



UNIVERSIDADE
ESTADUAL de LONDRINA

YURI RENAN BOVOLENTA

**PREDIÇÃO DOS ATRIBUTOS DO SOLO A PARTIR DE
ANÁLISE ESTATÍSTICA MULTIVARIADA BASEADA EM
ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO (NIRS)
E IMAGEM DIGITAL**

Londrina
2022

YURI RENAN BOVOLENTA

**PREDIÇÃO DOS ATRIBUTOS DO SOLO A PARTIR DE
ANÁLISE ESTATÍSTICA MULTIVARIADA BASEADA EM
ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO (NIRS)
E IMAGEM DIGITAL**

Tese apresentada ao Programa de Doutorado
Associado em Química UEL-UEPG-
UNICENTRO, como requisito parcial à obtenção
do título de Doutor em Química.

Orientadora: Profa. Dra. Suzana Lucy Nixdorf

Londrina
2022

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

Bovolenta, Yuri Renan.

Predição dos atributos do solo a partir de análise estatística multivariada baseada em espectroscopia de infravermelho próximo (nirs) e imagem digital / Yuri Renan Bovolenta. - Londrina, 2022.
107 f. : il.

Orientador: Suzana Lucy Nixdorf.

Tese (Doutorado em Química) - Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em Química, 2022.
Inclui bibliografia.

1. Análise multivariada - Tese. 2. Análise de solo - Tese. 3. Espectroscopia do infravermelho próximo - Tese. 4. Imagem Digital - Tese. I. Nixdorf, Suzana Lucy. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Química. III. Título.

CDU 54

YURI RENAN BOVOLENTA

**PREDIÇÃO DOS ATRIBUTOS DO SOLO A PARTIR DE
ANÁLISE ESTATÍSTICA MULTIVARIADA BASEADA EM
ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO (NIRS)
E IMAGEM DIGITAL**

Tese apresentada ao Programa de Doutorado
Associado em Química UEL-UEPG-
UNICENTRO, como requisito parcial à obtenção
do título de Doutor em Química.

BANCA EXAMINADORA

Orientadora: Prof.Dr^a Suzana Lucy Nixdorf
Universidade Estadual de Londrina - UEL

Prof.Dr^a Maria Josefa Santos Yabe
Universidade Estadual de Londrina - UEL

Prof. Dr^o Mario Henrique Montazzolli Killner
Universidade Estadual de Londrina - UEL

Prof. Dr^o Pedro Rodolfo Siqueira Vendrame
Universidade Estadual de Londrina - UEL

Prof. Graziela Moraes de Cesare Barbosa
Instituto de Desenvolvimento Rural do Paraná –
IDR-Paraná

Londrina, 26 de Maio de 2022

Dedico este trabalho a minha família, especialmente a meus pais Marcia e Valdir, meu irmão Yves, e minha esposa Elisiane, que sempre me apoiaram em todas as minhas escolhas, e a todos aqueles que estiverem ao meu lado durante a realização deste trabalho.

Agradecimentos

Agradeço à Profa. Dra. Suzana Lucy Nixdorf não só pela constante orientação neste trabalho, mas, sobretudo pela sua amizade e todos os ensinamentos durante este período, que possamos levar essa amizade e companheirismo para toda a vida.

Agradeço à minha família, em especial a meus pais, Marcia e Valdir, meu irmão Yves Rafael e meu sobrinho Ravi, por todo o amor, apoio e força que me deram para chegar até aqui, sem os esforços de vocês nada disso seria possível, serei eternamente grato por tudo que representam para mim.

Agradeço também à minha eterna companheira e esposa Elisiane pela paciência, força e por todo amor e atenção. Seu apoio e compreensão foram fundamentais para a conclusão desta tese de Doutorado.

Agradeço aos companheiros de laboratório, e amigos, em especial Tiago, Lycio e Alexandre e por toda a troca de conhecimentos, ajuda nas análises durante finais de semana, no período noturno e férias, e total apoio mesmo quando pensei em desistir. Agradeço por todos os momentos de alegria e companheirismo.

Agradeço à Universidade Estadual de Londrina pela oportunidade de desenvolvimento acadêmico e ao Departamento de Química por todo o amparo e estrutura, e a todos os docentes do Programa de Pós Graduação por todo o conhecimento e experiências adquiridos. Agradeço também aos órgãos de fomento, CNPq e FINEP pela estrutura relacionada à equipamentos fundamentais para conclusão desta tese.

“Um cientista no seu laboratório não é apenas um técnico: é, também, um menino colocado à frente de fenômenos naturais que impressionam como se de um conto de fadas se tratasse” (Marie Curie)

BOVOLENTA, Yuri Renan. **Predição dos atributos do solo a partir de análise estatística multivariada baseada em espectroscopia de infravermelho próximo (NIRS) e Imagem Digital**. 2022. 107 f. Tese (Doutorado em Química) – Centro de Ciências Exatas, Universidade Estadual de Londrina, Londrina, 2022.

RESUMO

Uma produtividade adequada depende da qualidade do solo para o desenvolvimento das culturas, que influencia e agrega valor ao agronegócio, responsável por 27,4% do Produto Interno Bruto brasileiro em 2021. Entretanto, as análises do solo, ainda são realizadas por métodos clássicos, que demandam demasiado tempo e utilizam grande quantidade de reagentes. Desta forma, este estudo objetiva avaliar o desempenho de métodos rápidos alternativos ao emprego de métodos primários de referência. Para isso, investigou-se a capacidade da espectroscopia no visível e infravermelho próximo (Vis-NIRS) associada a quimiometria, em efetuar a predição dos principais parâmetros analíticos do solo. A escolha desta técnica foi feita por não exigir muitas etapas no processo de preparo da amostra, de forma a preservar as características físicas e químicas das amostras, eliminar o uso de reagentes e agilizar as determinações rotineiras, cuja demanda é crescente nos laboratórios de solo. Para isso, buscou-se inicialmente determinar os valores dos principais parâmetros analíticos por suas respectivas técnicas primárias. Em seguida, os espectros Vis-NIR foram obtidos em um equipamento de bancada, pelo uso de 2,5 g de terra fina seca ao ar, disposta em cela de quartzo do tipo *spinning*. Os dados de toda a faixa espectral (400-2500 nm) das 227 amostras de solo foram utilizados para a criação de um banco de dados. Efetuou-se a segunda derivada como processamento dos dados dos espectros Vis-NIR. Inicialmente, aplicou-se a análise de componentes principais (PCA) aos resultados das técnicas primárias e do Vis-NIR, para avaliar o potencial de separar em grupos, levando em conta sua classificação de textura (areia, silte e argila). Em seguida, criou-se três modelos estatísticos empregando-se: regressão por componentes principais (PCR); regressão por mínimos quadrados parciais (PLS) e regressão por mínimos quadrados parciais modificados (PLSM), para verificar os níveis de correlação e predição de matéria orgânica (MO), granulometria, soma de bases (SB), pH, acidez (H+Al), P, S, Ca, Mg e Saturação por bases (V%). Houve correlação entre os resultados preditos pelos modelos estatísticos PLS, PCR e PLMS validados e os resultados das técnicas primárias, com coeficientes de determinação linear ($R^2 > 0,95$) para silte, areia e argila. Não houve diferença significativa em nível de 5%, quando comparados os coeficientes de variação de resultados primários, obtidos em ensaios de proficiência de laboratórios de referência em solo. Conclui-se que, a aplicação Vis-NIR é viável e contribui na análise dos principais parâmetros analíticos do solo, considerando a simplicidade de operação, redução de mão-de-obra, maior frequência analítica e menor quantidade de passivo ambiental gerado pela redução de resíduos em laboratórios de análises químicas e agronômicas.

Palavras-chave: quimiometria; métodos; modelos de regressão; predição.

BOVOLENTA, Yuri Renan. **Prediction of soil attributes from multivariate statistical analysis based on Near Infrared Spectroscopy (NIRS) and Digital Imaging**. 2022. 107 p. Doctoral thesis (PhD in Chemistry) - State University of Londrina, Londrina, 2022.

ABSTRACT

An adequate productivity depends on the quality of the soil for the development of crops, which influences and adds value to agribusiness, responsible for 27.4% of the Brazilian Gross Domestic Product in 2021. However, soil analyzes are still carried out by classical methods, that require too much time and use a large amount of reagents. Thus, this study aims to evaluate the performance of rapid alternative methods to the use of primary reference methods. For this, the ability of visible and near infrared spectroscopy (Vis-NIRS) associated with chemometrics to predict the main analytical parameters of the soil was investigated. This technique was chosen because it does not require sample preparation in order to preserve the physical and chemical characteristics of the samples, eliminate the use of reagents and speed up routine determinations whose demand is increasing in soil laboratories. For this, it was initially sought to determine the values of the main analytical parameters by their respective primary techniques. Then, the Vis-NIR spectra were obtained in a benchtop equipment, using 2.5 g of air-dried fine earth placed in a quartz spinning cell. Data from the entire spectral range (400-2500 nm) of the 227 soil samples were used to create a database. The second derivative was performed as data processing of the Vis-NIR spectra. Initially, principal component analysis (PCA) was applied to the results of the primary techniques and the Vis-NIR to evaluate the potential to separate into groups, taking into account their texture classification (sand, silt and clay). Then three statistical models were created using: regression by principal components (PCR); partial least squares regression (PLS) and modified partial least squares regression (PLSM), to verify the correlation and prediction levels of organic matter (OM), granulometry, sum of bases (SB), pH, acidity (H+Al), P, S, Ca, Mg and Base saturation (V%). There was a correlation between the results predicted by the validated PLS, PCR and PLMS statistical models and the results of the primary techniques, with coefficients of linear determination ($R^2 > 0.95$) for silt, sand and clay. There was no significant difference at the 5% level when comparing the coefficients of variation of primary results, obtained in proficiency tests of reference laboratories in soil. It is concluded that the Vis-NIR application is viable and contributes to the analysis of the main analytical parameters of the soil, considering the simplicity of operation, reduction of manpower, greater analytical frequency and less amount of environmental liability generated by the reduction of residues in chemical and agronomic analysis laboratories.

Key words: chemometrics; methods; regression models; prediction.

LISTA DE FIGURAS

Figura 1 - Esquema da instrumentação básica do NIR por refletância difusa ..	28
Figura 2 - Projeção das amostras no plano bidimensional das componentes principais 1 e 2 (gráfico de scores) referentes aos dados obtidos pelas técnicas primárias de 227 amostras de solo analisadas.....	39
Figura 3 - Projeção das amostras no plano bidimensional das componentes principais 1 e 2 (gráfico de scores) referente aos dados espectrais das 227 amostras de solo com relação às suas refletâncias difusas entre os comprimentos de onda de 400 a 2400 nm	41
Figura 4 - Espectros NIR de 227 amostras de solo	42
Figura 5 - Segunda derivada dos espectros NIR das 227 amostras de solo com diferentes parâmetros para macro nutrientes e textura com regiões de alta variação espectral em destaque	43
Figura 6 - Espectro NIR das amostras (A0 até A5) com adição de padrão de cátions metálicos.....	46
Figura 7 - Bandas do Espectro NIR de 1090 a 1105 nm das amostras (A0 até A5) com adição de padrão de cátions metálicos	46
Figura 8 - Bandas do Espectro NIR de 1350 a 1450 nm das amostras (A0 até A5) com adição de padrão de cátions metálicos	47
Figura 9 - Bandas do Espectro NIR de 2175 a 2235 nm das amostras (A0 até A5) com adição de padrão de cátions metálicos	48
Figura 10 - Comparação dos valores observados pelas técnicas primárias e preditos de Areia nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método. 50	51
Figura 11 - Comparação dos valores observados pelas técnicas primárias e preditos de Silte nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.....	51
Figura 12 - Comparação dos valores observados pelas técnicas primárias e preditos de Argila nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método	52

CAPÍTULO 2

Figura 1 - Representação de formas em diferentes resoluções	68
--	----

Figura 2 - Comprimento de onda do espectro visível e suas cores	69
Figura 3 - Cores complementares no disco de cores	70
Figura 4 - Representação dos modelos RGB e CMYK.....	71
Figura 5 - Representação do modelo HSV	72
Figura 6 - Representação do Modelo HSL	74
Figura 7 - Representação da câmara fotográfica utilizada na obtenção de imagens digitais. (a) projeto do sistema instrumental. (b) sistema instrumental montado.....	76
Figura 8 - Interface do programa de aquisição de dados de imagens digitais baseado no Sistema RGB tratadas pelo software lab-made ChemColour para smartphone com conversão para diferentes Sistemas de cores aplicada na determinação dos principais parâmetros de solo.....	77
Figura 9 - Projeção das amostras no plano bidimensional das componentes principais 1 e 2 (gráfico de scores) referentes aos dados das imagens digitais das amostras de solo, capturadas pelo smartphone e tratadas no Sistema DIA-ChemColour	78
Figura 10 - Comparação dos valores observados e preditos de matéria orgânica (MO) nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema DIA-ChemColour	82
Figura 11 - Comparação dos valores observados e preditos de areia nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema DIA-ChemColour	83
Figura 12 - Comparação dos valores observados e preditos de silte nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema DIA-ChemColour	84
Figura 13 - Comparação dos valores observados e preditos de argila nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema DIA-ChemColour	86
Figura 14 - Comparação dos valores observados e preditos de classificação de tipo de solo nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema DIA-ChemColour	87

LISTA DE TABELAS

Tabela 1 - Classificação baseada nos parâmetros físicos do solo	25
Tabela 2 - Variáveis selecionadas com maior variância entre os valores de refletância difusa nos espectros NIR das 227 amostras de solo analisadas	44
Tabela 3 - Modelo estatístico de regressão por componentes principais (PCR) com suas variações para os principais parâmetros do solo aplicado a 227 amostras de solo analisadas por NIR em duplicata	49
Tabela 4 - Modelo estatístico de regressão dos mínimos quadrados parciais (PLS) com suas variações para os principais parâmetros do solo aplicado a 227 amostras de solo analisadas por NIR em duplicata	49
Tabela 5 - Modelo estatístico de regressão dos mínimos quadrados parciais modificado (PLSM) das 227 amostras de solo em duplicata	50
Tabela 6 - Coeficiente de Variação (CV(%)) médio dos resultados amostrais obtidos dos modelos estatísticos Vis-NIRS comparados aos CVs dos ensaios de proficiência laboratoriais de análises agronômicas disponíveis no Brasil, obtidos por meio de técnicas primárias e ANOVA.....	53
Tabela 7 - Reagentes e tempo médio gasto para análise de 50 amostras	55

CAPITULO 2

Tabela 1 - Modelo estatístico de regressão por componentes principais (PCR) das 200 amostras de solo em duplicata utilizando o Sistema DIA-ChemColour	79
Tabela 2 - Modelo estatístico de regressão por mínimos quadrados parciais (PLS) das 200 amostras de solo em duplicata utilizando o Sistema DIA-ChemColour	80
Tabela 3 - Modelo estatístico de regressão por mínimos quadrados parciais modificado (PLSM) das 200 amostras de solo em duplicata utilizando o Sistema DIA-ChemColour	81

Tabela 4 - Coeficiente de Variação (CV(%)) médio dos resultados amostrais obtidos dos modelos estatísticos RGB utilizando o Sistema DIA-ChemColour comparados aos CVs dos ensaios de proficiência laboratoriais de análises agronômicas disponíveis no Brasil, obtidos por meio de técnicas primárias	89
---	----

LISTA DE ABREVIATURAS

EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
CEPEA	Centro de Estudos Avançados em Economia Aplicada
CNA	Confederação da Agricultura e Pecuária do Brasil
PIB	Produto Interno Bruto
ATP	Trifosfato de Adenosina
NIR	Infravermelho Próximo
MIR	Infravermelho médio
FIR	Infravermelho Distante
PCA	Análise de Componentes Principais
PC	Componentes Principais
PCR	Regressão por Componentes Principais
MRL	Regressão Linear Múltipla
PLS	Regressão por Mínimos Quadrados Parciais
NIRS	Espectroscopia de Infravermelho Próximo
PLSM	Regressão por Mínimos Quadrados Parciais Modificada
MP-AES	Espectrômetro de Emissão Atômica com Plasma Induzido por Micro-ondas
UV-Vis	Ultra violeta - Visível
M.O	Matéria Orgânica
H+Al	Acidez total do solo
SB	Soma de Bases
V%	Saturação por Bases
LAPA	Laboratório de Apoio a Pesquisa Agropecuária
SEC	Erro padrão de calibração
SECV	Erro padrão de validação cruzada
1-VR	Coefficiente de determinação para validação cruzada
R ²	Coefficiente de determinação
IAC	Instituto Agrônomo de Campinas
PAQLF	Programa de Análise de Qualidade de Laboratórios de Fertilidade
CCD	Dispositivos de Cargas Acopladas
DPI	Pontos por Polegada
CMOS	Semicondutores de Óxido Metálico Complementar

RGB	Vermelho, Verde e Azul
CMYK	Ciano, Magenta, Amarelo e Preto
HSV	Matiz, Saturação e Valor
HSL	Matiz, Saturação e Luminosidade
RMSEP	Erro Quadrático Médio de Predição

SUMÁRIO

	CAPÍTULO 1 - PREDIÇÃO DOS ATRIBUTOS DO SOLO A PARTIR DE ANÁLISE ESTATÍSTICA MULTIVARIADA BASEADA EM ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO (NIRS).....	16
1	INTRODUÇÃO	16
1.1	IMPORTÂNCIA DA AGRICULTURA	17
1.2	ANÁLISE DO SOLO	18
1.3	ESPECTROSCOPIA NO INFRAVERMELHO	21
1.4	QUIMIOMETRIA.....	23
1.4.1	Análise Multivariada.....	24
1.4.2	Análise de Componentes Principais	24
1.4.3	PCR	26
1.4.4	PLS.....	27
2	MATERIAL E MÉTODOS.....	29
2.1	AMOSTRAS	29
2.2	ANÁLISE DE MACRO NUTRIENTES	29
2.2.1	Análise de Fósforo (P)	30
2.2.2	Análise de Enxofre.....	30
2.2.3	Análise de Matéria Orgânica	31
2.2.4	Análise de pH e H+Al	31
2.2.5	Determinação de Granulometria.....	32
2.2.6	Soma de Bases (SB)	33
2.2.7	Saturação por Bases (V%)	33
2.2.8	Espectroscopia no Infravermelho Próximo (NIR).....	33
3	RESULTADOS.....	34
3.1	NIR	35
	REFERENCIAS.....	52

	CAPITULO 2 - DETERMINAÇÃO DE PARÂMETROS DO SOLO POR ANÁLISE DE IMAGEM DIGITAL	59
1	INTRODUÇÃO	62
1.1	IMAGEM DIGITAL.....	62
1.2	TEORIA DAS CORES.....	64
1.3	MODELO RGB E MODELO CMYK.....	65
1.4	MODELO HSV E MODELO HSL	67
2	MATERIAL E MÉTODOS	70
3	RESULTADOS	73
4	CONCLUSÃO	85
	REFERÊNCIAS	86
	APÊNDICE	88

Capítulo 1

Predição dos atributos do solo a partir de análise estatística multivariada baseada em espectroscopia de infravermelho próximo (NIRS)

1 INTRODUÇÃO

1.1 Importância da Agricultura

Nas últimas cinco décadas, o Brasil deixou de ser importador de alimentos para se tornar um dos mais importantes produtores e exportadores mundiais de alimentos (EMBRAPA, 2018). Estudos realizados pelo Centro de Estudos Avançados em Economia Aplicada (CEPEA, 2021) juntamente com a Confederação da Agricultura e Pecuária do Brasil (CNA) projetaram que em 2021 o agronegócio teria uma participação de 28% no Produto Interno Bruto brasileiro (PIB), um aumento de 7,5% em relação a 2019, um grande avanço considerando que em 1970 o agro contribuía com apenas 7,5% do PIB (FIELDVIEW, 2021). Com o objetivo de abastecer a crescente população urbana e reduzir os custos dos alimentos, o governo estimulou o desenvolvimento de novas políticas públicas, que visavam aumentar a produção e a produtividade agrícola, além de fomentar pesquisas na área e fornecer crédito rural (CHADDAD, 2016).

Em 2017 o setor do agronegócio contava com 19 milhões de pessoas empregadas, dentre elas quase a metade ocupada diretamente no campo, gerando um saldo de exportações de 71 bilhões de dólares, movimentando também o setor da agroindústria e de insumos agrícolas (BARROS, 2017).

As grandes culturas são as responsáveis pela maior parte do agronegócio brasileiro. Na safra de 2020/21 cerca de 68,9 milhões de hectares foram utilizados apenas no cultivo de soja e milho no País e a previsão é de um aumento de 3,6% para a safra de 2021/22, chegando a uma estimativa de produção de 288,61 milhões de toneladas de grãos (CONAB, 2021A). Sendo

o Brasil o maior produtor mundial de soja (135,4 milhões de toneladas) seguido pelos Estados Unidos (112,5 milhões de toneladas) (CONAB, 2021B).

1.2 Análise do Solo

A análise de solo é o primeiro requisito para atingir bons resultados em qualquer cultura. Através do resultado é que são prescritas as quantidades de corretivos e fertilizantes que devem ser aplicados para se atingir uma produção satisfatória e eficaz. Apesar da análise de solo já ser realizada desde o início do século passado, somente após a década de 60 é que no Brasil sua evolução se tornou mais expressiva com o uso de fertilizantes minerais (RAIJ, 2001).

A evolução no número de análises de solo foi expressiva. Segundo CANTARELLA et al. (1995) nos primeiros 10 anos (entre 1970 e 1980) o número de análises aumentou de 200 mil para 400 mil e, em 1989 chegou a mais de 700 mil. Atualmente, não há informações estatísticas oficiais sobre o número de análises anuais no País, mas o valor estimado é de mais de 4 milhões de amostras por ano.

O monitoramento dos atributos do solo é de extrema importância na agricultura porque garantem produções cada vez maiores nas mais diversas culturas. A matéria orgânica (que é formada por resíduos vegetais e animais) por exemplo é responsável por características químicas e físicas do solo, sendo um dos constituintes primário da coloração do mesmo, podendo influenciar até na absorção de calor, mantendo o solo com temperaturas mais elevadas (DEMATTE e GARCIA, 1999; DALMOLIN, 2002).

Já o cálcio e o magnésio têm importância na neutralização da acidez do solo, devido ao seu efeito de aumento de pH e a formação de óxidos básicos,

que auxiliam na diminuição do alumínio tóxico trocável, esses favorecem ainda no crescimento das raízes de plantas, aumentando a atividade de microrganismos benéficos as culturas, participando de diversas funções enzimáticas e disponibilizando diversos outros nutrientes. A deficiência de magnésio nas plantas pode acarretar a redução dos níveis de clorofila, comprometendo a realização da fotossíntese e o desenvolvimento da planta. (RAIJ et al., 1982; PAVAN et al., 1984)

Outro parâmetro do solo comumente monitorado é a acidez. Essa é uma das principais responsáveis pela redução na produtividade de solos tropicais (SOUSA & LOBATO, 2004). A acidez do solo é fracionada em três partes: a acidez ativa, que é atribuída pela atividade dos íons H^+ em solução aquosa e determinada pelo pH; a acidez trocável, que corresponde à presença de íons Al^{3+} disponíveis; e a acidez potencial, que corresponde à soma dos íons H^+ com a acidez trocável, medida pela mudança de pH em presença de uma solução tampão (EBELING et al., 2008). O fósforo é o segundo parâmetro que mais limita a produtividade dos solos. Apesar de muitas vezes estar presente em abundância, se apresenta ligado a compostos de alta energia de ligação, estando apenas uma pequena fração com baixa energia de ligação disponível para as plantas (SCHULZE, 1989). Sua disponibilidade é quantificada por diversos métodos. O método mais utilizado é o de Mehlich-1, que se dá pelo uso de solução extratora, o qual estima a quantidade deste elemento adsorvido com baixa energia, e pode ser absorvido pelas plantas (WALKER e SYERS, 1976; CROSS e SCHILESINGER, 1995). Nas plantas o fósforo é imprescindível para a formação do trifosfato de adenosina (ATP) que é responsável por fornecer

energia para sistemas vitais como a fotossíntese, divisão celular e carga genética (NETO et al. 2010).

Os níveis de cátions como Ca^{2+} , Mg^{2+} , K^+ , H^+ e Al^{3+} presentes no solo são representados pela Saturação por Bases (V%), baixos teores de V% podem indicar que há pequenas quantidades de Ca^{2+} , Mg^{2+} e K^+ presentes no solo, neste caso provavelmente o solo será ácido e poderá apresentar baixa produtividade. A maioria das culturas apresentam bons níveis de produtividade quando o solo apresenta entre 50 e 80% de V% (EMBRAPA, 2010).

Para classificação dos parâmetros físicos do solo é realizada a classificação dos solos em três tipos de acordo com a quantidade de areia, argila e silte presentes, o Ministério da Agricultura (MAPA, 2008) determina a classificação conforme a Tabela 01.

Tabela 01. Classificação baseada nos parâmetros físicos do solo.

	Areia (%)	Argila (%)
Tipo 1	>15	10-15
Tipo 2	15-70	16-35
Tipo 3	<15	>35*

*solos com menos de 35% de argila também podem ser enquadrados como Tipo 3 desde que apresentem menos de 15% de areia.

Solos tipo 1 e 2 são considerados arenosos e solos tipo 3 são considerados argilosos.

Os métodos de análise de solos empregados na atualidade são bastante difundidos, porém demandam demasiado tempo para execução das análises e utilizam uma série de metais potencialmente tóxicos. Por isso, se buscam novas técnicas que sejam mais rápidas e que não produzam tantos resíduos. Estudos

como o de Zhao (2020) e Coblinski (2021), buscam aplicar técnicas de Vis-NIR em conjunto com outras técnicas para quantificação da matéria orgânica, classificação pela textura do solo e a presença de minerais, demonstrando que técnicas não destrutivas podem auxiliar nas análises de solos.

1.3. Espectroscopia no Infravermelho

Com comprimento de onda localizado entre 780 e 100.000 nm o espectro infravermelho pode ser dividido em três faixas - Infravermelho Próximo (NIR) que compreende a região de 780 a 2.500 nm; o Infravermelho médio (MIR) que compreende a região de 2.500 a 5.000 nm; e o Infravermelho Distante (FIR) que compreende a região de 5.000 a 100.000 nm (SKOOG, 2002). Os limites de definição das faixas são comumente divergentes em diversas literaturas (PASQUINI, 2003; SOUZA, 2005).

As regiões MIR e FIR foram exploradas durante muitos anos no controle de qualidade da produção de materiais como: combustíveis, polímeros, lubrificantes e compostos orgânicos em geral, podendo ser ainda utilizados no estudo de gases, que apresentam momentos de dipolo permanentes (VALDERRAMA, 2005; NETO, 2005). A radiação infravermelha pode causar alterações nos modos vibracionais e rotacionais das moléculas. O campo elétrico da radiação infravermelha ao ser alternado interage com as moléculas, ocasionando alterações nos seus movimentos (BARBOSA, 2008; SKOOG, 2002). As vibrações podem ser classificadas em estiramento e deformação. A vibração de estiramento é a variação na distância entre os átomos ao longo do

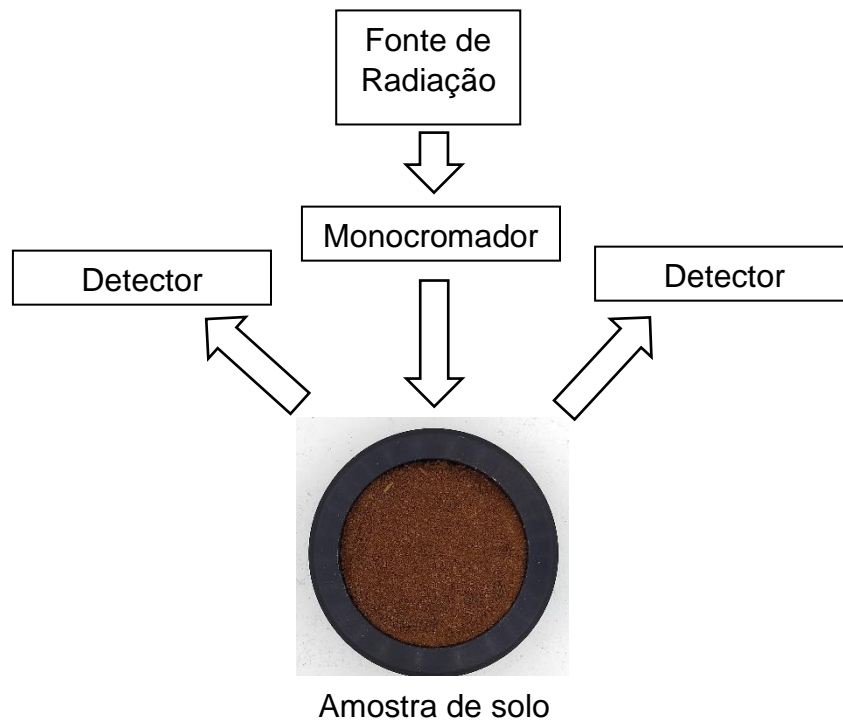
eixo de ligação de dois átomos; já a deformação, envolve uma variação no ângulo entre duas ligações (SKOOG, 2002).

A região do NIR recebeu mais atenção apenas na década de 70 na quantificação e identificação de moléculas orgânicas. Desde então, a espectroscopia no infravermelho próximo vem ganhando mais espaço (BURNS, 2001; BORIN, 2003). Nessa região, as bandas de absorção são de combinações de vibrações de estiramento e sobretons, que envolvem grupos com ligações polarizadas como C-H, N-H, O-H, S-H (PASQUINI, 2003). As bandas de absorção de NIR são muito fracas quando comparadas com as do infravermelho médio (de 10 a 1.000 vezes menores), o que seria uma desvantagem em relação ao MIR. No entanto, com o desenvolvimento de detectores de alta eficiência essa dificuldade foi superada (HONORATO, 2006; LIMA, 2009). O espectro do infravermelho próximo pode ser relacionado com parâmetros físicos, sendo capaz de distinguir amostras, de acordo com a sua textura superficial. As interações da radiação eletromagnética NIR podem resultar em diversos fenômenos como a reflexão, transmissão, refração e difração (COSTA FILHO, 2003).

O equipamento que gera os espectros NIR (Figura 1) é construído a partir de cinco elementos: uma fonte de radiação, um monocromador, um porta amostra, um detector de radiação e um dispositivo capaz de ler a resposta do detector. As fontes de radiação em geral são lâmpadas de halogênio e filamento de tungstênio. Os detectores mais utilizados, que apresentam alta taxa de detecção e velocidade de varredura são fabricados a partir de sulfeto de chumbo. O modo de aquisição dos espectros NIR pode variar de acordo com o estado físico das amostras, sendo o mais indicado para amostras de solo o modo de

reflectância difusa, no qual a radiação refletida ou espalhada pela superfície da amostra é medida (PASQUINI, 2003).

Figura 1. Esquema da instrumentação básica do NIR por refletância difusa



Fonte: Próprio autor (2022).

1.4 Quimiometria

A aplicação de métodos estatísticos, matemáticos e computacionais utilizados para planejar, investigar e fazer previsões é chamada de quimiometria (WOLD, 1995). A química analítica é a área mais beneficiada no desenvolvimento e estudo da quimiometria, associada a técnicas como o infravermelho, a espectroscopia de massas, a cromatografia e a ressonância magnética entre uma infinidade de técnicas, que oferecem um conjunto de dados caracterizados como multivariados (BRUNS, 1985). Com o surgimento da quimiometria na

década de 70, o processamento de um grande volume de dados de amostras com muitas variáveis, se tornou muito mais rápida, possibilitando um tratamento muito mais eficiente (NETO et al. 2006). A espectroscopia NIR é uma das técnicas dependentes da quimiometria, a qual auxilia na identificação e quantificação de diversos parâmetros em diferentes matrizes, uma vez que os espectros gerados carregam grande quantidade de informações, possibilitando desenvolvimento e sua aplicação em novos métodos (MCCLURE, 1994).

A escolha do método quimiométrico é feita com base na natureza do problema que se deseja resolver e do tipo de informação que se deseja obter. Os métodos que formam a quimiometria podem ser classificados em: reconhecimento de padrões, planejamento de experimentos, calibração multivariada e resolução multivariada de sinais. Para utilização desses métodos se faz necessária a correta interpretação dos dados, seleção de variáveis, remoção de *outliers* e pré-processamento das respostas instrumentais.

1.4.1 *Análise Multivariada*

1.4.2 *Análise de componentes principais*

O método de PCA (análise de componentes principais – do inglês *principal component analysis*) é um dos modelos matemáticos mais utilizados na análise de dados, permitindo a interpretação de um conjunto de dados multivariados, buscando identificar uma correlação entre as diversas variáveis, facilitando a interpretação do comportamento das amostras (FERREIRA, 2002; BEEBE et al., 1998; SABIN et al., 2004). Essa análise dos dados é utilizada quando a quantidade de dados que representam o sistema é muito grande, dificultando assim o tratamento deles, considerando-se que muitos desses dados podem não

ser relevantes para a análise quimiométrica. A análise por PCA reduz a quantidade de dados, deixando apenas os dados realmente úteis para a construção do modelo do sistema, eliminando desta forma uma boa parte dos ruídos experimentais, que não estão correlacionados com as informações da matriz de dados (MELLO, 1998; BARTHUS, 1999).

Para reduzir o número de dados, sem perder informações importantes relativas à perda de explicação da variância amostral do sistema, o modelo cria variáveis ortogonais entre si direcionadas em função da distribuição espacial dos dados, que são chamadas de Componentes Principais (PC). O método de PCA reduz os dados obtidos a partir dos dados iniciais de forma que continuem a representar as principais características (REIS, 1997; NAGATA et al., 2001). Segundo BEEBE et al. (1998) e SOUZA (2012), na PCA a matriz X é decomposta em um produto de duas matrizes (Equação 1).

$$X = TP^t + E \quad \text{Eq. (1)}$$

em que, t : scores; P : Pesos; E : Resíduos

Os scores são as coordenadas da amostra no sistema, formadas pelas componentes principais, sendo uma combinação linear das variáveis originais gerada para cada componente principal. Os pesos (*loadings*) são os cossenos dos ângulos entre as variáveis originais e as componentes principais. Basicamente os scores representam as relações de similaridade entre as amostras e os pesos a contribuição de cada variável na formação das componentes principais (PC) (WOLD, 1987; KAMAL, 1997).

1.4.3 PCR

A análise de regressão envolve a relação entre uma variável dependente e outras variáveis chamadas independentes, sendo representada por um modelo matemático chamado de modelo de regressão linear. Pelo fato do modelo de regressão múltipla (MRL) apresentar falhas, foi criado o modelo de regressão por componentes principais (PCR). Neste modelo, cada componente principal descreve uma fração da variação total contida nos dados e o modelo determina as variáveis independentes a serem incluídas na regressão. Desta forma é estabelecida uma relação linear entre os *scores* da matriz das variáveis e as concentrações das matrizes a serem preditas (PARREIRA, 2003; ROGGO, 2007).

O modelo PCR é construído em duas etapas. A primeira etapa utiliza apenas a matriz X na determinação da combinação linear das variáveis (Equação 2) enquanto, na segunda, as variáveis dependentes são empregadas apenas quando os coeficientes de regressão são estimados, utilizando a regressão linear múltipla para estabelecer a relação entre a matriz Y e a matriz de *scores* T (Equação 3).

$$XP = T \quad \text{Eq. (2)}$$

$$Y = Tb + f \quad \text{Eq. (3)}$$

A PCR pode apresentar dificuldades de predição quando o analito apresenta um sinal muito baixo, não influenciando de maneira significativa na composição das primeiras componentes principais (PC 1 e PC 2), necessitando

assim de um número maior de dados na construção do modelo. A regressão nesse modelo é feita nas matrizes de *scores*, obtidos das componentes principais e não diretamente das variáveis originais (PASTI et al., 1998).

1.4.4 PLS

No modelo de regressão por mínimos quadrados parciais (PLS) se busca o melhor ajuste para um conjunto de dados, de forma a minimizar a soma dos quadrados das diferenças entre o valor observado e os valores estimados (essas diferenças são chamadas resíduos).

O PLS é o mais utilizado em associação com o NIRS (BRERETON, 2003). Neste modelo de regressão, uma grande quantidade de informações relevantes ao conjunto de calibração é incluída. Assim, é possível lidar com muitas variáveis, mesmo apresentando ruídos e alta correlação de dados (BARKER, 2003) na construção de um modelo linear (Equação 4).

$$Y = XB + E \quad \text{Eq. (4)}$$

em que, Y é uma matriz de variáveis resposta, X é uma matriz de variáveis preditoras, B é uma matriz dos coeficientes de regressão e E é a matriz de ruídos para o modelo.

O modelo PLS extrai componentes que são capazes de absorver variâncias das covariáveis e obter correlações com as variáveis dependentes, aumentando a covariância entre os fatores da matriz de variáveis X e a matriz das variáveis preditoras (Y), maximizando assim a predição de novos resultados (WOLD et al., 2001; MARTENS e NAES, 1996). O modelo de PLS apresenta vantagens como a facilidade para: modelar regressões com múltiplas variáveis

resposta, identificar a presença de *outliers* e produzir grande covariância com as variáveis respostas, aumentando assim o poder de predição (THERMO, 2002).

O modelo de Regressão dos Mínimos Quadrados Parciais Modificado (PLSM) ou Kernel-PLS é uma extensão da técnica PLS utilizando modelagem não linear (GARCÍA-REIRIZ, 2010; WALCZAKA, 1996), considerando a perda evidenciada na capacidade preditiva do PLS quando os dados não são lineares (ALMEIDA, 2007). O modelo de PLSM utiliza de uma etapa prévia de linearização dos dados, e posteriormente utiliza desses dados para a calibração linear por PLS. A matriz de Kernel (K) é gerada com os dados originais (X) em um espaço gaussiano não linear (Equação 5).

$$K = \exp(-(|X_{caln} - X_{caln}, \dots, N_{cal}|)^2 / \sigma^2) \quad \text{Eq. 5}$$

na qual X_{caln} é o vetor que corresponde a cada linha da matriz de calibração dos dados centrados na média, N_{cal} é o número de amostras de calibração, sendo σ o parâmetro correspondente à largura da transformação gaussiana.

A partir da matriz K com os dados linearizados são construídos modelos PLS lineares entre K e o vetor Y de resposta, obtendo-se desta forma os coeficientes de regressão (ALMEIDA, 2007; GARCÍA-REIRIZ, 2010).

Deste modo, objetivou-se avaliar o desempenho de um método analítico empregando a espectroscopia Vis-NIR associada à quimiometria aplicado à quantificação dos principais parâmetros analíticos em diferentes tipos de solos, considerando ser uma determinação rápida, não destrutiva, de baixo custo e de reduzido impacto ambiental.

2 MATERIAL E MÉTODOS

2.1 Amostras

Foram analisadas 277 amostras de solo (perfil 00-20 cm) de 9 municípios distintos distribuídos nas Regiões Oeste de São Paulo, Norte do Paraná e Sul do Mato Grosso do Sul, coletadas empregando-se uma sonda amostradora de solo de 20 polegadas. Todas as amostras foram preparadas de acordo com as metodologias oficiais (RAIJ, 2001).

As amostras foram secas em estufa a temperatura de 40°C, para evitar aumentos significativos de alguns metais como Mn, Fe, (KHAN, 1978; LEGETT, 1983) P e S (PEVERILL et al., 1975; SEARLE, 1987) pelo potencial de oxirredução das frações orgânicas e mineral do solo. Posteriormente, as amostras foram trituradas em moinho do tipo martelo (TECNAL, TE-330) e peneirados em malha de 2 mm de abertura. Todas as análises primárias foram realizadas no Laboratório Solos & Plantas – Unidade de Assis –SP.

2.2 Análise de Macro Nutrientes

Para análise de macro nutrientes (Ca, Mg, K) foram utilizados 2,5 cm³ de solo, com adição em copo coletor (polipropileno) de 110 mL de 5,0 mL de água deionizada e esferas de moagem, posteriormente foram levadas à mesa agitadora orbital (TECNAL, TE-360) por 15 minutos a 220 rpm. As esferas de moagem foram então retiradas adicionando-se 2,5 cm³ de uma mistura (1:1 v/v) de resina de troca catiônica (Amberlite IR-120, DOW Chemical Company) e troca aniônica (Amberlite IRA-400, DOW Chemical Company) que permaneceram em agitação por mais 16 h. Em seguida, as amostras passaram por um lavador de

resina para separação das mesmas, sendo adicionados 50 mL de solução extratora (solução de cloreto de amônio $0,8 \text{ mol L}^{-1}$ em HCl $0,2 \text{ mol L}^{-1}$) agitando-se por mais 1 h. Após este preparo as amostras foram lidas em Espectrômetro de Emissão Atômica com Plasma Induzido por Micro-ondas (MP-AES 4200-Agilent) tendo o cálcio, magnésio e potássio sido quantificados nos comprimentos de onda 393,366; 285,213; 766,491 nm respectivamente. Os resultados foram expressos em $\text{mmol carga dm}^{-3}$.

2.2.1 *Análise de Fósforo (P)*

Em copo coletor (polipropileno) de 80 mL foram adicionados $2,5 \text{ cm}^3$ de solo e 25,0 mL de solução extratora de Mehlich 1 (solução contendo HCl $0,05 \text{ mol L}^{-1}$ e H_2SO_4 $0,0125 \text{ mol L}^{-1}$) agitando-se em mesa agitadora orbital (TECNAL, TE-360) por 10 minutos e filtrando-se em papel filtro qualitativo (faixa branca). Pipetou-se num tubo de ensaio 5,0 mL do filtrado e adicionou-se 5,0 mL de solução complexo B (molibdato de amônio $0,008 \text{ mol L}^{-1}$ e H_2SO_4 $0,125 \text{ mol L}^{-1}$). Agitou-se levemente de forma manual e esperou a cor se desenvolver por 45 minutos. Após decorrido o tempo, as leituras foram feitas em espectrofotômetro UV-Vis (METASH V-5000) no comprimento de onda de 660 nm. Os teores de fósforo foram determinados utilizando-se uma curva analítica de padrões. Os resultados foram expressos em mg dm^{-3} .

2.2.2 *Análise de Enxofre*

Empregou-se $10,0 \text{ cm}^3$ de solo juntamente com $0,5 \text{ cm}^3$ de carvão ativado adicionando-se 25,0 mL de solução extratora (fosfato monocálcico $0,01 \text{ mol L}^{-1}$)

em copo coletor (polipropileno) de 80 mL, agitando-se em mesa agitadora orbital (TECNAL, TE-360) a 220 rpm por 30 min. Filtrou-se em papel filtro qualitativo, tendo 10,0 mL do filtrado sido recolhido e adicionando-se 0,5 mL de solução semente ácida ($\text{HCl} + \text{K}_2\text{SO}_4$ 20 mg L^{-1}) e mais 0,5 cm^3 de cloreto de bário (ANIDROL, p.a.). As amostras foram agitadas manualmente e após 3 minutos a leitura foi realizada em espectrofotômetro (METASH V-5000) em comprimento de onda de 420 nm. Os teores de enxofre foram determinados utilizando uma curva analítica de padrões. Os resultados foram expressos em mg dm^{-3}

2.2.3 Análise de Matéria Orgânica

Foram utilizados 0,5 cm^3 de solo e adicionados 5,0 mL de solução extratora (solução de 0,667 mol L^{-1} de dicromato de sódio e 5 mol L^{-1} de H_2SO_4), agitou-se em mesa agitadora orbital (TECNAL, TE-360) a 220 rpm por 10 minutos. Aguardou-se por uma hora para digestão completa da amostra e adicionou-se 25,0 mL de água deionizada. Após 12 horas de descanso as amostras foram analisadas em espectrofotômetro (METASH V-5000) em 650 nm. Os teores de matéria orgânica foram determinados utilizando uma curva analítica de padrões tendo sido os resultados expressos em g dm^{-3} .

2.2.4 Análise de pH e H+Al

2.2.4.1 Análise de pH

Em um copo coletor (polipropileno de 110 mL) foram adicionados 10 cm^3 de solo e 25,0 mL de solução extratora (solução de CaCl_2 0,01 mol L^{-1}) agitando-se por

10 minutos a 220 rpm em mesa agitadora orbital (TECNAL, TE-360). Após 30 minutos as amostras foram lidas em pHmetro (TEC-11, TECNAL com eletrodos Digimed DME-CV1) previamente calibrado.

2.2.4.2 Análise de H+Al (Acidez)

Após a leitura de pH foram adicionados às mesmas amostras 5,0 mL de solução tampão SMP (pH 7), agitou-se por 10 minutos em mesa agitadora orbital (TECNAL, TE-360) a 220 rpm e após uma hora as amostras foram lidas novamente em pHmetro (TEC-11, TECNAL com eletrodos Digimed DME-CV1). Os valores de acidez foram obtidos por meio de uma curva analítica e os resultados foram expressos em mmol carga dm^{-3} .

2.2.5 Determinação de Granulometria

Foram utilizados 20,00 g de cada amostra de solo e adicionados 50,0 mL de solução de NaOH 0,1 mol L^{-1} para favorecer a dispersão pelo aumento de cargas negativas e maior repulsão entre partículas (MAURI et al. 2011) em um copo coletor (polipropileno) de 110 mL. Agitou-se em mesa agitadora por 30 minutos a 220 rpm. Após esse período, as amostras foram transferidas para uma proveta de 500 mL e completou-se o volume até o menisco com água deionizada. Agitou-se as soluções com agitador vertical de provetas (TECNAL, TE-167) para completa dissolução das partículas de argila e mediu-se a densidade utilizando densímetro (Humboldt 5+60 G/L – 152H). Após duas horas de repouso a densidade foi medida novamente. As soluções foram então peneiradas (Bronzinox 270 mesh), secas em estufa a 105°C e pesadas posteriormente, calculando assim a porcentagem de areia, argila e silte em cada uma das amostras. Os resultados obtidos foram expressos em g Kg^{-1} .

2.2.6. Soma de Bases (SB)

A soma de bases trocáveis (SB) é representada pela soma dos teores de cátions de cálcio, magnésio e potássio disponíveis (Equação 06) (EMBRAPA, 2010).

$$SB = Ca^{2+} + Mg^{2+} + K^+ \quad Eq. (6)$$

2.2.7 Saturação por Bases (V%)

A saturação por bases (V%) é expressa pela soma de bases (SB) multiplicada por cem em razão da capacidade de troca de cátions CTC (Equação 07) ($Ca^{2+} + Mg^{2+} + K^+ + H^+ + Al^{3+}$)

$$V\% = 100 \times SB / CTC \quad Eq. (7)$$

2.2.8 Espectroscopia no Infravermelho Próximo (NIR)

As amostras foram analisadas utilizando o espectrômetro NIR (FOSS XDS Rapid Content Analyser, Hillerod, Dinamarca). As leituras foram realizadas em comprimento de onda de 400 a 2498 nm, com intervalos de 2 nm. Cada espectro foi obtido no modo reflectância e convertido para absorbância [$\log(1/R)$], com uma média de 32 leituras por amostra em 30 segundos. Para a calibração foram utilizadas 227 amostras e para a validação foram utilizadas 50 amostras. As amostras com aproximadamente 2,5 g foram colocadas dentro da célula de quartzo (*spinning*) e levemente prensadas com suporte para minimizar

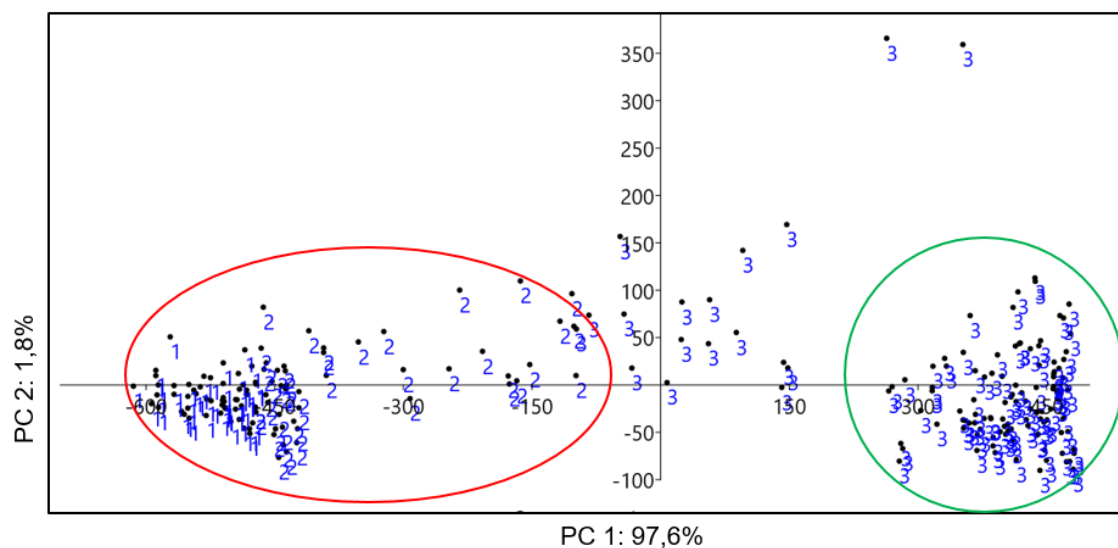
os efeitos produzidos pelo tamanho de partículas das amostras. Essas análises foram realizadas no Laboratório Multiusuários de Pesquisa Agropecuária (LAPA) na Universidade Estadual de Londrina (UEL).

Foram utilizados os tratamentos quimiométricos PCA, PCR, PLS e PLSM (WinISI™ FOSS, Dinamarca) buscando correlação entre os resultados das diversas técnicas e os espectros obtidos por NIR.

3 RESULTADOS

Utilizando os valores obtidos pelas técnicas primárias de todos os parâmetros (Tabela A1) foi realizada a análise de componentes principais (PCA). As amostras foram separadas em dois grandes grupos (Figura 2). Para melhor visualização as amostras foram nomeadas de acordo com sua classificação de textura.

Figura 2. Projeção das amostras no plano bidimensional das componentes principais 1 e 2 (gráfico de *scores*) referentes aos dados obtidos pelas técnicas primárias de 227 amostras de solo analisadas



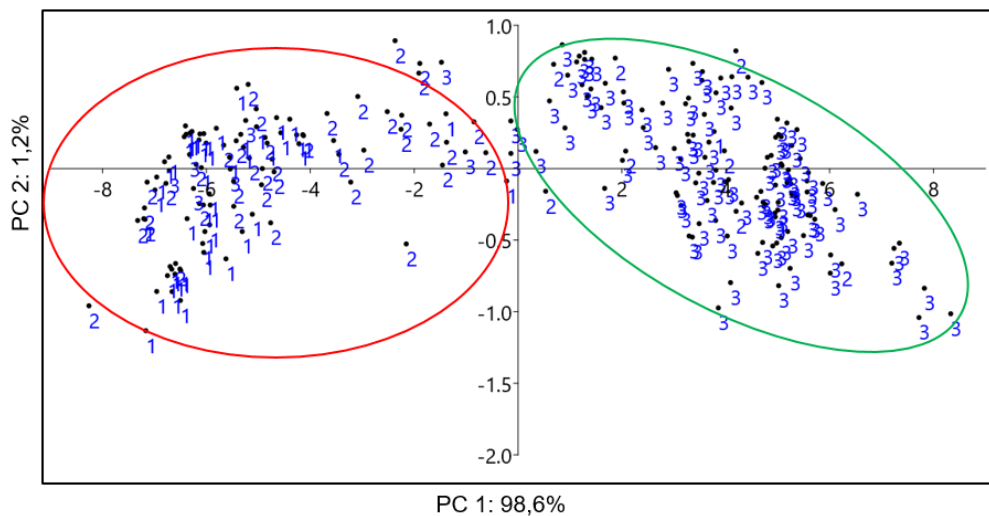
1 = Solo Tipo 1; 2 = Solo Tipo 2; 3 = Solo Tipo 3

As duas componentes principais explicam juntas 99,4% da variância total dos dados (97,6% para a PC1 e 1,8% para a PC2). Pode-se observar que o tratamento matemático aplicado sobre os dados das técnicas primárias permitiu a separação em dois grandes grupos solo predominantemente arenosos (Tipo 1 e 2) e solos de classificação em argilosos (Tipo 3), desta forma a análise de componentes principais indica que existe correlação e elevada variância entre os conjuntos de dados do sistema, sendo possível a aplicação de técnicas de análise multivariadas visando a construção de modelos preditivos para diferentes parâmetros das amostras.

3.1 NIR

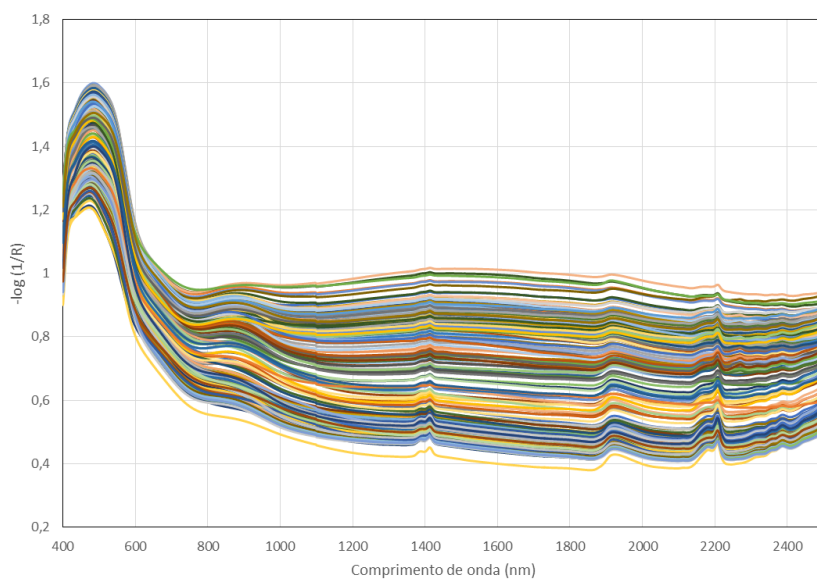
Primeiramente, testou-se a capacidade da técnica na separação de amostras de acordo com sua classificação de macro nutrientes e textura. Para isso, realizou-se a aquisição de dados espectrais Vis-NIR de 227 amostras, previamente analisadas pelas técnicas primárias. Realizou-se a Análise de componentes principais (PCA) dos dados espectrais referentes às amostras de solo (Figura 3), utilizando-se os valores de refletância difusa sem tratamento para todo o espectro Vis-NIR (400-2400 nm), observou-se a separação em dois grandes grupos.

Figura 3. Projeção das amostras no plano bidimensional das componentes principais 1 e 2 (gráfico de scores) referente aos dados espectrais das 227 amostras de solo com relação às suas refletâncias difusas entre os comprimentos de onda de 400 a 2400 nm.



1 = Solo Tipo 1; 2 = Solo Tipo 2; 3 = Solo Tipo 3

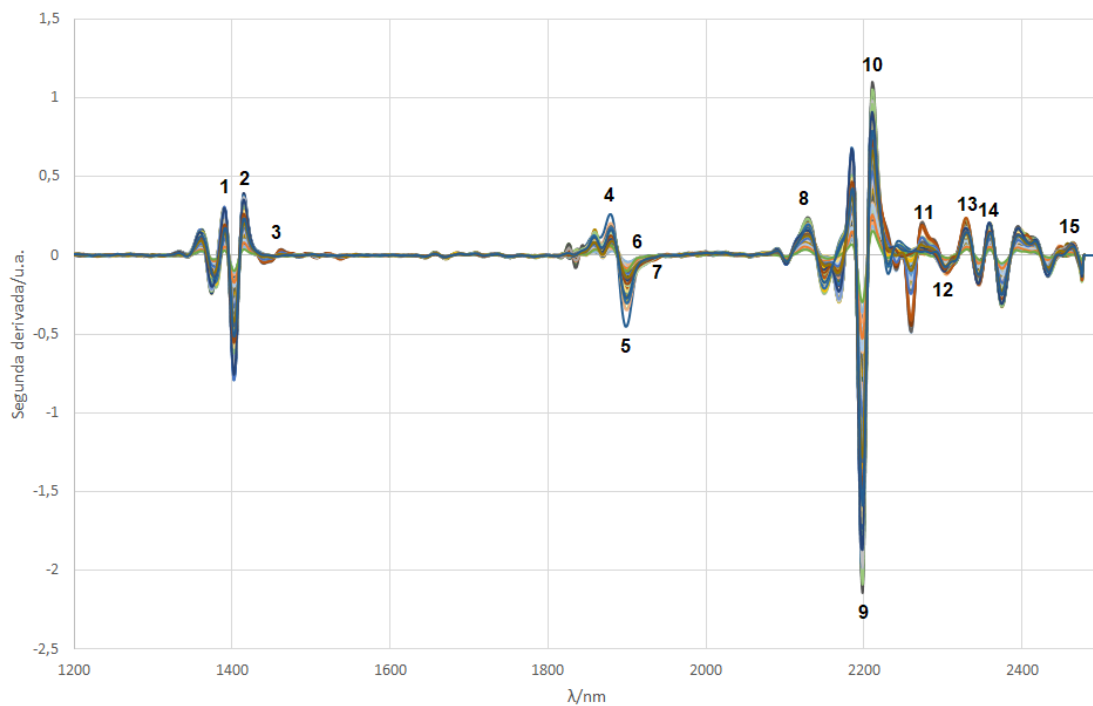
A PCA realizada utilizando as informações espectrais de todos os comprimentos de onda das leituras no NIR informou que duas componentes principais explicam 99,8% da variância total dos dados (98,6% para a PC1 e 1,2% para a PC2). Os espectros de infravermelho próximo sem tratamento, referente a estas amostras, estão apresentados na Figura 4.

Figura 4. Espectros NIR de 227 amostras de solo

Observando a Figura 4, todas as amostras de solo apresentaram perfis espectrais semelhantes, apresentando variações apenas nas intensidades relativas de algumas bandas como em 1390, 1410, 1950, 2180 e 2210 nm, o que já era esperado devido a semelhança entre as matrizes amostrais. As mudanças observadas no comprimento do visível (entre 400 e 700 nm) podem ser atribuídas à diferença de cor entre as amostras. Dos comprimentos de onda acima de 700 nm as diferenças podem ser atribuídas principalmente à presença de compostos orgânicos no solo (SILVA et al., 2020).

Realizou-se então, o tratamento estatístico (2ª derivada) nos espectros Vis-NIR (Figura 5) visando minimizar erros aleatórios na inclinação e a correção da linha de base.

Figura 5. Segunda derivada dos espectros NIR das 227 amostras de solo com diferentes parâmetros com regiões de alta variação espectral em destaque.



A partir da segunda derivada dos espectros Vis-NIR, foi possível observar as variáveis (bandas) que apresentaram maiores diferenças entre os valores de intensidade das amostras analisadas (Tabela 2).

Tabela 2. Variáveis selecionadas com maior variância entre os valores de refletância difusa nos espectros NIR das 227 amostras de solo analisadas

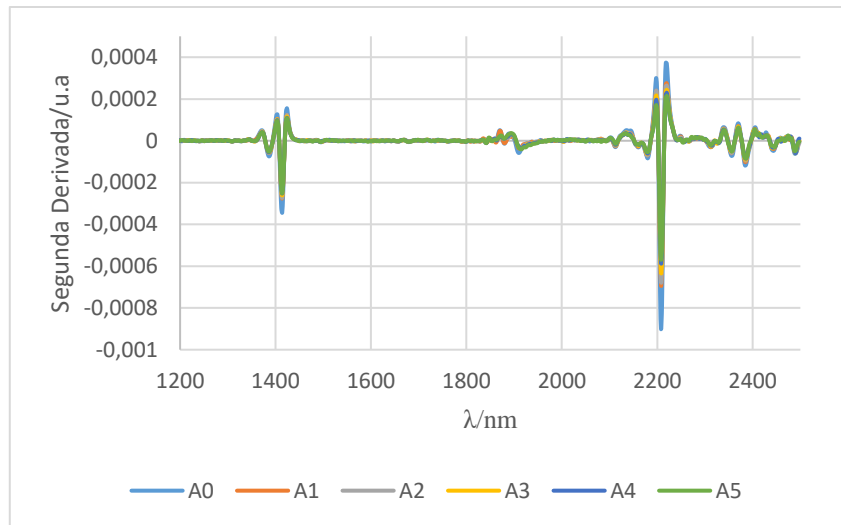
Região destacada no gráfico	Região espectral selecionada (nm)	NV	λ (nm)	Sobretom	Estiramento
1	1390 – 1406	8	1398	- -	C-H R-CH ₂
2	1410 – 1420	5	1410	-	O-H
3	1450 – 1460	5	1454	1° (R-OH)	O-H
4	1870 – 1890	10	1886	6° (C – Cl)	C-Cl
5	1894 – 1900	3	1896	2° COOH	C=O
6	1904 - 1920	8	1910	1° P-OH	O-H
7	1920 – 1930	5	1920	2° CONH	C=O O-H
8	2130 – 2140	5	2138	- -	C-H C=O
9	2188 – 2206	9	2196	- -	C-H C=O
10	2270- 2280	5	2272	- - -	O-H C-O R-Celulose
11	2280 – 2290	5	2286	- -	C-H R-OH
12	2300 – 2310	5	2308	-	C-H
13	2340 – 2350	5	2342	-	C-H
14	2350 – 2360	5	2354	2° Celulose	-
15	2470 – 2480	5	2474	1° (C-N)	C-N-C

NV = número de variáveis; λ = comprimento de onda; Resolução do espectrômetro NIR = 2,0 nm

Observando à Tabela 2, muitas regiões espectrais selecionadas são referentes à presença de compostos orgânicos responsáveis pelo teor de matéria orgânica no solo. A base espectral para predição de elementos metálicos, sem região de absorção no espectro Vis-NIRS é complexa e não compreendida completamente. Segundo Malley et al. (1999) provavelmente devido à sua interação química com grupos fortemente absorventes como O-H, complexos minerais, matéria orgânica e carbonato (CO_3^{2-}) podendo ser determinados pela técnica NIRS, o que explicaria a existência de correlação entre alguns dos parâmetros analisados.

Buscando observar possíveis mudanças de intensidade de bandas em diferentes concentrações de cátions metálicos foram adicionadas concentrações de padrões de Ca^{2+} (50 mg Kg^{-1} , 100 mg Kg^{-1} , 150 mg Kg^{-1} , 200 mg Kg^{-1} e 250 mg Kg^{-1}), Mg^{2+} (6,1 mg Kg^{-1} , 12,2 mg Kg^{-1} , 18,3 mg Kg^{-1} , 24,4 mg Kg^{-1} , 30,5 mg Kg^{-1}), K^+ (5,87 mg Kg^{-1} , 11,74 mg Kg^{-1} , 17,61 mg Kg^{-1} , 23,48 mg Kg^{-1} e 29,35 mg Kg^{-1}) respectivamente, a uma amostra com textura classificada como tipo 1 (solo predominantemente arenoso). O comportamento das bandas do espectro NIRS das 6 amostras nomeadas de A0, A1, A2, A3, A4 e A5 com concentração desses cátions modificada pode ser observado na Figura 6.

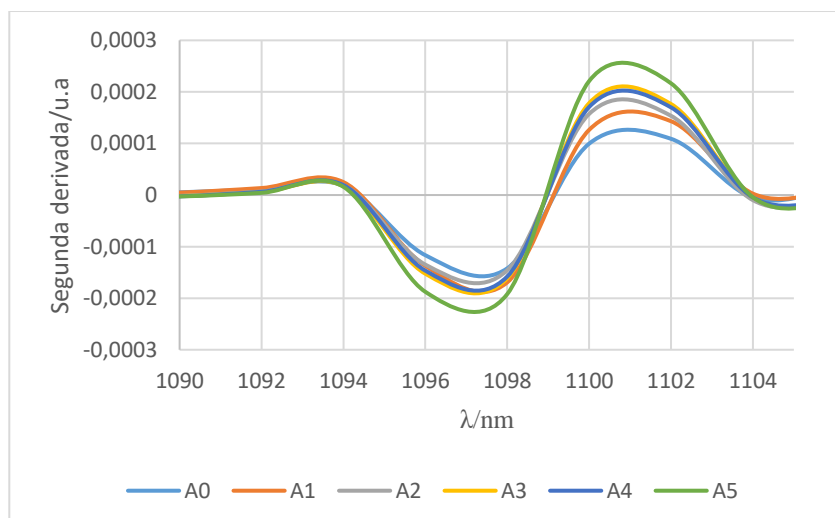
Figura 6. Espectro NIR das amostras (A0 até A5) com adição de padrão de cátions metálicos



Em que, A0 = amostra sem adição de padrão de cátions metálicos; A1, A2, A3, A4, e A5 = amostra com adição de padrão de cátions metálicos com concentração crescente.

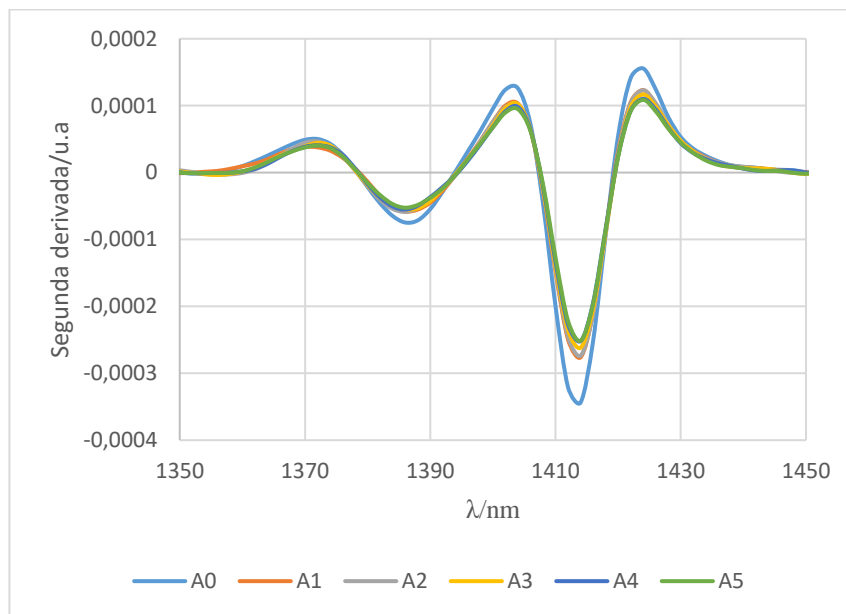
Observando a Figura 6, nota-se uma variação nas intensidades das bandas quando adicionados os padrões de cátions metálicos de concentração conhecida. Para uma melhor visualização nas bandas dos espectros nas regiões em que as alterações são perceptíveis foram ampliados (Figura 7 a 9).

Figura 7. Bandas do Espectro NIR de 1090 a 1105 nm das amostras (A0 até A5) com adição de padrão de cátions metálicos



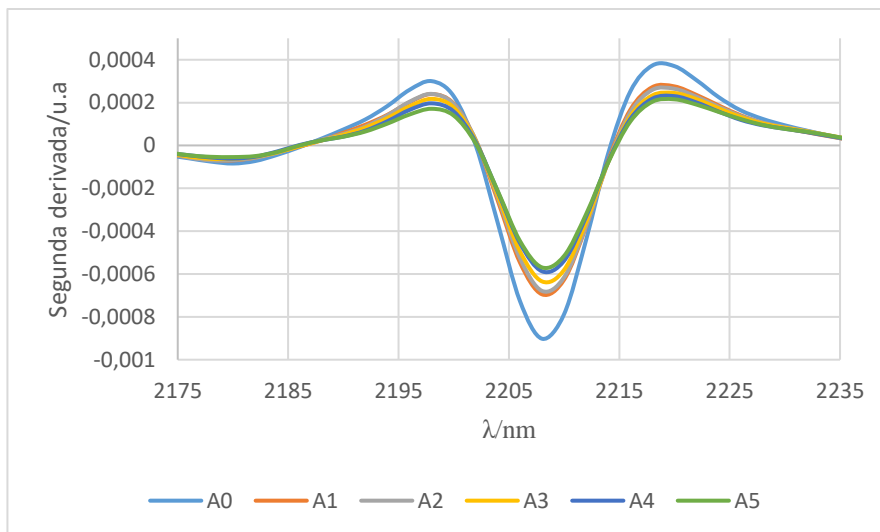
Entre os comprimentos de onda de 1090 e 1105 nm (Figura 7) pode-se observar um maior incremento na intensidade nas duas bandas da amostra que recebeu maior adição dos padrões de cátions metálicos (A5), indicando possível interação da radiação infravermelha com complexos metálicos adsorvidos pelo solo.

Figura 8. Bandas do Espectro NIR de 1350 a 1450 nm das amostras (A0 até A5) com adição de padrão de cátions metálicos



Na Figura 8 observa-se que as amostras cuja maior concentração de padrões de cátions metálicos adicionados (A5) tiveram suas bandas entre 1390 e 1430 nm atenuadas, enquanto a amostra sem adição cátions (A0) apresentou maior intensidade.

Figura 9. Bandas do Espectro NIR de 2175 a 2235 nm das amostras (A0 até A5) com adição de padrão de cátions metálicos



Na Figura 9 observou-se que novamente houve atenuação de bandas entre os comprimentos de onda de 2190 e 2225 nm nas amostras que receberam adição dos padrões de cátions metálicos em maior quantidade (amostra A5 em verde) quando comparadas à amostra sem adição (amostra A0 em ciano).

Desta forma, utilizando os espectros NIR das amostras de solo realizou-se uma calibração multivariada visando a determinação dos parâmetros pH, Ca, Mg, P, K, matéria orgânica, Acidez (H+Al), Soma de Base (SB), V%, Areia, Argila e Silte. Avaliou-se os resultados testando três modelos: regressão por componentes principais (PCR), regressão por mínimos quadrados parciais (PLS) e regressão por mínimos quadrados parciais modificado (PLSM), os quais estão apresentados nas Tabelas de 3 a 5 respectivamente. Observa-se diferentes números de amostras consideradas nos modelos para cada parâmetro devido aos *outliers* que são desconsiderados já na fase de construção dos modelos. Destaca-se que realizou-se a validação cruzada dos 3 modelos preditivos propostos, observando a exatidão e a precisão na previsão dos resultados, frente às técnicas analíticas primárias por validação cruzada (*cross-validation*).

Tabela 3. Modelo estatístico de regressão por componentes principais (PCR) com suas variações para os principais parâmetros do solo aplicado a 227 amostras de solo analisadas por NIR em duplicata

Constituinte	N	Média	SEC (%)	1-VR (%)	SECV (%)
pH	224	5,49	0,38	0,38	0,39
P	218	21,50 ¹	10,75	0,27	11,13
S	214	9,94 ¹	3,60	0,49	3,61
K	219	94,88 ²	31,69	0,46	31,76
Ca	222	39,41 ²	9,98	0,64	10,01
Mg	216	12,91 ²	3,45	0,66	3,47
H+Al	216	24,66 ²	4,86	0,67	4,95
SB	222	55,31 ²	13,29	0,67	13,30
V	223	67,27 ³	8,17	0,48	8,53
MO	216	25,70 ⁴	3,90	0,82	3,92
Areia	220	424,28 ⁴	71,61	0,95	73,23
Silte	221	140,35 ⁴	19,03	0,94	19,58
Argila	221	434,75 ⁴	56,09	0,95	57,64

N: número de amostras consideradas para a construção do modelo; **SEC:** Erro padrão de calibração; **1-VR:** Coeficiente de Determinação para Validação Cruzada; **SECV:** Erro padrão de Validação Cruzada. **1:** mg dm⁻³; **2:** mmol c dm⁻³; **3:** %, **4:** g dm⁻³

O modelo PCR (Tabela 3) apresentou coeficientes de determinação adequados para a textura (areia, argila e silte) 1-VR >0,94 e aceitáveis para a matéria orgânica (1-VR>0,80).

Tabela 4. Modelo estatístico de regressão dos mínimos quadrados parciais (PLS) com suas variações para os principais parâmetros do solo aplicado a 227 amostras de solo analisadas por NIR em duplicata

Constituinte	N	Média	SEC	1-VR	SECV
pH	224	5,49	0,31	0,48	0,35
P	220	21,92 ¹	8,23	0,47	9,79
S	224	9,98 ¹	3,61	0,50	3,62
K	218	95,47 ²	21,71	0,65	25,89
Ca	221	39,23 ²	7,16	0,76	8,16
Mg	217	12,95 ²	2,81	0,74	3,09
H+Al	218	24,69 ²	4,26	0,70	4,71
SB	221	55,05 ²	9,35	0,79	10,58
V	223	67,28 ³	6,12	0,65	7,06
MO	215	25,62 ⁴	3,07	0,85	3,54
Areia	220	426,99 ⁴	49,08	0,97	60,07
Silte	220	140,16 ⁴	13,25	0,96	15,73
Argila	219	434,42 ⁴	33,41	0,97	44,67

N: número de amostras consideradas para a construção do modelo; **SEC:** Erro padrão de calibração; **1-VR:** Coeficiente de Determinação para Validação Cruzada; **SECV:** Erro padrão de Validação Cruzada. **1:** mg dm⁻³; **2:** mmol c dm⁻³; **3:** %, **4:** g dm⁻³

Tabela 5. Modelo estatístico de regressão dos mínimos quadrados parciais modificado (PLSM) das 227 amostras de solo em duplicata.

Constituinte	N	Média	SEC	1-VR	SECV
pH	217	5,48	0,23	0,62	0,30
P	220	21,75 ¹	8,32	0,46	9,79
S	216	10,01 ¹	3,65	0,47	3,75
K	218	96,50 ²	16,23	0,64	27,46
Ca	222	39,45 ²	5,47	0,78	8,07
Mg	220	13,41 ²	1,97	0,78	3,12
H+Al	218	25,13 ²	3,87	0,71	4,96
SB	221	55,22 ²	7,20	0,80	10,71
V	220	67,38 ³	4,76	0,66	6,59
MO	214	25,57 ⁴	3,04	0,86	3,41
Areia	219	431,76 ⁴	41,22	0,98	48,23
Silte	217	137,21 ⁴	10,83	0,98	12,84
Argila	219	430,38 ⁴	28,24	0,98	36,20

N: número de amostras consideradas para a construção do modelo; **SEC:** Erro padrão de calibração; **1-VR:** Coeficiente de Determinação para Validação Cruzada; **SECV:** Erro padrão de Validação Cruzada. **1:** mg dm⁻³; **2:** mmol c dm⁻³; **3:** %, **4:** g dm⁻³

Pode-se observar que o modelo PLS (Tabela 4) apresentou coeficientes de determinação elevados para os parâmetros de areia, argila e silte, com 1-VR > 0,96. Já o modelo PLSM (Tabela 5) apresentou resultados superiores para os mesmos parâmetros 1-VR > 0,97. Valores aceitáveis também foram encontrados para predição de MO nos dois testes (1-VR>0,85). O erro padrão de calibração elevado para os demais parâmetros pode indicar que o modelo necessita de uma quantidade maior de amostras analisadas para que os erros de calibração sejam inferiores a 5%.

Pela leitura de 50 amostras de solo (não contidas na curva Vis-NIRS) em duplicada, totalizando 100 determinações realizou-se a validação externa, comparando os resultados preditos pelos modelos estatísticos e os resultados observados pela utilização das técnicas primárias. Obteve-se coeficientes de

determinação linear maiores que 90% ($R^2 > 0,90$) entre os valores observados pelas técnicas primárias e os valores preditos para os parâmetros de Areia, Silte e Argila (Figura 10 a 12). Os gráficos relativos aos demais parâmetros encontram-se no Apêndice deste trabalho (Figura A01 a A10).

Figura 10. Comparação dos valores observados pelas técnicas primárias e preditos de Areia nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

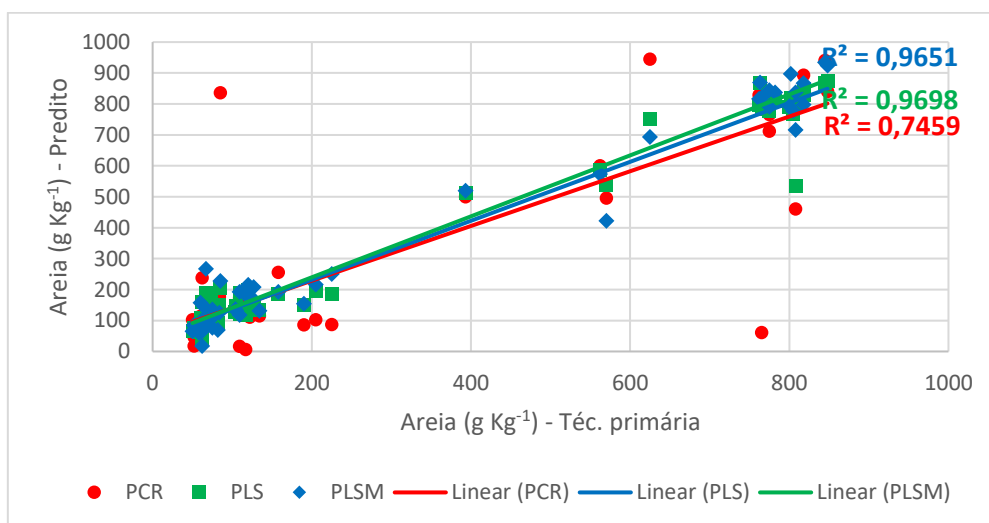


Figura 11. Comparação dos valores observados pelas técnicas primárias e preditos de Silte nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

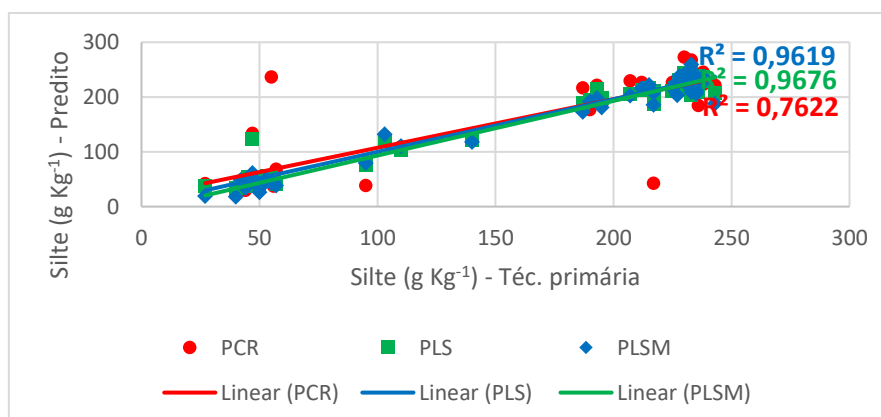
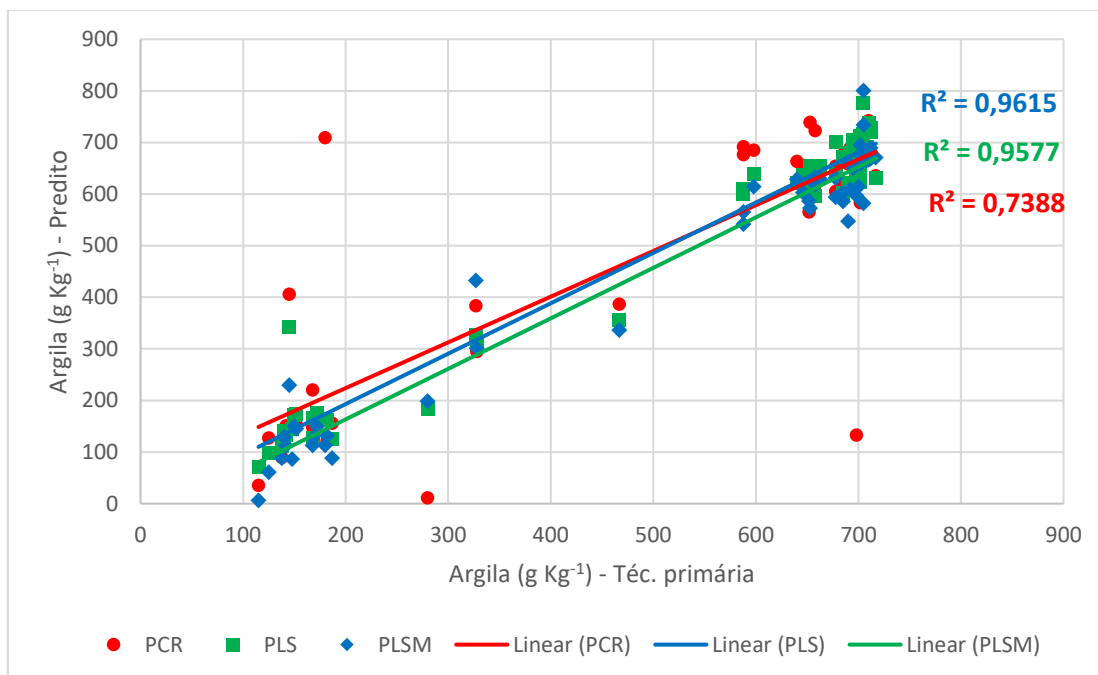


Figura 12. Comparação dos valores observados pelas técnicas primárias e preditos de Argila nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.



Segundo a análise de variância (ANOVA) não é possível concluir que haja diferença significativa ($p > 0,05$) entre três modelos quando comparados os valores preditos pelas técnicas primárias para predição de pH, S, Ca, Mg, H+Al, SB, V%, MO, Areia, Silte e Argila, sendo obtidos coeficientes de correlação de 0,65, 0,46, 0,78, 0,80, 0,70, 0,80, 0,65, 0,81, 0,98, 0,98, 0,98 respectivamente segundo Teste-t. Foi observada diferença significativa ($p < 0,05$) apenas para os modelos de predição de P (PLS e PLSM) e K (PLS), indicando baixo grau de exatidão na predição desses modelos.

A partir das Figuras 10 a 12, pode-se observar alto coeficiente de determinação linear ($R^2 > 0,95$) nas análises de textura, indicando que estes parâmetros podem ser preditos com grande precisão e exatidão pela técnica de Vis-NIRS utilizando os modelos PLS e PLSM para predição de areia, argila e silte.

Afim de verificar a qualidade de predição dos modelos obtidos, realizou-se o cálculo do coeficiente de variação (Equação 8) dos resultados gerados pelos três modelos.

$$CV = \frac{s}{x} \times 100 \quad \text{Eq. (8)}$$

Em que, CV: coeficiente de variação; s: desvio padrão; x: média

Comparou-se os coeficientes de variação dos três modelos construídos com os de dois ensaios de proficiência laboratoriais, do Instituto Agrônomo de Campinas (IAC) (IAC, 2021) e o Programa de Análise de Qualidade de Laboratórios de Fertilidade (PAQLF) (EMBRAPA, 2021). Os coeficientes de variação médios referentes aos modelos obtidos podem ser observados na Tabela 6.

Tabela 6. Coeficiente de Variação (CV(%)) médio dos resultados amostrais obtidos dos modelos estatísticos Vis-NIRS comparados aos CVs dos ensaios de proficiência laboratoriais de análises agronômicas disponíveis no Brasil, obtidos por meio de técnicas primárias e ANOVA.

	Coeficiente de Variação (%)					ANOVA	
	PCR	PLS	PLSM	IAC	PAQLF	F _{calculado}	valor-p
pH	4,23	4,00	3,92	6,44	5,90	1,07	0,362
P	29,54	31,10	33,78	37,10	25,90	0,57	0,635
S	28,24	22,72	23,13	39,79	37,28	3,94	0,009
K	29,49	27,22	26,83	27,00	22,13	0,12	0,944
Ca	19,00	17,90	22,20	24,86	15,93	0,56	0,638
Mg	18,41	17,51	17,67	27,13	18,33	2,11	0,100
H+Al	16,81	14,77	14,05	15,72	12,93	0,60	0,613
MO	14,87	10,35	11,14	13,42	12,63	1,55	0,202
Areia	31,04	17,54	20,07	12,07	6,70	4,06	0,008
Silte	12,72	6,96	9,64	30,21	24,83	12,72	1,58E-07
Argila	15,02	7,40	14,50	10,43	9,60	1,66	0,176
SB	17,60	15,72	20,65				
V%	9,00	9,06	9,35				

*F_{crítico}=2,65. **PCR**: Regressão por Componentes Principais; **PLS**: Regressão por Mínimos Quadrados Parciais; **PLSM**: Regressão por Mínimos Quadrados Parciais Modificada; **IAC**: Instituto Agrônomo de Campinas; **PAQLF**: Programa de Análise de Qualidade de Laboratórios de Fertilidade;

A ANOVA indicou que não houve diferença significativa (5,0%) entre as médias de coeficiente de variação (CV) para os parâmetros de pH, P, K, Ca, Mg, H+Al, MO e Argila ($F_{\text{calculado}} < F_{\text{crítico}}$) o que é reforçado pelo valor-p (em sua maioria $p > 0,05$).

O Ensaio de proficiência que conta com maior adesão dos laboratórios nacionais é o do IAC e a faixa de aceitação calculada para se obter valores analíticos aceitáveis é de um CV de até 20% (IAC 2021).

Apesar dos altos valores de coeficiente de variação para fósforo, enxofre e potássio os mesmos apresentaram valores próximos aos registrados pelos ensaios de proficiência, informando que até as técnicas primárias utilizadas apresentam uma alta variação interlaboratorial. Os valores de CV de soma de bases e V% não foram disponibilizados pelos ensaios de proficiência, pois são resultantes de cálculos de diversos parâmetros analíticos. O aumento do número de amostras no banco de dados utilizado para confecção dos modelos estatísticos, possivelmente, pode ocasionar diminuição dos valores de coeficiente de variância (CV), ressaltando a necessidade de se trabalhar com elevado número de amostras com composições distintas devido à alta complexidade existente na matriz analítica, demonstrando assim, o grande desafio para implementação de técnicas espectroscópicas em análises de matrizes complexas como o solo.

A viabilidade econômica torna a técnica NIRS ainda mais interessante por não necessitar de preparo prévio das amostras podendo ser executada por apenas um analista e não fazer uso de reagentes, possibilitando que a análise de diversos parâmetros seja feita em poucos minutos, o que levaria horas ou até mesmo dias pelos métodos primários (Tabela 7).

Tabela 7. Reagentes e tempo médio gasto para análise de 50 amostras pelas técnicas padrão.

Parâmetro	Reagentes	Tempo médio (horas)
pH e H+Al	Cloreto de Cálcio Trietanolamina Acetato de Cálcio Cromato de Potássio p-Nitrofenol	6
S	Fosfato monocálcico Cloreto de bário Ácido Clorídrico Sulfato de Potássio	5
P	Ácido Clorídrico Ácido Sulfúrico Molibdato de Amônio	5
K, Ca e Mg	Cloreto de Amônio Ácido Clorídrico	24
MO	Dicromato de Sódio Ácido Sulfúrico	14
Textura	Hidróxido de Sódio	8

4. CONCLUSÃO

Foi possível realizar a determinação dos parâmetros analíticos de matéria orgânica (MO), granulometria, soma de bases (SB), pH, acidez (H+Al), P, S, Ca, Mg e Saturação por bases (V%) por meio de modelos estatísticos de regressão linear multivariados gerados a partir da aquisição de espectros Vis-NIR.

Realizou-se a validação dos modelos preditivos com valores de coeficiente de determinação para validação cruzada (1-VR), SEC e SECV satisfatórios que resultaram em valores de coeficiente de variância (CV%) inferiores ou sem diferença significativa para parâmetros de pH, P, K, Ca, Mg, H+Al, MO e Argila aos apresentados pelos resultados dos ensaios de proficiência interlaboratoriais, realizados por meio de técnicas primárias, e coeficientes de determinação linear satisfatórios para análises agronômicas ($R^2 > 0,90$) para os parâmetros de Areia, Silte e Argila.

Os resultados obtidos demonstram o potencial de aplicação da técnica Vis-NIR em análises de rotina nos laboratórios agronômicos. Ressalta-se que, os modelos tendem a melhorar seu ajuste quanto maior o número de amostras, e pela técnica não exigir preparo de amostra, agregando no valor econômico no impacto ambiental.

REFERENCIAS

ALMEIDA, V. E. de. Algoritmo das projeções sucessivas associado ao Kernel-PLS para calibração multivariada não linear. Dissertação de Mestrado, Universidade Federal da Paraíba - João Pessoa, Paraíba 2007.

BARBOSA, L.C. de A. Espectroscopia no infravermelho na caracterização de compostos orgânicos. Universidade Federal de Viçosa – UFV, 2008.

BARKER, M.; RAYENS, W.; Partial Least Squares for Discrimination. Journal of Chemometrics. v. 17, 166, 2003.

BARROS, G. S. de C. MACRO/CEPEA: agronegócio soma 19 milhões de pessoas ocupadas, metade “dentro da porteira”. São Paulo: Cepea, 2017. Disponível em: <<https://www.cepea.esalq.usp.br/br/documentos/texto/macro-cepea-agronegocio-soma-19-milhoes-de-pessoas-ocupadas-metade-dentro-daporteira.aspx>>. Acesso em: 10 dez 2021.

BARTHUS, R.C. Aplicação de Métodos Quimiométricos para Análises de Controle de Qualidade de Óleos Vegetais utilizando Espectroscopias no Infravermelho e Raman. Dissertação de mestrado, Universidade Estadual de Campinas – UNICAMP. Campinas, São Paulo, 1999.

BEEBE, K.R.; PELL, R.J.; SEASHOLTZ, M.B. Chemometrics A Practical Guide. New York: John Wiley & Sons, 1998.

BORIN, A. Aplicação de quimiometria e espectroscopia no infravermelho no controle de qualidade de lubrificantes. Dissertação de mestrado, Universidade Estadual de Campinas- UNICAMP. Campinas, São Paulo, 2003.

BRETERON, R. G.; Chemometrics: Data Analysis for the Laboratory and Chemical Plant. John Wiley & Sons: England, 2003.

BRUNS, R., FAIGLE, J. Quimiometria. Química Nova. v. 8, 84-99. 1985.

BURNS, D.A.; CIURCZAK, E.W. Handbook of Near-Infrared Analysis. Second Edition, Revised and Expanded. Practical spectroscopy series volume 27. Marcel Dekker, Inc. New York, 2001.

CANTARELLA, H.; RAIJ, B. van; QUAGGIO, J.A. Situação da análise de solo no Brasil. In: REUNIÃO BRASILEIRA DE FERTILIDADE DO SOLO E NUTRIÇÃO DE PLANTAS, 21, 1994. Petrolina. Fertilizantes: insumo básico para agricultura e combate à fome. Anais. Petrolina: EMBRAPA-CPATSA/SBCS, p.9-33. 1995.

CEPEA – Centro de Estudos Avançados em Economia Aplicada. PIB do Agronegócio estabiliza no terceiro trimestre e setor cresce 10,79% de janeiro a setembro. Piracicaba – SP, 2021.

CHADDAD, F. The economics and organization of Brazilian agriculture: recent evolution and productivity gains. San Diego: Elsevier, 2016.

COBLINSKI, J.A.; INDA, A.V.; DEMATTÊ, J.A.M.; DOTTO, A.C.; GHOLIZADEH, A.; GIASSON, E. Identification of minerals in subtropical soils with different textural classes by VIS–NIR–SWIR reflectance spectroscopy. Catena, 2021.

CONAB - Companhia Nacional de Abastecimento: Estimativa indica aumento na produção de grãos na safra 2021/22, com previsão em 288,61 milhões de toneladas – Brasília-DF: 2021A. Disponível em: <<https://www.conab.gov.br/ultimas-noticias/4316-estimativa-indica-aumento-na-producao-de-graos-na-safra-2021-22-com-previsao-em-288-61-milhoes-de-toneladas>>. Acesso em: 10 dez 2021.

CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. Acompanhamento da Safra Brasileira de Grãos, Brasília, DF, v. 8, safra 2020/21, n. 8, oitavo levantamento, mai 2021 B.

CROSS, A.F.; SCHLESINGER, W.H. A literature review and evaluation of the Hedley fraction: Applications to the biogeochemical cycle of soil phosphorus in natural ecosystems. Geoderma, Amsterdam, v 64, p- 197-214, 1995.

DALMOLIN, R.S.D. Matéria orgânica e características físicas, químicas, mineralógicas e espectrais de Latossolos de diferentes ambientes. Tese de doutorado, Universidade Federal do Rio Grande do Sul, 2002.

DEMATTÊ, J.A.M.; GARCIA, G.J. Alteration of soil properties through a weathering sequence as evaluated by spectral reflectance. Soil Science Society of America Journal, Madison, v.63, p.327-342, 1999.

EBELING, A. G.; ANJOS, L.H.C.; PEREZ, D.V.; PEREIRA, M.G.; VALLADARES, G.S. Relação entre acidez e outros atributos químicos em solos com teores elevados de matéria orgânica. Solos e Nutrição de Plantas, Bragantina, 67 (2), 2008.

EMBRAPA- Empresa Brasileira de Pesquisa Agropecuária. Conceitos de fertilidade do solo e manejo adequado para as regiões tropicais / Carlos Cesar Ronquim – Campinas, São Paulo, 2010.

EMBRAPA- Empresa Brasileira de Pesquisa Agropecuária. Visão 2030: o futuro da agricultura brasileira. – Brasília, DF: Embrapa, 2018. 212 p. ISBN 978-85-7035-799- 1.

EMBRAPA- Empresa Brasileira de Pesquisa Agropecuária. PAQLF 2021. Relatório de Ensaio de Proficiência. Outubro de 2021.

FIELDVIEW; Disponível em:<<https://blog.climatefieldview.com.br/qual-e-a-participacao-do-agronegocio-no-pib-e-nas-exportacoes-brasileiras>>. Acesso em 12 dez 2021.

FERREIRA, M.M.C. Multivariate QSAR. Journal of the Brazilian Chemical Society, v. 13. 2002.

GARCÍA-REIRIZ, A.; DAMIANI, P.C.; OLIVIERI, A.C. Residual bilinearization combined with kernel-unfolded partial least-squares: a new technique for processing non-linear second-order data achieving the second-order advantage. Chemometrics and Intelligent Laboratory Systems 100, 127-135, 2010.

HONORATO, F.A. Previsão das propriedades das gasolinas do Nordeste empregando espectroscopia NIR/MID e transferência de calibração. Tese de doutorado, Universidade Federal de Pernambuco, Recife, 2006.

IAC- INSTITUTO AGRONOMICO DE CAMPINAS. Ensaio de Proficiência IAC para Laboratórios de Análise de Solo para fins Agrícolas. Relatório Nº 5/2021. Campinas, nov 2021.

KHAN, A.; SOLTANPOUR, P.N. Effect of wetting and drying on DTPA-extractable Fe, Zn, Mn, Na, Cu in soils. Communications in Soil Science and Plant Analysis, New York, v.9, p.193-202,1978.

KAMA-ELDIN, A.; ANDERSON, R. A multivariate Study of the Correlation Between Tocopherol Content and Fatty Acid Composition in Vegetable Oils. *Journal of the American Oil Chemists Society*, 74:375, 1997.

LIMA, K.M.G. et al. Sensores ópticos com detecção no infravermelho próximo e médio. *Química nova*, 32:1635, 2009.

LEGGET, E.G.; ARGYLE, P.D. The DTPA: Extractable Iron, Manganese, Copper and Zinc from neutral and calcareous soils dried under different conditions. *Soil Science Society of America Journal*, Madison, v.47, p.518-522, 1983.

MALLEY, D. F; YESMIN, L; WRAY, D; EDWARDS, S. Application of near-infrared spectroscopy in analysis of soil mineral nutrients. *Communications in Soil Science and Plant Analysis*, 30(7-8), 999–1012, 1999.

MAPA (MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO). Instrução Normativa Nº2, de 9 de Outubro de 2008. SECRETARIA DE POLÍTICA AGRÍCOLA 2008.

MARTENS, H.; NAES, T. *Multivariate Calibration*. John Wiley & Sons, New York, 1996.

MAURI, J.; Ruiz, H. A.; FERNANDES, R. B. A.; KER, J. C.; REZENDE, L. R. M. Dispersantes Químicos na Análise Granulométrica de Latossolos. *Revista Brasileira de Ciência do Solo* 35 (4). 2011.

McCLURE, W.F. Near-infrared spectroscopy: The Gian tis Runnin Strong. *Analytical Chemistry*. 66, 43A-53A. 1994.

MELLO, C.A. *Redes Neurais e Técnicas de Pré-processamento de Sinais em Espectroscopia de Reflectância no Infravermelho Próximo*. Tese de doutorado, Universidade Estadual de Campinas –UNICAMP. Campinas, São Paulo, 1998.

MORELLATO, S. A. Modelos de regressão PLS com erros heteroscedásticos. *Dissertação de Mestrado*, Universidade Federal de São Carlos- UFSCar. São Carlos, São Paulo, 2010.

NAGATA, N.; BUENO, M.I.M.S.; PERALTA-ZAMORA, P.G. Métodos matemáticos para correção de interferências espectrais e efeitos interelementos

na análise quantitativa por fluorescência de raios-X. Química Nova, Volume 24, No. 4, 531-539, 2001.

NETO, F. A.; GRAVINA, G.A.; SOUZA, N.O.S.; BEZERRA, A.A.C. Adubação fosfatada na cultura da soja na microrregião do alto médio Gurguéia. Revista Ciência Agronômica, v. 41, n. 2 p. 266-271, 2010.

NETO, W.B. Parâmetros de qualidade de lubrificantes e óleo de oliva através de espectroscopia vibracional, calibração multivariada, e seleção de variáveis. Tese de doutorado, Universidade Estadual de Campinas – UNICAMP. Campinas, São Paulo, 2005.

NETO, B.D.B., SCARMINIO, I., BRUNS, R. 25 anos de quimiometria no Brasil. Química Nova.v.29, p.1401-1406, 2006.

PARREIRA, T.F. Utilização de Métodos Quimiométricos em Dados de Natureza Multivariada. Dissertação de Mestrado, Universidade Estadual de Campinas – UNICAMP. Campinas, São Paulo, 2003

PASQUINI, C. Near infrared spectroscopy: fundamentals, practical aspects e analytical applications. Journal Brazilian Chemical Society, v. 14, No. 2, p. 198-219, 2003.

PASTI, I.; JOUAN-RIMBAUD D.; MASSART, D.L.; NOORD, O.E. Application of Fourier transform to multivariate calibration of near-infrared data. Analytica Chimica Acta. v. 368, p. 253-263, 1998.

PAVAN, M.A.; BINGHAM, F.T. & PRATT, P.F. Redistribution of exchangeable calcium, magnesium, and aluminum following lime or gypsum applications to a Brazilian Oxisol. Soil Science Society of American Journal. v. 48, p.33-38, 1984.

PEVERILL, K.I.; BRINER, G.P.; DOUGLAS, L.A. Changes in extrables sulphur and potassium levels in soil due to oven-drying and storage. Australian Journal Soil Research, East Melbourne, v.13, p.69-75, 1975.

RAIJ, B. van; ANDRADE, J. C. de; CANTARELLA, H.; QUAGGIO, J.A. Análise química para Avaliação da Fertilidade de Solos Tropicais. Campinas: Instituto Agronômico de Campinas, 2001.

RAIJ, B. VAN; CANTARELLA, H.; CAMARGO, A.P. DE & SOARES, E. Perdas de cálcio e magnésio durante cinco anos em ensaio de calagem. *Revista brasileira de Ciência do Solo*, 6:33-37, 1982.

REIS, M.M. Aplicação de Métodos Quimiométricos em Separação de Espectros e Reconhecimento de Padrões. Dissertação de mestrado, Universidade Estadual de Campinas- UNICAMP, Campinas, São Paulo, 1997.

ROGGO, Y.; CHALUS, P.; MAURER, L.; LEMA-MARTINEZ, C.; EDMOND, A.; JENT, N. A review of near infrared spectroscopy and chemometrics in pharmaceutical Technologies. *Journal of Pharmaceutical and Biomedical Analysis*, v.44, p.683-700, 2007.

SABIN, J. G.; FERRÃO, M. F.; FURTADO, J.C. Análise multivariada aplicada na identificação de fármacos antidepressivos. Parte II: Análise por componentes principais (PCA) e o método de classificação SIMCA. *Revista Brasileira de Ciências Farmacêuticas*, 2004.

SEARLE, PL.; SPARLING G.P. The effect of air-drying and storage on the amounts of sulfate and phosphate extracted from range of New Zealand topsoil. *Communications in Soil Science and Plant Analysis*, New York, v.18, n.2, p.725-734, 1987.

SILVA, J.C.A.; SIGNOR, D.; BRITO, A.M.S.S.; CERRI, C.E.P.; CAMARGO, P.B.; PEREIRA, C.F. Espectroscopia no Infravermelho Próximo e Análise de Componentes Principais para Investigação de Solos Submetidos a Diferentes Usos da Terra na Amazônia Oriental Brasileira. *Revista Virtual Química*, 2020.

SOUZA, A.M.; POPPI, R.J. Experimento didático de quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: um tutorial parte I. *Química Nova* 35, 223-229, 2012.

SOUZA, D.M. & LOBATO, E. Cerrado: Correção do solo e adubação. 2.ed. Brasília, Embrapa Informação Tecnológica, 416p, 2004.

SOUZA, E.R.; Determinação de parâmetros de qualidade de gasolina e estudo de contaminantes utilizando espectros do infravermelho próximo na fase vapor.

Tese de doutorado em Ciências, Universidade Estadual de Campinas – UNICAMP. Campinas, São Paulo, 2005.

SKOOG, D.A.; HLLER, F.J.; NIEMAN, T.A.; Princípios de Análise Instrumental. Capítulos 16 e 17. 5ª Edição. Editora Bookman. Porto Alegre, 2002.

THERMO, G. PLSplus IQ – User's Guide. Salem: USA.

VALDERRAMA, P. Avaliação de figuras de mérito em calibração multivariada na determinação de parâmetros de controle de qualidade em indústria alcooleira por espectroscopia no infravermelho próximo. Dissertação de mestrado, Universidade Estadual de Campinas – UNICAMP. Campinas, São Paulo, 2005.

WALCZAKA, B.; MASSART, D.L. The Radial basis functions – partial least squares approach as a flexible non-linear regression technique. *Analytical Chimica Acta* 331, 177-185, 1996.

WALKER, T.W.; SYERS, J.K. The fate of phosphorus during pedogenesis. *Geoderma, Holanda*, v 15, p. 01-19, 1976.

WOLD, S.; ESBENSEN, K.; GELADI, P. Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems*, 2:37, 1987.

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

ZHAO, D.; ARSHAD, M.; LI, N.; TRIANTAFILIS, J.; Predicting soil physical and chemical properties using vis-NIR in Australian cotton areas. *CATENA* 2020.

CAPÍTULO 2

DETERMINAÇÃO DE PARÂMETROS DO SOLO POR ANÁLISE DE IMAGEM DIGITAL

Resumo

DETERMINAÇÃO DE PARÂMETROS DO SOLO POR ANÁLISE DE IMAGEM DIGITAL

O cuidado com o solo reflete diretamente em aumento de produtividade, sendo imprescindível uma análise exata dos parâmetros do solo afim de se conhecer as necessidades de nutrientes de acordo com a cultura a ser implantada, resultando sempre em menores custos para o produtor. Os métodos empregados por laboratórios de análise agrônômica atualmente necessitam de um longo tempo de análise e geram quantidades expressivas de resíduos. Assim, este estudo tem como objetivo avaliar métodos rápidos que se apresentem como alternativa para os métodos primários de referência. Desta forma, investigou-se o uso de análise de imagem digital associada a quimiometria para efetuar a predição dos principais parâmetros do solo. Esta técnica utiliza apenas um aparelho de smartphone e um ambiente com luz controlada, diminuindo o preparo de amostra e eliminando o uso de reagentes. Inicialmente determinou-se os valores dos parâmetros analíticos por suas técnicas primárias de referência. Em seguida as imagens digitais foram obtidas com o uso de um smartphone de baixo custo acoplado a uma caixa de iluminação controlada, utilizando 2,5 g de terra fina seca ao ar, disposta em uma cela de quartzo do tipo spinning. As imagens digitais foram processadas no software *ChemColour* obtendo-se os parâmetros RGB e seus derivados de cada uma das amostras. Inicialmente aplicou-se a análise de componentes principais (PCA) aos resultados das imagens digitais para avaliar o potencial em separar as amostras em grupos de acordo com o tipo de solo. Posteriormente foram criados três modelos estatísticos PCR, PLS, PLSM, para avaliar a correlação e predição de pH, Ca, Mg, K, acidez (H+Al), soma de bases (SB), saturação por bases (V%), matéria orgânica (MO), areia, silte, argila e classificação de textura. Observou-se que não houve diferença significativa em nível de 5%, quando comparados os coeficientes de variação das técnicas primárias obtidos em ensaios de proficiência interlaboratoriais. Além disso, houve correlação entre os resultados obtidos pelas técnicas primárias e os resultados preditos utilizando os três modelos estatísticos, com coeficiente de determinação linear ($R^2 > 0,80$) para classificação de textura, argila, silte e areia. Conclui-se que a utilização da análise de imagem digital baseada nos parâmetros de RGB é uma alternativa para predição dos principais parâmetros analíticos do solo, sendo extremamente rápida a aquisição e tratamento dos dados, aumentando a capacidade analítica de laboratórios especializados além de não gerar resíduos.

Palavras-chave: química analítica, RGB, modelos estatísticos, métodos

Abstract

DETERMINATION OF SOIL PARAMETERS BY DIGITAL IMAGE ANALYSIS

Soil care directly reflects on increased productivity and an exact analysis of soil parameters is essential in order to know the nutrient needs according to the crop to be cultivated, always resulting in lower costs for the producer. The methods used by agronomic analysis laboratories currently require a long time of analysis and generate significant amounts of residues. So this study aims to evaluate rapid methods that are presented as an alternative to the primary reference methods. As shown we investigated the use of digital image analysis associated with chemometrics to predict the main soil parameters. This technique uses only a smartphone device and a controlled light environment, requiring no sample preparation and eliminating the use of reagents. Initially the values of the analytical parameters were determined by their primary reference techniques. Then the digital images were obtained using a low-cost smartphone coupled to a controlled lighting box using 2.5 g of air-dried fine earth, arranged in a spinning quartz cell. The digital images were processed in the ChemColour software obtaining the RGB parameters and their derivatives for each of the samples. Initially principal component analysis (PCA) was applied to the results of digital images to assess the potential to separate samples into groups according to soil type. Subsequently, three statistical models were created PCR, PLS, PLSM, to evaluate the correlation and prediction of pH, Ca, Mg, K, acidity (H+Al), sum of bases (SB), base saturation (V%), matter organic (OM), sand, silt, clay and texture classification. It was observed that there was no significant difference at the 5% level when comparing the coefficients of variation of the primary techniques obtained in interlaboratory proficiency tests. In addition there was a correlation between the results obtained by the primary techniques and the results predicted using the three statistical models with a linear determination coefficient ($R^2 > 0.80$) for classification of texture, clay, silt and sand. It is concluded that the use of digital image analysis based on RGB parameters is an alternative for the prediction of the main analytical parameters of the soil, being extremely fast the acquisition and treatment of the data, increasing the analytical capacity of specialized laboratories in addition to not generating residue.

Keywords: analytical chemistry, RGB, statistical models, methods

1. INTRODUÇÃO

1.1. Imagem Digital

O uso de imagem digital surgiu há mais de 70 anos utilizando-se do escaneamento de imagens analógicas por Russell A. Kirsch e sua equipe (NIST, 1957; KIRSCH, 1957 e KITAURA, 1975), as imagens apresentavam baixa resolução e se distorciam com relação a imagem original (KIRSCH, 2010; MITCHELL, 1994).

Com o desenvolvimento da tecnologia novos equipamentos surgiram como os dispositivos de cargas acopladas CCD (*Charge Coupled Device*), que utilizam silício e outros metais fotossensíveis, permitindo uma grande evolução nos scanners e a criação da câmera digital em 1975 (SONKA, 1993; KITAURA, 1975).

Uma imagem digital é captada utilizando-se câmeras ou sensores que captam a luz refletida. Essa é composta por uma matriz (Equação 1), na qual cada elemento dessa matriz é chamado de *pixel* (*Picture Element*). Para que a imagem digital se pareça com sua imagem original um número de *pixels* adequado é necessário (GOMES et al. 2008).

$$f(x, y) = \text{Cor no ponto } (x, y) \text{ Eq. (1)}$$

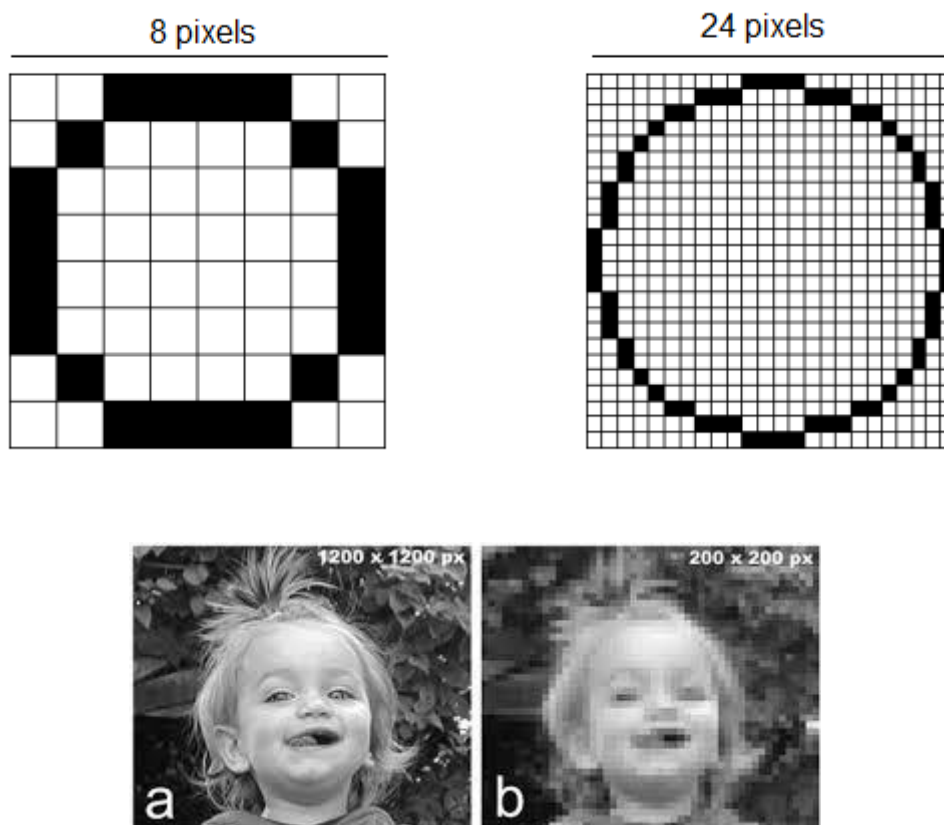
$$x \in [0, X] \text{ e } y \in [0, Y]$$

Realizando-se a divisão entre o número de *pixels* de uma imagem e seu tamanho real podemos encontrar seu valor de resolução, muitas vezes expresso em DPI (*dots per inch*, em português, pontos por polegadas) ou em pontos por centímetro (DPC). Diferentes aparelhos eletrônicos podem apresentar distintos

valores de resolução, impressoras comuns chegam a 600 DPI, monitores comuns a 72 DPI e scanners a 1200 DPI (SCURI, 2002). A câmera digital surgiu com o avanço dos dispositivos de cargas acopladas (CCD) e os semicondutores de óxido metálico complementar (CMOS) que são fotossensíveis. Ao receberem a luz refletida por um objeto convertem os fótons em elétrons (efeito fotoelétrico). Esses sensores atualmente apresentam baixo custo e permitem uma rápida aquisição de sinais (LUŠTICA, 2011).

Duas imagens de mesmo tamanho real com resoluções diferentes terão números de *pixels* distintos, sendo possível observá-las através da tela com tamanhos diferentes (Figura 1).

Figura 1. Representação de formas em diferentes resoluções



Fonte: adaptado de Scuri (2002) e <http://www.engesat.com.br/resolucao-o-que-e/>

Um computador armazena *bits*, composto por valores 0 e 1, para se formar a palavra *byte* (menor unidade de armazenamento) são necessários 8 *bits*. Um *byte* pode ser representado por até 2⁸ valores diferentes (variando de 0 a 255). (SCURI, 2002).

1.2. Teoria das cores

A cor de uma imagem é resultado do reflexo da luz que não foi absorvida pelos objetos contidos nessa imagem. O olho humano consegue captar comprimentos de onda entre 400 e 750 nm (Figura 2) (GONZALES, WOODS, 1992).

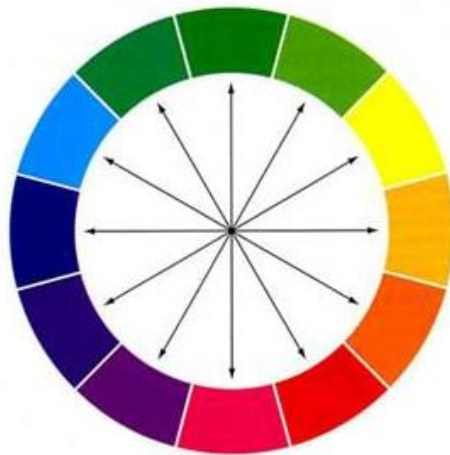
Figura 2. Comprimento de onda do espectro visível e suas cores

Cor	Comprimento de onda (nm)
Vermelho	625 – 740
Laranja	590 – 625
Amarelo	565 – 590
Verde	500 – 565
Ciano	485 – 500
Azul	440 – 485
Violeta	380 – 440

Fonte: adaptado de <https://rvq-sub.s bq.org.br/index.php/rvq/article/view/1152/608>

O sistema de cor percebido de um objeto é determinado pelo comprimento de onda absorvido e os demais comprimentos de onda que são refletidos, sendo chamados de comprimentos de onda complementares, que podem ser observados, de acordo com o disco de cores (Figura 3). Na situação em que a luz verde é absorvida, a cor observada será a cor oposta no disco de cores, no caso será a vermelha (MARTINS 2015).

Figura 3. Cores complementares no disco de cores



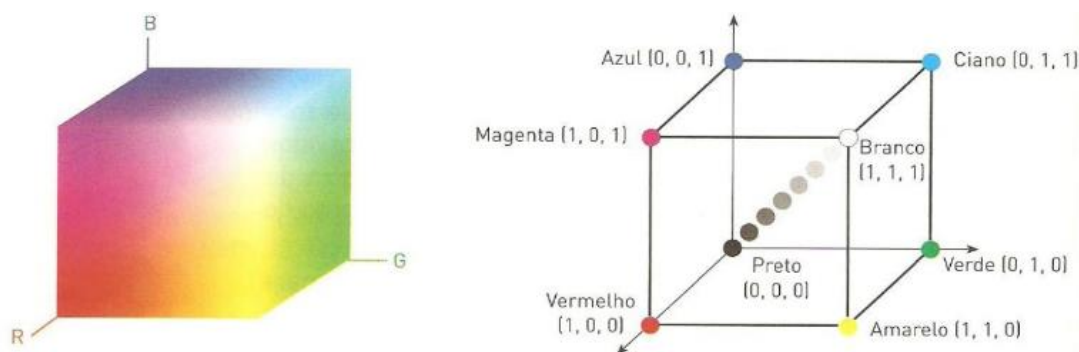
Fonte: <https://euquedesenhei.com/tudo-sobre-cor-disco-cromatico/>

1.3. Modelo RGB e modelo CMYK

Geralmente são utilizados três sensores, cada um com uma matriz, para armazenar as informações de uma imagem. A matriz R (*red*) contém os valores da cor vermelha, a matriz G (*green*) os valores da cor verde e a matriz B (*blue*) os valores da cor azul. Formando o que é conhecido como RGB e a combinação das matrizes (RGB) possibilita a aquisição de até 16 milhões de cores (GOMES et al. 2008, MALEKI 2004).

Além do modelo *RGB* (R, vermelho; G, verde; B, azul) existe também um modelo que considera as cores secundárias Ciano (C), Magenta (M) e Amarelo (Y), que é chamado de modelo *CMYK*. É muito comum seu uso em impressoras. Esse modelo é representado por uma espécie de cubo, contendo em cada vértice uma das cores do modelo (Figura 4).

Figura 4. Representação dos modelos *RGB* e *CMYK*



Fonte: <https://sites.google.com/site/aimciv/modelo-rgb>

A cor preta pode ser representada pela letra *K* e, por ser definida como a ausência de cores, recebe os valores de (0, 0, 0). No seu oposto no cubo está a cor branca (que pode ser representada pela letra *W*) no qual todos os valores alcançaram o seu máximo (255, 255, 255).

O cálculo dos parâmetros *CMY* e *K* dependem sempre dos valores de *RGB* e podem ser representados pelas Equações de 2 a 5 (WATANABE 2021).

$$K = 1 - \max\left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255}\right) \text{ Eq. (2)}$$

$$C = \left(\frac{1 - \frac{R}{255} - K}{1 - K}\right) \text{ Eq. (3)}$$

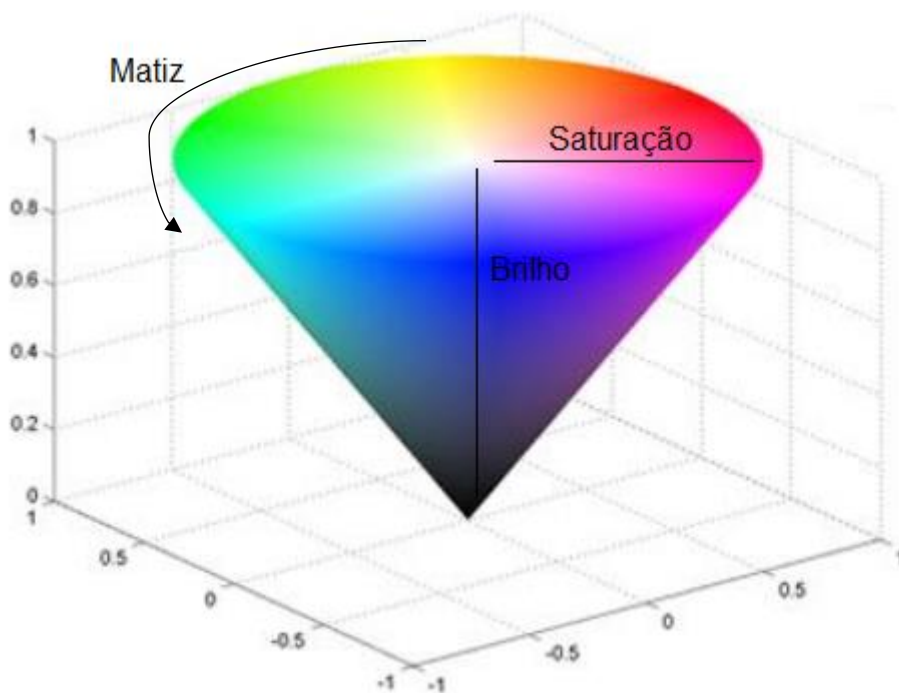
$$M = \left(\frac{1 - \frac{G}{255} - K}{1 - K}\right) \text{ Eq. (4)}$$

$$Y = \frac{1 - \frac{B}{255}k}{1-K} \text{ Eq. (5)}$$

1.4. Modelo HSV e modelo HSL

Além do *RGB* e *CMYK* outros modelos são considerados para a obtenção de parâmetros de cores, dentre eles o *HSV* (Matiz, Saturação e Valor, do inglês *Hue, Saturation, Value*) que é representado por um cone (Figura 5) onde a Matiz indica a cor, a Saturação indica a “pureza” da cor, quanto menor esse valor mais tom de cinza aparecerá na imagem e Valor indica o brilho da cor (CORRÊA 2014).

Figura 5. Representação do modelo HSV



Fonte: Adaptado de CORREA 2014.

Da mesma forma como no modelo *CMYK*, os parâmetros do modelo *HSV* também são obtidos por cálculos de conversão a partir do modelo *RGB* (WATANABE 2021).

$$V = \max \left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255} \right) \text{ Eq. (6)}$$

$$S = \begin{cases} 0 \\ V = 0 \frac{\Delta}{V} \\ V \neq 0 \end{cases} \text{ Eq. (7)}$$

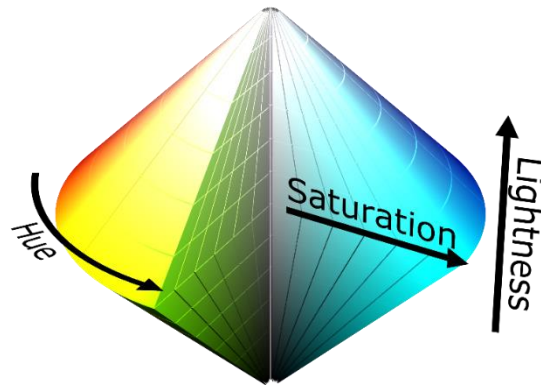
$$H = \begin{cases} 60^\circ \left(\frac{G-B}{\Delta} \text{ mod } 6 \right) \\ V = \frac{R}{255} 60^\circ \left(\frac{B-R}{\Delta} + 2 \right) \\ V = \frac{G}{255} 60^\circ \left(\frac{R-G}{\Delta} + 4 \right) \\ V = \frac{B}{255} \end{cases} \text{ Eq. (8)}$$

$$\Delta = V - \min \left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255} \right) \text{ Eq. (9)}$$

Em que, o mod6 é o dividendo com valor 6.

Outro sistema muito similar ao *HSV* é o *HSL* (Matiz, Saturação e Luminosidade, do inglês Hue, Saturation, Luminosity). Porém esse é representado por dois cones (Figura 06).

Figura 06. Representação do Modelo HSL



Fonte: <https://towardsdatascience.com/exploring-chromatic-storytelling-with-r-part-1-8e9ddf8d4187>

O parâmetro de luminosidade é muito semelhante ao brilho, variando entre uma escala de branco (máxima luminosidade) e preto (mínima luminosidade), podendo ser caracterizado também como o tom que uma cor assumiria se estivesse sendo representada apenas numa escala de preto e branco (GONZALEZ 1992). Como no modelo *HSV*, os parâmetros do modelo *HSL* também são obtidos por cálculos de conversão do modelo *RGB* (Equações 10-11).

$$L = \left(\frac{\max\left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255}\right) + \min\left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255}\right)}{2} \right) \text{ Eq.(10)}$$

$$S' = \begin{cases} 0 & \\ \Delta = \frac{\Delta}{1 - |2L - 1|} & \\ \Delta \neq 0 & \end{cases} \text{ Eq.(11)}$$

no qual delta (Δ) é o mesmo da Equação 9.

O uso de imagens digitais, no controle de qualidade de diversos produtos, apresenta crescimento principalmente na indústria, devido ao seu baixo custo de implementação, sendo utilizada para avaliar cores, além de outros parâmetros como a forma e textura dos objetos, sempre utilizando-se técnicas de análise multivariada (DU e SUN 2004). Gaião et al. (2006) utilizou imagem digital para determinar concentrações de reagentes em soluções aquosas, utilizando o sistema de cor RGB em *spot test*, com resultados similares aos das técnicas de referência.

Deste modo, objetivou-se avaliar o desempenho de um método analítico empregando a análise de imagem digital associada à quimiometria aplicado à quantificação dos principais parâmetros analíticos em diferentes tipos de solos, considerando ser uma determinação rápida, não destrutiva, de baixo custo e de reduzido impacto ambiental.

2. MATERIAL E MÉTODOS

2.1 Amostras

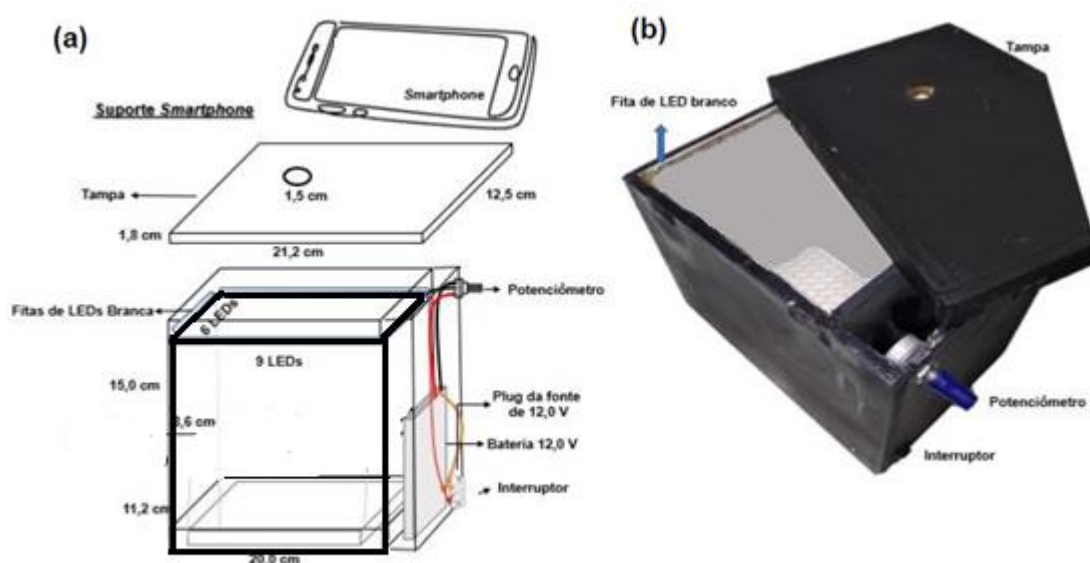
Foram utilizadas 200 amostras de solo (perfil 00-20 cm), que passaram pelo processo de secagem e moagem, conforme descrição feita no Capítulo I desta Tese.

2.2 Sistema Instrumental DIA-ChemColour

Aproximadamente 2,5 g de amostra foram inseridas em célula de quartzo (*spinning*) e inseridas em um conjunto composto por uma caixa preta, com

iluminação controlada. A caixa foi produzida em madeira de 4 mm de espessura (pintada de preto), com dimensões de 20 x 11,2 x 15 cm, e sua tampa apresentando dimensões de 21,2 x 12,5 x 1,8 cm de largura x comprimento x altura, respectivamente. Um orifício de 1,5 cm de diâmetro foi feito na tampa para o posicionamento da câmera do *smartphone* (Figura 7).

Figura 7. Representação da câmara fotográfica utilizada na obtenção de imagens digitais. (a) projeto do sistema instrumental. (b) sistema instrumental montado.



Fonte: Adaptado de Watanabe, 2021.

A parte interna da caixa foi pintada com tinta branca e fitas de LED branco (14,4 W/300 LEDs/5m – 6000k 12 V) instaladas na borda interna posicionadas a 0,5 cm do topo. Além disso um potenciômetro de 10 k Ω foi instalado para controlar a iluminação dentro da caixa (WATANABE, 2021).

O *smartphone* utilizado foi um Motorola, Moto G6 Play da versão de *hardware* P4 com 3 GB de memória RAM, 32 GB de memória ROM, resolução

máxima de tela de 1440x720, câmera frontal de 8 MP e traseira de 13 MP. A aquisição das imagens foi feita pelo aplicativo configurando a câmera para foco automático e *HDR* (do inglês, *High Dynamic Range*), que significa alto alcance dinâmico. Os arquivos foram armazenados em *jpg*. O tratamento das imagens digitais foi feito por meio do aplicativo *lab-made* desenvolvido em linguagem C# (C-Sharp) pelo Dr. Lycio Shinji Watanabe denominado de *ChemColour*.

O *software* possui opção de importar imagens digitais, e através de uma ferramenta, selecionar a área em que os parâmetros *RGB* serão determinados, conforme indicado na Figura 8.

Figura 8. Interface do programa de aquisição de dados de imagens digitais baseado no Sistema *RGB* tratadas pelo *software lab-made ChemColour para smartphone* com conversão para diferentes sistemas de cores aplicada na determinação dos principais parâmetros de solo

The screenshot displays the ChemColour software interface. On the left, there's an image of a soil sample with a circular selection tool. The main window shows a histogram of the selected area, with a red peak at approximately 185 on the R-axis. Below the histogram, a table lists color parameters for the selected area, which is highlighted with a red box:

Dados	
Novo	
Importar	
Salvar	
Excel	
RGB	
Apagar	
Salvar	
Backup	
Novos	
Hexadecimal	CMYK
Hex: #553323	C: 0
	Y: 0,23
	M: 0,4
	U: -0,05
	V: 0,09
	K: 0,67
RGB	YUV
R: 85	Y: 0,59
G: 51	V: 0,14
B: 35	
	HSV
	H: 19,2
	S: 0,59
	V: 0,14
	HSL
	H: 19,2
	S: 0,42
	L: 0,24
	CIE-L*a*b*
	x: 0,07
	Y: 0,06
	Z: 0,03
	L: 25,02
	a: 13,43
	b: 16,52

At the bottom of the interface, a status bar shows: Mouse Position: 201;104, pontual RGB: 0; 0; 0, Área: 31415,93, média RGB: 85;51;35, and 29/03/2022 23:26:13.

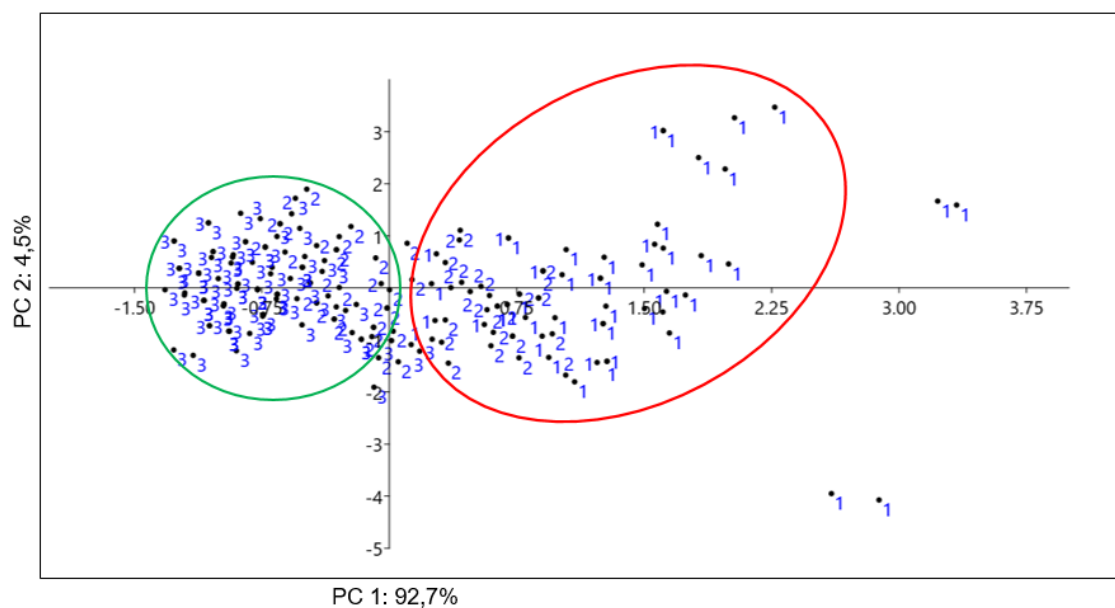
Fonte: *Software ChemColour* (WATANABE 2021).

Após a inserção da imagem digital, a área contendo a amostra pode ser selecionada, e os parâmetros *RGB* são determinados, além dos resultados calculados para outros sistemas como o *CMYK*, *HSL* e *HSV* (destacados em vermelho na Figura 7).

3. RESULTADOS

Com os valores das intensidades *RGB* das imagens digitais das amostras de solo, capturadas pelo smartphone e tratadas no Sistema *DIA-ChemColour* desenvolvido, realizou-se a análise de componentes principais (PCA) (Figura 9).

Figura 9. Projeção das amostras no plano bidimensional das componentes principais 1 e 2 (gráfico de *scores*) referentes aos dados das imagens digitais das amostras de solo, capturadas pelo smartphone e tratadas no Sistema *DIA-ChemColour*



As amostras foram classificadas de acordo com sua textura. As componentes principais 1 e 2 explicam juntas 97,2% da variância dos dados. É possível observar a formação de dois grandes grupos, de acordo com sua predominância de argila (solo classificado em tipo 3) e sua predominância de

areia (solo classificado em tipo 1 ou 2). A proximidade das amostras de mesma classificação no plano bidimensional, indicam que os parâmetros *RGB* permitem correlação com sua classificação de textura.

Para avaliar o potencial de predição da técnica de análise de imagem digital pelos parâmetros de RGB foram utilizados três modelos de regressão multivariada: PCR (Regressão por Componentes Principais), PLS (Regressão por Mínimos Quadrados Parciais) e PLSM (Regressão por Mínimos Quadrados Parciais Modificada), visando a determinação dos parâmetros pH, Ca, Mg, K, matéria orgânica, Acidez (H+Al), Soma de Bases (SB), V%, Areia, Argila, Silte e Textura. Utilizou-se um total de 200 amostras de solo (feitas em duplicata) com os parâmetros previamente determinados pelas técnicas primárias. A fim de realizar a validação dos modelos preditivos, realizou-se a leitura das imagens digitais de 40 amostras (não contidas na curva RGB) em duplicata, totalizando 80 determinações, observando os resultados, frente às técnicas analíticas primárias por validação cruzada.

O modelo PCR (Tabela 1) apresentou coeficientes de correlação linear (r) significativos para os parâmetros de Matéria Orgânica ($r > 0,80$), Areia ($r > 0,90$), Silte ($r > 0,90$), Argila ($r > 0,90$) e Textura ($r > 0,80$). Os coeficientes de determinação linear (R^2) encontrados para os parâmetros físicos do solo (areia, silte, argila e textura) foram superiores a 0,80.

Tabela 1. Modelo estatístico de regressão por componentes principais (PCR) das 200 amostras de solo em duplicata utilizando o Sistema *DIA-ChemColour*

PCR	N	r	R ²	RMSEP (%)
pH	182	0,2372	0,0562	0,4647
S	182	0,5183	0,2687	6,3577
K	182	0,4753	0,2259	39,5617
Ca	182	0,7487	0,5605	11,0914
Mg	182	0,7263	0,5275	4,6125
H+Al	182	0,6788	0,4607	6,7850
SB	182	0,7539	0,5685	15,3969
V	182	0,3884	0,1509	10,2195
MO	182	0,8164	0,6666	5,4974
Areia	182	0,9142	0,8358	134,1577
Silte	182	0,9110	0,8299	36,0183
Argila	182	0,9150	0,8373	98,3635
Textura	182	0,8991	0,8084	0,3587

N: número de amostras consideradas para a construção do modelo; r:coeficiente de correlação; R²:coeficiente de determinação; RMSEP: Erro quadrático médio de predição.

Tabela 2. Modelo estatístico de regressão por mínimos quadrados parciais (PLS) das 200 amostras de solo em duplicata utilizando o Sistema *DIA-ChemColour*

PLS	N	r	R ²	RMSEP (%)
pH	186	0,2518	0,0634	0,4630
S	186	0,5183	0,2687	6,3576
K	186	0,4757	0,2269	39,5521
Ca	186	0,7487	0,5606	11,0907
Mg	186	0,7263	0,5276	4,6121
H+Al	186	0,6808	0,4636	6,7670
SB	186	0,7540	0,5685	15,3960
V	186	0,3959	0,1567	10,1843
MO	186	0,8165	0,6667	5,4962
Areia	186	0,9150	0,8372	133,5840
Silte	186	0,9118	0,8315	35,8583
Argila	186	0,9158	0,8387	97,9506
Textura	186	0,8995	0,8092	0,3580

N: número de amostras consideradas para a construção do modelo; r:coeficiente de correlação; R²:coeficiente de determinação; RMSEP: Erro quadrático médio de predição.

Tabela 3. Modelo estatístico de regressão por mínimos quadrados parciais modificado (PLSM) das 200 amostras de solo em duplicata utilizando o Sistema *DIA-ChemColour*

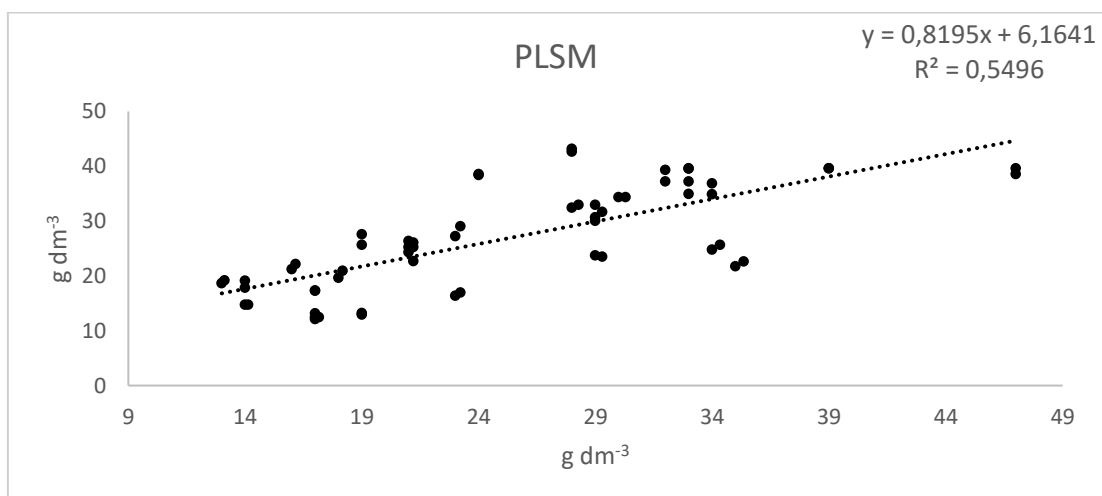
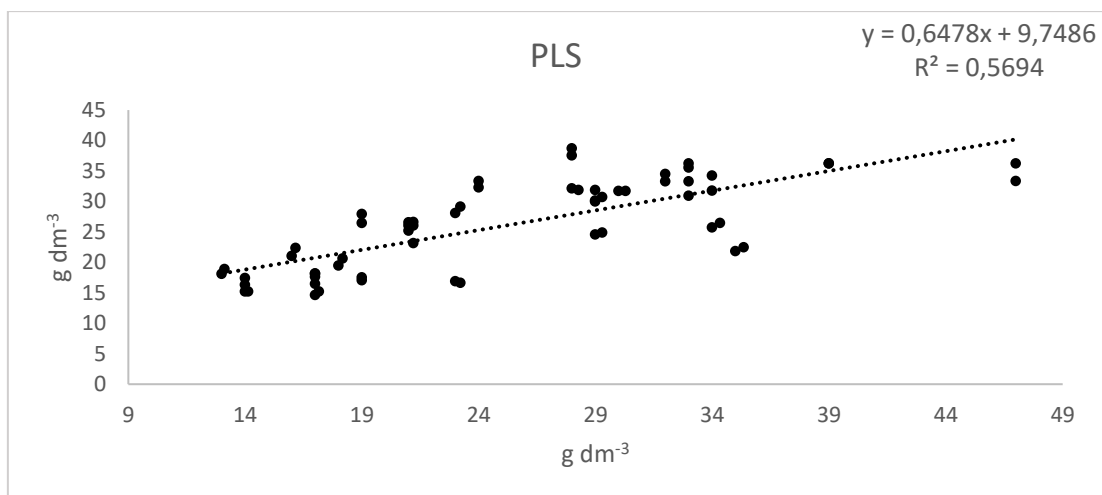
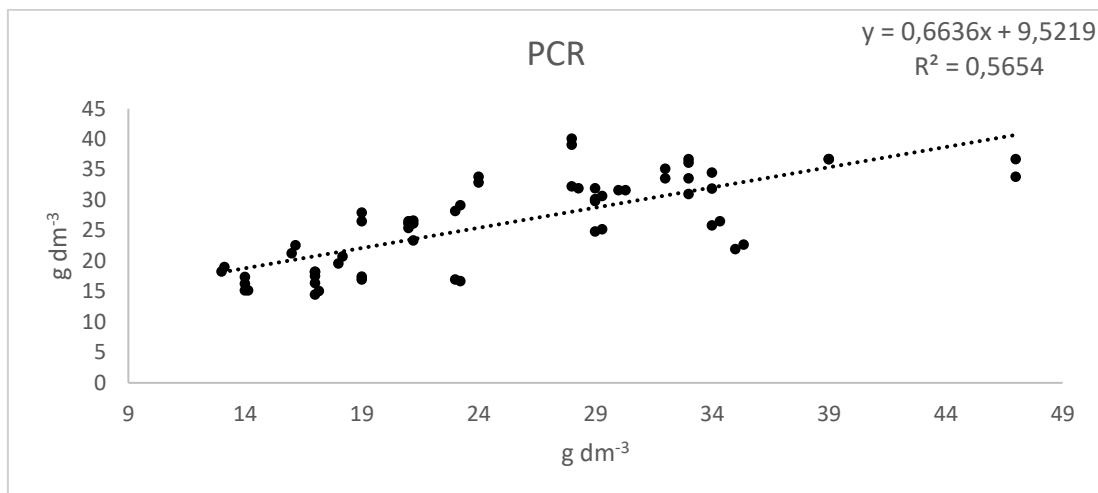
PLSM	N	r	R²	RMSEP (%)
pH	186	0,2375	0,0564	0,4647
S	186	0,5183	0,2687	6,3576
K	186	0,4752	0,2258	39,5632
Ca	186	0,7487	0,5606	11,0912
Mg	186	0,7263	0,5275	4,6124
H+Al	186	0,6789	0,4609	6,7839
SB	186	0,7540	0,5685	15,3966
V	186	0,3884	0,1509	10,2195
MO	186	0,8164	0,6666	5,4970
Areia	186	0,9143	0,836	134,1018
Silte	186	0,9111	0,8301	36,0040
Argila	186	0,9151	0,8375	98,3220
Textura	186	0,8991	0,8083	0,3588

N: número de amostras consideradas para a construção do modelo; **r:** coeficiente de correlação; **R²:** coeficiente de determinação; **RMSEP:** Erro quadrático médio de predição.

O modelo de Regressão por Mínimos Quadrados Parciais (PLS) (Tabela 2) apresentou parâmetros muito similares aos obtidos pelo modelo de Regressão por Mínimos Quadrados Parciais Modificado (PLSM) (Tabela 3).

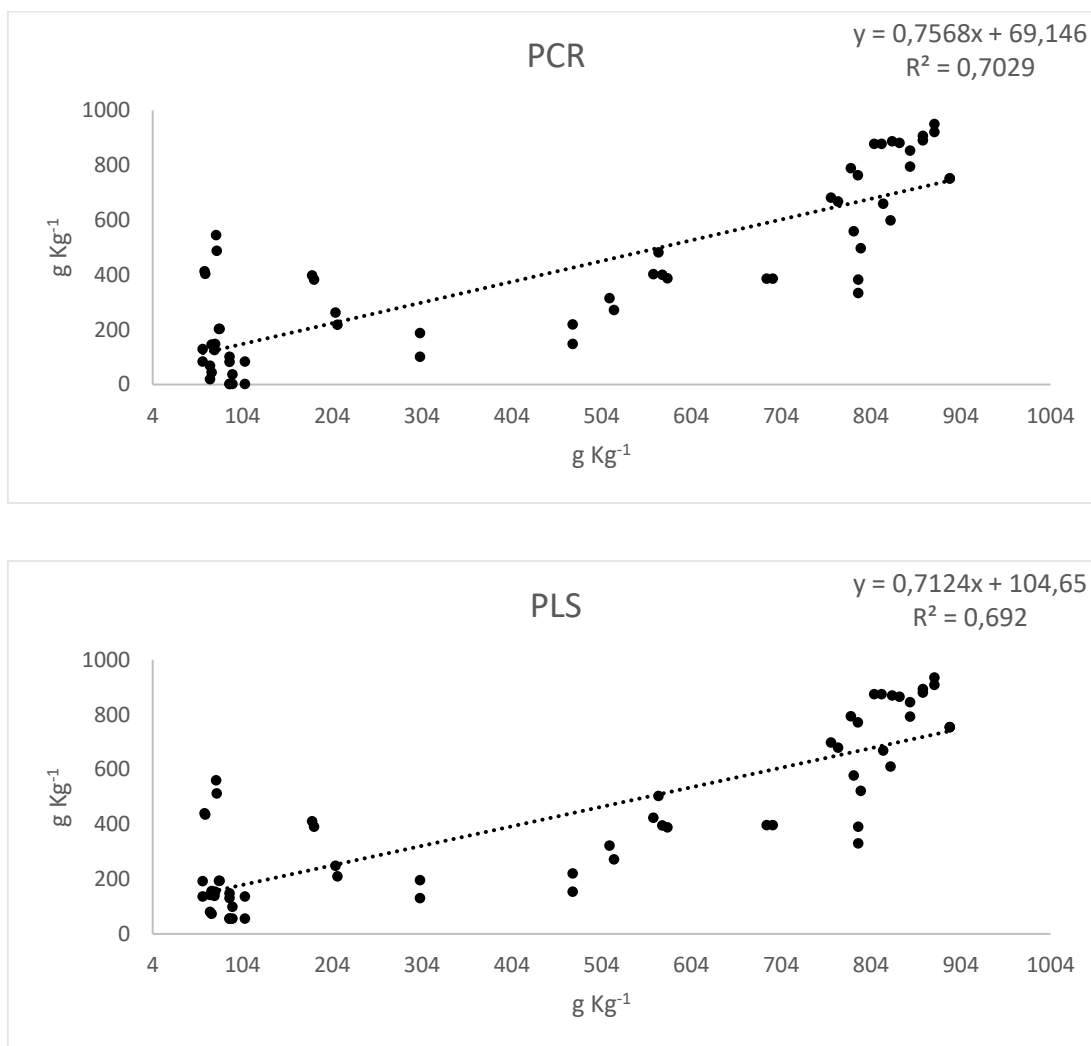
Realizou-se a comparação dos valores observados pelas técnicas primárias pelos valores preditos pelos modelos estatísticos e, pôde-se obter os coeficientes de determinação linear entre os valores preditos e observados para os parâmetros diretamente relacionados à cor do solo (Matéria Orgânica, Areia, Silte e Argila) (Figuras 10 a 14).

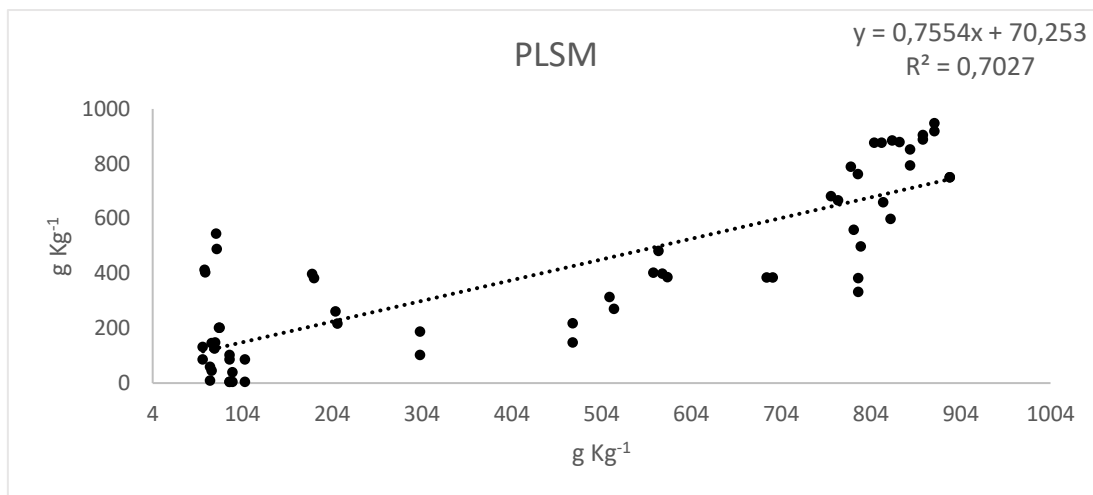
Figura 10. Comparação dos valores observados e preditos de matéria orgânica (MO) nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*



De acordo com o teste de análise de variância (ANOVA) (Tabela A2) não existe diferença significativa ($p > 0,05$) entre os resultados obtidos pelas técnicas primárias e os resultados previstos pelos três modelos estatísticos propostos, para predição do parâmetro de matéria orgânica, utilizando os dados adquiridos através do RGB, apresentando alto valor de correlação (correlação $> 0,80$ segundo teste t). O maior coeficiente de determinação para matéria orgânica observado foi de $R^2 = 0,5694$ no modelo PLS.

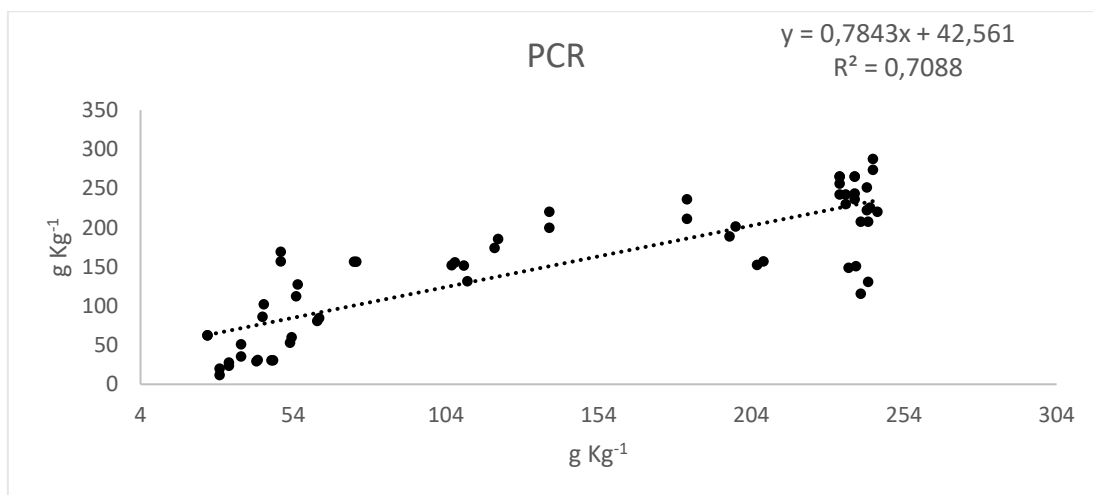
Figura 11. Comparação dos valores observados e preditos de areia nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*

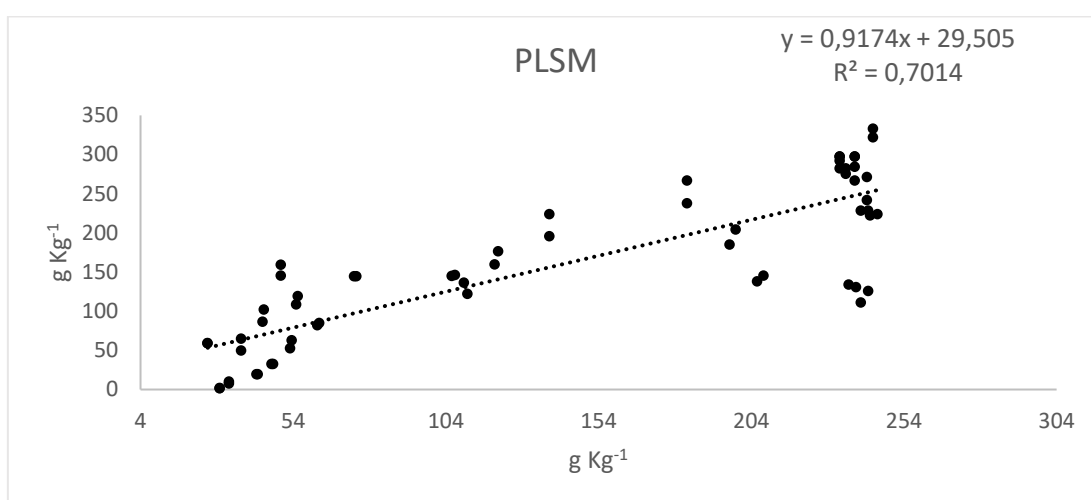
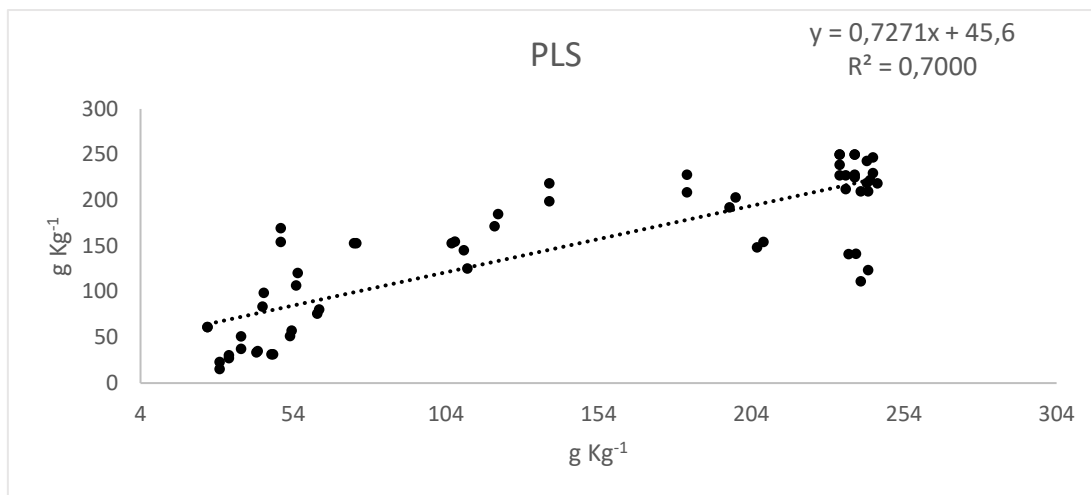




Segundo a ANOVA os três modelos não apresentaram diferença significativa quando comparados aos valores encontrados pelas técnicas primárias para o parâmetro de areia ($p > 0,9999$) com índices de correlação maiores que 90%. Nos modelos para predição de areia o maior coeficiente observado foi o do modelo de PCR ($R^2 = 0,7029$), seguido pelo modelo PLSM ($R^2 = 0,7027$) e o modelo de PLS ($R^2 = 0,6920$).

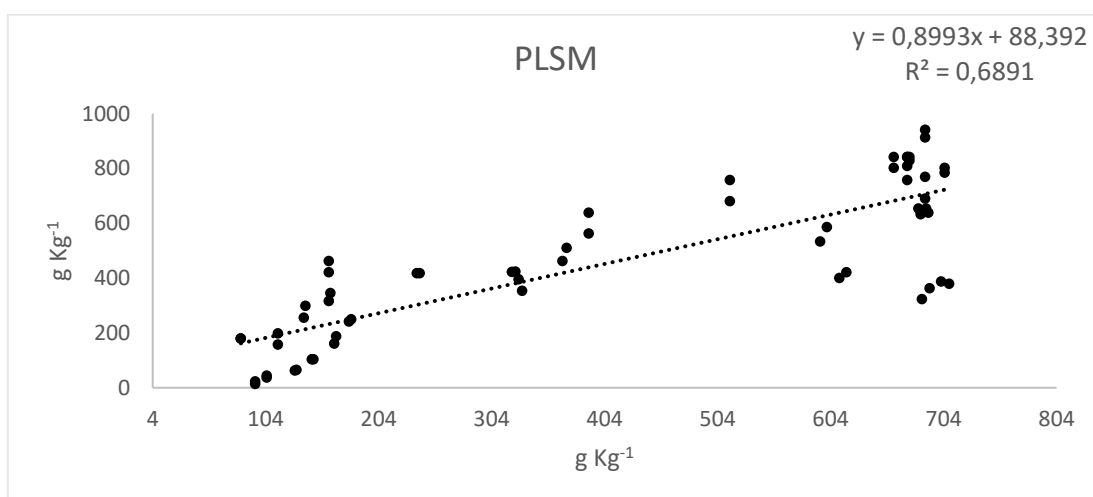
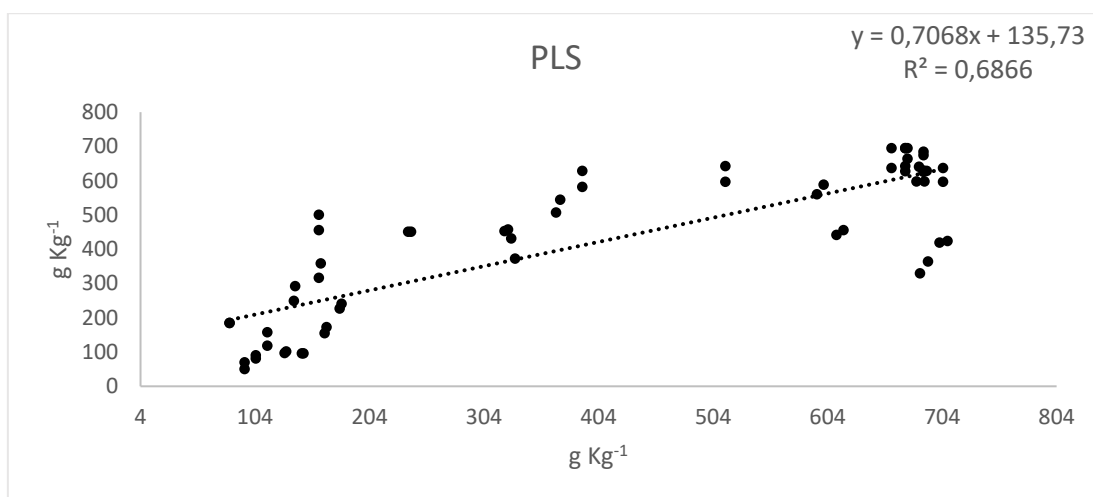
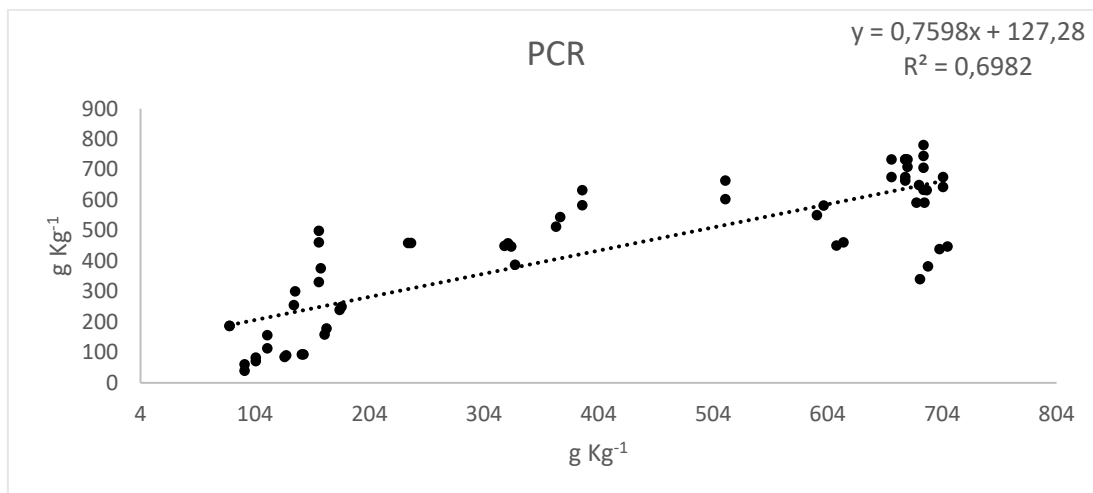
Figura 12. Comparação dos valores observados e preditos de silte nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*





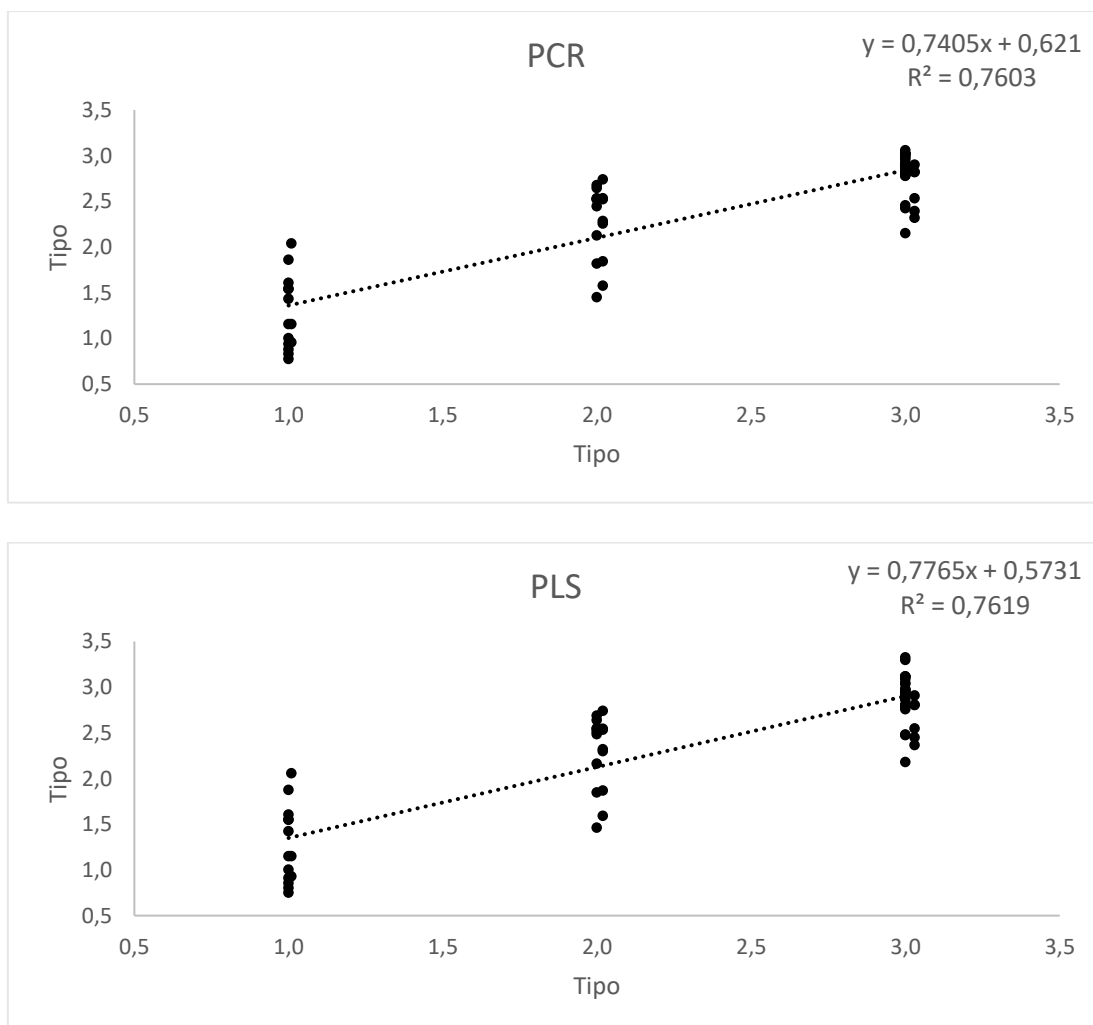
Os três modelos estatísticos construídos utilizando os parâmetros de *RGB* não apresentaram diferença significativa quando comparados aos valores encontrados nas técnicas primárias para o parâmetro silte ($p > 0,9999$), além de apresentarem alta correlação ($>90\%$) segundo o teste t. Para o parâmetro silte o maior coeficiente de determinação encontrado foi no modelo PCR ($R^2 = 0,7088$), seguido pelo modelo PLSM ($R^2 = 0,7014$) e PLS ($R^2 = 0,7000$).

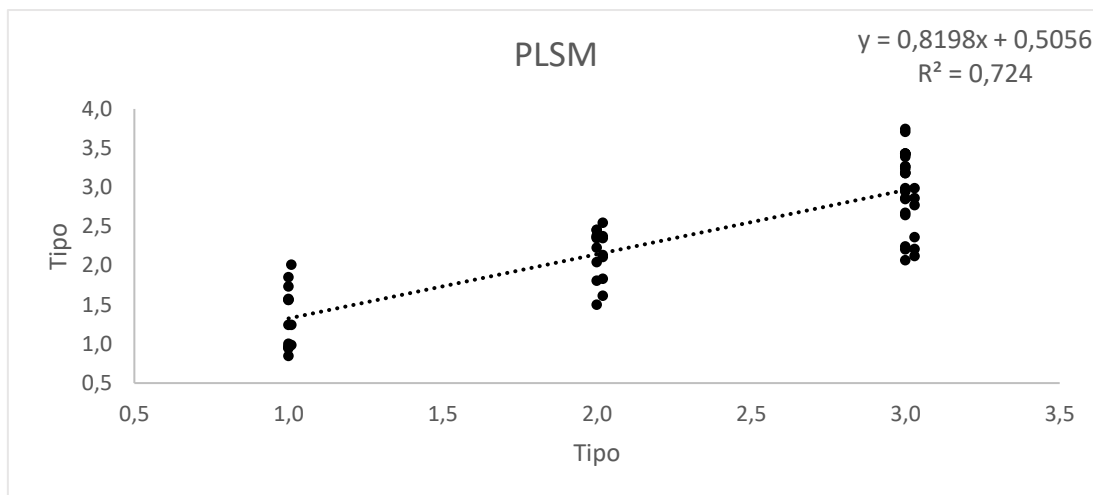
Figura 13. Comparação dos valores observados e preditos de argila nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*



Assim como os demais parâmetros de textura do solo, segundo o teste de análise de variância (ANOVA) não houve diferença significativa entre os valores encontrados nas técnicas primárias e os valores preditos pelos três modelos estatísticos ($p > 0,05$) para quantificação de argila, apresentando também alto valor de correlação entre os diferentes métodos de predição e as técnicas primárias ($> 90\%$). Para o parâmetro argila, o modelo que apresentou melhor coeficiente de determinação foi o PCR ($R^2 = 0,6982$), seguido pelo PLSM ($R^2 = 0,6891$) e o modelo PLS ($R^2 = 0,6866$).

Figura 14. Comparação dos valores observados e preditos de classificação de tipo de solo nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*





Foi possível calcular também através dos modelos estatísticos valores preditos para classificação de textura do solo em tipo 1 tipo 2 (solos predominantemente arenosos) e tipo 3 (solos predominantemente argilosos), obtendo-se um coeficiente de determinação de $R^2 = 0,7619$ no modelo de PLS, seguido pelos modelos PCR ($R^2 = 0,7603$) e PLSM ($R^2 = 0,7240$). Não houve diferença significativa quando comparados os valores encontrados pelos modelos estatísticos e a técnica primária ($p > 0,05$).

Observou-se grande similaridade nos três modelos estatísticos para todos os parâmetros, o que pode ser explicado devido à pequena quantidade de variáveis utilizadas na construção dos modelos, levando em conta apenas os valores de *RGB* e seus parâmetros secundários (*CMYK*, *HSV* e *HSL*). As demais figuras relativas aos outros parâmetros analisados como pH, Ca, Mg, K, H+Al, S, SB e V% se encontram no apêndice desse trabalho (Figuras A11 até A18).

Buscando verificar a qualidade de predição dos modelos obtidos, realizou-se o cálculo do coeficiente de variação (Equação 12).

$$CV = \frac{s}{x} \times 100 \quad \text{Eq. (12)}$$

Em que, CV: coeficiente de variação; s: desvio padrão; x: média

Comparou-se os coeficientes de variação dos principais ensaios de proficiência laboratoriais (IAC e PAQLF) com os coeficientes encontrados nos três modelos construídos (Tabela 4).

Tabela 4. Coeficiente de Variação (CV(%)) médio dos resultados amostrais obtidos dos modelos estatísticos RGB utilizando o Sistema *DIA-ChemColour* comparados aos CV dos ensaios de proficiência laboratoriais de análises agrônômicas disponíveis no Brasil, obtidos por meio de técnicas primárias

	CV (%)					ANOVA	
	PCR	PLS	PLSM	IAC	PAQLF	F	valor-P
pH	5,14	5,18	5,05	6,44	5,90	0,37	0,772
S	30,09	30,07	29,87	37,10	37,28	1,74	0,159
K	29,22	28,40	25,86	39,79	22,13	0,23	0,874
Ca	16,30	16,28	16,62	27,00	15,93	2,52	0,058
Mg	18,87	18,88	19,59	24,86	18,33	2,12	0,098
H+Al	17,95	17,72	18,77	27,13	12,93	0,48	0,693
M.O	12,37	12,18	15,14	15,72	12,63	1,54	0,203
Areia	39,83	32,36	39,64	13,42	6,70	3,98	0,008
Silte	23,69	21,75	31,60	12,07	24,83	2,53	0,058
Argila	23,09	21,67	28,76	30,21	9,60	4,79	0,003
SB	16,12	16,17	15,73				
V%	9,85	9,46	9,96				
Textura	11,16	11,50	12,14				

Observa-se que não houve diferença significativa (5,0%) entre as médias de coeficiente de variação (CV) para os parâmetros de pH, S,P, K, Ca, Mg, H+Al, MO e Silte ($F_{\text{calculado}} < F_{\text{crítico}}$) o que é reforçado pelo valor-p ($p > 0,05$).

Foi possível observar que os coeficientes de variação médios dos modelos estatísticos empregando o Sistema *DIA-ChemColour* são muito próximos ou até mesmo inferiores aos coeficientes de variação médios dos ensaios de proficiência interlaboratoriais do País (IAC e PAQLF), indicando uma boa exatidão dos três modelos estatísticos (PCR, PLS e PLSM) construídos utilizando os parâmetros de RGB.

4. CONCLUSÃO

É possível prever parâmetros analíticos de matéria orgânica (MO), granulometria (areia, silte e argila) e classificação de textura utilizando modelos estatísticos de regressão linear multivariados gerados a partir de aquisição dos parâmetros de *RGB* utilizando imagem digital por um aparelho de *smartphone*.

A predição se mostrou viável a partir da validação dos três modelos preditivos (PCR, PLS e PLSM) com valores de coeficientes de correlação e coeficientes de determinação satisfatórios, sem diferença significativa entre os valores preditos pelos modelos e os valores encontrados pelas técnicas primárias de todos os parâmetros descritos anteriormente. Os coeficientes de variação (CV%) se apresentaram inferiores aos apresentados pelos resultados dos ensaios de proficiência interlaboratoriais utilizando técnicas primárias.

Os resultados demonstram grande potencial de aplicação dos parâmetros de *RGB* pela análise de imagem digital captadas por *smartphones* em análises de rotina, como uma alternativa viável em laboratórios agrônômicos, diminuindo drasticamente o custo das análises – instrumental e operacional, o tempo de preparo das amostras e baixo impacto ambiental quando comparado as técnicas primárias utilizadas atualmente.

5. REFERÊNCIAS

DU, C.J.; SUN D.W. Recent developments in the applications of image processing techniques for food quality evaluation. *Food Science and Technology* v. 15, p. 230-249, 2004.

GAIÃO, E. N.; MARTINS, V. L.; LYRA, W. S.; ALMEIDA, L. F.; SILVA, E. C.; ARAÚJO, M. C. U.; Digital image-based titrations. *Analytica Chimica Acta*, v. 570, p. 283-290, 2006.

CORRÊA, J. H. G.; VEJA, G. U. S.; CUNHA, N. S.; SILVA, T. G. SILVA; MACIEL, P. D. Um modelo simples e parametrizável para classificação de cores no sistema RGB. IX CONNEPI, São Luís, Maranhão, 2014.

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. *Química Nova*, v. 31, No. 6, p. 1577-1581, 2008.

GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1992. ISBN 0201508036.

KIRSCH R. A., L. CAHN, C. RAY, G. H. URBAN, "Experiments in processing pictorial information with a digital computer", *Eastern Joint Comp. Conf.*, p. 221-229, 1957.

KIRSCH R. A., "Precision and accuracy in scientific imaging," *Journal of Research of the National Institute of Standards and Technology*, v. 115, no. 3, 2010.

KITAURA M.; YATA K., YAMADA S., to Minolta Camera Kabushiki Kaisha. Digital indicator for cameras. 3,895,875, 7-22-75, Cl. 356-226.000. 1975.

LUŠTICA A., "CCD and CMOS image sensors in new HD cameras," *Proceedings ELMAR*, Zadar, p. 133-136, 2011.

MARTINS, G.B.C; SUCUPIRA, R. R.; SUAREZ, P. A. Z. A Química e as cores. *Revista Virtual de Química*, v. 7, nº 4, 2015.

MALEKI, N.; SAFAVI, A.; SEDAGHATPOUR, F.; TALANTA. Single-step calibration, prediction and real samples data acquisition for artificial neural network using a CCD câmera. *Talanta*, v. 64, p. 830, 2004.

MITCHELL, W. J. *The reconfigured eye: visual truth in the post-photographic era*. Cambridge: MIT Press, 1994.

NIST. National Institute of Standards and Technology. "Fiftieth Anniversary of First Digital Image Marked". 1957. Disponível em: < <https://www.nist.gov/news-events/news/2007/05/fiftieth-anniversary-first-digital-image-marked>> Acesso em: 25 fev 2022.

SCURI, A. E. *Fundamentos da Imagem Digital*. Tecgraf, Rio de Janeiro, 2002.

Disponível em:

<<http://www.tecgraf.pucRio.br/ortoledo/teaching/cg1/apostila%20imagem%20digital>>. Acesso em: 2 mar 2022.

SONKA, M.; HLAVAC, V. e BOYLE, R. *Image Processing Analysis and Machine Vision*. London: Chapman & Hall, p. 25, 1993.

SONKA, M.; *Image Processing, Analysis, and Machine Vision*, PWD Publishing: New York, 1999.

WATANABE, L. S. *Desenvolvimento de Instrumentos e Metodologias Analíticas na detecção *in situ* de Adulterantes no Leite Bovino*. Tese de Doutorado. Universidade Estadual de Londrina – UEL. Londrina, Paraná, 2021.

6. APÊNDICE

Tabela A1. Parâmetros das 227 amostras analisadas pelas técnicas primárias

I.D	pH	P ¹	S ¹	K ²	Ca ²	Mg ²	H+Al ²	SB ²	V ⁴	M.O ³	Areia ⁵	Silte ⁵	Argila ⁵	Textura ⁶
1	5,4	8	9	1,9	44	20	35	65,9	65,3	29	58	234	708	3
2	5,8	10,8	8	5,3	54	22	27	81,3	75,1	32	62	233	705	3
3	5,4	11,1	11	1,8	43	18	36	62,8	63,6	28	85	227	688	3
4	5,4	10,3	6	0,8	44	19	34	63,8	65,2	26	82	228	690	3
5	5,5	12,3	8	1,1	39	19	24	59,1	71,1	28	75	233	692	3
6	5,4	10,5	8	1,4	40	17	34	58,4	63,2	28	78	230	692	3
7	5	7,2	9	2	29	13	38	44	53,7	29	62	236	702	3
8	5,7	9,8	12	1,2	58	20	30	79,2	72,5	27	62	236	702	3
9	5,4	8,8	8	3	41	19	33	63	65,6	31	68	232	700	3
10	5,7	10,7	8	2,3	36	17	22	55,3	71,5	29	62	233	705	3
11	5,7	11,3	8	1	58	17	30	76	71,7	28	55	237	708	3
12	5,8	8	9	0,8	55	19	26	74,8	74,2	31	52	236	712	3
13	5,2	8	9	1,1	42	17	36	60,1	62,5	29	58	234	708	3
14	5,4	8,2	9	1,4	33	15	26	49,4	65,5	29	50	238	712	3
15	6	6,8	9	1,3	64	17	24	82,3	77,4	30	52	238	710	3
16	5,8	6,4	10	4,4	73	17	25	94,4	79,1	30	60	235	705	3
17	5,7	5,7	11	3,2	65	24	30	92,2	75,5	34	65	235	700	3
18	5,7	17,8	21	3,1	61	19	29	83,1	74,1	29	322	153	525	3
19	5,7	39,8	23	3,3	53	14	29	70,3	70,8	27	325	153	522	3
20	5,5	31,9	8	2,8	57	22	31	81,8	72,5	29	190	182	628	3
21	5,4	22,3	9	3,5	43	17	34	63,5	65,1	28	188	187	625	3
22	5,4	18,6	13	1,9	47	18	35	66,9	65,7	28	185	180	635	3
23	5,1	26,4	9	3,8	38	12	39	53,8	58	29	182	173	645	3
24	5,1	29,2	6	1,9	39	13	37	53,9	59,3	28	158	192	650	3
25	5,8	38,3	17	1,9	56	17	31	74,9	70,7	28	155	190	655	3
26	5,1	24,7	8	3,5	39	11	36	53,5	59,8	28	150	192	658	3
27	4,8	13,7	23	1,4	30	12	46	43,4	48,5	27	145	195	660	3
28	6,5	24,3	12	2,7	62	24	18	88,7	83,1	26	225	180	595	3
29	6,3	30	15	2,4	69	29	22	100,4	82	31	150	200	650	3
30	6,1	16,9	11	3,1	65	24	21	92,1	81,4	30	142	198	660	3
31	6	33,6	14	4,8	57	22	23	83,8	78,5	34	158	190	652	3
32	6	29,9	22	2,2	56	20	25	78,2	75,8	32	165	185	650	3
33	5,8	23,1	15	2	51	20	25	73	74,5	29	162	178	660	3
34	5,4	25,3	11	2,7	52	14	34	68,7	66,9	27	225	187	588	3
35	5,4	19,6	16	2	44	17	35	63	64,3	29	145	207	648	3
36	5,2	18,9	12	2,3	38	14	36	54,3	60,1	30	148	194	658	3
37	4,9	14,5	20	1,8	31	13	42	45,8	52,2	26	142	206	652	3
38	5,7	21,3	13	1,8	44	17	29	62,8	68,4	28	125	200	675	3
39	5,6	20,5	36	1,8	38	14	32	53,8	62,7	27	122	193	685	3
40	5,6	18,9	13	1,8	47	20	28	68,8	71,1	30	118	202	680	3
41	4,7	28,5	20	1,8	28	9	53	38,8	42,3	29	132	193	675	3
42	5	19,9	15	2,7	37	11	43	50,7	54,1	30	122	198	680	3
43	4,9	40,2	8	2,2	42	12	42	56,2	57,2	31	100	210	690	3
44	4,9	36,3	18	2,5	43	14	51	59,5	53,8	31	98	207	695	3
45	4,6	20,8	25	1,2	32	9	62	42,2	40,5	32	105	217	678	3
46	5	26	12	2,4	44	14	45	60,4	57,3	32	92	210	698	3
47	4,8	30,9	12	2,1	36	11	50	49,1	49,5	50	120	195	685	3
48	4,6	24,1	14	2	34	10	54	46	46	53	118	192	690	3
49	6,1	26,2	20	1,8	59	21	26	81,8	75,9	55	122	203	675	3
50	5,3	22,5	18	2,6	51	18	34	71,6	67,8	46	100	218	682	3
51	5,2	17,1	16	2	50	15	37	67	64,4	30	102	200	698	3
52	5,7	20,2	19	2,7	55	18	29	75,7	72,3	30	120	192	688	3
53	5,4	18,7	19	2,1	51	18	32	71,1	69	30	122	206	672	3
54	5,5	17,3	16	1,3	58	19	33	78,3	70,4	29	128	212	660	3
55	5,6	22,8	16	1,9	72	27	32	100,9	75,9	31	85	217	698	3
56	6,1	30,7	2	0,9	30	12	16	42,9	72,8	14	762	56	182	2
57	6,3	32,8	2	0,9	25	9	12	34,9	74,4	14	782	50	168	2
58	6,4	27,8	3	1,3	28	12	14	41,3	74,7	13	800	48	152	2
59	6,3	51,2	3	0,7	26	11	13	37,7	74,4	13	805	45	150	1
60	6,3	55,1	4	1	27	7	11	35	76,1	14	822	43	135	1
61	5,3	28,1	3	0,8	19	7	15	26,8	64,1	15	782	53	165	2

62	5,2	62,1	6	0,9	20	6	20	26,9	57,4	16	805	45	150	1
63	5,7	27,8	4	1,2	19	7	14	27,2	66	11	765	55	180	2
64	5,4	38,5	6	1	22	6	19	29	60,4	13	775	53	172	2
65	6,3	30,6	6	1,4	38	14	12	53,4	81,7	18	762	56	182	2
66	5,9	45,7	7	1,6	28	9	12	38,6	76,3	16	818	42	140	1
67	5,7	51,6	4	1,5	20	7	15	28,5	65,5	12	800	48	152	2
68	5,6	29,8	5	0,9	21	9	13	30,9	70,4	12	810	45	145	1
69	5,5	11,1	10	3,6	44	15	22	62,6	74	21	688	74	238	2
70	5	22,4	13	4,9	39	10	31	53,9	63,5	27	620	90	290	2
71	5,6	31,3	4	2,4	16	5	13	23,4	64,3	15	892	26	82	1
72	5,2	35,3	6	2,1	22	6	19	30,1	61,3	16	812	46	142	1
73	5,6	37	5	1,7	25	6	15	32,7	68,6	16	875	30	95	1
74	5,2	43,5	4	1,7	20	5	17	26,7	61,1	19	862	33	105	1
75	5,9	56,1	4	1,6	29	7	12	37,6	75,8	15	890	28	82	1
76	5,9	6,3	3	1,9	20	9	12	30,9	72	16	912	23	65	1
77	5,5	32,9	5	2	25	5	16	32	66,7	20	862	33	105	1
78	4,9	39,2	5	2	19	5	22	26	54,2	17	875	30	95	1
79	4,7	34,9	5	1,5	15	6	21	22,5	51,7	19	895	25	80	1
80	6,1	31,1	4	3,1	48	9	11	60,1	84,5	14	882	28	90	1
81	5,4	61,4	4	1,9	26	7	16	34,9	68,6	17	890	25	85	1
82	5,3	47,1	4	2,2	22	5	16	29,2	64,6	17	892	26	82	1
83	6,1	14,5	7	2,5	51	19	16	72,5	81,9	21	572	106	322	2
84	6	24,5	10	4,1	75	24	18	103,1	85,1	30	502	123	375	3
85	5,8	13,4	13	2,5	69	20	20	91,5	82,1	25	512	118	370	2
86	6,1	19,5	11	6,7	97	34	20	137,7	87,3	35	325	170	505	3
87	4,5	18,7	20	4	33	15	53	52	49,5	34	155	213	632	3
88	5,3	54,8	6	12,2	52	23	35	87,2	71,4	39	168	207	625	3
89	5,3	49,1	8	12	87	33	41	132	76,3	40	240	190	570	3
90	4,8	15,9	18	3,6	33	14	42	50,6	54,6	28	170	208	622	3
91	4,7	8,4	17	3,2	30	13	44	46,2	51,2	29	208	197	595	3
92	4,7	17,2	16	2,9	33	14	44	49,9	53,1	34	182	206	612	3
93	5,1	10,3	22	1,3	49	17	35	67,3	65,8	27	210	198	592	3
94	6	45,7	7	4,1	39	8	15	51,1	77,3	23	798	47	155	2
95	6	54,8	7	2,4	36	9	14	47,4	77,2	23	782	53	165	2
96	5,9	22,6	9	2,5	29	7	15	38,5	72	20	795	47	158	2
97	5,7	49,3	8	2,3	33	6	16	41,3	72,1	21	818	42	140	1
98	6,4	41,6	9	3,5	42	9	11	54,5	83,2	20	755	60	185	2
99	5,9	26,8	5	2,1	23	10	15	35,1	70,1	20	778	53	169	2
100	6,5	25,5	5	1,9	38	15	12	54,9	82,1	21	776	50	174	2
101	5,6	18,2	9	2,7	29	10	19	41,7	68,7	21	668	78	254	2
102	5,7	24,3	14	4,1	50	13	21	67,1	76,2	31	528	112	360	2
103	5,1	11,6	15	3,8	42	12	33	57,8	63,7	34	393	140	467	3
104	4,8	9,5	15	2,8	32	11	40	45,8	53,4	31	323	160	517	3
105	5,3	14,5	14	3,8	53	10	27	66,8	71,2	31	418	135	447	3
106	5,6	22,9	10	6,3	53	15	25	74,3	74,8	30	365	148	487	3
107	5,2	9,6	15	2,7	39	12	28	53,7	65,7	25	430	128	442	3
108	5,4	8,6	11	4,1	49	14	28	67,1	70,6	30	370	140	490	3
109	4,9	22,6	11	2,5	28	8	27	38,5	58,8	18	570	103	327	2
110	5,4	15,5	10	4	40	13	24	57	70,4	23	513	113	374	3
111	6	18,1	7	4	60	17	22	81	78,6	36	101	207	692	3
112	6	14,4	14	4,1	59	18	22	81,1	78,7	34	93	210	697	3
113	5,5	13,2	17	3,1	49	14	29	66,1	69,5	36	78	212	710	3
114	5,7	13,9	20	3,7	58	16	27	77,7	74,2	37	81	205	714	3
115	5,6	20,9	9	2,9	45	13	29	60,9	67,7	37	106	207	687	3
116	5,1	10,5	30	1,7	33	9	36	43,7	54,8	23	110	213	677	3
117	5,8	15,3	13	4	58	16	26	78	75	34	68	215	717	3
118	5,6	20,9	15	4,9	59	17	29	80,9	73,6	37	71	217	712	3
119	6,2	15,2	41	4,8	80	19	19	103,8	84,5	37	73	225	702	3
120	4,5	4,4	4	0,8	8	3	39	11,8	23,2	16	763	50	187	2
121	4,9	3	5	1,5	9	5	22	15,5	41,3	12	780	46	174	2
122	4,7	2,7	5	0,7	8	4	29	12,7	30,5	14	773	55	172	2
123	4,2	3,8	5	0,6	3	1	40	4,6	10,3	12	770	53	177	2
124	4,4	2,6	4	0,4	5	2	31	7,4	19,3	11	776	55	169	2
125	5,3	13,6	7	1,2	17	6	17	24,2	58,7	16	815	43	142	1
126	5,2	9,3	5	1,5	15	5	17	21,5	55,8	16	818	44	138	1
127	5,7	6,5	4	1,7	16	6	14	23,7	62,9	15	830	42	128	1
128	5,4	5,4	4	1,5	16	5	15	22,5	60	15	828	40	132	1
129	6,4	11,1	5	2,1	43	8	11	53,1	82,8	18	845	33	122	1
130	6,3	5,4	5	1,8	33	8	11	42,8	79,6	17	858	34	108	1
131	5,5	16	7	1,3	21	7	14	29,3	67,7	15	832	40	128	1
132	5,7	6,2	5	1,7	20	4	13	25,7	66,4	13	848	27	125	1
133	5,5	8,8	5	1,8	21	7	15	29,8	66,5	16	855	35	110	1

134	5,4	28,1	9	2	18	7	16	27	62,8	17	852	33	115	1
135	4,7	41,3	8	1,6	13	4	21	18,6	47	14	812	46	142	1
136	5,9	9,4	8	2,4	55	21	24	78,4	76,6	33	75	230	695	3
137	5,7	15,1	13	4,2	60	19	27	83,2	75,5	46	84	238	678	3
138	5,4	13,6	13	3,6	52	16	34	71,6	67,8	47	84	238	678	3
139	5,3	11,8	9	3,1	44	11	34	58,1	63,1	44	104	233	663	3
140	5,4	11,8	13	5,2	45	13	32	63,2	66,4	46	109	233	658	3
141	5,2	9,1	11	3,3	47	13	37	63,3	63,1	45	117	230	653	3
142	5,4	11,8	62	4	63	18	31	85	73,3	44	107	233	660	3
143	5,8	13,2	12	5,4	71	26	25	102,4	80,4	47	107	233	660	3
144	5,5	12,5	17	3,8	55	16	28	74,8	72,8	41	122	230	648	3
145	5,4	18,4	38	3,3	53	15	34	71,3	67,7	45	127	227	646	3
146	6,4	27,4	7	2,5	69	32	17	103,5	85,9	45	67	243	690	3
147	6,4	21,7	9	2,5	77	29	18	108,5	85,8	64	109	233	658	3
148	6,4	21,9	7	3,3	75	33	17	111,3	86,7	47	67	243	690	3
149	6,1	16,4	9	2,8	60	25	22	87,8	80	44	67	243	690	3
150	6,3	27,4	5	1,9	67	27	19	95,9	83,5	42	127	227	646	3
151	6,3	29,3	5	1,9	58	22	20	81,9	80,4	42	152	220	628	3
152	6,6	26,5	7	1,5	78	33	17	112,5	86,9	40	137	225	638	3
153	6,5	20,1	9	1,8	81	29	17	111,8	86,8	44	134	226	640	3
154	6,2	20,9	6	2	60	22	20	84	80,8	43	127	227	646	3
155	6,3	46,4	6	1,5	68	28	20	97,5	83	41	84	238	678	3
156	5,9	17,2	5	2,5	56	24	24	82,5	77,5	43	104	233	663	3
157	5,8	12,4	13	4,1	60	23	27	87,1	76,3	38	102	235	663	3
158	5,2	14,3	20	3,8	55	13	35	71,8	67,2	37	61	245	694	3
159	5,3	16,6	7	4,9	55	15	31	74,9	70,7	25	391	150	459	3
160	5,6	10,4	5	4,1	38	10	24	52,1	68,5	23	513	120	367	2
161	5,4	11	5	4,9	40	14	26	58,9	69,4	23	416	142	442	3
162	5,9	7,6	8	2,2	46	23	22	71,2	76,4	30	190	212	598	3
163	5,4	21,9	8	1,2	39	18	35	58,2	62,4	32	205	207	588	3
164	5,5	13,2	8	0,8	45	23	35	68,8	66,3	28	208	207	585	3
165	5,3	18,9	8	1,2	43	20	36	64,2	64,1	30	120	230	650	3
166	4,1	23,9	43	2	10	2	38	14	26,9	11	808	47	145	1
167	5,4	21,4	22	2,4	33	9	20	44,4	68,9	21	738	62	200	2
168	5	32,7	13	2,3	20	4	19	26,3	58,1	14	848	37	115	1
169	5,4	41,7	18	2,7	25	5	17	32,7	65,8	18	792	50	158	2
170	4,8	27,4	9	1,6	14	4	23	19,6	46	20	830	42	128	1
171	4,9	61,8	8	1,5	21	4	23	26,5	53,5	22	800	48	152	2
172	5	52,6	10	1,7	18	6	21	25,7	55	18	818	44	138	1
173	4,4	52,1	12	1,7	13	2	26	16,7	39,1	17	828	42	130	1
174	4,8	33,8	9	1,9	18	4	23	23,9	51	14	808	47	145	1
175	5,9	47,7	6	1,6	32	10	12	43,6	78,4	17	830	42	128	1
176	5,1	39,2	6	2	19	5	21	26	55,3	16	815	45	140	1
177	6,1	14,3	7	4,1	61	17	19	82,1	81,2	31	102	236	662	3
178	5,3	17,2	6	3,2	26	10	24	39,2	62	21	740	65	195	2
179	5,4	46,8	14	4,9	41	16	27	61,9	69,6	26	515	123	362	2
180	5	15,5	12	3,1	28	12	28	43,1	60,6	25	458	134	408	3
181	4,9	20,5	25	2,6	26	10	32	38,6	54,7	21	562	110	328	2
182	5,4	18,7	7	2,7	41	13	24	56,7	70,3	23	625	95	280	2
183	5,3	27,6	10	3,2	45	12	27	60,2	69	25	595	103	302	2
184	6	30,8	10	2,4	38	12	14	52,4	78,9	15	775	55	170	2
185	5,2	63,6	16	1,6	26	5	20	32,6	62	13	788	54	158	2
186	5,8	27,2	5	1,9	29	6	14	36,9	72,5	18	808	50	142	1
187	5,9	36,1	8	2,6	29	7	13	38,6	74,8	15	775	57	168	2
188	6,2	23	5	2,5	37	9	13	48,5	78,9	17	805	50	145	1
189	6,4	33,9	7	2	37	9	11	48	81,4	15	800	52	148	1
190	4,8	15,8	10	6,6	39	14	41	59,6	59,2	29	472	138	390	3
191	5,1	21,9	10	4,5	38	10	30	52,5	63,6	23	465	140	395	3
192	5,3	31,3	12	5,2	41	12	23	58,2	71,7	22	562	113	325	2
193	5,2	27,1	8	3,4	28	8	22	39,4	64,2	19	708	74	218	2
194	5,6	53,5	10	1,9	31	8	16	40,9	71,9	13	760	62	178	2
195	5,3	56,1	6	2,7	15	5	17	22,7	57,2	11	830	42	128	1
196	5,8	39,2	5	2,2	24	8	15	34,2	69,5	16	785	55	160	2
197	5,2	37,4	6	2,1	20	8	17	30,1	63,9	14	830	42	128	1
198	4,7	61,1	8	3,2	12	4	25	19,2	43,4	15	795	53	152	2
199	5,3	45,7	7	2,2	18	6	16	26,2	62,1	13	800	50	150	1
200	5,7	41,7	5	1,8	27	8	17	36,8	68,4	18	790	55	155	2
201	5	51,4	6	3,1	15	5	23	23,1	50,1	17	810	50	140	1
202	5,3	9,7	17	3	50	18	34	71	67,6	36	70	242	688	3
203	5,3	8	26	5,9	50	18	33	73,9	69,1	39	90	238	672	3
204	5,3	8	20	3,4	51	19	32	73,4	69,6	33	70	242	688	3
205	5,3	9,5	48	2,9	62	19	33	83,9	71,8	41	58	244	698	3

206	5,4	11,1	19	2,3	58	19	31	79,3	71,9	38	80	240	680	3
207	5,5	4,1	15	3,4	60	19	28	82,4	74,6	34	70	242	688	3
208	5,3	9	18	2,9	49	17	36	68,9	65,7	34	62	246	692	3
209	5,6	4,8	16	2,1	55	19	28	76,1	73,1	35	75	240	685	3
210	5,4	13,5	26	3,3	52	20	29	75,3	72,2	30	70	242	688	3
211	5,5	11,7	15	6	47	20	26	73	73,7	32	90	238	672	3
212	5,2	9	17	2,2	46	15	36	63,2	63,7	32	80	240	680	3
213	5,3	7	24	3	50	20	33	73	68,9	33	85	240	675	3
214	5,1	6,4	10	3,1	28	9	24	40,1	62,6	17	740	65	195	2
215	5,7	8,6	11	2,8	59	25	20	86,8	81,3	25	552	116	332	2
216	4,6	6,7	14	2,1	12	9	32	23,1	41,9	17	658	87	255	2
217	5,3	12,6	10	2,1	30	15	20	47,1	70,2	19	772	60	168	2
218	5,2	14,8	4	1,6	20	8	17	29,6	63,5	15	840	42	118	1
219	5	5,8	4	1,1	22	8	24	31,1	56,4	22	802	50	148	1
220	5,8	19,9	4	1,3	33	15	15	49,3	76,7	16	862	36	102	1
221	5,4	12,1	5	1,2	26	11	17	38,2	69,2	16	868	32	100	1
222	5,1	6,8	4	1,2	20	9	19	30,2	61,4	13	865	33	102	1
223	5,3	13,6	4	1,5	19	8	13	28,5	68,7	14	828	42	130	1
224	5,1	10,8	4	1,1	20	7	21	28,1	57,2	16	860	35	105	1
225	5,3	9,5	6	1,1	29	12	20	42,1	67,8	22	845	40	115	1
226	5,3	13,7	5	1,4	21	7	16	29,4	64,8	20	888	30	82	1
227	5,1	7,4	11	2,2	29	13	40	44,2	52,5	26	52	236	712	3

As unidades de medida dos parâmetros estão representadas de acordo com os números sobrescritos: 1= mg dm^{-3} ; 2= mmol c dm^{-3} ; 3= g dm^{-3} ; 4= %; 5= g Kg^{-1} ; 6= Tipo

Figura A1. Comparação dos valores observados pelas técnicas primárias e preditos de pH nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método

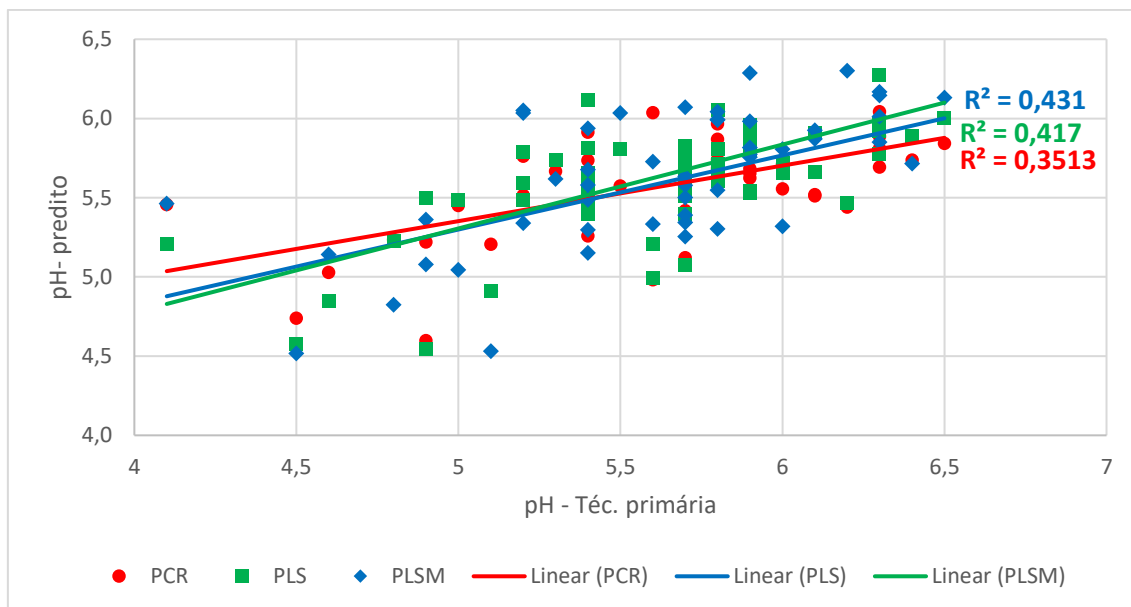


Figura A2. Comparação dos valores observados pelas técnicas primárias e preditos de Fósforo (P) nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

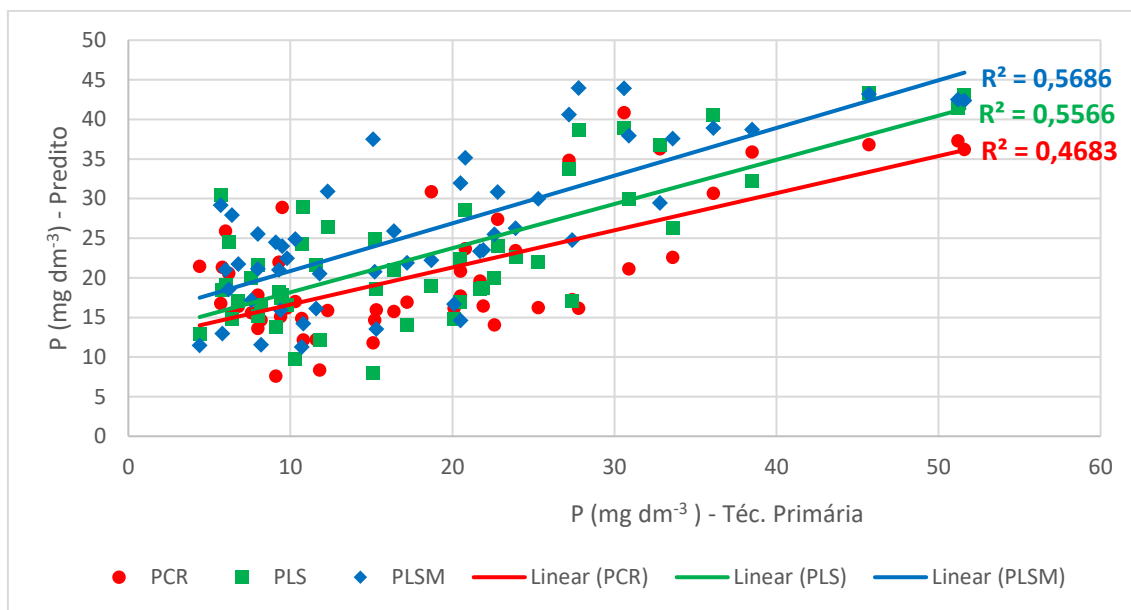


Figura A3. Comparação dos valores observados pelas técnicas primárias e preditos de Enxofre (S) nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

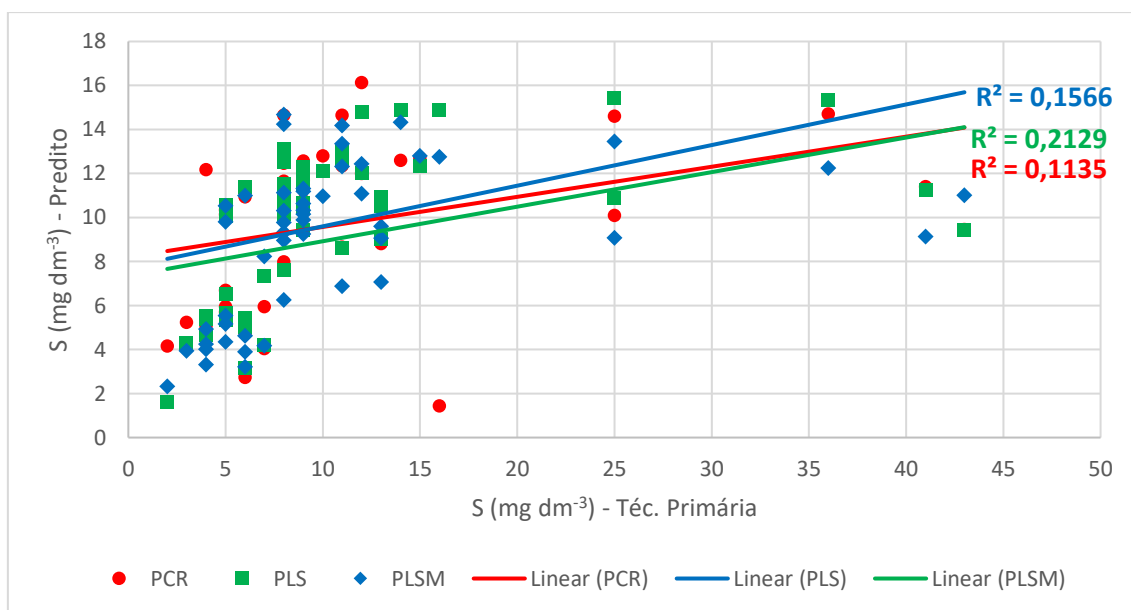


Figura A4. Comparação dos valores observados pelas técnicas primárias e preditos de Potássio (K) nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

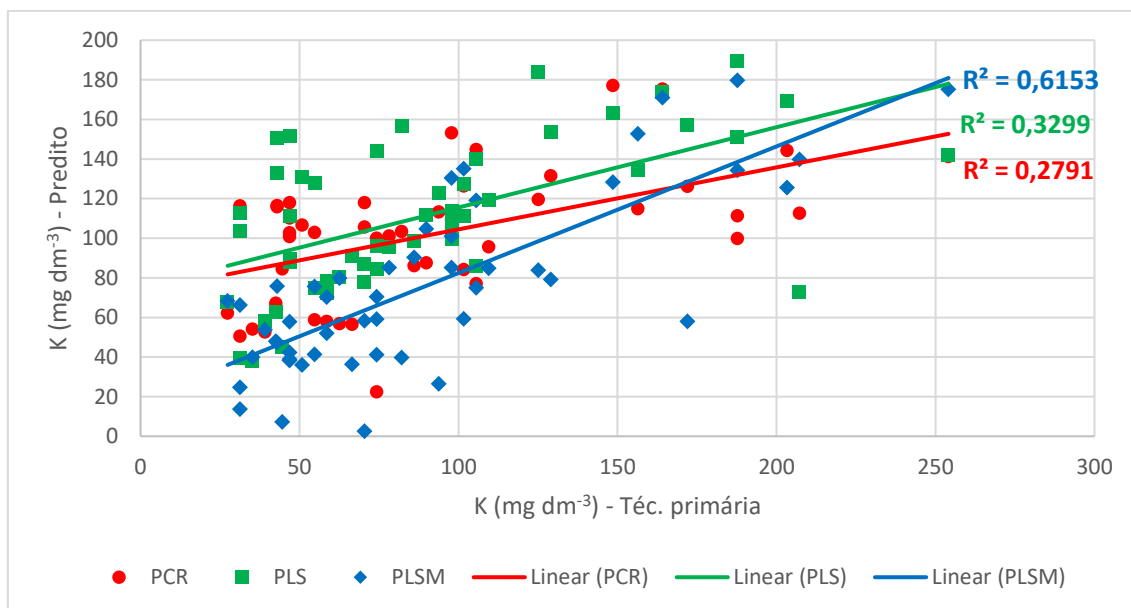


Figura A5. Comparação dos valores observados pelas técnicas primárias e preditos de Cálcio (Ca) nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

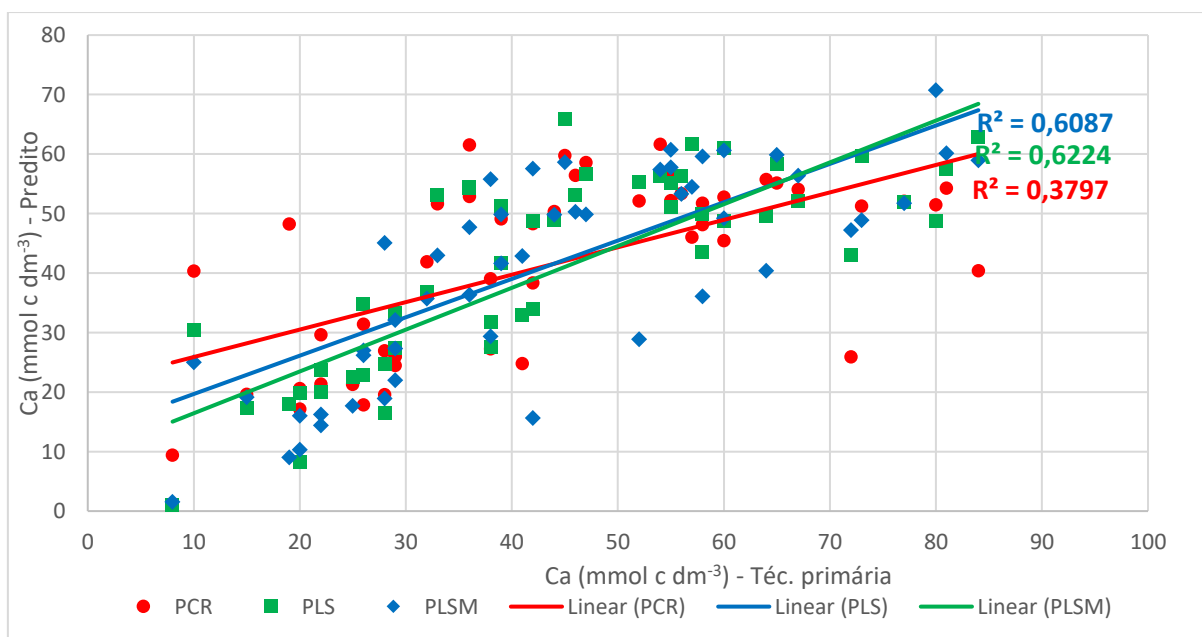


Figura A6. Comparação dos valores observados pelas técnicas primárias e preditos de Magnésio (Mg) nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

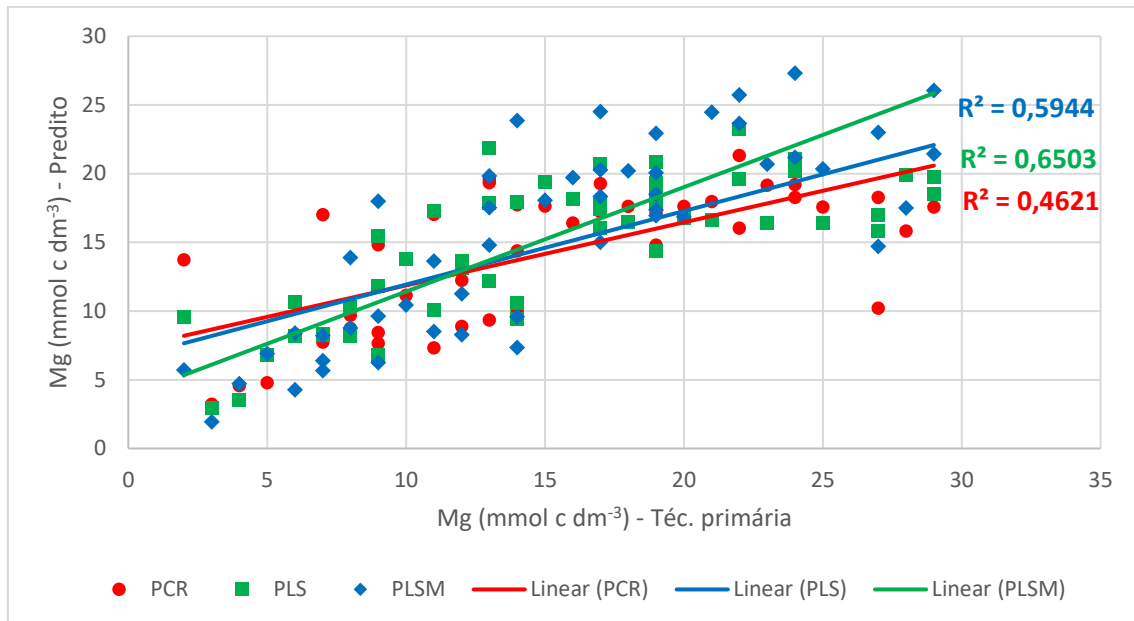


Figura A7. Comparação dos valores observados pelas técnicas primárias e preditos de Acidez (H+Al) nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

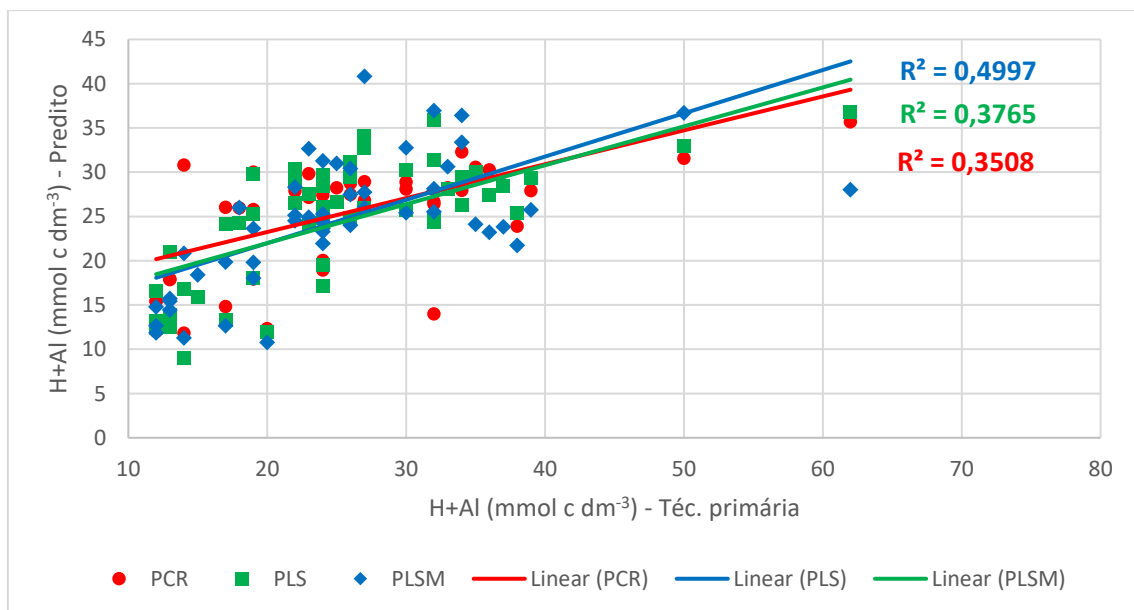


Figura A8. Comparação dos valores observados pelas técnicas primárias e preditos de Soma de Bases (SB) nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

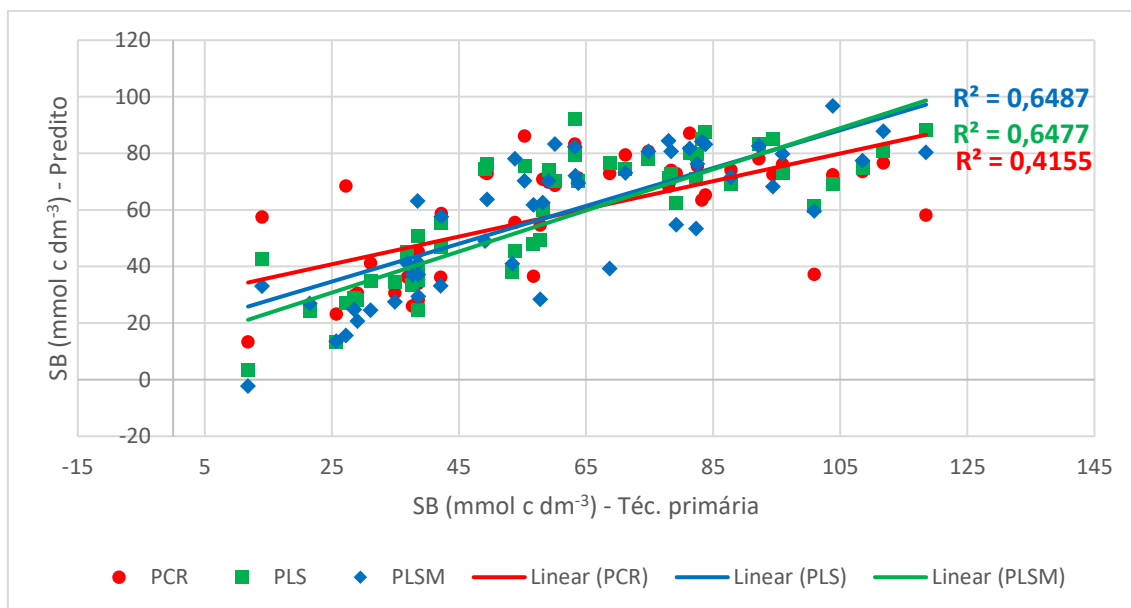


Figura A9. Comparação dos valores observados pelas técnicas primárias e preditos de V% nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

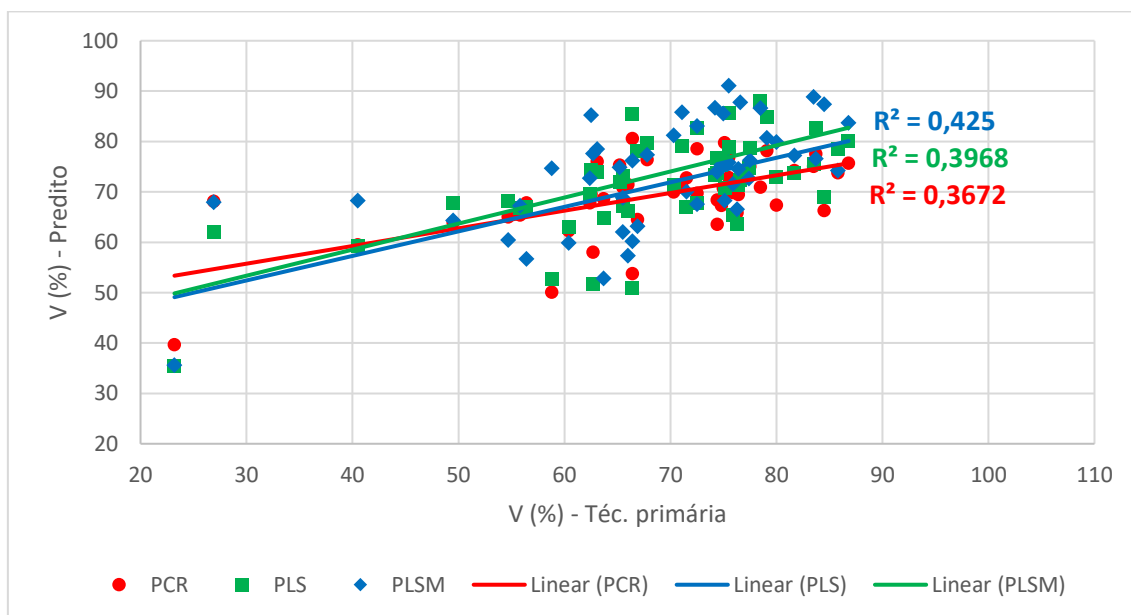


Figura A10. Comparação dos valores observados pelas técnicas primárias e preditos de Matéria Orgânica (M.O) nos modelos PCR, PLS e PLSM das 50 amostras de solo para validação do método.

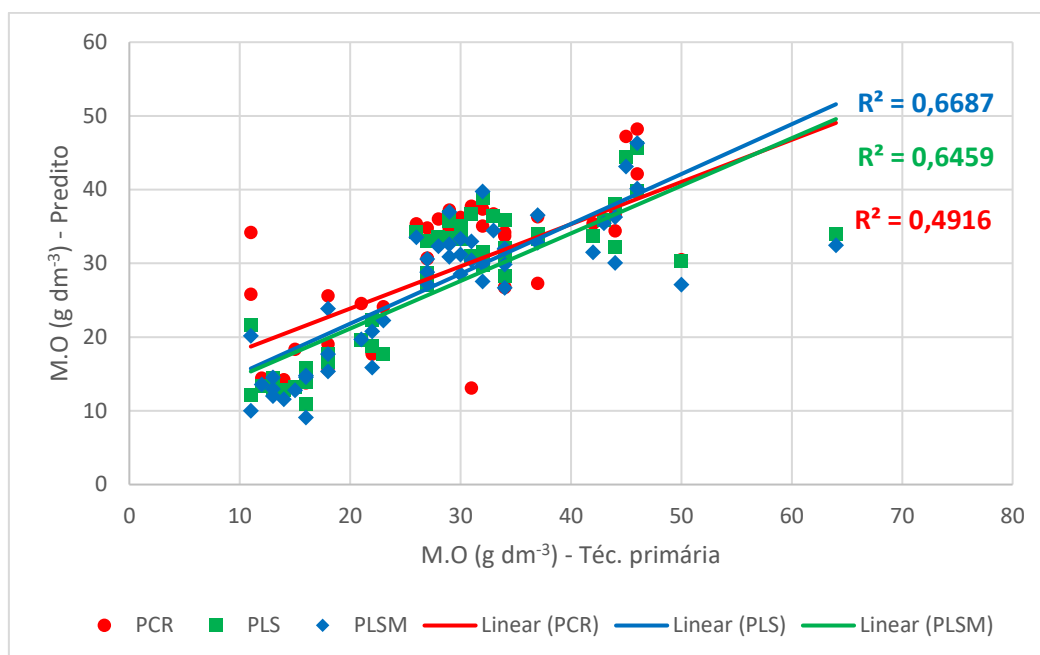


Tabela A2. ANOVA parâmetros preditos pelos modelos estatísticos e valores obtidos pelas técnicas primárias.

ANOVA						
	SQ	gl	MQ	F	valor-P	F crítico
MO	6,76E-09	3	2,25E-09	3,3E-11	1	2,627
Areia	9,79E-06	3	3,26E-06	3,37E-11	1	2,627
Silte	5,18E-07	3	1,73E-07	2,58E-11	1	2,627
Argila	1,08E-05	3	3,61E-06	6,87E-11	1	2,627
Tipo de solo	2,59E-11	3	8,65E-12	1,5E-11	1	2,627

SQ: Soma dos Quadrados; gl: graus de liberdade; MQ: média dos quadrados;

Figura A11. Comparação dos valores observados e preditos de pH nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*

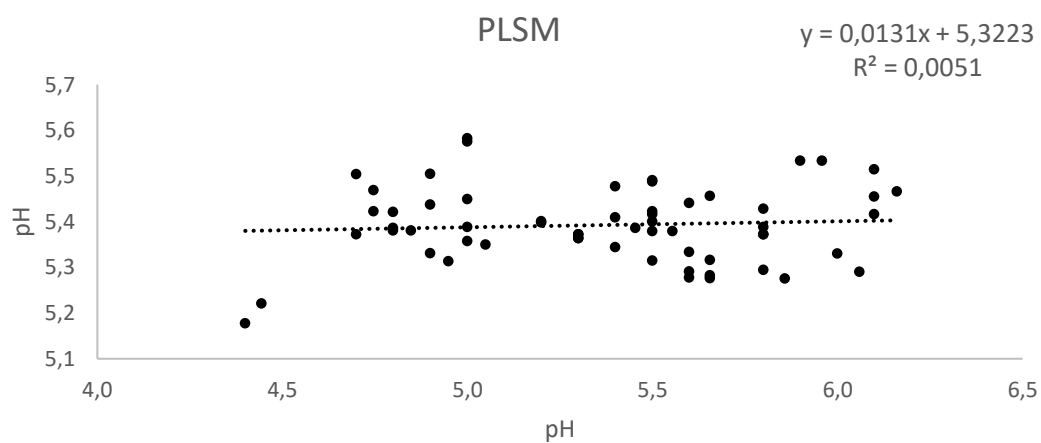
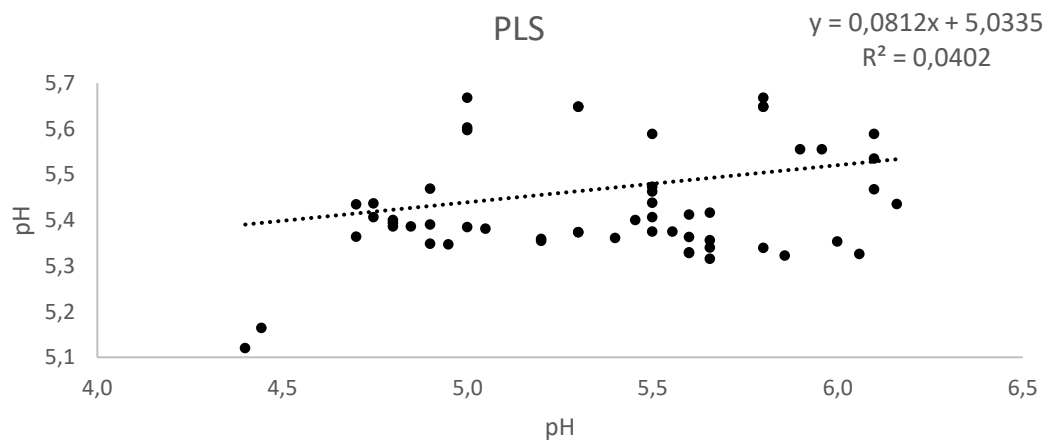
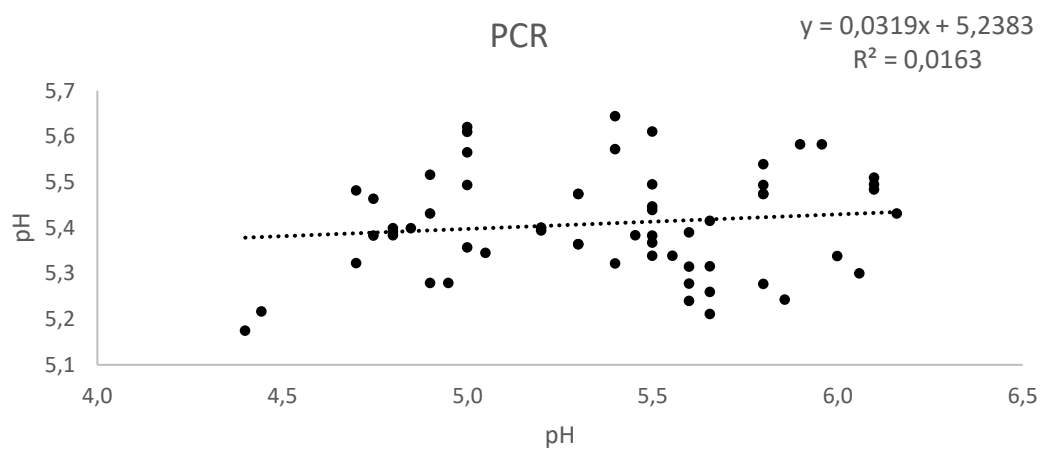


Figura A12. Comparação dos valores observados e preditos de S nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*

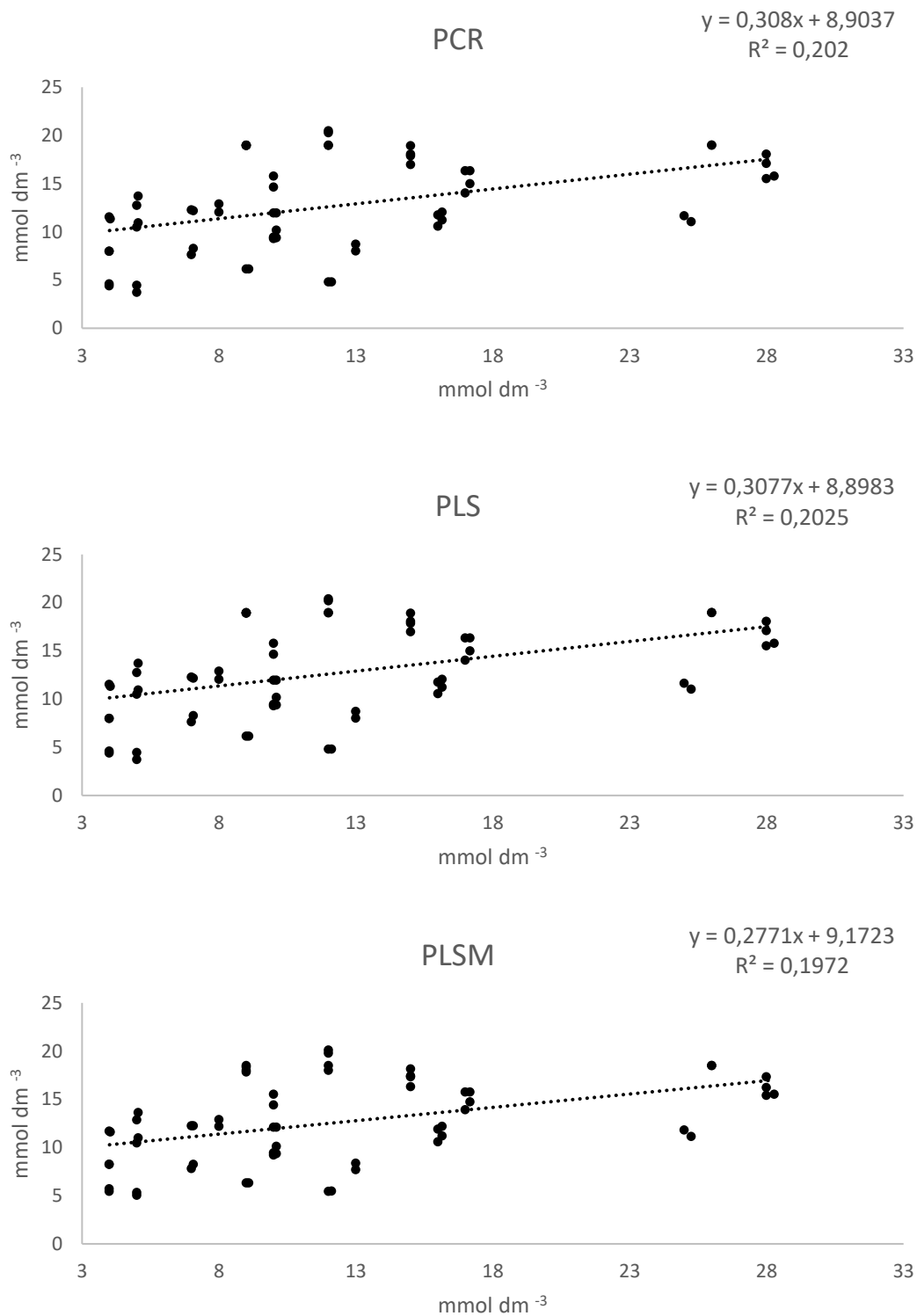


Figura A13. Comparação dos valores observados e preditos de K nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*

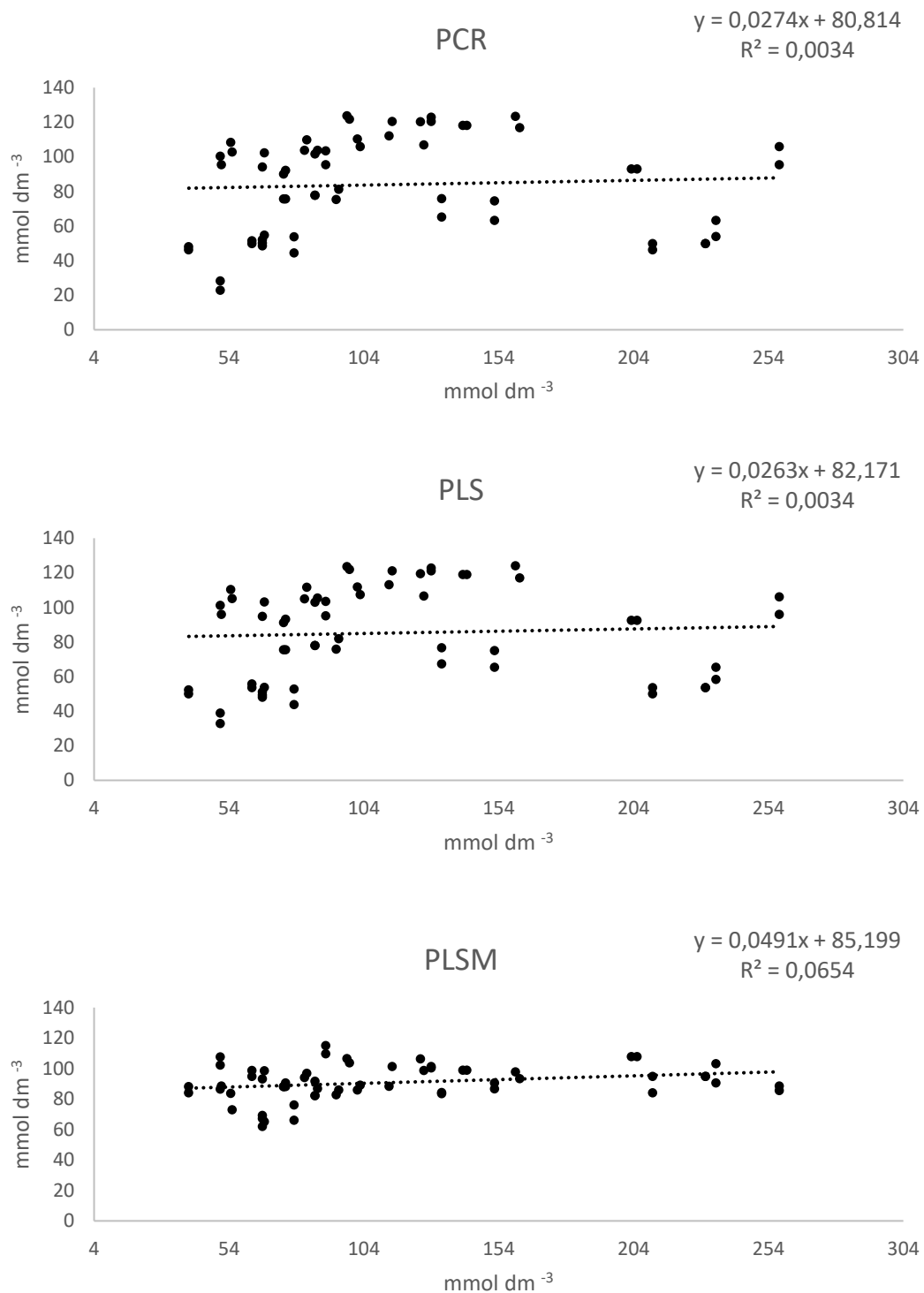


Figura A14. Comparação dos valores observados e preditos de Ca nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*

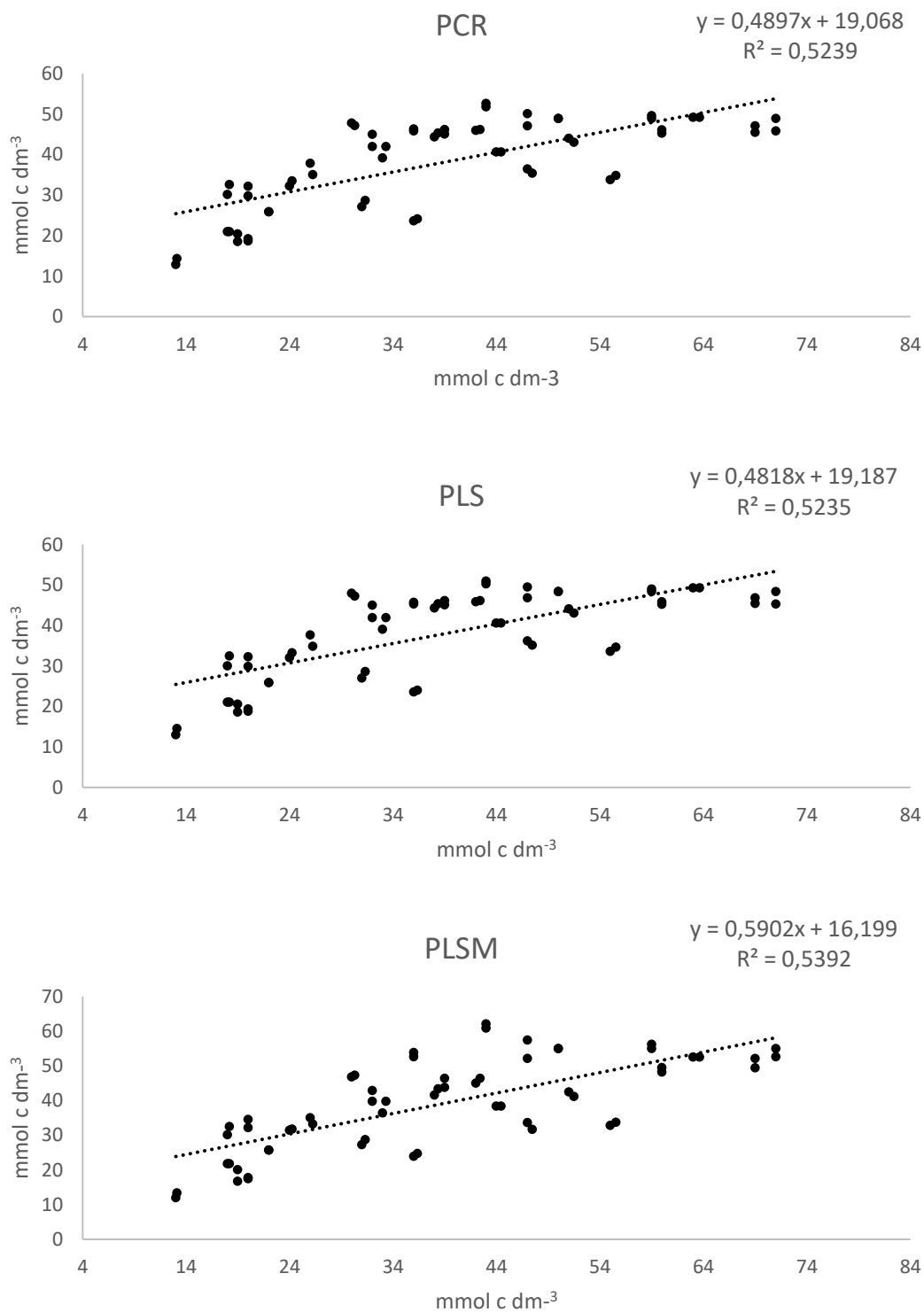


Figura A15. Comparação dos valores observados e preditos de Mg nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*

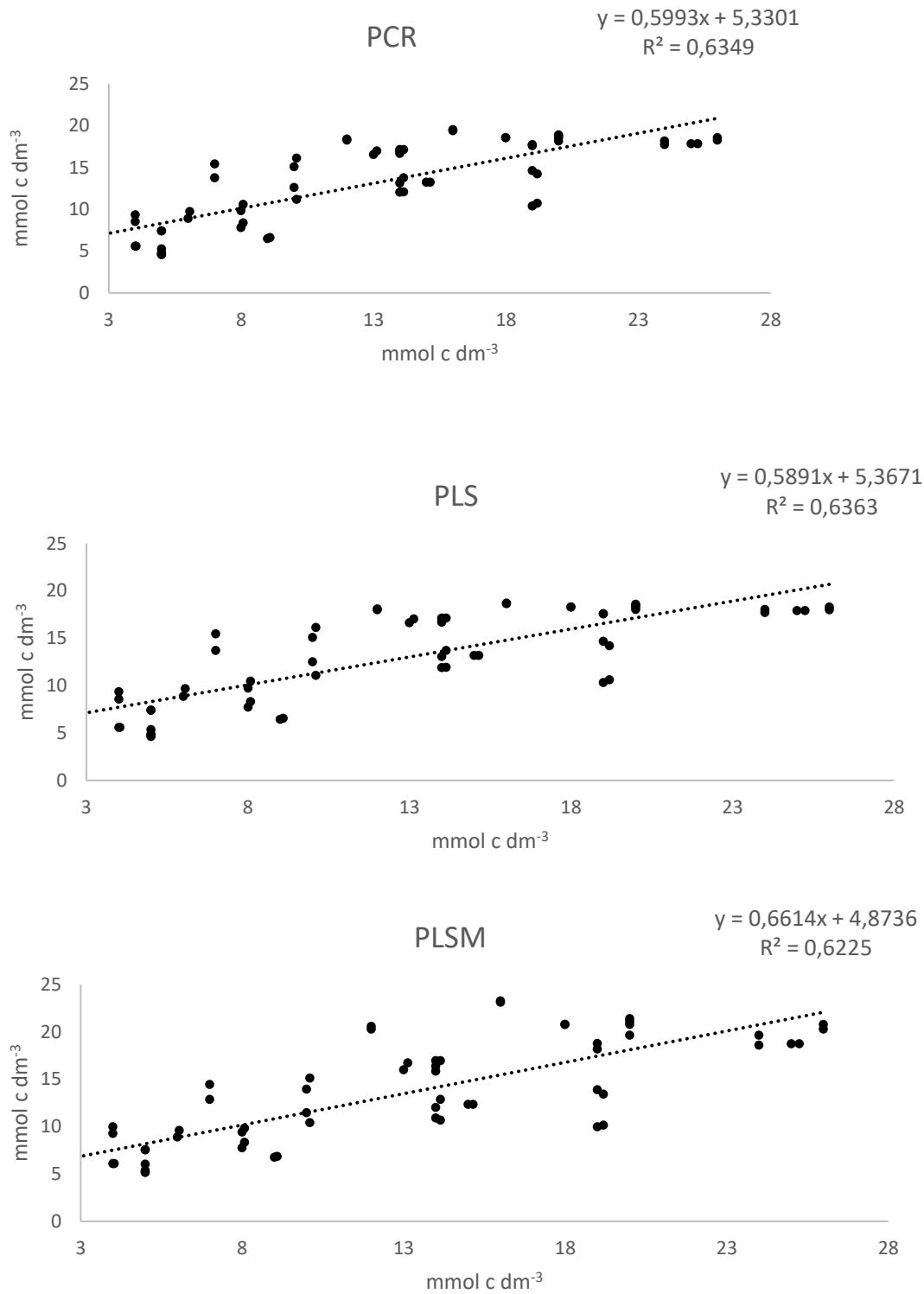


Figura A16. Comparação dos valores observados e preditos de H+Al nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*

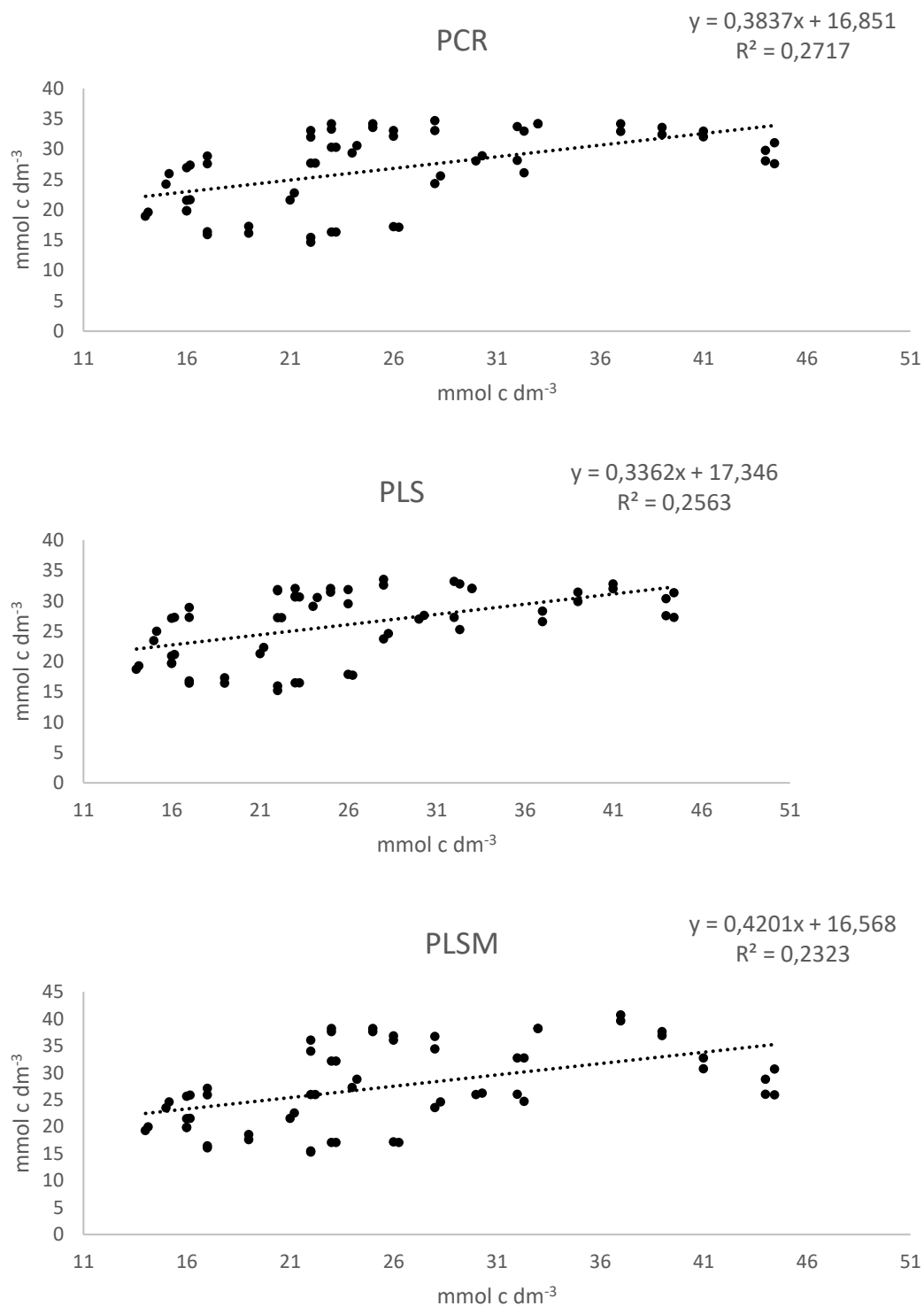


Figura A17. Comparação dos valores observados e preditos de SB nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*

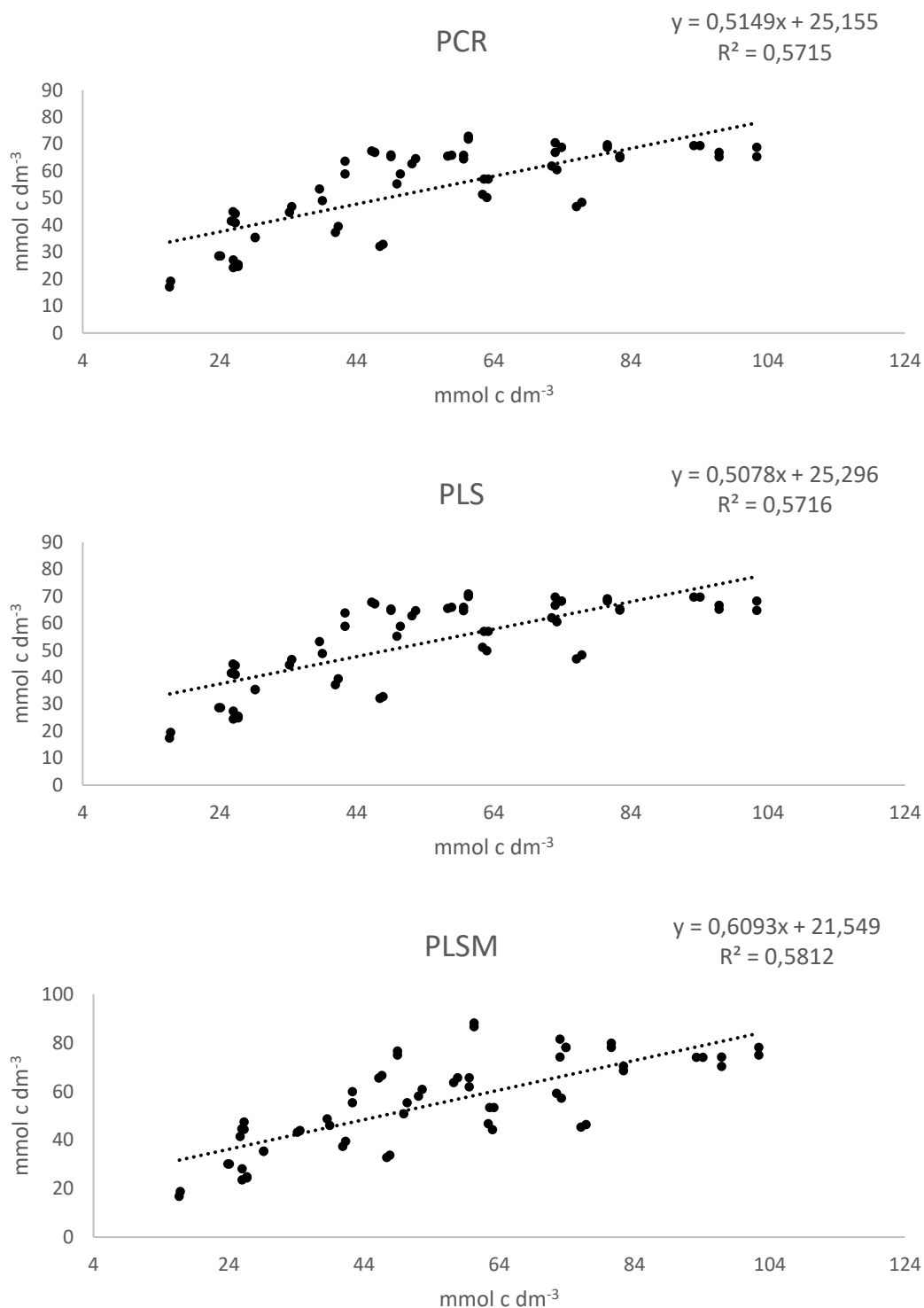


Figura A18. Comparação dos valores observados e preditos de V% nos modelos PCR, PLS e PLSM das 40 amostras de solo para validação do método utilizando o Sistema *DIA-ChemColour*

