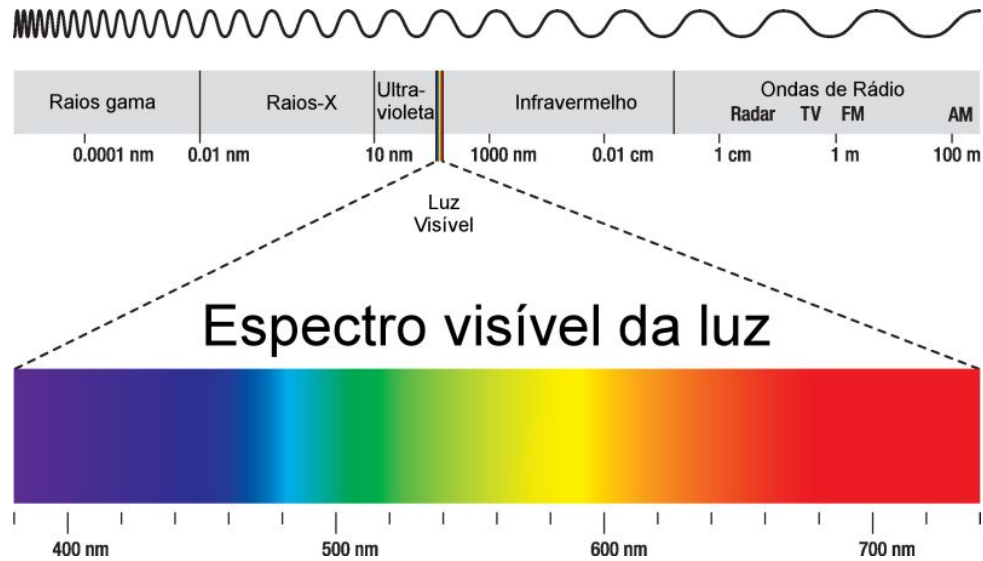


Figura 3 – Espectro eletromagnético.

Fonte: Info Escola <<http://www.infoescola.com/fisica/espectro-eletromagnetico/>>. Acesso em 26 dez. 2017

alguma tonalidade de cor, significa que o mesmo reflete preferencialmente um intervalo específico do espectro visível e absorve o restante da radiação. Por outro lado, se o objeto refletir todos os comprimentos de onda de forma igualitária, o mesmo apresentará a cor branca (GONZALEZ; WOODS, 2010; BHABATOSH; MAJUMDER, 2011).

As cores, de modo geral, podem ser distinguidas por meio de três parâmetros: (a) *brilho*, que corresponde à intensidade da radiação luminosa, (b) *matiz*, que corresponde ao comprimento de onda com maior predominância e (c) *saturação*, que indica a pureza da matiz ou da cor. A cor verde, por exemplo, apresenta saturação máxima. Já as cores lilás e violeta, embora apresentem o mesmo matiz, apresentam diferentes níveis de saturação, pois a cor lilás é resultado da combinação de violeta e branco (GONZALEZ; WOODS, 2010; BHABATOSH; MAJUMDER, 2011).

4.4.2 Modelos de cores

Os modelos de cores são sistemas tridimensionais que possibilitam a organização das cores de forma padronizada. Nesses sistemas de três dimensões cada ponto corresponde a uma cor específica. Os modelos de cores são utilizados principalmente no processamento de imagens e em dispositivos físicos, como monitores coloridos (HUGHES et al., 2013; BHABATOSH; MAJUMDER, 2011).

Os modelos de cores mais conhecidos e de maior aplicabilidade são o RGB (*Red, Green, Blue*), CMY (*Cyan Magenta Yellow*), YIQ (*Luminance In-phase Quadrature*), HSI (*Hue Saturation Intensity*) e YCbCr (GONZALEZ; WOODS, 2010; BHABATOSH; MAJUMDER, 2011).

4.4.2.1 Modelo RGB

O modelo RGB é o sistema de cor de ampla utilização principalmente em dispositivos de obtenção de imagens, como câmeras e *scanner*. A representação geométrica desse modelo é um cubo, Figura 4, onde cada vértice define uma cor. As cores preta e branca se localizam, respectivamente, nos vértices da origem e no mais distante. Os outros seis vértices correspondem às cores primárias, vermelho, verde e azul e às cores secundárias, amarelo, ciano e magenta. A linha que se projeta desde a origem, referente à cor preta até o ponto correspondente a cor branca, representa a escala cinza (GONZALEZ; WOODS, 2010; BHABATOSH; MAJUMDER, 2011).

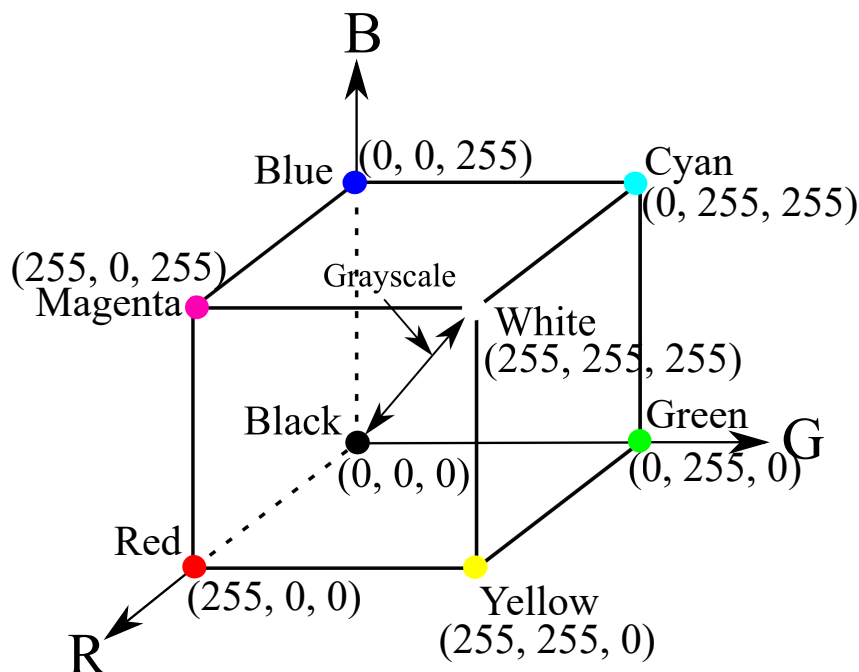


Figura 4 – Sistema RGB de cores.

O modelo RGB pode ser representado de duas formas: (a) *normalizado*, onde cada canal de cor, vermelho, verde e azul está contido dentro da faixa de 0 a 1 e (b) *não normalizado*, onde cada canal de cor possui 256 índices de cores, limitado ao intervalo, de números inteiros, de 0 a 255. Nesse último, as cores são formadas a partir da combinação

dos três canais. Dessa forma, um total de 16777216 (256^3) possibilidades de cores é possível. O sistema RGB não normalizado é mais utilizado em processamento de imagens (GELADI; GRAHN, 1996).

4.4.2.2 Modelo HSI

O modelo HSI corresponde às três componentes que caracterizam a cor, matiz, saturação e intensidade. Como já mencionado, esses parâmetros também são percebidos pelo ser humano. Dessa forma esse modelo de cor é largamente aplicado em sistemas de processamento de imagens baseados em cores inerentes à visão humana (GONZALEZ; WOODS, 2010).

De forma semelhante ao modelo RGB, o sistema HSI também pode ser representado como um sólido geométrico, Figura 5. A partir das representações geométricas, algumas equações podem ser deduzidas, Eqs. (4), (5), (6), com o intuito de converter o modelo RGB no sistema HSI (GONZALEZ; WOODS, 2010).

$$I = \frac{1}{3}(R + G + B) \quad (4)$$

$$S = 1 - \frac{3}{(R + G + B)}[\min(R, G, B)] \quad (5)$$

$$H = \cos^{-1} \left(\frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right) \quad (6)$$

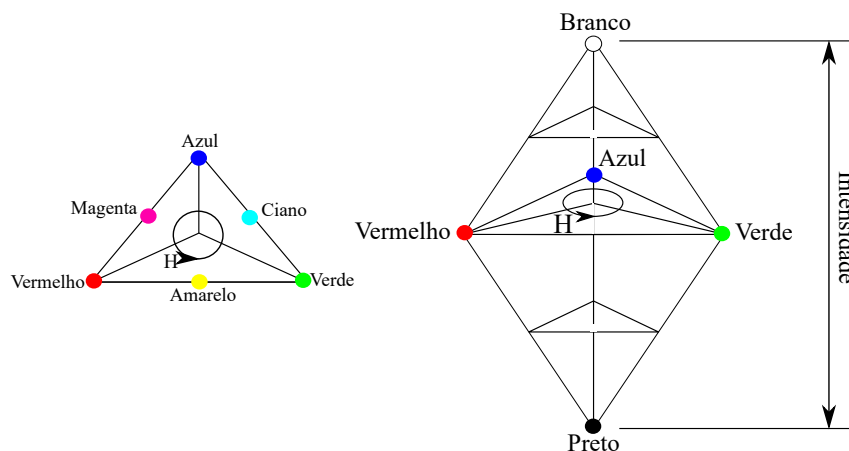


Figura 5 – Modelo de cores HSI.

4.5 Análise Multivariada de imagens

A Análise Multivariada de Imagens (MIA) consiste, primeiramente, na utilização ou na extração de informação multivariada de uma imagem, seguida da aplicação de ferramentas matemáticas e estatísticas com o intuito de classificar, realizar análises exploratórias ou até mesmo estimar algum parâmetro de interesse (GELADI et al., 1992). Previamente a extração de dados das imagens, realiza-se a seleção da região de interesse (ROI), isto é, a parcela da imagem com informações relevantes. Neste trabalho, após a seleção da região de interesse, extraíram-se histogramas de ocorrência de cor como etapa de extração de dados de imagens.

4.5.1 Histogramas de ocorrência de cor

Os histogramas de ocorrência de cor correspondem à distribuição de níveis de cor e suas proporções em uma imagem. Dessa forma, é possível observar para cada nível de cor a quantidade de pixels em uma imagem com tal intensidade. Um histograma de ocorrência pode ser ilustrado como um gráfico de barras, de modo que o eixo x representa os níveis de cinza e o eixo y a quantidade de pixels referente a cada nível (PASS; ZABIH, 1999). A Figura 6 apresenta um exemplo de um histograma de ocorrência de cor de uma imagem com dez níveis de cores.

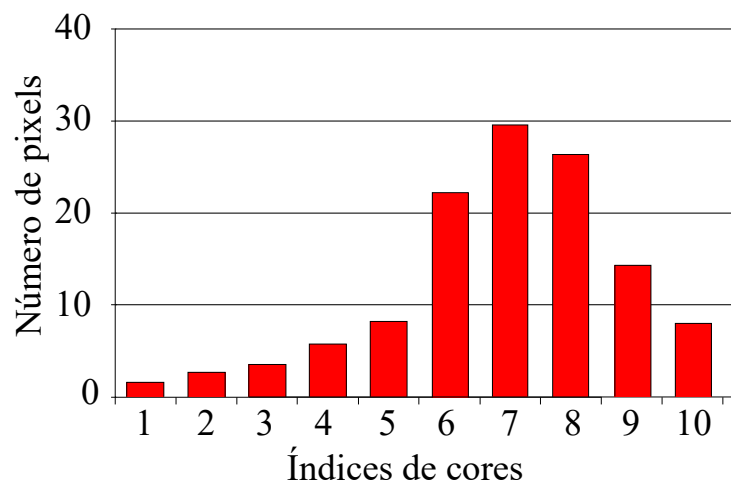


Figura 6 – Exemplo de histograma de uma imagem dispendo de dez níveis de cores.

4.6 Espectroscopia no infravermelho

A espectroscopia no infravermelho (IR) é uma técnica rápida, não destrutível e ainda apresenta facilidade na preparação da amostra. Essa técnica é largamente utilizada em determinações qualitativas e quantitativas de compostos orgânicos e inorgânicos. De modo geral, essa ferramenta é empregada principalmente na elucidação de substâncias orgânicas, especialmente, na detecção de grupos funcionais (HOLLER; SKOOG; CROUCH, 2009).

A região do infravermelho está compreendida entre a radiação visível e a radiação microondas, Figura 3. Essa região que se estende aproximadamente de 700 a 50000 nm (14290 a 200 cm^{-1}) é subdividida em três faixas: *a*) infravermelho próximo (NIR - *near infrared radiation*), *b*) infravermelho médio (MIR - *medium infrared radiation*) e *c*) infravermelho distante (FIR - *far infrared radiation*). A Tabela 1 apresenta o intervalo espectral de cada faixa (HOLLER; SKOOG; CROUCH, 2009).

Tabela 1 – Faixas espectrais do infravermelho.

Região	Número de onda (cm^{-1})	Comprimento de onda (nm)
NIR	12800 a 4000	780 a 2500
MIR	4.000 a 200	2500 a 5000
FIR	200 a 10	5000 a 100000

Fonte: (HOLLER; SKOOG; CROUCH, 2009)

Geralmente, a unidade utilizada para se referir a região do infravermelho é o número de onda, expressa em cm^{-1} . A conversão entre o comprimento de onda (λ) em número de onda pode ser obtido pela Eq. 7 (SUART, 2004).

$$\bar{\nu}(\text{cm}) = \frac{1}{\lambda(\text{cm})} \quad (7)$$

O número de onda, $\bar{\nu}$ (cm^{-1}), pode ser convertido em frequência, $\bar{\nu}$ (Hz), simplesmente multiplicando-o pela velocidade da luz, c (cm s^{-1}), Eq. (8) (SUART, 2004).

$$\bar{\nu}(\text{Hz}) = \nu c (\text{cm s}^{-1}) \quad (8)$$

A utilização da unidade número de onda é preferível devida a sua relação direta com a energia, Eq. (9)

$$E = h\nu \quad (9)$$

onde E é a energia e h é a constante de Planck (SUART, 2004).

4.6.1 Absorção de radiação na região do infravermelho

Os átomos nas moléculas, mesmo nos seus estados fundamentais, sempre estão em movimento. Essa movimentação corresponde aos modos de vibração dos átomos, na deformação e no estiramento das ligações, assim como pela rotação da molécula. Dessa maneira, quando as moléculas absorvem radiação no infravermelho elas são excitadas e dessa forma atingem estados vibracionais e rotacionais de maior energia. Em outras palavras, ocorre um aumento da amplitude do movimento vibracional das ligações, assim como da energia cinética da molécula (HOLLER; SKOOG; CROUCH, 2009; SUART, 2004).

4.6.2 Aplicações da espectroscopia no infravermelho

Cada ligação química apresenta um perfil de vibração. Entretanto, essa vibração pode ser influenciada pela estrutura da molécula. Desse modo, mesmo que uma molécula apresente uma mesma ligação que outra, os perfis espectrais podem ser diferentes devido a diferença de estrutura entre as moléculas. Nesse sentido, a espectroscopia no IR é uma boa alternativa para comparação de moléculas que aparentemente são idênticas, pois espectros de moléculas diferentes jamais serão iguais. Além disso, a espectroscopia no IR pode ser utilizada para elucidação de estruturas, especialmente de moléculas orgânicas (PAVIA *et al.*, 2010).

Dentro da região de infravermelho, a faixa de maior interesse para análise de compostos orgânicos, especialmente, na identificação de grupos funcionais, corresponde ao MIR (4000 a 400 cm^{-1}). Os espectros obtidos por MIR são de mais fácil interpretação, pois apresentam, de forma geral, picos bem definidos e de alta intensidade. Além disso, a varredura em MIR apresenta uma região conhecida como impressão digital. O perfil dessa região é bem característica para cada molécula. Assim, um composto pode ser identificado por meio da comparação entre seu espectro e um banco de dados espectral (HOLLER; SKOOG; CROUCH, 2009).

Além da comparação e elucidação de estruturas, a espectroscopia no infravermelho também é utilizada para análises quantitativas. Para tanto, é necessário a utilização de ferramentas multivariadas, como a calibração multivariada.

4.7 Calibração Multivariada

A calibração Multivariada é uma área da quimiometria que consiste na correlação de algum dado analítico de caráter multivariado com algum atributo de interesse (MARTENS; NÆS, 1991). Para tanto, primeiramente um modelo matemático é gerado relacionando o sinal analítico com alguma propriedade físico-química. Após essa etapa denominada de calibração, o modelo matemático construído é utilizado em um novo conjunto de amostras com o intuito de avaliar a aplicabilidade do mesmo, essa etapa é denominada de validação (BEEBE; KOWALSKI, 1987; KRAMER, 1998).

As etapas de calibração e validação utilizam conjuntos de amostras distintos selecionados de um grupo único de amostras. Dessa forma, com o intuito de evitar tendenciamentos na escolha das amostras, algumas metodologias podem ser empregadas na seleção de amostras. Dentre essas metodologias, destacam-se o algoritmo Kennard-Stone (KS) (KENNARD; STONE, 1969) e o algoritmo SPXY (*Sample set Partitioning based on joint $x - y$ distances*) (GALVÃO et al., 2005).

Posteriormente à separação das amostras entre calibração e validação, ferramentas de calibração multivariadas são empregadas. Dentre elas destacam-se a Regressão Linear Múltipla (MLR - *Multivariate Linear Regression*) (DRAPER; SMITH, 2014), regressão por componentes principais (PCR - *Principal Components Regression*) (BRERETON, 2003), regressão por mínimos quadrados parciais (PLSR - *Partial Least Squares Regression*) (WOLD; SJOSTROM, 2001) e a Máquina de Vetor Suporte por mínimos quadrados (LS-SVM - *Least Squares Support Vector Machines*) (SUYKENS et al., 2002).

4.7.1 Organização dos dados

Anteriormente a calibração multivariada, realiza-se a organização dos dados. Nesta etapa, organiza-se os dados em matrizes e vetores. Os dados analíticos das amostras, por exemplo, histogramas de ocorrência de cor, são dispostos em uma matriz. Nesta matriz cada linha corresponde a uma amostra e cada coluna a uma variável, Eq. (10). Já as

propriedades de interesse, como por exemplo, os teores de carbono orgânico do solo são dispostos como um vetor coluna, onde cada linha corresponde a uma amostra.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (10)$$

4.7.2 Pré-processamento de dados

O pré-processamento dos dados é uma etapa realizada anteriormente à aplicação dos métodos multivariados e consiste no aprimoramento dos dados analíticos como, por exemplo, na remoção de ruídos ou interferentes, com o intuito de otimizar a calibração. As técnicas de pré-processamento são operações matemáticas aplicadas ao sinal analítico. Dentre as técnicas mais utilizadas destacam-se o autoescalamento, centrar na média e a normalização (OTTO, 1999; RODIONOVA; POMERANTSEV, 2007).

Neste trabalho, somente o pré-processamento de dados: centrar na média foi utilizado.

4.7.2.1 Centrar na Média

O pré-processamento centrar na média é uma operação que altera as posições dos dados em direção ao centro dos dados, Figura 7. Esse método consiste primeiramente na determinação da média para cada variável, Eq. (11).

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{i,j} \quad (11)$$

onde n refere-se à quantidade de amostras e \bar{x}_j corresponde à média da variável.

Em seguida, realiza-se a diferença entre a variável inicial e o valor médio previamente calculado, Eq. (12) (OTTO, 1999).

$$\mathbf{xcm}_{i,j} = x_{i,j} - \bar{x}_{i,j} \quad (12)$$

onde $\mathbf{xcm}_{i,j}$ corresponde ao valor centrado na média específico à variável j e amostra i .

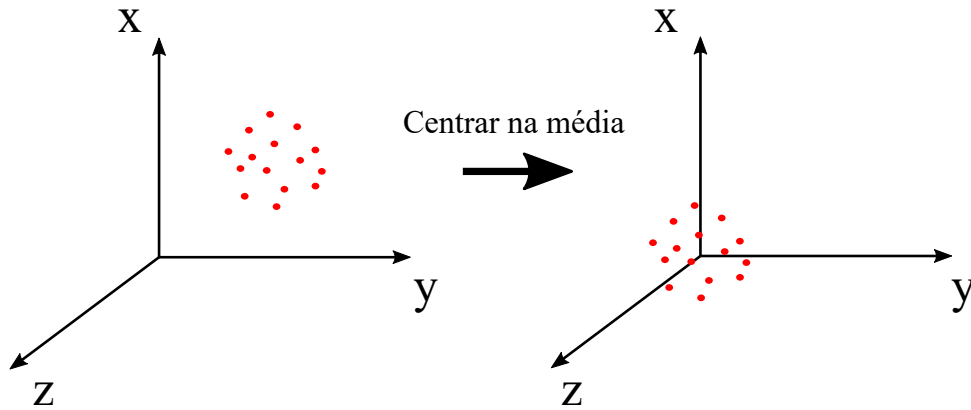


Figura 7 – Esquemática da etapa de pré-processamento centrar na média.

4.7.3 Algoritmo Kennard-Stone

O algoritmo Kennard-Stone (KS) (KENNARD; STONE, 1969) realiza a seleção de amostras representativas a partir dos dados analíticos do conjunto total de amostras (matriz \mathbf{X}). Esse algoritmo consiste na seleção de amostras em função das suas distâncias Euclidianas. Primeiramente, o par de amostras com maior distância Euclidiana entre si, Eq. (13), é determinado, posteriormente se determina e seleciona a amostra com a maior distância Euclidiana em relação a qualquer uma das primeiras amostras determinadas. A seleção do número de amostras especificado é obtido com a repetição do mecanismo descrito.

$$d_x(p, q) = \sqrt{\sum_{j=1}^J [x_p(j) - x_q(j)]^2}; \quad p, q \in [1, n] \quad (13)$$

onde $d_x(p, q)$ é a distância Euclidiana.

A finalidade do algoritmo KS é selecionar, a partir de um conjunto de n amostras, um subconjunto representativo, que por sua vez seja uniforme por toda a extensão da matriz \mathbf{X} , que contém as informações das amostras. Para tanto, o KS adota uma metodologia que primeiramente seleciona um par de amostras (p, q) , com a maior distância Euclidiana, $d_x(p, q)$, entre si, Eq. (13) e, em seguida, seleciona a amostra que apresenta a maior distância entre uma das amostras recém selecionadas. Esse processo se repete gradualmente até a seleção da quantidade de amostras desejadas.

4.7.4 Algoritmo SPXY

O algoritmo SPXY (GALVÃO et al., 2005) realiza a seleção de amostras representativas a partir dos seus respectivos dados analíticos, matriz \mathbf{X} , e do vetor \mathbf{y} que armazena os parâmetros de interesse. Esse algoritmo seleciona as amostras em função das suas distâncias Euclidianas. Primeiramente, o SPXY calcula as distâncias entre cada amostra, para a matriz \mathbf{X} , Eq. (14).

$$d_x(p, q) = \sqrt{\sum_{j=1}^J [x_p(j) - x_q(j)]^2}; \quad p, q \in [1, n] \quad (14)$$

onde $d_x(p, q)$ é a distância Euclidiana referente ao bloco \mathbf{X} .

Em seguida, o mesmo procedimento é realizado para o vetor \mathbf{y} , Eq. (15).

$$d_y(p, q) = \sqrt{(y_p - y_q)^2} = |y_p - y_q|; \quad p, q \in [1, n] \quad (15)$$

onde $d_y(p, q)$ é a distância Euclidiana referente ao bloco \mathbf{y} .

Posteriormente, realiza-se a soma das duas distâncias calculadas divididas pelos seus valores máximos, Eq. (16).

$$d_{xy}(p, q) = \frac{d_x(p, q)}{\max_{p, q \in [1, n]} d_x(p, q)} + \frac{d_y(p, q)}{\max_{p, q \in [1, n]} d_y(p, q)}; \quad p, q \in [1, n] \quad (16)$$

Assim, o SPXY identifica o par de amostras com a maior distância, $d_{xy}(p, q)$, entre si, e em seguida seleciona a amostra com a maior distância Euclidiana em relação a qualquer uma das primeiras amostras determinadas. A seleção do número de amostras especificado é obtido com a repetição do procedimento descrito.

4.7.5 Regressão Linear Múltipla

A Regressão Linear Múltipla (MLR - *Multivariate Linear Regression*) é um método de calibração multivariada de caráter linear bastante utilizado principalmente pela sua simplicidade. A Eq. (17) apresenta a correlação da matriz \mathbf{X} com o vetor \mathbf{y} . A Figura 8 apresenta a representação da MLR.

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e} \quad (17)$$

onde \mathbf{e} corresponde aos resíduos e \mathbf{b} corresponde aos coeficientes de regressão.

O vetor \mathbf{b} , Eq. (17), é determinado pelo método dos mínimos quadrados, Eq. (18).

$$\mathbf{b} = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{y} \quad (18)$$

Embora a MLR seja um método simples, ela só é recomendada para conjunto de dados com pouco ruído e que também não apresente multicolinearidade. Além disso, a aplicação da MLR é limitada a conjunto de dados com maior número de amostras do que variáveis, de outro modo a Eq. (18) não apresentaria solução, pois a matriz \mathbf{X} não seria invertível (MARTENS; NÆS, 1991).

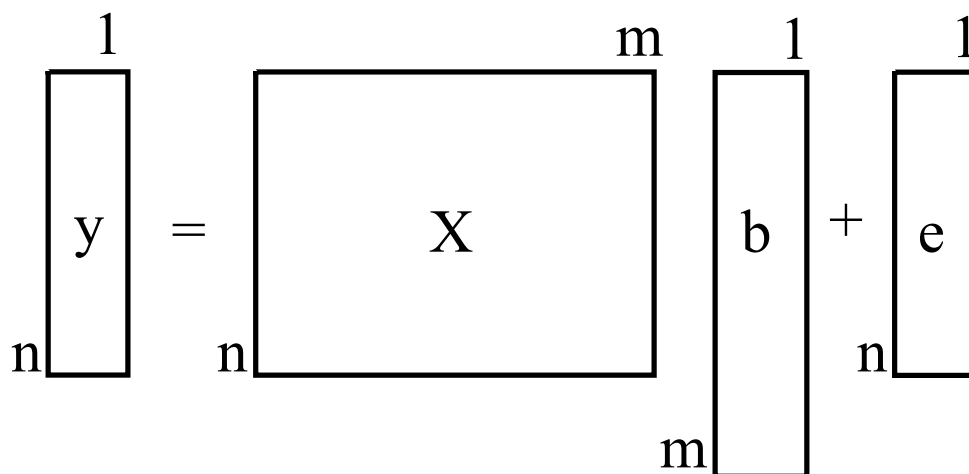


Figura 8 – Representação da MLR.

4.7.6 PLS

O método PLS (WOLD; SJOSTROM, 2001) é uma ferramenta multivariada utilizada em diversas aplicações. Este método relaciona a matriz \mathbf{X} e o vetor \mathbf{y} de modo a obter a máxima relação entre ambos. Para isso, a matriz \mathbf{X} e o vetor \mathbf{y} são decompostos em variáveis latentes (LVs), Eqs. (19) e (20). A Figura 9 apresenta a decomposição da matriz \mathbf{X} e do vetor \mathbf{y} .

$$\mathbf{X} = \mathbf{TP}^t + \mathbf{E} \quad (19)$$

$$\mathbf{y} = \mathbf{uq}^t + \mathbf{f} \quad (20)$$

onde \mathbf{T} e \mathbf{u} são respectivamente a matriz e o vetor de escores referentes a matriz \mathbf{X} e o vetor \mathbf{y} . \mathbf{P} e \mathbf{q} são os pesos dos blocos \mathbf{X} e \mathbf{y} , \mathbf{E} e \mathbf{f} são respectivamente a matriz e o vetor de resíduos da matriz \mathbf{X} e do vetor \mathbf{y} .

Cada coluna da matriz de escores, \mathbf{T} , referente à matriz \mathbf{X} , corresponde a uma variável latente (LV). Assim, após a decomposição, o PLS realiza a correlação dos escores da matriz \mathbf{X} e do vetor \mathbf{y} , \mathbf{T} e \mathbf{u} , Eq. (21). Desse modo, as estimações são realizadas por meio da Eq. (22)

$$\mathbf{u} = \mathbf{b}_h + \mathbf{t}_h \quad (21)$$

onde h corresponde ao número de LVs e \mathbf{b} aos coeficientes de regressão referentes ao número de LVs.

$$\mathbf{y} = \mathbf{T}\mathbf{b}\mathbf{q}^t \quad (22)$$

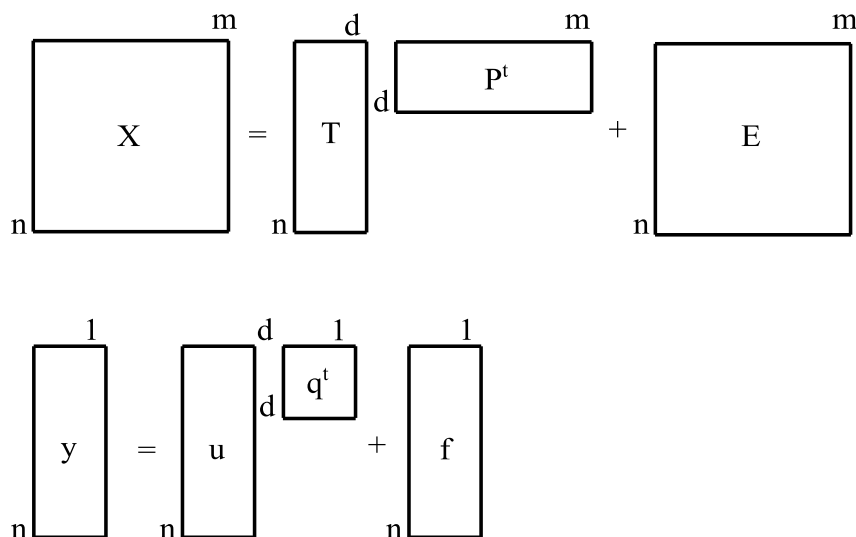


Figura 9 – Representação da decomposição da matriz \mathbf{X} e do vetor \mathbf{y} pelo método PLS

4.7.7 Seleção de LVs

A seleção do número de LVs a serem incluídas em um modelo PLS deve ser realizada de modo a evitar tanto o sobreajuste quanto o subajuste. Um modelo sobreajustado é eficaz na modelagem das amostras de calibração, porém é ineficiente na modelagem das amostras de validação. Já um modelo subajustado é pouco satisfatório para a modelagem das amostras de calibração. Diante disso, algum critério ou método de escolha de LVs deve ser utilizado. Uma ferramenta bastante utilizada nesse propósito é a validação cruzada (CV) que consiste na remoção de uma ou mais amostra do conjunto total, seguido da construção de um modelo de calibração. Em seguida, o modelo construído é empregado na estimação

de algum parâmetro de interesse da amostra ou grupo de amostras removidas. Todo esse mecanismo se repete até a estimação de todas as amostras pelo modelo (MARTENS; NÆS, 1991).

Após a previsão de todas as amostras, determina-se o erro médio quadrático da validação cruzada (RMSECV) para cada LV, Eq. (23)

$$RMSECV = \sqrt{\frac{\sum_{i=1}^n (y_{ref} - y_{prev})^2}{n_{cal}}} \quad (23)$$

onde y_{ref} e y_{prev} correspondem respectivamente aos valores de referência e estimados e n_{cal} o número de amostras de calibração.

O número de LVs adequado é selecionado, geralmente, de acordo com o menor valor de RMSECV.

4.7.8 LS-SVM

Introduzido na década de 60 por Vapnik (1963), a máquina de vetor de suporte (SVM - *Support Vector Machines*) corresponde a um notável método de classificação e também de regressão, aplicável em inúmeros problemas. Embora apresente diversas vantagens, a SVM não foi muito utilizada, já que a sua execução em ambientes computacionais é bastante complicada, principalmente se tratando de resoluções de equações não lineares (ESPINOZA; SUYKENS; MOOR, 2003). Entretanto, a partir dos fundamentos da SVM e agregando o método dos mínimos quadrados, Suykens e Vandewalle (1999) desenvolveram um novo método, denominado Máquinas de Vetor Suporte por mínimos quadrados (LS-SVM), que por sua vez, utiliza um conjunto de equações lineares e dessa forma, obtêm praticamente as mesmas soluções obtidas por SVM, com um tempo e custo computacional menores (LUTS et al., 2010).

A LS-SVM tem se mostrado superior quando comparado a técnicas como a MLR e PLS, graças a sua elevada capacidade de modelagem não linear, a sua considerável capacidade de generalização, e a sua robustez. Dessa forma não é necessário reduzir a dimensionalidade dos dados (NIAZI; GOODARZI; YAZDANIPOUR, 2008; FERRÃO et al., 2007; BALABIN; LOMAKINA, 2011).

Na regressão por LS-SVM é determinada a melhor relação entre x e y a partir da minimização de uma função custo (Q) que penaliza pesos e erros altos, Eq. (24)

$$Q = \frac{1}{2}w^t w + \frac{1}{2}\gamma \sum_{i=1}^N e_i^2 \quad (24)$$

onde γ corresponde aos pesos da função custo e deve ser otimizado pelo usuário.

A regressão por LS-SVM pode ser expressa pela Eq. (25)

$$y_i = w^t \phi(x_i) + b + e_i \quad (25)$$

onde w corresponde aos pesos, e aos erros de predições e ϕ a função de linearização das variáveis de entrada. Geralmente, a linearização dos dados originais é realizado pela função de base radial Kernel (RBF), Eq. (26)

$$K = \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right) \quad (26)$$

onde K é a função de base radial, σ é o grau de não linearidade que será modelada. Isto é, quanto maior o valor de σ , mais o modelo se ajustará para uma solução linear. Os parâmetros de γ e σ são otimizados por meio de uma superfície de resposta em função do erro de validação cruzada (RMSECV).

4.7.9 Seleção de variáveis

Os dados analíticos de caráter multivariado muitas vezes apresentam ruídos ou interferentes, assim como informações de pouca relevância ao modelo de calibração. Dessa forma, ferramentas de seleção de variáveis podem ser utilizadas a fim de minimizar os inconvenientes citados e ainda aprimorarem os modelos de calibração. Um método bastante utilizado nesse intuito é a o algoritmo das projeções sucessivas (SPA - *Successive Projections Algorithm*).

4.7.9.1 SPA

O SPA (GALVÃO et al., 2007; GALVÃO FILHO, 2010) seleciona variáveis de modo a minimizar a multicolinearidade presente nos dados analíticos. Para tanto, após a separação das amostras entre calibração e validação, divide-se o conjunto de calibração em dois novos subconjuntos, denominados *calibração* e *teste*.

Em seguida, o conjunto de calibração recentemente construído é novamente fragmentado em um número total de vetores igual ao número de variáveis. Cada vetor corresponderá a uma variável original da matriz de calibração. Posteriormente, novas matrizes serão construídas a partir do agrupamento desses vetores. Cada vetor será agrupado com os vetores que apresentarem as maiores projeções ortogonais entre si até a obtenção do tamanho limite de colunas iguais ao número de amostras.

Para cada matriz construída, um modelo de MLR será gerado. Em seguida, o conjunto de teste previamente isolado será utilizado para avaliar cada modelo de calibração obtido. Tal avaliação consiste na determinação do erro quadrático de predição (RMSEV), Eq. (27)

$$RMSEV = \sqrt{\frac{\sum_{i=1}^n (y_{ref} - y_{val})^2}{n_v}} \quad (27)$$

onde y_{ref} e y_{val} são, respectivamente, os valores de referência e previstos, e n_v o número de amostras do conjunto de teste.

O modelo que apresentar o menor RMSEV será selecionado e encaminhado a última etapa que consiste na eliminação de variáveis de menor relevância, utilizando como critério o teste F .

4.7.10 Avaliação dos modelos de calibração multivariada por meio de métodos estatísticos

Após a aplicação dos métodos de calibração multivariados seguido da geração dos modelos, realiza-se a avaliação dos mesmos por meio de alguns parâmetros estatísticos. Dentre esses parâmetros destacam-se o coeficiente de correlação de Pearson ao quadrado (r^2), o erro médio quadrático de calibração (RMSEC - *Root Mean Square Error of Calibration*), Eq. (28), o erro médio quadrático de previsão (RMSEP - *Root Mean Square Error of Prediction*), Eq. (29), e o Desvio Residual de Previsão (RPD - *Residual Prediction Deviation*), Eq. (30) (IGNE; HURBURGH, 2010; FEARN, 2002) calculado por meio da divisão entre o erro padrão de revisão (SEP - *Standard Error of Prediction*), Eq. (31), e o desvio padrão das amostras de validação (DP_{val})

$$RMSEC = \sqrt{\frac{\sum_{i=1}^n (y_{ref} - y_{cal})^2}{n_{cal}}} \quad (28)$$

onde y_{ref} e y_{cal} correspondem respectivamente aos valores de referência e previstos relativos a etapa de calibração, e n_{cal} refere-se ao total de amostras utilizadas na calibração

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (y_{ref} - y_{val})^2}{n_{val}}} \quad (29)$$

onde y_{ref} e y_{cal} correspondem respectivamente aos valores de referência e previstos relativos a etapa de validação, e n_{cal} refere-se ao total de amostras utilizadas na validação

$$RPD = \frac{DP_{val}}{SEP} \quad (30)$$

$$SEP = \sqrt{\frac{\sum_{i=1}^{n_{val}} (y_{ref} - y_{val} - bias)^2}{n_{val} - 1}} \quad (31)$$

onde n_{val} é o número de amostras de validação e $bias$ pode ser determinado pela Eq.(32).

$$bias = \frac{\sum_{i=1}^{n_{val}} (y_{ref} - y_{val})}{n_{val}} \quad (32)$$

Além dos parâmetros estatísticos mencionados, os modelos de calibração podem ser avaliados quanto a linearidade. Isto é, verificar se a relação entre o sinal analítico e as propriedades de interesse é linear.

4.7.10.1 Linearidade

A linearidade pode ser verificada por meio dos testes: *runs test* e Durbin-Watson teste (DURBIN; WATSON, 1950). O *runs test* consiste na determinação do número consecutivo de resíduos de mesmo sinal, denominado de *run*. Segundo esse teste, os dados são considerados como não-lineares caso um número expressivo de resíduos consecutivos de mesmo sinal for detectado, correspondendo a um evento estatístico pouco provável. Em contrapartida, caso o número de *runs* de sinais opostos sejam similares, a tendência dos dados será considerada como linear (CENTNER; NOORD; MASSART, 1998).

De acordo com o *runs test*, o número total de série de resíduos de sinal positivo e negativos consecutivos são denominadas, respectivamente, como n_+ e n_- . A determinação de não-linearidade dos dados é realizada por meio do cálculo de z , Eq. (33)

$$z = \frac{u - \mu + 0,5}{\sigma} \quad (33)$$

onde u corresponde ao número total de *runs* e μ e σ podem ser calculados, respectivamente, pelas Eqs. (34) e (35)

$$\mu = \frac{2n_+n_-}{n_+ + n_-} + 1 \quad (34)$$

$$\sigma^2 = \frac{2n_+n_-(2n_+n_- - n_+n_-)}{(n_+n_-)^2(n_+n_- - 1)} \quad (35)$$

O valor tabelado de $|z|$ à um nível de 95% confiança corresponde a 1,96. Nesse sentido, valores de z superiores a 1,96 correspondem a um comportamento não-linear dos dados (COOK; WEISBERG, 1982; CENTNER; NOORD; MASSART, 1998).

O teste de Durbin-Watson analisa a aleatoriedade dos resíduos. A hipótese nula (H_0) para esse teste consiste na inexistência de correlação entre os resíduos sucessivos. A rejeição da hipótese nula (H_1) consiste na correlação entre esses resíduos. O aceite ou a rejeição da hipótese nula é determinado por meio do cálculo do parâmetro d , Eq. (36) (CENTNER; NOORD; MASSART, 1998)

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (36)$$

onde e_i corresponde ao resíduo referente a amostra i e e_{i-1} ao resíduo da amostra anterior a i . Após a determinação de d , esse valor é comparado entre dois limites: inferior, d_l e superior, d_u . Caso o valor de d se encontre entre esses dois limites ($d_l < d < d_u$) o teste será considerado como inconclusivo. A hipótese nula é rejeitada, se o valor de d for menor que o limite inferior ($d < d_l$). Isso indica que há correlação entre os resíduos, evidenciando a presença de não linearidade nos dados. Por outro lado, a hipótese nula é aceita quando o valor de d for maior que o limite superior ($d > d_u$). Isto é, não há correlação entre os resíduos, tampouco evidências de que os dados apresentem comportamento não-linear (CENTNER; NOORD; MASSART, 1998).

Nesse estudo, os dois testes: *runs test* e Durbin-Watson test foram aplicados aos resíduos obtidos por quatro métodos: *partial response plot for LV1* (PRP) (y versus LV1), *residual plot* (RP) (e versus y), *plot the residuals versus latent variable e-LV1* (e versus LV1) e *Mallows augmented partial residual plot* (APaRP) ($e_{APaRP} + b_1LV_1 + b_{11}LV_1^2$ versus LV1).

4.8 Análise de solo por meio de MIA

A análise de solo por meio de MIA consiste inicialmente na aquisição de imagens digitais de amostras de solo secas e moídas. Em segundo lugar, de cada imagem, realiza-se a seleção da região de interesse (ROI), isto é, a região que contém a maior informação da imagem. De cada ROI, histogramas de ocorrência dos canais de cores RGB (item 4.4.2.1),

HSI (item 4.4.2.2) e escala cinza são extraídos, conforme descrito no item 4.5.1. Em seguida, os histogramas extraídos são organizados na forma de matriz (4.7.1) e pré-processados (4.7.2). Por último, realiza-se a correlação dos dados das imagens com os parâmetros de fertilidade do solo: COS (item 4.1), textura do solo (item 4.2) e os óxidos extraídos do solo (item 4.3) determinados por meio dos métodos de referência, utilizando os métodos de calibração multivariados: PLS (item 4.7.6), MLR (item 4.7.5) e LS-SVM (4.7.8). A Figura 10 apresenta um fluxograma que resume todas as etapas mencionadas.

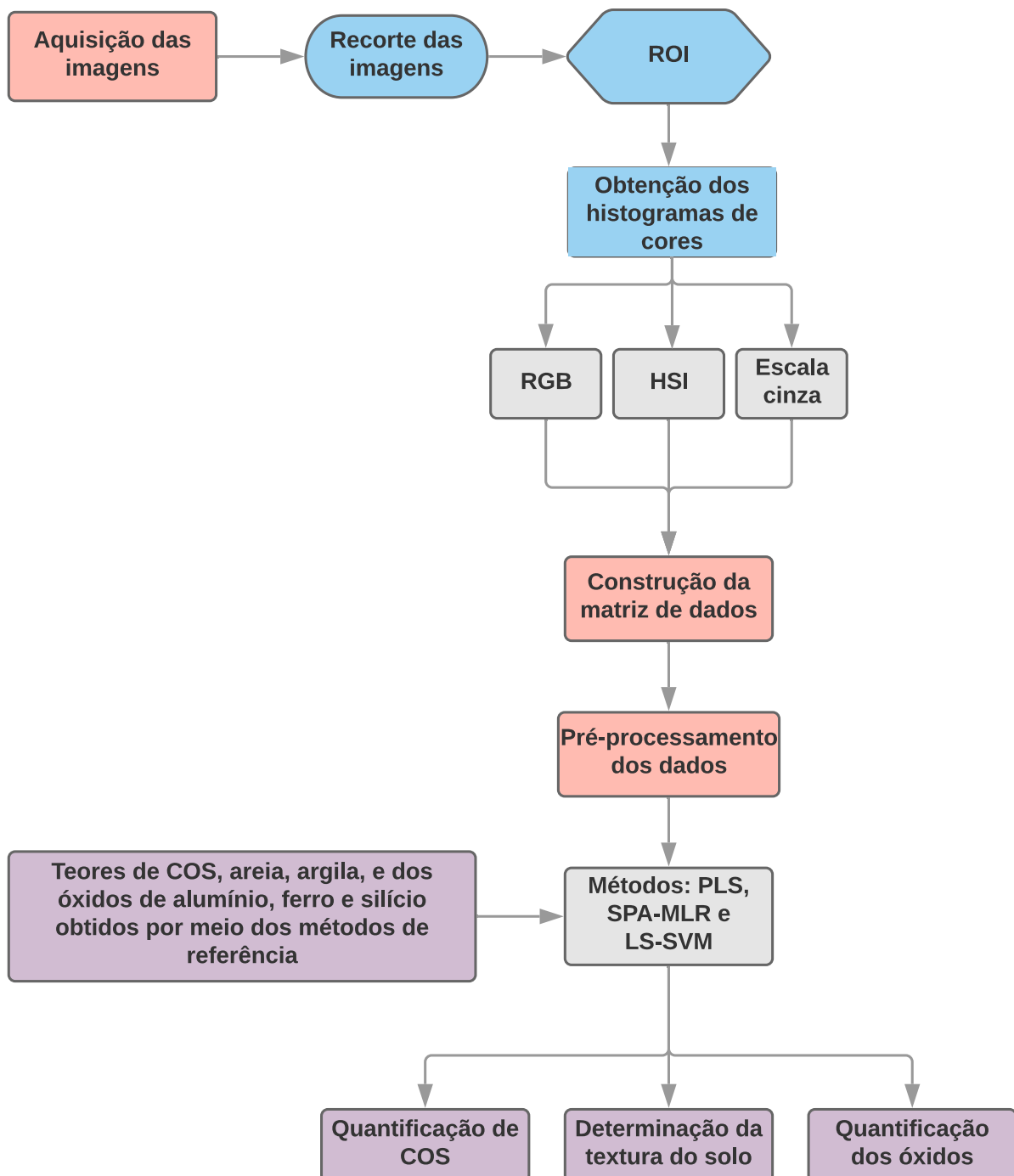


Figura 10 – Fluxograma da análise de solo por meio de imagens digitais.

4.9 Análise de solo por meio de espectros de MIR

A análise de solo por meio da espectroscopia de infravermelho médio (MIR) inicia-se na obtenção de espectros MIR das amostras de solo secas e moídas. Em seguida, os espectros são organizados na forma de matriz e pré-processados. Após a determinação dos parâmetros de fertilidade: COS (item 4.1), textura do solo (item 4.2) e dos óxidos de silício, ferro e alumínio extraídos do solo (item 4.3) pelos métodos de referência, realizou-se a correlação dos dados espectrais com os teores de cada parâmetro por meio dos métodos: PLS (item 4.7.6), MLR (item 4.7.5) e LS-SVM (4.7.8). A Figura 11 apresenta um fluxograma que resume todas as etapas descritas.

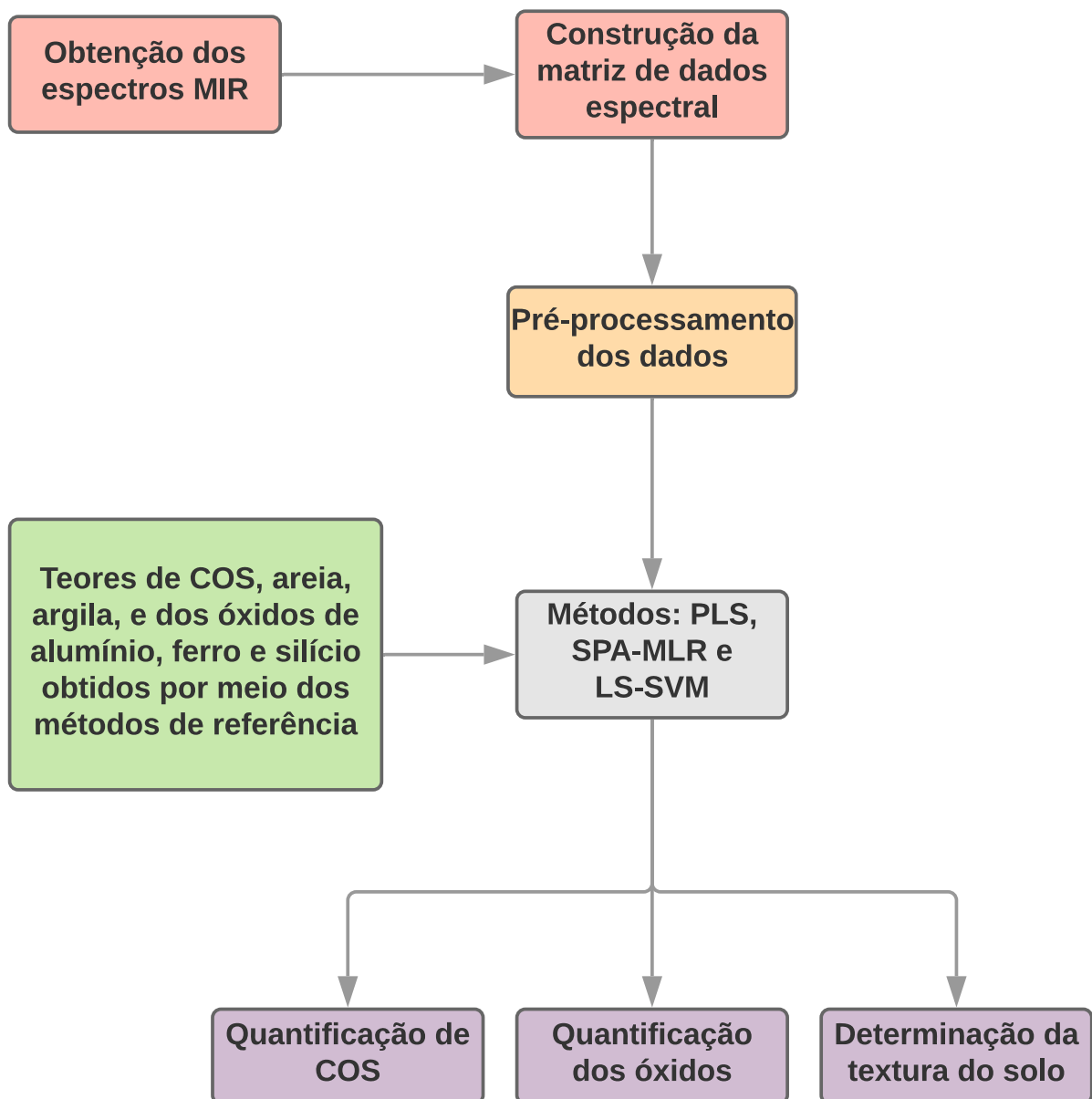


Figura 11 – Fluxograma da análise de solo por meio de imagens digitais.

5 Parte experimental

5.1 Amostragem

5.1.1 Coleta e tratamento das amostras

Neste trabalho, um total de 177 amostras foram coletadas em profundidade de até 10 cm nas cidades de Pirenópolis (15^o 51' 09", 48^o 57' 33"), Santo Antônio de Goiás (16^o 29' 8", 49^o 18' 32") e Brazabrantes (16^o 25' 47", 49^o 23' 13"), municípios do estado de Goiás; em Belém do São Francisco (8^o 45' 31", 38^o 57' 45"), no estado do Pernambuco; em Nova Canaã do Norte (10^o 33' 28", 55^o 57' 10"), no estado do Mato Grosso; em Campo Verde (15^o 32' 44", 55^o 9' 59"), no estado do Mato Grosso do Sul; em João Pessoa (07^o 06' 54", 34^o 51' 47"W), no estado da Paraíba; e em Palmas (10^o 12' 46", 48^o 21' 37"), no estado de Tocantins.

Após a coleta, as amostras de solo foram secas em estufa a 45 °C por 48 horas, moídas e peneiradas em malha de 2 mm.

5.2 Metodologia para quantificação do carbono orgânico do solo

O carbono orgânico do solo foi quantificado por meio do método Walkley-Black espectrofotométrico recentemente proposto por (SOUZA et al., 2016). Primeiramente, 0,2 g de amostra de solo seca e moída foram pesadas. Posteriormente, 4,0 mL de $K_2Cr_2O_7$ 0,167 mol L^{-1} , 8,0 mL de H_2SO_4 P.A. e 12 mL de água destilada foram adicionados. Após a agitação e arrefecimento, os sistemas foram mantidos em repouso por um período de 48 horas para decantação do solo. Em seguida, 5 mL do sobrenadante foram coletados e transferidos a um cubeta de vidro para a realização da leitura espectrofotométrica no comprimento de onda de 590 nm.

Um total de 10 amostras padrão de concentrações de Cr^{3+} entre 0,000 e 0,039 mol L^{-1} foram preparadas para construção da curva de calibração. Primeiramente, 4,0 mL de $K_2Cr_2O_7$ 0,167 mol L^{-1} e 8,0 mL de H_2SO_4 P.A. foram adicionados em um erlenmeyer de vidro. Após a agitação, volumes variados (0,5 a 7,5 mL) de $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ 0,5 mol L^{-1} em pH 0,3 foram adicionados em cada solução. Em seguida, água destilada foi adicionada até obtenção de volume final de 24 mL. A coleta de sobrenadante e leitura

em espectrofotômetro foi realizado da mesma forma como descrito para as amostras. A Tabela 2 apresenta os volumes adicionados de cada reagente para cada amostra padrão.

Tabela 2 – Condições de preparo das soluções padrão utilizadas na curva de calibração.

Solução padrão	Volumes adicionados para o preparo das soluções padrão (mL)	
	$Fe(SO_4)_2 \cdot 7H_2O$	H_2O destilada
1	0,00	12,00
2	0,50	11,50
3	1,50	10,50
4	2,50	9,50
5	3,00	9,00
6	4,50	7,50
7	5,50	6,50
8	6,00	6,00
9	6,50	5,50
10	7,50	4,50

5.3 Metodologia para quantificação da textura do solo

O método da pipeta (ROBINSON, 1922) foi utilizado como referência para análise textural do solo, uma vez que esse procedimento é considerado um método padrão no âmbito internacional e além disso, é o mais difundido em estudos granulométricos (PANSU; GAUTHEYROU, 2003).

Esse ensaio consiste na dispersão química e física da amostra seguido da determinação das frações granulométricas, areia, silte e argila. Primeiramente, pesaram-se 10 g de amostra de solo fina e seca. Em seguida, realizou-se a etapa química que consiste na adição de 10 mL de hidróxido de sódio 1 mol L^{-1} , 10 mL de hexametáfosfato de sódio 1 mol L^{-1} e 2 mL de peróxido de hidrogênio 9%. A dispersão física foi realizada utilizando ondas mecânicas oriundas de um processador ultra-sônico Hielscher UP400S (400W 24 kHz).

Após os dois processos de dispersão, a areia é retida por peneiramento em malha de 0,05 mm e quantificada por gravimetria. Após a coleta da mistura de solo que passou pela peneira, ajusta-se o volume final para 1,0 L em uma proveta com água destilada. A quantificação do teor de argila foi realizada retirando-se uma alíquota em uma profundidade de 5 cm, seguido da sua determinação por gravimetria. Por último determina-se o teor de silte, de forma indireta, descontando as porcentagens de areia e argila.

5.4 Metodologia para quantificação de fósforo, sódio e potássio disponíveis

Após a coleta e preparo do solo, os teores de fósforo, sódio e potássio foram determinados. Primeiramente, pesou-se 10 g de amostra de solo fina e seca. Em seguida, adicionou-se 100 mL de solução extratora Mehlich ($\text{HCl } 0,05 \text{ mol L}^{-1} + \text{H}_2\text{SO}_4 \text{ } 0,0125 \text{ mol L}^{-1}$) e agitou-se em mesa agitadora de movimento circular horizontal de rotação de 200 rpm durante 5 min. Após a decantação das amostras por um período de 15 h, coletou-se 25 mL do sobrenadante que foi utilizado para a determinação dos teores de fósforo sódio e potássio por meio da espectrofotometria de emissão ótica com plasma indutivamente acoplado (ICP-OES).

5.5 Metodologia para quantificação de micronutrientes disponíveis

A determinação dos micronutrientes: Zn, Cu, Fe e Mn foram determinados de modo similar a determinação do fósforo, sódio e potássio (Tópico 5.4). Entretanto, a extração foi realizada utilizando somente 50 mL de solução Mehlich. Após a agitação, decantação e coleta do sobrenadante, os teores de Zn, Cu, Fe e Mn foram determinados por meio de ICP-OES.

5.6 Metodologia para determinação de ferro, alumínio e silício em extrato sulfúrico

Após secagem, moagem e peneiragem das amostras, pesou-se 1.0 grama de solo em um balão de vidro de fundo chato. Em seguida, adicionou-se 20 mL de ácido sulfúrico (H_2SO_4) 9 mol L^{-1} , acoplou-se um condensador de refluxo ao balão e realizou-se o aquecimento utilizando uma chapa de aquecimento. Após o início da ebulição, o sistema foi mantido sob refluxo por 30 minutos. Posteriormente ao resfriamento, realizou-se a filtração e lavagem do conteúdo sólido utilizando água destilada. Em seguida, o filtrado e as águas de lavagem foram combinadas e avolumadas utilizando água destilada em um balão volumétrico de vidro. Os teores de Al e Fe foram determinados no filtrado utilizando a espectrofotometria de emissão ótica com plasma indutivamente acoplado (ICP-OES). O resíduo sólido obtido após a filtração foi utilizado para quantificação dos teores de silício. Primeiramente, esse resíduo foi transferido para um erlenmeyer e, em seguida

adicionou-se 4 mL de NaOH 7,5 mol L^{-1} . A mistura foi aquecida em placa aquecedora até o início da ebulição. Após o resfriamento, a mistura foi filtrada e avolumada em um balão volumétrico. Por último, realizou-se a quantificação da concentração de silício por meio da espectroscopia de absorção atômica.

5.7 Aquisição dos espectros de infravermelho

As varreduras espectrais na região do infravermelho médio (MIR) foram adquiridas em um espectrômetro Varian 660 FT-IR com transformada de Fourier. A faixa espectral utilizada foi de 4000 a 400 cm^{-1} , resolução de 8 cm^{-1} e 64 varreduras por leitura.

Como mencionado no item 4.3, o silício é um dos elementos mais abundantes no solo, sendo sua principal fonte o quartzo. Na espectroscopia de infravermelho médio, as frequências de absorção do quartzo estão presentes principalmente nas regiões entre 2000 a 600 cm^{-1} (SORIANO-DISLA *et al.*, 2014). A Tabela 3 apresenta as faixas de absorção dos óxidos de silício, assim como dos óxidos de alumínio e ferro.

Tabela 3 – Bandas de absorção referentes aos óxidos de silício, ferro e alumínio na espectroscopia de infravermelho médio (MIR) (SAIKIA; PARTHASARATHY *et al.*, 2010; IBRAHIM; HAMEED; JALBOUT, 2008; MADEJOVÁ, 2003; DJOMGOUE; NJOPWOUO *et al.*, 2013).

Bandas de absorção (cm^{-1})	Atribuição	Parâmetro
1080-1175	Estiramentos assimétrico Si-O	SiO ₂
780-800	Estiramentos simétrico Si-O	SiO ₂
695	Deformações angulares assimétricos Si-O	SiO ₂
464	Deformações angulares simétricos Si-O	SiO ₂
592-534	Deformações angulares Al-Si-O	SiO ₂
466-470	Deformações angulares Mg-Si-O	SiO ₂
676	Fe-O fora do plano de vibração	Fe ₂ O ₃
800	Estiramento de vibração do Fe(II)Fe(III)OH	Fe ₂ O ₃
890	Deformação angular OH ⁻ do Fe(III)AlOH	Fe ₂ O ₃
840	Deformação angular do OH ⁻ do MgAlOH	Al ₂ O ₃
920	Deformação angular OH ⁻ do Al ₂ OH	Al ₂ O ₃

De modo similar aos óxidos, o carbono orgânico do solo, componente da matéria orgânica, também apresenta bandas de absorção na faixa de infravermelho médio. Entretanto, a identificação dessas bandas é dificultada devido a baixa concentração de carbono orgânico no solo, assim como pela sobreposição de bandas de absorção de minerais (MADARI *et al.*, 2006). A Tabela 4 apresenta as frequências de absorção do carbono orgânico.

Tabela 4 – Bandas de absorção referentes ao carbono orgânico na espectroscopia de infravermelho médio (MIR) (STEVENSON, 1994; BAES; BLOOM, 1989; NIEMEYER; CHEN; BOLLAG, 1992; MADARI et al., 2006).

Banda de absorção (cm ⁻¹)	Atribuição
3380	Estiramento O-H de fenólico
3400-3300	Estiramento N-H
3030	Estiramento C-H de aromáticos
2940-2900	Estiramento C-H de alifático
2600	Estiramento O-H em carboxila (-COOH)
1725-1720	Estiramento da ligação C=O de carboxila (-COOH) e cetonas
1660-1630	Estiramento da ligação C=O em amidas e C=O em quinina
1620-1600	Estiramento C=C de aromáticos, Estiramento assimétrico e simétrico de -COO
1525	Estiramento C=C de aromático
1460-1450	Estiramento C-H de alifático Deformação OH,
1400-1390	Estiramento C-O de fenólico, Deformação C-H dos grupos CH ₃ e CH ₂ , Estiramento assimétrico COO-
1350	Estiramento simétrico COO- e a deformação angular -CH de alifáticos
1270	Estiramento C-OH de fenólico Estiramento C-O,
1280-1200	Deformação OH do grupo COOH, Estiramento C-O de éteres arílicos
1225	Estiramento C-O, deformação angular OH do grupo COOH
1170-950	Estiramento C-O de polissacarídeos
1170	Estiramento C-OH de OH alifático
1070	Estiramento C-C de grupos alifáticos
830 e 775	Deformação angular fora do plano CH de aromático

As partículas de argila são formadas principalmente por argilominerais (STENBERG et al., 2010), já a fração de areia é constituída especialmente por quartzo. Na espectroscopia de infravermelho médio, as bandas de absorção referentes ao quartzo e aos argilominerais são observados em 1050 cm⁻¹ e entre 900 a 400 cm⁻¹ (PIRIE; SINGH; ISLAM, 2005).

5.8 Aquisição das imagens digitais das amostras de solo

As imagens digitais foram adquiridas utilizando uma placa de Petri em que foram adicionadas 20 g de amostra de solo. Em seguida, a placa foi disposta no centro de um *scanner* de mesa HP Scanjet G4050, Figura 12.

Cada amostra foi digitalizada três vezes no sistema de cor RGB, com resolução de 600 dpi e formato TIFF (*Tagged Image File Format*). Desse modo, um total de 531



Figura 12 – Scanner de mesa utilizado para a aquisição das imagens de solo.

imagens foram obtidas. A Figura 13 apresenta seis imagens obtidas de amostras de solo diversificadas.

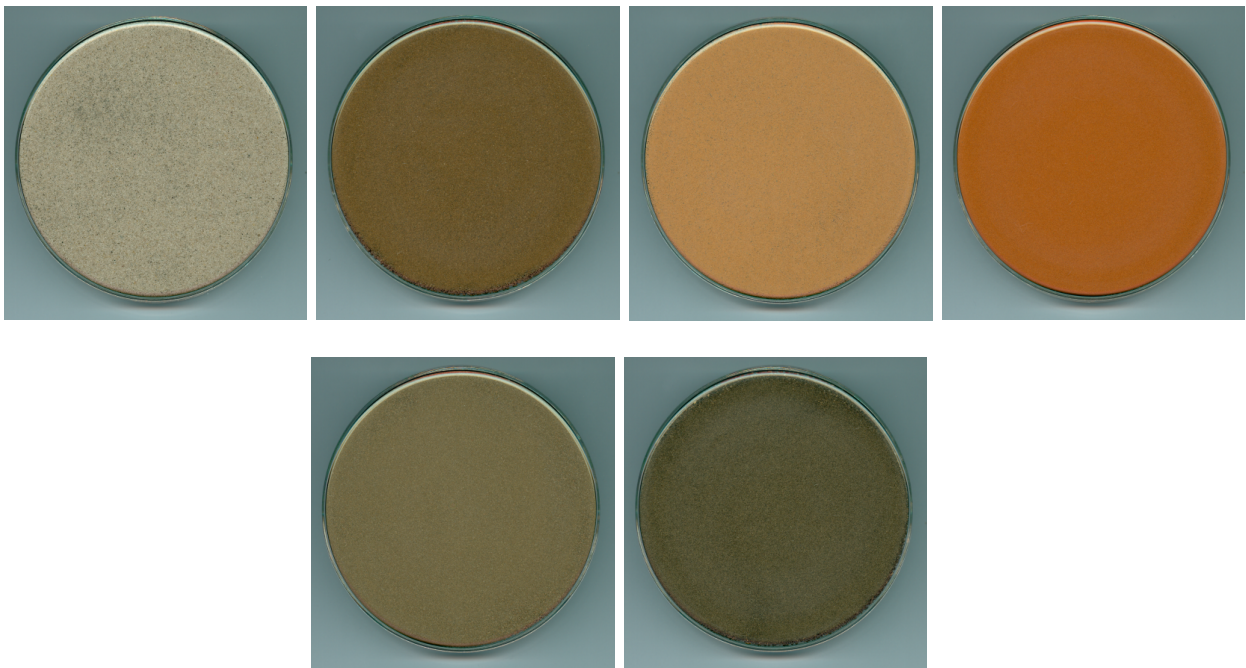


Figura 13 – Imagens digitais para seis amostras de solo (teores de COS entre 0,590 a 47,590 $g\ kg^{-1}$)

5.9 Análise de solo por meio de MIA

Posteriormente à obtenção das imagens (item 5.8), selecionou-se uma área central de cada imagem, referente a região de interesse (ROI), e realizou-se a conversão das mesmas

para os sistemas de cor: HSI e escala cinza. Em seguida, de cada ROI foram extraídos histogramas de ocorrência referentes aos espaços de cores RGB, HSI e escala cinza. A Figura 14 apresenta um exemplo da ROI selecionada para a primeira amostra de solo.

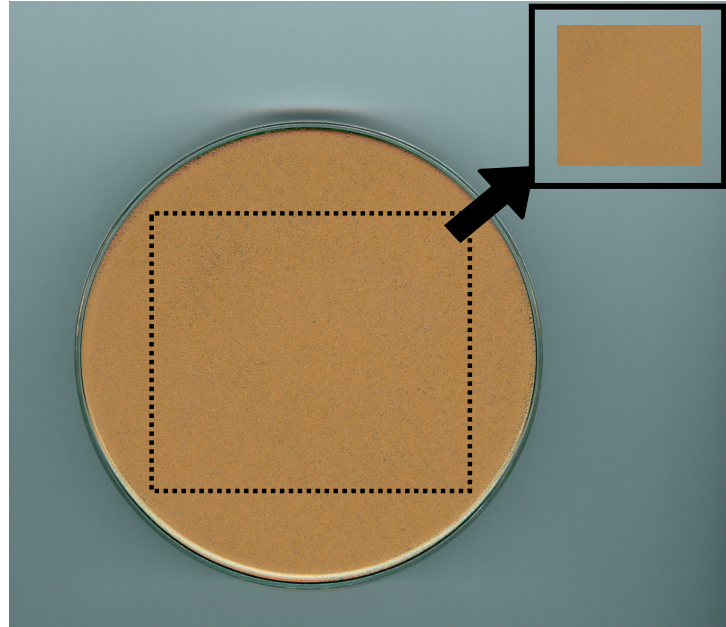


Figura 14 – Seleção de uma região de interesse da primeira imagem de amostra de solo obtida.

A Figura 15 apresenta todos os passos realizados para a obtenção dos histogramas de ocorrência de cor.

Após a obtenção dos histogramas de ocorrência de cor, os mesmos foram organizados em uma matriz de dados. Em seguida, os dados foram pré-processados. Por último, realizou-se a correlação dos dados das imagens com os parâmetros de fertilidade, por meio dos métodos de regressão: PLS, SPA-MLR e LS-SVM.

5.10 Análise de solo por meio de espectros de MIR

Após a secagem e a moagem das amostras de solo, obtiveram-se espectros de infravermelho médio (5.7). Em seguida, os espectros adquiridos foram organizados na forma de uma matriz. Após o pré-processamento, os dados espectrais foram correlacionados com os teores dos parâmetros de fertilidade por meio dos métodos de regressão: PLS, SPA-MLR e LS-SVM.

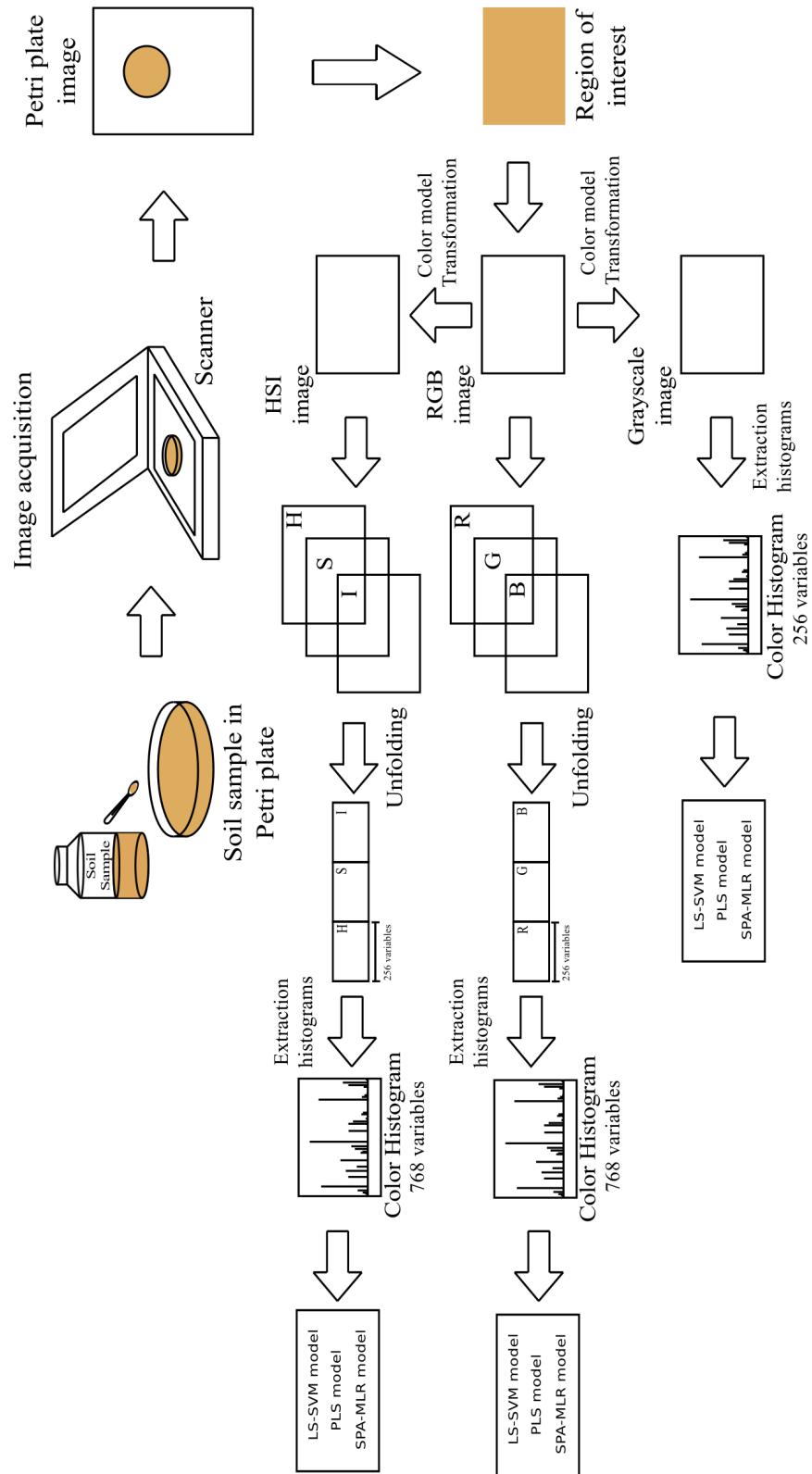


Figura 15 – Passos realizados para obtenção dos histogramas de ocorrência de cor de amostras de solo.

5.11 Calibração e Validação

As amostras foram classificadas em calibração e validação, utilizando o algoritmo SPXY (*Sample set Partitioning based on joint x-y distances*) (GALVÃO et al., 2005) definindo-se 70% das amostras (124) para calibração e 30% das amostras (53) para a validação. Após essa divisão, utilizaram-se os métodos de regressão multivariados PLS, MLR associado ao método de seleção de variáveis SPA, SPA-MLR e LS-SVM com o intuito de correlacionar histogramas de ocorrência de cor extraídos de imagens de amostras de solo com os seus teores de COS, argila e areia.

5.12 Manipulação e tratamento dos dados

Neste estudo, o software MATLAB™ R2014b, versão 8.4.0 (*Mathworks, Natick, EUA*), juntamente com os pacotes PLS Toolbox 2.0 (Eigenvector Research, EUA), LS-SVMlab Toolbox versão 1.8 (Suykens, Leuven, Belgium) (SUYKENS et al., 2002) e SPA Toolbox 1.0 (disponível em www.ele.ita.br/~kawakami/spa/) (ARAÚJO et al., 2001) foram utilizados para as transformações e tratamento das imagens, pré-processamento e redução da dimensionalidade dos dados e também para realizar as calibrações multivariadas.

6 Resultados e Discussões

6.1 Descrição estatística das propriedades das amostras de solo

A Tabela 5 apresenta uma descrição estatística de alguns parâmetros de fertilidade das 177 amostras. Já a distribuição desses parâmetros pode ser visualizada na Figura 16. A descrição completa dos parâmetros de fertilidade: COS, areia, silte e argila, assim como a classificação textural de todas as amostras encontra-se no [Apêndice A](#).

Tabela 5 – Atributos físicos e químicos das 177 amostras de solo.

Estatística descritiva	Granulometria (%)			COS (g kg ⁻¹)
	Argila	Silte	Areia	
Média	41,858 ± 21,776	11,012 ± 11,506	47,321 ± 25,968	14,263 ± 10,039
Mediana	44,900	5,300	41,138	13,081
Máximo	81,522	41,500	99,779	47,590
Mínimo	0,601	0,000	0,800	0,590

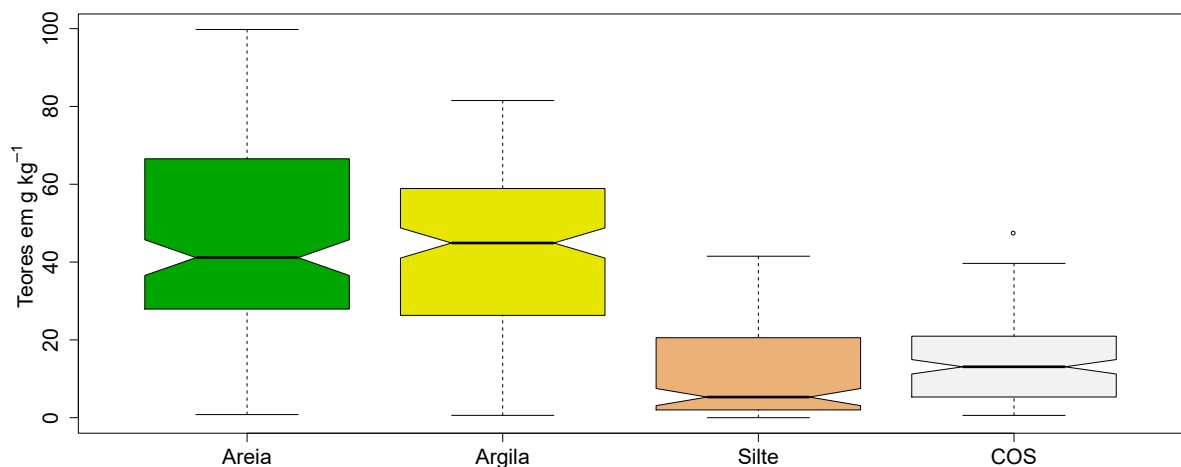


Figura 16 – Distribuição dos teores de carbono orgânico e granulometria do solo das 177 amostras.

A partir dos teores de argila, areia e silte, as amostras de solo foram classificadas em classes texturais. A Tabela 6 apresenta as proporções de amostras de solo para cada grupo textural. A Figura 17 apresenta graficamente a classificação textural das 177 amostras.

Tabela 6 – Proporções de amostras de solo para cada grupo textural.

Classificação textural	Número de amostras	Proporção (%)
Argila	93	52,5
Argilo-arenosa	11	6,5
Franco argilo arenosa	33	18,6
Franco arenosa	5	2,8
Franco argilosa	3	1,7
Areia franca	11	6,2
Areia	16	9,0
Franca	5	2,8



Figura 17 – Classificação textural de todas as amostras

6.2 Análise de solo por meio dos métodos de referência

Após a obtenção, moagem e do peneiramento das 177 amostras de solo, realizou-se a determinação dos teores de COS e da textura do solo, respectivamente, pelos métodos de referência Walkley-Black e Pipeta. Em seguida, obtiveram-se imagens digitais e espectros de infravermelho das amostras de solo peneiradas. Os teores de COS, argila e areia foram estimados por meio da análise multivariada de imagens (MIA) e por intermédio dos espectros de infravermelho.

6.2.1 Análise de solo por meio da MIA

As imagens de solo foram obtidas após a moagem e o peneiramento das amostras utilizando um *scanner* de mesa. Em seguida, os teores de COS e de textura do solo foram determinados, respectivamente, pelos métodos de referência Walkley-Black e Pipeta, e posteriormente pela análise multivariada de imagens. A Figura 18 apresenta o resumo das etapas utilizadas.

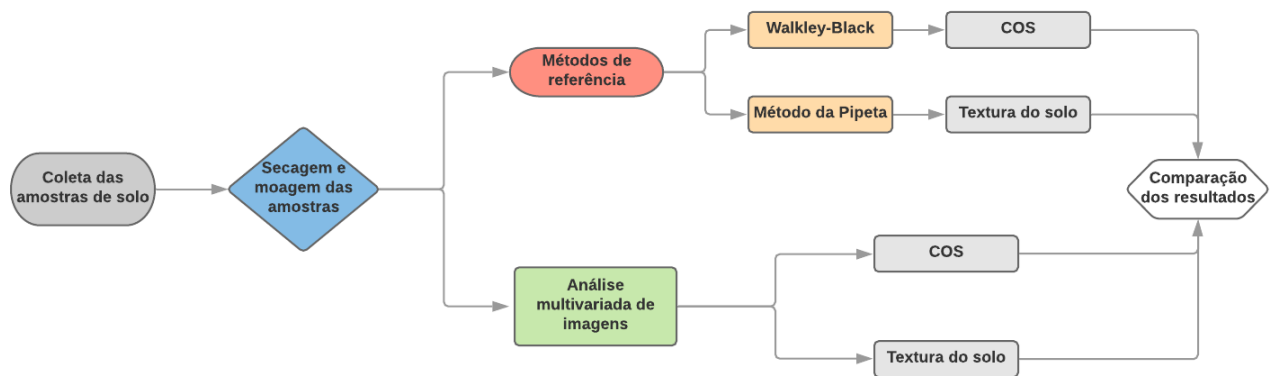


Figura 18 – Etapas utilizadas na quantificação do COS e da textura do solo pelo método de referência e pela MIA.

Após a aquisição das imagens digitais, uma região de interesse (ROI) foi selecionada de cada imagem. As ROIs, no sistema de cores RGB, foram convertidas para os sistemas HSI e escala de cinza. Posteriormente, extraíram-se os histogramas de ocorrência de cada sistema de cor. Em seguida, os histogramas de cores foram desdobrados e justapostos na forma de matrizes. A Figura 19 apresenta os desdobramentos dos histogramas de ocorrência dos modelos de cores RGB, HSI e escala cinza de todas as amostras.

Além das matrizes de dados com as informações dos histogramas RGB, HSI e escala cinza, mais quatro matrizes foram obtidas a partir da fusão dos dados: RGB + escala cinza, HSI + escala cinza, RGB + HSI e RGB + HSI + escala cinza. Desse modo, um total de 7 matrizes contendo os dados de imagens foram geradas. Em seguida, todas as matrizes foram centradas na média.

As 7 matrizes contendo os dados de imagens juntamente com os vetor \mathbf{y}_s , referente aos teores de COS, areia e argila, foram subdivididas entre calibração e validação utilizando a ferramenta SPXY. Desse modo, um total de 124 amostras foram designadas para calibração e 53 para validação. Por último utilizou-se os métodos de calibração multivariados

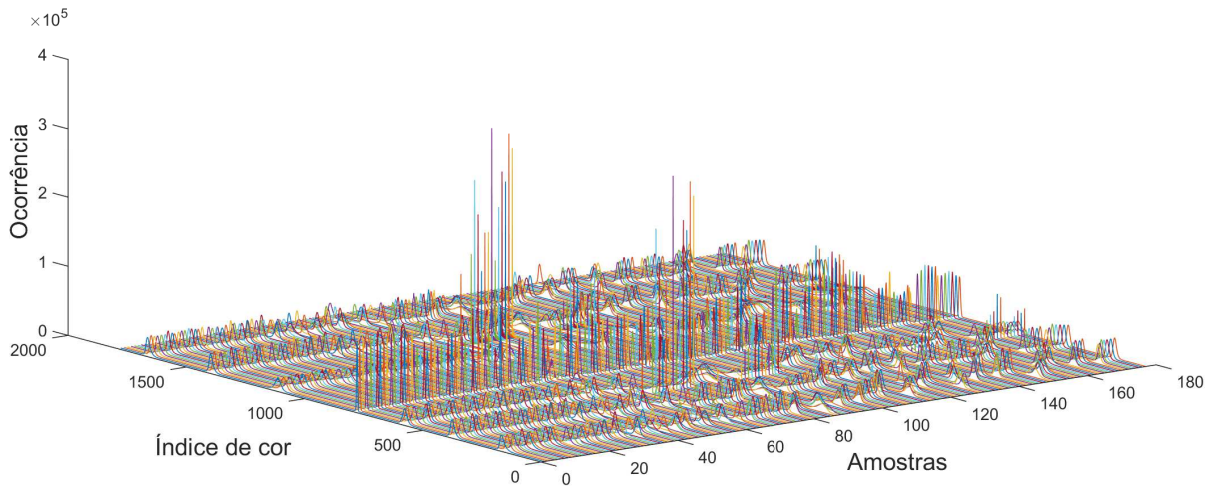


Figura 19 – Matriz de dados referentes aos histogramas de ocorrência dos modelos RGB, HSI e escala cinza de todas as imagens de solo.

PLS, SPA-MLR e LS-SVM para correlacionar as matrizes contendo os dados de imagens com os teores de COS, argila e areia. O número ideal de LVs para o método PLS foi selecionado de acordo com os menores valores de RMSEP. A Figura 20 apresenta todos os passos utilizados para estimação dos teores de COS, argila e areia por meio da análise multivariada de imagens.

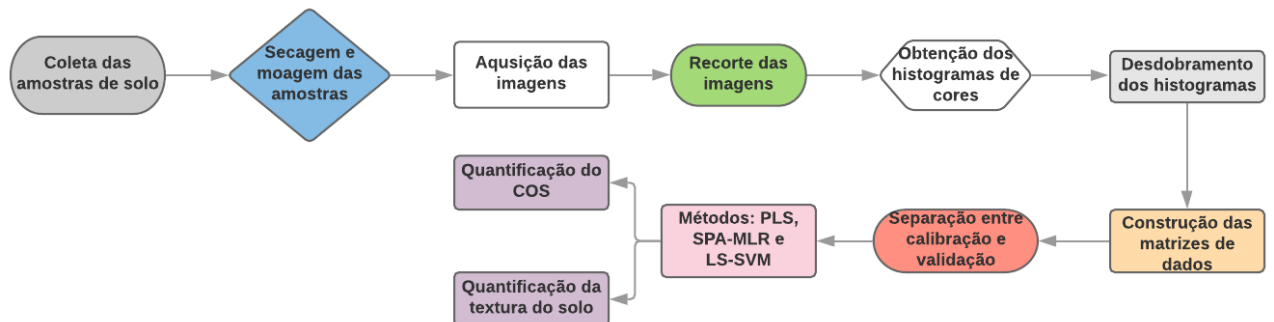


Figura 20 – Etapas utilizadas para estimação dos teores de COS, areia e argila em amostras de solo por MIA.

6.2.1.1 Quantificação do COS

As Tabelas 7, 8 e 9 apresentam os resultados obtidos, respectivamente, para os métodos PLS, SPA-MLR e LS-SVM na predição do COS por meio de imagens digitais.

Tabela 7 – Parâmetros estatísticos da calibração e validação obtidos utilizando o método PLS.

Matriz de dados de imagens	LVs	Calibração		Validação		
		r^2	RMSEC	r^2	RMSEP	RPD
RGB	20	0,7543	5,1	0,7679	4,5	2,1
Escala cinza	17	0,4784	7,3	0,3449	7,6	1,3
HSI	18	0,7296	5,3	0,6857	5,4	1,8
RGB + Escala cinza (RC)	22	0,8021	4,6	0,7961	4,4	2,2
HSI + Escala cinza (HC)	26	0,7775	4,9	0,7442	4,7	1,9
RGB + HSI (RH)	23	0,8543	4,0	0,8224	3,8	2,3
RGB + HSI + Escala cinza (RHC)	25	0,8721	3,7	0,8603	3,5	2,6

De acordo com a Tabela 7, os melhores resultados foram obtidos utilizando a matriz de dados que contém a fusão dos sistemas de cores RGB, HSI + escala cinza (RHC). Os valores de RMSEP de todos os modelos foram comparados por meio do teste-F à um nível de confiança de 95%. Segundo esse teste, somente a matriz de dados RGB + HSI apresentou RMSEP equivalente ao obtido pela matriz de dados: RGB + HSI + escala cinza.

A Figura 21 apresenta o gráfico das concentrações de COS do método de referência contra as concentrações de COS previstas pelo modelo PLS obtido utilizando a análise multivariada de imagens e a matriz de dados RHC (MIA-PLS-RHC).

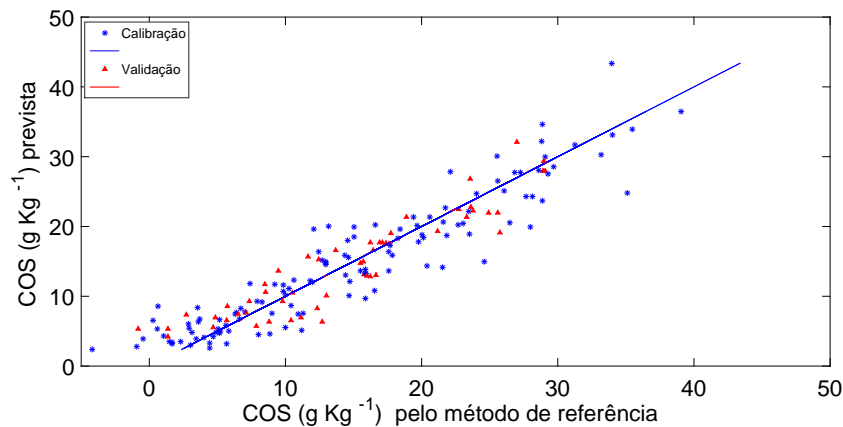


Figura 21 – Concentrações de COS de referência versus previstas pelo modelo MIA-PLS-RHC. Calibração (*) e validação (▲).

De acordo com a Tabela 8, diferente do modelo PLS, a matriz de dados que contém a fusão dos sistemas de cores RGB e escala cinza (RC) apresentou os melhores resultados. No entanto, com base no teste-F com 95% de confiança, todos os valores de RMSEP foram equivalentes, com a exceção das matrizes de dados: escala cinza e RGB + HSI.

Tabela 8 – Parâmetros estatísticos da calibração e validação obtidos utilizando o método SPA-MLR.

Matriz de dados de imagens	Número de variáveis	Calibração		Validação		
		r^2	RMSE	r^2	RMSE	RPD
RGB	16	0,5145	7,2	0,6924	5,7	1,7
Escala cinza	10	0,3473	8,2	0,4190	7,4	1,3
HSI	10	0,5177	7,0	0,6916	5,5	1,8
RGB + Escala cinza	27	0,6548	6,0	0,7351	5,0	1,9
HSI + Escala cinza	18	0,5845	6,5	0,5666	6,3	1,5
RGB + HSI	8	0,5008	7,3	0,5243	6,9	1,4
RGB + HSI + Escala cinza	23	0,6337	6,3	0,5628	5,8	1,5

O gráfico das concentrações de COS do método de referência contra as concentrações de COS previstas pelo modelo SPA-MLR obtido utilizando a análise multivariada de imagens e a matriz de dados RC (MIA-SPA-MLR-RC) está apresentado pela Figura 22.

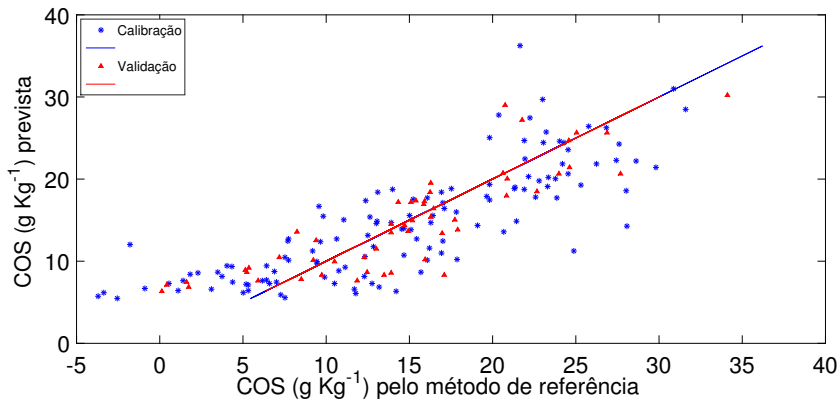


Figura 22 – Concentrações de COS de referência versus previstas pelo modelo MIA-SPA-MLR-RC. Calibração (*) e validação (▲).

Tabela 9 – Parâmetros estatísticos da calibração e validação obtidos utilizando o método LS-SVM.

Matriz de dados de imagens	Calibração		Validação			γ	σ^2
	r^2	RMSE	r^2	RMSE	RPD		
RGB	0,9499	2,5	0,9287	2,9	3,3	12,907	249,156
Escala cinza	0,4212	7,8	0,3629	8,0	1,2	7,888	3757,362
HSI	0,9253	2,9	0,8973	3,3	3,0	9,159	173,138
RGB + Escala cinza	0,9588	2,3	0,9375	2,9	3,3	15,792	305,133
HSI + Escala cinza	0,9307	2,9	0,8961	3,1	3,0	12,394	364,523
RGB + HSI	0,9658	2,1	0,9330	2,6	3,5	20,854	516,648
RGB + HSI + Escala cinza	0,9598	2,3	0,9376	2,5	3,7	13,462	519,993

Segundo a Tabela 9, a matriz de dados que contém a fusão dos três sistemas de cores RGB, HSI e escala cinza (RHC) apresentou os melhores resultados para a ferramenta LS-SVM. Entretanto, a comparação dos valores de RMSEP por meio do teste-F com 95%

de confiança, demonstrou que com exceção das matrizes de dados: escala cinza e HSI, todos os valores de RMSEP são equivalentes.

A Figura 23 apresenta o gráfico das concentrações de COS do método de referência contra as concentrações de COS previstas pelo modelo LS-SVM obtido utilizando a análise multivariada de imagens e a matriz de dados RHC (MIA-LS-SVM-RHC).

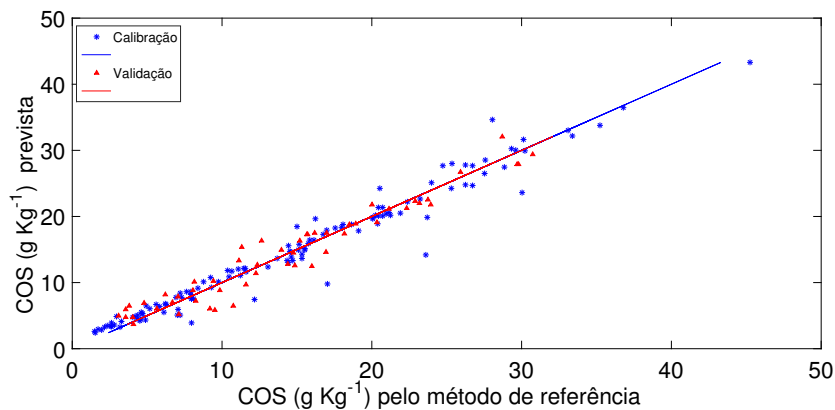


Figura 23 – Concentrações de COS de referência versus previstas pelo modelo MIA-LS-SVM-RHC. Calibração (*) e validação (▲).

O melhor resultado obtido pela ferramenta LS-SVM apresentou um RPD maior que 3.5. Segundo [Saeys, Mouazen e Ramon \(2005\)](#) um modelo de calibração com RPD maior que 3.0 pode ser classificado como excelente.

De acordo com as Tabelas 7, 8 e 9, o método LS-SVM apresentou os melhores resultados em comparação com os métodos PLS e SPA-MLR em qualquer sistema de cor ou fusão de dados, com exceção da matriz de dados referente a escala cinza. Os valores de RMSEP determinados para os três métodos em cada matriz de dados foram comparados por meio do teste-F com 95% de confiança. Segundo esse teste, os métodos PLS e SPA-MLR só apresentaram performance equivalente ao LS-SVM para a matriz de dados escala cinza. Isto é, em todas os outros sistemas de cores e suas fusões, o método LS-SVM apresentou uma maior precisão que o PLS e o SPA-MLR. Por outro lado, a precisão do método PLS foi superior ao método SPA-MLR em 4 matrizes de dados: RGB, HSI + escala cinza, RGB + HSI e RGB + HSI + escala cinza. No restante das matrizes de dados, os dois métodos apresentaram a mesma precisão.

A performance dos três métodos na predição do carbono orgânico do solo também foi comparada por meio do teste de randomização proposto por [Voet \(1994\)](#). De acordo com esse teste, os métodos PLS e SPA-MLR só apresentaram a mesma acurácia que o LS-SVM (p-valor > 0.05) na matriz de dados escala cinza. Em contrapartida, o método

SPA-MLR só apresentou a mesma acurácia que o método PLS para a matriz de dados: escala cinza e RGB + escala cinza. A superioridade notável da ferramenta LS-SVM ainda pode ser verificada comparando as Figuras 21, 22 e 23, pois os pontos para essa técnica estão mais próximos a linha de correlação.

Como já mencionado a ferramenta LS-SVM, de caráter não linear, apresentou os melhores resultados na estimação de COS. Dessa forma, realizou-se a investigação da não linearidade dos dados. Para tanto, dois métodos foram utilizados: *runs test* e Durbin-Watson test. Esses dois testes foram aplicados aos resíduos obtidos por quatro métodos: partial response plot for LV1 (PRP) (y versus LV1), residual plot (RP) (e versus y), plot the residuals versus latent variable e-LV1 (e versus LV1) and Mallows augmented partial residual plot (APaRP) ($e_{APaRP} + b_1LV_1 + b_{11}LV_1^2$ versus LV1). A Tabela 10 apresenta os resultados obtidos.

De acordo com a Tabela 10 os modelos MIA obtidos com as matrizes de dados RGB e RGB + escala cinza construídos com somente uma LV foram classificadas como não-lineares por todos os testes realizados. O restante dos modelos com somente um LV foram classificados como não lineares por todos os testes, com exceção do Durbin-Watson aplicado para RP. Os modelos construídos utilizando a matriz de dados HSI e escala cinza com mais de uma LV foram classificados como não-lineares por todos os testes, exceto o teste Durbin-Watson aplicado para RP, o qual foi inconclusivo. O modelo desenvolvido com a matriz de dados RGB com 20 LVs foi classificado como não-linear por dois *runs test* (RP e e-LV1) e por 3 testes Durbin-Watson (PRP, e-LV1 e APaRP for LV1). O modelo construído com a fusão dos dados RGB + escala cinza com 22 LVs foi classificado como não linear por 3 *runs test* (PRP, RP e APaRP for LV1) e por 2 testes Durbin-Watson (PRP e e-LV1). O modelo desenvolvido com a matriz de dados HSI e escala cinza com 26 LVs foi classificado como não-linear por todos os *runs tests* e por um teste Durbin-Watson (PRP).

Os modelos construídos pelas matrizes de dados: RGB + escala cinza e RGB + HSI + escala cinza com, respectivamente, 23 e 25 LVs foram classificados como inconclusivos por todos os testes Durbin-Watson, com exceção do PRP, o qual classificou ambos modelos como não-lineares. Além disso, esses modelos foram classificados como não-lineares por dois *runs tests* e lineares por outros dois.

Todos os modelos PLS utilizados para estimação de COS foram construídos utilizando um alto número de variáveis latentes (≥ 17). Geralmente, modelos com um número

Tabela 10 – Resultados obtidos na detecção da não linearidade dos dados por meio dos testes: *Runs* e Durbin-Watson

Modelo MIA	Plot	$ z $ (Runs tests)	Conclusão	d (Durbin-Watson)	Conclusão
RGB	(i) LV1- LV20 foram utilizadas no modelo ^a				
	PRP for LV1	5.12	Não-linear	1.02	Não-linear
	RP	1.20	Linear	1.68	Inconclusivo
	e-LV1	0.28	Linear	1.31	Não-linear
	APaRP for LV1	2.96	Não-linear	1.38	Não-linear
	(ii) Somente LV1 foi utilizada no modelo ^b				
	RP	7.09	Não-linear	1.52	Não-linear
	e-LV1	5.88	Não-linear	0.56	Não-linear
	APaRP for LV1	4.59	Não-linear	0.62	Não-linear
	Escala cinza	(i) LV1- LV17 foram utilizadas no modelo ^c			
PRP for LV1		2.87	Não-linear	1.51	Não-linear
RP		5.10	Não-linear	1.69	Inconclusivo
e-LV1		5.57	Não-linear	0.82	Não-linear
APaRP for LV1		6.40	Não-linear	0.93	Não-linear
(ii) Somente LV1 foi utilizada no modelo ^b					
RP		6.78	Não-linear	1.82	Linear
e-LV1		5.85	Não-linear	0.60	Não-linear
APaRP for LV1		7.65	Não-linear	0.60	Não-linear
HSI		(i) LV1- LV18 foram utilizadas no modelo ^d			
	PRP for LV1	3.32	Não-linear	1.17	Não-linear
	RP	2.56	Não-linear	1.83	Inconclusivo
	e-LV1	2.86	Não-linear	1.30	Não-linear
	APaRP for LV1	3.76	Não-linear	1.49	Não-linear
	(ii) Somente LV1 foi utilizada no modelo ^b				
	RP	6.61	Não-linear	2.22	Linear
	e-LV1	4.27	Não-linear	0.69	Não-linear
	APaRP for LV1	5.77	Não-linear	0.82	Não-linear
	RGB + escala cinza	(i) LV1- LV22 foram utilizadas no modelo ^e			
PRP for LV1		4.52	Não-linear	0.97	Não-linear
RP		2.11	Não-linear	1.73	Inconclusivo
e-LV1		0.15	Linear	1.41	Não-linear
APaRP for LV1		3.22	Não-linear	1.50	Inconclusivo
(ii) Somente LV1 foi utilizada no modelo ^b					
RP		7.32	Não-linear	1.57	Não-linear
e-LV1		5.58	Não-linear	0.59	Não-linear
APaRP for LV1		6.08	Não-linear	0.71	Não-linear
HSI + escala cinza		(i) LV1- LV26 foram utilizadas no modelo ^f			
	PRP for LV1	2.86	Não-linear	1.17	Não-linear
	RP	2.66	Não-linear	2.09	Linear
	e-LV1	2.86	Não-linear	1.63	Inconclusivo
	APaRP for LV1	3.40	Não-linear	1.77	Inconclusivo
	(ii) Somente LV1 foi utilizada no modelo ^b				
	RP	6.33	Não-linear	2.19	Linear
	e-LV1	4.55	Não-linear	0.79	Não-linear
	APaRP for LV1	5.44	Não-linear	0.87	Não-linear
	RGB + HSI	(i) LV1- LV23 foram utilizadas no modelo ^g			
PRP for LV1		3.09	Não-linear	1.29	Não-linear
RP		0.35	Linear	1.79	Inconclusivo
e-LV1		0.61	Linear	1.55	Inconclusivo
APaRP for LV1		2.74	Não-linear	1.64	Inconclusivo
(ii) Somente LV1 foi utilizada no modelo ^b					
RP		6.62	Não-linear	2.07	Linear
e-LV1		3.23	Não-linear	0.90	Não-linear
APaRP for LV1		5.72	Não-linear	0.90	Não-linear
RGB + HSI + escala cinza		(i) LV1- LV25 foram utilizadas no modelo ^h			
	PRP for LV1	3.27	Não-linear	1.27	Não-linear
	RP	1.61	Linear	1.84	Inconclusivo
	e-LV1	1.01	Linear	1.62	Inconclusivo
	APaRP for LV1	3.09	Não-linear	1.70	Inconclusivo
	(ii) Somente LV1 foi utilizada no modelo ^b				
	RP	6.93	Não-linear	2.02	Linear
	e-LV1	3.65	Não-linear	0.90	Não-linear
	APaRP for LV1	5.79	Não-linear	0.90	Não-linear

^a Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1.96$, $d_L = 1.52$ e $d_u = 1.99$;^b Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1.96$, $d_L = 1.74$ e $d_u = 1.76$;^c Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1.96$, $d_L = 1.56$ e $d_u = 1.95$;^d Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1.96$, $d_L = 1.55$ e $d_u = 1.95$;^e Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1.96$, $d_L = 1.50$ e $d_u = 2.02$;^f Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1.96$, $d_L = 1.45$ e $d_u = 2.07$;^g Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1.96$, $d_L = 1.49$ e $d_u = 2.03$;^h Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1.96$, $d_L = 1.46$ e $d_u = 2.06$;

significativo de LVs podem modelar pequenas tendências não lineares presentes nos dados, entretanto, isso pode prejudicar a robustez do modelo (CENTNER; NOORD; MASSART, 1998; SEASHOLTZ; KOWALSKI, 1993). Dessa forma, é justificável as alterações entre as conclusões dos testes de detecção de não linearidade para modelos construídos com somente uma LV e modelos com mais de uma LV.

Segundo os testes de investigação de não-linearidade, a correlação entre os dados das imagens e o carbono orgânico do solo apresenta tendências não-lineares. Embora, um modelo PLS com um número expressivo de LVs pode modelar uma fração desse comportamento não-linear, um método não linear como o LS-SVM apresentou melhor performance que o PLS em todas as matrizes de dados, com exceção da escala cinza.

A correlação entre a cor do solo e os teores de carbono orgânico já foram reportados na literatura (VISCARRA-ROSSEL et al., 2006a; STIGLITZ et al., 2017). Viscarra-Rossel et al. (2006a) quantificaram o COS por meio da correlação deste parâmetro com a cor do solo, medida por um espectrofotômetro. A melhor correlação, $r^2 = 0,66$, foi obtida utilizando a coordenada R do modelo de cor RGB. Stiglitz et al. (2017) correlacionaram a cor do solo, mensurada por um sensor de cor, com os teores de COS em solos secos e úmidos. A melhor correlação apresentou um r^2 de 0,7978 e foi obtida utilizando amostras secas. No método desenvolvido utilizando MIA, as melhores correlações apresentaram r^2 superiores a 0,93. Além disso, as imagens foram obtidas por um *scanner*, um equipamento de baixo custo.

Além da cor, o COS também apresenta relação com a luminosidade presente no solo (SCHULZE et al., 1993). Isso explica a boa correlação observada entre o COS e as matrizes de dados com informações do sistema de cor HSI. Esse sistema, como já mencionado corresponde, além da matiz e da saturação, ao parâmetro intensidade que representa, de modo geral, a quantidade de luminosidade presente. Schulze et al. (1993) relacionou a luminosidade determinada pela carta de Munsell com os teores de COS. Para amostras de pouca variação textural, obteve-se um $r^2 > 0,9$. Por outro lado, utilizando amostras de solo com maior variabilidade granulométrica, esse grupo de pesquisa obteve-se baixas correlações ($r^2 < 0,4$). No presente trabalho que foi desenvolvido utilizando MIA, embora amostras de grande variabilidade textural (0,601 a 81,522 % de argila) foram utilizadas, obtiveram-se boas correlações ($r^2 > 0,9$).

Diante do apresentado, a fusão dos dados referentes aos modelos de cores RGB, HSI e escala cinza, agrega uma maior quantidade de informação relevante do solo. Dessa forma, a aplicação dos métodos multivariados para essa fusão de dados resultaria em melhores

resultados em relação aos sistemas individuais de cores. Isso se confirma somente para o modelo PLS, que apresentou os melhores resultados para a matriz de dados RGB + HSI e para a fusão dos três modelos de cores. Entretanto para o modelo SPA-MLR, os resultados obtidos entre todas as matrizes de dados foram equivalentes a um nível de 95% de confiança, com exceção das matrizes de dados: escala cinza e RGB + HSI. De forma similar, o modelo LS-SVM apresentou resultados equivalentes para todos os modelos de cores e suas fusões, com exceção das matrizes de dados: escala cinza e HSI.

A quantificação do COS por imagens digitais apresentou várias vantagens em relação ao método considerado como padrão Walkley-Black. O método proposto além de apresentar baixo custo de implementação, por volta de US\$ 200,00, não utiliza reagentes químico em sua metodologia e não gera resíduos tóxicos. Além disso, a metodologia apresentada demonstra maiores operacionalidade e produtividade, pois a cada dia é possível analisar pelo menos 200 amostras, praticamente o dobro do que é analisado pelo método Walkley-Black no laboratório de Análise Agroambiental da Embrapa Arroz e Feijão.

6.2.1.2 Quantificação da textura do solo

De modo semelhante a quantificação do COS, realizou-se a determinação das frações texturais dos solo, argila e areia por meio de imagens digitais.

6.2.1.2.1 Quantificação dos teores de argila

As Tabelas 11, 12 e 13 apresentam os resultados obtidos na quantificação dos teores de argila por meio de imagens digitais, respectivamente, para os métodos PLS, SPA-MLR e LS-SVM.

Tabela 11 – Parâmetros estatísticos da calibração e validação na quantificação de argila utilizando o método PLS.

Modelo de Cor	LVs	Calibração		Validação		
		r^2	RMSE	r^2	RMSE	RPD
RGB	18	0,7690	10,6	0,7119	10,8	1,8
Escala cinza	8	0,4311	16,6	0,4743	17,4	1,4
HSI	14	0,7087	11,9	0,6382	12,1	1,6
RGB + Escala cinza	16	0,7811	10,3	0,7452	10,2	2,0
HSI + Escala cinza	15	0,7229	11,4	0,7081	11,6	1,8
RGB + HSI	10	0,6182	13,5	0,6966	11,8	1,7
RGB + HSI + Escala cinza	18	0,8242	9,1	0,7789	9,7	2,1

Segundo a Tabela 11, o modelo PLS com melhor desempenho foi obtido utilizando a fusão dos modelos de cores RGB, HSI e escala cinza. Entretanto, com base no teste-F com 95% de confiança, não houve diferença significativa entre os valores de RMSEP de todos os modelos, com exceção do modelo construído utilizando somente a escala cinza.

A Figura 24 apresenta o gráfico dos teores de argila do método de referência contra os teores de argila previstas para o melhor modelo PLS.

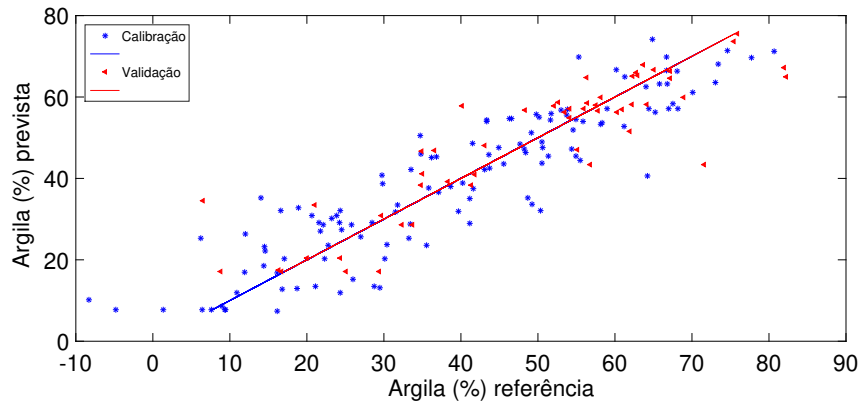


Figura 24 – Teores de argila de referência versus os teores de argila previstos pelo modelo PLS. Calibração (*) e validação (▲).

Tabela 12 – Parâmetros estatísticos da calibração e validação na quantificação de argila utilizando o método SPA-MLR.

Modelo de Cor	Número de variáveis	Calibração		Validação		
		r^2	RMSE	r^2	RMSE	RPD
RGB	15	0,5965	14,0	0,6205	12,2	1,6
Escala cinza	4	0,2248	19,4	0,5174	15,4	1,4
HSI	19	0,6346	13,3	0,5502	13,5	1,5
RGB + Escala cinza	21	0,7603	10,8	0,6988	11,1	1,8
HSI + Escala cinza	18	0,6546	12,8	0,5816	13,7	1,5
RGB + HSI	28	0,7701	10,5	0,6662	11,5	1,7
RGB + HSI + Escala cinza	20	0,7736	10,4	0,7484	10,5	1,9

De acordo com a Tabela 12, os melhores resultados foram obtidos utilizando a fusão dos modelos de cores RGB, HSI e escala cinza. Contudo, segundo o teste F com 95% de confiança, os modelos construídos com o espaço de cor RGB e as fusões de dados: RGB + escala cinza, RGB + HSI e RGB + HSI + escala cinza apresentaram valores de RMSEP estatisticamente iguais.

O gráfico dos teores de argila do método de referência contra os teores de argila previstos para o melhor modelo SPA-MLR está apresentado pela Figura 25.

De acordo com a Tabela 13, o método LS-SVM apresentou os melhores resultados na estimação dos teores de argila utilizando a fusão dos modelos de cores: RGB e escala

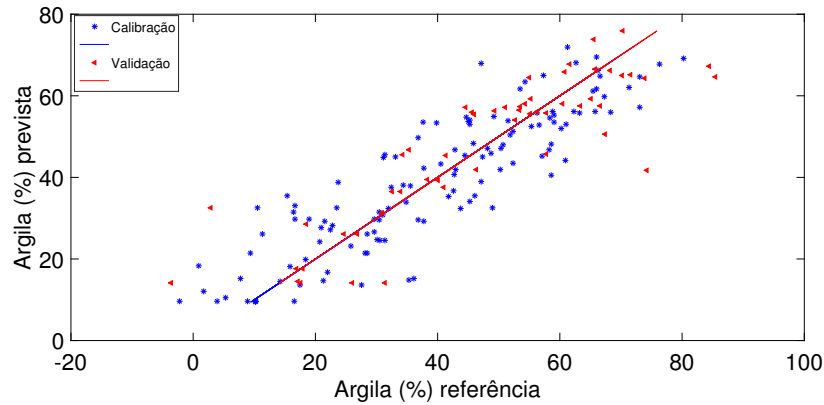


Figura 25 – Teores de argila de referência versus os teores de argila previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).

Tabela 13 – Parâmetros estatísticos da calibração e validação na quantificação de argila utilizando o método LS-SVM.

Modelo de cor	Calibração		Validação			γ	σ^2
	r^2	RMSE	r^2	RMSE	RPD		
RGB	0,9723	3,9	0,9204	5,7	3,4	14,847	160,902
Escala cinza	0,8929	7,7	0,8929	9,6	2,2	7,894	54,724
HSI	0,9350	6,0	0,8454	8,3	2,5	5,263	131,588
RGB + Escala cinza	0,9829	3,1	0,9618	4,0	5,1	21,602	240,466
HSI + Escala cinza	0,9669	4,2	0,9466	5,2	4,1	15,879	286,575
RGB + HSI	0,9476	5,3	0,9235	6,5	3,1	7,984	340,867
RGB + HSI + Escala cinza	0,9802	3,2	0,9575	4,4	4,6	26,459	515,159

cinza. Esta fusão apresentou um valor de RPD superior a 5, assim pode ser considerado como excelente segundo os critérios apontados por (SAEYS; MOUAZEN; RAMON, 2005). Os valores de RMSEP calculados para cada modelo foram comparados utilizando o teste- F a 95% de confiança. Segundo esse teste, somente os modelos construídos com as fusões de dados: RGB + escala cinza e RGB + HSI + escala cinza apresentaram valores de RMSEP estatisticamente iguais.

A Figura 26 apresenta o gráfico dos teores de argila do método de referência contra os teores de argila previstos para o melhor modelo LS-SVM.

Comparando os dados apresentados pelas tabelas 11, 12 e 13, nota-se a superioridade do método LS-SVM em relação aos métodos PLS e SPA-MLR em todos os modelos de cores e suas fusões. De acordo com o teste F com 95% de confiança, os valores de RMSEP para cada modelo de cor e suas fusões obtidos pelos métodos PLS e SPA-MLR não são estatisticamente iguais aos obtidos pelo método LS-SVM. Por outro lado, não houve diferença estatística entre os valores de RMSEP determinados pelos métodos PLS e SPA-MLR em todos os modelos de cor e suas fusões.

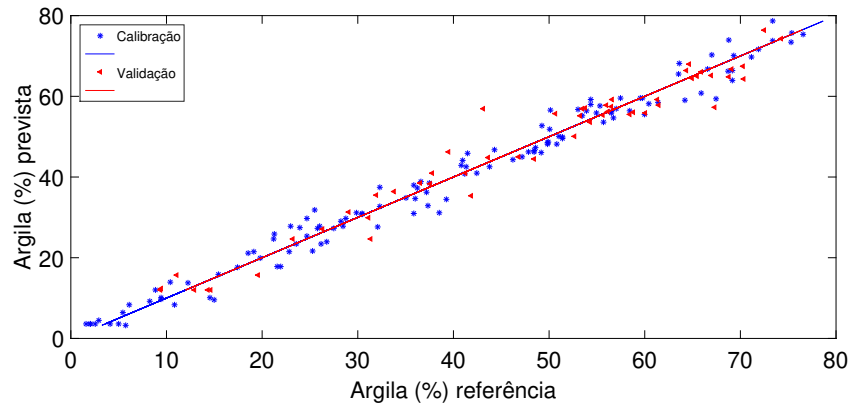


Figura 26 – Teores de argila de referência versus os teores de argila previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).

Com o intuito de comparar a performance dos três métodos de regressão na estimação de argila, aplicou-se o teste de randomização proposto por Voet (1994). Segundo esse teste, os métodos PLS e SPA-MLR não apresentaram a mesma acurácia que o método LS-SVM (p -valor $< 0,05$) em nenhum modelo de cor ou suas fusões. O desempenho excepcional da ferramenta LS-SVM também pode ser verificado comparando as figuras 24, 25 e 26. No gráfico de correlação obtido utilizando a ferramenta LS-SVM, nota-se uma maior proximidade dos pontos as linhas de correlação.

A não-linearidade dos dados também foi investigada utilizando dois métodos: *runs test* e Durbin-Watson test. Esses dois testes foram aplicados aos resíduos obtidos por quatro métodos: partial response plot for LV1 (PRP) (y versus LV1), residual plot (RP) (e versus y), plot the residuals versus latent variable e-LV1 (e versus LV1) and Mallows augmented partial residual plot (APaRP) ($e_{APaRP} + b_1LV_1 + b_{11}LV_1^2$ versus LV1). A Tabela 14 apresenta os resultados obtidos.

Segundo a Tabela 14 todos os modelos construídos com somente uma LV foram classificados como não lineares por todos os testes aplicados. Os *runs test* também apontaram que todos os modelos com mais de uma LV também são não-lineares. O teste Durbin-Watson aplicado aos resíduos RP foi inconclusivo para todos os modelos, com exceção do modelo construído com a matriz de dados RGB + HSI, que foi denominado como linear. O restante dos testes Durbin-Watson classificaram todos os modelos como não-lineares, com exceção do teste aplicado aos resíduos e-LV1 que foi inconclusivo para o modelo RGB + HSI e o teste aplicado aos resíduos APaRP que resultou inconclusivo para os modelos RGB e RGB + HSI.

Tabela 14 – Resultados obtidos na detecção da não linearidade dos dados por meio dos testes: *Runs* e Durbin-Watson

Modelo MIA	Plot	$ z $ (Runs tests)	Conclusão	d (Durbin-Watson)	Conclusão	
RGB	(i) LV1- LV18 foram utilizadas no modelo ^a					
		PRP for LV1	8,44	Não-linear	0,48	Não-linear
		RP	3,01	Não-linear	1,88	Inconclusivo
		e-LV1	3,32	Não-linear	1,42	Não-linear
		APaRP for LV1	4,18	Não-linear	1,66	Inconclusivo
	(ii) Somente LV1 foi utilizada no modelo ^b					
		RP	10,04	Não-linear	1,41	Não-linear
		e-LV1	9,93	Não-linear	0,06	Não-linear
		APaRP for LV1	8,14	Não-linear	0,97	Não-linear
	Escala cinza	(i) LV1- LV8 foram utilizadas no modelo ^c				
		PRP for LV1	8,34	Não-linear	1,33	Não-linear
		RP	6,63	Não-linear	1,84	Inconclusivo
		e-LV1	8,42	Não-linear	0,82	Não-linear
		APaRP for LV1	8,36	Não-linear	1,09	Não-linear
(ii) Somente LV1 foi utilizada no modelo ^b						
		RP	6,94	Não-linear	1,21	Não-linear
		e-LV1	9,02	Não-linear	0,28	Não-linear
		APaRP for LV1	9,03	Não-linear	0,39	Não-linear
HSI		(i) LV1- LV14 foram utilizadas no modelo ^d				
		PRP for LV1	8,42	Não-linear	0,52	Não-linear
		RP	3,53	Não-linear	1,85	Inconclusivo
		e-LV1	2,40	Não-linear	1,31	Não-linear
		APaRP for LV1	3,45	Não-linear	1,57	Não-linear
	(ii) Somente LV1 foi utilizada no modelo ^b					
		RP	10,67	Não-linear	1,35	Não-linear
		e-LV1	9,93	Não-linear	0,04	Não-linear
		APaRP for LV1	7,52	Não-linear	0,73	Não-linear
	RGB + escala cinza	(i) LV1- LV16 foram utilizadas no modelo ^e				
		PRP for LV1	8,73	Não-linear	0,54	Não-linear
		RP	2,86	Não-linear	1,78	Inconclusivo
		e-LV1	3,16	Não-linear	1,41	Não-linear
		APaRP for LV1	3,76	Não-linear	1,39	Não-linear
(ii) Somente LV1 foi utilizada no modelo ^b						
		RP	9,30	Não-linear	1,64	Não-linear
		e-LV1	9,65	Não-linear	0,21	Não-linear
		APaRP for LV1	8,72	Não-linear	0,68	Não-linear
HSI + escala cinza		(i) LV1- LV15 foram utilizadas no modelo ^f				
		PRP for LV1	8,44	Não-linear	0,53	Não-linear
		RP	3,11	Não-linear	1,79	Inconclusivo
		e-LV1	3,43	Não-linear	1,36	Não-linear
		APaRP for LV1	4,31	Não-linear	1,37	Não-linear
	(ii) Somente LV1 foi utilizada no modelo ^b					
		RP	9,34	Não-linear	1,56	Não-linear
		e-LV1	9,94	Não-linear	0,15	Não-linear
		APaRP for LV1	8,14	Não-linear	0,62	Não-linear
	RGB + HSI	(i) LV1- LV10 foram utilizadas no modelo ^g				
		PRP for LV1	8,34	Não-linear	0,69	Não-linear
		RP	2,66	Não-linear	1,98	Linear
		e-LV1	5,41	Não-linear	1,27	Inconclusivo
		APaRP for LV1	3,24	Não-linear	1,71	Inconclusivo
(ii) Somente LV1 foi utilizada no modelo ^b						
		RP	10,67	Não-linear	1,38	Não-linear
		e-LV1	9,93	Não-linear	0,05	Não-linear
		APaRP for LV1	7,54	Não-linear	0,95	Não-linear
RGB+ HSI + escala cinza		(i) LV1- LV18 foram utilizadas no modelo ^a				
		PRP for LV1	8,43	Não-linear	0,38	Não-linear
		RP	2,38	Não-linear	1,67	Inconclusivo
		e-LV1	3,32	Não-linear	1,41	Não-linear
		APaRP for LV1	4,52	Não-linear	1,36	Não-linear
	(ii) Somente LV1 foi utilizada no modelo ^b					
		RP	9,81	Não-linear	1,50	Não-linear
		e-LV1	9,63	Não-linear	0,11	Não-linear
		APaRP for LV1	7,84	Não-linear	0,83	Não-linear

^a Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,55$ e $d_u = 1,97$;^b Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,74$ e $d_u = 1,76$;^c Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,67$ e $d_u = 1,84$;^d Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,60$ e $d_u = 1,91$;^e Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,57$ e $d_u = 1,94$;^f Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,59$ e $d_u = 1,93$;^g Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,65$ e $d_u = 1,86$;

De acordo com os resultados dos testes de detecção de não-linearidade, a correlação entre os dados das imagens de solo e os teores de argila apresenta comportamento não-linear. Nesse sentido, é justificável que o método de caráter não-linear: LS-SVM apresente uma melhor performance que os métodos PLS e SPA-MLR.

6.2.1.2.2 Quantificação dos teores de areia

As Tabelas 15, 16 e 17 apresentam os resultados obtidos na quantificação dos teores de areia por meio de imagens digitais respectivamente para os métodos PLS, SPA-MLR e LS-SVM.

Tabela 15 – Parâmetros estatísticos da calibração e validação na quantificação de areia utilizando o método PLS.

Modelo de Cor	LVs	Calibração		Validação		
		r^2	RMSE	r^2	RMSE	RPD
RGB	18	0,7651	13,2	0,5979	14,3	1,5
Escala cinza	10	0,6051	17,0	0,4771	16,8	1,3
HSI	14	0,6744	15,2	0,4423	17,5	1,3
RGB + Escala cinza	19	0,8141	11,8	0,7204	11,7	1,9
HSI + Escala cinza	16	0,7529	13,5	0,7172	12,6	1,8
RGB + HSI	19	0,8350	11,0	0,7334	12,1	1,9
RGB + HSI + Escala cinza	20	0,8620	9,9	0,8439	10,2	2,5

De acordo com a Tabela 15, os melhores resultados na estimação dos teores de areia, foram obtidos utilizando a combinação dos modelos de cores RGB, HSI e escala cinza. No entanto, segundo o teste F com 95% de confiança, não houve diferença estatística entre os valores de RMSEP obtidos com os modelos que utilizaram as fusões de dados: RGB + escala cinza, HSI + escala cinza, RGB + HSI e RGB + HSI + escala cinza.

A Figura 27 apresenta o gráfico dos teores de areia do método de referência contra os teores de areia previstas para o melhor modelo PLS.

Tabela 16 – Parâmetros estatísticos da calibração e validação na quantificação de areia utilizando o método SPA-MLR.

Modelo de Cor	Número de variáveis	Calibração		Validação		
		r^2	RMSE	r^2	RMSE	RPD
RGB	7	0,4741	19,7	0,4942	15,4	1,4
Escala cinza	12	0,5483	18,2	0,5477	15,9	1,5
HSI	21	0,6276	16,3	0,4074	18,1	1,3
RGB + Escala cinza	23	0,7710	13,1	0,7148	11,5	1,9
HSI + Escala cinza	16	0,6718	15,5	0,5970	14,4	1,6
RGB + HSI	15	0,6507	16,0	0,6590	13,7	1,7
RGB + HSI + Escala cinza	18	0,7333	13,7	0,7350	12,4	1,9

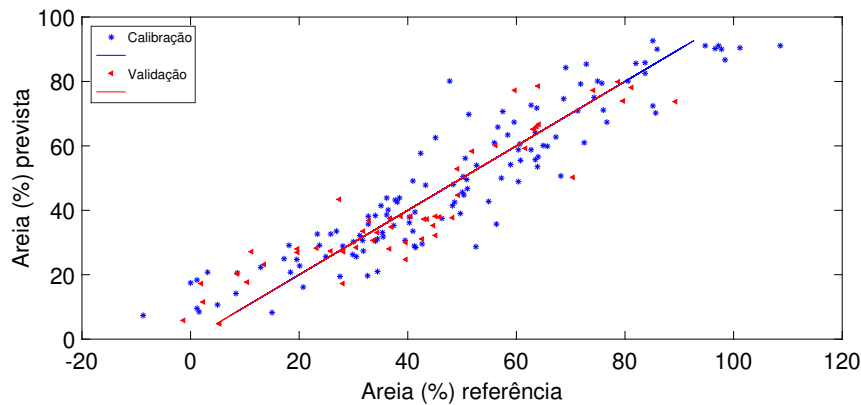


Figura 27 – Teores de areia de referência versus os teores de areia previstos pelo modelo PLS. Calibração (*) e validação (▲).

Segundo a Tabela 16, o modelo SPA com melhor performance na estimação de areia em amostras de solo foi obtido utilizando a combinação dos modelos de cores RGB e escala cinza. Porém, os valores de RMSEP calculados para os modelos obtidos utilizando as fusões de dados: RGB + escala cinza, HSI + escala cinza, RGB + HSI e RGB + HSI e escala cinza não apresentaram diferença significativa por meio do teste F a 95% de confiança.

O gráfico dos teores de areia do método de referência contra os teores de areia previstos para o melhor modelo SPA-MLR está apresentado pela Figura 28.

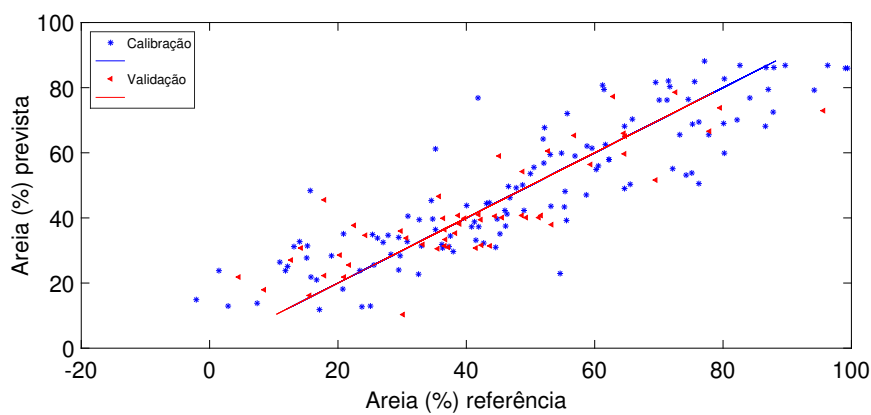


Figura 28 – Teores de areia de referência versus os teores de areia previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).

De acordo com a Tabela 17, os melhores resultados na estimação dos teores de areia foram obtidos utilizando a fusão HSI + escala cinza. O RPD calculado para esse modelo é superior a 4,5 e pode ser considerado como um modelo de calibração de excelentes previsões, de acordo com Saeys, Mouazen e Ramon (2005). Os valores de RMSEP também foram comparados por meio do teste F à um nível de 95% de confiança. Segundo esse

Tabela 17 – Parâmetros estatísticos da calibração e validação na quantificação de areia utilizando o método LS-SVM.

Modelo de cor	Calibração		Validação			γ	σ^2
	r^2	RMSE	r^2	RMSE	RPD		
RGB	0,9563	5,9	0,9295	6,2	3,5	22,473	235,772
Escala cinza	0,9706	4,8	0,7665	10,6	2,1	90,536	111,026
HSI	0,9384	7,1	0,9080	7,7	3,0	4,352	108,711
RGB + Escala cinza	0,9753	4,5	0,9251	5,9	3,6	14,353	241,126
HSI + Escala cinza	0,9736	4,5	0,9568	5,0	4,6	37,458	399,536
RGB + HSI	0,9628	5,4	0,9357	5,9	3,8	52,714	759,391
RGB + HSI + Escala cinza	0,9703	4,7	0,9452	5,9	4,1	76,697	1173,649

teste, com exceção do modelo obtido utilizando a escala cinza e o espaço de cor HSI, todos os modelos apresentaram valores de RMSEP sem diferença estatística.

A Figura 29 apresenta o gráfico dos teores de areia do método de referência contra os teores de areia previstos para o melhor modelo LS-SVM.

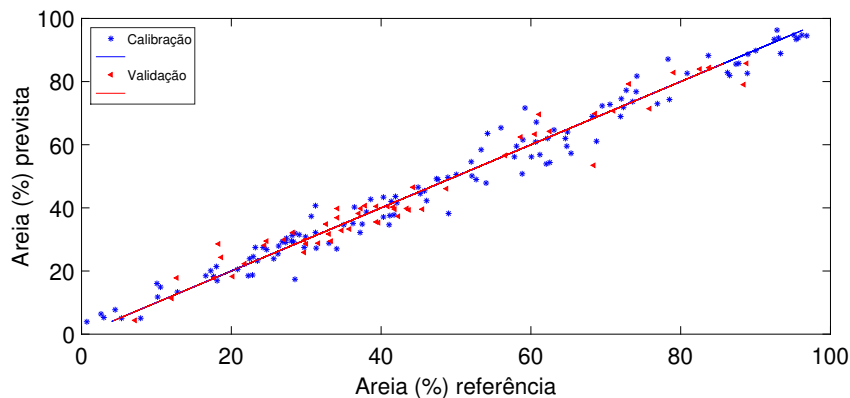


Figura 29 – Teores de areia de referência versus os teores de areia previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).

De acordo com as Tabelas 15, 16 e 17, o método LS-SVM apresentou melhor performance que os métodos PLS e SPA-MLR na estimação dos teores de areia independente de qual modelo de cor ou suas respectivas combinações. Os valores de RMSEP determinados para cada modelo de cor e suas fusões foram comparados entre os métodos: PLS, SPA-MLR e LS-SVM por meio do teste F. De acordo com esse teste, os valores de RMSEP calculados pelos métodos SPA-MLR e PLS não são estatisticamente iguais aos valores de RMSEP determinados pelo método LS-SVM. Isto é, o método LS-SVM apresenta a maior precisão entre os três métodos. Em contrapartida, os valores de RMSEP determinados pelos métodos SPA-MLR e PLS são estatisticamente iguais independente de qual modelo de cor ou fusão. Em outras palavras, os métodos PLS e SPA-MLR apresentam a mesma precisão na estimação de areia.

Os modelos de regressão utilizados para estimação de areia em solo também foram comparados utilizando o teste de randomização. De acordo com esse teste, os métodos SPA-MLR e PLS não apresentaram a mesma acurácia que o método LS-SVM (p-valor < 0.05). Segundo as Figuras 27, 28 e 29 é fácil notar a superioridade da técnica LS-SVM, pois os pontos encontram-se bem mais próximos as retas.

O comportamento não-linear dos dados também foi verificada por meio dois métodos: *runs test* e Durbin-Watson test. Ambos testes foram aplicados aos resíduos obtidos por quatro métodos: partial response plot for LV1 (PRP) (y versus LV1), residual plot (RP) (e versus y), plot the residuals versus latent variable e-LV1 (e versus LV1) and Mallows augmented partial residual plot (APaRP) ($e_{APaRP} + b_1LV_1 + b_{11}LV_1^2$ versus LV1). A Tabela 18 apresenta os resultados obtidos.

De acordo a Tabela 18 o *runs test* classificou todos os modelos como não lineares. Os testes Durbin-Watson classificaram todos os modelos com somente uma LV como não-lineares, com exceção do teste aplicado aos resíduos RP dos modelos RGB e RGB + HSI. Os testes Durbin-Watson aplicados aos modelos RGB + HSI e RGB + HSI + escala cinza ambos com mais de um LV1 foram inconclusíveis, com exceção do teste PRP para LV1 que classificou esses modelos como não-lineares. O modelo HSI + escala cinza foi classificado como não-linear por dois testes Durbin-Watson e inconclusíveis por outros dois testes. O modelo RGB + escala cinza foi classificado como linear por dois testes (RP e APaRP), não-linear pelo teste PRP para LV1 e inconclusivo pelo teste e-LV1. O modelo HSI foi classificado como não linear por dois testes (PRP para LV1 e APaRP), linear pelo teste RP e inconclusivo pelo teste APaRP. O modelo escala cinza foi classificado como não-linear por todos os testes, com exceção do teste RP que classificou esse modelo como linear. O modelo RGB foi classificado como linear por dois testes (RP e APaRP), não-linear pelo teste PRP para LV1 e inconclusivo para e-LV1.

Os modelos PLS construídos para predição dos teores de areia no solo foram construídos utilizando um número expressivo de LVs (≥ 10). Modelos PLS com um alto número de LVs podem, geralmente, modelar pequenos desvios de linearidade (CENTNER; NOORD; MASSART, 1998; SEASHOLTZ; KOWALSKI, 1993). Dessa forma, é esperado alterações entre as conclusões dos testes de verificação de não-linearidade dos dados entre modelos com somente uma LV e modelos com mais de uma LV. A correlação entre os dados referentes das imagens digitais do solo e os teores de areia apresenta desvios de linearidade. Nesse sentido, mesmo com a capacidade de modelagem de pequenos desvios de linearidade do

Tabela 18 – Resultados obtidos na detecção da não linearidade dos dados por meio dos testes: *Runs* e Durbin-Watson

Modelo MIA	Plot	$ z $ (Runs tests)	Conclusão	d (Durbin-Watson)	Conclusão	
RGB	(i) LV1- LV18 foram utilizadas no modelo ^a					
		PRP for LV1	8,44	Não-linear	0,72	Não-linear
		RP	2,96	Não-linear	2,23	Linear
		e-LV1	3,30	Não-linear	1,64	Inconclusivo
		APaRP for LV1	3,35	Não-linear	2,05	Linear
	(ii) Somente LV1 foi utilizada no modelo ^b					
		RP	8,44	Não-linear	1,56	Inconclusivo
		e-LV1	8,98	Não-linear	0,16	Não-linear
		APaRP for LV1	8,27	Não-linear	0,58	Não-linear
	Escala cinza	(i) LV1- LV10 foram utilizadas no modelo ^c				
		PRP for LV1	5,12	Não-linear	1,11	Não-linear
		RP	3,56	Não-linear	2,00	Linear
		e-LV1	4,97	Não-linear	1,18	Não-linear
		APaRP for LV1	4,37	Não-linear	1,40	Não-linear
(ii) Somente LV1 foi utilizada no modelo ^b						
		RP	7,30	Não-linear	1,03	Não-linear
		e-LV1	9,67	Não-linear	0,17	Não-linear
		APaRP for LV1	8,58	Não-linear	0,27	Não-linear
HSI		(i) LV1- LV14 foram utilizadas no modelo ^d				
		PRP for LV1	7,82	Não-linear	0,71	Não-linear
		RP	3,70	Não-linear	1,92	Linear
		e-LV1	5,72	Não-linear	1,27	Não-linear
		APaRP for LV1	3,66	Não-linear	1,78	Inconclusivo
	(ii) Somente LV1 foi utilizada no modelo ^b					
		RP	10,33	Não-linear	1,53	Não-linear
		e-LV1	9,20	Não-linear	0,05	Não-linear
		APaRP for LV1	9,07	Não-linear	0,39	Não-linear
	RGB + escala cinza	(i) LV1- LV19 foram utilizadas no modelo ^e				
		PRP for LV1	8,42	Não-linear	0,67	Não-linear
		RP	3,53	Não-linear	2,07	Linear
		e-LV1	3,01	Não-linear	1,68	Inconclusivo
		APaRP for LV1	5,70	Não-linear	2,10	Linear
(ii) Somente LV1 foi utilizada no modelo ^b						
		RP	7,44	Não-linear	1,30	Não-linear
		e-LV1	9,50	Não-linear	0,22	Não-linear
		APaRP for LV1	8,09	Não-linear	0,58	Não-linear
HSI + escala cinza		(i) LV1- LV16 foram utilizadas no modelo ^f				
		PRP for LV1	6,80	Não-linear	0,62	Não-linear
		RP	3,43	Não-linear	1,89	Inconclusivo
		e-LV1	3,44	Não-linear	1,42	Não-linear
		APaRP for LV1	4,12	Não-linear	1,90	Inconclusivo
	(ii) Somente LV1 foi utilizada no modelo ^b					
		RP	8,74	Não-linear	1,17	Não-linear
		e-LV1	9,78	Não-linear	0,12	Não-linear
		APaRP for LV1	9,65	Não-linear	0,24	Não-linear
	RGB + HSI	(i) LV1- LV19 foram utilizadas no modelo ^e				
		PRP for LV1	8,12	Não-linear	0,46	Não-linear
		RP	2,10	Não-linear	1,93	Inconclusivo
		e-LV1	2,10	Não-linear	1,62	Inconclusivo
		APaRP for LV1	5,23	Não-linear	1,93	Inconclusivo
(ii) Somente LV1 foi utilizada no modelo ^b						
		RP	9,81	Não-linear	1,55	Inconclusivo
		e-LV1	8,36	Não-linear	0,11	Não-linear
		APaRP for LV1	8,24	Não-linear	0,59	Não-linear
RGB+ HSI + escala cinza		(i) LV1- LV20 foram utilizadas no modelo ^g				
		PRP for LV1	7,48	Não-linear	0,47	Não-linear
		RP	3,31	Não-linear	1,96	Inconclusivo
		e-LV1	3,31	Não-linear	1,68	Inconclusivo
		APaRP for LV1	5,57	Não-linear	1,94	Inconclusivo
	(ii) Somente LV1 foi utilizada no modelo ^b					
		RP	9,60	Não-linear	1,36	Não-linear
		e-LV1	9,81	Não-linear	0,15	Não-linear
		APaRP for LV1	7,47	Não-linear	0,56	Não-linear

^a Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,55$ e $d_u = 1,97$;^b Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,74$ e $d_u = 1,76$;^c Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,64$ e $d_u = 1,86$;^d Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,60$ e $d_u = 1,91$;^e Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,53$ e $d_u = 1,98$;^f Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,59$ e $d_u = 1,93$;^g Os valores críticos de $|z|$ e d para cada *runs* e Durbin-Watson com α de 0,05 são, respectivamente, $z_{crit} = 1,96$, $d_L = 1,52$ e $d_u = 1,99$;

modelo PLS, a ferramenta LS-SVM, com capacidade de modelagem não-linear, apresentou melhores resultados que o PLS e o SPA-MLR em todas as matrizes de dados.

A relação entre as partículas físicas, areia, silte e argila e a cor do solo já foram apontadas na literatura (SÁNCHEZ-MARAÑÓN et al., 1997; TORRENT; BARRÓN, 1993; KONEN; BURRAS; SANDOR, 2003b). Segundo Konen, Burras e Sandor (2003b) a cor do solo apresenta alta correlação com os teores de argila e areia. Spielvogel, Knicker e Kögel-Knabner (2004) observaram que solos arenosos apresentam coloração mais escura que solos argilosos. Essa diferença de coloração é devida principalmente a constituição das frações de areia, argila e silte. As partículas de areia são formadas por quartzo e por isso refletem menos luz visível que as partículas de silte e argila, formadas especialmente por feldspato, mica, quartzo e flossilicato (CLARK, 1999; BLUME et al., 2016).

Murti e Satyanarayana (1971) correlacionaram os parâmetros de cor de Munsell: valor, matiz e saturação com algumas características do solo, dentre elas, os teores de argila. A melhor correlação, obtida utilizando o parâmetro valor referente a luminosidade do solo, obteve um r^2 de 0,60. Levin, Ben-Dor e Singer (2005) quantificaram os teores das partículas finas (argila e silte) por meio de imagens digitais. Após a obtenção das imagens e da região de interesse (ROI), as médias dos canais de cores: vermelho, verde e azul foram calculadas. Por meio desses valores, o índice de cor do solo vermelhidão (RI - *Redness index*) foi determinado por meio da Eq. (37)

$$RI = \frac{R^2}{(B \times G^3)} \quad (37)$$

onde R , G e B correspondem, respectivamente, aos valores médios dos canais de cores vermelho, verde e azul. O r^2 da correlação entre o índice de vermelhidão e as partículas finas foi de 0,81. No estudo desenvolvido por Levin, Ben-Dor e Singer (2005) somente um conjunto de 42 amostras foi utilizado. No presente estudo, um total de 177 amostras foram utilizadas. Além disso, os melhores modelos de predição dos teores de argila e areia obtiveram, respectivamente, r^2 de 0,9618 e 0,9568.

Diniz et al. (2006) quantificaram dentre outros parâmetros, os teores de argila do solo por meio do parâmetro da cor: valor, determinado pela carta de Munsell e por um colorímetro. Somente um total de 24 amostras foi utilizado. A correlação obtida entre a argila e o valor foi de $r^2 = 0,64$. Neste trabalho, o conjunto amostral foi mais expressivo e além disso, obteve-se um $r^2 > 0,96$.

Chung et al. (2012) determinaram as partículas físicas do solo: argila e areia por intermédio de histogramas de cor, extraídos de imagens digitais de amostras de solo. Após a extração dos histogramas, determinaram-se vários parâmetros, como a média, mediana e moda dos histogramas, a assimetria e o brilho. Em seguida, esses parâmetros foram correlacionados com os teores de argila e areia. O melhor modelo de predição de areia, obteve um r^2 de 0,35, RMSEC e RMSEP, respectivamente, de 10,3 % e 12,1 %. Já na predição de argila, o melhor model obteve um r^2 de 0,17, RMSEC e RMSEP, respectivamente, de 10,8 % e 8,4 %. No presente estudo, o melhor modelo de predição de argila, obteve um r^2 de 0,9618 e RMSEP de 4,0 %. Já na predição de areia, determinou-se um r^2 de 0,9568 e RMSEP de 5,0 %.

Aitkenhead et al. (2013) determinaram vários parâmetros do solo, dentre eles: os teores de areia e argila por meio dos valores dos canais dos sistemas de cores: RGB e CIELab. Tais valores foram obtidos por meio da conversão dos parâmetros da carta de Munsell: valor, matiz e saturação. A correlação entre os dados dos sistemas RGB e CIELab e os parâmetros do solo foram realizados utilizando redes neurais. Para os dados do modelo RGB, as melhores correlações entre os valores de referência e previstos obtiveram $r^2 = 0,584$ para argila e 0,524 para areia. Já para o modelo CIELab, obtiveram correlação com $r^2 = 0,661$ para argila e 0,590 para areia.

Recentemente, nosso grupo de pesquisa desenvolveu uma metodologia para análise textural do solo por meio de imagens digitais adquiridas por um microscópio estéreo (MORAIS et al., 2019). Para tanto, após a obtenção das imagens e extração dos histogramas de ocorrência de cor, os mesmos foram correlacionados com os teores de argila e areia utilizando o método de calibração PLS2. A correlação para os valores previstos versus de referência obteve r^2 de 0,8890 para argila e de 0,9307 para areia. Além disso, os valores de RMSEP calculados para a estimação da argila e areia foram, respectivamente, 4,8 e 5,1. Neste trabalho, resultados similares foram encontrados, mesmo utilizando um conjunto amostra de maior variabilidade (Tabela 5). Além disso, as imagens foram obtidas utilizando um *scanner* comercial que é de menor custo que um microscópio estéreo.

Perante o exposto, é justificável a correlação encontrada entre os histogramas de cor e os teores de argila e areia. As melhores correlações, tanto para areia e argila, foram obtidas, de forma geral, utilizando a fusão de modelos de cores. Isso, permitiu uma maior informação sobre a amostra e conseqüentemente melhores predições.

A determinação da granulometria do solo por meio de imagens digitais é uma metodologia simples, de baixo custo e com reduzido impacto ambiental. Além disso, esse método apresenta uma maior operacionalidade e rapidez que a metodologia considerada como padrão, método da Pipeta.

6.2.1.3 Classificação textural por meio de imagens digitais

As 177 amostras de solo foram classificadas em 8 grupos texturais, segundo suas proporções de argila, silte e areia quantificados pelo método de referência, método da Pipeta, Figura 17.

Em seguida, as amostras foram subdivididas entre calibração e validação utilizando o algoritmo Kennard-Stone (KENNARD; STONE, 1969). Posteriormente o método LS-SVM foi utilizado com o intuito de classificar as amostras do conjunto de validação em grupos texturais. A Tabela 19 apresenta a taxa de acerto na classificação textural utilizando imagens digitais em relação a classificação pelo método da Pipeta.

Tabela 19 – Taxas de acerto na classificação textural por meio de imagens.

Modelo de cor	Taxa de acerto (%)
RGB	90,6
Escala cinza	88,7
HSI	90,6
RGB + Escala cinza	90,6
HSI + Escala cinza	90,6
RGB + HSI	90,6
RGB + HSI + Escala cinza	90,6

De acordo com a Tabela 19, com exceção da escala cinza, todos os modelos de cores e suas fusões apresentaram taxa de acerto de 90,6%. Esse resultado é um indicativo da aplicabilidade do método de classificação textural por meio de imagens digitais.

Chung et al. (2012) predizeram os teores de argila e areia por meio de imagens digitais e em seguida classificaram as amostras de solo em classes texturais de acordo com o sistema de classificação adotado pelo Departamento de Agricultura dos Estados Unidos. O nível de acerto entre a classificação dada pelos resultados do método da pipeta e pelas imagens digitais foi de 48%. No presente trabalho, o nível de acerto foi bastante expressivo, acima de 94%. Diante disso, o método proposto além de não utilizar reagentes químicos e de não gerar resíduos tóxicos, apresenta uma alta produtividade de análise.

6.2.1.4 Determinação de fósforo, sódio, potássio e micronutrientes disponíveis

Utilizando um novo conjunto de 105 amostras de solo, realizou-se a determinação dos teores de fósforo, sódio, potássio assim como dos micronutrientes: Zn, Cu, Fe e Mn disponíveis. Em seguida, obtiveram-se imagens digitais de cada uma das 105 amostras de solo. De cada, ROIs foram selecionados e histogramas de ocorrência dos sistemas de cor RGB, HSI e escala cinza foram obtidos, desdobrados e justapostos na forma de matrizes. Os dados das imagens assim como os vetores contendo os dados dos elementos mencionados foram subdivididos entre calibração e validação utilizando o algoritmo SPXY. A correlação dos dados de imagens com as concentrações dos elementos determinados foi realizada utilizando os métodos: PLS, SPA-MLR e LS-SVM.

Não houve correlações significativas entre os dados das imagens e os teores de fósforo, sódio, potássio e micronutrientes disponíveis. Essa ausência de correlação é justificada devido a pouca ou nenhuma influência que esses componentes exercem na cor do solo.

6.2.1.5 Determinação dos óxidos de ferro, silício e alumínio em solo

Os teores de Fe_2O_3 , SiO_2 e Al_2O_3 de 52 amostras de solo das 177 do conjunto amostral foram quantificados após o ataque sulfúrico por meio de ICP e absorção atômica. A Tabela 20 apresenta a descrição estatística dos teores de COS, argila, areia e silte, assim como de Fe_2O_3 , SiO_2 e Al_2O_3 das 52 amostras.

Tabela 20 – Propriedades estatísticas das 52 amostras de solo.

Estatística descritiva	Constituintes texturais (%)			COS (g kg ⁻¹)	Fe_2O_3 (g kg ⁻¹)	SiO_2 (g kg ⁻¹)	Al_2O_3 (g kg ⁻¹)
	Argila	Silte	Areia				
Média	40,806	16,065	43,129	7,016	43,596	165,846	135,192
Desvio Padrão	21,304	10,311	26,940	6,581	31,280	89,289	72,331
Máximo	78,300	40,100	90,000	29,807	123,000	340,000	251,000
Mínimo	8,000	2,000	0,800	1,016	10,000	17,000	31,000

Após a obtenção das imagens de solo utilizando *scanner* das 52 amostras, selecionaram-se ROIs de cada imagem. Em seguida, histogramas de ocorrência de cor referente aos sistemas RGB, HSI e escala cinza foram obtidos, desdobrados e justapostos na forma de matrizes. A matriz de dados contendo a fusão dos histogramas dos três sistemas de cor (RGB, HSI e escala cinza), juntamente com os vetor \mathbf{y} s, referente aos teores de Fe_2O_3 , SiO_2 e Al_2O_3 , foram subdivididas entre calibração e validação utilizando a ferramenta SPXY. Desta forma, um total de 39 amostras foram designadas para calibração e 13

para validação. Posteriormente, utilizou-se os métodos de calibração multivariados: PLS, SPA-MLR e LS-SVM para correlacionar a matriz contendo os dados de imagens com os teores de Fe_2O_3 , SiO_2 e Al_2O_3 . O número ideal de LVs para o método PLS foi selecionado de acordo com os menores valores de RMSEP.

6.2.1.5.1 Determinação de Fe_2O_3 em amostras de solo

A Tabela 21 apresenta os resultados estatísticos obtidos na correlação dos dados de imagens com os teores de Fe_2O_3 em amostras de solo utilizando os métodos: PLS, SPA-MLR e LS-SVM.

Tabela 21 – Resultados estatísticos obtidos na estimação de Fe_2O_3 em solo por meio dos métodos: SPA-MLR, PLS e LS-SVM.

	Calibração					Validação			
	LV	Número de Variáveis	γ	σ^2	R^2_{cal}	RMSEC	R^2_{val}	RMSEP	RPD
SPA-MLR	-	14	-	-	0,9063	9,8	0,9423	7,5	3,6
PLS	3	-	-	-	0,5449	21,6	0,3524	22,9	1,2
LS-SVM	-	-	1601,289	31356,791	0,9648	6,2	0,9785	6,2	5,2

As Figuras 30, 31 e 32 apresentam os gráficos dos teores de Fe_2O_3 determinados pelo método de referência contra os teores de Fe_2O_3 previstos, respectivamente, para os modelos PLS, SPA-MLR e LS-SVM.

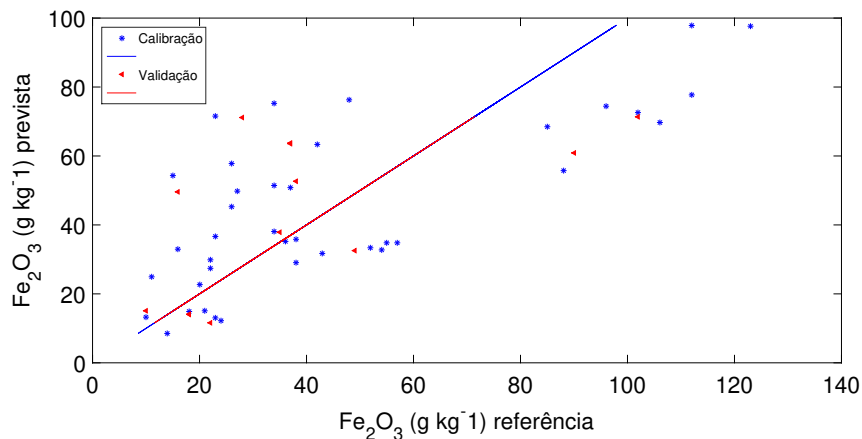


Figura 30 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo PLS. Calibração (*) e validação (▲).

Segundo a Tabela 21, o modelo LS-SVM apresentou melhor performance que os métodos PLS e SPA-MLR. Os valores de RPD, assim como r^2 determinados para o modelo LS-SVM foram superiores entre os métodos considerados. Entretanto, de acordo com o teste F a um nível de 95% de confiança, o valor de RMSEP calculado para o modelo

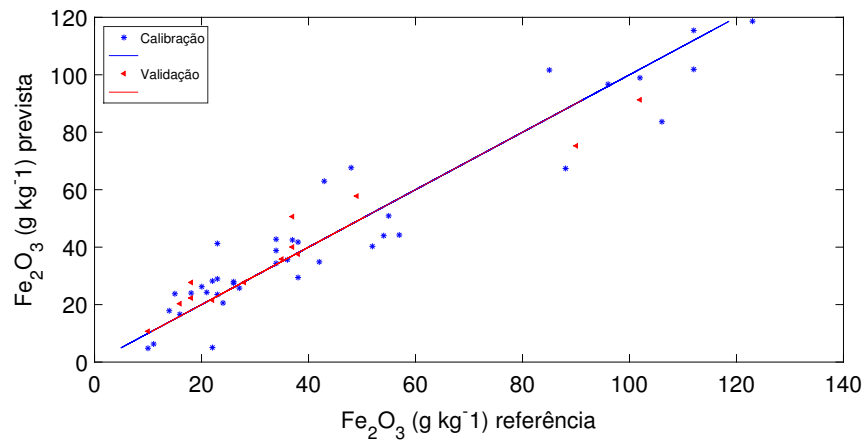


Figura 31 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).

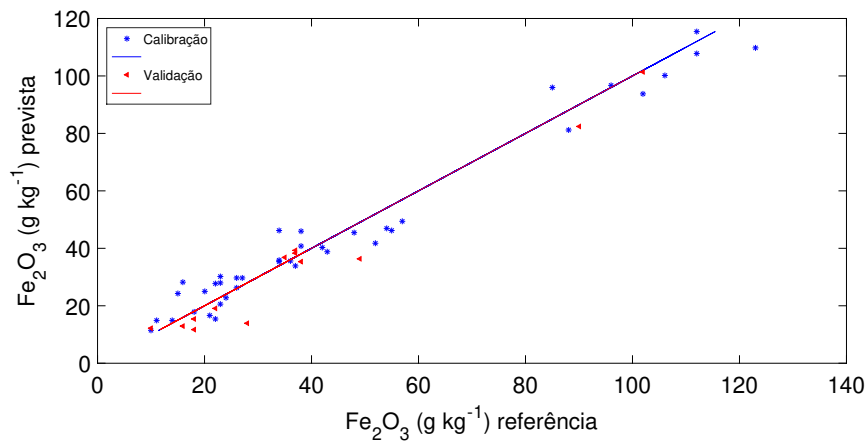


Figura 32 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).

LSSVM é estatisticamente igual ao valor de RMSEP determinado para o modelo SPA-MLR. Por outro lado, os valores de RMSEP determinados para os métodos SPA-MLR e LS-SVM são estatisticamente diferentes dos valores de RMSEP determinado para o modelo PLSR.

A relação entre os teores de Fe_2O_3 e a cor do solo já foi reportada em outros estudos (SCHWERTMANN, 1993; TORRENT et al., 1983; BARRÓN; TORRENT, 1986). Torrent, Schwertmann e Schulze (1980) quantificaram os teores de hematita de amostras de solo finas por meio do índice de vermelhidão estimados pela carta de Munsell. Utilizando um total de 11 amostras de solo, uma correlação com $r = 0.972$ foi obtido. Neste estudo, mesmo com um conjunto amostral maior (Tabela 20), correlações similares foram obtidas na estimação de Fe_2O_3 em solo com $r^2 > 0.96$ para calibração e validação. Barrón e Torrent (1986) quantificaram os teores de hematita em solo por meio da carta de Munsell e do sistema de cor CIELAB. As melhores correlações obtidas apresentaram $r^2 = 0.67$

para a carta de Munsell e $r^2 = 0.87$ para o sistema de cor CIELAB. Essas correlações foram menos significativas do que as encontradas neste estudo ($r^2 > 0.96$).

Levin, Ben-Dor e Singer (2005) quantificaram os teores de óxido de ferro livre em amostras de solo por meio de imagens digitais. Após a obtenção das imagens, os valores médio do sistema de cor RGB foram determinados e correlacionados com os teores de óxido de ferro livre. Essa correlação apresentou $r^2 = 0.89$, uma correlação menor do que a determinada neste trabalho. Moritsuka et al. (2014) determinaram os teores de ferro por meio da cor do solo medida por um leitor de cor. A correlação entre os parâmetros de cor e os teores de ferro apresentou um $r^2 = 0.77$, uma correlação inferior a obtida neste trabalho ($r^2 > 0.96$).

6.2.1.5.2 Determinação de SiO_2 em amostras de solo

A Tabela 22 apresenta os resultados estatísticos obtidos na correlação dos dados de imagens com os teores de SiO_2 em amostras de solo utilizando os métodos: PLS, SPA-MLR e LS-SVM.

Tabela 22 – Resultados estatísticos obtidos na estimaco de SiO_2 em solo por meio dos métodos: SPA-MLR, PLS e LS-SVM.

	Calibrao						Validao		
	LV	Nmero de variveis	γ	σ^2	R^2_{cal}	RMSEC	R^2_{val}	RMSEP	RPD
SPA-MLR	-	11	-	-	0,7004	48,8	0,6340	47,7	1,4
PLS	8	-	-	-	0,8224	37,6	0,7900	91,8	2,2
LS-SVM	-	-	8,383	237,376	0,9885	12,1	0,9516	23,4	3,7

As Figuras 33, 34 e 35 apresentam os grficos dos teores de SiO_2 determinados pelo mtodo de referncia contra os teores de SiO_2 previstos, respectivamente, para os modelos PLS, SPA-MLR e LS-SVM.

De acordo com a Tabela 22, a ferramenta LS-SVM apresentou os melhores resultados na estimaco de SiO_2 em amostras de solo em comparao com os mtodos PLS e SPA-MLR. Segundo o teste F a um nvel de 95% de confiana, houve diferena estatstica entre o valor de RMSEP determinado para o mtodo LS-SVM e os valores de RMSEP calculados para os mtodos PLS e SPA-MLR.

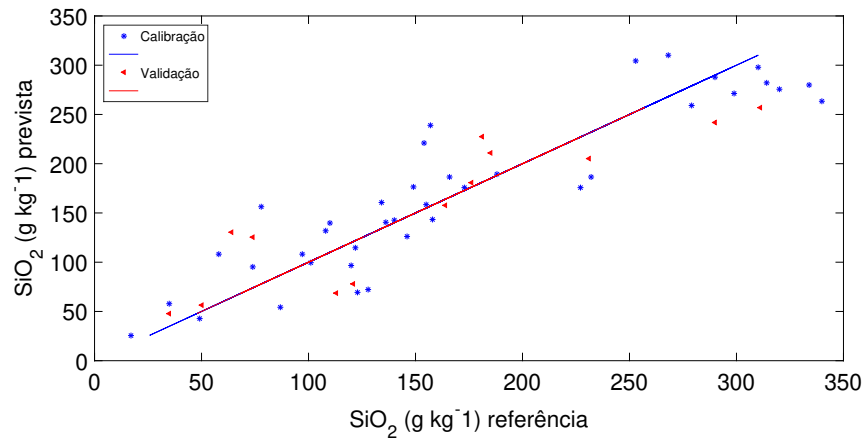


Figura 33 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo PLS. Calibração (*) e validação (▲).

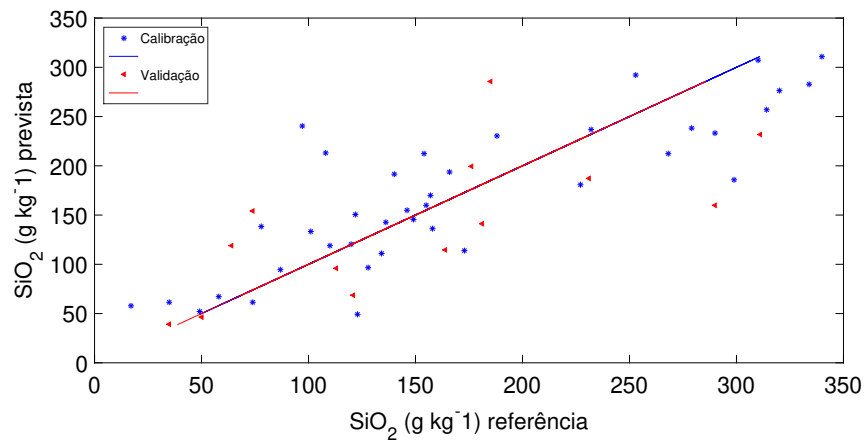


Figura 34 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).

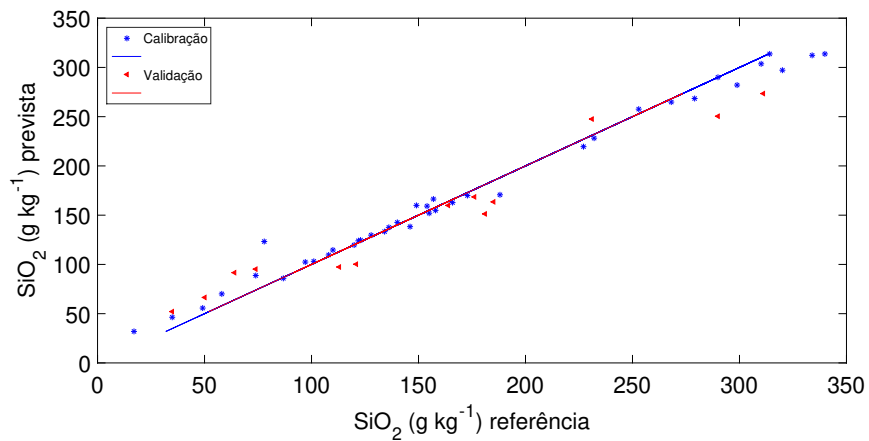


Figura 35 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).

6.2.1.5.3 Determinação de Al_2O_3 em amostras de solo

Os resultados obtidos na estimação de Al_2O_3 em amostras de solo por meio dos métodos: PLS, SPA-MLR e LS-SVM estão expostos na Tabela 23.

Tabela 23 – Resultados estatísticos obtidos na predição de Al_2O_3 por meio dos métodos SPA-MLR, PLSR e LS-SVMR.

	Calibração					Validação			
	LV	Número de Variáveis	γ	σ^2	R^2_{cal}	RMSEC	R^2_{val}	RMSEP	RPD
SPA-MLR	-	11	-	-	0,8154	31,6	0,7210	31,9	1,9
PLS	10	-	-	-	0,8758	25,9	0,8015	28,1	2,2
LS-SVM	-	-	14,609	290,888	0,9929	7,9	0,9460	13,1	4,7

As Figuras 36, 37 e 38 apresentam os gráficos dos teores de Al_2O_3 determinados pelo método de referência contra os teores de Al_2O_3 previstos, respectivamente, para os modelos PLS, SPA-MLR e LS-SVM.

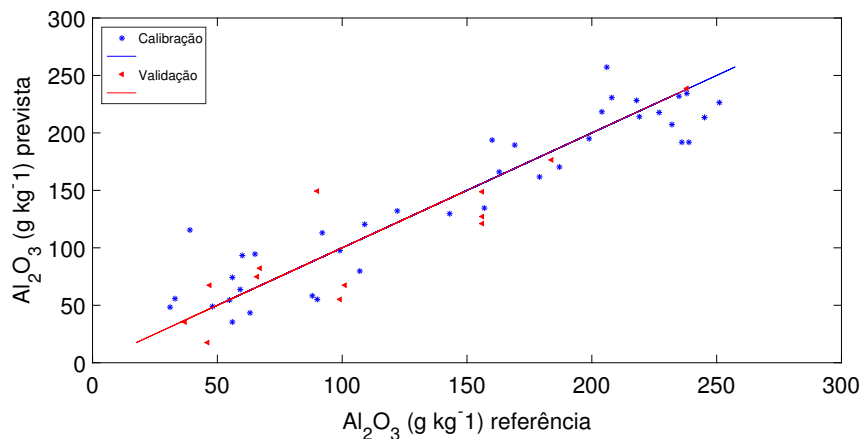


Figura 36 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo PLS. Calibração (*) e validação (▲).

De acordo com a Tabela 23, o método LS-SVM apresentou a melhor performance na predição de Al_2O_3 em amostras de solo. Segundo o teste F à um nível de confiança de 95%, os valores de RMSEP determinados para os métodos PLS e SPA-MLR são estatisticamente superiores ao valor de RMSEP calculado para o método LS-SVM. Os métodos SPA-MLR e PLS apresentaram performance semelhante na estimação de Al_2O_3 .

Diante dos resultados, a determinação de Fe_2O_3 , SiO_2 e Al_2O_3 em amostras de solo por meio de imagens digitais consiste em uma alternativa viável, de baixo custo, não-destrutiva e ainda de pouco impacto ambiental. Por meio dessa nova metodologia é possível determinar os teores de Fe_2O_3 , SiO_2 e Al_2O_3 de forma simultânea, o que garante um aumento considerável na produtividade de ensaio. Nesse procedimento não há utilização

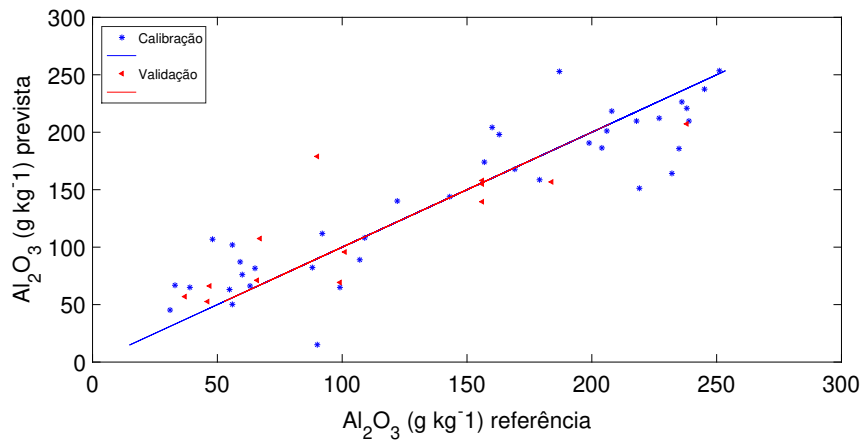


Figura 37 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).

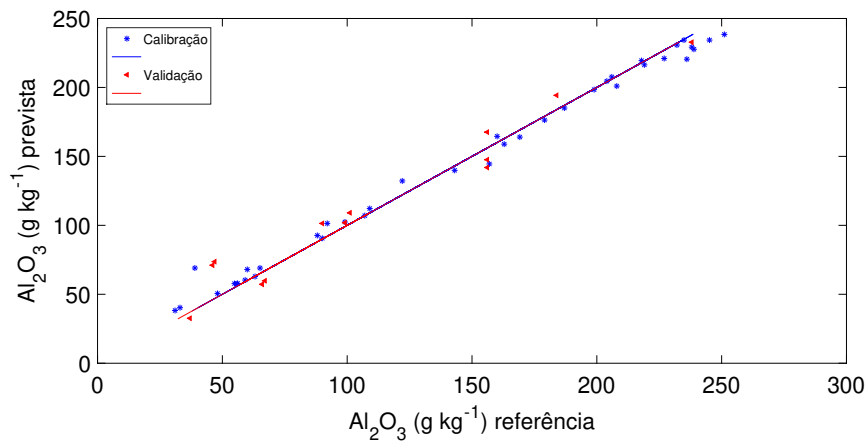


Figura 38 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).

de reagentes químicos. Dessa forma, além de promover um ambiente de trabalho mais seguro, com menor risco de acidente para o laboratorista, há uma redução significativa do resíduo gerado.

6.2.2 Análise de solo por meio de espectros MIR

Os espectros de infravermelho médio (MIR) foram obtidos após a secagem, moagem e peneiramento das amostras. Em seguida, os espectros MIR foram correlacionados com os teores de COS, argila e areia. A Figura 39 apresenta o resumo das etapas utilizadas.

Os espectros de cada amostra foram organizados em uma única matriz. A Figura 40 apresenta os espectros de todas as amostras.

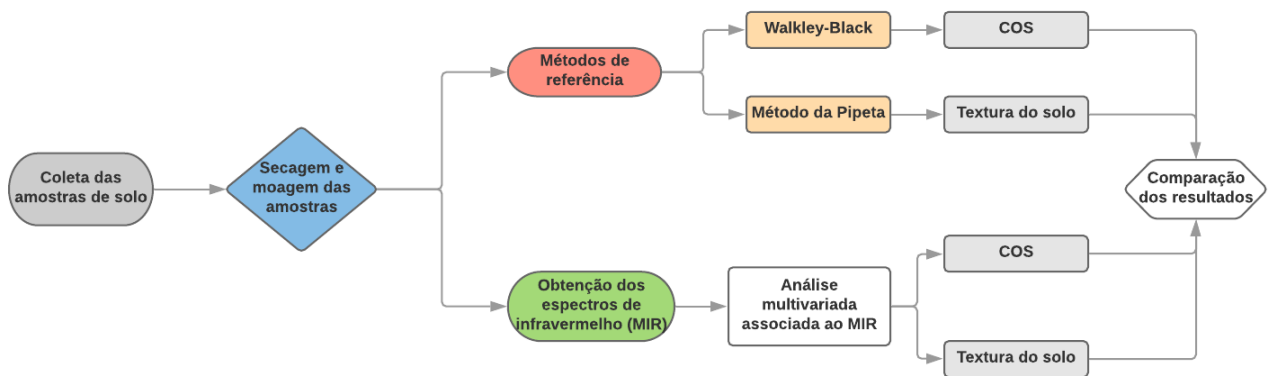


Figura 39 – Etapas utilizadas na quantificação do COS e da textura do solo pelo método de referência e por MIR.

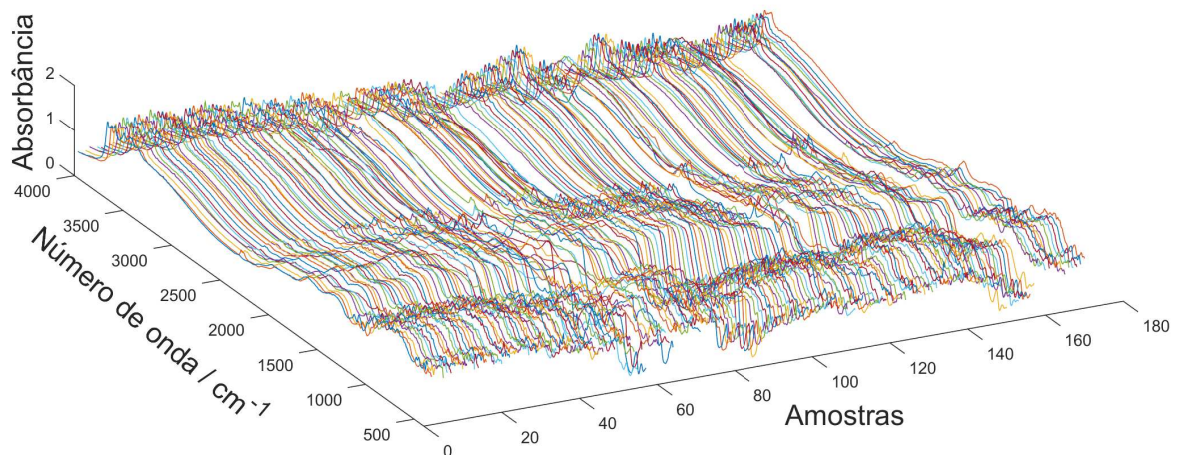


Figura 40 – Matriz de dados referentes aos espectros de infravermelho médio de todas as amostras de solo.

A matriz contendo os espectros foi centrada na média como etapa de pré-processamento. Em seguida, essa matriz contendo os dados espectrais das amostras juntamente com os vetores y_s , referente aos teores de COS, areia e argila, foram subdivididas entre calibração e validação utilizando a ferramenta SPXY. Assim, um total de 124 amostras foram designadas para calibração e 53 para validação. Por último, utilizou-se os métodos de calibração multivariados PLS, SPA-MLR e LS-SVM para correlacionar a matriz contendo os dados espectrais com os teores de COS, argila e areia. O número ideal de LVs para o método PLS foi selecionado de acordo com os menores valores de RMSEP. A Figura 41 apresenta todos os passos utilizados para estimação dos parâmetros de solo: COS, argila e areia por meio de MIR.

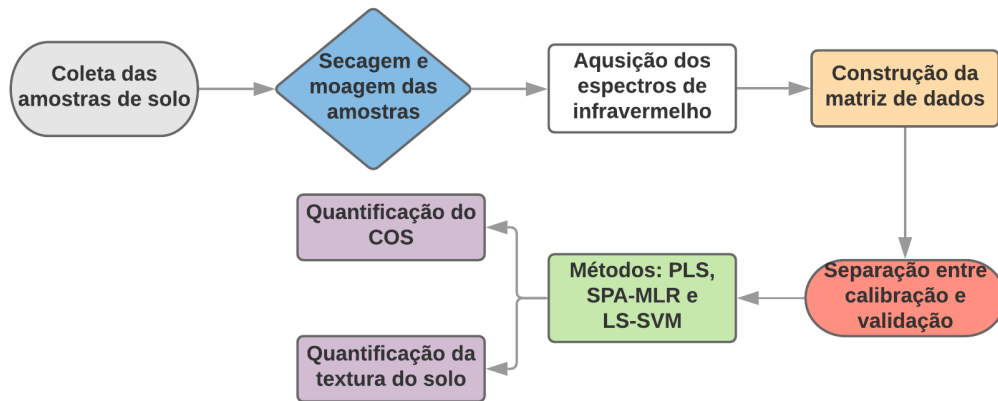


Figura 41 – Etapas utilizadas para estimação do COS e da textura em amostras de solo por MIR.

6.2.2.1 Quantificação do COS

A Tabela 24 apresenta os resultados obtidos para os métodos PLS, SPA-MLR e LS-SVM na predição dos teores de COS por meio de espectros de infravermelho médio (MIR).

Tabela 24 – Parâmetros estatísticos da calibração e validação na quantificação de COS utilizando os métodos SPA-MLR, PLS e LS-SVM.

	SPA-MLR	PLSR	LSSVMR
R^2_{cal}	0,8540	0,9155	0,9437
RMSEC	4,0	3,0	2,5
R^2_{val}	0,8060	0,8722	0,9057
RMSEP	4,6	3,8	3,3
Número de Variáveis	15	-	-
LV	-	12	-
γ	-	-	994,347
σ^2	-	-	8897,9
RPD	2,2	2,7	3,2

As Figuras 42, 43 e 44 apresentam os gráficos das concentrações de COS determinados experimentalmente pelo método de referência contra as concentrações de COS previstas, respectivamente, pelos modelos PLS, SPA-MLR e LSSVM obtidos utilizando espectros de MIR.

De acordo com a Tabela 24, a ferramenta LSSVM apresentou os melhores resultados na estimação do COS em comparação com os métodos PLS e SPA-MLR. Entretanto, de acordo com o teste F à 95% de confiança, não houve diferença entre os valores de RMSEP

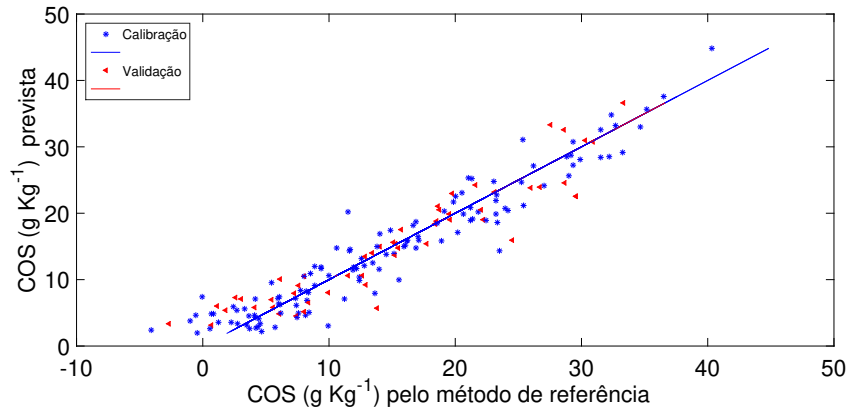


Figura 42 – Concentrações de COS de referência versus previstas pelo modelo PLS. Calibração (*) e validação (▲).

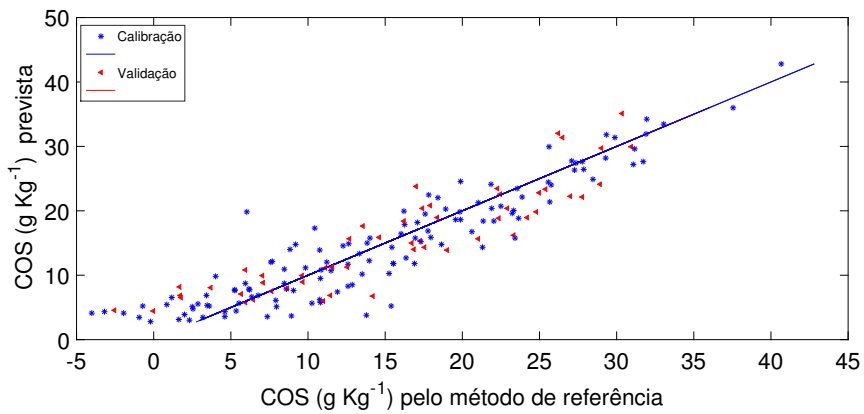


Figura 43 – Concentrações de COS de referência versus previstas pelo modelo SPA-MLR. Calibração (*) e validação (▲).

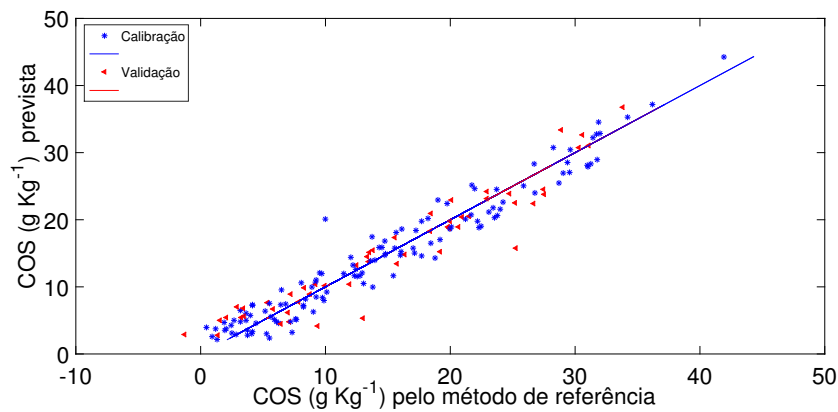


Figura 44 – Concentrações de COS de referência versus previstas pelo modelo LSSVM. Calibração (*) e validação (▲).

calculados para os métodos PLS e LSSVM. Em contrapartida, o valor de RMSEP calculado para o método SPA-MLR se difere estatisticamente dos valores de RMSEP determinados para os métodos LSSVM e PLS. O desempenho dos três métodos de calibração utilizados também foram comparados por meio do teste de randomização. Os resultados desse teste apontaram que os métodos LSSVM e PLS apresentam a mesma acurácia na estimação de COS utilizando os dados de MIR.

A determinação de carbono orgânico do solo por meio de MIR já foi reportada na literatura (MCCARTY *et al.*, 2002; MADARI *et al.*, 2005; MCCARTY; REEVES, 2006; LUDWIG *et al.*, 2008; KNOX *et al.*, 2015; CAMPBELL *et al.*, 2018). Vohland *et al.* (2014) determinaram COS de 60 amostras de solo por meio de MIR. A melhor correlação obteve um r^2 de 0,78 e um valor de RPD de 2,12. De forma semelhante, Bashagaluke *et al.* (2015) obtiveram um r^2 de 0,72 na correlação de espectros de MIR com os teores de COS utilizando o método PLS.

Utilizando espectros de MIR de 118 amostras de solo coletadas na Austrália, Viscarra-Rossel *et al.* (2006b) obtiveram r^2 de 0,72 na estimação de COS. Reeves e Smith (2009) obtiveram uma correlação com $r^2 = 0,58$ e RPD = 1,5 na determinação de COS utilizando espectros de MIR e o método PLS. De forma análoga, Campbell *et al.* (2018) obtiveram correlação com $r^2 = 0,65$ e RPD = 1,29. Neste estudo, a melhor correlação entre espectros de MIR e COS obteve $r^2 > 0,90$ para calibração e validação, assim como RPD superior a 3.

6.2.2.2 Quantificação da textura do solo

De forma similar a quantificação do COS, realizou-se a determinação das frações texturais dos solo, argila e areia por meio de espectros de infravermelho médio (MIR).

6.2.2.2.1 Quantificação dos teores de argila

A Tabela 25 apresenta os resultados obtidos na quantificação dos teores de argila por meio de espectros de infravermelho para os métodos PLS, SPA-MLR e LS-SVM.

As Figuras 45, 46 e 47 apresentam os gráficos dos teores de argila determinados experimentalmente pelo método de referência contra os teores de argila previstas, res-

Tabela 25 – Parâmetros estatísticos da calibração e validação na quantificação dos teores de argila utilizando os métodos SPA-MLR, PLS e LS-SVM.

	SPA-MLR	PLSR	LSSVMR
R^2_{cal}	0,9028	0,9223	0,9846
RMSEC	6,7	6,0	2,7
R^2_{val}	0,9032	0,9099	0,9634
RMSEP	7,1	6,6	4,4
Número de Variáveis	15	-	-
LV	-	11	-
γ	-	-	37,811
σ^2	-	-	533,940
RPD	3,1	3,3	5,1

pectivamente, pelos modelos PLS, SPA-MLR e LSSVM obtidos utilizando espectros de MIR.

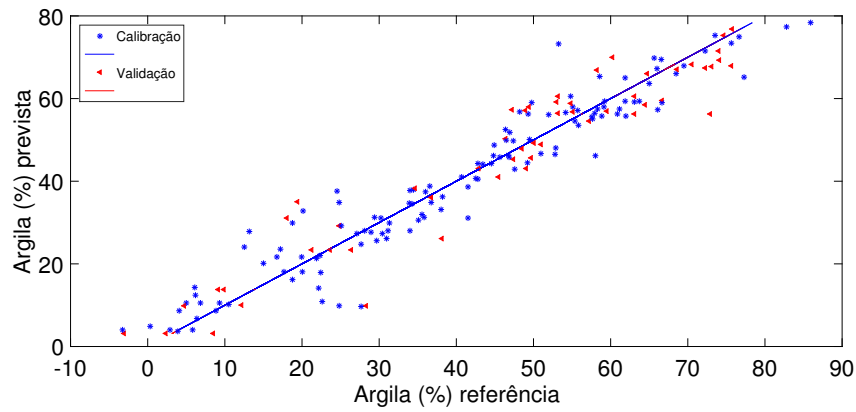


Figura 45 – Teores de argila de referência versus previstas pelo modelo PLS. Calibração (*) e validação (▲).

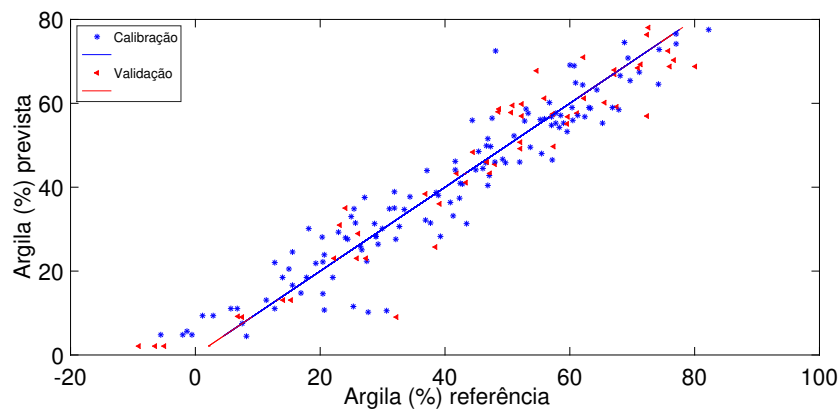


Figura 46 – Teores de argila de referência versus previstas pelo modelo SPA-MLR. Calibração (*) e validação (▲).

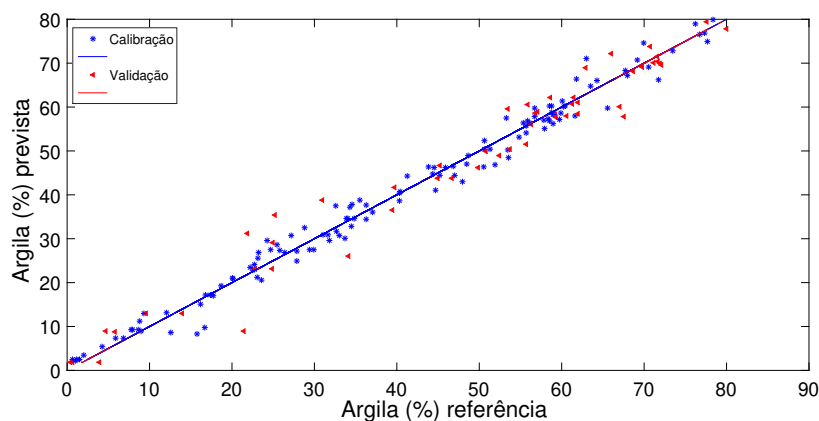


Figura 47 – Teores de argila de referência versus previstas pelo modelo LSSVM. Calibração (*) e validação (▲).

De acordo com a Tabela 25, o método LSSVM se destacou na estimação dos teores de argila em comparação com os métodos PLS e SPA-MLR. De acordo com o teste F à 95% de confiança, o valor de RMSEP calculado para o método LSSVM é diferente estatisticamente dos valores de RMSEP determinados para os métodos PLS e SPA-MLR. Além disso, não houve diferença estatística entre os valores de RMSEP calculados para os métodos SPA-MLR e PLS. O teste de randomização também foi utilizado para comparar o desempenho dos três métodos. Segundo esse teste, a ferramenta LSSVM apresenta maior acurácia que os métodos PLS e SPA-MLR na determinação de argila por meio de espectros de MIR. De acordo com as Figuras 45, 46 e 47, é notável a superioridade do método LS-SVM na determinação de argila em relação aos métodos SPA-MLR e PLS, pois para essa ferramenta há uma maior proximidade dos pontos a linha de correlação.

A determinação de argila por meio de MIR já foi reportada na literatura (VISCARRA-ROSSEL et al., 2006b; JANIK; SKJEMSTAD, 1995; JANIK; MERRY; SKJEMSTAD, 1998). (JANIK; SKJEMSTAD, 1995) correlacionaram espectros de MIR com os teores de argila utilizando o método PLS. Uma correlação com $r^2 = 0,87$ foi obtida. De forma similar, (JANIK; MERRY; SKJEMSTAD, 1998) obtiveram uma correlação com $r^2 = 0,79$. (VISCARRA-ROSSEL et al., 2006b) encontraram $r^2 = 0,82$ na correlação entre espectros de MIR e os teores de argila. De forma análoga, (CAMPBELL et al., 2018) obtiveram $r^2 = 0,69$. Neste trabalho, as melhores correlações entre os espectros de MIR e os teores de argilas obtiveram $r^2 > 0,96$ para calibração e validação.

6.2.2.2.2 Quantificação dos teores de areia

A Tabela 26 apresenta os resultados obtidos na quantificação dos teores de areia por meio de espectros de infravermelho médio (MIR) para os métodos PLS, SPA-MLR e LS-SVM.

Tabela 26 – Parâmetros estatísticos da calibração e validação na quantificação dos teores de areia utilizando os métodos SPA-MLR, PLS e LS-SVM.

	SPA-MLR	PLSR	LSSVMR
R^2_{cal}	0,8373	0,9076	0,9999
RMSEC	10,4	7,9	0,1
R^2_{val}	0,8882	0,9002	0,9767
RMSEP	8,5	8,3	4,2
Número de Variáveis	11	-	-
LV	-	12	-
γ	-	-	3979,5
σ^2	-	-	264,325
RPD	3,0	3,1	6,5

As Figuras 48, 49 e 50 apresentam os gráficos dos teores de areia determinados experimentalmente pelo método de referência contra os teores de areia previstas, respectivamente, pelos modelos PLS, SPA-MLR e LSSVM obtidos utilizando espectros de MIR.

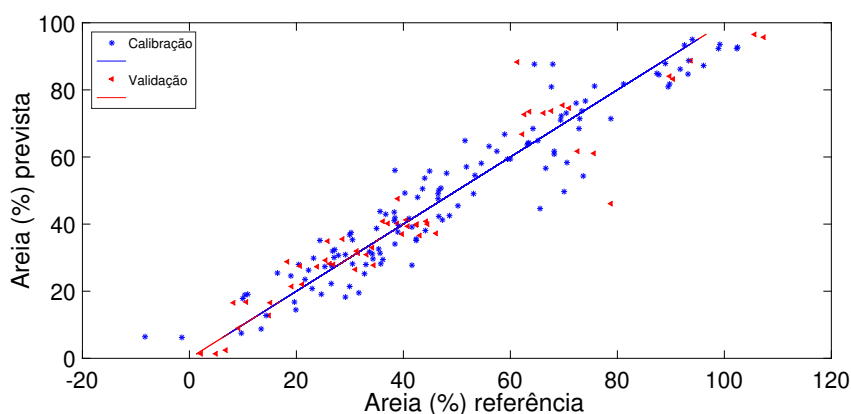


Figura 48 – Teores de areia de referência versus previstas pelo modelo PLS. Calibração (*) e validação (▲).

Segundo a Tabela 26, dentre os três métodos utilizados, a ferramenta LS-SVM apresentou os melhores resultados na estimação dos teores de areia. Os valores de RMSEP calculados para os três métodos foram comparados por meio do teste F à um nível de 95% de confiança. De acordo com esse teste, não houve diferença estatística entre os valores de

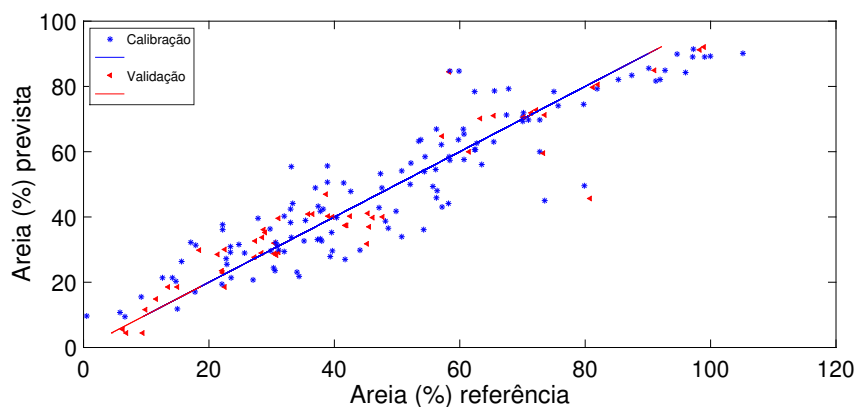


Figura 49 – Teores de areia de referência versus previstas pelo modelo SPA-MLR. Calibração (*) e validação (▲).

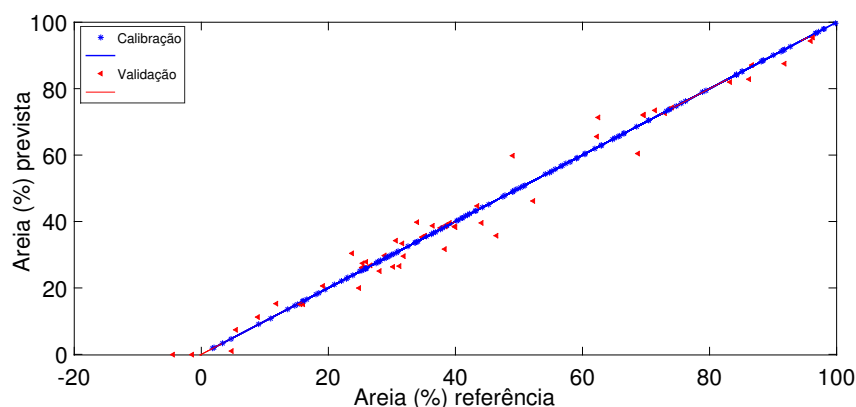


Figura 50 – Teores de areia de referência versus previstas pelo modelo LSSVM. Calibração (*) e validação (▲).

RMSEP determinados para os métodos PLS e SPA-MLR. Por outro lado, houve diferença estatística entre o valor de RMSEP calculado para o método LS-SVM e os valores de RMSEP determinados para os métodos PLS e SPA-MLR. O desempenho dos três métodos na estimação dos teores de areia por meio de espectros de MIR também foram comparados por meio do teste de randomização. De acordo com esse teste, os métodos PLS e SPA-MLR apresentam mesma acurácia, entretanto, ambos apresentam acurácia inferior que o método LS-SVM. O desempenho excepcional do método LS-SVM também pode ser observado comparando as Figuras 48, 49 e 50. Nota-se que o gráfico de correlação obtido utilizando LS-SVM, os pontos estão localizados mais próximos a linha de correlação.

A correlação entre espectros de MIR e os teores de areia já foram observados em outros estudos (VISCARRA-ROSSEL et al., 2006b; JANIK; SKJEMSTAD, 1995; JANIK; MERRY; SKJEMSTAD, 1998; TERHOEVEN-URSELMANS et al., 2010; PIRIE; SINGH; ISLAM, 2005). Pirie, Singh e Islam (2005) correlacionaram os espectros de MIR com teores de

areia utilizando o método de Regressão por Componentes Principais (PCR - *Principal Component Regression*). Uma correlação com $r^2 = 0,62$ e $RPD = 1,6$ foi obtida. Utilizando um total de 529 amostras de solo, [McCarty e Reeves \(2006\)](#) quantificaram os teores de areia, utilizando espectros de MIR. Dois terços das amostras foram utilizadas para calibração e um terço para validação. A correlação obtida para os dados de validação obteve um $r^2 = 0,76$. De forma similar, [Terhoeven-Urselmans et al. \(2010\)](#) obtiveram correlação com $r^2 = 0,64$ e $RPD = 1,59$ para o conjunto de validação. Neste estudo, correlações com $r^2 > 0,97$ foram obtidos para calibração e validação na estimação de areia por meio de MIR.

A determinação de argila e areia por meio de espectros de MIR consiste em uma alternativa (a) não destrutiva, permitindo novas análises futuras, (b) rápida, pois em poucos minutos é possível obter os resultados da análise, (c) de baixo impacto ambiental, pois praticamente não se produz resíduos.

6.2.2.3 Classificação textural por meio de espectros de MIR

Após a classificação das 177 amostras de solo em grupos texturais de acordo com os teores de argila, areia e silte determinados via método de referência, Figura 17, as amostras foram subdivididas entre calibração e validação utilizando o algoritmo Kennard-Stone ([KENNARD; STONE, 1969](#)). Em seguida, utilizou-se a ferramenta LS-SVM com o intuito de classificar as amostras do conjunto de validação em grupos texturais. Uma taxa de acerto de 90,2% foi obtido na classificação textural utilizando espectros de MIR em comparação com a classificação determinada pelo método de referência: método da Pipeta. Diante disso, a classificação textural por meio de MIR é uma alternativa valiosa e com boa aplicabilidade.

6.2.2.4 Determinação dos óxidos de ferro, silício e alumínio em solo

Posteriormente a determinação dos teores de Fe_2O_3 , SiO_2 e Al_2O_3 das 52 amostras de solo do conjunto total, obtiveram-se espectros de infravermelho médio (MIR) de cada amostra seca e moída. Em seguida, construiu-se uma matriz de dados contendo os espectros das 52 amostras. Essa matriz de dados, assim como os vetores y_s , correspondentes aos teores de Fe_2O_3 , SiO_2 e Al_2O_3 , foram subdivididas entre calibração e validação por meio do método SPXY. Dessa forma, um total de 39 amostras foram designadas para calibração

e 13 para validação. Por último, utilizou-se os métodos de calibração multivariados: PLS, SPA-MLR e LS-SVM para correlacionar os dados espectrais com os teores de Fe_2O_3 , SiO_2 e Al_2O_3 . A quantidade de variáveis latentes utilizadas no método PLS foi determinada de acordo com os menores valores de RMSEP.

6.2.2.4.1 Determinação de Fe_2O_3 em amostras de solo

A Tabela 27 apresenta os resultados estatísticos obtidos na correlação dos dados de espectros MIR com os teores de Fe_2O_3 em amostras de solo utilizando os métodos: PLS, SPA-MLR e LS-SVM.

Tabela 27 – Resultados estatísticos obtidos na estimação de Fe_2O_3 em solo por meio dos métodos SPA-MLR, PLS e LS-SVM.

	Calibração					Validação			
	LV	Number of Variable	γ	σ^2	R^2_{cal}	RMSEC	R^2_{val}	RMSEP	RPD
SPA-MLR	-	11	-	-	0,9151	9,3	0,9139	10,0	2,9
PLS	7	-	-	-	0,8778	11,1	0,8041	12,9	2,2
LS-SVM	-	-	29,061	816,116	0,9782	4,9	0,9725	5,6	5,2

As Figuras 51, 52 e 53 apresentam os gráficos dos teores de Fe_2O_3 determinados pelo método de referência contra os teores de Fe_2O_3 previstos, respectivamente, para os modelos PLS, SPA-MLR e LS-SVM.

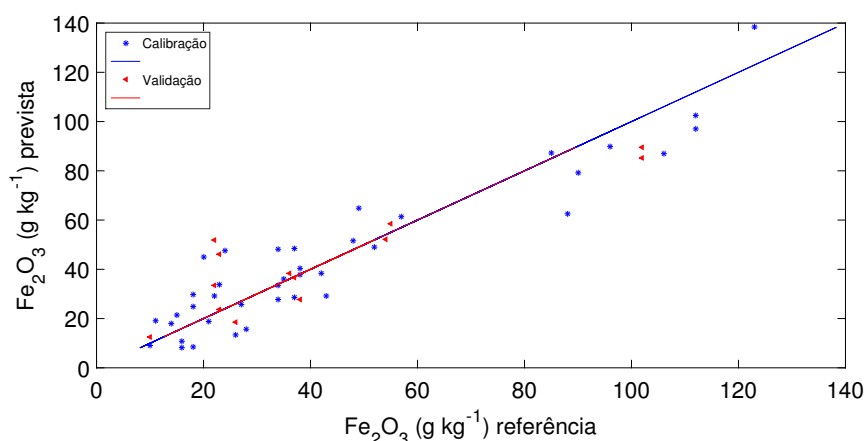


Figura 51 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo PLS. Calibração (*) e validação (\blacktriangle).

De acordo com a Tabela 27, os melhores resultados na estimação de Fe_2O_3 foram obtidos pela ferramenta LS-SVM. Esse método apresentou o maior valor de RPD, assim como a melhor correlação ($r^2 > 0,97$). Os valores de RMSEP calculados pelos três métodos foram comparados utilizando o teste F à um nível de 95% de confiança. De acordo com esse

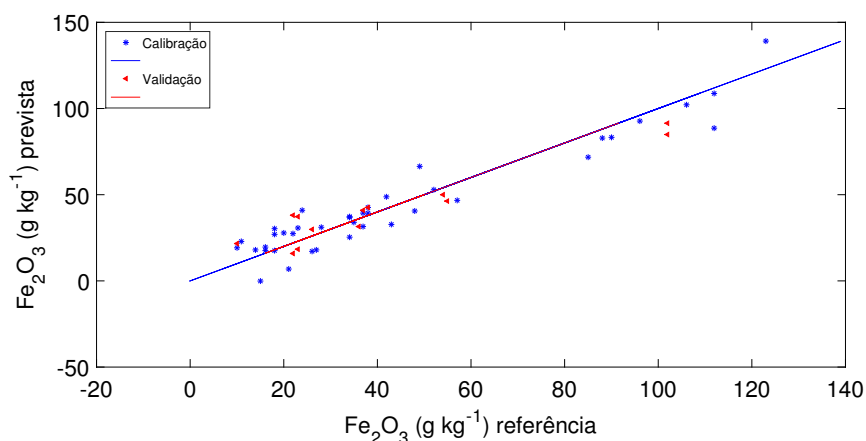


Figura 52 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).

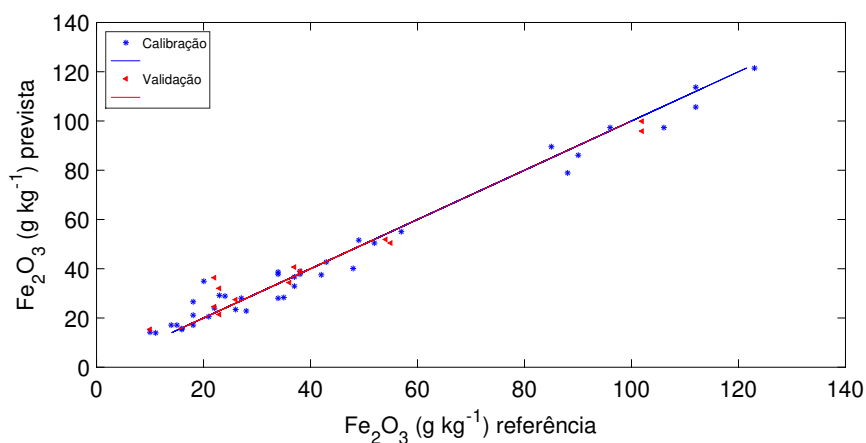


Figura 53 – Teores de Fe_2O_3 de referência versus os teores de Fe_2O_3 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).

teste, os valores de RMSEP calculados para os métodos SPA-MLR e PLS são superiores estatisticamente ao valor de RMSEP obtido pelo método LS-SVM.

A estimação de ferro por meio de espectros de infravermelho já foi apresentada em outros estudos (COZZOLINO; MORON, 2003; RAMAROSON et al., 2018). Cozzolino e Moron (2003) quantificaram os teores de ferro e outros metais em solo utilizando a espectroscopia de reflectância no infravermelho próximo (NIRS). A correlação entre os espectros e os teores de ferro apresentaram correlação de $r^2 = 0,92$. Neste estudo, uma correlação mais alta foi obtida para os dados de calibração e validação ($r^2 > 0,97$). Nanni e Demattê (2006) determinaram Fe_2O_3 em solo utilizando reflectância espectral. A correlação obtida apresentou $r^2 = 0,950$. Essa correlação é inferior a obtida neste trabalho ($r^2 > 0,97$). Ramaroson et al. (2018) quantificaram óxidos de ferro em solo por meio de NIRS. A correlação obtida, $r^2 = 0,80$, é menos significativa do que a obtida neste trabalho.

6.2.2.4.2 Determinação de SiO_2 em amostras de solo

A Tabela 28 apresenta os resultados estatísticos obtidos na correlação dos dados de espectros MIR com os teores de SiO_2 em amostras de solo utilizando os métodos: PLS, SPA-MLR e LS-SVM.

Tabela 28 – Resultados estatísticos obtidos na estimação de SiO_2 em solo por meio dos métodos SPA-MLR, PLS e LS-SVM.

	Calibração					Validação			
	LV	Número de Variáveis	γ	σ^2	R^2_{cal}	RMSEC	R^2_{val}	RMSEP	RPD
SPA-MLR	-	3	-	-	0,9208	23,7	0,9812	17,6	5,8
PLS	10	-	-	-	0,9713	14,3	0,9765	15,2	6,5
LS-SVM	-	-	54,822	1218,464	0,9882	9,3	0,9907	9,9	10,3

As Figuras 54, 55 e 56 apresentam os gráficos dos teores de SiO_2 determinados pelo método de referência contra os teores de SiO_2 previstos, respectivamente, para os modelos PLS, SPA-MLR e LS-SVM.

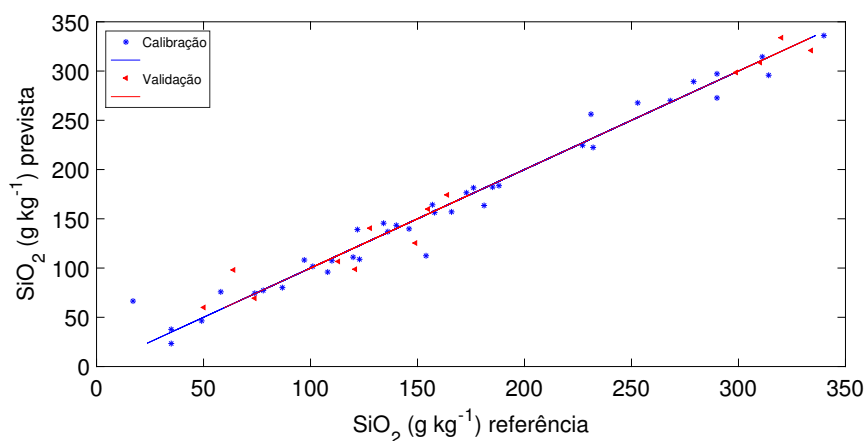


Figura 54 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo PLS. Calibração (*) e validação (▲).

De acordo com a Tabela 28, os melhores resultados na estimação de SiO_2 em amostras de solo foram obtidos pelo método LS-SVM. Os valores de RMSEP determinados para os métodos PLS e SPA-MLR são estatisticamente superiores ao valor de RMSEP obtido para o método LS-SVM de acordo com o teste F à um nível de 95% de confiança.

A correlação entre espectros de infravermelho e os teores de SiO_2 já foi reportada na literatura (NANNI; DEMATTÊ, 2006; DEMATTÊ; TERRA, 2014). Utilizando reflectância espectral, Nanni e Demattê (2006) quantificaram os teores de SiO_2 em solo. A correlação obtida apresentou $r^2 = 0,926$. Neste trabalho, uma melhor correlação, $r^2 > 0,98$, foi observada. Demattê e Terra (2014) determinaram SiO_2 em solo utilizando a análise por

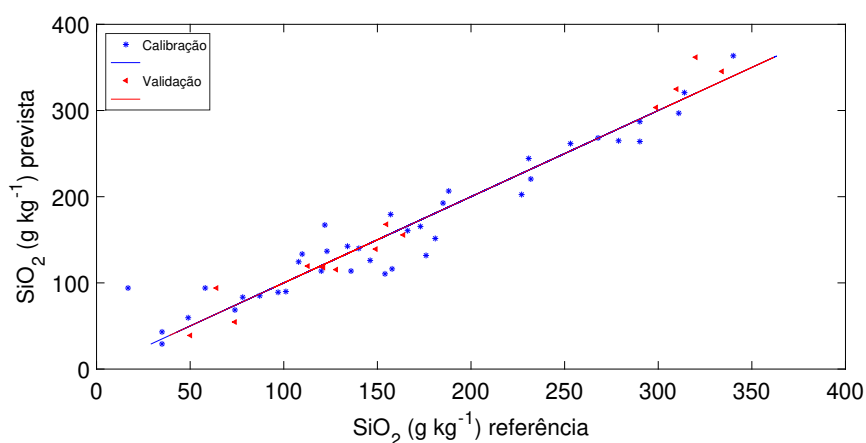


Figura 55 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo SPA-MLR. Calibração (*) e validação (▲).

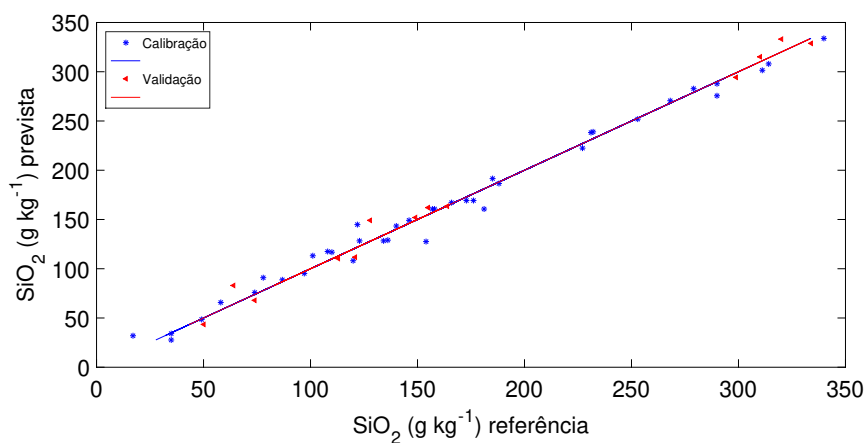


Figura 56 – Teores de SiO_2 de referência versus os teores de SiO_2 previstos pelo modelo LS-SVM. Calibração (*) e validação (▲).

espectroscopia de reflectância. A correlação encontrada, $r = 0,61$, foi inferior a verificada neste estudo ($r^2 > 0,98$). [Ramaroson et al. \(2018\)](#) quantificaram SiO_2 por meio de NIRS. As correlações obtidas, respectivamente, para calibração e validação foram 0,64 e 0,37. Neste estudo, correlações mais significativas foram encontradas tanto para calibração ($r^2 = 0,9882$) assim como para a validação ($r^2 = 0,9907$).

6.2.2.4.3 Determinação de Al_2O_3 em amostras de solo

A Tabela 29 apresenta os resultados estatísticos obtidos na correlação dos dados de espectros MIR com os teores de Al_2O_3 em amostras de solo utilizando os métodos: PLS, SPA-MLR e LS-SVM.

Tabela 29 – Resultados estatísticos obtidos na estimação de Al_2O_3 em solo por meio dos métodos SPA-MLR, PLS e LS-SVM.

	Calibração					Validação			
	LV	Número de Variáveis	γ	σ^2	R^2_{cal}	RMSEC	R^2_{val}	RMSEP	RPD
SPA-MLR	-	5	-	-	0,8494	27,6	0,9516	22,5	3,3
PLS	8	-	-	-	0,9412	17,3	0,9242	20,7	3,5
LS-SVM	-	-	60,097	1288,517	0,9753	11,4	0,9750	11,6	6,3

As Figuras 57, 58 e 59 apresentam os gráficos dos teores de Al_2O_3 determinados pelo método de referência contra os teores de Al_2O_3 previstos, respectivamente, para os modelos PLS, SPA-MLR e LS-SVM.

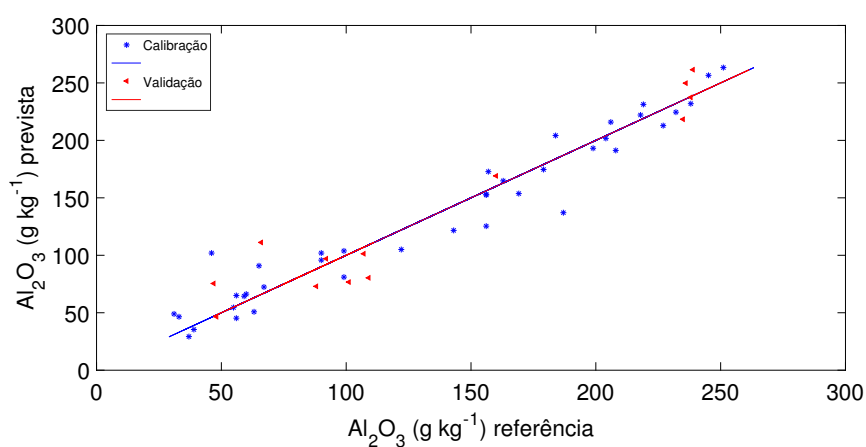


Figura 57 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo PLS. Calibração (*) e validação (\blacktriangle).

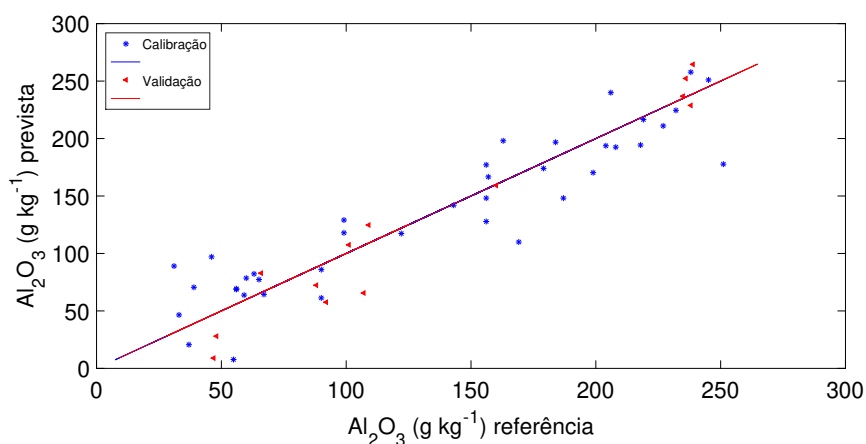


Figura 58 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo SPA-MLR. Calibração (*) e validação (\blacktriangle).

De acordo com a Tabela 29, o método LS-SVM apresentou os melhores resultados na estimação de Al_2O_3 em solo em comparação com os métodos SPA-MLR e PLS. Segundo o teste F à um nível de 95% de confiança, o valor de RMSEP obtido para o método

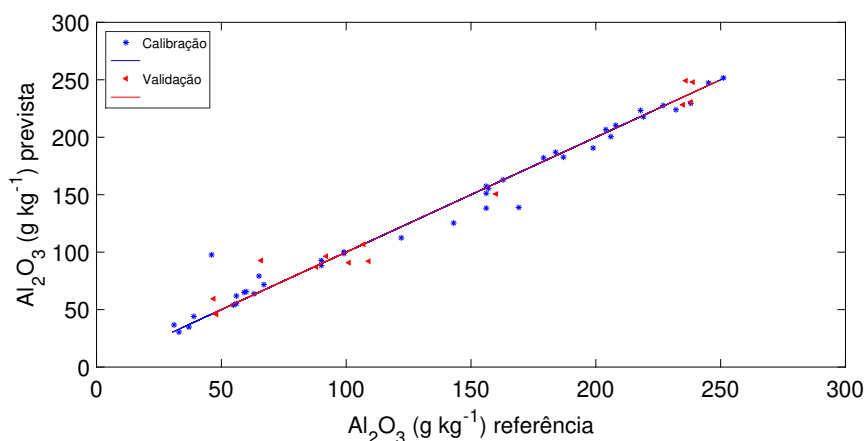


Figura 59 – Teores de Al_2O_3 de referência versus os teores de Al_2O_3 previstos pelo modelo LS-SVM. Calibração (*) e validação (\blacktriangle).

LS-SVM é inferior estatisticamente aos valores de RMSEP determinados para os métodos SPA-MLR e PLS.

Espectros de infravermelho já foram utilizados para quantificar os teores de alumínio em solo (RAMAROSON et al., 2018; DEMATTÊ; TERRA, 2014). Ramaroson et al. (2018) quantificaram os teores de Al_2O_3 em solo utilizando NIRS. As correlações para calibração e validação apresentaram r^2 , respectivamente, de 0,81 and 0,72. Além disso, o valor de RPD determinado para o conjunto de validação foi de 1,9. Neste estudo, observaram-se correlações mais significativas para calibração (0,9753) e validação (0,9750), assim como um valor de RPD superior (6,3). Os resultados apontados neste estudo também são melhores do que os verificados por Demattê e Terra (2014), que encontraram $r^2 < 0,6$ na determinação de Al_2O_3 em solo por meio de reflectância espectral.

6.2.3 Análise de solo por meio de MIA e MIR em relação aos métodos de referência

Este estudo propõe a quantificação dos parâmetros de fertilidade: carbono orgânico e textura do solo, assim como dos óxidos de ferro, silício e alumínio extraídos do solo por meio de imagens digitais e espectros de infravermelho médio. De modo a verificar a viabilidade das metodologias propostas, realizou-se uma comparação com os métodos de referência empregados em laboratórios de análise de solo, quanto aos critérios: operacionalidade, segurança, impacto ambiental e custo de implementação.

6.2.3.1 Operacionalidade das metodologias de análise de solo

A quantificação de COS por meio do método Walkley-Black envolve várias etapas sendo algumas laboriosas. Primeiramente, as amostras de solo secas e moídas são pesadas em erlenmeyer de vidro. Em seguida, adiciona-se solução de dicromato de potássio e ácido sulfúrico concentrado. Após a agitação e resfriamento, adiciona-se água e em seguida, mantém-se as amostras em repouso por dois dias. Após esse período, coleta-se o sobrenadante e realiza-se a leitura em espectrofotômetro. Esse procedimento, de difícil operacionalidade, leva mais de dois dias para a emissão do resultado de carbono orgânico do solo.

Já a determinação da textura do solo por meio do método da Pipeta consiste em uma metodologia bastante laboriosa. Após a pesagem da amostra de solo, adiciona-se os reagentes e realiza-se a agitação. Em seguida, realiza-se a peneiração da areia seguido da disposição do silte e argila em uma proveta de vidro. Após a agitação, o sistema é mantido em repouso em torno de 4 h. Em seguida, após a coleta de um volume específico do sobrenadante em um béquer, realiza-se a secagem em estufa durante 48 h. Por último, os béqueres contendo areia e argila são novamente pesados para determinação de seus teores no solo. Diante disso, o procedimento laboratorial do método da Pipeta consiste em múltiplas etapas e leva em média 3 dias para emissão dos resultados.

A quantificação dos óxidos de ferro, silício e alumínio extraídos do solo em laboratório inicia-se com a extração desses óxidos utilizando ácido sulfúrico em refluxo. Após a filtração, os óxidos de ferro e alumínio são quantificados em um equipamento de absorção atômica. Os teores de óxidos de silício são determinados utilizando o resíduo sólido obtido após a filtração. Após a extração com hidróxido de sódio, realiza-se a determinação dos óxidos de silício por meio da espectroscopia de absorção atômica. Essa metodologia envolve várias etapas e leva em média um dia para a geração dos resultados.

Como mencionado, os métodos de referência são laboriosos e de difícil operacionalidade e tardam em média de um a três dias para emissão dos resultados das análises. Já os métodos baseados em imagens digitais e espectros MIR envolvem poucas etapas e são de baixa laboriosidade e ainda de fácil operacionalidade. Dessa forma, é possível emitir os resultados das análises de solo em poucos minutos.

6.2.3.2 Segurança dos procedimentos de análise de solo

A determinação de COS por meio do método de referência Walkley-Black emprega em seu procedimento reagentes químicos de alta periculosidade, como o dicromato de potássio e o ácido sulfúrico concentrado. Ambos reagentes são utilizados em conjunto na quantificação do COS. Dessa forma, uma solução ácida de Cr^{6+} é gerada. O Cr^{6+} apresenta elevada toxicidade para o ser humano, podendo provocar inflamações, irritações e ainda é considerado como um possível carcinogênio (COHEN et al., 1993). Além disso, a utilização de ácido sulfúrico nesse procedimento ainda pode ocasionar queimaduras em contato com a pele assim como acidentes mais graves.

Já a quantificação dos teores de argila, silte e areia realizada por meio do método de referência: método da Pipeta, utiliza reagentes oxidantes e corrosivos, como o hidróxido de sódio e o peróxido de hidrogênio. Esses reagentes podem provocar inflamações, irritações e até mesmo queimaduras graves.

A análise dos óxidos de ferro, silício e alumínio em laboratório é realizada utilizando a extração sulfúrica, que consiste no emprego de ácido sulfúrico concentrado. Esse procedimento é de alto risco, pois o ácido concentrado é aquecido até fervura e mantido em refluxo por 30 minutos. Além do ácido sulfúrico, emprega-se o hidróxido de sódio que como já mencionado também pode causar acidentes.

Diante do exposto, durante a execução desses métodos de análise de solo, o laboratorista está exposto a um alto risco químico e pode ser vítima de um acidente de trabalho. Por outro lado, os métodos propostos para a análise dos parâmetros de fertilidade baseados em imagens digitais e espectros de MIR dispensam a utilização de qualquer reagente químico. Dessa forma, o ambiente de trabalho torna-se muito mais seguro para o analista e seus colegas de trabalho.

6.2.3.3 Impacto ambiental dos procedimentos de análise de solo

Como mencionado no item 6.2.3.2, a determinação de COS em laboratório por meio do método Walkley-Black utiliza principalmente: dicromato de potássio e ácido sulfúrico. Dessa forma, após a conclusão da análise, um resíduo ácido com a presença de Cr^{6+} é gerado. Em média, o volume de resíduo gerado por ensaio é de 28 mL, dessa forma, em uma análise de 100 amostras de solo, um volume de quase 3 L de resíduo é produzido.

Esse resíduo é altamente danoso ao meio ambiente, dessa forma, o mesmo deve ser tratado antes de seu descarte.

Posteriormente a análise textural do solo por meio do método da Pipeta, um volume de 1 L de resíduo básico é gerado em média por amostra. Anteriormente ao descarte, esse resíduo deve ser neutralizado de modo a minimizar o dano ambiental. De modo similar, a determinação dos óxidos de ferro, silício e alumínio extraídos do solo em laboratório gera em média 20 mL de resíduo ácido e 5 mL de resíduo básico por amostra. Esses resíduos também carecem de tratamento e por isso devem ser neutralizados antes do descarte.

Diante do mencionado, nota-se que os métodos de referência executados em laboratório geram um passivo ambiental bastante significativo, de acordo com o volume de amostras, que não pode ser descartado sem tratamento. Já as metodologias de análise de solo por meio de imagens não utilizam reagentes químicos. Dessa forma, o único resíduo gerado é a própria amostra de solo. Nesse sentido, a utilização dos métodos propostos contribuirá com a minimização do passivo ambiental e conseqüentemente da preservação ambiental.

6.2.3.4 Custo de implementação dos métodos de análise de solo

A implementação do método de referência: Walkley-Black requer uma série de vidrarias como béqueres, balão volumétrico, pipetas, erlenmeyers, tubos, assim como de equipamentos como: mesa agitadora, balança analítica, repipetador automático, bureta eletrônica e espectrofotômetro de absorção molecular na região do visível. Dessa forma, há um custo de implementação desse método, ao redor de R\$ 25 a 75 mil reais.

Já a implementação do método de referência: método da Pipeta, utilizado na determinação da textura do solo necessita de vidrarias como: provetas de 1 L, bastão de vidro, pipetas, béqueres, erlenmeyers, assim como de peneiras, balança analítica, processador ultrassônico e mesa agitadora. Dessa maneira, estima-se um custo de implementação desse método na faixa de R\$ 10 a 35 mil reais.

De forma similar, a determinação dos óxidos extraídos do solo por meio do método de referência demanda vidrarias como béqueres, erlenmeyers, funil, pipeta, coluna de refluxo, balão de fundo chato, assim como dos equipamentos: balança analítica, chapa de aquecimento, repipetador automático, espectrofotômetro de absorção atômica. Dessa

forma, estima-se um custo de implementação desse método na faixa de R\$ 125 a 200 mil reais.

Diante disso, nota que os métodos de referência apresentam custo de implementação elevados especialmente devido a a aquisição de equipamentos. Já os métodos propostos baseados em análise de imagens apresentam um baixo custo de implementação, pois só requerem um *scanner*, um equipamento acessível e de custo reduzido. Dessa forma, estima-se o custo de implementação desses métodos na faixa de R\$ 1 a 5 mil reais. Por outro lado, as metodologias propostas para análise do solo por meio de espectros de MIR apresentam um custo de implementação mais elevado devido a aquisição de um espectrofotômetro de infravermelho. Assim, estima-se um custo de implementação desses métodos na faixa de R\$ 30 a 100 mil reais.

6.2.4 Análise de solo por meio de MIA e MIR: comparação de resultados

Neste estudo, espectros de infravermelho médio (MIR) e imagens digitais foram utilizados para predição do carbono orgânico do solo, das partículas físicas do solo: areia, argila e silte, assim como dos teores de Fe_2O_3 , SiO_2 e Al_2O_3 . A análise de fertilidade do solo por meio dessas duas metodologias foram comparados quanto aos parâmetros: impacto ambiental, viabilidade, custo, rapidez de análise e operacionalidade.

Primeiramente, os melhores resultados obtidos pelas metodologias baseadas em imagens digitais e de espectros MIR foram comparados. Como mencionado, os melhores resultados obtidos na estimação dos parâmetros de fertilidade do solo foram obtidos utilizando a ferramenta LS-SVM. A Tabela 30 apresenta os melhores resultados obtidos na estimação dos teores de COS, areia, argila, Fe_2O_3 , SiO_2 e Al_2O_3 pelos dois métodos propostos.

Tabela 30 – Comparação de resultados na estimação de COS e da textura do solo por meio de imagens digitais e MIR.

Parâmetro	MIA					MIR				
	R_{cal}^2	RMSEC	R_{val}^2	RMSEP	RPD	R_{cal}^2	RMSEC	R_{val}^2	RMSEP	RPD
COS	0,9598	2,3	0,9376	2,5	3,7	0,9437	2,5	0,9057	3,3	3,2
Argila	0,9829	3,1	0,9618	4,0	5,1	0,9846	2,7	0,9634	4,4	5,1
Areia	0,9736	4,5	0,9568	5,0	4,6	0,9999	0,1	0,9767	4,2	6,5
Fe_2O_3	0,9648	6,2	0,9785	6,2	5,2	0,9782	4,9	0,9725	5,6	5,2
SiO_2	0,9885	12,1	0,9516	23,4	3,7	0,9882	9,3	0,9907	9,9	10,3
Al_2O_3	0,9929	7,9	0,9460	13,1	4,7	0,9753	11,4	0,9750	11,6	6,3

De acordo com a Tabela 30, observa-se que de modo geral, os resultados obtidos utilizando espectros de MIR e imagens digitais são similares. Entretanto, de acordo com o teste F à 95% de confiança, o valor de RMSEP obtido na estimação de COS pelo método de análise imagens é estatisticamente inferior ao valor de RMSEP determinado para o método que utiliza espectros MIR. Por outro lado, os valores de RMSEPs determinados na predição dos teores de SiO_2 e Al_2O_3 por meio de espectros de MIR são estatisticamente inferiores aos valores de RMSEPs calculados para o método baseado em imagens digitais. Não se observou diferença estatística entre os valores de RMSEP obtidos para ambos métodos na estimação de areia, argila e dos teores de Fe_2O_3 .

As metodologias baseadas na análise de imagens e em espectros de MIR não utilizam reagentes químicos e por isso, além de possibilitarem um ambiente de trabalho mais seguro aos laboratoristas, não há geração de resíduo tóxico passível de tratamento. Dessa forma, ambos métodos podem ser considerados como de baixo impacto ambiental. Além disso, a amostra de solo após ser analisada por esses métodos, não sofre nenhuma modificação, isto é, tratam-se de metodologias não destrutíveis.

Os procedimentos experimentais dos métodos de referência para a quantificação da textura e do COS levam em média mais de dois dias para a emissão dos resultados. As determinações dos teores de Fe_2O_3 , SiO_2 e Al_2O_3 pelos métodos de referência em laboratório levam geralmente um dia para a obtenção dos resultados. Já determinação de COS, textura e das concentrações de Fe_2O_3 , SiO_2 e Al_2O_3 por esses métodos não destrutíveis pode ser realizada de forma simultânea e em apenas alguns minutos, garantindo um aumento considerável na produtividade de ensaio.

Como mencionado, a análise de fertilidade do solo por meio de espectros de MIR é vantajosa, entretanto a mesma apresenta alto custo de implementação devido a aquisição de um espectrofotômetro de infravermelho, um equipamento de valor elevado. Além disso, ainda há custos relevantes para instalação e manutenção desse equipamento, pois o mesmo requer um ambiente de temperatura e umidade controlado. Dessa forma, a determinação dos parâmetros de fertilidade avaliados neste estudo por meio de imagens digitais é mais vantajosa graças a seu baixo custo de implementação, pois somente um *scanner* comercial e bastante acessível é necessário. Além disso, a instalação do *scanner* é bastante simples e sua manutenção é de baixo custo.

Além do maior custo do espectrofotômetro, a utilização desse equipamento requer mão de obra mais especializada devido a complexidade de funcionamento desse dispositivo.

Por outro lado, a aquisição de imagens utilizando um *scanner* é simples e praticamente não necessita de treinamento técnico. Diante do mencionado, nota-se que ambas metodologias são aplicáveis a rotina laboratorial com resultados satisfatórios. Entretanto, a metodologia baseada em análise de imagens é favorável de acordo com o anteriormente descrito.

7 Conclusões

Este trabalho propõe métodos de baixo custo para quantificação do carbono orgânico e da textura do solo, assim como dos teores de Fe_2O_3 , SiO_2 e Al_2O_3 extraídos do solo por meio de imagens digitais e espectros de infravermelho médio (MIR). A correlação entre os dados de imagens e os espectros de MIR, e os parâmetros de fertilidade foram realizados utilizando os métodos de calibração multivariados: PLS, SPA-MLR e LS-SVM. Dentre os três métodos, o LS-SVM apresentou de modo geral, os melhores resultados com valores de RPD superiores a 3,0 na estimação do COS, maior que 5,0 na estimação de argila, superior a 4,5 na estimação da areia, e superior a 3,5 na estimação dos teores de Fe_2O_3 , SiO_2 e Al_2O_3 .

A ferramenta LS-SVM também foi utilizada na classificação das amostras do conjunto de validação por meio de imagens e espectros infravermelho. Obtiveram-se taxas de acerto de mais de 94% e 90%, respectivamente, para imagens e MIR na classificação dessas amostras em relação a classificação dada pelo método da Pipeta.

Os métodos propostos são simples e não utilizam reagentes químicos. Dessa forma, possibilitam um ambiente mais seguro e de menor risco aos analistas. Além disso, essas alternativas podem ser classificadas como de baixo impacto ambiental, pois não há geração de resíduo tóxico. Ademais, essas metodologias não são destrutivas, logo as amostras de solo podem ser reutilizadas. O custo de implementação desses métodos é inferior aos dos métodos de referência.

A metodologia baseada em imagens digitais apresentou resultados semelhantes ao método que utiliza espectros de infravermelho. Entretanto, a determinação dos parâmetros de fertilidade do solo por meio de imagens é preferível pois apresenta um menor custo de implementação devido a utilização de equipamentos baratos e acessíveis, como o *scanner*. Além disso, a utilização de um *scanner* exige um menor nível técnico do que o uso de um espectrofotômetro de infravermelho.

Como perspectiva a este trabalho, espera-se aumentar o banco de dados de solo de modo a contemplar amostras de todas as regiões do país. Além disso, espera-se desenvolver um *software* que permita a utilização dessa tecnologia por outros usuários.

Referências¹

- Agência Brasil. *PIB do agronegócio cresce mais que conjunto da economia em 2019 e 2020*. 2019. <http://agenciabrasil.ebc.com.br/economia/noticia/2019-11/pib-do-agronegocio-cresce-mais-que-conjunto-da-economia-em-2019-e-2020>. Online; acessado 21 Novembro 2019. Citado na página 24.
- AITKENHEAD, M.; COULL, M.; TOWERS, W.; HUDSON, G.; BLACK, H. Prediction of soil characteristics and colour using data from the national soils inventory of scotland. *Geoderma*, Elsevier, v. 200, p. 99–107, 2013. Citado na página 86.
- AL-QADIRI, H. M.; AL-HOLY, M. A.; LIN, M.; ALAMI, N. I.; CAVINATO, A. G.; RASCO, B. A. Rapid detection and identification of pseudomonas aeruginosa and escherichia coli as pure and mixed cultures in bottled drinking water using fourier transform infrared spectroscopy and multivariate analysis. *Journal of agricultural and food chemistry*, v. 54, n. 16, p. 5749–5754, 2006. Citado na página 28.
- ALMEIDA, V. E. de; COSTA, G. B. da; FERNANDES, D. D. de S.; DINIZ, P. H. G. D.; BRANDÃO, D.; MEDEIROS, A. C. D. de; VÉRAS, G. Using color histograms and spa-lda to classify bacteria. *Analytical and bioanalytical chemistry*, Springer, v. 406, n. 24, p. 5989–5995, 2014. Citado na página 27.
- ALVES, G. M. *Método fundamentado em Processamento Digital de Imagens para Contagem Automática de Unidades Formadoras de Colônias*. 125 p. Tese (Doutorado) — Universidade Federal de São Carlos, 2006. Citado na página 27.
- ARAÚJO, M. C. U.; SALDANHA, T. C. B.; GALVAO, R. K. H.; YONEYAMA, T.; CHAME, H. C.; VISANI, V. The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometrics and Intelligent Laboratory Systems*, Elsevier, v. 57, n. 2, p. 65–73, 2001. Citado na página 64.
- BAES, A.; BLOOM, P. Diffuse reflectance and transmission fourier transform infrared (drift) spectroscopy of humic and fulvic acids. *Soil Science Society of America Journal*, v. 53, n. 3, p. 695–700, 1989. Citado 2 vezes nas páginas 15 e 60.
- BALABIN, R. M.; LOMAKINA, E. I. Support vector machine regression (SVR/LS-SVM)—an alternative to neural networks (ANN) for analytical chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data. *Analyst*, Royal Society of Chemistry, v. 136, n. 8, p. 1703–1712, 2011. Citado na página 49.
- BALLARD, D.; SKLANSKY, J. Tumor detection in radiographs. *Computers and Biomedical Research*, Elsevier, v. 6, n. 4, p. 299–321, 1973. Citado na página 27.
- BARBEDO, J. G. A.; KOENIGKAN, L. V.; SANTOS, T. T. Identifying multiple plant diseases using digital image processing. *Biosystems engineering*, Elsevier, v. 147, p. 104–116, 2016. Citado na página 27.
- BARRÓN, V.; TORRENT, J. Use of the kubelka—munk theory to study the influence of iron oxides on soil colour. *European Journal of Soil Science*, Wiley Online Library, v. 37, n. 4, p. 499–510, 1986. Citado 2 vezes nas páginas 34 e 90.

¹ De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

BARROS, J. A.; OLIVEIRA, F. M. d.; SANTOS, G. d. O.; WISNIEWSKI, C.; LUCCAS, P. O. Digital image analysis for the colorimetric determination of aluminum, total iron, nitrite and soluble phosphorus in waters. *Analytical Letters*, Taylor & Francis, v. 50, n. 2, p. 414–430, 2017. Citado na página 28.

BASHAGALUKE, J.; NSHOBOLE, N.; FATAKI, D.; MOCHOGE, B.; MUGWE, J.; WALANGULULU, J. Application of infrared technique in soil properties' characterization in south kivu province of dr congo. *International Research Journals*, v. 6, p. 58–67, 2015. Citado na página 98.

BAUMGARDNER, M. F.; KRISTOF, S.; JOHANNSEN, C. J.; ZACHARY, A. Effects of organic matter on the multispectral properties of soils. *Proc. Ind. Acad. Sci.*, v. 79, n. 3939, p. 413–422, 1969. Citado na página 31.

BEEBE, K. R.; KOWALSKI, B. R. An introduction to multivariate calibration and analysis. *Analytical Chemistry*, ACS Publications, v. 59, n. 17, p. 1007A–1017A, 1987. Citado na página 43.

BENEDETTI, L. P. dos S.; SANTOS, V. B. dos; SILVA, T. A.; FILHO, E. B.; MARTINS, V. L.; FATIBELLO-FILHO, O. A digital image-based method employing a spot-test for quantification of ethanol in drinks. *Analytical Methods*, Royal Society of Chemistry, v. 7, n. 10, p. 4138–4144, 2015. Citado na página 28.

BEUCHLE, R.; GRECCHI, R.; SHIMABUKURO, Y.; SELIGER, R.; EVA, H.; SANO, E.; ACHARD, F. Land cover changes in the brazilian cerrado and caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach. *Applied Geography*, v. 58, p. 116–127, 2015. Citado na página 24.

BHABATOSH, C.; MAJUMDER, D. *Digital image processing and analysis*. New Delhi: PHI Learning Private, 2011. Citado 3 vezes nas páginas 36, 37 e 38.

BHARATI, M.; MACGREGOR, J. Multivariate image analysis for real-time process monitoring and control. *Industrial & Engineering Chemistry Research*, ACS Publications, v. 37, n. 12, p. 4715–4724, 1998. Citado 2 vezes nas páginas 27 e 28.

BLUME, H.-P.; BRÜMMER, G. W.; HORN, R.; KANDELER, E.; KÖGEL-KNABNER, I.; KRETZSCHMAR, R.; STAHR, K.; WILKE, B.-M.; THIELE-BRUHN, S.; WELP, G. *Scheffer/Schachtschabel Soil Science*. 1. ed. Berlin Heidelberg: Springer-Verlag, 2016. 618 p. Citado na página 85.

BORIN, A.; FERRÃO, M. F.; MELLO, C.; CORDI, L.; PATACA, L. C.; DURÁN, N.; POPPI, R. J. Quantification of lactobacillus in fermented milk by multivariate image analysis with least-squares support-vector machines. *Analytical and bioanalytical chemistry*, Springer, v. 387, n. 3, p. 1105–1112, 2007. Citado na página 28.

BOTELHO, B. G. *Desenvolvimento e validação de métodos para análise direta de alimentos usando ferramentas quimiométricas, espectroscopia no infravermelho e imagens digitais*. 119 p. Tese (Tese de doutorado) — Universidade Federal de Minas Gerais, 2014. Citado na página 28.

BOTELHO, B. G.; DANTAS, K. C.; SENA, M. M. Determination of allura red dye in hard candies by using digital images obtained with a mobile phone and n-pls. *Chemometrics and Intelligent Laboratory Systems*, v. 167, p. 44–49, 2017. Citado na página 28.

BOTVINICK, E. H.; SHAMES, D.; HUTCHINSON, J. C.; ROE, B. B.; FITZPATRICK, M. Noninvasive diagnosis of a false left ventricular aneurysm with radioisotope gated cardiac blood pool imaging: differentiation from true aneurysm. *The American journal of cardiology*, Elsevier, v. 37, n. 7, p. 1089–1093, 1976. Citado na página 27.

BOUYOUCOS, G. J. Estimation of the Colloidal Material in Soils. *Science*, v. 64, n. 1658, p. 362, 1926. Citado na página 32.

BRERETON, R. G. *Chemometrics: data analysis for the laboratory and chemical plant*. New York: John Wiley & Sons, 2003. Citado na página 43.

BREUL, P.; GOUVERS, R. In field soil characterization: approach based on texture image analysis. *Journal of geotechnical and geoenvironmental engineering*, v. 132, p. 102–107, 2006. Citado na página 29.

BRONICK, C. J.; LAL, R. Soil structure and management: a review. *Geoderma*, v. 124, n. 1, p. 3–22, 2005. Citado na página 31.

BRUNET, D.; BARTHÈS, B. G.; CHOTTE, J.-L.; FELLER, C. Determination of carbon and nitrogen contents in alfisols, oxisols and ultisols from africa and brazil using nirs analysis: effects of sample grinding and set heterogeneity. *Geoderma*, Elsevier, v. 139, n. 1-2, p. 106–117, 2007. Citado na página 29.

CAMARGO, M. N.; KLAMT, E.; KAUFFMAN, J. H. *Classificação de solos usada em levantamentos pedológicos no Brasil*. Campinas: Sociedade Brasileira de Ciência do Solo, 1987. v. 12. 24 p. Citado na página 34.

CAMPBELL, P. M. d. M.; FILHO, E. I. F.; FRANCELINO, M. R.; DEMATTÊ, J. A. M.; PEREIRA, M. G.; GUIMARÃES, C. C. B.; PINTO, L. A. d. S. R. Digital soil mapping of soil properties in the “mar de morros” environment using spectral data. *Revista Brasileira de Ciência do Solo*, SciELO Brasil, v. 42, 2018. Citado 2 vezes nas páginas 98 e 100.

CATELANI, T. A.; BITTAR, D. B.; PEZZA, L.; PEZZA, H. R. Determination of amino acids in gym supplements using digital images and paper platform coupled to diffuse reflectance spectroscopy and usb device. *Talanta*, Elsevier, v. 196, p. 523–529, 2019. Citado na página 28.

CÉCILLON, L.; BARTHÈS, B. G.; GOMEZ, C.; ERTLEN, D.; GÉNOT, V.; HEDDE, M.; STEVENS, A.; BRUN, J.-J. Assessment and monitoring of soil quality using near-infrared reflectance spectroscopy (nirs). *European Journal of Soil Science*, Wiley Online Library, v. 60, n. 5, p. 770–784, 2009. Citado na página 29.

CENTNER, V.; NOORD, O. D.; MASSART, D. Detection of nonlinearity in multivariate calibration. *Analytica chimica acta*, Elsevier, v. 376, n. 2, p. 153–168, 1998. Citado 4 vezes nas páginas 52, 53, 74 e 83.

CHEVALLIER, S.; BERTRAND, D.; KOHLER, A.; COURCOUX, P. Application of pls-da in multivariate image analysis. *Journal of Chemometrics*, Chichester, Sussex, England: Wiley, c1987-, v. 20, n. 5, p. 221, 2006. Citado na página 28.

CHOODUM, A.; DAEID, N. N. Rapid and semi-quantitative presumptive tests for opiate drugs. *Talanta*, Elsevier, v. 86, p. 284–292, 2011. Citado na página 28.

- CHOODUM, A.; KANATHARANA, P.; WONGNIRAMAİKUL, W.; DAEID, N. N. Using the iphone as a device for a rapid quantitative analysis of trinitrotoluene in soil. *Talanta*, Elsevier, v. 115, p. 143–149, 2013. Citado na página 28.
- CHOODUM, A.; PARABUN, K.; KLAWACH, N.; DAEID, N. N.; KANATHARANA, P.; WONGNIRAMAİKUL, W. Real time quantitative colourimetric test for methamphetamine detection using digital and mobile phone technology. *Forensic science international*, Elsevier, v. 235, p. 8–13, 2014. Citado na página 28.
- CHOODUM, A.; SRIPROM, W.; WONGNIRAMAİKUL, W. Portable and selective colorimetric film and digital image colorimetry for detection of iron. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, Elsevier, v. 208, p. 40–47, 2019. Citado na página 28.
- CHUNG, S.-O.; CHO, K.-H.; CHO, J.-W.; JUNG, K.-Y.; YAMAKAWA, T. Soil texture classification algorithm using RGB characteristics of soil images. *Faculty of Agriculture Kyushu University*, v. 57, n. 2, p. 393–397, 2012. Citado 3 vezes nas páginas 29, 86 e 87.
- CLARK, R. N. Spectroscopy of rocks and minerals, and principles of spectroscopy. In: _____. *Manual of remote sensing*. [S.l.]: Wiley, 1999. v. 3, p. 3–58. Citado na página 85.
- CNA. *Agronegócio contribui para queda da inflação e geração de empregos em 2017*. 2017. <http://www.cnabrazil.org.br/noticias/agronegocio-contribui-para-queda-da-inflacao-e-geracao-de-empregos-em-2017>. Online; acessado 06 Dezembro 2017. Citado na página 24.
- COHEN, M. D.; KARGACIN, B.; KLEIN, C. B.; COSTA, M. Mechanisms of chromium carcinogenicity and toxicity. *Critical reviews in toxicology*, Taylor & Francis, v. 23, n. 3, p. 255–281, 1993. Citado na página 111.
- CONAB. *Acompanhamento da Safra Brasileira de Grãos 2019/20*. 2019. 25 p. Online; acessado 21 Novembro 2019. Disponível em: <<http://www.conab.gov.br>>. Citado na página 24.
- CONAB. *Primeiro levantamento da safra 2019/20 de grãos indica produção de 245 milhões de t*. 2019. Online; acessado 21 Novembro 2019. Disponível em: <<https://www.conab.gov.br/ultimas-noticias/3080-primeiro-levantamento-da-safra-2019-20-de-graos-indica-producao-de-245-8-milhoes-de-t>>. Citado na página 24.
- COOK, R. D.; WEISBERG, S. *Residuals and influence in regression*. New York: Chapman and Hall, 1982. Citado na página 53.
- CORÁ, J. E.; FERNANDES, C.; BERALDO, J. M. G.; MARCELO, A. V. Adição de areia para dispersão de solos na análise granulométrica. *R. Bras. Ci. Solo*, v. 33, n. 1, p. 255–262, 2009. Citado na página 33.
- COZZOLINO, D. An overview of the application of near infrared spectroscopy to analyze and monitor soil properties in south america. *Applied Spectroscopy Reviews*, Taylor & Francis, v. 50, n. 10, p. 859–867, 2015. Citado na página 30.
- COZZOLINO, D.; MORON, A. The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics. *The Journal of Agricultural Science*, Cambridge University Press, v. 140, n. 1, p. 65–71, 2003. Citado na página 105.

CRUZ, T. G. da. *Monitoramento por microscopia óptica e processamento digital de imagens do processo de conformação cerâmica por conformação com amidos comerciais*. 1–128 p. Tese (Doutorado) — Universidade Estadual Paulista, 2007. Citado na página 28.

DAMASCENO, D.; TOLEDO, T. G.; GODINHO, M. S.; SILVA, C. P. da; B de O. S.; OLIVEIRA, A. E. de. Análise multivariada de imagens na química: Um experimento para determinação do ph de águas potáveis. *Quim. Nova*, v. 38, n. 6, p. 836–841, 2015. Citado na página 28.

DAMASCENO, D.; TOLEDO, T. G.; SOARES, A. da S.; OLIVEIRA, S. B. de; OLIVEIRA, A. E. de. Compvis: a novel method for drinking water alkalinity and total hardness analyses. *Analytical Methods*, Royal Society of Chemistry, v. 8, n. 43, p. 7832–7836, 2016. Citado na página 28.

DAVIS, R. O. E.; BENNETT, H. H. et al. *Grouping of soils on the basis of mechanical analysis*. [S.l.]: US Dept. of Agriculture, 1927. Citado na página 32.

DEMATTE, J. A.; TERRA, F. da S. Spectral pedology: a new perspective on evaluation of soils along pedogenetic alterations. *Geoderma*, Elsevier, v. 217, p. 190–200, 2014. Citado 2 vezes nas páginas 106 e 109.

DINIZ, M. R. B. R. S.; PEDRON, D. F. de A.; BORKOWSKI, A. C. d. A. R.; MIGUEL, R. P. Medida da cor em solos do rio grande do sul com a carta de munsell e por colorimetria. *Ciência Rural*, SciELO Brasil, v. 36, n. 4, 2006. Citado na página 85.

DINIZ, P. H.; DANTAS, H. V.; MELO, K. D.; BARBOSA, M. F.; HARDING, D. P.; NASCIMENTO, E. C.; PISTONESI, M. F.; BAND, B. S.; ARAÚJO, M. C. Using a simple digital camera and spa-lda modeling to screen teas. *Analytical Methods*, Royal Society of Chemistry, v. 4, n. 9, p. 2648–2652, 2012. Citado na página 28.

DJOMGOUE, P.; NJOPWOUO, D. et al. Ft-ir spectroscopy applied for surface clays characterization. *Journal of Surface Engineered Materials and Advanced Technology*, Scientific Research Publishing, v. 3, n. 04, p. 275, 2013. Citado 2 vezes nas páginas 15 e 59.

DOURADO, T. A.; SILVA, L. F. S. d.; MARINHO, M. d. A. Performance of a reciprocal shaker in mechanical dispersion of soil samples for particle-size analysis. *Revista Brasileira de Ciência do Solo*, SciELO Brasil, v. 36, n. 4, p. 1131–1148, 2012. Citado na página 33.

DRAPER, N. R.; SMITH, H. *Applied regression analysis*. [S.l.]: John Wiley & Sons, 2014. v. 326. Citado na página 43.

DUCHESNE, C.; LIU, J.; MACGREGOR, J. Multivariate image analysis in the process industries: A review. *Chemometrics and Intelligent Laboratory Systems*, Elsevier, v. 117, p. 116–128, 2012. Citado 2 vezes nas páginas 27 e 28.

DURBIN, J.; WATSON, G. S. Testing for serial correlation in least squares regression: I. *Biometrika*, JSTOR, v. 37, n. 3/4, p. 409–428, 1950. Citado na página 52.

DZAYE, E. D.; TSANGOURI, E.; SPIESSENS, K.; SCHUTTER, G. D.; AGGELIS, D. G. Digital image correlation (dic) on fresh cement mortar to quantify settlement and shrinkage. *Archives of Civil and Mechanical Engineering*, Elsevier, v. 19, n. 1, p. 205–214, 2019. Citado na página 27.

EMBRAPA. *Trajectoria da agricultura brasileira*. 2018. <https://www.embrapa.br/visao/trajectoria-da-agricultura-brasileira>. Online; acessado 21 Novembro 2019. Citado na página 24.

ESPINOZA, M.; SUYKENS, J. A.; MOOR, B. D. Least squares support vector machines and primal space estimation. In: IEEE. *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*. [S.l.], 2003. v. 4, p. 3451–3456. Citado na página 49.

Estado de Minas. *Brasil deve assumir liderança na exportação de produtos agrícolas a partir de 2024*. 2015. 1 p. Disponível em: <http://www.em.com.br/app/noticia/economia/2015/11/09/internas_economia,705850/oferta-mundial-nas-maos-do-brasil.shtml>. Acesso em: 23 dez. 2015. Citado na página 24.

FAO. *Água na agricultura*. [S.l.: s.n.], 1998. 18 p. Citado na página 24.

FAO. *Status of the World's Soil Resources (SWSR)–Main Report*. [S.l.]: Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils, 2015. v. 650. Citado na página 24.

FEARN, T. Assessing calibrations: SEP, RPD, RER and R2. *NIR news*, v. 13, n. 6, p. 12–14, 2002. Citado na página 51.

FERRÃO, M. F.; MELLO, C.; BORIN, A.; MARETTO, D. A.; POPPI, R. J. LS-SVM: A new chemometric tool for multivariate regression. comparison of LS-SVM and PLS regression for determination of common adulterants in powdered milk by NIR spectroscopy. *Química Nova*, v. 30, n. 4, p. 852–859, 2007. Citado na página 49.

FIGUEIREDO, D. *Conceitos Básicos de Sensoriamento Remoto*. 2005. 32 p. Disponível em: <http://www.conab.gov.br/conabweb/download/SIGABRASIL/manuais/conceitos_sm.pdf>. Citado na página 27.

FILHO, J. P.; PETRUCI, J. F. da S.; CARDOSO, A. A. Development of a simple method for determination of no2 in air using digital scanner images. *Talanta*, Elsevier, v. 140, p. 73–80, 2015. Citado na página 28.

FIRDAUS, M. L.; ALWI, W.; TRINOVELDI, F.; RAHAYU, I.; RAHMIDAR, L.; WARSITO, K. Determination of chromium and iron using digital image-based colorimetry. *Procedia Environmental Sciences*, Elsevier, v. 20, p. 298–304, 2014. Citado na página 28.

FLOR, E. P. O.; CICERO, S. M.; NETO, J. B. F.; KRZYZANOWSKI, F. C. Avaliação de danos mecânicos em sementes de soja por meio da análise de imagens. *Revista Brasileira de Sementes*, v. 26, n. 1, p. 68–76, 2004. Citado na página 27.

FOCA, G.; MASINO, F.; ANTONELLI, A.; ULRICI, A. Prediction of compositional and sensory characteristics using rgb digital images and multivariate calibration techniques. *Analytica chimica acta*, Elsevier, v. 706, n. 2, p. 238–245, 2011. Citado na página 28.

FREITAS, M. R.; BARIGYE, S. J.; DARÉ, J. K.; FREITAS, M. P. Quantitative modeling of bioconcentration factors of carbonyl herbicides using multivariate image analysis. *Chemosphere*, v. 152, p. 190–195, 2016. Citado na página 28.

FREITAS, P. C. d.; CARVALHO, R. M. d.; SILVA, J. S. A.; PRADO, T.; DUARTE, E. R.; FRESCURA, V. L.; CHAVES, E. S. Extração assistida por ultrassom para determinação colorimétrica de ferro em solo: uma comparação com espectrometria de massa com plasma indutivamente acoplado. *Química Nova*, SciELO Brasil, v. 38, n. 4, p. 570–574, 2015. Citado na página 35.

GALVÃO FILHO, A. R. *Calibração multivariada empregando o algoritmo das projeções sucessivas com reamostragem e combinação de modelos*. 76 p. Dissertação (Mestrado) — Instituto Tecnológico de Aeronáutica, 2010. Citado na página 50.

GALVÃO, R. K. H.; ARAÚJO, M. C. U.; JOSÉ, G. E.; PONTES, M. J. C.; SILVA, E. C.; SALDANHA, T. C. B. A method for calibration and validation subset partitioning. *Talanta*, v. 67, n. 4, p. 736–740, 2005. ISSN 00399140. Citado 3 vezes nas páginas 43, 46 e 64.

GALVÃO, R. K. H.; ARAÚJO, M. C. U.; SILVA, E. C.; JOSÉ, G. E.; SOARES, S. F. C.; PAIVA, H. M. Cross-Validation for the Selection of Spectral Variables Using the Successive Projections Algorithm. *J. Braz. Chem. Soc.*, v. 18, n. 8, p. 1585–1589, 2007. Citado na página 50.

GEE, G. W.; BAUDER, J. W.; KLUTE, A. Particle-size analysis. *Methods of soil analysis. Part 1. Physical and mineralogical methods*, American Society of Agronomy, Inc., p. 383–411, 1986. Citado na página 32.

GELADI, P.; BENGTSSON, E.; ESBENSEN, K.; GRAHN, H. Image analysis in chemistry i. properties of images, greylevel operations, the multivariate image. *TRAC Trends In Analytical Chemistry*, Elsevier, v. 11, n. 1, p. 41–53, 1992. Citado 2 vezes nas páginas 27 e 40.

GELADI, P.; GRAHN, H. *Multivariate image analysis*. Chichester: John Wiley & Sons, 1996. 316 p. ISBN 0471930016. Citado 3 vezes nas páginas 35, 36 e 39.

GENRICH, D.; BREMNER, J. A reevaluation of the ultrasonic-vibration method of dispersing soils. *Soil Science Society of America Journal*, Soil Science Society of America, v. 36, n. 6, p. 944–947, 1972. Citado na página 33.

Globo Rural. *Brasil será maior exportador agrícola mundial em 2024, dizem OCDE e FAO*. 2015. 1 p. Disponível em: <<http://www.g1.globo.com/economia/agronegocios/noticia/2015/07/brasil-sera-maior-exportador-agricola-mundial-em-2024-dizem-ocde-e-fao.html>>. Acesso em: 23 dez. 2015. Citado na página 24.

GODINHO, M. S.; PEREIRA, R. O.; RIBEIRO, K. D. O.; SCHIMIDT, F.; OLIVEIRA, A. E. de; OLIVEIRA, S. B. de. Classificação de refrigerantes através de análise de imagens e análise de componentes principais (PCA). *Quim. Nova*, v. 31, n. 6, p. 1485–1489, 2008. Citado na página 28.

GOLDSTEIN, M. D.; NAGLER, M. Real time inspection of a large set of surface defects in metal parts. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Robotics and IECON'87 Conferences*. [S.l.], 1988. p. 184–190. Citado na página 27.

GOMES, M. S.; TREVIZAN, L. C.; NÓBREGA, J. A.; KAMOGAWA, M. Y. Uso de scanner em espectrofotometria de absorção molecular: aplicação em experimento didático

enfocando a determinação de ácido ascórbico. *Quim. Nova*, v. 31, n. 6, p. 1577–1581, 2008. Citado na página 28.

GONÇALVES, M. I. S.; VILAR, W. T.; MEDEIROS, E. P.; PONTES, M. J. C. An analytical method for determination of quality parameters in cotton plumes by digital image and chemometrics. *Computers and Electronics in Agriculture*, Elsevier, v. 123, p. 89–94, 2016. Citado na página 27.

GONZALEZ, R. C.; WOODS, R. E. *Processamento de Imagens Digitais*. 3. ed. São Paulo: Pearson Prentice Hall, 2010. Citado 4 vezes nas páginas 35, 37, 38 e 39.

HOGAN, R.; GOODWIN, M.; BOOTHMAN, N.; IAFOLLA, T.; PRETTY, I. A. Further opportunities for digital imaging in dental epidemiology. *Journal of dentistry*, Elsevier, v. 74, p. S2–S9, 2018. Citado na página 28.

HOLLER, F. J.; SKOOG, D. A.; CROUCH, S. R. *Princípios de análise instrumental*. sixth. Porto Alegre: Bookman, 2009. Citado 2 vezes nas páginas 41 e 42.

HUGHES, J. F.; DAM, A. V.; MCGUIRE, M.; SKLAR, D. F.; FOLEY, J. D.; FEINER, S. K.; AKELEY, K. *Computer Graphics Principles and Practice*. 3. ed. Reading, MA: Addison-Wesley, 2013. 1262 p. ISBN 9780321399526. Citado na página 37.

IBRAHIM, M.; HAMEED, A. J.; JALBOUT, A. Molecular spectroscopic study of river Nile sediment in the greater Cairo region. *Applied Spectroscopy*, SAGE Publications Sage UK: London, England, v. 62, n. 3, p. 306–311, 2008. Citado 2 vezes nas páginas 15 e 59.

IGNE, B.; HURBURGH, C. R. Local chemometrics for samples and variables: optimizing calibration and standardization processes. *Journal of Chemometrics*, Wiley Online Library, v. 24, n. 2, p. 75–86, 2010. Citado na página 51.

IGOE, D.; PARISI, A. Characterization of the corrosion of iron using a smartphone camera. *Instrumentation Science & Technology*, Taylor & Francis, v. 44, n. 2, p. 139–147, 2016. Citado na página 28.

JANIK, L. J.; MERRY, R. H.; SKJEMSTAD, J. Can mid infrared diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture*, CSIRO, v. 38, n. 7, p. 681–696, 1998. Citado 2 vezes nas páginas 100 e 102.

JANIK, L. J.; SKJEMSTAD, J. Characterization and analysis of soils using mid-infrared partial least-squares. 2. correlations with some laboratory data. *Soil Research*, CSIRO, v. 33, n. 4, p. 637–650, 1995. Citado 2 vezes nas páginas 100 e 102.

KENNARD, R. W.; STONE, L. A. Computer Aided Design Experiments. *Technometrics*, v. 11, n. 1, p. 137–148, 1969. Disponível em: <<http://www.jstor.org/stable/10.2307-1266219>>. Citado 4 vezes nas páginas 43, 45, 87 e 103.

KNOX, N.; GRUNWALD, S.; MCDOWELL, M.; BRULAND, G.; MYERS, D.; HARRIS, W. Modelling soil carbon fractions with visible near-infrared (vnir) and mid-infrared (mir) spectroscopy. *Geoderma*, Elsevier, v. 239, p. 229–239, 2015. Citado na página 98.

KOCHIAN, L. V. Cellular mechanisms of aluminum toxicity and resistance in plants. *Annual review of plant biology*, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 46, n. 1, p. 237–260, 1995. Citado na página 34.

KONEN, M.; BURRAS, C.; SANDOR, J. Organic carbon, texture, and quantitative color measurement relationships for cultivated soils in north central iowa. *Soil Science Society of America Journal*, Soil Science Society, v. 67, n. 6, p. 1823–1830, 2003. Citado na página 32.

KONEN, M.; BURRAS, C.; SANDOR, J. Organic carbon, texture, and quantitative color measurement relationships for cultivated soils in north central iowa. *Soil Science Society of America Journal*, Soil Science Society, v. 67, n. 6, p. 1823–1830, 2003. Citado na página 85.

KOVDA, V. *The bases of learning about soils*. Moscow: Nauka, 1973. v. 2. Citado na página 34.

KOZLOWSKI, C.; JEET, S.; BEYER, J.; GUERRERO, S.; LESCH, J.; WANG, X.; DEVOSS, J.; DIEHL, L. An entirely automated method to score dss-induced colitis in mice by digital image analysis of pathology slides. *Disease models & mechanisms*, v. 6, n. 3, p. 855–865, 2013. Citado na página 27.

KRAMER, R. *Chemometric techniques for quantitative analysis*. New York: Marcel Dekker, 1998. 1–203 p. ISBN 0824701984. Citado na página 43.

LADO, M.; PAZ, A.; BEN-HUR, M. Organic matter and aggregate-size interactions in saturated hydraulic conductivity. *Soil Science Society of America Journal*, v. 68, n. 1, p. 234–242, 2004. Citado na página 31.

LEDLEY, R. S. High-speed automatic analysis of biomedical pictures. *Science*, American Association for the Advancement of Science, v. 146, n. 3641, p. 216–223, 1964. Citado na página 27.

LEMES, M. A.; RABELO, D.; OLIVEIRA, A. E. de. A novel method to evaluate nanofluid stability using multivariate image analysis. *Analytical Methods*, v. 9, n. 39, p. 5826–5833, 2017. Citado na página 28.

LEVIN, N.; BEN-DOR, E.; SINGER, A. A digital camera as a tool to measure colour indices and related properties of sandy soils in semi-arid environments. *International Journal of Remote Sensing*, Taylor & Francis, v. 26, n. 24, p. 5475–5492, 2005. Citado 3 vezes nas páginas 29, 85 e 91.

LI, J.; SUN, Y.; CHEN, C.; SHENG, T.; LIU, P.; ZHANG, G. A smartphone-assisted microfluidic chemistry analyzer using image-based colorimetric assays for multi-index monitoring of diabetes and hyperlipidemia. *Analytica chimica acta*, Elsevier, v. 1052, p. 105–112, 2019. Citado na página 27.

LÓPEZ-GARCÍA, F.; ANDREU-GARCÍA, G.; BLASCO, J.; ALEIXOS, N.; VALIENTE, J.-M. Automatic detection of skin defects in citrus fruits using a multivariate image analysis approach. *Computers and Electronics in Agriculture*, Elsevier, v. 71, n. 2, p. 189–197, 2010. Citado na página 27.

LOPEZ-MOLINERO, A.; LIÑAN, D.; SIPIERA, D.; FALCON, R. Chemometric interpretation of digital image colorimetry. application for titanium determination in plastics. *Microchemical Journal*, Elsevier, v. 96, n. 2, p. 380–385, 2010. Citado na página 28.

LU, N. *Fractal Imaging*. San Diego: Academic Press, 1997. 412 p. ISBN 0124580106. Citado na página 36.

LUDWIG, B.; NITSCHKE, R.; TERHOEVEN-URSELMANS, T.; MICHEL, K.; FLESSA, H. Use of mid-infrared spectroscopy in the diffuse-reflectance mode for the prediction of the composition of organic matter in soil and litter. *Journal of Plant Nutrition and Soil Science*, Wiley Online Library, v. 171, n. 3, p. 384–391, 2008. Citado na página 98.

LUIZ, V. H. M.; LIMA, L. S.; ROSSINI, E. L.; PEZZA, L.; PEZZA, H. R. Paper platform for determination of bumetanide in human urine samples to detect doping in sports using digital image analysis. *Microchemical Journal*, Elsevier, 2019. Citado na página 27.

LUTS, J.; OJEDA, F.; PLAS, R. Van de; MOOR, B. D.; HUFFEL, S. V.; SUYKENS, J. A. A tutorial on support vector machine-based methods for classification problems in chemometrics. *Analytica Chimica Acta*, Elsevier, v. 665, n. 2, p. 129–145, 2010. Citado na página 49.

LYRA, W. da S.; ALMEIDA, L. F. D.; CUNHA, F. A. da S.; DINIZ, P. H. G. D.; MARTINS, V. L.; ARAÚJO, M. C. U. de. Determination of sodium and calcium in powder milk using digital image-based flame emission spectrometry. *Analytical Methods*, Royal Society of Chemistry, v. 6, n. 4, p. 1044–1050, 2014. Citado na página 28.

MA, J. F. Plant root responses to three abundant soil minerals: silicon, aluminum and iron. *Critical Reviews in Plant Sciences*, Taylor & Francis, v. 24, n. 4, p. 267–281, 2005. Citado na página 34.

MA, J. F.; TAMAI, K.; YAMAJI, N.; MITANI, N.; KONISHI, S.; KATSUHARA, M.; ISHIGURO, M.; MURATA, Y.; YANO, M. A silicon transporter in rice. *Nature*, Nature Publishing Group, v. 440, n. 7084, p. 688, 2006. Citado na página 34.

MA, T. S. Elemental analysis organic compounds. In: *Meyers, RA*. Orlando: Academic Press, 2001. (Encyclopedia of physical science and technology), p. 393–405. Citado na página 31.

MADARI, B. E.; III, J. B. R.; COELHO, M. R.; MACHADO, P. L.; DE-POLLI, H.; COELHO, R. M.; BENITES, V. M.; SOUZA, L. F.; MCCARTY, G. W. Mid-and near-infrared spectroscopic determination of carbon in a diverse set of soils from the brazilian national soil collection. *Spectroscopy Letters*, Taylor & Francis, v. 38, n. 6, p. 721–740, 2005. Citado 2 vezes nas páginas 29 e 98.

MADARI, B. E.; III, J. B. R.; MACHADO, P. L.; GUIMARÃES, C. M.; TORRES, E.; MCCARTY, G. W. Mid-and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two ferralsols. *Geoderma*, Elsevier, v. 136, n. 1-2, p. 245–259, 2006. Citado 3 vezes nas páginas 15, 59 e 60.

MADEJOVÁ, J. Ftir techniques in clay mineral studies. *Vibrational spectroscopy*, Elsevier, v. 31, n. 1, p. 1–10, 2003. Citado 2 vezes nas páginas 15 e 59.

MAPA. *Projeções do agronegócio-Brasil 2016/2017 a 2026/2027*. 8. ed. Brasília: Ministério da Agricultura, Pecuária e Abastecimento, 2017. 125 p. Citado na página 24.

MAPA. *Agropecuária Brasileira em Números*. 2019. <http://www.agricultura.gov.br/assuntos/politica-agricola/agropecuaria-brasileira-em-numeros>. Online; acessado 21 Novembro 2019. Citado na página 24.

MARCELO, M.; MARIOTTI, K.; ORTIZ, R.; FERRÃO, M.; ANZANELLO, M. Scott test evaluation by multivariate image analysis in cocaine samples. *Microchemical Journal*, Elsevier, v. 127, p. 87–93, 2016. Citado na página 28.

MARSCHNER, H. *Marschner's mineral nutrition of higher plants*. London: Academic press, 1995. Citado na página 34.

MARTENS, H.; NÆS, T. *Multivariate Calibration*. Chichester: John Wiley & Sons, 1991. 419 p. ISBN 0471909793. Citado 3 vezes nas páginas 43, 47 e 49.

MATICHENKOV, V.; CALVERT, D. Silicon as a beneficial element for sugarcane. *Journal American Society of Sugarcane Technologists*, v. 22, n. 2, p. 21–30, 2002. Citado na página 34.

MAURI, J.; RUIZ, H. A.; FERNANDES, R. B. A.; KER, J. C.; REZENDE, L. R. M. Dispersantes químicos na análise granulométrica de Latossolos. *R. Bras. Ci. Solo*, v. 35, n. 1, p. 1277–1284, 2011. Citado na página 33.

MCCARTY, G.; REEVES, J.; REEVES, V.; FOLLETT, R.; KIMBLE, J. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Science Society of America Journal*, Soil Science Society, v. 66, n. 2, p. 640–646, 2002. Citado na página 98.

MCCARTY, G. W.; REEVES, J. B. Comparison of near infrared and mid infrared diffuse reflectance spectroscopy for field-scale measurement of soil fertility parameters. *Soil Science*, LWW, v. 171, n. 2, p. 94–102, 2006. Citado 2 vezes nas páginas 98 e 103.

MCLEAN, E. O. Aluminum. In: _____. *Methods of Soil Analysis: Part II. Chemical methods*. Madison: American Society of Agronomy, 1965. p. 978–998. Citado na página 35.

MEBIUS, L. A rapid method for the determination of organic carbon in soil. *Analytica Chimica Acta*, Elsevier, v. 22, p. 120–124, 1960. Citado na página 31.

MELO, V.; FONTES, M.; NOVAIS, R.; SINGH, B.; SCHAEFER, C. Características dos óxidos de ferro de alumínio de diferentes classes de solos. *Revista Brasileira de Ciência do Solo*, v. 25, n. 1, p. 19–32, 2001. Citado na página 34.

MENESES, F. G. A. de. *Segmentação de imagens de arte rupestre utilizando o sistema de coordenadas estelares*. 1–58 p. Tese (Doutorado) — Universidade Federal Fluminense, 2013. Citado na página 28.

MORAIS, P. A. de O.; SOUZA, D. M. de; CARVALHO, M. T. de M.; MADARI, B. E.; OLIVEIRA, A. E. de. Predicting soil texture using image analysis. *Microchemical Journal*, Elsevier, 2019. Citado 2 vezes nas páginas 29 e 86.

MORITSUKA, N.; MATSUOKA, K.; KATSURA, K.; SANO, S.; YANAI, J. Soil color analysis for statistically estimating total carbon, total nitrogen and active iron contents in japanese agricultural soils. *Soil Science and Plant Nutrition*, Taylor & Francis, v. 60, n. 4, p. 475–485, 2014. Citado na página 91.

MORON, A.; COZZOLINO, D. Exploring the use of near infrared reflectance spectroscopy to study physical properties and microelements in soils. *Journal of Near Infrared Spectroscopy*, SAGE Publishing, v. 11, n. 2, p. 145–154, 2003. Citado na página 29.

MORRISON, I.; WILSON, A. The absorptiometric determination of silicon in water. part iii. method for determining the total silicon content. *Analyst*, v. 88, n. 1047, p. 446–455, 1963. Citado na página 35.

MURTI, G. K.; SATYANARAYANA, K. Influence of chemical characteristics in the development of soil colour. *Geoderma*, Elsevier, v. 5, n. 3, p. 243–248, 1971. Citado na página 85.

NANNI, M. R.; DEMATTÊ, J. A. M. Spectral reflectance methodology in comparison to traditional soil analysis. *Soil Science Society of America Journal*, Soil Science Society, v. 70, n. 2, p. 393–407, 2006. Citado 2 vezes nas páginas 105 e 106.

NETO, J. H. S.; PORTO, I. S.; SCHNEIDER, M. P.; SANTOS, A. M. dos; GOMES, A. A.; FERREIRA, S. L. Speciation analysis based on digital image colorimetry: Iron (ii/iii) in white wine. *Talanta*, Elsevier, v. 194, p. 86–89, 2019. Citado na página 28.

NIAZI, A.; GOODARZI, M.; YAZDANIPOUR, A. A comparative study between least-squares support vector machines and partial least squares in simultaneous spectrophotometric determination of cypermethrin, permethrin and tetramethrin. *Journal of the Brazilian Chemical Society*, SciELO Brasil, v. 19, n. 3, p. 536–542, 2008. Citado na página 49.

NIEMEYER, J.; CHEN, Y.; BOLLAG, J.-M. Characterization of humic acids, composts, and peat by diffuse reflectance fourier-transform infrared spectroscopy. *Soil Science Society of America Journal*, v. 56, n. 1, p. 135–140, 1992. Citado 2 vezes nas páginas 15 e 60.

NOCITA, M.; STEVENS, A.; WESEMAEL, B. van; AITKENHEAD, M.; WETTERLIND, J.; FRAC, M.; JEZIERSKA-TYS, S.; YAGUCHI, T. 10. chapter four-soil spectroscopy: An alternative to wet chemistry for soil monitoring. *Advances in Agronomy*, Elsevier, v. 132, p. 1–204, 2015. Citado na página 30.

NOVAIS, R.; ALVAREZ, V. V. H.; BARROS, N. F.; FONTES, R.; CANTARUTTI, R. B.; NEVES, J. (Ed.). *Fertilidade do solo*. [S.l.]: Sociedade Brasileira de Ciência do Solo, 2007. Citado na página 31.

OADES, J. The role of biology in the formation, stabilization and degradation of soil structure. *Geoderma*, v. 56, n. 1-4, p. 377–400, 1993. Citado na página 31.

OLIVEIRA, K. A.; DAMASCENO, D.; OLIVEIRA, C. R. de; SILVEIRA, L. A. da; OLIVEIRA, A. E. de; COLTRO, W. K. Dengue diagnosis on laser printed microzones using smartphone-based detection and multivariate image analysis. *Analytical Methods*, Royal Society of Chemistry, v. 8, n. 35, p. 6506–6511, 2016. Citado na página 27.

OLIVEIRA, L. F.; CANEVARI, N. T.; JESUS, A.; PEREIRA-FILHO, E. R. Uso de Imagens Digitais e Análise de Componentes Principais na Identificação dos Níveis de Cr (VI) em Amostras de Solos. *Revista Virtual de Química*, v. 5, n. 2, p. 169–178, 2013. Citado na página 28.

OTOMO, M. The spectrophotometric determination of aluminum with xylenol orange. *Bulletin of the Chemical Society of Japan*, The Chemical Society of Japan, v. 36, n. 7, p. 809–813, 1963. Citado na página 35.

- OTTO, M. *Chemometrics: statistics and computer application in analytical chemistry*. Weinheim: Wiley-VCH, 1999. 1–314 p. ISBN 352729628X. Citado na página 44.
- PACIORNIK, S.; YALLOUZ, A. V.; CAMPOS, R. C.; GANNERMAN, D. Scanner image analysis in the quantification of mercury using spot-tests. *Journal of the Brazilian Chemical Society*, SciELO Brasil, v. 17, n. 1, p. 156–161, 2006. Citado na página 28.
- PANSU, M.; GAUTHEYROU, J. *Handbook of Soil Analysis Mineralogical, organic and inorganic methods*. New York: Springer, 2003. 995 p. ISBN 9783540312109. Citado 3 vezes nas páginas 32, 33 e 57.
- PASS, G.; ZABIH, R. Comparing images using joint histograms. *Multimedia systems*, v. 7, n. 3, p. 234–240, 1999. Citado na página 40.
- PATACA, L. C. M. *Análises de mel e própolis utilizando métodos Quimiométricos de Classificação e Calibração*. 97 p. Tese (Tese de doutorado) — Universidade Estadual de Campinas, 2006. Citado na página 28.
- PAVIA, D. L.; LAMPMAN, G. M.; KRIZ, G. S.; VYVYAN, J. R. *Introdução à espectroscopia*. [S.l.]: Cengage Learning, 2010. Citado na página 42.
- PAWAR, P.; CORTÉS, C.; MURRAY, K.; KLEISSL, J. Detecting clear sky images. *Solar Energy*, Elsevier, v. 183, p. 50–56, 2019. Citado na página 27.
- PESSOA, K. D.; SUAREZ, W. T.; REIS, M. F. dos; FRANCO, M. d. O. K.; MOREIRA, R. P. L.; SANTOS, V. B. dos. A digital image method of spot tests for determination of copper in sugar cane spirits. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, Elsevier, v. 185, p. 310–316, 2017. Citado na página 28.
- PINTO, F.; CORRÊA, P.; QUEIROZ, D.; OLIVEIRA, A. Utilização de um sistema de visão artificial para estimativa do teor de umidade de frutos de café. *Revista Brasileira de Armazenamento*, v. 12, n. 6, p. 30–36, 2003. Citado na página 27.
- PIRIE, A.; SINGH, B.; ISLAM, K. Ultra-violet, visible, near-infrared, and mid-infrared diffuse reflectance spectroscopic techniques to predict several soil properties. *Australian Journal of Soil Research*, v. 43, n. 6, p. 713–721, 2005. Citado 2 vezes nas páginas 60 e 102.
- PLANTE, A. F.; CONANT, R. T.; STEWART, C. E.; PAUSTIAN, K.; SIX, J. Impact of soil texture on the distribution of soil organic matter in physical and chemical fractions. *Soil Science Society of America Journal*, Soil Science Society, v. 70, n. 1, p. 287–296, 2006. Citado na página 32.
- PRATS-MONTALBÁN, J.; JUAN, A. de; FERRER, A. Multivariate image analysis: A review with applications. *Chemometrics and Intelligent Laboratory Systems*, Elsevier B.V., v. 107, n. 1, p. 1–23, maio 2011. ISSN 01697439. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0169743911000529>>. Citado na página 27.
- RAMAROSON, V. H.; BECQUER, T.; SÁ, S. O.; RAZAFIMAHATRATRA, H.; DELARIVIÈRE, J. L.; BLAVET, D.; VENDRAME, P. R.; RABEHARISOA, L.; RAKOTONDRAZAFY, A. F. Mineralogical analysis of ferralitic soils in madagascar using nir spectroscopy. *Catena*, Elsevier, v. 168, p. 102–109, 2018. Citado 3 vezes nas páginas 105, 107 e 109.

REEVES, J. B.; SMITH, D. B. The potential of mid-and near-infrared diffuse reflectance spectroscopy for determining major-and trace-element concentrations in soils from a geochemical survey of north america. *Applied Geochemistry*, Elsevier, v. 24, n. 8, p. 1472–1481, 2009. Citado na página 98.

RIBEIRO, F. C.; OLIVEIRA, A. S.; ARAÚJO, A.; MARINHO, W.; SCHNEIDER, M. P.; PINTO, L.; GOMES, A. A. Detection oxidative degradation in lubricating oil under storage conditions using digital images and chemometrics. *Microchemical Journal*, Elsevier, 2019. Citado na página 28.

ROBINSON, G. W. A new method for the mechanical analysis of soils and other dispersions. *The Journal of Agricultural Science*, v. 12, p. 306–321, 7 1922. ISSN 1469-5146. Citado 2 vezes nas páginas 32 e 57.

RODIONOVA, O. Y.; POMERANTSEV, A. L. Chemometrics: achievements and prospects. *Russian Chemical Reviews*, v. 75, n. 4, p. 271–287, 2007. ISSN 0036-021X. Citado na página 44.

SAEYS, W.; MOUAZEN, A. M.; RAMON, H. Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy. *Biosystems engineering*, v. 91, n. 4, p. 393–402, 2005. Citado 3 vezes nas páginas 71, 77 e 81.

SAIKIA, B. J.; PARTHASARATHY, G. et al. Fourier transform infrared spectroscopic characterization of kaolinite from assam and meghalaya, northeastern india. *J. Mod. Phys*, v. 1, n. 4, p. 206–210, 2010. Citado 2 vezes nas páginas 15 e 59.

SALTER, P.; WILLIAMS, J. The influence of texture on the moisture characteristics of soils. *European Journal of Soil Science*, Wiley Online Library, v. 16, n. 1, p. 1–15, 1965. Citado na página 32.

SÁNCHEZ-MARAÑÓN, M.; DELGADO, G.; MELGOSA, M.; HITTA, E.; DELGADO, R. Cielab color parameters and their relationship to soil characteristics in mediterranean red soils. *Soil science*, LWW, v. 162, n. 11, p. 833–842, 1997. Citado na página 85.

SANTOS, A. D.; COSCIONE, A. R.; VITTI, A. C.; BOARETTO, A. E.; COELHO, A. M.; RAIJ, B. van; SILVA, C. A.; JUNIOR, C. H. A.; CARMO, C. A. F. S. do; ABREU, C. A. de; SILVA, C. R. da; GIANELLO, C.; ANDRADE, C. A. de; PÉREZ, D. V.; CASARINI, D. C. P.; SILVA, F. C. da; PRATA, F.; CARVALHO, F. d. C.; SANTOS, G. C. G. dos; CANTARELLA, H.; FERNANDES, H. M. G.; ANDRADE, J. a. C. de; QUAGGIO, J. A.; CHITOLINA, J. C.; CUNHA, L. M. S.; PAVAN, M. A.; ROSIAS, M. F. G. G.; TEDESCO, M. J.; MIYAZAWA, M.; ABREU, M. F. de; EIRA, P. A. da; HIGA, R. H.; MASSRUHá, S. M. F. S.; GOMES, T. F.; MURAOKA, T.; VIEIRA, W.; MELO, W. J. de; BARRETO, W. d. O. *Manual de análises químicas de solos , plantas e fertilizantes*. 2. ed. Brasília, DF: Embrapa, 2010. 627 p. ISBN 978-85-7383-430-7 1. Citado 2 vezes nas páginas 25 e 31.

SCHULZE, D.; NAGEL, J.; SCOYOC, G. van; HENDERSON, T.; BAUMGARDNER, M.; STOTT, D. Significance of organic matter in determining soil colors. In: J.M CIOLKOSZ, E. (Ed.). *Bigham*. Madison: Science Society of America, 1993. (Soil Color), p. 71–90. Citado na página 74.

SCHWERTMANN, U. Relations between iron oxides, soil color, and soil formation. In: BIGHAM, J.; CIOLKOSZ, E. (Ed.). *Soil color*. Madison: Soil Science Society of America, 1993. p. 51–69. Citado 2 vezes nas páginas 34 e 90.

SCHWERTMANN, U.; TAYLOR, R. M. Iron oxides. In: _____. *Minerals in soil environments*. 2. ed. Madison: Soil Science Society of America, 1989. p. 379–438. Citado na página 34.

SCOPEL, E.; TRIOMPHE, B.; AFFHOLDER, F.; SILVA, F. A. M. D.; CORBEELS, M.; XAVIER, J. H. V.; LAHMAR, R.; RECOUS, S.; BERNOUX, M.; BLANCHART, E.; MENDES, I. d. C.; TOURDONNET, S. d. Conservation agriculture cropping systems in temperate and tropical conditions, performances and impacts. a review. *Agronomy for Sustainable Development*, v. 33, p. 113–130, 2013. Citado na página 24.

SEASHOLTZ, M. B.; KOWALSKI, B. The parsimony principle applied to multivariate calibration. *Analytica Chimica Acta*, Elsevier, v. 277, n. 2, p. 165–177, 1993. Citado 2 vezes nas páginas 74 e 83.

SHEPHERD, K. D.; WALSH, M. G. Development of reflectance spectral libraries for characterization of soil properties. *Soil science society of America journal*, Soil Science Society, v. 66, n. 3, p. 988–998, 2002. Citado na página 29.

SILVA, S. H. M. G. D. *Processamento digital de imagens para identificação da sigatoka negra em bananais utilizando análise de componentes principais e redes neurais artificiais*. 83 p. Tese (Doutorado) — UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO, 2008. Citado na página 27.

SILVA, T. C. d. O.; GODINHO, M. S.; OLIVEIRA, A. E. Identification of pale lager beers via image analysis. *Latin American applied research*, v. 41, n. 2, p. 141–145, 2011. Citado na página 28.

SILVER, W. L.; NEFF, J.; MCGRODDY, M.; VELDKAMP, E.; KELLER, M.; COSME, R. Effects of soil texture on belowground carbon and nutrient storage in a lowland amazonian forest ecosystem. *Ecosystems*, Springer, v. 3, n. 2, p. 193–209, 2000. Citado na página 32.

SOARES, S.; LIMA, M. J.; ROCHA, F. R. A spot test for iodine value determination in biodiesel based on digital images exploiting a smartphone. *Microchemical Journal*, Elsevier, v. 133, p. 195–199, 2017. Citado na página 28.

SOARES, S.; TORRES, K. G.; PIMENTEL, E. L.; MARTELLI, P. B.; ROCHA, F. R. A novel spot test based on digital images for determination of methanol in biodiesel. *Talanta*, Elsevier, v. 195, p. 229–235, 2019. Citado na página 28.

SORIANO-DISLA, J. M.; JANIK, L. J.; VISCARRA-ROSSEL, R. A.; MACDONALD, L. M.; MCLAUGHLIN, M. J. The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. *Applied Spectroscopy Reviews*, v. 49, n. 2, p. 139–186, 2014. Citado 2 vezes nas páginas 29 e 59.

SOUZA, D. M. d.; MORAIS, P. A. d. O.; MATSUSHIGE, I.; ROSA, L. A. Development of alternative methods for determining soil organic matter. *Revista Brasileira de Ciência do Solo*, v. 40, 2016. Citado 2 vezes nas páginas 32 e 56.

- SOUZA, F. R. de; ALVES, G. L.; COLTRO, W. K. T. Capillary-driven toner-based microfluidic devices for clinical diagnostics with colorimetric detection. *Analytical chemistry*, v. 84, n. 21, p. 9002–9007, 2012. Citado na página 27.
- SOUZA, R. V. B.; CELLIGOI, A. Avaliação da condutividade hidráulica do solo em área agrícola e florestada na cidade de Londrina (PR) através do permeâmetro Guelph. *Boletim de Geografia*, v. 29, n. 2, p. 123–133, abr. 2012. ISSN 2176-4786. Citado na página 32.
- SPIELVOGEL, S.; KNICKER, H.; KÖGEL-KNABNER, I. Soil organic matter composition and soil lightness. *Journal of Plant Nutrition and Soil Science*, Wiley Online Library, v. 167, n. 5, p. 545–555, 2004. Citado na página 85.
- STENBERG, B.; ROSSEL, R. A. V.; MOUAZEN, A. M.; WETTERLIND, J. Visible and near infrared spectroscopy in soil science. In: _____. *Advances in agronomy*. [S.l.]: Academic Press, 2010. v. 107, p. 163–215. Citado 2 vezes nas páginas 29 e 60.
- STEVENSON, F. J. *Humus chemistry: genesis, composition, reactions*. 2. ed. New York: Wiley, 1994. 496 p. Citado 2 vezes nas páginas 15 e 60.
- STIGLITZ, R.; MIKHAILOVA, E.; POST, C.; SCHLAUTMAN, M.; SHARP, J. Using an inexpensive color sensor for rapid assessment of soil organic carbon. *Geoderma*, Elsevier, v. 286, p. 98–103, 2017. Citado na página 74.
- SUART, B. *Infrared spectroscopy: Fundamental and applications*. Chichester: John Wiley Sons, 2004. Citado 2 vezes nas páginas 41 e 42.
- SUMRIDDETKAJORN, S.; CHAITAVON, K.; INTARAVANNE, Y. Mobile-platform based colorimeter for monitoring chlorine concentration in water. *Sensors and Actuators B: Chemical*, Elsevier, v. 191, p. 561–566, 2014. Citado na página 28.
- SUN, C.; FIKSDAL, L.; HANSEN-BAUER, A.; RYE, M. B.; LEIKNES, T. Characterization of membrane biofouling at different operating conditions (flux) in drinking water treatment using confocal laser scanning microscopy (clsm) and image analysis. *Journal of membrane science*, v. 382, n. 1, p. 194–201, 2011. Citado na página 28.
- SUNOJ, S.; SUBHASHREE, S.; DHARANI, S.; IGATHINATHANE, C.; FRANCO, J.; MALLINGER, R.; PRASIFKA, J.; ARCHER, D. Sunflower floral dimension measurements using digital image processing. *Computers and Electronics in Agriculture*, Elsevier, v. 151, p. 403–415, 2018. Citado na página 27.
- SUYKENS, J. A.; GESTEL, T. V.; BRABANTER, J. D.; MOOR, B. D.; VANDEWALLE, J. *Least squares support vector machines*. Singapore: World Scientific, 2002. Citado 2 vezes nas páginas 43 e 64.
- SUYKENS, J. A.; VANDEWALLE, J. Least squares support vector machine classifiers. *Neural processing letters*, Springer, v. 9, n. 3, p. 293–300, 1999. Citado na página 49.
- TEIXEIRA, P. C.; DONAGEMMA, G. K.; FONTANA, A.; TEIXEIRA, W. G. *Manual de métodos de análise de solo*. 3. ed. Brasília: Embrapa Solos, 2017. 573 p. Citado 2 vezes nas páginas 34 e 35.

TERHOEVEN-URSELMANS, T.; VAGEN, T.-G.; SPAARGAREN, O.; SHEPHERD, K. D. Prediction of soil fertility properties from a globally distributed soil mid-infrared spectral library. *Soil Science Society of America Journal*, Soil Science Society, v. 74, n. 5, p. 1792–1799, 2010. Citado 2 vezes nas páginas 102 e 103.

THANAKIATKRAI, P.; YAODAM, A.; KITPIPIT, T. Age estimation of bloodstains using smartphones and digital image analysis. *Forensic science international*, Elsevier, v. 233, n. 1-3, p. 288–297, 2013. Citado na página 28.

TORRENT, J.; BARRÓN, V. Laboratory measurement of soil color: theory and practice. In: _____. *Soil color*. Madison: Soil Science Society of America, 1993. p. 21–33. Citado na página 85.

TORRENT, J.; SCHWERTMANN, U.; FECHTER, H.; ALFEREZ, F. Quantitative relationships between soil color and hematite content. *Soil Science*, LWW, v. 136, n. 6, p. 354–358, 1983. Citado 2 vezes nas páginas 34 e 90.

TORRENT, J.; SCHWERTMANN, U.; SCHULZE, D. Iron oxide mineralogy of some soils of two river terrace sequences in Spain. *Geoderma*, Elsevier, v. 23, n. 3, p. 191–208, 1980. Citado na página 90.

UEHARA, G. Acric properties and their significance to soil classification. In: *Eighth International Soil Classification Workshop: Classification, characterization and utilization of oxisols*. [S.l.]: EMBRAPA, 1988. p. 19–22. Citado na página 34.

VAPNIK, V. Pattern recognition using generalized portrait method. *Automation and remote control*, v. 24, p. 774–780, 1963. Citado na página 49.

VÁZQUEZ, M. A.; GALÁN, E.; ORTIZ, P.; ORTIZ, R. Digital image analysis and edx sem as combined techniques to evaluate salt damp on walls. *Construction and Building Materials*, v. 45, p. 95–105, 2013. Citado na página 27.

VIDAL, M.; AMIGO, J.; BRO, R.; BERG, F. V. D.; OSTRA, M.; UBIDE, C. Image analysis for maintenance of coating quality in nickel electroplating baths—real time control. *Analytica chimica acta*, Elsevier, v. 706, n. 1, p. 1–7, 2011. Citado 2 vezes nas páginas 27 e 28.

VILAR, W. T.; ARANHA, R. M.; MEDEIROS, E. P.; PONTES, M. J. Classification of individual castor seeds using digital imaging and multivariate analysis. *Journal of the Brazilian Chemical Society*, v. 26, n. 1, p. 102–109, 2015. Citado na página 27.

VISCARRA-ROSSEL, R.; FOUAD, Y.; WALTER, C. Using a digital camera to measure soil organic carbon and iron contents. *Biosystems Engineering*, Elsevier, v. 100, n. 2, p. 149–159, 2008. Citado na página 29.

VISCARRA-ROSSEL, R.; MINASNY, B.; ROUDIER, P.; MCBRACTNEY, A. Colour space models for soil science. *Geoderma*, v. 133, n. 3, p. 320–337, 2006. Citado 3 vezes nas páginas 31, 32 e 74.

VISCARRA-ROSSEL, R.; WALVOORT, D.; MCBRATNEY, A.; JANIK, L. J.; SKJEMSTAD, J. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, Elsevier, v. 131, n. 1-2, p. 59–75, 2006. Citado 4 vezes nas páginas 29, 98, 100 e 102.

VOET, H. van der. Comparing the predictive accuracy of models using a simple randomization test. *Chemometrics and intelligent laboratory systems*, Elsevier, v. 25, n. 2, p. 313–323, 1994. Citado 2 vezes nas páginas 71 e 78.

VOHLAND, M.; LUDWIG, M.; THIELE-BRUHN, S.; LUDWIG, B. Determination of soil properties with visible to near-and mid-infrared spectroscopy: Effects of spectral variable selection. *Geoderma*, Elsevier, v. 223, p. 88–96, 2014. Citado 2 vezes nas páginas 29 e 98.

WADT, P. (Ed.). *Práticas de conservação do solo e recuperação de áreas degradadas*. [S.l.]: EMBRAPA ACRE, 2003. Citado na página 24.

WALKLEY, A.; BLACK, I. A. An examination of the degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil science*, LWW, v. 37, n. 1, p. 29–38, 1934. Citado na página 31.

WILLARD, L. L. *Chemical equilibria in soils*. New York: John Wiley & Sons, 1979. Citado na página 34.

WILLIAMS, K.; MUNKVOLD, J.; SORRELLS, M. Comparison of digital image analysis using elliptic fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (*triticum aestivum* l.). *Euphytica*, v. 190, n. 1, p. 99–116, 2013. Citado na página 27.

WOLD, S.; SJOSTROM, M. PLS-regression : a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, v. 58, p. 109–130, 2001. Citado 2 vezes nas páginas 43 e 47.

WOLF, B. An improved universal extracting solution and its use for diagnosing soil fertility. *Communications in Soil Science and Plant Analysis*, Taylor & Francis, v. 13, n. 12, p. 1005–1033, 1982. Citado na página 35.

YU, H.; MACGREGOR, J. F. Multivariate image analysis and regression for prediction of coating content and distribution in the production of snack foods. *Chemometrics and Intelligent Laboratory Systems*, Elsevier, v. 67, n. 2, p. 125–144, 2003. Citado na página 28.

Apêndice A

Granulometria das partículas de solo (g kg^{-1}), classificação textural e teores de COS das 177 amostras de solo analisadas pelos métodos de referência (método da Pipeta e Walkley-Black).

Amostra	Areia (%)	Argila (%)	Silte (%)	Classificação textural	COS (g kg^{-1})
1	28,9	72,5	0,0	Argilosa	19,1
2	33,2	67,5	0,0	Argilosa	17,3
3	32,6	67,6	0,0	Argilosa	17,3
4	33,9	67,0	0,0	Argilosa	17,3
5	27,9	68,7	3,4	Argilosa	14,3
6	27,3	70,3	2,3	Argilosa	12,0
7	27,9	67,3	4,8	Argilosa	13,1
8	35,0	58,9	6,1	Argilosa	18,3
9	35,8	60,5	3,7	Argilosa	17,1
10	37,3	62,3	0,4	Argilosa	18,9
11	28,2	70,9	0,9	Argilosa	13,7
12	29,5	69,5	1,0	Argilosa	17,1
13	28,9	68,7	2,4	Argilosa	11,4
14	27,5	70,2	2,2	Argilosa	11,5
15	32,0	66,8	1,1	Argilosa	16,6
16	31,3	67,8	0,9	Argilosa	13,5
17	29,2	69,1	1,7	Argilosa	11,7
18	39,6	57,3	3,0	Argilosa	14,1
19	41,1	54,6	4,2	Argilosa	29,9
20	41,1	54,8	4,0	Argilosa	36,7
21	37,7	59,0	3,3	Argilosa	22,3
22	37,4	58,4	4,1	Argilosa	34,9
23	40,4	57,5	2,0	Argilosa	21,7
24	40,2	57,5	2,3	Argilosa	25,7
25	40,5	57,2	2,3	Argilosa	32,2
26	40,0	61,0	0,0	Argilosa	21,8
27	41,0	58,5	0,5	Argilosa	29,9
28	40,2	57,7	2,2	Argilosa	22,8
29	40,4	59,6	0,0	Argilosa	23,2
30	38,0	59,4	2,5	Argilosa	19,3
31	42,3	55,7	2,0	Argilosa	26,6

Continua na próxima página

Amostra	Areia (%)	Argila (%)	Silte (%)	Classificação textural	COS (g kg ⁻¹)
32	41,5	56,7	1,8	Argilosa	22,3
33	41,2	56,9	1,9	Argilosa	25,0
34	40,3	57,8	1,9	Argilosa	31,6
35	23,1	80,5	0,0	Argilosa	14,7
36	60,3	37,8	1,9	Argilo arenosa	30,1
37	68,5	34,1	0,0	Franco argilo arenosa	15,9
38	65,9	35,9	0,0	Franco argilo arenosa	11,6
39	65,6	37,5	0,0	Franco argilo arenosa	18,5
40	67,6	35,5	0,0	Franco argilo arenosa	15,9
41	66,5	37,8	0,0	Franco argilo arenosa	12,8
42	73,6	29,7	0,0	Franco argilo arenosa	13,5
43	36,6	60,9	2,5	Argilosa	11,3
44	41,6	58,3	0,1	Argilosa	29,7
45	25,9	72,3	1,9	Argilosa	16,7
46	62,9	38,7	0,0	Argilo arenosa	11,7
47	74,7	23,5	1,7	Franco argilo arenosa	20,8
48	74,1	26,6	0,0	Franco argilo arenosa	21,5
49	43,2	56,3	0,5	Argilosa	35,8
50	70,4	30,5	0,0	Franco argilo arenosa	13,9
51	73,8	20,4	5,8	Franco argilo arenosa	22,5
52	34,1	59,3	6,6	Argilosa	34,2
53	33,7	58,7	7,6	Argilosa	37,6
54	22,0	81,5	0,0	Argilosa	11,8
55	35,5	56,6	7,9	Argilosa	29,6
56	36,2	56,5	7,3	Argilosa	39,7
57	84,2	7,4	8,4	Areia franca	20,4
58	85,1	7,6	7,3	Areia franca	6,7
59	99,8	0,6	0,0	Areia	3,9
60	2,0	77,9	20,0	Argilosa	3,9
61	73,2	26,8	0,0	Areia franca	9,0
62	64,7	30,5	4,8	Franco argilo arenosa	14,3
63	38,7	61,0	0,3	Argilosa	29,4
64	54,8	34,4	11,4	Franco argilo arenosa	10,1
65	36,9	60,6	2,5	Argilosa	34,7
66	57,4	37,4	5,2	Argilo arenosa	14,6
67	25,5	58,0	16,5	Argilosa	10,6
68	31,0	47,3	21,7	Argilosa	32,4

Continua na próxima página

Amostra	Areia (%)	Argila (%)	Silte (%)	Classificação textural	COS (g kg ⁻¹)
69	30,2	46,7	23,1	Argilosa	28,3
70	31,0	45,5	23,5	Argilosa	32,6
71	49,3	44,4	6,3	Argilo arenosa	20,7
72	43,1	44,9	12,0	Argilosa	26,2
73	36,8	43,1	20,1	Argilosa	28,4
74	28,2	47,1	24,7	Argilosa	19,0
75	27,5	49,2	23,2	Argilosa	21,0
76	29,1	50,4	20,6	Argilosa	21,0
77	23,9	47,9	28,2	Argilosa	19,6
78	27,8	50,6	21,6	Argilosa	20,9
79	29,4	49,4	21,3	Argilosa	21,3
80	27,8	53,6	18,6	Argilosa	22,5
81	92,7	4,0	3,3	Areia	11,2
82	91,8	6,0	2,2	Areia	6,4
83	91,2	6,0	2,8	Areia	4,6
84	90,1	8,0	1,9	Areia	3,6
85	90,0	8,0	2,0	Areia	2,6
86	88,4	8,0	3,6	Areia franca	2,8
87	96,7	2,0	1,3	Areia	23,8
88	97,9	1,0	1,1	Areia	9,4
89	96,7	1,0	2,3	Areia	3,4
90	97,0	1,0	2,0	Areia	1,5
91	97,2	1,0	1,8	Areia	0,8
92	98,0	1,0	1,0	Areia	0,6
93	98,0	1,0	1,0	Areia	0,8
94	88,1	8,1	3,8	Areia franca	15,0
95	88,5	8,0	3,5	Areia franca	7,9
96	85,1	10,1	4,8	Areia franca	6,5
97	85,2	12,1	2,7	Areia franca	4,8
98	84,2	12,1	3,7	Areia franca	3,6
99	84,3	12,1	3,6	Areia franca	3,0
100	79,4	16,2	4,4	Franco arenosa	21,4
101	78,9	16,2	4,9	Franco arenosa	13,1
102	76,5	20,2	3,3	Franco argilo arenosa	10,0
103	74,8	22,2	3,0	Franco argilo arenosa	6,1
104	74,5	22,2	3,3	Franco argilo arenosa	4,2
105	73,7	22,2	4,1	Franco argilo arenosa	3,4

Continua na próxima página

Amostra	Areia (%)	Argila (%)	Silte (%)	Classificação textural	COS (g kg ⁻¹)
106	62,0	20,1	17,9	Franco argilo arenosa	5,5
107	60,4	28,2	11,4	Franco argilo arenosa	4,1
108	66,5	28,2	5,3	Franco argilo arenosa	2,3
109	45,2	22,2	32,6	Franca	17,7
110	41,9	24,3	33,8	Franca	11,6
111	33,6	26,3	40,1	Franca	6,8
112	30,0	30,4	39,6	Franco argilosa	2,1
113	26,0	32,5	41,5	Franco argilosa	2,7
114	56,8	25,0	18,2	Franco argilo arenosa	20,7
115	55,7	27,0	17,3	Franco argilo arenosa	16,0
116	58,0	24,9	17,1	Franco argilo arenosa	7,8
117	54,2	29,1	16,7	Franco argilo arenosa	3,7
118	50,9	31,3	17,8	Franco argilo arenosa	3,4
119	56,6	27,0	16,4	Franco argilo arenosa	3,1
120	44,2	29,1	26,7	Franco argilosa	3,2
121	49,7	18,4	31,9	Franca	1,9
122	46,4	22,8	30,8	Franca	1,7
123	62,4	30,2	7,4	Franco argilo arenosa	6,5
124	61,8	32,2	6,0	Franco argilo arenosa	4,8
125	59,2	34,3	6,5	Franco argilo arenosa	3,4
126	59,0	34,3	6,7	Franco argilo arenosa	2,3
127	47,8	46,4	5,8	Argilo arenosa	1,8
128	49,0	46,4	4,6	Argilo arenosa	1,2
129	75,6	14,1	10,3	Franco arenosa	19,4
130	75,5	16,1	8,4	Franco arenosa	13,7
131	65,1	26,3	8,6	Franco argilo arenosa	7,7
132	55,1	34,4	10,5	Franco argilo arenosa	5,3
133	50,7	38,5	10,8	Argilo arenosa	4,1
134	47,5	40,5	12,0	Argilo arenosa	2,3
135	48,0	40,8	11,2	Argilo arenosa	2,1
136	49,0	40,6	10,4	Argilo arenosa	1,5
137	15,1	48,8	36,1	Argilosa	23,5
138	22,8	46,8	30,4	Argilosa	19,4
139	22,0	48,8	29,2	Argilosa	14,8
140	19,5	50,8	29,7	Argilosa	8,3
141	21,0	50,9	28,1	Argilosa	8,5
142	18,5	52,8	28,7	Argilosa	5,6

Continua na próxima página

Amostra	Areia (%)	Argila (%)	Silte (%)	Classificação textural	COS (g kg ⁻¹)
143	26,5	41,1	32,4	Argilosa	25,7
144	25,8	42,7	31,5	Argilosa	15,8
145	30,7	42,7	26,6	Argilosa	10,3
146	21,3	45,1	33,6	Argilosa	6,9
147	26,0	44,6	29,4	Argilosa	5,1
148	25,0	44,6	30,4	Argilosa	5,0
149	25,3	46,6	28,1	Argilosa	2,5
150	50,3	37,1	12,6	Argilo arenosa	47,6
151	62,9	30,4	6,6	Franco argilo arenosa	26,3
152	76,2	19,8	3,9	Franco arenosa	15,0
153	89,6	8,0	2,4	Areia	9,5
154	91,6	7,1	1,4	Areia	5,8
155	91,6	8,5	0,0	Areia	4,2
156	84,3	11,9	3,8	Areia franca	5,5
157	70,4	27,0	2,6	Franco argilo arenosa	3,7
158	2,0	59,9	38,1	Argilosa	18,8
159	0,8	68,3	30,9	Argilosa	8,5
160	3,3	68,2	28,5	Argilosa	7,3
161	4,7	57,7	37,6	Argilosa	23,3
162	0,8	76,4	22,8	Argilosa	7,0
163	1,8	76,3	21,9	Argilosa	1,7
164	18,0	76,0	6,0	Argilosa	1,0
165	14,7	57,0	28,3	Argilosa	30,6
166	16,3	57,0	26,7	Argilosa	18,9
167	16,5	59,0	24,5	Argilosa	16,0
168	16,3	61,0	22,7	Argilosa	11,4
169	16,1	61,0	22,9	Argilosa	8,6
170	15,9	61,0	23,1	Argilosa	6,4
171	16,6	57,6	25,8	Argilosa	29,8
172	16,0	59,5	24,5	Argilosa	20,4
173	13,6	65,7	20,7	Argilosa	15,3
174	12,4	74,1	13,5	Argilosa	12,4
175	9,0	78,1	12,9	Argilosa	7,4
176	8,6	78,3	13,1	Argilosa	4,7
177	10,9	72,0	17,1	Argilosa	3,7