

UFRRJ

**INSTITUTO DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM
AGRONOMIA - CIÊNCIA DO SOLO**

DISSERTAÇÃO

**Predição de C, N e P do solo por Espectroscopia
vis-NIR**

Marcela Lopes Lázaro

2018



**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA
CIÊNCIA DO SOLO**

PREDIÇÃO DE C, N E P DO SOLO POR ESPECTROSCOPIA VIS-NIR

MARCELA LOPES LÁZARO

Sob a Orientação da Professora
Érika Flávia Machado Pinheiro

e Coorientação do Pesquisador
Gustavo de Mattos Vasques

Dissertação submetida como requisito parcial para obtenção do grau de **Mestra**, no Curso de Pós-Graduação em Agronomia – Ciência do Solo, Área de Concentração em Biologia do Solo.

Seropédica, RJ
Agosto de 2018

Universidade Federal Rural do Rio de Janeiro
Biblioteca Central/Seção de Processamento Técnico

Ficha catalográfica elaborada
Com os dados fornecidos pelo(a) autor(a)

L431p	<p>Lázaro, Marcela Lopes, 1991- Predição de carbono, nitrogênio e fósforo no solo por espectroscopia de reflectância vis-NIR/Marcela Lopes Lázaro. – Seropédica, 2018. 52 f. : il.</p> <p>Orientadora: Érika Flávia Machado Pinheiro. Dissertação (Mestrado). – – Universidade Federal Rural do Rio de Janeiro, Programa de Pós-Graduação em Agronomia – Ciência do Solo, 2018.</p> <p>1. Sensoriamento próximo. 2. Agricultura orgânica. 3. Matéria orgânica. I. Pinheiro, Érika Flávia Machado, 1975-, orient. II. Universidade Federal Rural do Rio de Janeiro. Programa de Pós-Graduação em Agronomia – Ciência do Solo III. Título.</p>
-------	--

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

É permitida a cópia parcial ou total desta Tese, desde que seja citada a fonte.



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA (CIÊNCIAS DO SOLO)



**HOMOLOGAÇÃO DE DISSERTAÇÃO DE MESTRADO Nº 24 / 2024 - CPGACS
(12.28.01.00.00.00.27)**

Nº do Protocolo: 23083.017410/2024-06

Seropédica-RJ, 02 de abril de 2024.

UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DO SOLO

MARCELA LOPES LÁZARO

Dissertação submetida como requisito parcial para obtenção do grau de **Mestra**, no Programa de Pós-graduação em Agronomia-Ciência do Solo, Área de Concentração em Manejo do Solo e Qualidade Ambiental.

DISSERTAÇÃO APROVADA EM: 30/08/2018.

Gustavo de Mattos Vasques. Dr. Embrapa Solos
(Coorientador)

Marcos Bacis Ceddia. Dr. UFRRJ

Fabrcio Terra. Dr. UFVJM

(Assinado digitalmente em 04/04/2024 08:52)
MARCOS BACIS CEDDIA
PROFESSOR DO MAGISTERIO SUPERIOR
DATS (11.39.00.35)
Matrcula: 1220296

(Assinado digitalmente em 03/04/2024 10:47)
FABRÍCIO DA SILVA TERRA
ASSINANTE EXTERNO
CPF: 818.150.690-15

(Assinado digitalmente em 02/04/2024 21:06)
GUSTAVO DE MATTOS VASQUES
ASSINANTE EXTERNO
CPF: 084.272.437-07

Visualize o documento original em <https://sipac.ufrj.br/public/documentos/index.jsp> informando seu número: **24**, ano: **2024**, tipo: **HOMOLOGAÇÃO DE DISSERTAÇÃO DE MESTRADO**, data de emissão: **02/04/2024** e o código de verificação: **912df98619**

DEDICATÓRIA

À memória da minha mãe, Virgínia Campos Lopes, dedico.

AGRADECIMENTOS

A Deus, pela experiência da vida;

À minha família, em especial a minha avó Thereza de Jesus C. Lopes e ao meu pai Marcelo G. Lázaro pela criação, dedicação, cuidados e apoio incondicional, inclusive para meu ingresso na Pós-Graduação;

Aos meus orientadores Érika Flávia Machado Pinheiro e Gustavo M. Vasques pela apresentação a uma nova área da Ciência do solo através da proposta, desenvolvimento e acompanhamento deste trabalho, além do conhecimento e conselhos transmitidos, paciência e confiança;

Ao professor Marcos Bacis Ceddia, ao pesquisador José Guilherme Marinho Guerra e a Maria Gabriela da Mata, bem como a Embrapa Agrobiologia e Sistema Integrado de Produção Agroecológica e a todos os estagiários e funcionários envolvidos no desenvolvimento, montagem e manutenção do módulo experimental, realização e descrição de análises;

À Embrapa Solos e ao pesquisador Silvio L. Tavares pelo espaço físico, equipamentos e infra-estrutura cedidos para o desenvolvimento deste trabalho;

Ao estagiário Leandro Ribeiro dos Santos, pela ajuda em parte das leituras espectrais do solo;

Ao colega Dr. Jean Michel Bueno pelo auxílio com a linguagem estatística R;

Aos professores que fizeram parte da minha jornada, que me inspiraram e acrescentaram bastante à minha trajetória acadêmica;

Ao CNPQ pela concessão da bolsa;

E a todos e todas que participaram, direta ou indiretamente, da construção deste trabalho, mesmo não sendo citados, foram importantes para a sua realização.

RESUMO

LÁZARO, Marcela Lopes. **Predição de carbono, nitrogênio e fósforo no solo por espectroscopia de reflectância vis-NIR**. 2020. 52p. Dissertação (Mestrado em Agronomia, Ciência do Solo). Instituto de Agronomia, Departamento de Solos, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2018.

A espectroscopia de reflectância na região do visível e infravermelho próximo (vis-NIR) vem sendo utilizada como método para predição de atributos do solo, sejam eles químicos, físicos, biológicos ou mineralógicos, apresentando resultados rápidos, sem gerar resíduos e não destrutivos quando comparados aos métodos convencionais de análise. O objetivo geral deste trabalho foi avaliar a acurácia da espectroscopia vis-NIR, aliada ou não a diferentes técnicas de pré-processamentos espectrais, na predição dos teores de C, N e P do solo. Este trabalho foi realizado em amostras de terra de um Planossolo Háplico, sob produção orgânica intensiva de hortaliças, no Sistema Integrado de Produção Agroecológica (SIPA), no município de Seropédica (RJ). Foram coletadas 294 amostras de terra, georreferenciadas num grid de 5x5m, na profundidade de 0-20 cm para a leitura espectral no espectrorradiômetro Field Spec 4. Os espectros, referentes a cada uma das amostras, foram gerados, exportados e analisados quantitativamente através de regressão por mínimos quadrados parciais (PLS), treinados por validação cruzada, validados com dados externos. Paralelamente, com a finalidade de melhorar as predições, foram testados seis pré-processamentos: absorvância, suavização, derivadas por Savitzky-Golay, variação normal padrão, correção multiplicativa do sinal e correção do contínuo. Os modelos de predição para C, N e P apresentaram valores médios de R^2 de, 0,82; 0,88 e 0,78; RMSE médios de 1,16 g kg⁻¹; 0,09 g kg⁻¹; 22,26 mg g⁻¹ e RPIQ médios de 4,00; 5,34 e 4,46, respectivamente. Os resultados mostraram que a espectroscopia de reflectância vis-NIR é uma técnica que possui acurácia, sendo promissora para predição de C, N e P do solo inclusive sem a necessidade do uso de pré-processamentos espectrais.

Palavras-chave: Sensoriamento próximo. Agricultura orgânica. Matéria orgânica.

ABSTRACT

LÁZARO, Marcela Lopes. **Prediction of soil carbon, nitrogen and phosphorus by vis-NIR spectroscopy**. 2020. 52p. Dissertation (Master in Agronomy, Soil Science) Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2018.

Visible and Near infrared (vis-NIR) reflectance spectroscopy has been used as a method to predict chemical, physical, biological or mineralogical soil attributes with fast and non-destructive results comparing to conventional methods of analysis in Brazil. The aim of this work was to evaluate and determine the accuracy of vis-NIR spectroscopy with and without spectral pre-processing techniques, to predict of C, N and P soil contents. This work was carried out in the SIPA (Integrated System of Agroecological Production) in Seropédica – RJ, Brazil. A total of 294 soil samples were collect at 20 cm depth for chemical characterization of C, N and P and spectral readings, that were made on the Field Spec 4 spectroradiometer. The spectra for each sample were generated and quantitatively analysed by Partial Least Square regression, trained by cross validation and the best models were select and then applied to new soil samples in SIPA (external data). In parallel, six spectral pre-processes were tested to improve predictions: absorbance, smoothing, Savitzky-Goay derivatives, standard normal variation, multiplicative correction of the signal and continuum removal. The prediction models for C, N and P presented mean values of R^2 of 0.82; 0.88 and 0.78; mean RMSE of 1.16 g kg⁻¹; 0.09 g kg⁻¹; 22.26 mg g⁻¹ and RPIQ of 4.00; 5.34 and 4.46 for C, N and P, respectively. The results suggest that vis-NIR reflectance spectroscopy is a promising technique for the prediction of soil C, N and P and without the need for spectral pre-processing.

Keywords: Proximal sensing. Organic agriculture. Organic matter.

LISTA DE ABREVIACÕES E SÍMBOLOS

Vis	=	Visível (<i>Visible</i>)
NIR	=	Infravermelho próximo (<i>Near infra-red</i>)
SIPA	=	Sistema integrado de produção agroecológica
TFSA	=	Terra fina seca ao ar
MOS	=	Matéria orgânica do solo
PP	=	Pré-processamento
ERD	=	Espectroscopia de reflectância difusa
REF	=	Refletância
ABS	=	Absorbância
SMO	=	Suavização (<i>Smoothing</i>)
SGD	=	Derivadas Savitzky-Golay (<i>Savitzky-Golay derivative</i>)
SNV	=	Varição normal padrão (<i>Standard normal variate</i>)
MSC	=	Correção multiplicativa de sinal (<i>Multiplicative scatter correction</i>)
CRR	=	Correção do contínuo (<i>Continuum removal</i>)
PLS	=	Mínimos quadrados parciais (<i>Partial least square</i>)
NC	=	Número de componentes PLS
Q1	=	1º quartil
Q3	=	3º quartil
R ²	=	Coefficiente de determinação
RMSE	=	Raiz quadrada do erro médio (<i>Root mean square error</i>)
RPIQ	=	Razão da performance interquartil (<i>Ratio of performance to interquartile distance</i>)

ÍNDICE DE FIGURAS

- Figura 1.** Ilustração da luz incidindo em uma amostra sólida e as frações de energia refletidas de forma difusa pela mesma e captadas por um sensor (Adaptado de STEMBERG, VISCARRA ROSSEL, MOURAZEN 2010).5
- Figura 2.** Espectro eletromagnético com detalhe para as porções do visível (350 a 750 nm), NIR - infravermelho próximo (750 a 2500 nm), além do MIR – infravermelho médio (2500 a 25000nm) e FIR – infravermelho distante (25000 a 100000 nm).6
- Figura 3.** Curvas espectrais típicas de solos nas porções vis (A) e NIR (B) do espectro eletromagnético. (Adaptado de VISCARRA ROSSEL et al, 2011).6
- Figura 4.** Número de publicações científicas abordando o uso de técnicas espectroscópicas Vis-NIR em solos mundiais e em solos brasileiros (Dados retirados do banco de dados Scopus em dezembro de 2019).8
- Figura 5.** Diagrama representando o processo de treinamento de um modelo quimiométrico. (Adaptado de FERREIRA, 1999).10
- Figura 6.** Localização do módulo experimental de produção orgânica intensiva de hortaliças.12
- Figura 7.** Organização espacial das sub-áreas do módulo experimental compostas por capineira, gliricídia, canteiros a céu aberto e telados com hortícolas e área de compostagem orgânica (C.O.) (Adaptado de MATA, 2012).13
- Figura 8.** Croqui dos pontos de coleta do módulo experimental georreferenciados. Todos os pontos (vermelhos e azuis) representam as amostras do grupo 2, coletadas em 2013 e os pontos em vermelho representam as amostras do grupo 1, coletadas em 2014.14
- Figura 9.** Da esquerda para a direita: amostras de TFSA colocadas e niveladas em placas de Petri; estufa; conjunto de equipamentos necessários às leituras espectrais, sendo A- espectrorradiômetro Field Spec 4; B- disco-mesa giratória suporte para as placas de Petri; C- computador acoplado ao espectrorradiômetro e D - placa de spectralon.16
- Figura 10.** Esquema mostrando o funcionamento da validação cruzada por leave-one-out.17
- Figura 11.** Fluxograma das etapas realizadas incluindo as análises químicas laboratoriais, assim como, as espectrais e quimiométricas desenvolvidas no presente trabalho.19
- Figura 12.** Histogramas para os dados de C, N e P do solo das amostras dos anos de 2013 e 2014, respectivamente da esquerda para a direita.21
- Figura 13.** Frações totais de areia, argila e silte do solo do módulo de produção de hortaliças.22
- Figura 14.** Comportamento espectral total e médio (Med), com \pm desvio padrão (DP), para o módulo como um todo (B, H) e para cada área do módulo experimental sendo elas: Capineira (C, I); Gliricídia (D, J); Canteiros (E, L) e Telados (F, M) para os anos de 2013 e 2014.23
- Figura 15.** Comportamento espectral médio dos solos para cada área do módulo experimental para os anos de 2013 (A), 2014 (B) e as médias espectrais do módulo como um todo para as amostras referentes aos dois anos.24
- Figura 16.** Curvas espectrais médias de REF espectral e após a adição dos pré-processamentos espectrais de ABS, SMO, SGD, SNV, MSC, CRR.26
- Figura 17.** Amostras identificadas como outliers em vermelho para o conjunto de dados de treinamento (trei) e validação (val) para C e somente validação (val) para N e P.27
- Figura 18.** Gráficos de componentes PLS e RMSE para os treinamentos dos modelos preditivos de C.28

Figura 19. Treinamento dos modelos de predição de C a partir de dados de REF, ABS, SMO, SGD, SNV, MSC e CRR.29

Figura 20. Validação dos modelos preditivos de C para os dados de REF, ABS, SMO, SGD, SNV, MSC e CRR, para os anos de 2013 e 2014.32

Figura 21. Gráficos de componentes PLS e RMSE para os treinamentos dos modelos preditivos de N.34

Figura 22. Treinamento dos modelos de predição de N a partir de dados de REF, ABS, SMO, SGD, SNV, MSC e CRR.35

Figura 23. Validações dos modelos preditivos de N para os dados de REF, ABS, SMO, SGD, SNV, MSC, CRR para os anos de 2013 e 2014.37

Figura 24. Gráficos de componentes PLS e RMSE para os treinamentos dos modelos preditivos de P.39

Figura 25. Treinamento dos modelos de predição de P a partir de dados de REF, ABS, SMO, SGD, SNV, MSC e CRR.40

Figura 26. Validações dos modelos preditivos de C para os dados de REF, ABS, SMO, SGD, SNV, MSC, CRR para os anos de 2013 e 2014.42

ÍNDICE DE TABELAS

- Tabela 1.** Principais trabalhos no país abordando espectroscopia vis-NIR para predição de C, N e P do solo nos últimos 5 anos.9
- Tabela 2.** Conjunto de dados particionado em treinamento e validação para as amostras referentes ao ano de 2014 e conjuntos globais de dados para as amostras de 2013 e 2014.16
- Tabela 3.** Estatística descritiva dos teores de C, N e P do solo, para os anos de 2013 e 2014.20
- Tabela 4.** Frações granulométricas do solo referentes a cada área (capineira, gliricidia, canteiros e telados) do módulo experimental.21
- Tabela 5.** Reflectância média (REF) \pm 1 desvio padrão (DP) e matéria orgânica do solo (MOS) para as subáreas do módulo experimental.25
- Tabela 6.** Parâmetros estatísticos obtidos a partir do treinamento e validação dos modelos preditivos de C, bem como o número de componentes PLS escolhido para modelo.30
- Tabela 7.** Validações dos modelos preditivos de C para os dados de REF, ABS, SMO, SGD, SNV, MSC, CRR para os anos de 2014 e 2013.33
- Tabela 8.** Parâmetros estatísticos obtidos a partir do treinamento e validação dos modelos preditivos, bem como o número de componentes PLS escolhido e utilizado para a predição de C, N e P do solo.36
- Tabela 9.** Validações dos modelos preditivos de N para os dados de REF, ABS, SMO, SGD, SNV, MSC, CRR para os anos de 2014 e 2013.38
- Tabela 10.** Parâmetros estatísticos obtidos a partir do treinamento e validação dos modelos de predição das concentrações de P do solo.41
- Tabela 11.** Validações dos modelos preditivos de P para os dados de REF, ABS, SMO, SGD, SNV, MSC, CRR para os anos de 2014 e 2013.43

SUMÁRIO

1. INTRODUÇÃO1
2. REVISÃO DE LITERATURA3
 - 2.1. Matéria Orgânica do Solo3
 - 2.2. Sensoriamento Próximo do Solo4
 - 2.3. Espectroscopia de Reflectância Difusa4
 - 2.4. Espectroscopia vis-NIR5
 - 2.4.1. Vis-NIR e MOS em solos brasileiros7
 - 2.5. Quimiometria9
 - 2.5.1. Pré-processamento de curvas espectrais10
 - 2.5.2. Regressão por Mínimos quadrados parciais (PLSR)11
3. MATERIAL E MÉTODOS12
 - 3.1. Caracterização da Área de Estudo12
 - 3.2. Amostragem e Caracterização do Solo13
 - 3.3. Análises Químicas do Solo14
 - 3.4. Estatística Descritiva dos Dados15
 - 3.5. Análises Granulométricas do Solo15
 - 3.6. Análises Espectrais do Solo15
 - 3.7. Pré-Processamento dos Dados Espectrais do Solo16
 - 3.8. Modelagem Quimiométrica de C, N e P do Solo16
 - 3.8.1. Particionamento dos dados16
 - 3.8.2. Análise multivariada17
 - 3.8.3. Construção e validação dos modelos de predição PLSR17
 - 3.8.4. Avaliação dos modelos de predição PLSR18
 - 3.8.5. Remoção de outliers18
4. RESULTADOS E DISCUSSÃO20
 - 4.1. Estatística Descritiva20
 - 4.2. Dados Granulométricos21
 - 4.3. Dados Espectrais22
 - 4.4. Pré-Processamento dos Dados Espectrais25
 - 4.5. Modelagem26
 - 4.5.1. Identificação de outliers26
 - 4.5.2. Treinamento e validação dos modelos de predição de C27
 - 4.5.3. Treinamento e validação dos modelos de predição de N33
 - 4.5.4. Treinamento e validação dos modelos de predição de P38
5. CONCLUSÕES44
6. REFERÊNCIAS BIBLIOGRÁFICAS45

1. INTRODUÇÃO

O solo tem grande influência sobre a produtividade agrícola nutrindo e fomentando toda a vida existente no planeta (AL e MOLDENHAUER, 1987; SPARKS, 1988). O Brasil, por suas características climáticas e ambientais, apresenta solos favoráveis à agricultura, com pouco mais de 7,5% do seu território, o que corresponde a, aproximadamente, 64.000.000 ha, destinados a esse fim (EMBRAPA TERRITORIAL, 2016; NASA, 2017). É sabido, no entanto, que a agricultura sob o cultivo convencional é uma das grandes responsáveis pela degradação das propriedades químicas e físicas do solo. Em contrapartida, há um aumento no estudo de sistemas de manejo que sejam capazes de balancear as necessidades do solo e das culturas além de focar na sustentabilidade dos recursos naturais e na produtividade das culturas a longo prazo (LAL & PIERCE, 1991).

Nesse contexto, técnicas agroecológicas de manejo vêm ganhando espaço por partirem de um movimento social que implica na gestão da agricultura a partir de um modelo sustentável, no qual não se utiliza substâncias sintéticas e haja o respeito pelo meio ambiente e pelos ciclos naturais, com ênfase nas complexas relações existentes entre pessoas, cultivos, solo e a água, ao invés da maximização da produção de uma atividade em especial (VEIGA, 2003; ASSIS, 2006). Com base nesses princípios, em 1993 foi criado o Sistema Integrado de Produção Agroecológica (SIPA), em uma parceria da Embrapa Solos, Embrapa Agroecologia, Pesagro-RJ, Universidade Federal Rural do Rio de Janeiro (UFRRJ) e Colégio Técnico da UFRRJ, para o desenvolvimento e incentivo do estudo, pesquisa e extensão em agroecologia e agricultura orgânica. Dentro do SIPA, sob um Planossolo, foi instalado um módulo de produção orgânica intensiva de hortaliças, como protótipo de uma produção em escala real visando estreitar os estudos em agroecologia de forma integrada com o produtor orgânico, desde o cultivo até a venda dos produtos. Também é objetivo do módulo gerar conhecimentos que contribuam para a sustentabilidade dos sistemas agroecológicos e, entre outros aspectos, para o entendimento da dinâmica da matéria orgânica do solo (MOS) através do monitoramento de nutrientes utilizando fontes de base vegetal (MATA, 2012).

A MOS é um atributo chave para avaliação da qualidade do solo devido à sua capacidade de influenciar suas propriedades químicas, físicas e biológicas e, com isso, ser sensível às modificações de uso e manejo. Adicionada ao solo, a matéria orgânica proporciona a entrada de carbono (C) e nutrientes como fósforo (P), potássio (K), enxofre (S), cálcio (Ca), magnésio (Mg) e, principalmente, nitrogênio (N), que não provém da fração inorgânica do solo. A quantificação da MOS e dos teores de nutrientes no solo é feita, convencionalmente, através de análises químicas laboratoriais de forma lenta e onerosa, gerando resíduos tóxicos provenientes de reagentes químicos (SOUSA JUNIOR et al, 2011).

Nas últimas décadas, o uso de técnicas de sensoriamento próximo vem ganhando destaque na análise do solo, especialmente as técnicas espectroscópicas aliadas às análises quimiométricas para a quantificação da MOS e de outros atributos do solo. Para o estudo envolvendo MOS, grande parte dos trabalhos científicos envolvem as faixas espectrais do visível (vis) e do infravermelho próximo (NIR) do espectro eletromagnético, localizadas entre 350-750 nm e 750-2500 nm, respectivamente (COZZOLINO & MORÓN; VISCARRA ROSSEL et. al., 2006). Nessas faixas do espectro são observadas relações da interação da energia eletromagnética incidida e refletida na amostra de solo com sua mineralogia, umidade, granulometria e material orgânico. Essa interação é detectada por sensores e apresentada em diferentes comprimentos de onda, que podem ser processados, avaliados qualitativamente, comparados e submetidos a análises quimiométricas, ou seja, analisados quantitativamente. Todo esse processo ocorre de forma rápida, ambientalmente segura, econômica e não destrutiva

à amostra (ARAÚJO; DEMATTÊ; BELLINASO, 2013). Além disso, está sendo desenvolvido equipamentos vis-NIR portáteis com a possibilidade de análise no campo, sem a necessidade de coleta e transporte de amostras de terra para o laboratório, tornando o processo menos oneroso. Dessa forma, há a possibilidade e o interesse em utilizar a espectroscopia vis-NIR como ferramenta para a determinação da MOS, especialmente em um contexto agroecológico coerente, ou seja, sem a necessidade de uso e descarte de reagentes químicos sintéticos.

Diante do exposto, a hipótese deste trabalho é que a espectroscopia de reflectância vis-NIR pode ser utilizada para a predição razoavelmente boa ($R^2 > 0,6$) dos teores de C, N e P do solo. Sendo assim, o objetivo geral deste trabalho foi avaliar a acurácia da espectroscopia vis-NIR, aliada ou não a diferentes técnicas de pré-processamentos espectrais, na predição dos teores de C, N e P de solo identificado como Planossolo Háplico, sob sistema orgânico de produção de hortaliças em Seropédica, Rio de Janeiro.

2. REVISÃO DE LITERATURA

2.1. Matéria Orgânica do Solo

O solo pode ser definido como corpo natural dinâmico que ocupa a maior parte do manto da Terra e que é constituído por materiais minerais e orgânicos (EMBRAPA, 2009). A porção orgânica é chamada de matéria orgânica do solo e, por sua vez, pode ser definida como material resultante da deposição de resíduos de plantas e animais, vivos ou mortos, em vários estágios de desenvolvimento (THENG, 1989).

A partir da entrada desses resíduos orgânicos de diversas naturezas há uma série de transformações biológicas, químicas e físicas contínuas formando, gradativamente, um sistema complexo de substâncias com características e funções variadas (CAMARGO et al. 1999). Essas substâncias são chamadas de húmicas e não húmicas, de acordo com seu grau de complexidade. As substâncias não húmicas são as constituídas por estruturas biologicamente ativas e quimicamente bem definidas, como as alifáticas e aromáticas de baixo peso molecular. Essas são utilizadas como substratos para os microrganismos e por isso têm rotatividade relativamente rápida no solo (ANDERSON, 1979). Por outro lado, substâncias húmicas são aquelas constituídas por macromoléculas, mais estabilizadas e persistentes no sistema, quando comparada às anteriores. Como as substâncias húmicas são mais resistentes à decomposição, acabam ficando ali acumuladas, através de diversos mecanismos, formando então a principal reserva orgânica dos solos.

A decomposição da MOS depende principalmente das condições climáticas e edáficas do ambiente, da comunidade de decompositores e da qualidade do solo a ser decomposto. (ANDERSON & SWIFT, 1983). Em solos tropicais, como é o caso dos solos brasileiros, o processo de decomposição da MOS é muito intenso, de forma que a umidade e a temperatura são os fatores que mais influenciam as taxas de mineralização no solo, ou seja, o processo onde as substâncias orgânicas são convertidas em substâncias inorgânicas (STANFORD et al. 1973)

A matéria orgânica humificada influencia as várias funções e processos biológicos, físicos e químicos que ocorrem no solo como a ciclagem de nutrientes, complexação de elementos tóxicos, estruturação das partículas, infiltração e retenção de água, susceptibilidade à erosão, além de ser fonte primária de nutrientes às plantas, especialmente de C, N, P, S e micronutrientes (GREGORICHTH et al,1994; REEVES, 1997; TAM et al, 1998; SHEPHERD et al, 2002; CONFORTI et al, 2013). Nesse sentido, a MOS é um atributo decisivo para a definição e monitoramento da qualidade do solo, além de ocupar uma posição central na manutenção dos ambientes agrícolas (FRIGHETTO, 2000).

Em solos sob vegetação natural, a MOS encontra-se em equilíbrio dinâmico, com teores praticamente constantes ao longo do tempo (D'ANDRÉA et al., 2004). Porém a transformação das áreas de vegetação natural em áreas de cultivo agrícola implica em mudanças na estrutura e no funcionamento natural do solo, alteração na dinâmica da matéria orgânica nesses ambientes, influenciando em novas taxas de adição e perdas de MOS no sistema (EBELING et al., 2008; NUNES et al., 2011). Dessa forma, um novo equilíbrio é atingido conforme as características do sistema de manejo adotado e das condições edafoambientais da área (FREITAS et al., 2000).

Alguns sistemas de cultivo agrícola, chamados sustentáveis, consideram a conservação e ampliação da biodiversidade além da produção através de princípios ecológicos, socioeconômicos e agronômicos (ALTIERI, 1989; ASSIS, 2006). Dessa forma, diversas técnicas alternativas são utilizadas como a adubação orgânica, através do uso da matéria orgânica decomposta como fonte de nutrientes; cobertura vegetal do solo; rotação de culturas vegetais e etc. Nesse contexto, há nesses sistemas sustentáveis, entre outros fatores, a intenção de aumentar a concentrações e a qualidade da MOS.

2.2. Sensoriamento Próximo do Solo

Em uma produção agrícola é importante monitorar os fatores que influenciam o ambiente de cultivo para que sejam tomadas as estratégias de manejo, independentemente de qual seja o sistema adotado. Sabida a importância da MOS, sua quantificação através das concentrações de C, N e P pelos métodos tradicionais de laboratório demanda tempo; uso de reagentes químicos tóxicos, como por exemplo ácido sulfúrico, ácido clorídrico, molibdato de amônio, que nem sempre são descartados de forma ambientalmente segura (EMBRAPA, 2009). Além disso, essas análises geralmente são caras, demoradas e restringem o número de amostras representativas de uma região (VISCARRA ROSSEL, 2011).

A partir da década de 1970, começaram a ser desenvolvidos métodos de análises utilizando sensores, que têm a capacidade de medir parâmetros do solo de forma rápida, com uma boa relação custo benefício dando origem a dados de qualidade (MCBRATNEY et al., 2006; SIQUEIRA et al., 2010). Além disso, para a análise com sensores não é necessário o uso de reagentes químicos, o que é seguro para o meio ambiente e proporciona segurança condições de trabalho para os analisadores (ROSSEL e BEHREN, 2010).

Os sensores podem ser divididos conforme à proximidade ao alvo, que nesse caso passa a ser o solo. Enquanto os sensores remotos obtêm informações sem que haja contato físico, os sensores próximos obtêm sinais quando o detector está em contato direto ou a uma distância de no máximo a 2 metros do solo (VISCARRA ROSSEL e MCBRATNEY, 1998; VISCARRA ROSSEL et al., 2010). As vantagens do sensoriamento próximo é que suas análises podem ser realizadas em superfície e subsuperfície e com o tempo, esses sensores estão se tornando menores, mais rápidos, mais precisos, mais eficientes em energia e mais inteligentes (VISCARRA ROSSEL et al., 2011).

As técnicas mais pesquisadas para sensoriamento proximal de solo incluem métodos óptico (reflectância espectral), elétrico (condutividade elétrica) e eletroquímico (elétrons ions seletivos) (GEBBERS & ADAMCHUK, 2010; ADAMCHUCK et al., 2015). Os eletroquímicos mais comuns utilizam membranas íon-seletivas para detectar a atividade de íons como hidrogênio potássio e nitrato. A condutividade elétrica aparente do solo pode ser utilizada como indicador de características como salinidade, textura, umidade, densidade, lixiviação (CORWIN & LESCH, 2005). A reflectância espectral, em diversos comprimentos de onda, pode ser utilizada para determinação de cor, classe de solos, composição mineralógica, teor de umidade, granulometria e, com o auxílio de técnicas quimiométricas, a predição do teor de MOS (EPIPHANIO et al., 1992; VISCARRA ROSSEL et al., 2008).

2.3. Espectroscopia de Reflectância Difusa

A espectroscopia de reflectância difusa (ERD) é uma técnica que, a partir de dados espectroscópicos da energia, registra o reflexo da luz ou de ondas ou partículas de uma superfície opaca através de um sensor. O princípio da ERD é fundamentado na Lei de Beer-Lambert, que conceitua que quando uma energia eletromagnética incide sobre um alvo, parte dos vários comprimentos de onda incididos sobre ele são absorvidos por essa amostra, parte atravessa sua superfície e parte é refletida de forma difusa, ou seja, espalhada em muitos ângulos (STONER & BAUMGARDNER, 1986). Essa energia é registrada por um sensor e decomposta em diferentes comprimentos de onda, de forma contínua ao longo do espectro eletromagnético, onde ficam definidas suas feições, que são conhecidas como curva espectral ou assinatura espectral (NOVO, 1992; MORAES, 2002). Uma ilustração de todo esse processo pode ser observada na Figura 1.

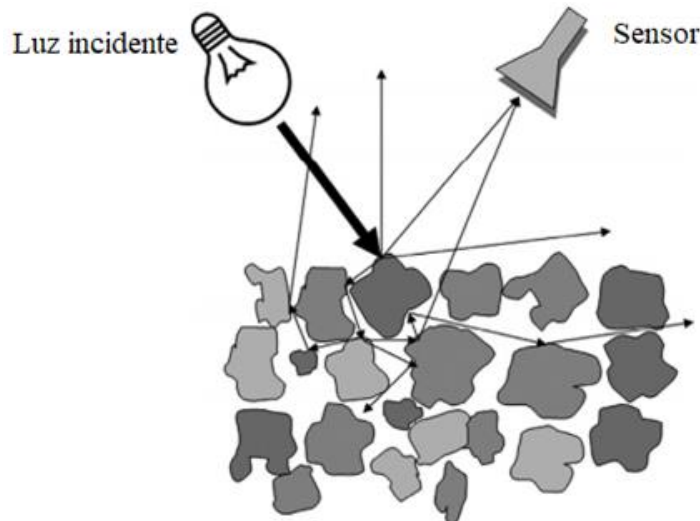


Figura 1. Ilustração da luz incidindo em uma amostra sólida e as frações de energia refletidas de forma difusa pela mesma e captadas por um sensor (Adaptado de STEMBERG, VISCARRA ROSSEL, MOURAZEN 2010).

O solo é um alvo opaco, tem capacidade intrínseca de absorver e refletir luz incidente. Dessa forma, conforme variam as moléculas presentes na amostra de solo, varia a resposta das mesmas à radiação, formando bandas que são registradas no espectro eletromagnético gerando curvas espectrais específicas (MARTIN-NETO et al., 1996; DALMOLIN et al., 2005). Dessa forma, a partir de uma leitura de ERD é possível obter informações sobre atributos e características do solo como a presença de minerais (VISCARRA-ROSSEL et al., 2006); teores de argila, areia, silte (BEN-DOR et al., 2008); carbono orgânico e matéria orgânica (MOUAZEN et al., 2007; DEMATTÊ., 2011).

As medições dos atributos do solo acontecem de forma direta ou indireta. Para medições diretas, as relações são baseadas em fenômenos físicos que afetam a refletância da luz em uma parte específica do espectro. Para medições indiretas, as relações são realizadas a partir de um domínio finito e os efeitos combinados de vários atributos podem estar relacionados a uma determinada característica do solo. Exemplos de medições diretas e indiretas são o teor de água e a MOS, respectivamente. Para a MOS, suas medições são possíveis graças às medições indiretas, a partir das funções de pedotransferência, que predizem certas propriedades do solo a partir de outras obtidas mais facilmente (BOUMA, 1989; BUDIMAN, 2003).

2.4. Espectroscopia vis-NIR

Para o estudo de MOS, as faixas do espectro eletromagnético mais comumente utilizada são as do visível (vis\visible - 350 a 750 nm) e infravermelho próximo (NIR\near infrared - 750 a 2.500 nm). Essas porções estão representadas na Figura 2.

A espectroscopia de reflectância vis-NIR tem sido apontada alternativa na análise de solos, pela rapidez, não utilização de reagentes químicos e não destrutividade à amostra (VISCARRA ROSSEL et al., 2006; DEMATTÊ et al., 2006; BEN-DOR et al., 2008; VASQUES et al., 2008). Dessa forma, são feitas medições simultâneas e repetíveis, o que proporciona significativa vantagem sobre as medições através das análises convencionais em laboratório (PASQUINI, 2018).

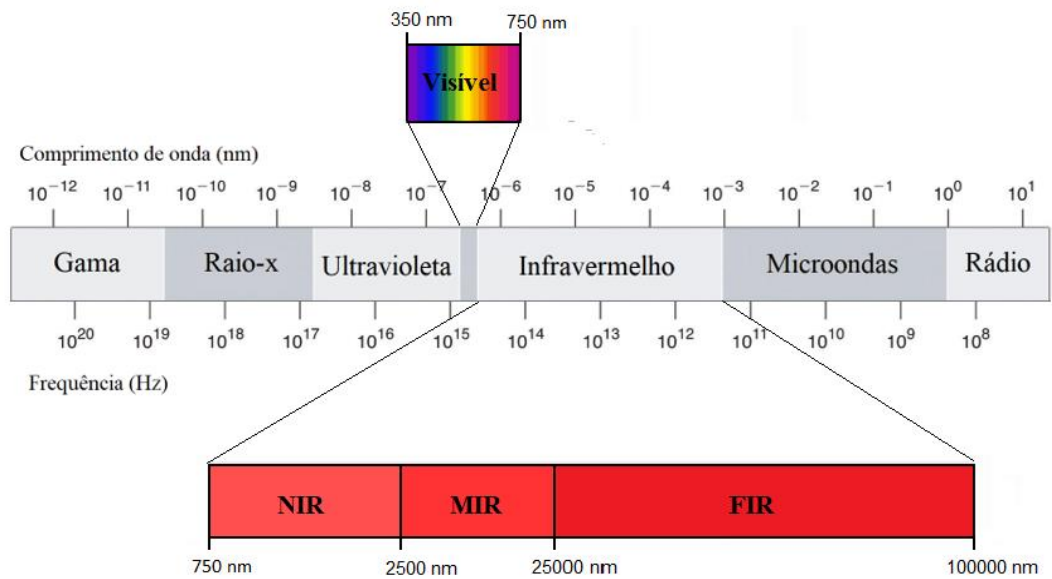


Figura 2. Espectro eletromagnético com detalhe para as porções do visível (350 a 750 nm), NIR - infravermelho próximo (750 a 2500 nm), além do MIR – infravermelho médio (2500 a 25000nm) e FIR – infravermelho distante (25000 a 100000 nm).

Essas análises são possíveis porque ao absorver a radiação vis-NIR, as moléculas que compõem a amostra de solo têm seus elétrons externos promovidos do estado fundamental à excitação. A diferença de energia entre esses dois estados ou a alteração na amplitude da vibração, alongamento ou flexão molecular é emitida e registrada pelo sensor através de curvas espectrais com feições facilmente distinguíveis de absorção e reflectância. Essas feições vis-NIR estão relacionadas aos atributos do solo como água livre e argila mineral; matéria orgânica e minerais não coloridos, como óxidos de ferro, carbonatos e sais e podem ser observadas na Figura 3 (STONER; BAUMGARDNER, 1981; BEN DOR et al., 1999; DEMATTÊ, 2002).

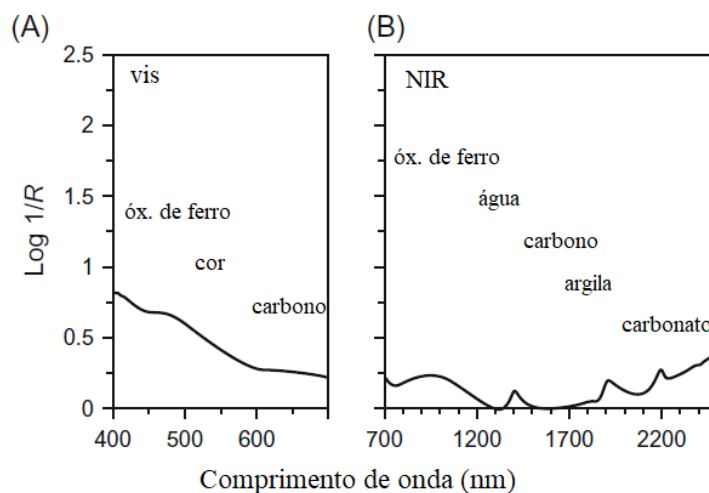


Figura 3. Curvas espectrais típicas de solos nas porções vis (A) e NIR (B) do espectro eletromagnético. (Adaptado de VISCARRA ROSSEL et al, 2011).

As medições a partir de dados espectrais vis-NIR podem ser diretas ou indiretas. Para medições diretas, as relações são baseadas em fenômenos físicos que afetam a refletância da luz em uma parte específica do espectro, como por exemplo, a mineralogia do solo ou o teor de

água. Para as medições indiretas, as relações são determinadas por uma combinação de vários atributos relacionados a uma determinada característica, como por exemplo a MOS, devido às conotações, flexões e alongamento dos grupamentos de NH, OH e CH (VISCARRA-ROSSEL et al, 2006). As absorções no vis-Nir das moléculas orgânicas, dependem dos grupamentos funcionais, chamados cromóforos, que contêm elétrons de valência de baixa excitação energia. Dessa forma, são fornecidas informações sobre a estrutura dos grupos funcionais presentes na matéria orgânica, bem como sobre a natureza de suas ligações químicas e sua reatividade (STEVENSON, 1994). Além disso, a presença de SOM afeta a cor do solo e, portanto, poderia estar diretamente relacionado à região visível do espectro (BAUMGARDNER et al, 1970). Dessa forma, as curvas espectrais do solo podem ser utilizadas tanto para o monitoramento da qualidade, fertilidade e manejo do solo de áreas ambientais e agrícolas, através das análises qualitativas (SHEPHERD & WALSH, 2007).

A partir do final da década de 1990, com o avanço da agricultura de precisão, muitos estudos têm sido realizados usando a ERD Vis-NIR também para análises quantitativas de atributos do solo (DEMATTE e GARCIA, 1999; VISCARRA ROSSEL et al., 2006). Os atributos quantificados variaram entre mineralogia, química e física do solo (VISCARRA ROSSEL et al., 2006; GENU & DEMATTÊ, 2011), levantamentos de solos (DEMATTE et al., 2004; VISCARRA ROSSEL e WEBSTER, 2011), mapeamento digital de solos (VISCARRA ROSSEL & MCBRATNEY, 2008), avaliação da qualidade do solo (BEN-DOR et al., 2009; MOUAZEN et al., 2005), agricultura de precisão (VISCARRA ROSSEL & CHEN, 2011), e especialmente para determinação de C e MOS (VISCARRA ROSSEL et al, 2006; VASQUES et al, 2009).

2.4.1. Vis-NIR e MOS em solos brasileiros

No Brasil, do ponto de vista geográfico, predominam os solos sob domínio climático tropical, que ocorrem na faixa que se estende desde o extremo norte do Estado de Roraima (aproximadamente 5 graus N) até a cidade de São Paulo (aproximadamente 23,5 graus ao sul do equador). Mais ao sul desta latitude de São Paulo ocorrem solos sob domínio climático subtropical, que possuem características diferentes quando comparados aos outros solos distribuídos pelo Brasil (compreendendo parte do estado de São Paulo, Paraná, Santa Catarina e Rio Grande do Sul).

Apesar da crescente disponibilidade dos equipamentos utilizados para a realização das análises espectrais Vis-NIR, associado a maior demanda, oportunidades de pesquisa e aplicação da técnica, poucos trabalhos dessa técnica em solos brasileiros têm sido publicados, quando em comparação com os estudos feitos e publicados em solos do resto do mundo. Uma pesquisa no banco de dados de periódicos *Scopus* mostra essa relação entre as publicações a respeito desse tema e pode ser observado na Figura 4.

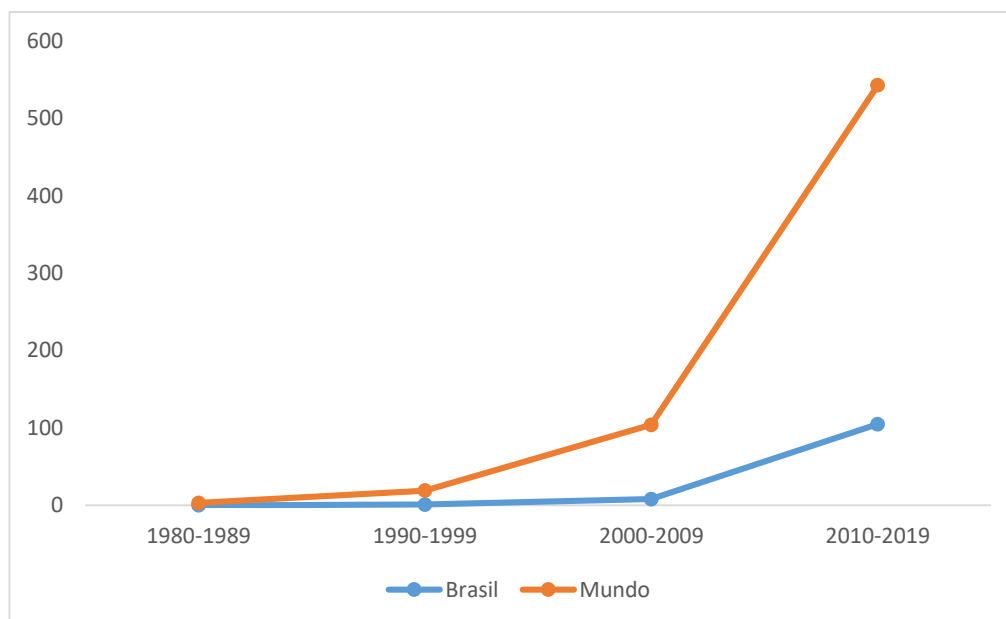


Figura 4. Número de publicações científicas abordando o uso de técnicas espectroscópicas Vis-NIR em solos mundiais e em solos brasileiros (Dados retirados do banco de dados Scopus em dezembro de 2019).

Os estudos referentes aos últimos 5 anos que tratam a respeito da predição de C, N, P a partir de dados espectroscópicos de solos brasileiros foram compilados, conforme mostra a Tabela 1.

Terra tem se dedicado à pesquisa de dados de uma Biblioteca espectral de solos que compreende amostras de terra de regiões de produção de cana-de açúcar nos estados de Mato Grosso do Sul, Minas Gerais, São Paulo e Goiás. Seus últimos trabalhos focaram em promover uma melhor acurácia nas predições de C orgânico, seja com o teste e escolha dos melhores métodos de análise multivariada e pré-processamentos espectrais (TERRA, 2015), como com trabalhos inovadores abordando a fusão de bandas pertencentes ao vis-Nir e ao Mir para criação de melhores preditores de C orgânico em comparação com análises dessas respostas espectrais feitas separadamente (TERRA, 2019).

Na região Norte do Brasil, Almeida e colaboradores indicam que em áreas da bacia do Acre somente sejam usadas quantificações de N por espectroscopia NIR em caso de grandes volumes de processamento de dados, pois o método nesse caso não substituiu a exatidão obtida através das análises elementar. Araújo pesquisou solos Amazônicos, mas de Terra preta do índio no Pará, ou seja, com grandes quantidades antrópicas de MOS. Foi comparada a precisão e eficiência dos métodos de vis-NIR e MIR. Apesar de obterem um bom ajuste em ambos os sensores, os modelos de previsão baseados nos dados do MIR superaram os com base em dados vis – NIR nesses tipos de solos. Além disso, ambos subestimaram o conteúdo de C desses solos. Por outro lado, Pinheiro e colaboradores obtiveram em solos da Amazônia alta aplicabilidade nas técnicas espectroscópicas vis-NIR em relação aos conteúdos de C e N e enquanto encontraram limitações aos modelos de P daqueles solos.

Em clima subtropical, Moura-Bueno e Dotto exploram técnicas de pré-processamentos espectrais e análises multivariadas na predição de C, com o espectro vis-NIR completo ou reduzido, nos estados do Rio Grande de Sul e Paraná. Felix (2016) observou a influência da granulometria das amostras de terra do Paraná, com melhores resultados em partículas menores que de 0,2mm. Também no Paraná, Carra (2019), obteve bons ajustes dos modelos preditivos C e N para solos sob diferentes condições climáticas e usos de terra.

Tabela 1. Principais trabalhos no país abordando espectroscopia vis-NIR para predição de C, N e P do solo nos últimos 5 anos.

Referências	Local	N*	Atributo	PP	R ²	RMSE	RPD
Carra, 2019	PR	431	C	PCA	0,93	4,78	0,71
Terra, 2019	SP, MG, MS, GO	1259	C	SMV	0,69	3,38	-
Moura-Bueno, 2019	RS	841	C	PLSR	0,71	0,58	-
				MLR	0,71	0,56	-
				SNV	0,63	0,64	-
				RF	0,62	0,63	-
Dotto, 2018	SC	595	C	PLSR	0,74	0,58	-
				PCR	0,73	0,59	-
				MLR	0,74	0,59	-
				SVM	0,78	0,53	-
				RF	0,76	0,58	-
				BMA	0,74	0,59	-
				WAPLS	0,65	0,62	-
				GPR	0,73	0,61	-
ANN	0,69	0,64	-				
Pinheiro, 2017	AM	434	C	PLSR	0,71	5,69	1,84
Dotto, 2017	SC	299	C	PLSR	0,88	0,34	-
				SVM	0,86	0,36	-
Félix, 2016	PR	214	C	PLSR	0,57	-	1,55
Terra, 2015	SP, MG, MS, GO	1254	C	SVM	0,65	0,16	-
Araújo, 2015	PA		C	PLSR	0,90 0,05	1,50 6,40	- -
Carra, 2019	PR		N	PCA	0,93	4,78	0,71
Félix, 2016	PR	214	N	PLSR	0,68	-	1,92
Almeida et al. 2016	AC	168	N	PLSR	0,84		
Carra, 2019	PR	431	P	PCA	0,685	0,83	6,2
Pinheiro, 2017	AM	434	P	PLSR	0,11	1,98	1,05
Félix, 2016	PR	214	P	PLSR	0,37	-	1,26
Terra, 2015	SP, MG, MS, GO	1254	P	SVM	0,24	0,37	

* N - Número de amostras.

2.5. Quimiometria

A quimiometria pode ser definida como uma área da química que, através de ferramentas matemáticas, físicas e computacionais, tem o objetivo de planejar ou selecionar condições ótimas de medidas e experimentos, extraindo o máximo de informação química relevante (FERREIRA et al, 1999). A quimiometria teve origem a partir da popularidade do uso da análise multivariada, para quantificar uma variável de interesse, com o tratamento e predição de dados químicos através de modelos estatísticos.

Os modelos estatísticos determinam a relação entre propriedades medidas e concentrações químicas a partir de dois conjuntos de variáveis. Dessa forma, é desenvolvido então um modelo de treinamento para um conjunto de amostras cujas propriedades são conhecidas. Uma vez estabelecido, esse modelo de treinamento, ele é utilizado para classificar novas amostras, com concentrações desconhecidas da mesma propriedade de um novo

conjunto, chamado de validação, para então determinar quanto das propriedades medidas estão relacionadas de fato às concentrações químicas (SANCHEZ & KOWALSKI, 1988; BRERETON, 2003). Para estudos com espectrometria, por exemplo, os modelos determinam o quanto as curvas espectrais das amostras relacionam-se com informações químicas do solo, como C, N e P. Um esquema representando as etapas para a criação de um modelo pode ser visto na Figura 5.

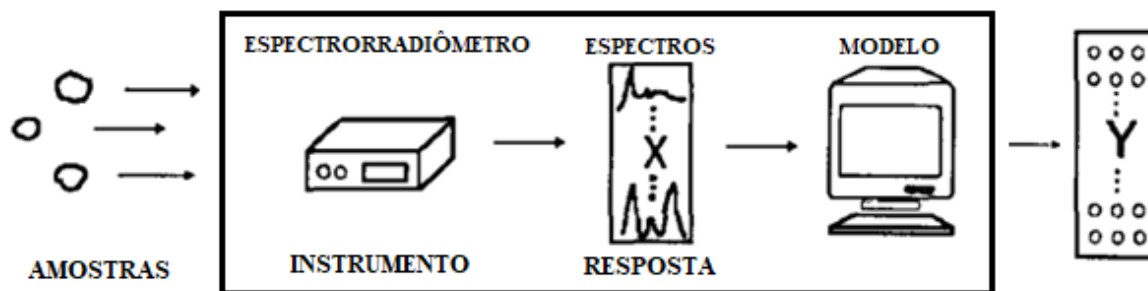


Figura 5. Diagrama representando o processo de treinamento de um modelo quimiométrico. (Adaptado de FERREIRA, 1999).

Para a criação dos modelos estatísticos, são necessárias inúmeras análises multivariadas, ou seja, a regressão por mínimos quadrados parciais (PLSR, do inglês Partial Least Squares Regression) é uma das técnicas mais comuns para a calibração e predição de modelos espectrais de solos (MCCARTY et al, 2002).

2.5.1. Pré-processamento de curvas espectrais

A aquisição de dados espectrais não fornece apenas informações relevantes sobre a presença, ausência e concentração de compostos químicos. Dependendo da técnica de aquisição de espectros, equipamentos, condições experimentais, acessórios utilizados, uma grande quantidade de informação relativa aos fenômenos físicos, erros aleatórios e sistemáticos também estarão presentes no conjunto de dados adquiridos. Essas interferências podem causar não-linearidades entre os espectros e as concentrações dos componentes de interesse resultando em ruído aleatório, desvio de linha de base e efeito de espalhamento nos espectros, o que pode afetar a robustez dos modelos de calibração posteriores. Dessa forma, para que essas informações não encubram a informação que está verdadeiramente relacionada com a propriedade que se deseja estudar, uma série de ferramentas matemáticas, chamadas de pré-processamentos espectrais, são utilizadas (SIMÕES, 2008). As ferramentas matemáticas das técnicas de pré-processamento espectral operam nas amostras e são aplicadas às colunas da matriz de dados, para cada variável. São exemplos desse conjunto de técnicas a normalização, correção de linha de base, suavização, derivadas, padronização pelo desvio padrão, correção multiplicativa de sinal (KOWALSKI; BEEBE, 1987; FERREIRA, 2015).

Um artifício comumente utilizado para destacar e facilitar a identificação das feições de absorção nos espectros é a transformação dos valores de reflectância para absorbância, e com isso melhorar a linearidade dos espectros. Essa propriedade mede a capacidade dos materiais em absorver radiações eletromagnéticas em frequências específicas (HOLLAS, 2004). Para aumentar os picos quimicamente mais relevantes nos espectros e reduzir os efeitos, como os desvios da linha de base e a curvatura geral, as técnicas mais comumente utilizadas incluem correção de dispersão multiplicativa (MSC) (GELADI et al., 1985) e a variação normal padrão (SNV) (BARNES et al., 1989). Esses efeitos são provocados pela não-homogeneidade das

amostras, decorrentes principalmente de diferença de granulometria, causados pelo espalhamento de luz.

Para reduzir o ruído aleatório nos sinais espectrais, ou seja, os sinais contínuos e sem relação com atributos mensuráveis, o método de suavização é comumente utilizado. Ele atua criando espectros médios para diminuir a relação sinal / ruído através do filtro de média móvel, que calcula o valor espectral de pontos considerando um intervalo definido a partir de um ponto médio (TSAI; PHILPOT, 1998).

Métodos de derivadas realizam correção de linha de base e aumentam os sinais fracos, além de remover tanto os aditivos e efeitos multiplicativos nos espectros através derivação numérica de um vetor que inclui um passo de suavização. O primeiro derivado remove a linha de base e elimina efeitos de fundo; a segunda derivada remove a linha de base e a linear (SAVITZKY & GOLAY, 1964). A técnica de remoção contínua é usada para isolar características de absorção devidos a processos diferentes daqueles de interesse em espectros de refletância difusa (CLARK & ROUSH, 1984).

Embora alguns trabalhos na literatura já propuseram ferramentas de otimização para pré-processamentos, a escolha do pré-processamento mais adequado para um determinado conjunto de dados é um processo de tentativa e erro. (DEVOS; DUPONCHEL, 2011; JARVIS; GOODACRE, 2005; FERREIRA, 2015).

2.5.2. Regressão por Mínimos quadrados parciais (PLSR)

Proposta inicialmente por H. Wold (2001), PLSR é uma técnica de análise de dados multivariados utilizada para relacionar uma ou mais variáveis resposta (Y) com diversas variáveis independentes (X), com base no uso de fatores. PLSR é especialmente empregado quando existem muitas variáveis preditoras altamente colineares (COZZOLINO & MORON, 2003; VISCARRA ROSSEL et al., 2006; CONFORTI et al. 2013).

O PLSR implica em encontrar um conjunto de vetores base (componentes principais) para os dados espectrais e um conjunto separado de vetores base para os dados de concentração e, em seguida, relacioná-los um com o outro. O primeiro componente principal corresponde aquele que descreve a máxima quantidade de variância das amostras. Quando toda a variância de um conjunto de amostras não puder ser explicada por apenas um componente principal, um segundo componente principal perpendicular ou ortogonal ao primeiro será utilizado, e assim por diante de forma que o modelo irá selecionar o menor número possível de variáveis independentes com o máximo de informações contidas em cada variável, e altamente correlacionadas umas com as outras (ADAMS, 1995). Dessa forma, após a modelagem, teoricamente, a matriz dos quadrados dos resíduos deverá conter apenas a variância não explicada associada ao ruído.

Essa é uma das técnicas estatísticas multivariadas mais comuns para calibração e predição de atributos dos solos devido a facilidade na interpretação dos seus resultados (REEVES, MCCARTY et al., 2002; COZZOLINO; MORÓN, 2006; VISCARRA ROSSEL; MCBRATNEY, 2006).

3. MATERIAL E MÉTODOS

3.1. Caracterização da Área de Estudo

O estudo foi realizado no módulo experimental de produção orgânica intensiva de hortaliças (módulo experimental), localizado no Sistema Integrado de Produção Agroecológica (SIPA), conhecido como 'Fazendinha Agroecológica - Km 47', no município de Seropédica, estado do Rio de Janeiro Brasil, entre as coordenadas 22°46'S, 43°41'W (Figura 6). O SIPA tem por objetivo desenvolver atividades de pesquisa, ensino e extensão em agroecologia e agricultura orgânica, em um esforço conjunto de pesquisadores da Embrapa Agrobiologia, Embrapa Solos, PESAGRO-RJ e de professores e pesquisadores da Universidade Federal Rural do Rio de Janeiro (UFRRJ) e do Colégio Técnico da UFRRJ (CTUR) (ABBOUD et al., 2005).

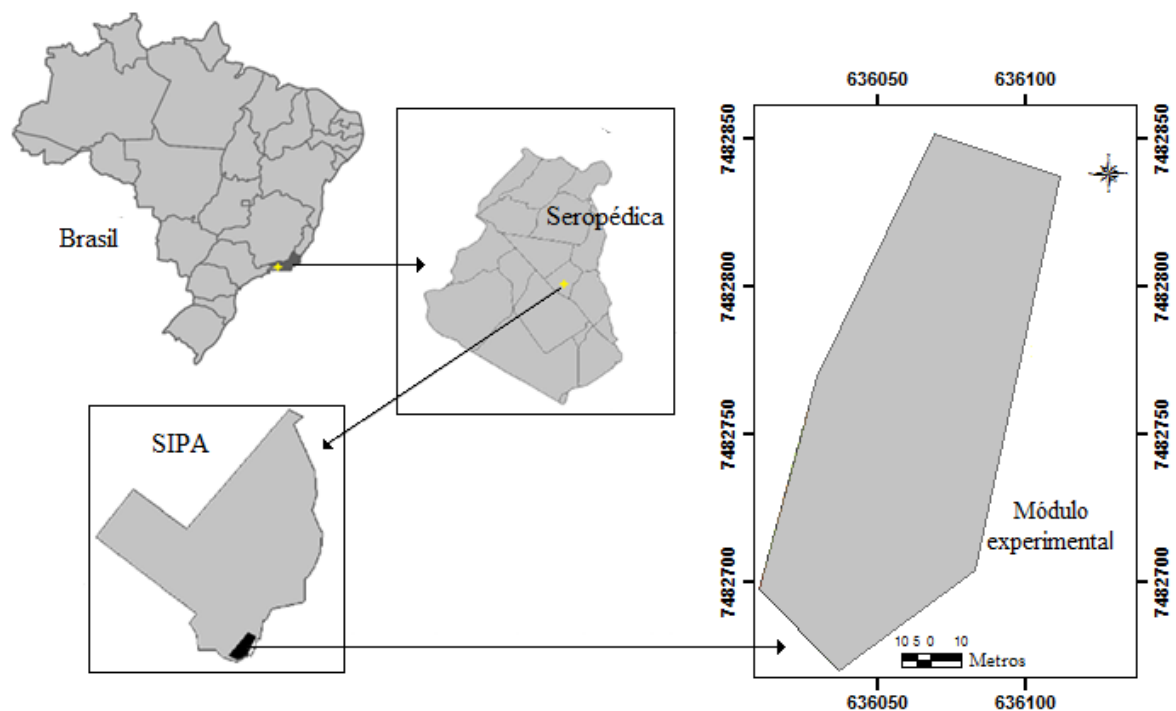


Figura 6. Localização do módulo experimental de produção orgânica intensiva de hortaliças.

O módulo experimental, por sua vez, foi criado em 2010 com a proposta de ser um sistema de produção orgânica intensiva de hortaliças e tem 1,06 hectare de área total. Ele está localizado em relevo plano com altitude média de 33 m e clima predominante do tipo Aw, caracterizado como quente e úmido (KÖPPEN, 1948), com valores médios de temperatura de 24,5° C e precipitação pluviométrica anual de 1.213 mm (DIAS, 2007).

A organização espacial e o uso da terra das subáreas do módulo pode ser vista na Figura 7 e é composta por área de hortaliças a pleno sol (canteiros: 3.578m²); cultivo de hortaliças protegidas por telados, com redução de 30% de radiação incidente (telados: 527 m²); cultura de capim (*Pennisetum purpureum*) (capineira: 3982 m²); leguminosas arbustivas (*Gliricidia sepium*) (gliricídia: 694m²) além de uma área destinada à compostagem orgânica (C.O.), bordas das culturas e ruas de tráfego (1842 m²) (MATA, 2012).

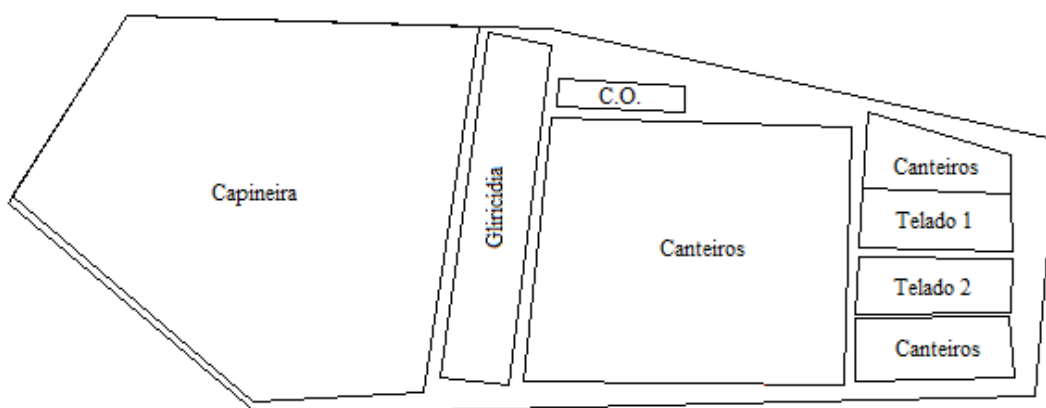


Figura 7. Organização espacial das subáreas do módulo experimental compostas por capineira, gliricídia, canteiros a céu aberto e telados com hortícolas e área de compostagem orgânica (C.O.) (Adaptado de MATA, 2012).

As áreas do módulo experimental representam, em sua prática de manejo, duas regiões principais, com objetivos distintos. São elas: região de produção de hortaliças, com interesse comercial (Canteiros e Telados) e região de produção de biomassa vegetal (Capineira e Gliricidia), que têm suas palhadas utilizadas como cobertura morta nas culturas de hortaliças. Dessa forma, há importação e exportação de matéria orgânica de uma área para a outra dentro do próprio módulo. Segundo Ronquim (2010), essa prática proporciona, a longo prazo, uma melhora nas condições físicas, biológicas e químicas de solo tropicais, de forma econômica e ecológica e é uma das técnicas agroecológicas adotadas no SIPA.

3.2. Amostragem e Caracterização do Solo

O solo do módulo experimental foi classificado como Planossolo Háplico Distrófico arênico (DIAS, 2007) e tem sido amostrado e monitorado, anualmente, desde 2011 para que seja avaliada a qualidade do solo, mediante implantação de técnicas agroecológicas de manejo. Para garantir que a amostragem fosse feita sempre nos mesmos locais ao longo do tempo, a área foi identificada em um grid de, aproximadamente, 5 x 5 metros através de GPS com correção diferencial (GPS TRIMBLE PRO XP), exceto na área de capineira, que foi amostrada em um grid de 10 x 10 metros para abranger as variações e características do solo. As coletas de amostras de terra foram realizadas com trado tipo holandês, na profundidade de 0 a 20 cm, em cada um dos pontos georreferenciados, com 133 amostras destinadas à análise de C (grupo 1) e 294 amostras destinadas às análises de N e P (grupo 2, correspondente ao número total de amostras do módulo).

As coletas foram realizadas nos meses de agosto, nos anos de 2013 e 2014. Ao longo do tempo, algumas amostras foram perdidas ou tiveram sua forma original danificada no processo de armazenamento e, por esses motivos, foram inutilizadas na elaboração deste trabalho. Na Figura 8 é possível observar os pontos de amostragem para do módulo experimental para os grupos 1 e 2 (A) e as amostras que não se perderam e então foram de fato utilizadas para o trabalho, sendo o total de 244 e 271 para os anos de 2013 (B) e 2014 (B), respectivamente.

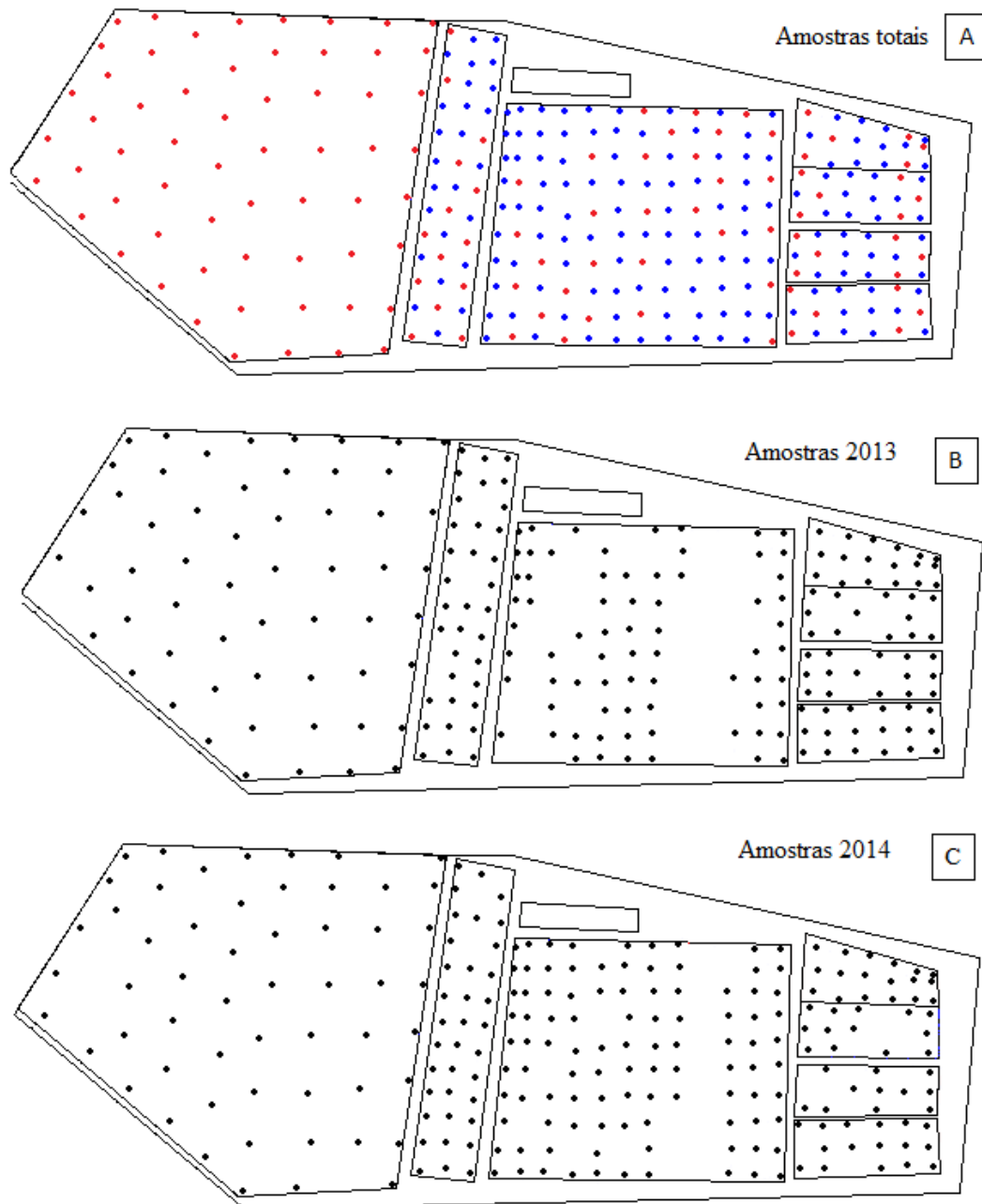


Figura 8. Croqui dos pontos de coleta do módulo experimental georreferenciados. Todos os pontos (vermelhos e azuis) representam as amostras do grupo 2, coletadas em 2013 e os pontos em vermelho representam as amostras do grupo 1, coletadas em 2014.

3.3. Análises Químicas do Solo

As amostras de terra foram secas ao ar, destorroadas, passadas por peneira de malha de 2 mm para a obtenção da terra fina seca ao ar (TFSA). Uma parte das amostras foi direcionada às análises químicas, que ocorreram no Laboratório de Química Agrícola na Embrapa Agrobiologia e no Laboratório de Matéria Orgânica do Solo, do Departamento de Solos da UFRRJ. Foi realizada a análise dos teores de carbono orgânico total pelo método Walkley & Black modificado (EMBRAPA, 2009). O teor de nitrogênio total pelo método Kjeldahl

(SILVA, 2009) e o teor de fósforo disponível trocável através do método de Mehlich 1, com a determinação do fósforo por colorimetria (MEHLICH, 1953).

Para comparação com as reflectâncias e absorbâncias espectrais, a MOS foi estimada a partir do cálculo de van Bemmelen, que assume teor no solo de 58%, pela seguinte equação:

$$\text{MOS} = C \times 1,724 \quad (1)$$

3.4. Estatística Descritiva dos Dados

A estatística descritiva (mínimos, máximos, 1º e 3º quartis, média, mediana, desvio padrão, assimetria e curtose) dos teores de C, N e P foi calculada e utilizada para resumir e avaliar a magnitude da dispersão do conjunto de dados. O conjunto de dados foi submetido à análise visual de histogramas para avaliar a normalidade dos dados.

3.5. Análises Granulométricas do Solo

As frações granulométricas do solo interferem nos teores de MOS e também nas feições de absorção e intensidade de reflectância geradas nas suas curvas espectrais. Dessa forma, uma parte das amostras de terra foi encaminhada ao Laboratório de Física do Solo da UFRRJ para determinação da granulometria através da realização do método da Pipeta (EMBRAPA, 1997). As quantidades de areia, argila e silte foram obtidas e utilizadas para auxiliar na interpretação da matéria orgânica e dos dados espectrais do solo.

3.6. Análises Espectrais do Solo

Uma parte das amostras de TFSA foi encaminhada para o Laboratório de Contaminantes e Resíduos da Embrapa Solos, onde as amostras foram homogeneizadas e colocadas em placas de Petri (90 mm de diâmetro) formando uma camada de aproximadamente 1 cm de espessura. Em seguida, as partículas foram niveladas dentro das placas, para minimizar os efeitos da rugosidade, e então colocadas na estufa por 12h a 45-50°C, para eliminar o efeito da umidade.

As leituras espectrais das amostras foram obtidas pelo equipamento FieldSpec 4 (Malvern Panalytical B.V., Almelo, Holanda) que registra a reflectância difusa do solo na região Vis-NIR do espectro eletromagnético na faixa espectral de 350 a 2500 nm. As placas de Petri com as amostras foram colocadas sobre uma mesa giratória (ASD Turntable), girando a 22 rotações por minuto, para que as leituras espectrais fossem feitas com a amostra em movimento. Foi realizada uma leitura espectral por amostra/placa que, por sua vez, foi obtida através da média de 100 varreduras de repetição interna feita pelo instrumento, para a aquisição de um espectro representativo. As leituras das amostras de terra foram feitas em conjuntos de dez e, antes do escaneamento de cada conjunto, foi obtida uma varredura do padrão Spectralon® branco (LabSphere, North Sutton, NH, EUA) com 100% de reflectância. As etapas do processamento das amostras e o conjunto de instrumentos utilizados para as leituras espectrais podem ser visualizados na Figura 9.



Figura 9. Da esquerda para a direita: amostras de TFSA colocadas e niveladas em placas de Petri; estufa; conjunto de equipamentos necessários às leituras espectrais, sendo A- espectrorradiômetro Field Spec 4; B- disco-mesa giratória suporte para as placas de Petri; C- computador acoplado ao espectrorradiômetro e D - placa de spectralon.

3.7. Pré-Processamento dos Dados Espectrais do Solo

Os dados de reflectância (REF) das amostras coletadas em 2014 e 2013 foram submetidos a seis pré-processamentos, de forma a avaliar qual a estratégia mais adequada para o pré-processamento dos dados, sem remover informações espectrais importantes e evidenciando as que têm ligação com as concentrações de C, N e P. Os pré-processamentos utilizados foram: suavização por média móvel simples com janela de 9 bandas (SMO), derivação Savitzky-Golay com polinômio de 1ª ordem e janela de 9 bandas (SGD), variação normal padrão (SNV), correção multiplicativa de sinal (MSC) e correção do contínuo (CRR).

3.8. Modelagem Quimiométrica de C, N e P do Solo

3.8.1. Particionamento dos dados

O conjunto de dados referente à amostra de terra coletas em 2014 foi dividido ao acaso em conjunto de treinamento (~70% das amostras) e conjunto de validação (~30%; das amostras), por meio da função ‘sample’ do software estatístico R (R CORE TEAM, 2017), que faz uma seleção aleatória dos dados. Os dados de 2013 não sofreram particionamento, conforme pode ser observado na Tabela 2.

Tabela 1. Conjunto de dados particionado em treinamento e validação para as amostras referentes ao ano de 2014 e conjuntos globais de dados para as amostras de 2013 e 2014.

Atributos	2013		2014	
	Conjunto Global	Conjunto Global	Treinamento	Validação
Carbono	123	122	86	36
Nitrogênio	244	271	190	81
Fósforo	244	271	190	81

Dessa forma, para o grupo 1 ficaram 122 amostras no conjunto global, 86 e 36 amostras nos conjuntos de treinamento e validação, respectivamente. Para as amostras do conjunto 2 ficaram 271 amostras no conjunto global, 190 e 81 para os conjuntos de treinamento e de validação, respectivamente. As amostras do ano de 2013 ficaram com 123 amostras para o grupo 1 e 244 para o grupo 2.

3.8.2. Análise multivariada

Foi utilizada regressão por mínimos quadrados parciais (PLSR) nos dados de treinamento e validação das amostras do ano de 2014 e apenas validação para os dados de 2013, ambos a partir de dados espectrais pré-processados (ABS, SMO, SGD, SNV, MSC e CRR), além dos dados de reflectância (REF) com o objetivo de alcançar o melhor ajuste entre os teores de C, N e P e dados espectrais Vis-NIR através de suas variáveis latentes extraídas por PLSR.

3.8.3. Construção e validação dos modelos de predição PLSR

Para a construção dos modelos de predição, foi realizado um treinamento do algoritmo a partir dos dados do conjunto de treinamento das amostras de 2014. Para a seleção do número de variáveis predictoras, ou componentes principais, foi utilizada a validação cruzada, através da técnica de *leave-one-out*. Nesse processo, um conjunto menor foi formado com 10 amostras, sendo uma delas removida para que tivesse sua concentração de C, N ou P predita a partir das outras nove amostras remanescentes. Esses testes (treinamentos) foram repetidos 10 vezes, cada vez com uma nova amostra removida e uma nova predição a partir das restantes. Para cada teste foi atribuído um erro quadrático médio (RMSE) em relação ao valor predito e ao, de fato, obtido. O número de componentes principais responsável por apresentar o menor erro de predição nesse processo foi escolhido e incorporado ao modelo para então pudesse ser aplicado às amostras externas ao conjunto de treinamento (amostras dos conjuntos de validação). Uma representação de como funciona a validação cruzada por *leave-one-out* pode ser observada na Figura 10.

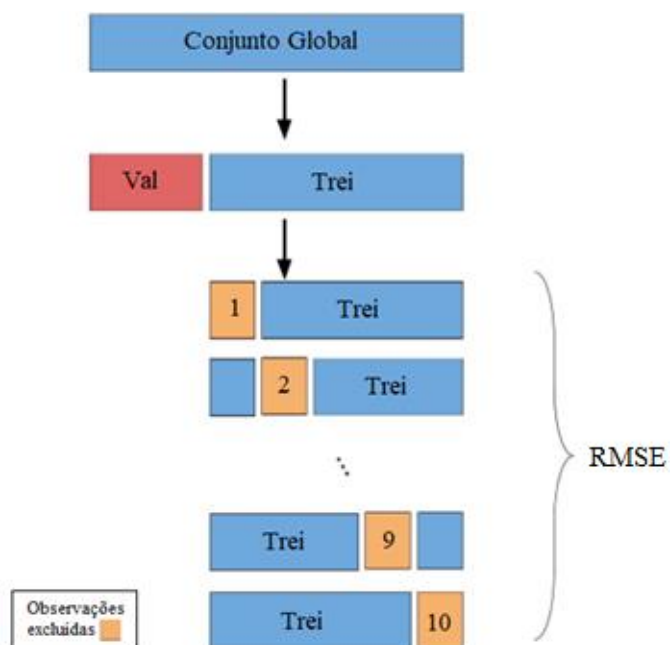


Figura 10. Esquema mostrando o funcionamento da validação cruzada por *leave-one-out*.

Após o treinamento dos modelos, os mesmos foram aplicados às amostras referentes ao conjunto de validação de 2014, para verificar o quanto de fato eles refletiram os teores de C, N e P do solo no módulo experimental, desta vez com as amostras que não estavam presentes no treinamento. Foi gerado um modelo de predição para cada atributo do solo (C, N, P), separadamente, por pré-processamento (ABS, SMO, SGD, SNV, MSC e CRR) e para os dados brutos de reflectância (REF) totalizando 42 modelos treinados e validados para os dados de 2014.

No processo de modelagem, há uma etapa de validação onde novas amostras, do mesmo solo em estudo, avaliam a capacidade preditiva dos modelos, diferentes da validação cruzada e validação externa. Nesse caso, para avaliar a aplicabilidade dos modelos preditivos de C, N e P no módulo experimental, os modelos foram validados também nas amostras referentes à coleta de 2013 simulando uma aplicação na prática rotineira do módulo como, por exemplo, para o monitoramento temporal dos teores desses nutrientes no solo.

3.8.4. Avaliação dos modelos de predição PLSR

A avaliação da performance dos modelos de predição foi realizada de acordo com seus ajustes e acurácias através do cálculo dos seguintes índices: coeficiente de determinação – R^2 – equação (2); raiz do erro quadrático médio – RMSE – equação (3); e razão do desempenho e intervalo interquartil – RPIQ – equação (4).

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

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

$$RPIQ = \frac{(Q3 - Q1)}{RMSE} \quad (4)$$

Onde \hat{y} é o valor predito; y é o valor observado; \bar{y} é a média dos valores observados; n é o número de amostras; Q1 e Q3 correspondem aos valores referentes ao primeiro e terceiro quartis, respectivamente.

3.8.5. Remoção de outliers

O termo outlier é utilizado para designar amostras anômalas que podem estar presentes nos conjuntos de construção e validação de modelos. No presente trabalho, foi feita a etapa de treinamento e validação dos modelos a partir de dados brutos de reflectância e a partir desses modelos foi feita a identificação de outliers através da análise visual das amostras que estavam em posições muito aleatórias às demais, em relação à correlação linear entre dados preditos *versus* dados observados. Essas amostras foram retiradas dos conjuntos as quais pertenciam (treinamento e validação) e então novos modelos para REF foram realizados, além dos modelos para ABS, SMO, SGD, SNV, MSC e CRR. Um fluxograma resumido de todas as etapas realizadas nesse trabalho é apresentado na Figura 11.

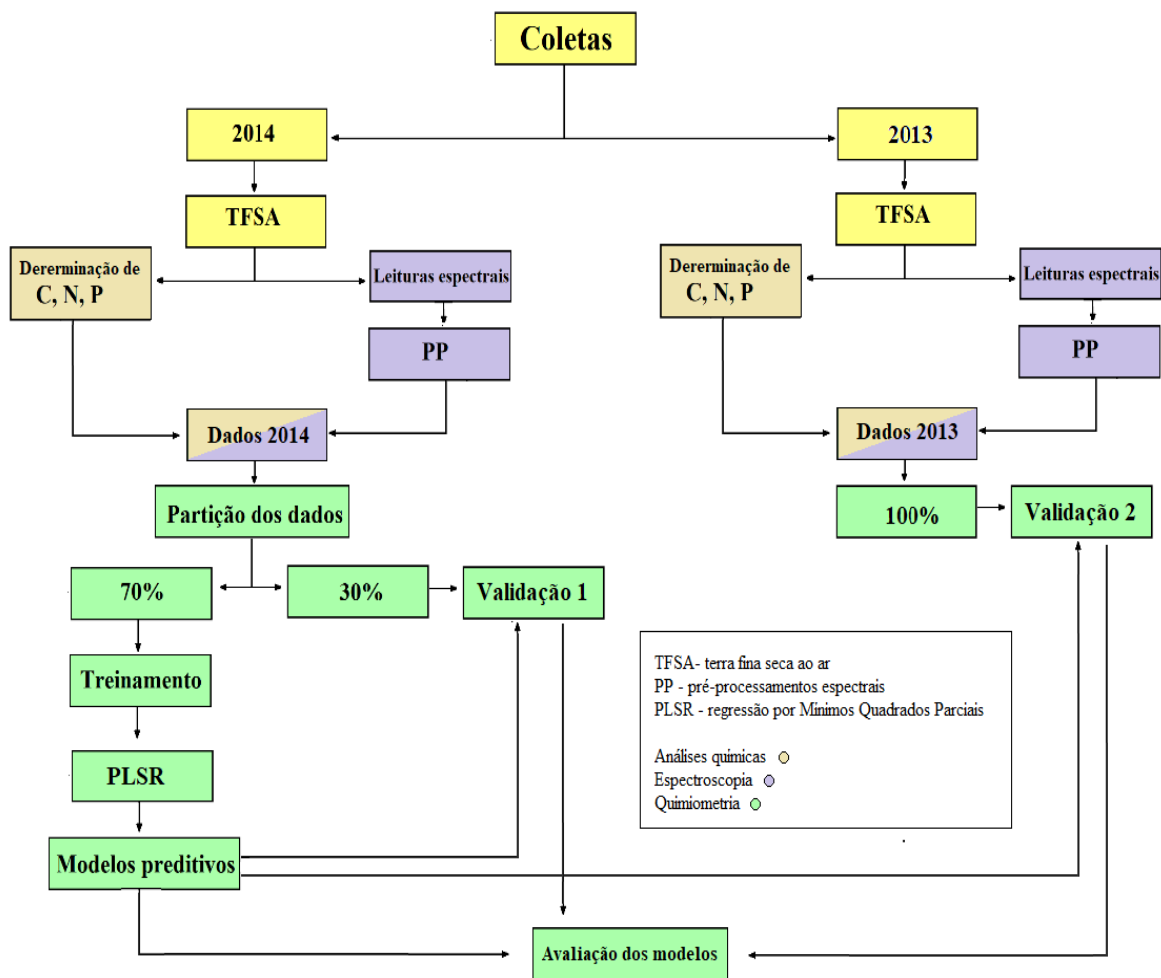


Figura 11. Fluxograma das etapas realizadas incluindo as análises químicas laboratoriais, assim como, as espectrais e quimiométricas desenvolvidas no presente trabalho.

4. RESULTADOS E DISCUSSÃO

4.1. Estatística Descritiva

Os resultados das análises químicas dos teores de C, N e P do solo, para as mostras anos de 2013 e 2014, foram descritos por Mata (2016) e estão apresentados na Tabela 3, através dos valores médios obtidos para cada área do módulo experimental.

Tabela 2. Estatística descritiva dos teores de C, N e P do solo, para os anos de 2013 e 2014.

Observações	2013			2014		
	C (.kg ⁻¹)	N (g.kg ⁻¹)	P (mg.l ⁻¹)	C (g.kg ⁻¹)	N (g.kg ⁻¹)	P (mg.l ⁻¹)
	123	244	244	122	271	271
Mínimo	0,90	0,10	2,48	0,34	0,00	2,03
Máximo	18,00	1,80	257,57	16,35	1,80	257,57
1° Quartil	4,67	0,70	27,66	4,46	0,60	37,33
3° Quartil	8,47	1,10	109,23	8,92	1,00	106,85
Média	6,89	0,87	76,59	6,77	0,84	79,62
Mediana	6,38	0,80	72,42	6,09	0,80	79,22
DP	3,16	0,30	55,38	3,01	0,30	54,59
Assimetria	0,74	0,32	0,78	0,59	0,22	0,80
Curtose	0,96	-0,21	0,32	0,04	0,25	0,56

Os valores de C variaram de 0,90 a 18,00 g.kg⁻¹ em 2013 e de 0,34 a 16,35 g.kg⁻¹ em 2014 e apresentaram um pequeno decréscimo médio de 6,89 g.kg⁻¹ para 6,69 g.kg⁻¹ dos valores médios nesse período de tempo. Os valores de N variaram de 0,10 a 1,80 g.kg⁻¹ em 2013 e de 0,00 a 1,80 g.kg⁻¹ em 2014 e também apresentaram pequeno decréscimo dos valores médios de 0,87 para 0,84 g.kg⁻¹ nesse período de tempo. Para os valores de P, os valores encontrados variaram de 2,48 a 257,57 mg.l⁻¹, apresentando média de 76,86 mg.l⁻¹ em 2013. Em 2014, a variação de foi 2,03 a 257,5 mg.l⁻¹, com média de 79,62 mg.l⁻¹, indicando uma grande variação desse elemento em comparação com os dados de P de 2013.

Os valores de assimetria e curtose indicam a distribuição dos valores em torno do ponto central e seu achatamento, respectivamente. Esses valores podem ser observados também graficamente através dos histogramas apresentados na Figura 12.

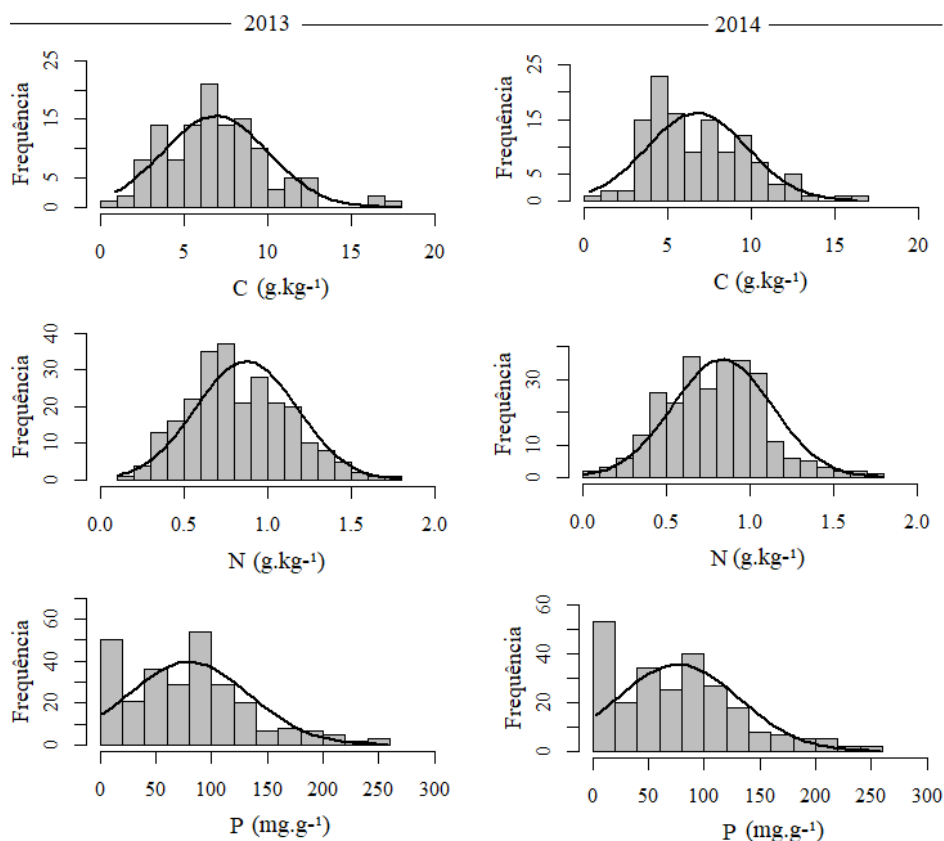


Figura 12. Histogramas para os dados de C, N e P do solo das amostras dos anos de 2013 e 2014, respectivamente da esquerda para a direita.

4.2. Dados Granulométricos

Para uma interpretação mais abrangente sobre o comportamento da fração orgânica do solo, representada neste estudo por C, N e P, foram utilizados os dados granulométricos do módulo experimental, obtidos por Mata (2012) e apresentados na Tabela 4.

Tabela 3. Frações granulométricas do solo referentes a cada área (capineira, gliricidia, canteiros e telados) do módulo experimental.

Sub-Área	Areia grossa (g.kg ⁻¹)	Areia fina (g.kg ⁻¹)	Areia total (g.kg ⁻¹)	Argila (g.kg ⁻¹)	Silte (g.kg ⁻¹)
Capineira	681,9	177,8	859,7	97,6	42,7
Gliricídia	674,5	207,1	881,6	36,5	81,8
Canteiros	610,7	223,1	833,8	86,3	79,8
Telados	564,3	210,6	774,9	127,4	97,7

A fração areia grossa predomina em todas as áreas do módulo experimental seguida da fração de areia fina. As áreas da capineira, canteiros e telados apresentaram teores maiores de argila quando em comparação aos teores de silte. Os valores encontrados foram de 9,76 g.kg⁻¹, 8,63 g.kg⁻¹ e 12,74 g.kg⁻¹ de argila e 4,27 g.kg⁻¹, 7,98 g.kg⁻¹ e 7,54 g.kg⁻¹ de silte para Capineira e Telados, respectivamente. Por outro lado, a área da Gliricídia apresentou maior concentração de silte que de argila com valores de 3,65 g.kg⁻¹ e 8,18 g.kg⁻¹ para argila e silte, respectivamente.

Quando esses dados são trabalhados de forma a representar a área do módulo como um todo (Figura 13), é possível observar que a fração de areia total é extremamente alta (84 g.kg⁻¹) e não proporciona o surgimento de mecanismos para que a matéria orgânica das culturas vegetais fique ali retida e preservada. Logo, em solos como esse, de textura arenosa, a matéria orgânica lábil, ou seja, a mais facilmente decomponível, tem maior importância na dinâmica da MOS nesses sistemas (FELLER, 1997).

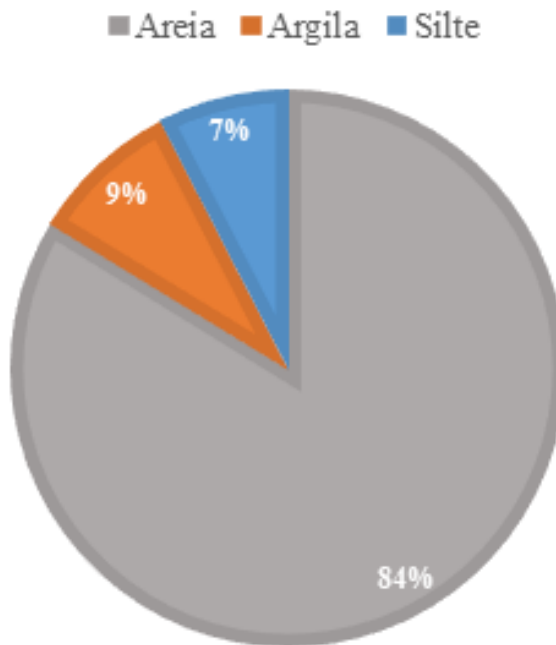


Figura 13. Frações totais de areia, argila e silte do solo do módulo de produção de hortaliças.

4.3. Dados Espectrais

As amostras de solo foram lidas nas porções vis-NIR do espectro eletromagnético (350 a 2500 nm). As curvas espectrais médias obtidas para os anos de 2013 e 2014, para cada área do módulo estão apresentadas na Figura 14, comparadas graficamente na Figura 15 e descritas na Tabela 5. Além disso, o comportamento espectral do solo foi analisado qualitativamente graças à influência composição mineralógica, teor de umidade, teor de matéria orgânica e granulometria no comportamento espectral dos solos (STONER; BAUMGARDNER, 1981; BEN-DOR et al. 1997; DEMATTÊ, 2002).

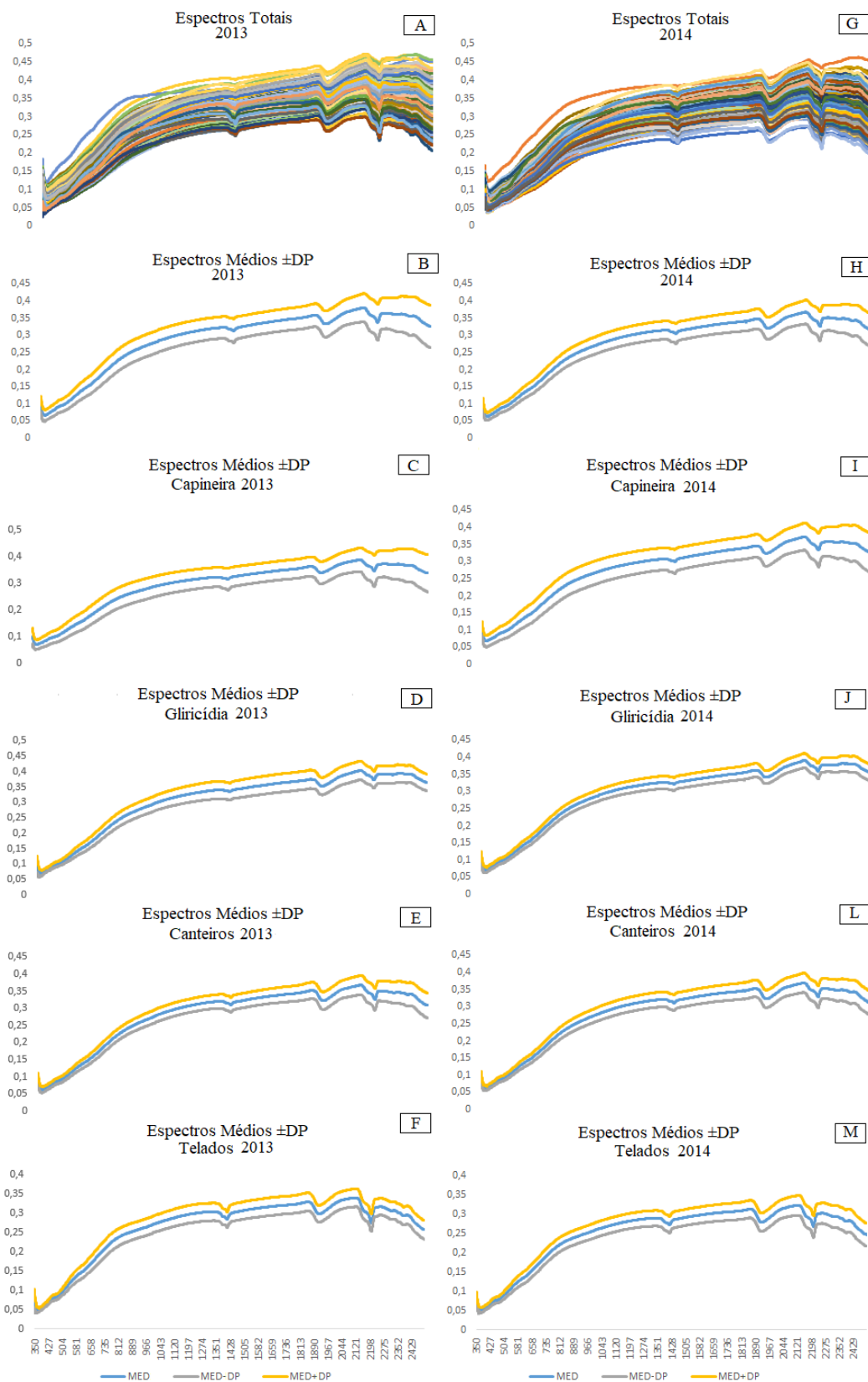


Figura 14. Comportamento espectral total e médio (Med), com \pm desvio padrão (DP), para o módulo como um todo (B, H) e para cada área do módulo experimental sendo elas: Capineira (C, I); Glicírdia (D, J); Canteiros (E, L) e Telados (F, M) para os anos de 2013 e 2014.

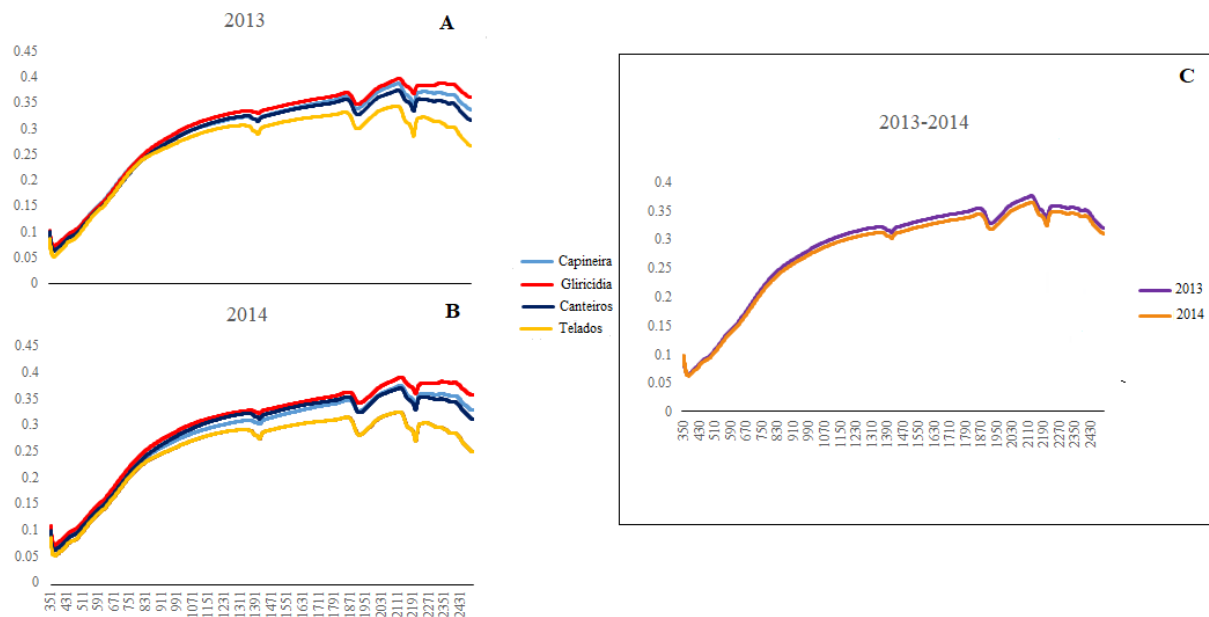


Figura 15. Comportamento espectral médio dos solos para cada área do módulo experimental para os anos de 2013 (A), 2014 (B) e as médias espectrais do módulo como um todo para as amostras referentes aos dois anos.

As amostras de todas as áreas do módulo experimental apresentaram assinaturas espectrais semelhantes para os anos de 2013 e 2014, com picos de absorção nas mesmas regiões (1400, 1900, 2200 nm). Os picos de absorção nos comprimentos de onda próximos de 1.400 nm representam o primeiro sobretom de estiramento das ligações O-H, associadas às moléculas de água ou metal hidratado (VISCARRA ROSSEL & CHEN, 2011). Da mesma forma, a absorção no comprimento de onda 1.900 nm está relacionada à combinação de estiramentos e flexões das ligações O-H, presentes na molécula H₂O (CLARK et al., 1990). Também apareceram bandas de absorção referentes à presença de argilominerais (GROVE et al, 1992); caulinita (2200 nm), e óxido de ferro, que é indicado por uma concavidade presente entre 850 a 900 nm (EPIPHANIO, 1992). Os picos de absorção entre 1.350-1.450 e entre 1.850-2.030 nm podem estar relacionados a pequenas diferenças de umidade das amostras, mesmo após a secagem em estufa, na etapa de preparação. Em relação aos grupamentos ligados à fração da MOS, pôde-se observar a formação de feições referentes a presença de C-H, O-H, N-H (1390 nm), C-O, C-N, O-H, N-H, S-H (1900 nm), C-H, N-H (2200 nm) e C-O, O-H, S-H (2350 nm) (STEVENS, 2008).

Mais especificamente no caso da textura do solo, a influência no vis-NIR ocorre quanto à intensidade da reflectância ao longo de toda a assinatura espectral (albedo), uma vez que solos com maior teor de argila tendem a apresentar maior agregação das partículas que os compõem, o que por sua vez ocasiona menor reflexão da energia eletromagnética incidente (BAUMGARDNER et al., 1985; BELLINASSO et al., 2010).

Tabela 4. Reflectância média (REF) \pm 1 desvio padrão (DP) e matéria orgânica do solo (MOS) para as subáreas do módulo experimental.

Área do módulo	REF	REF + DP	REF - DP	MOS (g.kg ⁻¹)
	2013			
Capineira	0,29	0,33	0,26	10,42
Glicícidia	0,30	0,33	0,28	10,08
Canteiros	0,28	0,30	0,26	13,77
Telados	0,27	0,29	0,25	15,28
2014				
Capineira	0,28	0,32	0,25	9,38
Glicícidia	0,30	0,31	0,28	8,84
Canteiros	0,28	0,31	0,26	13,74
Telados	0,25	0,27	0,23	15,22

REF – reflectância; DP – desvio padrão; MOS – matéria orgânica do solo.

De todas as áreas do módulo experimental, os telados apresentaram menor reflectância média em 2013 e 2014, respectivamente (0,27 e 0,25nm), maior conteúdo de MOS (15,28 e 15,22 g.kg), o maior teor de argila do módulo (12,74%) e morfologia mais horizontalizada das curvas espectrais. Todos esses aspectos confirmam o que foi dito por Demattê (2012), que quanto menor a reflectância espectral, maior o conteúdo de MOS e de argila da amostra de solo. Da mesma forma, porém no sentido inverso, a área de glicícidia apresentou a menor concentração de argila (3,64%), o menor teor de MOS (10,08 g.kg) e a curva espectral média mais ascendente, quando comparada às demais. Características medianas foram encontradas nas áreas de capineira e canteiros com teores de MOS de 10,42 e 13,77 g.kg para o ano 2013 e 9,38 e 13,74 para o ano de 2014 e reflectância média de 0,28 e 0,29nm para o ano de 2013 e 0,28 para ambas as áreas no ano de 2014.

Para ambos os anos analisados, os telados apresentaram maiores valores médios de reflectância enquanto que a glicícidia apresentou os menores. A capineira e os canteiros apresentaram valores equivalentes.

Como o estudo foi feito com amostras do mesmo solo, descartam-se as influências da textura, mineralogia e conteúdo de óxidos sobre as variações entre as curvas espectrais. Logo, essas variações foram atribuídas às diferenças de manejo entre as áreas do módulo, que podem ser relacionadas aos teores de MOS. A MOS, por sua vez, absorve maior radiação eletromagnética, atenuando a intensidade de reflectância dos solos (DALMOLIN et al., 2005).

Pinheiro (2017) e Genu (2010) encontraram feições espectrais nas curvas de solos arenosos e tropicais, na Amazônia e em São Paulo, semelhantes com as obtidas no presente trabalho. Esses solos foram classificados como neossolo quartzarênico e argissolo vermelho-amarelo distrófico arênico, que também tem grande porção de areia em superfície.

4.4. Pré-Processamento dos Dados Espectrais

Os dados espectrais (Vis-NIR) foram pré-processados com objetivo de melhorar a acurácia dos modelos de regressão. As curvas espectrais médias dos dados brutos de reflectância (REF) e pré-processados matematicamente para dados de absorbância (ABS), suavização (SMO), Derivadas Savitzky-Golay (SGD), Variação Normal Padrão (SNV), Correção Multiplicativa de Sinal (MSC) e Correção do Contínuo (CRR) estão apresentados na Figura 16.

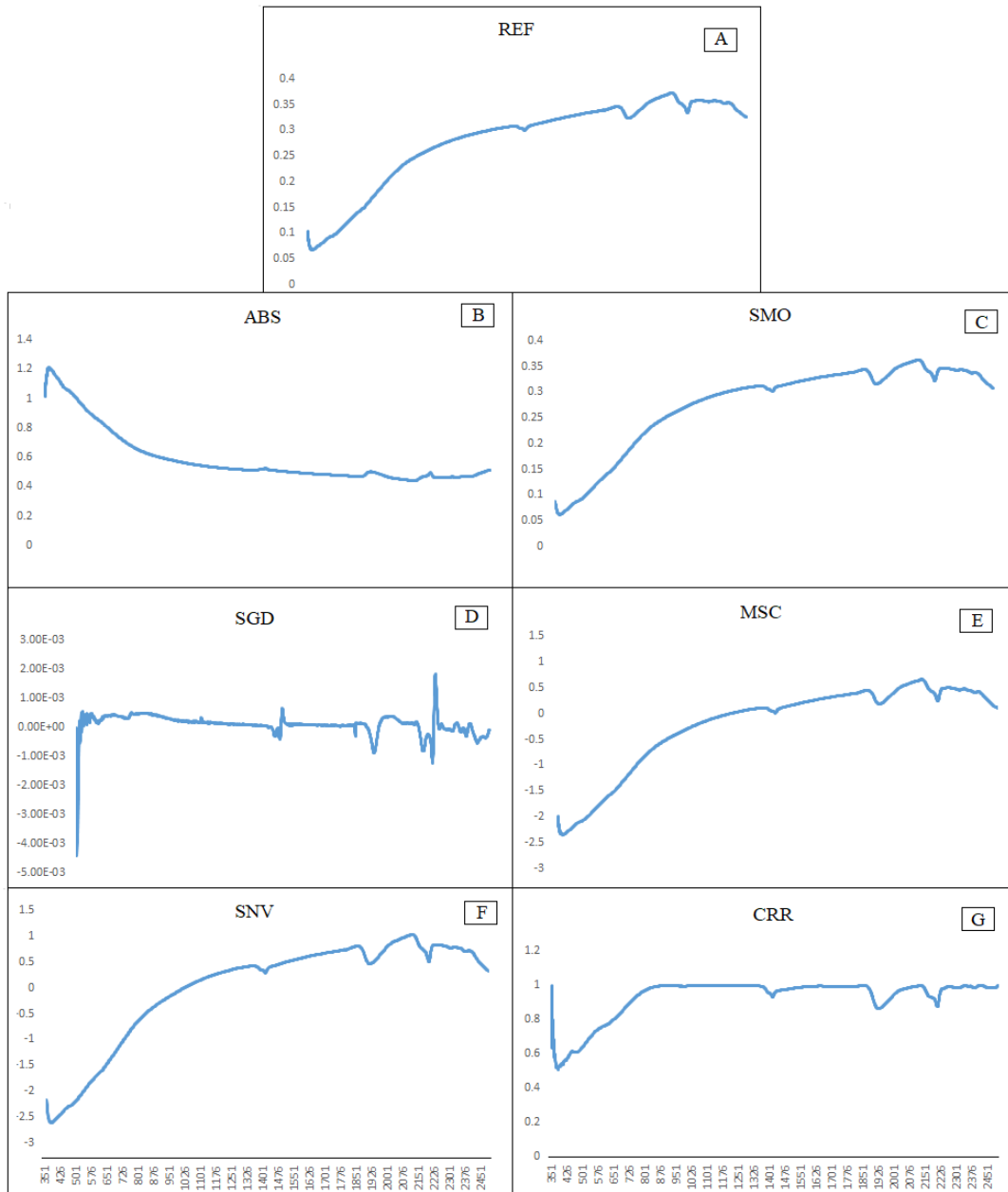


Figura 16. Curvas espectrais médias de REF espectral e após a adição dos pré-processamentos espectrais de ABS, SMO, SGD, SNV, MSC, CRR.

4.5. Modelagem

4.5.1. Identificação de outliers

Os outliers são os dados que apresentam comportamento anômalo em relação aos demais da série, de forma que podem prejudicar a interpretação de resultados obtidos por algoritmos e análises de dados, como é o caso das análises multivariadas. Para esse estudo, as amostras foram identificadas como outliers a partir de um primeiro teste de modelagem PLSR, pela técnica de *leave-one-out*, com os dados brutos de refletância vis-NIR e as concentrações de C, N e P das amostras de terra de 2014 do módulo. A partir de então, os dados mais afastados à correlação linear gerada foram retirados dos conjuntos de dados de treinamento e validação, conforme pode ser visto em vermelho na Figura 17. Esses novos conjuntos (sem outliers) foram

submetidos a todas as etapas posteriores de pré-processamentos espectrais e modelagem a partir de treinamento e validações.

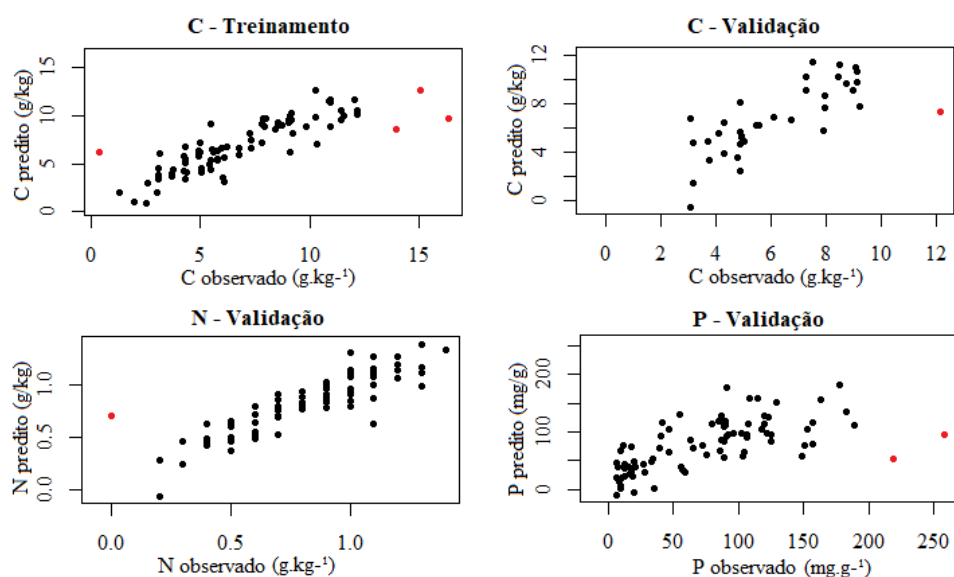


Figura 17. Amostras identificadas como outliers em vermelho para o conjunto de dados de treinamento (trei) e validação (val) para C e somente validação (val) para N e P.

Para os modelos teste de predição de C, foram considerados outliers 4 amostras do conjunto de treinamento e 3 do conjunto de validação, o que corresponde a 4,1% do número total de amostras. Para o modelo de N, apenas uma amostra do conjunto de validação foi considerada outlier, o que representa 0,37% das amostras totais enquanto que para o modelo de P, duas amostras do conjunto de validação foram consideradas outliers, representando 0,47 das amostras totais.

4.5.2. Treinamento e validação dos modelos de predição de C

O treinamento dos modelos de predição de C e o número de componentes PLS capazes de prever concentrações químicas das amostras a partir de dados espectrais com o menor RMSE atribuído variou de 3 a 12 e podem ser observados graficamente em vermelho na Figura 18 e descrito na Tabela 6.

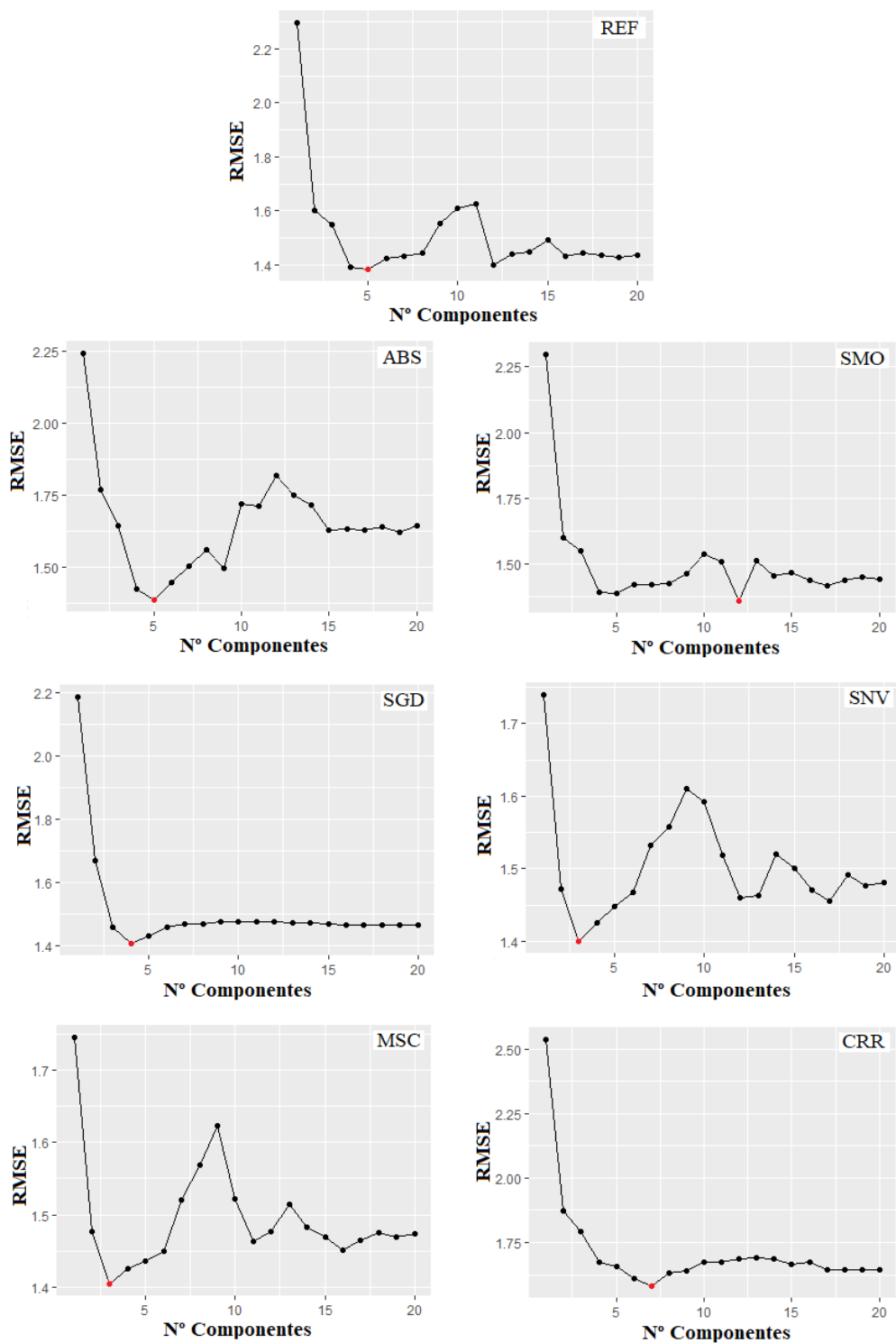


Figura 18. Gráficos de componentes PLS e RMSE para os treinamentos dos modelos preditivos de C.

Após o treinamento dos modelos, foi gerado gráfico de regressão linear representando a diferença entre as concentrações de C observadas e previstas pelo modelo, onde o modelo ideal teria todos os pontos sobre a reta. Os gráficos para os modelos de teste para dados de REF e pré-processados (ABS, SMO, SGD, SNV, MSC, CRR) estão na Figura 19 e dados na Tabela 6.

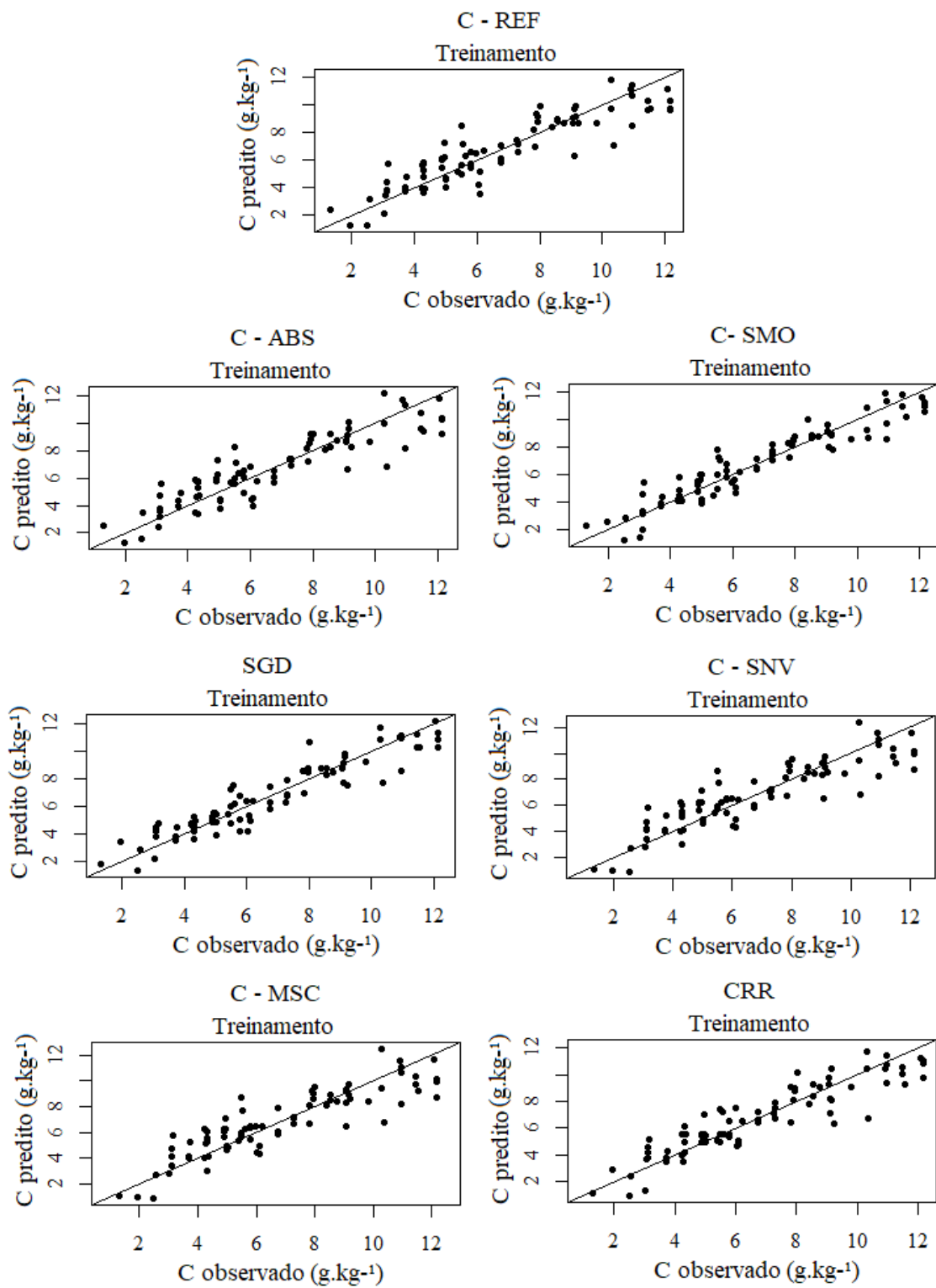


Figura 19. Treinamento dos modelos de predição de C a partir de dados de REF, ABS, SMO, SGD, SNV, MSC e CRR.

Tabela 5. Parâmetros estatísticos obtidos a partir do treinamento e validação dos modelos preditivos de C, bem como o número de componentes PLS escolhido para modelo.

Treinamento - C				
PP	R ²	RMSE	RPIQ	NC
REF	0,81	1,23	3,71	5
ABS	0,80	1,24	3,69	5
SMO	0,89	0,92	4,95	12
SGD	0,88	0,97	4,71	4
SNV	0,78	1,31	3,49	3
MSC	0,78	1,31	3,48	3
CRR	0,83	1,15	3,98	7

REF: reflectância; ABS: absorvância; SMO: suavização; SGD: derivadas Savitzky-Golay; SNV: derivação normal padrão; MSC: correção multiplicativa de sinal; CRR: remoção do contínuo; R²: coeficiente de determinação; RMSE: raiz quadrada do erro médio; RPIQ: relação de desempenho para distâncias interquartis; NC: número de componentes PLS.

O R² é uma medida que descreve a quantidade de variabilidade nos dados que é explicada pelo modelo de regressão ajustado. Dessa forma, quanto maior o valor de R² obtido, melhor o modelo nesse aspecto. Os valores de R² para os treinamentos dos modelos preditivos de C variaram de 0,78 a 0,89; o que indica que todos os modelos tiveram ajustes de bons a muito bons. Comparando com os valores obtidos para o modelo com os dados de REF (sem pré-processamento), alguns algoritmos retiraram dados importantes para o ajuste, principalmente os que corrigiram a linearidade (ABS R²= 0,80) e o espalhamento da luz (SNV e MSC, ambos R²= 0,78). Por outro lado, o modelo SMO apresentou melhor R² (0,89), indicando que as amostras sofreram interferências aleatórias, o que provocou ruídos e conseqüentemente um ajuste prejudicado. Quando o SMO foi aplicado aos dados, houve um aumento do ajuste dos modelos de predição de 0,08%.

O RMSE representa a qualidade do ajuste de um modelo e indica o quanto as respostas variam em torno das suas predições. Dessa forma, quanto mais o RMSE for próximo de zero, melhor a qualidade do ajuste. Os valores de RMSE para os treinamentos dos modelos de predição de C variaram de 0,92 (SMO) a 1,31 (SNV e MSC).

O RPIQ representa o desempenho dos modelos de previsão, através do espalhamento entre os quartis 1 e 3. Os valores de RPIQ para os treinamentos dos modelos de predição de C variaram de 4,95 a 3,48 para SMO e MSC, respectivamente.

Após o treinamento dos modelos, para saber se eles são capazes de predizer o c em amostras externas, ou seja, que não participaram do processo de treinamento, os modelos obtidos a partir dos dados de REF e de todos os pré-tratamentos (ABS, SMO, SGD, SNV, SMO e CRR) foram usados utilizados nos dados do conjunto de validação de 2014 (validação 1) e em seguida para o conjunto global de dados referentes às amostras de 2013 (validação 2). A validação 2 foi realizada com o objetivo de testar uma segunda avaliação desses modelos quimiométricos para monitorar o C do solo do módulo, de forma prática. Esses dados podem ser encontrados na Figura 20 e na Tabela 7.

Para a validação 1, os valores de R² variaram de 0,65 a 0,72; sendo o menor valor obtido para os modelos SMO e MSC e o maior para o modelo SGD. Para RMSE houve variação de

1,57 a 1,98, para os modelos REF e CRR, respectivamente. E para os valores de RPIQ, houve variação entre 3,48 e 4,95; para os modelos MSC e SMO. Esses dados indicam que os modelos para validação 1 apresentaram ajustes razoavelmente bons de bons.

Para a validação 2, os valores de R^2 variaram de 0,35 a 0,43; sendo o menor valor obtido para os modelos SMO e o maior para os modelos SNV e MSC. Para RMSE houve variação de 2,41 a 2,77, para os modelos SNV e SMO, respectivamente. Para os valores de RPIQ, houve variação entre 1,37 e 1,58; para os modelos SMO e SNV. Esses modelos apresentaram baixo potencial preditivo. Isso pode ter acontecido devido a uma transição no manejo do módulo que ocorreu nesse ano, variando as entradas e saídas de C em comparação com o ano de 2014.

O modelo que obteve a melhor performance foi o SNV, que corrige os efeitos negativos provocados pela não-homogeneidade das amostras. Isso pode ter acontecido por se trataram de amostras coletadas em outro período de tempo das demais utilizadas para treinamento e validação 1. Com o tempo em armazenamento, é possível que as amostras tenham ficado não homogêneas em relação à granulometria, com migração das partículas finas para o fundo. Assim, durante a aquisição de dados espectrais pode ter havido espalhamento da luz.

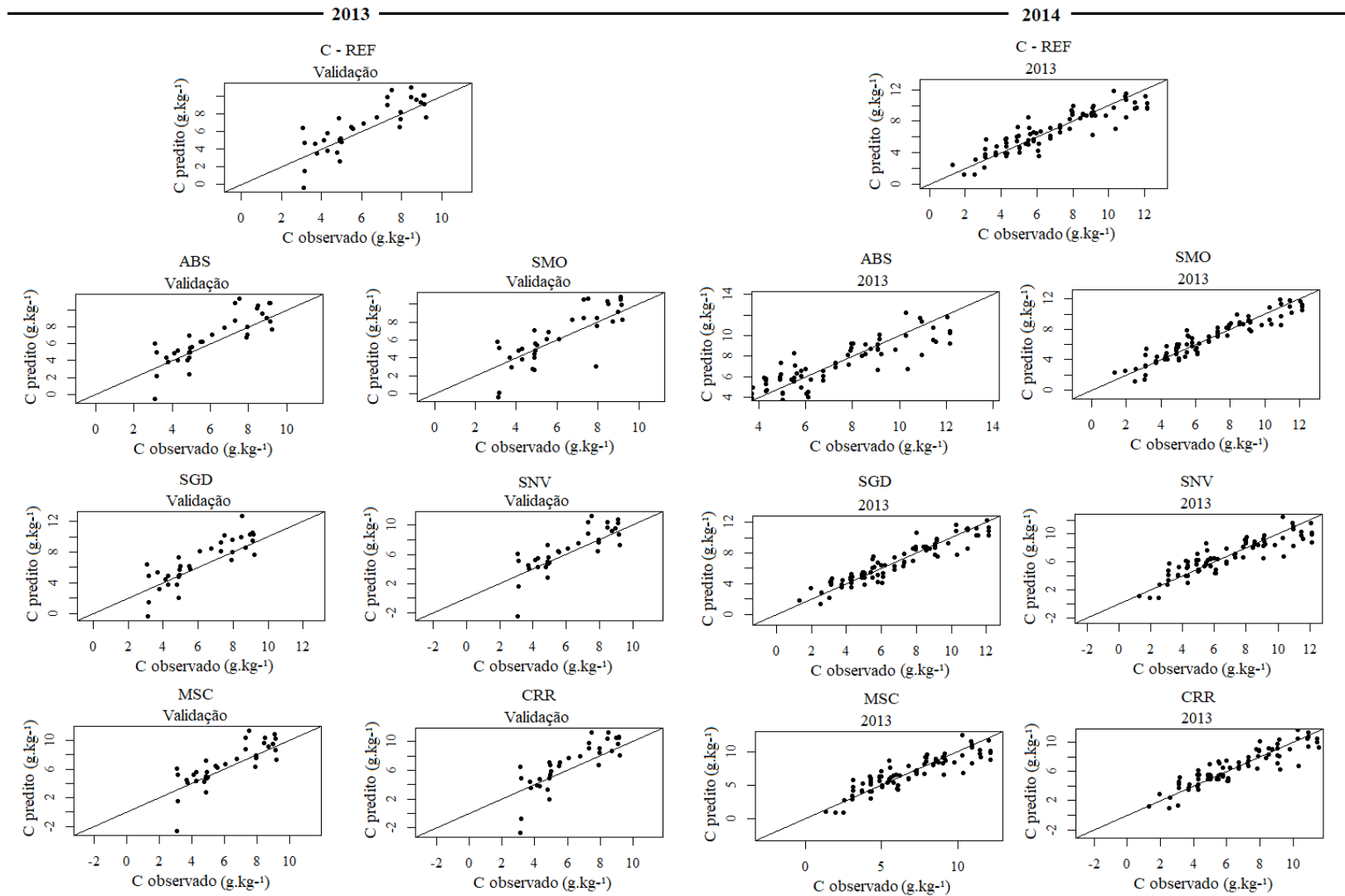


Figura 20. Validação dos modelos preditivos de C para os dados de REF, ABS, SMO, SGD, SNV, MSC e CRR, para os anos de 2013 e 2014.

Tabela 6. Validações dos modelos preditivos de C para os dados de REF, ABS, SMO, SGD, SNV, MSC, CRR para os anos de 2014 e 2013.

C - Validação 1 – 2014			
PP	R²	RMSE	RPIQ
REF	0,71	1,57	3,71
ABS	0,70	1,64	3,69
SMO	0,65	1,79	4,95
SGD	0,72	1,70	4,71
SNV	0,66	1,74	3,49
MCS	0,65	1,76	3,48
CRR	0,69	1,98	3,98
C - Validação 2 – 2013			
PP	R²	RMSE	RPIQ
REF	0,41	2,49	1,53
ABS	0,39	2,51	1,51
SMO	0,35	2,77	1,37
SGD	0,37	2,59	1,47
SNV	0,43	2,41	1,58
MSC	0,43	2,42	1,57
CRR	0,39	2,52	1,52

REF: reflectância; ABS: absorvância; SMO: suavização; SGD: derivadas Savitzky-Golay; SNV: derivação normal padrão; MSC: correção multiplicativa de sinal; CRR: remoção do contínuo; R²: coeficiente de determinação; RMSE: raiz quadrada do erro médio; RPIQ: relação de desempenho para distâncias interquartis; NC: número de componentes PLS.

Comparando o modelo REF com os melhores modelos para cada parâmetro estatístico avaliado em ambas as validações, os valores de R², o RMSE e o RPIQ apresentaram pouca diferença entre eles. Dessa forma, em uma análise corriqueira de rotina, ou até qualitativa, para avaliação do aumento ou decréscimo de C entre as áreas de importação e exportação do módulo, a modelagem poderia ser realizada apenas a partir dos modelos dos dados brutos de REF, sem adição de ferramentas de pré-processamento espectral.

4.5.3. Treinamento e validação dos modelos de predição de N

O treinamento dos modelos de predição de N e o número de componentes PLS capazes de prever concentrações químicas das amostras a partir de dados espectrais com o menor RMSE atribuído variou de 5 a 20 e podem ser observados graficamente na Figura 21.

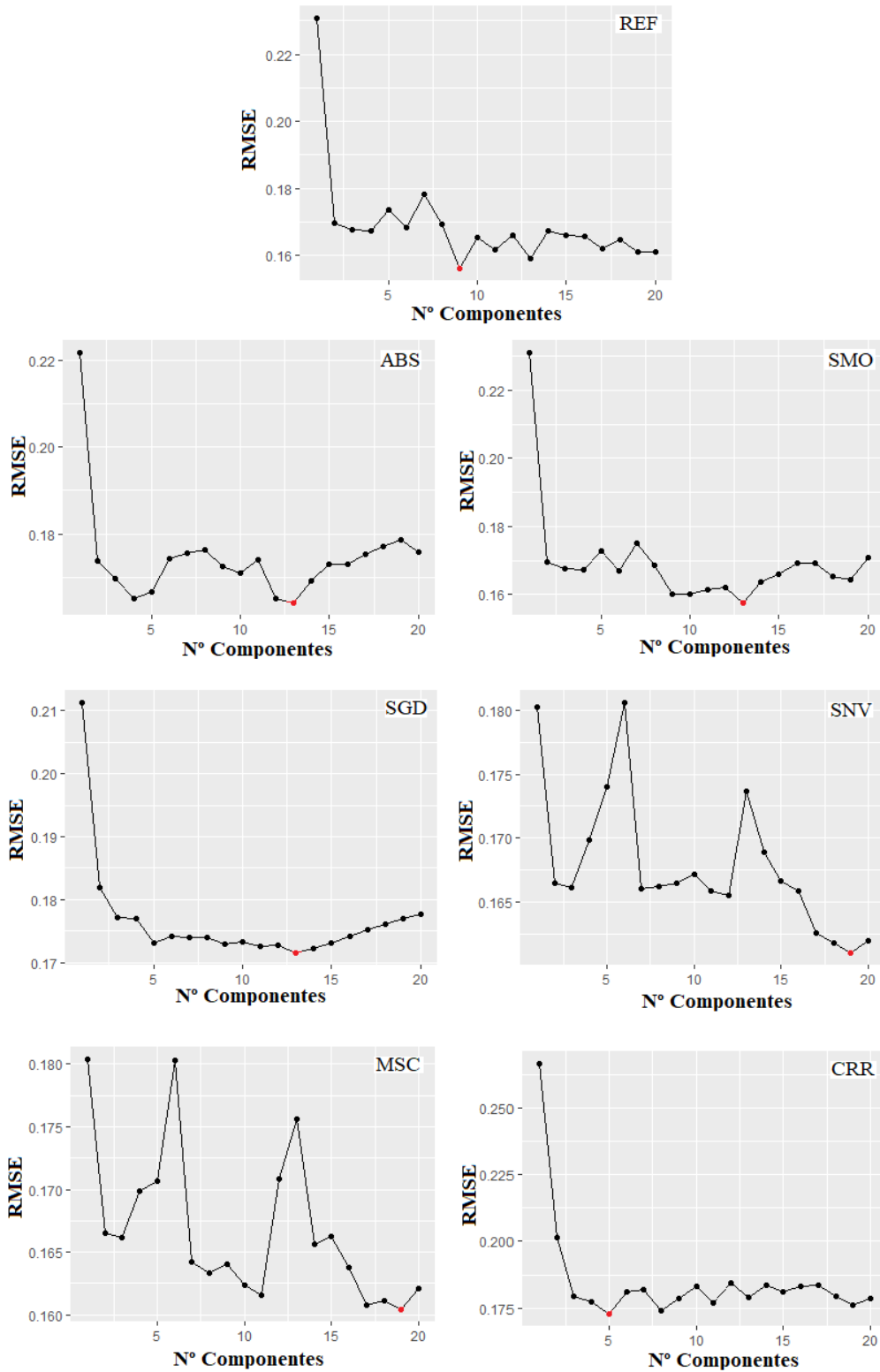


Figura 21. Gráficos de componentes PLS e RMSE para os treinamentos dos modelos preditivos de N.

Em seguida, foram criados os modelos de teste para os dados brutos de reflectância (REF) e os pré-processados espectralmente (ABS, SMO, SGD, SNV, MSC, CRR), apresentados na Figura 22 e descrito na Tabela 8.

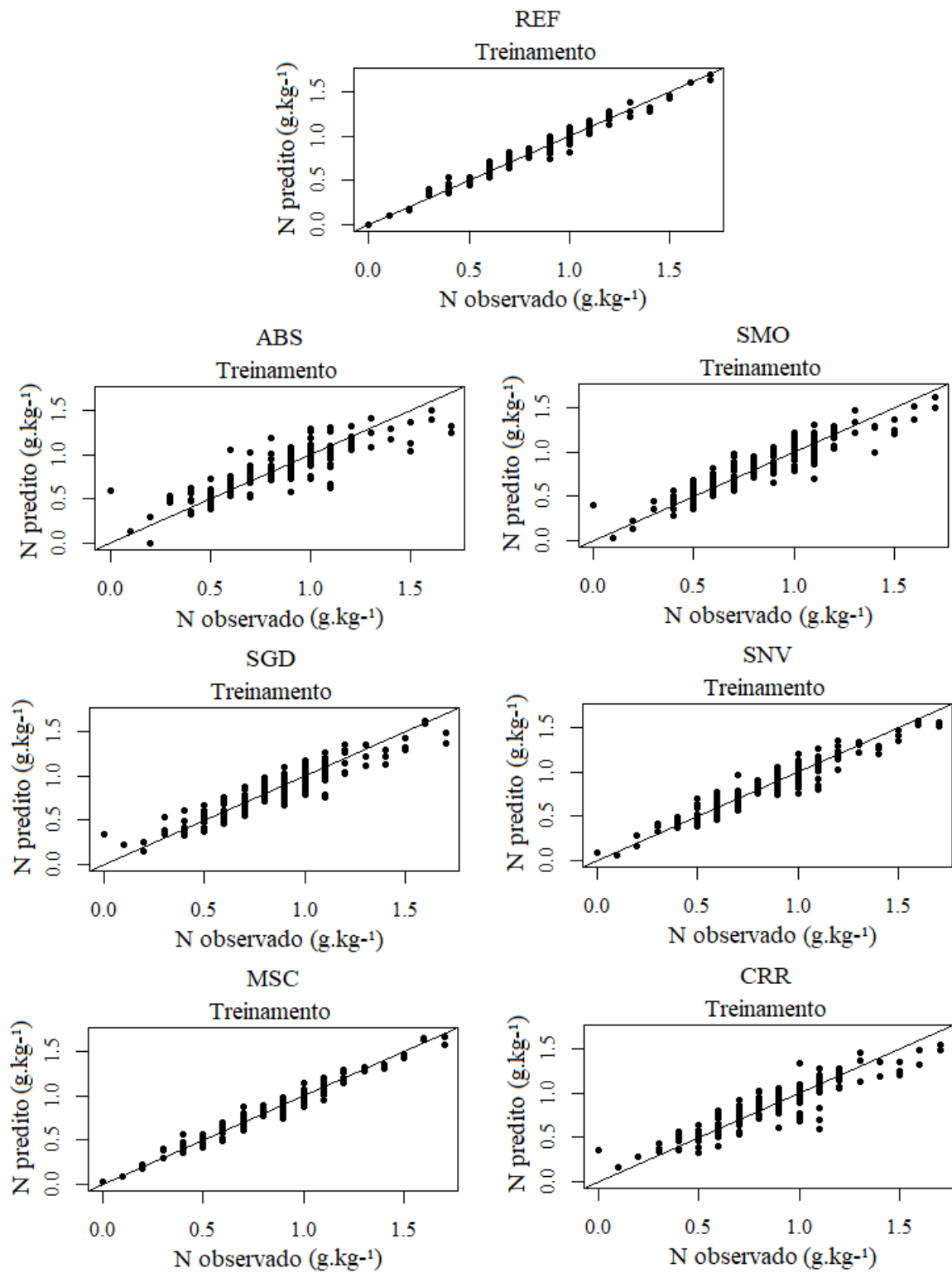


Figura 22. Treinamento dos modelos de predição de N a partir de dados de REF, ABS, SMO, SGD, SNV, MSC e CRR.

Tabela 7. Parâmetros estatísticos obtidos a partir do treinamento e validação dos modelos preditivos, bem como o número de componentes PLS escolhido e utilizado para a predição de C, N e P do solo.

Treinamento – N				
PP	R²	RMSE	RPIQ	NC
REF	0,80	0,13	2,26	9
ABS	0,88	0,10	2,89	13
SMO	0,81	0,13	2,28	13
SGD	0,99	0,03	8,68	13
SNV	0,99	0,03	9,34	19
MSC	0,99	0,03	9,99	19
CRR	0,73	0,15	1,99	5

REF: Reflectância; ABS: Absorbância; SMO: Suavização; SGD: Derivadas Savitzky-Golay; SNV: Derivação normal padrão; MSC: Correção multiplicativa de sinal; R²: Coeficiente de determinação; RMSE: Raiz quadrada do erro médio; RPIQ: Relação de desempenho para distâncias interquartis; NC: número de componentes PLSR.

Os valores de R² para os treinamentos dos modelos preditivos de N variaram de 0,73 (CRR) a 0,99 (SGD, SNV e MSC); o que indica que todos os modelos tiveram ajustes de bons a excelentes. Os valores de RMSE variaram de 0,03 (SGD, SNV e MSC) a 0,15 (CRR). Os valores de RPIQ variaram de 1,99 (CRR) a 9,99 (MSC).

Embora não seja regra, nutrientes como N, Ca e Mg, os quais apresentam fontes minerais com atividade espectral na região vis-NIR costumam apresentar determinações mais assertivas (BEN-DOR et al 1997).

Para a validação 1, os valores de R² variaram de 0,65 a 0,81 para os modelos de SGD e REF, respectivamente. Os valores de RMSE variaram de 0,13 (REF e SMO) a 0,18 (SGD) enquanto que o RPIQ variou de 2,39 (SGD) a 3,34 (REF).

Para a validação 2, os valores de R² variaram de 0,43 (CRR) a 0,61 (SMO); os valores de RMSE variaram de 0,20 (ABS e SMO) a 0,23 (CRR) enquanto que o RPIQ variou de 1,70 (CRR) a 4,95 (ABS).

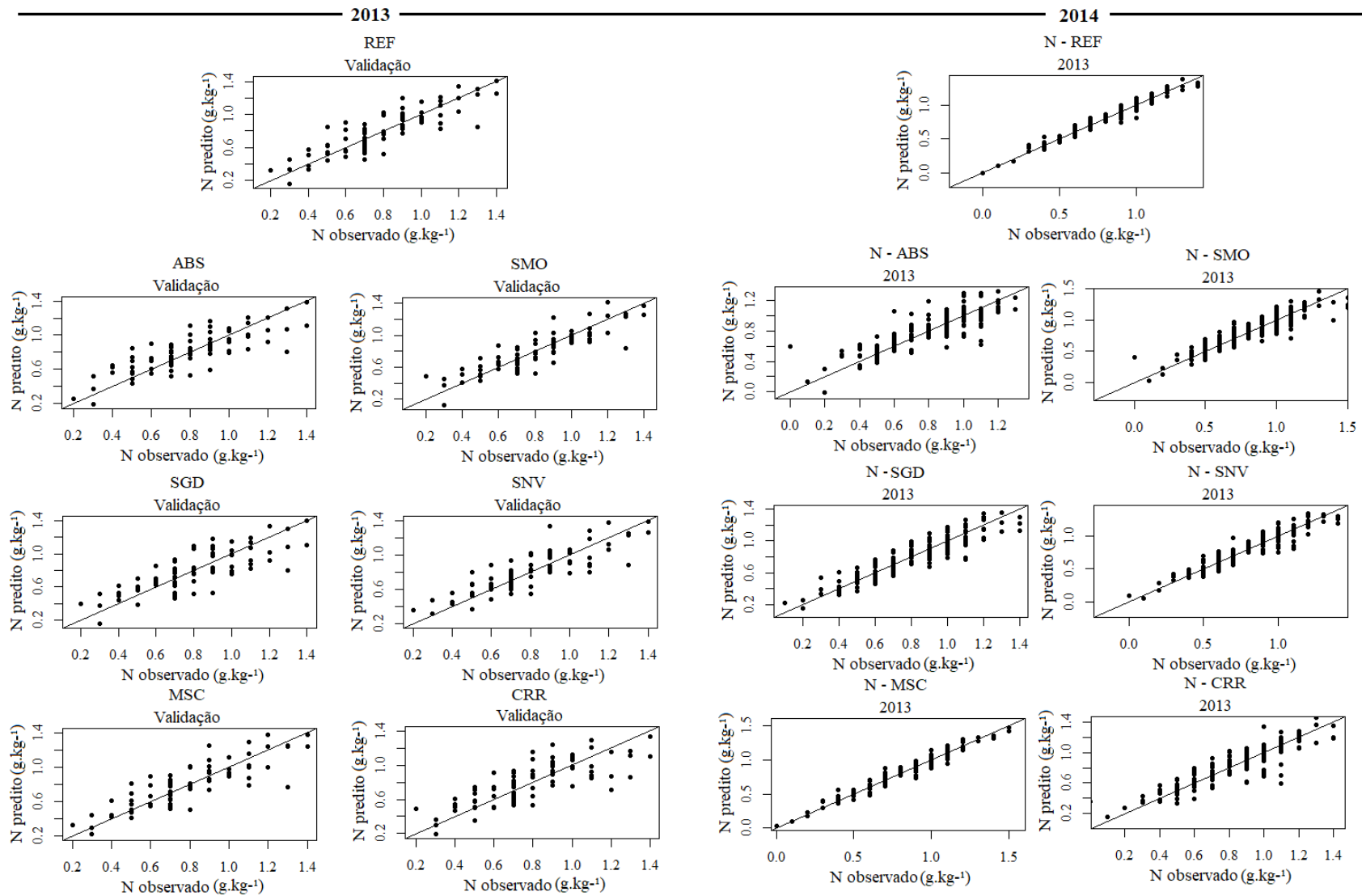


Figura 23. Validações dos modelos preditivos de N para os dados de REF, ABS, SMO, SGD, SNV, MSC, CRR para os anos de 2013 e 2014.

Tabela 9. Validações dos modelos preditivos de N para os dados de REF, ABS, SMO, SGD, SNV, MSC, CRR para os anos de 2014 e 2013.

N - Validação – 2014			
PP	R²	RMSE	RPIQ
REF	0,81	0,13	3,34
ABS	0,77	0,14	3,01
SMO	0,80	0,13	3,28
SGD	0,65	0,18	2,39
SNV	0,75	0,15	2,92
MCS	0,75	0,15	2,84
CRR	0,71	0,16	2,70
N - Validação – 2013			
PP	R²	RMSE	RPIQ
REF	0,56	0,22	1,85
ABS	0,58	0,20	1,98
SMO	0,61	0,20	1,95
SGD	0,56	0,21	1,93
SNV	0,53	0,22	1,82
MSC	0,59	0,21	1,95
CRR	0,43	0,23	1,70

REF: Reflectância; ABS: Absorbância; SMO: Suavização; SGD: Derivadas Savitzky-Golay; SNV: Derivação normal padrão; MSC: Correção multiplicativa de sinal; R²: Coeficiente de determinação; RMSE: Raiz quadrada do erro médio; RPIQ: Relação de desempenho para distâncias interquartis.

4.5.4. Treinamento e validação dos modelos de predição de P

O treinamento dos modelos de predição de P e o número de componentes PLS capazes de prever concentrações químicas das amostras a partir de dados espectrais com o menor RMSE atribuído, variou de 5 a 19 e podem ser observados graficamente em vermelho na Figura 24 e descritos, junto com os modelos na Tabela 10.

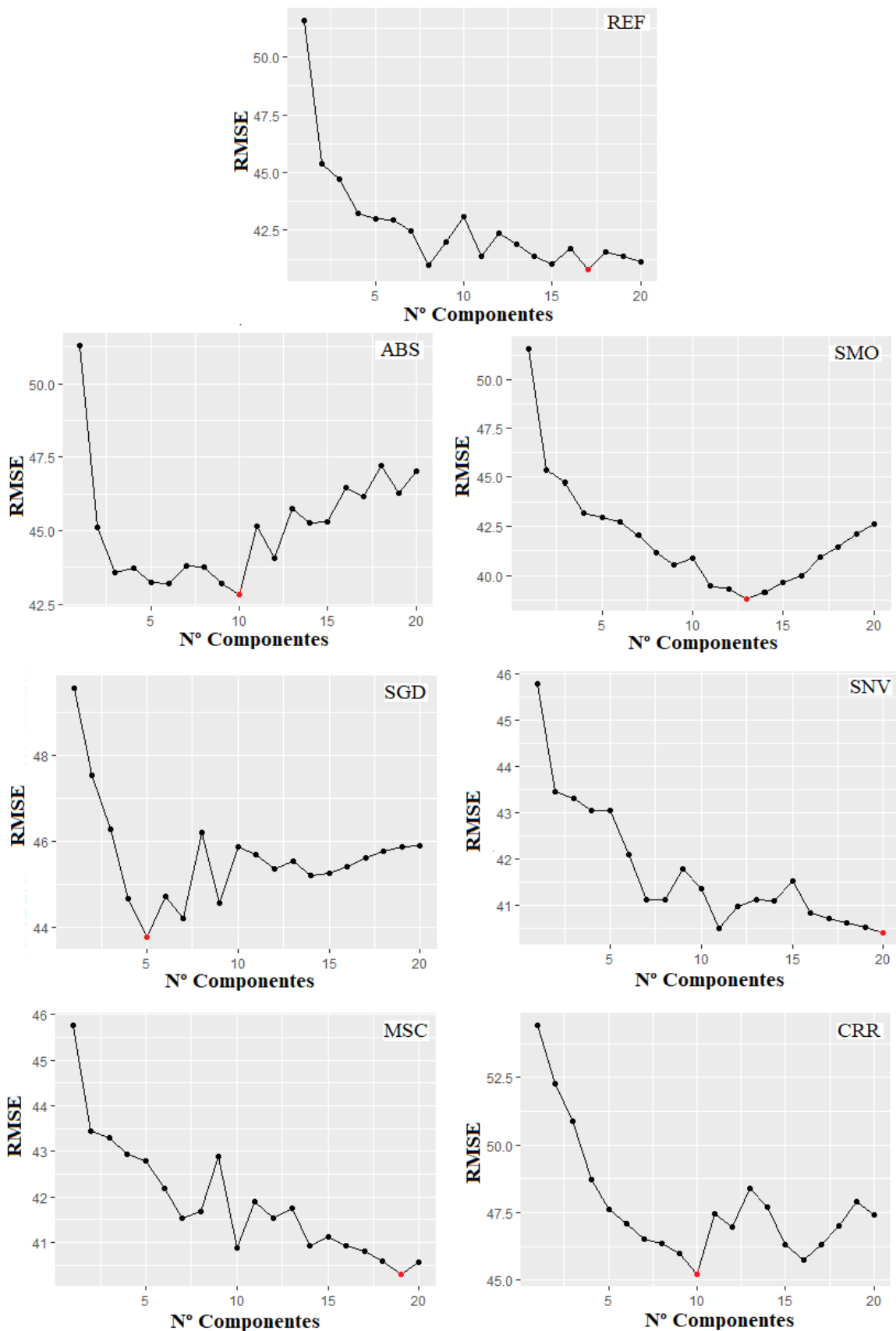


Figura 24. Gráficos de componentes PLS e RMSE para os treinamentos dos modelos preditivos de P.

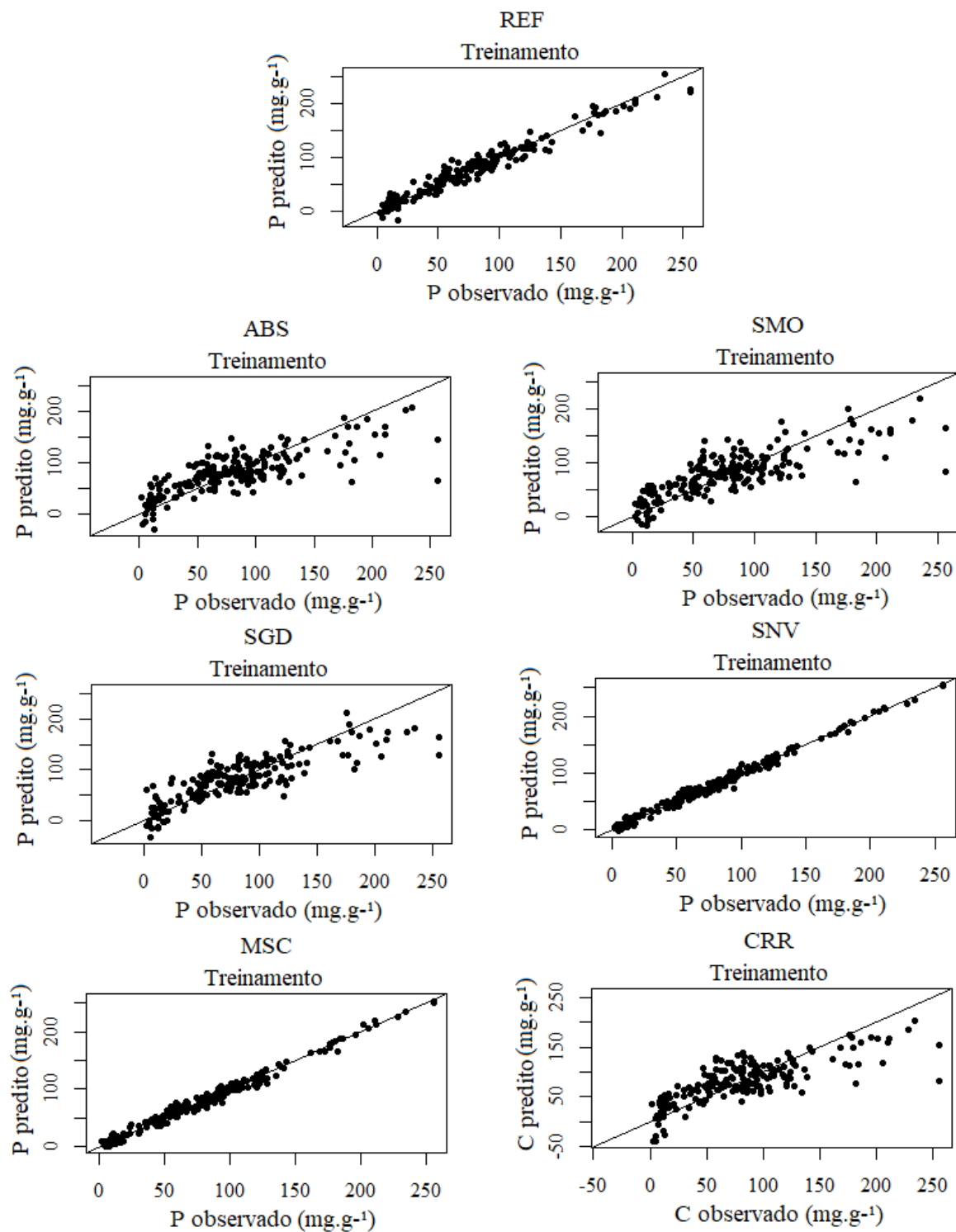


Figura 25. Treinamento dos modelos de predição de P a partir de dados de REF, ABS, SMO, SGD, SNV, MSC e CRR.

Tabela 10. Parâmetros estatísticos obtidos a partir do treinamento e validação dos modelos de predição das concentrações de P do solo.

Treinamento – P				
PP	R²	RMSE	RPIQ	NC
REF	0,95	12,37	4,90	17
ABS	0,59	34,35	1,76	10
SMO	0,62	33,04	1,83	13
SGD	0,70	29,71	2,04	5
SNV	0,99	5,93	10,00	20
MSC	0,98	6,81	8,89	19
CRR	0,61	33,63	1,80	9

REF: Reflectância; ABS: Absorbância; SMO: Suavização; SGD: Derivadas Savitzky-Golay; SNV: Derivação normal padrão; MSC: Correção multiplicativa de sinal; CRR: correção do contínuo; R²: Coeficiente de determinação; RMSE: Raiz quadrada do erro médio; RPIQ: Relação de desempenho para distâncias interquartis; NC: número de componentes PLSR.

Para os treinamentos dos modelos de predição de P os valores de R² apresentaram variação de 0,59 (ABS) a 0,99 (SNV), o que indica que o ajuste dos modelos também variou entre razoavelmente bons a excelentes. Os valores de RMSE variaram de 5,93 (SNV) a 34,35 (ABS). Os valores de RPIQ variaram de 1,76 (CRR) a 10,00 (MSC).

Apesar de nutrientes como o P apresentarem pouca atividade direta nos espectros de solo, os bons modelos de predição gerados através da sua correlação com outros parâmetros do solo (IZNAGA et al. 2014). Como grande parte de absorvância desse solo provavelmente se deu graças a presença de MOS, essa correlação ficou mais evidente, em contraste com outros estudos que obtiveram ajustes de modelos de predição sem qualidade.

Para a validação 1, os valores de R² variaram de 0,50 a 0,60 para os modelos de ABS e SMO, respectivamente. Os valores de RMSE variaram de 32,21 (MSC) a 36,96 (ABS), enquanto o RPIQ variou de 2,33 (ABS) a 2,45 (MSC).

Para a validação 2, os valores de R² variaram de 0,43 (SGD) a 0,56 (SNV); os valores de RMSE variaram de 37,15 (SNV) a 55,02 (MSC) enquanto que o RPIQ variou de 1,48 (MSC) a 2,20 (SNV).

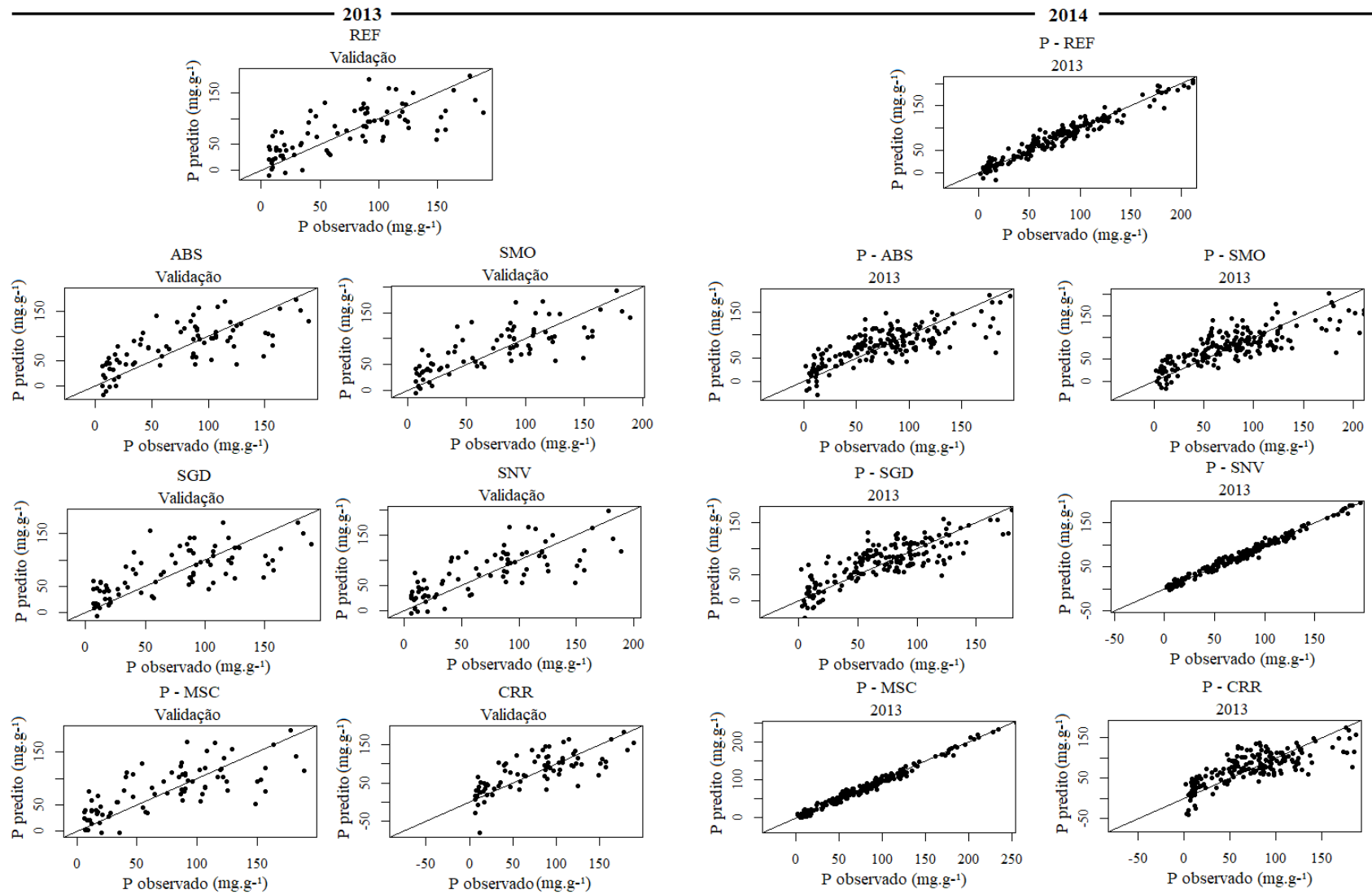


Figura 26. Validações dos modelos preditivos de C para os dados de REF, ABS, SMO, SGD, SNV, MSC, CRR para os anos de 2013 e 2014.

Tabela 11. Validações dos modelos preditivos de P para os dados de REF, ABS, SMO, SGD, SNV, MSC, CRR para os anos de 2014 e 2013.

P - Validação – 2014			
PP	R²	RMSE	RPIQ
REF	0,53	35,84	2,40
ABS	0,50	36,96	2,33
SMO	0,60	32,93	2,62
SGD	0,51	36,08	2,39
SNV	0,57	34,06	2,53
MCS	0,54	32,21	2,45
CRR	0,56	35,49	2,43
P - Validação – 2013			
PP	R²	RMSE	RPIQ
REF	0,52	39,29	2,08
ABS	0,47	40,22	2,03
SMO	0,49	40,55	2,01
SGD	0,43	41,60	1,96
SNV	0,56	37,12	2,20
MSC	0,53	55,02	1,48
CRR	0,50	39,12	2,08

REF: reflectância; ABS: absorvância; SMO: suavização; SGD: derivadas Savitzky-Golay; SNV: derivação normal padrão; MSC: correção multiplicativa de sinal; CRR: correção do contínuo; R²: Coeficiente de determinação; RMSE: Raiz quadrada do erro médio; RPIQ: Relação de desempenho para distâncias interquartis.

5. CONCLUSÕES

As técnicas de espectroscopia de reflectância na região do visível e infravermelho próximo (vis-NIR) se mostraram eficazes para a análise qualitativa e para a predição das concentrações de C, N e P num Planossolo Háplico, situado em Seropédica (RJ). Apenas utilizando os dados espectrais de reflectância (REF) pode-se fazer o monitoramento do módulo de produção orgânica de hortaliças, de uma forma rápida e acurada de avaliação das áreas de importação e exportação de matéria orgânica do solo e, conseqüente tomada de decisão em relação ao manejo a ser implantado ou modificado. Os pré-processamentos que melhor otimizarão as predições de C, N e P no Planossolo Háplico, que possui elevado conteúdo de areia, são os que eliminam o espalhamento da luz (MSC e SNV). Isso se deve, provavelmente, à alta reflectância do quartzo, presente na fração areia desse solo.

Finalizando, através desse estudo ficou claro também que nem todas as amostras de terra coletadas para a predição dos atributos avaliados são necessárias para se obter uma boa acurácia, já que parte das amostras foi perdida e ainda assim os modelos apresentaram bom ajuste na predição dos teores de C, N e P no Planossolo Háplico.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- ABBOUD, A. C.; RIBEIRO, R. L. D.; ESPINDOLA, J. A. Doze anos de Fazendinha Agroecológica do km 47: uma parceria entre UFRRJ, Embrapa Agrobiologia e PESAGRO – Rio: Concurso Nacional de Sistematização de Experiências sobre Agroecologia e Agriculturas Alternativas. **Documentos**, 31 p. Embrapa Agrobiologia. 2005.
- ADAMCHUK, V.; ALLRED, B.; DOOLITTLE, J.; GROTE, K.; VISCARRA ROSSEL, R. **Tools for Proximal Soil Sensing**. 2015.
- ADAMS, M. J. **Chemometrics in analytical spectroscopy**. RSC, Analytical spectroscopy monographs. Cambridge, UK: The Royal Society of Chemistry. 1995.
- ALMEIDA, G.; SILVA, L.; MARCHÃO, R.; WADT, P.; SOUZA, A.; OLIVEIRA, L. Métodos Kjeldahl, Elementar e o Potencial da Espectroscopia do Infravermelho Próximo para Determinação de Nitrogênio em Solos da Bacia do Acre. **Biota Amazônia** 6. 38-43, 2016.
- ALTIERI, M. A.; HECHT, S. B. **Agroecology and Small Farm Development**. Boca Raton: CRC Press. 1989.
- ANDERSON, D. W. Processes of humus formation and transformation in soils of the Canadian Great Plains. **J. Soil Sci.** 30, 77-84, 1979.
- ANDERSON J. M.; SWIFT M. J. Decomposition in tropical forests. In: SUTTON S. C.; WHITMORE T. C.; CHADWICK A. C. (eds), **Tropical Rain Forest: Ecology and Management**. Oxford, UK: Blackwell, pp. 287-310, 1983.
- ARAÚJO, S. R.; DEMATTÊ, J. A. M.; BELLINASSO, H. Analysing the effects of applying agricultural lime to soils by VNIR spectral sensing: a quantitative and quick method. **International Journal of Remote Sensing**, v. 34, n. 13, p. 4570-4584, 2013.
- ARAÚJO, S. R.; SÖDERSTRÖM, M.; ERIKSSON, J.; ISENDAHL, C.; STENBORG, P.; DEMATTÊ, J. M. Determining soil properties in Amazonian Dark Earths by reflectance spectroscopy. **Geoderma**, 237-238, 308-317, 2015.
- ASSIS, R. L.; ROMEIRO, A. R. O processo de conversão de sistemas de produção de hortaliças convencionais para orgânicos. **RAP** 41(5), 863-85. 2007.
- BARNES, R. J.; DHANOA, M. S.; LISTER, S. J. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. **Applied Spectroscopy**, 43(5), 772-777, 1989.
- BAUMGARDNER, M. F.; KRISTOF, S.; JOHANNSEN, C. J.; ZACHARY, A. Effects of organic matter on the multispectral properties of soils. **Proceedings Indiana Academy of Science**, Brookville, v. 79, p. 413-422, 1970.
- BAUMGARDNER, M. F.; STONER, E. R.; SILVA, L. F.; BIEHL, L. L. Reflective properties of soils. In: BRADY, N. (Ed), **Advances in Agronomy**, 38. New York: Academic Press, pp. 1-44. 1985.
- BEEBE, K. R.; KOWALSKI, B. R. An Introduction to multivariate calibration and analysis. **Analytical Chemistry** 59:17, A1007-A1017. 1987.

BELLINASO, H.; DEMATTE, J. A. M.; ROMEIRO, S. A. Soil spectral library and its use in soil classification. **Rev. Bras. Ciênc. Solo**, Viçosa, v. 34, n. 3, p. 861-870, 2010.

BEN-DOR, E. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400-2500 nm) during a controlled decomposition process. **Remote Sensing of Environment**, 61(1), 1-15, 1997.

BEN-DOR, E.; IRONS, J. A.; EPEMA, A. Soil Spectroscopy. In: RENCZ, A. (Ed.), **Manual of Remote Sensing**, Third edition. New York: J. Wiley & Sons, Inc. pp. 111-188. 1999.

BEN-DOR, E.; TAYLOR, R. G.; HILL, J.; DEMATTÊ, J. A. M.; WHITING, M. L.; CHABRILLAT, S.; SOMMER, S. Imaging Spectrometry for Soil Applications. **Advances in Agronomy**, Volume 97, 321-392, 2008.

BEN-DOR, E.; CHABRILLAT, S.; DEMATTÊ, J. A. M.; TAYLOR, G. R.; HILL, J.; WHITING, M. L.; SOMMER, S. Using Imaging Spectroscopy to study soil properties. **Remote Sensing of Environment**, 113, S38-S55, 2009.

BOUMA, J. Using soil survey data for quantitative land evaluation. **Adv. Soil Sci.**, 9:177-213, 1989.

BRERETON, R. G. **Chemometrics: Data Analysis for the Laboratory and Chemical Plant**. Chichester: Wiley, 489 pp. *Journal of Analytical Chemistry - J ANAL CHEM-ENGL TR.* 60. 994-996. 10.1007/s10809-005-0223-6. 2003.

BUDIMAN, M.; MC BRATNEY, A. B.; MENDONÇA-SANTOS, M. L.; DOS SANTOS, H. G. Revisão sobre Funções de Pedotransferência (PTFs) e Novos Métodos de Predição de Classes e Atributos do Solo. **Documentos** n. 45. Rio de Janeiro: Embrapa Solos. 50 p. 2003.

CAMARGO, A. P.; MARIN, F. R.; SENTELHAS, P. C.; PICINI, A. G. Ajuste da equação de Thornthwaite para estimar a evapotranspiração potencial em climas áridos e superúmidos, com base na amplitude térmica diária. **Revista Brasileira de Agrometeorologia**, v. 7, p. 252-257, 1999.

CARRA, J. B.; FABRIS, M.; DIECKOW, J.; BRITO, O. R.; VENDRAME, P. R. S.; MACEDO DOS SANTOS TONIAL, L. Near-infrared spectroscopy coupled with chemometrics tools: a rapid and non-destructive alternative on soil evaluation. **Communications in Soil Science and Plant Analysis**, 1-14, 2019.

CLARK, R. N.; ROUSH, L. Reflectance Spectroscopy Quantitative Analysis Techniques for Remote Sensing Applications. **Journal of Geophysical Research**, 89, 6329-6340, 1984.

CLARK, R. N.; KING, T. V. V.; KLEJWA, M.; SWAYZE, G. A.; VERGO, N. High spectral resolution reflectance spectroscopy of minerals. **Journal Geophysical Research**, v. 95, p. 12653-12680, 1990.

CONFORTI, M.; BUTTAFUOCO, G.; LEONE, A. P.; AUCELLI, P. P. C.; ROBUSTELLI, G.; SCARCIGLIA, F. Studying the relationship between water-induced soil erosion and soil organic matter using Vis-NIR spectroscopy and geomorphological analysis: A case study in southern Italy. **Catena**, 110, 44-58, 2013.

CORWIN, D. L.; LESCH, S. M. Apparent soil electrical conductivity measurements in agriculture. **Comput. Electron. Agric.** 46, 11-43, 2005.

COZZOLINO, D.; MORÓN, A. The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics. **Journal of Agricultural Science.** 140. 65-71, 2003.

COZZOLINO, D.; MORÓN, A. The potential of near-infrared reflectance spectroscopy chemometrics to predict soil organic carbon fractions. **Soil and Tillage Research**, v. 85, p.78-85, 2006.

D'ANDRÉA, A. F.; SILVA, M. L. N.; CURI, N.; GUILHERME, L. R. G. Estoque de carbono e nitrogênio e formas de nitrogênio mineral em um solo submetido a diferentes sistemas de manejo. **Pesq. Agropec. Bras.**, 39:179-186, 2004.

DALMOLIN, R. S. D.; GONÇALVES, C. N.; KLAMT, E.; DICK, D. P. Relações entre os constituintes do solo e seu comportamento espectral. **Ciência Rural**, Santa Maria, v. 35, n. 2, p. 481- 489, 2005.

DEMATTE, J. A. M.; GARCIA, G. J. Alteration of soil properties through a weathering sequence as evaluated by spectral reflectance. **Soil Sci. Soc. Am. J.**, 63:327-342, 1999.

DEMATTE, J. A. M. Characterization and discrimination of soils by their reflected electromagnetic energy. **Pesquisa Agropecuária Brasileira**, Brasília, v. 37, n. 10, p. 1445-1458, 2002.

DEMATTE, J. A. M.; TOLEDO, A. M. A.; SIMOES, M. S. Metodologia para reconhecimento de três solos por sensores: laboratorial e orbital. **Rev. Bras. Ciênc. Solo**, Viçosa, v. 28, n. 5, p. 877-889, Oct. 2004.

DEMATTE, J. A. M.; SOUSA, A. A.; ALVES, M. C.; NANNI, M. R.; FIORIO, P. R.; CAMPOS, R. C. Determining soil water status and other soil characteristics by spectral proximal sensing. **Geoderma**, 135, 179-195, 2006.

DEMATTE, J. A. M.; BORTOLETTO, M. A. M.; VASQUES, G. M.; RIZZO, R. Quantificação de matéria orgânica do solo através de modelos matemáticos utilizando colorimetria no sistema Munsell de cores. **Bragantia**, 70(3), 590-597, 2011.

DEVOS, O.; DUPONCHEL, L. Parallel genetic algorithm co-optimization of spectral pre-processing and wavelength selection for PLS regression. **Chemometrics and Intelligent Laboratory Systems**, 107(1), 50-58, 2011.

DIAS, J. E. **Monitoramento do uso da terra e dos níveis de nutrientes do solo no Sistema Integrado de Produção Agroecológica utilizando geoprocessamento.** 2007. Tese (Doutorado em Fitotecnia). Departamento de Fitotecnia, Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro 111p. Seropédica, RJ, 2007.

DOTTO, A. C.; DALMOLIN, R. S. D.; GRUNWALD, S.; TEN CATEN, A.; PEREIRA FILHO, W. Two preprocessing techniques to reduce model covariables in soil property predictions by Vis-NIR spectroscopy. **Soil and Tillage Research**, 172, 59-68, 2017.

DOTTO, A. C.; DALMOLIN, R. S. D.; TEN CATEN, A.; GRUNWALD, S. A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. **Geoderma**, 314, 262-274, 2018.

EBELING, A. G.; ANJOS, L. H. C.; PEREIRA, M. G.; PINHEIRO, E. F. M.; VALLADARES, G. S. Substâncias húmicas e relação com atributos edáficos. **Bragantia**, p. 70:157, v.65, 2008.

EMBRAPA. Levantamento semidetalhado dos solos da área do Sistema Integrado de Produção Agroecológica (SIPA) – km 47 – Seropédica, RJ. **Boletim de Pesquisa**, n. 5. Rio de Janeiro: Embrapa Solos. CD-ROM. 1999.

EMBRAPA. **Sistema Brasileiro de Classificação de Solos – SiBCS** (3ª edição). Rio de Janeiro: Centro Nacional de Pesquisa de Solos. 2013.

EIPHANIO, J. C. N.; FORMAGGIO, A. R.; VALERIANO, M. M.; OLIVEIRA, J. B. **Comportamento espectral de solos do Estado de São Paulo**. São José dos Campos: INPE (INPE-5424-PRP-172). 131p. 1992.

FELIX, J. C.; VENDRAME, P. R. S.; MARCHÃO, R. L.; OLIVEIRA, J. F. DE; GUIMARÃES, M. DE F.; BROSSARD, M.; BRITO, O. R. Predição de fósforo, carbono e nitrogênio em solos de basalto, por meio de espectroscopia NIR. **Pesquisa Agropecuária Brasileira**, 51(9), 1405-1416, 2016.

FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S., VOLPE, P. L O. **Chemometrics I – Multivariate calibration, a tutorial**. *Quim. Nova*, 22, 724-731, 1999.

FERREIRA, M. M. C. **Quimiometria** - Conceitos, Métodos e Aplicações, 1ª ed., Campinas: Editora Unicamp, 2015.

FREITAS, P. L. DE; BLANCANEUX, P.; GAVINELLI, E.; LARRÉ-LARROUY, M. C.; FELLER, C. Nível e natureza do estoque orgânico de latossolos sob diferentes sistemas de uso e manejo. **Pesquisa Agropecuária Brasileira**, 35(1), 157-170, 2000.

GEBBERS, R.; ADAMCHUK, V. I. Precision Agriculture and Food Security. **Science**, 327(5967), 2010.

GELADI, P.; MACDOUGALL, D.; MARTENS, H. Linearization and scatter-correction for near-infrared reflectance spectra of meat. **Appl. Spectrosc.**, 39:491-500, 1985.

GENÚ, A. M.; DEMATTÊ, J. A. M. Informações espectroeletrônicas e topográficas na determinação de teores de componentes do solo. **Bragantia**, 69:157-164, 2010.

GREGORICH E. G.; CARTER M. R.; ANGERS, D. A.; MONREAL, C. M.; ELLERT, B. H. Towards a minimum data set to assess soil organic matter quality in agricultural soils. **Canadian Journal of Soil Science** 74, 36-385, 1994.

GROVE, C. I.; PAYLOR II, E. D.; HOOK, S. J. **Laboratory Reflectance Spectra of 160 Minerals, 0.4 to 2.5 Micrometers**. California, USA.: Jpl Publication 92-2, NASA – National Aeronautics And Space Administration, Jet Propulsion Laboratory. 1992.

HOLLAS, J. M. **Modern Spectroscopy**. Wiley, 2004.

JARVIS, R. M.; GOODACRE, R. Genetic algorithm optimization for pre-processing and variable selection of spectroscopic data. **Bioinformatics**, 21(7), 860-868, 2005.

LAL, R.; PIERCE, F. J. Soil management for sustainability. Ankeny, **Soil Water Conservation Society**, 1991. p.1-5.

MATA, M. G. F. **Qualidade do solo e avaliação microeconômica de um módulo experimental de produção orgânica intensiva de hortaliças**. 2012. 75f. Dissertação (Mestrado em Agronomia, Ciência do Solo). Instituto de Agronomia, Departamento de Solos, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2012.

MATA, M. G. F. **Qualidade do solo de um sistema orgânico de produção intensiva de hortaliças em Seropédica-RJ**. 2016. 116f. Tese (Doutorado em Agronomia, Ciência do Solo). Instituto de Agronomia, Departamento de Solos, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2016.

MCBRATNEY, A. B.; MINASNY, B.; VISCARRA ROSSEL, R. Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. **Geoderma**, 136(1-2), 272-278, 2006.

MCCARTY, G. W.; REEVES, J. B.; REEVES, V. B., FOLLETT, R. F.; KIMBLE, J. M. Mid-Infrared and Near-Infrared Diffuse Reflectance Spectroscopy for Soil Carbon Measurement. **Soil Science Society of America Journal**, 66(2), 640, 2002.

MEHLICH, A. **Determination of P, Ca, Mg, K, Na and NH₄ by North Carolina Soil Testing Laboratories**. Raleigh: University of North Carolina, 1953.

MORAES, E. C. **Fundamentos de Sensoriamento Remoto**. Capítulo 1. São José dos Campos – SP: INPE, 2002.

MOUAZEN, A. M.; DE BAERDEMAEKER, J.; RAMON, H. Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer. **Soil and Tillage Research**, 80(1-2), 171-183, 2005.

MOUAZEN, A. M.; KAROUI, R.; DECKERS, J.; DE BAERDEMAEKER, J.; RAMON, H. Potential of visible and near-infrared spectroscopy to derive colour groups utilising the Munsell soil colour charts. **Biosystems Engineering**, 97(2), 131-143, 2007.

MOURA-BUENO, J. M.; DALMOLIN, R. S. D.; TEN CATEN, A.; DOTTO, A. C.; DEMATTÊ, J. A. M. Stratification of a local VIS-NIR-SWIR spectral library by homogeneity criteria yields more accurate soil organic carbon predictions. **Geoderma**, 337, 565-581, 2019.

NOVO, E. M. L. M. **Sensoriamento remoto: princípios e aplicações**. São Paulo: Edgard Blücher, 1992. 308p.

NUNES, R. S.; LOPES, A. L. C.; SOUSA, D. M. G.; MENDES, I. C. Sistemas de manejo e os estoques de carbono e nitrogênio em Latossolo de Cerrado com a sucessão soja-milho. **R. Bras. Ci. Solo**, 35:1407-1419, 2011.

PASQUINI, C. Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. **Analytica Chimica Acta**, 1026, 8-36, 2018.

PINHEIRO, E.; CEDDIA, M.; CLINGENSMITH, C.; GRUNWALD, S.; VASQUES, G. Prediction of Soil Physical and Chemical Properties by Visible and Near-Infrared Diffuse Reflectance Spectroscopy in the Central Amazon. **Remote Sensing**, 9(4), 293, 2017.

R CORE TEAM. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. 2017. URL <https://www.R-project.org/>.

REEVES, D. W. The role of soil organic matter in maintaining soil quality in continuous cropping systems. **Soil and Tillage Research**, 43(1-2), 131-167, 1997.

REEVES, J.; MCCARTY, G.; MIMMO, T. The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. **Environmental Pollution**, 116, S277-S284, 2002.

RIGHETTO, R. T. S.; VALARINI, P. J. Indicadores biológicos e bioquímicos da qualidade do solo: manual técnico. **Documentos**, 21. 198p. Jaguariúna: Embrapa Meio Ambiente, 2000.

RONQUIM, C. C. Conceitos de fertilidade do solo e manejo adequado para as regiões tropicais. **Boletim de Pesquisa e Desenvolvimento**. Campinas: Embrapa Monitoramento por Satélite. 26 p. 2010.

ROSSEL, R. A. V.; MCBRATNEY, A. B. Diffuse Reflectance Spectroscopy as a Tool for Digital Soil Mapping. **Digital Soil Mapping with Limited Data**, 165-172.

ROSSEL, R. A. V.; CHEN, C. Digitally mapping the information content of visible–near infrared spectra of surficial Australian soils. **Remote Sensing of Environment**, 115(6), 2011.

SANCHEZ, E.; KOWALSKI, B. R. J. **Chemometrics** 2, 265. 1988.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least-Squares Procedures. **Analytical Chemistry**, 36, 1627-1639, 1964.

SIMÕES, S. S. **Desenvolvimento de métodos validados para a determinação de captopril usando espectrofotometria NIR e calibração multivariada**. 2008. Tese (Doutorado em Ciências) Departamento de Química, Universidade Federal da Paraíba, João Pessoa, 2008.

SHEPHERD, K. D.; WALSH, M. G. Development of reflectance spectral libraries for characterization of soil properties. **Soil Science Society of America Journal**, Madison, v. 66, p. 988-998, 2002.

SHEPHERD, M. A.; HARRISON, R.; WEBB, J. Managing soil organic matter - implications for soil structure on organic farms. **Soil Use and Management**, 18, 284-292, 2006.

SHEPHERD, K. D.; WALSH, M. G. Infrared Spectroscopy - Enabling an Evidence-Based Diagnostic Surveillance Approach to Agricultural and Environmental Management in Developing Countries. **Journal of Near Infrared Spectroscopy**, 15(1), 1-19, 2007.

SILVA, F. C. **Manual de Análises Químicas de Solos, Plantas e Fertilizantes**. Brasília-DF: Embrapa Informação Tecnológica. 627p. 2009.

SIQUEIRA, D. S.; MARQUES, J.; PEREIRA, G. T. The use of landforms to predict the variability of soil and orange attributes. **Geoderma**, 155(1-2), 55-66, 2010.

- SOUSA JUNIOR, J. G.; DEMATTÊ, J. A. M.; ARAÚJO, S. R. Modelos espectrais terrestres e orbitais na determinação de teores de atributos dos solos: potencial e custos. **Bragantia**, v. 7, n. 3, p. 610-621, 2011.
- STANFORD, G.; FRERE, M. H.; SCHWANIGER, D. H. Temperature coefficient of soil nitrogen mineralization. **Soil Science**. 115: 321-23, 1973.
- STENBERG, B.; VISCARRA ROSSEL, R. A.; MOUAZEN, A. M.; WETTERLIND. Visible and near infrared spectroscopy in soil science. **Advances in Agronomy**, v. 107, p. 163-215, 2010.
- STEVENS, A.; WESEMAEL, B.; BARTHOLOMEUS, H.; ROSILLON, D.; TYCHON, B.; BEN-DOR, E. Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. **Geoderma**. 144: 395-404. 2008.
- STEVENSON, F. J. **Humus chemistry: genesis, composition, reactions**. 2. ed. New York, USA: John Willey, 496 pp, 1994.
- STONER, E. R.; BAUMGARDNER, M. F. Characteristic variations in reflectance of surface soils. **Soil Science Society of America Journal**, Madison, v. 45, p.1161-1165, 1981.
- STONER, E. R.; BAUMGARDNER, M. F. Data acquisition through remote sensing. In: BEATTY, M. T. Planning the uses and management of land. Madison. **Soil Science Society of America**, 1986. p. 159-185.
- TAM, N. F. Y., WONG, Y. S. Variations of Soil Nutrient and Organic Matter Content in a Subtropical Mangrove Ecosystem. **Water, Air, & Soil Pollution** 103, 245-261, 1998.
- TERRA, F. S.; DEMATTÊ, J. A. M.; VISCARRA ROSSEL, R. A. Spectral libraries for quantitative analyses of tropical Brazilian soils: Comparing vis-NIR and mid-IR reflectance data. **Geoderma**, 255-256, 81-93, 2015.
- TERRA, F. S.; DEMATTÊ, J. A. M.; VISCARRA ROSSEL, R. A. Proximal spectral sensing in pedological assessments: vis-NIR spectra for soil classification based on weathering and pedogenesis. **Geoderma**, 318, 123-136, 2018.
- TERRA, F. S.; VISCARRA ROSSEL, R. A.; DEMATTÊ, J. A. M. Spectral fusion by Outer Product Analysis (OPA) to improve predictions of soil organic C. **Geoderma**, 335, 35-46, 2019.
- THENG, B. K. G.; TATE, K. R.; SOLLINS, P. Constituents of organic matter in temperate and tropical soils. In: COLEMAN, D. C.; OADES, J. M.; UEHARA, G., eds. **Dynamics of soil organic matter in tropical ecosystems**. Hawaii: Niftal Project, p. 5-32. 1989.
- TSAI, F.; PHILPOT, W. Derivative analysis of hyperspectral data. **Remote Sensing of Environment**, 66(1): 41-51, 1998.
- VASQUES, G. M.; GRUNWALD, S.; SICKMAN, J. O. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. **Geoderma**, 146(1-2), 14-25, 2008.

VASQUES, G. M.; GRUNWALD, S.; SICKMAN, J. O. Modeling of soil organic carbon fractions using visible–near-infrared spectroscopy. **Soil Sci. Soc. Am. J.** 73:176-184, 2009.

VEIGA, J. E. da. A. Agricultura no mundo moderno: diagnóstico e perspectivas. In TRIGUEIRO, A (Org.). **Meio ambiente no século 21: especialistas falam da questão ambiental nas suas áreas de conhecimento.** Rio de Janeiro: Sextante. p. 199-213. 2003.

VISCARRA ROSSEL, R. A.; MCBRATNEY, A. B. Laboratory evaluation of a proximal sensing technique for simultaneous measurement of soil clay and water content. **Geoderma**, 85(1), 19-39, 1998.

VISCARRA ROSSEL, R. A.; WALVOORT, D. J. J.; MCBRATNEY, A. B.; JANIK, L. J.; SKJEMSTAD, J. O. Visible, near infrared, mid-infrared or combined use reflectance spectroscopy for simultaneous assessment of various soil properties. **Geoderma**, 131, 59-75, 2003.

VISCARRA ROSSEL, R. A.; MCBRATNEY, A. B. Diffuse Reflectance Spectroscopy as a Tool for Digital Soil Mapping. **Digital Soil Mapping with Limited Data**, 165-172, 2006.

VISCARRA ROSSEL, R. A. ParLeS: Software for chemometric analysis of spectroscopic data. **Chemometrics and Intelligent Laboratory Systems**, 90(1), 72-83, 2008.

VISCARRA ROSSEL, R. A.; BEHRENS, T. Using Data Mining to Model and Interpret Soil Diffuse Reflectance Spectra. **Geoderma**, 158, 46-54, 2010.

VISCARRA ROSSEL, R. A.; ADAMCHUK, V. I.; SUDDUTH, K. A.; MCKENZIE, N. J.; LOBSEY, C. Proximal Soil Sensing: An Effective Approach for Soil Measurements in Space and Time. **Advances in Agronomy**, 243-291, 2011.

VISCARRA ROSSEL, R. A.; CHAPPELL, A.; DE CARITAT, P.; MCKENZIE, N. J. On the soil information content of visible-near infrared reflectance spectra. **European Journal of Soil Science**, 62(3), 442-453, 2011.

VISCARRA ROSSEL, R. A.; WEBSTER, R. Discrimination of Australian soil horizons and classes from their visible-near infrared spectra. **European Journal of Soil Science**, 62(4), 637-647, 2011.

VISCARRA ROSSEL, R. A.; CHEN, C. Digitally mapping the information content of visible-near infrared spectra of surficial Australian soils. **Remote Sensing of Environment**, 115(6), 1443-1455, 2011.

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