

UNESP – UNIVERSIDADE ESTADUAL PAULISTA
CÂMPUS DE ARARAQUARA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

MAURÍLIO GUSTAVO NESPECA

Aplicação de espectroscopia no infravermelho próximo e análise multivariada para
identificação e quantificação de hidrocarbonetos totais do petróleo em solo.

Araraquara
2018

MAURÍLIO GUSTAVO NESPECA

Aplicação de espectroscopia no infravermelho próximo e análise multivariada para identificação e quantificação de hidrocarbonetos totais do petróleo em solo.

Tese apresentada ao Instituto de Química,
Universidade Estadual Paulista, como parte
dos requisitos para obtenção do título de
Doutor em Química

Orientador: Prof. Dr. Rodrigo Sequinel
Coorientador: Prof. Dr. Danilo Luiz Flumignan

Araraquara
2018

FICHA CATALOGRÁFICA

N457a Nespeca, Maurilio Gustavo
Aplicação de espectroscopia no infravermelho próximo e análise multivariada para identificação e quantificação de hidrocarbonetos totais do petróleo em solo / Maurilio Gustavo Nespeca. – Araraquara : [s.n.], 2018
216 f. : il.

Tese (doutorado) – Universidade Estadual Paulista, Instituto de Química
Orientador: Rodrigo Sequinel
Coorientador: Danilo Luiz Flumignan

1. Hidrocarbonetos. 2. Monitoramento ambiental.
3. Espectroscopia de infravermelho. 4. Análise multivariada.
5. Cromatografia a gás. I. Título.


CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: "Aplicação de espectroscopia no infravermelho próximo e análise multivariada para identificação e quantificação de hidrocarbonetos totais do petróleo em solo"

AUTOR: MAURILIO GUSTAVO NESPECA


ORIENTADOR: RODRIGO SEQUINEL

Aprovado como parte das exigências para obtenção do Título de Doutor em QUÍMICA, pela Comissão Examinadora:



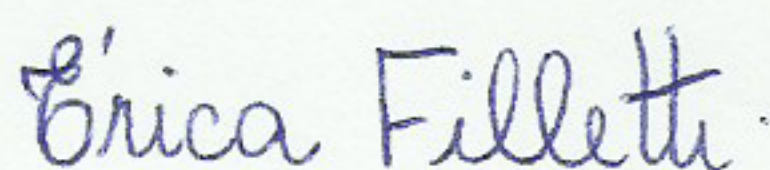
Prof. Dr. RODRIGO SEQUINEL

Departamento de Engenharias e Exatas / Universidade Federal do Paraná - UFPR - Palotina



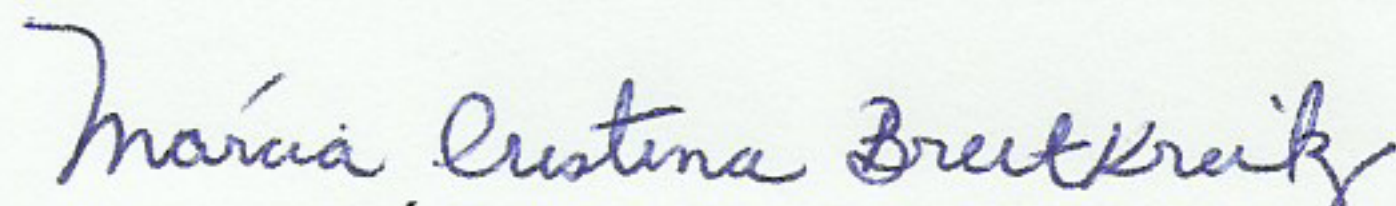
Prof.ª Dr.ª EDILENE CRISTINA FERREIRA

Departamento de Química Analítica / Instituto de Química - UNESP - Araraquara



Prof.ª Dr.ª ERICA REGINA FILLETTI NASCIMENTO

Departamento de Físico-Química / Instituto de Química - UNESP - Araraquara



Prof.ª Dr.ª MÁRCIA CRISTINA BREITKREITZ

Departamento de Química Analítica / Instituto de Química - UNICAMP - Campinas



Prof. Dr. HERON DOMINGUEZ TORRES DA SILVA

Departamento de Ciências Exatas e da Terra / Universidade Federal de São Paulo - UNIFESP - Diadema

Araraquara, 24 de agosto de 2018

CURRICULUM LATTES

1. Identificação

Maurílio Gustavo Nespeca

Nome em citações bibliográficas: NESPECA, M. G.

2. Formação acadêmica/titulação

2014 – 2018. Doutorado em Química

Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, Instituto de Química.

Orientador: Prof. Dr. Rodrigo Sequinel

Coorientador: Prof. Dr. Danilo Luiz Flumignan

Bolsista do: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

2011 – 2013. Mestrado em Química

Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, Instituto de Química.

Orientador: Prof. Dr. José Eduardo de Oliveira.

Bolsista do: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

2005 – 2009. Graduação em Química Licenciatura.

Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, Campus Bauru.

Orientador: Prof. Dr. Manoel Lima de Menezes.

3. Formação Complementar

2017 - 2017

Curso de Matlab - Aplicações em Bioprocessos. (Carga horária: 20h).

Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP.

2017 - 2017

Independent Component Analysis (ICA) Applications. (Carga horária: 6h).

Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP.

2017 - 2017

Seleção de variáveis. (Carga horária: 6h).

Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP.

2017 - 2017

MATLAB Onramp. Curso online. (Carga horária: 2h).

The MathWorks, Inc.

2017 - 2017

Quimiometria aplicada a espectroscopias que obedecem a lei de Beer-Lambert.

(Carga horária: 40h). Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP.

2017 - 2017

A microbiologia e a bioquímica da cerveja: aspectos práticos de fabricação.

(Carga horária: 8h). Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP.

2013 – 2013

Treinamento teórico e prático do equipamento IQT-LM.

(Carga horária: 16h). ICR3 Científica Comércio e Representações Ltda, ICR3, Brasil.

4. Atuação Profissional

2018. Professor efetivo de Química na rede de ensino do Estado de São Paulo.

(Carga horária: 10 h semanais)

5. Projetos de pesquisa

2014 – Atual: Aplicação de espectroscopia no infravermelho e análise multivariada para identificação e quantificação de hidrocarbonetos totais de petróleo em solo e água.

Descrição: Os métodos existentes para determinação de TPH apresentam diversos obstáculos: demandam muito tempo e volume de solvente na preparação das amostras, as análises cromatográficas são morosas e a quantificação por espectroscópica no infravermelho médio não é seletiva. Visto o grande número de contaminações referentes a postos de combustíveis e os riscos ambientais gerados por estes, é de extrema importância o desenvolvimento de metodologias que possam identificar e quantificar os contaminantes de forma rápida, com baixo custo e que atenda os limites de quantificação estabelecidos na Decisão de Diretoria nº 195-2005-E, de 23 de novembro de 2005. Sabendo-se que a espectroscopia no infravermelho associada à quimiometria torna-se uma poderosa ferramenta analítica, este trabalho propõe o desenvolvimento de uma metodologia analítica para quantificação e identificação de hidrocarbonetos em água e solo sem nenhum preparo de amostra.

2013 – 2013: Determinação de TPH e BTEX em solo e água subterrânea de áreas contaminadas por postos de combustíveis.

Descrição: Este projeto tem como objetivo determinar a presença de hidrocarbonetos em solo e água subterrânea provenientes da contaminação por postos de combustíveis. Serão implantados métodos de quantificação de TPH (hidrocarbonetos totais do petróleo na faixa da gasolina, diesel e óleo lubrificante) e BTEX (benzeno, tolueno, etilbenzeno e xileno) por cromatografia gasosa, baseados no método EPA 8015C.

Financiador: Fundação para o Desenvolvimento da UNESP - Bolsa.

2011 – 2013: Desenvolvimento de modelos quimiométricos associados à espectroscopia no infravermelho para determinação de parâmetros físico-químicos do óleo diesel.

Descrição: Projeto de mestrado. Tendo em vista que o óleo diesel é o derivado energético do petróleo mais consumido no Brasil, torna-se imprescindível o desenvolvimento de métodos de monitoramento capazes de avaliar a qualidade deste combustível de forma rápida, econômica e não destrutiva. Deste modo, este trabalho teve como objetivo o desenvolvimento de modelos quimiométricos com capacidade de estimar os parâmetros físico-químicos do óleo diesel (mistura B5) através de análises espectroscópicas nas regiões do infravermelho médio e próximo. Os valores de referência foram obtidos por meio de ensaios realizados no Cempeq (Centro de Monitoramento e Pesquisa da Qualidade de Combustíveis, Biocombustíveis, Petróleo e Derivados) utilizando-se métodos ABNT e ASTM, estabelecidos na Resolução ANP Nº 65 de 2011.

2009 – 2009: Avaliação da hidrólise ácida e alcalina da semente de abacate para produção de etanol.

Descrição: Realizou-se o estudo da hidrólise ácida e alcalina de sementes de abacate para produção de etanol como trabalho de Monografia de Conclusão de Curso. Avaliou-se a eficiência da hidrólise ácida/alcalina em um pré-tratamento ao processo de hidrólise enzimática.

2008 – 2009: Estudo teórico da solvatação aquosa e especiação de actínídeos do ciclo do tório.

Descrição: Projeto de iniciação científica. Neste trabalho, estudamos a primeira camada de hidratação do íon Th(IV) com a presença de carbonatos através de uma abordagem químico-quântica aplicando o método de dinâmica molecular mesclado com métodos ab initio. Financiador: Fundação de Amparo à Pesquisa do Estado de São Paulo - Bolsa.

6. Projetos de extensão

2006 – 2007: Cursinho Pré-Vestibular Ferradura Mirim.

Atuação como professor de Química e Matemática. (Carga horária: 4h semanais)

Coordenador: Prof. Dr. Antonio Francisco Marques

7. Revisor de periódico

2017. FUEL

2018. Chemical Science International Journal

2018. Journal of Environmental and Occupational Science

8. Idiomas

Inglês

Compreende Bem, Fala Razoavelmente, Lê Bem, Escreve Razoavelmente.

Espanhol

Compreende Razoavelmente, Fala Pouco, Lê Bem, Escreve Pouco.

9. Produção Bibliográfica

Artigos completos

1. NESPECA, MAURÍLIO GUSTAVO; MUNHOZ, JOÃO FERNANDO VILLARRUBIA LOPES ; FLUMIGNAN, DANILO LUIZ ; DE OLIVEIRA, JOSÉ EDUARDO . Rapid and sensitive method for detecting adulterants in gasoline using ultra-fast gas chromatography and Partial Least Square Discriminant Analysis. FUEL, v. 215, p. 204-211, 2018.

2. NESPECA, M. G.; HATANAKA, R. R. ; FLUMIGNAN, DANILO LUIZ ; DE OLIVEIRA, JOSÉ EDUARDO . Rapid and Simultaneous Prediction of Eight Diesel Quality Parameters through ATR-FTIR Analysis. Journal of Analytical Methods in Chemistry, v. 2018, p. 1-10, 2018.

3. NESPECA, M. G.; Piassalonga, G. B. ; OLIVEIRA, J. E. . Infrared spectroscopy and multivariate methods as a tool for identification and quantification of fuels and lubricant oils in soil. ENVIRONMENTAL MONITORING AND ASSESSMENT, v. 190, p. 1-12, 2018.
4. RODRIGUES, CAROLINE VARELLA ; NESPECA, MAURÍLIO GUSTAVO ; SAKAMOTO, ISABEL KIMIKO ; DE OLIVEIRA, JOSÉ EDUARDO ; AMÂNCIO VARESCHE, MARIA BERNADETE ; MAINTINGUER, SANDRA IMACULADA . Bioconversion of crude glycerol from waste cooking oils into hydrogen by sub-tropical mixed and pure cultures. INTERNATIONAL JOURNAL OF HYDROGEN ENERGY, 2018.
5. NESPECA, MAURÍLIO GUSTAVO; RODRIGUES, CAROLINE VARELLA ; SANTANA, KAMILI OLIVEIRA ; MAINTINGUER, SANDRA IMACULADA ; DE OLIVEIRA, JOSÉ EDUARDO . Determination of alcohols and volatile organic acids in anaerobic bioreactors for H₂ production by near infrared spectroscopy. INTERNATIONAL JOURNAL OF HYDROGEN ENERGY, p. 20480-20493, 2017.
6. TORQUATO, LILIAN D.M. ; PACHIEGA, RENAN ; CRESPI, MARISA S. ; NESPECA, MAURÍLIO GUSTAVO ; DE OLIVEIRA, JOSÉ EDUARDO ; MAINTINGUER, SANDRA I. . Potential of biohydrogen production from effluents of citrus processing industry using anaerobic bacteria from sewage sludge. Waste Management (Elmsford), v. 59, p. 181-193, 2017.
7. RODRIGUES, C. V. ; SANTANA, K. O. ; NESPECA, M. G. ; OLIVEIRA, J. E. ; Maintinguer, S. I. Aplicação de glicerol bruto pré-tratado em processos biológicos de geração de H₂. REVISTA CIÊNCIA & TECNOLOGIA: FATEC-JB, v. 8, p. 2016, 2016.
8. SANTANA, K. O. ; RODRIGUES, C. V. ; NESPECA, M. G. ; OLIVEIRA, J. E. ; Maintinguer, S. I. Produção biológica de gás hidrogênio a partir de glicerol bruto e esgoto sanitário. REVISTA CIÊNCIA & TECNOLOGIA: FATEC-JB, v. 8, p. 1, 2016.
9. RODRIGUES, CAROLINE VARELLA ; SANTANA, KAMILI OLIVEIRA ; NESPECA, MAURÍLIO GUSTAVO ; EDUARDO DE OLIVEIRA, JOSÉ ; MAINTINGUER, SANDRA IMACULADA . Crude glycerol by transesterification process from used cooking oils: Characterization and potentialities on hydrogen bioproduction. International Journal of Hydrogen Energy, v. 41, p. 14641-14651, 2016.

Resumos expandidos publicados em anais de congressos

1. NESPECA, M. G.; RODRIGUES, CAROLINE VARELLA ; SANTANA, KAMILI OLIVEIRA ; DE OLIVEIRA, JOSÉ EDUARDO . Determinação de metanol em biorreatores anaeróbios por NIRS-SVM. In: VIII Simpósio de Microbiologia Aplicada, 2017, Rio Claro. Anais do VIII Simpósio de Microbiologia Aplicada. Araras: Revista Ciência, Tecnologia & Ambiente, 2017. v. 6. p. 319-324.

Apresentações de Trabalho

1. Rodrigues, C. V. ; Alcaraz F.A.R. ; Santana, K. O. ; NESPECA, M. G. ; Maintinguer, S. I. . Biosistema integrado aplicado na produção de bio-H₂ a partir de resíduos agroindustriais.. 2017. (Apresentação de Trabalho/Simpósio).
2. Santana, K. O. ; Rodrigues, C. V. ; NESPECA, M. G. ; Pires, L. O. ; Maintinguer, S. I. . Esgoto sanitário aplicado na codigestão de resíduo da indústria de biodiesel para produção de hidrogênio e ácidos orgânicos.. 2017. (Apresentação de Trabalho/Simpósio).
3. NESPECA, M. G.; Rodrigues, C. V. ; Santana, K. O. ; OLIVEIRA, J. E. . Determinação de metanol em biorreatores anaeróbios por NIRS-SVM. 2017. (Apresentação de Trabalho/Simpósio).
4. NESPECA, M. G.; SANTOS, L. K. ; OLIVEIRA, J. E. . Optimization of BTEX extraction from soil using sonication and ethanol as solvent. 2017. (Apresentação de Trabalho/Outra).
5. NESPECA, M. G.; MUNHOZ, J. F. V. L. ; HATANAKA, R. R. ; OLIVEIRA, J. E. . Detection of gasoline adulteration using FTIR and GA-PLS-DA. 2017. (Apresentação Oral).
6. NESPECA, M. G.; HATANAKA, R. R. ; OLIVEIRA, J. E. . Avaliação do uso de seleção de variáveis na determinação de metanol em gasolina por infravermelho próximo e método de regressão PLS. 2016. (Apresentação de Trabalho/Outra).
7. Santana, K. O. ; Rodrigues, C. V. ; NESPECA, M. G. ; OLIVEIRA, J. E. ; Pires, L. O. ; Maintinguer, S. I. . Produção biológica de gás hidrogênio a partir de glicerol bruto e esgoto sanitário. 2016. (Apresentação de Trabalho/Congresso).
8. SANTANA, KAMILI OLIVEIRA ; RODRIGUES, CAROLINE VARELLA ; NESPECA, MAURÍLIO GUSTAVO ; OLIVEIRA, J. E. ; Pires, L. O. ; MAINTINGUER, SANDRA IMACULADA . Biological hydrogen production from raw glycerol and wastewater. 2016. (Apresentação de Trabalho/Congresso).
9. RODRIGUES, CAROLINE VARELLA ; SANTANA, KAMILI OLIVEIRA ; NESPECA, M. G. ; OLIVEIRA, J. E. ; SAKAMOTO, I. K. ; MAINTINGUER, SANDRA IMACULADA. Potencial biohydrogen production with crude glycerol by mixed cultures and Enterobacter sp. 2016. (Apresentação de Trabalho/Congresso).
10. Rodrigues, C. V. ; Santana, K. O. ; NESPECA, M. G. ; OLIVEIRA, J. E. ; Maintinguer, S. I. . Aplicação de glicerol bruto pré-tratado em processos biológicos de geração de H₂. 2016. (Apresentação de Trabalho/Congresso).
11. NESPECA, M. G.; OLIVEIRA, J. E. . Predição de Parâmetros Físico-Químicos de Óleo Diesel Através de Modelos Quimiométricos e Espectroscopia no Infravermelho Próximo. 2013. (Apresentação de Trabalho/Congresso).

10. Participação em bancas de comissões julgadoras

1. Comissão avaliadora dos painéis da III Escola de Inverno em Quimiometria (EIQ 2017). 2017. Universidade Estadual Paulista Júlio de Mesquita Filho.

11. Participação em eventos/congressos/simpósios/oficinas/workshops

1. III Escola de Inverno de Quimiometria. 2017.
2. VIII Simpósio de Microbiologia Aplicada. 2017.
3. Entrepreneurship Workshop for Scientists and Engineers.Smart Energy. 2016.
4. Estímulo a Pesquisa Competitiva na FMB: Busca de Apoio Financeiro Adicional às Agências de Fomento. 2016.
5. VII Workshop de Quimiometria.Avaliação do uso de seleção de variáveis na determinação de metanol em gasolina por infravermelho próximo e método de regressão PLS. 2016.
6. Comitê Brasileiro de Tintas - PROJETO ABNT NBR 16388. 2015
7. Seminário Técnico de Cromatografia HPLC/GC/preparação de amostras. 2015.
8. Leitura e produção de textos técnicos e acadêmicos. 2014.
9. Encontro Nacional de Química Analítica. 2013.
10. 8º Fórum de Debates sobre qualidade e uso de combustíveis - os novos combustíveis do mercado brasileiro e suas implicações. 2012.
11. 41ª Semana da Química, no Instituto de Química da Unesp. 2011.
12. V Semana da Química do Centro-Oeste Paulista. 2008.
13. II Ciclo de Histórias de Professores: A experiência vivida em sala de aula. 2007.
14. IV Semana da Química da Unesp. 2007.
15. V Semana da Química do Centro-Oeste Paulista. 2007.
16. IV Semana da Química do Centro-Oeste Paulista. 2006.
17. II Semana de Química Ambiental. 2005.

Dedico este trabalho à minha família, amigos e à minha amada esposa Mariana.

AGRADECIMENTOS

À minha família, pela educação, criação, incentivo e convivência.

À minha amada esposa pelo apoio, paciência, companheirismo e amor, que facilitam a superação das dificuldades e direcionam melhor a vida.

Aos grandes amigos com quem tive oportunidade de conviver durante estes ano em Araraquara, e aos amigos avareenses que sempre me apoiaram.

A todos os colegas do Cempeqc.

À CAPES e Fundunesp.

Ao meu orientador e coorientador pelos ensinamentos na vida científica.

À Prof^a. Dr^a. Sandra pelas oportunidades proporcionadas em vários projetos.

Ao Prof. Dr. Zelão pela orientação inicial neste projeto.

À doutoranda e amiga Cássia Bonifácio, ao Prof. Dr. Giovanni Penni e ao Eng. Filipe Bandeca pela disponibilização das amostras de solo que permitiram o desenvolvimento deste trabalho.

E, finalmente, às bandas e aos projetos musicais, nos quais posso descarregar as tensões com prazer.

“O amor pela ciência deve ser um dinamismo psíquico autógeno. No estado de pureza alcançado por uma psicanálise do conhecimento objetivo, a ciência é a estética da inteligência.”

Gaston Bachelard ¹

RESUMO

Segundo dados, de 2017, da Companhia Ambiental do Estado de São Paulo (CETESB), os postos de combustíveis são responsáveis pela contaminação ambiental de 72% das 5942 áreas contaminadas cadastradas no estado de São Paulo. A contaminação de solos e águas subterrâneas por combustíveis fósseis causam imensos danos ambientais devido à alta toxicidade dos constituintes destes combustíveis, além de apresentarem propriedades carcinogênicas e elevada permanência no solo. O monitoramento ambiental em áreas de risco potencial de contaminação, como postos de combustíveis, é realizado através da análise de hidrocarbonetos totais do petróleo (TPH), entre outros hidrocarbonetos individuais. Estas análises são realizadas através de métodos cromatográficos em fase gasosa que requerem etapas de extração com solventes halogenados, purificação e, muitas vezes, pré-concentração. O elevado custo e o tempo demandado para a quantificação de TPH em solo tornam-se barreiras para o crescimento do monitoramento e acompanhamento de processos de remediação das áreas contaminadas. Desta forma, este trabalho objetivou o desenvolvimento de métodos analíticos mais rápidos e de menor custo para a identificação e quantificação de TPH em solo através da espectroscopia na região do infravermelho próximo (NIR). Três métodos NIR foram desenvolvidos: (i) sem preparo de amostra; (ii) após extração por hexano; e (iii) após extração por etanol. Os modelos de classificação foram desenvolvidos pelo método de análise discriminante por mínimos quadrados parciais (PLS-DA) e os modelos de calibração por mínimos quadrados parciais (PLS). No desenvolvimento dos modelos, nove diferentes pré-processamentos e seleção de variáveis pelo algoritmo genético (GA) foram avaliados. Os modelos foram desenvolvidos usando-se amostras de solo fortificado com os contaminantes (gasolina, diesel e óleo lubrificantes) e validados com amostras de solo contaminado, adquiridas no monitoramento ambiental de postos de combustíveis. O modelo PLS-DA proporcionou 100% de classificações corretas das amostras contaminadas sem a necessidade de preparo de amostra, enquanto a predição da concentração dos analitos tornou-se possível através dos modelos PLS após a etapa de extração com etanol. Como objetivos secundários deste trabalho, foram desenvolvidos métodos de quantificação das diferentes frações de TPH e dos compostos benzeno, tolueno, etilbenzeno e xilenos (BTEX) por cromatografia em fase gasosa ultrarrápida com detector de ionização por chama (UFGC-FID). Além do desenvolvimento dos métodos UFGC-FID, otimizou-se o processo de extração dos analitos por sonicação através de planejamentos experimentais, avaliando-se diferentes solventes, tempo de sonicação, agitação por vórtex e volume de solvente. Os métodos UFGC-FID proporcionaram análises de 3 a 17 vezes mais rápidas que o método cromatográfico de acordo com a norma EPA 8015. Ao final deste trabalho, os métodos desenvolvidos e o método EPA 8015 foram comparados segundo aspectos analíticos, ambientais e financeiros. De forma geral, os três métodos apresentaram mesma exatidão; o método cromatográfico (EPA 8015) foi o método mais preciso; o método UFGC-FID foi o de menor investimento inicial e de menor tempo para retorno financeiro; e, finalmente, o método NIR após extração com etanol foi o mais sensível, rápido, favorável à química verde e de menor custo por análise.

Palavras-chave: hidrocarbonetos totais do petróleo; contaminação de solo; espectroscopia no infravermelho próximo; modelos de classificação; modelos de calibração; cromatografia em fase gasosa ultrarrápida;

ABSTRACT

According to the Environmental Company of the State of São Paulo (CETESB), the gas stations are responsible for the environmental contamination of 72% of the 5942 contaminated areas registered in the state of São Paulo. Contamination of soils and groundwater by fossil fuels causes immense environmental damages due to their high toxicity, carcinogenic properties and high permanence in the soil. Environmental monitoring in areas of potential contamination risk, such as gas stations, is carried out through the analysis of total petroleum hydrocarbons (TPH), among other individual compounds. These analyzes are performed by gas chromatographic methods which require some sample preparation steps, such as extraction with halogenated solvents, purification, and often preconcentration. The high cost and time demanded for the quantification of TPH in the soil become barriers for the growth of the monitoring program and remediation processes of the contaminated areas. Therefore, this work aimed at the development of faster and lower cost analytical methods for the identification and quantification of TPH in soil through near-infrared spectroscopy (NIR). Three NIR methods were developed: (i) without sample preparation; (ii) after hexane extraction; and (iii) after extraction with ethanol. The classification models were developed by partial least squares discriminant analysis (PLS-DA) method and the calibration models by partial least squares (PLS) method. In the development of the models, we evaluated nine different preprocessing and selection of variables by the genetic algorithm (GA). The models were developed using soil samples fortified with contaminants (gasoline, diesel and lubricant oil) and validated with samples of contaminated soil acquired in the environmental monitoring of gas stations. The PLS-DA model provided 100% correct classifications without sample preparation, while the prediction of the concentration of the analytes was possible by PLS models after the ethanol extraction. As secondary objectives of this work, we developed quantification methods for the different fractions of TPH and benzene, toluene, ethylbenzene, and xylenes (BTEX) compounds by ultrafast gas chromatography with flame ionization detector (UFGC-FID). In addition to the UFGC-FID methods, the TPH extraction by sonication was optimized through experimental design, evaluating different solvents, sonication time, agitation and solvent volume. The UFGC-FID methods provided analyzes 3 to 17 times faster than the chromatographic method according to EPA 8015. At the end of this work, the developed methods and the EPA 8015 method were compared according to analytical, environmental and financial aspects. In general, the three methods presented the same accuracy; the EPA 8015 method was the most precise; the UFGC-FID method presented lower initial investment and less time for financial return; and the NIR method after ethanol extraction was the most sensitive, fast, favorable to green chemistry and presented the lowest cost per analysis.

Keywords: total petroleum hydrocarbons; soil contamination; near-infrared spectroscopy; classification models; calibration models; ultrafast gas chromatography.

LISTA DE FIGURAS

Figura 1 – Faixa de hidrocarbonetos abrangida por diferentes métodos analíticos.....	37
Figura 2 – Organograma das principais etapas desenvolvidas no presente trabalho.....	68
Figura 3 – Gráfico de loadings (Δ) e scores (\diamond) resultante da PCA das propriedades dos solos.	93
Figura 4 – Espectros NIR das amostras de solo obtidos por refletância difusa.	94
Figura 5 – Gráfico de escores gerado a partir da PCA dos espectros médios dos solos pré-processados por SNV.....	96
Figura 6 – Relação entre as áreas dos ésteres presentes nas amostras de diesel.	101
Figura 7 – Espectros NIR das 30 amostras dos contaminantes (gasolinas, óleos diesel e óleos lubrificantes).....	103
Figura 8 – Gráfico de escores gerado a partir da PCA dos espectros dos contaminantes. ...	105
Figura 9 – Avaliação do tempo de incubação na extração de compostos GRO por headspace.....	106
Figura 10 – Gráfico de resíduos da curva analítica para GRO pelo método EPA 5021/8015.....	108
Figura 11 – Curva analítica da fração ORO em hexano abrangendo todo intervalo de mediação.....	115
Figura 12 – Gráficos de efeitos principais gerados pela ANOVA para cada DOE.....	120
Figura 13 – Espectro médio do conjunto de amostras de solo fortificado. As regiões inseridas nos retângulos foram excluídas previamente à análise dos pré-processamentos.	133
Figura 14 – Tendência significativa na predição de fluoreno em etanol em amostras externas.....	154
Figura 15 – Relação entre desvio padrão relativo e concentração de GRO para os três conjuntos amostrais.....	156
Figura 16 – Cartas de controle das medidas de GRO em solo fortificado (dois níveis de concentração).....	158
Figura 17 – Cartas de controle das medidas de criseno em solo e o-xileno em etanol.....	158
Figura 18 – Relação entre desvio padrão relativo (20 replicatas) e a concentração de todos os analitos nos três conjuntos de dados.	159
Figura 19 – Relação entre valores de recuperação e concentração de TPH para o método NIR com extração por etanol.	164

Figura 20 – Espectro das soluções dos contaminantes em etanol e das amostras de validação extraídas por etanol.....	164
Figura 21 – Espectros dos solventes n-hexano utilizados no desenvolvimento dos modelos e na extração das amostras de validação.....	165
Figura 22 – Relação entre o faturamento obtido de acordo com cada método analítico e o número de amostra ao longo dez anos.	177

LISTA DE TABELAS

Tabela 1 – Valores de prevenção e intervenção de contaminantes em solo.....	33
Tabela 2 – Estudos na literatura relacionados à determinação de TPH de em solo envolvendo a espectroscopia NIR.	55
Tabela 3 – Condições cromatográficas para caracterização das amostras de combustíveis. ...	71
Tabela 4 – Condições cromatográficas para determinação da fração ORO dos óleos lubrificantes por cromatografia gasosa ultrarrápida.	73
Tabela 5 – Condições cromatográficas (EPA 8015C) para determinação das frações GRO e DRO.	74
Tabela 6 – Condições operacionais do headspace.	75
Tabela 7 – Condições cromatográficas iniciais para avaliação da separação dos compostos GRO pelas colunas PH SE54 e PH RTX-1.	76
Tabela 8 – Condições cromatográficas iniciais para cada coluna na avaliação da separação dos compostos do TPH Mix.	77
Tabela 9 – Condições definidas para determinação das frações GRO, DRO e ORO por.....	77
Tabela 10 – Níveis dos fatores do planejamento experimental de cada grupo de compostos.	78
Tabela 11 – Lista de materiais, equipamentos e softwares utilizados no desenvolvimento do trabalho.....	90
Tabela 12 – Resultados das análises química, físicas e elementares das amostras de solo. ...	92
Tabela 13 – Atribuições das bandas espectrais das amostras de solo.	95
Tabela 14 – Parâmetros de separação do GRO Mix segundo o método ASTM D5759.....	97
Tabela 15 – Parâmetros das curvas analíticas e perfil composicional das amostras de gasolina.....	98
Tabela 16 – Parâmetros cromatográficos da análise dos grupos TPH, PAH e ésteres metílicos	99
Tabela 17 – Parâmetros das curvas analíticas e perfil composicional das amostras de diesel.	100
Tabela 18 – Concentração da faixa de compostos ORO das amostras de óleo lubrificante.	102
Tabela 19 – Atribuições das bandas espectrais das amostras de contaminantes.	104
Tabela 20 – Parâmetros cromatográficos da separação do GRO Mix pelo método EPA 8015.	107

Tabela 21 – Parâmetros das curvas analíticas dos métodos de referência para GRO e DRO.	107
Tabela 22 – Parâmetros cromatográficos da análise do TPH Mix pelo método EPA 8015.	108
Tabela 23 – Parâmetros cromatográficos das análises do GRO Mix e TPH Mix pelos métodos UFGC.	110
Tabela 24 – Parâmetros das curvas analíticas para determinação de GRO em hexadecano.	111
Tabela 25 – Parâmetros das curvas analíticas para determinação de GRO em etanol.	112
Tabela 26 - Parâmetros das curvas analíticas desenvolvidas em diclorometano e hexano para quantificação das frações DRO e ORO.	114
Tabela 27 – Parâmetros analíticos das curvas para determinação das faixas DRO e ORO em hexano.	116
Tabela 28 – Respostas do planejamento fatorial completo para análise de BTEX e GRO.	117
Tabela 29 – Respostas do planejamento fatorial completo para análise de DRO.	118
Tabela 30 – Valores de p obtidos na ANOVA de cada planejamento experimental.	119
Tabela 31 – Condições determinadas através da ANOVA para melhores recuperações.	121
Tabela 32 – Repetibilidade da extração dos compostos BTEX e fatores de correção.	121
Tabela 33 – Repetibilidade da extração das faixas DRO e ORO.	122
Tabela 34 – Avaliação da exatidão dos métodos determinação das frações TPH e BTEX por extração ultrassônica e subsequente análise UFGC.	123
Tabela 35 – Determinação de BTEX e das frações TPH das “amostras reais” por UFGC.	124
Tabela 36 – Avaliação inicial do uso de diferentes pré-processamentos para os modelos PLS-DA.	127
Tabela 37 – Avaliação de pré-processamentos na análise do conjunto de validação (II) através dos modelos PLS-DA.	128
Tabela 38 – Resultados da predição dos conjuntos de validação (I) pelos modelos PLS-DA com e sem seleção de variáveis.	129
Tabela 39 – Resultados em porcentagem da predição das amostras reais pelos modelos PLS-DA.	131
Tabela 40 – Resultados da classificação do conjunto de validação (II) pelos modelos.	131

Tabela 41 – Resultados em porcentagem da predição das amostras reais pelos modelos KNN e SVM.	132
Tabela 42 – Avaliação dos pré-processamentos para os conjuntos de amostras de solos fortificados e soluções de hexano e etanol com os contaminantes.	136
Tabela 43 – Parâmetros analíticos dos modelos PLS das frações TPH desenvolvidos com os espectros completos e com variáveis selecionadas. Os valores de RMSE estão na unidade mg kg^{-1}	140
Tabela 44 – Avaliação da redução dos valores de RMSEP pelo uso de GA através de teste F.	142
Tabela 45 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de amostras de solo fortificado.	144
Tabela 46 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de soluções dos contaminantes em hexano.	147
Tabela 47 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de soluções dos contaminantes em etanol.	150
Tabela 48 – Porcentagens de FP e FN gerados pelos modelos de calibração na análise do conjunto de replicatas.	160
Tabela 49 – Resultados de validação dos métodos de determinação de TPH com e sem preparo de amostra.	163
Tabela 50 – Médias dos erros relativos resultantes da validação dos métodos desenvolvidos.	166
Tabela 51 – Valores de referência e preditos pelos modelos das frações de TPH para as amostras de solo coletado para monitoramento ambiental.	167
Tabela 52 – Parâmetros analíticos dos métodos convencionais e desenvolvidos para determinação frações de TPH.	170
Tabela 53 – Frequência analítica e volume de solvente necessários para a análise de 100 amostras em cada método proposto e em duas situações hipotéticas (i) 100% amostras contaminadas; e (ii) 22% amostras contaminadas.	173
Tabela 54 – Descrição dos custos para análise das frações de TPH pelo método cromatográfico (EPA), método cromatográfico ultrarrápido (UFGC) e método espectroscópico NIR.	174
Tabela 55 – Avaliação do tempo necessário para recuperação do investimento inicial e o custo anual para o contratante de acordo com cada método analítico.	176
Tabela 56 – Pontuação (1 a 5) dos fatores considerados na comparação dos métodos analíticos.	177

LISTA DE ABREVIATURAS E SIGLAS

ANOVA	análise de variância
API	American Petroleum Institute
ASTM	American Society for Testing and Materials
ATR	refletância total atenuada
BTEX	benzeno, tolueno, etilbenzeno e xilenos
CETESB	Companhia Ambiental do Estado De São Paulo
DPR	desvio padrão relativo
DRIFT	refletância difusa na região do infravermelho com transformada de Fourier
DRO	compostos da faixa do diesel
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
EMM	modelo de mistura estendida
EPA	Environmental Protection Agency
EPO	ortogonalização de parâmetros externos
FN	falso negativo
FP	falso positivo
FT	transformada de Fourier
GA	algoritmo genético
GC-FID	cromatografia gasosa com detector de ionização de chama
GC-MS	cromatografia gasosa com detector de espectrometria de massa
GC-PID	cromatografia gasosa com detector de fotoionização
GLSW	ponderação por mínimos quadrados generalizada
GRO	compostos da faixa da gasolina
LD	limite de detecção
LQ	limite de quantificação
MIR	infravermelho médio
MRC	material de referência certificado
MSC	correção multiplicativa de espalhamento
NAS	sinal analítico líquido
NIR	infravermelho próximo
ORO	compostos da faixa dos óleos lubrificantes
OSC	correção ortogonal de sinal
PAH	hidrocarbonetos policíclicos aromáticos
PLS	mínimos quadrados parciais
RBCA	ações corretivas baseada no risco
RMSEP	raiz da média quadrática dos erros
RPD	desvio percentual relativo
SCF	fluido supercrítico
SNV	padronização normal de sinal
SPME	microextração em fase sólida
SWIR	região de alta frequência do infravermelho médio
TECP	penetrômetro de cone de extração térmica
TPH	hidrocarbonetos totais do petróleo
UFGC	cromatografia gasosa ultrarrápida
Vis	região visível do espectro eletromagnético
VL	variável latente

SUMÁRIO

1.	INTRODUÇÃO.....	22
2.	OBJETIVOS.....	24
3.	REVISÃO BIBLIOGRÁFICA.....	25
3.1.	Solo.....	25
3.1.1.	Poluição do solo	26
3.1.2.	Contaminação de solo por hidrocarbonetos.....	27
3.1.3.	Processo de investigação de contaminação de solo.....	30
3.1.4.	Remediação de áreas contaminadas.....	34
3.2.	Hidrocarbonetos totais do petróleo – TPH.....	35
3.2.1.	Métodos analíticos para determinação de TPH	38
3.2.2.	Métodos de extração	39
3.2.3.	Métodos de quantificação	44
3.3.	Quimiometria.....	59
3.3.1	Importância da Quimiometria na aplicação de NIR	59
3.3.2.	Métodos de regressão multivariada	60
3.3.3.	Pré-processamentos	62
3.3.4.	Seleção de variáveis.....	64
3.3.5.	Reconhecimento de padrões por PLS-DA	65
3.3.6.	Validação de métodos multivariados.....	67
4.	METODOLOGIA.....	68
4.1.	Conjunto amostral.....	68
4.1.1.	Amostras de solo	69
4.1.2.	Amostras de combustíveis e óleos lubrificantes.....	69
4.2.	Métodos de referência.....	73
4.3.	Análises espectroscópicas e quimiométricas	79
4.3.1.	Aquisição dos espectros NIR	79
4.3.2.	Conjuntos amostrais.....	80
4.3.3.	Modelos de classificação	81
4.3.4.	Modelos de calibração	82
4.4.	Determinação dos parâmetros analíticos	83
4.4.1.	Parâmetros analíticos cromatográficos	83

4.4.2.	Parâmetros analíticos espectroscópicos multivariados	86
4.5.	Comparação entre os métodos analíticos	90
4.6.	Lista de materiais, equipamentos e softwares	90
5.	RESULTADOS E DISCUSSÃO.....	92
5.1.	Caracterização das amostras de solo e contaminantes.....	92
5.1.1.	Análises físico-químicas e composicionais dos solos	92
5.1.2.	Caracterização das amostras de combustível e óleos lubrificantes	96
5.2.	Métodos de referência.....	106
5.2.1.	Métodos determinativos por cromatografia gasosa convencional	106
5.2.2.	Métodos determinativos por cromatografia gasosa ultrarrápida	109
5.2.3.	Avaliação do método de extração por ultrassom.....	117
5.2.4.	Análise das amostras de monitoramento ambiental	123
5.3.	Modelos quimiométricos.....	125
5.3.1.	Análises prévias.....	125
5.3.2.	Modelos de classificação	126
5.3.3.	Modelos de calibração	133
5.4.	Comparativo entre os métodos convencionais e desenvolvidos	169
5.4.1.	Parâmetros analíticos	169
5.4.2.	Avaliação de frequência analítica e consumo de solvente.....	170
5.4.3.	Comparativo de custos estimados	173
5.4.4.	Comparação geral entre os métodos analíticos	177
6.	CONCLUSÕES	179
	REFERÊNCIAS	181
	ANEXO A – Características dos combustíveis e óleos lubrificantes	196
	ANEXO B – Cromatogramas.....	198
	ANEXO C – Resultados quimiométricos	203

1. INTRODUÇÃO

Apesar do desenvolvimento de tecnologias para o uso de combustíveis renováveis, a sociedade atual ainda é altamente dependente de combustíveis derivados de petróleo.² Devido aos riscos desses combustíveis para o meio ambiente e a saúde, a Companhia Ambiental do Estado de São Paulo (CETESB) configura os postos de combustíveis como estabelecimentos de risco. Embora a gasolina e o diesel sejam constituídos principalmente por hidrocarbonetos saturados, eles representam um alto risco para a saúde humana, já que a concentração de hidrocarbonetos aromáticos, substâncias com propriedades cancerígenas, pode atingir até 50% do volume de gasolina e 40% do volume de diesel.³⁻⁵ De acordo com informações da CETESB, os postos de gasolina são os principais responsáveis pela contaminação ambiental no estado de São Paulo, representando aproximadamente 72% das 5942 áreas contaminadas registradas pelo monitoramento ambiental.⁶

Em vista do risco ambiental, o licenciamento de postos de combustíveis depende da análise ambiental para verificar a contaminação do combustível no solo e nas águas subterrâneas. A verificação é realizada por meio de análise dos hidrocarbonetos totais de petróleo (TPH), na qual a concentração não pode exceder 1000 mg kg⁻¹ no solo e 600 µg L⁻¹ em água subterrânea, segundo a regulamentação brasileira.⁷ Dada a alta complexidade da composição dos derivados de petróleo, os TPH são divididos em frações de acordo com o número de carbonos na cadeia. Hidrocarbonetos contendo 6 a 10 átomos de carbono são atribuídos à faixa da gasolina (GRO), 10 a 28 carbonos à faixa do diesel (DRO); e acima de 28 carbonos à faixa dos óleos lubrificantes (ORO).⁸

Os métodos comumente utilizados para determinação de TPH empregam a espectroscopia vibracional na região do infravermelho médio (EPA 418.1) e cromatografia gasosa com detector de ionização por chama (EPA 8015) ou de espectrometria de massas (EPA 8270). O método EPA 418.1 é baseado na quantificação de TPH pelo número de onda referente à ligação entre carbono e hidrogênio do grupo CH₃ (2950 cm⁻¹),⁹ portanto o método não é seletivo, podendo quantificar como TPH outras matérias orgânicas presentes na matriz. O método foi removido pela EPA (Environmental Protection Agency) devido aos riscos ao meio ambiente por usar Freon ou tetraclorometano como solvente de extração, entretanto, ainda foi usado muito tempo em alguns países.¹⁰ Posteriormente, a mesma agência adaptou o método EPA 418.1 para o método EPA 8440, no qual os hidrocarbonetos são extraídos por dióxido de carbono supercrítico em amostras sólidas, contudo é um método moroso e não seletivo.¹¹ Como alternativa, a ASTM (American Society for Testing and Materials) desenvolveu um método

similar para determinação de TPH em água (ASTM D7678) utilizando cicloexano como solvente, porém a seletividade continuou sendo um obstáculo.¹² A limitação destes métodos encontra-se na modelagem univariada, a qual baseia-se apenas na banda de estiramento de ligação CH₃ (métodos EPA) ou deformação angular de grupos CH₂ (método ASTM). Portanto, compostos de alta toxicidade, como benzeno, naftaleno, antraceno, criseno, entre outros cuja cadeia carbônica não apresenta grupos CH₃ e CH₂, são negligenciados por estes métodos espectroscópicos. Já os métodos cromatográficos para determinação de TPH e suas frações são seletivos e possuem limites de detecção na ordem microgramas por quilo. Todavia, o tempo de análise para cada grupo de hidrocarbonetos é cerca de 40 minutos, de acordo com os métodos EPA 8015 e 8270,^{8,13} e os procedimentos de extração são trabalhosos e requerem uso de diclorometano para obter extrações eficientes (métodos EPA 3510, 3540 e 3550).¹⁴⁻¹⁶

A espectroscopia na região do infravermelho próximo com transformada de Fourier (FT-NIR) destaca-se como uma das técnicas mais utilizadas atualmente na análise fingerprinting de misturas complexas. A técnica FT-NIR possui a maioria das características desejáveis em métodos analíticos, sendo: rápida (um minuto ou menos por amostra); não destrutiva; não invasiva; adaptável para uso on-line e at-line; sensível a ligações C-H; N-H; S-H e O-H, o que torna uma técnica bastante abrangente; aplicável a amostras gasosas, líquidas e sólidas; e requer mínima ou nenhuma preparação de amostra.^{17,18} Contudo, tal técnica produz uma enorme quantidade de informação, muitas vezes complexa, tornando necessário o uso de ferramentas mais sofisticadas para tratar os dados e extrair informações relevantes.¹⁹ Nesse contexto, a quimiometria é indispensável para a análise de espectros no infravermelho próximo.

No campo da quimiometria, os dados multivariados podem ser analisados com diferentes finalidades, entre elas análise exploratória, classificação de amostras e predição de propriedades.²⁰ As associações de técnicas espectroscópicas com ferramentas quimiométricas são relativamente simples e permitem análises rápidas, precisas, com exatidão comparáveis aos métodos convencionais e com uso de pequenas quantidades de amostra.²¹ Com o emprego da quimiometria, a espectroscopia na região do infravermelho próximo torna-se uma técnica promissora na quantificação de TPH em amostras de solo, possibilitando análises por refletância difusa, sem necessidade de preparação de amostra, ou após extração realizada por um solvente e aquisição espectral por transfletância.¹⁸

Na literatura, alguns estudos mostram resultados promissores para quantificação de TPH por técnicas espectroscópicas no infravermelho médio e próximo,^{10,22} enquanto outros não obtiveram bons resultados utilizando NIR e vis-NIR devido à complexidade do conjunto amostral, especialmente, pela variação de umidade no solo.²³⁻²⁵ Sendo assim, o conjunto de

calibração deve ser bem planejado prevendo o máximo de interferências possíveis, como umidade e textura do solo, variação entre combustíveis, exatidão e precisão do método de referência, entre outras possíveis fontes de variação. Visto o grande número de interferências, o bom desempenho do modelo de predição pode depender do uso de ferramentas quimiométricas mais sofisticadas para contornar a falta de linearidade e aumentar a seletividade e sensibilidade do método analítico.

2. OBJETIVOS

Visto o grande número de contaminações referentes a postos de combustíveis, os riscos ambientais gerados por hidrocarbonetos do petróleo e a demanda por métodos analíticos rápidos e econômicos, este trabalho teve como objetivo principal o desenvolvimento de métodos para identificação e quantificação de TPH e suas frações em solo através da técnica FT-NIR e métodos multivariados.

Além do objetivo principal, propôs-se o desenvolvimento de métodos cromatográficos ultrarrápidos para determinação das frações de TPH (GRO, DRO e ORO) e dos compostos individuais BTEX (benzeno, tolueno, etilbenzeno e xilenos) e avaliou-se a otimização do processo de extração dos contaminantes por agitação combinada à sonicação utilizando solventes não halogenados.

3. REVISÃO BIBLIOGRÁFICA

3.1. Solo

O solo é um dos componentes críticos e mais importantes do ecossistema e de recursos ambientais devido à sua enorme influência no fornecimento de nutrientes às plantas, participação no ciclo do carbono e na filtração da água, além de ser essencial para a produção de alimentos e habitação para os seres humanos.²⁶⁻²⁸

Como resultado da ação de fatores geológicos, topográficos, climáticos, e antropogênicos, o solo é um meio heterogêneo cujas variáveis, como composição e teor de minerais, matéria orgânica, água, ar e organismos vivos, tornam o solo um meio tão complexo.^{29,30}

Além das funções intrínsecas do solo relacionadas à sustentação da vida, manutenção e proteção do ciclo da água e de nutrientes, o aumento das atividades antropogênicas torna o solo responsável pela manutenção da atividade sócio econômica e, principalmente, diminui o impacto em bacias hidrográficas causado por poluentes, como agrotóxicos e resíduos industriais.^{29,31}

O solo é constituído por três fases: (i) sólida, formada de material rochoso local ou transportado e material orgânico; (ii) líquida, constituída por solução aquosa de substâncias orgânicas e inorgânicas; e (iii) gasosa, de composição variável de acordo com a biomassa envolvida.²⁹

As propriedades físicas do solo, como textura, densidade, estrutura e permeabilidade, dependem da espécie mineralógica que deu origem e dos mecanismos de intemperismo e transporte. Conforme a influência destes fatores, o solo apresentará diferentes proporções das frações granulométricas: areias (0,053–2,000 mm); silte (0,002–0,053 mm); argila (0,0001–0,0020 mm); e colóides (<0,0001 mm).²⁹

O tipo de material e a granulometria influenciam nas propriedades físico-químicas do solo, nos mecanismos de atenuação e transporte de poluentes. Assim, a mobilidade da água e de poluentes dependem da forma, do tamanho e da compactação das partículas que constituem o solo (fatores que governam a permeabilidade do solo). Além das características físicas, as propriedades químicas e biológicas também influenciam no transporte de poluentes através de mecanismos de adsorção, absorção, precipitação, oxidação, troca iônica e neutralização de íons.²⁹

Apesar da heterogeneidade e complexidade do solo, as informações sobre suas propriedades são essenciais para tomar decisões sobre a qualidade, preservação e manejo ambiental, além da segurança dos alimentos produzidos por meio de seus recursos.^{28,32} Visto que a degradação do solo por monoculturas, poluição e urbanização crescem juntamente à necessidade do aumento da produção de alimentos, torna-se crucial a utilização de metodologias que otimizem a obtenção de informações sobre a qualidade do solo para tomadas de decisões e, assim, manter sua preservação.^{26,27,33}

3.1.1. Poluição do solo

Embora o solo atue como um filtro de poluentes para os lençóis freáticos e bacias hidrográficas, sua capacidade é limitada devido ao efeito cumulativo da deposição de poluentes, alterando a qualidade do solo.²⁹ Além da poluição atenuar a capacidade das funções do solo para com o ecossistema, afeta diretamente na produção de culturas e, conseqüentemente, na saúde humana.³¹

A preocupação da sociedade e das autoridades com a poluição do solo vem crescendo constantemente, contudo, ainda existem dificuldades técnicas e barreiras políticas/econômicas para acessar sua qualidade, assim como sua preservação e remediação.³⁴ A falta de controle e monitoramento adequado do solo, principalmente em países em crescimento urbano, agrícola e industrial, podem levar a conseqüências irreversíveis para a ciclagem de nutrientes e da água.^{29,35} Como os solos frequentemente reagem lentamente de forma perceptível a mudanças físicas e químicas, a detecção de mudanças na qualidade pode ser difícil antes da ocorrência de danos irreversíveis, tornando necessário o monitoramento constante de suas propriedades.³¹

Ainda que a maior parte das contaminações aconteça de forma localizada – contaminação por depósito de resíduos, descartes incorretos, uso agrícola indevido, vazamentos ou derramamentos – quando o poluente atinge a superfície do solo, este pode ser lixiviado pelas águas de infiltração até alcançar águas subterrâneas e, então, ser transportado para outras regiões através do fluxo das águas.²⁹

Devido às abrangentes atividades antropogênicas com potenciais riscos de contaminação, muitas substâncias tóxicas, mesmo em baixos níveis de concentração, devem ser monitoradas no solo. Conforme a atividade realizada no local, um grupo ou mais de substâncias são determinados e as concentrações são comparadas com os valores orientadores para a definição da qualidade do solo. No Brasil, os valores orientadores para as substâncias perigosas são especificados na Decisão de Diretoria Nº 256/2016/E, de 22 de novembro de 2016,

realizado pela Companhia Ambiental do Estado de São Paulo (CETESB).³⁶ Dentre as substâncias monitoradas, enquadram-se metais pesados; hidrocarbonetos aromáticos voláteis (BTEX); hidrocarbonetos policíclicos aromáticos (PAH); organoclorados; fenóis; ésteres ftálicos; dioxinas; furanos; e hidrocarbonetos totais do petróleo (TPH), os quais abrangem hidrocarbonetos de 6 a 40 átomos de carbono.^{6,36}

3.1.2. Contaminação de solo por hidrocarbonetos

Embora exista uma busca constante por fontes de energia alternativa e processos químicos ambientalmente amigáveis, a economia atual ainda é altamente dependente dos produtos derivados do petróleo para o transporte e em processos industriais.^{2,37} Esses produtos, incluindo gasolina, diesel e óleos lubrificantes, são manipulados várias vezes desde a produção até o mercado final.^{38,39} Durante o transporte, armazenamento e uso, são eventualmente liberados acidentalmente no meio ambiente, ocasionando poluição local e difusa dos solos e das águas subterrâneas.⁴⁰ Este tipo de contaminação ambiental é generalizado e crescente devido à expansão do número de instalações e processos que são potenciais fontes de contaminação.^{41,42} Em consequência do uso expansivo de combustíveis fósseis, os principais contaminantes em solo em regiões urbanas e industriais são alcanos, hidrocarbonetos aromáticos e substâncias como piridina, fenóis, acetona, entre outros solventes.⁴³

Óleos lubrificantes minerais apresentam maiores riscos para o meio ambiente que combustíveis como gasolina e diesel, apesar da baixa mobilidade de hidrocarbonetos de cadeias extensas, principais constituintes destes produtos. Grande parte dos óleos lubrificantes novos possui alta concentração de hidrocarbonetos voláteis e com maior solubilidade em água. Enquanto, o longo tempo de uso destes produtos no motor pode proporcionar a formação de novos compostos que contenham metais pesados como chumbo, cromo, cádmio, ou outras substâncias tóxicas como hidrocarbonetos clorados, naftalenos e óxidos de enxofre.^{35,42}

A contaminação por hidrocarbonetos em solo tem sido de grande preocupação para a ciência ambiental devido à sua persistência no meio ambiente e aos efeitos adversos para o ecossistema.^{44,45} Os hidrocarbonetos, principalmente aromáticos, impactam negativamente as características biológicas, físicas e químicas do solo.⁴⁶ A presença desses no solo pode alterar significativamente propriedades como microestrutura, permeabilidade, pH, nitrogênio total e fósforo total.⁴⁷ A natureza fitotóxica dos hidrocarbonetos e a imobilidade de nutrientes afetam diretamente as plantas na região contaminada.⁴⁸

Segundo o Instituto Americano do Petróleo (API), concentrações de TPH abaixo de 10.000 mg kg⁻¹ no solo não afetam o crescimento da maioria das plantas e não expõe risco de lixiviação para águas subterrâneas. Entretanto, em alguns estados dos EUA, o valor limite de TPH em solo é equivalente a 100 mg kg⁻¹, o qual é utilizado como parâmetro de qualidade para contaminações por petróleo ou combustíveis fósseis.⁴⁹

Os hidrocarbonetos do petróleo podem estar presentes nos solos em diversas formas.⁴⁷ Em solos saturados de hidrocarbonetos por longo tempo, ocorre a formação de estruturas complexas entre as partículas do solo e os contaminantes. Uma porção de TPH é adsorvida ou absorvida pelas micropartículas do solo de forma fraca ou forte, dependendo da cadeia carbônica e da composição da matéria orgânica no solo, enquanto outra porção é presa em microporos do solo, o que dificulta a extração dos contaminantes.⁴⁷

A mobilidade de TPH no solo é controlada diretamente ou indiretamente por diversos fatores como a composição da mistura de TPH, textura do solo, teor de água, matéria orgânica e minerais no solo.⁴⁷ Os hidrocarbonetos mais voláteis apresentam maior dinâmica e são distribuídos de forma não uniforme no solo. Já os hidrocarbonetos de cadeia maior (semivoláteis) apresentam maior estabilidade, podendo acumular por longos períodos e se degradarem para hidrocarbonetos voláteis durante o tempo.⁵⁰ Uma das causas do acúmulo de hidrocarbonetos de cadeia longa é a alta hidrofobicidade, a qual diminui sua mobilidade e, principalmente, torna-os menos disponíveis para biodegradação devido à menor sorção biológica.⁵¹

Os hidrocarbonetos aromáticos apresentam maior mobilidade no solo que alcanos devido à maior solubilidade em água e, por conseguinte, maior interação com a matéria orgânica disponível no solo.⁴¹ A gasolina e o diesel contêm, respectivamente, até 35 %(v/v) e 11 %(v/v) de compostos aromáticos em sua composição, o que torna ainda mais preocupante a contaminação ambiental devido à solubilidade e toxicidade destes compostos.^{3,52} Além disso, os biocombustíveis adicionados à gasolina (27 %(v/v) de etanol) e ao diesel (10 %(v/v) de biodiesel) aumentam a solubilidade de compostos aromáticos, aumentando ainda mais a mobilidade de compostos BTEX e PAH no solo.^{53,54}

A exposição de TPH ao ser humano pode afetar a atividade cerebral e seu desenvolvimento, além de provocar náusea, desorientação, confusão mental, problemas de fala e distúrbios de memória. A exposição a altos níveis pode causar extrema debilitação, perda de equilíbrio, coma, convulsão e até morte. Exposições de longo período podem acarretar alterações na capacidade neurofisiológica ou psicológica e induzir a câncer de pulmão, pele e bexiga, além de outros efeitos carcinogênicos.¹⁰ Dentre os hidrocarbonetos do petróleo, os

compostos aromáticos apresentam maior risco à biota e aos seres humanos devido às suas propriedades carcinogênicas e mutagênicas. Os compostos BTEX podem levar a diversas doenças no sistema sanguíneo e linfático como diminuição de células brancas, linfoma de Hodgkins, anemia aplástica, leucemia aguda, anormalidade na medula óssea e síndrome mielodisplásica.⁵⁵ Em consequência da alta toxicidade dos compostos aromáticos (BTEX e PAH), os valores orientadores para prevenção do solo para estes compostos são menores que 1 mg kg^{-1} , no Brasil. Já para o benzeno, o valor para prevenção do solo é igual a $0,002 \text{ mg kg}^{-1}$ e para o etilbenzeno e xilenos é equivalente a $0,03 \text{ mg kg}^{-1}$.³⁶

Além dos riscos à saúde humana apresentados pelos hidrocarbonetos do petróleo, os microrganismos do solo são muito sensíveis a qualquer perturbação do ecossistema, e sua diversidade e atividade são rapidamente alteradas por tais perturbações.⁵¹ Estudos mostram que a presença de hidrocarbonetos do petróleo no solo na faixa de concentração de $8\text{--}34.600 \text{ mg kg}^{-1}$ diminui a taxa de crescimento microbiano, a produção de uréase, entre outras atividades metabólicas, com o aumento da concentração de TPH.^{40,56}

Visto que os combustíveis fósseis e os óleos lubrificantes são as principais fontes de hidrocarbonetos alifáticos e aromáticos em regiões urbanizadas, os postos de combustíveis tornam-se fontes potenciais de contaminação ambiental.⁴⁴ As principais causas de contaminação do solo em postos de combustíveis estão presentes no abastecimento dos tanques subterrâneos e, sobretudo, em falhas nos tubos e tanques subterrâneos.^{53,57} Apesar da estocagem de combustíveis em tanques subterrâneos reduzir os riscos de incêndio e explosão, este armazenamento leva a maiores riscos de contaminação ambiental devido ao processo de corrosão dos tanques. Além disso, tanques subterrâneos apresentam maior dificuldade na inspeção da integridade que tanques aéreos e pequenos vazamentos podem levar a centenas de litros perdidos para ambiente antes que sejam notados.⁵⁸ Muitas vezes, as contaminações causadas por vazamentos de tanques subterrâneos são notadas apenas após sua remoção para substituição por tanques novos.⁵⁹

O principal motivo dos vazamentos em tanques subterrâneos é a corrosão nos pontos de solda causada por diversos fatores do solo, como salinidade, acidez, umidade, concentração de oxigênio, etc.⁵³ “O Procedimento para Licenciamento Ambiental de Postos e Sistemas Retalhistas de Combustíveis”, de 2018, desenvolvido e publicado pela CETESB, estabelece que tanques com mais de 15 anos devem ser substituídos por novos a fim de reduzir os riscos de contaminação.⁶⁰ A partir de 2001, os tanques de armazenamento passaram a ser desenvolvidos com paredes duplas, sendo a parede externa não metálica, e apresentam um sensor de vazamento no espaço intersticial entre as paredes.⁵³

A partir de 2002, a CETESB tem divulgado as áreas contaminadas cadastradas no estado de São Paulo. Dentre as 5942 áreas cadastradas até 2017, os postos de combustíveis são responsáveis por 72% dos registros, seguidos das atividades industriais com 19%, das atividades comerciais com 5%, das instalações para destinação de resíduos com 3% e dos casos de acidentes, agricultura e fonte de contaminação de origem desconhecida com 1%. Por outro lado, o crescente número de áreas contaminadas por postos de combustíveis é resultante da implementação do Programa de Licenciamento de Postos de Combustíveis que se iniciou em 2001, com a publicação da Resolução CONAMA No 273, de 2000.^{6,61} Através deste programa, a aquisição ou renovação da licenciamento de postos de combustíveis obriga a análise ambiental dos grupos de contaminantes TPH, BTEX e PAH.⁷

De acordo com os dados apresentados pela CETESB, os principais grupos de contaminantes remetem às atividades industriais e, principalmente, à revenda de combustíveis. Dentre os contaminantes determinados no estado de São Paulo de 2001 a 2017, os principais foram combustíveis automotivos (4543 constatações); solventes orgânicos aromáticos – BTEX (4049 constatações); hidrocarbonetos policíclicos aromáticos – PAH (2494 constatações); metais pesados (1101 constatações) e TPH (578 constatações).⁶ Ressalva-se que as constatações para BTEX, PAH e TPH podem estar relacionadas, respectivamente, à contaminação por gasolina, diesel, e óleo lubrificante.

3.1.3. Processo de investigação de contaminação de solo por hidrocarbonetos.

Em todo o mundo, protocolos nacionais ou estaduais foram desenvolvidos para a avaliação de locais contaminados, estabelecendo sistemas de apoio à decisão para avaliar a necessidade de remediação do solo contaminado. A maioria dos protocolos propõe modelos de decisão baseados nos riscos à saúde humana, ao ecossistema e à segurança alimentar. A avaliação de risco costuma ser específica ao local em questão e depende do uso futuro da área contaminada. Independentemente do objetivo futuro do solo, as decisões são apoiadas por uma investigação preliminar. Dependendo do grau de contaminação e do grupo de contaminantes, são realizadas investigações mais intensas e detalhadas. A abordagem inicial permite que os dados de cada estágio sejam avaliados e usados no planejamento das etapas de investigação detalhada e na remediação. Em cada nível progressivo da investigação, a avaliação baseia-se em informações mais específicas do local e, portanto, é mais complexa, demorada e mais custosa.⁶²

A estratégia de avaliação de riscos ambiental foi primeiramente desenvolvida, em 1995, pela agência ambiental americana Environmental Protection Agency (EPA)⁶³ e, no mesmo ano, adaptada e homologada pela American Society for Testing and Materials (ASTM), através da norma ASTM E1739.⁶⁴ Essa abordagem, conhecida como ações corretivas baseada no risco (risk-based corrective action – RBCA), avalia os graus de contaminação e os riscos de exposição e transporte dos contaminantes em três etapas de investigação. A primeira etapa de avaliação (tier 1) envolve análises de varredura, nas quais quantifica-se a concentração de TPH no solo, entre outros possíveis contaminantes. Na segunda etapa (tier 2), são realizadas análises mais detalhadas sobre os contaminantes, como a determinação dos compostos individuais dos grupos PAH e BTEX. Finalmente, na terceira etapa (tier 3), são realizadas determinações de biomarcadores específicos para planejamento das estratégias de remediação da área contaminada.⁴⁶ A vantagem desta abordagem está no estudo de casos específicos que, portanto, pode otimizar significativamente os custos envolvidos na investigação da contaminação e nos processos de remediação.⁵³ O valor de TPH é um parâmetro útil na avaliação de risco, pois possibilita uma avaliação rápida do grau de contaminação e fornece informações úteis para o processo de remediação.⁶⁵

No Brasil, a estratégia RBCA foi adaptada em 2000, pela Câmara Ambiental do Comércio de Derivados de Petróleo, para monitorar e remediar a contaminação ambiental por postos de revenda de combustíveis, distribuidoras e outras áreas onde manipulem ou armazenem outros compostos derivados de petróleo.⁵³ Atualmente, a estratégia é descrita detalhadamente no “Procedimento para a Proteção da Qualidade do Solo e das Águas Subterrâneas”, desenvolvido pela CETESB e aprovado pela Decisão de Diretoria nº 038/2017/c, de 07 fevereiro de 2017.⁶⁶ O modelo brasileiro da análise de risco segue o mesmo protocolo desenvolvido inicialmente pela EPA, o qual avalia o meio contaminado, a fonte de contaminação, grupos de compostos químicos envolvidos, mecanismos de transporte dos poluentes e receptores potenciais (fauna e flora).

No Brasil, o programa de licenciamento de postos de combustíveis tem colaborado intensamente com a identificação e remediação de áreas contaminadas. Das 4284 áreas contaminadas por postos de combustíveis registradas até 2017, 9% encontraram-se em processo de investigação; 14% com risco ambiental confirmado; 27% em processo de remediação; 19% reabilitadas para uso; 1% foram reutilizadas devido ao baixo risco ambiental; e 30% em processo de monitoramento dos níveis de contaminação após a remediação.⁶ O licenciamento ambiental de estabelecimentos que envolvam armazenamento e abastecimento de combustíveis automotivos, assim como troca de óleo lubrificante e lavagem de veículos devem seguir os

Quadros de Exigências para o Licenciamento Ambiental,⁶⁷ e as licenças ambientais devem ser renovadas a cada cinco anos.⁶⁰

Dentre as exigências para o licenciamento ambiental, são requeridas as análises de TPH, BTEX e PAH em solo. A amostragem de solo deve ser realizada para cada tanque de armazenamento de combustível (exceto de etanol), efetuando-se três sondagens em cada até a profundidade de 5 metros, quando possível; e nas unidades de abastecimento e filtros de diesel, deve ser realizada uma sondagem até 2 metros de profundidade. O método de sondagem deve ser realizado de acordo com a geologia e a hidrogeologia local. Para cada metro, uma amostra é retirada através de um liner e dividida em duas alíquotas. A primeira alíquota é utilizada para a análise de gases logo após a amostragem e a segunda alíquota é mantida no liner e armazenada sob refrigeração (abaixo de 4°C). Após todas amostras terem a primeira alíquota avaliada, a segunda alíquota da amostra que apresentou maior concentração de gases será transferida para um saco plástico impermeável autosselante (preferencialmente de polietileno) de um litro de capacidade, para análise de TPH, BTEX e PAH em laboratório.⁷ Em seguida, os valores dos contaminantes são comparados com os valores dos “níveis aceitáveis baseados no risco” (NABR), estabelecidos pela Decisão de Diretoria nº 256/2016/e, de 22 de novembro de 2016, da CETESB.³⁶ Os valores NABR são estabelecidos para cada composto individual em três níveis: (i) valor de referência de qualidade, o qual indica o valor máximo aceitável de contaminação; (ii) valor de prevenção, utilizado para indicar a qualidade de um solo capaz de sustentar as suas funções primárias, porém que exige um monitoramento para avaliar possíveis impactos e alertar o uso agrícola na área afetada; e (iii) valor de intervenção, indicativo de riscos potenciais diretos ou indiretos à saúde humana.⁶⁸ Os valores NABR de intervenção são divididos conforme as áreas em avaliação: agrícola (limites inferiores); residencial (limites intermediários); e industrial (limites superiores).³⁶ Apesar de TPH não estar presente na tabela de valores NABR da Decisão de Diretoria nº 256/2016/e, o valor limite de intervenção deste grupo em solo é fixado em 1000 mg kg⁻¹ no Sistema de Licenciamento de Postos IV, da CETESB.⁷ Os valores NABR dos compostos BTEX, PAH e TPH estão dispostos na Tabela 1.

Tabela 1 – Valores de prevenção e intervenção de contaminantes em solo.^{7,36}

Substância	Concentração (mg kg ⁻¹)			
	Prevenção	Intervenção		
		Agrícola	Residencial	Industrial
TPH	-	1000	1000	1000
benzeno	0,002	0,02	0,08	0,2
etilbenzeno	0,03	0,2	0,6	1,4
tolueno	0,9	5,6	14	80
xilenos	0,03	12	3,2	19
antraceno	0,3	2300	4600	10000
benzo(a)antraceno	0,2	1,6	7	22
benzo(b)fluoranteno	0,7	2	7,2	25
benzo(k)fluoranteno	0,8	27	75	240
benzo(g,h,i)perileno	0,5	-	-	-
benzo(a)pireno	0,1	0,2	0,8	2,7
criseno	1,6	95	600	1600
dibenzo(a,h)antraceno	0,2	0,3	0,8	2,9
fenantreno (3,4)	3,6	15	40	95
indeno(1,2,3-c,d)pireno	0,4	3,4	8	30
naftaleno	0,7	1,1	1,8	5,9

Embora TPH seja um indicativo útil para avaliar a contaminação de solos, é de extrema importância avaliar este parâmetro em conjunto com os compostos aromáticos individuais para acessar de fato o risco ambiental ocasionado pelo contato de combustíveis com o solo. A grande diferença entre os limites de intervenção de TPH e dos compostos aromáticos individuais pode gerar falsos positivos ou negativos para contaminação. A concentração de fenantreno no solo tem sido utilizada como indicativo da contaminação no solo por PAH,⁶⁹ contudo, outros compostos apresentam limites muito inferiores ao fenantreno, como benzo(a)pireno e dibenzo(a,h)antraceno, podendo levar a falsos negativos para contaminação de solo por diesel, por exemplo. O mesmo é válido para a identificação de contaminação através de compostos cujos limites são elevados comparados aos limites de TPH, como os limites de intervenção para antraceno. Em muitos trabalhos, é comum apresentar apenas a concentração de grupos de compostos, como a somatória dos compostos BTEX ou PAH.⁷⁰ Desta forma, torna-se essencial a determinação de TPH e dos compostos aromáticos individuais em solo para avaliar o risco ambiental.^{41,71}

Dada a grande importância e as dificuldades encontradas na investigação de áreas contaminadas, o Decreto 59.263, de 05 de junho de 2013, da Assembleia Legislativa do Estado de São Paulo, o qual regulamenta a Lei nº 13.577, de 8 de julho de 2009, estabeleceu a obrigatoriedade da atualização contínua do cadastro de áreas contaminadas e reabilitadas;

determinou a criação de instrumentos econômicos para financiar a investigação e remediação, além de apoiar as futuras iniciativas para a revitalização de regiões industriais e comerciais desativadas ou abandonadas.⁶

3.1.4. Remediação de áreas contaminadas

Antes da remediação dos meios impactados, é essencial realizar uma avaliação completa dos riscos apresentados pelos contaminantes ao meio ambiente e à saúde humana para criar estratégias e escolher a técnica mais adequada para remediação. Esta avaliação envolve tanto a identificação da composição química como a determinação da área afetada pelos contaminantes acima do valor de intervenção.⁴⁶

A remediação de solos contaminados por combustíveis fósseis e óleos lubrificantes é dificultada pela natureza dos hidrocarbonetos principalmente de longa cadeia, como alta hidrofobicidade e baixa volatilidade.⁴³ Assim, torna-se ainda mais desafiador a descontaminação de solos afetados por diesel e óleos lubrificante que aqueles contaminados por gasolina.⁴ Os processos de recuperação da qualidade do solo costumam ser muito custosos devido aos métodos analíticos para avaliação, às técnicas de remediação e ao grande volume de solo de contaminado.⁶²

A seleção da técnica de remediação é extremamente dependente da força de ligação dos TPH com a matriz sólida, assim como o coeficiente de partição dos contaminantes com o solvente utilizado.^{47,72} A remoção ou degradação de TPH no solo pode ser realizada por processos físicos, químicos, biológicos ou pela associação de diferentes processos. As técnicas de remediação mais comuns são extração de vapores do solo, recuperação de produto livre, bioventilação, extração com solventes, incineração, torres de aeração, adsorção em carvão ativado, biorremediação, fitorremediação, entre outros.⁵³ De acordo com os dados disponibilizados pela CETESB, as técnicas mais utilizadas no estado de São Paulo para remediação de solo são extração por bombeamento e tratamento; remoção de solo/resíduo; extração de vapores; biorremediação e cobertura de resíduo/solo contaminado.⁶ A biorremediação, na qual agentes biológicos são utilizados para a degradação de hidrocarbonetos, tem se mostrado a técnica de remediação com melhor custo-benefício, além de ser ambientalmente amigável.⁴⁸ Vários estudos têm mostrado que a combinação de técnicas física, químicas e biológicas, ou mais de uma técnica biológica, são mais eficientes e podem reduzir o tempo e os custos do processo de remediação.^{35,55,73,74}

Apesar do número crescente de identificações de áreas contaminadas no estado de São Paulo, o número de investigações e recuperações são, incontestavelmente, dependentes dos custos requeridos pelos métodos analíticos e processos de remediação.⁵⁹ Portanto, o desenvolvimento e a implementação de métodos analíticos de menores custos são primordiais para o crescimento do monitoramento de áreas com potencial risco de contaminação por vazamento e/ou derramamento de combustíveis fósseis e óleos lubrificantes minerais.

3.2. Hidrocarbonetos totais do petróleo – TPH

Os hidrocarbonetos presentes no petróleo bruto e nos produtos derivados, como gasolina e diesel, abrangem centenas de diferentes compostos e, apesar da complexidade de suas composições, podem ser categorizados em três principais grupos: alifáticos; aromáticos; e compostos contendo heteroátomos como nitrogênio, enxofre e oxigênio. Grande parte da composição do petróleo é constituída de cicloalcanos, também conhecidos por naftenos, e por compostos cuja estrutura é a combinação de naftenos e anéis aromáticos.⁶² A proporção destes compostos no petróleo bruto depende de vários fatores relacionados ao seu histórico geológico.⁷⁵

O desenvolvimento dos processos de refinação proporcionou a separação de várias frações do petróleo, obtendo-se diversos produtos como o gás natural, gás liquefeito do petróleo (GLP), nafta, gasolina, querosene, diesel, graxa, asfalto, entre outros.⁷⁶ Embora, haja o desenvolvimento e aprimoramento constante de biocombustíveis e processos de obtenção destes, atualmente, o transporte rodoviário é extremamente dependente dos derivados do petróleo (gasolina e diesel).⁷⁷

A composição da gasolina e do diesel já é bem caracterizada desde o desenvolvimento de técnicas modernas de cromatografia associada a detectores de espectrometria de massa (GC-MS).⁷⁸ A gasolina é um produto gerado a partir de muitos fluxos de refinaria. Em sua composição, existem mais de 200 componentes individuais no intervalo de ebulição de hidrocarbonetos com 4 a 12 átomos de carbono. Estes hidrocarbonetos são constituídos de 40% a 70% de alcanos lineares, ramificados e cíclicos; de alcenos em baixas concentrações que variam até 10%; e de 20 a 50% de compostos aromáticos, sendo majoritários os compostos benzeno, tolueno, etilbenzeno e xilenos.⁵

Analogamente à gasolina, o diesel de uso rodoviário é constituído principalmente por hidrocarbonetos alifáticos e aromáticos que variam de 8 a 28 átomos de carbono, e a

concentração de compostos aromáticos pode variar de 10% a 40% (v/v).⁵ A composição do diesel, formulada por hidrocarbonetos de massa molecular superior aos constituintes da gasolina, proporciona a este combustível maior ponto de fulgor, densidade, temperaturas de destilação e teor de resíduos que a gasolina.⁷⁹

Os óleos lubrificantes são essenciais para o funcionamento de motores, não apenas automotivos. Sua composição varia conforme sua aplicação e as condições de trabalho do motor, como faixa de temperatura mínima e máxima. Basicamente são constituídos por hidrocarbonetos parafínicos (alcanos de cadeia linear) que variam normalmente entre 20 e 45, ou mais, átomos de carbono. A base parafínica é enriquecida com outros compostos cíclicos, aromáticos e não hidrocarbonetos para atender as propriedades físico-químicas como densidade, viscosidade, lubricidade, ponto de fulgor etc. Os óleos lubrificantes podem conter de 10% a 30% (v/v) de compostos aromáticos policíclicos em sua composição. Além do enriquecimento citado, são inseridos aditivos para aprimorar seu desempenho e prolongar a capacidade de lubricidade, como detergentes metálicos, antioxidantes, modificadores de fricção, desemulsificantes, melhoradores de viscosidade, entre outros.⁵

A composição destes combustíveis e dos óleos lubrificantes automotivos depende da peculiaridade da composição do petróleo utilizado, dos processos envolvidos em cada refinaria e da demanda do mercado (como aumento do teor de biocombustíveis para atender a demanda).^{5,76,80}

Os hidrocarbonetos policíclicos aromáticos (PAH) além de estarem presentes no diesel e nos óleos lubrificantes, também são fontes destes contaminantes a combustão incompleta de carvão mineral e combustíveis, incêndios florestais, combustões industriais, sistemas de aquecimento doméstico e emissões vulcânicas.⁸¹ De acordo com suas propriedades físico-químicas e comportamentos em meio biológico, os PAH podem ser classificados em compostos de baixa massa molecular (2 a 3 anéis) e de alta massa molecular (4 a 7 anéis). Os hidrocarbonetos de baixa massa molecular apresentam efeito tóxico agudo, conseqüente de sua alta biodisponibilidade, enquanto os de alta massa molecular são mais persistentes no meio ambiente devido à menor volatilidade e maior interação com partículas finas do solo.⁸²⁻⁸⁴ Já os hidrocarbonetos monoaromáticos, como os compostos BTEX, e o naftaleno possuem maior solubilidade e menor adsorção nas partículas de solo, portanto, maior mobilidade no solo, por outro lado, são mais susceptíveis à biodegradação devido a suas estruturas mais simples.⁸⁵

Visto a complexidade da composição dos combustíveis fósseis e dos óleos lubrificantes, novamente o parâmetro TPH surge como um indicativo útil do grau de contaminação

ambiental.^{65,86} Para a maioria dos casos, o uso do parâmetro TPH é suficiente para verificar a ocorrência de contaminação e a eficiência do processo de remediação.⁸⁷ A definição de TPH é complexa, pois abrange uma extensa faixa de hidrocarbonetos e é totalmente dependente do método analítico utilizado em sua quantificação.^{5,47} A determinação de TPH pode ser realizada por diferentes técnicas e métodos analíticos convencionais, os quais também possibilitam uso de diferentes métodos de extração e purificação (clean up). Deste modo, o valor de TPH de uma mesma amostra pode variar consideravelmente de um método para outro.⁴⁹

A Figura 1 ilustra as faixas de hidrocarbonetos conforme o tamanho das cadeias carbônicas, abrangidas por diferentes métodos EPA para determinação de TPH.⁵ Por exemplo, o método gravimétrico abrange hidrocarbonetos acima de 14 carbonos devido à etapa de evaporação do solvente extrator, enquanto a determinação por cromatografia gasosa (método EPA 8015) requer uma análise para os compostos voláteis (C₄ a C₁₀) e outra para os semivoláteis (C₁₀ a C₄₀) em decorrência dos diferentes métodos extrações utilizados para cada fração.⁵

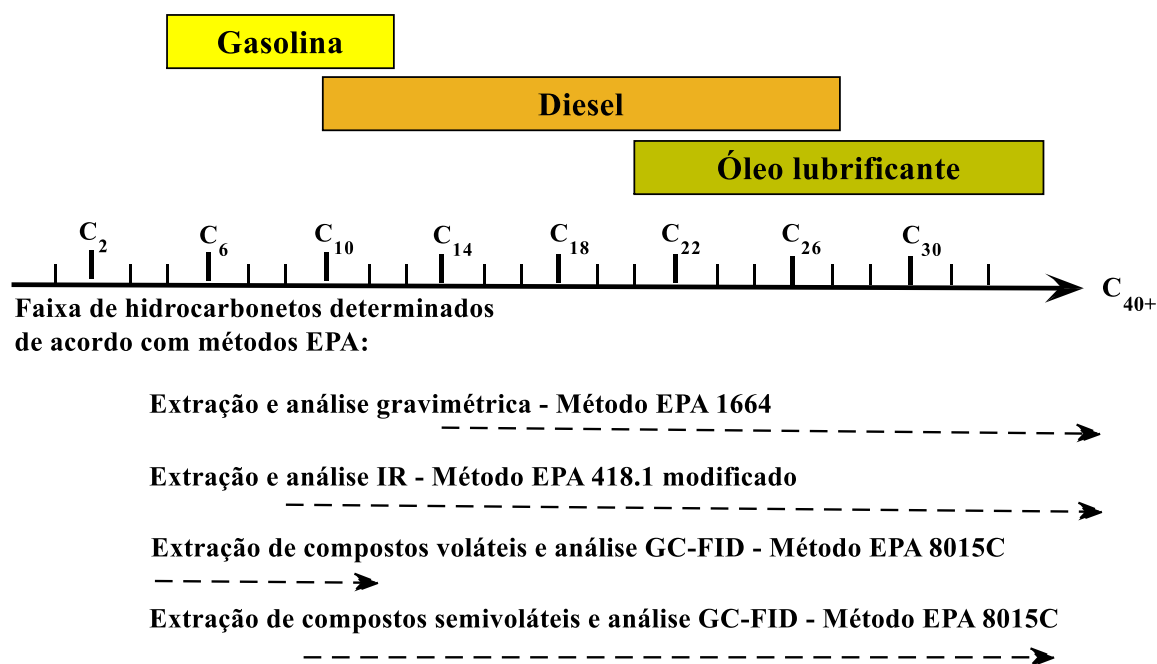


Figura 1 – Faixa de hidrocarbonetos abrangida por diferentes métodos analíticos.
 FONTE: Adaptado de Weisman (1998).⁵

3.2.1. Métodos analíticos para determinação de TPH

A avaliação da qualidade do solo através da concentração de TPH é uma das tarefas mais solicitadas, porém uma das menos padronizadas.⁵⁰ A possibilidade do uso de diferentes métodos analíticos para a quantificação de TPH pode resultar em divergências nos valores quantificados. Caso o método utilizado não seja o mais adequado para determinada situação ou seja usado de forma incorreta, a concentração de TPH obtida pode representar materiais naturais do solo, como raízes e folhas, e levar a falsos positivos para contaminação.⁴⁹

As metodologias tradicionais são baseadas em análise gravimétrica (método EPA 1664); cromatografia em fase gasosa (métodos EPA 8015 e EPA 8270) e; espectroscopia na região do infravermelho médio (método EPA 418.1 e EPA 8440).^{8,9,11,13,88} E os métodos de extração convencionais são baseados em de sonicação (método EPA 3550); Soxhlet (método EPA 3540); fluido pressurizado (método EPA 3545); fluido supercrítico (método EPA 3560); micro-ondas (método EPA 3546); e extração de voláteis por headspace (método EPA 5021) ou purge-and-trap (método EPA 5030).^{89,90}

Com o desenvolvimento de instrumentações portáteis, novos métodos com diferentes técnicas vêm sendo desenvolvidos para a determinação de TPH e PAH em campo. Entre elas, destacam-se a espectroscopia nas regiões do visível (vis), infravermelho próximo (NIR), região de alta frequência do infravermelho médio (SWIR) e infravermelho médio (MIR); espectroscopia Raman; e imunoenaios.^{46,50,62}

A portabilidade de instrumentos analíticos possibilita a análise de passivos ambientais em campo. Quanto mais próximo o instrumento estiver da amostra, mais representativos serão os dados e mais rapidamente as informações poderão ser obtidas sobre o local.⁹¹ Análises espectroscópicas, principalmente as baseadas em transformadas de Fourier, possibilitam quantificações rápidas e de baixo custo. Desta forma, essas técnicas apresentam-se cada vez mais promissoras para a avaliação de locais que exigem a análise de grande quantidade de amostras.³³ Por outro lado, os métodos alternativos ainda necessitam de aprimoramento para tornar a quantificação de contaminantes mais precisa, robusta e reprodutível.⁹²

Além dos métodos de quantificação de TPH e suas frações, métodos de preparo de amostras estão em constante desenvolvimento para a redução do tempo de extração e para o uso de solventes extratores menos nocivos ao meio ambiente e/ou em menores volumes, de forma que atendam aos baixos limites de quantificação exigidos para a qualidade do solo.^{93,94} Nas próximas seções, serão detalhados os métodos convencionais e alternativos de extração e quantificação de TPH.

3.2.2. Métodos de extração

3.2.2.1. Soxhlet

A extração de hidrocarbonetos através de um aparato Soxhlet é baseada na recirculação de um solvente de baixo ponto de ebulição (abaixo de 70 °C) e com elevado coeficiente de partição para os analitos. Usualmente, são utilizados diclorometano, tetracloreto, clorofórmio, hexano, acetona, ou misturas destes solventes.¹⁵ Por ser uma técnica de extração quase exaustiva, torna-se possível a determinação de contaminantes em baixas concentrações no solo (na ordem de partes por bilhão) e é uma das técnicas mais usuais para extração de compostos não voláteis ou semivoláteis.⁵⁰ A extração por Soxhlet é principalmente recomendada para análises de amostras úmidas de solo e quantidades superiores a 3 g de amostra.³⁹ Por outro lado, esta técnica apresenta várias desvantagens, como: (i) necessidade de grande volume de solvente (pelo menos 100 mL), sendo desfavorável à Química Verde; (ii) longo tempo de extração (cerca de 24 h), o que torna inadequado para um grande número de amostras; (iii) o alto conteúdo de umidade em amostras de solo (acima de 15% da massa) pode aumentar a variabilidade na extração devido à fraca interação entre solventes apolares e as superfícies de solo hidratadas; e (iv) a temperatura de extração diminui a recuperação de PAH de cadeias menores.^{39,84,87}

3.2.2.2. Sonicação

A extração por sonicação, ou extração ultrassônica, é um método fundamentado no efeito de cavitação causado por uma onda acústica que passa pela fase líquida com uma frequência superior a 16 kHz.⁵⁰ O tratamento ultrassônico produz bolhas de cavitação que geram temperatura e pressão interna, fazendo com que o solvente extrator se propague de forma mais eficaz e colida com as partículas da matriz. Como resultado, a matriz pode ser quebrada e expor mais área de superfície para extração pelo solvente.⁷⁴ Este é um método econômico e altamente eficiente, mas requer uma avaliação da energia aplicada, pois a degradação excessiva das partículas do solo pode resultar em um aumento na absorção de hidrocarbonetos, principalmente PAH, devido ao aumento da superfície da fase sólida e, assim, diminuindo a eficiência da extração.

Após a extração ultrassônica, há a necessidade de uma etapa de purificação da fase líquida por centrifugação e/ou filtração, normalmente por coluna preenchida por sílica gel ou Florisil (sílica gel carregada com íons magnésio), e uma etapa de pré-concentração, quando necessária.^{16,50}

Muitas vezes, a sonicação é preferível à extração por Soxhlet devido à rapidez (3 a 30 min), praticidade, volume reduzido de solvente (10 a 100 mL, dependendo da concentração dos analitos), e eficiência muitas vezes comparável à extração por Soxhlet.⁸⁷ Entretanto, vários fatores influenciam a capacidade de extração de hidrocarbonetos por sonicação.

Estudos mostram que o tipo de solvente extrator utilizado, a razão entre o solvente e massa de solo e o tempo de extração são efeitos significativos, enquanto a influência da potência ultrassônica apresenta pouca influência na recuperação de hidrocarbonetos em lodo. O estudo realizado por Guerin (1999),⁸⁴ no qual avaliou-se a capacidade de diferentes solventes e misturas para a extração de PAH por sonicação, mostrou que a mistura dos solventes diclorometano e acetona (1:1 v/v) resultou na melhor recuperação dos analitos (96%), enquanto o diclorometano puro recuperou 71%; a mistura hexano e acetona (3:1 v/v) recuperou 36%; e hexano puro resultou em 16% de recuperação. Hu e colaboradores (2016) avaliaram o uso de diferentes solventes para extração ultrassônica de TPH em lodo.⁷⁴ Dentre os solventes testados, a butanona apresentou melhores resultados de recuperação (>80% de recuperação) que os solventes ciclohexano e acetato de etila. Outros fatores também influenciam a capacidade extratora de TPH por sonicação, como: (i) a concentração dos contaminantes e a natureza do solo;⁵⁰ (ii) o tamanho da cadeia carbônica, quanto maior a cadeia menor a eficiência da extração;⁴⁷ e (iii) o período de interação entre os contaminantes e as partículas de solo, quanto mais antiga a contaminação, maior a dificuldade na extração.⁸⁴

Risdon et al (2008) desenvolveram uma metodologia para extração de TPH de solos através de sonicação sequencial⁹⁵ e tem sido utilizada em alguns estudos de contaminação por TPH e PAH.⁹⁶⁻⁹⁸ A metodologia é quase exaustiva, chegando a limites de detecção de 10 mg kg⁻¹ de TPH (diesel) em solo, em contrapartida, envolve várias etapas e o tempo necessário para extração é de pelo menos 40 min por amostra.⁹⁵

3.2.2.3. Agitação mecânica

Apesar de não haver um método oficial para extração de hidrocarbonetos do solo por agitação mecânica, é um método comumente utilizado devido à simplicidade, rapidez, baixo custo, pequeno volume de solvente necessário e, muitas vezes, à eficiência próxima à extração

ultrassônica.⁵⁰ Não obstante, vários fatores influenciam na capacidade extratora da técnica, assim como na extração por sonicção ou Soxhlet. Schwab e colaboradores (1999)³⁹ realizaram um estudo comparativo entre a extração de TPH por agitação e Soxhlet e, a partir dos resultados obtidos, puderam chegar às seguintes conclusões: (i) a elevada umidade de solos diminui a capacidade de extração, porém o uso de solventes mais polares podem otimizar a extração; (ii) quanto maior o teor de argila na textura solo, menor é a recuperação de PAH; (iii) a repetibilidade dos resultados é reduzida com o aumento do teor de água no solo; (iv) a extração por agitação foi equivalente à extração por Soxhlet quando a massa de amostra utilizada foi de 1 g a 3 g; (v) a mistura de diclorometano e acetona (1:1 v/v) tornou a extração mais eficiente devido à presença de água nas amostras. Complementando este estudo, Hu e colaboradores (2016)⁷⁴ determinaram que o tipo de solvente extrator é fator mais significativo para a extração de TPH por agitação, seguido pelos fatores velocidade de agitação, tempo de extração e razão solvente/lodo.

3.2.2.4. *Micro-ondas*

A extração por micro-ondas (radiação de micro-ondas com frequência de 300 MHz a 300 GHz) é um método mais rápido, econômico e com maior repetibilidade que o procedimento por Soxhlet.⁹⁹ A combinação do aquecimento com a agitação mecânica permite que uma extração eficiente de hidrocarbonetos seja realizada dentro de um curto período (10 a 15 min) e com um baixo consumo de solvente (30 mL). Sistemas modernos de micro-ondas, voltados ao preparo de amostras, envolvem a secagem preliminar de amostras e a evaporação do extrato. Na extração por micro-ondas, os solventes devem ser selecionados pela sua permissividade, que determina o grau de absorção da energia de micro-ondas. As desvantagens deste método incluem o maior custo em equipamento; a necessidade de purificação adicional do extrato; redução da recuperação com o aumento do teor de água; e a limitação da quantidade de amostra de solo de aproximadamente 1 g.^{50,99}

3.2.2.5. *Fluído pressurizado*

A técnica de extração por fluído pressurizado requer temperaturas elevadas (100 °C a 180 °C) e altas pressões (1500 psi a 2000 psi) para a extração de compostos semivoláteis do solo. De acordo com o método EPA 3545, a extração abrange a faixa de concentração de

250 $\mu\text{g kg}^{-1}$ a 12.500 $\mu\text{g kg}^{-1}$ para semivoláteis e o tempo de extração é de apenas 10 min. Em contrapartida, os equipamentos apresentam um custo relativamente elevado e a amostra deve ser secada por 48 h previamente à análise. Embora o método EPA 3545 recomende o uso de solução extratora composta por diclorometano e acetona ou hexano e acetona, a água em seu estado supercrítico (250 °C e 60-100 atm) pode ser utilizada como solvente extrator para PAH, porém o extrato aquoso limita a possibilidade de técnicas analíticas a serem utilizadas na quantificação.⁵⁰

3.2.2.6. *Fluído supercrítico*

A extração com fluído supercrítico (SCF) é um método que utiliza compostos no estado supercrítico como solventes. Os solventes utilizados com maior frequência são CO₂, H₂O e hidrocarbonetos gasosos, como etano, propano e propileno. Dentre os possíveis solventes, o CO₂ no estado supercrítico encontra o maior uso devido ao fato de necessitar condições mais brandas para atingir o estado supercrítico (31 °C e 74 bar) e à maior compatibilidade ambiental que hidrocarbonetos gasosos.⁵⁰ Miede e colaboradores (1998)¹⁰⁰ apresentaram resultados de comparação entre os procedimentos de extração por Soxhlet e SCF. Demonstrou-se que a recuperação dos PAH em solos foi semelhante em ambos os métodos, portanto, a capacidade de dissolução dos fluídos supercríticos foi comparável a dos solventes orgânicos. Embora a técnica SCF seja mais complexa em questões operacionais, seus resultados apresentam boa reprodutibilidade, seletividade e possibilita a aplicação em métodos on-line devido baixa de perda de compostos voláteis.¹⁰¹

3.2.2.7. *Headspace*

Os métodos de extração citados previamente não são adequados para análise de solos contaminados por gasolina devido à perda de hidrocarbonetos voláteis (<C₁₀) nos processos de agitação, sonicação, alta pressurização ou aquecimento. Desta forma, os compostos voláteis são normalmente analisados através dos métodos EPA 5021,¹⁰² o qual descreve a extração automatizada por headspace, e EPA 5030,¹⁰³ pelo método purge-and-trap. A automatização da extração dos compostos da gasolina é essencial para atingir a repetibilidade e a exatidão necessárias.

A extração por headspace envolve a adição de uma solução modificadora (solução aquosa saturada de NaCl, para promover o efeito salt-in-out) em um vial contendo a amostra de solo e o aquecimento junto à agitação do vial durante 50 min para atingir o equilíbrio entre a fase líquida e gasosa.¹⁰² Apesar do longo tempo de equilíbrio, a extração por headspace é seguida por análise cromatográfica, assim, as extrações seguintes podem ser realizadas durante o tempo entre uma análise e outra.

Estudos já avaliaram a associação da técnica headspace com a técnica de microextração em fase sólida (SPME) para análise de BTEX, PAH e TPH em solo.¹⁰⁴⁻¹⁰⁶ A associação de ambas as técnicas mostrou-se adequada para análise de compostos voláteis com boa precisão, linearidade e sem efeitos significativos para o teor de água no solo,¹⁰⁵ contudo, a técnica apresentou baixa sensibilidade para compostos semivoláteis, principalmente para PAH com cinco ou seis anéis aromáticos, inclusive em extração à temperaturas de 95 °C.¹⁰⁴ Desta forma, a técnica limita-se a substâncias com pressão de vapor suficiente para haver concentração na fase gasosa e adsorção na fibra SPME, especialmente, quando a amostragem é realizada à temperatura ambiente.¹⁰⁶ Além da baixa sensibilidade para compostos de maior ponto de ebulição, o método de extração desenvolvido necessitou de 2 h para o equilíbrio térmico e mais 40 min para a adsorção dos analitos à fibra SPME.¹⁰⁶

3.2.2.8. *Purge-and-trap*

Semelhante à técnica headspace, na extração por purge-and-trap os compostos mais voláteis são purgados para a fase gasosa em um vial através de agitação e aquecimento, porém, os compostos volatilizados são transferidos através de um gás inerte (normalmente hélio) para uma coluna preenchida com resina para adsorção e concentração dos analitos.⁶² Em consequência do “aprisionamento” dos analitos na coluna, esta técnica permite limites de detecção (LD) na ordem de 5 µg kg⁻¹, em contrapartida do LD da técnica headspace de 0,1 mg kg⁻¹.¹⁰⁷

Os parâmetros que afetam a sensibilidade da técnica purge-and-trap incluem o tempo e temperatura para a purga da amostra; purga a seco; tempo de dessorção dos analitos na resina; e acondicionamento da resina. A maioria dos fabricantes destes sistemas recomenda 11 min de purga ou 440 mL de gás de purga através da amostra. Tipicamente, a extração de BTEX em solo por esta técnica leva de 40 a 48 min.¹⁰⁸ De forma análoga à extração por headspace, o tempo necessário para a extração das amostras pode ser aproveitado entre as injeções cromatográficas. As grandes desvantagens destas técnicas encontram-se na estreita faixa de

hidrocarbonetos abrangida pela extração e, principalmente, no elevado custo de sistemas de amostragem automática.

3.2.3. Métodos de quantificação

3.2.3.1. *Gravimetria*

Os métodos gravimétricos são geralmente utilizados para a quantificação de hidrocarbonetos de alto peso molecular extraídos de solos contaminados por um solvente apolar de baixo ponto de ebulição. De acordo com o procedimento gravimétrico do método EPA 1664,⁸⁸ as amostras de solo são peneiradas para homogeneização e secas em estufa a 105 °C por 12 h. Então, os hidrocarbonetos são extraídos com n-hexano e o extrato obtido é purificado através de um sorvente (sílica gel ou óxido de alumínio) para remover materiais biogênicos. O solvente é evaporado e o resíduo resultante é pesado para determinação da concentração de TPH. Caso o contaminante do solo seja petróleo bruto ou graxa, a etapa de purificação não é realizada. O limite de detecção para TPH em solo é estimado em cerca de 50 mg kg⁻¹.⁵

Para laboratórios com poucos recursos, este método pode ser o mais ideal para análise de varredura de TPH no solo devido ao baixo investimento necessário.⁵⁷ Por outro lado, a determinação gravimétrica apresenta diversas desvantagens como: (i) não especificidade, pois não fornecem informações sobre a composição do resíduo obtido; (ii) potencial para falsos positivos ou negativos, dependendo da fração de hidrocarbonetos contida no solo; (iii) longo tempo necessário para a evaporação completa do solvente (pelo menos 60 min); e (iv) baixa capacidade extratora do n-hexano para hidrocarbonetos de alto peso molecular.^{5,46,57,87}

3.2.3.2. *Cromatografia gasosa*

Dentre as técnicas analíticas utilizadas para a análise de TPH em solo, a cromatografia gasosa com detector de ionização de chama (GC-FID) ou detector de espectrometria de massa (GC-MS) são as mais usuais em laboratórios devido à sensibilidade, seletividade (GC-MS), precisão, linearidade, reprodutibilidade e capacidade de abranger uma grande faixa de hidrocarbonetos.^{46,50,62,97} A cromatografia gasosa é uma técnica de separação baseada na distribuição de substâncias entre as fases estacionária e móvel. A separação da mistura é

realizada em colunas capilares contendo um sorvente específico fixado a um suporte. Durante o arraste da mistura por um gás inerte (hélio, hidrogênio ou nitrogênio), os compostos interagem com a fase estacionária, ocorrendo a separação destes através da combinação de vários fenômenos dependentes de diversos fatores como temperatura; fluxo da fase móvel; composição e estrutura da fase estacionária; e estrutura molecular dos analitos. Após a eluição completa dos compostos separados ou não pela coluna cromatográfica, os analitos atingem o detector, o qual gera um sinal de intensidade, intrínseco de cada tipo de detector, que é finalmente utilizado na obtenção do cromatograma. Neste, os compostos são representados por picos conforme seus tempos de retenção e, para relacionar a concentração dos analitos com os sinais instrumentais, são utilizadas as áreas dos picos integrados ou suas respectivas alturas.⁷⁸

A análise de TPH por GC-FID requer que a temperatura da coluna seja gradualmente aumentada para que os compostos sejam separados de acordo com seus pontos de ebulição, pois apresentam interações semelhantes com a fase estacionária.⁴⁶ Dado que os derivados do petróleo são constituídos de centenas de compostos, a separação destes é extremamente difícil e requer colunas cromatográficas de longo comprimento (100 m), o que torna o tempo de análise muito longo.⁷⁸ Desta forma, a quantificação de TPH é realizada pela somatória das áreas dos compostos eluídos, através de uma coluna capilar de 30 m, em determinada faixa de tempo de retenção. De acordo com o método EPA 8015,⁸ os hidrocarbonetos são separados em três faixas: (i) compostos da faixa da gasolina (GRO), determinados pelos tempos de retenção dos compostos 3-metilpentano e 1,2,4-trimetilbenzeno (C_6 a C_{10}); (ii) compostos da faixa do diesel (DRO), pelos tempos de retenção do decano e octacosano (C_{10} a C_{28}); e (iii) compostos da faixa dos óleos lubrificantes (ORO), que abrangem os tempos de retenção de C_{28} a C_{40} . Desta forma, o valor de TPH quantificado pelo método EPA 8015 corresponderá à faixa estabelecida pelos tempos de retenção dos compostos tidos como referência. Devido aos TPH serem constituídos por compostos voláteis e semivoláteis, são necessárias duas análises GC-FID para cada amostras a fim de abranger as três faixas de contaminantes, pois, uma vez que os compostos GRO são voláteis, estes devem ser extraídos pelos métodos headspace ou purge-and-trap, enquanto os compostos semivoláteis (DRO e ORO) podem ser extraídos pelos demais métodos, preferencialmente por Soxhlet, sonicação, fluido supercrítico ou micro-ondas.⁸ Apesar da necessidade de duas análises e dois métodos de extração por amostra, quando a fonte de contaminação é desconhecida, o método cromatográfico gasoso apresenta a maior abrangência e especificidade dentre os métodos convencionais, o que possibilita detectar a fonte de contaminação e avaliar o processo de degradação dos hidrocarbonetos.

Diversos detectores podem ser acoplados a cromatógrafos gasosos conforme sua aplicação. O detector tipo FID é o mais utilizado para a determinação de TPH devido à sua simplicidade, robustez e sensibilidade à hidrocarbonetos. Na câmara do detector, uma chama de hidrogênio é utilizada para a ionização dos compostos que eluíram pela coluna até o detector. As partículas ionizadas diminuem a resistência entre dois eletrodos, aumentando a corrente medida na proporção da concentração total do analito.⁵⁰ Esta mudança de corrente é registrada como sinal analítico e a unidade usual é milivolt. A relação entre a concentração e o sinal instrumental é conhecida através da análise de padrões analíticos com concentrações conhecidas, ou seja, pelo desenvolvimento de curvas analíticas. A quantificação das frações de TPH, segundo o método EPA 8015,⁸ pode ser realizada através de curvas analíticas desenvolvidas com materiais de referência certificados (MRC) que contenham os analitos utilizados na determinação dos tempos de retenção das diferentes faixas, ou desenvolvidas com amostras de combustíveis sem aditivos, por exemplo gasolina e diesel tipo A, no Brasil. Dependendo da matriz e o do método de extração, o método GC-FID pode alcançar limite de detecção de 10 mg kg⁻¹ para TPH em solo.⁵ Apesar de FID não ser um detector seletivo, é algumas vezes utilizado para determinação de PAH individuais e os limites de detecção são tipicamente 330 µg kg⁻¹ em solo,⁵⁰ porém a resolução cromatográfica deve ser aumentada consideravelmente de forma a evitar coeluições.

Outros detectores também utilizados na verificação de contaminação de solos por combustíveis são os detectores de fotoionização (PID) e espectrometria de massas. A aplicação de PID é específica para compostos com duplas ligações conjugadas, incluindo anéis aromáticos. Normalmente, este detector é utilizado em conjunto com o FID, pois apresenta alta sensibilidade para os compostos BTEX, chegando a limites de detecção próximos a 5 µg kg⁻¹ em solo, enquanto o sinal de TPH é detectado pelo FID. Apesar da EPA descrever um método para análise de BTEX por GC-PID, método EPA 8020,¹⁰⁹ este método é raramente utilizado devido à limitação de sua aplicação em amostras que contenham hidrocarbonetos de alto peso molecular, pois a sensibilidade do detector pode ser altamente prejudicada.⁵⁰ Para análise de hidrocarbonetos individuais, o detector frequentemente utilizado é o MS. De forma extremamente simplista, a espectrometria de massas é uma técnica analítica baseada na ionização de compostos seguida pela separação de íons pela a taxa de massa/carga e, enfim, a determinação da abundância relativa dos íons. Durante a ionização, os compostos usualmente se fragmentam em diversos íons e, devido a estabilidade química de cada fragmento, os tipos de íons formados e suas abundâncias tornam-se uma “impressão digital” do composto inicial. Desta maneira, a combinação da cromatografia gasosa com a espectrometria de massas

possibilita a separação de misturas e a identificação de compostos individuais de acordo com o padrão de fragmentação.⁵⁰ Melhoras significativas na determinação de PAH por GC-MS surgiram após a determinação pelo modo de monitoramento do íon selecionado (SIM mode). Através deste modo, os limites de detecção de PAH individuais são muitos inferiores em relação a outros métodos e variam entre 5 $\mu\text{g kg}^{-1}$ e 50 $\mu\text{g kg}^{-1}$ em solo, conforme o analito, o que torna este método o mais adequado para a determinação de PAH em monitoramentos ambientais.⁵⁰

Embora a cromatografia gasosa seja a técnica mais utilizada para a quantificação de TPH e suas frações devido às características apresentadas, esta técnica e os métodos utilizados apresentam algumas desvantagens: (i) altos custos de equipamentos, requisitos operacionais e manutenções preventivas e corretivas;^{46,59} (ii) necessidade de solventes apolares de baixo ponto de ebulição como diclorometano e hexano, os quais apresentam riscos à saúde humana e ao meio ambiente;⁹⁷ (iii) valores de desvio padrão relativo (DPR) costumam ser maiores que 30% para concentrações de TPH inferiores a 500 mg kg^{-1} ;¹¹⁰ (iv) a variância na reprodutibilidade é maior para grupos de compostos (TPH) que para compostos individuais;¹¹⁰ (v) o modo de integração influencia diretamente nos valores de LD e LQ e nos parâmetros da curva analítica;¹¹⁰ e (vi) longo tempo de análise cromatográfica (40 min), o que torna inadequado para análise de grande volume de amostra.⁴⁶

Inovações instrumentais, como o controle eletrônico de pressão da fase móvel, detectores FID de alta frequência e MS em tempo de voo, e softwares que facilitam a integração de cromatogramas e a revalidação de método, permitiram a substituição de colunas capilares convencionais por colunas de dimensões reduzidas a fim de obter um alto ganho no tempo de análise mantendo a mesma resolução entre os picos cromatográficos.^{93,111} A cromatografia gasosa ultrarrápida (UFGC) surgiu então com a utilização de colunas capilares de dimensões reduzidas (2-10 m de comprimento e 0,1-0,05 mm de diâmetro interno) e revestidas por uma resistência que permite taxas de aquecimento acima de 60 $^{\circ}\text{C min}^{-1}$.¹¹¹ O fluxo, além dos demais parâmetros cromatográficos, tem grande importância para compensar a perda de resolução.⁹³ Para manter a eficiência na separação dos compostos, o gás hidrogênio é utilizado como fase móvel devido à menor altura de pratos teóricos proporcionada por uma larga faixa de velocidade linear do gás de arraste.¹¹² Embora a UFGC já tenha sido utilizada para a separação de hidrocarbonetos voláteis ($\text{C}_5\text{-C}_8$)¹¹³ e análise de gasolina,^{114,115} encontra-se atualmente na literatura apenas um trabalho sobre sua aplicação em análise ambiental de solos.⁹¹ Robbart Jr. (2001) desenvolveu um método para a análise de PAH, extraídos do solo através de um penetrômetro de cone de extração térmica (TECP) e determinados por UFGC-MS. De acordo com o método desenvolvido, a análise cromatográfica é realizada em 5 min através de uma

coluna Restek's MXT-5 (15 m de comprimento, 0,28 mm d.i. e 0,5 μm de espessura de filme de 5% difenil/95% dimetilpolisiloxano). Os picos não resolvidos foram deconvoluídos para determinação dos PAH individuais.⁹¹

Nas últimas décadas, uma variedade de sistemas GC-MS portáteis vem sendo desenvolvidos. Seu peso pode variar entre 16 kg a 60 kg e, para alguns modelos, as análises podem ser realizadas em menos de 10 min. Contudo, relata-se que a portabilidade prejudica principalmente a sensibilidade do método, além do poder de separação e identificação dos compostos.⁴⁶

3.2.3.3. *Cromatografia líquida de alta eficiência*

A cromatografia líquida de alta eficiência (HPLC) é a técnica analítica cromatográfica preferida para o estudo de compostos de alto ponto de ebulição, incluindo PAH, pois não requer o uso de alta temperatura, evitando a degradação de compostos termoinstáveis. As características de HPLC envolvem o uso de alta pressão; uma grande variedade de fases estacionárias, constituídas de sorventes de tamanho de partícula entre 1,8-5 μm ; e a possibilidade de gradiente de eluição; características que permitem a separação de misturas complexas de forma eficiente e rápida. Três tipos de detectores são comumente utilizados para análise de PAH: fluorescência (FLD); UV-vis; e arranjo de diodos (DAD). O limite de detecção para PAH em extratos de solo pode variar entre 0,01 $\mu\text{g L}^{-1}$ e 0,1 $\mu\text{g L}^{-1}$, quando utilizados os detectores FLD/UV ou FLD/DAD.⁵⁰

As desvantagens da determinação de PAH por HPLC encontram-se no elevado custo dos equipamentos, manutenção e operação; na necessidade de purificação dos extratos; ajuste de diversos parâmetros operacionais; coeluição de isômeros; e o longo tempo de análise.⁵⁰

3.2.3.4. *Espectroscopia no infravermelho médio*

Apesar da determinação de TPH através da espectroscopia no infravermelho médio (MIR) por métodos convencionais não ser específica, é muito utilizada devido à rapidez e ao baixo custo de análise.⁵⁰ A radiação infravermelha (IR) abrange a faixa do espectro eletromagnético entre os comprimentos de onda de 0,78 mm e 1,00 mm (número de onda entre 4000-400 cm^{-1}). A espectroscopia no infravermelho baseia-se na absorção de radiação infravermelha por moléculas que convertem a radiação em energia de vibração molecular. A

frequência de uma absorção depende das massas relativas dos átomos ligados, das constantes de força das ligações e da geometria dos átomos, deste modo, cada frequência de vibração pode ser associada a um tipo específico de ligação química.¹¹⁶ A radiação IR absorvida pelas moléculas resulta em vibrações de deformações angulares e axiais (estiramentos) e somente vibrações que levam à alteração do momento dipolo da molécula são observadas no espectro de infravermelho. Desta forma, são observadas no espectro MIR a vibrações de ligações moleculares como O-H, C-H, N-H, C=O, C-N, N-O ou C=C. Assim, o estiramento de algumas ligações simétricas, nas quais os elétrons são igualmente compartilhados, não dá origem a vibrações ativas no infravermelho, por exemplo, o estiramento em moléculas N₂.³³

O número teórico de frequências de absorção raramente é observado, pois as bandas correspondentes são acompanhadas de vibrações harmônicas (múltiplos de uma dada frequência) e de combinações de frequências de absorção.¹¹⁶ Desta forma, mesmo compostos de estruturas simples apresentam diversos sinais analíticos, o que torna o espectro MIR de uma substância pura único, ou uma impressão digital do composto (fingerprint).³³

Os métodos MIR convencionais para determinação de TPH são baseados na relação entre a concentração de hidrocarbonetos extraídos por um solvente aprótico e os sinais das bandas de absorção das ligações C-H dos grupos CH₃ ou CH₂. O procedimento do primeiro método MIR para análise de TPH em solo (método EPA 418) foi constituído de três etapas: extração dos contaminantes por agitação utilizando o solvente 1,1,2-tricloro-1,2,2-trifluoroetano (Freon 113); purificação em uma coluna de sílica gel; e quantificação dos hidrocarbonetos através da banda de CH₃ em 2950 cm⁻¹.⁵ Devido aos riscos ambientais relacionados ao uso de Freon, o método EPA 418 foi modificado (método EPA 418.1) e o solvente foi substituído por tetraclorometano, pois também não apresenta sinais no número de onda 2950 cm⁻¹.⁹ Posteriormente, a EPA desenvolveu o método EPA 8440, no qual os hidrocarbonetos são extraídos por fluído supercrítico (CO₂), sem causar riscos ao meio ambiente, porém o método continuou a ser não específico, visto que qualquer composto com absorção em 2950 cm⁻¹ é quantificado como TPH e, além disso, sua aplicação é inadequada para análise de compostos voláteis. Como uma alternativa para determinação de TPH em águas subterrâneas, a ASTM desenvolveu o método ASTM D7678, similar ao EPA 418.1, com as principais diferenças no solvente utilizado (ciclohexano) e na banda de absorção utilizada na quantificação (deformação angular do grupo CH₂ em 1370-1380 cm⁻¹).¹² Apesar das alterações, o método ASTM continuou a apresentar a não especificidade e, além disso, o solvente ciclohexano apresenta menor capacidade extratora para TPH que outros solventes, como diclorometano e hexano.⁷⁴

A abordagem univariada na determinação de TPH por espectroscopia IR tornam os espectros ricos em informação em sinais analíticos não seletivos para as diferentes frações de TPH; e proporcionam tendências tanto positivas como negativas nos valores quantificados. As tendências negativas podem ser resultantes desde a baixa eficiência na extração; perda de hidrocarbonetos voláteis; remoção de PAH com 5 ou 6 anéis aromáticos durante a purificação pela sílica gel; até as diferenças de absorvidade molar entre os padrões de calibração e os contaminantes na matriz.^{38,86} Já a tendência positiva é normalmente causada pela quantificação de matéria orgânica não removida no processo de purificação.⁴⁵ Para a atenuar a tendência negativa na determinação de TPH, alguns pesquisadores avaliaram a extração de hidrocarbonetos por sonicação previamente à análise MIR. Embora a eficiência da extração tenha sido otimizada, os solventes avaliados foram Freon 113 e tetracloroetano.^{87,117,118}

Nos métodos convencionais, o modo de aquisição dos espectros é realizado através da transmitância, ou seja, o feixe passa através da amostra em uma cubeta e segue para o detector. Neste modo, o desvio da radiação é mínimo e, portanto, a sensibilidade da técnica é grande, chegando a limites de detecção de 10 mg kg⁻¹ para TPH em solo. Outros modos de medição ocorrem através da transfletância, onde a radiação IR passa pela amostra, é refletida em um espelho e passa novamente pela amostra antes de chegar ao detector;¹⁸ e da refletância total atenuada (ATR), na qual ocorrem diversas reflexões da radiação IR na amostra através de um cristal com alto índice de refração.¹¹⁹ Em ambos os modos de medição, é possível utilizar um pequeno volume de amostra sem grandes perdas na sensibilidade e o processo de ambientação entre as análises é simplificado. Apesar da praticidade e eficiência da técnica ATR, ainda há a necessidade de etapa de extração por solvente e poucos trabalhos são encontrados na literatura sobre análise de TPH em água ou solo.^{119,120}

A espectroscopia no infravermelho teve imenso avanço com o desenvolvimento de espectrômetros capazes de medir em poucos segundos uma grande faixa espectral através de interferômetros e converter o sinal obtido para espectros IR através de transformadas de Fourier (FT).¹¹⁶ Além dos avanços na velocidade, precisão e resolução proporcionada por estes espectrômetros FT-IR; o desenvolvimento instrumental para otimizar a sensibilidade de análises realizadas por refletância difusa, onde o feixe de radiação chega à amostra e é refletido de forma difusa, tornou possível a análise de propriedades físico-químicas e composicionais de sólidos sem a necessidade de preparo de amostra.¹²¹ Embora a espectroscopia de refletância difusa na região do infravermelho com transformada de Fourier (DRIFT) tenha sido muito estudada e aplicada com sucesso na determinação de propriedades do solo,⁶² seu uso na quantificação de TPH ainda é muito pouco avaliado. Os recentes estudos realizados mostraram

uma baixa exatidão devido às diferentes texturas dos solos e à distribuição heterogênea dos hidrocarbonetos no solo.^{22,23}

3.2.3.5. Metodologias para determinação de BTEX e PAH em campo

Além das metodologias explanadas anteriormente, é importante citar os métodos de determinação de PAH em campo por imunoenaios, espectroscopia de fluorescência e espectroscopia Raman. Os imunoenaios apresentam a praticidade da portabilidade e são baseados na ligação química seletiva entre anticorpos e constituintes específicos do petróleo. Os princípios dos métodos de imunoenaios variam conforme a metodologia utilizada, que podem ser através do uso de enzimas que, ao se ligarem a compostos PAH, alteram a densidade ótica da solução; ou através de reagentes marcadores fluorescentes; ou marcadores electrocatalíticos. Os imunoenaios são ideais para análise de grandes quantidades de amostras, porém são afetados pela textura do solo e tempo de contaminação.^{5,46}

A espectroscopia de fluorescência é uma técnica eficiente para análise quantitativa e qualitativa de compostos aromáticos em solo por apresentarem propriedade luminescente através da excitação por radiação ultravioleta (UV). A determinação de BTEX e PAH através desta técnica apresenta várias vantagens como limite de detecção próximo a 1 mg kg^{-1} ; não destrutivo; reprodutivo; rapidez; seletividade para compostos aromáticos; e portabilidade; porém, não é aplicável a hidrocarbonetos alifáticos.⁵⁰

Métodos quantitativos para PAH em solo têm sido desenvolvidos através da espectroscopia Raman. Esta espectroscopia vibracional é baseada no espalhamento inelástico da luz pelo material analisado a partir uma fonte de luz monocromática. A rapidez e a portabilidade dos instrumentos são grandes vantagens proporcionadas pela técnica, entretanto, grandes cuidados devem ser considerados a fim de obter boa repetibilidade, pois a precisão da técnica é altamente afetada pelo direcionamento da fonte de luz e por fontes externas de fluorescência.⁴⁶

3.2.3.6. Espectroscopia nas regiões do infravermelho próximo e visível

Frequentemente, as regiões do infravermelho próximo ($4000-12.500\text{ cm}^{-1}$; $700-2500\text{ nm}$) e visível ($12.500-15.000\text{ cm}^{-1}$; $400-750\text{ nm}$) são abrangidas por um mesmo espectrômetro pelo uso comum da fonte de radiação (lâmpadas de tungstênio ou halogênio) e do detector (usualmente, InGaAs). Inicialmente aplicada na agricultura durante a década de 60, a espectroscopia no infravermelho próximo (NIR) baseia-se na absorção de radiação infravermelha por moléculas que convertem a energia eletromagnética em energia de vibração molecular, assim como na região MIR.¹¹⁶ Todavia, as bandas de absorção na região NIR são resultantes de combinações ou sobretons das vibrações fundamentais na faixa do infravermelho médio, podendo ser resultantes da combinação de duas ou mais bandas de transição fundamental. Desta forma, os espectros NIR são pouco úteis na elucidação de estruturas moleculares, contudo, muito aplicáveis em determinações quantitativas através de análises multivariadas.¹²²

A natureza da espectroscopia vibracional pode ser explicada através do modelo do oscilador anarmônico. Neste modelo, considera-se que dois átomos ligados a uma distância de equilíbrio terão a menor energia potencial no estado fundamental. Quando o estado vibracional é aumentado, a energia potencial aumenta de forma quantizada e os níveis energéticos não são equidistantes. A energia potencial é dependente da força de ligação e da massa reduzida dos átomos envolvidos. Portanto, quanto maior a diferença entre as massas dos átomos, maior a anarmocidade e a intensidade do sinal analítico.¹²¹ A região NIR é, principalmente, sensível às ligações entre os grupos de átomos C-H, N-H, O-H e S-H.¹⁸ A transição entre o estado vibracional fundamental e o primeiro nível energético é chamada de transição fundamental e são observadas na região MIR, enquanto as transições para os níveis subsequentes são chamadas de sobretons e suas intensidades diminuem com o aumento dos níveis energéticos. Diferentemente do espectro MIR, o espectro NIR é resultado quase exclusivamente de sobretons e bandas de combinação, o que dificulta a interpretação de espectros devido a sobreposição das bandas, além de outros fenômenos como as ressonâncias de Fermi e de Darling-Denninson.¹²¹ A complexidade e a riqueza de informação dos espectros NIR tornam essencial o uso da quimiometria para extrair as informações úteis e correlacionar com as propriedades de interesse.^{123,124}

A luz visível (vis) abrange apenas uma pequena parte da radiação eletromagnética na faixa de comprimento de onda entre $0,4$ e $0,78\text{ mm}$. A interação da radiação visível com o solo produz transição eletrônicas nas moléculas, principalmente devido a processos conhecidos

como efeito de campo cristalino e transferência de carga. Geralmente, as transições eletrônicas produzem bandas de absorção amplas que controlam a cor do material capturada pelo olho humano. Embora a resposta espectral na região do visível seja fraca, é possível obter informações quantitativas a partir dos espectros.³³

Nas últimas duas décadas, a aplicação de espectroscopia NIR e vis-NIR na determinação de parâmetros do solo e monitoramento de sua qualidade tem sido cada vez mais estudada e utilizada.^{32,125,126} Dentre os parâmetros do solo correlacionados com os espectros vis-NIR estão nitrogênio total (TN), carbono orgânico (SOC), teor de água, textura, capacidade de troca de cátions, pH, atividade microbiana, entre outros.^{124,127-129} Além da determinação dos parâmetros por modelos de calibração multivariada, a espectroscopia NIR também foi aplicada na classificação de solos conforme suas classificações e regiões.^{27,130}

A intensa busca pela aplicação da espectroscopia vis-NIR na análise de solos está embasada na ausência ou mínimo preparo de amostra, rapidez, baixo custo, opção de portabilidade e, principalmente, na sensibilidade por compostos com grupos C-H, N-H e O-H.^{131,132} Assim como na espectroscopia MIR, os espectros podem ser obtidos pelos modos de transmitância e de refletância difusa, o qual é mais aplicado para análise em campo por dispensar o preparo de amostra.¹³³ Diferentemente do modo de transmitância, a análise por refletância difusa não obedece à lei de Beer-Lambert devido ao efeito de espalhamento da radiação incidida na amostra. As mudanças na direção dos fótons não dependem somente do tamanho e forma das partículas do sólido, como também de seu comprimento de onda, da direção da luz incidente e das mudanças do índice de refração. O espalhamento tem impacto direto na absorvância, pois, quanto mais fótons são espalhados, mais absorvidos ou dissipados pelo meio eles serão.^{92,134} Na análise de solos, o espalhamento tem efeito direto na intensidade média do espectro, ou seja, quanto maior o espalhamento, menos radiação refletida chega ao detector.

Para ressaltar a importância dada atualmente à técnica vis-NIR para análise de solos, uma pesquisa sem precedentes foi realizada por Rossel et al. (2016) junto a 45 cientistas de 35 instituições para a criação de uma biblioteca de espectros vis-NIR e parâmetros físico-químicos de 23.631 amostras de solo, abrangendo 92 países e 7 continentes.²⁶ Para isso, foi criado um protocolo para a aquisição dos espectros e as informações físico-químicas e composicionais junto aos espectros foram transmitidas ao banco de dados. Após os tratamentos matemáticos necessários para padronização dos espectros, desenvolveram modelos de regressão para diversos parâmetros. Os coeficientes de correlação dos parâmetros determinados variaram de 0,68 (teor de areia) a 0,92 (SOC), o que mostra ser muito promissor a aplicação de vis-NIR

utilizando uma biblioteca global de espectros.²⁶ Outro grande avanço para análise de solos foi apresentado recentemente pelo desenvolvimento de um sistema de amostragem (penetrômetro) acoplado a um espectrômetro vis-NIR, o qual permite a determinação da propriedade do solo sem a necessidade de coleta ou preparo de amostra.¹³⁵

As limitações e desvantagens da técnica vis-NIR encontram-se muitas vezes na baixa intensidade das bandas relacionadas à propriedade de interesse e às interferências proporcionadas pelo solo, como textura, matéria orgânica e, principalmente, umidade.⁴⁶ Desvantagens como a necessidade da criação de um banco de dados local podem ser superadas através da transferência de calibração, utilizada para compatibilizar dois bancos de dados por meio de algoritmos, evitando novos esforços desnecessários.¹²⁴

Apesar dos grandes esforços no desenvolvimento de métodos NIR ou vis-NIR para acessar a qualidade do solo de forma rápida e de baixo custo, o número de trabalhos voltados a análises de TPH, BTEX e PAH em solo ainda é relativamente pequeno.^{62,97} Grande parte dos trabalhos envolvem conjuntos amostrais pouco representativos, número limitado de amostras e são avaliados para concentrações acima dos limites de intervenção.^{10,62,132} De acordo com Cozzolino (2015),²⁷ o número pequeno de publicações nesta área de pesquisa pode estar relacionado à falta de conhecimento e compreensão sobre esta técnica e fundamentos da quimiometria.

Na Tabela 2, encontram-se os estudos na literatura direcionados à determinação de hidrocarbonetos do petróleo através da técnica NIR ou em conjunto com outras regiões do espectro eletromagnético. Os primeiros estudos foram realizados na década de 90¹³⁶⁻¹³⁸ e, somente após 11 anos, foram publicados novos estudos relacionados à contaminação de solos por TPH. Os possíveis motivos deste gap podem estar relacionados ao desenvolvimento de espectrômetros portáteis e no aprimoramento e redução de custo de espectrômetros vis-NIR durante os anos 2000.¹³⁹ Os primeiros estudos mostraram uma forte tendência para aplicação de NIR na quantificação de gasolina, diesel e óleo lubrificante em solo, sem preparo de amostra, contudo, os modelos desenvolvidos foram pouco representativos, utilizando apenas um tipo de solo e de cada contaminante e, principalmente, um número pequeno de amostra. Um conjunto pouco representativo pode levar tanto a modelos com ajustes superestimados como subestimados.

Tabela 2 – Estudos na literatura relacionados à determinação de TPH de em solo envolvendo a espectroscopia NIR.

Referência	Técnica	Compostos de interesse	Contaminação	nº de amostras	Faixa de concentração (mg kg ⁻¹)	Método de regressão	Pré-processamento espectral	Resultados
Stallard et al (1996) Ref. ¹³⁶	NIR	TPH	Fort. com óleo lubrificante	17	0 - 60.000	PLS	-	R ² cv (0,99)
Zwanziger e Forster (1998) Ref. ¹³⁷	NIR	TPH	Fort. com gasolina e diesel	2 conj. de 25 amostras	1000 - 50.000	PLS	C.M.	R ² cv (0,99)
Malley et al (1999) Ref. ¹³⁸	NIR	TPH	Diesel	26	0,001 - 6,26	PLS e MLR	1ª der. + s.	R ² val (0,68 - 0,72); RPD (1,76 - 1,82)
Chackraborty et al (2010) Ref. ¹⁴⁰	vis-NIR	TPH	Petróleo	46	43,3 - 48.188,0	PLS e BRT	1ª der. + s.	R ² val (0,64); RPD (1,70)
Forrester et al (2010) Ref. ¹⁴¹	FT-NIR	TPH	Fort. com soluções de diesel em ciclohexano	172	0 - 100.000	PLS	-	R ² cv (0,81)
Sorak et al (2012) Ref. ¹⁴²	NIR portátil	TPH	Fort. com soluções de diesel em DCM	70	0 - 70.000	PLS	C.B.; eH ₂ O	R ² cv (0,96)
	NIR portátil	TPH	Fort. com soluções de gasolina, diesel e petróleo em DCM	34	-	PCA	C.B.; eH ₂ O	Discriminação conforme contaminação
Chackraborty et al (2012) Ref. ¹⁴³	vis-NIR-SWIR	TPH	Fort. com diesel, óleo lub. e petróleo	52	4000 - 28.000	PLS e MLR	1ª der.	R ² val (0,84 - 0,94); RPD (2,50 - 3,97)
	vis-NIR-SWIR	TPH	Fort. com diesel, óleo lub. e petróleo	45	1000 - 30.000	LDA, SVM e RF	1ª der.	33% a 40% de classificações corretas
Forrester et al (2013) Ref. ²³	FT-NIR	TPH	Fort. com diesel	199	0 - 15.000	PLS	s.v.m.	R ² cv (0,84); RPDcv (2,4)
Okparanma e Mouazen (2013) Ref. ⁶⁹	vis-NIR	fenantreno	Fort. com diesel	150	0,58 - 2,49	PLS	1ª der.	R ² val (0,83); RPD (2,32)
Schwartz et al (2013) Ref. ¹⁰	vis-NIR-SWIR	TPH	Fort. com gasolina, diesel e querosene	15 conj. de 48 amostras	100 - 5000	PLS e PCR	1ª der., MSC, s. e c.r.	conjuntos separados: R ² val (0,92 - 0,99); RPD (4 - 23)

Tabela 2 – Estudos na literatura relacionados à determinação de TPH de em solo envolvendo a espectroscopia NIR. (continuação).

Referência	Técnica	Compostos de Interesse	Contaminação	nº de amostras	Faixa de concentração (mg kg ⁻¹)	Método de regressão	Pré-processamento espectral	Resultados
Okparanma e Mouazen (2013) Ref. ²⁵	vis-NIR	PAH	Fort. com diesel	150	1,083 -2,86	PLS	N + 1ª der. + s.	R ² val (0,89); RPD (2,75)
	vis-NIR	Diesel	Fort. com diesel	150	30.000 - 150.000	PCA	N + 1ª der. + s.	Discriminação conforme contaminação
Okparanma et al (2014) Ref. ⁹⁶	vis-NIR	PAH	Fort. com padrões de PAH	137	0 - 17	PLS	N + 1ª der. + s.	R ² val (0,89); RPD (3,12)
Chackraborty et al (2015) Ref. ⁵⁹	vis-NIR	TPH	Petróleo	108	0 - 3.260.000	PLS, PSR e RF	1ª der. + s	R ² val (0,73); RPD (1,96)
Webster et al (2016) Ref. ²²	FT-MIR-NIR	TPH	Diesel	67	0 - 60.000	PLS	C.B.	R ² val (0,99); RPD (8 - 13)
Douglas et al (2018a) Ref. ⁹⁸	vis-NIR portátil	Alcanos	Petróleo	85	11,54 - 767,00	PLS e RF	N + 1ª der. + s.	R ² val (0,36 - 0,58); RPD (1,29 - 1,59)
	vis-NIR portátil	PAH	Petróleo	85	0,175 - 321,8	PLS e RF	N + 1ª der. + s.	R ² val (0,56 - 0,71); RPD (1,55 - 1,99)
Douglas et al (2018b) Ref. ⁹⁷	vis-NIR portátil	TPH	Petróleo	85	16 - 666	PLS e RF	N + 1ª der. + s.	R ² val (0,54 - 0,68); RPD (1,51 - 1,85)

Abreviações: Fort. = fortificação; C.M. = centramento na média; 1ª der. = primeira derivada; s. = suavização; N. = normalização; C.B. = correção de linha base por algoritmo polinomial; eH₂O = exclusão da banda da água (5448–4686 cm⁻¹); s.v.m. = seleção de variáveis manual; c.r. = continuum removal; PLS = mínimos quadrados parciais; PCR = regressão por componentes principais; RF = floresta aleatória; PSR = regressão penalizada por splines; BRT = árvore de regressão amplificada; PCA = análise de componentes principais; e LDA = análise discriminante linear.

Na primeira metade dos anos de 2010, a maior parte dos estudos foram baseados na fortificação de amostras com o contaminante de interesse, porém, muitas vezes os modelos foram desenvolvidos com poucas amostras^{10,140,142,143} ou validados apenas com amostras do conjunto de calibração (validação cruzada).^{23,141,142} Além disso, alguns trabalharam com um intervalo de medição muito superior aos limites de intervenção²⁵ ou com concentrações distribuídas de forma assimétrica,¹⁴³ o que pode levar a amostras com alta influência (leverage) na modelagem e tornar a predição de amostras desconhecidas tendenciosa.⁶² Os trabalhos que apresentaram conjuntos amostrais representativos e melhores correlações entre os valores preditos e os valores medidos, de forma representativa, foram os de Okparanma e Mouazen (2013) e Okparanma et al (2014) na quantificação de fenantreno e PAH total na região vis-NIR.^{25,69}

Além das análises quantitativas, três trabalhos empregaram análises discriminantes por métodos supervisionados¹⁴³ e não supervisionados.^{25,142} De acordo com o método desenvolvido por Chakraborty e colaboradores (2012)¹⁴³, foi possível discriminar corretamente 100% das amostras contaminadas por petróleo das amostras não contaminadas.

Os estudos realizados na segunda metade da mesma década foram mais representativos, ou seja, com maior número de amostras, as quais não foram fortificadas; e utilizaram conjunto de amostra externo para a validação. Exceto pelo trabalho de Webster et al (2016)²², no qual o conjunto amostral utilizado não possuiu concentrações com distribuição simétrica, os demais estudos apresentaram baixo coeficiente de determinação (R^2) e desvio percentual relativo (RPD) em decorrência de diversos fatores como: maiores interferências no solo heterogêneo e úmido; diferenças composicionais dos contaminantes para diferentes amostras; e uso de instrumentação portátil, a qual possui menor sensibilidade e resolução.

Observa-se na Tabela 2, o pequeno número de pré-processamentos avaliados diante de uma grande quantidade de tratamentos matemáticos que podem auxiliar na redução de ruídos espectrais, aumentando a razão sinal-ruído do método analítico. Além de pré-processamentos, algoritmos de seleção de variáveis têm se mostrado muito úteis na análise de misturas complexas, pois podem aumentar a sensibilidade, a seletividade e a relação linear entre as variáveis espectrais e a propriedade de interesse.¹⁴⁴

A maior barreira ao acesso às informações do solo através de análises NIR encontra-se na superação das interferências referentes à natureza de cada amostra de solo. Como o solo é uma mistura complexa de materiais, é difícil atribuir características específicas do espectro a componentes químicos específicos.²⁸ Além disso, devido ao efeito de espalhamento, a parte útil dos espectros de absorção, relacionada à concentração elementar, é relativamente pequena

quando comparada ao efeito de espalhamento.⁹² Várias propriedades do solo estão sujeitas à variação no espaço e tempo, como umidade, rugosidade do solo (irregularidades na superfície do solo), textura, assim como o teor de matérias orgânicas originados de vegetação e culturas de microrganismos.^{31,145}

A presença de matéria orgânica pode proporcionar tendências positivas aos valores preditos, assim como em métodos MIR, ou até mesmo prejudicar a representatividade do conjunto de calibração se houver grandes diferenças nos teores de matéria orgânica entre as amostras de calibração e as amostras “desconhecidas”.^{28,146} Além da matéria orgânica, os minerais no solo têm a capacidade de alterar o espalhamento da radiação incidente proporcionando alterações no espectro obtido.¹⁴⁷

De acordo com vários autores, a umidade tem sido relatada como a principal interferência na predição de diversos parâmetros do solo.^{22,124,126,148} Devido à massa reduzida proporcionada pelo grupo O-H, as bandas relacionadas à água são intensas e acabam sobrepondo bandas de interesse na região NIR. A presença de água em alta concentração (acima de 15% da massa) diminui o efeito de espalhamento, propagando a radiação mais profundamente no solo. A redução no sinal médio do espectro não é linear e é independente do tipo de solo.^{22,126} O efeito da umidade se intensifica em solos argilosos em decorrência da maior superfície de contato para adsorção de água pelas partículas do solo.¹²⁴

Além da influência do teor de argila na intensificação da interferência da água devido à sua propriedade higroscópica,¹⁴⁹ a textura do solo tem efeito direto no espalhamento da radiação incidente sobre o solo analisado, alterando toda faixa espectral NIR adquirida por refletância difusa.²⁸ Agregados de argila causam um efeito de escudo sobre a radiação, diminuindo a intensidade das bandas referentes aos compostos TPH, enquanto em solos arenosos, as moléculas de TPH formam filmes entre as partículas de areia, proporcionando sinais de TPH mais intensos no espectro.²² A profundidade, na qual o solo foi coletado, é outro fator que pode influenciar na quantificação dos hidrocarbonetos, pois diferentes camadas de solo frequentemente exibem propriedades físicas, químicas e biológicas distintas. Por exemplo, as camadas superficiais dos solos costumam apresentar maior teor de matéria orgânica.¹²⁷

Uma forma de eliminar interferentes, como água e variáveis granulométricas, é através do preparo de amostras, porém deve ser realizado com cautela, seguindo protocolos, para não introduzir mais erros aos modelos de predição.¹²⁴ Duas estratégias comumente utilizadas para reduzir os interferentes são por meio da secagem prévia das amostras; e pela separação granulométrica das frações do solo; ou ambas em conjunto. Embora, estas estratégias sejam

úteis para a determinação de carbono orgânico (SOC) e nitrogênio total (TN),¹⁵⁰ além de outras propriedades do solo, seu uso é limitado às frações TPH de alto peso molecular.²² Uma abordagem alternativa para aprimorar os ajustes de modelos para predição de SOC foi apresentada por Nocita et al (2013),¹⁵¹ na qual as amostras foram divididas conforme seu teor de água para a modelagem, contudo, isto requer a secagem das amostras posteriormente às aquisições espectrais, levando cerca de 24 h de acordo com procedimentos bem estabelecidos.¹⁵²

Além das interferências citadas, a capacidade preditiva dos modelos é altamente influenciada pelas variabilidades instrumentais e pela forma de aquisição dos espectros. É de extrema importância verificar o posicionamento da fonte de radiação, a qualidade do background e a presença de ruídos instrumentais, assim como o posicionamento da amostra no espectrômetro ou vice-versa.¹⁵³

Visto a grande quantidade de variações que podem interferir na qualidade dos modelos de predição, é essencial realizar medições em replicatas, utilizando a média espectral de diversas varreduras, para reduzir os efeitos da heterogeneidade do solo e do espalhamento difuso da radiação.²² Ainda que o conjunto amostral seja representativo para os tipos de solo, a robustez dos modelos também é dependente da representatividade dos contaminantes, uma vez que diferentes hidrocarbonetos apresentam diferentes absorvidades molares e sinais espectrais.⁸⁶

Várias formas de contornar as interferências na predição de propriedades do solo por análises NIR vêm sendo desenvolvidas através de métodos quimiométricos, como pré-processamentos por filtros multivariados, métodos de regressão não lineares e seleção de variáveis. Estas abordagens serão apresentadas nas próximas seções.

3.3. Quimiometria

3.3.2. Importância da quimiometria na utilização da espectroscopia NIR

Embora a faixa espectral do infravermelho próximo tenha sido descoberta primeiro que a região do infravermelho médio, a aplicação da espectroscopia NIR teve seu crescimento após o trabalho desenvolvido por Karl Norris, em 1965, devido à necessidade de métodos estatísticos multivariados para correlacionar as informações ricas, porém complexas, dos espectros NIR com a propriedade de interesse.¹²¹ A quimiometria, disciplina que envolve modelos estatísticos

voltados ao tratamento de dados químicos, tem o papel de extrair informações de dados multivariados, muitas vezes complexos, em informações úteis e aplicá-las na predição de propriedades de interesse. Através da quimiometria é possível realizar análises qualitativas por métodos de reconhecimento de padrão ou análises exploratórias; quantitativas por métodos de regressão multivariada; e otimização de processos através do planejamento de experimentos.

Com o avanço computacional, a análise de dados constituídos de um grande número de amostras e variáveis tornou-se mais prática e permitiu a popularização da espectroscopia NIR.¹³⁹ Diferentemente dos métodos MIR univariados para análise de TPH, os métodos multivariados permitem o uso de várias bandas espectrais, diminuindo a possibilidade de quantificações errôneas devido à presença de matéria orgânica, e eliminam a necessidade do uso de solventes apróticos, como Freon 113 e tetraclorometano.¹²⁰ Além disso, quando utilizado um conjunto amostral representativo em relação às variáveis naturais do solo e aos tipos de contaminantes, a análise multivariada pode superar as limitações da técnica NIR causadas pelos interferentes. Para isto, as variações externas devem ser controladas e as abordagens quimiométricas devem ser aplicadas e validadas da forma mais adequada possível.

3.3.2. Métodos de regressão multivariada

Como pode ser observado na Tabela 2, o método de regressão por mínimos quadrados parciais (PLS) é o método mais utilizado no desenvolvimento de modelos de predição de TPH e PAH e também na determinação de parâmetros do solo.²⁶ A eficiência e a facilidade da aplicação e da interpretação dos resultados gerados pelo método PLS o tornam o método mais comum em diversas áreas de pesquisa.^{92,154}

Com o desenvolvimento instrumental dos espectrômetros, a resolução espectral teve grande aumento e, conseqüentemente, aumentou o número de variáveis. Mais que na região MIR, as bandas espectrais no NIR são constituídas por diversas variáveis e um mesmo grupo funcional costuma apresentar mais de um sinal no espectro em decorrência de sobretons e bandas de combinação. Desta forma, grande parte das n variáveis espectrais (x) são colineares, ou seja, apresentam a mesma informação química e, frequentemente, apenas poucas delas contém as informações relacionadas à propriedade de interesse.⁹² Métodos de regressão multivariada, como PCR (regressão por componentes principais) e PLS, surgem então com o papel de comprimir as n dimensões do espectro em variáveis não colineares (ortogonais) e, assim, extrair informações úteis da matriz de dados. Diferentemente do método PCR, as variáveis latentes obtidas pelo método PLS são construídas de forma a incluir o(s) vetor(es) y

durante a modelagem, portanto, determinam coeficientes de regressão a partir da correlação entre a matriz \mathbf{X} e o vetor \mathbf{y} (propriedade de interesse). A variância explicada nos blocos \mathbf{X} e \mathbf{y} é aumentada com a introdução de novas variáveis latentes (VLs), contudo, acréscimo de VLs deve ser realizado com cautela para evitar adição de variáveis não informativas e proporcionar um sobreajuste do modelo de predição.¹⁰ Usualmente, o número de variáveis latentes é determinado na etapa de validação cruzada, na qual amostras do conjunto de calibração são retiradas da modelagem e testadas no modelo desenvolvidos. A validação cruzada pode ser executada retirando-se uma amostra por vez ou em blocos de amostras. Muitas vezes, os resultados obtidos a partir da validação cruzada são superestimados devido ao uso das mesmas amostras na calibração, portanto, é essencial a verificação da capacidade preditiva do modelo com um conjunto amostral externo ao da modelagem.

Outro importante fator na habilidade preditiva dos modelos é o número de amostras utilizado na modelagem. O número necessário irá variar de acordo com a complexidade da amostra, número de variáveis explicativas e a quantidade de interferentes presentes no tipo de amostra.¹²⁴ Por exemplo, segundo Ramirez e colaboradores (2014),¹⁴⁹ o número mínimo de amostras para o desenvolvimento de um modelo robusto para determinação de teor de argila pela técnica vis-NIR foi definido como 200. A escolha das amostras para constituir o conjunto de calibração é uma importante etapa para tornar o modelo representativo. Normalmente, o conjunto amostral é dividido em 67% para calibração e 33% para validação e muitos trabalhos têm utilizado algoritmos, como Kennard-Stone, para separar os conjuntos amostrais de forma representativa.^{98,125,133,149}

Apesar das vantagens do método PLS, como ampla aplicabilidade, simplicidade, rapidez e praticidade, ele também apresenta algumas limitações. O principal requerimento de sua aplicação baseia-se na relação linear entre as variáveis x e a propriedade de interesse. Como já citado anteriormente, as diversas interferências provenientes dos componentes do solo não apresentam uma relação linear com os espectros NIR adquiridos por refletância difusa e, conseqüentemente, a exatidão do modelo PLS é prejudicada pelas interferências. A adição de uma etapa de extração com solvente pode contornar os efeitos dos interferentes, proporcionando uma relação linear entre as variáveis espectrais e a concentração dos analitos, a qual obedece a lei de Beer-Lambert nos modos de transmitância ou transfletância.¹²⁰ Outros fatores que influenciam negativamente os modelos PLS são a heterocedasticidade dos resíduos (a variância não é constante em todo intervalo de medição) e a distribuição assimétrica dos valores y , pois valores extremos de concentração costumam apresentar alta influência (leverage) na regressão.⁹² Como solução para a assimetria de valores y , normalmente, aplica-se a

transformação logarítmica ao vetor y ou fraciona-se o conjunto amostral de acordo com a concentração.²²

Diversos métodos de regressão podem ser utilizados para contornar o problema de não linearidade entre os espectros NIR e as propriedades do solo. Os métodos de regressão não linear já utilizados na determinação de TPH em solo foram floresta aleatória (RF),^{59,97,98} regressão penalizada por splines (PSR)⁵⁹ e árvore de regressão amplificada (BRT).¹⁴⁰ Em ambos os trabalhos desenvolvidos por Douglas et al (2018), os modelos RF apresentam melhores resultados de validação comparados aos modelos PLS.^{97,98} Dentre os métodos comumente utilizados na predição de propriedades do solo, podem ser citados redes neurais artificiais (ANN) e máquinas de vetores suporte (SVM).

Alguns trabalhos mostram que o método ANN foi superior ao PLS para alguns parâmetros do solo,¹⁵⁵ enquanto outros relatam a superioridade do método SVM.^{156,157} Por outro lado, há estudos que mostram que PLS foi o método mais adequado, inclusive para predição de carbono orgânico.^{32,158} Apesar da vantagem destes métodos em solucionar relações não lineares, eles podem gerar modelos superestimados caso o conjunto amostral não for representativo e ainda não estão disponíveis na maioria dos softwares fornecidos pelos fabricantes dos espectrômetros.⁹²

3.3.3. Pré-processamentos

Uma estratégia muito utilizada na análise de dados multivariados para contornar problemas de não linearidades, interferências e ruídos instrumentais é a aplicação de transformadas matemáticas previamente à modelagem, conhecidas como pré-processamentos. Estas transformações, quando utilizadas devidamente, podem otimizar a relação linear entre as variáveis espectrais e a propriedade de interesse, possibilitando o uso de PLS para regressão e otimizando o ajuste do modelo de predição.

Como citado anteriormente, as diferenças granulométricas do solo alteram o sinal médio do espectro de refletância difusa em NIR devido aos efeitos causados pelo espalhamento da radiação. Esta alteração na intensidade média pode ser corrigida através de pré-processamentos que normalizam as intensidades dos espectros, aproximando suas magnitudes e eliminando (ou reduzindo) o efeito de espalhamento da luz. Os pré-processamentos mais comuns para este fim são normalização pela norma; correção multiplicativa de espalhamento (MSC); e padronização normal de sinal (SNV). Embora os dois últimos sejam muito usuais na análise de solos, não

foram aplicados nos trabalhos desenvolvidos para determinação de TPH, exceto pelo trabalho de Schwartz et al (2013).¹⁰

Outro problema comum aos espectros NIR é o deslocamento de linha base ocasionado por variações instrumentais. Usualmente, a correção de linha base é realizada utilizando um espectro de referência, algoritmos polinomiais de correção e, principalmente, derivadas. Observa-se na Tabela 2 que a primeira derivada é o pré-processamento mais usual entre os estudos realizados. Embora seja eficaz na correção de linha base, o espectro derivado apresenta um sinal muito menos intenso do que o original e cada banda passa a ser representada por dois lobos, gerando uma redução da razão sinal-ruído.¹⁵⁹ Além do deslocamento de linha base, é muito comum os espectros NIR de sólidos apresentarem grande quantidade de ruído. Normalmente, o ruído é reduzido através da média de várias varreduras consecutivas da mesma amostra, porém, ainda pode ser necessário o uso de suavização ou alisamento dos espectros através de algoritmos. O método mais comum para esta finalidade é a suavização de Savitzky-Golay. A redução de variações espectrais em virtudes de variações instrumentais ou da reflexão difusa geralmente otimizam a razão sinal-ruído, contudo, deve ser realizada cautelosamente para não eliminar sinais espectrais referentes às propriedades de interesse.¹⁵⁹

A capacidade preditiva de modelos PLS é altamente dependente da centragem dos dados na média, pois, desta forma, a variância nos blocos \mathbf{X} e \mathbf{Y} será melhor representada pelas variáveis latentes.¹⁶⁰ O resultado deste pré-processamento é apenas uma translação dos eixos para o valor médio de cada um deles sem alterar a estrutura dos dados originais.¹⁵⁹ Similarmente à centragem na média, o autoescalamento realiza a centragem na média seguida pela divisão de cada coluna da matriz pela sua variância. O objetivo deste pré-processamento é equalizar a magnitude das variáveis, dando mesmo peso para cada variável. Geralmente, evita-se a utilização do autoescalamento em dados NIR devido à grande quantidade de ruídos espectrais.¹⁵⁹ Por outro lado, pode ser de grande utilidade visto que, com o aumento da frequência, as bandas tornam-se menos intensas porém não menos importantes.

Os filtros multivariados, embora muito promissores ao aumento da razão sinal-ruído de dados espectrais NIR, foram negligenciados por muitos anos no desenvolvimento de métodos multivariados para acessar a qualidade de solo e apenas recentemente ganharam maior destaque na análise de solos.⁹² O objetivo destes pré-processamentos, como correção ortogonal de sinal (OSC), ponderação por mínimos quadrados generalizada (GLSW), ortogonalização de parâmetros externos (EPO), e modelo de mistura estendida (EMM), é remover ou atenuar as variáveis da matriz \mathbf{X} que não estão correlacionadas com o vetor \mathbf{y} através do princípio de que as variáveis x ortogonais à variável y não são explicativas.¹⁶¹ Como resultado da remoção destas

variáveis, a complexidade do modelo é reduzida e menos variáveis latentes são necessárias para obter uma alta variância explicada.¹⁵⁹ Com a redução de variáveis pouco explicativas, a sensibilidade e a seletividade do modelo são otimizadas.

Alguns estudos têm mostrado melhoras na capacidade preditiva de propriedades do solo através da aplicação do pré-processamento EPO.^{92,135,146-148,156,158,162,163} De acordo com o autores, o principal objetivo de sua aplicação é remover a variância causada pela diferença no teor de água entre as amostras. Exemplificando sua aplicação, no estudo realizado por Ge, Morgan e Ackerson (2014), foi desenvolvido um modelo PLS com amostras de solo secas para determinação de SOC através de análise vis-NIR e, posteriormente, aplicaram o modelo às amostras analisadas em campo. Após o uso de EPO, a influência da umidade na predição de SOC foi removida e a análise de amostras úmidas foi possível.¹⁶² Em outro estudo realizado pelo mesmo grupo de pesquisa, os valores de RMSEP de modelos de predição de teor de argila foram reduzidos em 79% após o uso de EPO.¹³⁵ Não menos importante, porém menos popular, o filtro multivariado OSC também se mostrou eficaz e equivalente ao EPO na eliminação do efeito da água na predição de matéria orgânica em solo por análises vis-NIR, segundo o estudo realizado por Franceschini et al (2018).¹⁴⁷

Embora estes pré-processamentos tenham se mostrado eficazes na redução de interferências na predição de propriedades do solo através da espectroscopia NIR, suas aplicações em solos contaminados por hidrocarbonetos do petróleo ainda não foram avaliadas, de acordo com a literatura.

3.3.4. Seleção de variáveis

A espectroscopia NIR é caracterizada por múltiplos sobretons de uma determinada vibração molecular que, inevitavelmente, conterão a mesma informação química. Devido ao grande número de variáveis espectrais, a colinearidade pode ser um problema na sensibilidade e seletividade do modelo caso o composto de interesse apresente sinais de baixa intensidade.¹⁶⁴ Apesar de métodos de regressão como PLS reduzirem a dimensionalidade da matriz de dados, o ruído instrumental e a alta colinearidade podem ser introduzidos às variáveis latentes, prejudicando a capacidade preditiva do modelo.

Muitos estudos mostram que melhores ajustes aos modelos podem ser alcançados através do uso de variáveis mais representativas que utilizando toda faixa espectral.¹⁶⁴ Além disso, a seleção de variáveis informativas reduz a complexidade do modelo, aumentando sua robustez, e pode proporcionar um melhor entendimento químico sobre o sistema.^{165,166}

A seleção de variáveis pode ser realizada com base no conhecimento espectral (abordagem manual) ou através de algoritmos que buscam variáveis que fornecem o erro mínimo de predição ao modelo. Dentre uma grande variedade de métodos de seleção de variáveis, alguns dos métodos mais populares são o método de seleção de intervalos, como o mínimos quadrados parciais por intervalo (FiPLS) e o mínimos quadrados parciais por exclusão (BiPLS); e o algoritmo genético (GA), uma técnica que emprega um método probabilístico, de busca não-localizada, que manipula cordas binárias constituídas das variáveis experimentais codificadas. Detalhes sobre esses métodos de seleção de variáveis podem ser encontrados no artigo de revisão de Xiaobo et al (2010).¹⁶⁵

Como estes métodos de seleção de variáveis são fundamentados na redução dos erros gerados a partir da validação cruzada, as variáveis selecionadas podem não ser representativas para um conjunto de amostras externas e, conseqüentemente, a capacidade preditiva do modelo pode ser reduzida.¹⁵⁹ Portanto, torna-se crucial a validação do modelo com amostras externas após a seleção de variáveis.

Apesar das vantagens proporcionadas pela seleção de variáveis em dados espectrais obtidos na região NIR por refletância difusa, o número de estudos que envolvem a predição de parâmetros da qualidade de solo são reduzidos.¹⁴⁵ No estudo realizado por Vohland et al (2017), diferentes abordagens de seleção de variáveis na região vis-NIR-SWIR foram avaliadas para 8 parâmetros do solo, incluindo carbônico orgânico. Os resultados obtidos a partir da validação de um conjunto amostral externo mostraram que as variáveis selecionadas pelo algoritmo genético proporcionaram melhores ajustes para todos os modelos de predição.¹⁴⁵ Entre os trabalhos desenvolvidos para determinação de TPH/PAH em solo por análises NIR ou vis-NIR, apenas o estudo realizado por Sorak et al (2012) realizou a seleção de variáveis espectrais, onde excluiu-se manualmente a região de absorção da água ($5448-4686\text{ cm}^{-1}$).¹⁴² Visto a grande quantidade de interferentes em espectros de solo, o uso de algoritmos de seleção de variáveis pode tornar os modelos de predição mais ajustados e robustos para a determinação de contaminantes em solo.

3.3.5. Reconhecimento de padrões por PLS-DA

Diferentemente de análises exploratórias, onde padrões entre as propriedades das amostras ou suas respostas instrumentais são apenas visualizados graficamente, os métodos de reconhecimento de padrão buscam classificar as amostras de forma supervisionada (classes pré-estabelecidas) ou não supervisionadas (classificação de acordo apenas com suas medidas

experimentais). Métodos probabilísticos de classificação possuem a vantagem de proporcionar a probabilidade de uma amostra pertencer a uma classe. Estes métodos são baseados na estatística Bayesian, que assume o conhecimento da função de densidade de probabilidade de cada classe, e consideram que as variáveis satisfazem uma distribuição normal e usam a homogeneidade da matriz de variância-covariância das classes para definir e posicionar a região de limite para cada classe.¹⁵⁹

A análise discriminante por mínimos quadrados parciais (PLS-DA) é um método de reconhecimento de padrão supervisionado e probabilístico baseado no algoritmo PLS, porém, ao invés de utilizar a propriedade de interesse como vetor dependente (\mathbf{y}), são usadas as classes de forma codificada.¹⁶⁷ O método PLS maximiza a relação entre a variável dependente e os escores e, conseqüentemente, as variáveis latentes (LV) representam as direções que melhor discriminam as classes. Quando a aplicação envolve a discriminação de apenas duas classes, o método PLS-1 é empregado, caso contrário o método PLS-2 é utilizado. Analogamente ao método de regressão PLS, o número de variáveis latentes deve ser determinado de forma a proporcionar o melhor ajuste e evitar sobreajuste do modelo.^{168,169} Mais detalhes sobre o método PLS-DA podem ser encontrados no artigo de Brereton e Lloyd (2014).¹⁶⁷

Dos 16 estudos realizados sobre determinação de TPH em solo através de análises NIR (Tabela 2), apenas 3 avaliaram o uso de métodos de reconhecimento de padrão. Dois estudos realizam o reconhecimento de padrão pela análise de componentes principais (PCA). No trabalho de Sorak et al (2012), observou-se a separação de dois grupos, amostras contaminadas e não contaminadas através de três componentes principais, mas não foi possível discriminar amostras contaminadas com diferentes produtos (gasolina, diesel e óleo lubrificante).¹⁴² Okparanma e Mouazen (2013c) realizaram a análise exploratória de amostras de solo úmidas e secas, contaminadas e não contaminadas com diesel. Em duas componentes principais, observaram apenas a separação do conjunto de amostras contaminadas e úmidas das demais amostras.²⁵ Em 2012, Chakraborty e colaboradores¹⁴³ avaliaram a discriminação de solos conforme o tipo de contaminantes presente (diesel, óleo lubrificante e petróleo) através dos métodos supervisionados LDA, SVM e RF, porém o número de classificações corretas foi muito baixo (nº de acertos: LDA – 40%; RF – 37%; e SVM – 33%).

Assim como os métodos de regressão, os métodos de classificação são susceptíveis às interferências do solo e sua habilidade preditiva depende do uso de pré-processamentos adequados, da eliminação ou redução de ruídos espectrais e de conjuntos de treinamento representativos.

Como a identificação preliminar da fonte de contaminação pode evitar danos maiores ao meio ambiente e auxiliar na estratégia de remediação, o desenvolvimento de modelos de classificação de contaminantes em solo – como gasolina, diesel e óleo lubrificante – pode ser de grande utilidade para o monitoramento ambiental e cabível de estudos de novas abordagens quimiométricas que proporcionem melhores capacidades preditivas.

3.3.6. Validação de métodos multivariados

A validação de métodos analíticos é uma etapa crucial para avaliar e controlar os limites do método desenvolvido/implementado previamente a sua aplicação na análise de amostras “desconhecidas”. A validação é realizada através da avaliação de parâmetros analíticos (ou figuras de mérito) que representam a exatidão, precisão, linearidade, tendência, seletividade, sensibilidade, robustez, entre outros. Na maioria dos métodos vis-NIR para análises de solos, são considerados apenas os parâmetros raiz da média quadrática dos erros de predição (RMSEP), coeficiente de determinação (R^2) e diferença percentual relativa (RPD). Estes parâmetros são indicativos da similaridade entre os valores preditos pelo modelo e os valores de referência, ou seja, a exatidão do método multivariado. O valor de R^2 também é utilizado para avaliar a linearidade entre os valores preditos e os de referência, diferentemente de seu uso em métodos univariados, que indica a linearidade entre a concentração dos analitos e o sinal instrumental. Embora estes parâmetros sejam bons indicativos do ajuste do modelo, parâmetros como precisão, sensibilidade e seletividade são essenciais para o conhecimento das limitações dos métodos de predição.

A ausência de muitos parâmetros analíticos nos estudos sobre qualidade de solo e a técnica vis-NIR pode ser em decorrência da morosidade experimental (acesso à precisão), indisponibilidade de amostras “branco” (acesso à sensibilidade analítica, razão sinal-ruído e limites de quantificação e detecção) e indisponibilidade destes parâmetros na maioria dos softwares comerciais. Diferentemente dos métodos univariados, a seletividade e a sensibilidade analítica, entre outros parâmetros, são dependentes de um conjunto de variáveis e podem ser determinados segundo o conceito de sinal analítico líquido (NAS).¹⁵⁹ Neste conceito, o espectro medido é considerado como dois vetores ortogonais entre si: um deles é exclusivo do sinal analítico (vetor NAS); e o outro é o vetor dos interferentes, que contém a informação de todos os outros constituintes da amostra e variações externas. Os cálculos envolvidos no conceito NAS para os métodos de regressão PLS e PCR estão bem estabelecido na literatura científica e podem ser encontrados nas referências.^{159,170,171}

4. METODOLOGIA

A metodologia do presente trabalho resume-se em três grandes etapas: (i) aquisição e caracterização de amostras de solo e contaminantes (gasolina, diesel e óleo lubrificante); (ii) implementação de métodos de referência para determinação dos contaminantes nas amostras de solo; e (iii) desenvolvimento de métodos de classificação e quantificação dos contaminantes em solo através de espectroscopia NIR. As grandes etapas foram divididas em seções, as quais serão descritas detalhadamente a seguir e podem ser visualizadas no organograma da Figura 2.

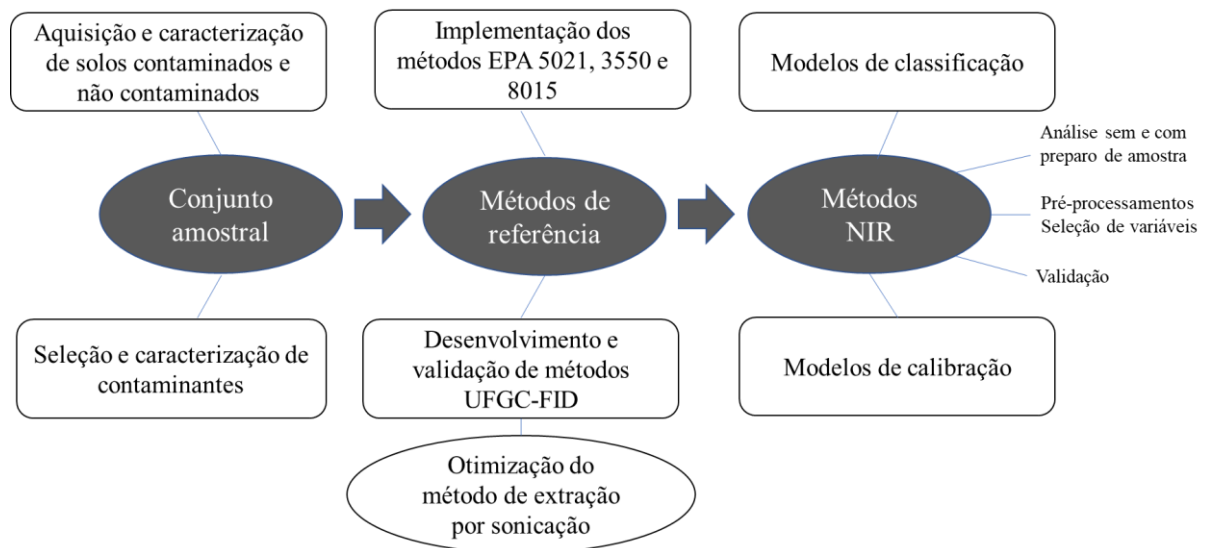


Figura 2 – Organograma das principais etapas desenvolvidas no presente trabalho.

4.1. Conjunto amostral

A primeira etapa do desenvolvimento do projeto deu-se com a construção do conjunto amostral constituído de 8 amostras de solo (5 amostras não contaminadas e 3 contaminadas); e 10 amostras de cada tipo de contaminante (gasolina, diesel e óleo lubrificante). A aquisição e a caracterização das amostras estão detalhadas a seguir.

4.1.1. Amostras de solo

Cinco amostras de solo não contaminado foram cedidas pelo Laboratório de Análise Ambiental, da Universidade Estadual de Maringá – UEM, pela MSc. Cássia Maria Bonifácio, as quais foram coletadas na região Norte-Central do Paraná, entre as latitudes 22°32'30" S e 23°36'18" S e longitudes 51°22'42" W e 52°12'30" W, em zona rural.

Três amostras de solo contaminado foram cedidas pelo Prof. Dr. Giovanni Chaves Penner (UFPA) e pelo Eng. Philipe Bandeca (Bandeca Engenharia Ambiental, Ribeirão Preto). As amostras foram coletadas no monitoramento ambiental em postos de combustível, localizados na cidade Ribeirão Preto – SP (21° 10' 39" S, 47° 48' 37" W).

Após a aquisição das amostras de solo, a caracterização foi realizada de acordo com o Manual de Métodos de Análise de Solo, desenvolvido pela Embrapa.¹⁵² Os parâmetros determinados foram: densidade aparente; densidade de partícula; granulometria (argila total, silte total, areia total e areia grossa); pH; umidade total; e matéria orgânica. Dentre os parâmetros citados, a granulometria e a classe morfopedológica das amostras de solo coletadas no Paraná foram determinadas pelo Laboratório de Análise Ambiental da UEM.

Além destes parâmetros, realizou-se a análise elementar das amostras por espectroscopia de fluorescência de raio-X por energia dispersiva (EDX) no laboratório de monitoramento de combustíveis do Cempecq (Centro Multidisciplinar de Pesquisa em Combustíveis, Biocombustíveis, Petróleo e Derivados). O espectrômetro utilizado foi um EDX-800, da marca Shimadzu, e a varredura abrangeu os elementos de sódio a urânio.

As amostras de solo também foram analisadas no espectrômetro FT-NIR, modelo Nicolet 6700 (Thermo Fisher Scientific), para identificação dos principais grupos funcionais presentes nas amostras. Foram utilizadas 5 alíquotas de cada amostra e a média dos espectros, adquiridos em uma resolução de 4 cm⁻¹ e 32 varreduras, foi utilizada na identificação dos grupos funcionais.

Com a finalidade de buscar padrões existentes nas amostras, realizaram-se duas análises exploratória através de PCA, a primeira utilizando os parâmetros físico-químicos e a outra os espectros NIR.

4.1.2. Amostras de combustíveis e óleos lubrificantes

O primeiro passo para a construção do conjunto amostral de contaminantes foi a seleção de 10 amostras representativas de gasolina e 10 amostras de óleo diesel a partir de 266 amostras

de gasolina e 380 de óleo diesel cedidas pelo laboratório de monitoramento do Cempeqc. Para tal, efetuou-se uma Análise Hierárquica de Agrupamento (HCA) através do método de Ward, e os parâmetros físico-químicos das amostras foram utilizados para análise após o autoescalamento. Os parâmetros físico-químicos foram determinados pelo laboratório de monitoramento de combustíveis do Cempeqc de acordo com normas ASTM e ABNT, especificadas nas resoluções ANP nº40 de 2013 (gasolina) e ANP nº50 de 2013 (óleo diesel). O conjunto amostral foi dividido em dez grupos pelo ajuste do grau de similaridade e, então, uma amostra de cada grupo foi selecionada. Finalmente, foram selecionadas 5 amostras de gasolina tipo comum, 3 tipo premium e 2 tipo aditivada; e 5 amostras de diesel S500 e 5 amostras diesel S10. Os valores dos parâmetros físico-químicos referentes às amostras selecionadas estão dispostos na Tabela A1, no Anexo A.

Para obter um conjunto de amostras de óleo lubrificante representativo, buscou-se adquirir amostras de diferentes marcas, graus de viscosidade, categoria API, origem (mineral ou semissintética) e aplicação (carros leves, pesados ou motores 4 tempos). Além destes fatores, fez parte da composição do conjunto amostral uma amostra de óleo lubrificante usado após 10.000 km rodados em um carro leve movido a etanol combustível. As informações detalhadas de cada amostra de óleo lubrificante adquirida encontram-se na Tabela A2.

4.1.2.1. Caracterização das amostras de combustíveis e óleos lubrificantes

As amostras de gasolina, diesel e óleos lubrificante foram analisadas por cromatografia gasosa para determinação das frações GRO, DRO e ORO, e dos compostos BTEX e PAH. O equipamento utilizado foi um cromatógrafo a gás com detector de ionização por chama (GC-FID), modelo Trace GC Ultra (Thermo Scientific), equipado com injetor split/splitless e autoamostrador Triplus. A quantificação das frações GRO e DRO foi realizada por calibração externa conforme a norma EPA 8015C.

A fração GRO de cada amostra foi definida pela soma das áreas dos picos cromatográficos com tempo de retenção entre o 3-metilpentano e o 1,2,4-trimetilbenzeno. Os tempos de retenção foram apontados pela análise de um material de referência certificado (MRC) “GRO Mix” (Supelco, Lot. LC01984), composto por: 3-metilpentano; benzeno; 2,2,4-trimetilpentano; tolueno; etilbenzeno; m-xileno; o-xileno; 1,2,4-trimetilbenzeno; e naftaleno, diluídos em metanol na concentração individual de 2000 $\mu\text{g L}^{-1}$, aproximadamente.

De acordo com a EPA 8015C, a calibração externa pode ser realizada com padrões analíticos ou com gasolina sem aditivos, denominada no Brasil como “gasolina tipo A”, assim,

optou-se pela segunda alternativa por questões de disponibilidade, custo e maior representatividade em relação às amostras de gasolina comercial. Desta forma, a soma das áreas entre o 3-metilpentano e o 1,2,4-trimetilbenzeno da amostra de gasolina A (GasA) foi dada como fração GRO igual a 100% e, então, foi usada como base para determinação da fração GRO nas amostras de gasolina Comum, Aditivada e Premium.

A curva analítica foi desenvolvida com 5 soluções de gasolina A em metanol, analisadas em triplicata, e a faixa de medição foi de 4,50 % (m/m) a 11,17 % (m/m). Então, as amostras de gasolina foram diluídas em metanol na concentração média de 8,46 % (m/m) para análise cromatográfica.

Em consequência do GRO Mix apresentar compostos individuais em concentração muito baixa, preparou-se uma solução de padrões BTEX [benzeno 99.9% (Aldrich); tolueno 99.5% (Vetec); etilbenzeno 99.8% (Acros Organics); m-xileno 99% (Acros Organics); e o-xileno 97% (Aldrich)] em metanol e desenvolveram-se três curvas analíticas de 5 pontos cada em triplicata, abrangendo a faixa de concentração de 0,01% a 4,16% (m/m), aproximadamente. Na ausência de um padrão individual de naftaleno, este foi quantificado através do fator de resposta determinado pela razão entre as áreas deste composto em relação ao benzeno. Tal relação foi calculada com base nos resultados da análise do MRC “GRO Mix” no qual os dois componentes estavam contidos.

O método cromatográfico utilizado para caracterização das amostras de gasolina foi baseado na norma ASTM D5769,¹⁷² adaptada para uso de detector FID e com algumas alterações na programação da temperatura do forno para melhor separação. As condições cromatográficas estão dispostas na Tabela 3.

Tabela 3 – Condições cromatográficas para caracterização das amostras de combustíveis.

Parâmetro	GRO e BTEX	DRO e PAH
	ASTM D5769 (adaptada)	ASTM D6209 (adaptada)
Vazão do gás de arraste (hélio)	2,5 mL min ⁻¹	7,0 mL min ⁻¹
Razão de Split	20:1	Splitless
Temperatura do injetor	250 °C	250 °C
Temperatura do detector	300 °C	300 °C
Programação da temperatura do forno:	50 °C (0 min) 2 °C min ⁻¹ até 140 °C (0 min) 50 °C min ⁻¹ até 300 °C (5 min)	60 °C (2 min) 7 °C min ⁻¹ até 290 °C (30 min) 10 °C min ⁻¹ até 320 °C (7 min)
Coluna cromatográfica	60 m x 0,25 mm x 1,0 µm dimetilpolissiloxano (Thermo Fisher Scientific)	100 m x 0,25 mm x 0,5 µm dimetilpolissiloxano (Agilent Technology J&W)

A fração DRO das amostras de diesel foi determinada de forma análoga à fração GRO. O método cromatográfico da norma ASTM D6209, desenvolvido para análise de PAH por GC/MS, foi utilizado como ponto de partida para caracterização das amostras de diesel. O método foi avaliado e adaptado para uso de uma coluna de maior comprimento e detector FID. As condições cromatográficas e as colunas capilares utilizadas são mostradas na Tabela 3. Os tempos de retenção inicial e final foram determinados pela análise do MRC “TPH MIX” (RTC, Lot. 016538), composto por 17 hidrocarbonetos alifáticos, de 8 a 40 átomos de carbono na cadeia, diluídos em diclorometano, e a quantificação foi realizada por calibração externa usando-se uma amostra de diesel sem aditivos (diesel tipo A). A faixa de medição da curva analítica foi de 0,31 %(m/m) a 0,89 %(m/m) e as amostras foram diluídas em diclorometano na concentração média de 0,59 %(m/m).

As concentrações dos compostos PAH nas amostras de diesel foram determinadas por curvas analíticas desenvolvidas com o MRC “PAH Mix” (Supelco, Lot. XA12140V), composto por 13 hidrocarbonetos: acenaftileno; fluoreno; fenantreno; antraceno; pireno; benzo(a)antraceno; criseno; benzo(b)fluoranteno; benzo(k)fluoranteno; benzo(a)pireno; dibenzo(a,h)antraceno; benzo(g,h,i)perileno; e indeno(1,2,3-CD)pireno, diluídos em diclorometano na concentração de $500 \mu\text{g mL}^{-1}$, aproximadamente. As curvas analíticas foram desenvolvidas com 5 pontos em triplicata e intervalo de medição de $22,3 \mu\text{g g}^{-1}$ a $181,1 \mu\text{g g}^{-1}$. As amostras de diesel foram diluídas em diclorometano na concentração de $5000 \mu\text{g g}^{-1}$, aproximadamente.

Os ésteres presentes no diesel foram determinados utilizando uma mistura dos quatro ésteres metílicos mais comuns presentes em biodiesel comercial (palmitato, oleato, linoleato e estearato de metila). A proporção entre os quatro ésteres foi utilizada para prever a matéria-prima que originou o biodiesel adicionado ao diesel.

A fração ORO dos óleos lubrificantes não foi determinada através da cromatografia gasosa convencional, pois a eluição da maior parte dos compostos ORO ocorreu juntamente ao drift presente nos cromatogramas durante a última rampa de aquecimento. Como a ocorrência de drift é comum em altas taxas de aquecimento e altas temperaturas finais, optou-se determinar a fração ORO por cromatografia gasosa ultrarrápida por não haver presença de drift devido ao aquecimento resistivo e maior resistência da coluna a sangramentos. Os parâmetros cromatográficos estão dispostos na Tabela 4.

Tabela 4 – Condições cromatográficas para determinação da fração ORO dos óleos lubrificantes por cromatografia gasosa ultrarrápida.

Parâmetro	valor
Vazão do gás de arraste (hidrogênio)	3,0 mL min ⁻¹
Razão Split	10:1
Temperatura do injetor	275 °C
Temperatura do detector	280 °C
Volume de injeção	0,1 µL
Programação da temperatura do módulo ultrarrápido:	60 °C (0,3 min) 80 °C min ⁻¹ a 300 °C (1,7 min)
Coluna cromatográfica	PH SE54 (5 m x 0,32 mm x 0,05 µm) 5% fenil em dimetil-polissiloxano (Thermo Scientific)

A curva analítica para determinação de ORO foi desenvolvida com 5 pontos em triplicata, utilizando o MRC “TPH Mix” na faixa de concentração 76 µg g⁻¹ a 1200 µg g⁻¹. As amostras de óleo lubrificante foram diluídas em diclorometano na concentração aproximada de 195 µg g⁻¹ e cada amostra foi analisada em triplicata.

Após as análises cromatográficas, as amostras dos contaminantes foram analisadas espectroscopicamente na região NIR para identificação dos grupos funcionais e para realizar uma análise exploratória por PCA para verificação de possíveis padrões espectrais.

4.2. Métodos de referência

Os métodos de referência foram implementados neste projeto com três objetivos: (i) quantificar a concentração das frações de TPH nas amostras de solo adquiridas; (ii) conhecer e comparar os parâmetros analíticos dos métodos convencionais com os métodos desenvolvidos; e (iii) avaliar e otimizar os métodos de extração.

Foram implementados os métodos de extração EPA 5021A,¹⁰² para extração da fração GRO por headspace, e EPA 3550C,¹⁶ para extração da fração DRO e ORO por sonicação. A quantificação das frações TPH deu-se através do método cromatográfico EPA 8015C⁸ e pelos métodos cromatográficos ultrarrápidos desenvolvidos neste projeto.

4.2.1. Métodos EPA

Os métodos cromatográficos convencionais foram desenvolvidos através de um cromatógrafo a gás Trace GC Ultra (Thermo Fisher Scientific), equipado com injetor split/splitless, detector FID, coluna cromatográfica Agilent Technology (30 m x 0,53 mm x 1,5 µm; 5% fenil em metilsilicone) e autoamostrador Triplus (injeção líquida e headspace).

Na Tabela 5, estão apresentadas as condições cromatográficas utilizadas para cada fração e a coluna cromatográfica de acordo com o método EPA 8015C. Uma rampa de aquecimento foi adicionada ao final da programação do método para DRO para garantir a eluição completa dos compostos de maior peso molecular presentes no diesel. Pelos mesmos motivos apresentados na caracterização dos óleos lubrificantes, o método GC-FID convencional não foi implementado para quantificação da fração ORO.

Tabela 5 – Condições cromatográficas (EPA 8015C) para determinação das frações GRO e DRO.

Parâmetros	GRO	DRO
Vazão do gás de arraste (hélio)	5 mL min ⁻¹	5 mL min ⁻¹
Razão de split	Splitless	6:1
Volume de amostra injetado	500 µL	1 µL
Temperatura do injetor	200 °C	200 °C
Temperatura do detector	340 °C	340 °C
Programação da temperatura do forno	45 °C (1 min)	45 °C (3 min)
	5 °C min ⁻¹ até	12 °C min ⁻¹ até
	100 °C (0 min)	275 °C (12 min)
	8 °C min ⁻¹ até	*20 °C min ⁻¹ até
	275 °C (5 min)	*300 °C (5 min)
Tempo de análise	39 min	40 min
Coluna cromatográfica	30 m x 0,53 mm x 1,5 µm 5% fenil em metilpolissilixano (Agilent Technology J&W)	

* Temperaturas adicionadas à programação do método EPA 8015.

Previamente ao desenvolvimento das curvas analíticas para quantificação da fração GRO, o método de extração EPA 5021 foi avaliado para proporcionar maior repetibilidade e capacidade de extração. De acordo com método EPA 5021, 2 g de amostra sólida são inseridos em um vial de 22 mL com 5 mL de solução NaCl 36% (m/v), então, a amostra é agitada e aquecida durante 50 min para equilíbrio das fases.¹⁰² As condições operacionais do headspace estão descritas na Tabela 6.

Tabela 6 – Condições operacionais do headspace.

Parâmetro	Condição
Volume amostrado pela seringa	500 μL
Temperatura de incubação	85,0 $^{\circ}\text{C}$
Tempo de incubação	50 min
Velocidade de agitação	250 rpm
Tempo de agitação ligada	5 s
Tempo de agitação desligada	2 s

Considerando que o tempo de incubação de 50 min, recomendado pelo método EPA 5021, é demasiadamente longo, o método de extração foi avaliado usando-se tempo de incubação reduzido. Os tempos 5 min; 10 min; 15 min; 20 min; 30 min; e 40 min foram testados pela análise de seis amostras de gasolina em solo, na concentração próxima a 1000 mg kg^{-1} , em triplicata. O tempo que gerou menor desvio padrão relativo e maior área foi utilizado no desenvolvimento da curva analítica.

A curva analítica para quantificação de GRO foi desenvolvida com 5 pontos em triplicata real, na faixa de concentração de 672 mg kg^{-1} a 6090 mg kg^{-1} , resultantes da fortificação de alíquotas do solo (S1) com gasolina A. O intervalo de integração para quantificação da fração GRO foi definido pelos tempos de retenção dos compostos 3-metilpentano e 1,2,4-trimetilbenzeno.

Desenvolveu-se a curva analítica para DRO através de 5 pontos em triplicata de soluções de diesel A em diclorometano, na faixa de concentração de 111 mg kg^{-1} a 1161 mg kg^{-1} . Os tempos de retenção foram definidos pela eluição dos compostos decano e octacosano e, então, foram utilizados para determinação da fração DRO.

As misturas e as soluções mãe foram fortificadas com surrogate para avaliação da recuperação dos analitos. Os surrogates utilizados foram: isobutilbenzeno para a fração GRO; e o-terfenil para a fração DRO.

Os critérios de aceitação para a curva analítica e para a quantificação das frações TPH foram baseados no guia EPA 8000C.¹⁷³ A curva analítica foi aceita quando apresentou $R^2 \geq 0,99$ e $\text{DPR} \leq 20\%$ para cada ponto; e a extração quando apresentou valor de recuperação entre 70% a 130% do valor teórico.

4.2.2. Desenvolvimento dos métodos cromatográficos ultrarrápidos

Um cromatógrafo Trace GC Ultra (Thermo Fisher Scientific), equipado com injetor split/splitless, detector FID, módulo ultrarrápido e autoamostrador AS3000 (injeção líquida), foi utilizado para o desenvolvimento dos métodos cromatográficos ultrarrápidos (UFGC). Duas colunas capilares de dimensões reduzidas foram avaliadas na separação dos analitos: (i) PH SE54 (5 m x 0,32 mm x 0,05 μm ; 5% difenil em dimetilpolissiloxano); e (ii) PH RTX-1 (10 m x 0,32 mm x 3,0 μm ; 100% dimetilpolissiloxano), ambas da marca Thermo Fisher Scientific.

O autoamostrador AS3000 trabalha apenas no modo de injeção líquida, portanto, propôs-se a análise da fração GRO em solução etanólica e em solução de hexadecano devido ao alto fator de retenção do solvente pela coluna cromatográfica.

Uma rápida avaliação da performance de cada coluna na separação dos compostos GRO foi realizada através da análise de uma solução preparada com os seguintes analitos: etanol, hexano, benzeno, tolueno, etilbenzeno, o-xileno, m-xileno, p-xileno, decano e isobutilbenzeno. As condições cromatográficas utilizadas na avaliação prévia estão na Tabela 7.

Tabela 7 – Condições cromatográficas iniciais para avaliação da separação dos compostos GRO pelas colunas PH SE54 e PH RTX-1.

Parâmetro	Coluna Cromatográfica	
	PH SE54 5 m x 0,32 mm x 0,05 μm	PH RTX-1 10 m x 0,32 mm x 3,0 μm
Vazão do gás de arraste (hidrogênio)	2,0 mL min ⁻¹	3,5 mL min ⁻¹
Razão Split	10:1	6:1
Temperatura do injetor	250 °C	250 °C
Temperatura do detector	275 °C	275 °C
Volume de injeção	0,5 μL	0,5 μL
Programação da temperatura do módulo ultrarrápido:	30 °C (0,5 min)	40 °C (0,5 min)
	50 °C min ⁻¹ a	50 °C min ⁻¹ a
	60 °C (0 min)	60 °C (0 min)
	200 °C min ⁻¹ a	200 °C min ⁻¹ a
	150 °C (0 min)	150 °C (0,3 min)
	300 °C min ⁻¹ a	300 °C min ⁻¹ a
	300 °C (0,6 min)	290 °C (1,2 min)
Tempo de análise	2,7 min	3,3 min

A separação dos compostos DRO e ORO também foi avaliada com ambas colunas cromatográficas. A análise foi realizada pela injeção do MRC “TPH Mix” com a adição de o-terfenil. Os parâmetros cromatográficos utilizados na avaliação das colunas estão na Tabela 8.

Tabela 8 – Condições cromatográficas iniciais para cada coluna na avaliação da separação dos compostos do TPH Mix.

Parâmetro	Coluna Cromatográfica	
	PH SE54 5 m x 0,32 mm x 0,05 µm	PH RTX-1 10 m x 0,32 mm x 3,0 µm
Vazão do gás de arraste (hidrogênio)	3,0 mL min ⁻¹	3,5 mL/min
Razão Split	10:1	6:1
Temperatura do injetor	250 °C	250 °C
Temperatura do detector	275 °C	275 °C
Volume de injeção	0,1 µL	0,5 µL
Programação da temperatura do módulo ultrarrápido:	60 °C (0,5 min)	250 °C (0 min)
	50 °C min ⁻¹ a	150 °C min ⁻¹ a
	300 °C (1,5 min)	290 °C (6,0 min)
Tempo de análise	6,8 min	6,3 min

Após a definição da melhor coluna cromatográfica, foram realizados reajustes nos métodos cromatográficos. Na Tabela 9, encontram-se as condições definidas para separação dos compostos BTEX e das frações GRO, DRO e ORO, independente do solvente.

Tabela 9 – Condições definidas para determinação das frações GRO, DRO e ORO por cromatografia gasosa ultrarrápida.

Parâmetro	GRO e BTEX	DRO e ORO
	PH RTX-1 10 m x 0,32 mm x 3,0 µm	PH SE54 5 m x 0,32 mm x 0,05 µm
Vazão do gás de arraste (hidrogênio)	2,5 mL min ⁻¹	3,0 mL min ⁻¹
Razão Split	6:1	10:1
Temperatura do injetor	275 °C	275 °C
Temperatura do detector	280 °C	280 °C
Volume de injeção	0,5 µL	0,1 µL
Programação da temperatura do módulo ultrarrápido:	40 °C (0,5 min)	60 °C (0,3 min)
	30 °C min ⁻¹ a	
	120 °C (0 min)	
	200 °C min ⁻¹ a	
	300 °C (0,5 min)	300 °C (1,7 min)
Tempo de análise	4,6 min	5,0 min

Definidos os parâmetros cromatográficos e os tempos de retenção dos analitos, foram desenvolvidas as curvas analíticas para as três frações TPH e para os compostos BTEX. O intervalo de medição das frações foi de 100 mg kg⁻¹ a 10.000 mg kg⁻¹ e dos compostos BTEX de 1 mg kg⁻¹ a 1000 mg kg⁻¹. As curvas analíticas referentes à gasolina foram desenvolvidas com 14 pontos e as referentes ao diesel e óleo lubrificante com 10 pontos. As soluções foram preparadas com os mesmos combustíveis e surrogates utilizados nas curvas dos métodos convencionais e a curva para ORO foi preparada com a amostra O4 devido ao seu maior teor da fração ORO.

As curvas analíticas foram validadas por figuras de mérito, como linearidade, sensibilidade, razão sinal-ruído, limites de detecção e quantificação, intervalo dinâmico linear, intervalo de confiança, precisão e precisão intermediária, a qual foi realizada através de um estudo de estabilidade das soluções armazenadas à -19°C, durante 15 dias. A exatidão dos métodos UFGC foi avaliada juntamente ao método de extração através do parâmetro recuperação.

4.2.3. Avaliação do método de extração por sonicação

Após alguns testes preliminares de extração segundo o método EPA 3550, foram propostas algumas adaptações no procedimento. De acordo com o método, 2 g de amostra são medidos em um vial e, adicionados em sequência, 2 g de sulfato de sódio e 10 mL de solvente. A amostra é sonicada por 3 min, filtrada a vácuo e transferida para uma coluna de 20 mm de diâmetro interno, preenchida de sílica gel e com lã de vidro no fundo para a purificação. As adaptações envolveram o uso de vial com tampa rosqueável e septo PTFE; 4 g de amostra; agitação por vórtex previamente à sonicação; transferência da solução extratora para uma seringa de plástico contendo sulfato de sódio, sílica gel (tamanho de partícula de 10 µm a 40 µm e ligante CaSO₄) e algodão; e filtração por filtro de seringa de polietersulfona (PES), de porosidade 5 µm. O equipamento para sonicação utilizado foi um ultrassom THORNTON, modelo T14, de potência máxima de 100 Watts.

Os parâmetros do método de extração foram avaliados através de um planejamento experimental (DOE) para cada grupo e fração de compostos (BTEX, GRO e DRO) e avaliaram-se os seguintes fatores: tempo de ultrassom (U); agitação por vórtex (A); volume de solvente (V); e tipo de solvente (S). O planejamento para os compostos BTEX e GRO foi do tipo fatorial completo em 2 níveis e 3 replicatas no ponto central para cada solvente, totalizando em 22 experimentos. Para a fração DRO, o fator “agitação” também foi categórico e os pontos centrais foram analisados em duplicatas, assim, o número de experimentos foi igual a 24. Os valores mínimos, máximos e centrais de cada fator nos planejamentos experimentais estão dispostos na Tabela 10.

Tabela 10 – Níveis dos fatores do planejamento experimental de cada grupo de compostos.

Parâmetro	BTEX e GRO	DRO
Tempo de ultrassom	5 min; 10 min; 15 min	5 min; 12,5 min; 20 min
Agitação por vórtex	0 s; 30 s; 60 s	sem; com
Volume de solvente	5,0 mL; 7,5 mL; 10,0 mL	5,0 mL; 7,5 mL; 10,0 mL
Solvente	etanol; hexadecano	diclorometano; hexano

As respostas das extrações foram avaliadas através da Análise de Variância (ANOVA) e as equações geradas foram utilizadas para predição das melhores condições de extração. Por fim, a repetibilidade do método otimizado foi determinada através da análise de sete amostras de solo fortificadas em mesma concentração.

Havendo apontado as condições dos métodos de extração e cromatográficos, as amostras de solo obtidas do monitoramento ambiental foram analisadas para determinação das frações de TPH e BTEX.

4.3. Análises espectroscópicas e quimiométricas

4.3.1. Aquisição dos espectros NIR

A aquisição dos espectros utilizados neste projeto foi realizada através de um espectrômetro IR com transformada de Fourier, modelo Nicolet 6700 FTIR (Thermo Fisher Scientific), equipado com fonte de lâmpada de tungstênio e detector de InGaAs, presente no acessório de aquisição de espectros NIR, Smart NIR Integration Sphere. Os interferogramas foram convertidos a espectros NIR (absorbância por número de onda), na faixa espectral de 10.000 cm^{-1} a 4000 cm^{-1} , via software Omnic (Thermo Fisher Scientific).

As análises NIR de soluções e misturas líquidas deram-se pelo modo de transletância, utilizando uma cubeta de aço inox de caminho ótico de 0,5 mm. Os espectros NIR resultaram da média de 22 varreduras na resolução de 4 cm^{-1} , totalizando em 3112 variáveis espectrais. As amostras sólidas foram analisadas no modo de refletância difusa diretamente em vials de 22 mL, portanto, o número de varreduras foi aumentado para 32 por espectro com o objetivo de reduzir a quantidade de ruído espectral.

Para reduzir o efeito da heterogeneidade das amostras de solo na análise por refletância difusa, avaliou-se o número de replicatas necessário para obter um espectro representativo. Para tal, preparou-se um conjunto de 50 amostras fortificadas com a gasolina G2 na faixa de concentração 132 mg kg^{-1} a 3920 mg kg^{-1} . Cada amostra foi analisada em sextuplicata, homogeneizando-se a mistura a cada aquisição de espectro. Para avaliar o número de replicatas necessário, desenvolveu-se um modelo PLS com as médias das sextuplicatas e outro com as médias das triplicatas. Os ajustes resultantes dos dois modelos foram comparados e o menor número de replicatas foi utilizado caso não houvesse diferença significativa entre os ajustes.

4.3.2. Conjuntos amostrais

Para alcançar o objetivo do presente trabalho – identificar e quantificar os hidrocarbonetos oriundos de combustíveis e óleos lubrificantes em solo através de análise NIR, com ou sem preparo de amostra – três conjuntos de 300 amostras foram preparados abrangendo a faixa de concentração de TPH de, aproximadamente, 100 mg kg^{-1} a $10.000 \text{ mg kg}^{-1}$.

O primeiro conjunto constituiu de 300 amostras de solo fortificado com os contaminantes. A fortificação foi realizada com os 30 contaminantes e os 5 solos não contaminados de forma que cada solo fosse fortificado pelo mesmo contaminante duas vezes no conjunto amostral. Além das 300 amostras preparadas, 5 espectros de cada solo não contaminado foram adicionados como “brancos” ao conjunto, totalizando em 325 amostras. As amostras foram analisadas espectroscopicamente em triplicata, contudo, apenas as médias das réplicas foram utilizadas na modelagem.

Os outros dois conjuntos constituíram-se de soluções dos contaminantes em solventes extratores, um preparado em hexano e o outro em etanol. Preparou-se 10 soluções com cada contaminante e os solventes puros foram analisados 20 vezes para fornecer os “brancos” para os conjuntos amostrais, totalizando em 320 amostras para cada conjunto. Devido à homogeneidade das soluções, apenas um espectro NIR foi adquirido por amostra.

Cada conjunto amostral foi dividido em dois subconjuntos: (i) de calibração, com dois terços das amostras; e (ii) de validação, com um terço das amostras. As amostras de calibração foram selecionadas pelo algoritmo Onion, o qual busca amostras com menor covariância através da distância da média dos valores de y para, então, obter um conjunto amostral representativo.^{174,175} A execução do algoritmo foi realizada para cada analito antes da modelagem para assegurar que o conjunto de calibração abrangeu a maior faixa de concentração.

As amostras de calibração foram utilizadas no desenvolvimento dos modelos, enquanto as amostras de validação foram responsáveis pela verificação dos ajustes. Dado que estas amostras de validação derivaram de um mesmo conjunto que as amostras de calibração, prepararam-se outros três conjuntos para validação dos modelos. Cada conjunto foi formado por 27 amostras de solo fortificado, sendo 9 amostras fortificadas com cada tipo de contaminante em três níveis de concentração – nível inferior, superior e próximo ao limite de intervenção para TPH. Através destes conjuntos de validação, verificou-se a influência de possíveis interferentes provenientes do processo de extração na predição pelos modelos quimiométricos.

Por fim, os modelos desenvolvidos foram validados pela predição das 3 amostras de solo contaminado (amostras “reais”), cujos valores de referência das frações TPH foram determinados pelo método cromatográfico ultrarrápido. Portanto, os modelos de calibração e classificação foram validados por três diferentes conjuntos: (I) amostras de validação preparadas juntamente às amostras de calibração; (II) amostras de solo fortificado que passaram pelo processo de preparo de amostra; e (III) amostras “reais”.

4.3.3. Modelos de classificação

Os modelos de classificação foram desenvolvidos através do método PLS-DA executado via software Matlab associado ao pacote PLS toolbox. As classes foram codificadas como 0 (branco); 1 (gasolina); 2 (diesel); e 3 (óleo lubrificante). A matriz \mathbf{Y} foi autoescalada e diversos pré-processamentos foram avaliados para a matriz \mathbf{X} , entre eles estão: centragem na média; autoescalamento; correção de linha base por peso de mínimos quadrados; 1ª derivada; 2ª derivada; correção de espalhamentos múltiplos (MSC); padronização normal de sinal (SNV); correção por sinal ortogonal (OSC); ponderação por mínimos quadrados gerais (GLSW); ortogonalização por parâmetro externo (EPO); e modelo de mistura estendida (EMM). Os pré-processamentos foram avaliados pela variância explicada das matrizes \mathbf{X} e \mathbf{Y} e pelos parâmetros dos modelos gerados, como coeficiente de correlação e valores de RMSE dos conjuntos de calibração, validação cruzada (venetian blind) e validação externa.

Além dos pré-processamentos, a aplicação do algoritmo genético (GA) para seleção de variáveis foi avaliada a fim de obter modelos com melhores capacidades preditivas. Para aumentar as probabilidades de determinar as variáveis mais adequadas para a predição das classes, o algoritmo genético foi executado com os valores máximos permitidos pelo PLS toolbox para tamanho de população (256 populações) e número de gerações (200 gerações). Além destes parâmetros, a configuração do GA constituiu de taxa de mutação de 0,5%; duplo cruzamento; início com 30% das variáveis; janelas de uma variável; método de regressão PLS com número máximo de 10 VL; validação cruzada contínua com divisões de 5 em 5 amostras e número de iterações iguais a 1. As variáveis selecionadas pelo GA foram utilizadas para desenvolver o modelo de predição e para uma segunda execução do GA, com as mesmas configurações da primeira, a fim de reduzir ainda mais a covariância na matriz \mathbf{X} .

Após a avaliação dos pré-processamentos e da seleção de variáveis, definiu-se o número de variáveis latentes (VL) baseando-se nos valores de RMSEC, RMSECV, RMSEP e variância

explicada das matrizes \mathbf{X} e \mathbf{Y} , de forma a minimizar os erros de predição e evitar o sobreajuste do modelo.^{168,169,176}

A capacidade preditiva dos modelos PLS-DA foi avaliada através da porcentagem de verdadeiros e falsos positivos gerados a partir dos três conjuntos de validação.

Ademais, outros métodos de classificação, como k-vizinho mais próximo (KNN) e máquina de vetor suporte (SVM), foram avaliados de forma singela para proporcionar alternativas ao método PLS-DA. A matriz \mathbf{X} dos três conjuntos amostrais foram pré-processadas com SNV e centragem na média. O número de vizinhos para os modelos KNN foi escolhido pelo menor erro de classificação para o conjunto de validação cruzada. E os modelos SVM foram desenvolvidos pelo algoritmo LIBSVM, tipo nu-SVC, função kernel radial basis, viés de probabilidade igual a 0,5 e compressão da matriz \mathbf{X} pelo método PLS. Os modelos foram aplicados aos três conjuntos de validação e avaliados pela porcentagem de classificações corretas.

4.3.4. Modelos de calibração

Através do método de regressão PLS, foram desenvolvidos modelos para predição das frações de TPH (GRO, DRO e ORO) e dos seguintes analitos: benzeno; tolueno; etilbenzeno; o-xileno; m,p-xilenos; naftaleno; acenaftileno; fluoreno; fenantreno; antraceno; pireno; benzo(a)antraceno; criseno; e benzo(b)fluoranteno. Os pré-processamentos avaliados nos modelos de classificação foram testados de forma similar nos modelos de calibração, assim como a execução do algoritmo genético para a seleção de variáveis. A redução de erros proporcionada pelo uso de GA foi avaliada por um teste F para verificar se a otimização da capacidade preditiva foi significativa ou não.

Os conjuntos de calibração e validação (I) foram utilizados para determinação de diversas figuras de mérito, como: exatidão; linearidade; precisão; tendência; razão sinal-ruído; seletividade; sensibilidade; limites de detecção e quantificação; erro relativo médio; diferença percentual relativa; e presença de erros sistemáticos. A forma de determinação de cada figura de mérito está descrita detalhadamente na seção 4.4.2.

Os modelos desenvolvidos para predição das frações de TPH foram aplicados aos conjuntos de validação (II) e de amostras reais. A partir dos resultados obtidos, definiu-se o método NIR mais adequado para o uso em amostras externas.

4.4. Determinação dos parâmetros analíticos

Dado que a determinação das frações de TPH por cromatografia gasosa é realizada por método de regressão univariada e, por espectroscopia NIR, regressão multivariada, os parâmetros analíticos de cada método estão descritos separadamente em duas subseções.

4.4.1. Parâmetros analíticos cromatográficos

4.4.1.1. Resolução cromatográfica

O primeiro parâmetro avaliado dos métodos cromatográficos foi a separação dos picos devido à sua grande importância na identificação dos analitos e, conseqüentemente, na quantificação correta das frações de TPH e analitos individuais. A separação dos picos foi avaliada pelos valores de resolução (R_s) entre os picos eluídos, calculados de acordo com a equação 1:

$$R_s = \frac{\sqrt{N}}{4} \cdot \frac{(\alpha-1)}{\alpha} \cdot \frac{k'}{(k'+1)} \quad (\text{Eq. 1})$$

onde N é o número de pratos teóricos; α , o fator de seletividade; e k' , o fator de retenção de cada analito. O fator de retenção foi calculado a partir do valor de tempo de retardamento de cada método, determinado teoricamente pelos valores de volume da coluna e fluxo de gás de arraste. Já o número de pratos teóricos (N) e o fator de seletividade (α) foram calculados segundo as equações 2 e 3, respectivamente:

$$N = 16 \cdot \left(\frac{t_r}{w}\right)^2 \quad (\text{Eq. 2})$$

$$\alpha = \frac{k'_2}{k'_1} \quad (\text{Eq.3})$$

onde t_r é o tempo de retenção; e w a largura do pico. Os valores de resolução foram utilizados para comparação da seletividade entre os métodos EPA e os métodos cromatográficos ultrarrápidos desenvolvidos.

4.4.1.2. Linearidade

Após a definição dos parâmetros instrumentais cromatográficos, as curvas analíticas foram desenvolvidas por mínimos quadrados e os parâmetros analíticos foram determinados.

A linearidade dos métodos foi avaliada através do valor dos coeficientes de correlação de Pearson (r); de determinação (R^2); e de determinação ajustado (R^2_{ajustado}); e, finalmente, pela distribuição dos resíduos no gráfico de regressão.

De acordo com o guia de validação da Agência Nacional de Vigilância Sanitária (ANVISA) e da EPA,^{8,177} o coeficiente de correlação deve ser igual ou maior que 0,990 para o ajuste dos dados ser considerado ideal para a regressão linear, sendo este critério adotado neste trabalho.

A faixa linear dinâmica da curva analítica, a qual indica se todo intervalo de medição se comporta de forma linear, foi avaliada através dos limites de confiança da linha de regressão. Não havendo nenhum ponto da curva fora dos limites de confiança (95%), a faixa linear dinâmica abrangeu todo intervalo de medição. Os valores de limite de quantificação não foram incluídos da faixa linear dinâmica.

4.4.1.3. Sensibilidade

A sensibilidade dos métodos cromatográficos foi avaliada por seis parâmetros que se relacionam entre si: inclinação da curva analítica, S ; sensibilidade analítica, γ ; inverso da sensibilidade analítica, γ^{-1} ; razão sinal-ruído, S/N ; limite de detecção, LD ; e limite de quantificação, LQ .¹⁷⁰ Para a determinação destes parâmetros, exceto a inclinação da curva analítica, 5 replicatas do branco (solvente) foram analisadas para obter o valor do ruído instrumental.

A inclinação da curva analítica, também chamada de sensibilidade, foi determinada pelo coeficiente angular da reta de regressão. A sensibilidade analítica foi obtida pela razão entre a inclinação (S) e o desvio padrão do branco (s), e o inverso desse valor foi calculado para conhecer a menor diferença de concentração que pode ser distinguida na faixa linear dinâmica.

A razão sinal-ruído foi determinada para os pontos de menor e maior concentração da curva. Os limites de detecção (LD) e quantificação (LQ) foram calculados segundo as equações 4 e 5:¹⁷⁸

$$LD = 3,3 \cdot \frac{s}{S} \quad (\text{Eq. 4})$$

$$LQ = 10 \cdot \frac{s}{S} \quad (\text{Eq. 5})$$

4.4.1.4. Precisão

Para os métodos cromatográficos convencionais, a precisão foi verificada apenas pelo desvio padrão relativo das replicatas dos pontos da curva. Já a precisão dos métodos ultrarrápidos (UFGC), a avaliação foi realizada pelos parâmetros: repetibilidade; precisão intermediária; intervalo de confiança; e coeficiente de variação do tempo de retenção (CV_{t_R}). Para determinação da repetibilidade e do intervalo de confiança (equação 6), foram realizadas 10 replicatas nos pontos inferior e superior e 20 replicatas no ponto intermediário, próximo à concentração de interesse (1000 mg kg^{-1}). Na equação 6, \bar{x} é o valor médio da concentração de ponto da curva e t_{n-1} é o valor crítico de t para $(n - 1)$ graus de liberdade.

$$\text{intervalo de confiança da média} = \bar{x} \pm t_{n-1} \cdot \frac{s}{\sqrt{n}} \quad (\text{Eq. 6})$$

Os valores de CV_{t_R} foram determinados pelo desvio padrão do tempo de retenção do surrogate para as 20 replicatas no ponto intermediário.

A precisão intermediária foi obtida pela análise de 7 amostras de mesma concentração, armazenadas à -19°C , em um período de 15 dias. Esta análise permitiu o estudo da estabilidade das soluções armazenadas por 15 dias após o preparo de amostra.

4.4.1.5. Exatidão

A exatidão dos métodos UFGC foi avaliada considerando-se o processo de extração e a quantificação pelas curvas analíticas desenvolvidas. Para cada grupo de analitos, fortificou-se amostras de solo em três níveis de concentração. As amostras utilizadas na fortificação foram a G2, D2, O4 e uma solução BTEX em etanol. Os métodos de extração foram executados com as condições otimizadas e os extratos foram analisados subsequentemente pelos métodos UFGC. Finalmente, a exatidão foi avaliada pelos valores de recuperação, determinados segundo a equação 7.

$$\text{Recuperação (\%)} = \left(\frac{\text{valor determinado}}{\text{valor teórico}} \right) \times 100 \quad (\text{Eq. 7})$$

Os limites de aceitação para recuperação de TPH foram baseados no guia EPA Method 8000C,¹⁷³ cujo intervalo de recuperação sugerido na ausência de valores definidos é de 70% a 130%, e na norma ASTM D7678,¹² a qual estabelece o mesmo intervalo para determinação de TPH, extraído de água subterrânea, e analisado por IR.

4.4.2. Parâmetros analíticos espectroscópicos multivariados

Como citado anteriormente, as figuras de mérito de modelos univariados e multivariados são calculados de formas distintas devido à natureza dos dados e dos métodos de regressão. Certas figuras de mérito de modelos multivariados são calculadas de acordo com o conceito do sinal do analítico líquido, NAS. O objetivo deste conceito é determinar o sinal instrumental referente apenas ao analito em meio ao sinal total gerado pelo instrumento. Desta forma, os parâmetros analíticos dependentes do valor $\|nas\|$ são: seletividade; razão sinal-ruído; sensibilidade analítica; e limites de detecção e quantificação. O conceito NAS encontra-se bem estabelecido na literatura e os cálculos para determinação dos valores escalares de **nas** para os métodos de regressão PLS e PCR podem ser encontrados nas referências.^{159,170,171}

4.4.2.1. Linearidade

A linearidade dos modelos de calibração foi avaliada através dos valores de R^2 , r , intercepto e tendência dos conjuntos de calibração e validação. Ainda, os gráficos de resíduo pela concentração e de valores de referência pelos valores preditos (também conhecido por gráfico de ajuste) foram utilizados para avaliar a linearidade.

Diferentemente dos métodos univariados, não há valores mínimos definidos para o coeficiente de determinação ou correlação para definir a linearidade do modelo de calibração multivariado. Portanto, foram considerados modelos com alta linearidade aqueles que apresentaram coeficiente de correlação do conjunto de validação (r_{val}) maior ou igual a 0,99.

A tendência (bias) foi um importante parâmetro para indicar se o modelo apresentou tendência em prever resultados com valores acima ou abaixo do valor tido como real. A tendência média foi calculada pela soma das diferenças entre o valor de referência e o valor predito, dividido pelo número de amostras (equação 8).

$$\text{bias} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (\text{Eq.8})$$

A significância da tendência dos modelos foi verificada através de um teste t , no nível de confiança de 95%, utilizando o conjunto de validação. Para isto, determinou-se o valor do desvio padrão dos erros de validação (SDV) pela equação 9 e em seguida o valor de t_{bias} pela equação 10.¹⁷⁹

$$\text{SDV} = \sqrt{\frac{\sum [(y_i - \hat{y}_i) - \text{bias}]^2}{n_v - 1}} \quad (\text{Eq. 9})$$

$$t_{\text{bias}} = \frac{|\text{bias}| \sqrt{n_v}}{\text{SDV}} \quad (\text{Eq. 10})$$

Se o valor de t_{bias} calculado foi maior que o valor de t crítico para $(n_v - 1)$ graus de liberdade, o modelo de predição apresentou erros sistemáticos significativos.

4.4.2.2. Sensibilidade

A sensibilidade dos modelos quimimétricos pôde ser avaliada através diversos parâmetros analíticos, como razão sinal-ruído, limites de detecção e quantificação, inclinação da curva de regressão do modelo pseudo-univariado, sensibilidade analítica e o inverso da sensibilidade analítica.

Para determinação da razão sinal-ruído, foram utilizadas 12 replicatas do branco para os modelos desenvolvidos com as soluções e 15 replicatas para os modelos desenvolvidos com os solos. A razão sinal-ruído foi determinada para cada amostra pela divisão do valor $\| \text{nas} \|$ da amostra pelo desvio padrão do branco.

O valor de sensibilidade dos modelos foi obtido pela equação 11 ^{159,170}:

$$\widehat{\text{SEN}} = \frac{1}{\|b\|} \quad (\text{Eq. 11})$$

onde $\|b\|$ é a norma do vetor dos coeficientes de regressão resultante do modelo de calibração. Para proporcionar uma melhor interpretação dos valores de sensibilidade, estes foram convertidos em sensibilidade analítica, dividindo-os pelo desvio padrão do branco, e depois

foram convertidos no inverso na sensibilidade analítica, para conhecer a menor variação de concentração dos analitos distinguível pelos modelos.

Os valores de LD e LQ foram determinados pelas mesmas equações utilizadas nas curvas analíticas univariadas (equações 4 e 5), porém, ao invés da inclinação, utilizou-se o valor de $\widehat{S\hat{E}N}$ como divisor.^{159,170}

4.4.2.3. Seletividade

Seletividade de métodos espectroscópicos multivariados refere-se à quantidade do sinal instrumental relacionada ao analito de interesse. A seletividade dos modelos foi obtida através da média das seletividades das amostras individuais determinadas de acordo com a equação 12:¹⁷⁰

$$\widehat{S\hat{E}L}_i = \frac{nas_i}{\|\mathbf{x}_i\|} \quad (\text{Eq. 12})$$

onde $\|\mathbf{x}_i\|$ é a norma das variáveis espectrais da amostra i .

4.4.2.4. Precisão

A precisão dos modelos foi avaliada através da análise do conjunto de replicatas, constituído de 28 amostras analisadas em nonoplicata mais 3 amostras, incluindo o branco, em 20 replicatas, totalizando em 312 espectros. As 30 amostras e o branco foram utilizados na determinação da precisão de acordo com a equação 13:

$$\text{precisão} = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (\hat{y}_{ij} - \hat{y}_i)^2}{n(m-1)}} \quad (\text{Eq. 13})$$

onde \hat{y}_{ij} é o valor da concentração predita para a replicata j da amostra i ; \hat{y}_i é a concentração média da amostra i ; n é o número de amostras e m é número de replicatas.¹⁷⁰

Além da determinação da precisão pela equação 13, a precisão foi avaliada pelos valores de DPR de cada amostra do conjunto de replicatas em relação à concentração dos analitos. Ainda, a presença de erros sistemáticos foi avaliada através de cartas de controle construídas pela análise das 20 replicatas obtidas de duas amostras de cada conjunto amostral.

4.4.2.5. Exatidão

A avaliação da exatidão dos modelos foi realizada através dos parâmetros obtidos pelo conjunto de calibração: RMSEC, RMSECV, erro relativo médio (equação 14), e diferença relativa percentual (RPD, equação 15); e pelos parâmetros obtidos dos três conjuntos de validação: RMSEP (equação 16), erro relativo médio,¹⁸⁰ RPD,¹⁸¹ coeficiente de correlação e recuperação.

$$\text{Erro relativo médio (\%)} = \left(\frac{\sum_{i=1}^n \left(\frac{\hat{y}_i - y_i}{y_i} \right)}{n} \right) \times 100 \quad (\text{Eq. 14})$$

$$RPD = \frac{\text{desvio padrão das concentrações do conjunto amostral}}{RMSE} \quad (\text{Eq. 15})$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n_v} (y_i - \hat{y}_i)^2}{n_v}} \quad (\text{Eq. 16})$$

Os parâmetros analíticos obtidos pelos conjuntos de calibração foram utilizados como indicativos da adequação do método de regressão, dos pré-processamentos e da seleção de variáveis para a correlação dos sinais instrumentais com a concentração dos analitos. Já a capacidade preditiva dos modelos desenvolvidos foi avaliada através dos parâmetros obtidos pelos conjuntos de validação.

Através dos valores dos erros relativos foi possível avaliar a exatidão em relação à concentração dos analitos, enquanto o erro relativo médio (equação 15) foi um indicativo da exatidão do modelo como um todo.

O parâmetro RPD foi utilizado para avaliar os erros do modelo em relação à faixa de trabalho utilizada. De acordo com Williams e Norris (2001), modelos com valores de RPD acima de 2,5 são considerados com exatidão aceitável e acima de 10,0 excelente.¹³⁴

Finalmente, a exatidão de cada modelo foi ponderada através dos valores de recuperação obtidos da análise do conjunto de validação (II) e dos valores preditos das amostras “reais”.

4.5. Comparação entre os métodos analíticos

Os métodos de determinação de TPH desenvolvidos neste projeto foram comparados entre si e com o método EPA 3550/8015 através dos parâmetros analíticos dos modelos de regressão (exatidão, precisão, linearidade e sensibilidade), frequência analítica, geração de resíduos e custos para implementação dos métodos e realização das análises. Os valores dos equipamentos foram cotados em 2017, através de fornecedores locais, na região de Araraquara. Os valores das vidrarias, solventes e padrões analíticos foram pesquisados em lojas online, em 2018.

4.6. Lista de materiais, equipamentos e softwares

Os materiais e equipamentos utilizados no desenvolvimento deste trabalho estão listados na Tabela 11.

Tabela 11 – Lista de materiais, equipamentos e softwares utilizados no desenvolvimento do trabalho.

Software	Marca	Versão
Matlab	MathWorks	2013a
PLS toolbox	Eigenvector Research Inc.	7.3.1.
OMNIC	Thermo Fisher Scientific	8.0
ChromQuest	Thermo Fisher Scientific	5.0
Excel	Microsoft	2013
Minitab	Minitab Inc.	17
Equipamento	Marca	Modelo
GC-FID	Thermo Fisher Scientific	Trace GC Ultra
Autoamostrador	Thermo Fisher Scientific	Triplus
UFGC-FID	Thermo Fisher Scientific	Trace GC Ultra
Autoamostrador	Thermo Fisher Scientific	AS 3000
Espectrômetro FTIR	Thermo Fisher Scientific	Nicolet 6700
Acessório NIR	Thermo Fisher Scientific	Smart NIR Integration Sphere
Ultrassom	Thornton	T14
Agitador de tubos	Phoenix	AP 56
Coluna cromatográfica	Marca	Dimensões
RTx-1 (100% dimetilpolissiloxano)	Thermo Fisher Scientific	60 m x 0,25 mmDI x 1,0 µm (fase)
DB-1 (100% dimetilpolissiloxano)	Agilent Technology J&W	100 m x 0,25 mmDI x 0,5 µm (fase)
DB-5 (5% fenil em metilpolissiloxano)	Agilent Technology J&W	30 m x 0,53 mmDI x 1,5 µm (fase)
PH RTx-1	Thermo Fisher Scientific	10 m x 0,32 mmDI x 3,0 µm (fase)
PH SE54	Thermo Fisher Scientific	5 m x 0,32 mmDI x 0,05 µm (fase)

(Continua...)

Tabela 11 – Lista de materiais, equipamentos e softwares utilizados no desenvolvimento do trabalho.
(Continuação)

Solventes	Marca	Lote
n-hexano - PA ACS 99%	Merk	K44106667
n-hexano - PA ACS 99%	Êxodo Científica	H28565RA
etanol - PA ACS 99,5%	Êxodo Científica	AE25885RA
diclorometano - PA ACS 99,9%	Sigma-Aldrich	SZBDO37MV
n-hexadecano - PA 99%	Vetec	DCBC4835
Padrões analíticos	Marca	Lote
o-terfenil - PA 99%	Sigma-Aldrich	MKBJ3853V
isobutilbenzeno - PA 99%	Sigma-Aldrich	BCBG6704V
GRO Mix - PA 99,9%	Sigma-Aldrich	LC01984
TPH Mix - PA 99,9%	RTC	16538
EPA 525 PAH Mix A - PA 91,6% a 99,9%	Sigma-Aldrich	XA12140V
Insumos (preparo de amostra)	Marca	Lote
sulfato de sódio - PA ACS 99,0%	Sigma-Aldrich	MKBP3690V
sílica gel (mesh 10-40 µm, ligante CaSO ₄)	Sigma	56H1290
filtro para seringa Xtra PES-500/25	Chromafil	-

5. RESULTADOS E DISCUSSÃO

5.1. Caracterização das amostras de solo e contaminantes

5.1.1. Análises físico-químicas e composicionais dos solos

A caracterização das amostras de solos utilizadas no desenvolvimento e validação dos métodos a serem apresentados foram realizadas de acordo com o Manual de Métodos de Análise de Solo, desenvolvido pela EMBRAPA (1997).¹⁵² Os resultados da caracterização estão apresentados na Tabela 12. As amostras obtidas do monitoramento ambiental de postos de combustíveis estão discriminadas pelas letras “AR” (amostra “real”) e as amostras coletadas em regiões rurais pela letra “S”. As amostras S1, S2 e S3 foram coletadas em mesmo local, porém em diferentes profundidades, assim como as amostras S4 e S5. As análises granulométricas e as classes destas amostras foram realizadas pelo Laboratório de Solos da UEM. As amostras S1, S2 e S3 foram classificadas como latossolo vermelho e as amostras S4 e S5 como nitossolo vermelho.

Tabela 12 – Resultados das análises química, físicas e elementares das amostras de solo.

Amostra	Densidade		Granulometria			
	Aparente	Partícula	Argila total	Silte total	Areia total	Areia Grossa
Unidade	g/cm ³	g/cm ³	%			
AR1	1,13	3,76	23,39	27,78	48,83	20,72
AR2	0,99	2,48	15,73	40,64	43,63	11,54
AR3	1,26	2,51	56,77	8,67	34,55	32,51
S1	1,35	2,70	13,63	3,98	82,40	13,33
S2	1,25	2,51	18,70	3,90	77,40	48,24
S3	1,35	2,99	18,42	3,91	77,66	53,94
S4	1,09	2,72	9,95	49,36	40,69	34,67
S5	1,16	2,57	7,45	37,45	47,82	37,38

Amostra	Análise química			Análise elementar			
	pH	Umidade Atual	Matéria Orgânica	Si	Al	Fe	Ti
Unidade	% (m/m)		% (m/m)	%			
AR1	7,34	25,68	12,7	28,24	0,00	63,55	7,05
AR2	6,41	19,21	11,8	27,25	0,00	63,92	7,63
AR3	7,16	7,09	8,2	31,36	57,80	8,93	1,41
S1	6,50	0,52	2,4	57,20	37,63	3,90	1,03
S2	5,83	0,86	2,6	55,62	39,59	3,87	0,83
S3	6,34	0,64	2,9	49,96	44,69	4,25	0,96
S4	6,07	1,80	7,2	63,97	0,00	30,12	3,96
S5	6,34	1,77	6,4	59,98	0,00	35,15	3,93

Para tornar mais clara a relação entre as amostras de solo e suas características, efetuou-se uma PCA para visualização gráfica. As variáveis foram autoescaladas para proporcionar mesmo peso para todas previamente à modelagem. Utilizando 3 componentes principais, o modelo PCA teve 89% da variância explicada dos dados. A Figura 3 representa o gráfico de loadings e scores para as duas primeiras componentes principais resultantes da PCA. Como esperado, as amostras S formaram dois agrupamentos conforme suas similaridades, enquanto as amostras AR tiveram scores bem distintos das demais amostras. Nota-se que as amostras AR1 e AR2 apresentaram altos teores de umidade atual, matéria orgânica, ferro e titânio em relação às demais. As amostras S1, S2 e S3 se diferenciaram das outras amostras, principalmente, devido ao alto teor de areia total e areia grossa, enquanto a amostra AR3 pelo elevado teor de argila total. As amostras S4 e S5 apresentaram propriedades físico-químicas intermediárias entre as amostras e seu agrupamento foi resultante, principalmente, devido aos baixos teores de argila total e alumínio, pH mais básico, maior teor de silte total e valores intermediários de Fe. Portanto, pode-se concluir que as amostras S formaram dois grupos por similaridade, enquanto as amostras AR foram distintas entre si e das demais.

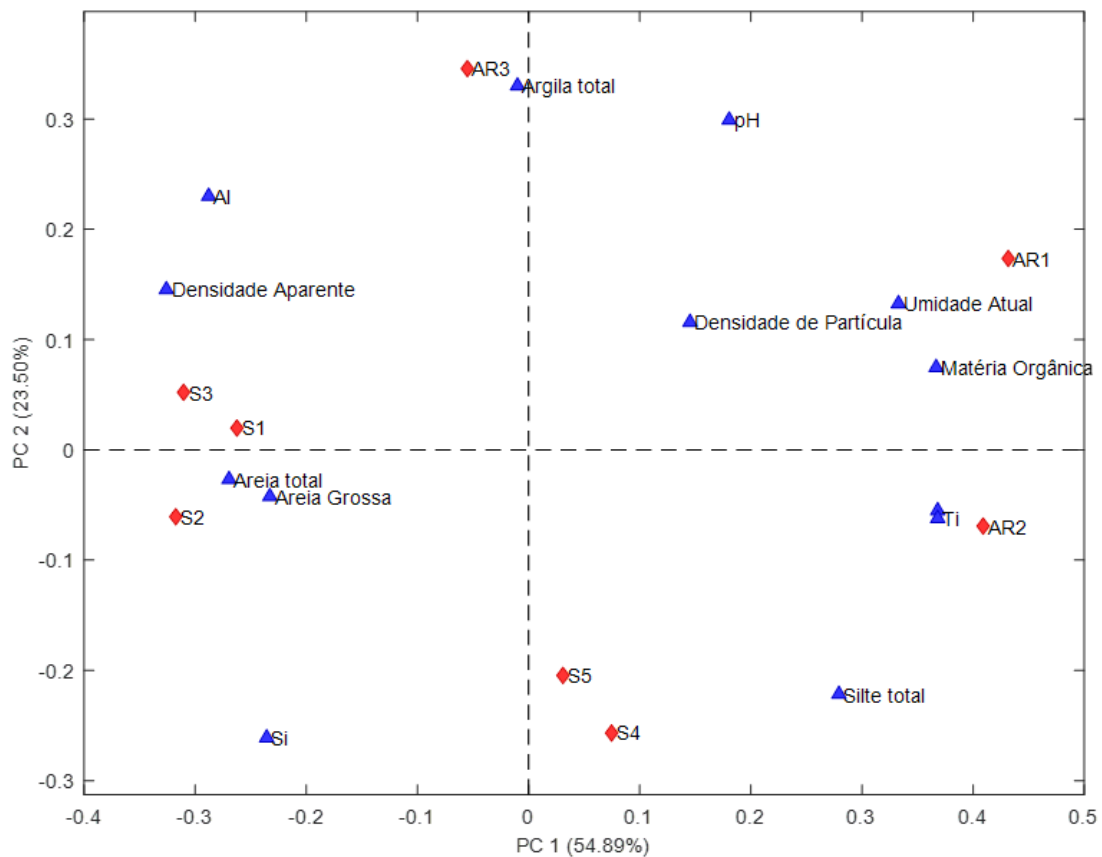


Figura 3 – Gráfico de loadings (Δ) e scores (◇) resultante da PCA das propriedades dos solos.

5.1.2. Espectros NIR das amostras de solo

As amostras de solo não contaminado e de solo contaminado foram analisadas no espectrômetro FT-NIR em quintuplicata e suas médias foram obtidas a fim de realizar uma análise exploratória dos perfis espectroscópicos das amostras. A análise de amostras sólidas por refletância difusa resulta da radiação NIR absorvida pela amostra como também difundida para o ambiente. A radiação espalhada para o ambiente ocasiona em deslocamentos de linha base dos espectros NIR. Como pode ser observado na Figura 4, as amostras AR1 e AR2 possuem linhas de base mais elevadas devido à menor radiação resultante da menor compactação do solo no vial. Consequentemente, menor quantidade de radiação NIR alcança o detector e o sinal espectral médio é elevado. Condizente à literatura, as amostras com maior teor de argila (AR3, S1, S2 e S3) apresentaram espectros com menor intensidade média.²⁸ Diante deste fato, é visível a necessidade do uso de pré-processamentos para corrigir o deslocamento de linha base e normalizar os espectros antes da análise multivariada.

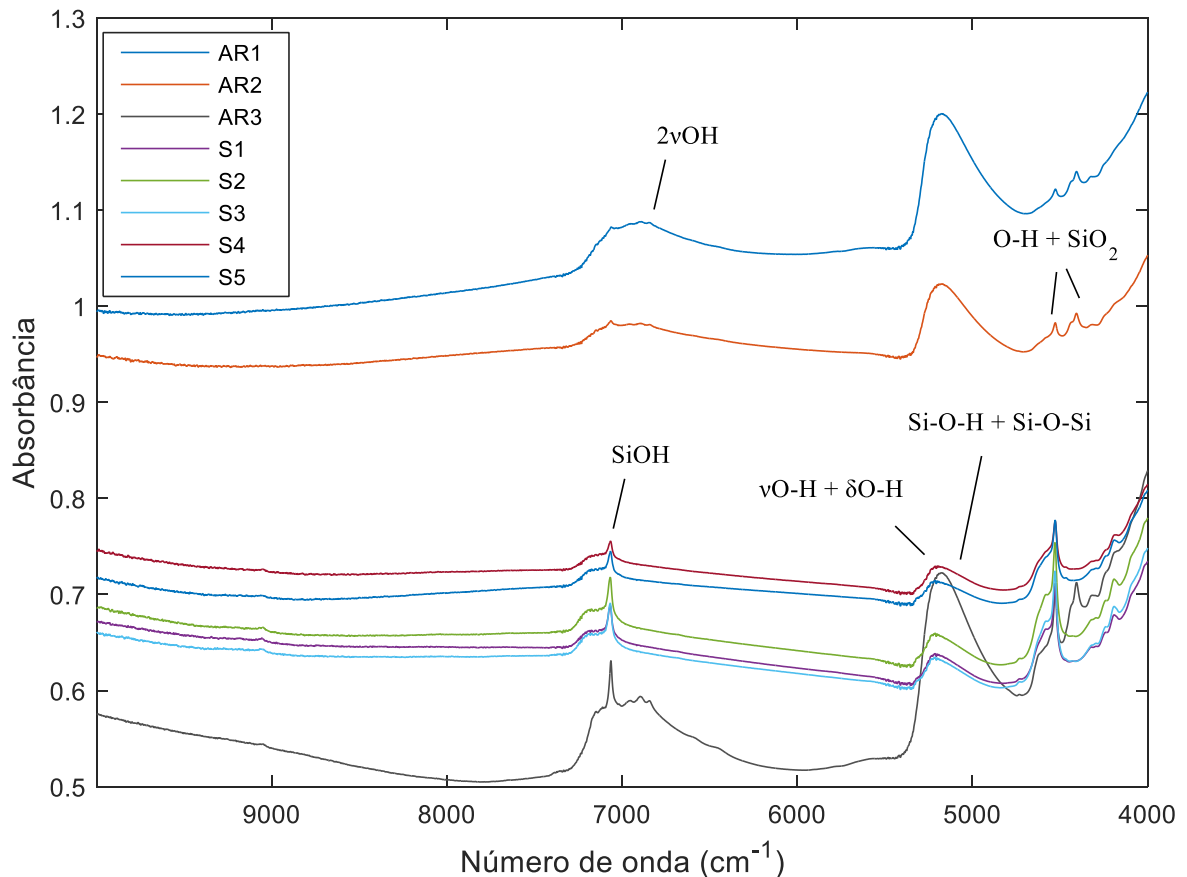


Figura 4 – Espectros NIR das amostras de solo obtidos por refletância difusa.

Segundo a Tabela 12, o principal elemento presente nos solos é o silício. Este elemento encontra-se nos solos majoritariamente como sílica (óxido de silício, SiO_2) e silanol (ácido silícico, SiH_3OH).¹⁸² A sílica e o silanol apresentam sinais característicos na região NIR, contudo, suas bandas são influenciadas pela presença de água devido a seus caracteres hidrofílicos.¹⁸³ O primeiro sobretom do estiramento OH do grupo silanol é em aproximadamente 7200 cm^{-1} , porém, esta banda é alargada e deslocada para número de onda menor no silanol hidratado.¹⁸³ O estiramento do grupo OH em combinação com o estiramento do grupo SiO_2 geram bandas de 4100 cm^{-1} , 4450 cm^{-1} e 4520 cm^{-1} (Figura 4). A atribuição das principais bandas nos espectros das amostras de solo estão apresentadas na Tabela 13.¹⁸³ As vibrações resultantes do estiramento das ligações foram representadas por “v” e da deformação angular por “ δ ”. Os níveis vibracionais foram representados pelos números prefixos às letras gregas, por exemplo, primeiro sobretom representado pelo número “2”, segundo sobretom pelo número “3”, e assim por diante. O tipo de vibração molecular não informada na literatura não foi representada por letras gregas.

Tabela 13 – Atribuições das bandas espectrais das amostras de solo.¹⁸³

Número de onda (cm^{-1})	Atribuição
4191	C-H aromático
4239	combinação $2\nu\text{CH}_2 + \delta\text{CH}_2$
4323	$3\delta\text{C-H}$
4406	Combinação O-H + SiO_2
4526	
5174	Combinação Si-O-H + Si-O-Si
5213	Combinação $\nu\text{O-H} + \delta\text{O-H}$
6889	$2\nu\text{O-H}$
7062	SiOH

A similaridade entre os espectros das amostras foi avaliada através de uma PCA. Os espectros médios foram normalizados pelo pré-processamento SNV e, subsequentemente, centrados na média. O gráfico de scores (Figura 5), gerado a partir da PCA com 3 PCs, mostra a separação entre as amostras AR e S no eixo da PC1. Utilizando gráfico de loadings (não apresentado), observou-se que as variáveis no quadrante positivo da PC1 estiveram relacionadas às bandas de absorção das ligações C-H ($4190\text{-}4400\text{ cm}^{-1}$) e das ligações O-H ($4650\text{-}6900\text{ cm}^{-1}$). Ou seja, a separação das classes se deu pela diferença de teor de água e matéria orgânica. A separação da amostra S5 das demais amostras rurais se relacionou às variáveis da banda do SiOH em 7062 cm^{-1} . Apesar da similaridade das características das amostras S4 e S5, o silício pode estar presente nas amostras na forma de óxido ou silanol e,

portanto, proporcionar diferentes intensidades na referente banda vibracional. Além disso, o padrão de agrupamento observado nos gráficos de escores das Figuras 3 e 5 mostrou que os espectros NIR refletiram as propriedades físico-químicas em seus agrupamentos por PCA.

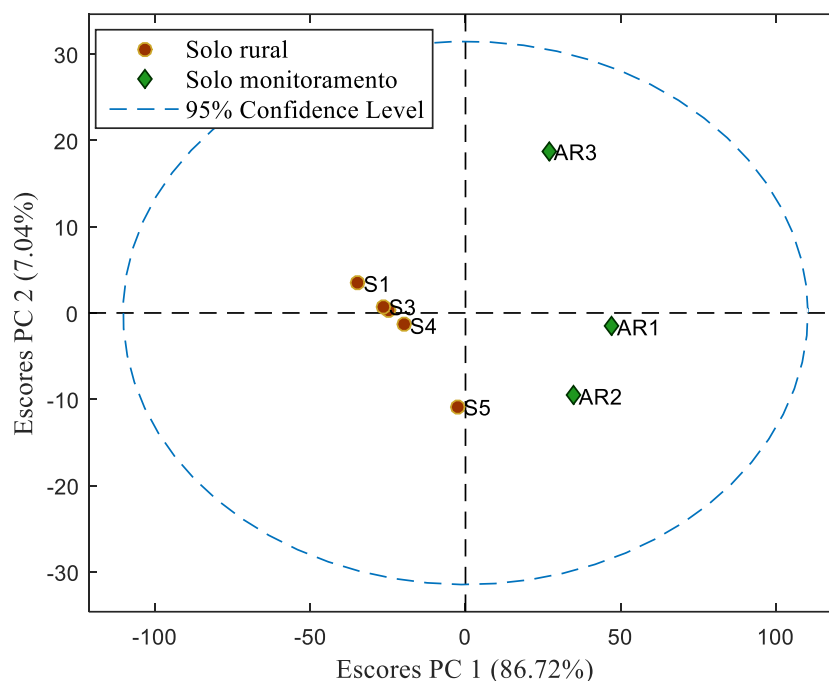


Figura 5 – Gráfico de escores gerado a partir da PCA dos espectros médios dos solos pré-processados por SNV.

5.1.2. Caracterização composicional das amostras de combustível e óleos lubrificantes

Combustíveis fósseis e óleos lubrificantes minerais ou semissintéticos são uma mistura complexa de hidrocarbonetos e outros compostos, além de aditivos melhoradores de desempenho. Portanto, o uso de apenas uma amostra para cada tipo de contaminante não seria representativa para a realidade da contaminação de solos provenientes de postos de combustíveis.

Após a seleção de amostras representativas de uma região do estado de São Paulo, estas amostras foram analisadas cromatograficamente para determinação das frações GRO, DRO e ORO, além de analitos comuns à gasolina, como benzeno, tolueno, etilbenzeno e xilenos (BTEX), e ao óleo diesel, como os hidrocarbonetos policíclicos aromáticos. Os valores determinados das frações e dos analitos individuais foram utilizados posteriormente para o desenvolvimento dos modelos de calibração.

5.1.2.1. Análise cromatográfica das amostras de gasolina

A caracterização composicional das amostras de gasolina foi realizada pelo método cromatográfico adaptado da norma ASTM D5769. A separação dos compostos da respectiva faixa foi avaliada pela análise do padrão GRO Mix e o tempo de retenção (t_R), largura (w), fator de seletividade (α), número de pratos teóricos (N), fator de retenção (k') e resolução (R_s) de cada analito estão apresentados na Tabela 14.

Tabela 14 – Parâmetros de separação do GRO Mix segundo o método ASTM D5759.

Composto	t_R (min)	w (min)	α	N	k'	R_s
3-metilpentano	10,178	0,158		66394	7,64	
benzeno	13,412	0,428	1,359	15712	10,38	7,56
2,2,4-trimetilpentano	15,523	0,330	1,173	35403	12,18	6,40
tolueno	20,045	0,508	1,315	24912	16,01	8,90
etilbenzeno	27,363	0,517	1,388	44819	22,23	14,15
m-xileno	28,030	0,525	1,025	45609	22,79	1,27
o-xileno	29,967	0,578	1,072	43008	24,44	3,35
1,2,4-trimetilbenzeno	38,407	0,523	1,293	86285	31,60	16,14
naftaleno	48,325	0,158	1,266	1496751	40,02	62,77

O método de separação apresentou valores de resolução acima de 3,0, exceto entre os picos do m-xileno e etilbenzeno. Os valores de k' são elevados para os últimos cinco analitos, contudo, amostras de gasolina são misturas complexas que necessitam de análises cromatográficas com alta eficiência. De acordo com os valores de resolução, o método mostrou-se adequado para a caracterização das amostras. Os cromatogramas da análise do GRO Mix (Figura B1) da amostra GasA (Figura B2) estão presentes no Anexo B.

Nas mesmas condições cromatográficas, foram desenvolvidas as curvas analíticas e determinada a concentração dos compostos BTEX, naftaleno e GRO das amostras de gasolina. Os isômeros meta- e para-xilenos coeluíram formando apenas um pico cromatográfico, portanto, ambos foram quantificados conjuntamente pela curva desenvolvida com m-xileno. A quantificação de naftaleno foi realizada através do fator de resposta entre naftaleno e benzeno. O fator de resposta entre os dois analitos foi igual a 1,03.

Os resultados da quantificação da fração GRO de compostos BTEX individuais estão dispostos na Tabela 15. De acordo com os resultados obtidos, houve uma grande diferença das concentrações de BTEX das amostras de gasolina Premium (G1, G6 e G8) em relação às gasolinas Comum e Aditivada. A concentração de tolueno nas amostras Premium foi

consideravelmente maior, enquanto os outros compostos apresentaram concentração até 30 vezes menor que nas demais gasolinas.

Todas gasolinas apresentaram uma quantidade significativa de compostos com menos de 6 átomos de carbono (área integrada entre o pico do solvente e o 3-metilpentano, Área < C₆H₁₄). As amostras, incluindo a gasolina A (GasA), apresentaram média de 12% ± 2% da área total do cromatograma associada a picos com tempo de retenção precedente à faixa GRO (excluindo a área do pico do etanol presente nas gasolinas) e, apesar da amostra G2 apresentar 16% de área, este valor não foi considerado anômalo pelo teste de Grubbs (Z = 2,26; Z_{tabelado} = 2,23). Portanto, o erro relacionado à diferença da composição de hidrocarbonetos de cadeia menor não é significativo na quantificação por calibração externa. Todavia, a metodologia negligencia uma fração de contaminantes em potencial.

Tabela 15 – Parâmetros das curvas analíticas e perfil composicional das amostras de gasolina.

Parâmetro	Benzeno	Tolueno	Etilbenzeno	m,p-xilenos
R ²	0,9996	0,9996	0,9993	0,9992
LD (ug/g)	0,45	0,28	0,17	0,78
LQ (ug/g)	1,35	0,84	0,51	2,37
Amostra	Concentração %(m/m)			
G1	0,10 ± 0,00	23,85 ± 2,14	0,10 ± 0,0	0,13 ± 0,01
G2	0,69 ± 0,02	1,56 ± 0,08	0,41 ± 0,02	1,28 ± 0,09
G3	0,79 ± 0,02	2,89 ± 0,09	0,86 ± 0,03	2,91 ± 0,09
G4	0,78 ± 0,01	1,91 ± 0,04	0,48 ± 0,01	1,45 ± 0,04
G5	0,62 ± 0,02	2,01 ± 0,11	0,56 ± 0,03	1,91 ± 0,11
G6	0,11 ± 0,00	12,77 ± 0,82	0,11 ± 0,00	0,13 ± 0,00
G7	0,65 ± 0,04	3,00 ± 0,30	0,92 ± 0,10	3,33 ± 0,40
G8	0,12 ± 0,00	10,01 ± 0,27	0,11 ± 0,00	0,13 ± 0,00
G9	0,75 ± 0,03	2,42 ± 0,14	0,71 ± 0,05	2,39 ± 0,19
G10	0,64 ± 0,03	2,64 ± 0,04	0,81 ± 0,06	2,83 ± 0,06
GasA	0,53 ± 0,01	2,28 ± 0,08	0,90 ± 0,04	2,85 ± 0,14
Parâmetro	o-xileno	Naftaleno	GRO	Área < C ₆ H ₁₄
R ²	0,9990	0,9996	0,9904	—
LD (ug/g)	0,35	0,45	—	—
LQ (ug/g)	1,05	1,35	—	—
Amostra	Concentração %(m/m)			% Área
G1	0,07 ± 0,00	0,08 ± 0,00	78,51 ± 5,34	10,13
G2	0,50 ± 0,03	0,16 ± 0,01	61,92 ± 1,81	16,43
G3	1,20 ± 0,04	0,24 ± 0,01	77,65 ± 1,85	12,59
G4	0,57 ± 0,02	0,14 ± 0,00	75,86 ± 1,25	11,47
G5	0,81 ± 0,04	0,22 ± 0,01	66,42 ± 2,72	14,19
G6	0,14 ± 0,00	0,19 ± 0,12	71,68 ± 3,59	12,01
G7	1,40 ± 0,17	0,23 ± 0,03	76,44 ± 4,74	10,95
G8	0,10 ± 0,00	< LD	75,25 ± 1,72	11,75
G9	0,99 ± 0,08	0,17 ± 0,06	73,90 ± 3,10	13,35
G10	1,17 ± 0,03	0,23 ± 0,08	78,60 ± 6,45	11,29
GasA	1,21 ± 0,06	0,25 ± 0,02	100	10,18

5.1.2.2. Análise cromatográfica das amostras de diesel

A separação e quantificação da concentração da fração DRO (C₁₀ – C₂₈) e dos compostos PAH foi realizada através da adaptação do método ASTM D6209 a fim de obter uma maior resolução para detecção por FID. A separação dos padrões analíticos e o perfil cromatográfico da amostra de diesel (D2) podem ser visualizados nas Figuras B3, B4, B5 e B6, presentes no Anexo B. Na Tabela 16, encontram-se parâmetros cromatográficos da separação dos padrões analíticos.

Tabela 16 – Parâmetros cromatográficos da análise dos grupos TPH, PAH e ésteres metílicos pelo método ASTM D6902 adaptado.

TPH Mix / PAH Mix / Ésteres - Método Desenvolvido						
Composto	t _r (min)	w (min)	α	N	k'	Rs
Octano	14,95	0,230		67600	3,42	
Decano	19,46	0,162	1,390	230874	4,75	27,84
Dodecano	23,57	0,225	1,256	175580	5,96	18,27
Acenaftileno	24,76	0,25	1,059	156943	6,31	4,76
Fluoreno	27,39	0,177	1,123	383140	7,09	14,86
Tetradecano	27,49	0,300	1,004	134347	7,12	0,33
Hexadecano	30,95	0,172	1,144	518065	8,14	20,12
Fenantreno	31,08	0,185	1,005	451584	8,18	0,70
Antraceno	31,28	0,215	1,007	338670	8,24	0,93
Palmitato de metila	32,63	0,130	1,048	1008016	8,64	10,38
Octadecano	34,10	0,732	1,050	34722	9,07	2,01
Oleato de metila	35,31	0,125	1,039	1276719	9,43	9,68
Linoleato de metila	35,40	0,102	1,003	1927197	9,46	0,88
Estearato de metila	35,71	0,163	1,010	767935	9,55	1,90
Pireno	36,49	0,347	1,024	176933	9,78	2,25
Eicosano	37,13	0,187	1,019	630793	9,97	3,42
Docosano	40,59	0,227	1,103	511572	10,99	15,24
Benzo(a)antraceno	42,56	0,420	1,053	164295	11,57	4,69
Criseno	42,82	0,285	1,007	361180	11,65	0,91
Tetracosano	45,18	0,290	1,060	388344	12,35	8,14
Benzo(b)fluoranteno	51,41	0,423	1,149	236338	14,19	14,73
Hexacosano	51,62	0,427	1,004	233830	14,25	0,49
Benzo(k)fluoranteno	51,70	0,370	1,002	312354	14,27	0,21
Benzo(a)pireno	54,67	0,442	1,062	244779	15,15	6,73
Octacosano	61,19	0,582	1,127	176862	17,08	11,20
Dibenzo(a,h)antraceno	69,23	0,942	1,139	86419	19,45	8,54
Benzo(g,h,i)perileno	69,47	0,197	1,004	1989675	19,52	1,22
Indeno(1,2,3-CD)pireno	71,63	0,718	1,033	159243	20,16	3,01

As resoluções abaixo de 2,0 ocorreram entre isômeros ou hidrocarbonetos de diferentes funções. Os últimos três analitos a eluir apresentaram tempo de retenção superior ao octacosano, portanto não houve a necessidade de maior resolução para estes, pois não são quantificados junto à fração DRO. Os parâmetros das curvas analíticas e as concentrações dos analitos determinados são mostrados na Tabela 17.

Tabela 17 – Parâmetros das curvas analíticas e perfil composicional das amostras de diesel.

Parâmetro	Naftaleno	Acenaftileno	Fluoreno	Fenantreno	Antraceno
R ²	0,9990	0,9928	0,9916	0,9923	0,9910
LD (ug/g)	0,10	0,19	0,78	2,23	1,61
LQ (ug/g)	0,31	0,56	2,36	6,75	4,87
Amostra	Concentração %(m/m)				
D1	0,71 ± 0,04	0,71 ± 0,01	0,87 ± 0,02	0,62 ± 0,02	0,33 ± 0,01
D2	0,77 ± 0,03	0,56 ± 0,01	0,99 ± 0,16	0,82 ± 0,01	0,49 ± 0,01
D3	0,61 ± 0,00	0,43 ± 0,01	0,50 ± 0,01	0,38 ± 0,01	0,23 ± 0,01
D4	0,59 ± 0,01	0,45 ± 0,03	0,56 ± 0,03	0,49 ± 0,03	0,34 ± 0,02
D5	0,67 ± 0,01	0,46 ± 0,02	0,55 ± 0,03	0,43 ± 0,02	0,27 ± 0,02
D6	0,57 ± 0,00	0,56 ± 0,02	0,89 ± 0,04	0,85 ± 0,04	0,42 ± 0,02
D7	0,76 ± 0,02	0,71 ± 0,01	0,95 ± 0,06	0,59 ± 0,01	0,31 ± 0,01
D8	0,59 ± 0,03	0,51 ± 0,01	0,70 ± 0,06	0,56 ± 0,01	0,44 ± 0,02
D9	0,44 ± 0,01	0,47 ± 0,01	0,65 ± 0,01	0,58 ± 0,02	0,37 ± 0,01
D10	0,73 ± 0,02	0,70 ± 0,06	0,92 ± 0,08	0,58 ± 0,05	0,41 ± 0,04
DA	0,69 ± 0,01	0,49 ± 0,00	0,66 ± 0,01	0,54 ± 0,01	0,38 ± 0,01
Parâmetro	Pireno	Benzo(a)antraceno	Criseno	Benzo(b)fluoranteno	DRO
R ²	0,9906	0,9984	0,9933	0,9958	0,9958
LD (ug/g)	3,86	2,69	2,27	62,09	
LQ (ug/g)	11,70	8,16	6,88	188,16	
Amostra	Concentração %(m/m)				
D1	0,17 ± 0,00	0,16 ± 0,01	0,08 ± 0,01	0,13 ± 0,00	86,68 ± 0,97
D2	0,38 ± 0,00	0,22 ± 0,02	0,11 ± 0,00	0,09 ± 0,02	93,27 ± 0,22
D3	0,32 ± 0,01	0,35 ± 0,02	0,23 ± 0,04	0,29 ± 0,01	68,08 ± 1,29
D4	0,33 ± 0,00	0,32 ± 0,01	0,17 ± 0,01	0,24 ± 0,02	74,91 ± 3,12
D5	0,34 ± 0,01	0,31 ± 0,01	0,15 ± 0,12	0,25 ± 0,01	73,19 ± 3,36
D6	0,37 ± 0,01	0,23 ± 0,01	0,10 ± 0,01	0,10 ± 0,00	88,71 ± 2,40
D7	0,11 ± 0,00	0,12 ± 0,01	0,07 ± 0,06	0,08 ± 0,00	79,10 ± 1,74
D8	0,34 ± 0,06	0,42 ± 0,10	0,16 ± 0,06	0,10 ± 0,02	79,15 ± 0,00
D9	0,44 ± 0,01	0,34 ± 0,00	0,17 ± 0,04	0,21 ± 0,00	75,56 ± 1,99
D10	0,23 ± 0,00	0,14 ± 0,02	0,06 ± 0,01	0,06 ± 0,00	92,62 ± 3,57
DA	0,35 ± 0,02	0,37 ± 0,02	0,10 ± 0,11	0,30 ± 0,03	100,00

As concentrações de todos PAH analisados apresentaram-se abaixo de 1 % (m/m) nas amostras de diesel, sendo os analitos de maior e menor concentração, respectivamente, fluoreno e criseno. As amostras de diesel apresentaram média da fração DRO igual a $81\% \pm 9\%$ (m/m).

A relação entre as áreas dos picos dos ésteres é um indicativo da matriz utilizada na produção do biodiesel adicionado ao diesel. O linoleato de metila é um dos produtos originados da reação de transesterificação do óleo de soja com metanol, o palmitato e estearato de metila estão principalmente relacionados ao sebo bovino e o oleato de metila origina-se de ambas matrizes.¹⁸⁴ Estas matérias-primas são as mais utilizadas na produção de biodiesel no Brasil⁷⁷ e, portanto, espera-se que maior concentração de linoleato de metila na amostra de diesel esteja relacionada à adição de biodiesel de soja, enquanto, palmitato e estearato de metila à adição de biodiesel de sebo. Dependendo da viabilidade econômica, é comum ocorrer a mistura de biodiesel de ambas matrizes no diesel comercializado. A Figura 6 mostra a relação entre as áreas dos quatro ésteres identificados na análise cromatográfica. Os ésteres relacionados ao sebo mostraram-se evidentes principalmente nas amostras D3 e D5, enquanto as demais amostras apresentaram maior área referente ao linoleato de metila. A amostra D5, provavelmente, possui uma mistura de biodiesel de ambas matrizes. Portanto, as amostras de diesel selecionadas para o desenvolvimento dos métodos foram representativas em relação à fonte de biodiesel.

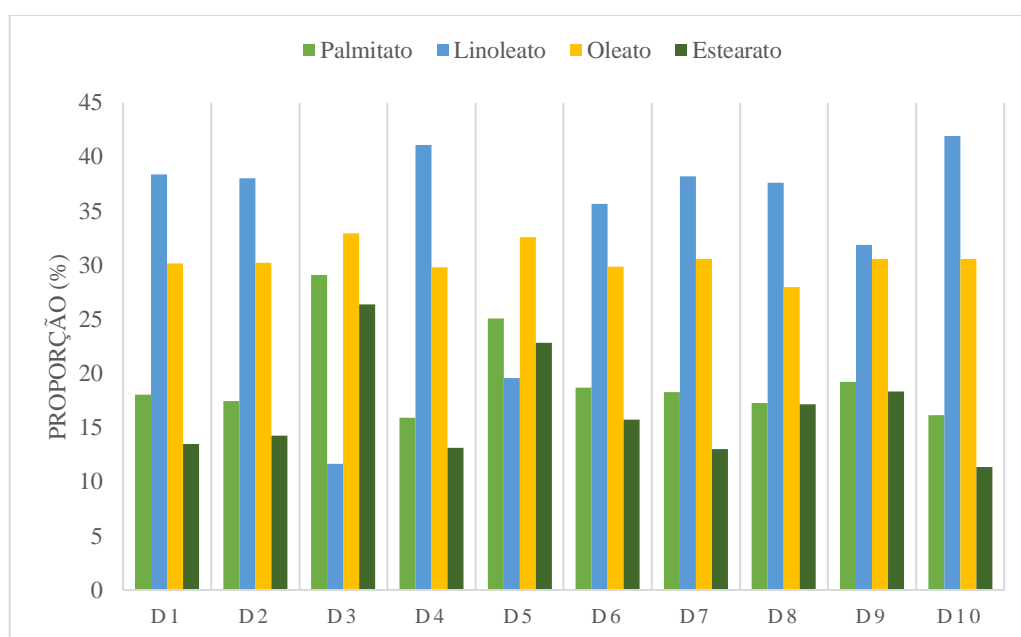


Figura 6 – Relação entre as áreas dos ésteres presentes nas amostras de diesel.

5.1.2.3. Análise cromatográfica dos óleos lubrificantes

A fração ORO dos óleos lubrificantes não foi determinada através da cromatografia gasosa convencional, pois a eluição da maior quantidade dos compostos ORO ocorreu juntamente ao drift presente nos cromatogramas durante a última rampa de aquecimento. Desta forma, a fração ORO foi determinada pelo método UFGC cujos parâmetros cromatográficos estão descritos seção 4.1.2.1. O cromatograma da análise do MRC TPH Mix está disponível no Anexo B (Figura B7) e os parâmetros cromatográficos relacionados serão apresentados na discussão dos métodos cromatográficos ultrarrápidos.

Os resultados da determinação da fração ORO das dez amostras de óleo lubrificante estão presentes na Tabela 18. Observou-se que apenas as amostras O4, O5 e O6 apresentaram a fração de compostos ORO acima de 80 %(m/m). Os compostos eluídos fora da faixa ORO tiveram tempo de retenção dentro da faixa DRO, isto é, a composição dos óleo lubrificantes apresentaram até 44% em massa de compostos semivoláteis.

Tabela 18 – Concentração da faixa de compostos ORO das amostras de óleo lubrificante.

Conc. %(m/m)	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
ORO	56,08 ± 4,250	69,31 ± 2,470	58,81 ± 0,500	86,18 ± 2,308	83,50 ± 1,220	80,48 ± 2,272	66,84 ± 0,389	75,20 ± 0,866	68,92 ± 1,243	59,65 ± 0,355

As concentrações dos analitos e das frações de TPH determinados no procedimento de caracterização das amostras de combustíveis e óleos lubrificantes foram utilizadas nos modelos de predição destes analitos.

5.1.2.4. Análises NIR das amostras de combustíveis e óleos lubrificantes

Uma vez que os combustíveis fósseis e os óleos lubrificantes minerais são misturas complexas compostas majoritariamente por hidrocarbonetos, esperou-se os espectros NIR das amostras dos contaminantes fossem ricas em informações espectrais devido à sensibilidade para ligações C-H. Embora a região NIR seja complexa em consequência de combinações de vibrações moleculares e sobretons, entre outros fenômenos que geram deslocamento de bandas, notou-se algumas bandas características de grupos funcionais. Na Figura 7, estão dispostos os espectros das trinta amostras de contaminantes e as principais bandas vibracionais. A atribuição das demais bandas identificadas nos espectros está disponível na Tabela 19.

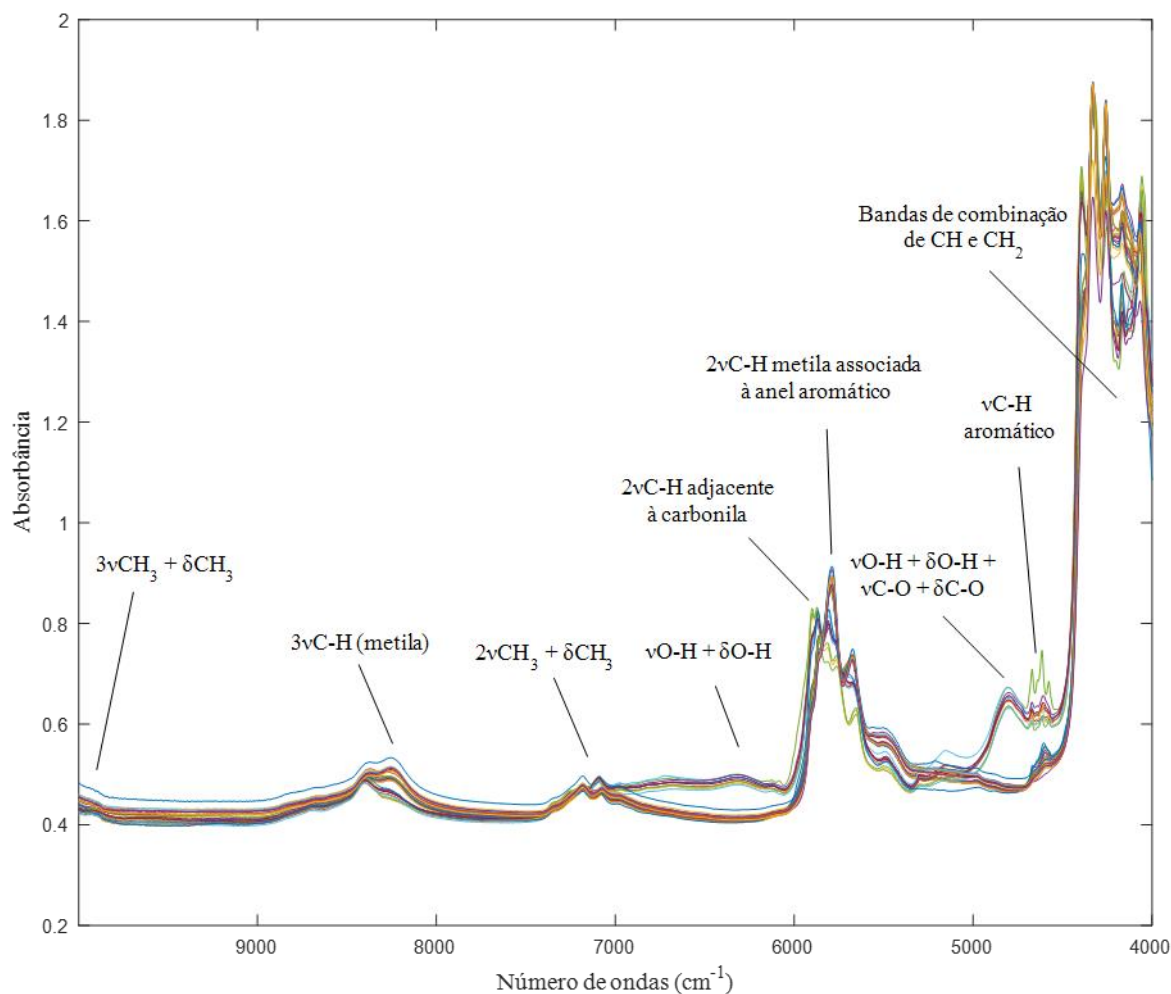


Figura 7 – Espectros NIR das 30 amostras dos contaminantes (gasolinas, óleos diesel e óleos lubrificantes).

A região espectral entre $4000\text{-}4400\text{ cm}^{-1}$ apresentou bandas mais intensas devido à combinação de transições no nível fundamental do estiramento e de deformações angulares entre átomos de carbono e hidrogênio provenientes de hidrocarbonetos aromáticos e alifáticos. As bandas entre $4576\text{-}4669\text{ cm}^{-1}$ foram resultantes do estiramento da ligação C-H de anéis aromáticos. A amostra que apresentou bandas mais intensas nesse intervalo foi a amostra G1, pois o teor de tolueno é superior em relação às demais (Tabela 15). A região entre $5654\text{-}5874\text{ cm}^{-1}$ apresentou um conjunto de bandas oriundas do primeiro sobretom do estiramento e deformação angular do grupo CH aromático e alifático. Nesta região, é possível distinguir as bandas de absorção de diferentes compostos aromáticos, como BTEX.¹⁸³ A combinação do estiramento e da deformação angular do CH_3 resultou nas bandas próximas à 7100 cm^{-1} (primeiro sobretom) e à 9900 cm^{-1} (segundo sobretom). A banda em 8395 cm^{-1} é um indicativo qualitativo do teor de hidrocarbonetos ramificados, dado que esta banda se refere ao estiramento da ligação C-H (segundo sobretom).

Tabela 19 – Atribuições das bandas espectrais das amostras de contaminantes.¹⁸³

Número de onda (cm ⁻¹)	Atribuição
4055	combinação $\delta\text{CC}+\nu\text{CH}$ de anel aromático
4157	combinação $\nu\text{CH} + \delta\text{CH}$ de anel aromático
4252	combinação $\nu\text{CH} + \delta\text{CH}$ de anel aromático
4322	combinação $2\nu\text{CH}_2$ e δCH_2
4393	combinação $\nu\text{C-H} + \delta\text{C-H}$
4576	$\nu\text{C-H}$ aromático
4614	$\nu\text{C-H}$ aromático
4669	$\nu\text{C-H}$ aromático
4805	combinação $\nu\text{O-H} + \delta\text{O-H} + \nu\text{C-O} + \delta\text{C-O}$
5144	combinação $\nu\text{O-H}$ e $\delta\text{O-H}$
5654	$2\nu\text{C-H}$ metila associada à anel aromático
5764	$2\nu\text{C-H}$ metila associada à anel aromático
5874	$2\nu\text{C-H}$ hidrocarboneto ramificado
5899	$2\nu\text{C-H}$ adjacente à carbonila
6314	combinação $\nu\text{O-H}$ e $\delta\text{O-H}$
6674	$2\nu\text{O-H}$
7090	combinação $2\nu\text{CH}_3$ e δCH_3 $2\nu\text{O-H}$ (álcool alifático)
8395	$3\nu\text{C-H}$ (metila)
8703	$3\nu\text{C-H}$ aromático
9899	combinação $3\nu\text{CH}_3$ e δCH_3

As principais diferenças espectrais entre os grupos de contaminantes estiveram relacionadas às bandas vibracionais do O-H, proveniente do etanol presente na gasolina; às bandas do estiramento do CH adjacente à carbonila, oriundos do biodiesel dos óleos diesel; e ao segundo sobretom do estiramento do CH₃, que apresentou maior intensidade nas amostras de óleo lubrificante. Essas diferenças espectrais indicaram a possibilidade da classificação do tipo de contaminante presente na matriz utilizando os espectros NIR e métodos de reconhecimento de padrão.

Além da observação das bandas de absorção dos contaminantes, realizou-se uma PCA para avaliar as similaridades entre os espectros das amostras. O modelo PCA foi desenvolvido com os dados normalizados pelo pré-processamento MSC e, subsequentemente, centrados na média. Observou-se a disposição das amostras no gráfico de scores com 3 PCs (Figura 8). No eixo da PC3, houve um agrupamento das amostras de gasolina premium (G1, G6 e G8) distante das demais amostras, enquanto a gasolina A (GasA) se isolou das demais no eixo da PC2. As amostras de diesel formaram um agrupamento denso resultante da similaridade espectral entre elas. Nos eixos da PC1 e PC2, as amostras O1 e O3 se aproximam das amostras de diesel,

resultado do baixo teor de ORO, enquanto a amostra O9 se aproxima do diesel no eixo da PC3, podendo estar relacionado à baixa viscosidade da amostra O9 (SAE 10W30).

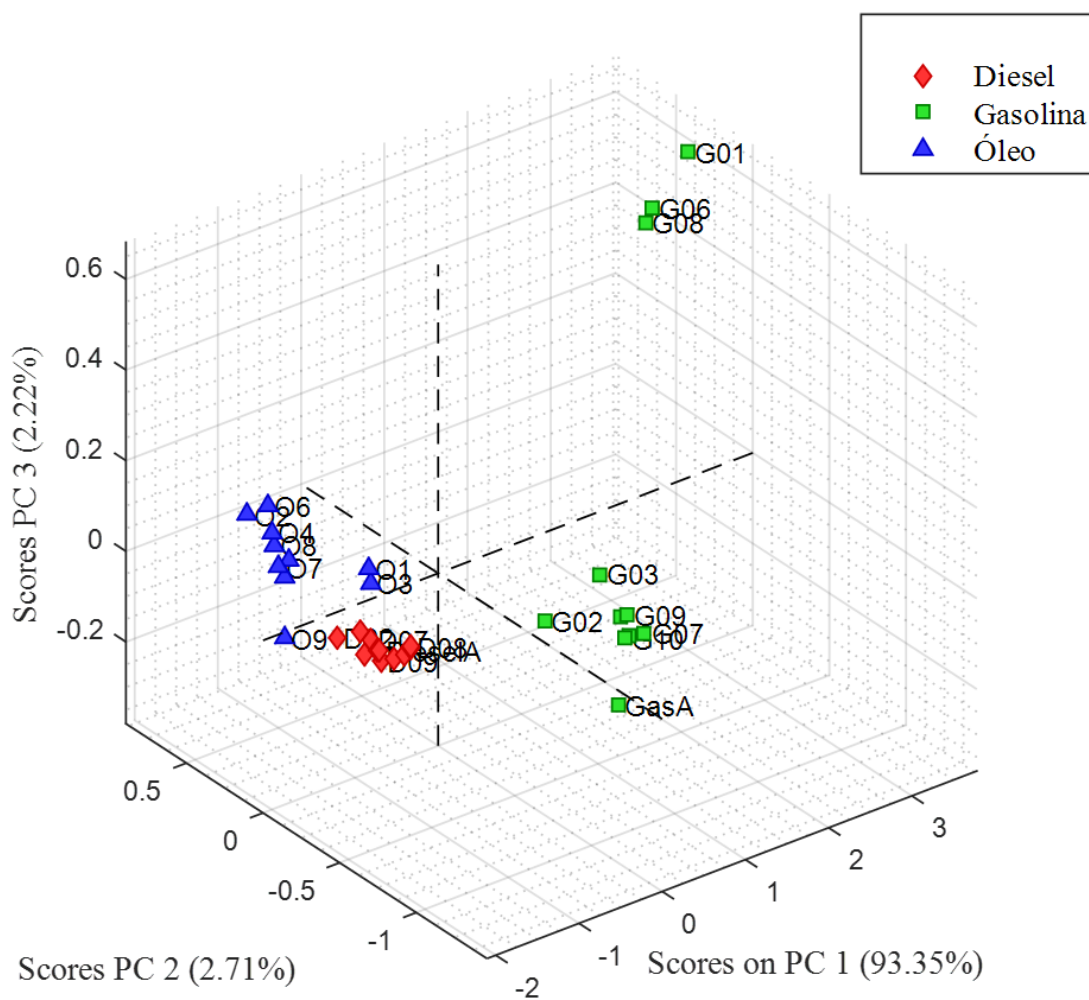


Figura 8 – Gráfico de escores gerado a partir da PCA dos espectros dos contaminantes.

5.2. Métodos de referência

5.2.1. Métodos determinativos por cromatografia gasosa convencional

Previamente ao desenvolvimento da curva analítica pelo método EPA 8015, avaliou-se o tempo de incubação, a fim de reduzir o tempo de 50 min, recomendado pelo método EPA 5021. Os valores de área média normalizada e DPR das análises GRO para cada tempo de incubação podem ser observados na Figura 9. Notou-se que 5 min de incubação geraram maior sinal, consequente de maior extração dos compostos GRO, e menor DPR, proporcionando maior repetibilidade para o método determinativo.

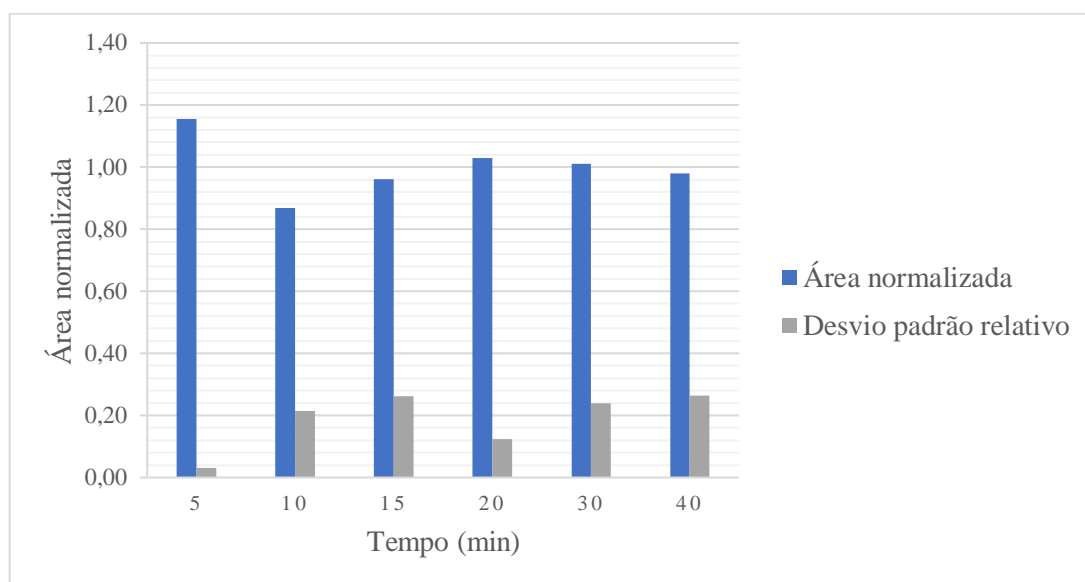


Figura 9 – Avaliação do tempo de incubação na extração de compostos GRO por headspace.

Em relação ao tempo total de análise, o tempo de retenção do último analito da faixa GRO (naftaleno) mostrou que o tempo de 39 min, conforme o método, era desnecessário. Desta maneira, aumentou-se a taxa de aquecimento da segunda rampa para $18\text{ }^{\circ}\text{C min}^{-1}$ até $275\text{ }^{\circ}\text{C}$, a qual foi mantida por 2 min para eluição de possíveis analitos de maior peso molecular, e o método passou a ser executado em 25 min. O cromatograma do MCR GRO mix analisado pelo método EPA 8015C está representado na Figuras B8 (Anexo B) e os parâmetros cromatográficos na Tabela 20.

Tabela 20 – Parâmetros cromatográficos da separação do GRO Mix pelo método EPA 8015.

Composto	t_R (min)	w (min)	α	N	k'	R_s
3-metilpentano	2,319	0,163		3226	0,75	
benzeno	3,385	0,190	2,070	5078	1,56	5,61
2,2,4-trimetilpentano	3,794	0,220	1,199	4759	1,87	1,86
tolueno	5,537	0,293	1,705	5702	3,18	5,94
etilbenzeno	8,114	0,220	1,611	21763	5,13	11,71
m-xileno	8,339	0,313	1,033	11334	5,30	0,72
o-xileno	9,058	0,370	1,102	9590	5,84	1,94
1,2,4-trimetilbenzeno	12,133	0,330	1,397	21627	8,17	9,32
naftaleno	16,934	0,210	1,444	104040	11,79	22,86

Após a determinação da resolução dos picos, a curva analítica para quantificação de GRO foi desenvolvida. Os parâmetros analíticos da curva analítica estão dispostos na Tabela 21. Segundo a norma EPA 8015, o valor de LD para GRO é de 5 mg kg^{-1} , portanto as curvas analíticas atenderam aos critérios das normas EPA, exceto pelo valor de R^2 abaixo de 0,99.

Tabela 21 – Parâmetros das curvas analíticas dos métodos de referência para GRO e DRO.

Parâmetro	GRO	DRO
Intervalo de medição (mg kg^{-1})	672 a 6090	111 a 1161
Razão sinal/ruído	143	0,15
R^2	0,9178	0,9976
R^2 ajustado	0,8904	0,9969
r	0,9580	0,9988
Inclinação	1612997	348195
Intercepto	$6,74 \cdot 10^8$	$-1,61 \cdot 10^7$
RMSE (mg kg^{-1})	750	23
Faixa linear dinâmica (mg kg^{-1})	672 a 6090	111 a 1161
LD (mg kg^{-1})	0,04	20
LQ (mg kg^{-1})	0,12	60
Sensibilidade	1612997	348195
Sensibilidade analítica (mg kg^{-1})	83,49	0,17
γ^{-1} (mg kg^{-1}) ⁻¹	0,01	6,01
Precisão		
DPR menor concentração (%)	4,00	6,30
DPR maior concentração (%)	2,58	7,01

A baixa linearidade obtida na regressão pôde ser avaliada através da curva analítica na Figura 10, resultante da análise de 5 pontos em triplicatas reais. Observou-se que o primeiro ponto apresentou tendência negativa, a qual pode ter sido ocasionada por uma baixa extração

em menores níveis de concentração, e os resíduos apresentaram uma tendência decrescente a partir do segundo ponto da curva até ser negativa no ponto 5. Dado que a linearidade do método de extração por headspace depende do equilíbrio entre a fase líquida e a gasosa, a área dos picos cromatográficos passam a não aumentar quando ocorre a saturação da fase gasosa. Portanto, é possível que tenha ocorrido a saturação do headspace a partir de 3000 mg kg⁻¹ de GRO através das configurações utilizadas no método de extração.

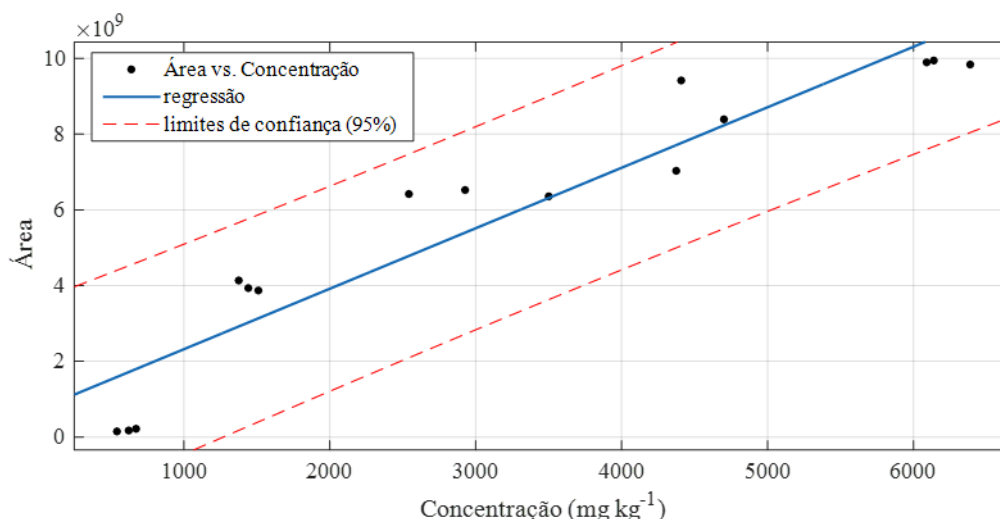


Figura 10 – Gráfico de resíduos da curva analítica para GRO pelo método EPA 5021/8015.

Quanto à separação dos hidrocarbonetos alifáticos do TPH Mix, o método EPA 8015 proporcionou boas resoluções (acima de 4,5). O cromatograma da análise de TPH Mix pelo método convencional pode ser observado na Figura B9 (Anexo B) e os parâmetros cromatográficos na Tabela 22.

Tabela 22 – Parâmetros cromatográficos da análise do TPH Mix pelo método EPA 8015.

Composto	t _R (min)	w (min)	A	N	k'	Rs
octano	6,275	0,191		17328	3,74	
decano	9,514	0,210	1,654	32947	6,19	15,45
dodecano	12,161	0,186	1,323	68638	8,19	14,25
tetradecano	14,445	0,202	1,211	81546	9,91	11,29
hexadecano	16,472	0,178	1,154	136499	11,44	11,37
octadecano	18,296	0,138	1,120	282602	12,82	13,25
eicosano	19,954	0,152	1,098	276949	14,07	10,93
docosano	21,469	0,261	1,081	107983	15,22	5,80
tetracosano	22,961	0,179	1,074	264250	16,35	8,35
hexacosano	25,010	0,262	1,095	145792	17,89	7,82
octacosano	28,096	0,672	1,130	27942	20,23	4,59
triacontano	32,897	0,819	1,179	25815	23,85	5,86
dotriacontano	37,128	0,558	1,134	70837	27,05	7,58
tetratriacontano	41,662	0,839	1,127	39453	30,47	5,40
hexatriacontano	48,197	1,063	1,162	32892	35,41	6,15
octatriacontano	57,791	1,511	1,205	23415	42,66	6,35
tetracontano	71,993	2,248	1,252	16405	53,39	6,32

Todos os analitos apresentaram separação adequada, porém houve a necessidade de expandir o tempo de permanência da temperatura final em mais 35 min para eluir os quatro últimos analitos. A curva analítica para quantificação de DRO apresentou R^2 e r iguais a 0,99 e valores de RMSE relativamente baixos (Tabela 21). Porém o método não apresentou boa sensibilidade, levando a altos valores de LQ (60 mg kg^{-1}) e γ^{-1} igual a 6 mg kg^{-1} . O problema de sensibilidade pode ser contornado pelo aumento do volume de amostra injetada no cromatógrafo, de forma a não saturar a coluna capilar. Todavia, o método atendeu aos critérios recomendados pela EPA.

5.2.2. Métodos determinativos por cromatografia gasosa ultrarrápida

A cromatografia gasosa ultrarrápida (UFGC) visa diminuir drasticamente o tempo de análise, economizando gases e o consumo de energia. Ao contrário da GC convencional, em que o forno é aquecido por gradiente, na UFGC o aquecimento é feito por fio de resistência acoplado diretamente na coluna, possibilitando maiores taxas de aquecimento, com menor gasto energético. Assim a diminuição do tempo total de análise também favorece o aumento da frequência analítica.

5.2.2.1. Separação das frações TPH e dos compostos BTEX

A partir de análises prévias, verificou-se a impossibilidade do uso de uma mesma coluna cromatográfica para a análise das frações voláteis e semivoláteis (pela avaliação das colunas ultrarrápidas disponíveis). O uso da coluna de menor comprimento (PH SE54) proporcionou baixa resolução entre os primeiros analitos da fração da gasolina e os primeiros eluentes apresentaram tempo de retenção muito próximo ao solvente, tornando inadequado o uso desta coluna para determinação de GRO e BTEX. Já a coluna de maior comprimento e espessura de filme (PH RTX-1) gerou bandas largas ao picos dos hidrocarbonetos acima de C_{10} e o tempo de análise necessário para abranger toda a faixa de hidrocarbonetos foi incondizente com a característica de UFGC. Por outro lado, as colunas PH RTX-1 e PH SE54 proporcionaram boas separações, respectivamente, para os compostos GRO e DRO/ORO.

Os parâmetros cromatográficos obtidos a partir do métodos UFGC definidos estão presentes na Tabela 23 e os cromatogramas dos materiais de referência TPH Mix e GRO Mix

estão apresentados, respectivamente, nas Figuras B7 e B10, no Anexo B. A resolução dos picos revelou a separação adequada dos analitos para a determinação de TPH e BTEX em amostras solo e, em relação ao método EPA 8015, o tempo de análise foi 4,5 vezes menor para GRO e 17 vezes menor para DRO/ORO.

Tabela 23 – Parâmetros cromatográficos das análises do GRO Mix e TPH Mix pelos métodos UFGC.

GRO Mix - Coluna PH RTX-1						
Composto	t_R (min)	w (min)	α	N	k'	Rs
3-metilpentano	1,355	0,117		2146	3,21	
benzeno	1,805	0,097	1,435	5540	4,61	4,64
2,2,4-trimetilpentano	2,047	0,158	1,163	2686	5,36	1,53
tolueno	2,513	0,097	1,270	10739	6,81	4,80
etilbenzeno	3,155	0,07	1,293	32503	8,81	9,17
m-xileno	3,207	0,082	1,018	24473	8,97	0,63
o-xileno	3,317	0,085	1,038	24365	9,31	1,29
1,2,4-trimetilbenzeno	3,587	0,097	1,090	21880	10,15	2,78
naftaleno	3,863	0,055	1,085	78930	11,01	5,02
TPH Mix - Coluna PH SE54						
Composto	t_R (min)	w (min)	α	N	k'	Rs
decano	0,422	0,085		394	2,15	
dodecano	0,785	0,070	2,261	2012	4,86	5,19
tetradecano	1,110	0,095	1,499	2184	7,28	3,42
hexadecano	1,400	0,077	1,297	5289	9,44	3,77
octadecano	1,663	0,063	1,208	11149	11,41	4,17
eicosano	1,903	0,077	1,157	9773	13,20	3,12
docosano	2,123	0,050	1,124	28846	14,84	4,40
tetracosano	2,332	0,060	1,105	24170	16,40	3,48
hexacosano	2,525	0,102	1,088	9805	17,84	1,89
octacosano	2,705	0,068	1,075	25318	19,18	2,65
triacontano	2,880	0,098	1,068	13818	20,49	1,79
dotriacontano	3,057	0,140	1,064	7629	21,81	1,26
tetratriacontano	3,248	0,113	1,065	13219	23,23	1,69
hexatriacontano	3,455	0,263	1,066	2761	24,78	0,79
octatriacontano	3,727	0,272	1,082	3004	26,80	1,00
tetracontano	4,127	0,377	1,111	1917	29,79	1,06

5.2.2.2. Desenvolvimento de curvas analíticas e validação do método UFGC

Após definição das colunas e métodos cromatográficos, foram desenvolvidas as curvas analíticas como descrito no item 4.2.2. Os parâmetros das curvas analíticas para determinação de GRO extraídos com hexadecano e etanol estão, respectivamente, nas Tabelas 24 e 25.

Tabela 24 – Parâmetros das curvas analíticas para determinação de GRO em hexadecano.

Parâmetro	GRO	benzeno	tolueno	etilbenzeno	m-xileno	o-xileno	isobutilbenzeno
Linearidade							
Intervalo de medição (mg kg ⁻¹)	126 a 13599	0,6 a 832,2	0,8 a 1027,2	0,9 a 1144,2	0,8 a 1105,2	0,8 a 1092,2	0,9 a 1209,3
R ²	0,9859	0,9995	0,9993	0,9990	0,9990	0,9990	0,9986
R ² ajustado	0,9848	0,9995	0,9992	0,9990	0,9990	0,9989	0,9985
r	0,9929	0,9997	0,9996	0,9995	0,9995	0,9995	0,9993
Inclinação	29712	24777	20524	18345	18641	18828	14325
Intercepto	4443485	-78053	-58581	-71596	-72579	-67636	-54498
RMSE (mg kg ⁻¹)	537	6,0	8,7	11,3	10,9	11,1	14,6
Faixa linear dinâmica (mg kg ⁻¹)	126 a 13599	0,6 a 832,2	0,8 a 1027,2	0,9 a 1144,2	0,8 a 1105,2	0,8 a 1092,2	0,9 a 1209,3
LD (mg kg ⁻¹)	38	0,04	0,06	0,20	0,15	0,06	0,19
LQ (mg kg ⁻¹)	117	0,12	0,19	0,60	0,46	0,18	0,58
Sensibilidade	29712	24897	20690	18527	18823	19017	14507
Sensibilidade analítica (mg kg ⁻¹)	0,09	86,2	52,1	16,7	21,9	57,0	17,5
γ^{-1} (mg kg ⁻¹) ⁻¹	11,11	0,01	0,02	0,06	0,05	0,02	0,06
Razão Sinal/Ruído (mín)	16	73	114	35	28	84	54
Razão Sinal/Ruído (máx)	1070	71296	53337	19075	24180	62181	21106
Precisão							
DPR menor concentração (%)	28,01	16,10	8,45	8,52	6,66	12,12	5,41
DPR maior concentração (%)	30,50	6,95	7,83	8,77	8,99	9,20	10,83
Repetibilidade (%)	56,09	3,00	3,97	4,35	4,43	4,58	5,27
CV t _R (%)	0,03	0,15	0,10	0,09	0,07	0,07	0,04
Precisão intermediária^a							
DPR baixa concentração	29,15% (394)	3,67% (8,4)	5,54% (13,2)	6,51% (14,5)	6,09% (14,4)	7,34% (16,3)	8,80% (17,0)
DPR alta concentração	5,27% (1659)	3,86% (64,7)	5,77% (93,9)	7,30% (110,9)	7,65% (108,0)	7,75% (108,8)	9,27% (123,7)
Intervalo de confiança							
C inferior (mg kg ⁻¹)	126 ± 36	4,0 ± 0,1	5,1 ± 0,1	6,0 ± 0,1	5,2 ± 0,1	5,1 ± 0,1	7,0 ± 0,1
C intermediária (mg kg ⁻¹)	1315 ± 246	69,8 ± 11,6	91,8 ± 15,6	105,2 ± 17,8	102,2 ± 17,3	101,5 ± 17,3	113,2 ± 19,6
C superior (mg kg ⁻¹)	13599 ± 2721	833,8 ± 41,3	1035,2 ± 57,8	1156,9 ± 72,3	1117,8 ± 71,5	1104,8 ± 72,4	1228,0 ± 94,8

^a Os valores entre parênteses referem-se à concentração (mg kg⁻¹) das respectivas amostras analisadas.

Tabela 25 – Parâmetros das curvas analíticas para determinação de GRO em etanol.

Parâmetro	GRO	benzeno	tolueno	etilbenzeno	m-xileno	o-xileno	isobutilbenzeno
Linearidade							
Intervalo de medição (mg kg ⁻¹)	99 a 9926	0,5 a 854,0	0,6 a 1138,6	0,6 a 1125,7	0,6 a 1151,6	0,7 a 1164,5	0,6 a 1099,8
R ²	0,9999	0,9999	1,0000	1,0000	1,0000	1,0000	0,9984
R ² ajustado	0,9999	0,9999	1,0000	1,0000	1,0000	1,0000	0,9983
r	0,9999	0,9999	1,0000	1,0000	1,0000	1,0000	0,9992
Inclinação	74483	108888	103352	106512	105027	105836	104035
Intercepto	2689920	44032	50905	15398	31623	25389	390775
RMSE (mg kg ⁻¹)	35	1,6	2,1	2,0	2,3	2,3	14,3
Faixa linear dinâmica (mg kg ⁻¹)	99 a 9926	0,5 a 854,0	0,6 a 1138,6	0,6 a 1125,7	0,6 a 1151,6	0,7 a 1164,5	0,6 a 1099,8
LD (mg kg ⁻¹)	15	0,00	0,05	0,16	0,01	0,01	0,12
LQ (mg kg ⁻¹)	46	0,01	0,16	0,49	0,03	0,04	0,36
Sensibilidade	74483	108888	103352	106512	105027	105836	104035
Sensibilidade analítica (mg kg ⁻¹)	0,22	784	63	20	328	227	28
γ^{-1} (mg kg ⁻¹) ⁻¹	4,55	0,001	0,016	0,049	0,003	0,004	0,036
Razão Sinal/Ruído (mín)	25	430	59	15	223	166	21
Razão Sinal/Ruído (máx)	2181	666407	71393	22794	376193	263269	29660
Precisão							
DPR menor concentração (%)	0,27	5,94	15,74	17,92	9,06	2,38	1,40
DPR maior concentração (%)	5,15	0,48	0,49	0,49	0,54	0,53	0,46
Repetibilidade (%)	1,94	0,66	0,57	0,72	0,69	0,72	0,79
CV t _R (%)	0,04	0,15	0,12	0,12	0,10	0,07	0,04
Precisão intermediária^a							
DPR baixa concentração	2,42% (1059)	0,66% (6,4)	0,40% (9,1)	0,54% (9,2)	0,97% (9,2)	0,46% (9,6)	0,30% (9,3)
DPR alta concentração	1,65% (1377)	1,00% (80,0)	0,68% (106,2)	0,61% (109,7)	0,65% (110,7)	0,63% (113,0)	0,63% (109,9)
Intervalo de confiança							
C inferior (mg kg ⁻¹)	68 ± 10	0,14 ± 0,02	0,44 ± 0,10	0,60 ± 0,10	0,38 ± 0,04	0,49 ± 0,01	3,00 ± 0,01
C intermediária (mg kg ⁻¹)	875 ± 8	71,29 ± 0,33	96,94 ± 0,40	96,47 ± 0,49	98,70 ± 0,48	99,75 ± 0,52	95,00 ± 0,54
C superior (mg kg ⁻¹)	9897 ± 33	849,10 ± 2,89	1134,15 ± 3,97	1121,49 ± 3,96	1146,80 ± 4,40	1158,98 ± 4,36	1067,85 ± 3,56

^a Os valores entre parênteses referem-se à concentração (mg kg⁻¹) das respectivas amostras analisadas.

As curvas analíticas desenvolvidas com ambos solventes apresentaram coeficientes de determinação e correlação maiores ou iguais a 0,99. Os métodos ultrarrápidos exibiram menor sensibilidade que os convencionais, pois o comprimento reduzido da coluna PH RTX-1 não permitiu o modo de injeção splitless devido à saturação do solvente e, por conseguinte, sobreposição dos picos dos analitos pela larga banda do solvente. A menor sensibilidade acarretou menor razão S/R e menor capacidade de diferenciar pequenas concentrações, por

outro lado, os valores de LD e LQ apresentaram-se abaixo do intervalo de medição das curvas analíticas.

As concentrações determinadas por curvas analíticas com solventes distintos mostraram a influência do tipo de solvente na intensidade do sinal instrumental no modo split. Quando a amostra é inserida no injetor split, a solução é vaporizada e apenas parte dela é arrastada para a coluna a fim de evitar a saturação da coluna. O volume da nuvem da solução vaporizada depende das propriedades físico-químicas de cada solvente, como peso molecular, densidade, interações intermoleculares, e das condições cromatográficas, como dimensões do liner, temperatura do injetor, volume de amostra injetado e pressão total no injetor. Desta forma, torna-se complexa a predição do volume efetivo de amostra inserida na coluna, tratando-se do modo split.¹⁸⁵ As diferentes razões entre área e concentração proporcionaram curvas analíticas de diferentes inclinações, assim, o mesmo valor de área levou à quantificação de maior concentração pela curva em hexadecano e menor concentração em etanol.

Os valores de DPR da curva preparada com hexadecano revelaram a baixa precisão do método gerada pela baixa repetibilidade da injeção. Isto pode ser justificado pela viscosidade dinâmica do hexadecano ser 2,7 vezes maior que do etanol ($\mu_{\text{hexadecano}} = 2,66 \text{ mPa.s}$; $\mu_{\text{etanol}} = 0,98 \text{ mPa.s}$).^{186,187} A repetibilidade pode ser melhorada pelo ajuste da velocidade em que a seringa coleta a solução do vial e a injeta no cromatógrafo, entretanto, o autoamostrador AS3000 é limitado quanto à configuração das velocidades de amostragem e mais ajustes não puderam ser realizados.

A precisão intermediária obtida pelo solvente hexadecano foi influenciada pela baixa repetibilidade, levando a valores muito superiores aos valores obtidos com etanol. A variação de 2% da concentração de GRO em etanol mostrou a possibilidade do armazenamento de amostras a $-19 \text{ }^\circ\text{C}$, durante 15 dias, sem perdas significativas. Os compostos BTEX apresentaram precisão intermediária em etanol inferiores ou iguais a 1%, portanto, períodos de armazenamentos mais longos ainda podem ser avaliados.

Os intervalos de confiança obtidos pela análise de 10 replicatas para o primeiro e último ponto e 20 replicatas para o ponto intermediário apresentaram valores relativamente altos para o hexadecano e baixos para o etanol, justificados pela mesma razão da repetibilidade. A concentração próxima à concentração alvo para GRO apresentou valores de intervalo de confiança $\pm 8 \text{ mg kg}^{-1}$ com o solvente etanol.

Assim como para os compostos GRO e BTEX, foram desenvolvidas curvas com diferentes solventes para quantificação das frações DRO e ORO. Uma vez ensejada a substituição do diclorometano por um solvente não-halogenado, apenas os principais parâmetros das curvas analíticas preparadas com diclorometano foram avaliados e estão dispostos na Tabela 26.

Tabela 26 - Parâmetros das curvas analíticas desenvolvidas em diclorometano e hexano para quantificação das frações DRO e ORO.

Faixa de compostos	DRO		ORO	
	Diclorometano	Hexano	Diclorometano	Hexano
Parâmetro				
Intervalo de medição (mg kg ⁻¹)	149 a 10412	115 a 9053	172 a 10666	119 a 11579
R ²	0,9999	1,0000	0,9909	0,9966
RMSE (mg kg ⁻¹)	42	23	355	242
LD (mg kg ⁻¹)	27	23	3	31
LQ (mg kg ⁻¹)	83	70	9	94
Sensibilidade	59573	24592	40751	24053
γ^{-1} (mg kg ⁻¹) ⁻¹	8,31	7,01	0,93	9,38
Precisão				
DPR menor concentração (%)	5,97	1,63	0,24	4,82
DPR maior concentração (%)	0,91	1,80	1,59	1,57

As curvas geradas apresentaram-se lineares ($R^2 \geq 0,99$), com faixa linear dinâmica de duas ordens de grandeza. Os limites de detecção e quantificação mostraram que a sensibilidade é suficiente para a aplicação do método e os valores de DPR apresentaram-se abaixo de 5% para o ponto inferior da curva, exceto para DRO em diclorometano. A maior sensibilidade das curvas preparadas em diclorometano foram consequência do maior volume de expansão deste solvente em relação ao hexano.¹⁸⁵

Realizou-se um teste F para verificar se a redução dos valores de RMSE utilizando hexano como solvente foi significativa. Se a razão $RMSE_1^2 / RMSE_2^2$, onde $RMSE_1$ é o maior valor, for maior que o valor de F tabelado para 10 graus de liberdade no nível de confiança de 95% ($F_{\text{tabelado}} = 2,98$), os erros das curvas em hexano foram significativamente menores que os erros da curva em diclorometano.¹⁸⁸ O valor de F calculado para DRO foi igual a 3,33 e para ORO igual a 2,15, portanto, os erros foram significativamente menores utilizando hexano como solvente apenas para a fração DRO.

Visto que os parâmetros analíticos das curvas preparadas em hexano são similares aos parâmetros das curvas em diclorometano, apesar da menor sensibilidade, preparou-se novas curvas para as frações DRO e ORO com surrogate e essas foram validadas.

A curva para ORO foi preparada com a amostra O4, a qual apresentou maior teor da fração ORO, e então a faixa de medição foi dividida em dois intervalos a fim de aumentar a linearidade e eliminar a tendência negativa dos pontos mais diluídos, como pode ser notado na Figura 11.

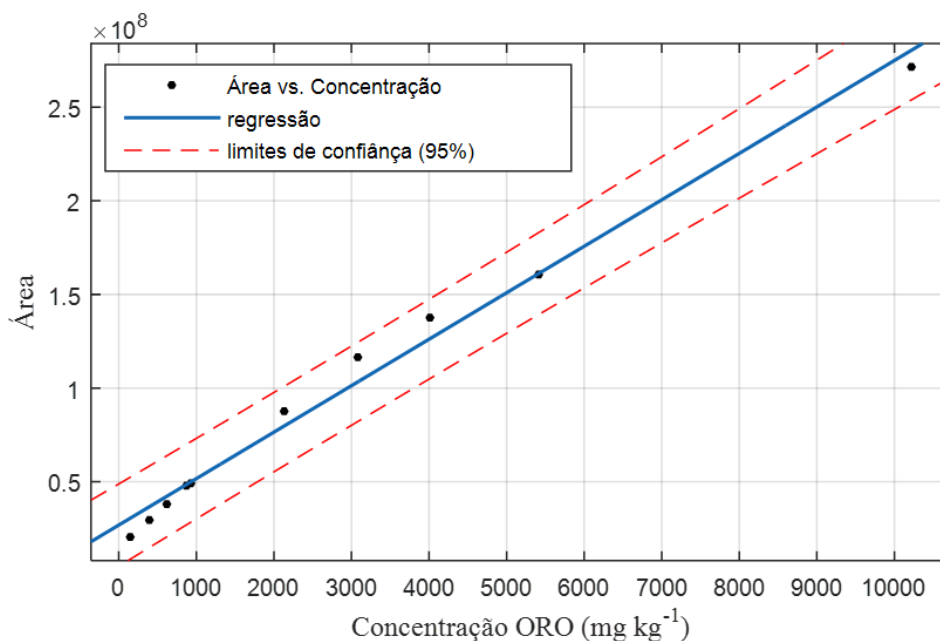


Figura 11 – Curva analítica da fração ORO em hexano abrangendo todo intervalo de mediação.

Os parâmetros analíticos avaliados das curvas para as frações DRO e ORO estão dispostos na Tabela 27. As curvas analíticas para ambas as frações de TPH e para o surrogate (o-terfenil) apresentaram-se lineares, com coeficientes de determinação e correlação acima de 0,990. A curva para a fração ORO foi dividida em duas faixas lineares dinâmicas (ORO₁ e ORO₂): a primeira com 6 pontos (147 - 2132 mg kg⁻¹) e a segunda com 5 pontos (2132 - 10218 mg kg⁻¹). As curvas analíticas mostraram-se sensíveis e com limites de detecção e quantificação abaixo do valor limite de prevenção (1000 mg kg⁻¹), contudo, as curvas para ORO apresentaram sensibilidade inferior à curva para DRO, podendo ser resultado da combustão incompleta de hidrocarbonetos de cadeias longas e ramificadas, caracterizadas na seção 5.1.2.4. O inverso da sensibilidade analítica (γ^{-1}) revelou que são distinguíveis concentrações de 13 mg kg⁻¹ e 50 mg kg⁻¹ para as frações DRO e ORO, respectivamente.

Tabela 27 – Parâmetros analíticos das curvas para determinação das faixas DRO e ORO em hexano.

Parâmetro	Fração DRO		Fração ORO		
	DRO	o-terfenil	ORO ₁	ORO ₂	o-terfenil
Linearidade					
Intervalo de medição (mg kg ⁻¹)	151 a 9625	41,9 a 2670,2	147 a 2132	2132 a 10218	56,1 a 3906,9
R ²	0,9998	0,9996	0,9975	0,9968	0,9967
R ² ajustado	0,9998	0,9995	0,9968	0,9958	0,9963
r	0,9999	0,9998	0,9987	0,9984	0,9984
Inclinação	46838	58535	33656	22204	53654
Intercepto	7759659	457432	16856184	44370762	3448356
RMSE (mg kg ⁻¹)	45	18,8	39	207	72,8
Faixa linear dinâmica (mg kg ⁻¹)	151 a 9625	41,9 a 2670,17	147 a 2132	2132 a 10218	56,1 a 3906,9
LD (mg kg ⁻¹)	42	4,2	109	-	2,8
LQ (mg kg ⁻¹)	126	12,7	331	-	8,6
Sensibilidade	46838	58535	33656	22204	53654
Sensibilidade analítica (mg kg ⁻¹)	0,08	0,78	0,03	0,02	1,17
γ^{-1} (mg kg ⁻¹) ⁻¹	13	1,3	33	50	0,9
Razão Sinal/Ruído (mín)	26	35	18	-	81
Razão Sinal/Ruído (máx)	771	2076	-	244	4505
Precisão					
DPR menor concentração (%)	27,11	2,89	12,07	-	1,24
DPR maior concentração (%)	0,22	0,27	-	1,05	0,43
Repetibilidade (%)	4,66	0,46	7,14	-	2,36
CV t _R (%)	0,07	0,07	0,12	-	0,12
Precisão intermediária^a					
DPR	9,35% (977)	- -	19,95% (707)	- -	- -
Intervalo de confiança					
C inferior (mg kg ⁻¹)	163 ± 164	36,3 ± 2,3	107 ± 165	-	4,9 ± 1,9
C intermediária (mg kg ⁻¹)	890 ± 83	241,4 ± 2,8	958 ± 109	-	334,5 ± 5,9
C superior (mg kg ⁻¹)	9559 ± 39	2638,7 ± 13,0	-	10228 ± 288	3799,6 ± 37,4

^a Os valores entre parênteses se referem à concentração (mg kg⁻¹) das respectivas amostras analisadas.

Quanto à precisão, os pontos de menor concentração apresentaram altos valores de DPR: 27% para a fração DRO e 12% para ORO. Não obstante, os valores de repetibilidade calculados através da análise de 20 replicatas em concentração próxima ao valor de interesse (1000 mg kg⁻¹) foram iguais a 5% e 7% para as frações DRO e ORO, respectivamente. A retenção dos analitos se apresentou precisa, dado que os coeficientes de variação (CV t_R) foram inferiores à 0,2%. A precisão intermediária, calculada com 7 replicatas analisadas em um intervalo de 15 dias, apresentou valores acima dos valores determinados para a fração GRO em

etanol. A baixa precisão intermediária pode ser resultado da maior volatilidade do hexano em relação ao etanol, levando a maiores valores de DPR.

Como resultado da baixa precisão nos pontos de menor concentração, os intervalos de confiança para estes pontos foram elevados, mas para os valores próximos à concentração de interesse, o intervalo de confiança representou cerca de 10% do valor de concentração.

A exatidão dos métodos UFGC para determinação das frações TPH e compostos BTEX foi avaliada juntamente ao processo de extração, portanto, será discutida ao final da seção seguinte, onde serão decorridos os resultados dos métodos de extração.

5.2.3. Avaliação do método de extração por ultrassom

As respostas dos planejamentos experimentais obtidas pelas extrações e análises cromatográficas subsequentes dos grupos GRO e DRO estão presentes, respectivamente, nas Tabelas 28 e 29. A extração dos compostos BTEX foi avaliada pelos valores de recuperação do isobutilbenzeno, uma vez que as correlações entre os valores de recuperação dos analitos e do surrogate foram de pelo menos 0,97.

Tabela 28 – Respostas do planejamento fatorial completo para análise de BTEX e GRO.

Amostra	Fatores [†]				Recuperação (%)							
	U	V	A	S [‡]	B	T	E	m-X	o-X	isobutilb.	GRO	surrogate
SG1	0	0	0	-1	60	65	65	67	65	65	315	101
SG2	-1	1	0	-1	40	44	44	45	44	44	236	99
SG3	1	-1	0	-1	39	44	45	47	46	50	192	77
SG4	0	-1	-1	-1	37	41	41	42	41	42	276	149
SG5	-1	0	-1	-1	25	27	26	27	27	26	324	140
SG6	1	0	1	-1	33	37	37	39	38	38	181	79
SG7	-1	-1	1	-1	42	46	46	48	47	47	249	126
SG8	0	1	1	-1	52	57	58	60	59	58	253	105
SG9	0	0	0	-1	27	31	31	33	32	32	183	101
SG10	1	1	-1	-1	39	45	45	47	46	46	196	102
SG11	0	0	0	-1	39	44	45	47	46	46	172	100
SG12	-1	1	-1	1	2	2	2	2	2	8	14	60
SG13	-1	0	1	1	53	59	60	59	57	73	14	27
SG14	0	0	0	1	55	63	64	62	60	80	5	16
SG15	1	-1	-1	1	23	26	27	26	25	40	7	21
SG16	1	1	-1	1	84	93	96	92	88	112	0	23
SG17	1	-1	1	1	13	14	14	13	12	23	6	20
SG18	-1	-1	0	1	75	86	89	87	85	103	10	20
SG19	0	0	0	1	11	11	11	11	9	24	7	18
SG20	1	1	1	1	32	34	35	34	32	52	2	23
SG21	0	-1	1	1	12	12	13	12	11	22	8	19
SG22	0	0	0	1	117	136	141	138	134	157	1	18

[†] Tempo de ultrassom (U); volume de solvente (V); agitação por vórtex (A); e solvente (S). [‡] Etanol (-1); hexadecano (1).

Tabela 29 – Respostas do planejamento fatorial completo para análise de DRO.

Amostra	FATORES [†]				Recuperação (%)	
	U	V	A	S [‡]	DRO	Surrogate
SD1	1	1	1	1	136	147
SD2	0	0	1	-1	87	60
SD3	1	-1	-1	1	65	47
SD4	-1	-1	1	-1	69	50
SD5	1	1	-1	1	38	49
SD6	-1	1	-1	1	16	22
SD7	-1	1	1	1	91	105
SD8	1	1	-1	-1	73	51
SD9	0	0	-1	-1	1	21
SD10	0	0	1	1	68	79
SD11	-1	1	1	-1	85	65
SD12	-1	1	-1	-1	7	27
SD13	0	0	1	1	92	93
SD14	1	-1	-1	-1	24	20
SD15	1	1	1	-1	52	49
SD16	1	-1	1	-1	98	68
SD17	0	0	-1	1	9	23
SD18	1	-1	1	1	66	90
SD19	-1	-1	1	1	83	61
SD20	0	0	-1	-1	6	20
SD21	0	0	1	-1	55	54
SD22	0	0	-1	1	19	36
SD23	-1	-1	-1	-1	82	54
SD24	-1	-1	-1	1	4	12

[†]Tempo de ultrassom (U); volume de solvente (Vs); agitação por vórtex (A); e solvente (S). [‡] Diclorometano (-1); hexano (1).

Para cada planejamento, observou-se a presença de uma amostra anômala (SG22 para BTEX; SG1 para GRO; e SD8 para DRO) e estas foram excluídas da modelagem. Os valores de p determinados pela Análise de Variância (ANOVA) e os ajustes de cada modelo estão presentes na Tabela 30. Valores de p abaixo do valor crítico (0,05) mostraram quais fatores ou interações foram significativas e se o modelo gerado foi suficiente para explicar as respostas. Desta forma, a extração dos compostos BTEX não pôde ser explicada pela ANOVA visto que o modelo linear apresentou valores de p acima de 0,05 e R^2 igual a 0,6039. Já a extração de compostos GRO apresentou modelo linear com valor de p abaixo de 0,05, R^2 igual a 0,9634 e fatores significativos: tempo de ultrassom; e solvente. Para a extração dos compostos DRO, foram significativos o fator volume de solvente e as interações entre solvente e tempo de ultrassom e entre solvente e agitação.

Tabela 30 – Valores de p obtidos na ANOVA de cada planejamento experimental.

Modelo/Fator(es) †	valor p		
	BTEX	GRO	DRO
Modelo	0,64	0,00	0,00
Linear	0,90	0,00	0,00
U	0,88	0,02	0,85
V	0,84	0,73	0,32
A	0,91	0,28	0,00
S	0,37	0,00	0,19
Interações em 2 níveis	0,34	0,60	0,12
U*V	0,11	0,84	0,73
U*A	0,29	0,54	0,88
U*S	0,86	0,17	0,04
V*A	0,40	0,45	0,06
V*S	0,97	0,86	0,76
A*S	0,66	0,45	0,03
Falta de ajuste	0,46	0,01	0,20
Ajuste			
R ²	0,6039	0,9634	0,8498
Desvio padrão	27,78	12,74	20,77

† Tempo de ultrassom (U); volume de solvente (V); agitação por vórtex (A); e solvente (S).

Através do gráfico de efeitos principais (Figura 12), observou-se a influência de cada fator na extração. O tempo de ultrassom proporcionou maiores extrações de BTEX e DRO nos valores extremos, contudo, a fração GRO teve maior extração apenas no tempo mínimo. Diferenças no volume extração praticamente não influenciaram a extração, apesar de menores volumes dificultarem a filtração devido a maior concentração de solo suspenso na solução.

A agitação por vórtex prévia à sonicação exibiu comportamento distinto para cada grupo de compostos. As extrações de GRO e BTEX apresentaram comportamentos opostos, sendo a maior extração de GRO nos tempos extremos (0 s e 60 s) e de BTEX no tempo intermediário (30 s). A diminuição da recuperação de BTEX após 30 s de agitação pode ser decorrente da maior volatilização destes grupo de compostos voláteis. Já para os compostos semivoláteis, como DRO, a agitação proporcionou melhores recuperações de forma significativa.

O tipo solvente foi um fator significativo na extração dos compostos GRO, contudo, a grande diferença entre os valores de recuperação entre os solventes pode ter sido consequência da baixa repetibilidade do método UFGC com hexadecano. Os compostos BTEX apresentaram recuperação superior com hexadecano e a fração DRO com hexano, porém não foram significativas. Os três gráficos da parte inferior da Figura 12 mostram a relação do solvente com os demais fatores para a extração de DRO. As interações U*S e V*S revelaram características opostas na extração com os diferentes solventes, pois a extração por hexano

apresentou melhores resultados com maior tempo de sonicação e maior volume de solvente, enquanto, por diclorometano as melhores recuperações ocorreram nas condições opostas.

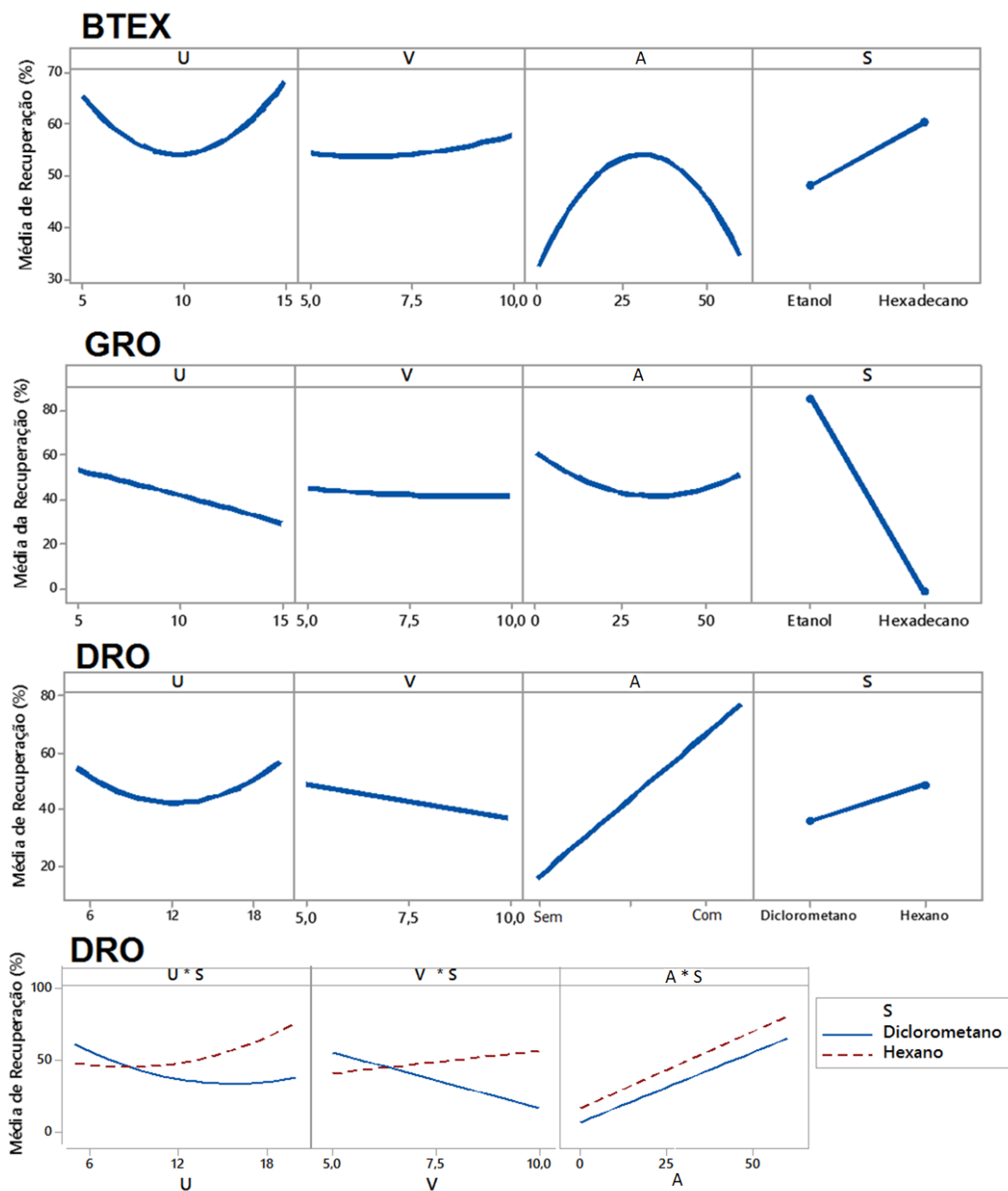


Figura 12 – Gráficos de efeitos principais gerados pela ANOVA para cada DOE.

As equações dos modelos foram utilizadas para predição das condições ótimas para extração de cada grupo de compostos. As condições determinadas para as melhores recuperações estão apresentadas na Tabela 31.

Tabela 31 – Condições determinadas através da ANOVA para melhores recuperações.

	BTEX		GRO		DRO	
	etanol	hexadecano	etanol	hexadecano	hexano	diclorometano
Solvente						
Ultrassom (min)	15	15	10	5	12,5	5
Volume de solvente (mL)	10	10	7,5	5	10	5
Tempo de agitação (s)	31	25	7	0	60	60
Recuperação predita (%)	85,7	95,8	100,0	34,6	100,0	95,3

A repetibilidade da extração nas condições ótimas previstas pelo modelo foi avaliada para os compostos BTEX (Tabela 32). Observou-se que os compostos de menor cadeia (mais voláteis) apresentaram menores valores de recuperação e houve uma alta correlação entre a recuperação dos compostos e a do surrogate (isobutilbenzeno), possibilitando o uso de fatores de correção para a quantificação de BTEX, considerando granulometria e umidade similares para extração de novas amostras.

Tabela 32 – Repetibilidade da extração dos compostos BTEX e fatores de correção.

Solvente extrator: hexadecano						
	Recuperação (%)					
	benzeno	tolueno	etilbenzeno	o-xileno	m-xileno	isobutilbenzeno
Replicata 1	32	39	43	41	41	59
Replicata 2	35	43	47	46	46	72
Replicata 3	43	54	62	61	63	91
Replicata 4	46	58	66	66	67	96
Replicata 5	29	36	40	38	38	61
Replicata 6	27	34	40	38	39	63
Replicata 7	31	40	46	44	45	71
Média	35	43	49	48	48	73
desvio padrão	7	9	11	11	12	15
Correlação com surrogate	0,9342	0,9633	0,9786	0,9813	0,9852	
Solvente extrator: etanol						
	Recuperação (%)					
	benzeno	tolueno	etilbenzeno	m-xileno	o-xileno	isobutilbenzeno
Replicata 1	55	71	77	77	81	85
Replicata 2	47	62	69	69	73	75
Replicata 3	59	76	84	83	88	92
Replicata 4	54	70	76	76	80	87
Replicata 5	57	74	82	82	86	95
Replicata 6	60	77	85	85	90	93
Replicata 7	54	71	79	79	83	88
Média	55	72	79	79	83	88
desvio padrão	4	5	5	5	6	7
Correlação com surrogate	0,9150	0,9354	0,9430	0,9451	0,9411	
FATOR DE CORREÇÃO						
Solvente	benzeno	tolueno	etilbenzeno	m-xileno	o-xileno	
Etanol	1,59	1,22	1,11	1,11	1,06	
Hexadecano	2,11	1,69	1,50	1,54	1,53	

A repetibilidade do método de extração para as faixas DRO e ORO foi avaliada apenas o hexano como solvente extrator, uma vez que se desejou a substituição do diclorometano na extração. Foram preparadas sete replicatas para uma das faixas utilizando as amostras DieA e O4. Os valores de recuperação das replicatas estão presentes na Tabela 33.

Tabela 33 – Repetibilidade da extração das faixas DRO e ORO.

	Recuperação (%)			
	DRO	Surrogate	ORO	Surrogate
Replicata 1	100	55	107	11
Replicata 2	118	52	119	15
Replicata 3	105	86	130	22
Replicata 4	116	48	124	35
Replicata 5	111	156	100	30
Replicata 6	102	97	88	22
Replicata 7	103	60	222	41
Média	108	79	127	25
desvio padrão	7	39	44	11
correlação com o surrogate		-0,0502		0,6433

Diferentemente da faixa GRO e dos compostos BTEX, a recuperação das faixas DRO e ORO não apresentou correlação com a recuperação do surrogate, o que não torna possível a determinação de fatores de correção para determinação destas faixas. Os baixos valores de recuperação do surrogate (o-terfenil) podem ter sido em decorrência da menor interação intermolecular entre o analito e o solvente apolar (hexano) que entre solventes com grupos receptores de elétrons, como diclorometano e acetona.³⁹ Os elevados valores de recuperação de ORO podem ter sido ocasionados extração de matéria orgânica do solo, a qual não foi retida eficientemente na etapa de purificação.

5.2.3.1. Determinação da exatidão dos métodos UFGC

A exatidão dos métodos UFGC foi avaliada considerando-se o processo de extração e a quantificação pelas curvas analíticas desenvolvidas. Para cada grupo de analitos, fortificou-se amostras de solo em três níveis de concentração. As amostras utilizadas na fortificação foram a G2, D2, O4 e uma solução BTEX em etanol. Os métodos de extração foram executados conforme as condições otimizadas e os extratos foram analisados subsequentemente pelos métodos cromatográficos. A exatidão para cada fração e compostos BTEX foi avaliada pelos

valores de recuperação, os quais estão presentes na Tabela 34, juntamente aos respectivos valores de concentração.

Tabela 34 – Avaliação da exatidão dos métodos determinação das frações TPH e BTEX por extração ultrassônica e subsequente análise UFGC.

Solvente	Etanol						Hexano	
	GRO	benzeno	tolueno	etilbenzeno	m-xileno	o-xileno	DRO	ORO
Recuperação I	70% (871)	102% (36,6)	102% (45,2)	103% (50,3)	104% (48,6)	105% (48,0)	97% (625)	131% (1310)
Recuperação II	111% (3853)	97% (41,4)	100% (51,1)	102% (56,9)	102% (54,9)	103% (54,3)	105% (2037)	113% (4181)
Recuperação III	98% (7747)	101% (74,7)	99% (92,2)	98% (102,7)	98% (99,2)	99% (98,0)	104% (9293)	101% (7235)

‡ Os valores entre parênteses se referem à concentração (mg kg^{-1}) das respectivas amostras analisadas.

Os limites de aceitação para recuperação de TPH foram baseados no guia EPA 8000¹⁷³ e na norma ASTM D7678.¹² Desta forma, apenas a recuperação dos compostos ORO na concentração de 1310 mg kg^{-1} permaneceu fora do limite estabelecido e os compostos GRO na concentração de 871 mg kg^{-1} próximo ao limite. Portanto, pode-se dizer que os métodos de extração e determinação são exatos para faixa de trabalho avaliada.

5.2.4. Análise das amostras de monitoramento ambiental para determinação de TPH e BTEX

As amostras de solo coletadas em postos de combustíveis para o monitoramento ambiental (“amostras reais”) foram analisadas de acordo com os métodos de extração e cromatográficos ultrarrápidos desenvolvidos. As amostras de solo foram divididas em cinco parcelas e, de cada parcela, retirou-se uma alíquota para análise. Os valores das frações TPH e BTEX das amostras foram resultantes da média dos valores obtidos de cada alíquota e os valores de BTEX foram multiplicados pelos fatores de correção obtidos na seção anterior. Os resultados das análises cromatográficas das amostras reais estão dispostos na Tabela 35. Para calcular as médias e os desvios padrão, os valores abaixo do limite de detecção (< LD) não foram considerados no cálculo.

Tabela 35 – Determinação de BTEX e das frações TPH das “amostras reais” por UFGC.

Amostra	Concentração (mg kg ⁻¹)								Rec. (%)*
	benzeno	tolueno	etilbenzeno	m,p-xilenos	o-xileno	GRO	DRO	ORO	
AR I 1	<LD	<LD	<LD	0,05	<LD	<LD	1024	1807	73
AR I 2	<LD	<LD	<LD	0,05	<LD	<LD	997	2594	75
AR I 3	<LD	<LD	<LD	0,11	<LD	<LD	7458	2011	72
AR I 4	<LD	<LD	<LD	<LD	<LD	<LD	1224	4488	43
AR I 5	<LD	<LD	<LD	<LD	<LD	<LD	1038	1252	81
AR I (média)	<LD	<LD	<LD	0,07 ± 0,03	<LD	<LD	2348 ± 2858	2430 ± 1246	69
AR II 1	<LD	<LD	<LD	0,13	<LD	56	1737	<LD	73
AR II 2	<LD	0,65	<LD	<LD	<LD	<LD	1894	764	41
AR II 3	<LD	<LD	<LD	<LD	<LD	<LD	1229	<LD	72
AR II 4	<LD	<LD	<LD	<LD	<LD	<LD	1944	463	99
AR II 5	<LD	0,58	<LD	<LD	<LD	<LD	2261	829	89
AR II (média)	<LD	0,62 ± 0,05	<LD	0,13	<LD	56	1813 ± 378	685 ± 195	75
AR III 1	<LD	<LD	0,21	0,82	<LD	93	1635	4593	76
AR III 2	<LD	<LD	<LD	0,53	<LD	33	1992	4606	58
AR III 3	<LD	<LD	<LD	0,33	<LD	70	1625	3777	76
AR III 4	<LD	<LD	<LD	0,26	<LD	52	1945	4691	76
AR III 5	<LD	<LD	<LD	0,29	<LD	80	1921	4477	61
AR III (média)	<LD	<LD	0,21	0,45 ± 0,23	<LD	66 ± 24	1823 ± 179	4429 ± 372	69

*Valores de recuperação para o surrogate, isobutilbenzeno.

Observou-se na Tabela 35 que a maioria das médias apresentou valores elevados de desvio padrão. Dado que a validação do método apresentou boa repetibilidade, este fato foi em decorrência da heterogeneidade das amostras AR I e AR II. A análise cromatográfica mostrou que as amostras reais apresentaram contaminantes das frações DRO e ORO acima do limite especificado pela CETESB (1000 mg kg⁻¹) e compostos da faixa GRO em baixas concentrações.

A contaminação das amostras AR I e AR III abre duas possibilidades para suas causas: (1) contaminação devido ao vazamento de óleo diesel e óleo lubrificante; e (2) contaminação por óleo lubrificante e subsequente degradação dos hidrocarbonetos de cadeia longa. Dois fatos sugerem a segunda hipótese: (i) a caracterização das amostras de óleo lubrificante revelou que uma parcela considerável dos compostos, até 46% (m/m), apresentou tempo de retenção dentro da faixa DRO; e (ii) os hidrocarbonetos se degradam nas condições naturais do solo para hidrocarbonetos de cadeia menor. Portanto, nesta perspectiva, o óleo lubrificante é o contaminante mais provável das amostras AR I e AR III e o óleo diesel da amostra AR II.

Os valores médios determinados nesta seção foram utilizados na validação dos modelos quimiométricos em respeito à predição das frações TPH em amostras “reais”.

5.3. Modelos quimiométricos

5.3.1. Análises prévias

A análise espectroscópica de amostras sólidas por refração difusa sofre grande influência da heterogeneidade da amostra. Para reduzir a quantidade de ruídos espectrais, é comum realizar várias varreduras, ou escaneamentos, da mesma amostra e utilizar a média das varreduras para obtenção do espectro médio. Entretanto, a área da amostra suscetível à radiação NIR pode não ser representativa para a amostra toda. Existem dispositivos que otimizam a representatividade através da rotação da amostra em uma bandeja, porém, apenas a camada inferior da amostra continua a representar a totalidade da amostra. Desta forma, procurou-se tornar o espectro médio mais representativo através de homogeneizações das amostras de forma manual entre as aquisições de replicatas.

Para evitar uma aquisição demasiada de replicatas antes da análise de todo o conjunto amostral, avaliou-se o número mínimo de replicatas necessário para obter modelos com capacidade preditiva equivalente aos modelos desenvolvidos com até seis replicatas. Preparou-se um conjunto de 50 amostras (5 amostras preparadas com cada amostra S) fortificadas com a gasolina G2 na faixa de concentração 132-3920 mg kg⁻¹. As amostras foram analisadas em sextuplicada e, então, desenvolveu-se modelos PLS com as médias de sextuplicadas e triplicatas para comparação dos ajustes. Para reduzir a quantidade de ruído na modelagem, excluiu-se as regiões pouco informativas (10.000-7500 cm⁻¹ e 7000-5500 cm⁻¹) e os pré-processamentos 1^a derivada e centragem na média foram aplicados previamente à regressão.

Os valores de RMSEP e r_{val} (coeficiente de correlação do conjunto de validação) foram obtidos através do conjunto de validação constituído de um terço do conjunto amostral. O modelo desenvolvido com sextuplicatas apresentou RMSEP e r_{val} iguais a, respectivamente, 379 mg kg⁻¹ e 0,9227 e o com triplicatas iguais a 374 mg kg⁻¹ e 0,9252, portanto não houve diferença significativa ($F_{tabelado} = 2,27$; $F_{calculado} = 1,03$) e o número de três replicatas por amostras foi definido para as análises posteriores.

5.3.2. Modelos de classificação

5.3.2.1. Avaliação dos pré-processamentos

Diversos pré-processamentos foram avaliados no desenvolvimento dos modelos de classificação para identificação do tipo de contaminante presente no solo ou no solvente extrator. Primeiramente, os pré-processamentos foram avaliados através do desenvolvimento de modelos PLS-DA com as 213 amostras de calibração e, posteriormente, verificados com o conjunto de validação para cada conjunto amostral. A escolha do pré-processamento mais adequado para a modelagem foi baseada nos valores de variância explicada nos blocos \mathbf{X} e \mathbf{Y} e nos coeficientes de correlação gerados a partir do conjunto de validação. Dado que o método de classificação PLS-DA é fundamentado no método de regressão PLS, modelos PLS-DA mais ajustados proporcionam coeficientes de correlação mais próximos de 1 e tornam os agrupamentos das classes mais distantes entre si estatisticamente. Uma segunda avaliação dos pré-processamentos foi realizada no conjunto de predição constituído de 27 amostras analisadas em triplicata, portanto, 81 espectros NIR.

Os dados presentes na Tabela 36 foram adquiridos da primeira avaliação dos pré-processamentos descrita acima. O número de variáveis latentes foi definido para o modelo com menores erros de classificação para a calibração e validação cruzada (venetian blind, 10 splits). O elevado número de variáveis latentes pode ser justificado pela grande complexidade das amostras, inclusive das soluções dos contaminantes nos solventes extratores, pois, tratando-se de grupos de analitos com grupos funcionais similares, tornou-se necessário o uso de muitas variáveis latentes para explicar a variância relacionada às três faixas de TPH presentes nas amostras. Tipicamente, os filtros multivariados, como OSC, GLSW, EPO e EMM, costumam gerar variâncias explicadas equivalentes ou maiores utilizando menos variáveis latentes que outros pré-processamentos, pois atenuam a influência de variáveis x menos correlacionadas com o bloco \mathbf{Y} . Entretanto, quando mais de duas classes são classificadas, a correlação entre as variáveis x e y é avaliada para várias classes conjuntamente e, conseqüentemente, mais variáveis latentes podem ser necessárias para explicar as variâncias das matrizes de dados. Ainda assim, o filtro multivariado EMM proporcionou predições com coeficientes de correlação similares aos outros pré-processamentos utilizando um número de reduzido de variáveis latentes.

Os espectros das amostras sem extração por solvente apresentaram grande deslocamento de linha base decorrente da dispersão da radiação NIR resultante da heterogeneidade dos solos. Deste modo, os pré-processamentos utilizados, além do autoescalamento e centragem na média,

proporcionaram melhores coeficientes de correlação após a correção da linha base, exceto a 1ª e 2ª derivada. A derivação dos espectros NIR é útil para a correção de linha base, entretanto, ocorre a atenuação do sinal analítico quando aplicada, comprometendo a sensibilidade do método. Portanto, o uso de 2ª derivada não foi adequado para os conjuntos de espectros analisados.

Tabela 36 – Avaliação inicial do uso de diferentes pré-processamentos para os modelos PLS-DA.

Sem extração							
Pré-processamento	#V.L.	Variância Explicada (%)		Coeficiente de correlação (validação)			
		Bloco X	Bloco Y	Branco	Gasolina	Diesel	Óleo lub.
autoescalamento	11	99,99	70,91	0,75	0,97	0,81	0,82
C.M.	11	99,99	70,82	0,75	0,96	0,81	0,82
C.L.B.	12	99,64	85,40	0,89	0,98	0,84	0,87
1ª der.	10	91,99	86,40	0,84	0,95	0,86	0,85
2ª der.	9	71,17	84,65	0,78	0,84	0,80	0,76
MSC	13	99,78	91,52	0,93	0,98	0,90	0,90
SNV	13	99,74	86,76	0,93	0,98	0,86	0,83
OSC	12	100,00	85,77	0,92	0,98	0,87	0,84
GLSW	9	55,57	91,68	0,92	0,98	0,91	0,91
EPO	12	99,15	91,43	0,92	0,98	0,91	0,92
EMM	8	33,00	99,67	0,93	0,99	0,93	0,93
Extração com hexano							
Pré-processamento	#V.L.	Variância Explicada (%)		Coeficiente de correlação (validação)			
		Bloco X	Bloco Y	Branco	Gasolina	Diesel	Óleo lub.
autoescalamento	11	99,91	88,33	0,87	0,95	0,88	0,95
C.M.	11	99,86	90,48	0,87	0,95	0,91	0,96
C.L.B.	11	98,74	91,74	0,86	0,96	0,92	0,97
1ª der.	11	93,37	93,50	0,80	0,96	0,88	0,97
2ª der.	8	75,57	91,34	0,83	0,93	0,86	0,93
MSC	8	99,20	83,21	0,84	0,94	0,88	0,91
SNV	8	99,18	84,25	0,85	0,94	0,89	0,91
OSC	8	100,00	91,50	0,85	0,94	0,88	0,91
GLSW	7	72,58	88,31	0,88	0,94	0,91	0,94
EPO	7	99,13	94,42	0,85	0,96	0,92	0,97
EMM	3	99,60	99,29	0,86	0,98	0,93	0,98
Extração com etanol							
Pré-processamento	#V.L.	Variância Explicada (%)		Coeficiente de correlação (validação)			
		Bloco X	Bloco Y	Branco	Gasolina	Diesel	Óleo lub.
autoescalamento	12	99,94	82,35	0,78	0,92	0,95	0,89
C.M.	10	99,94	78,46	0,68	0,91	0,95	0,88
C.B.	12	99,00	90,21	0,84	0,94	0,97	0,95
1ª der.	12	81,11	91,81	0,77	0,90	0,96	0,93
2ª der.	5	30,62	79,13	0,71	0,88	0,92	0,89
MSC	12	99,67	91,48	0,83	0,95	0,97	0,95
SNV	12	99,66	90,99	0,85	0,95	0,97	0,95
OSC	12	100,00	89,97	0,81	0,94	0,97	0,95
GLSW	10	57,94	91,83	0,84	0,95	0,96	0,95
EPO	12	96,78	91,83	0,84	0,96	0,96	0,95
EMM	7	24,22	99,64	0,81	0,95	0,98	0,97

Para cada conjunto de dados, houve pré-processamentos com resultados de validação similares, como o MSC e o filtro EPO para a análise sem preparo de amostra; a correção de linha base (polinomial) e EMM para a análise após extração com hexano; e o SNV e EMM para a análise em etanol. Desta forma, estes pré-processamentos foram avaliados através dos resultados gerados a partir do conjunto de validação (II), dado que os interferentes presentes nos solos são considerados neste conjunto de validação. Nesta avaliação, os melhores pré-processamentos da primeira avaliação foram reavaliados utilizando não apenas os espectros completos, mas também as variáveis selecionadas pelo algoritmo genético executado uma vez (GA1) e duas vezes consecutivas (GA2). Comparou-se o número de classificações corretas e amostras classificadas erroneamente como não contaminadas (branco), ou seja, resultados falsos positivos (Tabela 37).

Tabela 37 – Avaliação de pré-processamentos na análise do conjunto de validação (II) através dos modelos PLS-DA.

Método	Pré-processamento	Classificações corretas (%)			Classificados como branco (%)		
		Completo	GA1	GA2	Completo	GA1	GA2
Sem extração	MSC	9,9	16,0	22,2	56,8	49,4	40,7
	EPO	23,5	22,2	42,0	33,3	42,0	21,0
Ext. c/ hexano	BC	34,6	9,9	49,4	65,4	90,1	39,5
	EMM	27,2	32,1	55,6	64,2	67,9	0,0
Ext. c/ etanol	SNV	33,3	33,3	33,3	0,0	0,0	0,0
	EMM	61,7	34,6	60,5	0,0	2,5	0,0

Embora os pré-processamentos tenham gerados resultados similares na primeira avaliação, a análise de amostras fortificadas, com ou sem extração, revelou melhoras significativas através dos filtros multivariados. Para o conjunto de dados sem extração, o filtro EPO proporcionou, aproximadamente, o dobro de classificações corretas e a metade de falsos positivos. O filtro EMM forneceu melhores predições para os conjuntos de dados com extração e geraram resultados sem falsos positivos. Desta maneira, os respectivos filtros multivariados foram definidos para cada conjunto de dados.

5.3.2.2. Performance dos modelos PLS-DA

Havendo definido os pré-processamentos, avaliou-se a performance dos modelos PLS-DA desenvolvidos com os espectros NIR completos e com as variáveis selecionadas pelo algoritmo genético executado uma (GA1) e duas vezes (GA2).

Na Tabela 38, estão apresentados o pré-processamento, o número de variáveis x e as variáveis latentes utilizados no desenvolvimento de cada modelo. Os resultados de sensibilidade, seletividade e coeficiente de correlação foram obtidos a partir da predição do conjunto de validação (I) de cada conjunto de dados. A sensibilidade é um índice, de zero a um, de classificações corretas para determinada classe enquanto a seletividade é um índice de amostras classificadas erroneamente em outras classes. Desta forma, os resultados de validação mostram que a seleção de variáveis reduziu a capacidade preditiva dos modelos, dado que houve decréscimo dos valores de sensibilidade, seletividade e coeficientes de correlação. A redução da variância explicada no bloco Y , apesar do aumento no bloco X , pode ter levado à diminuição da capacidade preditiva dos modelos.

Tabela 38 – Resultados da predição dos conjuntos de validação (I) pelos modelos PLS-DA com e sem seleção de variáveis.

	Sem extração			Extração com hexano			Extração com etanol		
	Completo	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
Pré-processamento	EPO (2PC)			EMM			EMM		
#variáveis x	3112	1033	351	3112	877	232	3112	771	247
#V.L.	12	12	12	3	3	3	7	6	5
Var. exp. X (%)	99,15	99,26	99,40	35,35	40,72	53,78	24,22	25,90	33,76
Var. exp. Y (%)	91,43	90,20	87,44	96,99	94,38	87,77	99,64	97,65	92,24

Classe	Sem extração								
	Sensibilidade			Seletividade			Coeficiente de correlação		
	Completo	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
Branco	1,00	1,00	1,00	1,00	1,00	1,00	0,92	0,90	0,87
Gasolina	1,00	1,00	1,00	1,00	1,00	1,00	0,98	0,98	0,98
Diesel	1,00	0,97	0,94	0,99	0,97	0,97	0,91	0,89	0,89
Óleo lub.	1,00	1,00	1,00	0,99	1,00	1,00	0,92	0,93	0,93

Classe	Extração com hexano								
	Sensibilidade			Seletividade			Coeficiente de correlação		
	Completo	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
Branco	0,71	0,86	0,57	1,00	1,00	1,00	0,88	0,84	0,78
Gasolina	1,00	1,00	1,00	1,00	1,00	1,00	0,98	0,97	0,89
Diesel	0,97	0,94	1,00	0,99	0,99	0,95	0,94	0,92	0,88
Óleo lub.	1,00	1,00	1,00	1,00	1,00	1,00	0,98	0,98	0,95

Classe	Extração com etanol								
	Sensibilidade			Seletividade			Coeficiente de correlação		
	Completo	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
Branco	0,86	0,43	0,57	1,00	1,00	1,00	0,81	0,79	0,67
Gasolina	1,00	0,97	0,97	1,00	0,99	0,92	0,95	0,91	0,83
Diesel	1,00	1,00	1,00	1,00	1,00	1,00	0,98	0,97	0,92
Óleo lub.	1,00	1,00	1,00	1,00	1,00	0,97	0,97	0,96	0,92

Embora os resultados de predição dos conjuntos de validação (I) mostrarem uma redução da eficiência dos modelos após o uso de GA, os resultados da classificação das amostras de validação (II), as quais foram submetidas aos processos de extração, presentes na

Tabela 37, revelam que a seleção de variáveis proporcionou maior número de classificações corretas, além de reduzir o número de amostras classificadas erroneamente como não contaminadas. Portanto, o uso de GA tornou o modelo PLS-DA mais eficaz para predição de amostras externas, as quais apresentam mais interferentes.

5.3.2.3. *Classificação das amostras reais pelos modelos PLS-DA*

De acordo com as análises cromatográficas das amostras coletadas do monitoramento ambiental de postos de combustíveis, “amostras reais”, as amostras estavam contaminadas principalmente com hidrocarbonetos das faixas DRO e ORO, nas quais a faixa DRO está como majoritário na amostra ARII. A presença das duas faixas nas amostras de solo pode ser decorrente da contaminação de óleos lubrificantes, pois estes costumam apresentar hidrocarbonetos na faixa DRO e, além disso, hidrocarbonetos de cadeias maiores degradam para cadeias menores nas condições ambientais das matrizes. Portanto, esperou-se que os modelos de predição classificassem as amostras como contaminadas por óleo lubrificante ou diesel. Para as análises com extração, foram realizadas extrações das amostras em triplicata e cada solução extraída foi analisada espectroscopicamente em triplicata, resultando em 9 espectros por amostra. Para a análise sem extração, cinco porções de cada amostra foram analisadas em triplicata, resultando em 15 espectros por amostra.

Os resultados da classificação das amostras reais foram dispostos como porcentagem de espectros preditos por classe e estão presentes na Tabela 39. Utilizando a faixa espectral completa, o método sem extração mostrou ser mais adequado, dado que não houve classificação de amostras contaminadas como “branco”, entretanto, a amostra ARII foi classificada estritamente como solo contaminado por gasolina. A presença da faixa ORO foi identificada principalmente através do método com extração por etanol, enquanto os demais métodos classificaram nenhuma ou uma pequena parcela dos espectros como amostras contaminadas por óleo lubrificante.

Visto que as “amostras reais” apresentam valores de TPH acima do limite de intervenção, apenas o método sem extração desenvolvido com a faixa espectral completa não apresentou falsos negativos para contaminação. Portanto, este método foi o mais adequado para a análise das “amostras reais”.

Tabela 39 – Resultados em porcentagem da predição das amostras reais pelos modelos PLS-DA.

Amostra	Classe	Sem extração			Extração com hexano			Extração com etanol		
		Completo	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
ARI	Branco	0,0	100,0	60,0	11,1	88,9	44,4	100,0	100,0	100,0
	Gasolina	0,0	0,0	0,0	77,8	0,0	11,1	0,0	0,0	0,0
	Diesel	100,0	0,0	40,0	0,0	0,0	44,4	0,0	0,0	0,0
	óleo lub.	0,0	0,0	0,0	11,1	11,1	0,0	0,0	0,0	0,0
ARII	Branco	0,0	6,7	20,0	0,0	22,2	33,3	0,0	0,0	0,0
	Gasolina	100,0	93,3	0,0	100,0	66,7	33,3	11,1	0,0	0,0
	Diesel	0,0	0,0	80,0	0,0	0,0	33,3	0,0	11,1	11,1
	óleo lub.	0,0	0,0	0,0	0,0	11,1	0,0	88,9	88,9	88,9
ARIII	Branco	0,0	40,0	0,0	0,0	0,0	11,1	0,0	22,2	0,0
	Gasolina	0,0	0,0	0,0	100,0	88,9	33,3	11,1	0,0	0,0
	Diesel	100,0	60,0	100,0	0,0	0,0	55,6	0,0	0,0	0,0
	óleo lub.	0,0	0,0	0,0	0,0	11,1	0,0	88,9	77,8	100,0

5.3.2.4. Classificação pelos métodos KNN e SVM

Além do método PLS-DA, os métodos KNN e SVM foram avaliados para a identificação das faixas de TPH em solo. Os modelos KNN para os métodos com extração foram desenvolvidos com 4 vizinhos e o método sem extração com 5 vizinhos. Todos os modelos KNN e SVM resultaram em sensibilidades e seletividades iguais a 1,00 para os conjuntos de validação (I), ou seja, apenas resultados verdadeiros positivos e sem falsos positivos.

Para o conjunto de validação (I), todos os métodos foram equivalentes na identificação da classe de contaminante, contudo, a análise do conjunto de validação (II) gerou uma baixa porcentagem de classificações corretas e alta de falsos negativos para contaminação, exceto a predição das amostras sem extração pelo método SVM (Tabela 40).

Tabela 40 – Resultados da classificação do conjunto de validação (II) pelos modelos

Classificações	KNN			SVM		
	Sem extração	Extração etanol	Extração com etanol	Sem extração	Extração etanol	Extração com etanol
Corretas (%)	40,7	2,5	33,3	67,9	0,0	37,0
Branco (%)	14,8	66,7	0,0	0,0	55,6	0,0

A validação do método sem extração apresentou melhores resultados através do modelo SVM, contudo, o método PLS-DA proporcionou melhores resultados para os métodos com

extração (Tabela 38). Já o método KNN não apresentou resultados de classificação melhores para nenhum dos métodos desenvolvidos.

Os modelos KNN e SVM foram aplicados ao conjunto de espectros das “amostras reais” e os resultados das predições estão dispostos em porcentagem na Tabela 41. Embora, o modelo SVM tenha apresentado melhores resultados para o conjunto de validação sem extração, as amostras reais foram classificadas como contaminadas por gasolina. Já os métodos com extração resultaram em 100% de classificações para a classe de óleo lubrificante pelo método SVM. A predição pelos modelos KNN resultaram em classificações condizentes com os resultados cromatográficos para o método com extração com etanol.

Tabela 41 – Resultados em porcentagem da predição das amostras reais pelos modelos KNN e SVM.

Amostra	Classe	KNN			SVM		
		Sem extração	Extração c/ hexano	Extração c/ etanol	Sem extração	Extração c/ hexano	Extração c/ etanol
ARI	branco	0,0	55,6	0,0	0,0	0,0	0,0
	gasolina	0,0	44,4	0,0	100,0	0,0	0,0
	diesel	0,0	0,0	33,3	0,0	0,0	0,0
	óleo lub.	100,0	0,0	66,7	0,0	100,0	100,0
ARII	branco	13,3	33,3	0,0	0,0	0,0	0,0
	gasolina	0,0	0,0	0,0	100,0	0,0	0,0
	diesel	0,0	0,0	100,0	0,0	0,0	0,0
	óleo lub.	86,7	66,7	0,0	0,0	100,0	100,0
ARIII	branco	0,0	66,7	0,0	0,0	0,0	0,0
	gasolina	100,0	0,0	0,0	100,0	0,0	0,0
	diesel	0,0	0,0	100,0	0,0	0,0	0,0
	óleo lub.	0,0	33,3	0,0	0,0	100,0	100,0

5.3.2.5. Síntese dos resultados dos modelos de classificação

A avaliação dos pré-processamentos mostrou que, embora diferentes pré-processamentos possam gerar ajustes similares para a modelagem, a predição de amostras externas apresentou melhores resultados com modelos PLS-DA pré-processados com os filtros multivariados EPO ou EMM (Tabela 37).

A seleção de variáveis pelo algoritmo genético reduziu o ajuste dos modelos PLS-DA para as amostras de validação (I) (Tabela 38), porém aumentou as classificações corretas das amostras de validação (II) (Tabela 37). Para a análise das amostras “reais”, o uso de GA proporcionou melhores resultados apenas para o método com extração com etanol (Tabela 39).

A análise das “amostras reais” mostrou que o método sem extração é o mais apropriado para a identificação da fonte de contaminação pelo método PLS-DA (Tabela 39).

A avaliação de outros métodos de classificação das amostras reais apontou que o método KNN foi adequado para o método com extração com etanol, enquanto, o método SVM para ambos os métodos com extração por solvente (Tabela 41).

5.3.3. Modelos de calibração

5.3.3.1. Avaliação dos pré-processamentos

Primeiramente, avaliou-se o pré-processamento mais adequado para cada conjunto de dados. Para o conjunto amostral de solos fortificados, excluíram-se as regiões espectrais pouco informativas e ruidosas ($10.000-8000\text{ cm}^{-1}$ e $7000-6000\text{ cm}^{-1}$) previamente à avaliação dos pré-processamentos (Figura 13).

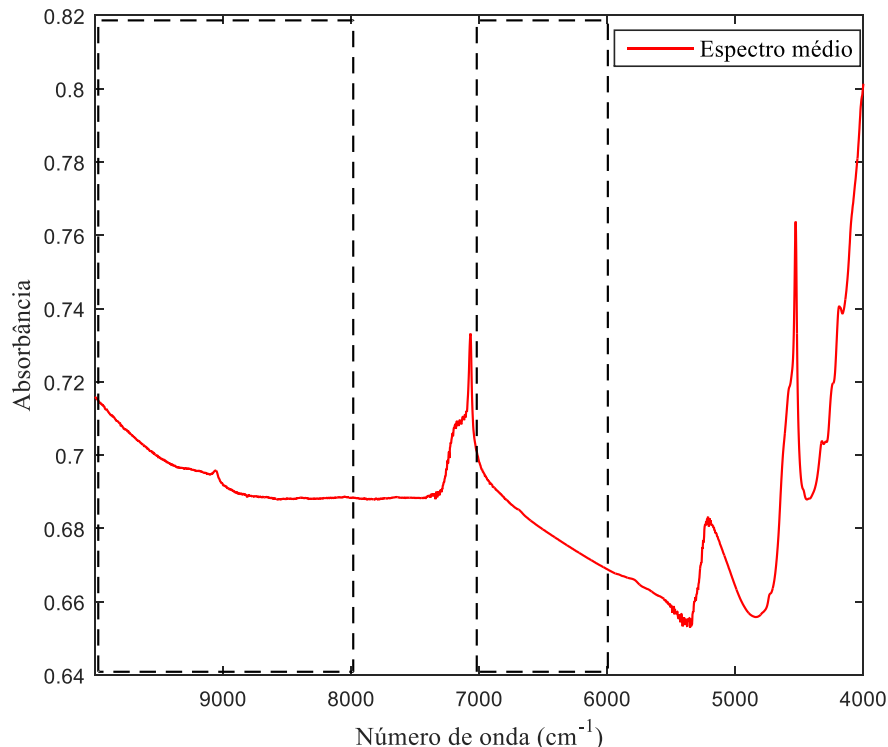


Figura 13 – Espectro médio do conjunto de amostras de solo fortificado. As regiões inseridas nos retângulos foram excluídas previamente à análise dos pré-processamentos.

Os modelos de predição para ORO em etanol foram desenvolvidos apenas com o intervalo de medição entre 0 mg kg⁻¹ e 3814 mg kg⁻¹ devido à baixa solubilidade dos óleos lubrificantes neste solvente. Como evidência da baixa solubilidade, o modelo desenvolvido com o intervalo de medição completo (0 - 7953 mg kg⁻¹) apresentou R²_{cal} igual a 0,48 e o modelo com menor intervalo de medição, R²_{cal} igual a 0,91.

Na Tabela 42, encontram-se o número de variáveis latentes (VL) utilizado na modelagem, a porcentagem de variância capturada para os blocos **X** e **y**, os valores de RMSEC, RMSEP e os coeficientes de determinação do conjunto de calibração (R²_{cal}) e validação (R²_{val}). Os parâmetros da validação foram obtidos a partir da predição do conjunto de validação (I).

Usualmente, pré-processamentos de autoescalamiento não são aplicados a dados espectrais devido à grande quantidade de ruídos presentes na matriz **X**, entretanto, os espectros NIR são constituídos por absorções de diferentes magnitudes em decorrência de sobretons – níveis mais energéticos apresentam menor intensidade – e anarmonicidades diferentes para cada grupo de átomos. Na Tabela 42, observa-se que os pré-processamentos autoescalamiento e centragem na média (C.M.) proporcionaram resultados similares para a maioria dos analitos. O autoescalamiento gerou melhores resultados para os modelos ORO em hexano, GRO em etanol e DRO em etanol; e resultados inferiores para os modelos DRO em solo fortificado, DRO em hexano e ORO em etanol. Os melhores resultados são justificados pelo maior peso dado às variáveis de menor intensidade que estavam relacionadas ao vetor **y**, enquanto os resultados inferiores foram em decorrência do aumento do ruído, diminuindo da razão S/N.

Quanto aos pré-processamentos utilizados para correção de linha base, o pré-processamento correção de linha base por peso de mínimos quadrados (CLB) gerou melhores resultados de validação que os modelos gerados com 1^a e 2^a derivadas. Estes dois últimos pré-processamentos são eficazes na correção de linha base, porém atenuam o sinal instrumental. Conseqüentemente, a variância explicada no bloco **X** é reduzida, tornando o modelo menos adequado para predição de amostras externas ao conjunto de calibração. O uso da 2^a derivada resultou em maiores valores de RMSEP para todas as frações de TPH.

Os pré-processamentos MSC e SNV, utilizados para normalização e corrigir variações provenientes do espalhamento do feixe de radiação NIR, também levaram a um alinhamento das linhas bases. Ambos os pré-processamentos apresentaram resultados de validação próximos, pois promovem correções espectrais similares. Uma vez que os espectros das amostras de solo são resultados da reflexão difusa, os pré-processamentos MSC e SNV foram os mais adequados para a análise NIR sem preparo de amostra. A análise NIR dos conjuntos amostrais de soluções em hexano ou etanol sofreram menos interferências relacionadas ao

espalhamento da radiação, pois os espectros foram adquiridos através do modo de transletância. Desta forma, os pré-processamentos MSC e SNV, apesar de terem apresentado resultados promissores de validação dos conjuntos amostrais de soluções, não foram tão significativos como na análise direta das amostras de solo.

Variações espectrais não relacionadas ao vetor y podem ser atenuadas ou eliminadas através dos pré-processamentos baseados no princípio de ortogonalização do bloco X . Os pré-processamentos testados, OSC, GLSW, EPO e EMM, aumentaram a variância explicada em X para a maioria dos analitos, porém alguns modelos apresentaram diminuição da variância explicada em y . A comparação dos valores de RMSEC e RMSEP gerados pelos modelos pré-processados com estes algoritmos revela um sobreajuste dos modelos de calibração, uma vez que houve uma grande redução dos valores de RMSEC e um aumento dos valores de RMSEP. A redução da habilidade preditiva dos modelos com solos fortificados através destes pré-processamentos pode ser justificada pela grande quantidade de ruído espectral e baixa precisão relacionada à reflexão difusa, pois impossibilitam a identificação de variâncias relacionadas e não relacionadas ao vetor y . Por outro lado, estes pré-processamentos apresentaram resultados plausíveis para os conjuntos amostrais de soluções, mas apenas os modelos GRO em hexano, DRO em hexano e ORO em etanol apresentam melhores resultados. Optou-se por utilizar apenas a centragem na média para o modelo DRO em hexano, dado que os resultados de validação foram próximos e não houve um aumento da demanda computacional.

Os pré-processamentos escolhidos como ponto de partida para a seleção de variáveis estão em **negrito** na Tabela 42. Após a seleção de variáveis pelo algoritmo GA, os pré-processamentos foram avaliados novamente e o modelo com melhores resultados de predição do conjunto de validação teve as figuras de mérito (sensibilidade, seletividade, precisão etc.) determinadas.

Tabela 42 – Avaliação dos pré-processamentos para os conjuntos de amostras de solos fortificados e soluções de hexano e etanol com os contaminantes.

Solos fortificados							
GRO							
Pré-processamento	#V.L.	% Var. Exp. X	% Var. Exp. Y	RMSEC	RMSEP	R ² cal	R ² val
autoescalamamento	9	100,00	88,53	1293	2086	0,8853	0,6987
C.M.	9	100,00	88,63	1288	2087	0,8863	0,6992
C.L.B.	10	99,97	94,15	924	2070	0,9415	0,7054
1ª der.	10	99,59	94,71	879	2122	0,9471	0,6913
2ª der.	10	98,41	99,47	277	2218	0,9947	0,6592
MSC	11	99,93	95,68	794	1785	0,9568	0,7803
SNV	12	99,93	95,76	786	1743	0,9576	0,7909
OSC	12	100,00	95,93	770	2042	0,9593	0,7138
GLSW	9	99,62	88,68	1285	2087	0,8868	0,6993
EPO	12	99,96	96,17	747	2050	0,9617	0,7121
EMM	4	50,27	99,48	275	2181	0,9948	0,6815
DRO							
Pré-processamento	#V.L.	% Var. Exp. X	% Var. Exp. Y	RMSEC	RMSEP	R ² cal	R ² val
autoescalamamento	10	100,00	92,15	1537	2484	0,9215	0,7646
C.M.	10	100,00	92,10	1541	2395	0,9210	0,7840
C.L.B.	9	99,94	91,70	1580	2365	0,9170	0,7934
1ª der.	6	99,15	90,70	1673	2625	0,9070	0,7391
2ª der.	5	97,93	91,84	1566	3005	0,9184	0,6664
MSC	9	99,87	93,53	1395	1985	0,9353	0,8605
SNV	8	99,63	91,59	1590	2076	0,9159	0,8364
OSC	9	100,00	91,14	1632	2214	0,9114	0,8127
GLSW	10	100,00	91,20	1626	2220	0,9120	0,8121
EPO	9	100,00	91,47	1601	2327	0,9147	0,7962
EMM	4	99,71	98,93	567	2744	0,9893	0,7415
ORO							
Pré-processamento	#V.L.	% Var. Exp. X	% Var. Exp. Y	RMSEC	RMSEP	R ² cal	R ² val
autoescalamamento	7	99,99	76,64	1406	1658	0,7664	0,6352
C.M.	7	99,99	76,42	1413	1652	0,7642	0,6393
C.L.B.	10	99,96	95,10	644	1481	0,9510	0,7141
1ª der.	3	94,82	72,42	1528	1527	0,7242	0,7239
2ª der.	3	92,97	77,49	1381	1894	0,7749	0,5195
MSC	6	99,17	78,03	1364	1488	0,7803	0,7239
SNV	10	99,95	85,88	1093	1498	0,8588	0,7021
OSC	7	100,00	76,78	1402	1740	0,7678	0,5955
GLSW	7	100,00	76,78	1402	1740	0,7678	0,5955
EPO	6	99,61	76,53	1410	1711	0,7653	0,6106
EMM	4	38,63	99,73	151	1821	0,9973	0,5654

Continua.

Tabela 42 – Avaliação dos pré-processamentos para os conjuntos de amostras de solos fortificados e soluções de hexano e etanol com os contaminantes. (Continuação)

Soluções em hexano							
GRO							
Pré-processamento	#V.L.	% Var. Exp. X	% Var. Exp. Y	RMSEC	RMSEP	R ² cal	R ² val
autoescalamiento	10	99,93	99,46	205	402	0,9946	0,9837
C.M.	9	99,90	98,95	286	403	0,9895	0,9821
C.L.B.	9	98,17	99,76	136	400	0,9976	0,9820
1ª der.	6	85,86	99,04	274	461	0,9904	0,9741
2ª der.	8	60,74	99,91	86	1411	0,9991	0,7397
MSC	9	99,54	99,86	105	386	0,9986	0,9832
SNV	9	99,53	99,81	120	386	0,9981	0,9835
OSC	9	100,00	99,87	101	370	0,9987	0,9845
GLSW	11	100,00	99,87	101	379	0,9987	0,9835
EPO	10	100,00	99,86	106	376	0,9986	0,9837
EMM	4	99,99	99,34	227	406	0,9934	0,9804
DRO							
Pré-processamento	#V.L.	% Var. Exp. X	% Var. Exp. Y	RMSEC	RMSEP	R ² cal	R ² val
autoescalamiento	11	99,80	99,82	158	1324	0,9982	0,8680
C.M.	9	99,79	97,95	529	1214	0,9795	0,8845
C.L.B.	9	99,21	99,68	210	1254	0,9968	0,8730
1ª der.	6	95,47	97,18	621	1304	0,9718	0,8632
2ª der.	4	80,43	90,65	1131	1928	0,9065	0,6896
MSC	9	97,38	99,52	256	1241	0,9952	0,8777
SNV	9	97,38	99,52	256	1241	0,9952	0,8777
OSC	7	100,00	99,69	170	1218	0,9979	0,8809
GLSW	9	100,00	97,50	585	1211	0,9750	0,8834
EPO	10	100,00	99,61	233	1211	0,9961	0,8817
EMM	4	99,99	99,49	265	1281	0,9949	0,8660
ORO							
Pré-processamento	#V.L.	% Var. Exp. X	% Var. Exp. Y	RMSEC	RMSEP	R ² cal	R ² val
autoescalamiento	8	99,90	96,67	579	757	0,9667	0,9392
C.M.	8	99,87	97,16	536	923	0,9716	0,9112
C.L.B.	6	98,95	95,72	657	908	0,9572	0,9133
1ª der.	6	93,71	98,70	362	1176	0,9870	0,8529
2ª der.	7	80,15	99,64	190	1591	0,9964	0,7375
MSC	9	99,25	99,63	194	800	0,9963	0,9332
SNV	9	99,24	99,63	194	803	0,9963	0,9326
OSC	7	100,00	99,77	152	812	0,9977	0,9308
GLSW	8	100,00	95,74	656	941	0,9574	0,9076
EPO	10	100,00	99,84	126	862	0,9984	0,9215
EMM	3	99,99	98,35	408	886	0,9835	0,9172

Continua.

Tabela 42 – Avaliação dos pré-processamentos para os conjuntos de amostras de solos fortificados e soluções de hexano e etanol com os contaminantes. (Continuação)

Soluções em etanol							
Pré-processamento	GRO						
	#V.L.	% Var. Exp. X	% Var. Exp. Y	RMSEC	RMSEP	R ² cal	R ² val
autoescalamiento	11	99,85	99,86	64	560	0,9986	0,9480
C.M.	9	99,78	99,05	169	731	0,9905	0,9102
C.L.B.	9	99,21	99,66	100	643	0,9966	0,9283
1ª der.	6	82,71	95,61	362	1273	0,9561	0,5907
2ª der.	3	33,35	86,09	644	1541	0,8609	0,4014
MSC	8	98,53	99,21	154	702	0,9921	0,9037
SNV	8	98,52	99,21	154	700	0,9921	0,9043
OSC	6	100,00	99,20	209	781	0,9853	0,8855
GLSW	8	100,00	97,87	252	749	0,9787	0,9029
EPO	7	100,00	97,57	269	738	0,9757	0,9045
EMM	2	99,98	95,34	373	765	0,9534	0,8666
DRO							
Pré-processamento	#V.L.	% Var. Exp. X	% Var. Exp. Y	RMSEC	RMSEP	R ² cal	R ² val
autoescalamiento	11	99,69	99,90	75	320	0,9990	0,9832
C.M.	10	99,68	99,88	83	354	0,9988	0,9798
C.L.B.	9	99,06	99,80	107	384	0,9980	0,9757
1ª der.	10	74,01	99,91	72	830	0,9991	0,8873
2ª der.	4	19,85	96,73	429	1921	0,9673	0,3772
MSC	9	98,72	99,90	77	356	0,9990	0,9804
SNV	9	98,72	99,90	77	354	0,9990	0,9806
OSC	11	100,00	99,94	57	355	0,9994	0,9796
GLSW	11	100,00	99,93	63	356	0,9993	0,9794
EPO	10	100,00	99,92	69	363	0,9992	0,9784
EMM	4	100,00	99,66	139	413	0,9966	0,9719
ORO							
Pré-processamento	#V.L.	% Var. Exp. X	% Var. Exp. Y	RMSEC	RMSEP	R ² cal	R ² val
autoescalamiento	9	99,96	90,67	274	590	0,9067	0,5981
C.M.	8	99,97	94,95	201	569	0,9495	0,6214
C.L.B.	8	98,72	94,91	202	554	0,9491	0,6418
1ª der.	6	69,00	93,86	222	704	0,9386	0,4271
2ª der.	5	31,15	97,82	133	793	0,9782	0,2873
MSC	8	99,85	94,54	209	545	0,9454	0,6506
SNV	8	99,84	94,55	209	542	0,9455	0,6551
OSC	5	100,00	63,58	541	539	0,6358	0,6727
GLSW	6	100,00	69,08	499	531	0,6908	0,6755
EPO	6	100,00	67,46	512	525	0,6746	0,6800
EMM	4	100,00	98,01	127	550	0,9801	0,6471

5.3.3.2. Avaliação da seleção de variáveis

Os espectros NIR são constituídos, normalmente, por alguns milhares de variáveis devido à alta resolução dos espectrômetros atuais. Muitas das variáveis espectrais estão correlacionadas entre si por fazerem parte de uma mesma banda espectral e, também, devido aos grupos funcionais apresentarem respostas em mais de uma região do infravermelho próximo, isto é, no primeiro, segundo e terceiro sobretom. A grande quantidade de covariância na matriz \mathbf{X} pode ofuscar variáveis que estão correlacionadas com a propriedade de interesse, o que diminui a capacidade preditiva do modelo quimiométrico. Portanto, a redução da covariância na matriz \mathbf{X} pode tornar o modelo mais ajustado e robusto.

Na seção anterior, foram discutidos os pré-processamentos utilizados para diminuir covariância no bloco \mathbf{X} , tais como OSC, GLSW, EPO e EMM, que objetivaram reduzir o peso de variáveis pouco correlacionadas com o vetor \mathbf{y} . Outro caminho para reduzir a covariância é através da seleção de variáveis.¹⁶⁴ A remoção de variáveis, onde o ruído domina a informação relacionada ao analito, frequentemente leva a uma melhor exatidão do método analítico e já é uma técnica bem aceita.^{165,189} A seleção de variáveis pode ser realizada através do conhecimento espectral (seleção manual) ou através de algoritmos que buscam minimizar o erro de predição, tais como o algoritmo genético (GA), interval Partial Least Square (iPLS), algoritmo de projeções sucessivas (SPA), entre outros. O algoritmo genético foi escolhido para efetuar a seleção de variáveis devido ao tipo de dados da matriz \mathbf{X} e à forma de busca por melhores variáveis relacionadas à propriedade de interesse. Dado que o GA é uma técnica de otimização heurística populacional que emprega um processo probabilístico de busca não local, manipulando cordas binárias (duas em duas variáveis), torna-se ideal para análise de espectros cujas variáveis vizinhas são altamente correlacionadas. De acordo com a literatura, além da melhor compreensão do sistema químico proporcionada pelo GA, a calibração de misturas de compostos com espectros quase idênticos tem sido realizada com sucesso após uso do algoritmo.^{164,165,190}

Sendo o GA um algoritmo probabilístico, utilizou-se o tamanho de população (número de modelos) e o número de gerações máximos permitidos pelo software PLS toolbox para aumentar a probabilidade de se obter o modelo com melhor capacidade preditiva. Os erros dos modelos foram avaliados através dos valores de RMSECV determinados a partir da predição de grupos de 5 amostras. As variáveis selecionadas pela primeira execução do GA (GA1) foram utilizadas para o desenvolvimento do modelo PLS-GA1 e para a segunda execução do GA (GA2). O número de variáveis, o pré-processamento e os parâmetros analíticos resultantes das

modelagens estão dispostos na Tabela 43 para apenas as frações de TPH, para os demais analitos (BTEX e PAHs) estão dispostos na Tabela C1, Anexo C.

Tabela 43 – Parâmetros analíticos dos modelos PLS das frações TPH desenvolvidos com os espectros completos e com variáveis selecionadas. Os valores de RMSE estão na unidade mg kg^{-1} .

Analito	GRO										
	Conjunto	Solos fortificados				Soluções em hexano			Soluções em etanol		
		Espectro	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1
#variáveis		3112	1818	498	76	3112	790	247	3112	759	199
Pré-processamento		MSC	SNV	SNV	SNV	OSC	C.M.	Autoes.	Autoes.	Autoes.	Autoes.
#V.L.		9	12	8	11	9	11	9	11	11	12
Variância X (%)		99,70	99,93	99,60	99,89	100,00	99,92	99,94	99,85	99,87	99,89
Variância y (%)		95,74	95,76	93,03	97,87	99,87	99,87	99,82	99,86	99,85	99,92
RMSEC		788	786	1008	557	101	100	118	64	67	48
RMSECV		1453	1530	1179	841	651	408	281	754	353	192
RMSEP		1967	1743	1612	1298	370	208	146	560	334	148
R ² cal		0,9574	0,9576	0,9303	0,9787	0,9987	0,9987	0,9982	0,9986	0,9985	0,9992
R ² cv		0,8561	0,8431	0,9047	0,9516	0,9480	0,9795	0,9903	0,8252	0,9584	0,9879
R ² val		0,7327	0,7909	0,8181	0,8844	0,9845	0,9950	0,9975	0,9480	0,9795	0,9946
Analito	DRO										
	Conjunto	Solos fortificados				Soluções em hexano			Soluções em etanol		
		Espectro	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1
#variáveis		3112	1811	517	154	3112	728	206	3112	706	183
Pré-processamento		MSC	MSC	MSC	MSC	C.M.	OSC	OSC	Autoes.	Autoes.	Autoes.
#V.L.		7	9	9	12	9	4	4	11	12	12
Variância X (%)		99,21	99,87	99,81	99,84	99,79	100,00	100,00	99,69	99,72	99,74
Variância y (%)		94,85	93,53	97,42	99,66	97,95	99,99	99,83	99,90	99,98	99,96
RMSEC		1245	1395	880	322	529	45	155	75	37	50
RMSECV		2749	2086	1665	1076	1336	887	594	423	231	151
RMSEP		2766	1985	1590	788	1214	725	368	320	189	129
R ² cal		0,9485	0,9353	0,9742	0,9966	0,9795	0,9999	0,9983	0,9990	0,9998	0,9996
R ² cv		0,7511	0,8562	0,9088	0,9620	0,8755	0,9441	0,9746	0,9683	0,9907	0,9962
R ² val		0,7195	0,8605	0,9193	0,9778	0,8845	0,9583	0,9887	0,9832	0,9944	0,9975
Analito	ORO										
	Conjunto	Solos fortificados				Soluções em hexano			Soluções em etanol		
		Espectro	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1
#variáveis		3112	1918	426	108	3112	757	220	3112	785	169
Pré-processamento		C.M.	MSC	MSC	MSC	Autoes.	Autoes.	Autoes.	EPO (1 PC)	Autoes.	Autoes.
#V.L.		7	6	9	9	8	9	8	6	11	11
Variância X (%)		99,97	99,17	99,90	99,91	99,90	99,92	99,89	100,00	99,98	99,97
Variância y (%)		78,79	78,03	81,47	83,12	96,67	99,00	99,34	67,46	97,81	98,25
RMSEC		1340	1364	845	691	579	318	258	512	133	118
RMSECV		1550	1518	1188	1040	1200	720	496	777	513	375
RMSEP		1809	1488	1334	950	757	475	368	525	366	208
R ² cal		0,7879	0,7803	0,9156	0,9437	0,9667	0,9900	0,9934	0,6746	0,9781	0,9825
R ² cv		0,7185	0,7297	0,8346	0,8723	0,8577	0,9525	0,9772	0,3134	0,6739	0,8291
R ² val		0,5651	0,7239	0,7641	0,8882	0,9392	0,9762	0,9873	0,6800	0,8629	0,9517

Na análise dos solos fortificados, primeiramente, foi realizada uma seleção manual das variáveis – a mesma realizada na seção anterior – e o GA1 foi executado a partir das variáveis não excluídas. O número de variáveis selecionadas em cada execução não ultrapassou 30% das variáveis iniciais, dado que foram utilizadas 30% das variáveis totais para dar início à seleção pelo algoritmo.

Na Tabela 43, nota-se que o autoescalamiento das variáveis foi o pré-processamento mais adequado para a predição de GRO em hexano e ORO em etanol. Com a redução de variáveis pouco ou não correlacionadas com o vetor \mathbf{y} , o autoescalamiento proporcionou maior peso para bandas pouco intensas, como o terceiro sobretom das vibrações C-H de metilas e a região espectral próxima à 4600 cm^{-1} , relacionada a vibrações C-H de anéis aromáticos. Como já discutido na seção 5.1.2.4., as vibrações $3\nu\text{C-H}$ de metila foram mais presentes nas amostras de óleo lubrificante e $\nu\text{C-H}$ aromático nas amostras de gasolina, assim, o autoescalamiento forneceu melhores resultados após o uso de GA.

A seleção de variáveis quase não alterou os valores de variância explicada no bloco \mathbf{X} , contudo, houve um aumento considerável na variância explicada no bloco \mathbf{y} , principalmente para predição de ORO em solos fortificados e em etanol. Este aumento é resultante da maior correlação entre a concentração da fração ORO e as variáveis espectrais. Apesar da redução do número de variáveis x , o número de VL aumentou para a maioria dos modelos. Visto que as amostras são misturas complexas com diversos sinais analíticos, a diminuição da covariância e do ruído na matriz \mathbf{X} permitiu que mais sinais analíticos fossem comprimidos e representados através de um maior número de VL.

Para o conjunto de calibração, apenas o modelo de predição de GRO em hexano não foi otimizado pelo GA (Tabela 43). Devido ao GA buscar variáveis que proporcionem menores valores de RMSECV, todos os modelos apresentaram uma redução significativa ($F_{\text{calculado}}$ acima de 1,45) para a validação cruzada. Observou-se então os valores de RMSEP e R^2_{val} para avaliar a melhora na capacidade preditiva para amostras externas. Para todos os modelos, o GA proporcionou uma redução nos valores de RMSEP e um aumento nos valores de R^2_{val} , portanto, a seleção de variáveis, tanto a manual como pelo GA, tornou os modelos mais adequados para a predição de amostras externas.

Embora a aplicação do GA tenha resultado em modelos mais ajustados, cada execução do algoritmo levou cerca de 4 h utilizando um computador com processador Intel Core i7 3632QM 2,20 GHz e 8,0 Gb de memória RAM. Desta forma, foi importante avaliar se a melhora proporcionada pelo algoritmo foi significativa ou não, pois, havendo a necessidade de

alimentação dos modelos com novas amostras e revalidação periódica destes, o tempo demandado pela seleção de variáveis terá um peso na frequência analítica. Esta avaliação foi realizada através de teste F com os valores de RMSEP e os resultados para cada analito estão apresentados na Tabela 44. Para 40 amostras de validação e nível de confiança de 95%, o valor de F é igual a 1,70. Os valores de F calculado maiores que o valor de F tabelado mostram que a seleção de variáveis pelo GA reduziu significativamente os valores de RMSEP dos modelos gerados com o espectro completo. A primeira execução do GA não reduziu significativamente os erros de predição apenas para os analitos GRO e tolueno em solos fortificados e pireno e benzo(a)antraceno em soluções de etanol, contudo, o uso sequencial de duas execuções do GA reduziu significativamente os erros para todos os analitos.

Tabela 44 – Avaliação da redução dos valores de RMSEP pelo uso de GA através de teste F.

Analito	F calculado					
	Solos fortificados		Soluções em hexano		Soluções em etanol	
	Completo/ GA1	Completo/ GA2	Completo/ GA1	Completo/ GA2	Completo/ GA1	Completo/ GA2
GRO	1,49	2,30	3,18	6,41	2,80	14,24
DRO	3,03	12,31	2,81	10,86	2,87	6,17
ORO	1,84	3,63	2,54	4,22	2,06	6,40
benzeno	2,51	5,18	3,08	8,80	3,20	10,69
tolueno	1,62	4,90	3,71	8,05	2,60	6,00
etilbenzeno	1,84	4,18	4,14	7,35	3,14	6,87
m,p-xilenos	1,85	3,50	3,08	8,08	2,08	6,27
o-xileno	1,93	4,38	3,37	7,09	2,03	8,28
naftaleno (GRO)	1,95	4,65	2,13	7,61	2,37	8,88
naftaleno (DRO)	2,37	8,31	3,54	16,64	2,65	7,51
acenaftileno	3,57	14,04	2,95	6,90	4,13	13,60
fluoreno	4,09	14,52	3,58	6,65	2,19	6,82
fenantreno	4,09	14,52	2,16	6,37	2,40	3,68
antraceno	2,65	7,34	1,94	4,65	3,78	5,31
pireno	2,50	6,34	2,10	3,36	1,47	1,86
benzo(a)antraceno	3,87	8,41	2,25	6,95	1,64	5,39
criseno	3,11	5,37	2,62	6,73	1,87	7,29
benzo(b)fluoranteno	2,39	4,09	1,75	5,43	2,37	9,63
F tabelado	1,70 (nível de confiança 95% e 39 graus de liberdade)					

O teste F revelou que a redução dos erros proporcionada pela seleção de variáveis por GA foi significativa para a predição de amostras externas ao conjunto de calibração, desta forma, o tempo demandado pela execução do algoritmo foi compensado pelo ajuste do modelo.

5.3.3.3. Avaliação das figuras de mérito

Definidas as variáveis e os pré-processamentos a serem utilizados para obtenção dos modelos PLS com melhor capacidade preditiva, as demais figuras de mérito dos modelos finais foram determinadas. Nas seguintes Tabelas (Tabelas 44, 45 e 46), estão presentes os parâmetros analíticos determinados através dos conjuntos de calibração, validação (I) e replicatas para cada analito nos diferentes conjuntos amostrais.

De acordo com a norma ASTM E1655, o número de variáveis latentes (VL) deve ser menor que $(n - 6)/6$ [seção 25.1.4.1]¹⁷⁹, onde n é o número de amostras de calibração. Portanto, sendo n igual a 80, o número máximo de VL foi igual a 12. Embora 12 VL pareça uma quantidade demasiada de variáveis explicativas, as amostras apresentam grande nível de complexidade devido à variabilidade de sinais analíticos de diferentes grupos funcionais e moleculares. Segundo Wise et al. (2006), uma boa forma de determinar o número de VL para o modelo de calibração é através da adição de VL que aumente pelo menos 2% da variância explicada.¹⁷⁶ Normalmente, a variância explicada de modelo desenvolvido com os dados autoescalados apresenta-se distribuída em mais variáveis latentes.¹⁵⁹ Desta forma, alguns modelos desenvolvidos apresentaram até 13 VL.

Todos os modelos PLS foram isentos de amostras anômalas (outliers), dado que os conjuntos amostrais foram constituídos de amostras de solos ou solventes fortificados com os contaminantes. As amostras foram preparadas utilizando uma balança analítica calibrada e não houve ocorrência de erros de transcrição de valores numéricos.

Os intervalos de medição das frações de TPH variaram de 0 mg kg⁻¹ (brancos) até, em média, 12.260 mg kg⁻¹. A concentração dos demais analitos foi relativa à concentração destes nas amostras de gasolina ou diesel. Dentre os analitos individuais, o tolueno apresentou intervalo de medição mais amplo (0-3724,7 mg kg⁻¹ em solo) devido a seu elevado teor determinado nas amostras de gasolina premium.

Tabela 45 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de amostras de solo fortificado.

Analito	GRO	benzeno	tolueno	etilbenzeno	m,p-xilenos	o-xileno	naftaleno (GRO)
#variáveis espectrais	76	108	88	105	81	81	120
#variáveis latentes	11	12	11	11	11	11	11
% Variância bloco X	99,89	99,94	99,89	99,95	99,94	99,95	99,92
% Variância bloco y	97,87	98,69	97,59	96,39	96,27	96,47	97,77
Conjunto de calibração							
C máxima (mg kg ⁻¹)	14561,3	136,8	3724,7	171,1	619,2	260,0	43,2
RMSEC (mg kg ⁻¹)	557,2	3,9	86,4	6,5	23,4	9,4	1,4
RMSECV (mg kg ⁻¹)	841,1	7,9	160,7	11,3	38,1	15,2	2,8
Tendência (cal)	2,2.10 ⁻¹¹	-3,2.10 ⁻¹⁴	4,0.10 ⁻¹²	9,7.10 ⁻¹²	-4,3.10 ⁻¹¹	9,7.10 ⁻¹²	-9,4.10 ⁻¹²
Tendência (CV)	3,0.10 ¹	6,1.10 ⁻¹	8,5	4,5.10 ⁻²	-1,1	-7,1.10 ⁻²	-5,9.10 ⁻²
R ² cal	0,9787	0,9869	0,9759	0,9639	0,9627	0,9647	0,9777
R ² ajustado	0,9781	0,9865	0,9769	0,9632	0,9618	0,9644	0,9776
R ² CV	0,9516	0,9474	0,9208	0,8924	0,9005	0,9073	0,9164
Razão S/N mínimo	0,2	1,7	1,6	0,6	0,1	1,0	1,0
Razão S/N máximo	30,8	55,7	42,3	42,6	30,9	41,8	38,6
Seletividade média	2,9.10 ⁻²	1,1.10 ⁻²	7,2.10 ⁻³	9,0.10 ⁻³	8,1.10 ⁻³	7,9.10 ⁻³	7,1.10 ⁻³
Intercepto	9,1.10 ⁻³	9,1.10 ⁻³	1,2.10 ⁻²	2,1.10 ⁻⁴	1,9.10 ⁻⁴	1,8.10 ⁻⁴	1,3.10 ⁻⁴
Inclinação	1,4.10 ⁻⁵	8,7.10 ⁻⁴	4,3.10 ⁻⁵	1,6.10 ⁻⁵	3,7.10 ⁻⁶	8,5.10 ⁻⁶	4,9.10 ⁻⁵
Sensibilidade	5,7.10 ⁻²	3,0.10 ⁻²	2,5.10 ⁻²	5,7.10 ⁻⁴	4,6.10 ⁻⁴	4,4.10 ⁻⁴	4,8.10 ⁻⁴
Sens. analítica (mg kg ⁻¹) ⁻¹	8,5	1,4.10 ¹	6,0	8,4	5,9	7,8	8,3
Inv. sens. analit. (mg kg ⁻¹)	1,2.10 ⁻¹	7,2.10 ⁻²	1,7.10 ⁻¹	1,2.10 ⁻¹	1,7.10 ⁻¹	1,3.10 ⁻¹	1,2.10 ⁻¹
LD (mg kg ⁻¹)	0,4	0,2	0,6	0,4	0,6	0,4	0,4
LQ (mg kg ⁻¹)	1,2	0,7	1,7	1,2	1,7	1,3	1,2
Erro relativo médio (%)	90,4	39,2	117,1	104,1	228,5	122,1	27,6
RPD	4,6	4,4	3,5	3,1	3,2	3,3	3,5
Conjunto de validação							
C máxima (mg kg ⁻¹)	14216,8	128,3	2748,8	149,9	524,1	217,6	42,1
RMSEP	1298,5	9,9	161,5	11,7	45,6	16,8	2,6
Tendência (val)	186,2	0,2	6,8	1,1	2,7	2,1	0,6
t bias	0,96	0,13	0,28	0,59	0,39	0,79	1,43
R ² val	0,8844	0,9137	0,8929	0,8961	0,8607	0,8892	0,9388
r val	0,9404	0,9559	0,9449	0,9466	0,9277	0,9430	0,9689
Erro relativo médio (%)	220,1	79,7	170,8	82,0	219,8	187,8	44,3
RPD	3,0	3,4	3,1	3,1	2,7	3,1	4,0
Conjunto de replicatas							
Precisão (mg kg ⁻¹)	697,1	8,1	317,0	9,4	41,2	16,1	4,1
DPR mín. (%)	3,4	3,3	15,5	4,3	5,9	4,9	4,9
DPR máx. (%)	1564,5	8702,6	730,4	591,4	1093,9	2809,1	363,1
DRP médio > intervenção (%)	23,0	45,9	190,9	51,2	71,3	65,8	104,9
*DPR 20 replicatas C baixa	34,8% (1119)	55,9% (11,3)	359,1% (41,6)	41,3% (12,4)	82,7% (41,9)	44,8% (17,2)	52,7% (3,4)
*DPR 20 replicatas C alta	4,4% (11681)	3,7% (118,7)	61,2% (434,8)	6,7% (129,6)	10,5% (437,1)	8,2% (180,1)	12,3% (36,0)

* Os valores entre parênteses representam as concentrações dos analitos em mg kg⁻¹.

Continua

Tabela 45 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de amostras de solo fortificado. (Continuação).

Analito	DRO	naftaleno (DRO)	acenaftileno	fluoreno	fenantreno	antraceno
#variáveis espectrais	154	87	97	143	143	126
#variáveis latentes	12	9	11	12	12	9
% Variância bloco X	99,84	99,81	99,77	99,82	99,82	99,78
% Variância bloco y	99,66	98,53	99,31	99,64	99,64	98,74
Conjunto de calibração						
C máxima (mg kg ⁻¹)	19881,6	163,1	146,3	211,7	174,1	103,8
RMSEC (mg kg ⁻¹)	322,1	5,2	3,1	3,2	3,2	2,9
RMSECV (mg kg ⁻¹)	1075,7	9,0	6,5	9,6	9,6	5,6
Tendência (cal)	-2,9.10 ⁻¹⁰	2,9.10 ⁻¹²	5,6.10 ⁻¹²	1,7.10 ⁻¹²	1,7.10 ⁻¹²	-6,7.10 ⁻¹²
Tendência (CV)	-4,4.10 ¹	3,5.10 ⁻²	5,8.10 ⁻¹	-3,2.10 ⁻¹	-3,2.10 ⁻¹	5,0.10 ⁻¹
R ² cal	0,9966	0,9853	0,9931	0,9964	0,9964	0,9874
R ² ajustado	0,9964	0,9847	0,9931	0,9963	0,9903	0,9872
R ² CV	0,9620	0,9552	0,9702	0,9670	0,9670	0,9533
Razão S/N mínimo	1,3	0,3	0,5	3,2	1,6	1,7
Razão S/N máximo	76,8	55,0	14,9	115,6	59,1	63,1
Seletividade média	7,8.10 ⁻³	1,5.10 ⁻²	4,2.10 ⁻³	7,5.10 ⁻³	8,0.10 ⁻³	7,9.10 ⁻³
Intercepto	6,3.10 ⁻⁵	6,0.10 ⁻⁵	3,0.10 ⁻⁴	7,4.10 ⁻⁵	8,8.10 ⁻⁵	8,8.10 ⁻⁵
Inclinação	1,3.10 ⁻⁷	1,3.10 ⁻⁵	6,2.10 ⁻⁶	1,1.10 ⁻⁵	1,5.10 ⁻⁵	2,5.10 ⁻⁵
Sensibilidade	7,0.10 ⁻⁴	7,0.10 ⁻⁴	1,3.10 ⁻²	5,9.10 ⁻⁴	6,2.10 ⁻⁴	6,4.10 ⁻⁴
Sens. analítica (mg kg ⁻¹) ⁻¹	2,1.10 ¹	1,8.10 ¹	1,1.10 ²	2,8.10 ¹	1,3.10 ¹	1,4.10 ¹
Inv. sens. analit. (mg kg ⁻¹)	4,9.10 ⁻²	5,5.10 ⁻²	9,3.10 ⁻³	3,6.10 ⁻²	7,6.10 ⁻²	6,9.10 ⁻²
LD (mg kg ⁻¹)	0,2	0,2	0,0	0,1	0,3	0,2
LQ (mg kg ⁻¹)	0,5	0,6	0,1	0,4	0,8	0,7
Erro relativo médio (%)	50,5	19,7	29,2	14,5	17,6	19,3
RPD	5,1	4,8	5,8	5,5	4,8	4,5
Conjunto de validação						
C máxima (mg kg ⁻¹)	17954,3	147,7	130,0	179,6	170,2	85,5
RMSEP	788,5	8,0	5,7	8,3	8,3	4,5
Tendência (val)	166,6	0,7	1,3	2,5	2,5	1,9
t bias	1,40	0,54	2,38	2,05	1,58	3,07
R ² val	0,9778	0,9654	0,9756	0,9773	0,9773	0,9707
r val	0,9888	0,9825	0,9877	0,9886	0,9886	0,9852
Erro relativo médio (%)	120,8	36,5	62,0	39,2	45,6	36,4
RPD	4,0	5,3	6,3	6,2	4,6	5,4
Conjunto de replicatas						
Precisão (mg kg ⁻¹)	1658,8	12,3	14,0	17,2	22,0	8,2
DPR mín. (%)	5,8	3,0	8,8	6,0	6,0	4,3
DPR máx. (%)	1807,4	686,7	2378,9	455,2	455,2	641,9
DRP médio > intervenção (%)	30,1	27,6	31,4	82,4	117,5	79,9
*DPR 20 replicatas C baixa	79,1% (1082)	65,9% (8,9)	137,3% (8,8)	66,2% (10,8)	66,2% (7,7)	129,8% (4,2)
*DPR 20 replicatas C alta	5,8% (19881)	6,35% (163,0)	8,8% (119,9)	5,6% (211,7)	5,6% (174,0)	6,0% (103,8)

* Os valores entre parênteses representam as concentrações dos analitos em mg kg⁻¹.

Continua

Tabela 45 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de amostras de solo fortificado. (Continuação).

Analito	pireno	benzo(a)antraceno	criseno	benzo(b)fluoranteno	ORO
#variáveis espectrais	121	143	116	100	1918
#variáveis latentes	10	10	9	10	6
% Variância bloco X	99,78	99,84	99,76	99,84	99,17
% Variância bloco y	97,12	97,5	97,22	95,96	78,03
Conjunto de calibração					
C máxima (mg kg ⁻¹)	91,7	71,3	45,7	56,6	10892,7
RMSEC (mg kg ⁻¹)	3,7	3,0	1,6	2,4	690,5
RMSECV (mg kg ⁻¹)	7,3	5,8	2,9	4,4	1040,2
Tendência (cal)	3,7.10 ⁻¹²	3,7.10 ⁻¹²	9,4.10 ⁻¹³	-2,5.10 ⁻¹²	1,3.10 ⁻⁹
Tendência (CV)	4,8.10 ⁻¹	2,6.10 ⁻¹	8,4.10 ⁻²	1,1.10 ⁻¹	19,7
R ² cal	0,9712	0,9750	0,9722	0,9596	0,9437
R ² ajustado	0,9703	0,9740	0,9712	0,9588	0,9418
R ² CV	0,8883	0,9090	0,9059	0,8666	0,8723
Razão S/N mínimo	0,0	0,9	1,3	0,7	0,6
Razão S/N máximo	32,8	28,3	43,1	29,0	18,5
Seletividade média	6,6.10 ⁻³	5,9.10 ⁻³	6,7.10 ⁻³	5,3.10 ⁻³	9,7.10 ⁻³
Intercepto	8,6.10 ⁻⁵	9,9.10 ⁻⁵	1,1.10 ⁻⁴	1,1.10 ⁻⁴	5,0.10 ⁻⁵
Inclinação	2,1.10 ⁻⁵	2,6.10 ⁻⁵	5,0.10 ⁻⁵	3,0.10 ⁻⁵	2,3.10 ⁻⁷
Sensibilidade	4,7.10 ⁻⁴	5,0.10 ⁻⁴	4,9.10 ⁻⁴	3,8.10 ⁻⁴	7,3.10 ⁻⁴
Sens. analítica (mg kg ⁻¹) ⁻¹	9,3	7,5	8,8	6,2	5,0
Inv. sens. analit. (mg kg ⁻¹)	1,1.10 ⁻¹	1,3.10 ⁻¹	1,1.10 ⁻¹	1,6.10 ⁻¹	2,0.10 ⁻¹
LD (mg kg ⁻¹)	0,4	0,4	0,4	0,5	0,7
LQ (mg kg ⁻¹)	1,1	1,3	1,1	1,6	2,0
Erro relativo médio (%)	31,5	31,6	26,7	44,2	129,1
RPD	3,0	3,3	3,3	2,8	2,8
Conjunto de validação					
C máxima (mg kg ⁻¹)	80,3	67,9	43,1	53,5	10117,3
RMSEP	5,0	4,2	2,8	4,5	949,6
Tendência (val)	0,1	0,8	0,7	0,9	163,6
t bias	0,14	1,22	1,94	1,31	1,15
R ² val	0,9528	0,9553	0,9358	0,8915	0,8882
r val	0,9761	0,9774	0,9674	0,9442	0,9424
Erro relativo médio (%)	48,9	64,3	69,8	87,3	157,2
RPD	4,5	4,5	3,8	2,9	2,9
Conjunto de replicatas					
Precisão (mg kg ⁻¹)	9,1	7,9	4,4	6,4	1112,0
DPR mín. (%)	9,5	7,7	9,4	10,9	6,6
DPR máx. (%)	512,7	1663,7	12272,8	10464,7	1315,7
DRP médio > intervenção (%)	162,9	179,3	179,1	229,8	30,1
*DPR 20 replicatas C baixa	314,4% (2,1)	964,9% (2,0)	1307,8% (1,0)	181,5% (1,6)	79,1% (1021)
*DPR 20 replicatas C alta	11,4% (80,3)	11,8% (46,4)	16,3% (24,0)	22,3% (19,3)	5,8% (10893)

* Os valores entre parênteses representam as concentrações dos analitos em mg kg⁻¹.

Tabela 46 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de soluções dos contaminantes em hexano.

Analito	GRO	benzeno	tolueno	etilbenzeno	m,p-xilenos	o-xileno	naftaleno (GRO)
#variáveis espectrais	247	204	116	178	194	183	181
#variáveis latentes	9	13	12	10	12	11	12
% Variância bloco X	99,94	99,95	99,96	99,93	99,94	99,94	99,93
% Variância bloco y	99,82	99,85	99,78	99,29	99,62	99,23	99,76
Conjunto de calibração							
C máxima (mg kg ⁻¹)	11823,8	96,4	2820,2	121,7	425,6	176,7	34,2
RMSEC (mg kg ⁻¹)	118,3	0,9	19,4	2,1	5,2	3,1	0,4
RMSECV (mg kg ⁻¹)	281,0	3,6	62,9	6,2	20,6	8,8	1,7
Tendência (cal)	1,0.10 ⁻¹²	7,2.10 ⁻¹³	-3,1.10 ⁻¹¹	6,9.10 ⁻¹³	1,9.10 ⁻¹²	6,8.10 ⁻¹³	-3,5.10 ⁻¹³
Tendência (CV)	-1,3.10 ¹	7,8.10 ⁻²	-1,4.10 ⁻²	-1,5.10 ⁻¹	-2,1	-3,7.10 ⁻¹	2,7.10 ⁻²
R ² cal	0,9982	0,9985	0,9978	0,9929	0,9962	0,9923	0,9976
R ² ajustado	0,9982	0,9985	0,9965	0,9928	0,9962	0,9922	0,9976
R ² CV	0,9903	0,9768	0,9768	0,9397	0,9460	0,9411	0,9464
Razão S/N mínimo	7,6	-1,1	8,0	3,2	6,0	4,5	6,3
Razão S/N máximo	147,2	192,1	133,7	58,8	115,5	66,5	110,5
Seletividade média	7,5.10 ⁻³	4,0.10 ⁻³	2,7.10 ⁻³	3,1.10 ⁻³	2,6.10 ⁻³	2,8.10 ⁻³	3,8.10 ⁻³
Intercepto	7,6.10 ⁻²	6,8.10 ⁻³	6,6.10 ⁻²	5,2.10 ⁻²	4,8.10 ⁻²	5,1.10 ⁻²	5,3.10 ⁻²
Inclinação	1,1.10 ⁻⁴	6,7.10 ⁻³	2,7.10 ⁻⁴	4,9.10 ⁻³	1,2.10 ⁻³	3,2.10 ⁻³	1,9.10 ⁻²
Sensibilidade	3,0.10 ⁻¹	1,6.10 ⁻¹	1,1.10 ⁻¹	1,2.10 ⁻¹	1,0.10 ⁻¹	1,1.10 ⁻¹	1,4.10 ⁻¹
Sens. analítica (mg kg ⁻¹) ⁻¹	3,3.10 ¹	4,6.10 ¹	1,8.10 ¹	1,1.10 ¹	2,1.10 ¹	1,2.10 ¹	2,2.10 ¹
Inv. sens. analit. (mg kg ⁻¹)	3,0.10 ⁻²	2,2.10 ⁻²	5,5.10 ⁻²	9,0.10 ⁻²	4,7.10 ⁻²	8,2.10 ⁻²	4,6.10 ⁻²
LD (mg kg ⁻¹)	0,1	0,1	0,2	0,3	0,2	0,3	0,2
LQ (mg kg ⁻¹)	0,3	0,2	0,6	0,9	0,5	0,8	0,5
Erro relativo médio (%)	42,2	82,5	177,7	178,9	334,7	146,6	37,3
RPD	10,0	6,6	6,6	4,0	4,1	4,0	4,3
Conjunto de validação							
C máxima (mg kg ⁻¹)	9282,6	94,9	2242,7	109,1	394,7	165,7	27,5
RMSEP	146,3	3,6	57,8	4,6	16,5	7,4	1,3
Tendência (val)	0,5	-0,2	8,7	1,0	0,8	0,5	0,3
t bias	0,02	0,31	0,72	1,35	0,29	0,39	1,44
R ² val	0,9975	0,9821	0,9809	0,9705	0,9678	0,9643	0,9671
r val	0,9988	0,9910	0,9904	0,9852	0,9838	0,9820	0,9834
Erro relativo médio (%)	70,0	153,5	497,7	170,0	537,2	421,9	74,6
RPD	19,1	6,9	7,2	5,7	5,6	5,3	5,4
Conjunto de replicatas							
Precisão (mg kg ⁻¹)	358,8	5,0	98,2	7,7	30,3	12,2	2,1
DPR mín. (%)	3,3	5,5	5,6	6,8	12,7	8,4	4,8
DPR máx. (%)	793,9	902,4	2028,7	5625,3	1039,8	1638,5	658,9
DRP médio > intervenção (%)	13,3	62,8	100,4	107,5	45,8	94,7	73,4
*DPR 20 replicatas C baixa	35,8% (1075)	35,4% (12,1)	58,9% (27,1)	48,8% (7,2)	60,5% (22,2)	69,9% (8,7)	36,7% (2,7)
*DPR 20 replicatas C alta	5,5% (6703)	5,5 (75,1)	32,8% (169,0)	12,3% (44,6)	16,2% (138,2)	15,0% (54,5)	13,2% (17,0)

* Os valores entre parênteses representam as concentrações dos analitos em mg kg⁻¹.

Continua

Tabela 46 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de soluções dos contaminantes em hexano. Continuação.

Analito	DRO	naftaleno	acenaftileno	fluoreno	fenantreno	antraceno
#variáveis espectrais	206	169	175	212	201	226
#variáveis latentes	4	6	3	3	5	3
% Variância bloco X	100	100	93,75	99,99	100	99,99
% Variância bloco y	99,83	99,87	99,99	99,61	99,82	99,61
Conjunto de calibração						
C máxima (mg kg ⁻¹)	14539,3	125,6	117,0	157,4	125,0	64,6
RMSEC (mg kg ⁻¹)	154,8	1,0	0,3	2,2	1,2	1,0
RMSECV (mg kg ⁻¹)	594,0	3,8	4,5	8,9	6,1	3,1
Tendência (cal)	-5,7.10 ⁻¹	1,1.10 ⁻³	6,7.10 ⁻¹³	3,0.10 ⁻²	-4,3.10 ⁻³	-6,3.10 ⁻³
Tendência (CV)	3,3.10 ¹	-1,3.10 ⁻¹	3,6.10 ⁻¹	-2,1.10 ⁻¹	2,0.10 ⁻¹	1,3.10 ⁻¹
R ² cal	0,9983	0,9987	0,9999	0,9961	0,9982	0,9961
R ² ajustado	0,9982	0,9987	0,9999	0,9961	0,9982	0,9961
R ² CV	0,9746	0,9832	0,9692	0,9419	0,9524	0,9644
Razão S/N mínimo	7,6	6,8	34,6	2,8	6,8	4,8
Razão S/N máximo	134,7	142,3	581,4	56,5	110,4	78,8
Seletividade média	1,3.10 ⁻²	9,9.10 ⁻³	3,3.10 ⁻¹	1,1.10 ⁻²	1,0.10 ⁻²	1,4.10 ⁻²
Intercepto	9,2.10 ⁻³	5,8.10 ⁻³	1,0	8,2.10 ⁻³	9,0.10 ⁻³	1,0.10 ⁻²
Inclinação	9,1.10 ⁻⁶	7,8.10 ⁻⁴	1,4.10 ⁻¹	7,5.10 ⁻⁴	1,0.10 ⁻³	2,2.10 ⁻³
Sensibilidade	3,4.10 ⁻²	2,2.10 ⁻²	3,5	2,7.10 ⁻²	3,0.10 ⁻²	3,5.10 ⁻²
Sens. analítica (mg kg ⁻¹) ⁻¹	3,2.10 ¹	3,1.10 ¹	1,2.10 ²	1,2.10 ¹	2,4.10 ¹	1,9.10 ¹
Inv. sens. analit. (mg kg ⁻¹)	3,1.10 ⁻²	3,3.10 ⁻²	8,4.10 ⁻³	8,4.10 ⁻²	4,2.10 ⁻²	5,3.10 ⁻²
LD (mg kg ⁻¹)	0,1	0,1	0,0	0,3	0,1	0,2
LQ (mg kg ⁻¹)	0,3	0,3	0,1	0,8	0,4	0,5
Erro relativo médio (%)	54,2	31,5	8,6	48,7	30,4	51,7
RPD	6,3	7,6	5,7	4,1	4,6	5,3
Conjunto de validação						
C máxima (mg kg ⁻¹)	13110,0	114,5	109,2	144,7	99,3	62,7
RMSEP	368,4	2,9	3,7	5,6	4,6	3,1
Tendência (val)	15,7	0,3	0,2	0,7	0,9	0,3
t bias	0,26	0,56	0,36	0,80	1,19	0,59
R ² val	0,9887	0,9906	0,9864	0,9749	0,9720	0,9722
r val	0,9943	0,9953	0,9932	0,9874	0,9859	0,9860
Erro relativo médio (%)	110,5	74,5	132,3	113,2	163,8	142,9
RPD	9,5	10,0	7,1	6,3	5,7	5,6
Conjunto de replicatas						
Precisão (mg kg ⁻¹)	1079,1	7,8	7,0	10,6	9,5	4,6
DPR mín. (%)	5,2	6,5	6,2	7,4	6,4	5,6
DPR máx. (%)	989,2	367,0	1445,9	795,7	1641,5	1247,7
DRP médio > intervenção (%)	29,1	26,1	39,5	87,6	56,0	28,2
*DPR 20 replicatas C baixa (%)	49,6% (1000)	50,6% (9,6)	64,9% (8,9)	44,3% (12,0)	56,5% (7,4)	50,2% (3,9)
*DPR 20 replicatas C alta (%)	10,2% (14539)	9,9% (114,5)	9,7% (109,2)	9,7% (144,7)	13,4% (90,4)	7,5% (64,6)

* Os valores entre parênteses representam as concentrações dos analitos em mg kg⁻¹.

Continua

Tabela 46 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de soluções dos contaminantes em hexano. Continuação.

Analito	Pireno	benzo(a)antraceno	criseno	benzo(b)fluoranteno	ORO
#variáveis espectrais	215	227	225	217	220
#variáveis latentes	4	8	6	6	8
% Variância bloco X	100	100	100	100	99,89
% Variância bloco y	96,15	99,55	98,56	99,72	99,34
Conjunto de calibração					
C máxima (mg kg ⁻¹)	68,4	53,6	27,3	32,6	14739,3
RMSEC (mg kg ⁻¹)	2,8	0,8	0,7	0,4	258,4
RMSECV (mg kg ⁻¹)	3,9	3,3	1,6	1,8	495,7
Tendência (cal)	-1,3.10 ⁻²	-1,5.10 ⁻³	-2,4.10 ⁻³	-1,3.10 ⁻³	1,3.10 ⁻¹¹
Tendência (CV)	1,7.10 ⁻²	2,5.10 ⁻¹	-1,3.10 ⁻¹	-8,8.10 ⁻²	5,7
R ² cal	0,9615	0,9955	0,9856	0,9972	0,9934
R ² ajustado	0,9610	0,9954	0,9854	0,9971	0,9933
R ² CV	0,9262	0,9192	0,9367	0,9450	0,9772
Razão S/N mínimo	-1,7	4,8	1,2	7,5	2,2
Razão S/N máximo	36,1	119,2	50,6	124,4	61,7
Seletividade média	6,5.10 ⁻³	2,2.10 ⁻³	5,1.10 ⁻³	6,1.10 ⁻³	5,0.10 ⁻³
Intercepto	7,7.10 ⁻³	2,6.10 ⁻³	5,2.10 ⁻³	7,6.10 ⁻³	6,8.10 ⁻²
Inclinação	1,5.10 ⁻³	7,2.10 ⁻⁴	2,5.10 ⁻³	2,8.10 ⁻³	6,7.10 ⁻⁵
Sensibilidade	2,3.10 ⁻²	8,3.10 ⁻³	1,6.10 ⁻²	2,1.10 ⁻²	2,2.10 ⁻¹
Sens. analítica (mg kg ⁻¹) ⁻¹	7,4	2,4.10 ¹	1,1.10 ¹	2,7.10 ¹	1,3.10 ¹
Inv. sens. analit. (mg kg ⁻¹)	1,4.10 ⁻¹	4,1.10 ⁻²	8,9.10 ⁻²	3,7.10 ⁻²	7,9.10 ⁻²
LD (mg kg ⁻¹)	0,4	0,1	0,3	0,1	0,3
LQ (mg kg ⁻¹)	1,4	0,4	0,9	0,4	0,8
Erro relativo médio (%)	190,4	46,1	79,5	51,7	95,1
RPD	3,7	3,5	4,0	4,2	6,4
Conjunto de validação					
C máxima (mg kg ⁻¹)	53,0	53,2	24,8	30,8	11085,7
RMSEP	3,9	2,3	1,2	1,7	368,5
Tendência (val)	1,3	0,0	-0,3	0,1	-123,6
t bias	2,15	0,12	1,54	0,29	2,20
R ² val	0,9347	0,9746	0,9609	0,9487	0,9873
r val	0,9668	0,9872	0,9802	0,9740	0,9936
Erro relativo médio (%)	271,9	223,0	101,0	211,5	188,4
RPD	3,7	6,0	5,0	4,2	8,4
Conjunto de replicatas					
Precisão (mg kg ⁻¹)	5,3	5,1	2,5	3,6	603,7
DPR mín. (%)	6,1	11,2	13,8	13,1	4,5
DPR máx. (%)	3101,7	888,8	1383,7	579,1	662,4
DRP médio > intervenção (%)	77,2	28,4	33,9	37,4	10,5
*DPR 20 replicatas C baixa	41,2% (1,4)	93,1% (1,5)	236,6% (0,9)	71,6% (1,0)	86,0% (831)
*DPR 20 replicatas C alta	15,5% (36,2)	19,7% (22,1)	21,1% (9,7)	21,6% (9,6)	6,4% (10119)

* Os valores entre parênteses representam as concentrações dos analitos em mg kg⁻¹.

Tabela 47 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de soluções dos contaminantes em etanol.

Analito	GRO	benzeno	tolueno	etilbenzeno	m,p-xilenos	o-xileno	naftaleno (GRO)
#variáveis espectrais	199	183	170	146	175	152	140
#variáveis latentes	12	12	10	11	13	13	12
% Variância bloco X	99,89	99,77	99,84	99,89	99,91	99,91	99,9
% Variância bloco y	99,92	99,91	98,81	99,44	99,81	99,75	99,83
Conjunto de calibração							
C máxima (mg kg ⁻¹)	7894,5	79,8	1492,77	79,8	279,0	115,8	22,4
RMSEC (mg kg ⁻¹)	48,3	0,5	24,0	1,3	2,5	1,2	0,2
RMSECV (mg kg ⁻¹)	192,1	2,8	110,9	4,0	13,7	4,6	0,9
Tendência (cal)	4,5.10 ⁻¹¹	-2,8.10 ⁻¹³	-2,5.10 ⁻¹²	8,9.10 ⁻¹⁴	-3,9.10 ⁻¹³	-5,7.10 ⁻¹³	-8,4.10 ⁻¹⁴
Tendência (CV)	-1,2.10 ¹	1,6.10 ⁻¹	-8,3	-1,1.10 ⁻¹	-1,2	-2,1.10 ⁻¹	-1,9.10 ⁻²
R ² cal	0,9992	0,9991	0,9881	0,9944	0,9981	0,9975	0,9983
R ² ajustado	0,9992	0,9991	0,9991	0,9943	0,9981	0,9975	0,9983
R ² CV	0,9879	0,9697	0,7497	0,9497	0,9446	0,9623	0,9656
Razão S/N mínimo	9,1	10,9	3,1	3,5	7,4	5,7	5,3
Razão S/N máximo	190,5	180,0	69,6	57,2	94,6	77,4	94,1
Seletividade média	1,1.10 ⁻²	9,6.10 ⁻³	2,9.10 ⁻³	5,5.10 ⁻³	4,5.10 ⁻³	4,7.10 ⁻³	5,5.10 ⁻³
Intercepto	7,3.10 ⁻²	1,1.10 ⁻¹	6,3.10 ⁻²	7,0.10 ⁻²	7,3.10 ⁻²	6,9.10 ⁻²	5,4.10 ⁻²
Inclinação	1,9.10 ⁻⁴	2,0.10 ⁻²	5,5.10 ⁻⁴	1,0.10 ⁻²	2,8.10 ⁻³	6,5.10 ⁻³	3,4.10 ⁻²
Sensibilidade	3,3.10 ⁻¹	3,1.10 ⁻¹	1,2.10 ⁻¹	1,7.10 ⁻¹	1,6.10 ⁻¹	1,6.10 ⁻¹	1,6.10 ⁻¹
Sens. analítica (mg kg ⁻¹) ⁻¹	4,0.10 ¹	3,3.10 ¹	9,9	1,1.10 ¹	1,8.10 ¹	1,4.10 ¹	1,9.10 ¹
Inv. sens. analit. (mg kg ⁻¹)	2,5.10 ⁻²	3,0.10 ⁻²	1,0.10 ⁻¹	8,9.10 ⁻²	5,6.10 ⁻²	6,9.10 ⁻²	5,3.10 ⁻²
LD (mg kg ⁻¹)	0,1	0,1	0,3	0,3	0,2	0,2	0,2
LQ (mg kg ⁻¹)	0,2	0,3	1,0	0,9	0,6	0,7	0,5
Erro relativo médio (%)	15,7	34,0	117,1	86,6	86,6	65,9	15,8
RPD	9,0	5,6	2,0	4,2	4,2	5,2	5,4
Conjunto de validação							
C máxima (mg kg ⁻¹)	7750,9	63,2	1234,3	75,6	258,4	108,5	18,2
RMSEP	148,3	2,2	65,7	2,9	15,6	5,3	0,9
Tendência (val)	-18,0	0,6	-1,1	-0,2	-0,9	0,8	0,3
t bias	0,75	1,62	0,10	0,34	0,36	0,92	2,42
R ² val	0,9946	0,9818	0,9431	0,9679	0,9510	0,9670	0,9636
r val	0,9973	0,9908	0,9711	0,9838	0,9752	0,9833	0,9816
Erro relativo médio (%)	50,0	77,6	404,8	139,7	499,4	302,4	72,5
RPD	13,3	7,0	4,0	5,6	3,9	4,7	4,9
Conjunto de replicatas							
Precisão (mg kg ⁻¹)	270,9	3,5	88,4	4,6	16,8	7,5	1,3
DPR mín. (%)	3,3	3,6	17,0	6,5	8,0	7,1	6,2
DPR máx. (%)	1253,9	912,9	1020,9	763,5	846,0	1324,4	2504,5
DRP médio > intervenção (%)	6,5	112,6	104,8	107,9	105,2	110,9	42,8
*DPR 20 replicatas C baixa	45,3% (438)	95,7% (4,9)	105,8% (11,1)	223,2% (2,9)	639,2% (9,0)	796,1% (3,6)	76,6% (1,1)
*DPR 20 replicatas C alta	7,9% (4913)	27,5% (6,1)	19,3% (1492,8)	39,9% (5,9)	26,9% (8,1)	48,4% (4,6)	16,0% (5,1)

* Os valores entre parênteses representam as concentrações dos analitos em mg kg⁻¹.

Continua

Tabela 47 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de soluções dos contaminantes em etanol. (Continuação).

Analito	DRO	naftaleno (DRO)	acenaftileno	fluoreno	fenantreno	antraceno
#variáveis espectrais	183	188	197	188	219	185
#variáveis latentes	12	13	12	6	13	13
% Variância bloco X	99,74	99,76	99,75	100	99,76	99,74
% Variância bloco y	99,96	99,97	99,95	99,95	99,93	99,94
Conjunto de calibração						
C máxima (mg kg ⁻¹)	10298,0	73,6	68,5	103,6	98,2	52,1
RMSEC (mg kg ⁻¹)	50,2	0,3	0,3	0,5	0,5	0,3
RMSECV (mg kg ⁻¹)	150,6	2,1	1,3	3,5	2,6	1,5
Tendência (cal)	4,3.10 ⁻¹¹	1,1.10 ⁻¹⁴	1,2.10 ⁻¹³	-9,4.10 ⁻⁵	1,4.10 ⁻¹³	-5,7.10 ⁻¹⁴
Tendência (CV)	-5,3	-1,5.10 ⁻¹	-6,8.10 ⁻²	-2,3.10 ⁻¹	8,1.10 ⁻²	2,8.10 ⁻²
R ² cal	0,9996	0,9997	0,9995	0,9995	0,9993	0,9994
R ² ajustado	0,9995	0,9997	0,9995	0,9995	0,9993	0,9993
R ² CV	0,9962	0,9872	0,9936	0,9748	0,9813	0,9803
Razão S/N mínimo	11,5	20,2	11,7	26,5	11,4	8,7
Razão S/N máximo	194,7	312,5	200,3	510,6	250,4	150,0
Seletividade média	2,4.10 ⁻²	1,8.10 ⁻²	1,9.10 ⁻²	1,4.10 ⁻³	1,4.10 ⁻²	1,6.10 ⁻²
Intercepto	1,5.10 ⁻¹	1,1.10 ⁻¹	1,3.10 ⁻¹	3,2.10 ⁻³	1,2.10 ⁻¹	1,1.10 ⁻¹
Inclinação	2,5.10 ⁻⁴	2,5.10 ⁻²	3,2.10 ⁻²	5,7.10 ⁻⁴	2,3.10 ⁻²	3,8.10 ⁻²
Sensibilidade	6,0.10 ⁻¹	4,7.10 ⁻¹	5,1.10 ⁻¹	1,3.10 ⁻²	4,3.10 ⁻¹	4,1.10 ⁻¹
Sens. analítica (mg kg ⁻¹) ⁻¹	4,3.10 ¹	7,4.10 ¹	4,4.10 ¹	1,0.10 ²	4,6.10 ¹	2,9.10 ¹
Inv. sens. analit. (mg kg ⁻¹)	2,3.10 ⁻²	1,3.10 ⁻²	2,3.10 ⁻²	9,7.10 ⁻³	2,2.10 ⁻²	3,4.10 ⁻²
LD (mg kg ⁻¹)	0,1	0,0	0,1	0,0	0,1	0,1
LQ (mg kg ⁻¹)	0,2	0,1	0,2	0,1	0,2	0,3
Erro relativo médio (%)	13,9	5,9	8,1	8,9	16,2	13,9
RPD	15,8	8,9	12,2	6,3	7,3	7,0
Conjunto de validação						
C máxima (mg kg ⁻¹)	9416,0	70,1	64,5	92,2	68,3	49,3
RMSEP	128,7	1,7	1,0	2,8	2,4	1,3
Tendência (val)	-5,4	-0,1	0,0	-0,9	0,3	0,0
t bias	0,26	0,48	0,09	2,07	0,79	0,03
R ² val	0,9975	0,9922	0,9967	0,9888	0,9796	0,9902
r val	0,9987	0,9961	0,9984	0,9944	0,9897	0,9951
Erro relativo médio (%)	36,9	43,4	22,7	51,0	46,3	29,3
RPD	19,1	11,3	17,5	9,0	7,0	9,4
Conjunto de replicatas						
Precisão (mg kg ⁻¹)	329,8	3,2	2,7	4,7	3,9	2,2
DPR mín. (%)	2,1	2,7	2,5	4,8	3,0	4,0
DPR máx. (%)	698,6	3059,2	1046,8	2496,7	1470,4	3778,4
DRP médio > intervenção (%)	15,3	27,1	14,7	17,1	7,5	13,8
*DPR 20 replicatas C baixa	61,2% (514)	126,6% (4,2)	35,8% (4,2)	48,6% (5,1)	71,4% (3,7)	51,3% (2,0)
*DPR 20 replicatas C alta	36,1% (1093)	55,9% (6,9)	36,0% (6,9)	29,9% (11,0)	47,3% (10,4)	32,9% (5,2)

* Os valores entre parênteses representam as concentrações dos analitos em mg kg⁻¹.

Continua

Tabela 47 – Parâmetros analíticos dos modelos de calibração desenvolvidos com o conjunto de soluções dos contaminantes em etanol. Continuação.

Analito	pireno	benzo(a)antraceno	criseno	benzo(b)fluoranteno	ORO
#variáveis espectrais	231	148	175	133	169
#variáveis latentes	7	10	12	11	11
% Variância bloco X	99,5	98,78	97,67	100	99,97
% Variância bloco y	96,13	99,94	99,93	99,41	98,25
Conjunto de calibração					
C máxima (mg kg ⁻¹)	50,3	49,7	20,1	24,6	5705,1
RMSEC (mg kg ⁻¹)	1,9	0,2	0,1	0,4	118,5
RMSECV (mg kg ⁻¹)	2,8	2,3	0,9	1,3	374,8
Tendência (cal)	-7,0.10 ⁻¹⁴	2,9.10 ⁻¹³	3,5.10 ⁻¹³	9,3.10 ⁻⁵	-4,1.10 ⁻¹¹
Tendência (CV)	-1,2.10 ⁻¹	4,6.10 ⁻²	1,4.10 ⁻²	-2,3.10 ⁻²	-5,4.10 ¹
R ² cal	0,9613	0,9994	0,9992	0,9941	0,9825
R ² ajustado	0,9608	0,9994	0,9992	0,9940	0,9923
R ² CV	0,9195	0,9335	0,9552	0,9536	0,8291
Razão S/N mínimo	1,5	12,0	11,3	4,7	0,9
Razão S/N máximo	34,0	216,0	194,7	64,9	64,1
Seletividade média	2,0.10 ⁻²	1,4.10 ⁻¹	1,2.10 ⁻¹	2,0.10 ⁻⁴	2,6.10 ⁻³
Intercepto	2,9.10 ⁻¹	1,1.10 ⁻³	2,7.10 ⁻¹	7,6.10 ⁻⁴	1,2.10 ⁻²
Inclinação	1,9.10 ⁻²	3,8.10 ⁻⁴	1,9.10 ⁻¹	3,6.10 ⁻⁴	4,0.10 ⁻⁵
Sensibilidade	6,3.10 ⁻¹	3,4.10 ⁻³	8,5.10 ⁻¹	2,1.10 ⁻³	5,9.10 ⁻²
Sens. analítica (mg kg ⁻¹) ⁻¹	9,1	3,7.10 ¹	4,0.10 ¹	1,4.10 ¹	1,5.10 ¹
Inv. sens. analit. (mg kg ⁻¹)	1,1.10 ⁻¹	2,7.10 ⁻²	2,5.10 ⁻²	7,2.10 ⁻²	6,5.10 ⁻²
LD (mg kg ⁻¹)	0,4	0,1	0,1	0,2	0,2
LQ (mg kg ⁻¹)	1,1	0,3	0,3	0,7	0,6
Erro relativo médio (%)	82,5	16,3	16,3	46,6	107,4
RPD	3,5	3,9	4,7	4,6	3,9
Conjunto de validação					
C máxima (mg kg ⁻¹)	42,4	39,1	19,5	24,2	4573,2
RMSEP	3,7	1,4	0,7	0,8	207,7
Tendência (val)	-0,2	0,1	0,0	0,3	-13,4
t bias	0,40	0,64	0,11	2,12	0,46
R ² val	0,8823	0,9791	0,9773	0,9801	0,9517
r val	0,9393	0,9895	0,9886	0,9900	0,9755
Erro relativo médio (%)	115,0	131,7	116,1	129,8	42,5
RPD	2,9	6,9	6,6	6,4	5,6
Conjunto de replicatas					
Precisão (mg kg ⁻¹)	2,1	4,0	1,8	2,5	683,5
DPR mín. (%)	5,9	9,4	8,3	14,0	9,2
DPR máx. (%)	873,5	612,9	784,7	1441,3	840,8
DRP médio > intervenção (%)	18,1	49,7	40,1	26,6	23,4
*DPR 20 replicatas C baixa	279,7% (1,0)	240,7% (1,0)	427,2% (0,5)	669,2% (0,8)	38,2% (693)
*DPR 20 replicatas C alta	41,0% (4,5)	141,9% (2,8)	81,9% (1,2)	827,4% (1,3)	13,6% (1085)

* Os valores entre parênteses representam as concentrações dos analitos em mg kg⁻¹.

A exatidão dos modelos foi avaliada através dos parâmetros RMSE, erro relativo médio, coeficiente de correlação do conjunto de validação (r_{val}) e diferença percentual relativa (RPD). Este último parâmetro é um indicativo útil para avaliar a exatidão quando se trata de diferentes intervalos de medição, pois o RPD mostra a relação entre os erros de predição e a distribuição dos valores de concentração, ou seja, o desvio padrão dos valores medidos do conjunto amostral. A diferença percentual relativa (RPD) foi obtida através da razão entre o desvio padrão das concentrações do conjunto de calibração ou validação e o valor de RMSECV ou RMSEP obtido pelo modelo. Valores de RPD acima de 2,5 mostram que o modelo apresenta exatidão aceitável em relação ao intervalo de medição e valores acima de 10 são considerados excelentes para métodos alternativos.¹⁸¹ Desta maneira, todos os modelos apresentaram valores de RPD aceitáveis ou excelentes, exceto o modelo de predição de tolueno em etanol (RPD = 2,0). A média dos valores de RPD para o conjunto de solos fortificados, soluções em hexano e soluções em etanol foram respectivamente 4,0, 5,3 e 6,3. Os menores valores de RPD para o conjunto de solos fortificados foram resultados da heterogeneidade das amostras e do maior ruído nos espectros NIR. A diferença entre os valores obtidos nos conjuntos de soluções em hexano e em etanol mostra que o uso de um solvente com sinais analíticos mais distintos dos analitos proporciona uma melhor exatidão na predição da concentração dos analitos. Pelos mesmos motivos, os maiores valores de RMSEP e erros relativos médios referem-se ao conjunto de solos fortificados e os menores ao conjunto de soluções em etanol. Exceto o modelo de predição de ORO em solo, todos os modelos apresentaram r_{val} acima de 0,9277. Portanto, todos os modelos apresentaram exatidão aceitável. Os gráficos de correlação entre os valores medidos e os valores preditos para cada analito pode ser observado no Anexo C.

Atualmente, o método de regressão PLS é o método mais utilizado e bem sucedido na predição de propriedades que apresentam boa linearidade com o sinal instrumental.⁶² Assim, a linearidade dos modelos foi avaliada através dos parâmetros coeficiente de determinação (R^2) e tendência. Para os conjuntos de calibração, as amostras de solo fortificado apresentaram menores valores de tendência, por outro lado, para o conjunto de validação resultaram em maiores tendências, significando que a predição de amostras externas apresenta uma tendência sistemática. A tendência positiva pode proporcionar um número maior de falsos positivos para contaminação de solos, não obstante, torna-se mais preocupante resultados falsos negativos devido ao risco ambiental proporcionado por estes grupos de analitos. A significância das tendências obtidas pela predição do conjunto de validação foi verificada através de um teste t de Student. Se o valor calculado de t_{bias} for superior ao valor de t_{tabelado} (bicaudal) para 39 graus de liberdade a um nível de confiança de 95% (2,02), houve tendência significativa na predição

de amostras externas. Portanto, os modelos que apresentaram tendência significativa foram: (i) em solo – acenaftileno, fluoreno e antraceno; (ii) em hexano – pireno e ORO; e (iii) em etanol – naftaleno (GRO), fluoreno e benzo(b)fluoranteno. Exceto as predições de ORO em hexano e fluoreno em etanol, os modelos apresentaram tendências positivas. A tendência na predição pode ser verificada no gráfico de y preditos versus os respectivos resíduos (Figura 14). Neste gráfico, observa-se também a proporcionalidade dos resíduos em relação à concentração. Os resíduos são relativamente maiores em concentrações menores, mas este fato será melhor discutido adiante.

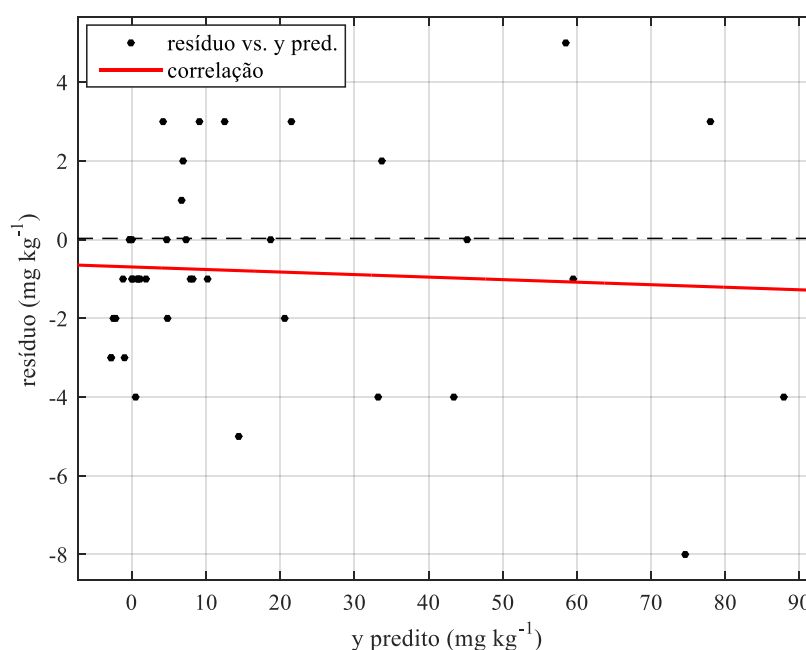


Figura 14 – Tendência significativa na predição de fluoreno em etanol em amostras externas.

Através da teoria de sinal analítico líquido (NAS), determinou-se os parâmetros analíticos razão S/N; seletividade; $R^2_{ajustado}$; intercepto; inclinação; sensibilidade e limites de detecção e quantificação. Utilizando o vetor $\|nas\|$, construiu-se um modelo de regressão pseudo-univariado para determinar a inclinação, o intercepto e $R^2_{ajustado}$. Estes dois últimos parâmetros são bons indicativos da linearidade do modelo, pois um modelo bem ajustado deve apresentar intercepto próximo a zero e $R^2_{ajustado}$ próximo a um. Dado que o valor de $R^2_{ajustado}$ foi calculado para um modelo de regressão pseudo-univariado por mínimos quadrados com dois coeficientes (duas variáveis explicativas), o valor de $R^2_{ajustado}$ foi próximo ao valor de R^2_{cal} .

Para modelos univariados, a inclinação equivale à sensibilidade do método determinativo, enquanto para os modelos multivariados a sensibilidade é inversamente proporcional à norma do vetor de regressão, \mathbf{b} , cujos valores são determinados para cada

variável instrumental, x . Como a sensibilidade é calculada de formas diferentes para modelos univariados e multivariados, a sensibilidade analítica (γ) e o inverso da sensibilidade analítica (γ^{-1}) são parâmetros mais adequados para fins de comparação e aplicação dos métodos. Assim como a exatidão, os menores valores de sensibilidades foram proporcionados pelo conjunto de solos fortificados e os maiores pelo conjunto de soluções etílicas. Os valores médios de γ^{-1} para as frações de TPH para os conjuntos de solos fortificados, soluções em hexano e soluções em etanol foram, respectivamente, $0,12 \text{ (mg kg}^{-1}\text{)}^{-1}$, $0,05 \text{ (mg kg}^{-1}\text{)}^{-1}$ e $0,04 \text{ (mg kg}^{-1}\text{)}^{-1}$. Estes valores de γ^{-1} mostram que o método de determinação das frações de TPH em soluções etílicas é 3 vezes mais sensível que a determinação por análise direta de solos.

Diferentemente da seletividade em métodos cromatográficos, a qual indica o grau de separação de picos cromatográficos, a seletividade em métodos NIR mostra o quanto da resposta instrumental refere-se ao analito. Portanto, esta figura de mérito deve ser determinada pelo conceito NAS. Uma vez que cada amostra possui um valor nas, a seletividade do modelo foi considerada a média das seletividades das amostras. Os valores médios de seletividade para as frações de TPH para os conjuntos de solos fortificados, soluções em hexano e soluções em etanol foram, respectivamente, 1,5%, 0,9% e 1,2%. Notou-se que a similaridade espectral entre o solvente e os analitos refletiu na seletividade do modelo, pois o uso de hexano como solvente proporcionou a mais baixa seletividade, enquanto os solos fortificados proporcionaram a maior.

Os valores das razões S/N e os limites de detecção e quantificação foram calculados utilizando o desvio padrão de 12 replicatas do branco (solos ou solventes não fortificados). Assim como a seletividade, a razão S/N foi determinada para cada amostra e apenas os valores mínimos e máximos foram apresentados nas Tabelas 45 a 47. Estes valores foram influenciados por três fatores: (i) precisão do método – quanto maior a precisão, menor é o desvio padrão do branco e, por consequência, menor é o ruído; (ii) seletividade – quanto maior a seletividade, maior é o sinal referente ao analito; e (iii) sensibilidade – quanto maior a sensibilidade, maior é a razão S/N máxima. Os conjuntos de solos fortificados, soluções em hexano e soluções em etanol apresentaram os valores médios de razão S/N mínimo para as frações de TPH, respectivamente, 0,70, 5,79 e 7,17 e de razão S/N máximo, respectivamente, 42,05, 140,98 e 192,61. Uma vez que o conjunto de solos fortificados apresentou maior seletividade, pode-se dizer que a sensibilidade e a precisão foram os fatores predominantes para maior razão S/N. Como os valores de LD e LQ são diretamente proporcionais ao ruído e inversamente proporcionais à sensibilidade, os modelos que apresentaram os limites mais baixos para detecção e quantificação foram aqueles desenvolvidos com as soluções em etanol, seguido pelos desenvolvidos com as soluções em hexano e, por fim, com os solos fortificados.

A precisão dos modelos foi determinada pela análise dos conjuntos de replicatas e avaliada pelos parâmetros analíticos precisão (equação 13); DPR de 9 replicatas de 30 amostras em três níveis de concentração (inferior, intermediário e superior); e DPR de 20 replicatas de 2 amostras, no nível de concentração próximo ao limite de intervenção para TPH e no nível superior de concentração. Quando observados os valores de precisão e RMSEP dos modelos, nota-se que os valores de precisão são próximos ou até superiores aos valores de RMSEP. Esta é uma evidência de que os modelos são pouco precisos considerando toda faixa de trabalho, porém, apesar da baixa precisão, a exatidão dos modelos foi aceitável. Na Figura 15, estão plotados os valores de DPR das 30 amostras analisadas em nonoplicata, para cada conjunto amostral, versus a concentração de GRO. O gráfico da Figura 14 foi ampliada, omitindo os maiores valores DPR, para permitir uma melhor observação da variação de DPR pela concentração. Ao avaliar a precisão em relação à concentração, tornou-se claro que os elevados valores de precisão estão relacionados às baixas concentrações. Para os conjuntos de solos fortificados e de soluções em hexano, o DPR tornou-se próximo a 20% acima de 1400 mg kg⁻¹ e para o conjunto de soluções em etanol acima de 1000 mg kg⁻¹. Esta relação entre o DPR e a concentração manteve-se para as demais frações de TPH, enquanto para os demais analitos variou conforme a faixa de trabalho.

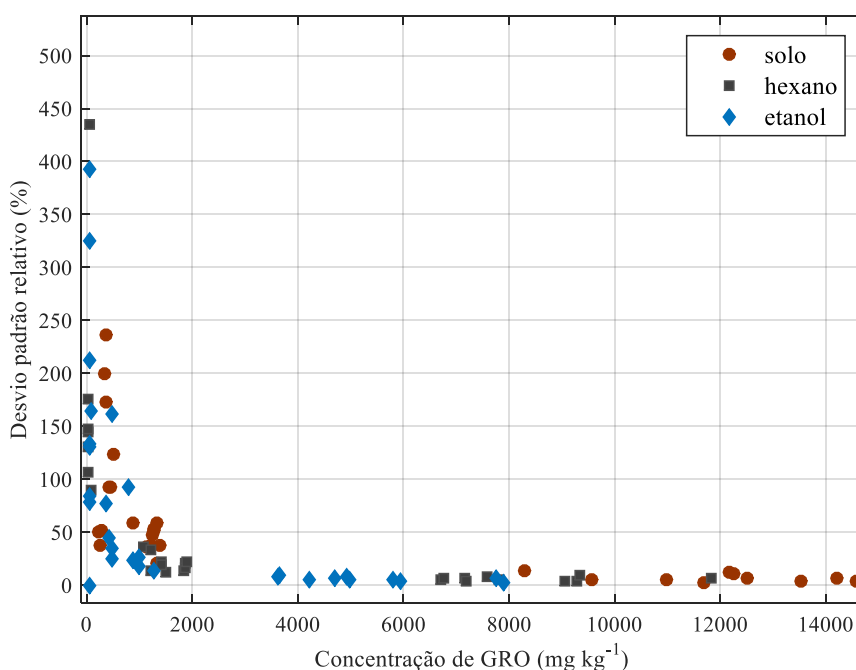


Figura 15 – Relação entre desvio padrão relativo e concentração de GRO para os três conjuntos amostrais.

A Figura 14 e a Figura 15 mostraram que as baixas concentrações tiveram grande peso na exatidão e na precisão dos modelos. Isto é reflexo da baixa sensibilidade e, principalmente, da baixa seletividade dos modelos, dado que nenhum dos modelos apresentou seletividade maior que 2,4%, exceto pelos modelos de predição de acenaftileno em hexano (33,3%), benzo(a)antraceno em etanol (13,7%) e criseno em etanol (12,1%). Os maiores valores de seletividade destes modelos pode ter sido em decorrência da seleção de variáveis pelo GA, o qual ocasionou um sobreajuste para o conjunto de calibração, levando a maiores valores de seletividade sem gerar melhoras proporcionais para o conjunto de validação.

Visto a influência das baixas concentrações nos valores de DPR, calculou-se o DPR médio das amostras com concentração acima do limite de intervenção para TPH (1000 mg kg⁻¹). Para as frações de TPH, os valores médios resultantes dos conjuntos de solos fortificados, soluções em hexano e soluções em etanol foram, respectivamente, 27,7%, 21,2% e 10,9%. Portanto, apenas os modelos gerados com as soluções em etanol apresentaram DPR abaixo de 20%, conforme recomenda-se a EPA 8000.¹⁷³

Finalmente, a precisão foi avaliada pelo o conjunto de amostras analisadas em 20 replicatas para verificação de erros sistemáticos. Para isto, foram plotados gráficos de cartas de controle, onde os limites superior e inferior foram 3 vezes o desvio padrão das medidas. Na Figura 16, estão representadas as cartas de controle das determinações de GRO em solo nas concentrações de 1119 mg kg⁻¹ e 11.681 mg kg⁻¹. Apesar da tendência positiva nas observações, nenhuma medida ultrapassou os limites das cartas de controle e, como pode ser observado, a ausência de erros sistemáticos pôde ser comprovada pela aleatoriedade das medidas. Foram plotadas cartas de controle para todos os analitos e verificou-se a ausência de erros sistemáticos para todos os casos. Embora alguns modelos tenham gerado valores de DPR demasiadamente elevados para o conjunto de 20 replicatas, também não apresentaram erros sistemáticos, como pode ser observado na Figura 17. Nesta Figura, estão representadas as cartas de controle da determinação de criseno em solo e o-xileno em etanol. A determinação destes analitos geraram DPR iguais a, respectivamente, 1307,8% e 791,6%; porém, sem haver medidas anômalas ou sistemáticas.

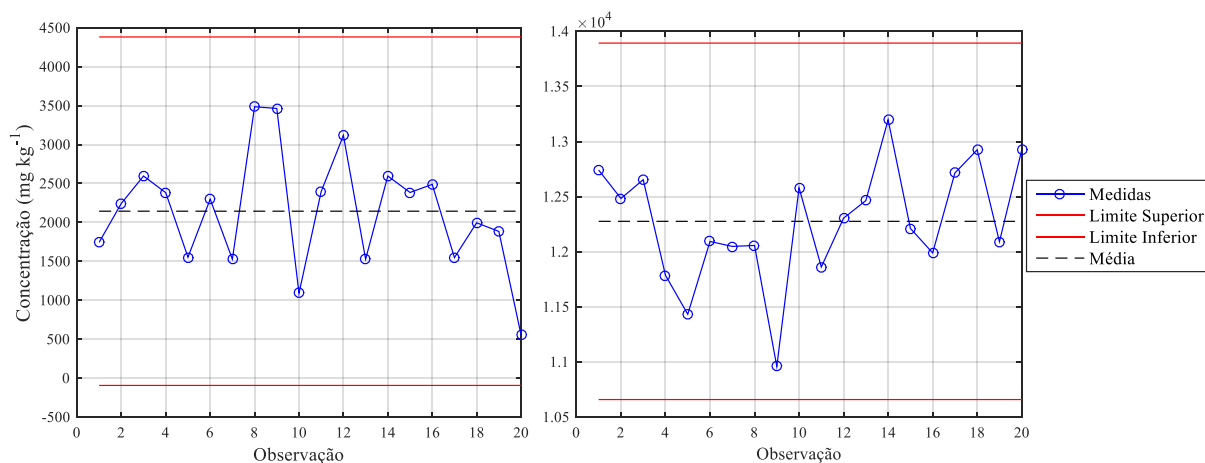


Figura 16 – Cartas de controle das medidas de GRO em solo fortificado (dois níveis de concentração).

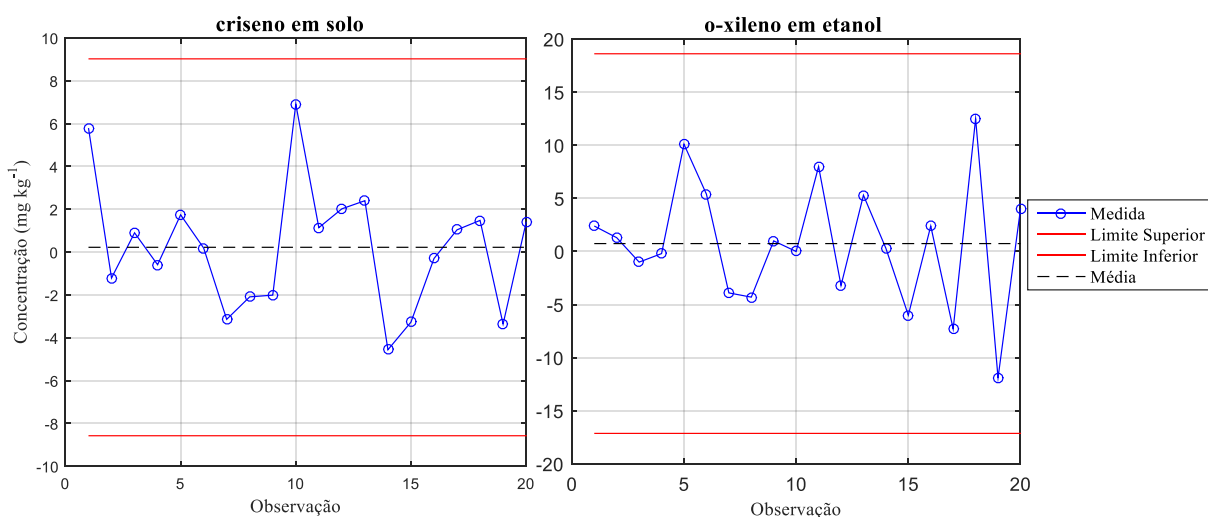


Figura 17 – Cartas de controle das medidas de crisenol em solo e o-xileno em etanol.

O gráfico da Figura 18, onde estão os valores de DPR das 20 replicatas para os três conjuntos amostrais e de todos analitos, mostra que o DPR aproxima-se de 20% para a maioria dos analitos a partir de 45 mg kg⁻¹; e os modelos de soluções em hexano apresentaram menores valores de DPR em relação aos outros, seguido pelos modelos de soluções em etanol.

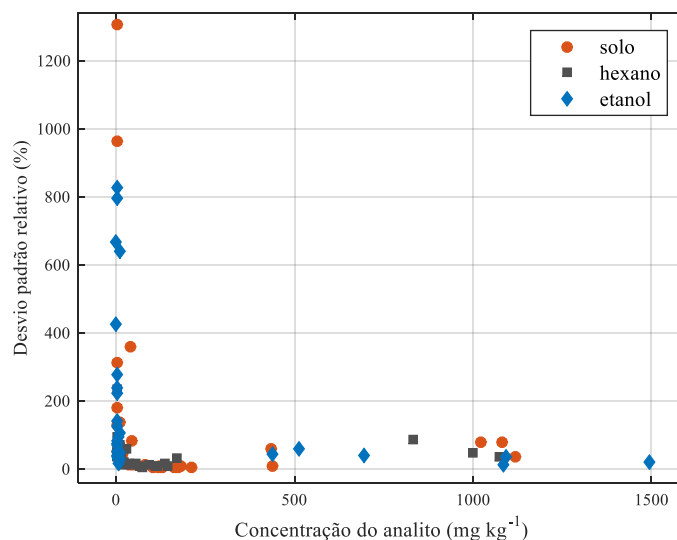


Figura 18 – Relação entre desvio padrão relativo (20 replicatas) e a concentração de todos os analitos nos três conjuntos de dados.

O conjunto de 270 espectros (30 amostras analisadas em nonoplicata) também foi utilizado para avaliar os riscos de os modelos de calibração gerarem resultados falsos positivos ou negativos para contaminação, conforme os limites de concentração estabelecidos pela CETESB para cada analito. Na Tabela 48, estão apresentados os limites de prevenção e intervenção (agrícola, residencial e industrial) e a porcentagem de amostras que indicaram valores de FP e FN. Considerando-se que resultados FN colocam mais em risco o meio ambiente e, normalmente, não levam a uma verificação do resultado em métodos já bem estabelecidos, os modelos que geraram maior porcentagem de FN foram julgados como menos adequados para aplicação. Assim sendo, a determinação de GRO foi mais adequada pelos modelos de soluções em etanol, enquanto, as frações DRO e ORO foram determinadas de forma mais apropriada pelos modelos de soluções em hexano.

Ao final da Tabela 48, estão as médias dos resultados FP e FN por conjunto amostral. Ponderando-se todos analitos, o conjunto de soluções em hexano gerou o menor número de resultados FP (9%) e FN (6%), embora a exatidão, precisão, linearidade e sensibilidade tenham sido inferiores em relação aos modelos empregando soluções em etanol.

Tabela 48 – Porcentagens de FP e FN gerados pelos modelos de calibração na análise do conjunto de replicatas.

GRO		Solos fortificados		Soluções em hexano		Soluções em etanol	
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	1000	10	7	0	4	7	0
II	1000	10	7	0	4	7	0
III	1000	10	7	0	4	7	0
IV	1000	10	7	0	4	7	0
benzeno							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	0,002	0	21	0	24	16	20
II	0,02	0	21	0	24	17	20
III	0,08	0	21	0	20	17	21
IV	0,2	0	21	0	19	10	17
tolueno							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	0,9	0	33	5	24	0	29
II	5,6	0	34	18	21	7	18
III	14	7	25	18	20	15	11
IV	80	19	10	12	0	14	9
etilbenzeno							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	0,03	0	19	3	16	0	24
II	0,2	0	20	8	13	3	20
III	0,6	3	17	16	2	6	15
IV	1,4	4	17	13	3	17	10
m,p-xilenos							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	0,03	0	21	4	19	0	30
II	12	15	14	16	2	25	5
III	3,2	13	16	16	2	15	13
IV	19	21	6	18	2	21	4
o-xileno							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	0,03	0	18	4	21	0	28
II	12	23	6	15	2	16	5
III	3,2	12	10	15	1	14	14
IV	19	24	1	18	1	9	1
naftaleno (GRO)							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	0,7	10	14	13	2	16	7
II	1,1	17	11	11	3	17	9
III	1,8	17	5	10	3	17	3
IV	5,9	16	1	12	0	10	0

Continua.

Tabela 48 – Percentagens de FP e FN gerados pelos modelos de calibração na análise do conjunto de replicatas. (Continuação).

DRO		Solos fortificados		Soluções em hexano		Soluções em etanol	
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	1000	20	3	17	2	6	5
II	1000	20	3	17	2	6	5
III	1000	20	3	17	2	6	5
IV	1000	20	3	17	2	6	5
naftaleno (DRO)							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	0,7	0	16	12	2	13	4
II	1,1	0	16	14	2	15	3
III	1,8	0	18	12	2	13	4
IV	5,9	16	6	8	8	7	3
fenantreno							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	3,6	9	14	12	5	16	5
II	15	29	0	14	0	3	0
III	40	7	2	0	0	1	0
IV	95	6	3	4	4	0	3
antraceno							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	0,3	0	15	14	6	6	1
II	2300	0	0	0	0	0	0
III	4600	0	0	0	0	0	0
IV	10000	0	0	0	0	0	0
benzo(a)antraceno							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	0,2	0	17	18	5	8	13
II	1,6	13	14	21	4	17	5
III	7	27	0	18	0	9	0
IV	22	4	0	3	1	6	2
criseno							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	1,6	26	6	14	4	26	1
II	95	0	0	0	0	0	0
III	600	0	0	0	0	0	0
IV	1600	0	0	0	0	0	0
benzo(b)fluoranteno							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	0,7	8	14	14	5	13	13
II	2	31	2	20	2	11	1
III	7,2	18	1	9	2	4	3
IV	25	6	3	0	11	0	0
ORO							
Limite	C (mg kg ⁻¹)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
I	1000	36	1	10	0	27	17
II	1000	36	1	10	0	27	17
III	1000	36	1	10	0	27	17
IV	1000	36	1	10	0	27	17
		Solos fortificados		Extração com hexano		Extração com etanol	
		FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
Média		11	10	9	6	10	8

Limite I: Prevenção

Limite II: Intervenção região agrícola

Limite III: Intervenção região residencial

Limite IV: Intervenção região industrial

5.3.3.4. Validação dos métodos NIR – conjunto de validação (II)

A validação dos modelos quimiométricos realizada com as soluções dos contaminantes em solventes foi crucial para conhecimento de algumas limitações dos modelos, como sensibilidade, seletividade, precisão e exatidão. Contudo, a validação dos modelos utilizando estas soluções, cujo preparo foi realizado como as soluções de calibração, superestima a capacidade preditiva dos modelos, uma vez que interferentes oriundos da matriz não estão presentes nas soluções de validação. Desta forma, os modelos PLS foram validados através do conjunto de validação (II), cujas amostras foram preparadas posteriormente e passaram pelo processo de extração. Para cada método desenvolvido (sem preparo de amostra; extração por hexano; e extração por etanol), foram preparadas nove amostras: três abaixo do limite de intervenção; três próximas ao limite; e três acima do limite. As amostras foram preparadas e armazenadas à -4 °C por dois dias para melhor homogeneização dos contaminantes no solo. O procedimento de extração foi realizado como definido para cada solvente e as análises espectroscópicas foram realizadas em triplicata. Os resultados da validação dos métodos NIR estão apresentados na Tabela 49.

De acordo com os valores de recuperação, os três métodos apresentaram baixa eficiência, dado que os métodos sem extração, extração com hexano e extração com etanol resultaram em, respectivamente, 37%, 11% e 56% de amostras com valores de recuperação dentro do limite de 70% a 130%. Quando consideramos apenas as amostras com concentração acima de 1000 mg kg⁻¹, os percentuais de amostras dentro do limite de recuperação passam a ser 48%, 5% e 75%. Os valores inconsistentes de recuperação para as amostras de baixa concentração foram consequências da baixa sensibilidade, seletividade e precisão dos métodos desenvolvidos.

Tabela 49 – Resultados de validação dos métodos de determinação de TPH com e sem preparo de amostra.

Amostra	Sem extração				Extração com hexano				Extração com Etanol				
	Concentração TPH (mg kg ⁻¹)			Recup. (%)	Concentração TPH (mg kg ⁻¹)			Recup. (%)	Concentração TPH (mg kg ⁻¹)			Recup. (%)	
	medido	predito	s		medido	predito	s		medido	predito	s		
GI-1	420	54	1023	13	422	94	235	22	324	850	143	262	
GI-2	327	-214	469	-66	312	454	407	145	422	1290	291	306	
GI-3	426	156	121	37	411	327	118	80	323	1142	76	353	
GII-1	2188	2502	465	114	2112	39	272	2	1812	1861	178	103	
GII-2	2124	2176	664	102	2080	129	347	6	1904	2077	290	109	
GII-3	2375	2439	222	103	2288	132	64	6	2192	2174	155	99	
GIII-1	7185	7040	341	98	6955	711	373	10	5606	4668	234	83	
GIII-2	8834	7987	494	90	8432	691	228	8	6710	4515	94	67	
GIII-3	8945	8126	724	91	8630	85028	234	985	7323	6087	116	83	
DI-1	447	148	1049	33	513	228	1316	44	377	1041	106	276	
DI-2	442	55	919	13	513	401	531	78	499	5326	273	1066	
DI-3	499	1316	299	264	569	-1071	1511	-188	376	1563	8	416	
DII-1	1911	4091	617	214	2217	989	228	45	2101	1895	164	90	
DII-2	2079	1410	1148	68	2330	361	549	15	2228	2475	194	111	
DII-3	1930	3319	214	172	2160	-426	1843	-20	2555	2009	98	79	
DIII-1	11630	15466	688	133	13422	7403	469	55	6580	8456	213	128	
DIII-2	11546	14720	651	127	13256	6540	889	49	7946	9191	196	116	
DIII-3	10224	12159	1313	119	11814	5376	835	46	8424	7958	343	94	
OI-1	1026	1692	603	165	1474	907	632	62	1474	1041	106	71	
OI-2	1164	714	314	61	1107	1406	444	127	1107	5326	273	481	
OI-3	1384	1259	144	91	776	1487	79	192	776	1563	8	201	
OII-1	4239	2507	602	59	3908	172	504	4	3908	1895	164	48	
OII-2	3010	1306	758	43	3718	656	223	18	3718	2475	194	67	
OII-3	2516	1684	925	67	3223	822	645	25	3223	2009	98	62	
OIII-1	10813	5338	1077	49	7580	764	102	10	7580	8456	213	112	
OIII-2	10897	7936	651	73	8704	-230	528	-3	8704	9191	196	106	
OIII-3	10858	4004	354	37	7197	944	569	13	7197	7958	343	111	
Amostras dentro do limite (70-130%)				37%					11%				
Amostras dentro do limite (70-130%) acima de 1000 mg kg ⁻¹				48%					5%				

s: desvio padrão de triplicatas

Na Figura 19, a qual exibe a relação entre a recuperação e a concentração de TPH para o método com extração por etanol, observa-se valores de recuperação dentro do limite especificado apenas acima de 1500 mg kg⁻¹. O método NIR com extração por etanol, além de apresentar melhores valores de recuperação, gerou resultados com menores valores de desvio padrão (s), sendo mais preciso, enquanto o método sem extração apresentou maiores valores de desvio padrão devido à heterogeneidade das amostra de solo (Tabela 49).

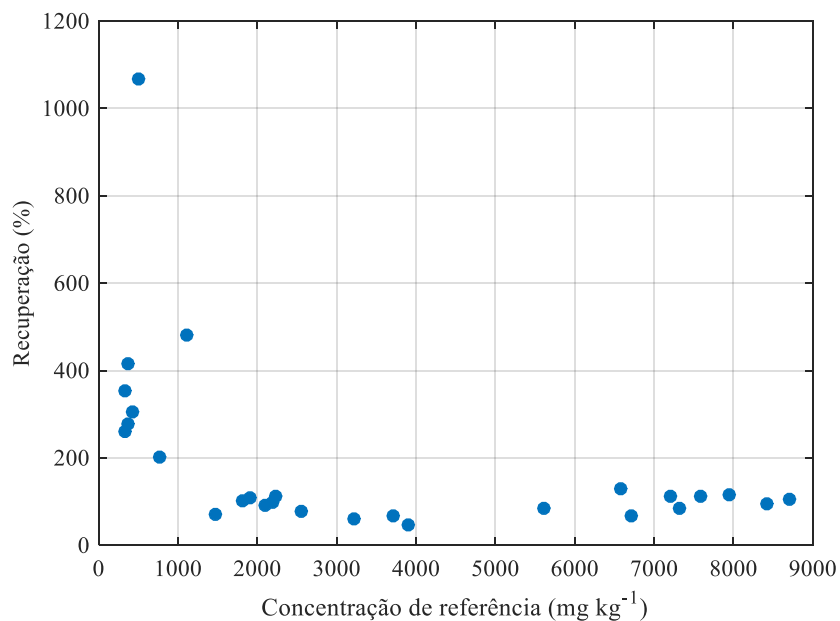


Figura 19 – Relação entre valores de recuperação e concentração de TPH para o método NIR com extração por etanol.

A principal interferência na determinação de TPH pelo método com extração por etanol foi a presença de água no solo. Apesar da etapa de purificação com sulfato de sódio e sílica gel antes das análises NIR, o etanol possui forte interação com a água e, portanto, resíduos de água geraram sinais espectrais próximos a 5000 cm^{-1} e 7000 cm^{-1} , como pode ser observado na Figura 20.

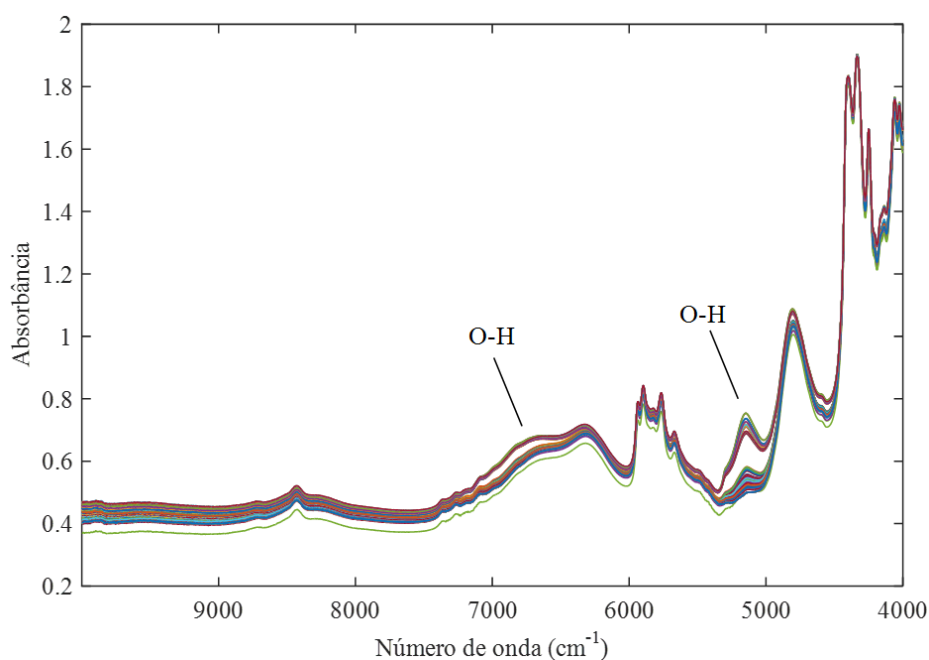


Figura 20 – Espectro das soluções dos contaminantes em etanol e das amostras de validação extraídas por etanol

Embora o hexano proporcione uma melhor extração de hidrocarbonetos em relação ao etanol, os resultados de validação mostraram uma baixa recuperação mesmo para altas concentrações. Um fator que pode ter contribuído para estes resultados insatisfatórios foi a mudança de marca do solvente entre a construção do modelo e a validação do método. As soluções utilizadas no desenvolvimento dos modelos foram preparadas com n-hexano, da marca Merk, com pureza maior ou igual a 99,0%, enquanto as amostras de validação foram extraídas com n-hexano, da marca Êxodo Científica, com pureza maior ou igual a 98,5%. Como as principais impurezas destes solventes são isômeros de cadeia do n-hexano, diferenças espectrais na região NIR são visivelmente notáveis, principalmente nas bandas de combinação (Figura 21).

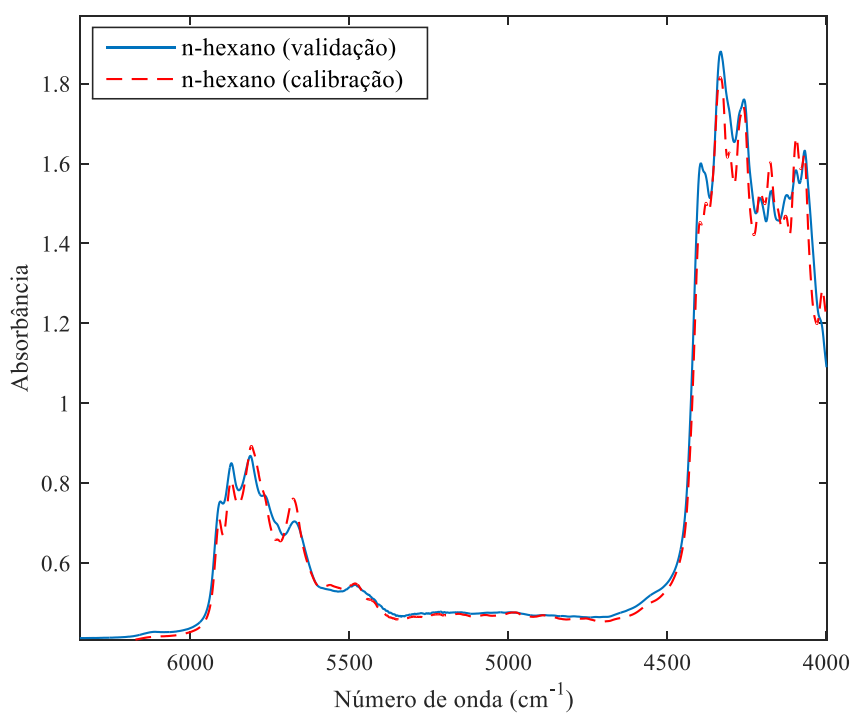


Figura 21 – Espectros dos solventes n-hexano utilizados no desenvolvimento dos modelos e na extração das amostras de validação.

A fim de corrigir as diferenças espectrais entre as soluções em n-hexano de diferentes purezas, realizou-se a subtração entre os espectros puros de ambos solventes e a diferença foi somada às amostras extraídas com o novo solvente. Apesar dos espectros das amostras de validação ficarem similares aos espectros das soluções de calibração, a presença de 0,5% (v/v) de impurezas a mais pode representar cerca de 5000 mg kg⁻¹ de interferentes nas amostras de validação, levando a valores de recuperação incongruentes. Portanto, apesar da correção

espectral, tornou-se clara a importância da utilização de solventes de mesma qualidade para a faixa de concentrações dos analitos estudada neste trabalho.

Como já discutido na seção 5.3.3.2., a seleção de variáveis pelo método GA pode sobreajustar os modelos de calibração, tornando-os menos adequados para conjuntos de amostras externas. Porém, os modelos desenvolvidos neste trabalho, após o uso de GA, geraram melhores resultados para o conjunto de validação (I).

Nesta seção, o conjunto amostral foi composto de amostras com maior número de interferentes devido ao processo de extração. Desta forma, a seleção de variáveis também foi avaliada para este conjunto de amostras através dos erros relativos dos valores preditos. Como pode ser notado na Tabela 50, os valores de erro relativo médio mostraram que a aplicação de seleção de variáveis foi uma ferramenta eficaz na atenuação da presença de interferentes, tornando os modelos mais seletivos, exceto para a predição de ORO sem extração por solvente.

Tabela 50 – Médias dos erros relativos resultantes da validação dos métodos desenvolvidos.

Fração TPH	Erro Relativo Médio (%)						
	Sem Extração			Extração com hexano		Extração com etanol	
	Espectro completo	Exclusão manual	GA	Espectro completo	GA	Espectro completo	GA
GRO	82	52	40	187	161	336	72
DRO	188	117	68	128	98	202	174
ORO	43	49	205	101	69	68	69

5.3.3.5. Predição das amostras de monitoramento ambiental

A terceira e última etapa de validação dos modelos de calibração foi a predição das frações de TPH nas amostras de solo coletadas para monitoramento ambiental. Das três amostras de solo (ARI, ARII e ARIII), três alíquotas foram retiradas para extração e as soluções resultantes foram analisadas em triplicata no espectrômetro. Na Tabela 51, estão dispostos os valores de referência, determinados pelo método cromatográfico ultrarrápido, os valores preditos pelos modelos e, ao fim, os valores de RMSEP para cada método.

Tabela 51 – Valores de referência e preditos pelos modelos das frações de TPH para as amostras de solo coletado para monitoramento ambiental.

	Concentração média (mg/kg)							
	Referência		Sem extração		Extração com hexano		Extração com etanol	
	medido	s	predito	s	predito	s	predito	s
GRO								
ARI	< LD	-	-8471	2747	-3093	275	-1766	12
ARII	56	-	6698	1327	-2197	325	-233	141
ARIII	66	24	6738	491	-3371	250	554	265
DRO								
ARI	2348	2858	90174	6572	2098	635	471	111
ARII	1813	378	36990	5634	2219	843	1524	93
ARIII	1823	179	84499	1234	3723	463	1703	61
ORO								
ARI	2430	1246	-289	685	3751	252	3993	413
ARII	685	195	5663	441	2902	440	2155	331
ARIII	4429	372	7817	172	6265	437	2527	342
RMSEP			42151		2117		1303	

Os resultados apresentados pela metodologia sem extração foram, em grande maioria, demasiadamente elevados e incoerentes com os valores de referência. Na seção 5.1.1. (análises físico-químicas e composicionais dos solos), a PCA dos espectros dos solos mostrou que os solos utilizados na construção dos modelos e os solos de monitoramento apresentaram espectros muito distintos, separando os grupos de amostras em apenas uma PC e a amostra ARIII das demais na segunda PC. Portanto, os erros de predição podem ser justificados pela grande diferença espectral entre as amostras, portanto, a metodologia sem preparo de amostra não foi adequada através de amostras de solos fortificados e distintos das amostras “reais”.

Quanto à metodologia de determinação de TPH após extração, não houve FP para contaminação por gasolina e apenas um FP para contaminação por óleo lubrificante (ARII) para ambos os solventes. O método determinativo após extração com etanol apresentou uma determinação FN para DRO na amostra ARI, por outro lado, o valor de referência desta amostra possui um desvio padrão muito elevado, podendo levar à dupla interpretação.

Na determinação dos valores de referência, o etanol foi utilizado como solvente extrator para fração GRO, enquanto o hexano foi utilizado para outras frações. Observou-se que o etanol gerou resultados com menores desvios padrões que o hexano. Além da melhor precisão, os valores de RMSEP, dispostos na Tabela 51, mostram que os resultados obtidos pelo método com extração por etanol foram os mais próximos aos valores de referência. Portanto, o uso de

etanol para extração de TPH proporcionou resultados mais precisos e exatos que o uso de hexano e, principalmente, que análise de amostras de solo sem preparo.

5.3.3.6. *Síntese do desenvolvimento e validação dos modelos de calibração*

A avaliação dos pré-processamentos na modelagem para predição de TPH com os espectros completos mostrou que cada conjunto amostral e cada analito requer um pré-processamento adequado. Os pré-processamentos de normalização, MSC e SNV, foram mais adequados para o conjunto amostral de solos fortificados; pré-processamentos de ortogonalização foram melhores para GRO em hexano (OSC) e ORO em etanol (EPO); centragem na média para DRO em hexano; e autoescalamamento foi o mais adequado para os demais modelos.

A seleção de variáveis foi executada pelo algoritmo genético (GA) em duas fases. A primeira execução do GA não reduziu significativamente os erros de validação apenas para os analitos GRO e tolueno em solos fortificados; e pireno e benzo(a)antraceno em soluções de etanol. Já a seleção de variáveis realizada por duas execuções sequenciais do GA reduziu significativamente os erros para todos os analitos, nos três conjuntos amostrais.

Todos modelos de calibração, desenvolvidos após o uso de GA, mostraram exatidão aceitável e boa linearidade, mas baixa sensibilidade, seletividade e precisão. Os valores de DPR tornaram-se abaixo de 20% apenas acima de 1400 mg kg^{-1} . Apesar da baixa precisão, a ausência de erros sistemáticos foi comprovada através de cartas de controle para todos os analitos. Considerando-se todos os parâmetros analíticos avaliados, os modelos gerados a partir de soluções em etanol apresentam os melhores resultados de calibração e validação.

A análise de amostras preparadas posteriormente àquelas utilizadas no desenvolvimento dos modelos revelou que apenas uma pequena porcentagem dos resultados apresentou recuperação entre 70% e 130%. Para concentrações acima de 1000 mg kg^{-1} , 75% das amostras validadas para método com extração por etanol apresentaram-se dentro dos limites de recuperação. Durante esta avaliação, notou-se a importância do uso de solventes de mesma qualidade e a influência da água nos espectros das soluções etílicas.

A predição das amostras “reais” mostrou que a análise direta das amostras solo gerou resultados incoerentes com os valores de referência devido à grande diferença entre as amostras de calibração e as amostras “reais”. Já os métodos com extração por solvente proporcionaram

resultados preditos próximos aos valores de referência. Nesta etapa, tornou-se evidente a melhor exatidão e precisão dos métodos NIR com extração por etanol que os demais.

5.4. Comparativo entre os métodos convencionais e desenvolvidos

As vantagens e limitações dos métodos convencionais e desenvolvidos foram avaliadas em três aspectos: (i) parâmetros analíticos dos modelos de regressão; (ii) frequência analítica e consumo de solvente; e (iii) custos dos equipamentos e operação.

5.4.1. Parâmetros analíticos

Na Tabela 52, estão presentes os parâmetros analíticos dos métodos convencionais (EPA) e dos métodos desenvolvidos para determinação das frações TPH. Dentre os métodos apresentados, os métodos UFGC apresentaram maior linearidade (R^2 ajustado) e menores erros de ajuste (RMSE) e DRP, exceto para a fração DRO. Os altos valores de RPD (acima de 69,6) foram consequência do grande intervalo de medição das curvas analíticas e dos baixos valores de RMSE. A baixa sensibilidade comparada aos demais métodos foi devida ao pequeno volume de amostra suportado pelas colunas UFGC, diferentemente da alta sensibilidade do método EPA para GRO causada pela injeção splitless de 500 mL de amostra gasosa.

Quanto aos métodos espectroscópicos, as análises NIR sem preparo de amostra apresentaram RMSE até 6 vezes maior que as análises NIR de soluções; e valores de RPD e sensibilidade analítica inferiores. A precisão (DPR) deste método foi próxima aos valores dos métodos NIR após extração.

Os modelos gerados com as soluções em etanol apresentaram valores de R^2 ajustado, r e RMSE próximos aos valores dos métodos UFGC, exceto para a fração ORO.

De forma geral, as curvas analíticas geradas pelos métodos UFGC apresentaram os melhores valores dos parâmetros analíticos, seguido pelos métodos NIR com soluções em etanol e os métodos GC (EPA).

Tabela 52 – Parâmetros analíticos dos métodos convencionais e desenvolvidos para determinação frações de TPH.

Parâmetro	GC (EPA)	UFGC	NIR Solo fortificado	NIR Soluções em hexano	NIR Soluções em etanol
GRO					
R ² ajustado	0,8904	0,9999	0,9781	0,9982	0,9992
r	0,9580	0,9999	0,9893	0,9991	0,9996
RMSE (mg kg ⁻¹)	750	35	557	118	48
Sensib. analít. (mg kg ⁻¹) ⁻¹	83,49	0,22	8,5	33	40
RPD	3,0	92,2	4,6	10,0	9,0
DPR menor concentração (%)	4,0	2,4	34,8	35,8	45,3
	(672)	(1059)	(1119)	(1075)	(438)
DPR maior concentração (%)	2,6	1,7	4,4	5,5	7,9
	(6090)	(1377)	(11681)	(6703)	(4913)
DRO					
R ² ajustado	0,9969	0,9998	0,9964	0,9982	0,9995
r	0,9988	0,9999	0,9983	0,9992	0,9998
RMSE (mg kg ⁻¹)	23	45	322	155	50
Sensib. analít. (mg kg ⁻¹) ⁻¹	0,17	0,08	21	32	43
RPD	16,9	69,6	5,1	6,3	15,8
DPR menor concentração (%)	6,3	27,1	65,9	49,6	61,2
	(111)	(151)	(8,9)	(1000)	(514)
DPR maior concentração (%)	7,0	0,2	6,4	10,2	36,1
	(1161)	(9625)	(163)	(14.539)	(1093)
ORO					
R ² ajustado		0,9968	0,9418	0,9933	0,9923
r		0,9987	0,9714	0,9967	0,9912
RMSE (mg kg ⁻¹)		39	690	258	118
Sensib. analít. (mg kg ⁻¹) ⁻¹		0,03	5	13	15
RPD		69,7	2,8	6,4	3,9
DPR menor concentração (%)		12,07	79,1	86,0	38,2
		(147)	(1021)	(831)	(693)
DPR maior concentração (%)		19,95	5,8	6,4	13,6
		(707)	(10.893)	(10.119)	(1085)

5.4.2. Avaliação de frequência analítica e consumo de solvente.

De acordo com os dados da Cetesb para ano de 2017, 72% das 5942 áreas contaminadas cadastradas no estado de São Paulo foram provenientes de postos de combustíveis. ⁶ Destes 72%, apenas 20% já foram reabilitadas para reuso, enquanto os 80% estão em processo de investigação, remediação ou monitoramento para encerramento, ou seja, ainda estão sendo realizadas análises ambientais de 3427 áreas contaminadas.

No mesmo ano, o número de postos revendedores de combustíveis, no estado de São Paulo, cadastrados na ANP foi igual a 15.519 postos.¹⁹¹ Deste modo, considerando que as 3427 áreas contaminadas sejam de postos cadastrados na ANP, pelo menos cerca de 22% dos postos no estado de São Paulo estiveram contaminados no ano de 2017.

Segundo a CETESB,⁶ os compostos BTEX são os principais compostos presentes em áreas contaminadas (4049 constatações), seguidos pelos PAH (2494 constatações). Se os compostos BTEX representarem a gasolina e, os PAH, o diesel e o óleo lubrificante, pode-se dizer 62% das contaminações foram causadas por gasolina e 38% pelos outros dois contaminantes.

Os valores apresentados nesta seção foram utilizados para simular a aplicação dos métodos convencionais e dos métodos aqui desenvolvidos a fim de comparação. Para simplificar a explanação, consideramos a análise de 100 amostras de solo, cada uma coletada em um diferente posto de combustível. Dado que 22% dos postos estão em áreas contaminadas, supomos que 78 amostras analisadas não estão contaminadas.

Utilizando os métodos convencionais de acordo com a EPA, todas as 100 amostras serão analisadas duas vezes: uma para determinação de GRO; e outra para determinação de DRO e ORO. Para a determinação de GRO, a extração por headspace (EPA 5021) demanda 50 min para a primeira amostra e a extração das demais ocorre durante as análises cromatográficas. Segundo o método EPA 8015,⁸ cada análise cromatográfica requer 45 min por amostra, considerando-se o tempo de resfriamento do forno entre as análises. Para a determinação das frações DRO e ORO, o tempo de análise cromatográfica é o mesmo, porém, o tempo de extração e purificação foi igual a 2 min por amostra, aproximadamente. De acordo com a EPA 3550C¹⁶, a extração de analitos acima de 20 mg kg⁻¹ necessita de 10 mL de solvente (diclorometano). Portanto, a determinação das frações de TPH das 100 amostras requer 157,5 h e 1 L de diclorometano, conforme os métodos convencionais.

Através dos métodos cromatográficos ultrarrápidos desenvolvidos, as amostras também devem ser analisadas por dois métodos, porém cada análise requer 10 min por amostra, considerando a estabilização do sistema entre as análises. As três frações de TPH são extraídas por sonicação, porém a fração GRO deve ser extraída com etanol (7,5 mL por amostra) e as frações DRO e ORO com hexano (10 mL por amostra). O tempo demandado para extração e purificação de acordo com o método utilizado é de aproximadamente 4 min por amostra. Assim o tempo total para análise das 100 amostras é igual a 46,7 h e o volume de solvente de 0,75 L de etanol e 1 L de hexano.

De acordo com os resultados da validação dos métodos NIR desenvolvidos, a análise direta das amostras de solo pode ser utilizada apenas qualitativamente, sendo útil na identificação da fonte de contaminação, enquanto as análises quantitativas devem ser realizadas após extração com etanol. Se o solo estiver contaminado com óleo lubrificante na concentração acima de 5700 mg kg^{-1} , a extração deve ser realizada com hexano. Como em apenas uma análise todas as frações são avaliadas, as análises espectroscópicas requerem cerca de 3 min por amostra e a extração 4 min, totalizando em 11,7 h e 0,75 mL de etanol.

Se o método NIR for utilizado qualitativamente previamente às análises cromatográficas ou espectroscópicas (quantitativas), o tempo de análise e o gasto de solvente podem ser reduzidos consideravelmente. A partir dos dados citados no início da seção, criamos uma situação hipotética com as 100 amostras para exemplificar. Considerando que 22% dos postos estão cadastrados como áreas contaminadas, 78 das 100 amostras estão não contaminadas. Assim, o tempo de análise gasto para avaliar qualitativamente seria de 2 min por amostra, totalizando em 3,3 h. Sabendo-se que 68% das contaminações são causadas por vazamento de gasolina, o tempo de análise pelos métodos convencionais seria reduzido para 17,5 h e o solvente para 0,07 L de diclorometano; pelos métodos ultrarrápidos, o tempo seria reduzido para 5,1 h e os solventes para 0,11 L de etanol e 0,07 L de hexano; e pelo método NIR, o tempo necessário seria de 1,8 h e solvente seria 0,17 L de etanol.

Na Tabela 53, estão dispostos os valores de frequência analítica e o volume de solvente para a análise das 100 amostras, nas duas situações hipotéticas: (i) análise completa das 100 amostras; e (ii) análise NIR prévia para avaliação qualitativa e, subsequente, análise quantitativa apenas das 22 amostras contaminadas. Na primeira situação, o método UFGC é 3,4 vezes mais rápido que o método convencional e o método NIR é 13,5 vezes mais rápido. E, na segunda situação, o método UFGC é 4,1 vezes mais rápido que o método convencional e o método NIR é 11,6 vezes mais rápido. Quanto ao uso de NIR previamente ao método convencional, o tempo de análise é reduzido em 7,6 vezes e, quando comparamos ao método convencional, com a aplicação dos métodos NIR qualitativo e quantitativo, a redução do tempo gasto é de 87,5 vezes e do consumo de solvente de 5,9 vezes, além da substituição de diclorometano por etanol.

Tabela 53 – Frequência analítica e volume de solvente necessários para a análise de 100 amostras em cada método proposto e em duas situações hipotéticas (i) 100% amostras contaminadas; e (ii) 22% amostras contaminadas.

	Frequência analítica (amostras/horas)	Tempo de análise (h/100 amostras)	V_{solvente} (L/100 amostras)
(i) 100% amostras contaminadas			
GC (EPA)	0,6	157,5	1,00 (diclorometano)
UFGC	2,1	46,7	0,75 (etanol) e 1,00 (hexano)
NIR (quantit.)	8,6	11,7	0,75 (etanol)
(ii) 22% amostras contaminadas			
NIR (qualit.) + GC (EPA)	4,8	20,8	0,07 (diclorometano)
NIR (qualit.) + UFGC	19,6	5,1	0,11 (etanol) e 0,07 (hexano)
NIR (qualit.) + NIR (quantit.)	55,6	1,8	0,17 (etanol)

5.4.3. Comparativo de custos estimados

Os custos das análises das frações de TPH pelo método cromatográfico EPA e pelos métodos desenvolvidos, UFGC e NIR, foram estimados considerando apenas os custos diretos (equipamentos, insumos e analista). Os valores dos equipamentos foram cotados com fornecedores locais, em 2017, e os valores dos insumos foram obtidos pelo website da Sigma-Aldrich. O cálculo do valor da hora dos equipamentos foi realizado considerando a depreciação dos equipamentos em 10 anos, utilizando-os durante 8 h por dia, apenas em dias úteis, e com dedicação exclusiva para análise de TPH. Para obter o valor da hora, o valor do equipamento foi dividido pelo número máximo de amostras que podem ser analisadas em 10 anos.

Para os métodos cromatográficos, levaram-se em consideração duas análises por amostra: uma para a faixa GRO e outra para as faixas DRO e ORO. A descrição detalhada dos valores estimados dos custos para cada método está disponível na Tabela 54.

Tabela 54 – Descrição dos custos para análise das frações de TPH pelo método cromatográfico (EPA), método cromatográfico ultrarrápido (UFGC) e método espectroscópico NIR.

Método GC (EPA)				
Descrição do custo	Valor unitário	Valor por hora	Quantidade por amostra	Custo por amostra
Ultrassom	R\$ 3.197,51	R\$ 0,16	0,005 h	R\$ 0,00
Bomba de vácuo	R\$ 4.371,21	R\$ 0,22	0,12 h	R\$ 0,03
Filtro de Buchner	R\$ 16,80	R\$ 0,00	0,03 h	R\$ 0,00
Kitassato	R\$ 33,81	R\$ 0,00	0,03 h	R\$ 0,00
Coluna de vidro	R\$ 54,20	R\$ 0,00	0,09 h	R\$ 0,00
Na ₂ SO ₄	R\$ 267,00		0,002 kg	R\$ 0,53
Sílica gel	R\$ 2.420,00		0,0001 kg	R\$ 0,24
Diclorometano	R\$ 227,00		0,01 mL	R\$ 2,27
Pipeta de Pasteur descartável	R\$ 0,07		1 un.	R\$ 0,07
Vial 20 mL	R\$ 1,57		2 un.	R\$ 3,14
Vial 2 mL	R\$ 1,11		1 un.	R\$ 1,11
GC-FID + autoamostrador HS	R\$ 350.000,00	R\$ 17,28	1,5 h	R\$ 25,92
Coluna cromatográfica	R\$ 2.224,00	R\$ 0,11	1,5 h	R\$ 0,16
Septo	R\$ 8,56	R\$ 0,00	0,01 un.	R\$ 0,00
Microseringa	R\$ 349,19	R\$ 0,02	0,0004	R\$ 0,00
Nitrogênio (m ³)	R\$ 56,50		2700 mL	R\$ 0,15
Ar sintético (m ³)	R\$ 30,00		36000 mL	R\$ 1,08
Hélio (m ³)	R\$ 140,00		9900 mL	R\$ 1,39
Hidrogênio (m ³)	R\$ 58,75		2700 mL	R\$ 0,16
MRC TPH mix	R\$ 227,00		0,0001 mL	R\$ 0,02
MRC GRO mix	R\$ 289,00		0,0001 mL	R\$ 0,03
Padrão o-terfenil (25 g)	R\$ 354,00		0,001 g	R\$ 0,35
Padrão isobutilbenzeno (200 mL)	R\$ 628,00		0,002 mL	R\$ 1,26
Manutenção preventiva	R\$ 4.674,82	R\$ 0,23	1,5 h	R\$ 0,35
Manutenção corretiva	R\$ 5.813,69	R\$ 0,29	1,5 h	R\$ 0,43
Analista (ano)	R\$ 39.647,30	R\$ 20,02	1,6 h	R\$ 32,03
Total				R\$ 70,73
Método UFGC				
Descrição do custo	Valor unitário	Valor por hora	Quantidade por amostra	Custo por amostra
Ultrassom	R\$ 3.197,51	R\$ 0,16	0,0375 h	R\$ 0,01
Seringa desc.	R\$ 0,48		1	R\$ 0,48
Filtro de seringa PTFE	R\$ 7,92		1	R\$ 7,92
Na ₂ SO ₄	R\$ 267		0,004 kg	R\$ 1,07
Sílica gel	R\$ 2.420		0,002 kg	R\$ 4,84
Etanol	R\$ 227		0,0075 L	R\$ 1,70
n-Hexano	R\$ 240		0,010 L	R\$ 2,40
Pipeta de Pasteur descartável	R\$ 0,07		2	R\$ 0,15
Vial 22 mL	R\$ 1,57		2	R\$ 3,14
Vial 2 mL	R\$ 1,11		2	R\$ 2,22
UFGC-FID + autoamostrador liq.	R\$ 157.063,40	R\$ 7,75	0,33 h	R\$ 2,56
Septo	R\$ 8,56	R\$ 0,00	0,01	R\$ 0,00
Microseringa	R\$ 349,19	R\$ 0,02	0,0004	R\$ 0,00
Coluna PH SE54	R\$ 14.837,21	R\$ 0,73	0,17 h	R\$ 0,12
Coluna RTx-1	R\$ 19.732,18	R\$ 0,97	0,17 h	R\$ 0,17
Nitrogênio (m ³)	R\$ 56,50		600 mL	R\$ 0,03
Ar Sintético (m ³)	R\$ 30		8000 mL	R\$ 0,24
Hidrogênio (m ³)	R\$ 58,75		2910 mL	R\$ 0,17
MRC TPH mix	R\$ 227		0,00003 mL	R\$ 0,01
MRC GRO mix	R\$ 289,00		0,00003 mL	R\$ 0,03
Padrão o-terfenil (25 g)	R\$ 354,00		0,001 g	R\$ 0,35
Padrão isobutilbenzeno (100 mL)	R\$ 314,00		0,002 mL	R\$ 0,63
Manutenção preventiva	R\$ 4.674,82	R\$ 0,23	0,33 h	R\$ 0,08
Manutenção corretiva	R\$ 5.813,69	R\$ 0,29	0,33 h	R\$ 0,09
Analista (ano)	R\$ 39.647,30	R\$ 20,02	0,47 h	R\$ 9,41
Total				R\$ 37,82

Continua.

Tabela 54 – Descrição dos custos para análise das frações de TPH pelo método cromatográfico (EPA), método cromatográfico ultrarrápido (UFGC) e método espectroscópico NIR. (Continuação).

Método NIR				
Descrição do custo	Valor unitário	Valor por hora	Quantidade por amostra	Custo por amostra
Ultrassom	R\$ 3.197,51	R\$ 0,16	0,0375 h	R\$ 0,01
Seringa descartável	R\$ 0,48		1	R\$ 0,48
Filtro de seringa PTFE	R\$ 7,92		1	R\$ 7,92
Na ₂ SO ₄	R\$ 267		0,002 kg	R\$ 0,53
Sílica gel	R\$ 2.420		0,001 kg	R\$ 2,42
Etanol	R\$ 227		0,0075 L	R\$ 1,70
Pipeta de Pasteur descartável	R\$ 0,07		1	R\$ 0,07
Vial 22 mL	R\$ 1,57		1	R\$ 1,57
Vial 2 mL	R\$ 1,11		1	R\$ 1,11
Espectrômetro NIR	R\$ 282.648,00	R\$ 13,95	0,05 h	R\$ 0,70
Manutenção preventiva	R\$ 2.688,79	R\$ 0,13	0,05 h	R\$ 0,01
Manutenção corretiva	R\$ 3.788,74	R\$ 0,19	0,05 h	R\$ 0,01
Analista (ano)	R\$ 39.647,30	R\$ 20,02	0,12 h	R\$ 2,40
Total				R\$ 18,93

Notou-se que os maiores custos para a análise pelo método convencional concentraram-se no tempo do analista e no cromatógrafo com autoamostrador equipado com sistema headspace. Estes custos apresentaram-se elevados em decorrência do número máximo de amostras analisadas, ou seja, o valor da hora trabalhada e o uso do equipamento foram diluídos em poucas amostras devido ao longo tempo de análise.

O método UFGC apresentou, aproximadamente, metade do custo por amostra que o método convencional. A substituição do método de extração headspace pela extração por solvente permitiu o uso de um autoamostrador mais simples, reduzindo o valor do equipamento pela metade. Além disso, a maior frequência analítica proporcionada pelo método reduziu o custo relacionado ao equipamento e ao analista. Os maiores pesos para o custo final da análise foram a hora trabalhada pelo analista, o filtro de seringa e a sílica gel.

A determinação das frações de TPH pelo método NIR proporcionou um custo por amostra 3,7 vezes menor que o método convencional. Devido à elevada frequência analítica, o custo do equipamento por amostra tornou-se muito baixo. Além da análise espectroscópica não exigir insumos, a permanência do equipamento ligado à corrente elétrica durante 24 h é recomendada pelos fabricantes para reduzir os efeitos da umidade. Portanto, os custos da análise pelo método NIR concentraram-se no preparo de amostra.

Para estimar o tempo necessário para obter o retorno financeiro da implementação do método analítico, foram realizadas algumas considerações. O valor do investimento inicial foi considerado o custo para aquisição dos equipamentos e insumos necessários para análise de

1000 amostras. O valor cobrado pela análise foi tido como o valor do custo mais 100% deste valor. O número de amostras analisadas anualmente foi determinado pelo número aproximado de postos de combustíveis no raio de 70 km de Araraquara (1004 postos) e a análise de cada amostra realizada em triplicata, totalizando em 3012 amostras por ano.

Os itens levados em consideração para cálculo do tempo para o retorno financeiro estão apresentados na Tabela 55. Segundo a frequência analítica, o método GC (EPA) permite o número máximo análises por ano igual a 1608, enquanto os demais métodos estão limitados ao número de amostras coletadas na região, 3012 amostras. Considerando o valor cobrado por análise e o número de amostras analisadas anualmente, o método UFGC proporciona o retorno do investimento em menor tempo, 2,4 anos, seguido pelo método GC (EPA), 3,7 anos, e por fim o método NIR, 6,1 anos.

Em contraste com o tempo necessário para recuperação do investimento, há o custo anual para o contratante da prestação de serviços. O custo anual para as 3012 análises pelo método UFGC é cerca de 2 vezes maior que o método NIR e, aproximadamente, 2 vezes menor que o método convencional (analisando-se 1608 amostras). Portanto, a análise das frações de TPH pelo método NIR apresenta maior vantagem econômica do ponto de vista do contratante.

Tabela 55 – Avaliação do tempo necessário para recuperação do investimento inicial e o custo anual para o contratante de acordo com cada método analítico.

Item	GC-FID	UFGC-FID	NIR
Investimento inicial	R\$ 422.821,83	R\$ 275.793,60	R\$ 346.669,34
Custo da análise	R\$ 70,73	R\$ 37,82	R\$ 18,93
Valor cobrado por análise (custo + 100%)	R\$ 141,45	R\$ 75,63	R\$ 37,86
Nº ensaios para abater investimento	5978	7293	18311
Nº máximo de ensaios por ano	1608	5422	21641
Nº de análises na região*	3012	3012	3012
Tempo para pagar o investimento (anos)	3,7	2,4	6,1
Custo para o contratante (anual)	R\$ 426.061,39	R\$ 227.797,96	R\$ 114.046,36

*Monitoramento dos 1004 postos de combustíveis da região de Araraquara - análises em triplicata.

Através do gráfico da Figura 22, observa-se que o número de amostras é um forte limitante para a escolha do método analítico. Quanto maior for o número de amostras ao longo dos anos, o faturamento pelo método NIR torna-se consideravelmente maior que dos demais métodos. Porém, se o número de amostras analisadas por ano for entre 9000 e 2000 amostras, o maior faturamento é obtido pelo método UFGC, e se for menor que 2000, torna-se vantajoso o uso do método GC (EPA).

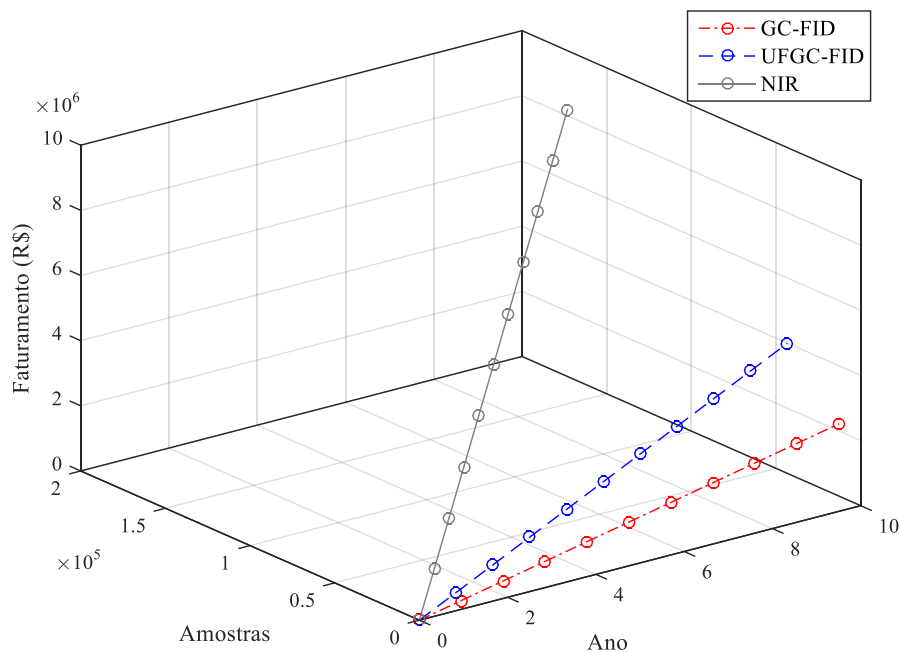


Figura 22 – Relação entre o faturamento obtido de acordo com cada método analítico e o número de amostra ao longo dez anos.

5.4.4. Comparação geral entre os métodos analíticos

Com o objetivo de otimizar a visão geral dos pontos positivos e negativos de cada método, selecionaram-se alguns fatores importantes para fins de comparação. Estes fatores tiveram seus valores normalizados pela média e o maior valor foi utilizado como referência para a maior pontuação (5 pontos). A pontuação dos fatores dos três métodos avaliados (GC, UFGC e NIR) estão presentes na Tabela 56.

Tabela 56 – Pontuação (1 a 5) dos fatores considerados na comparação dos métodos analíticos.

Método	Exatidão			Sensibilidade			Precisão		
	GRO	DRO	ORO	GRO	DRO	ORO	GRO	DRO	ORO
GC	*****	*****	-	*****	*	-	*****	*****	-
UFGC	*****	*****	*****	*	*	*	*****	***	***
NIR	*****	*****	*****	**	*****	*****	*	*	*

Método	Frequência analítica	Volume de solvente	Química Verde	Custo por análise	Investimento inicial	Tempo para retorno financeiro	Custo anual para o contratante
GC	*	***	*	*****	*****	***	*****
UFGC	**	*****	**	***	**	**	***
NIR	*****	**	****	*	****	*****	**

De forma geral, os três métodos apresentaram mesma exatidão; o método GC (EPA) foi o método mais preciso; o método UFGC foi o de menor investimento inicial e de menor tempo para retorno financeiro; e, finalmente, o método NIR após extração com etanol foi o mais sensível, rápido, verde e de menor custo por análise.

6. CONCLUSÕES

O presente trabalho demonstrou o desenvolvimento e validação de métodos analíticos para identificação e quantificação dos hidrocarbonetos totais do petróleo provenientes da contaminação de solo por postos de combustíveis. Além da complexidade composicional dos combustíveis e óleos lubrificantes derivados do petróleo, o solo é uma matriz complexa que gera grande quantidade de interferências na análise NIR, principalmente, pela aquisição espectral no modo de refletância difusa.

Na etapa de desenvolvimento e validação dos métodos de referência UFGC e otimização do método extração, concluiu-se a possibilidade da análise das frações de TPH de forma muito mais rápida (até 17 vezes mais rápida que o método EPA 8015) e baseada nos princípios da química verde, através do uso de menor volume de solventes não halogenados, como etanol e hexano. Apesar do método UFGC apresentar-se linear, preciso e exato, necessitou-se o uso de duas colunas cromatográficas, uma para a fração GRO e outra para as frações DRO e ORO.

Os modelos de reconhecimento de padrão mostraram-se aptos à identificação de contaminação em solo por diferentes frações de TPH. O modelo PLS-DA (sem preparo de amostra) não apresentou falsos positivos para contaminação nas amostras “reais”, entretanto, a origem do contaminante não foi totalmente identificada de forma correta. O modelo KNN (extração com etanol) e os modelos SVM (extração com hexano ou etanol) também não geraram falsos positivos e, além disso, proporcionaram porcentagens maiores de classificações corretas em relação às origens das contaminações. Em acordo à literatura, a aplicação do pré-processamento EPO aos espectros NIR (sem preparo de amostra) gerou melhores ajustes ao modelo de predição em consequência da atenuação dos interferentes presentes nos solos.

Quanto aos modelos de calibração, diferentes frações e analitos exigiram diferentes pré-processamentos para tornar o modelo mais ajustado. Os pré-processamentos de normalização, MSC e SNV, foram mais adequados para o conjunto amostral de solos fortificados; pré-processamentos de ortogonalização foram melhores para GRO em hexano (OSC) e ORO em etanol (EPO); centragem na média para DRO em hexano; e autoescalamamento foi o mais adequado para os demais modelos. A seleção de variáveis foi executada pelo algoritmo genético (GA) em duas fases. A execução sequencial do GA reduziu significativamente os erros para todos os analitos, nos três conjuntos amostrais. Todos modelos GA-PLS mostraram exatidão aceitável e boa linearidade, contudo, baixa sensibilidade, seletividade e precisão. Considerando-se todos os parâmetros analíticos avaliados, os modelos gerados a partir de

soluções em etanol apresentaram os melhores resultados de calibração e validação. A predição das amostras “reais” mostrou que a análise direta das amostras de solo gerou resultados incoerentes com os valores de referência em decorrência da grande diferença entre as amostras de calibração e as amostras “reais”. Já os métodos com extração por solvente proporcionaram predições próximas aos valores de referência, tornando-se evidente a melhor exatidão e precisão dos métodos NIR após extração por etanol.

Finalmente, a avaliação entre as características de cada método (GC convencional, UFGC e NIR após extração por etanol) mostrou exatidão similar para os três métodos; entretanto, o método GC (EPA) foi o mais preciso; o método UFGC foi o de menor investimento inicial e de menor tempo para retorno financeiro; e o método NIR foi o mais sensível, rápido, verde e de menor custo por análise. Desta forma, as barreiras presentes no processo de monitoramento ambiental, decorrentes dos custos e do tempo exigidos pelas análises convencionais, podem ser superadas através da aplicação da espectroscopia NIR.

REFERÊNCIAS

- 1 BACHELARD, G. **A formação do espírito científico: contribuição para uma psicanálise do conhecimento**. Rio de Janeiro: Contraponto, 1996. 316 p.
- 2 BP. **Statistical Review of World Energy**. London 2016. (Report). Disponível em: <<https://www.bp.com/content/dam/bp/pdf/energy-economics/statistical-review-2016/bp-statistical-review-of-world-energy-2016-full-report.pdf>>. Acessado em: 31 maio 2017.
- 3 TODD, G. D.; CHESSIN, R. L.; COLMAN, J. **Toxicological profile for total petroleum hydrocarbons (TPH)**. Atlanta: Agency for Toxic Substances and Disease Registry, 1999. pg 51.
- 4 WANG, S.Y. et al. Bioremediation of diesel and lubricant oil-contaminated soils using enhanced landfarming system. **Chemosphere**, v. 164, p. 558–567, 2016.
- 5 WEISMAN, W.H. **Analysis of petroleum hydrocarbons in environmental media**. Amherst: Total Petroleum Hydrocarbon Criteria Working Group Series, 1998. 98 p.
- 6 FILHO, G. A. et al. Texto explicativo - Relação de áreas contaminadas e reabilitadas no Estado de São Paulo. São Paulo, 2017. Disponível em: <<http://cetesb.sp.gov.br/areas-contaminadas/wp-content/uploads/sites/17/2018/01/Texto-explicativo.pdf>>. Acesso em: 02 jan. 2018
- 7 COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO. **Sistema de licenciamento de postos IV - Procedimento para identificação de passivos ambientais em estabelecimentos com sistema de armazenamento subterrâneo de combustíveis (SASC)**. São Paulo, 2016. 8 p.
- 8 US ENVIRONMENTAL PROTECTION AGENCY. **Method 8015C**: nonhalogenated organics by gas chromatography. Washington DC, 2007. 36 p.
- 9 US ENVIRONMENTAL PROTECTION AGENCY. **Method 418.1**: petroleum hydrocarbons (spectrophotometric, infrared). Washington DC, 1978. 3 p.
- 10 SCHWARTZ, G.; BEN-DOR, E.; ESHEL, G. Quantitative assessment of hydrocarbon contamination in soil using reflectance spectroscopy: A “multipath” approach. **Applied Spectroscopy**, v. 67, n. 11, p. 1323–1331, 2013.
- 11 US ENVIRONMENTAL PROTECTION AGENCY. **Method 8440**: Total recoverable petroleum hydrocarbons by infrared spectrophotometry. Washington DC, 1996. 8 p.
- 12 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM D7678**: Standard test method for total petroleum hydrocarbons (TPH) in water and wastewater with solvent extraction using mid-IR laser. West Conshohocken, 2011. 9 p.

- 13 US ENVIRONMENTAL PROTECTION AGENCY. **Method 8270D**: semivolatile organic compounds by gas chromatography/mass spectrometry (GC/MS). Washington DC, 2007. 72 p.
- 14 US ENVIRONMENTAL PROTECTION AGENCY. **Method 3510C**: separatory funnel liquid-liquid extraction. Washington DC, 1996. 8 p.
- 15 US ENVIRONMENTAL PROTECTION AGENCY. **Method 3540C**: soxhlet extraction. Washington DC, 1996. 8 p.
- 16 US ENVIRONMENTAL PROTECTION AGENCY. **Method 3550**: ultrasonic extraction. Washington DC, 2007. 17 p.
- 17 JACOBI, H. F.; MOSCHNER, C. R.; HARTUNG, E. Use of near infrared spectroscopy in monitoring of volatile fatty acids in anaerobic digestion. **Water Science & Technology**, v. 60, n. 2, p. 339, 2009.
- 18 PASQUINI, C. Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society**, v. 14, n. 2, p. 198–219, 2003.
- 19 BLANCO, M.; VILLARROYA, I. NIR spectroscopy: A rapid-response analytical tool. **Trends in Analytical Chemistry**, v. 21, n. 4, p. 240–250, 2002.
- 20 FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, P. L. O. Chemometrics I: multivariate calibration, a tutorial. **Química Nova**, 22, 724–731. 1999.
- 21 HATANAKA, R. R. et al. Development and validation of an environmentally friendly attenuated total reflectance in the mid-infrared region method for the determination of ethanol content in used engine lubrication oil. **Talanta**, v. 109, p. 191–196, 2013.
- 22 WEBSTER, G. T. et al. Rapid prediction of total petroleum hydrocarbons in soil using a hand-held mid-infrared field instrument. **Talanta**, v. 160, p. 410–416, 2016.
- 23 FORRESTER, S.T. et al. Total Petroleum Hydrocarbon Concentration Prediction in Soils Using Diffuse Reflectance Infrared Spectroscopy. **Soil Science Society of America Journal**, v. 77, n. 2, p. 450, 2013.
- 24 CHAKRABORTY, S. et al. Analysis of petroleum contaminated soils by spectral modeling and pure response profile recovery of n-hexane. **Environmental Pollution**, v. 190, p. 10–18, 2014.
- 25 OKPARANMA, R.N.; MOUAZEN, A.M. Combined effects of oil concentration, clay and moisture contents on diffuse reflectance spectra of diesel-contaminated soils. **Water, Air, and Soil Pollution**, v. 224, n. 5, 2013.
- 26 VISCARRA ROSSEL, R.A. et al. A global spectral library to characterize the world's soil. **Earth-Science Reviews**, v. 155, n. 1, p. 198–230, 2016.

- 27 COZZOLINO, D. An Overview of the Application of Near Infrared Spectroscopy to Analyze and Monitor Soil Properties in South America. **Applied Spectroscopy Reviews**, v. 50, n. 10, p. 859–867, 2015.
- 28 DOTTO, A.C. et al. Two preprocessing techniques to reduce model covariables in soil property predictions by Vis-NIR spectroscopy. **Soil and Tillage Research**, v. 172, n. 3, p. 59–68, 2017.
- 29 COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO. **Qualidade do solo**. São Paulo, 2018. Disponível em: <<http://cetesb.sp.gov.br/solo/qualidade-do-solo/>>. Acesso em: 21 jun. 2018.
- 30 SENA, M.M. et al. Avaliação do uso de métodos quimiométricos em análise de solos. **Química Nova**, v. 23, n. 4, p. 547–556, 2000.
- 31 BÜNEMANN, E.K.; et al. Soil quality – A critical review. **Soil Biology and Biochemistry**, v. 120, n. 9, p. 105–125, 2018.
- 32 XU, S. et al. Comparison of multivariate methods for estimating selected soil properties from intact soil cores of paddy fields by Vis–NIR spectroscopy. **Geoderma**, v. 310, n. 9, p. 29–43, 2018.
- 33 NOCITA, M. et al. Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring. **Advances in Agronomy**, v. 132, p. 139–159, 2015.
- 34 O’ROURKE, S.M. et al. Synergistic Use of Vis-NIR, MIR, and XRF Spectroscopy for the Determination of Soil Geochemistry. **Soil Science Society of America Journal**, v. 80, n. 4, p. 888, 2016.
- 35 ESCOBAR-ALVARADO, L.F. et al. Hydrocarbon Degradation and Lead Solubility in a Soil Polluted with Lead and Used Motor Oil Treated by Composting and Phytoremediation. **Bulletin of Environmental Contamination and Toxicology**, v. 100, n. 2, p. 280–285, 2018.
- 36 COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO. **Decisão de Diretoria no 256/2016/E**. São Paulo, 2016. 4 p.
- 37 GALLOTTA, F.D.C.; CHRISTENSEN, J.H. Source identification of petroleum hydrocarbons in soil and sediments from Iguazu River Watershed, Paraná, Brazil using the CHEMSIC method (CHEMometric analysis of Selected Ion Chromatograms). **Journal of Chromatography A**, v. 1235, p. 149–158, 2012.
- 38 DOUGLAS, G. et al. The use of hydrocarbon analyses for environmental assessment and remediation. **Journal of Soil Contamination**, v. 3, n. 10, p. 197–216, 1992.
- 39 SCHWAB, A.P. et al. Extraction of petroleum hydrocarbons from soil by mechanical shaking. **Environmental Science and Technology**, v. 33, n. 11, p. 1940–1945, 1999.
- 40 FAUSTORILLA, M.V. et al. Determination of total petroleum hydrocarbons in australian groundwater through the improvised gas chromatography–flame ionization detection technique. **Journal of Chromatographic Science**, v. 55, n. 8, p. 775–783, 2017.

- 41 PIÑEDO, J. et al. Assessment of soil pollution based on total petroleum hydrocarbons and individual oil substances. **Journal of Environmental Management**, v. 130, p. 72–79, 2013.
- 42 DİNDAR, E. et al. Biodegradation of used engine oil in a wastewater sludge-amended agricultural soil. **Turkish Journal of Agriculture and Forestry**, v. 40, p. 631–641, 2016.
- 43 WU, H. et al. Persulfate oxidation for the remediation of petroleum hydrocarbon-contaminated soils. **Polish Journal of Environmental Studies**, v. 25, n. 2, p. 851–857, 2016.
- 44 CAETANO, M.O. et al. A compact remediation system for the treatment of groundwater contaminated with BTEX and TPH. **Environmental Technology**, v. 38, n. 11, p. 1408–1420, 2017.
- 45 PAÍGA, P. et al. Determination of total petroleum hydrocarbons in soil from different locations using infrared spectrophotometry and gas chromatography. **Chemical Papers**, v. 66, n. 8, p. 711–721, 2012.
- 46 OKPARANMA, R.N.; MOUAZEN, A.M. Determination of total petroleum hydrocarbon (TPH) and polycyclic aromatic hydrocarbon (PAH) in soils: A review of spectroscopic and nonspectroscopic techniques. **Applied Spectroscopy Reviews**, v. 48, n. 6, p. 458–486, 2013.
- 47 KWON, M.J. et al. Effects of soil micro-particles and micro-pores on petroleum hydrocarbons released from contaminated soils during solvent extraction with ultrasound. **Water, Air, & Soil Pollution**, v. 227, n. 9, p. 357, 2016.
- 48 HAGHOLLAHI, A. et al. The effect of soil type on the bioremediation of petroleum contaminated soils. **Journal of environmental management**, v. 180, p. 197–201, 2016.
- 49 AMERICAN PETROLEUM INSTITUTE. **Frequently Asked Questions About TPH Analytical Methods for Crude Oil. 2001**. Disponível em: <http://www.api.org/~media/Files/EHS/Environmental_Performance/Frequently-Asked-Questions-About-TPH-Analytical-Methods-for-Crude-Oil.pdf>. Acesso em: 21 jun. 2018.
- 50 PIKOVSKII, Y.I. et al. Laboratory analytical methods for the determination of the hydrocarbon status of soils (a review). **Eurasian Soil Science**, v. 50, n. 10, p. 1125–1137, 2017.
- 51 DINDAR, E. et al. Variations of soil enzyme activities in petroleum-hydrocarbon contaminated soil. **International Biodeterioration & Biodegradation**, v. 105, p. 268–275, 2015.
- 52 WANG, Y. et al. The impact of fuel compositions on the particulate emissions of direct injection gasoline engine. **Fuel**, v. 166, p. 543–552, 2016.
- 53 MARANHÃO, D.; MARQUES, T.; TEIXEIRA, A. Procedimentos de investigação e avaliação da contaminação em postos de combustíveis, utilizando metodologias de análise de risco: aplicação da ACBR em estudo de caso na RMS. 2007. 121 f. Monografia – (Especialização em Gerenciamento de Tecnologias Ambientais) – Universidade Federal da Bahia, Salvador, 2007.

- 54 AGÊNCIA NACIONAL DO PETRÓLEO GÁS NATURAL E BIOCOMBUSTÍVEIS. Percentual obrigatório de biodiesel. 2018a. Disponível em: <<http://www.anp.gov.br/noticias/4333-percentual-obrigatorio-de-biodiesel-passa-para-10>>. Acesso em: 7 jul. 2018.
- 55 EKPERUSI, O.A.; AIGBODION, I.F. Bioremediation of heavy metals and petroleum hydrocarbons in diesel contaminated soil with the earthworm: *Eudrilus eugeniae*. **SpringerPlus**, v. 4, n. 1, p. 540, 2015.
- 56 GUO, H. et al. Effects of petroleum contamination on soil microbial numbers, metabolic activity and urease activity. **Chemosphere**, v. 87, n. 11, p. 1273–1280, 2012.
- 57 VILLALOBOS, M. et al. An improved gravimetric method to determine total petroleum hydrocarbons in contaminated soils. **Water, Air, and Soil Pollution**, v. 194, n. 1–4, p. 151–161, 2008.
- 58 RITCHIE, L. et al. A novel sensor for monitoring leakage of petroleum and other liquid hydrocarbons into soil environments. **Journal of Environmental Monitoring**, v. 2, n. 2, p. 193–196, 2000.
- 59 CHAKRABORTY, S. et al. Development of a hybrid proximal sensing method for rapid identification of petroleum contaminated soils. **Science of the Total Environment**, v. 514, p. 399–408, 2015.
- 60 COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO. **Procedimento para Licenciamento Ambiental de Postos e Sistemas Retalhistas de Combustíveis**. São Paulo, 2018. 7 p.
- 61 CONSELHO NACIONAL DO MEIO AMBIENTE. RESOLUÇÃO CONAMA nº 273, de 29 de novembro de 2000. Brasília, 2000. 10 p.
- 62 HORTA, A. et al. Potential of integrated field spectroscopy and spatial analysis for enhanced assessment of soil contamination: A prospective review. **Geoderma**, v. 241–242, p. 180–209, 2015.
- 63 US ENVIRONMENTAL PROTECTION AGENCY. **Use of risk-based decision-making in UST corrective action programs**. Washington DC, 1995. 16 p.
- 64 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM E1739**: Standard Guide for Risk-Based Corrective Action. West Conshohocken, 2015. 53 p.
- 65 NASCIMENTO, A.R. et al. Avaliação do desempenho analítico do método de determinação de TPH (total petroleum hydrocarbon) por detecção no infravermelho. **Eclética Química**, v. 33, n. 1, p. 35–42, 2008.
- 66 COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO. **Decisão de Diretoria no 038/2017/C**. São Paulo, 2017. 65 p.

- 67 COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO. **Quadros de Exigências para o Licenciamento Ambiental de Postos e Sistemas Retalhistas de Combustíveis**. São Paulo, 2018. 16 p.
- 68 COMPANHIA DE TECNOLOGIA DE SANEAMENTO AMBIENTAL. **Decisão de Diretoria no 195-2005- E, de 23 de novembro de 2005**. São Paulo, 2005. 5 p.
- 69 OKPARANMA, R.N.; MOUAZEN, A.M. Visible and near-infrared spectroscopy analysis of a polycyclic aromatic hydrocarbon in soils. **The Scientific World Journal**, v. 2013, p. 160360, 2013.
- 70 GENNADIEV, A.N. et al. Factors and features of the hydrocarbon status of soils. **Eurasian Soil Science**, v. 48, n. 11, p. 1193–1206, 2015.
- 71 BOLL, E.S. et al. Chemical fingerprinting of hydrocarbon-contamination in soil. Environmental science. **Processes & Impacts**, n. 17, p. 606–618, 2015.
- 72 LIN, Y. et al. Diesel degradation in soil catalyzed by the addition of a bioagent. **International Journal of Environmental Science and Technology**, v. 13, n. 2, p. 551–560, 2016.
- 73 SANTOS, E.V. DOS et al. Application of electrochemical technologies to treat polluted soil by diesel. **Chemical Engineering Transactions**, v. 41, n. Special Issue, p. 157–162, 2014.
- 74 HU, G. et al. Oil recovery from petroleum sludge through ultrasonic assisted solvent extraction. **Journal of Environmental Science and Health, Part A**, v. 51, n. 11, p. 921–929, 2016.
- 75 PEJICIC, B.; Boyd, L.; et al. Direct quantification of aromatic hydrocarbons in geochemical fluids with a mid-infrared attenuated total reflection sensor. **Organic Geochemistry**, v. 55, p. 63–71, 2013.
- 76 SPEIGHT, J.G. **The chemistry and technology of petroleum**. Boca Raton: CRC Press, 2006. 955 p.
- 77 AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS. **Anuário estatístico brasileiro do petróleo, gás natural e biocombustíveis: 2016**. Rio de Janeiro: Agência Nacional do Petróleo Gás Natural e Biocombustíveis, 2016. 265 p.
- 78 GROB, R.L.; BARRY, E.F. **Modern Practice of Gas Chromatography**. Hoboken: John Wiley & Sons, Inc., 2004. 1037 p.
- 79 AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS. **Panorama da Qualidade dos Combustíveis 2011**. Brasília: Agência Nacional do Petróleo Gás Natural e Biocombustíveis, 2012. 32 p.

- 80 AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS. **ANP anuncia medidas para garantir a continuidade do abastecimento e para inibir preços abusivos dos combustíveis**. 2018. Disponível em: <<http://www.anp.gov.br/noticias/4497-anp-anuncia-medidas-para-garantir-a-continuidade-do-abastecimento-e-para-inibir-precos-abusivos-dos-combustiveis>>. Acesso em: 8 jul. 2018.
- 81 ELCOROARISTIZABAL, S. et al. Chemometric determination of PAHs in aerosol samples by fluorescence spectroscopy and second-order data analysis algorithms. **Journal of Chemometrics**, v. 28, n. 4, p. 260–271, 2014.
- 82 JOHNSEN, A.R., WICK, L.Y., HARMS, H. Principles of microbial PAH-degradation in soil. **Environmental Pollution**, v. 133, n. 1, 71–84, 2005.
- 83 OLOLADE, I. A. et al. Spatial trends of petroleum hydrocarbons in water and sediments. **Central European Journal of Chemistry**, v. 7, n. 1, p. 83–89, 2008.
- 84 GUERIN, T.F. The extraction of aged polycyclic aromatic hydrocarbon (PAH) residues from a clay soil using sonication and a Soxhlet procedure: a comparative study. **Journal of Environmental Monitoring**, v. 1, n. 1, p. 63–67, 1999.
- 85 COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO **Relatório de Qualidade das Águas Superficiais**. Disponível em: <<http://cetesb.sp.gov.br/aguas-interiores/wp-content/uploads/sites/12/2013/11/Apêndice-D-Significado-A>>. Acesso em: 13 jun. 2018.
- 86 SCHWARTZ, G. et al. Quantitative analysis of total petroleum hydrocarbons in soils: comparison between reflectance spectroscopy and solvent extraction by 3 certified laboratories. **Applied and Environmental Soil Science**, v. 2012, 2012.
- 87 COUTO, M.N. et al. An improved method for the determination of petroleum hydrocarbons from soil using a simple ultrasonic extraction and fourier transform infrared spectrophotometry. **Petroleum Science and Technology**, v. 32, n. 4, p. 426–432, 2013.
- 88 US ENVIRONMENTAL PROTECTION AGENCY. **Method 1664**: n-hexane extractable material (HEM; oil and grease) and silica gel treated n-hexane extractable material (SGT-HEM; non-polar material) by extraction and gravimetry. Washington DC, 2010. 35 p.
- 89 US ENVIRONMENTAL PROTECTION AGENCY. **Method 3500C**: organic extraction and sample preparation. Test Methods for Evaluating Solid Waste, Physical/Chemical Methods. Washington DC, 2007. 19 p.
- 90 US ENVIRONMENTAL PROTECTION AGENCY (1996). **Method 5000**: sample preparation for volatile organic compounds. Washington DC, 1996. 11 p.
- 91 ROBBAT JR., A. Environmental applications of thermal extraction cone penetrometry and ultrafast gas chromatography/mass spectrometry. **Field analytical chemistry and technology**, v. 5, p. 60–68, 2001.
- 92 GOBRECHT, A.; ROGER, J. M.; BELLON-MAUREL, V. Major issues of diffuse reflectance NIR spectroscopy in the specific context of soil carbon content estimation: a review. **Advances in Agronomy**. v. 123, n. 1, p. 145–175, 2014.

- 93 BICCHI, C. et al. High-speed gas chromatography with direct resistively-heated column (ultra fast module-GC)-separation measure (S) and other chromatographic parameters under different analysis conditions for samples of different complexities and volatilities. **Journal of Chromatography A**, v. 1071, n. 1–2, p. 3–12, 2005.
- 94 NAMIESNIK, J.; WARDENCKI, W. Solventless sample preparation techniques in environmental analysis. **HRC Journal of High Resolution Chromatography**, v. 23, n. 4, p. 297–303, 2000.
- 95 RISDON, G.C. et al. Development of an analytical procedure for weathered hydrocarbon contaminated soils within a UK risk-based framework. **Analytical Chemistry**, v. 80, n. 18, p. 7090–7096, 2008.
- 96 OKPARANMA, R.N. et al. Analysis of petroleum-contaminated soils by diffuse reflectance spectroscopy and sequential ultrasonic solvent extraction-gas chromatography. **Environmental Pollution**, v. 184, p. 298–305, 2014.
- 97 DOUGLAS, R.K. et al. Rapid prediction of total petroleum hydrocarbons concentration in contaminated soil using vis-NIR spectroscopy and regression techniques. **Science of the Total Environment**, v. 616–617, p. 147–155, 2018.
- 98 DOUGLAS, R.K. et al. Rapid detection of alkanes and polycyclic aromatic hydrocarbons in oil-contaminated soil with visible near-infrared spectroscopy. **European Journal of Soil Science**, n. Especial issue, p. 1–11, 2018.
- 99 LETELLIER, M. et al. Focused microwave-assisted extraction of polycyclic aromatic hydrocarbons and alkanes from sediments and source rocks. **Organic Geochemistry**, v. 30, n. 11, p. 1353–1365, 1999.
- 100 MIÈGE, C. et al. Optimization and validation of solvent and supercritical-fluid extractions for the trace-determination of polycyclic aromatic hydrocarbons in sewage sludges by liquid chromatography coupled to diode-array and fluorescence detection. **Journal of Chromatography A**, v. 823, n. 1–2, p. 219–230, 1998.
- 101 CURRENT, R.W.; TILOTTA, D.C. Determination of total petroleum hydrocarbons in soil by on-line supercritical fluid extraction-infrared spectroscopy using a fiberoptic transmission cell and a simple filter spectrometer. **Journal of Chromatography A**, v. 785, n. 1–2, p. 269–277, 1997.
- 102 US ENVIRONMENTAL PROTECTION AGENCY. **Method 5021A**: volatile organic compounds in various sample matrices using equilibrium headspace analysis. Washington DC, 2003. 25 p.
- 103 US ENVIRONMENTAL PROTECTION AGENCY. **Method 5030C**: purge-and-trap for aqueous samples. Washington DC, 2003. 28 p.
- 104 LLOMPART, M. et al. Headspace solid phase microextraction (HSSPME) for the determination of volatile and semivolatile pollutants in soils. **Talanta**, v. 48, n. 2, p. 451–459, 1999.

105 CAM, D. et al. Determination of polycyclic aromatic hydrocarbons in sediment using solid-phase microextraction with gas chromatography-mass spectrometry. **Journal Of Chromatographic Science**, v. 38, n. 2, p. 55–60, 2000.

106 HAVENGA, W.J.; ROHWER, E.R. Chemical characterization and screening of hydrocarbon pollution in industrial soils by headspace solid-phase microextraction. **Journal of Chromatography A**, v. 848, n. 1–2, p. 279–295, 1999.

107 MANURA, J. J.; OVERTON, S. **Note 39: comparison of sensitivity of headspace GC, purge and trap thermal desorption and direct thermal extraction techniques for volatile organics**. 1999. Disponível em: <<http://www.sisweb.com/referenc/applnote/app-39.htm>>. Acesso em: 9 jul. 2018.

108 CHANG, I.L.; JI, Z. (2000). Optimized analysis of gasoline (BTEX) in water and soil using gc / fid with purge and trap application note. **Agilent Technologies Application Note** 228-324, 6 p.

109 US ENVIRONMENTAL PROTECTION AGENCY. **Method 8020A**: Aromatic volatile organics by gas chromatography. Washington DC, 1994. 14 p.

110 BECKER, R. et al. Quantification of total petrol hydrocarbons (TPH) in soil by IR-spectrometry and gas chromatography - conclusions from three proficiency testing rounds. **Accreditation and Quality Assurance**, v. 7, n. 7, p. 286–289, 2002.

111 BICCHI, C. et al. Direct resistively heated column gas chromatography (Ultrafast module-GC) for high-speed analysis of essential oils of differing complexities. **Journal of Chromatography A**, v. 1024, n. 1–2, p. 195–207, 2004.

112 SEQUINEL, R. et al. Cromatografia gasosa ultrarrápida: uma visão geral sobre parâmetros, instrumentação e aplicações. **Química Nova**, v. 33, n. 10, p. 2226–2232, 2010.

113 DEURSEN, M.M. Van et al. Evaluation of time-of-flight mass spectrometric detection for fast gas chromatography. **Journal of Chromatography A**, v. 878, n. 2, p. 205–213, 2000.

114 MIRANDA, N.T. et al. Ultrafast gas chromatography method with direct injection for the quantitative determination of benzene, toluene, ethylbenzene, and xylenes in commercial gasoline. **Journal of Separation Science**, v. 40, n. 7, p. 1508–1515, 2017.

115 NESPECA, M.G. et al. Rapid and sensitive method for detecting adulterants in gasoline using ultra-fast gas chromatography and Partial Least Square Discriminant Analysis. **Fuel**, v. 215, 2018.

116 SILVERSTEIN, R. M.; WEBSTER, F. X.; KIEMLE, D. J. Espectrometria no infravermelho. In: _____. **Identificação espectrométrica de compostos orgânicos**. 7. ed. Rio de Janeiro: LTC, 2010. Cap. 2, p. 70-104.

117 IDODO-UMEH, G.; OGBEIBU, A.E. Bioaccumulation of the Heavy Metals in Cassava Tubers and Plantain Fruits Grown in Soils Impacted with Petroleum and Non-Petroleum Activities. **Research Journal of Environmental Sciences**, v. 4, n. 1, p. 33–41, 2010.

- 118 LISTE, H.H.; FELGENTREU, D. Crop growth, culturable bacteria, and degradation of petrol hydrocarbons (PHCs) in a long-term contaminated field soil. **Applied Soil Ecology**, v. 31, n. 1–2, p. 43–52, 2006.
- 119 VERSHININ, V.I.; PETROV, S. V. The estimation of total petroleum hydrocarbons content in waste water by IR spectrometry with multivariate calibrations. *Talanta*, v. 148, p. 163–169, 2016.
- 120 NESPECA, M.G.; PIASSALONGA, G.B.; DE OLIVEIRA, J.E. Infrared spectroscopy and multivariate methods as a tool for identification and quantification of fuels and lubricant oils in soil. **Environmental Monitoring and Assessment**, v. 190, n. 2, 2018.
- 121 BURNS, D.A.; CIURCZAK, E.W. (2009). **Handbook of near-infrared analysis**, 3rd ed. Boca Raton: CRC Press, 2009. 836 p.
- 122 SIESLER, H.W. et al. **Near-infrared spectroscopy: Principles, instruments, applications**. Weinheim: Wiley-VCH. 2002, 355 p.
- 123 GEMPERLINE, P. **Practical guide to chemometrics**. 2nd ed. Boca Raton: CRC Press, 2006. 520 p.
- 124 SORIANO-DISLA, J.M. The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. **Applied Spectroscopy Reviews**, v. 49, n. 2, p. 139–186, 2014.
- 125 ROUDIER, P. et al. Evaluation of two methods to eliminate the effect of water from soil vis–NIR spectra for predictions of organic carbon. **Geoderma**, v. 296, p. 98–107, 2017.
- 126 ROBERTS, J.J.; COZZOLINO, D. Wet or dry? The effect of sample characteristics on the determination of soil properties by near infrared spectroscopy. **Trends in Analytical Chemistry**, v. 83, p. 25–30, 2016.
- 127 JIANG, Q. et al. Estimation of soil organic carbon and total nitrogen in different soil layers using VNIR spectroscopy: Effects of spiking on model applicability. **Geoderma**, v. 293, p. 54–63, 2017.
- 128 JI, W. et al. Accounting for the effects of water and the environment on proximally sensed vis-NIR soil spectra and their calibrations. **European Journal of Soil Science**, v. 66, n. 3, p. 555–565, 2015.
- 129 DOTTO, A.C. et al. A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. **Geoderma**, v. 314, n. 11, p. 262–274, 2018.
- 130 VASQUES, G.M. et al. Soil classification using visible/near-infrared diffuse reflectance spectra from multiple depths. **Geoderma**, v. 223–225, n. 1, p. 73–78, 2014.

- 131 WIJEWARDANE, N. K.; GE, Y; MORGAN, C. L. S., N.K.. Prediction of soil organic and inorganic carbon at different moisture contents with dry ground VNIR: a comparative study of different approaches. **European Journal of Soil Science**, v. 67, n. 5, p. 605–615, 2016.
- 132 LUCE, M. St. et al. Rapid determination of soil organic matter quality indicators using visible near infrared reflectance spectroscopy. **Geoderma**, v. 232–234, p. 449–458, 2014.
- 133 CLAIROTTE, M. et al. National calibration of soil organic carbon concentration using diffuse infrared reflectance spectroscopy. **Geoderma**, v. 276, p. 41–52, 2016.
- 134 WILLIAMS, P.; NORRIS, K. **Near Infrared Technology in the Agricultural and Food Industries**. St. Paul: American Association of Cereal Chemist, 2001. 296 p.
- 135 ACKERSON, J.P. et al. Penetrometer-mounted VisNIR spectroscopy: Application of EPO-PLS to in situ VisNIR spectra. **Geoderma**, v. 286, p. 131–138, 2017.
- 136 STALLARD, B.R. et al. Near-IR reflectance spectroscopy for the determination of motor oil contamination in sandy loam. **Applied Spectroscopy**, v. 50, n. 3, p. 334–338, 1996.
- 137 ZWANZIGER, H.W.; FÖRSTER, H. Near infrared spectroscopy of fuel contaminated sand and soil. I. preliminary results and calibration study. **Journal of Near Infrared Spectroscopy**, v. 6, n. 1–4, p. 189–197, 1998.
- 138 MALLEY, D.F. et al. Analysis of diesel fuel contamination in soils by near-infrared reflectance spectrometry and solid phase microextraction-gas chromatography. **Soil and Sediment Contamination**, v. 8, n. 4, p. 481–489, 1999.
- 139 OZAKI, Y. Near-infrared spectroscopy--its versatility in analytical chemistry. **Analytical sciences : the international journal of the Japan Society for Analytical Chemistry**, v. 28, n. 6, p. 545–63, 2012.
- 140 CHAKRABORTY, S. et al. Rapid Identification of Oil-Contaminated Soils Using Visible Near-Infrared Diffuse Reflectance Spectroscopy. **Journal of Environment Quality**, v. 39, n. 4, p. 1378, 2010.
- 141 SORAK, D. et al. New developments and applications of handheld raman, mid-infrared, and near-infrared spectrometers. **Applied Spectroscopy Reviews**, v. 47, n. 2, p. 83–115, 2012.
- 142 CHAKRABORTY, S. et al. Spectral reflectance variability from soil physicochemical properties in oil contaminated soils. **Geoderma**, v. 177–178, p. 80–89, 2012.
- 143 FORRESTER, S. et al. An infrared spectroscopic test for total petroleum hydrocarbon (TPH) contamination in soils. **Proc. 19th World Congress of Soil Science**, 1–6 August 2010, Brisbane, Australia, n. 8, p. 13–16, 2010.
- 144 NAES, T. et al. **A user-friendly guide to multivariate calibration and classification**. Chichester: NIR Publications, 2002. 354 p.

- 145 VOHLAND, M. et al. Quantification of soil properties with hyperspectral data: Selecting spectral variables with different methods to improve accuracies and analyze prediction mechanisms. **Remote Sensing**, v. 9, n. 11, p. 1–24, 2017.
- 146 KUANG, B.; MOUAZEN, A.M. Non-biased prediction of soil organic carbon and total nitrogen with vis-NIR spectroscopy, as affected by soil moisture content and texture. **Biosystems Engineering**, v. 114, n. 3, p. 249–258, 2013.
- 147 FRANCESCHINI, M.H.D. et al. Effects of external factors on soil reflectance measured on-the-go and assessment of potential spectral correction through orthogonalisation and standardisation procedures. **Soil and Tillage Research**, v. 177, n. October 2017, p. 19–36, 2018.
- 148 MINASNY, B. et al. Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. **Geoderma**, v. 167–168, p. 118–124, 2011.
- 149 RAMIREZ-LOPEZ, L. et al. Sampling optimal calibration sets in soil infrared spectroscopy. **Geoderma**, v. 226–227, n. 1, p. 140–150, 2014.
- 150 AN, X. et al. Eliminating the interference of soil moisture and particle size on predicting soil total nitrogen content using a NIRS-based portable detector. **Computers and Electronics in Agriculture**, v. 112, p. 47–53, 2015.
- 151 NOCITA, M. et al. Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy. **Geoderma**, v. 199, p. 37–42, 2013.
- 152 EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Manual de Métodos de Análise de Solo**. Rio de Janeiro: Empresa brasileira de pesquisa agropecuária, 1997. 212 p.
- 153 BEN-DOR, E. et al. Reflectance measurements of soils in the laboratory: Standards and protocols. **Geoderma**, v. 245–246, p. 112–124, 2015.
- 154 VISCARRA, R.A. et al. Determining the composition of mineral-organic mixes using UV-vis-NIR diffuse reflectance spectroscopy. **Geoderma**, v. 137, n. 1–2, p. 70–82, 2006.
- 155 KUANG, B. et al. Comparison between artificial neural network and partial least squares for on-line visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content. **Soil and Tillage Research**, v. 146, n. PB, p. 243–252, 2015.
- 156 TERRA, F.S. et al. Spectral libraries for quantitative analyses of tropical Brazilian soils: Comparing vis-NIR and mid-IR reflectance data. **Geoderma**, v. 255–256, p. 81–93, 2015.
- 157 MORELLOS, A. et al. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. **Biosystems Engineering**, v. 152, p. 1–13, 2016.
- 158 WIJEWARDANE, N.K.; GE, Y.; MORGAN, C.L.S. Moisture insensitive prediction of soil properties from VNIR reflectance spectra based on external parameter orthogonalization. **Geoderma**, v. 267, p. 92–101, 2016.

159 FERREIRA, M.M.C. **Quimiometria - Conceitos, Métodos e Aplicações**. Campinas: Editora Unicamp, 2015. 496 p.

160 KUMAR, N. et al. Chemometrics tools used in analytical chemistry: An overview. **Talanta**, v. 123, p. 186–199, 2014.

161 EIGENVECTOR RESEARCH INC. **Advanced Preprocessing: Multivariate Filtering. 2013**. Disponível em: <http://wiki.eigenvector.com/index.php?title=Advanced_Preprocessing:_Multivariate_Filtering>. Acessado em: 15 mar. 2017.

162 GE, Y. et al. VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact. **Geoderma**, v. 221–222, p. 61–69, 2014.

163 ACKERSON, J.P. et al. Predicting clay content on field-moist intact tropical soils using a dried, ground VisNIR library with external parameter orthogonalization. **Geoderma**, v. 259, p. 196–204, 2015.

164 VOHLAND, M. et al. Determination of soil properties with visible to near- and mid-infrared spectroscopy: Effects of spectral variable selection. **Geoderma**, v. 223–225, n. 1, p. 88–96, 2014.

165 XIAOBO, Z. et al. Variables selection methods in near-infrared spectroscopy. **Analytica Chimica Acta**, v. 667, n. 1–2, p. 14–32, 2010.

166 MEHMOOD, T. et al. A review of variable selection methods in Partial Least Squares Regression. **Chemometrics and Intelligent Laboratory Systems**, v. 118, p. 62–69, 2012.

167 BRERETON, R.G.; LLOYD, G.R. Partial least squares discriminant analysis: Taking the magic away. **Journal of Chemometrics**, v. 28, n. 4, p. 213–225, 2014.

168 HAWKINS, D.M. The Problem of Overfitting. **Journal of Chemical Information and Computer Sciences**, v. 44, n. 1, p. 1–12, 2004.

169 ANIBAL, C. V. Di; Callao, M.P.; et al. 1H NMR variable selection approaches for classification. A case study: the determination of adulterated foodstuffs. **Talanta**, v. 86, n. 1, p. 316–323, 2011.

170 VALDERRAMA, P.; BRAGA, J.W.B. Estado da arte de figuras de mérito de calibração multivariada. **Química Nova**, v. 32, n. 5, p. 1278–1287, 2009.

171 BRO, R.; ANDERSEN, C.M. Theory of net analyte signal vectors in inverse regression. **Journal of Chemometrics**, v. 17, n. 12, p. 646–652, 2003.

172 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM D5769**: standard test method for determination of benzene, toluene, and total aromatics in finished gasolines by gas chromatography/mass spectrometry. West Conshohocken, 2015. 7p.

173 US ENVIRONMENTAL PROTECTION AGENCY. **Method 8000C**: determinative chromatographic separations. Washington DC, 2003. 66 p.

- 174 SOUSA, A.G. et al. A multivariate approach for high throughput pectin profiling by combining glycan microarrays with monoclonal antibodies. **Carbohydrate Research**, v. 409, p. 41–47, 2015.
- 175 SHRESTHA, S. et al. Classification of different tomato seed cultivars by multispectral visible-near infrared spectroscopy and chemometrics. **Journal of Spectral Imaging**, v. 5, n. 4, p. 1–9, 2016.
- 176 WISE, B.M. et al. Chemometrics tutorial for PLS toolbox and solo. Wenatchee: Eigenvector Research, Inc., 2006. 420 p.
- 177 INSTITUTO NACIONAL DE METROLOGIA. **Orientação sobre validação de métodos analíticos: DOQ-CGCRE-008**. 3ª ed. Rio de Janeiro: Coordenação Geral de Acreditação, 2010. 20 p.
- 178 AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA. **Resolução da Diretoria Colegiada - RDC No 166, de 24 de julho de 2017**. Brasília, 2017. 22 p.
- 179 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM E1655**: standard practices for infrared multivariate quantitative analysis. West Conshohocken, 2012. 29 p.
- 180 FLUMIGNAN, D.L. et al. Multivariate calibrations in gas chromatographic profiles for prediction of several physicochemical parameters of brazilian commercial gasoline. **Chemometrics and Intelligent Laboratory Systems**, v. 92, n. 1, p. 53–60, 2008.
- 181 LOMBORG, C.J. et al. Near infrared and acoustic chemometrics monitoring of volatile fatty acids and dry matter during co-digestion of manure and maize silage. **Bioresource Technology**, v. 100, n. 5, p. 1711–1719, 2009.
- 182 JONES, L.H.P.; HANDRECK, K.A. Silica in soils, plants, and animals. **Advances in Agronomy**, v. 19, n. C, p. 107–149, 1967.
- 183 WORKMAN, J.; WEYER, L. Practical Guide to Interpretive Near-Infrared Spectroscopy. 2007.
- 184 KNOTHE, G. et al. **Biodiesel handbook**. 2nd ed. Urbana: AOCS Press, 2010. 485 p.
- 185 RESTEK. **Solvent expansion calculator**. 2017. Disponível em: <http://www.restek.com/images/calcs/calc_backflash.htm>. Acesso em: 1 abr. 2017.
- 186 VISCOPIEDIA. **Ethanol**. 2017. Disponível em: <<http://www.viscopedia.com/viscosity-tables/substances/ethanol/>>. Acesso em: 1 abr. 2017.
- 187 DORTMUND DATA BANK. **Dynamic viscosity of hexadecane**. 2017. Disponível em: <http://www.ddbst.com/en/EED/PCP/VIS_C516.php>. Acesso em: 1 abr. 2017.
- 188 ROCHA, W.F.C. et al. Chemometric techniques applied for classification and quantification of binary biodiesel/diesel blends. **Analytical Letters**, v. 45, n. 16, p. 2398–2411, 2012.
- 189 MEHMOOD, T. et al. A review of variable selection methods in partial least squares regression. **Chemometrics and Intelligent Laboratory Systems**, v. 118, p. 62–69, 2012.

190 WANG, L.; Liu, E.; et al. Novel methodologies for automatically and simultaneously determining BTEX components using FTIR spectra. **Talanta**, v. 144, p. 1104–1110, 2015.

191 AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS. **Postos**. 2017. Disponível em: <<http://www.anp.gov.br/postos/consulta.asp>>. Acessado em: 15 jun. 2018.

ANEXO A – CARACTERÍSTICAS DOS COMBUSTÍVEIS E ÓLEOS LUBRIFICANTES

Tabela A1 - Propriedades físico-químicas das amostras de gasolina e óleo diesel selecionadas.

GASOLINA													
Amostra	Tipo	Densidade (kg/m ³)	Teor Alcoólico (% v/v)	Volatilidade				Octanagem		Composição			
				T 10 (°C)	T 50 (°C)	T 90 (°C)	Tf (°C)	RON	MON	olefinas (% v/v)	saturados (% v/v)	aromáticos (% v/v)	benzeno (% v/v)
G1	Premium	0,7354	25	55,5	74,2	116,6	189,9	97,5	88,4	0,0	50,2	25,3	0,00
G2	Comum	0,7325	25	54,5	68,1	129,3	176,2	98,9	87,0	4,1	54,1	16,7	0,65
G3	Comum	0,7192	25	49,4	64,2	123,1	171,0	95,2	83,6	6,5	61,8	7,1	0,64
G4	Comum	0,7357	26	52,3	68,9	145,3	185,8	98,0	83,9	9,7	53,7	9,8	0,62
G5	Comum	0,7420	27	55,7	71,9	155,8	200,8	96,7	84,1	6,5	55,8	12,0	0,65
G6	Premium	0,7226	26	55,2	72,8	114,2	176,9	97,4	87,0	0,0	65,5	10,1	0,90
G7	Aditivada	0,7426	25	55,8	72,5	161,9	200,6	99,8	84,3	12,6	47,8	13,4	0,60
G8	Premium	0,7173	25	55,4	72,5	116,6	193,1	98,6	87,3	0,0	69,1	6,5	0,00
G9	Comum	0,7326	25	54,0	69,4	146,9	191,5	96,8	84,1	7,3	56,8	10,9	0,64
G10	Aditivada	0,7415	25	55,9	71,8	152,2	196,8	97,3	83,9	9,7	52,1	12,9	0,62

ÓLEO DIESEL											
Amostra	Tipo	Densidade (kg/m ³)	Teor de Biodiesel (% v/v)	Volatilidade					Ponto de fulgor (°C)	Índice de Cetano	Enxofre total (% v/v)
				T 10 (°C)	T 50 (°C)	T 85 (°C)	T 90 (°C)	T 95 (°C)			
D1	S500	849	4,3	164	278	341	354	373	18	45	1375
D2	S500	845	5,5	170	281	352	369	392	44	46	320
D3	S500	844	6,0	158	261	351	375	385	43	43	303
D4	S10	840	6,2	191	263	319	329	346	51	47	5,9
D5	S500	845	6,5	165	267	351	371	391	44	44	313
D6	S10	843	6,6	206	277	333	343	362	53	45	<LQ*
D7	S500	852	6,8	176	286	351	370	378	42	45	419
D8	S10	836	7,0	184	257	317	331	354	49	45	<LQ*
D9	S10	843	7,2	207	281	333	343	361	56	45	<LQ*
D10	S10	835	7,4	179	255	313	326	346	47	45	<LQ*

* Abaixo do limite de quantificação.

Tabela A2 - Informações das amostras de óleo lubrificantes adquiridas.

Amostra	Marca Comercial	Grau de viscosidade SAE	Categoria API	Origem	Aplicação
O1	Extron Super	50	SF	mineral	veículos leves
O2	Lubrioil Super Max	40	SF/CF	mineral	veículos leves ^a
O3	Shell Rimula R3X	15w40	SI	mineral	veículos pesados
O4	VR Extra Mold	40	SF	mineral	veículos leves
O5	GT-Oil ATF	Tipo A		mineral	sistemas hidráulicos de veículos
O6 ^b	Ipiranga F1 Master	15W40	SI	semissintético	veículos leves
O7	GP Lubrax	20W50	SJ	mineral	motores a 4 tempos
O8	Lubrax	25W50	SL/SJ	mineral	veículos leves
O9	Ipiranga F1 Master Plus	10W30	SL	mineral	veículos leves
O10	Shell Helix HX5	15W50	SL/CF	semissintético	veículos leves ^a

^a Lubrificação de motores a álcool, gasolina, GNV e motores a diesel aspirados (serviços leves).

^b Óleo lubrificante usado após 10.000 km rodados em veículo leve.

ANEXO B – CROMATOGRAMAS

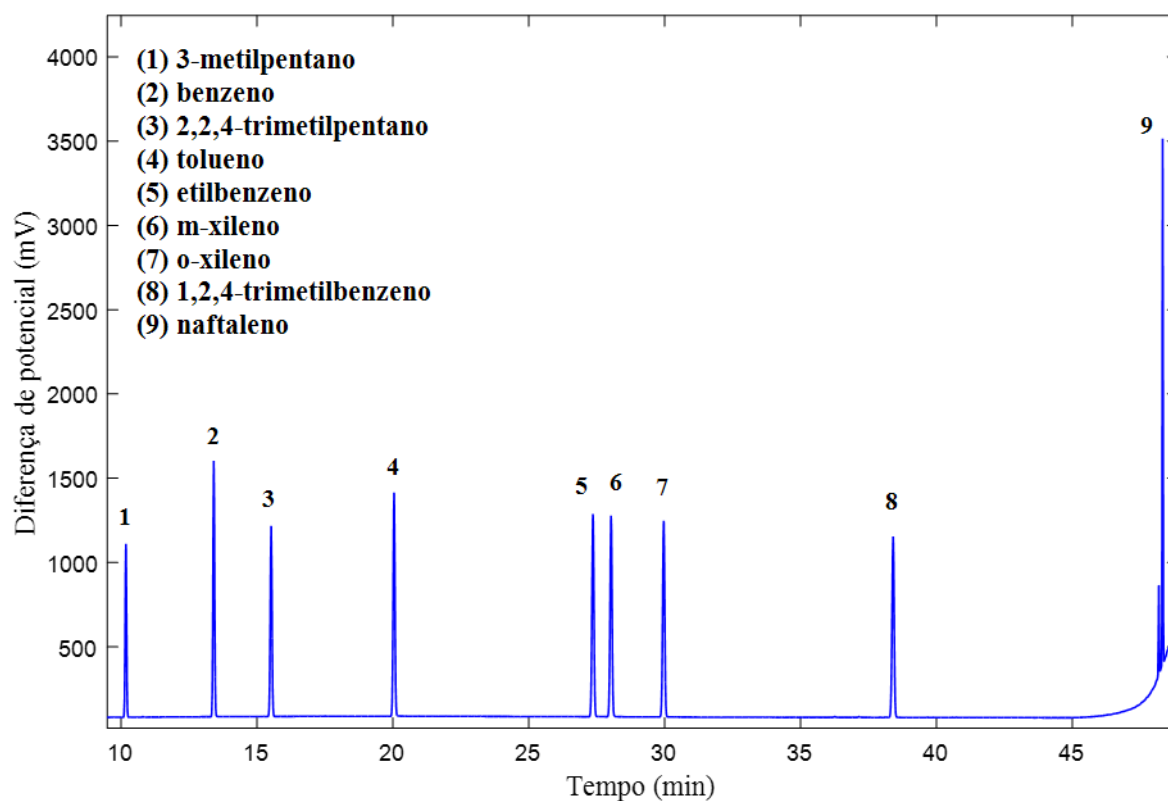


Figura B1 – Cromatogramas da análise do Mix GRO segundo o método ASTM D5769.

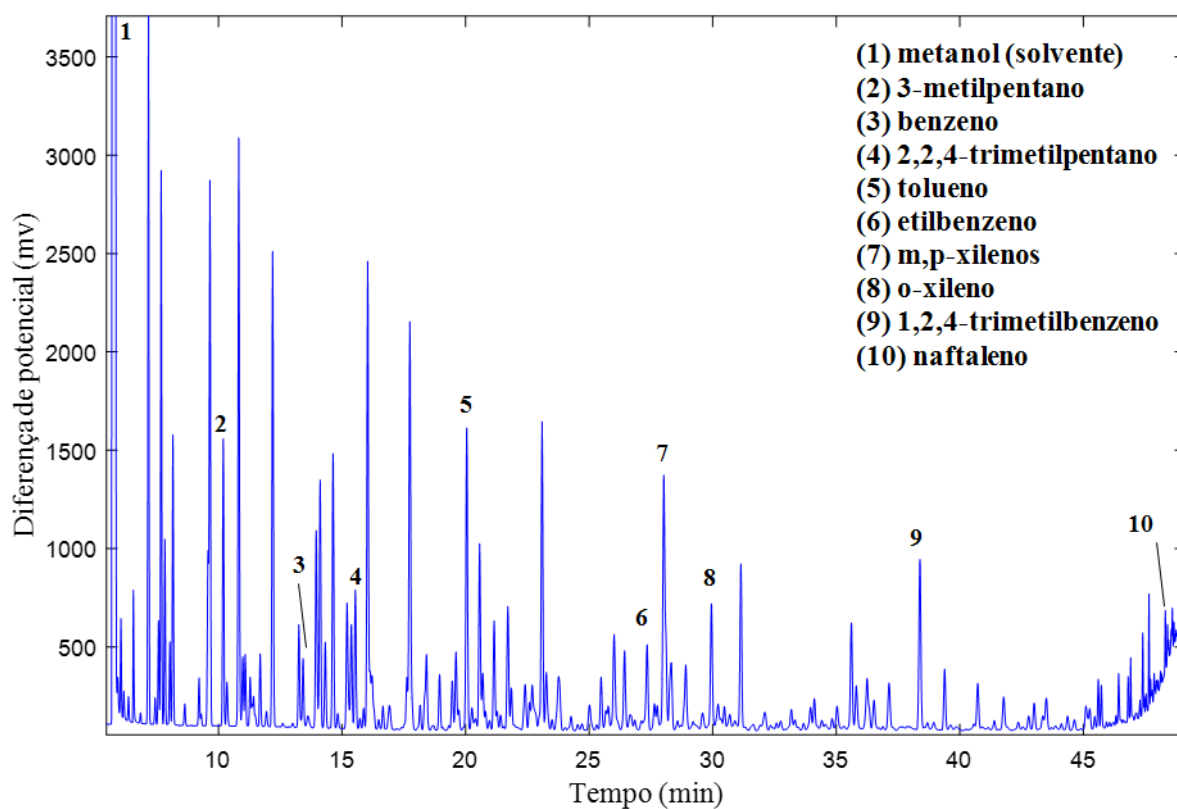


Figura B2 – Cromatograma da amostra de gasolina A (GasA).

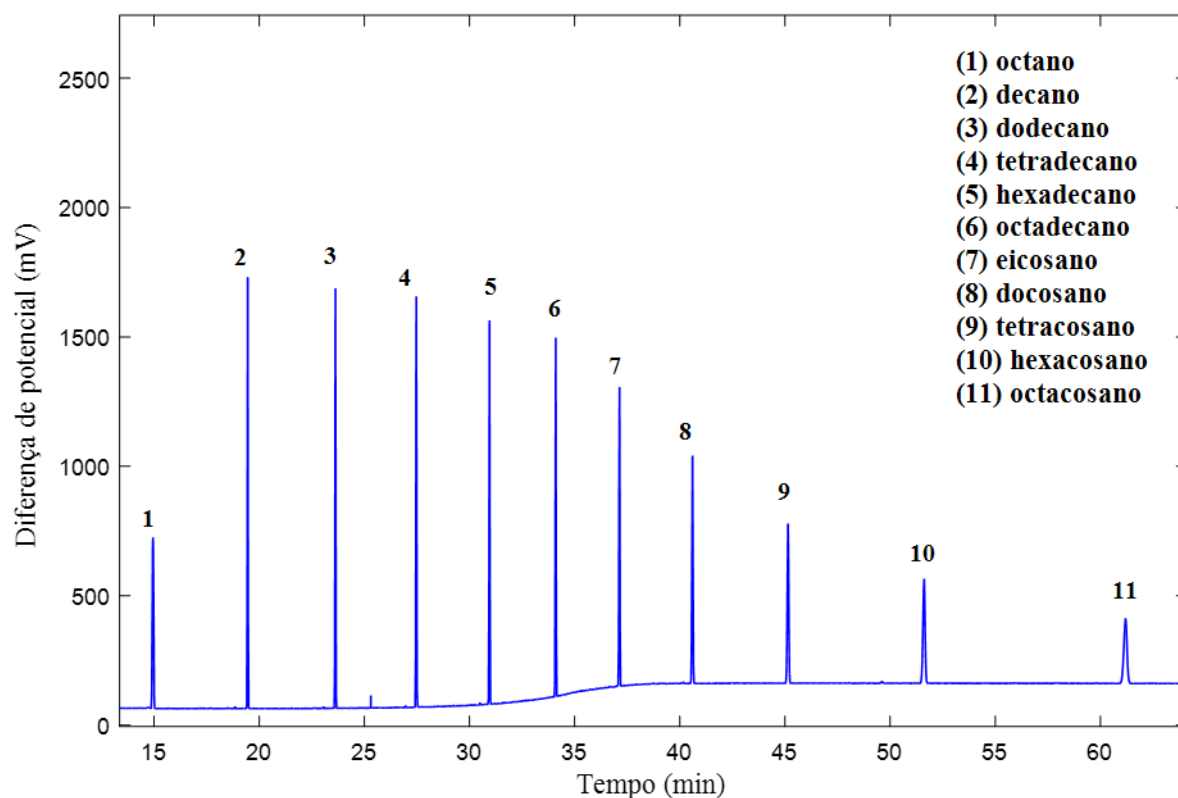


Figura B3 – Cromatograma do TPH Mix analisado pelo método ASTM D6209 adaptado.

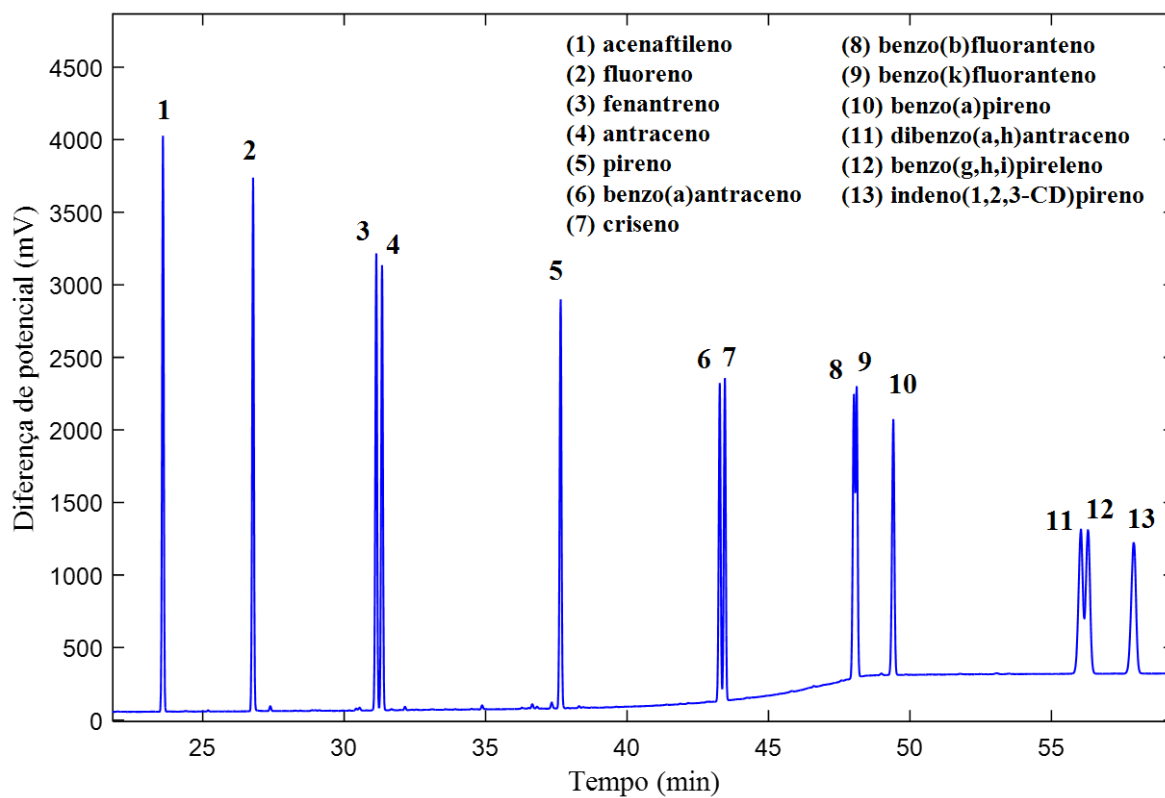


Figura B4 – Cromatograma do PAH Mix analisado pelo método ASTM D6209 adaptado.

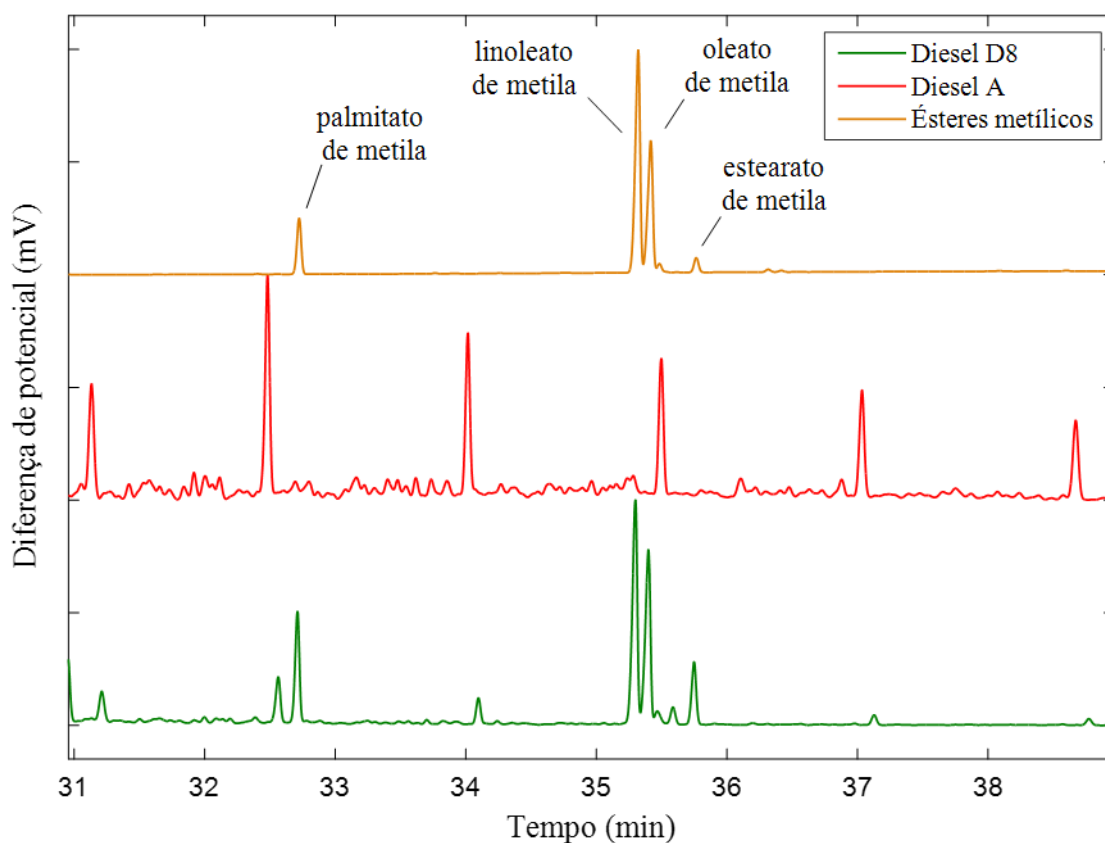


Figura B5 – Cromatogramas das amostras D8, DA e mistura de ésteres metílicos analisadas pelo método ASTM D6209 adaptado.

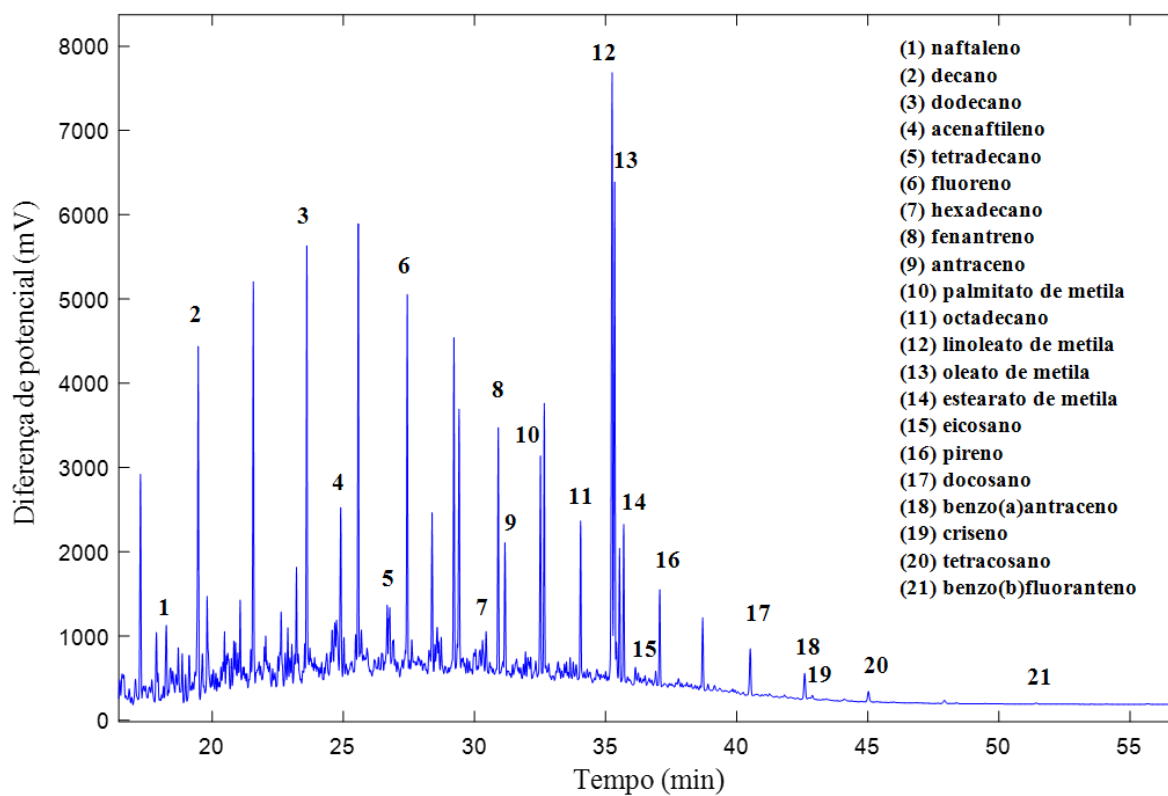


Figura B6 – Cromatograma da amostra D2 analisada pelo método ASTM D6209 adaptado.

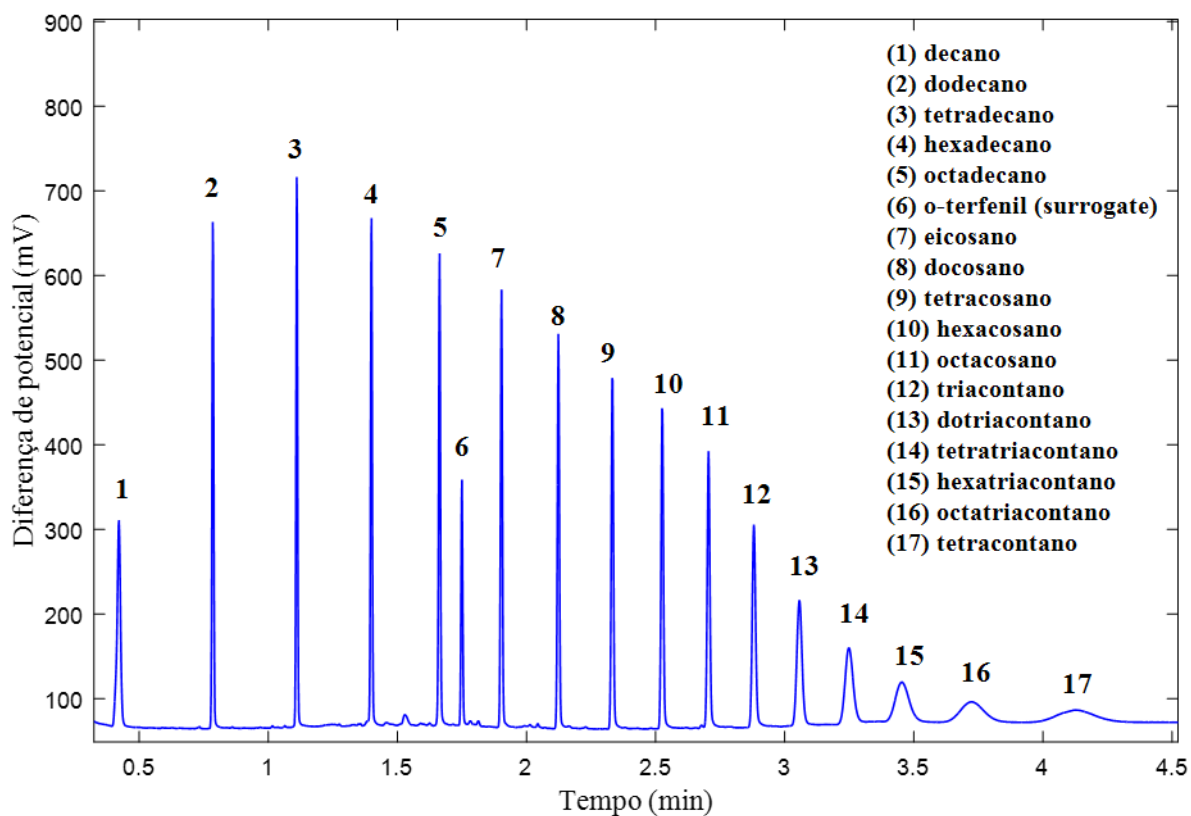


Figura B7 – Cromatograma da análise do TPH Mix pelo método UFGC-FID.

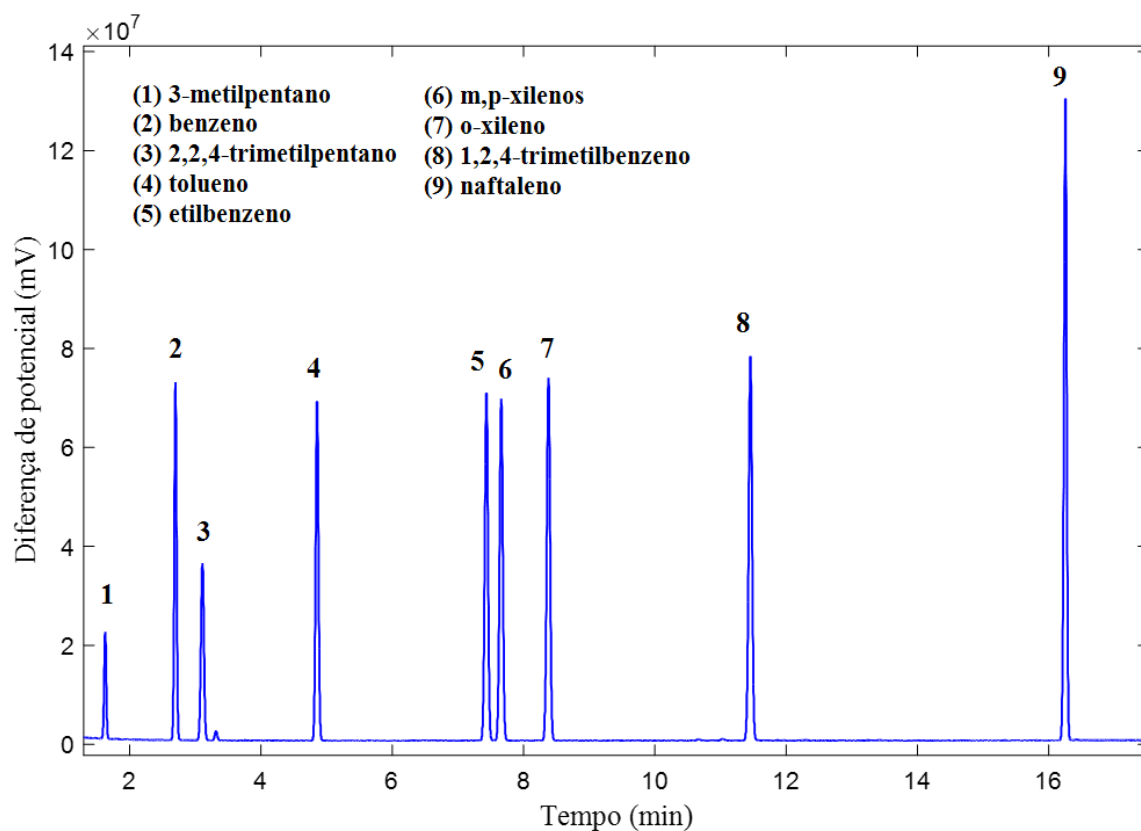


Figura B8 – Cromatograma do GRO Mix analisado pelo método EPA 8015.

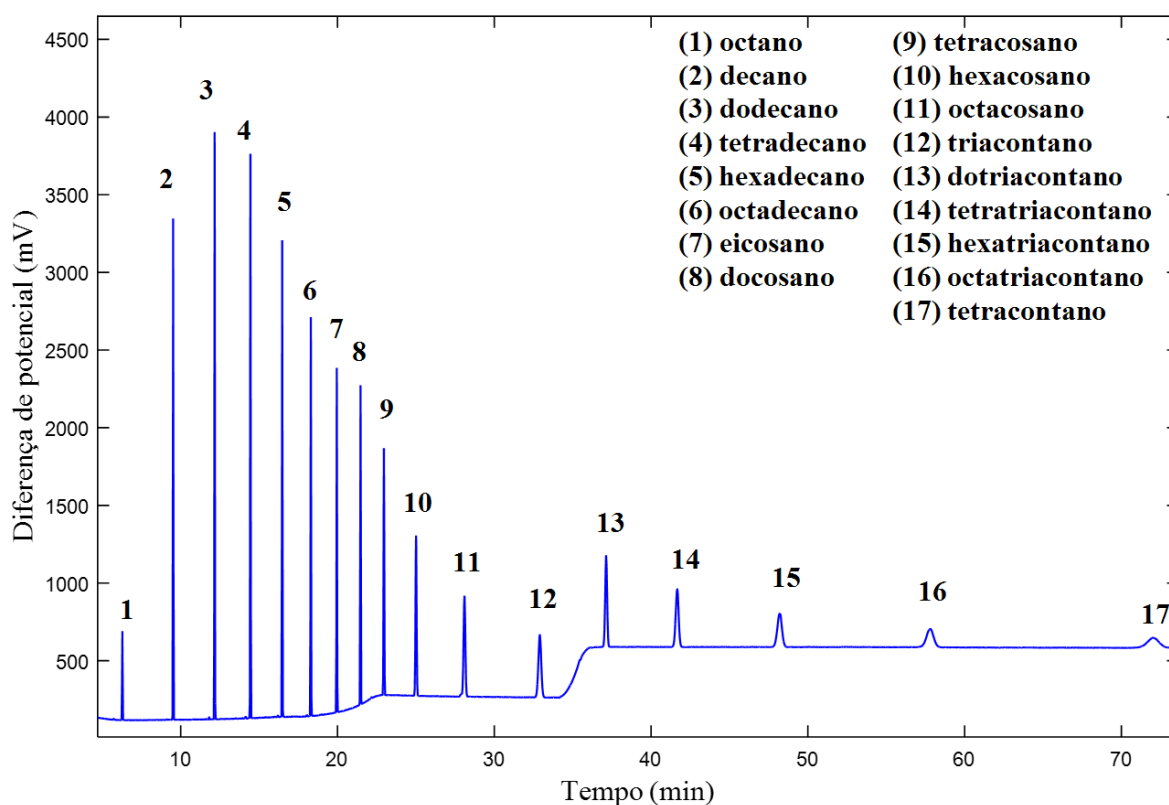


Figura B9 – Cromatograma do TPH Mix analisado pelo método EPA 8015 expandido.

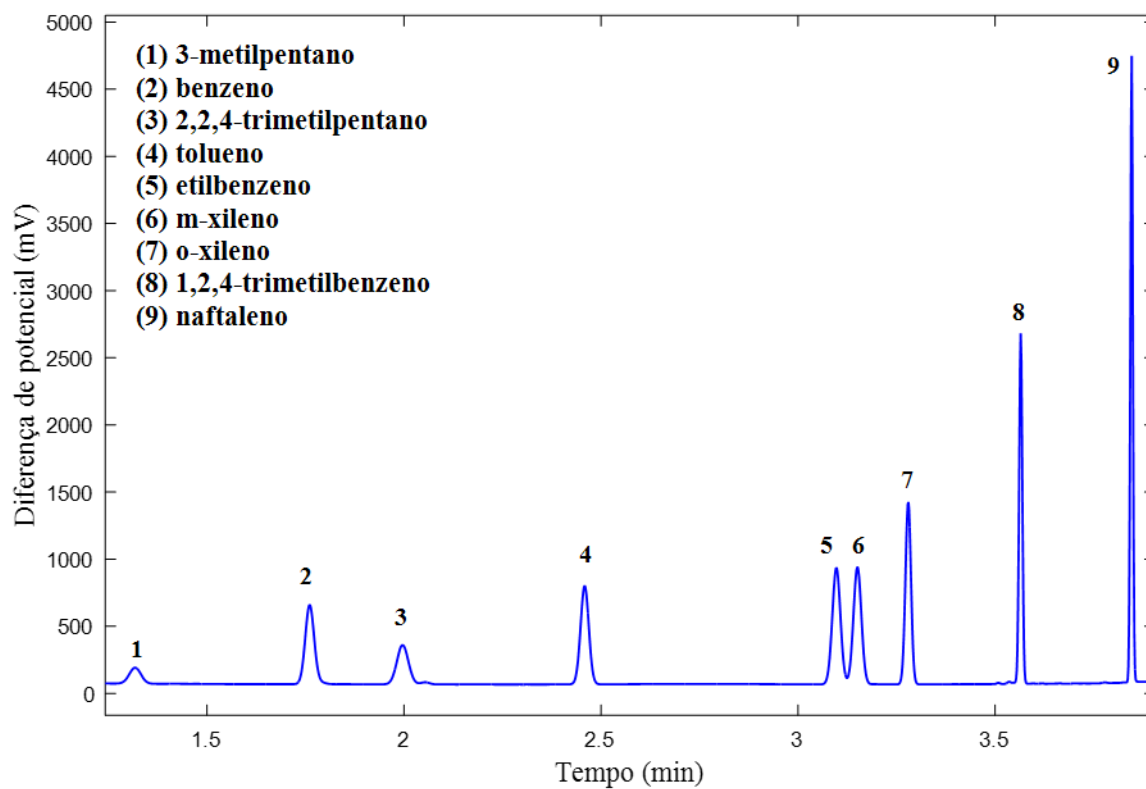


Figura B10 – Cromatograma da análise do GRO Mix pelo método Ultrarrápido otimizado.

ANEXO C – RESULTADOS QUIMIOMÉTRICOS

Tabela C1 – Parâmetros analíticos dos modelos PLS desenvolvidos com os espectros completos e com variáveis selecionadas. Os valores de RMSE estão na unidade de concentração mg kg⁻¹.

Análito	Benzeno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1911	390	108	3112	731	204	3112	709	183
Pré-processamento	MSC	SNV	SNV	SNV	Autoes.	Autoes	Autoes	Autoes., OSC	GLS, Autoes.	GLS, Autoes.
#V.L.	9	11	12	12	11	13	13	3	10	12
Variância X (%)	99,70	99,92	99,93	99,94	99,93	99,95	99,95	88,51	99,31	99,77
Variância y (%)	93,93	94,09	98,41	98,69	99,28	99,92	99,85	99,93	99,45	99,91
RMSEC	8,4	8,3	4,3	3,9	2,0	0,6	0,9	0,3	1,2	0,5
RMSECV	16,5	12,6	9,7	7,9	8,7	5,4	3,6	7,9	4,3	2,8
RMSEP	22,5	17,9	14,2	9,9	10,6	6,0	3,6	7,1	4,0	2,2
R ² cal	0,9393	0,9409	0,9841	0,9869	0,9928	0,9992	0,9985	0,9995	0,9945	0,9991
R ² CV	0,7722	0,8670	0,9214	0,9474	0,8594	0,9467	0,9768	0,7444	0,9288	0,9697
R ² val	0,5269	0,7066	0,8238	0,9137	0,8125	0,9443	0,9821	0,8182	0,9407	0,9818
Análito	Tolueno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1696	432	88	3112	674	116	3112	679	170
Pré-processamento	SNV	1 ^a der, OSC, C.M.	SNV	SNV	Autoes.	Autoes	Autoes	Autoes.	Autoes.	Autoes.
#V.L.	10	5	11	11	10	11	12	10	10	10
Variância X (%)	99,69	90,97	99,87	99,89	99,93	99,95	99,96	99,83	99,84	99,84
Variância y (%)	79,06	91,45	94,14	97,59	98,30	99,44	99,78	95,48	97,68	98,81
RMSEC	254,8	162,8	134,8	86,4	53,5	30,7	19,4	46,8	33,5	24,0
RMSECV	496,2	498,1	322,8	160,7	179,3	115,8	62,9	221,2	180,0	110,9
RMSEP	357,6	320,2	280,8	161,5	164,0	85,2	57,8	160,9	99,8	65,7
R ² cal	0,7906	0,9145	0,9414	0,9759	0,9830	0,9944	0,9978	0,9548	0,9768	0,9881
R ² CV	0,2690	0,3166	0,6642	0,9208	0,8125	0,9225	0,9768	0,1687	0,4253	0,7497
R ² val	0,5082	0,6304	0,6835	0,8929	0,8466	0,9572	0,9809	0,6601	0,8934	0,9431
Análito	Etilbenzeno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1911	375	105	3112	734	178	3112	682	146
Pré-processamento	MSC	MSC	MSC	MSC	Autoes.	Autoes	Autoes	Autoes.	Autoes.	Autoes.
#V.L.	7	8	11	11	11	11	10	10	12	11
Variância X (%)	99,47	99,88	99,94	99,95	99,93	99,94	99,93	99,87	99,90	99,89
Variância y (%)	80,91	81,74	96,08	96,39	98,89	99,12	99,29	97,45	99,71	99,44
RMSEC	15,0	14,7	6,8	6,5	2,6	2,3	2,1	2,7	0,9	1,3
RMSECV	22,6	18,5	12,5	11,3	14,0	9,9	6,2	10,1	7,0	4,0
RMSEP	23,8	22,4	17,6	11,7	12,5	6,1	4,6	7,5	4,3	2,9
R ² cal	0,8091	0,8174	0,9608	0,9639	0,9889	0,9912	0,9929	0,9745	0,9971	0,9944
R ² CV	0,5819	0,7116	0,8666	0,8924	0,6809	0,8417	0,9397	0,6453	0,8384	0,9497
R ² val	0,5664	0,6073	0,7589	0,8961	0,7757	0,9469	0,9705	0,7789	0,9341	0,9679

Continua.

Tabela C1 – Parâmetros analíticos dos modelos PLS desenvolvidos com os espectros completos e com variáveis selecionadas. Os valores de RMSE estão na unidade de concentração mg kg⁻¹.
(Continuação)

Análito	m,p-xilenos									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1925	373	81	3112	706	194	3112	808	175
Pré-processamento	MSC	1ª der, OSC	MSC	MSC	Autoes.	Autoes	Autoes	GLS, Autoes.	GLS, Autoes.	Autoes
#V.L.	7	9	12	11	11	10	12	11	11	13
Variância X (%)	99,46	99,76	99,94	99,94	99,93	99,93	99,94	98,40	99,47	99,91
Variância y (%)	78,00	91,75	97,56	96,27	98,80	98,31	99,62	99,21	99,51	99,81
RMSEC	56,7	34,7	18,9	23,4	9,3	11,0	5,2	5,1	4,0	2,5
RMSECV	85,6	70,5	47,5	38,1	52,2	35,2	20,6	30,8	20,6	13,7
RMSEP	85,3	78,6	62,8	45,6	47,0	26,8	16,5	39,0	27,0	15,6
R ² cal	0,7800	0,9175	0,9756	0,9627	0,9880	0,9831	0,9962	0,9921	0,9951	0,9981
R ² CV	0,5156	0,6649	0,8459	0,9005	0,6231	0,8342	0,9460	0,7157	0,8788	0,9446
R ² val	0,5539	0,6543	0,7445	0,8607	0,7481	0,9185	0,9678	0,5858	0,8395	0,9510
Análito	o-xileno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1904	352	81	3112	720	183	3112	677	152
Pré-processamento	MSC	SNV	MSC	MSC	Autoes.	Autoes	Autoes	Autoes.	Autoes.	Autoes
#V.L.	7	11	11	11	11	12	11	10	13	13
Variância X (%)	99,46	99,91	99,94	99,95	99,93	99,94	99,94	99,86	99,90	99,91
Variância y (%)	78,17	90,61	96,23	96,47	98,85	99,61	99,23	97,52	99,83	99,75
RMSEC	23,3	15,3	9,7	9,4	3,8	2,2	3,1	3,7	1,0	1,2
RMSECV	34,7	26,1	20,4	15,2	21,4	15,5	8,8	12,3	8,4	4,6
RMSEP	35,3	33,4	25,4	16,8	19,6	10,7	7,4	15,3	10,7	5,3
R ² cal	0,7817	0,9061	0,9623	0,9647	0,9885	0,9961	0,9923	0,9752	0,9983	0,9975
R ² CV	0,5308	0,7283	0,8337	0,9073	0,6285	0,8054	0,9411	0,7315	0,8797	0,9623
R ² val	0,5510	0,6089	0,7563	0,8892	0,7448	0,9239	0,9643	0,6302	0,8608	0,9670
Análito	Naftaleno (GRO)									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1932	430	120	3112	735	181	3112	693	140
Pré-processamento	MSC	MSC	MSC	MSC	Autoes.	Autoes	Autoes	Autoes.	Autoes.	Autoes
#V.L.	7	7	12	11	10	10	12	11	13	12
Variância X (%)	99,48	99,74	99,93	99,92	99,93	99,94	99,93	99,86	99,90	99,90
Variância y (%)	82,80	75,08	98,17	97,77	98,01	98,80	99,76	99,34	99,93	99,83
RMSEC	4,0	4,8	1,3	1,4	1,0	0,8	0,4	0,4	0,1	0,2
RMSECV	5,7	5,4	3,8	2,8	3,1	2,6	1,7	2,1	1,5	0,9
RMSEP	5,6	5,4	4,0	2,6	3,7	2,5	1,3	2,7	1,8	0,9
R ² cal	0,8280	0,7508	0,9817	0,9777	0,9801	0,9880	0,9976	0,9934	0,9993	0,9983
R ² CV	0,6498	0,6805	0,8402	0,9164	0,8211	0,8736	0,9464	0,7992	0,8972	0,9656
R ² val	0,7128	0,7314	0,8484	0,9388	0,7509	0,8837	0,9671	0,6595	0,8482	0,9636

Continua.

Tabela C1 – Parâmetros analíticos dos modelos PLS desenvolvidos com os espectros completos e com variáveis selecionadas. Os valores de RMSE estão na unidade de concentração mg kg⁻¹.
(Continuação)

Analito	Naftaleno (DRO)									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1803	370	87	3112	751	169	3112	703	188
Pré-processamento	MSC	MSC	MSC	MSC	OSC	OSC	OSC	Autoes.	Autoes.	Autoes.
#V.L.	8	9	9	9	6	6	6	12	13	13
Variância X (%)	99,29	99,87	99,85	99,81	100,00	100,00	100,00	99,70	99,74	99,76
Variância y (%)	96,34	92,72	96,24	98,53	99,27	99,49	99,87	99,96	99,98	99,97
RMSEC	8,1	11,5	8,2	5,2	2,4	2,0	1,0	0,4	0,3	0,3
RMSECV	19,6	16,9	12,5	9,0	11,9	8,3	3,8	4,9	3,1	2,1
RMSEP	23,2	18,0	15,1	8,0	11,7	6,2	2,9	4,7	2,9	1,7
R ² cal	0,9634	0,9272	0,9624	0,9853	0,9927	0,9949	0,9987	0,9996	0,9998	0,9997
R ² CV	0,7903	0,8433	0,9133	0,9552	0,8343	0,9189	0,9832	0,9300	0,9721	0,9872
R ² val	0,6980	0,8191	0,8794	0,9654	0,8506	0,9566	0,9906	0,9416	0,9797	0,9922
Analito	Acenafileno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	2061	390	97	3112	673	175	3112	681	197
Pré-processamento	MSC	MSC	MSC	MSC	C.M.	OSC	Autoes., OSC	Autoes.	Autoes.	Autoes.
#V.L.	8	9	11	11	8	6	3	11	10	12
Variância X (%)	99,20	99,84	99,84	99,77	99,77	100,00	93,75	99,69	99,76	99,75
Variância y (%)	97,18	93,43	98,86	99,31	93,58	99,99	99,99	99,82	99,73	99,95
RMSEC	6,3	9,6	4,0	3,1	6,4	0,2	0,3	0,7	0,8	0,3
RMSECV	20,3	15,8	10,5	6,5	12,6	7,1	4,5	4,2	2,5	1,3
RMSEP	21,2	17,1	11,2	5,7	9,8	5,7	3,7	3,6	1,8	1,0
R ² cal	0,9718	0,9343	0,9886	0,9931	0,9358	0,9999	0,9999	0,9982	0,9973	0,9995
R ² CV	0,7073	0,8229	0,9215	0,9702	0,7593	0,9253	0,9692	0,9314	0,9765	0,9936
R ² val	0,6612	0,7942	0,9210	0,9756	0,8873	0,9618	0,9864	0,9587	0,9901	0,9967
Analito	Fluoreno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1868	508	143	3112	738	212	3112	765	188
Pré-processamento	MSC	MSC	MSC	MSC	OSC	OSC	OSC	OSC	OSC	OSC
#V.L.	9	9	10	12	6	3	3	7	6	6
Variância X (%)	99,23	99,85	99,83	99,82	99,77	99,99	99,99	100,00	100,00	100,00
Variância y (%)	98,38	92,95	98,45	99,64	93,64	99,18	99,61	95,72	99,50	99,95
RMSEC	6,7	14,0	6,6	3,2	4,7	3,3	2,2	4,0	0,7	0,5
RMSECV	27,0	20,6	13,6	9,6	17,1	11,6	8,9	6,3	4,5	3,5
RMSEP	31,4	21,8	15,6	8,3	14,3	7,6	5,6	7,2	4,9	2,8
R ² cal	0,9838	0,9295	0,9845	0,9964	0,9832	0,9918	0,9961	0,9662	0,9989	0,9995
R ² CV	0,7377	0,8480	0,9346	0,9670	0,7792	0,8981	0,9419	0,9169	0,9576	0,9748
R ² val	0,6333	0,8279	0,9158	0,9773	0,8378	0,9569	0,9749	0,9164	0,9619	0,9888

Continua.

Tabela C1 – Parâmetros analíticos dos modelos PLS desenvolvidos com os espectros completos e com variáveis selecionadas. Os valores de RMSE estão na unidade de concentração mg kg⁻¹.
(Continuação)

Análito	Fenantreno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1868	508	143	3112	714	201	3112	727	219
Pré-processamento	MSC	MSC	MSC	MSC	C.M.	OSC	OSC	Autoes.	Autoes.	Autoes.
#V.L.	9	9	10	12	8	3	5	8	11	13
Variância X (%)	99,23	99,85	99,83	99,82	99,76	99,99	100,00	99,60	99,71	99,76
Variância y (%)	98,38	92,95	98,45	99,64	93,30	99,89	99,82	96,37	99,82	99,93
RMSEC	6,7	14,0	6,6	3,2	7,3	0,9	1,2	3,5	0,8	0,5
RMSECV	27,0	20,6	13,6	9,6	12,7	9,4	6,1	6,2	3,9	2,6
RMSEP	31,4	21,8	15,6	8,3	11,7	7,9	4,6	4,7	3,0	2,4
R ² cal	0,9838	0,9295	0,9845	0,9964	0,9330	0,9989	0,9982	0,9637	0,9982	0,9993
R ² CV	0,7377	0,8480	0,9346	0,9670	0,8004	0,8899	0,9524	0,8874	0,9552	0,9813
R ² val	0,6333	0,8279	0,9158	0,9773	0,8012	0,9110	0,9720	0,9243	0,9685	0,9796
Análito	Antraceno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1775	442	126	3112	754	226	3112	688	185
Pré-processamento	MSC	MSC	MSC	MSC	C.M.	OSC	OSC	C.M.	Autoes.	Autoes.
#V.L.	9	8	9	9	9	4	3	7	13	13
Variância X (%)	99,29	99,82	99,81	99,78	99,79	100,00	99,99	99,59	99,76	99,74
Variância y (%)	97,28	88,23	97,02	98,74	97,34	99,96	99,61	96,66	99,95	99,94
RMSEC	4,2	8,7	4,4	2,9	2,6	0,3	1,0	1,9	0,2	0,3
RMSECV	14,4	11,8	7,9	5,6	6,7	4,4	3,1	3,0	2,1	1,5
RMSEP	12,1	9,1	7,4	4,5	6,7	4,8	3,1	3,0	1,5	1,3
R ² cal	0,9728	0,8823	0,9702	0,9874	0,9734	0,9996	0,9961	0,9666	0,9995	0,9994
R ² CV	0,6786	0,7870	0,9045	0,9533	0,8377	0,9307	0,9644	0,9213	0,9620	0,9803
R ² val	0,7786	0,8588	0,9327	0,9707	0,8590	0,9275	0,9722	0,9432	0,9859	0,9902
Análito	Pireno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1	GA2
#variáveis	3112	1868	433	121	3112	793	215	3112	738	231
Pré-processamento	MSC	MSC	MSC	MSC	C.M.	OSC	OSC	Autoes.	Autoes.	Autoes.
#V.L.	7	7	9	10	8	4	4	8	7	7
Variância X (%)	99,22	99,72	99,83	99,78	99,75	100,00	100,00	99,51	99,50	99,50
Variância y (%)	87,25	76,14	92,74	97,12	93,72	98,60	96,15	92,70	94,15	96,13
RMSEC	7,8	10,7	5,9	3,7	3,6	1,7	2,8	2,1	2,4	1,9
RMSECV	13,6	14,4	10,3	7,3	7,1	4,9	3,9	4,1	3,6	2,8
RMSEP	12,6	10,5	8,0	5,0	7,1	4,9	3,9	5,0	4,2	3,7
R ² cal	0,8725	0,7614	0,9274	0,9712	0,9372	0,9860	0,9615	0,9560	0,9415	0,9613
R ² CV	0,6179	0,5765	0,7792	0,8883	0,7579	0,8849	0,9262	0,8314	0,8709	0,9195
R ² val	0,6989	0,7974	0,8788	0,9528	0,7441	0,8924	0,9347	0,7741	0,8475	0,8823

Continua.

Tabela C1 – Parâmetros analíticos dos modelos PLS desenvolvidos com os espectros completos e com variáveis selecionadas. Os valores de RMSE estão na unidade de concentração mg kg⁻¹.
(Continuação)

Analito	Benzo(a)antraceno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Modelo	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1
#variáveis	3112	1911	450	143	3112	791	227	3112	708	148
Pré-processamento	MSC	MSC	MSC	MSC	Autoes.	OSC	OSC	OSC, C.M.	OSC, C.M.	OSC, C.M.
#V.L.	11	12	12	10	9	4	8	6	9	10
Variância X (%)	99,53	99,90	99,89	99,84	99,79	100,00	100,00	97,83	98,74	98,78
Variância y (%)	99,19	97,60	98,19	97,50	96,94	99,74	99,55	82,00	98,92	99,94
RMSEC	1,7	2,9	2,6	3,0	2,0	0,6	0,8	3,7	0,7	0,2
RMSECV	10,8	10,7	7,4	5,8	6,9	3,9	3,3	5,0	3,6	2,3
RMSEP	12,2	10,2	6,2	4,2	6,0	4,0	2,3	3,2	2,5	1,4
R ² cal	0,9919	0,9760	0,9819	0,9750	0,9694	0,9974	0,9955	0,8227	0,9930	0,9994
R ² CV	0,6828	0,6884	0,8483	0,9090	0,6448	0,8842	0,9192	0,6840	0,8300	0,9335
R ² val	0,6088	0,7180	0,8993	0,9553	0,8380	0,9236	0,9746	0,8855	0,9305	0,9791
Analito	Criseno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Modelo	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1
#variáveis	3112	1854	425	116	3112	746	225	3112	731	175
Pré-processamento	MSC	MSC	MSC	MSC	Autoes.	OSC	OSC	OSC, Autoes.	OSC, Autoes.	OSC, Autoes.
#V.L.	7	8	9	9	9	6	6	5	12	12
Variância X (%)	99,23	99,83	99,80	99,76	99,78	100,00	100,00	92,90	98,63	97,67
Variância y (%)	83,80	78,69	94,54	97,22	96,14	98,65	98,56	81,75	99,94	99,93
RMSEC	3,8	4,4	2,2	1,6	1,2	0,7	0,7	1,9	0,1	0,1
RMSECV	6,2	7,2	4,5	2,9	4,0	2,3	1,6	2,5	1,6	0,9
RMSEP	6,5	5,5	3,7	2,8	3,2	2,0	1,2	2,0	1,5	0,7
R ² cal	0,8380	0,7869	0,9454	0,9722	0,9614	0,9865	0,9856	0,8185	0,9994	0,9992
R ² CV	0,5837	0,4598	0,7793	0,9059	0,5833	0,8711	0,9367	0,6689	0,8699	0,9552
R ² val	0,6412	0,7321	0,8898	0,9358	0,7361	0,8936	0,9609	0,8305	0,9160	0,9773
Analito	Benzo(b)fluoranteno									
	Solos fortificados				Soluções em hexano			Soluções em etanol		
	Modelo	Completo	Manual	GA1	GA2	Completo	GA1	GA2	Completo	GA1
#variáveis	3112	1832	411	100	3112	707	217	3112	698	133
Pré-processamento	MSC	MSC	MSC	MSC	Autoes.	OSC	OSC	GLSW	GLSW	GLSW
#V.L.	9	11	10	10	9	6	6	5	11	11
Variância X (%)	99,48	99,89	99,86	99,84	99,78	100,00	100,00	100,00	100,00	100,00
Variância y (%)	91,17	92,75	93,30	95,96	96,12	99,67	99,72	73,85	99,57	99,41
RMSEC	3,6	3,2	3,1	2,4	1,5	0,4	0,4	2,9	0,4	0,4
RMSECV	8,7	8,5	6,2	4,4	4,6	2,9	1,8	3,5	2,3	1,3
RMSEP	9,1	6,9	5,9	4,5	4,1	3,1	1,7	2,6	1,7	0,8
R ² cal	0,9117	0,9275	0,9330	0,9596	0,9612	0,9967	0,9972	0,7385	0,9957	0,9941
R ² CV	0,4947	0,5271	0,7338	0,8666	0,6299	0,8581	0,9450	0,6327	0,8413	0,9536
R ² val	0,5243	0,7196	0,7959	0,8915	0,7003	0,8359	0,9487	0,8040	0,9221	0,9801

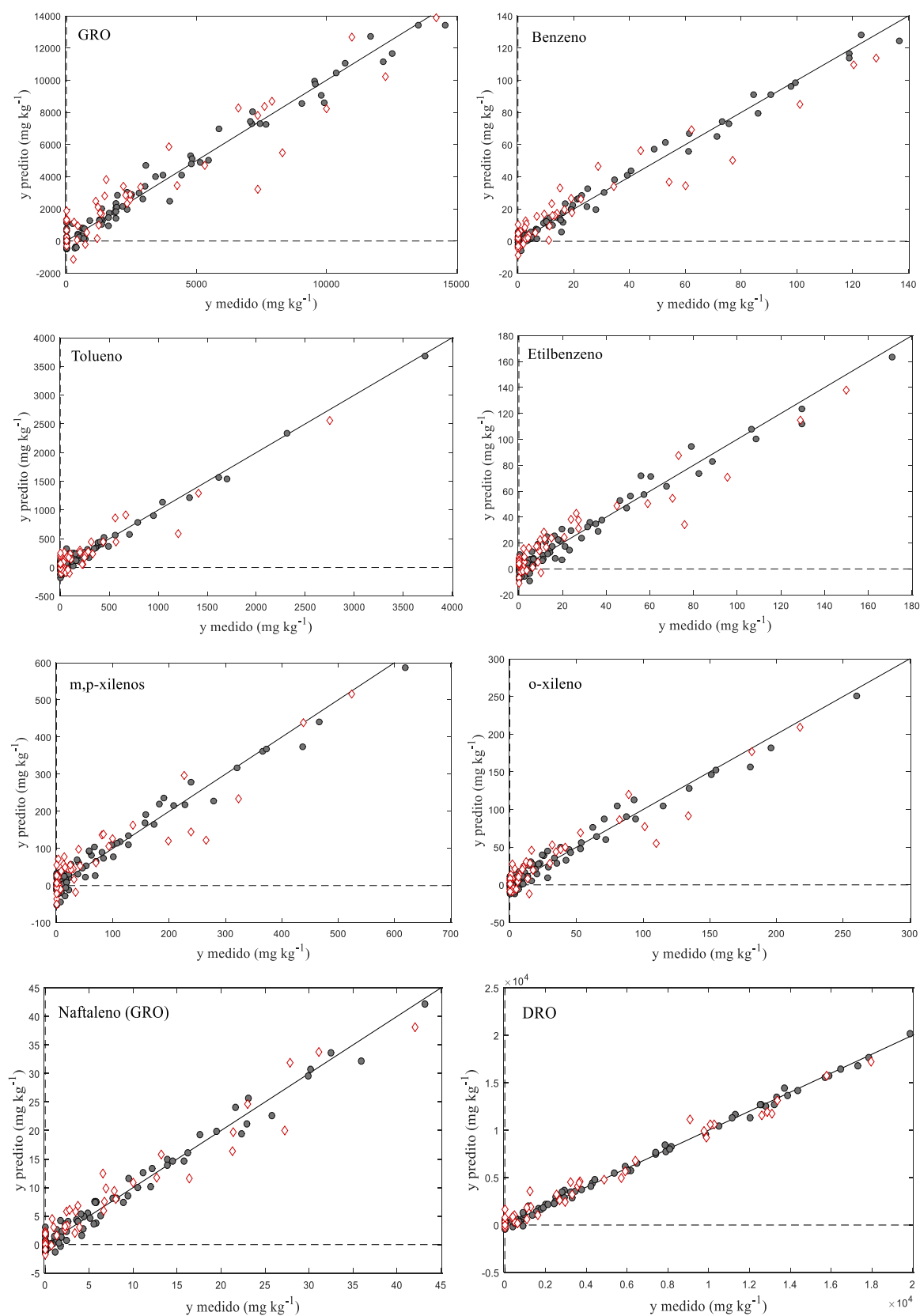


Figura C1.1 – Gráficos de ajuste dos modelos PLS para quantificação das frações TPH e analitos individuais determinados pela análise NIR sem preparo de amostra. Os círculos representam as amostras de calibração, os losangos as amostras de validação externa e a linha preta o ajuste perfeito.

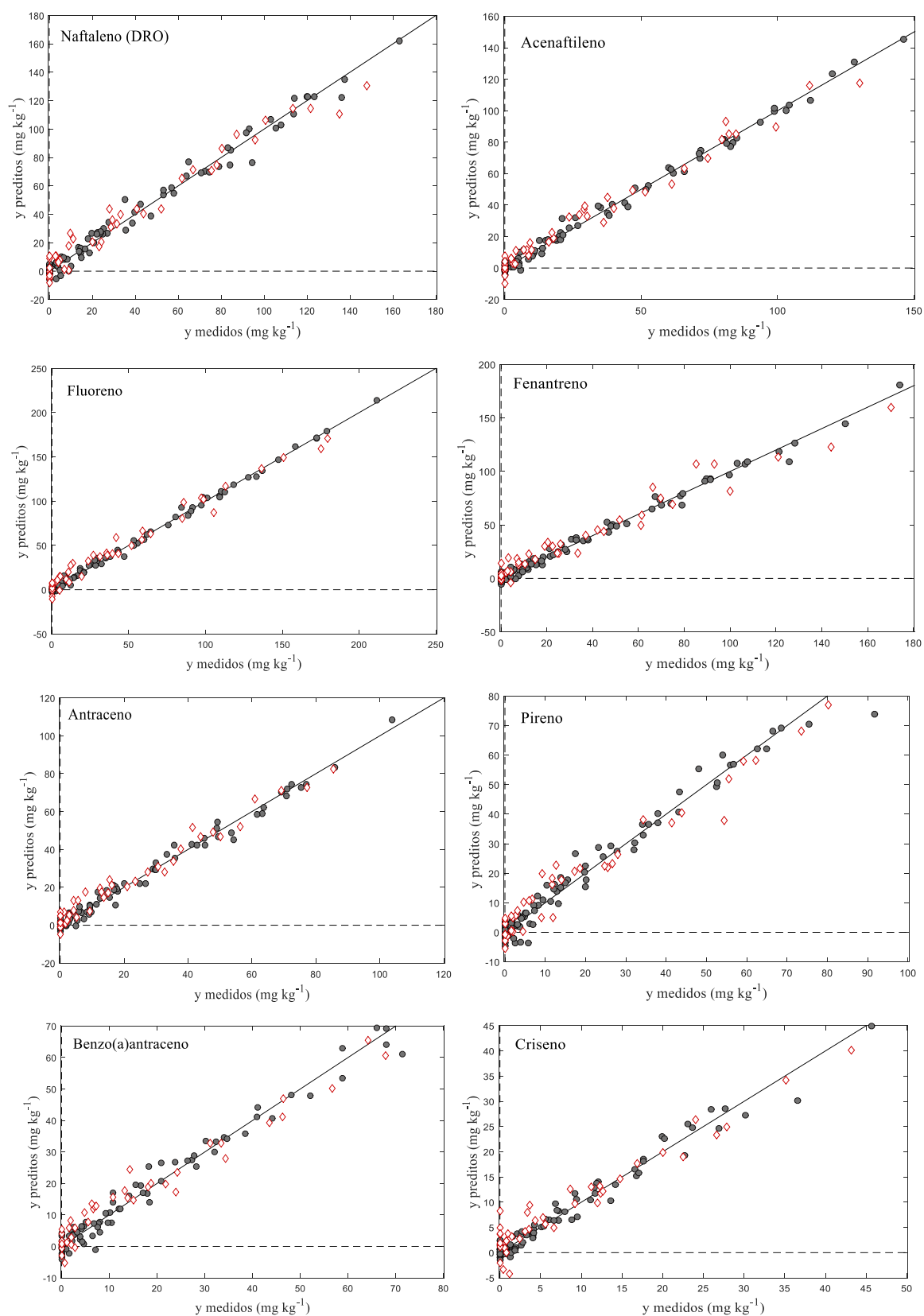


Figura C1.2 – Gráficos de ajuste dos modelos PLS para quantificação das frações TPH e analitos individuais determinados pela análise NIR sem preparo de amostra. Os círculos representam as amostras de calibração, os losangos as amostras de validação externa e a linha preta o ajuste perfeito.

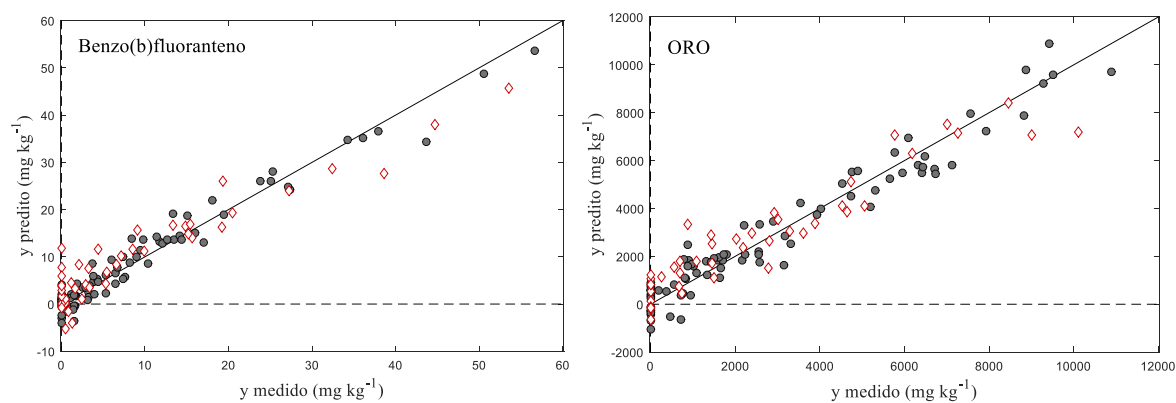


Figura C1.3 – Gráficos de ajuste dos modelos PLS para quantificação das frações TPH e analitos individuais determinados pela análise NIR sem preparo de amostra. Os círculos representam as amostras de calibração, os losangos as amostras de validação externa e a linha preta o ajuste perfeito.

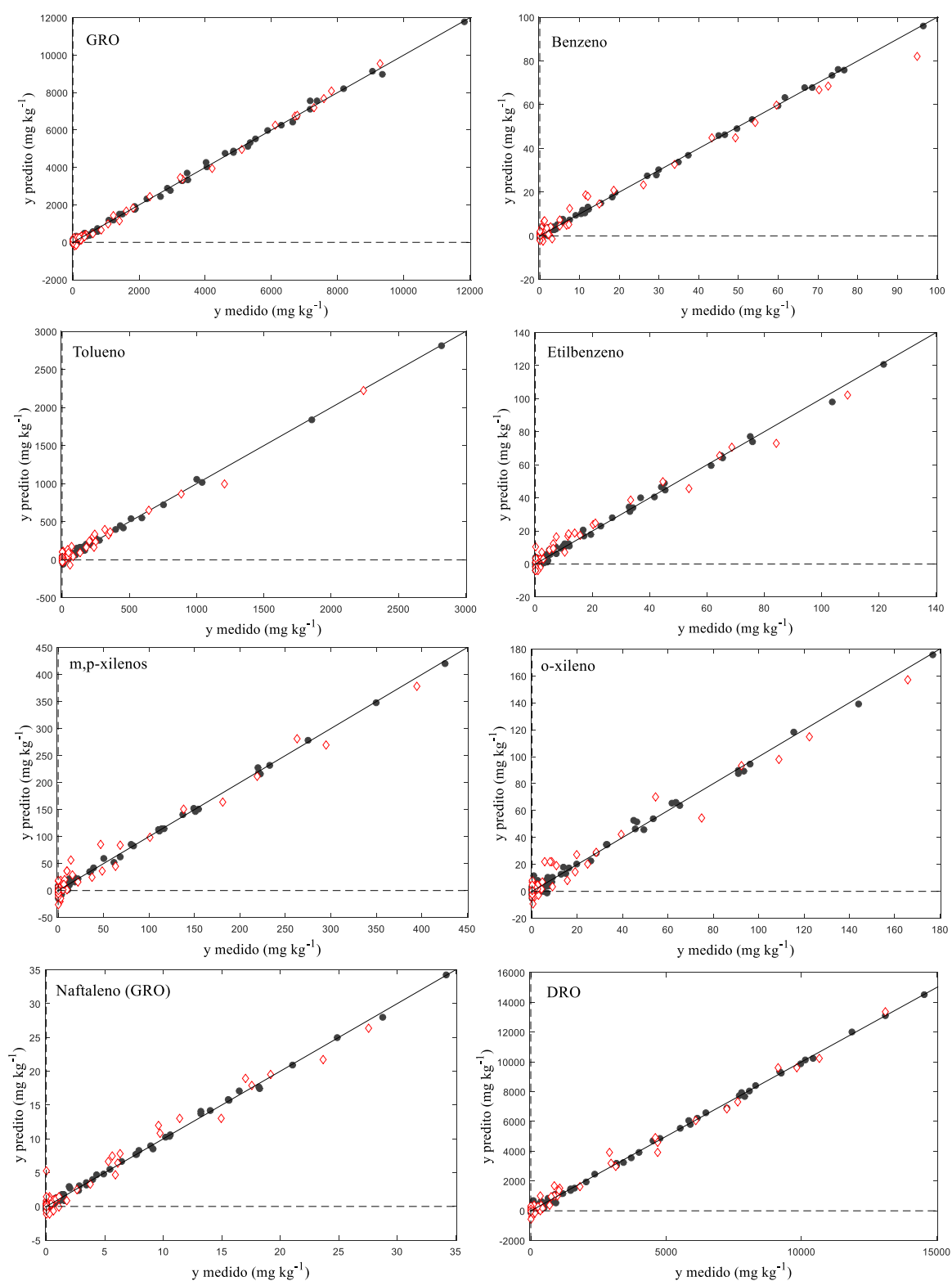


Figura C2.1 – Gráficos de ajuste dos modelos PLS para quantificação das frações TPH e analitos individuais determinados pela análise NIR de soluções em hexano. Os círculos representam as amostras de calibração, os losangos as amostras de validação externa e a linha preta o ajuste perfeito.

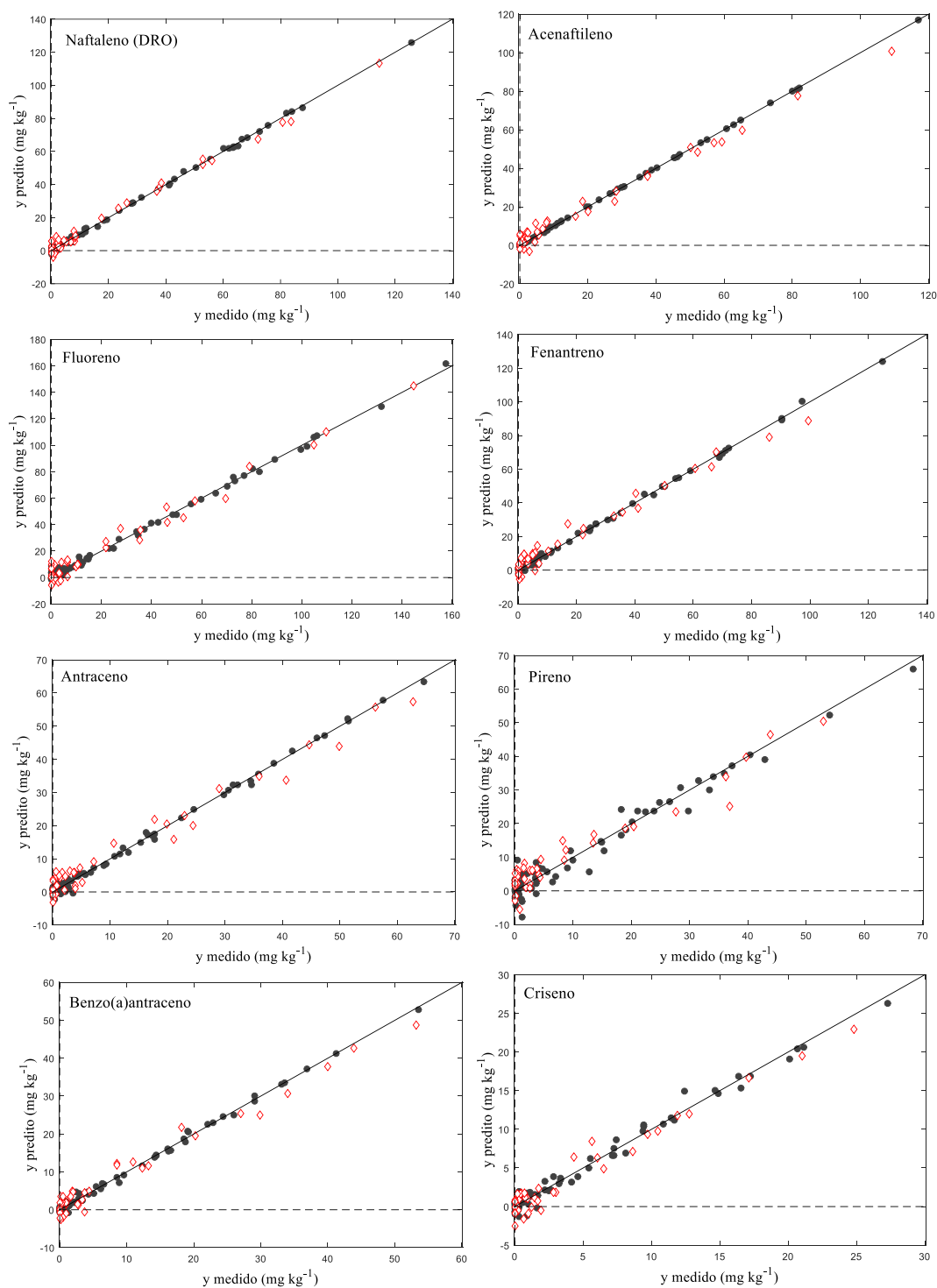


Figura C2.2 – Gráficos de ajuste dos modelos PLS para quantificação das frações TPH e analitos individuais determinados pela análise NIR de soluções em hexano. Os círculos representam as amostras de calibração, os losangos as amostras de validação externa e a linha preta o ajuste perfeito.

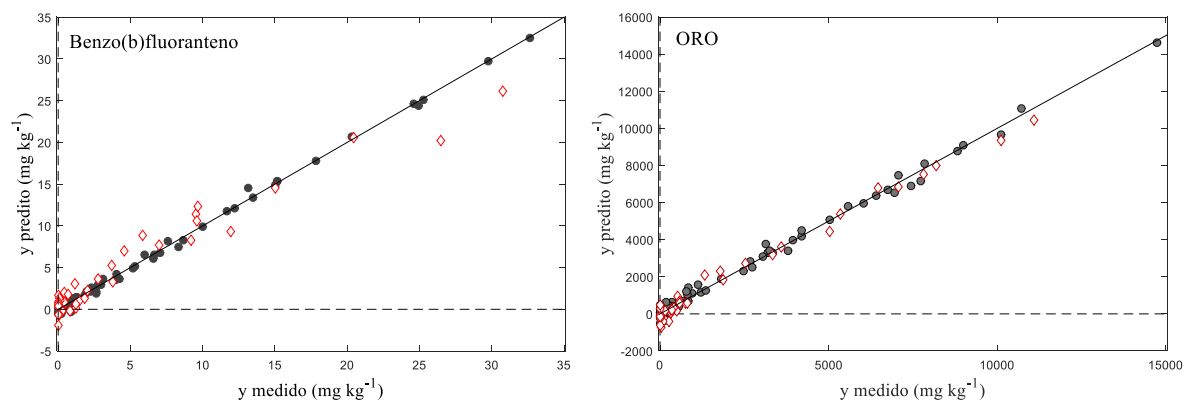


Figura C2.3 – Gráficos de ajuste dos modelos PLS para quantificação das frações TPH e analitos individuais determinados pela análise NIR de soluções em hexano. Os círculos representam as amostras de calibração, os losangos as amostras de validação externa e a linha preta o ajuste perfeito.

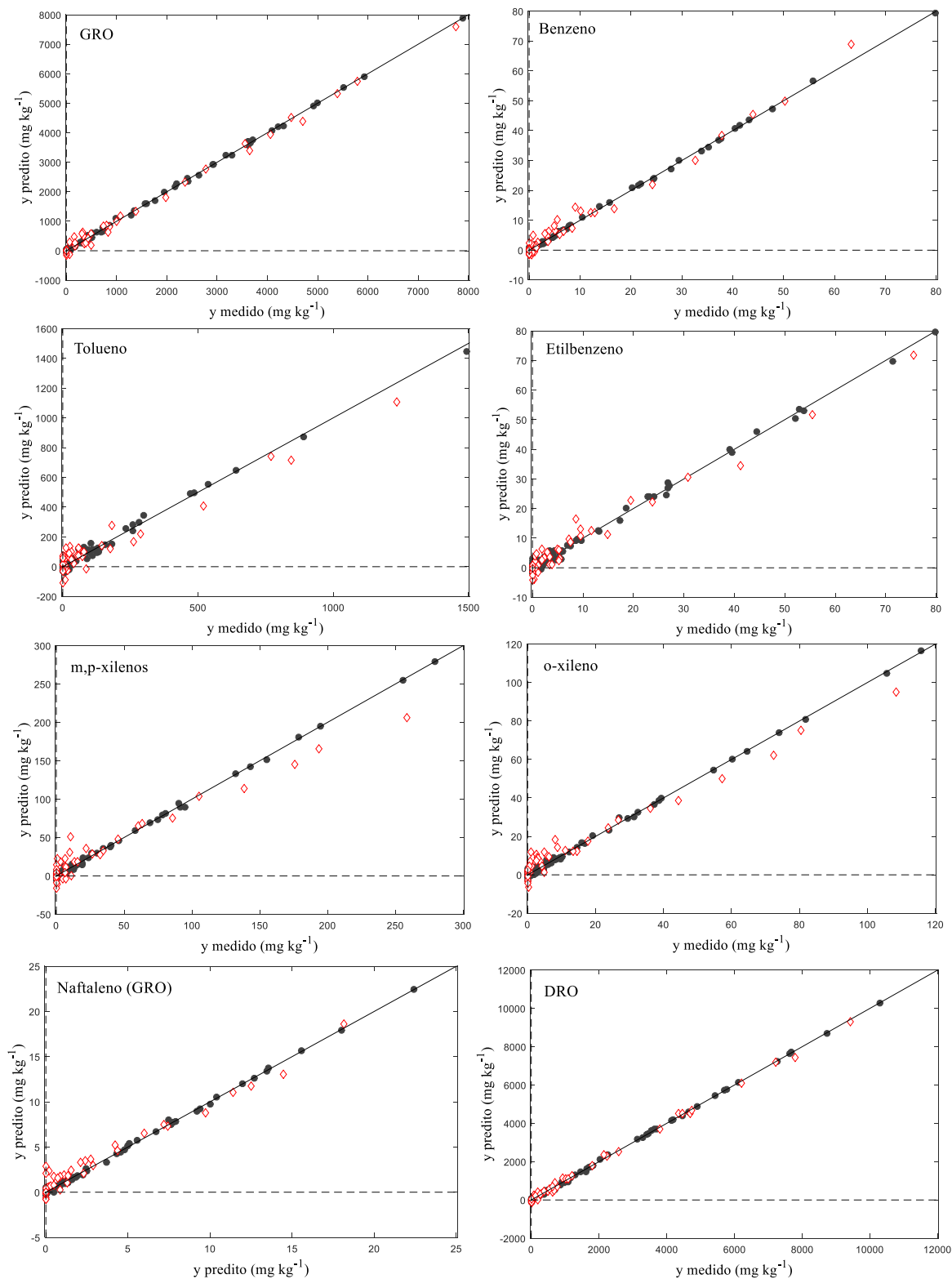


Figura C3.1 – Gráficos de ajuste dos modelos PLS para quantificação das frações TPH e analitos individuais determinados pela análise NIR de soluções em etanol. Os círculos representam as amostras de calibração, os losangos as amostras de validação externa e a linha preta o ajuste perfeito.

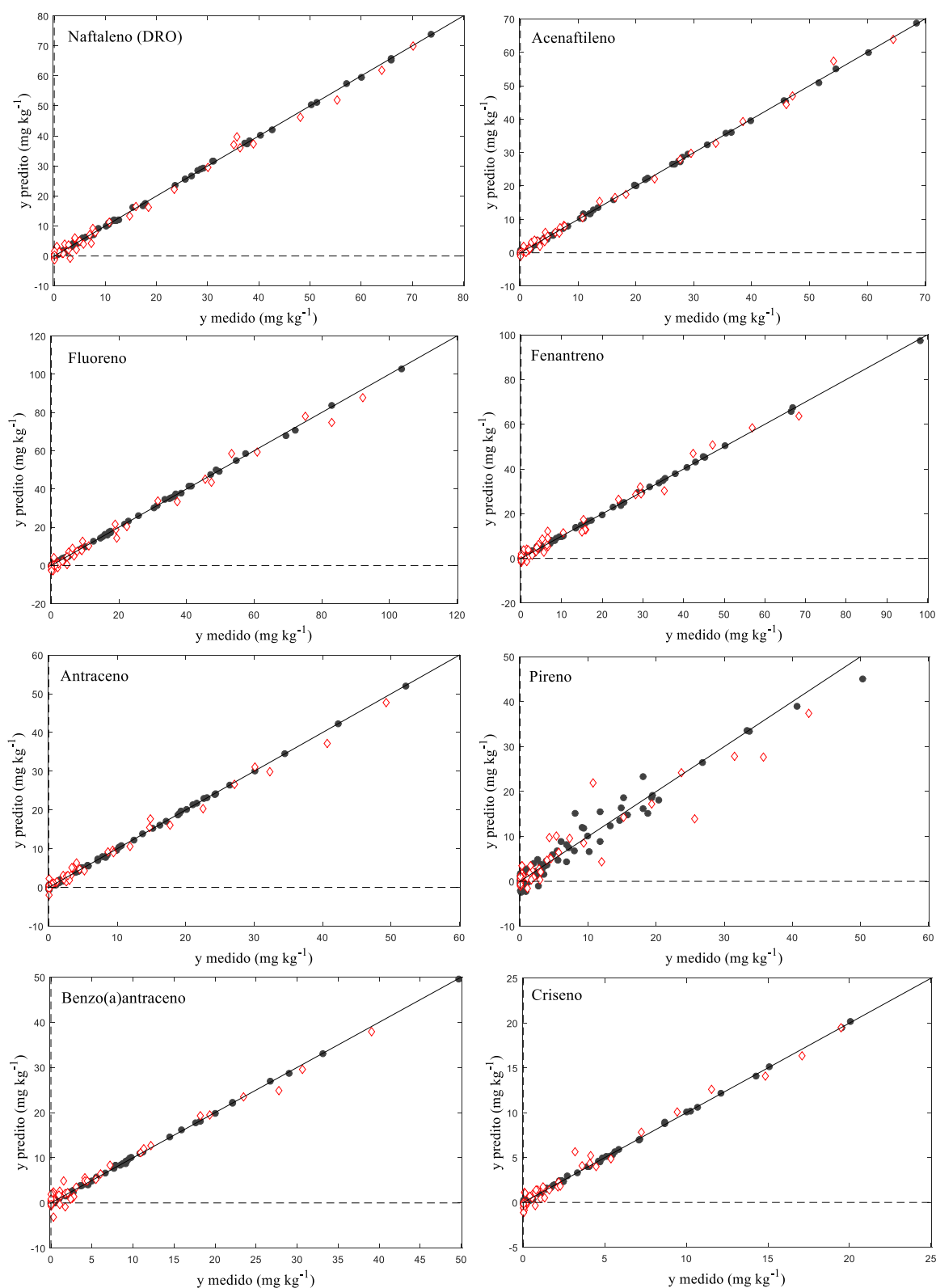


Figura C3.2 – Gráficos de ajuste dos modelos PLS para quantificação das frações TPH e analitos individuais determinados pela análise NIR de soluções em etanol. Os círculos representam as amostras de calibração, os losangos as amostras de validação externa e a linha preta o ajuste perfeito.

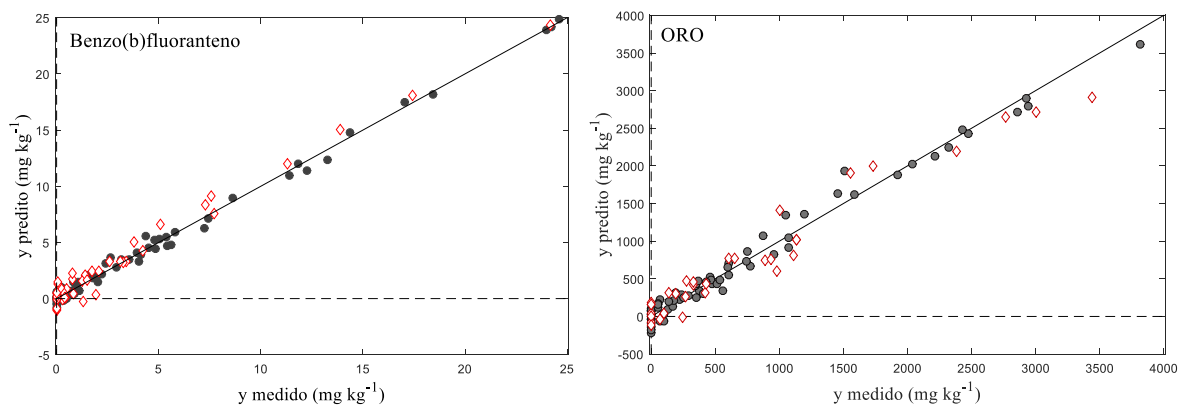


Figura C3.3 – Gráficos de ajuste dos modelos PLS para quantificação das frações TPH e analitos individuais determinados pela análise NIR de soluções em etanol. Os círculos representam as amostras de calibração, os losangos as amostras de validação externa e a linha preta o ajuste perfeito.