



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA  
DEPARTAMENTO DE QUÍMICA FUNDAMENTAL

LUAN CÁSSIO BARBOSA PATRÍCIO

**ANÁLISE COMPARATIVA DE GASOLINAS AUTOMOTIVAS PARA FINS  
FORENSES UTILIZANDO ENSAIOS NORMATIZADOS, CURVAS DE DESTILAÇÃO  
E ESPECTROSCOPIA NO INFRAVERMELHO ASSOCIADAS A TÉCNICAS  
QUIMIOMÉTRICAS**

Recife  
2021

LUAN CÁSSIO BARBOSA PATRÍCIO

**ANÁLISE COMPARATIVA DE GASOLINAS AUTOMOTIVAS PARA FINS  
FORENSES UTILIZANDO ENSAIOS NORMATIZADOS, CURVAS DE DESTILAÇÃO  
E ESPECTROSCOPIA NO INFRAVERMELHO ASSOCIADAS A TÉCNICAS  
QUIMIOMÉTRICAS**

Dissertação apresentada à Coordenação do Curso de Pós-Graduação em Química da Universidade Federal de Pernambuco, como requisito parcial à obtenção do grau de mestre em Química

Área de concentração: Química Analítica

Orientador: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Maria Fernanda Pimentel Avelar;

2º Orientador: Prof<sup>º</sup>. Dr. Neirivaldo Cavalcante Silva

Co-Orientador: Dr<sup>ª</sup>. Rafaella de Figueiredo Sales

Recife  
2021

Catálogo na fonte  
Bibliotecária Nataly Soares Leite Moro, CRB4-1722

P314a Patrício, Luan Cássio Barbosa  
Análise comparativa de gasolinas automotivas para fins forenses utilizando ensaios normatizados, curvas de destilação e espectroscopia no infravermelho associadas a técnicas quimiométricas / Luan Cássio Barbosa Patrício. – 2021.  
113 f.: il., fig., tab.

Orientadora: Maria Fernanda Pimentel Avelar.  
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Química, Recife, 2021.  
Inclui referências e apêndices.

1. Química analítica. 2. Forense. 3. Gasolina. 4. Quimiometria. I. Avelar, Maria Fernanda Pimentel (orientadora). II. Título.

543                      CDD (23. ed.)                      UFPE- CCEN 2022 - 96

# **LUAN CÁSSIO BARBOSA PATRÍCIO**

ANÁLISE COMPARATIVA DE GASOLINAS AUTOMOTIVAS PARA FINS FORENSES UTILIZANDO ENSAIOS NORMATIZADOS, CURVAS DE DESTILAÇÃO E ESPECTROSCOPIA NO INFRAVERMELHO ASSOCIADAS A TÉCNICAS QUIMIOMÉTRICAS.

Dissertação apresentada ao Programa de Pós-Graduação no Departamento de Química Fundamental da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Química.

Aprovada em: 27/12/2021

## **BANCA EXAMINADORA**

---

**Profa. Maria Fernanda Pimentel Avelar (Orientadora)**  
Universidade Federal de Pernambuco

---

**Profa. Claudete Fernandes Pereira**  
Universidade Federal de Pernambuco

---

**Profa. Fernanda Araújo Honorato**  
Universidade Federal de Pernambuco

*Dedico este trabalho a todos que contribuíram com o desenvolvimento desta pesquisa e àqueles que fizeram e fazem parte da minha vida.*

## AGRADECIMENTOS

A toda minha família, em especial aos meus pais Severino e Lúcia, e meus irmãos Katielle e Klébson, pelo apoio e suporte, durante esse período de dupla atividade, trabalhando como Analista Ambiental de Engenharia Química e ao mesmo tempo realizando minhas atividades do mestrado em Química.

A todos os meus amigos, por estarem presentes, ajudando sempre que preciso. Em especial a Ulisses Braga, Luiz Paulo Guerra, Nielline Andrade, Emilse Fernandez, Erika Rodrigues, Arthur Velloso e meus amigos da Secretaria Executiva de Meio Ambiente do Paulista, os quais muitas vezes ouviram meus lamentos e minhas alegrias durante este processo, me auxiliando sempre que precisei.

À professora Dr<sup>a</sup>. Maria Fernanda Pimentel, ao professor Dr<sup>o</sup>. Neirivaldo Cavalcante da Silva, à Dr<sup>a</sup>. Rafaella Figueiredo Sales e à perita criminal Dr<sup>a</sup> Livia Rodrigues e Brito por me orientarem na realização deste trabalho, por todas as conversas, discussões, por todo conhecimento passado e pela paciência e incentivo.

A Maria Eduarda Fernandes da Silva, sem a qual eu não poderia ter desenvolvido este projeto, por todo suporte, todas as discussões e desafios que passamos juntos para poder realizar um trabalho com qualidade e rigor científico.

A todos os meus amigos e colegas que compõem o Laboratório de Combustíveis (LAC) do Instituto de Petróleo, Gás e Energia (LITPEG) da Universidade Federal de Pernambuco (UFPE), por todos os momentos compartilhados, em especial ao gerente técnico João Rangel e toda equipe de analistas do laboratório, os quais me ajudaram na coleta das amostras e em diversas outras etapas deste trabalho.

A todos os professores da Pós-graduação em Química do Departamento de Química Fundamental (DQF) da UFPE com os quais tive a oportunidade de aprender sobre ciência, tecnologia e tantos outros assuntos.

E, por fim, a todas as pessoas que direta ou indiretamente contribuíram com este trabalho e/ou me deram força, apoio e suporte durante sua realização.

## RESUMO

A Polícia Científica do estado de Pernambuco, Brasil, vem se deparando com casos forenses que demandam reconhecer se amostras de gasolina são de mesma origem, como, por exemplo, situações em que é preciso saber se uma amostra apreendida em um caminhão é a mesma de um tanque. Ensaio físico-químico convencional para análise de gasolina podem ser utilizados nesses casos na tentativa de comparar as amostras. Outra opção é empregar a espectroscopia no infravermelho, o que tornaria o trabalho da polícia científica mais fácil e rápido, podendo inclusive viabilizar análises *in situ* com instrumentos portáteis. Nesse contexto, este trabalho tem como objetivo avaliar o potencial de utilização dos métodos convencionais e da espectroscopia na região do infravermelho (médio - MIR e próximo - NIR), usando equipamentos de bancada e portátil, para comparação de pares de amostras de gasolinas para fins forenses. A metodologia desenvolvida considera as incertezas de medição relacionadas aos ensaios convencionais e às medidas espectroscópicas. As incertezas para as medidas espectrais foram estimadas considerando quatro principais fontes: variação na temperatura, variação no equipamento em termos de posicionamento do comprimento de onda, repetibilidade e precisão intermediária. A partir destas, foi estimada a incerteza combinada e expandida, conforme o ISO GUM - *Guide to the Expression of Uncertainty in Measurement*. Foram coletadas 54 amostras de gasolina de 9 regiões do estado de Pernambuco, as quais foram fracionadas e destinadas para análises pelos ensaios normatizados pela Agência Nacional de Petróleo, Gás Natural e Biocombustíveis - ANP, obtenção das curvas de destilação e análises espectroscópicas na região do infravermelho próximo e médio. Os espectros foram obtidos em triplicata por meio de dois instrumentos distintos, sendo um espectrofotômetro de bancada (SPECTRUM 400, da Perkin Elmer, com aquisição de espectros na faixa de 10000 a 4000  $\text{cm}^{-1}$  no NIR e de 4000 a 400  $\text{cm}^{-1}$  no MIR) e um equipamento portátil (MicroNIR, da Viavi, com aquisição na faixa de 11012 a 5965  $\text{cm}^{-1}$ ), resultando em um total de 54 espectros médios em cada equipamento. Considerando as incertezas expandidas estimadas, foram gerados aleatoriamente 49 espectros virtuais para cada amostra dos ensaios convencionais, cada curva de destilação e para cada um dos espectros médios, totalizando 13500 amostras para compor o estudo, sendo 2700 amostras dos ensaios convencionais, 2700 das curvas de destilação, 2700 espectros no equipamento NIR portátil, 2700 do equipamento de bancada no NIR e 2700 do equipamento de bancada no MIR. Em cada caso, as amostras foram organizadas em pares pela distância Euclidiana, em seguida os dados dos ensaios convencionais e das curvas de destilação foram pré-processados com autoescalonamento e os dados espectroscópicos com SNV para então serem comparados em pares de amostras (cada um contendo a amostra real e as respectivas amostras virtuais) em uma análise de componentes principais (PCA). Foi possível visualizar a separação clara na PCA em 97% dos pares das análises convencionais, o restante foi parcialmente separado. Para as curvas de destilação, percebeu-se a separação em 100% dos pares. No caso dos espectros obtidos com o equipamento portátil NIR, 13,9% dos pares não foram separados, 14,7% foram parcialmente separados e 71,4% foram completamente separados. Já para os dados espectrais NIR adquiridos com equipamento de bancada, 20,5% dos pares não foram separados, 14% foram parcialmente separados e 65,5% foram completamente separados. Para o conjunto de espectros obtidos no MIR, 29,3% dos pares não foram separados, 54,5% foram parcialmente separados e 16,1% foram completamente separados. Os resultados obtidos mostram o potencial das estratégias desenvolvidas para discriminação das amostras de gasolina.

**Palavras-chave:** forense; gasolina; incerteza; quimiometria; separação.

## ABSTRACT

In Pernambuco State, Brazil, the police had faced forensic cases such as fuel theft and adulteration often require identification of gasoline sample as to origins, to decide whether or not samples seized at different locations came from the same fuel. Conventional physicochemical analysis for gasoline can be applied to try to solve these cases. Nevertheless, near (NIR) and middle (MIR) infrared spectroscopy can facilitate forensic work, especially due to the possibility of carrying out in situ analysis using handheld NIR devices. These devices require less instrumentation and time than needed for conventional physicochemical analyses and facilitate use in the field. In this context, this work aims to evaluate the potential of conventional physicochemical analyses and NIR and MIR spectroscopy, comparing a handheld NIR device and a benchtop NIR spectrometer, to differentiate one individual gasoline sample from another. The methodology considers measurement uncertainties related to conventional analyses and spectroscopic measurements. The uncertainties for the spectral measurements were estimated considering four main sources: temperature variation, equipment variation in terms of wavelength positioning, repeatability and intermediate precision. From these, the combined and expanded uncertainty was estimated, according to the ISO GUM - Guide to the Expression of Uncertainty in Measurement. Fifty-five gasoline samples were collected from 9 regions of the state of Pernambuco, Brazil, which were fractionated and destined for analysis by tests standardized by the ANP, obtaining distillation curves and spectroscopic analysis in the near and mid-infrared region. The spectra were obtained in triplicate using two different instruments, a benchtop spectrophotometer (SPECTRUM 400, by Perkin Elmer, with acquisition of spectra in the range of 10000 to 4000  $\text{cm}^{-1}$  in the NIR and from 4000 to 400  $\text{cm}^{-1}$  in the MIR) and a handheld instrument (MicroNIR, from Viavi, with acquisition of spectra in the range from 11012 to 5965  $\text{cm}^{-1}$ ), resulting in a total of 54 average spectra in each instrument. Considering the estimated expanded uncertainties, 49 virtual spectra were randomly generated for each sample of the conventional analyses, each distillation curve and for each of the average spectra, totaling 13,500 samples to compose the study, 2700 of which were from the conventional analyses, 2700 of the distillation curves, 2700 spectra on handheld NIR instrument, 2700 spectra on benchtop instrument on NIR and 2700 on benchtop instrument on MIR. In each case, the samples were arranged in pairs by their Euclidean distance, then the data sets from conventional analyses and distillation curves were pre-processed with autoscaling and the spectroscopic data sets with SNV to then be compared in pairs of samples (each containing the real sample and the respective virtual samples) in the principal component analysis (PCA). It was possible to visualize a clear separation in the PCA in 97% of the pairs for the conventional analyses data sets, the rest were partially separated. For the distillation curves data sets, a separation in 100% of the pairs was noticed. In the case of the spectra obtained with the portable NIR equipment, 13.9% of the pairs were not separated, 14.7% were partially separated and 71.4% were completely separated. As for the NIR spectral data acquired with bench equipment, 20.5% of the pairs were not separated, 14% were partially separated and 65.5% were completely separated. For the set of spectra obtained in the MIR, 29.3% of the pairs were not separated, 54.5% were partially separated and 16.1% were completely separated. The results obtained showed the potential of the developed strategies to discriminate gasoline samples.

**Keywords:** *chemometrics; forensic; fuels; uncertainty; separation.*

## LISTA DE FIGURAS

Figura 1 –	Especificações das gasolinas comum e premium conforme Resolução ANP N° 807, de 23 de janeiro de 2020. ....	24
Figura 2 –	Esquema de processos de fabricação de gasolina em refinarias, pelos métodos tradicionais.....	25
Figura 3 –	Gráfico de energia potencial para o modelo do oscilador anarmônico.....	30
Figura 4 –	Representação de análise por ATR. ....	32
Figura 5 –	Esquema de um Diagrama de Ishikawa para identificar as fontes de incerteza em resultado de medida. ....	34
Figura 6 –	Áreas de conhecimento que compõem a Quimiometria. ....	36
Figura 7 –	Exemplo de análise de componentes principais, apresentando o conjunto de dados em suas coordenadas originais e eixos da PC1 e PC2; .....	40
Figura 8 –	Esquema do fracionamento de amostras de gasolinas e destinação para análises espectroscópicas. ....	45
Figura 9 –	Representação do equipamento portátil MicroNIR 1700 (a) e equipamento acoplado a acessório de transmitância (b). ....	46
Figura 10 –	Aparato para a realização de análises de gasolina pelo MicroNIR. ....	46
Figura 11 –	Aparato para realização de análises na região do NIR no SPECTRUM 400.....	47
Figura 12 –	Acessório de reflectância total atenuada acoplado no equipamento utilizado nas análises de gasolina na região do MIR. ....	48
Figura 13 –	Diagrama de espinha de peixe com as fontes de incerteza da medida de absorvância dos espectros de infravermelho.....	50
Figura 14 –	Esquema indicando a distribuição das amostras utilizadas para os cálculos da repetibilidade e da precisão intermediária.....	51
Figura 15 –	Representação de uma banda, o valor máximo de absorvância e seus parâmetros utilizados na estimativa da incerteza. ....	53
Figura 16 –	Representação esquemática do cálculo das possíveis distâncias Euclidianas entre todos os pares de amostras. ....	56
Figura 17 –	Scores em PC1 e PC2 para todos os 54 grupos de amostras dos ensaios normalizados. ....	58

Figura 18 –	Scores para o par de amostras mais próximas em PC1xPC2 (a) e em PC1xPC3(b), considerando os métodos convencionais .....	58
Figura 19 –	<i>Loadings</i> das amostras em PC1.....	59
Figura 20 –	<i>Scores</i> em PC1xPC2 para dois pares de amostras acompanhados dos respectivos <i>loadings</i> em PC1. ....	60
Figura 21 –	<i>Scores</i> para o par de amostras completamente separadas em PC1xPC2 (a) e em PC1xPC3(b). ....	61
Figura 22 –	<i>Loadings</i> em PC1 para um par de amostras distintas.....	61
Figura 23 –	Exemplo de uma curva de destilação obtida (a) e todas as 54 curvas de destilação das amostras analisadas em laboratório (b).....	62
Figura 24 –	Curva de destilação de uma amostra real e suas respectivas 49 curvas obtidas virtualmente. ....	63
Figura 25 –	PCA com todo conjunto de dados (54 grupos de amostras). ....	64
Figura 26 –	<i>Scores</i> para o segundo par de amostras mais próximas em PC1xPC2 (a) e em PC1xPC7(b).....	64
Figura 27 –	<i>Loadings</i> em PC1 para o par de amostras mais próxima. ....	65
Figura 28 –	<i>Scores</i> em PC1xPC2 para o 3º par de amostras mais próximas (a) e o 500º par de amostras mais próximas (b).....	66
Figura 29 –	<i>Loadings</i> em PC1 para o 3º par de amostras mais próximas (a) e o 500º par de amostras mais próximas (b).....	66
Figura 30 –	Espectros médios brutos das 54 amostras analisadas.....	67
Figura 31 –	Espectros pré-processados com MSC (a), correção de linha de base (b) e SNV + deslocamento no eixo y (c). ....	68
Figura 32 –	Incertezas referentes à precisão intermediária, ou reprodutibilidade (a), à repetibilidade (b), à variação de temperatura (c) e ao equipamento (d).....	69
Figura 33 –	Incerteza expandida relativa (a) e incerteza expandida absoluta (b) em cada comprimento de onda para os espectros do MicroNIR com as contribuições das 4 fontes de incerteza. ....	70
Figura 34 –	Incerteza expandida relativa (a) e incerteza expandida absoluta (b) em cada comprimento de onda para os espectros do MicroNIR sem a contribuição da fonte de incerteza do equipamento. ....	71

Figura 35 –	Exemplo de espectro real e um dos seus respectivos espectros virtuais obtido com intervalo $n=4$ . .....	72
Figura 36 –	Scores em PC1 e PC2 para todos os 54 grupos de amostras.....	72
Figura 37 –	<i>Scores</i> em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, obtidas utilizando o equipamento portátil, que não foram distinguidas visualmente pela PCA. ....	73
Figura 38 –	<i>Loadings</i> em PC1 (a) e PC2 (b) para um par de amostras que não puderam ser distinguidas pela PCA. ....	74
Figura 39 –	<i>Scores</i> em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras parcialmente distinguidas visualmente pela PCA. ....	74
Figura 40 –	<i>Loadings</i> em PC1 (a) e PC2 (b) para um par de amostras parcialmente distinguidas pela PCA. ....	75
Figura 41 –	<i>Scores</i> em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras completamente distinguidas visualmente pela PCA. ....	75
Figura 42 –	<i>Loadings</i> em PC1 (a) e PC2 (b) para um par de amostras totalmente distinguidas pela PCA. ....	76
Figura 43 –	Espectros médios brutos das 54 amostras analisadas com o instrumento NIR de bancada. ....	77
Figura 44 –	Representação dos 54 espectros médios na região NIR adquiridos com equipamento de bancada pré-processados com SNV e deslocados no eixo y.....	78
Figura 45 –	Incertezas estimadas dos espectros NIR adquiridos com um equipamento de bancada, referentes à precisão intermediária, ou reprodutibilidade (a), à repetibilidade (b), à variação de temperatura (c) e ao equipamento (d)..	79
Figura 46 –	Incertezas estimadas dos espectros NIR adquiridos com um equipamento de bancada, referentes à precisão intermediária, ou reprodutibilidade (a), à repetibilidade (b), à variação de temperatura (c) e ao equipamento (d) excluída a região entre 10000 $\text{cm}^{-1}$ e 8400 $\text{cm}^{-1}$ . ....	80
Figura 47 –	Incerteza expandida dos espectros NIR adquiridos com um equipamento de bancada, relativa (a) e incerteza expandida absoluta (b) em cada número de onda. ....	80

Figura 48 –	Exemplo de espectro real na região NIR adquirido com equipamento de bancada e um dos seus respectivos espectros virtuais. ....	81
Figura 49 –	Scores em PC1 e PC2 para a PCA aplicada a todos os 54 grupos de amostras do conjunto de dados de espectros NIR adquiridos com equipamento de bancada. ....	82
Figura 50 –	Scores em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, que não foram distinguidas visualmente pela PCA. ....	83
Figura 51 –	<i>Loadings</i> em PC1 para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, que não puderam ser distinguidas pela PCA. ....	83
Figura 52 –	Scores em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, parcialmente distinguidas visualmente pela PCA. ....	84
Figura 53 –	<i>Loadings</i> em PC1 para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, parcialmente distinguidas pela PCA. ....	85
Figura 54 –	<i>Scores</i> em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, completamente distinguidas visualmente pela PCA. ....	85
Figura 55 –	<i>Loadings</i> em PC1 (a) e PC2 (b) para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, totalmente distinguidas pela PCA. ....	86
Figura 56 –	Espectros médios brutos das 54 amostras analisadas na região do MIR. ..	87
Figura 57 –	Representação dos 54 espectros médios na região do MIR pré-processados com SNV e deslocados no eixo y. ....	88
Figura 58 –	Incertezas dos espectros MIR referentes à precisão intermediária, ou reprodutibilidade (a), à repetibilidade (b), à variação de temperatura (c) e ao equipamento (d). ....	89

Figura 59 –	Incertezas dos espectros MIR referentes à precisão intermediária, ou reprodutibilidade (a), à repetibilidade (b), à variação de temperatura (c) e ao equipamento (d) excluída a região entre 4000 e 3650 $\text{cm}^{-1}$ .....	90
Figura 60 –	Incerteza expandida relativa dos espectros MIR (a) e incerteza expandida absoluta (b).....	90
Figura 61 –	Exemplo de espectro real e um dos seus respectivos espectros virtuais na região do MIR. ....	91
Figura 62 –	<i>Scores</i> em PC1 e PC2 para todos os 54 grupos de amostras obtidas na região do MIR. ....	92
Figura 63 –	<i>Scores</i> em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, com espectros adquiridos no MIR, que não foram distinguidas visualmente pela PCA. ....	93
Figura 64 –	<i>Loadings</i> em PC1 para um par de amostras com espectros adquiridos no MIR que não foram distinguidas pela PCA. ....	93
Figura 65 –	<i>Scores</i> em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, com espectros adquiridos no MIR, que foram parcialmente distinguidas visualmente pela PCA. ....	94
Figura 66 –	<i>Loadings</i> em PC1 (a) e PC2 (b) para um par de amostras, com espectros adquiridos no MIR, parcialmente distinguidas pela PCA. ....	95
Figura 67 –	<i>Scores</i> em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras com espectros adquiridos no MIR que foram completamente distinguidas visualmente pela PCA. ....	95
Figura 68 –	<i>Loadings</i> em PC1 para um par de amostras com espectros adquiridos no MIR completamente distinguidas pela PCA. ....	96

## LISTA DE TABELAS

Tabela 1 –	Métodos de referências para os parâmetros físico-químicos de análise de gasolina monitorados pelo PMQC. ....	27
Tabela 2 –	Regiões do infravermelho. ....	29
Tabela 3 –	Valores de incertezas para as grandezas medidas pelos métodos normatizados estimadas segundo o ISO GUM. ....	49
Tabela 4 –	Resultados de análises físico-químicas para uma amostra de cada região analisada. ....	57

## LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
AEAC	Álcool etílico anidro combustível
ANP	Agência Nacional de Petróleo, Gás Natural e Biocombustíveis
ASTM	American Society for Testing and Materials
ATR	Reflectância Total Atenuada
CVA	Análise de Variável Ortogonal Canônica
FBP	Ponto final de ebulição
FIR	<i>Far Infrared</i> (Espectroscopia no Infravermelho distante)
GC/MS	Cromatografia gasosa acoplada a espectrometria de massas
HCA	Análise Hierárquica de agrupamentos
IAD	Índice antidetonante
IBP	Ponto inicial de ebulição
KS	Algoritmo de Kennard-Stone
LAC	Laboratório de Combustíveis da Universidade Federal de Pernambuco
LDA	Análise Discriminante Linear
LSSVM	Máquina de vetor de suporte de mínimos quadrados
MANOVA	Análise Multivariada de Variância
MicroNIR	Espectrômetro NIR portátil
MIR	<i>Mid Infrared</i> (Espectroscopia no Infravermelho Médio)
MLP	<i>Multilayer Perception</i>
MON	<i>Motor Octane Number</i>
MTBE	Metil-terc-butil éter

NBR	Norma Brasileira de Regulamentação
NIR	<i>Near Infrared</i> (Espectroscopia no Infravermelho Próximo)
OCVA	Análise de Variável Ortogonal Canônica
PC	Componente Principal
PCA	Análise por Componentes Principais
PLS	Regressão pelo Método dos Mínimos Quadrados Parciais
PLS-DA	Análise Discriminante pelo Método dos Mínimos Quadrados Parciais
QDA	Análise Discriminante Quadrática
RDA	Análise Discriminante Regularizada
REPE	Repetibilidade
REPRO	Reprodutibilidade
RF	Floresta Aleatória
RON	<i>Research Octane Number</i>
SEC	Erro padrão de calibração (do inglês <i>standard error of calibration</i> )
SIMCA	Modelagem Independente e Flexível por Analogia de Classes
SNV	Varição Normal Padrão ( <i>Standard Normal Variate</i> )
SVM	Máquinas de Vetores de Suporte
UFPE	Universidade Federal de Pernambuco

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>20</b>
1.1	OBJETIVOS .....	22
1.1.1	Geral.....	22
1.1.2	Específicos.....	22
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA .....</b>	<b>23</b>
2.1	GASOLINA .....	23
2.1.1	Composição química e características .....	23
2.1.2	Classificação das gasolinas .....	24
2.1.3	Processos de fabricação de gasolinas.....	25
2.2	PARÂMETROS FÍSICO-QUÍMICOS DAS GASOLINAS .....	26
2.2.1	Ensaio normatizados ANP .....	26
2.2.2	Massa específica .....	27
2.2.3	Teor de etanol.....	27
2.2.4	Curvas de destilação .....	28
2.3	ESPECTROSCOPIA NA REGIÃO DO INFRAVERMELHO .....	29
2.3.1	A absorção no infravermelho.....	29
2.4	INCERTEZA DE MEDIÇÃO .....	32
2.4.1	Erros e incertezas .....	32
2.4.2	Estimativa da incerteza .....	33
2.4.3	Estudos de Repetibilidade, reprodutibilidade e precisão intermediária ...	34
2.5	TÉCNICAS QUIMIOMÉTRICAS MULTIVARIADAS DE ANÁLISE .....	35
2.5.1	Quimiometria .....	35
2.5.2	Organização dos dados e Técnicas de Pré-processamento.....	37
2.5.2.1	Centragem na média .....	38

2.5.2.2	Padronização Normal de Sinal.....	38
<b>2.5.3</b>	<b>Análise exploratória dos dados.....</b>	<b>39</b>
2.5.3.1	Análise de Componentes Principais (PCA).....	39
2.6	USO DE ESPECTROSCOPIA NO INFRAVERMELHO ASSOCIADA À QUIMIOMETRIA PARA CLASSIFICAÇÃO DE GASOLINAS .....	40
<b>3</b>	<b>METODOLOGIA.....</b>	<b>44</b>
3.1	ANÁLISES FÍSICO-QUÍMICAS E ESPECTROSCÓPICAS DAS GASOLINAS.....	44
<b>3.1.1</b>	<b>Obtenção das amostras.....</b>	<b>44</b>
<b>3.1.2</b>	<b>Análises convencionais (ANP) e curvas de destilação.....</b>	<b>44</b>
<b>3.1.3</b>	<b>Análises por espectroscopia no infravermelho.....</b>	<b>44</b>
3.1.3.1	Aquisição dos espectros no infravermelho próximo em equipamento portátil.	45
3.1.3.2	Aquisição dos espectros no infravermelho próximo em equipamento de bancada .....	47
3.1.3.3	Aquisição dos espectros no infravermelho médio em equipamento de bancada .....	48
3.2	CÁLCULO DE INCERTEZAS E GERAÇÃO DE AMOSTRAS VIRTUAIS .....	48
<b>3.2.1</b>	<b>Análises convencionais (ANP).....</b>	<b>49</b>
<b>3.2.2</b>	<b>Curvas de destilação .....</b>	<b>49</b>
<b>3.2.3</b>	<b>Análises por espectroscopia no infravermelho.....</b>	<b>50</b>
3.2.3.1	Especificação do mensurando e identificação das fontes de incerteza .....	50
3.2.3.2	Quantificação dos componentes da incerteza .....	51
3.2.3.3	Cálculo da incerteza combinada .....	54
<b>3.2.4</b>	<b>Geração de amostras virtuais .....</b>	<b>54</b>
3.3	ANÁLISE DE COMPONENTES PRINCIPAIS.....	55

<b>4</b>	<b>RESULTADOS E DISCUSSÃO.....</b>	<b>56</b>
4.1	ENSAIOS NORMALIZADOS.....	56
<b>4.1.1</b>	<b>Análises Convencionais .....</b>	<b>56</b>
<b>4.1.2</b>	<b>Geração de amostras virtuais e Análise de Componentes Principais (PCA).....</b>	<b>57</b>
4.2	CURVAS DE DESTILAÇÃO.....	62
<b>4.2.1</b>	<b>Curvas de destilação .....</b>	<b>62</b>
<b>4.2.2</b>	<b>Cálculo de incertezas e geração de amostras virtuais.....</b>	<b>63</b>
<b>4.2.3</b>	<b>Análise de Componentes Principais .....</b>	<b>63</b>
4.3	ESPECTROSCOPIA NIR EM INSTRUMENTO PORTÁTIL .....	67
<b>4.3.1</b>	<b>Espectros .....</b>	<b>67</b>
<b>4.3.2</b>	<b>Cálculo de incertezas .....</b>	<b>69</b>
<b>4.3.3</b>	<b>Geração de amostras virtuais .....</b>	<b>71</b>
<b>4.3.4</b>	<b>Análise de Componentes Principais .....</b>	<b>72</b>
4.4	ESPECTROSCOPIA NO NIR EM INSTRUMENTO DE BANCADA .....	77
<b>4.4.1</b>	<b>Espectros .....</b>	<b>77</b>
<b>4.4.2</b>	<b>Cálculo de incertezas .....</b>	<b>78</b>
<b>4.4.3</b>	<b>Geração de amostras virtuais .....</b>	<b>81</b>
<b>4.4.4</b>	<b>Análise de Componentes Principais (PCA) .....</b>	<b>82</b>
4.5	ESPECTROSCOPIA NO MIR.....	87
<b>4.5.1</b>	<b>Espectros .....</b>	<b>87</b>
<b>4.5.2</b>	<b>Cálculo de incertezas .....</b>	<b>88</b>
<b>4.5.3</b>	<b>Geração de amostras virtuais .....</b>	<b>91</b>
<b>4.5.4</b>	<b>Análise de Componentes Principais (PCA) .....</b>	<b>92</b>
<b>5</b>	<b>CONCLUSÕES.....</b>	<b>97</b>

5.1	PERSPECTIVAS FUTURAS.....	98
	<b>REFERÊNCIAS.....</b>	<b>100</b>
	<b>APÊNDICE A – ROTINA USADA NO MATLAB PARA A GERAÇÃO DE AMOSTRAS VIRTUAIS.....</b>	<b>106</b>
	<b>APÊNDICE B – ESPECTROS VIRTUAIS DE TESTES COM DIFERENTES VALORES PARA INTERVALOS DE VARIAÇÃO LINEAR DO COEFICIENTE RANDÔMICO DA INCERTEZA DOS ESPECTROS COM EQUIPAMENTO PORTÁTIL .....</b>	<b>109</b>
	<b>APÊNDICE C – ESPECTROS VIRTUAIS DE TESTES COM DIFERENTES VALORES PARA INTERVALOS DE VARIAÇÃO LINEAR DO COEFICIENTE RANDÔMICO DA INCERTEZA DOS ESPECTROS NIR COM EQUIPAMENTO DE BANCADA .....</b>	<b>110</b>
	<b>APÊNDICE D – ESPECTROS VIRTUAIS DE TESTES COM DIFERENTES VALORES PARA INTERVALOS DE VARIAÇÃO LINEAR DO COEFICIENTE RANDÔMICO DA INCERTEZA DOS ESPECTROS MIR.....</b>	<b>112</b>

## 1 INTRODUÇÃO

A utilização da gasolina como combustível automotivo é parte essencial do modo de vida atual, o qual exige grandes deslocamentos diários para os quais o uso desse combustível fóssil ainda se mantém indispensável. Neste contexto, o Brasil se destaca devido ao domínio da tecnologia de exploração e refino do petróleo, de onde provém a gasolina. Em 2019, conforme dados da Agência Nacional de Petróleo, Gás Natural e Biocombustíveis (ANP), 26,9% da produção brasileira de derivados energéticos de petróleo foi de gasolina, o que equivale ao derivado energético com o segundo maior volume produzido no país, atrás apenas da produção de óleo diesel (ANP, 2020).

Segundo a mesma agência, o consumo de combustíveis no Brasil em 2020 reduziu cerca de 6% em comparação com o ano de 2019 (venda total de 131,7 bilhões de litros), o que, no caso da gasolina, corresponde a uma queda de 6,1% (venda total de 35,8 bilhões de litros), resultado dos grandes impactos causados pela pandemia do novo coronavírus (ANP, 2021). Apesar disso, as proporções deste mercado no Brasil ainda se mantêm gigantescas, o que, aliado à grande dimensão territorial do país, traz consigo a necessidade de controle minucioso dos produtos que são produzidos e comercializados em todo território nacional, para tanto, exigindo atuação conjunta de diversos órgãos, tais quais polícias federal, estaduais e municipais, órgãos de controle de qualidade, instituições de pesquisa e ensino, bem como a própria sociedade civil.

Parte dos esforços dispendidos por essas instituições se deve ao frequente problema de falsificação e adulteração de gasolinas, seja pelo comércio ilegal em zonas de fronteiras, através de países como a Venezuela, Peru, Bolívia e Paraguai (SILVA, 2013), ou pela adulteração em solo brasileiro para os mais diversos fins. As consequências são distorções no mercado de distribuição e revenda de combustíveis, com efeitos negativos tanto para o consumidor quanto para a concorrência (ALVES, 2013).

A fim de monitorar a qualidade dos combustíveis comercializados no Brasil, a ANP instituiu na década de 1990 o Programa de Monitoramento da Qualidade de Combustíveis (PMQC), o qual tem a finalidade de acompanhar os indicadores da qualidade de combustíveis e identificar aqueles que não atendem às normas (ANP, 2021). Em 2020, foram coletadas mais de 75.000 amostras de combustíveis e realizadas cerca de 679.000 análises.

As ações de controle de qualidade utilizadas pela ANP, incluindo as resultantes do PMQC, bem como fiscalização e demais formas de atuação, vêm se mostrando eficazes em reduzir a

probabilidade de comercialização de combustíveis fora das especificações pelos varejistas de combustíveis automotivos (MENESES, 2019).

Embora as ações da ANP sejam fundamentais para o controle de qualidade dos combustíveis no país, órgãos de polícia também se fazem imprescindíveis, em especial em situações de flagrantes de crimes e nas regiões fronteiriças. Em Pernambuco, a Polícia Científica do estado vem se deparando com casos que demandam identificar se as amostras de gasolina apreendidas são de uma mesma fonte ou não. Como exemplo, podemos citar casos de roubos ou situações em que é preciso saber se uma amostra apreendida em um caminhão é a mesma de uma que está armazenada em um tanque, o que exige a análise por meio de instrumentos existentes na instituição, como espectrômetros de infravermelho de bancada. Daí a importância de se desenvolver métodos que sejam aplicados à realidade da instituição e que possam ser utilizados com a instrumentação que possuem.

No que diz respeito ao uso da espectroscopia de infravermelho, há alguns trabalhos na literatura que descrevem, por exemplo, a utilização da espectroscopia no infravermelho próximo (NIR – *Near Infrared*) para a classificação de gasolinas empregando diversas técnicas quimiométricas de classificação como Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA), Análise Discriminante Linear (LDA), Modelagem Independente e Flexível por Analogia de Classes (SIMCA) e outras (BALABIN, 2008, 2010; AL GHOUTI *et. al*, 2008).

Foram também realizados estudos utilizando curvas de destilação de gasolinas com a finalidade de identificar a origem de diversas amostras (ALEME, 2008), a influência de adulterações no perfil de destilação (DE OLIVEIRA, 2004), e para o desenvolvimento de um modelo capaz de identificar amostras de gasolinas adulteradas de não adulteradas, utilizando, para tal, curvas de destilação associadas às técnicas de Análise de Componentes Principais (PCA) e PLS-DA (MENDES, 2013; LIMA, 2019). Entretanto, não foram encontrados trabalhos que avaliam o potencial da espectroscopia no infravermelho, bem como dos métodos de referência recomendados pela ANP, para comparar pares de amostras de gasolinas, considerando suas respectivas incertezas de medição. Neste contexto, faz-se necessário o desenvolvimento de metodologias simples e confiáveis, que possibilitem realizar uma análise comparativa entre amostras de gasolina, demanda que tem surgido na área de química forense. Na presente dissertação, foram utilizados tanto dados provenientes de análises de qualidade estabelecidas pela ANP, como pontos da destilação, teor de álcool e massa específica, quanto espectros na região do infravermelho próximo e médio, a fim de

desenvolver métodos de fácil aplicação e sensíveis às diferenças de composição química de amostras de gasolina de diferentes procedências, com o intuito de elucidar casos forenses.

## 1.1 OBJETIVOS

### 1.1.1 Geral

Avaliar o potencial de reconhecer pares de amostras de gasolinas de diferentes procedências, em casos forenses, utilizando os métodos convencionais de análise normatizados pela ANP e a espectroscopia no infravermelho médio e próximo (com instrumentos de bancada e portátil) associados a técnicas quimiométricas.

### 1.1.2 Específicos

- Analisar dados obtidos através dos métodos convencionais de controle de qualidade da gasolina, considerando suas respectivas incertezas, a fim de avaliar o potencial de separação, aplicando técnica de reconhecimento de padrões não supervisionada;
- Desenvolver um método para obtenção de incertezas associadas a espectros de absorvância na região do infravermelho médio e próximo para amostras de gasolina;
- Desenvolver uma estratégia para geração de espectros virtuais na região do infravermelho, com base nas incertezas obtidas para espectros de gasolinas;
- Avaliar o potencial de separar pares de gasolinas usando os dados espectroscópicos e técnica de reconhecimento de padrões não supervisionada.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 GASOLINA

#### 2.1.1 Composição química e características

A gasolina é um combustível constituído basicamente por hidrocarbonetos parafínicos (hidrocarbonetos lineares, sem ligações duplas ou triplas), olefínicos (ou alcenos: hidrocarbonetos insaturados com pelo menos uma ligação dupla), naftênicos (naftenos ou cicloparafinas, caracterizados pela presença de anéis sem ligações duplas) e aromáticos (caracterizados pela presença do anel benzênico em sua estrutura). Apresenta faixa de ebulição entre 30°C e 220°C, sendo a maior parte das cadeias de 5 a 10 átomos de carbono, além de possuir contaminantes de compostos de enxofre, oxigênio, nitrogênio e metais em baixas quantidades (DO BRASIL *et. al*, 2012; VIDAL, 1999).

A gasolina pode ser utilizada como combustível para aviação e para automóveis, cada uma com características específicas. Em meados de 2020, começaram a ser exigidos novos padrões para as gasolinas comercializadas no Brasil, estabelecidos pela Resolução ANP N° 807, de 23 de janeiro de 2020, a qual determina as especificações de gasolinas de uso automotivo e as obrigações quanto ao controle de qualidade a serem atendidas. A Figura 1 mostra as novas especificações adotadas, o que inclui parâmetros como pontos de destilação com 10%, 50% e 90% evaporados e ponto final, além de teor de etanol e outros (ANP, 2020). Essas especificações têm como objetivo estabelecer um padrão de gasolina que atenda aos requisitos de qualidade de eficiência de utilização na combustão nos veículos.

Figura 1 – Especificações das gasolinas comum e premium conforme Resolução ANP N° 807, de 23 de janeiro de 2020.

CARACTERÍSTICA	UNIDADE	LIMITE				MÉTODO	
		Gasolina Comum		Gasolina Premium		ABNT NBR	ASTM
		A	C	A	C		
Cor	-	(1)				visual	
Aspecto	-	(2)				14954	D4176 (3)
Teor de Etanol Anidro Combustível (EAC)	% volume	(4)	(5)	(4)	(5)	13992	D5501 (6)
Massa específica a 20 °C, min. (22)	kg/m <sup>3</sup>	(7)	715,0	(7)	715,0	7148 14065	D1298 D4052
Destilação (8)						9619	D86 D7345 (9)
10% evaporados, máx.	°C	65,0					
50% evaporados (22)		77,0 a 120,0	Máx. 80,0	77,0 a 120,0	Máx. 80,0		
90% evaporados, máx.		190,0					
PFE, máx.		215,0					
Resíduo, máx.	% volume	2,0					
N° de Octano Motor - MON, min. (10)	-	-	82,0	-	anotar	-	D2700
N° de Octano Pesquisa - RON, min. (10)(22)	-	-	93,0 (11)	-	97,0	-	D2699
Pressão de Vapor a 37,8 °C (12)	kPa	45,0 a 62,0	69,0 (máx.)	45,0 a 62,0	69,0 (máx.)	14149 16306	D4953 D5191 D5482 D6378
Goma Atual Lavada, máx.	mg/100 mL	5				14525	D381
Período de Indução a 100 °C, min. (13)	min.	-	360	-	360	14478	D525
Corrosividade ao Cobre a 50 °C, 3h, máx.	-	1				14359	D130
Teor de Enxofre, máx. (10)(14)	mg/kg	-	50	-	50	- - - - -	D2622 D3120 D5453 D6920 D7039 D7220
Benzeno, máx. (15)(16)	% volume	-	1,0	-	1,0	15289 15441	D3606 D5443 D6277 D6729 D6730
Teor de Silício	mg/kg	Anotar				-	D7757
							AAS
							ICP-AES
Hidrocarbonetos: (15)(17)						14932	D1319
Aromáticos, máx.	% volume	-	35	-	35		
Olefinicos, máx.		-	25	-	25		
Saturados		Anotar					
Teor de Metanol, máx (18)(19)	% volume	0,5				16041	-
Chumbo, máx. (18)	g/L	0,005				-	D3237 D5059
Fósforo, máx. (18)	mg/L	1,3				-	D3231

Fonte: ANP, 2020.

### 2.1.2 Classificação das gasolinas

Segundo a resolução ANP n° 40 de 2013, as gasolinas podem ser classificadas de acordo com a presença de etanol, podendo ser do tipo A (produzida pelas refinarias e entregue às companhias distribuidoras sem a adição de álcool etílico anidro combustível – AEAC) ou do tipo C

(gasolinas comercializadas, as quais possuem AEAC, por exigência da ANP, conforme Portaria nº 75 do Ministério da Agricultura).

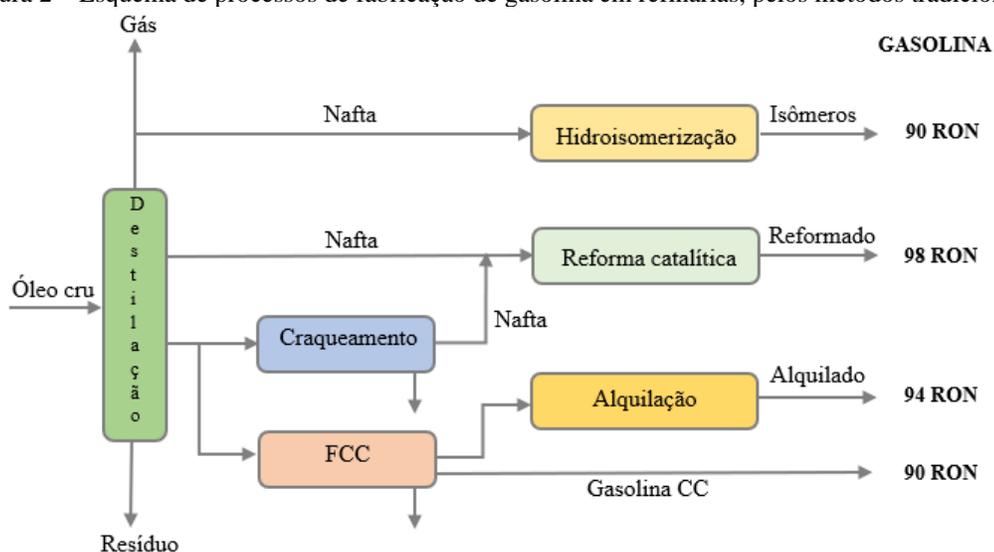
Quanto às propriedades antidetonantes, as gasolinas podem ser divididas em gasolina comum (aquelas que possuem índice de octanagem abaixo de 90% e não possuem aditivos), aditivada (também possui índice de octanagem abaixo de 90%, no entanto, possui aditivos que auxiliam na limpeza do motor) e premium (possui índice de octanagem acima de 90% e aditivos, proporcionando maior eficiência na combustão).

### 2.1.3 Processos de fabricação de gasolinas

A depender da origem e características do petróleo utilizado na fabricação da gasolina, bem como do processo de fabricação empregado e do uso final (automotiva ou para aviação), a gasolina resultante pode apresentar características bem distintas, com diferentes composições químicas.

A gasolina pode ser fabricada por refinarias, formuladores ou petroquímicas, sendo dois os processos de fabricação mais comuns. O primeiro é o método tradicional, em que a gasolina é obtida através da destilação fracionada do petróleo e de processos subsequentes de craqueamento catalítico (quebra de moléculas pesadas derivadas do petróleo em moléculas menores, em condições específicas de temperatura e pressão), isomerização, polimerização e alquilação, como mostrado na Figura 2.

Figura 2 – Esquema de processos de fabricação de gasolina em refinarias, pelos métodos tradicionais.



Fonte: O autor, baseado em DO BRASIL *et. al*, 2012.

Já o segundo método é o processo de formulação em petroquímicas e formuladores, em que diversas correntes de hidrocarbonetos líquidos são combinadas para produção de gasolina, as quais podem ser provenientes de fases intermediárias do processo ou adquiridas de outros produtores (SPEIGHT, 1999; SILVA, 2017).

O uso dos processos citados acima e mostrados na Figura 2 visam melhorar a qualidade da gasolina, aumentando o seu poder antidetonante, tendo em vista que as gasolinas provenientes de destilação direta possuem alto índice de hidrocarbonetos parafínicos (com alta tendência para detonação). Por outro lado, as resultantes do processo de reforma catalítica possuem alto teor de isoparafinas e hidrocarbonetos aromáticos e as que passaram por craqueamento (térmico ou catalítico) contêm elevado teor de hidrocarbonetos olefínicos, todos com baixa tendência para detonação, resultando em gasolinas com alto índice de octano e, portanto, maior qualidade (DO BRASIL *et. al*, 2012; PETROBRAS, 2020).

Embora as formas de produção sejam variadas, para a ANP, não há distinção entre gasolina formulada e refinada, visto que ambas devem atender aos padrões exigidos pela agência, a diferença está apenas na sua origem. As diferenças de composição química, no entanto, são um indicativo dos diferentes processos e origens.

## 2.2 PARÂMETROS FÍSICO-QUÍMICOS DAS GASOLINAS

### 2.2.1 Ensaios normatizados ANP

Através do PMQC da ANP, periodicamente, são feitas análises de combustíveis comercializados em todo o território nacional. As especificações a serem atendidas são aquelas estabelecidas na Resolução ANP N° 807, de 23 de janeiro de 2020, a qual determina que a avaliação da qualidade inclui parâmetros como aspecto visual e coloração, massa específica a 20°C, teor de etanol e metanol, destilação a 10, 50 e 90% evaporado, destilação no ponto final de ebulição (FBP), destilação residual, teor de enxofre e presença de marcador (TOSATO, 2020).

Para os parâmetros que necessitam de análise físico-química, existem os métodos de referência, como os apresentados na Tabela 1.

Tabela 1 – Métodos de referências para os parâmetros físico-químicos de análise de gasolina monitorados pelo PMQC.

<b>Característica</b>	<b>Unidade</b>	<b>Método de referência</b>
<b>Massa específica a 20 °C</b>	Kg/m <sup>3</sup>	NBR 7148, NBR 14065 / D1298, D4052
<b>Teor de etanol anidro combustível (EAC)</b>	% vol.	NBR 13992
<b>10% evaporado, máx.</b>	°C	NBR 9619 / D86
<b>50% evaporado, máx.</b>	°C	NBR 9619 / D86
<b>90% evaporado, máx.</b>	°C	NBR 9619 / D86
<b>FBP, máx.</b>	°C	NBR 9619 / D86

Fonte: O autor.

### 2.2.2 Massa específica

Devido à diversidade de composição química das gasolinas, a depender do petróleo de origem, processo de fabricação e outros fatores, pode existir uma gama enorme de valores para a massa específica (medida em kg/m<sup>3</sup>). As condições de estocagem podem afetar diretamente este parâmetro, mas também a presença de adulterantes pode ser indicada pelo valor de massa.

Os padrões determinados pela Resolução ANP N° 807, de 23 de janeiro de 2020, para massa específica de gasolinas são obtidos segundo os critérios estabelecidos pela norma brasileira ABNT NBR 14065.

### 2.2.3 Teor de etanol

As gasolinas tipo C comercializadas no Brasil devem ter adição de EAC, o que é regulado pela Portaria n° 75 do Ministério da Agricultura. Esta característica torna as gasolinas brasileiras facilmente detectadas em comparação com gasolinas provenientes de outros países, as quais, em sua maioria, não possuem adição de álcool.

Tal adição se deve à necessidade de baratear o combustível, à diminuição da emissão de monóxido de carbono (gás poluente) emitido pela combustão incompleta da gasolina e pelo aumento da octanagem do combustível. Isso porque o etanol atua como um antidetonante para a gasolina devido ao seu poder calorífico ser menor (RODRIGUES, 2017; PERRY, 1999), o que

substitui o uso do metil-terc-butil éter (MTBE), antes usado como antidetonante em gasolinas, mas que possui alta toxicidade e gera maior impacto aos componentes do motor de automóveis.

Atualmente, o percentual exigido de etanol anidro nas gasolinas brasileiras é de 27% na gasolina comum e de 25% na gasolina premium.

#### **2.2.4 Curvas de destilação**

Cada composto tem um ponto de ebulição específico, que se trata da temperatura, a uma dada pressão, em que a pressão de vapor do composto se iguala à pressão atmosférica e assim permite com que moléculas da fase líquida partam para a fase vapor. Neste processo, a temperatura permanece constante até que todo líquido seja transformado em vapor (SMITH, VAN NESS, ABBOTT, 2007). A gasolina, no entanto, se trata de uma mistura de diversos compostos químicos, cada um com ponto de ebulição específico, sendo assim, não existe para ela um único ponto de ebulição, mas sim uma curva de destilação, a qual é dada pela temperatura em função do percentual destilado em volume.

Nestas curvas, são adotadas as definições de “ponto inicial de ebulição” (IBP), correspondente à temperatura quando cai a primeira gota de condensado. E o “ponto final de ebulição” (FBP), que corresponde à maior temperatura observada na destilação, que não necessariamente é igual à destilação de 100% do volume devido às perdas e resíduos existentes no fundo do balão de destilação (CHASIB, KARTHIK, 2017; TAKESHITA, 2006).

Para fins de análise da qualidade das gasolinas comercializadas no Brasil, são utilizados apenas alguns pontos da curva de destilação, os quais correspondem às temperaturas a 10, 50 e 90% de evaporado. Estes pontos analisados como critério exigido pela ANP se referem à performance do motor com o combustível, visto que, se a gasolina for muito volátil, as temperaturas obtidas nos pontos determinados da curva serão inferiores aos especificados, o que pode alterar a eficiência do combustível. Assim como, se a gasolina for muito pesada, os valores da curva serão maiores, podendo ocasionar a formação de resíduos que geram depósitos nos componentes do motor, prejudicando a eficiência da combustão.

## 2.3 ESPECTROSCOPIA NA REGIÃO DO INFRAVERMELHO

### 2.3.1 A absorção no infravermelho

Na região do infravermelho, a faixa espectral pode ser dividida em 3 regiões: infravermelho próximo (NIR, do inglês *Near Infrared*), médio (MIR, do inglês *Mid Infrared*) e distante (FIR, do inglês *Far Infrared*). Na Tabela 2, podem ser vistas as faixas de comprimento de onda (e de número de onda) correspondentes a cada uma destas três regiões.

Tabela 2 – Regiões do infravermelho.

Faixa espectral	Comprimento de onda (nm)	Número de onda (cm <sup>-1</sup> )
NIR	780 – 2.500	12.800 – 4.000
MID	2.500 – 50.000	4.000 – 200
FAR	50.000 – 1.000.000	200 – 10

Fonte: Adaptado de Skoog, Holler e Crouch (2017).

Trata-se de uma radiação que possui baixa energia associada, que, ao interagir com a matéria, não permite a transição entre níveis eletrônicos, visto que tais transições exigem uma quantidade maior de energia. Esta radiação, no entanto, possibilita transições que requerem menor energia, tais quais as transições vibracionais e as rotacionais (PASQUINI, 2003; 2018). Dentre as formas de vibração podem ser citadas as vibrações de estiramento (que envolvem uma variação na distância entre os átomos ao longo do eixo da ligação química) e as deformações (que envolvem variações no ângulo entre duas ligações).

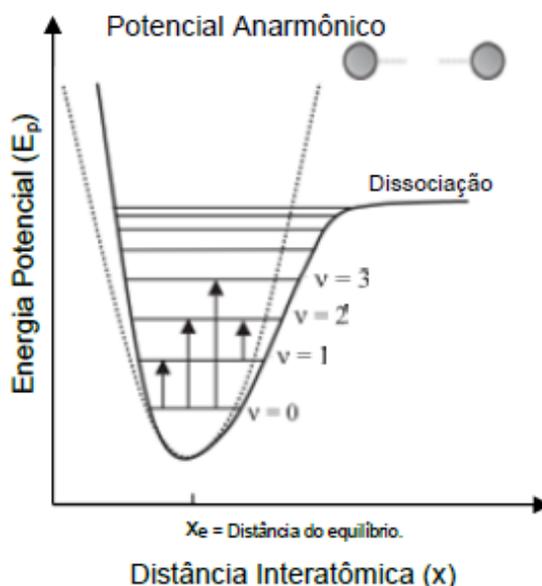
Essas transições dependem das massas dos átomos e das ligações químicas existentes, portanto, são únicas para cada tipo de ligação, podendo ser utilizadas para identificar grupos característicos das substâncias analisadas (PASQUINI, 2003; 2018). Para que essas transições vibracionais ocorram ao absorver radiação infravermelha, é necessário que aconteça uma variação no momento de dipolo decorrente do modo vibracional.

O modelo do oscilador harmônico, proposto pela mecânica clássica pode ser utilizado para explicar o mecanismo das vibrações citadas. Este associa as vibrações moleculares de estiramento a um sistema composto de duas massas ligadas por uma mola, fornecendo uma representação aproximada das vibrações das ligações químicas. Entretanto, este modelo se distancia da realidade e não explica alguns fenômenos observados na região espectral. Isso porque, uma molécula não pode

assumir todos os estados de energia, mas apenas valores discretos, os estados quânticos vibracionais (números quânticos  $v = 0, 1, 2, \dots$ ). Além do mais, a diferença de energia entre níveis vizinhos não é a mesma. As transições entre o estado fundamental e o nível quântico imediatamente superior são chamadas de transições fundamentais (PASQUINI, 2003; SKOOG, 2017).

O modelo que melhor descreve o movimento vibracional de uma molécula é o do oscilador anarmônico, o qual considera que as variações de energia entre níveis vizinhos não são necessariamente iguais, além de diminuir com o aumento do número quântico, como pode ser visto na Figura 3. No caso de um oscilador anarmônico são permitidas transições nas quais a variação do número quântico é maior que 1 (sobretens).

Figura 3 – Gráfico de energia potencial para o modelo do oscilador anarmônico.



Fonte: Silva, 2013.

Desta forma, fenômenos como os sobretens (bandas cujos valores de frequência são aproximadamente múltiplos inteiros dos valores das vibrações normais) e as bandas de combinações (resultado de combinações lineares de múltiplos de frequências normais), presentes nos espectros na região do infravermelho próximo, são mais bem explicados (PASQUINI, 2003; SILVA, 2013).

Como mostra a Tabela 2, a região do infravermelho próximo corresponde àquela entre os comprimentos de onda de 780 e 2500 nm. Os espectros obtidos no NIR possuem bandas de absorção atribuídas a sobretens e combinações entre vibrações de ligações de átomos de H e de C, O, N

e S. As bandas são em geral largas, de baixa intensidade (de 10 a 1000 vezes menor que no infravermelho médio) e resultantes de muitas sobreposições (PASQUINI, 2003).

A região NIR foi por muito tempo pouco explorada, visto que a interpretação dos espectros acaba sendo mais difícil devido a efeitos aditivos e multiplicativos (desvios de deslocamentos e de linha de base dos espectros), à sobreposição de informações derivadas de reflexão, refração, espalhamento e às informações vibracionais das ligações. Isso se deve às próprias características das amostras analisadas, tais quais tamanho, ao caminho óptico utilizado e à variação nas condições ambientais (como pressão e temperatura), especialmente em amostras sólidas (WORKMAN; WEYER, 2009). Sendo assim, com o avanço da tecnologia e do desenvolvimento da Quimiometria (essencial para interpretação das informações químicas contidas nos espectros, em especial as relacionadas às bandas de baixas intensidades), esta espectroscopia passou a ser mais explorada (PASQUINI, 2003).

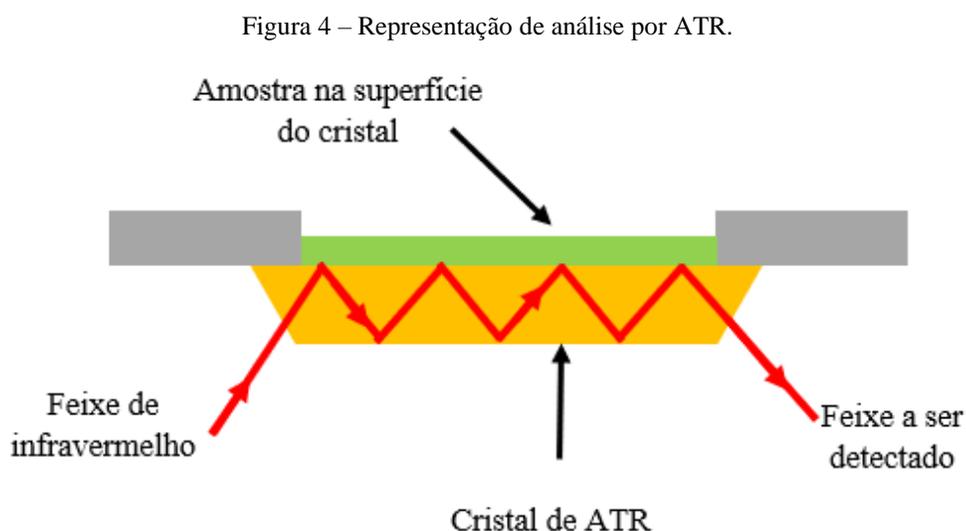
Uma vantagem da espectroscopia NIR está relacionada aos equipamentos, que são, em geral, mais baratos e robustos que os utilizados na região do infravermelho médio. Além disso, estes podem ser mais facilmente miniaturizados em instrumentos portáteis, o que facilita a obtenção dos espectros *in situ*, sem a necessidade de deslocar grandes instrumentos para o local da análise ou de deslocar amostras para o laboratório.

Já na região entre 2.500 e 50.000 nm, que corresponde ao infravermelho médio, são observadas, principalmente, bandas relacionadas a transições fundamentais, que são em geral mais estreitas e intensas que as bandas encontradas no espectro NIR. Também são notadas bandas de sobretons e combinações de grupos/moléculas envolvendo átomos pesados (PASQUINI, 2003; SILVA, 2013).

Devido às absorções mais intensas no MIR, os percursos óticos das cubetas utilizadas devem ser pequenos. Por esta razão, a técnica de reflexão total atenuada (ATR, do inglês *attenuated total reflection*) é muito empregada. Esta técnica se baseia na diferença de densidade entre dois meios os quais terão contato com a radiação incidente. Ao passar de um meio mais denso para um menos denso, parte da radiação sofre reflexão, sendo esta fração refletida dependente do ângulo de incidência. A partir de um certo ponto (ângulo crítico) a reflexão é total. No entanto, o feixe refletido se comporta como se penetrasse no meio menos denso antes de ser refletido e essa radiação penetrante é conhecida como onda evanescente. Em alguns casos, a onda evanescente é absorvida

pelo meio menos denso, o que causa atenuação do feixe de radiação, fenômeno este chamado de reflectância total atenuada (SKOOG, 2017).

A análise por esta técnica consiste, em geral, em múltiplas reflexões da radiação infravermelha no interior de cristais com alto índice de refração, ocorrendo interação dessa radiação com a amostra localizada na superfície do cristal, como mostra a Figura 4.



Fonte: O autor.

São utilizados cristais de diversos tipos, como os de seleneto de zinco (ZnSe), silício (Si), germânio (Ge), safira e um cristal misto de brometo de tálio/iodeto de tálio. Um cristal de diamante também pode ser utilizado, mas neste caso o tamanho é bem reduzido (e conseqüentemente o número de reflexões), por conta do alto custo.

## 2.4 INCERTEZA DE MEDIÇÃO

### 2.4.1 Erros e incertezas

Por mais rigorosas que sejam as medições de uma determinada análise, é inevitável a existência de erros, isto porque os instrumentos e métodos de análises utilizados estão sujeitos aos efeitos de inúmeros fatores que podem ocasionar um desvio do valor medido em relação ao valor tido como verdadeiro. Tais fatores podem ser: condições ambientais, como a temperatura, umidade e pressão; fatores referentes à não idealidade do comportamento físico ou químico dos reagentes e reações, o que pode ocasionar erros devido à instabilidade de espécies químicas, reações paralelas,

reações incompletas; erros pessoais que estão relacionados às limitações do analista; e ainda características intrínsecas dos equipamentos utilizados, como a resolução, sensibilidade e estabilidade.

Segundo o Vocabulário Internacional de Termos Básicos e Genéricos em Metrologia (INMETRO, 2012), incerteza é um parâmetro associado ao resultado de uma medição que caracteriza a dispersão de valores que precisam ser razoavelmente atribuídas ao mensurando. Ou seja, a incerteza reflete a falta de conhecimento do valor exato de uma medição, indicando uma estimativa deste valor (EURACHEM/CITAC, 2002).

É importante saber diferenciar erro de incerteza. O erro representa a distância entre o valor medido e o valor tido como verdadeiro, portanto, é uma forma numérica de dizer qual a variação em relação a este último. Já a incerteza é um parâmetro não negativo representado em forma de faixa em torno do valor medido, dentro da qual se garante (com a probabilidade desejada) que o valor real referente àquele medido se encontra. Este é, portanto, um indicador de qualidade, pois através do seu uso é possível avaliar o nível de confiança de uma determinada medição, bem como identificar as principais fontes de incerteza (GULLBERG, 2012; BUCKMANN, SARKIS, 2002).

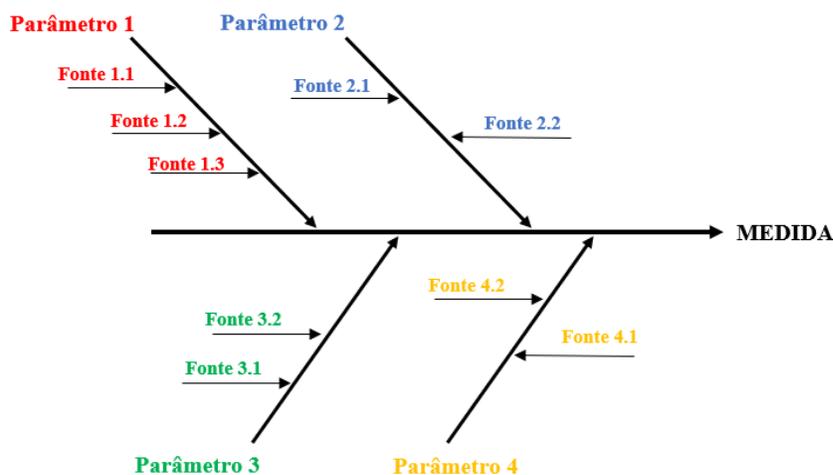
Sendo assim, os erros não representam as incertezas, mas são uma forma de tentar mensurar a incerteza associada a um determinado experimento.

#### **2.4.2 Estimativa da incerteza**

Dois abordagens são comumente utilizadas para a estimativa de incertezas. São elas a *top-down*, a qual é baseada em estudos interlaboratoriais, e a *bottom-up*, que consiste em identificar e quantificar as fontes de incerteza relevantes (OLIVEIRA, AGUIAR, 2009; WERNIMONT, 1985).

A estimativa com base na abordagem *bottom-up*, parte-se do princípio de que a incerteza de uma dada medição é obtida pela combinação de diversas fontes de incertezas, as quais precisam ser identificadas para, posteriormente serem estimadas e, por fim, combinadas. Tais fontes de incerteza podem ser identificadas utilizando-se o diagrama de Ishikawa, ou diagrama de causa e efeito, conforme mostrado na Figura 5, o qual é usado para determinar os parâmetros dos quais a medida depende e as fontes de incerteza que influenciam cada um desses parâmetros.

Figura 5 – Esquema de um Diagrama de Ishikawa para identificar as fontes de incerteza em resultado de medida.



Fonte: O autor.

Definidas e quantificadas as fontes de incertezas, a incerteza pode então ser obtida, considerando a contribuição de cada uma das fontes. Neste momento, se faz necessário o uso dos desvios padrão associados a fontes de incertezas, os quais são usados como contribuição da incerteza, essas são as chamadas incertezas padronizadas. Estas contribuições podem ser combinadas e expressa em termo de intervalo por uma incerteza expandida (EURACHEM/CITAC, 2002).

### 2.4.3 Estudos de Repetibilidade, reprodutibilidade e precisão intermediária

Duas das componentes das incertezas presentes em todas as medidas em química analítica podem ser obtidas pelos chamados estudos de R&R (repetibilidade e reprodutibilidade), os quais visam esclarecer o quanto de variação é gerada pelo mesmo operador e por vários operadores. Tais conceitos são de extrema importância para melhor entendimento das causas de variabilidades em experimentos e medições.

A repetibilidade representa a variação nas medições mantendo-se as mesmas condições de operação (mesmo operador, mesmo instrumento de medição, em curto intervalo de tempo). Nesse caso, sob uma mesma condição, o operador realiza várias medições de uma mesma amostra, usando o mesmo instrumento. A repetibilidade mede, portanto, quanta variabilidade no sistema de medição é causada pelo dispositivo de medição (ESCALANTE, 2010; GUTIÉRREZ, 2009).

Já a reprodutibilidade, mede o grau de concordância entre os resultados das medições quando há variação das condições de medida, tais quais diferentes operadores, laboratórios,

condições ambientais (temperatura, pressão, etc.) ou outros. É, portanto, comumente utilizada em estudos interlaboratoriais (EURACHEM/CITAC, 2002).

Assim como a reprodutibilidade, a precisão intermediária indica o efeito da variação de determinados eventos, sendo a mais representativa da variabilidade dos resultados em um mesmo laboratório (EURACHEM/CITAC, 2002; MAROTO *et.al*, 1999). A reprodutibilidade não pode ser calculada quando se considera apenas um laboratório, no entanto, a precisão intermediária sim. Por isso, é comum o uso desta última, (MAROTO *et.al*, 1999).

O uso desses parâmetros para estimar a incerteza de medição é fundamental, já que leva em consideração as variações advindas de erros associados a mudanças de condições, o que torna a análise mais robusta e representativa da realidade.

As contribuições dessas fontes de incertezas, juntamente com as demais fontes, determinadas em cada caso a ser estudado, são utilizadas para obter a incerteza combinada, utilizando a equação (1), em que  $u$  é a incerteza combinada e  $u_i$  ( $i = 1, 2 \dots n$ ) representa cada uma das fontes que contribuem para a incerteza, as quais podem ser dadas em valores relativos.

$$u = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2} \quad (1)$$

Em seguida, a incerteza combinada é corrigida com um fator de abrangência  $k$ , escolhido com base no nível de confiança desejado, a fim de obter a incerteza expandida, que é de fato considerada a incerteza dos dados estudados e que será utilizada como a faixa em torno do valor medido dentro da qual se garante encontrar o valor real da medição, conforme mostra a equação (2).

$$ue = u * k \quad (2)$$

## 2.5 TÉCNICAS QUIMIOMÉTRICAS MULTIVARIADAS DE ANÁLISE

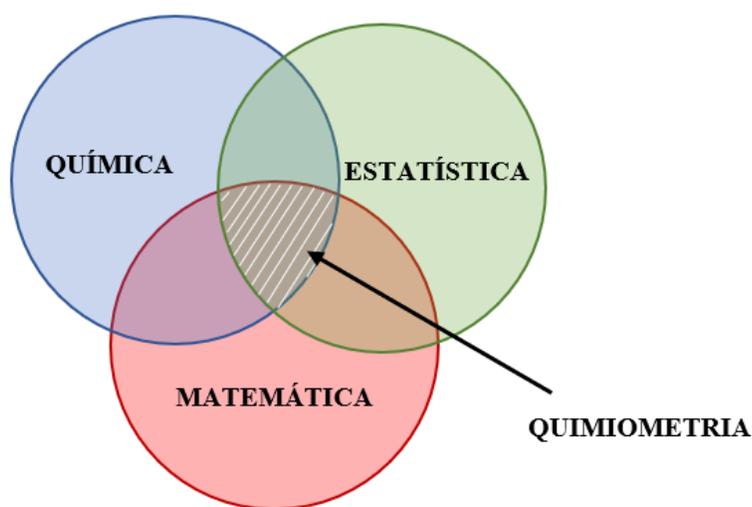
### 2.5.1 Quimiometria

Já na segunda metade do século XX, com o desenvolvimento de novas técnicas de análise e disseminação do uso de instrumentos como espectrômetros e cromatógrafos em análises de rotina nos laboratórios, a geração de dados analíticos de natureza química atingiu um patamar nunca visto.

No entanto, boa parte desses dados não era devidamente explorada, devido à falta de procedimentos capazes de organizar e extrair informações relevantes de um volume tão grande de dados, visto que os instrumentos utilizados não apresentam as informações de natureza química de maneira explícita, exigindo interpretação dos resultados apresentados (FERREIRA, 2015; OTTO, 2017).

Foi necessário aliar conhecimentos de matemática e estatística à química, aproveitando-se do uso de computadores modernos e capazes de processarem os conjuntos de dados estudados. Surgiu, então, uma área da química hoje conhecida como Quimiometria a qual consiste na aplicação de ferramentas matemáticas e estatísticas a dados de origem química a fim de extrair o máximo de informação relevante (KOWALSKI, 1975; BEEBE, 1998). O diagrama apresentado na Figura 6 mostra como a Quimiometria é uma área multidisciplinar cuja aplicação envolve diversos tipos de conhecimento.

Figura 6 – Áreas de conhecimento que compõem a Quimiometria.



Fonte: O autor.

O uso de técnicas quimiométricas possibilita um planejamento e seleção de experimentos de forma otimizada, o que resulta em uma análise profunda do conjunto de dados analíticos estudado, obtendo assim maior extração de informação química útil. Além disso, permite um aumento da confiabilidade das conclusões, tendo em vista o uso de ferramentas estatísticas avançadas, as quais são fundamentais na interpretação de dados multivariados (em que são analisadas múltiplas variáveis por amostra/observação, em um dado experimento).

A quantidade de técnicas quimiométricas é enorme, o que inclui desde técnicas de pré-processamento dos dados, a métodos de calibração (tais quais calibração univariada e multivariada), e métodos supervisionados e não supervisionados de reconhecimento de padrões.

### 2.5.2 Organização dos dados e Técnicas de Pré-processamento

Em determinado estudo, após realizados os experimentos, o conjunto de dados obtido precisa ser organizado de modo a permitir a aplicação de técnicas quimiométricas. Desta maneira, os dados são organizados em forma de uma matriz  $\mathbf{X}$ , em que cada linha corresponde a uma amostra e cada coluna a uma variável (FERREIRA, 2015), como mostra a Equação (3). No caso de espectros, cada coluna corresponde aos valores de absorbância em um determinado comprimento de onda.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & \dots & x_{1J} \\ x_{21} & x_{22} & \dots & \dots & x_{2J} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ x_{I1} & x_{I2} & \dots & \dots & x_{IJ} \end{bmatrix} \quad (3)$$

Antes de iniciar as análises dos dados, eles precisam ser previamente visualizados (por meio de gráficos) a fim de verificar como estão distribuídos e identificar se existem amostras destoantes das demais, com aspecto destinto das outras, ou muito deslocadas em relação ao conjunto de dados. Para tais amostras com um comportamento diferente, é necessário cuidado nas análises, exigindo em algumas situações que elas sejam excluídas.

Em seguida, quando necessário, os dados são pré-processados para que sejam reduzidas contribuições indesejadas que não estão relacionadas à propriedade de interesse das amostras estudadas. Isso pode ser devido ao fato de que os sinais medidos pelos instrumentos são resultados da somatória entre o sinal verdadeiro da amostra e o ruído aleatório existente. Este ruído pode causar efeitos visíveis nos dados. Outra situação, no caso de espectros de infravermelho, por exemplo, é o deslocamento no eixo das ordenadas (deslocamento da linha de base) (BEEBE; PELL; SEASHOLTZ, 1998; FERREIRA, 2015). Sendo assim, o uso de estratégias matemáticas para minimizar o efeito dessas contribuições indesejadas não altera as informações significativas contidas nos dados, e permite a realização de uma análise mais fidedigna, visto que serão retiradas as contribuições de variabilidade não relacionadas à informação química de interesse.

Os pré-processamentos podem ser aplicados tanto às amostras, ou seja, a cada linha da matriz de dados  $\mathbf{X}$ , quanto às variáveis, ou seja, a cada coluna dessa matriz. Dentre as técnicas de pré-processamento, além da centragem na média, podem ser citadas as técnicas de alisamento (pelo método da média, pelo método de Savitzky-Golay, com filtros de Fourier), as técnicas de correção da linha de base, como Correção Multiplicativa de Espalhamento (MSC), Padronização Normal de Sinal (SNV). A seguir, serão detalhadas as técnicas de centragem na média e SNV, aplicadas no desenvolvimento deste trabalho.

### 2.5.2.1 Centragem na média

A centragem na média é uma técnica de pré-processamento aplicada às variáveis, a qual envolve uma translação de eixo em relação ao valor médio. A operação matemática consiste em obter o valor médio ( $\bar{x}_j$ ) de cada coluna da matriz  $\mathbf{X}$ , como mostra a Equação (4), e então subtrair esse valor de cada um dos valores da coluna ( $x_{ij}$ ), obtendo assim o valor centrado na média ( $x_{ij}^{CM}$ ), como mostra a Equação (5) (FERREIRA, 2015; GEMPERLINE, 2006).

$$\bar{x}_j = \frac{1}{I} \sum_{i=1}^I x_{ij} \quad (4)$$

$$x_{ij}^{CM} = x_{ij} - \bar{x}_j \quad (5)$$

### 2.5.2.2 Padronização Normal de Sinal

A transformação SNV é uma técnica de pré-processamento aplicada às amostras, a qual tem a função de corrigir tanto efeitos aditivos quanto multiplicativos resultantes do espalhamento da radiação causado por fenômenos físicos. Como exemplo podemos citar mudanças no caminho óptico e variações na temperatura, pressão e outros fatores, os quais são dependentes dos comprimentos de onda e são fontes de variabilidade irrelevantes para o problema, já que não têm a ver com a composição química das amostras e, sendo assim, precisam ser removidos.

Esta transformação consiste em autoescalar cada linha da matriz original de dados (FERREIRA, 2015). Ou seja, para cada linha da matriz  $\mathbf{X}$ , é calculada a média ( $\bar{x}_i$ ) e esta é subtraída de

todos os valores da respectiva linha ( $x_{ij}$ ), sendo o resultado dividido pelo seu desvio padrão ( $\sigma_i$ ), como é representado pela Equação (6).

$$x_{ij}^{SNV} = \frac{(x_{ij} - \bar{x}_i)}{\sigma_i} \quad (6)$$

### 2.5.3 Análise exploratória dos dados

Após os pré-processamentos, os dados passam agora para a etapa de análise, cujo objetivo é utilizar estratégias matemáticas para representar os dados em gráficos bidimensionais e tridimensionais para que então seja possível identificar, ou não, a existência de padrões. Esses métodos podem ser classificados em dois grupos: métodos supervisionados e não supervisionados.

Os métodos não supervisionados, também chamados de análise exploratória, são aqueles em que os dados são analisados sem a necessidade de se construir uma regra preditiva. São exemplos a Análise de Componentes Principais (PCA) e a Análise Hierárquica de Agrupamentos (HCA). Já os métodos supervisionados são aqueles em que os dados são dispostos em classes (uma ou mais de uma), sendo para tanto construídos modelos de classificação. Ou seja, os métodos supervisionados, diferentemente dos não supervisionados, requerem conhecimento prévio sobre a classificação das amostras estudadas (BEEBE; PELL; SEASHOLTZ, 1998; FERREIRA, 2015).

#### 2.5.3.1 Análise de Componentes Principais (PCA)

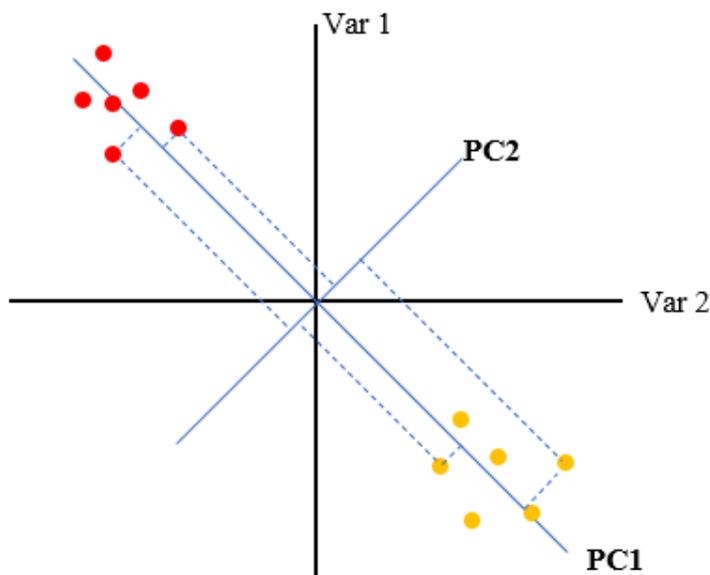
A Análise de Componentes Principais (ou PCA, do inglês *Principal Component Analysis*) tem como objetivo reduzir a dimensionalidade do espaço de um conjunto de dados multivariados ao projetá-los em um espaço de menor dimensão, por meio de uma transformação linear ortogonal que cria um novo sistema de coordenadas. Esta redução é chamada de “compressão dos dados”, o que se consegue fazer ao realizar combinações lineares das variáveis originais, agrupando as que fornecem informações semelhantes (BRERETON, 2003; FERREIRA, 2015).

O novo sistema de coordenadas é obtido no sentido da maior variância dos dados, sendo o eixo na direção que descreve a máxima variância dos dados originais a componente principal 1 (PC1). Já a componente principal 2 (PC2) é aquela na direção de máxima variância dos dados ortogonal à PC1. A PC3 é na direção da máxima variância sendo ortogonal às duas PCs anteriores

e assim por diante, em que o número de Componentes Principais é, no máximo, igual ao número de variáveis.

A Figura 7 mostra imagens que representam a utilização da PCA. É possível ver o conjunto de dados com suas variáveis originais Var 1 e Var 2, assim como as duas componentes principais PC1 e PC2. As coordenadas dos dados projetados nos novos eixos (PCs) são chamadas de scores, e o cosseno dos ângulos entre o eixo das componentes principais e as variáveis originais são chamados de peso ou *loadings* (BEEBE; PELL; SEASHOLTZ, 1998; BRERETON, 2003).

Figura 7 – Exemplo de análise de componentes principais, apresentando o conjunto de dados em suas coordenadas originais e eixos da PC1 e PC2;



Fonte: Adaptado de Ferreira, 2015.

## 2.6 USO DE ESPECTROSCOPIA NO INFRAVERMELHO ASSOCIADA À QUIMIOMETRIA PARA CLASSIFICAÇÃO DE GASOLINAS

Quanto à utilização da espectroscopia na região do infravermelho associada a técnicas quimiométricas com a finalidade de classificar e/ou discriminar amostras de gasolinas, alguns trabalhos foram encontrados na literatura.

Silva e colaboradores (2015) utilizaram a espectroscopia na região do infravermelho médio (MIR) para o desenvolvimento de modelos multivariados, através das técnicas de PLS-DA e SIMCA, para classificar gasolinas brasileiras e venezuelanas (SILVA *et.al*, 2015). Em uma das estratégias utilizadas foi possível classificar amostras como sendo provenientes do Brasil, da

Venezuela ou de nenhum destes países, metodologia esta que pode ser aplicada para outros problemas forenses que envolvam classificação química de um produto.

Em um trabalho de 2013, Khanmohammadi e seu grupo desenvolveram uma metodologia para classificação de amostras de gasolinas de refinarias (em aprovadas e não aprovadas, conforme os requisitos do Controle de Qualidade da indústria), utilizando espectroscopia de Reflectância Total Atenuada associada ao método de Análise Discriminante Quadrática (QDA) (KHANMOHAMMADI et al., 2013). Já em 2018, Maldonado e colaboradores apresentaram um sensor portátil de baixo custo desenvolvido para identificação de gasolinas puras e com aditivos e adulterantes, com tomada de espectros na região do infravermelho médio. Aos *scores* da PCA obtida através dos dados coletados pelo sensor, foi aplicada uma Análise Multivariada de Variância (MANOVA). Os resultados demonstraram que a técnica foi capaz de discriminar entre gasolina, óleo diesel e os álcoois metanol, etanol, n-propanol, n-butanol e n-hexanol e que tem potencial para ser aplicada na indústria de combustíveis (MALDONADO et al., 2018).

Também podem ser encontrados alguns trabalhos utilizando espectroscopia na região do infravermelho próximo. Por exemplo, foram utilizadas as técnicas de Análise Discriminante Linear (LDA), Modelagem Independente e Flexível por Analogia de Classes (SIMCA) e uma rede neural do tipo multicamadas de neurônios (MLP – *Multilayer Perception*) para classificar 382 amostras de gasolinas pela origem (provenientes de refinarias ou de petroquímicas), concluindo que todos os métodos foram efetivos na classificação de gasolinas, mas o uso de MLP foi o que apresentou melhores resultados (BALABIN, SAFIEVA, 2008). Em 2010, Balabin e colaboradores desenvolveram um trabalho com o objetivo de comparar o uso de diversas técnicas multivariadas para classificação de gasolinas usando a espectroscopia NIR. Foram utilizadas as técnicas de LDA, QDA, SIMCA, Análise Discriminante Regularizada (RDA), PLS, KNN, MLP e outras, chegando à conclusão de que as técnicas de KNN, SVM e PNN (*probabilistic neural network*) foram as melhores em termos de classificação, enquanto o uso de redes neurais foi o que gerou os piores resultados.

Ferreiro-González *et. al* (2015) descreveram o uso de espectroscopia de massas por heads-pace e espectroscopia no NIR aliadas às técnicas quimiométricas de LDA e Análise Hierárquica de Agrupamentos (HCA) para classificação de gasolinas pela octanagem (duas classes foram estudadas: uma com octanagem de 95 e outra com 98). Os melhores resultados foram obtidos aplicando LDA, o que permitiu a completa discriminação das amostras de gasolina. Assim como este trabalho, um estudo mais recente, desenvolvido por Barea-Sepúlveda e colaboradores (2021) também

apresentou uma comparação entre espectroscopia de massas por *headspace* e espectroscopia no NIR para discriminar gasolinas de acordo com seu índice de octano (RON). As técnicas quimiométricas utilizadas foram HCA, Máquinas de Vetores de Suporte (SVM) e Floresta Aleatória (RF). A maior acurácia na classificação foi obtida com o uso da RF aplicada aos dados fundidos (espectroscopia de massas por *headspace* + espectroscopia no NIR) em detrimento do uso apenas dos dados NIR.

Recentemente, foi utilizado um espectrômetro NIR portátil para a discriminação de diversos tipos de óleos e combustíveis, incluindo gasolina, nafta, diesel, querosene e óleos crus. Para os espectros NIR de nafta, gasolina, diesel e querosene, PLS-DA permitiu a identificação de qualquer um desses produtos com sensibilidade, especificidade e precisão de 100% (SANTOS *et. al*, 2021). Já Correia e colaboradores (2017) realizaram um estudo da aplicação de um espectrômetro NIR portátil para o controle de qualidade de combustíveis (dentre eles, a gasolina), utilizando técnicas quimiométricas como PLS e PLS-DA.

Outros trabalhos trataram da classificação de gasolinas utilizando técnicas quimiométricas em dados provenientes de diversos métodos espectrométricos, tais quais: Análise direta em espectrometria de massa em tempo real (DART-MS), associada a técnicas quimiométricas como PLS-DA e PCA (BARNETT, 2018; ZHAO *et. al*, 2021); e cromatografia gasosa acoplada a espectroscopia de massas (GC/MS) associada a métodos de PCA, LDA, Análise de Variável Canônica (CVA) e Análise de Variável Ortogonal Canônica (OCVA) (PETRACO *et. al*, 2008). Este último revelou que a variabilidade na população amostral foi suficiente para distinguir todas as amostras umas das outras conhecendo seus grupos. Os autores ainda informam que os resultados deste estudo inicial ajudaram a desenvolver procedimentos para a aplicação de análise multivariada para trabalhos em casos de destroços de incêndio. Outros trabalhos também utilizaram cromatografia gasosa ou GC/MS, os quais fizeram uso de PCA, redes neurais artificiais, SIMCA e outras técnicas (DOUBLE, 2003; RUDNEV, 20011; FLUMIGNAN *et. al*, 2007, 2008, 2010).

A espectroscopia Raman também foi aplicada a amostras de gasolina para fins de classificação quanto à origem. Neste trabalho, foram utilizadas 128 amostras de gasolina fornecidas por três refinarias diferentes e pertencentes a três marcas diferentes. As estratégias quimiométricas aplicadas aos dados foram PCA, Máquina de vetor de suporte de mínimos quadrados (LSSVM) e LDA. Os resultados mostraram que a espectroscopia Raman é um meio eficaz de classificar amostras de gasolina quanto à marca e à origem (LI, 2012).

Além desses trabalhos diretamente voltados para a classificação/discriminação de gasolinas, trabalhos com fins forenses de identificação e quantificação de adulterantes em amostras de gasolinas utilizando espectroscopia no infravermelho e quimiometria também vem sendo desenvolvidos. Mabood e colaboradores, em trabalho publicado em 2017 (MABOOD et al, 2017), utilizaram a espectroscopia NIR associada a técnica PLS-DA e PLS com a finalidade de detectar e quantificar adulterações de gasolinas Super Premium 95 com gasolinas Premium 91, o que resultou em um método simples, rápido e capaz de quantificar menos de 2% da adulteração de etanol em gasolina. O trabalho de Vempapatu e Kanaujia, de 2017, faz uma ampla revisão sobre os métodos utilizados para a identificação de adulteração em combustíveis derivados do petróleo (incluindo gasolina), tratando do uso de métodos físico-químicos, cromatográficos e espectroscópicos, assim como as estratégias quimiométricas associadas a esses métodos (VEMPAPATU e KANAUIA, 2017).

Em comparação aos trabalhos descritos, o presente trabalho, que objetivou utilizar a espectroscopia na região do infravermelho associada a técnicas quimiométricas com a finalidade de reconhecer amostras de gasolinas de diferentes procedências, apresenta aspectos inovadores ao avaliar o potencial de separação de amostras de gasolinas em pares com o objetivo de elucidar casos forenses. Foram utilizados dados espectrais obtidos tanto na região do infravermelho médio, quanto na região do infravermelho próximo, fazendo uso de um equipamento de bancada e um portátil. Outro aspecto inovador está relacionado à estimativa das incertezas nas medidas espectrais e ao desenvolvimento de uma metodologia para geração computacional de espectros com base nas incertezas em cada comprimento de onda.

### **3 METODOLOGIA**

#### **3.1 ANÁLISES FÍSICO-QUÍMICAS E ESPECTROSCÓPICAS DAS GASOLINAS**

##### **3.1.1 Obtenção das amostras**

As amostras utilizadas foram cedidas pelo Laboratório de Combustíveis da UFPE (LAC), através do PMQC, correspondentes a gasolinas comuns de postos que comercializam combustíveis no estado de Pernambuco.

Para o PMQC, o estado de Pernambuco é dividido em 9 regiões, das quais são coletadas diversas amostras para análises dos combustíveis de acordo com metodologia estabelecida pela ANP. Para este trabalho, foram selecionadas aleatoriamente 6 amostras de cada uma das 9 regiões do estado, as quais tiveram seus conteúdos distribuídos parte para as análises normatizadas de rotina, conforme normativas da ANP, e parte fracionada para análises espectroscópicas.

Tais amostras foram reservadas em geladeiras em temperatura em torno de 3,0 °C até o momento das análises.

##### **3.1.2 Análises convencionais (ANP) e curvas de destilação**

Cada amostra foi analisada conforme métodos de referência exigidos pela ANP (Tabela 1) a partir dos quais foi possível obter dados referentes à massa específica, ao teor de etanol anidro combustível e à curva de destilação.

Para fins de análise da qualidade de gasolinas, a ANP considera, como já citado, apenas alguns pontos da curva de destilação. No entanto, os destiladores atmosféricos utilizados (OptiDist, da PAC-Herzog) coletam em seu sistema todos os pontos da curva, o que permitiu a utilização de todos estes pontos da curva de destilação.

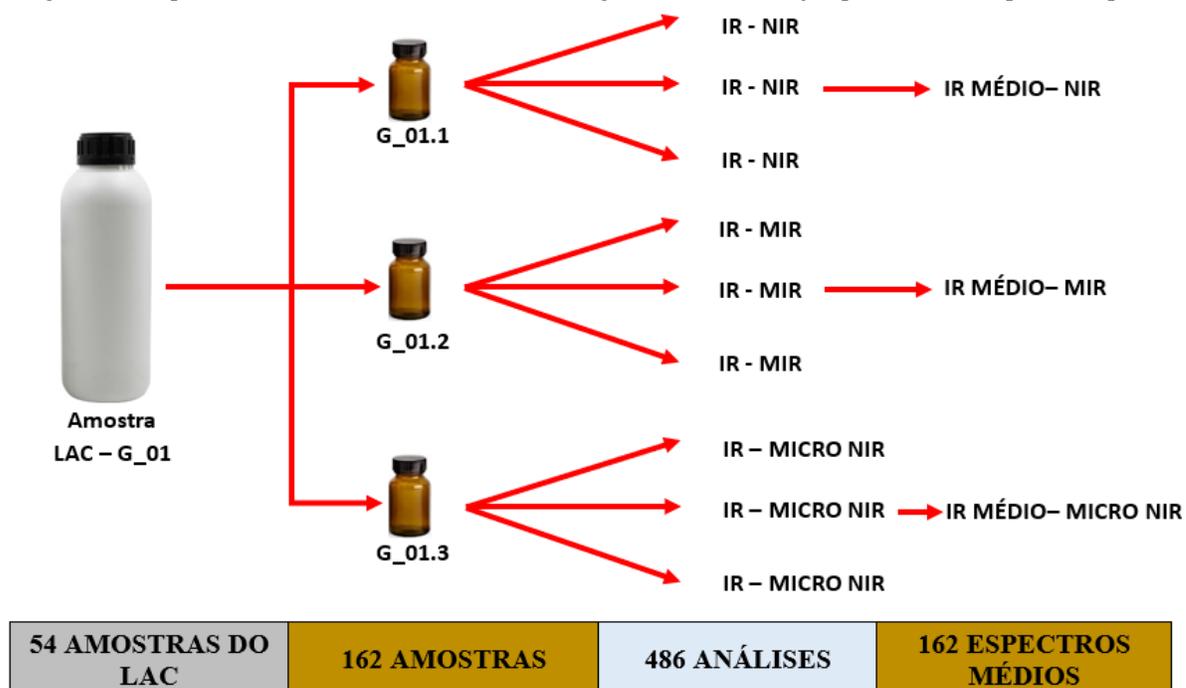
##### **3.1.3 Análises por espectroscopia no infravermelho**

Parte do conteúdo das 54 amostras coletadas pelo LAC nas 9 regiões de Pernambuco foi destinada para fracionamento em 3 grupos de frascos os quais, posteriormente, foram utilizados para aquisição dos espectros na região do infravermelho próximo (em dois tipos de instrumento, um portátil e um de bancada) e médio (em instrumento de bancada).

Todas as amostras dos 3 grupos de frascos de gasolina foram analisadas em triplicata, como mostra esquema da Figura 8, totalizando, ao final das análises, 486 espectros obtidos. Para cada

grupo de triplicatas foi calculado um espectro médio, o que resultou em um total de 162 espectros médios a serem analisados, sendo 54 para cada um dos 3 equipamentos utilizados.

Figura 8 – Esquema do fracionamento de amostras de gasolinas e destinação para análises espectroscópicas.



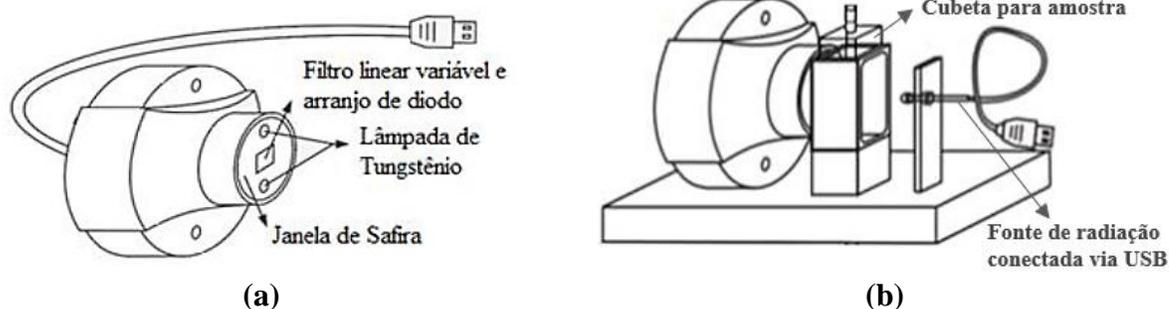
Fonte: O autor.

### 3.1.3.1 Aquisição dos espectros no infravermelho próximo em equipamento portátil

Foram obtidos espectros na região NIR utilizando o instrumento portátil MicroNIR™ PRO 1700 (da VIAVI Solutions Inc.), o qual opera na faixa de 11012 a 5965  $\text{cm}^{-1}$  (908,1 a 1676,2 nm), com resolução espectral de 6,25 nm, média de 50 varreduras e tempo de integração de 5.000  $\mu\text{s}$ .

Esse equipamento é composto por duas lâmpadas de tungstênio como fonte de radiação interna, um filtro variável linear (LVF), usado como elemento dispersivo, e um detector de 128 elementos (InGaAs). A Figura 9 mostra uma representação deste equipamento, assim como o acessório em que se acopla para a realização das medidas (PAIVA et al., 2015).

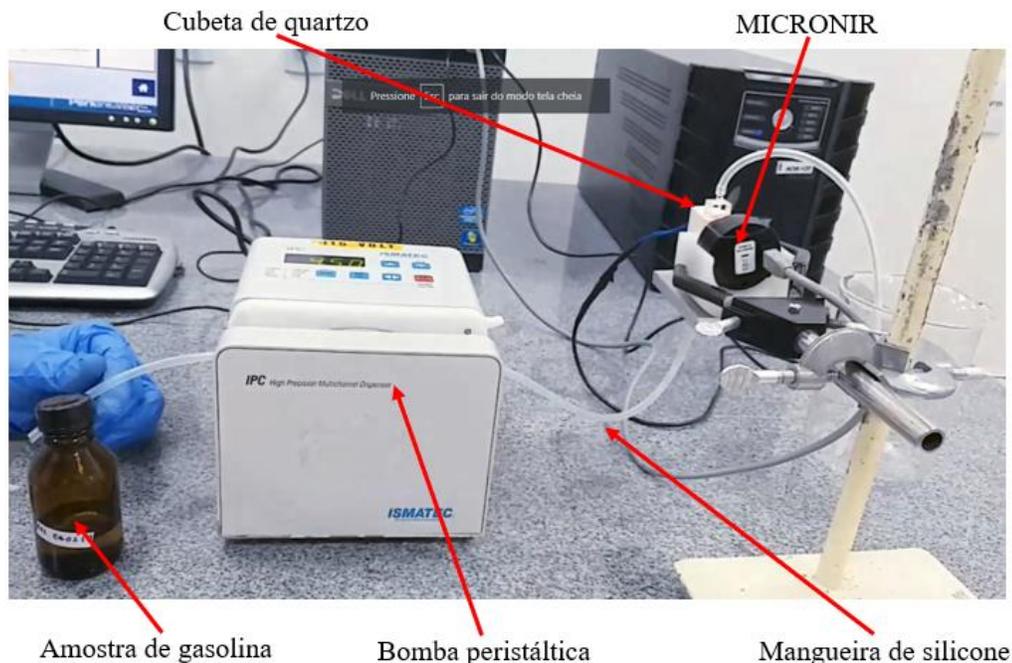
Figura 9 – Representação do equipamento portátil MicroNIR 1700 (a) e equipamento acoplado a acessório de transmitância (b).



Fonte: Adaptado de Paiva *et. al* (2015).

Para a realização das medidas, foi montado o aparato constando do equipamento, acessório de aquisição de espectros por transmitância, ambos conectados a um computador via cabo USB, e cubeta de quartzo com caminho óptico de 10 mm. As amostras de gasolinas a serem analisadas preenchem a cubeta de quartzo por meio de uma mangueira de silicone conectada a uma bomba peristáltica (Ismatec IPC – 4). Esse sistema é apresentado na Figura 10.

Figura 10 - Aparato para a realização de análises de gasolina pelo MicroNIR.



Fonte: O autor.

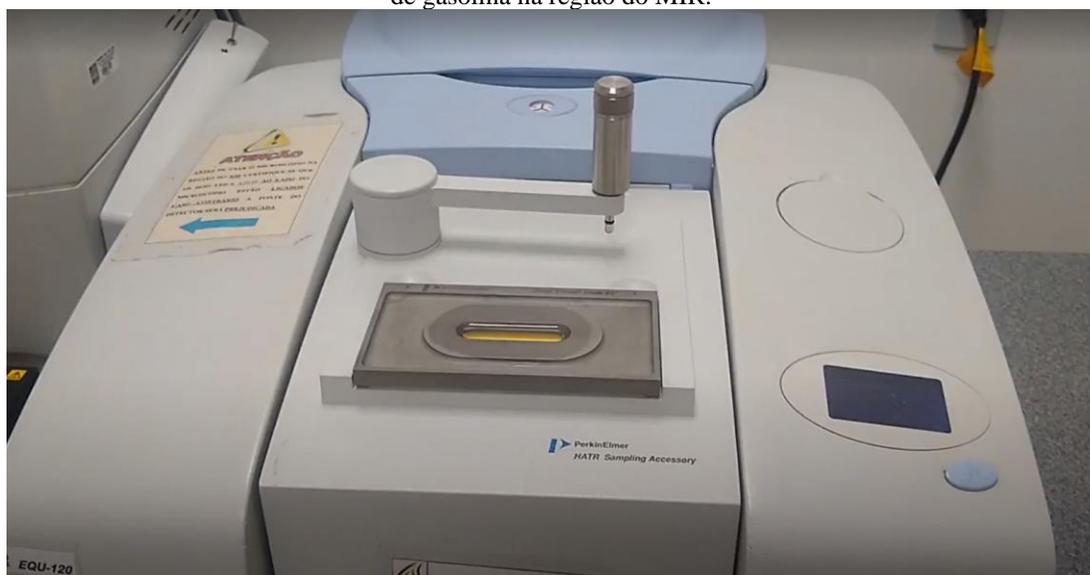
Antes da realização das medições propriamente ditas das gasolinas, é necessária aquisição do espectro de background (“dark”), correspondente a 0% de transmitância, obtido através da



### 3.1.3.3 Aquisição dos espectros no infravermelho médio em equipamento de bancada

Os espectros na região do MIR foram obtidos em um instrumento de bancada, o SPECTRUM 400, da Perkin Elmer, com Transformada de Fourier (FT-IR), com aquisição de espectros na faixa de  $4000$  a  $650\text{ cm}^{-1}$ , resolução de  $4\text{ cm}^{-1}$ , média de 16 varreduras e medidas de absorvância a cada  $1\text{ cm}^{-1}$ . Empregou-se o acessório de reflectância total atenuada horizontal (HATR) com cristal de seleneto de zinco, como mostra a Figura 12.

Figura 12 – Acessório de reflectância total atenuada acoplado no equipamento utilizado nas análises de gasolina na região do MIR.



Fonte: O autor.

Para a realização das medições, a amostra de gasolina foi colocada sob o cristal de seleneto de zinco utilizando uma pipeta de Pasteur (para cada amostra foi utilizada uma pipeta nova, não contaminada com outras amostras). Finalizada a aquisição dos espectros, a amostra foi retirada com o auxílio de um papel toalha (com a finalidade de retirar o excesso de gasolina e não danificar o cristal) e o cristal foi limpo com um algodão embebido de acetona. Em seguida, nova amostra foi posta sob o cristal de seleneto de zinco para repetição deste procedimento.

## 3.2 CÁLCULO DE INCERTEZAS E GERAÇÃO DE AMOSTRAS VIRTUAIS

Neste trabalho, foi utilizada uma estratégia computacional para a geração de amostras virtuais (considerando as análises convencionais, curvas de destilações e espectros no infravermelho),

ou seja, amostras obtidas de forma artificial, que representam a possível variabilidade considerando as incertezas de medição.

Para que isso fosse possível, as amostras geradas virtualmente deveriam estar compreendidas nos intervalos estimados de acordo com as incertezas das respectivas técnicas de medição. Assim, inicialmente, foram realizadas as estimativas de incertezas para posterior uso no cálculo das amostras virtuais.

### 3.2.1 Análises convencionais (ANP)

Por se tratar de métodos normatizados pela ANP, já bem estabelecidos e estudados, assim como pelo fato de o Laboratório de combustíveis da UFPE ter acreditação pelo INMETRO para esses ensaios segundo a NBR ISO 17025, as análises já contavam com as incertezas estimadas.

A Tabela 3 mostra os valores de incertezas para cada uma das grandezas medidas pelos métodos normatizados, estimadas de acordo com o ISO GUM (BIPM, 2008).

Tabela 3 – Valores de incertezas para as grandezas medidas pelos métodos normatizados estimadas segundo o ISO GUM.

<b>Grandeza</b>	<b>Unidade</b>	<b>Incerteza expandida (95%)</b>
<b>10% evaporado, máx. (D10)</b>	°C	0,9
<b>50% evaporado, máx. (D50)</b>	°C	0,8
<b>90% evaporado, máx. (D90)</b>	°C	1,4
<b>TPF, máx. (FBP)</b>	°C	2,1
<b>Resíduo da destilação (RES)</b>	mL	0,3
<b>Massa específica a 20 °C (ME)</b>	Kg/m <sup>3</sup>	0,55
<b>Teor de etanol anidro combustível (EAC) (T.ET)</b>	% vol.	1,0

Fonte: O autor.

### 3.2.2 Curvas de destilação

Conforme mostrado na seção anterior (incertezas para as análises convencionais), os valores de incerteza para os pontos da curva de destilação, determinados pela ANP como padrão para o controle de qualidade de gasolinas já estavam definidos. No entanto, para a realização deste

trabalho, também foram utilizadas as curvas de destilação por completo, o que inclui pontos para os quais não foram estimadas as incertezas.

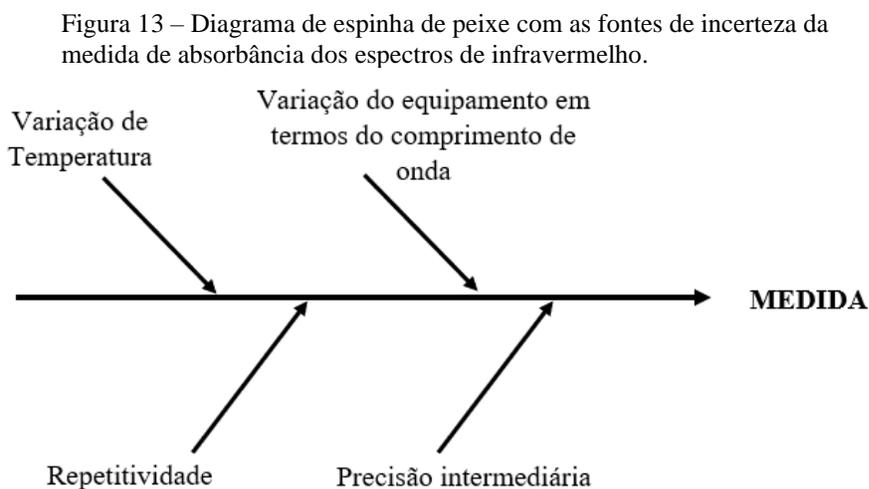
Desta maneira, foi estabelecida uma estratégia para estimativa da incerteza em cada ponto da curva. Para os pontos de 0 a 49% de volume destilado, foi utilizada a incerteza equivalente a 10% de destilado. Para os valores entre 50 e 90%, foi feita uma aproximação por meio de uma função de primeiro grau entre os valores de incerteza já calculados para esses dois pontos, respectivamente de 0,8 e 1,4 °C. Para os pontos entre 91 e 99% de destilados, foi utilizada a aproximação para o valor de incerteza a 90% de destilado.

### 3.2.3 Análises por espectroscopia no infravermelho

Foi utilizado o Guia EURACHEM/CITAC (2002) para determinação de incertezas em medições analíticas, cuja metodologia determina que o processo de estimativa da incerteza de medição é dividido em 4 etapas: especificação do mensurando, identificação das fontes de incerteza, quantificação dos componentes de incerteza e cálculo da incerteza combinada.

#### 3.2.3.1 Especificação do mensurando e identificação das fontes de incerteza

Sabendo que o mensurando para o qual se deseja obter as incertezas é a absorvância em cada comprimento de onda, foi realizado um *brainstorming* para a definição das principais fontes de incerteza que afetam a medição. Foram elas a repetibilidade, a precisão intermediária, a variação de temperatura da amostra e a variação do equipamento em termos do comprimento de onda, como mostrado no diagrama de espinha de peixe (ou Diagrama de Ishikawa) da Figura 13.

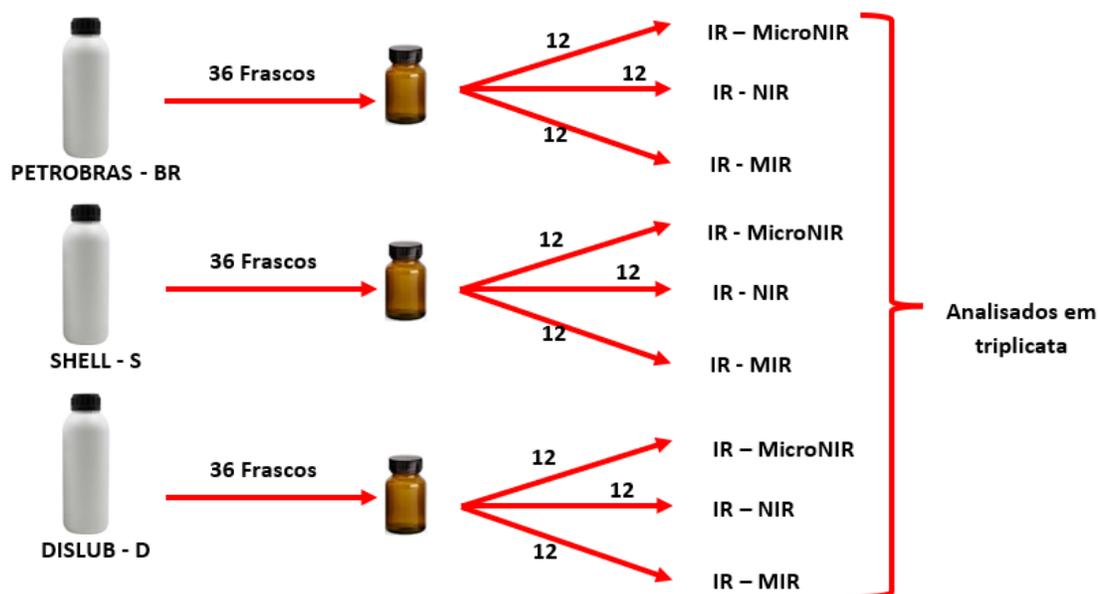


Fonte: O autor.

### 3.2.3.2 Quantificação dos componentes da incerteza

Cada uma das fontes foi quantificada. Para o cálculo da **repetibilidade e precisão intermediária**, foi utilizado um conjunto de dados composto por amostras de gasolinas provenientes de 3 marcas comerciais distintas (Petrobras – BR, Shell – S e Dislub – D). Tais amostras foram divididas em 108 frascos (36 para cada tipo de gasolina), os quais foram armazenados em uma geladeira em temperatura em torno de 3°C, a fim de que fossem conservados. A Figura 14 mostra um esquema de distribuição das amostras utilizadas.

Figura 14 – Esquema indicando a distribuição das amostras utilizadas para os cálculos da repetibilidade e da precisão intermediária.



Fonte: O autor.

Ao longo de 12 dias, foram analisadas 3 amostras de cada tipo de combustível (BR, S e D), uma amostra para cada um dos 3 equipamentos, totalizando 9 amostras analisadas por dia. Os espectros foram adquiridos em triplicata para cada amostra, totalizando 9 medidas diárias por equipamento, ou seja, 27 espectros obtidos em um dia.

Os espectros foram inicialmente pré-processados usando SNV e, em seguida, foram realizados os seguintes cálculos para a obtenção das incertezas. Para a estimativa da precisão intermediária foram calculadas médias das triplicatas dos espectros de cada amostra, o espectro médio, dado por  $\bar{X}$ , e estimado o desvio padrão em relação a essas médias para cada uma das marcas de gasolina usando a equação (7), em que  $X_i$  representa cada uma das triplicatas.

$$PI_{(BR, S \text{ ou } D)} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} \quad (7)$$

Os valores calculados para cada marca foram então combinados utilizando a equação (8), chegando então ao valor da precisão intermediária final ( $u_1$ ), na qual  $\gamma_i$  ( $i = 1, 2, 3$ ) representam os graus de liberdade das replicatas.

$$u_1 = \sqrt{\frac{\gamma_1 * PI_{BR}^2 + \gamma_2 * PI_S^2 + \gamma_3 * PI_D^2}{\gamma_1 + \gamma_2 + \gamma_3}} \quad (8)$$

Já no caso da contribuição da repetibilidade foram calculados inicialmente os desvios padrões das triplicatas,  $S_i$ , em cada conjunto de dados (BR, S e D), usando a equação (9).

$$S_i = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} \quad (9)$$

Isso permitiu obter a repetibilidade para cada uma das marcas de gasolinas pela seguinte equação:

$$REPE_{(BR, S \text{ ou } D)} = \sqrt{\frac{\gamma_1 * S_1^2 + \dots + \gamma_n * S_n^2}{\gamma_1 + \dots + \gamma_n}} \quad (10)$$

Por fim, chegou-se aos valores de repetibilidade final ( $u_2$ ) ao combinar as 3 contribuições calculadas no passo anterior, usando, para tanto, a equação (11).

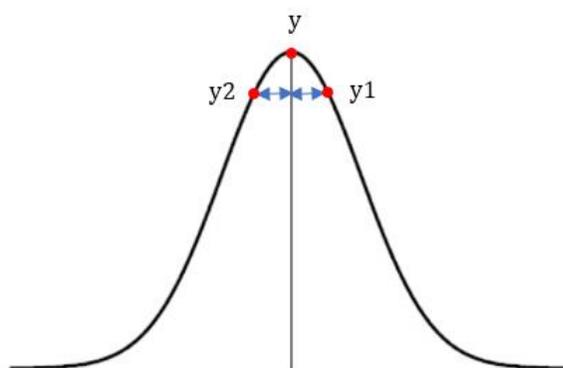
$$u_2 = \sqrt{\frac{\gamma_1 * REPE_{BR}^2 + \gamma_2 * REPE_S^2 + \gamma_3 * REPE_D^2}{\gamma_1 + \gamma_2 + \gamma_3}} \quad (11)$$

Em relação à contribuição da **incerteza devido à variação de temperatura** ( $u_3$ ), foi utilizado um novo conjunto de dados, com 45 amostras das marcas BR, S e D, sendo destinadas 15 amostras para cada um dos 3 equipamentos de infravermelho. Os espectros foram obtidos com as amostras submetidas a diferentes temperaturas, entre 12 e 20°C, utilizando um termômetro de

mercúrio para a medição da temperatura. Para este conjunto de dados, foi calculado o desvio padrão e foram geradas incertezas para todos os comprimentos de onda do MicroNIR e números de onda dos equipamentos MIR e NIR de bancada.

A **incerteza da variação do equipamento em termos do comprimento de onda** ( $u_4$ ) foi estimada selecionando-se os máximos das bandas mais representativas dos espectros, sendo uma banda representada como mostrado na Figura 15.

Figura 15 – Representação de uma banda, o valor máximo de absorbância e seus parâmetros utilizados na estimativa da incerteza.



Fonte: O autor.

Para a estimativa, foi então adaptada a forma de avaliar o erro na posição do comprimento de onda descrita na ASTM E 1655 (*Standard Practices for Infrared Multivariate Quantitative Analysis*). Adaptando esta metodologia, foi então calculado o valor de incerteza utilizando a equação (12),  $up_i$  ( $i = 1, 2... n$ , em que  $n$  é o número de bandas selecionadas) a qual é calculada para cada máximo selecionado. Tal equação leva em consideração a variação do valor de absorbância nos comprimentos de onda adjacentes ao comprimento de onda do máximo de absorbância da banda e resulta em um valor adimensional, tratando-se, portanto, de uma incerteza relativa.

$$up_i = \frac{(y - y1) + (y - y2)}{y} \quad (12)$$

Os valores obtidos para cada máximo de absorbância selecionados foram então combinados para obter a incerteza da variação do equipamento em termos do comprimento de onda ( $u_4$ ), como mostra a equação (13).

$$u_4 = \sqrt{up_1^2 + up_2^2 + \dots + up_n^2} \quad (13)$$

### 3.2.3.3 Cálculo da incerteza combinada

As contribuições das 4 fontes de incertezas ( $u_1$ ,  $u_2$ ,  $u_3$  e  $u_4$ ) foram combinadas, usando a equação (14), obtendo-se assim a incerteza combinada relativa ( $uc$ ). Para realizar este cálculo, as incertezas de cada uma das fontes foram transformadas em incertezas relativas ao dividir o seu valor pelo valor da absorvância do respectivo espectro médio das amostras utilizadas no cálculo ( $y_i$ ) em cada comprimento de onda.

$$uc = \sqrt{\left(\frac{u_1}{y_1}\right)^2 + \left(\frac{u_2}{y_2}\right)^2 + \left(\frac{u_3}{y_3}\right)^2 + u_4^2} \quad (14)$$

Em seguida, foi obtida a incerteza relativa expandida ( $u$ ), usando um fator de abrangência igual a 2, que representa 95% de confiança, conforme o Guia EURACHEM/CITAC (2002):

$$u = 2 * uc \quad (15)$$

Este é o valor que é considerado como a incerteza do espectro de infravermelho para cada comprimento de onda, sendo, portanto, representado por um vetor linha, em que cada elemento corresponde ao valor de incerteza no respectivo comprimento de onda.

### 3.2.4 Geração de amostras virtuais

As amostras virtuais foram obtidas calculando valores aleatórios em torno de cada uma das medições (referentes às obtidas pelos métodos normalizados), a cada um dos pontos da curva de destilação e aos valores de absorvância para cada comprimento de onda dos espectros de infravermelho), dentro das faixas de incertezas previamente estimadas.

Para cada amostra real foram geradas 49 amostras virtuais nos conjuntos de dados dos métodos normalizados e nas curvas de destilação. O mesmo procedimento foi feito para os espectros no infravermelho, tendo sido geradas 49 amostras virtuais para cada um dos 54 espectros médios de cada conjunto (MicroNIR, NIR, MIR). A rotina em linguagem MATLAB é apresentada no Apêndice A.

No caso dos espectros de infravermelho, a existência de diferentes incertezas para cada comprimento de onda poderia gerar espectros virtuais com mudanças abruptas nos valores de absorvância, o que não corresponderia aos espectros das amostras reais. Sendo assim, com o intuito de suavizar essas mudanças, a seguinte estratégia foi adotada para a geração de cada amostra virtual: inicialmente, um vetor de mesmo tamanho do espectro da amostra real foi criado e dividido em intervalos iguais ( $n$ ); para os pontos extremos de cada intervalo, foram adotados valores randômicos entre 0 e 1; em seguida, uma variação linear entre esses pontos foi estabelecida, a fim de se gerar mudanças gradativas dentro dos intervalos; o vetor gerado foi então multiplicado (elemento a elemento) pelo vetor contendo as incertezas expandidas ( $uc*2$ ) para cada comprimento de onda, resultando em um vetor de valores aleatórios ( $va$ ).

Com a finalidade de gerar espectros virtuais com valores de absorvância tanto maiores, quanto menores que os respectivos valores dos espectros reais, dentro da incerteza calculada, foi subtraído o vetor contendo as incertezas expandidas do espectro real (o que representa o menor valor de absorvância dentro da incerteza), o qual foi, por fim, adicionado ao vetor de valores aleatórios ( $va$ ), levando à geração do espectro virtual.

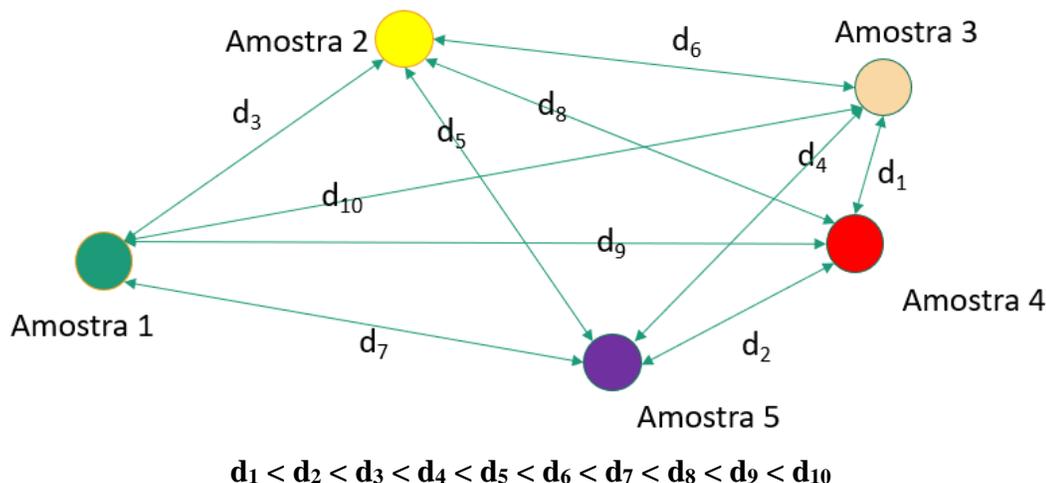
Foram realizados testes utilizando diversos valores de intervalos e feita inspeção visual a fim de selecionar o intervalo que gerasse espectros uniformes e correspondentes à realidade. O Apêndice B apresenta o código em linguagem MATLAB desenvolvido para a realização desses testes.

### 3.3 ANÁLISE DE COMPONENTES PRINCIPAIS

A comparação entre as amostras de gasolina foi realizada por pares, cada par sendo representado pelas medidas reais dos parâmetros físico-químicos ou espectros de infravermelho e os respectivos 49 valores obtidos virtualmente para cada parâmetro ou variável espectral.

Com o intuito de selecionar os pares de amostras para serem analisados por PCA, foram calculadas as distâncias Euclidianas entre todos os pares possíveis de amostras reais, e esses foram então organizados em ordem crescente de distância. A Figura 16 mostra uma representação esquemática do procedimento.

Figura 16 – Representação esquemática do cálculo das possíveis distâncias Euclidianas entre todos os pares de amostras.



Fonte: O autor.

Em seguida foi realizada uma análise de componentes principais para os pares sistematicamente selecionados de acordo com a distâncias calculadas, utilizando o MATLAB® (*Math Works, Inc., Natick, MA*) e o pacote de ferramentas PLS-Toolbox versão 8.6.2, que funciona em ambiente MATLAB.

Para cada PCA, os dados foram pré-processados utilizando uma centragem na média e auto-escalamento para os parâmetros físico-químicos (ensaios normalizados e curvas de destilação) e centragem na média e *SNV* para os dados espectrais.

Inicialmente foram analisados os pares cujas respectivas amostras estão mais próximas entre si (menores distâncias Euclidianas) até se encontrar um par com distinção das amostras.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 ENSAIOS NORMALIZADOS

#### 4.1.1 Análises Convencionais

Os resultados obtidos dos ensaios normalizados são referentes a diferentes métodos de análise, cada um com uma interpretação química individual. Diferentemente de um espectro, em que cada comprimento de onda é uma variável, ou de uma curva de destilação, na qual as variáveis são os volumes destilados, não é possível fazer uma análise visual através de gráfico com os valores para as análises dos ensaios convencionais.

As amostras das 9 regiões de Pernambuco analisadas nesse estudo apresentaram valores dentro das especificações exigidas pela ANP, como mostram os exemplos da Tabela 4.

Tabela 4 – Resultados de análises físico-químicas para uma amostra de cada região analisada.

Região da amostra	Variáveis						
	D10 (°C)	D50 (°C)	D90 (°C)	FBP (°C)	RES (mL)	ME (Kg/m3)	T. ET. (% vol.)
01	53,0	69,5	137,2	184,2	1,0	733,80	27
02	52,8	70,7	158,6	198,7	1,1	744,90	27
03	54,1	70,3	144,4	186,2	1,5	733,60	28
04	52,5	69,8	145,5	185,7	0,8	733,70	28
05	54,0	70,2	147,5	191,6	1,0	735,40	27
06	53,0	69,2	140,4	185,6	1,0	732,10	29
07	55,0	71,0	147,7	189,0	1,0	740,20	28
08	55,2	71,8	150,1	196,7	1,0	748,10	27
09	52,1	69,3	138,0	183,3	1,0	734,40	28

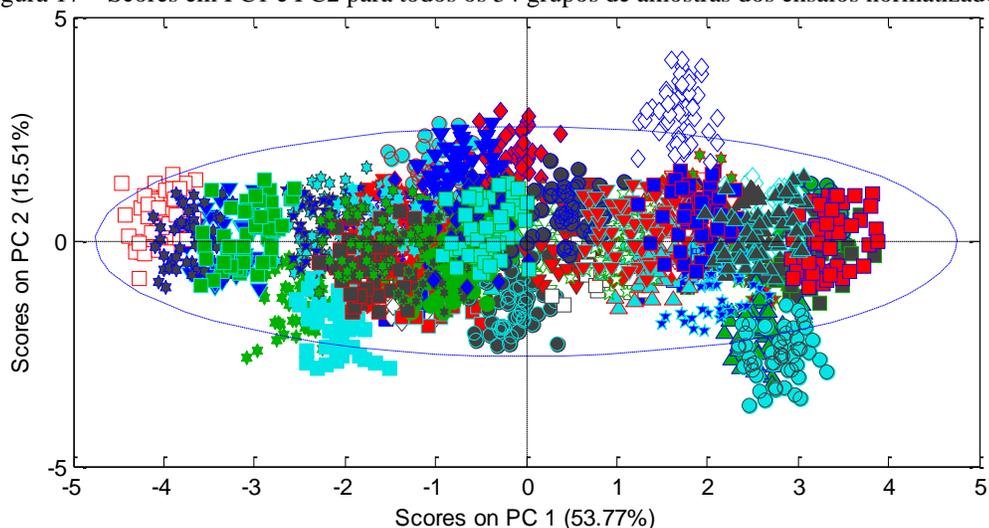
Fonte: O autor.

#### 4.1.2 Geração de amostras virtuais e Análise de Componentes Principais (PCA)

Após a geração das amostras virtuais pela estratégia descrita, foi obtido um total de 2700 amostras, sendo 54 grupos (cada grupo é composto por uma amostra real e 49 amostras virtuais). Esses dados foram pré-processados com auto-escalamento para então serem analisados por meio de uma PCA.

A Figura 17 mostra os scores em PC1 e PC2 para todas as 2700 amostras, através da qual não é possível distinguir os grupos de amostras, devido à grande variabilidade dos dados, o que também é percebido ao se analisar os scores em outras PCs.

Figura 17 – Scores em PC1 e PC2 para todos os 54 grupos de amostras dos ensaios normalizados.

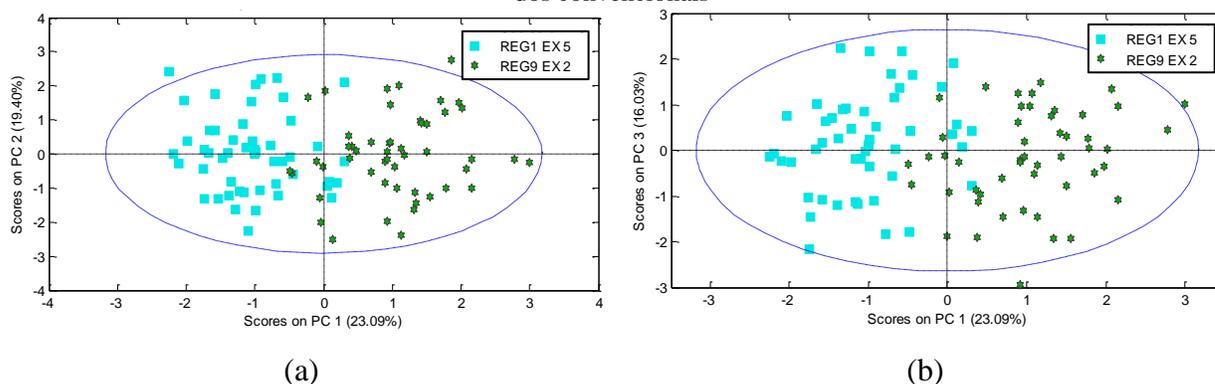


Fonte: O autor.

Os 54 grupos foram então analisados em pares a fim de investigar o potencial para separar as amostras. A quantidade possível de pares a serem formados foi de 1431 (combinação 2 a 2 das 54 amostras), os quais foram organizados pela ordem crescente de distâncias Euclidianas entre as respectivas amostras reais de cada grupo.

A Figura 18 mostra os scores em PC1, PC2 e PC3 para o par de amostras mais próximas.

Figura 18 – Scores para o par de amostras mais próximas em PC1xPC2 (a) e em PC1xPC3(b), considerando os métodos convencionais

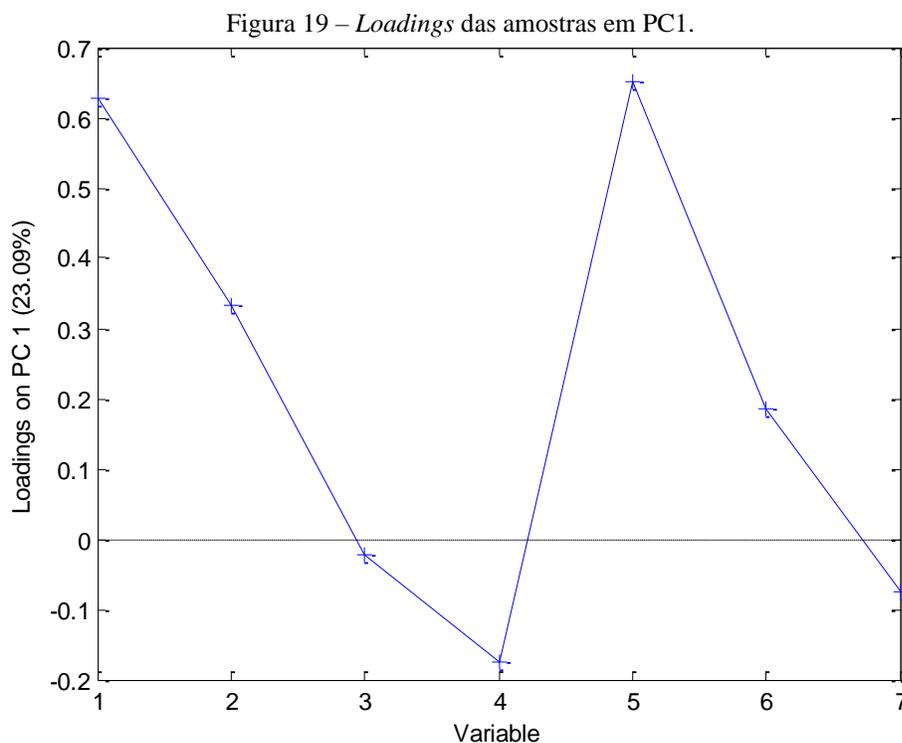


Fonte: O autor.

Esta PCA para o par de amostras mais próximas mostra que é possível identificar uma certa distinção entre os dois grupos de amostras, no entanto, não é possível inferir com certeza que as amostras reais são distintas, visto que a separação pelos scores da PCA não é total, o que se repete

ao analisar até 7 PCs. Apesar disso, há um forte indicativo de que as amostras são de fato diferentes, visto que é perceptível a tendência de separação.

Neste caso, os *loadings* indicam que as variáveis que mais contribuíram para a variabilidade das amostras em PC1 foram o resíduo da destilação (RES, variável 5) e o ponto de destilação a 10% de volume destilado (D10, variável 1), como mostra a Figura 19.

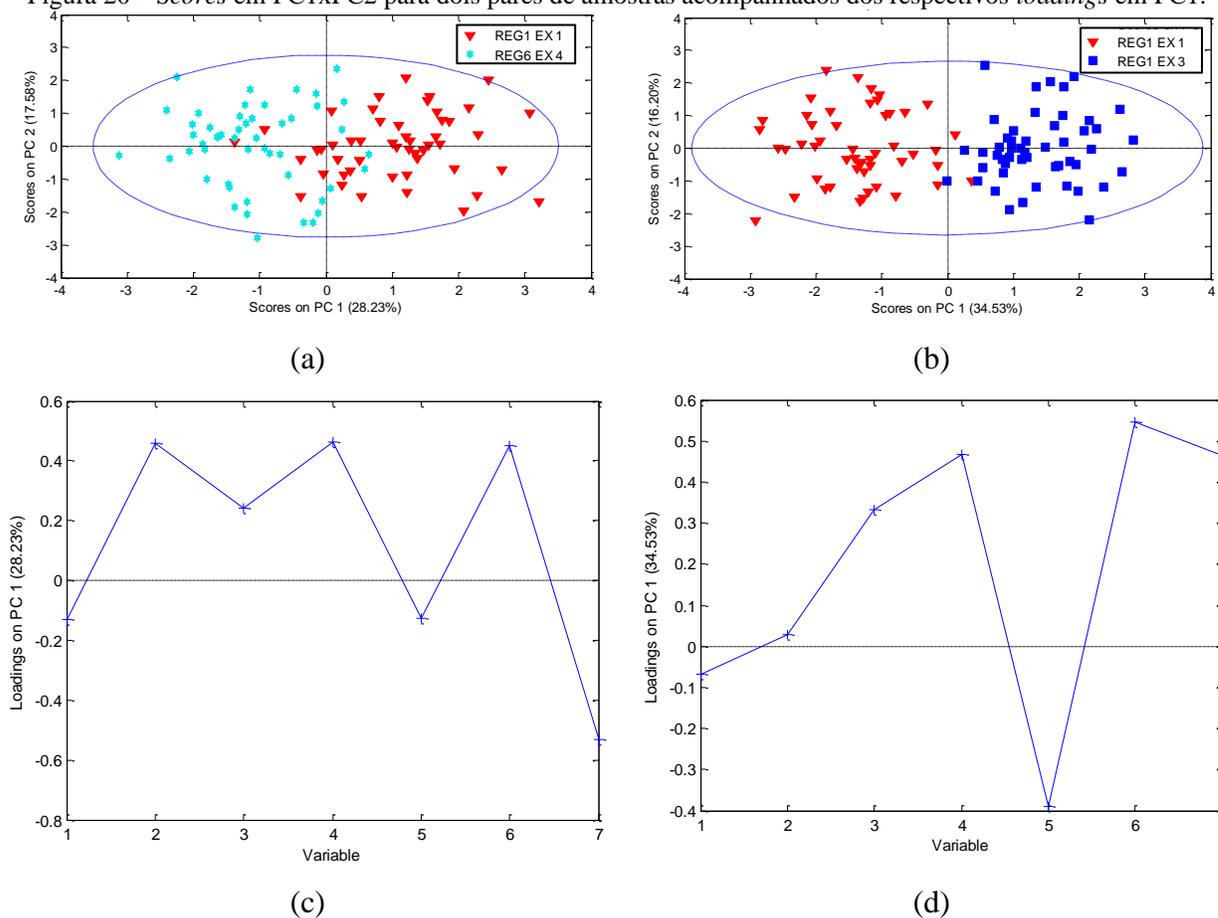


Fonte: O autor.

Este mesmo padrão de distinção entre os grupos, indicando fortemente serem amostras com origens distintas, mas sem separação total por meio das análises dos scores das amostras nas PCs, é visto nos 43 pares de amostras mais próximas.

A Figura 20 mostra os *scores* em PC1 e PC2 para outros dois pares de amostras, pelos quais não é possível identificar separação total entre os dois grupos. Também são apresentados os respectivos *loadings* em PC1, os quais mostram que a variabilidade em PC1 para o primeiro par (amostras dos grupos REG1 EX1 e REG6 EX4) se deve principalmente às variáveis 7 (T.ET), 6 (ME), 2 (D50) e 4 (FBP), enquanto para o segundo par (amostras dos grupos REG1 EX1 e REG1 EX3) se deve principalmente às variáveis 6 (ME), 4 (FBP) e 5 (RES).

Figura 20 – Scores em PC1xPC2 para dois pares de amostras acompanhados dos respectivos *loadings* em PC1.

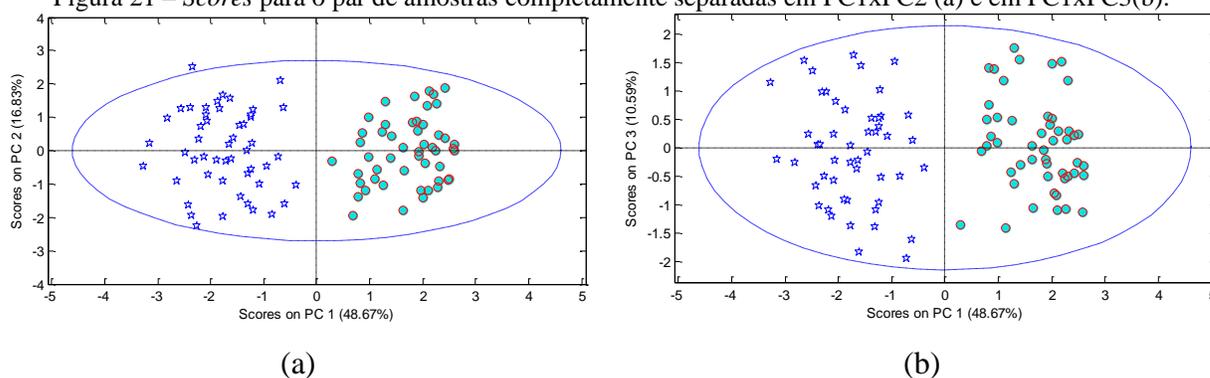


Fonte: O autor.

Estes 43 pares para os quais não houve separação total representam 3% do total de pares possíveis, enquanto para os outros 97% dos pares, pode-se observar uma clara separação ao se analisar os *scores* em PC1 e PC2.

A Figura 21 mostra um par para o qual houve separação total das amostras numa análise visual dos *scores* em PC1xPC2 e PC1xPC3, o que leva a concluir que as amostras são de fato distintas, ou seja, possuem diferentes origens.

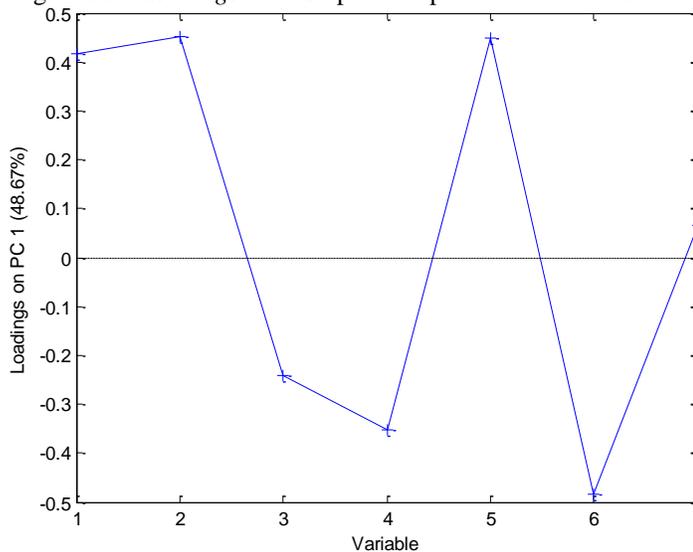
Figura 21 – Scores para o par de amostras completamente separadas em PC1xPC2 (a) e em PC1xPC3(b).



Fonte: O autor.

Os *loadings* em PC1 para este par de amostras é apresentado na Figura 22.

Figura 22 – Loadings em PC1 para um par de amostras distintas.



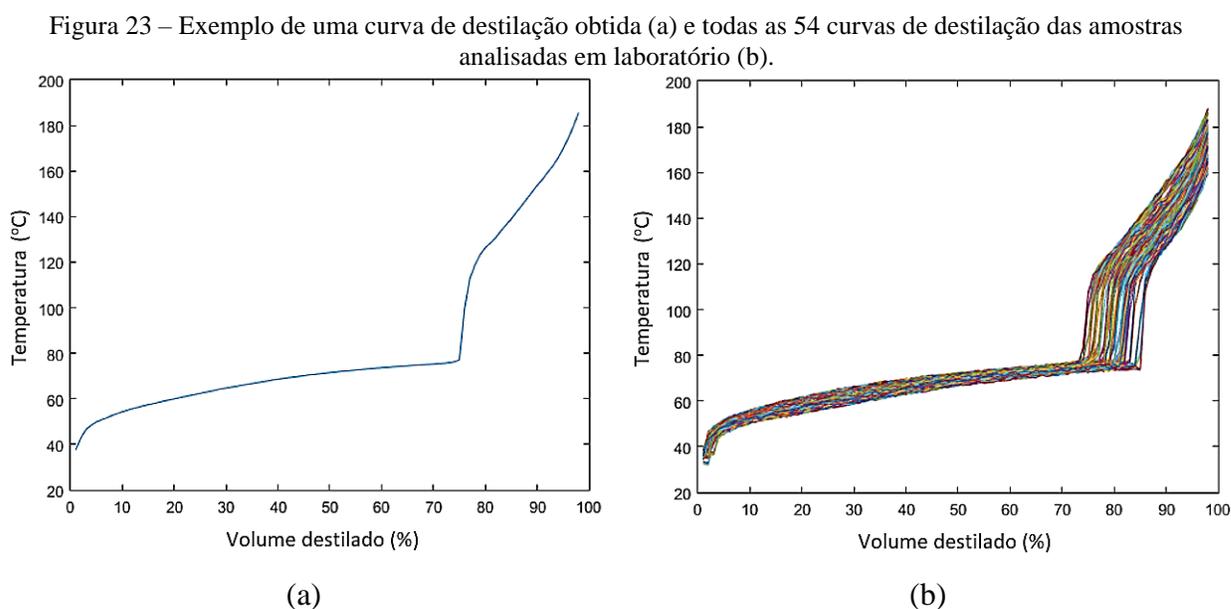
Fonte: O autor.

Ao se comparar os *loadings* deste par de amostras com os outros já apresentados, nota-se que não há um padrão indicando as variáveis que mais contribuem com a variabilidade dos dados nas PCs, isso se deve ao fato de que as amostras em cada par de grupos são separadas pelos *scores* da PCA devido a diferentes variáveis, o que pode mudar a depender da composição química das amostras. Portanto, as variáveis que mais contribuem para a distinção de determinado par de amostras podem não ser as mesmas para outros pares.

## 4.2 CURVAS DE DESTILAÇÃO

### 4.2.1 Curvas de destilação

As curvas de destilação obtidas apresentaram o padrão esperado. Entre 50 e 85° C ocorre a destilação do etanol presente na gasolina, devido a isso, a curva de destilação apresenta uma mudança brusca de temperatura, como pode ser visto na Figura 23 (a), que representa uma das curvas obtidas neste trabalho. A Figura 23 (b) apresenta todas as 54 curvas de destilação das amostras disponíveis para estudo.



Fonte: O autor.

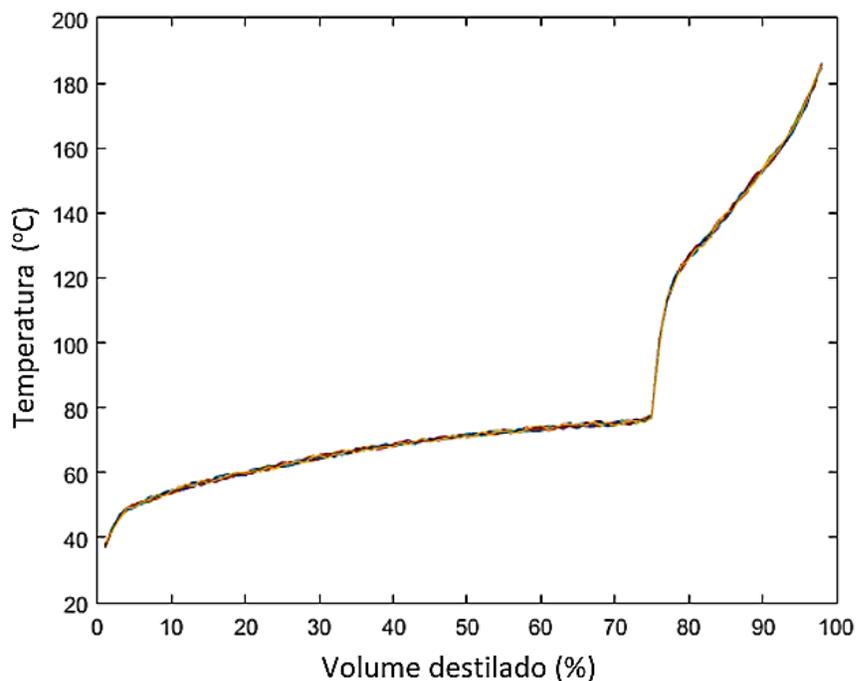
É possível perceber uma variação do percentual de destilado no ponto de mudança abrupta de temperatura, o que se deve à diferença de quantidade de álcool anidro presente na gasolina bem como à sua composição química, que pode afetar diretamente a curva de destilação, visto que amostras compostas por uma maior quantidade de constituintes leves apresentariam pontos de destilação mais baixos. De acordo com Almeida (2002), gasolinas provenientes do processo de destilação fracionada e aquelas originárias de produção em formuladores ou petroquímica também podem apresentar diferenças em suas curvas de destilação devido à composição química.

Importante ressaltar que todas as amostras analisadas estão em conformidade com os parâmetros da ANP para os pontos de destilação, que estabelece os limites de 65 °C, 80 °C e 190 °C para 10%, 50% e 90% de destilado, respectivamente, como pode ser visto nas curvas apresentadas.

#### 4.2.2 Cálculo de incertezas e geração de amostras virtuais

Aplicando a estratégia de calcular valores aleatórios em torno de cada uma das medições da curva de destilação, dentro das faixas de incertezas previamente estimadas, foram obtidas as curvas de destilação virtuais. A Figura 24 mostra um conjunto de curvas de destilação formado pela amostra original e suas 49 amostras virtuais.

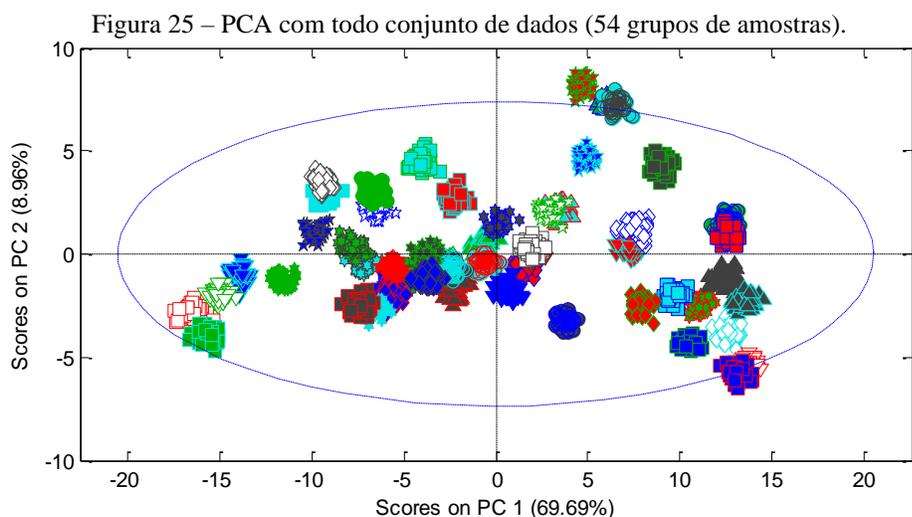
Figura 24 – Curva de destilação de uma amostra real e suas respectivas 49 curvas obtidas virtualmente.



Fonte: O autor.

#### 4.2.3 Análise de Componentes Principais

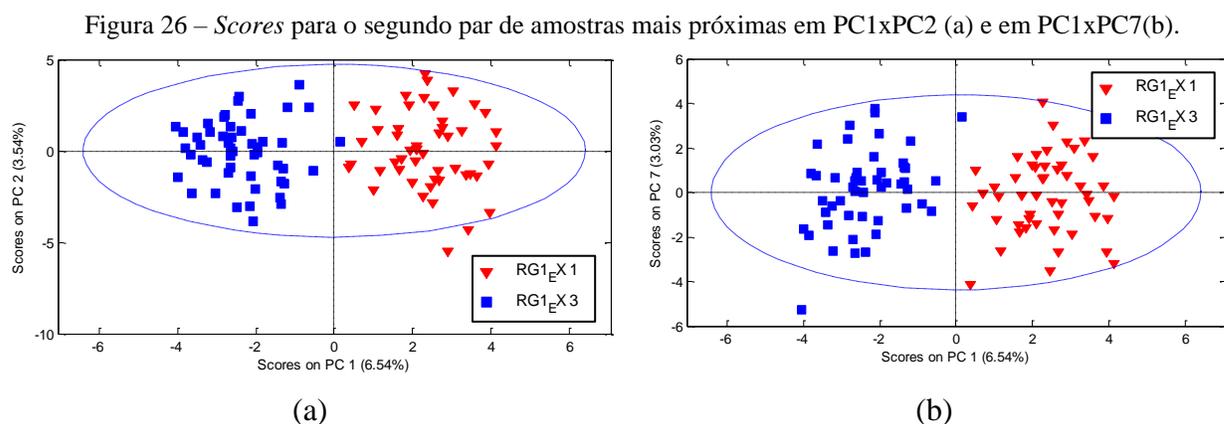
Após o cálculo das distâncias Euclidianas entre os 1431 pares de amostras, foi realizada uma Análise de Componentes Principais. Os espectros foram previamente pré-processados com autoescalonamento. Ao fazer uma PCA com todos os dados, como mostra a Figura 25, nota-se claramente a existência de diversos grupos bem separados (alguns deles relativamente próximos entre si), o que dá um indicativo de que as amostras tenderão a se separar mais facilmente ao serem analisadas em pares.



Fonte: O autor.

Em seguida, foram feitas as análises para os pares de amostras. Foi possível identificar uma clara separação, em pelo menos uma PC, para todos os pares calculados, o que confirma a eficiência do uso de curvas de destilação para separar amostras de gasolinas com base nos *scores* da PCA.

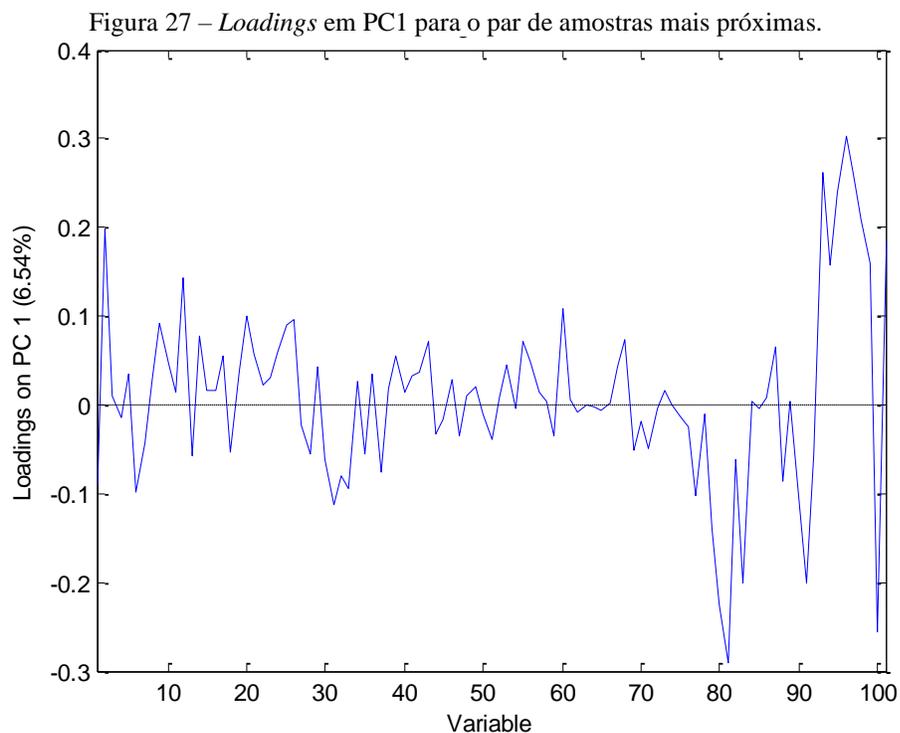
Para o par de amostras mais próximas entre si, foi necessário utilizar pelo menos 7 PCs para identificar separação completa entre os grupos. A Figura 26 mostra os scores em PC1xPC2 e PC1xPC7 para o par de amostras mais próximas, percebe-se que com os scores em PC1xPC2 os dois conjuntos foram separados, mas não por completo. No entanto, os scores de PC1xPC7 mostram uma clara separação entre os dois grupos.



Fonte: O autor.

Os *loadings* em PC1 mostram que, neste caso, as variáveis que mais contribuíram para a variabilidade foram aquelas entre 80 e 100% de volume de destilado, o que indica que as amostras

tendem a se diferenciar principalmente devido às características químicas dos compostos da sua fração mais pesada, como é mostrado na Figura 27.

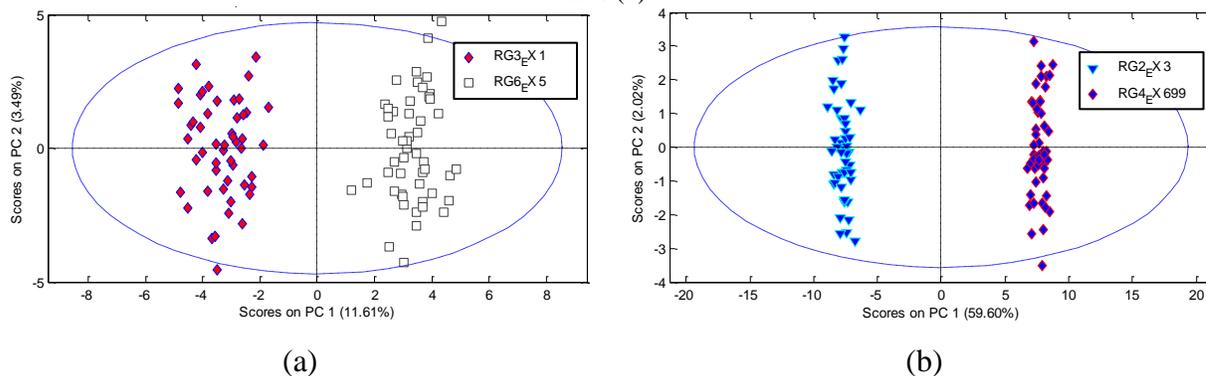


Fonte: O autor.

Para os demais 1430 pares, 99,93% do total, foi possível perceber uma clara separação entre os dois conjuntos de amostras já nas primeiras PCs, alguns com uma separação mais modesta, enquanto outros com uma separação mais evidente. Quanto maior a distância Euclidiana entre as amostras do par, mais distantes os grupos ficam a serem representados nas PCs.

Na Figura 28 são apresentados os scores em PC1xPC2 das amostras do 3º par na ordem crescente de distâncias Euclidianas (a), em que os grupos estão modestamente separados, mas são claramente distintos, e do 500º par na mesma ordem (b), com um distanciamento bem mais considerável entre os grupos.

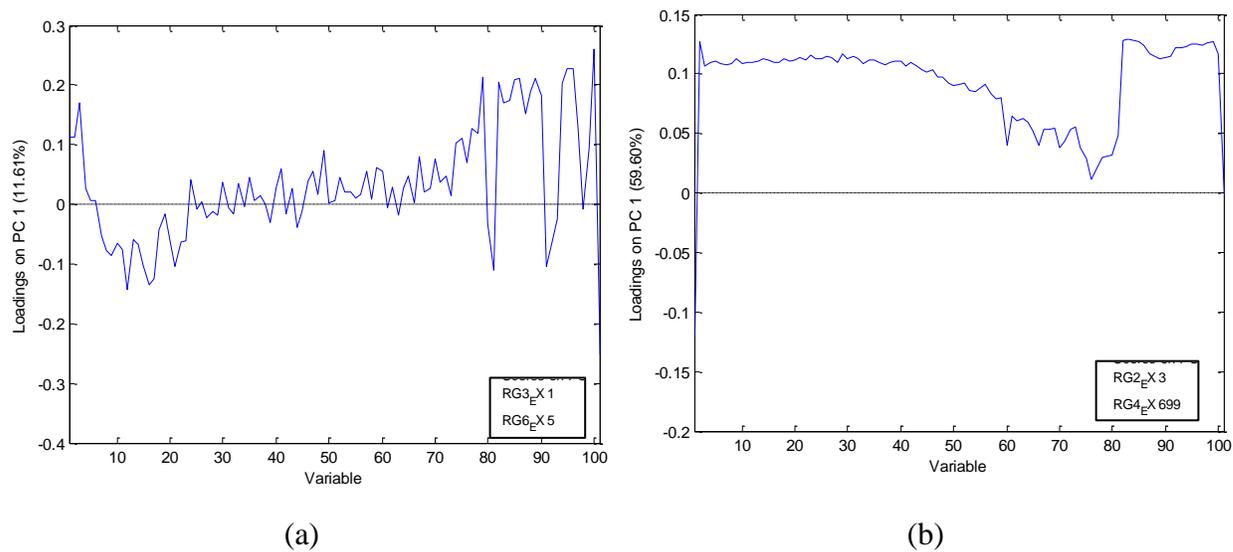
Figura 28 – Scores em PC1xPC2 para o 3º par de amostras mais próximas (a) e o 500º par de amostras mais próximas (b).



Fonte: O autor.

Os *loadings* em PC1 para ambos os pares analisados são mostrados na Figura 29.

Figura 29 – Loadings em PC1 para o 3º par de amostras mais próximas (a) e o 500º par de amostras mais próximas (b).

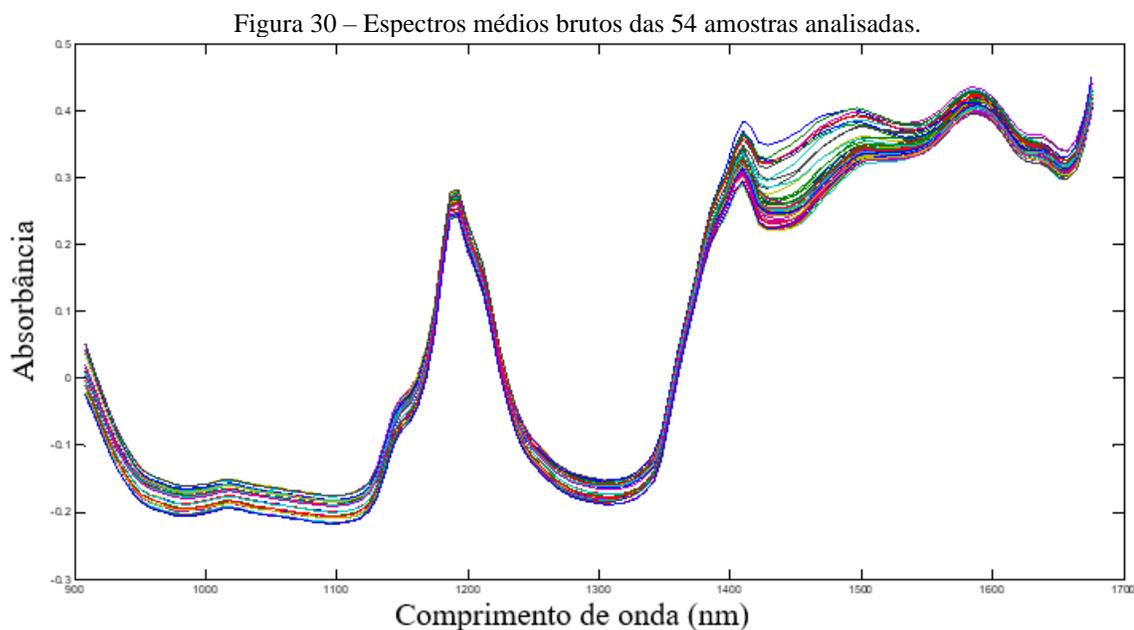


Fonte: O autor.

## 4.3 ESPECTROSCOPIA NIR EM INSTRUMENTO PORTÁTIL

### 4.3.1 Espectros

Das 162 análises realizadas no instrumento portátil da Viavi, o MicroNIR, foram obtidos 54 espectros médios brutos, os quais são apresentados na Figura 30.

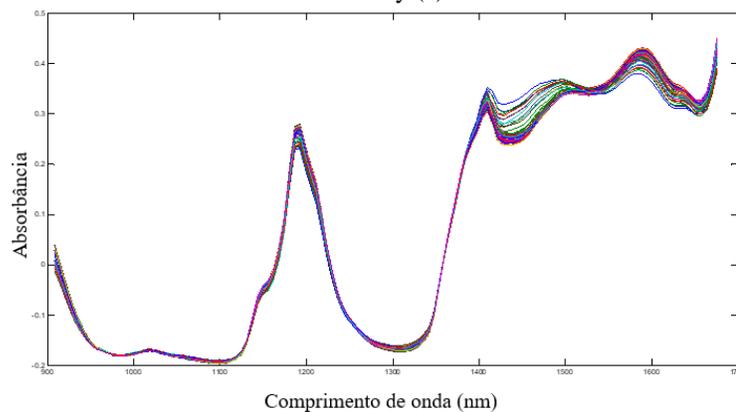


Podem ser visualizadas algumas bandas características da gasolina nesses espectros, tais quais as bandas na região entre 800 e 1100 nm, referente ao terceiro sobretom da ligação C-H dos hidrocarbonetos; entre 1100 e 1200 nm, referente ao segundo sobretom da ligação C-H; assim como as bandas entre 1300 e 1400 nm, que se referem às combinações das vibrações fundamentais do grupo C-H e o sobretom das vibrações de estiramento O-H na região entre 1420 e 1600 nm.

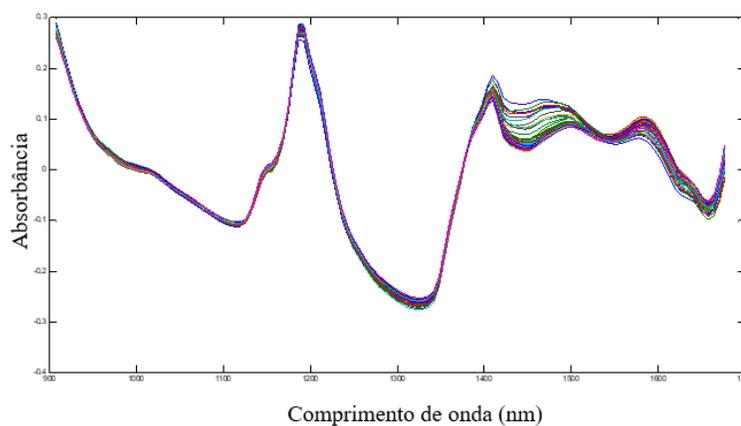
Os espectros obtidos apresentam valores negativos de absorbância em diversos comprimentos de onda, sendo assim, foi realizado um deslocamento desses espectros ao longo do eixo y para que todos os valores de absorbância dos espectros fossem maiores ou iguais a zero, a fim de não gerar erros nos cálculos das incertezas a serem realizados posteriormente.

Foram testados 3 pré-processamentos para os espectros: MSC, Correção de linha de base e SNV. A Figura 31 mostra os espectros pré processados.

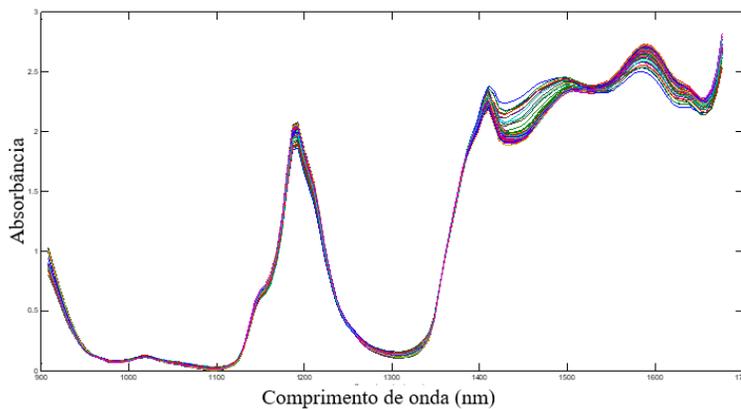
Figura 31 – Espectros pré-processados com MSC (a), correção de linha de base (b) e SNV + deslocamento no eixo y (c).



(a)



(b)



(c)

Fonte: O autor.

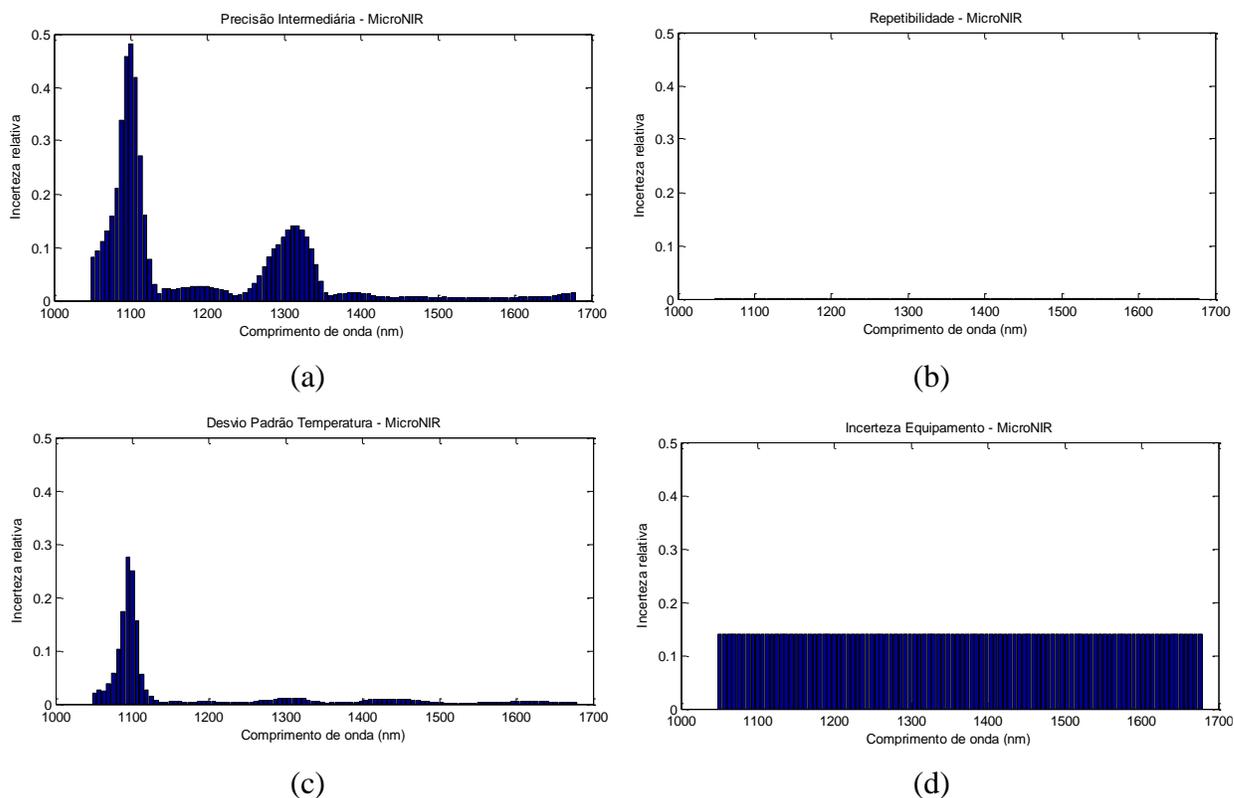
Os resultados obtidos utilizando MSC e SNV foram semelhantes, no entanto, optou-se pelo pré-tratamento com SNV e um deslocamento dos espectros pré-processados ao longo do eixo y para deixar todos os valores de absorvância positivos.

### 4.3.2 Cálculo de incertezas

Com os dados pré-processados, foi utilizada a metodologia descrita para a obtenção das incertezas em cada comprimento de onda dos espectros. Desta maneira, foram encontrados os valores de contribuição da incerteza para cada uma das quatro fontes: repetibilidade, reprodutibilidade (precisão intermediária), variação de temperatura da amostra e variação do equipamento em termos do comprimento de onda.

A Figura 32 mostra os resultados obtidos, onde é possível visualizar a dimensão das contribuições de cada fonte de incerteza. Os valores para a repetibilidade (Figura 32 (b)) foram os menores (na ordem de  $10^{-3}$ ), enquanto os maiores valores foram obtidos para os pontos em torno do comprimento de 1100 nm para a precisão intermediária (Figura 32 (a)), o qual chega a atingir um valor de incerteza relativa próximo a 0,5. No entanto, é importante notar que as incertezas ao longo do comprimento de onda provenientes do equipamento (Figura 32 (d)) apresentaram os maiores valores médios, todos acima de 0,1.

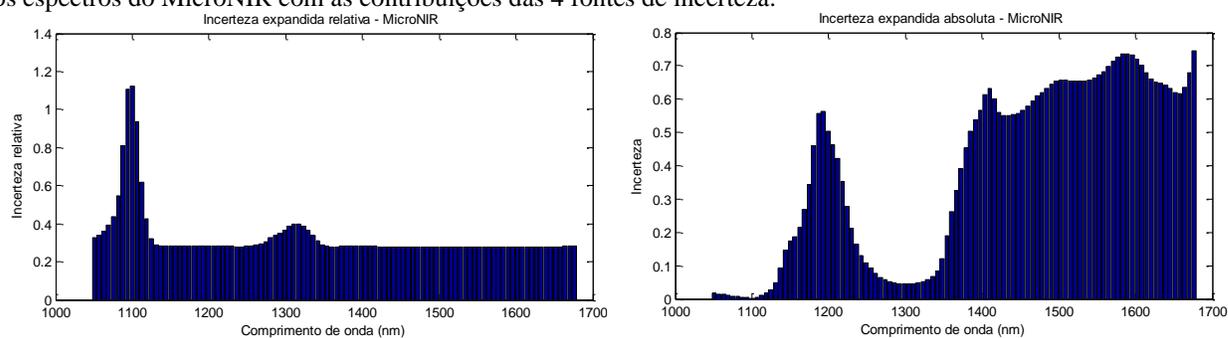
Figura 32 – Incertezas referentes à precisão intermediária, ou reprodutibilidade (a), à repetibilidade (b), à variação de temperatura (c) e ao equipamento (d).



Fonte: O autor.

As incertezas expandidas relativa e absoluta estimadas estão mostradas na Figura 33. Nota-se que o impacto da incerteza referente ao equipamento foi considerável para a incerteza combinada, o que, muito provavelmente, se deve ao fato de que o instrumento utilizado realiza a medida de absorção em um número muito reduzido de comprimentos de onda (apenas 125), com uma diferença de mais de 6 nm entre pontos consecutivos, o que acabou por superestimar as incertezas pela estratégia utilizada, resultando em valores altos inclusive nas regiões de ruído dos espectros. Estes resultados não estariam coerentes com a alta reprodutibilidade esperada no posicionamento dos comprimentos de onda, considerando equipamentos a base de filtros, sem partes móveis, como é o caso do MicroNIR.

Figura 33 – Incerteza expandida relativa (a) e incerteza expandida absoluta (b) em cada comprimento de onda para os espectros do MicroNIR com as contribuições das 4 fontes de incerteza.



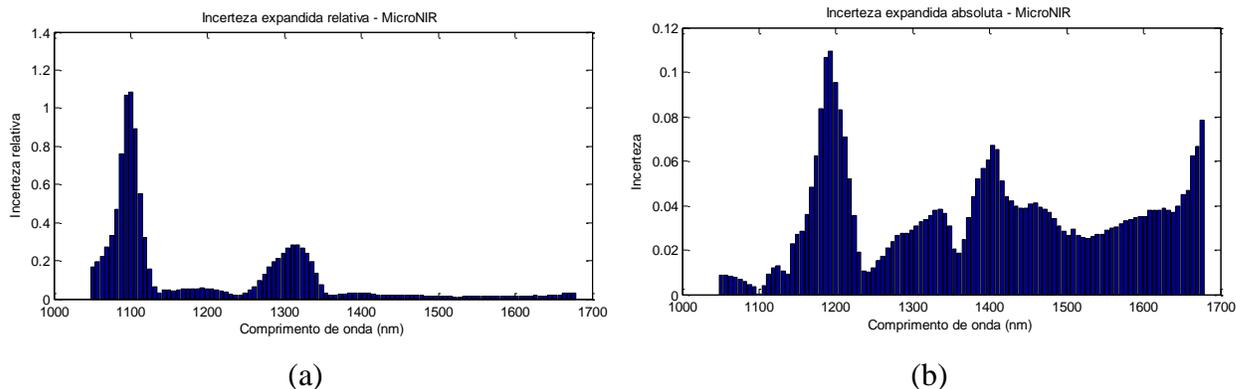
(a)

(b)

Fonte: O autor.

Optou-se, então, por eliminar a fonte de incerteza referente ao equipamento, sendo a incerteza expandida para os comprimentos de onda dos espectros do MicroNIR obtidas pela combinação de 3 fontes de incerteza: repetibilidade, reprodutibilidade (precisão intermediária) e variação de temperatura da amostra. A Figura 34 mostra os resultados obtidos.

Figura 34 – Incerteza expandida relativa (a) e incerteza expandida absoluta (b) em cada comprimento de onda para os espectros do MicroNIR sem a contribuição da fonte de incerteza do equipamento.



Fonte: O autor.

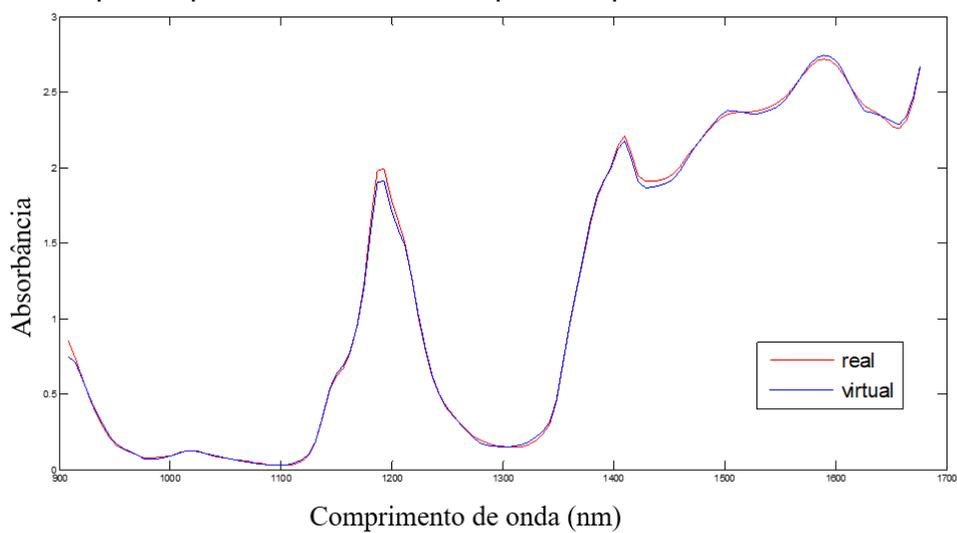
Ao se visualizar as incertezas expandidas relativas, Figura 33(a) e Figura 34(a), percebe-se uma clara diferença na forma das incertezas ao longo do comprimento de onda. No segundo caso, já não se nota mais os altos valores nas regiões de ruído, como no primeiro caso. As incertezas obtidas da segunda maneira, cujos valores das bandas com importância química são da ordem de  $10^{-2}$ , são bem menores que aquelas superestimadas ao se utilizar a contribuição do equipamento, cujos valores são na ordem de  $10^{-1}$ .

### 4.3.3 Geração de amostras virtuais

De acordo com a estratégia descrita para a geração das amostras virtuais, foram testados 3 diferentes intervalos ( $n$ ) de variáveis para o vetor contendo valores entre 0 e 1. Os testes foram feitos com os intervalos de 2, 4 e 31.

O uso de intervalos muito pequenos poderia gerar espectros com aparência anormal, com quebras bruscas, enquanto o uso de intervalos muito grandes tenderia a diminuir a aleatoriedade dos valores de absorvância ao longo do comprimento de ondas. Os 3 intervalos analisados geraram espectros visualmente fidedignos, sem variações abruptas nos valores de absorvância, semelhantes aos espectros das amostras reais, como pode ser visto no Apêndice B. Dessa maneira, foi escolhido o intervalo intermediário (4 pontos) para a geração de todas as amostras virtuais para os dados do MicroNIR, para o qual é apresentado um exemplo de espectro virtual gerado na Figura 35.

Figura 35 – Exemplo de espectro real e um dos seus respectivos espectros virtuais obtido com intervalo  $n=4$ .

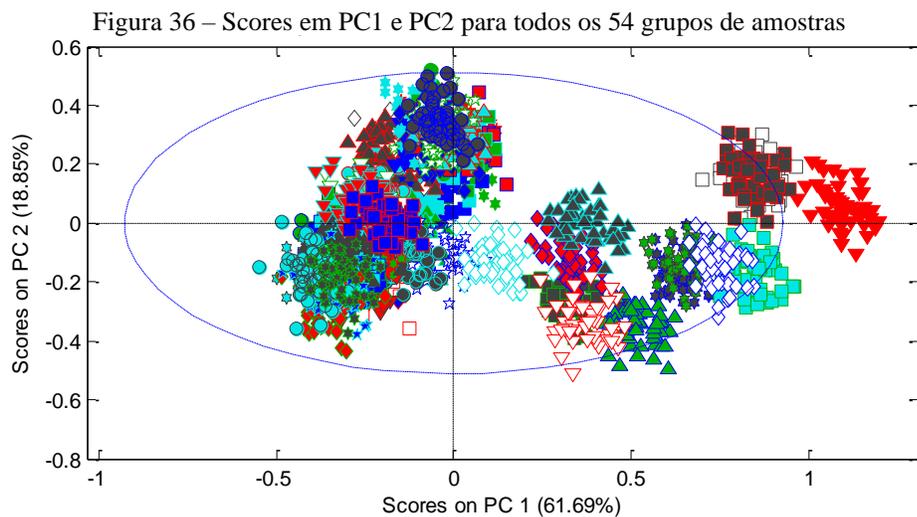


Fonte: O autor.

Dessa forma, foi obtido um total de 2700 amostras, sendo 54 grupos (cada grupo composto por uma amostra real e as respectivas 49 amostras virtuais).

#### 4.3.4 Análise de Componentes Principais

O conjunto de dados foi pré-processado com centragem na média e SNV para em seguida ser analisado por PCA, cujos scores em PC1xPC2 são mostrados na Figura 36. Percebe-se que os grupos estão, em sua grande maioria, bem misturados, indistinguíveis, sem a formação de classes específicas claramente distintas umas das outras.



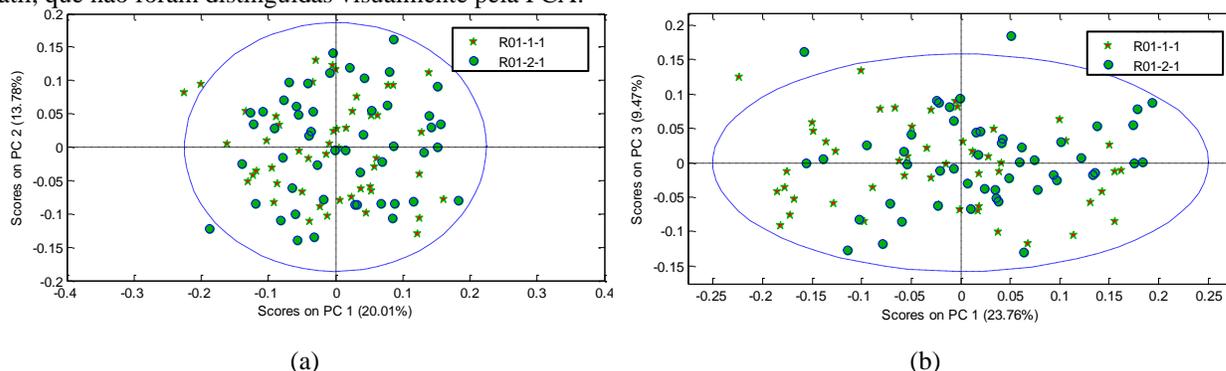
Fonte: O autor.

Na região central da Figura 36 as amostras estão misturadas a tal ponto de não ser possível identificar muitos dos grupos individuais de amostras (cada grupo está identificado por um mesmo símbolo). Já na região à direita, alguns grupos apresentam maior separação em relação aos outros, o que indica que estes grupos tenderão a se distinguir mais facilmente de outros ao serem analisados em pares.

Após essa visualização preliminar de todos os 54 grupos de amostras em uma PCA, esse conjunto de dados foi analisado aos pares. O total de pares de grupos de amostras formados foi de 1431, os quais, ao serem analisados por PCA, apresentaram 3 padrões distintos: alguns dos pares não puderam ser distinguidos, alguns puderam ser distinguidos parcialmente e outros puderam ser distinguidos por completo.

A Figura 37 mostra um exemplo de par cujas amostras não puderam ser distinguidas por meio de uma análise visual dos scores da PCA, mesmo ao se analisar mais de 5 PCs, o que aconteceu em cerca de 199 pares, 13,9% dos 1431 pares. Nota-se que as amostras dos dois grupos possuem scores tais que as tornam homogêneas, ou seja, não é possível dizer que estas amostras reais de cada grupo possuem origens distintas por meio da metodologia utilizada.

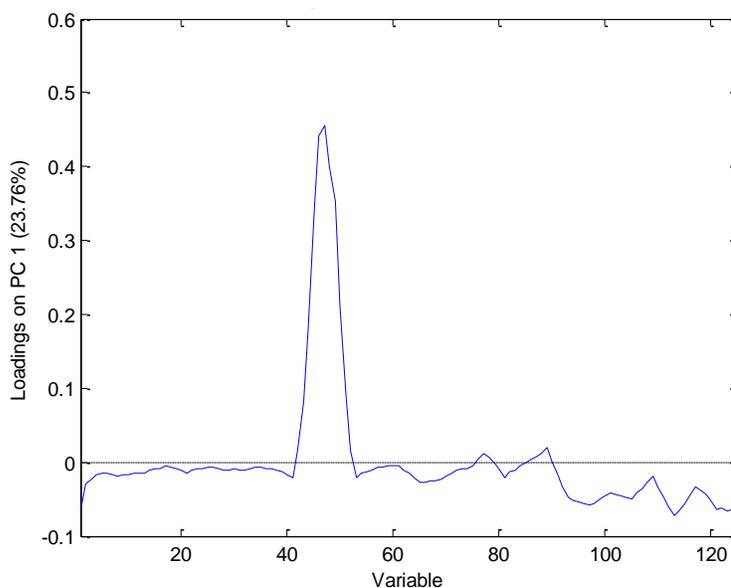
Figura 37 – Scores em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, obtidas utilizando o equipamento portátil, que não foram distinguidas visualmente pela PCA.



Fonte: O autor.

A Figura 38 mostra os *loadings*, em PC1 que explicam 23,76% da variabilidade das amostras. As variáveis que mais contribuíram para a variabilidade das amostras foram as que representam a banda de absorção na região entre 1100 e 1300 nm (variáveis 40 a 55), referente ao segundo sobretom da ligação C-H.

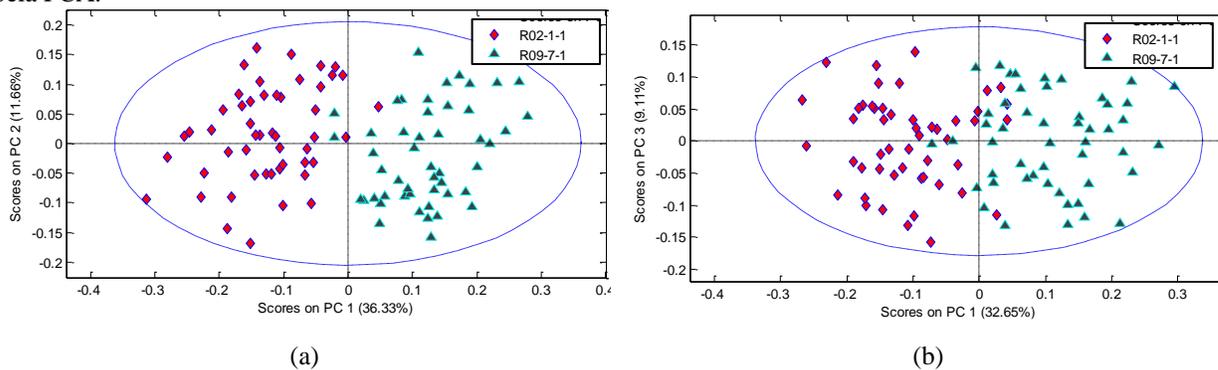
Figura 38 – *Loadings* em PC1 (a) e PC2 (b) para um par de amostras que não puderam ser distinguidas pela PCA.



Fonte: O autor.

Já a Figura 39 mostra um exemplo de par cujas amostras puderam ser parcialmente distinguidas. Pelos scores em PC1xPC2 e PC1xPC3 é possível identificar uma tendência de separação entre os grupos, no entanto, parte das amostras estão em uma região intermediária, pertencendo a ambos os grupos, o que impossibilita afirmar que as amostras reais de cada grupo são de fato diferentes, no entanto, o baixo nível de homogeneidade entre os grupos dá um indicativo de que as amostras são distintas. 210 pares de grupos de amostras se apresentaram em condição semelhante, o que representa cerca 14,67% do total.

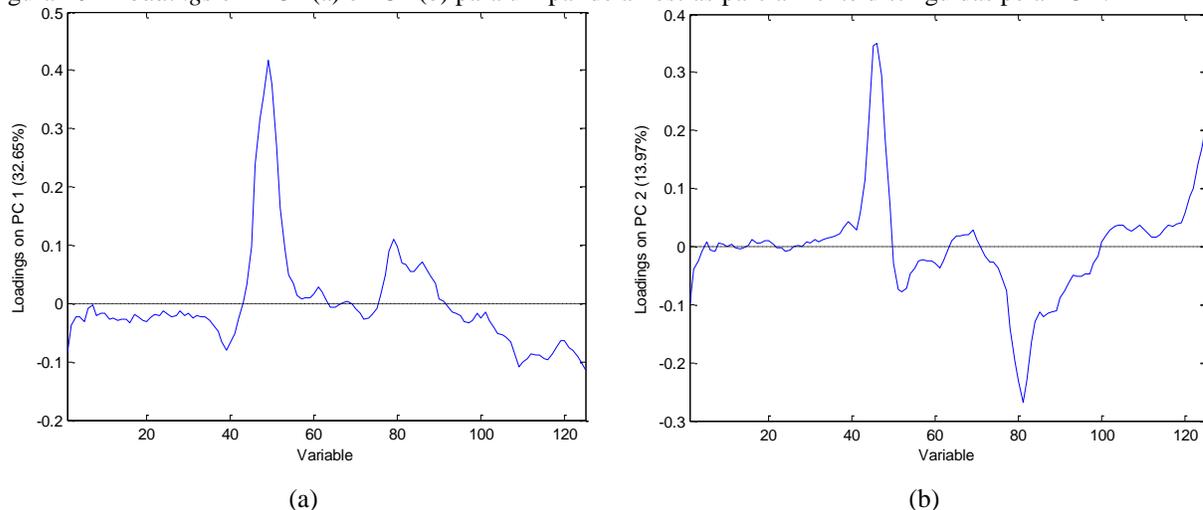
Figura 39 – *Scores* em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras parcialmente distinguidas visualmente pela PCA.



Fonte: O autor.

Os *loadings* em PC1 e PC2, apresentados na Figura 40, mostram que a principal fonte de variabilidade também foi devido a banda relacionada ao segundo sobretom da ligação C-H na região entre 1100 e 1300 nm (variáveis entre 40 e 60).

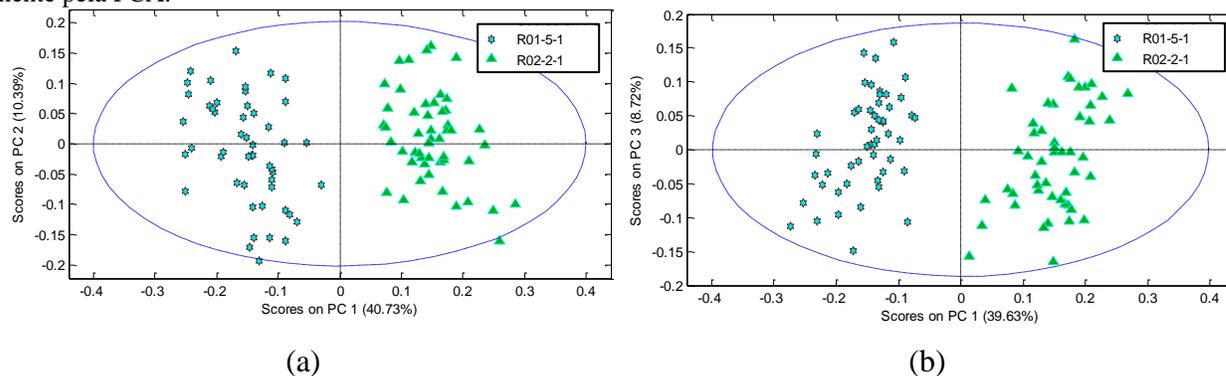
Figura 40 – *Loadings* em PC1 (a) e PC2 (b) para um par de amostras parcialmente distinguidas pela PCA.



Fonte: O autor.

Por fim, a Figura 41 mostra os scores em PC1xPC2 e PC1xPC3 de um exemplo do terceiro caso, em que os grupos puderam ser completamente distinguidos já ao se analisar os scores em PC1xPC2, o que leva à conclusão de que estas amostras são de fato distintas. O mesmo aconteceu em 1022 pares, 71,42% do total, o que mostra que a metodologia utilizada é efetiva em diferenciar amostras analisadas aos pares.

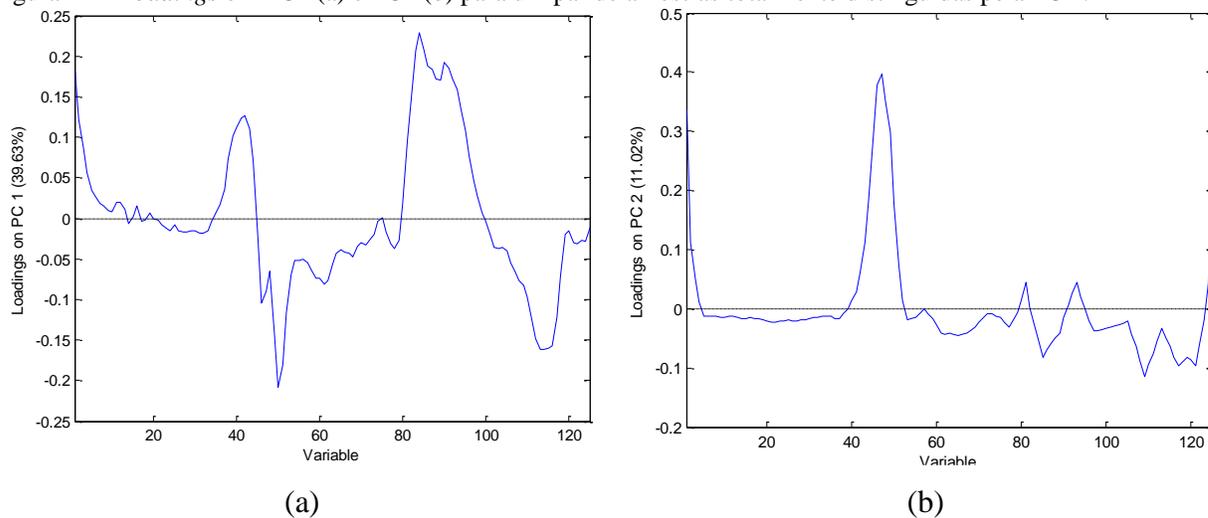
Figura 41 – *Scores* em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras completamente distinguidas visualmente pela PCA.



Fonte: O autor.

Neste caso, os *loadings* em PC1, apresentados na Figura 42(a), responsável pela separação das amostras, indicam as variáveis entre os comprimentos de onda de 1400 e 1500 nm (entre as variáveis 80 e 100), referentes sobretudo das vibrações de estiramento O-H. Já na PC2 se deve ao pico do segundo sobretudo da ligação C-H, como mostra a Figura 42(b).

Figura 42 – *Loadings* em PC1 (a) e PC2 (b) para um par de amostras totalmente distinguidas pela PCA.

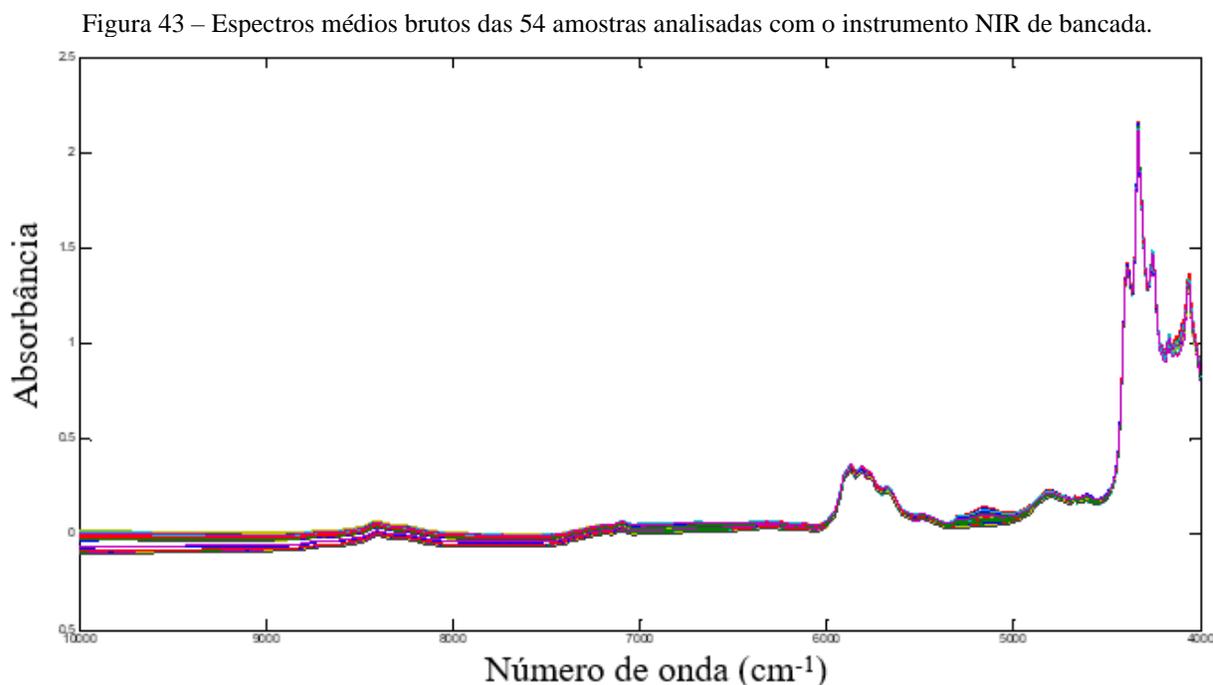


Fonte: O autor.

## 4.4 ESPECTROSCOPIA NO NIR EM INSTRUMENTO DE BANCADA

### 4.4.1 Espectros

Das 162 análises realizadas no instrumento de bancada, SPECTRUM 400, foram obtidos 54 espectros médios brutos na faixa de 10000 a 4000  $\text{cm}^{-1}$ , os quais são apresentados na Figura 43.



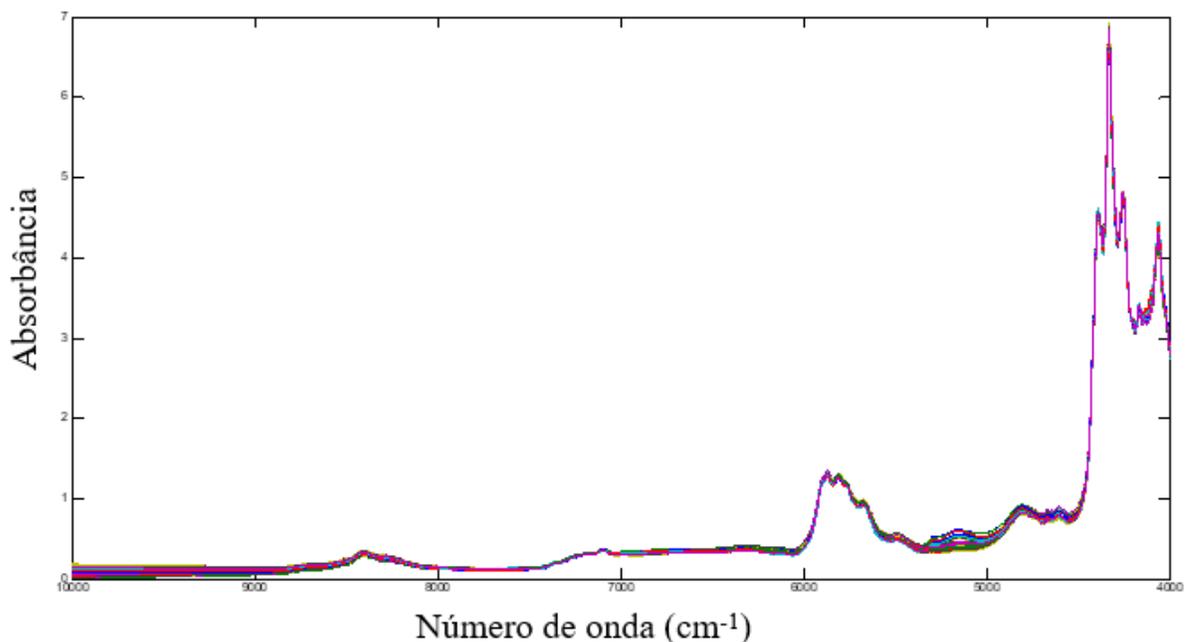
Fonte: O autor.

É possível notar regiões características da gasolina, tais quais as faixas entre 7862 e 8847  $\text{cm}^{-1}$ , referente ao segundo sobretom do modo de estiramento da ligação C-H dos hidrocarbonetos; entre 6536 e 7521  $\text{cm}^{-1}$ , relacionada com a combinação de modos de estiramento e deformação angular da ligação C-H; entre 5362 e 6233  $\text{cm}^{-1}$ , referente ao primeiro sobretom de estiramento da ligação C-H; entre 4737 e 5135  $\text{cm}^{-1}$ , referente à combinação dos modos de estiramento e deformação angular da ligação O-H do etanol; assim como a região entre 4135 e 4500  $\text{cm}^{-1}$ , referente à combinação de modos de estiramento e deformação angular da ligação C-H (OLIVEIRA *et.al* 2004).

Para eliminar efeitos indesejáveis de variação de linha de base, foram testados 3 pré-processamentos: MSC, correção de linha de base e SNV. Destes, o que apresentou melhores resultados foi o SNV, no entanto, os espectros permaneciam com diversos valores negativos de absorvância. Sendo assim, a fim de lidar com esta situação, a qual poderia afetar os cálculos das análises

seguintes, foi feito um deslocamento dos espectros ao longo do eixo y, o que tornou todos os valores de absorvância positivos. A Figura 44 mostra os 54 espectros médios pré-processados com SNV e deslocados ao longo do eixo y.

Figura 44 – Representação dos 54 espectros médios na região NIR adquiridos com equipamento de bancada pré-processados com SNV e deslocados no eixo y.



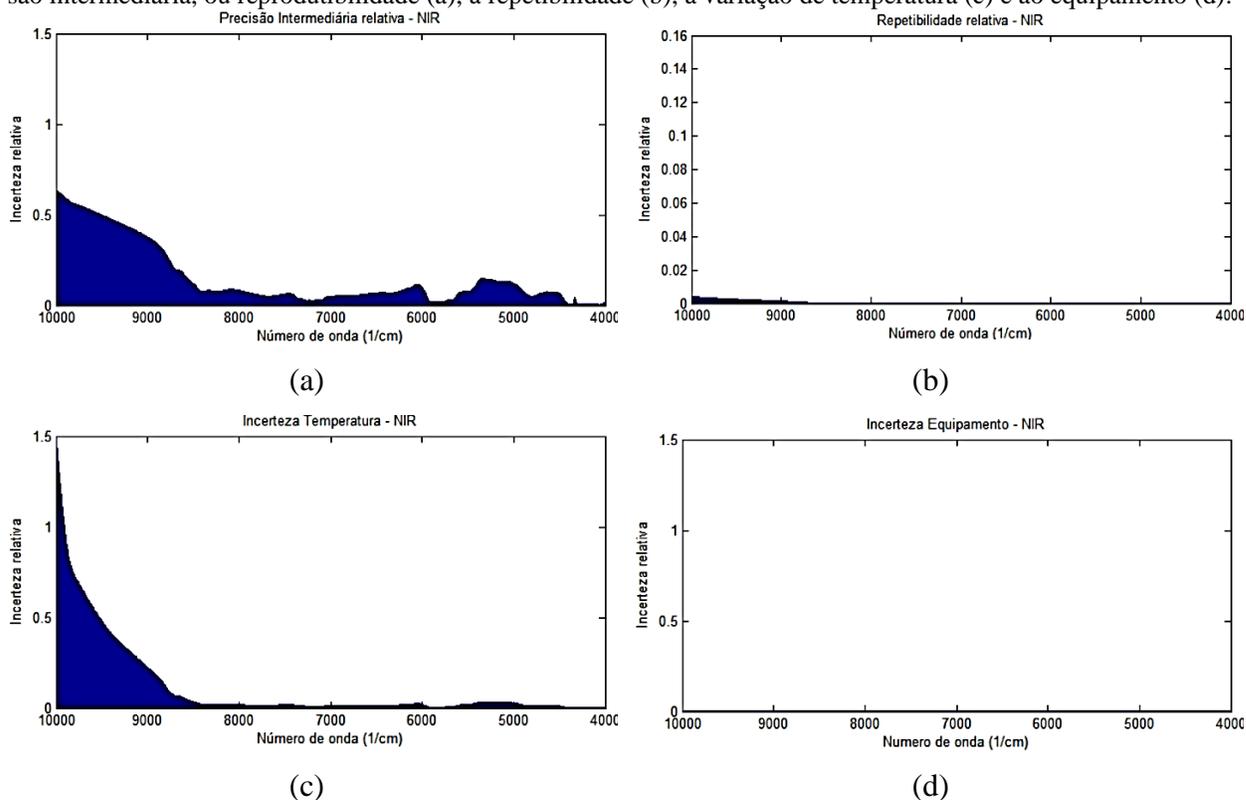
Fonte: O autor.

#### 4.4.2 Cálculo de incertezas

Além do conjunto de dados a ser analisado por PCA, também foram pré-processados com SNV (e deslocamento ao longo do eixo y) os conjuntos de dados utilizados para o cálculo das incertezas em cada número de onda, relativas a cada uma das 4 fontes: repetibilidade, reprodutibilidade (precisão intermediária), variação de temperatura da amostra e variação do equipamento em termos do número de onda.

A estes dados foi aplicada a metodologia já descrita, o que resultou nas incertezas relativas das fontes citadas, as quais são apresentadas na Figura 45.

Figura 45 – Incertezas estimadas dos espectros NIR adquiridos com um equipamento de bancada, referentes à precisão intermediária, ou reprodutibilidade (a), à repetibilidade (b), à variação de temperatura (c) e ao equipamento (d).

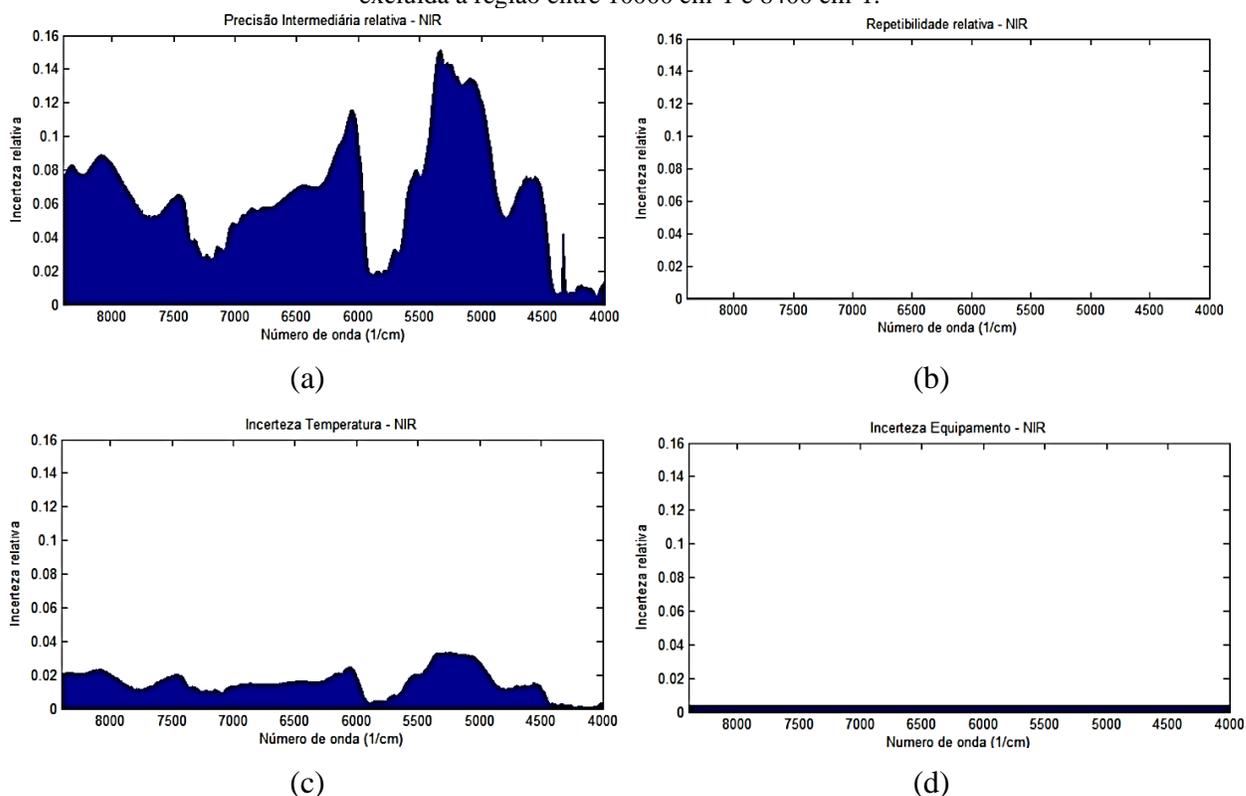


Fonte: O autor.

As incertezas resultantes apresentaram valores baixos, em especial aqueles referentes à repetibilidade e ao equipamento, os quais são na ordem de  $10^{-3}$ . No entanto, é perceptível que para a precisão intermediária, a repetibilidade e a incerteza pela variação de temperatura, os resultados obtidos para a região entre  $10000 \text{ cm}^{-1}$  e aproximadamente  $8400 \text{ cm}^{-1}$  são mais altos em comparação com os obtidos para os demais números de onda. Muito provavelmente isso se deve em maior parte à variabilidade apresentada pela variação da linha de base dos espectros nessa região, a qual não foi possível excluir por completo após os pré-processamentos.

Estes valores de incerteza para a região entre  $10000 \text{ cm}^{-1}$  e  $8400 \text{ cm}^{-1}$ , considerados superestimados, têm o potencial de afetar a análise posterior das amostras no que diz respeito à geração das amostras virtuais. Portanto, optou-se por excluir essa região dos espectros e dos valores de incerteza calculados. A Figura 46 apresenta os resultados para as incertezas das 4 fontes após exclusão da região superestimada, os quais de fato foram utilizados no estudo.

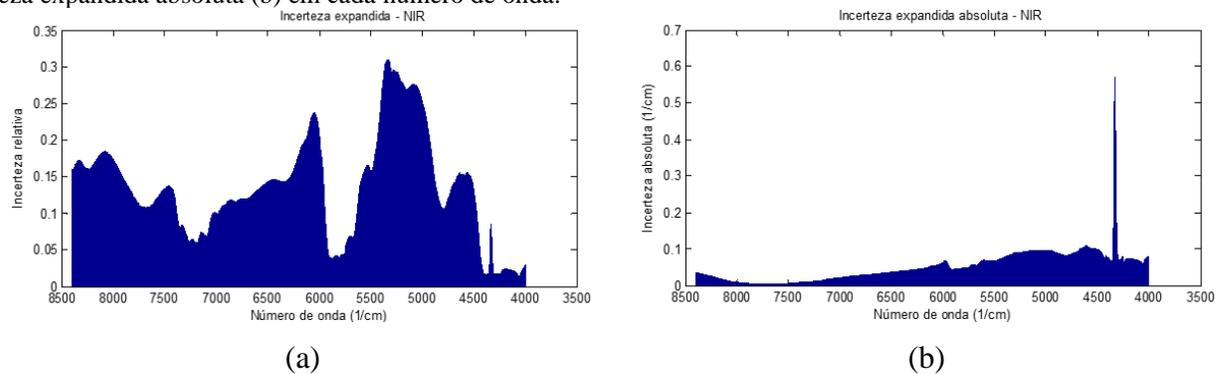
Figura 46 – Incertezas estimadas dos espectros NIR adquiridos com um equipamento de bancada, referentes à precisão intermediária, ou reprodutibilidade (a), à repetibilidade (b), à variação de temperatura (c) e ao equipamento (d) excluía a região entre 10000  $\text{cm}^{-1}$  e 8400  $\text{cm}^{-1}$ .



Fonte: O autor.

Essas incertezas foram combinadas e expandidas, seguindo a metodologia apresentada no ISO GUM, o que resultou nas incertezas expandidas relativa e absoluta, mostradas na Figura 47.

Figura 47 – Incerteza expandida dos espectros NIR adquiridos com um equipamento de bancada, relativa (a) e incerteza expandida absoluta (b) em cada número de onda.



Fonte: O autor.

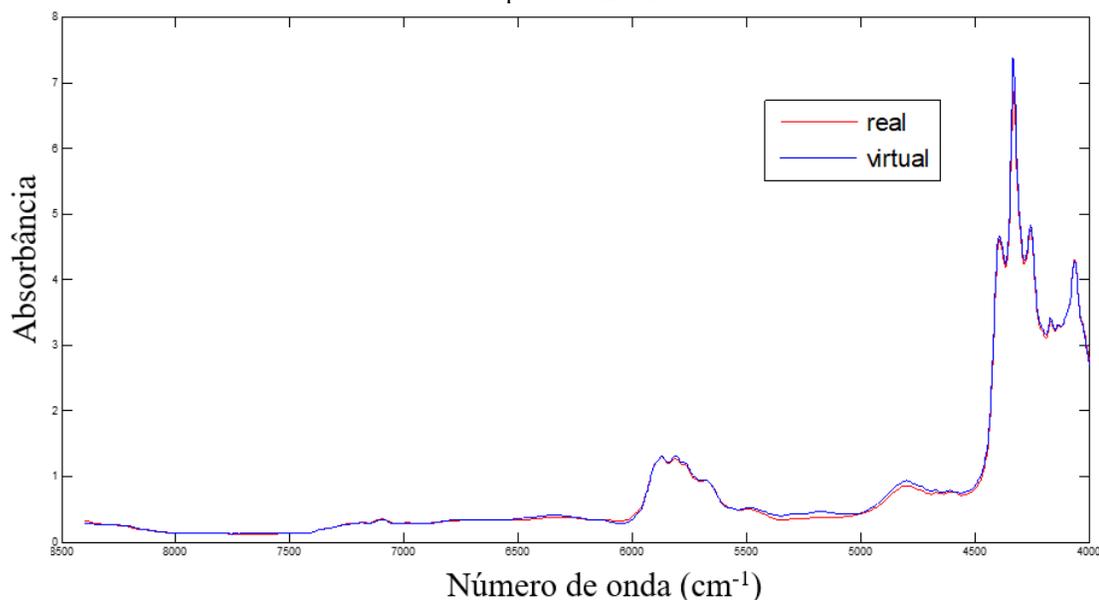
### 4.4.3 Geração de amostras virtuais

Para a geração das amostras virtuais, foram testados 5 diferentes intervalos (n) ao longo do número de onda, dentro dos quais há uma variação linear entre os valores randômicos gerados para o primeiro ponto do intervalo e para o último ponto do intervalo. Os testes foram feitos com os intervalos de 50, 75, 100, 150 e 200. Os melhores resultados foram obtidos com o intervalo de 200 pontos, o qual gerou espectros virtuais homogêneos, em comparação com os espectros reais, sem apresentar quebras bruscas dos valores de absorbância.

No APÊNDICE C, é possível visualizar exemplos de espectros virtuais gerados ao se utilizar os demais intervalos. Em intervalos muito baixos, os espectros virtuais obtidos se apresentaram pouco homogêneos em comparação com os espectros reais, com aumento ou diminuição bruscas de valores de absorbância em pontos próximos, o que os tornaram visualmente diferentes dos espectros reais obtidos em laboratório.

A Figura 48 mostra um exemplo de espectro real na região do NIR obtido por meio de análise em laboratório com o espectrofotômetro SPECTRUM 400, da Perkin Elmer, e um dos respectivos espectros virtuais gerados com o intervalo de 200 pontos.

Figura 48 – Exemplo de espectro real na região NIR adquirido com equipamento de bancada e um dos seus respectivos espectros virtuais.



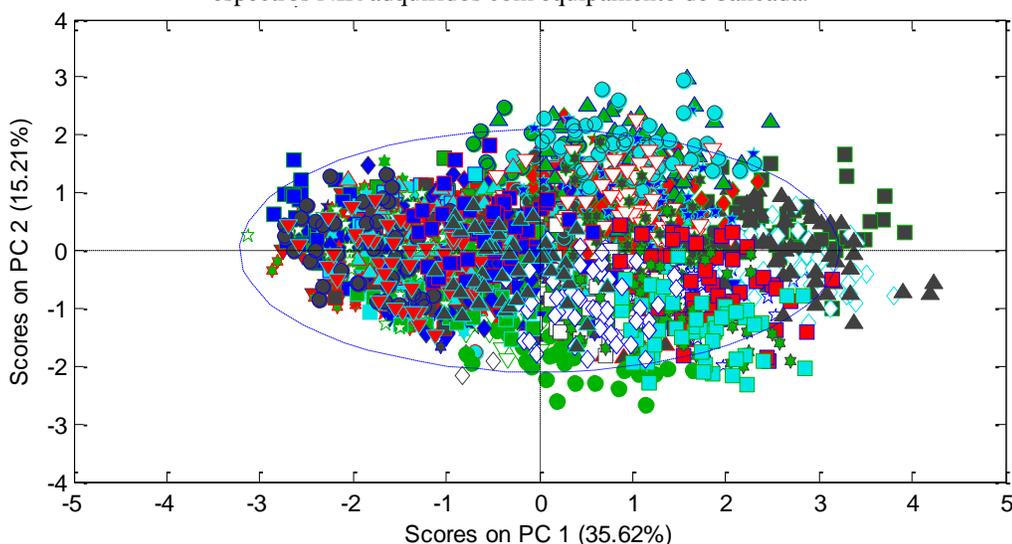
Fonte: O autor.

#### 4.4.4 Análise de Componentes Principais (PCA)

O conjunto de dados gerados, composto por 2700 espectros, sendo 54 grupos, cada um com 50 amostras (em que uma é a amostra real e as outras 49 são amostras geradas virtualmente), foi então pré-processado com uma centragem na média e SNV para em seguida ser realizada a PCA.

O resultado da PCA com todos os dados é apresentado na Figura 49, por meio da qual é possível perceber que não há distinção clara entre os grupos de amostras. Isso mostra que não é possível distinguir nenhum dos grupos estudados quando analisado por PCA o conjunto total de dados, devido à grande variabilidade das amostras.

Figura 49 – Scores em PC1 e PC2 para a PCA aplicada a todos os 54 grupos de amostras do conjunto de dados de espectros NIR adquiridos com equipamento de bancada.



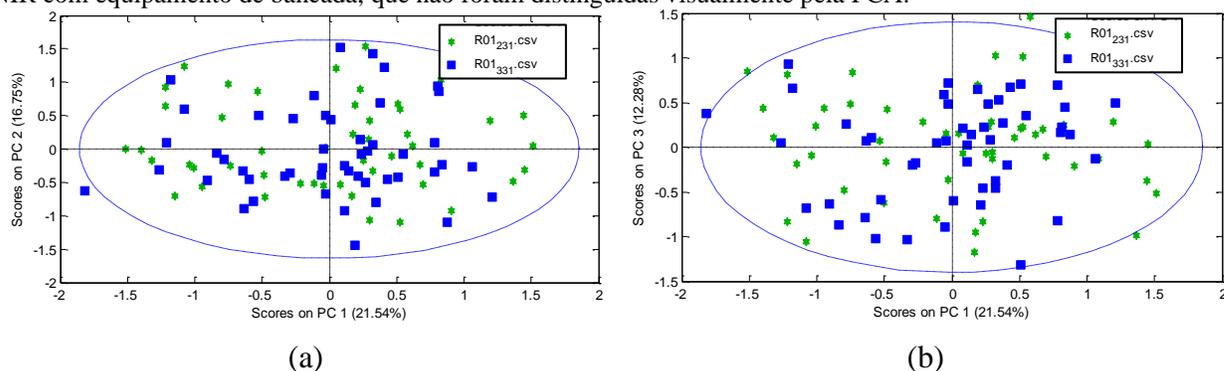
Fonte: O autor.

Os grupos de amostras foram então estudados aos pares a fim de analisar o potencial de distinção visual das amostras. O total de pares formados pelos 54 grupos de amostras foi de 1431, os quais, ao serem analisados por PCA, apresentaram 3 padrões distintos, assim como no estudo realizado com o MicroNIR: alguns dos pares não puderam ser distinguidos, alguns puderam ser distinguidos parcialmente e outros puderam ser distinguidos por completo.

A Figura 50 mostra um exemplo de par cujas amostras não puderam ser distinguidas por meio de uma análise visual dos scores da PCA, mesmo ao se analisar mais de 5 PCs, o que aconteceu em cerca de 294 pares, 20,54% dos 1431 pares totais. Percebe-se que as amostras de ambos

os grupos possuem scores tais que as tornam homogêneas, portanto, nestes casos não é possível dizer que as amostras reais de cada grupo possuem origens distintas.

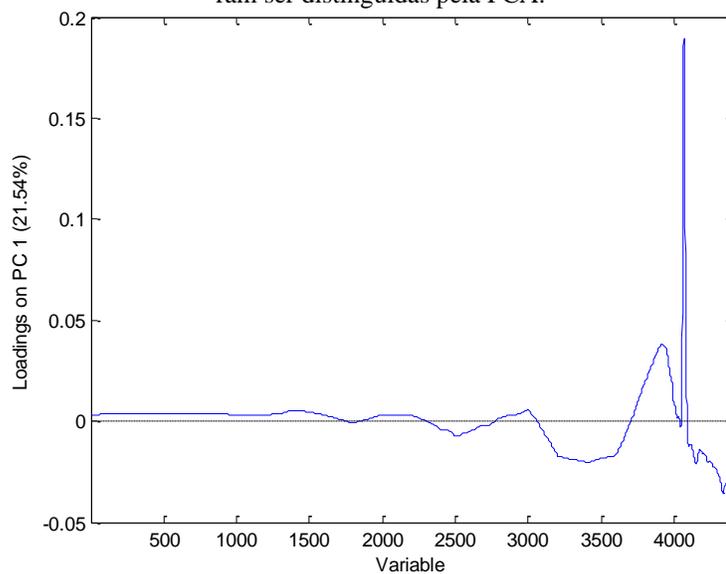
Figura 50 – Scores em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, que não foram distinguidas visualmente pela PCA.



Fonte: O autor.

A Figura 51 apresenta os *loadings* em PC1, que explica 21,54% da variabilidade das amostras. A variável que mais contribuiu para esta variabilidade foi a do número de onda de  $4333\text{ cm}^{-1}$  (variável número 4067), o pico de maior absorvância em todos os espectros, situado na região correspondente à combinação de modos de estiramento e deformação angular da ligação C-H.

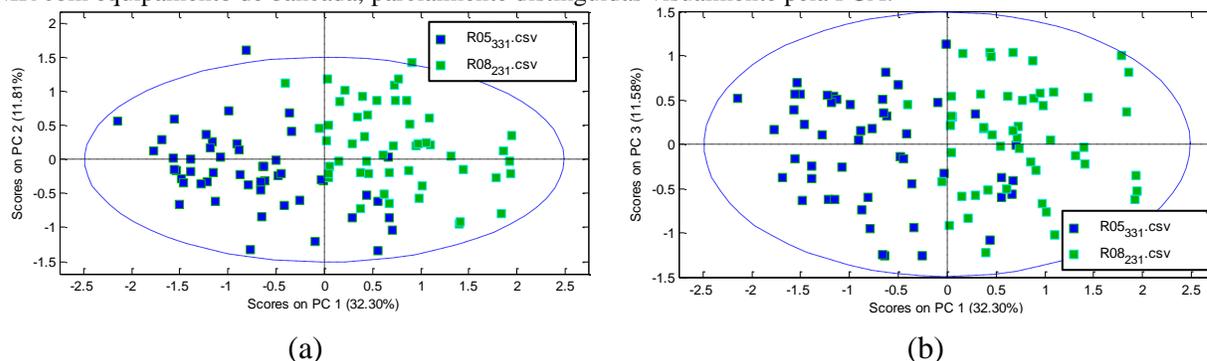
Figura 51 – *Loadings* em PC1 para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, que não puderam ser distinguidas pela PCA.



Fonte: O autor.

A segunda situação se refere aos pares que foram parcialmente separados, como mostram os *scores* em PC1xPC2 e PC1xPC3 apresentados na Figura 52. Por meio de tais *scores* percebe-se que os grupos não estão completamente separados, no entanto, há uma tendência de separação, visto que apenas uma pequena parcela de amostras está numa região intermediária, podendo pertencer a ambos os grupos. Sendo assim, não é possível afirmar com que as amostras reais destes dois grupos são de fato distintas. No entanto, esse baixo nível de homogeneidade e a tendência de separação indicam que muito provavelmente são amostras distintas. Cerca de 200 pares apresentaram este comportamento, o que representa 13,98% do total.

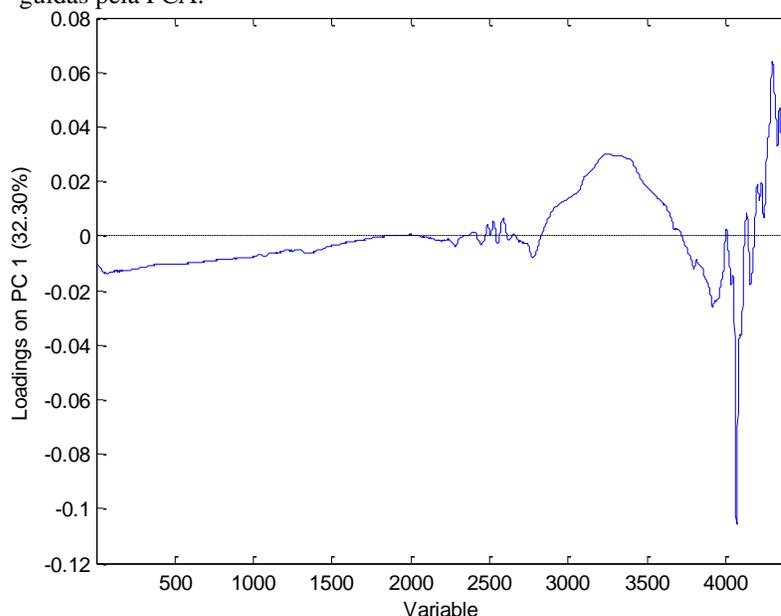
Figura 52 – Scores em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, parcialmente distinguidas visualmente pela PCA.



Fonte: O autor.

Os *loadings* em PC1, apresentados na Figura 53, mostram que, mais uma vez, a principal fonte de variabilidade nesta PC, responsável pela tendência de separação, se deve ao número de onda de  $4333\text{ cm}^{-1}$  (pico de maior absorvância dos espectros) e também a região em torno de  $5362$  e  $6233\text{ cm}^{-1}$ , referente ao primeiro sobretom de estiramento da ligação C-H.

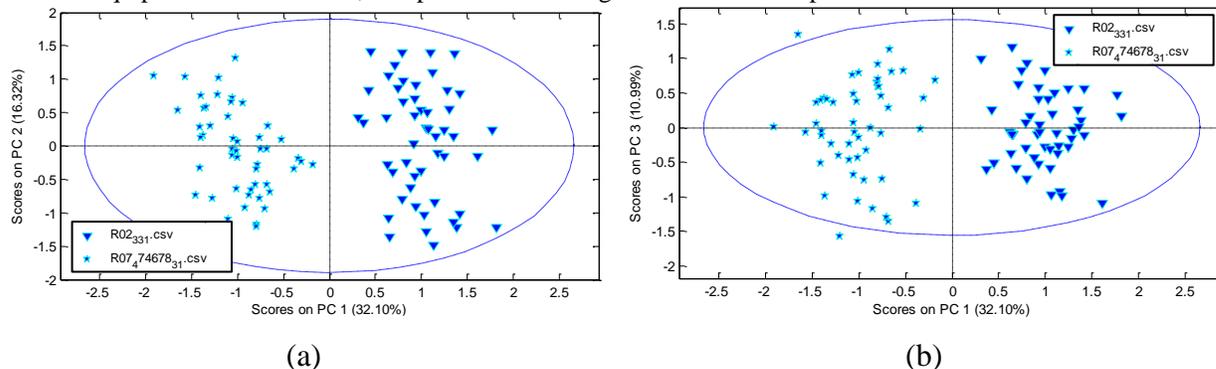
Figura 53 – *Loadings* em PC1 para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, parcialmente distinguidas pela PCA.



Fonte: O autor.

Já para os pares que puderam ser completamente distinguidos, na Figura 54 é mostrado um exemplo (*scores* em PC1xPC2 e PC1xPC3). É notável a separação completa entre os grupos, com 32,1% da variabilidade das amostras explicada por PC1, o que leva à conclusão de que as amostras reais são de fato de origens distintas. Este mesmo comportamento aconteceu para 937 pares, cerca de 65,48% do total de pares, o que mostra que a metodologia utilizada é efetiva nesta proporção em diferenciar amostras analisadas em pares.

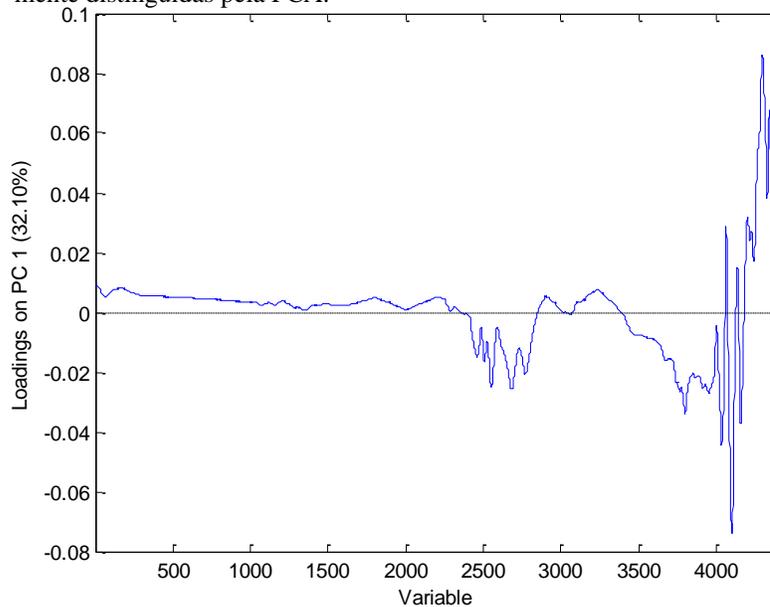
Figura 54 - *Scores* em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, completamente distinguidas visualmente pela PCA.



Fonte: O autor.

Os *loadings* em PC1 são apresentados na Figura 55, e mostram que as variáveis que mais contribuíram para a variabilidade, foram aquelas na região entre 4135 e 4500  $\text{cm}^{-1}$ , referente à combinação de modos de estiramento e deformação angular da ligação C-H.

Figura 55 – *Loadings* em PC1 (a) e PC2 (b) para um par de amostras, cujos espectros foram adquiridos no NIR com equipamento de bancada, totalmente distinguidas pela PCA.



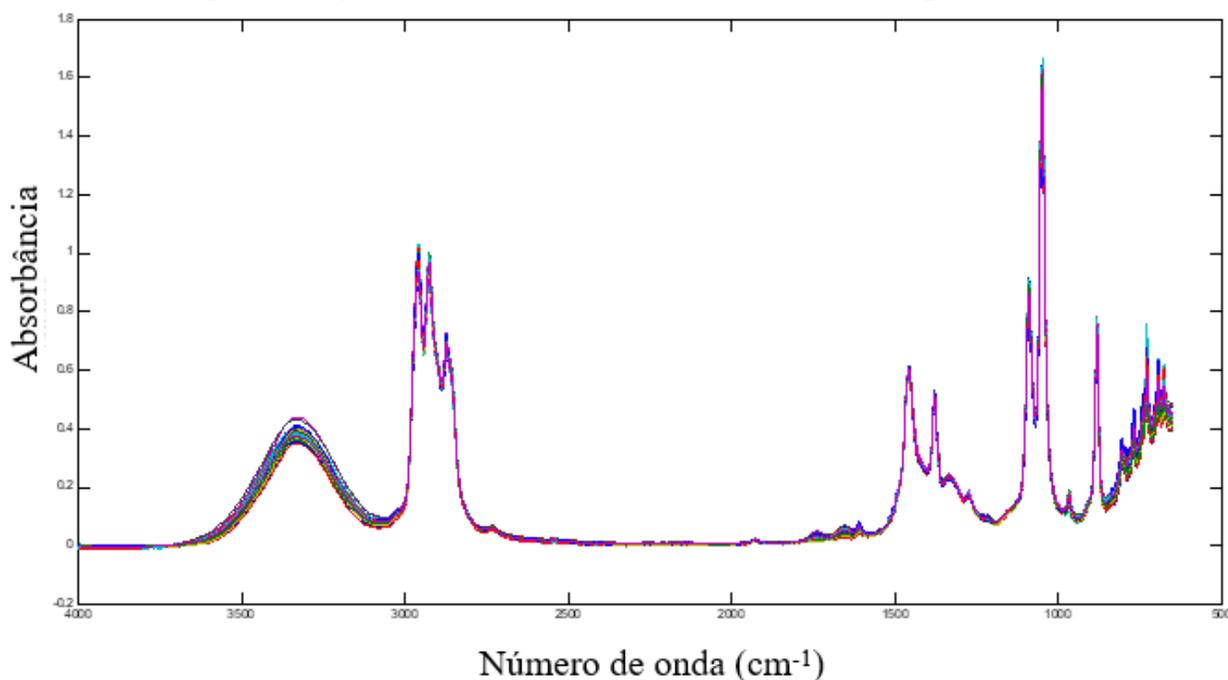
Fonte: O autor.

## 4.5 ESPECTROSCOPIA NO MIR

### 4.5.1 Espectros

Das 162 análises realizadas na região do MIR no instrumento de bancada, SPECTRUM 400, foram obtidos 54 espectros médios na faixa de 4000 a 650  $\text{cm}^{-1}$ , os quais são apresentados, de forma bruta, na Figura 56.

Figura 56 – Espectros médios brutos das 54 amostras analisadas na região do MIR.



Fonte: O autor.

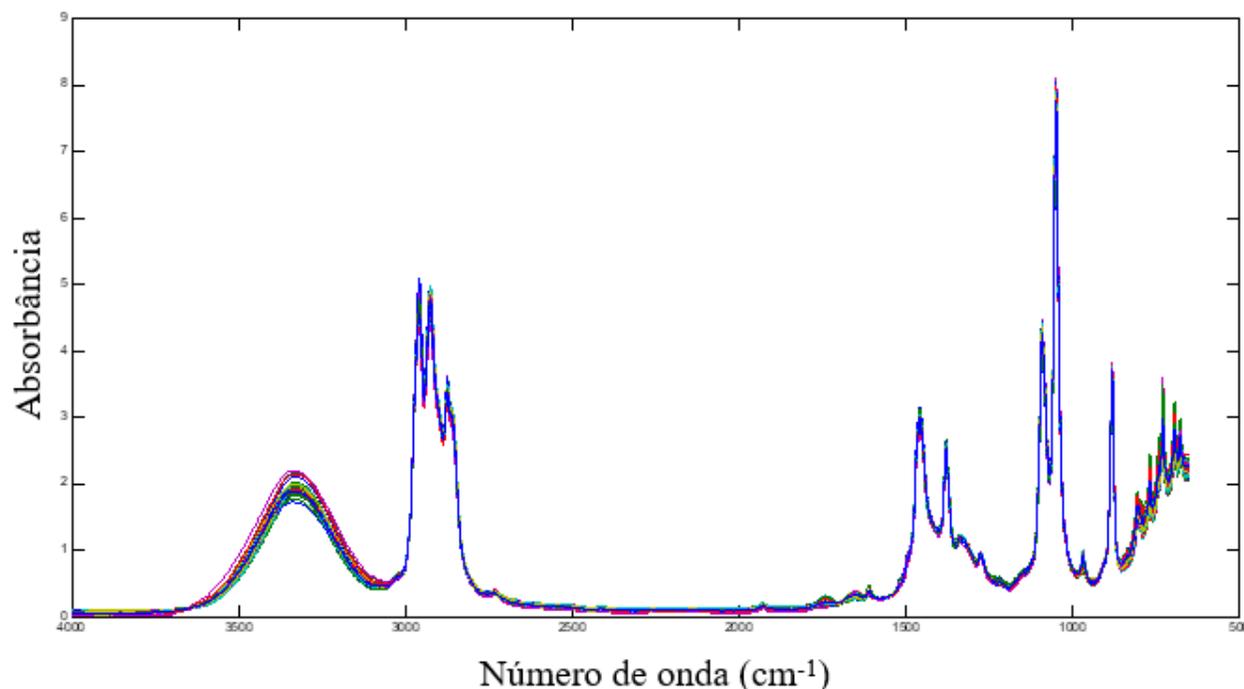
Percebe-se as bandas características de gasolinas brasileiras em espectros na região do MIR. No intervalo entre 3662 e 3101  $\text{cm}^{-1}$ , é possível notar a banda larga de estiramento da ligação O-H. Já na região entre 3020 e 2802  $\text{cm}^{-1}$  estão as bandas características de estiramentos das ligações C-H. Na região de menores números de onda estão as bandas referentes às deformações angulares das ligações C-H (intervalo entre 1500 e 1200  $\text{cm}^{-1}$ ) e ao estiramento C-O entre 1200 e 1100  $\text{cm}^{-1}$  (SILVA, 2013; CORSETTI *et. al*, 2015).

Os espectros brutos foram então pré-processados, para isso, foram testados 3 pré-processamentos: MSC, correção de linha de base e SNV. Destes, o que apresentou melhores resultados foi o SNV. No entanto, os espectros apresentam diversos valores negativos de absorvância, o que exigiu a realização de um deslocamento dos espectros ao longo do eixo y para tornar todos os valores

de absorvância positivos, tendo em vista que a existência de valores negativos poderia afetar os cálculos a serem realizados posteriormente.

A Figura 57 mostra os 54 espectros médios pré-processados com SNV e deslocados ao longo do eixo y.

Figura 57 – Representação dos 54 espectros médios na região do MIR pré-processados com SNV e deslocados no eixo y.



Fonte: O autor.

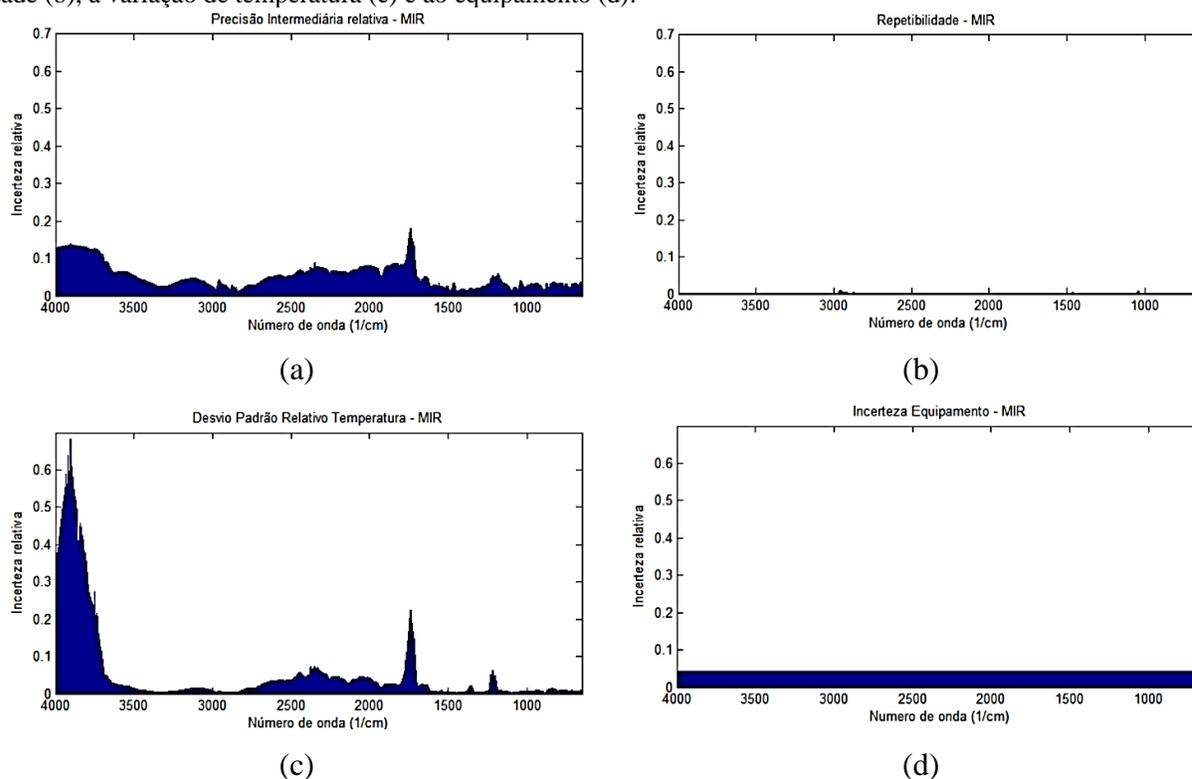
#### 4.5.2 Cálculo de incertezas

A fim de realizar o estudo seguindo a metodologia proposta, foi inicialmente realizado o cálculo da incerteza em cada número de onda do espectro. Dessa forma, foi necessário estimar as incertezas relativas a cada uma das 4 fontes: repetibilidade, reprodutibilidade (precisão intermediária), variação de temperatura da amostra e variação do equipamento em termos do número de onda.

Para o cálculo da incerteza relativa a cada uma dessas fontes, foram utilizados conjuntos de dados específicos, os quais também foram pré-processados com SNV e deslocados ao longo do eixo y.

Em seguida, as incertezas foram estimadas de acordo com a metodologia descrita, resultando nos valores mostrados na Figura 58.

Figura 58 – Incertezas dos espectros MIR referentes à precisão intermediária, ou reprodutibilidade (a), à repetibilidade (b), à variação de temperatura (c) e ao equipamento (d).



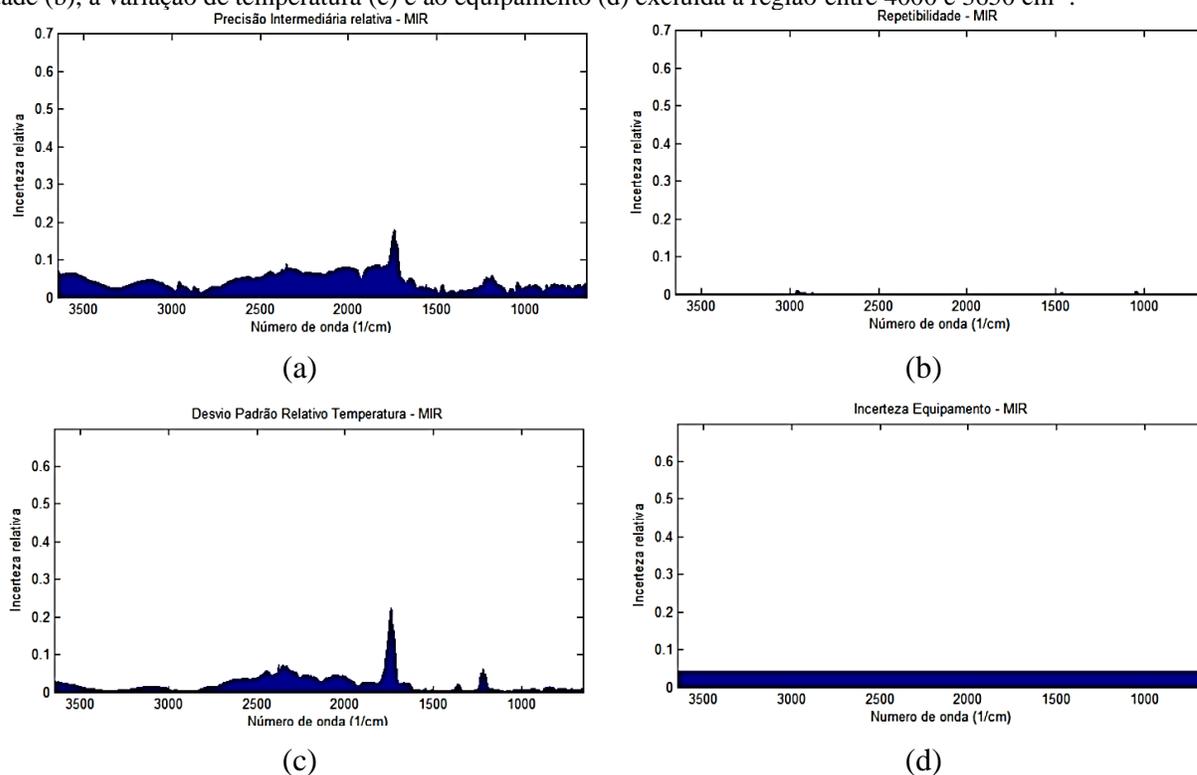
Fonte: O autor.

É perceptível a existência de valores de incerteza altos nas regiões de números de onda maiores, entre 4000 e 3650  $\text{cm}^{-1}$ , para os casos da precisão intermediária (Figura 58 (a)) e da incerteza devido à variação de temperatura (Figura 58 (b)). Isso não era esperado, visto que esta região não apresenta bandas importantes, sendo composta principalmente por ruído, o que leva à conclusão de que estes altos valores de incerteza foram superestimados devido à variabilidade apresentada pela variação da linha de base dos espectros nessa região, a qual não pode ser excluída por completo após os pré-processamentos, assim como aconteceu no caso do estudo feito na região do NIR utilizando o SPECTRUM 400.

Considerando que o uso de valores superestimados poderia ter grande impacto na estimativa da incerteza expandida pela metodologia proposta, optou-se por excluir a região entre 4000 e 3650

$\text{cm}^{-1}$  dos espectros e dos valores de incerteza calculados. A Figura 59 apresenta os resultados após a exclusão dessa região.

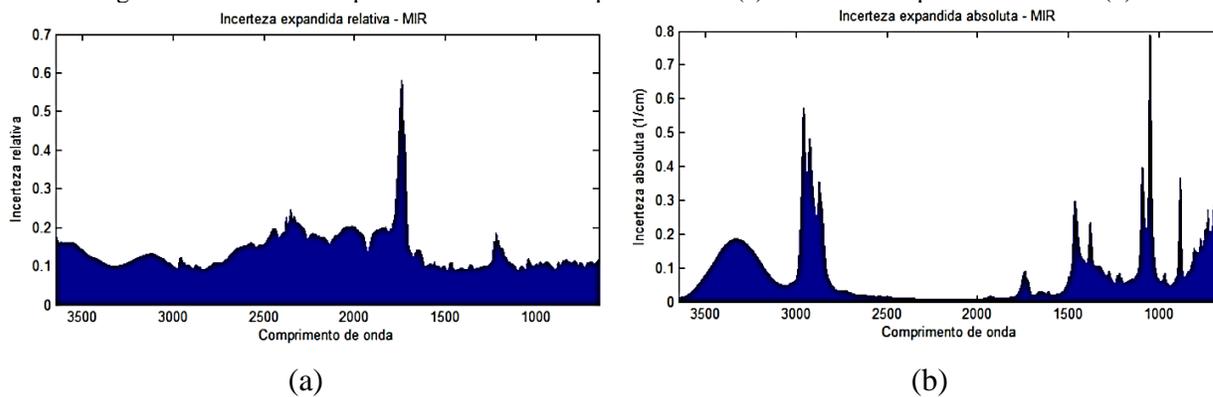
Figura 59 – Incertezas dos espectros MIR referentes à precisão intermediária, ou reprodutibilidade (a), à repetibilidade (b), à variação de temperatura (c) e ao equipamento (d) excluía a região entre 4000 e 3650  $\text{cm}^{-1}$ .



Fonte: O autor.

Posteriormente, as incertezas das 4 fontes foram utilizadas para o cálculo da incerteza combinada e, por fim, da incerteza expandida. A Figura 60 mostra os valores de incertezas expandidas relativa (a) e absoluta (b) obtidos para cada número de onda.

Figura 60 – Incerteza expandida relativa dos espectros MIR (a) e incerteza expandida absoluta (b).



Fonte: O autor.

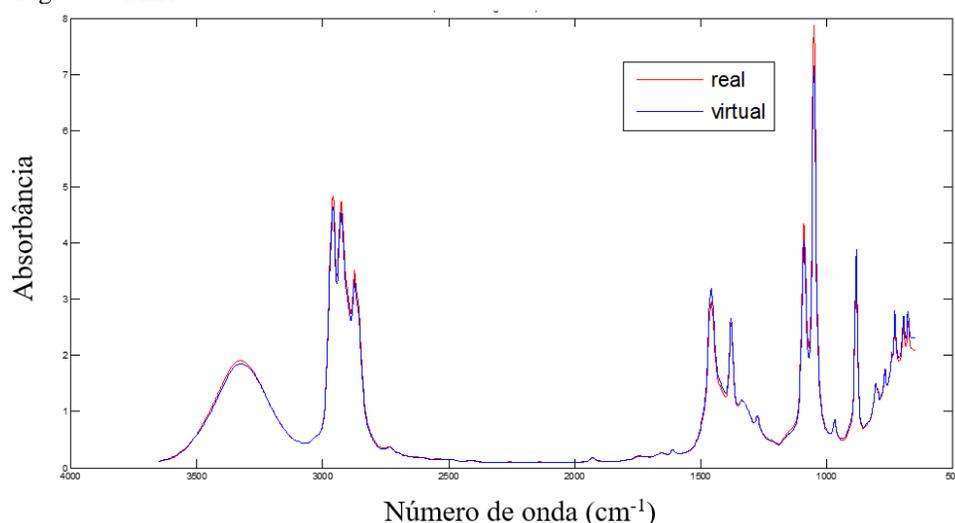
### 4.5.3 Geração de amostras virtuais

Com os valores de incertezas expandidas absolutas, foram geradas as 49 amostras virtuais para cada uma das 54 amostras reais seguindo a metodologia descrita. Foram testados 5 diferentes intervalos (n) ao longo do número de onda. Os testes foram feitos com os intervalos de 05, 25, 50, 67 e 134 pontos. Os melhores espectros virtuais obtidos foram aqueles gerados utilizando-se intervalo de 134 pontos, o qual gerou espectros virtuais homogêneos, em comparação com os espectros reais, sem apresentar quebras bruscas dos valores de absorvância.

No APÊNDICE D, é possível visualizar exemplos de espectros virtuais gerados ao se utilizar os demais intervalos.

A Figura 61 mostra um exemplo de espectro real na região do MIR obtido por meio de análise em laboratório com o espectrofotômetro SPECTRUM 400, da Perkin Elmer, e um dos respectivos espectros virtuais gerados com o intervalo de 134 pontos. É possível notar que o espectro virtual apresentado possui forma muito semelhante ao seu espectro real e uma aparência homogênea.

Figura 61 – Exemplo de espectro real e um dos seus respectivos espectros virtuais na região do MIR.



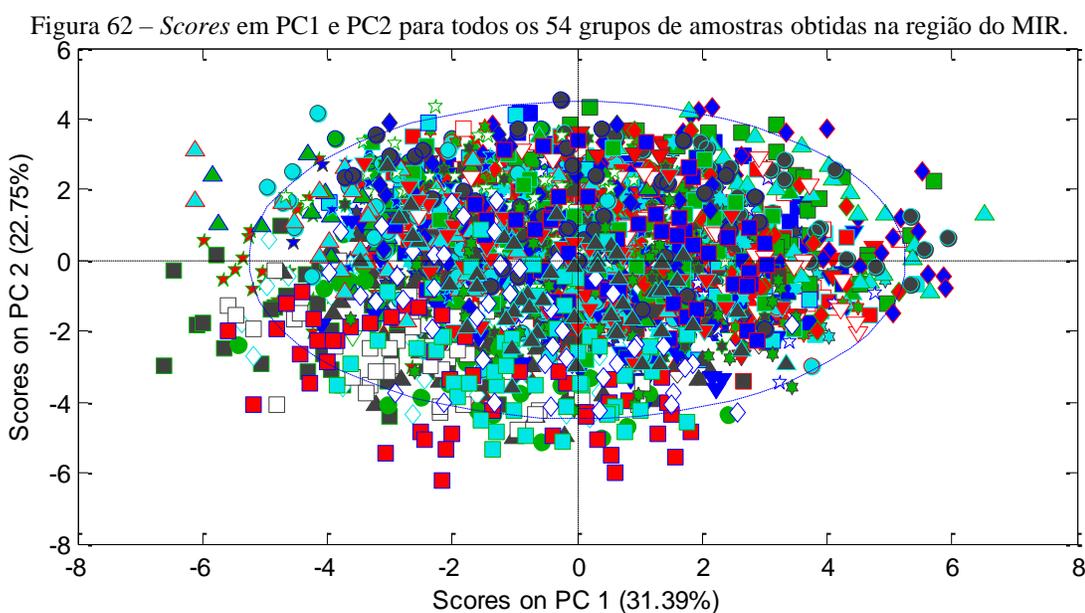
Fonte: O autor.

Como resultado, foram obtidos um total de 2700 amostras (espectros), divididos em 54 grupos, cada um composto por 1 amostra real e suas respectivas 49 amostras virtuais. Desta maneira, as amostras de um dado grupo são na verdade variações de uma medida real (obtida em laboratório) dentro das incertezas calculadas.

#### 4.5.4 Análise de Componentes Principais (PCA)

O conjunto de dados gerados foi então pré-processado com uma centragem na média e SNV e em seguida analisados através da PCA.

O resultado da PCA com todos os dados é apresentado na Figura 62, por meio da qual é possível perceber que não há distinção clara entre os grupos de amostras. Isso mostra que não é possível distinguir nenhum dos grupos estudados quando analisados por PCA ao mesmo tempo.

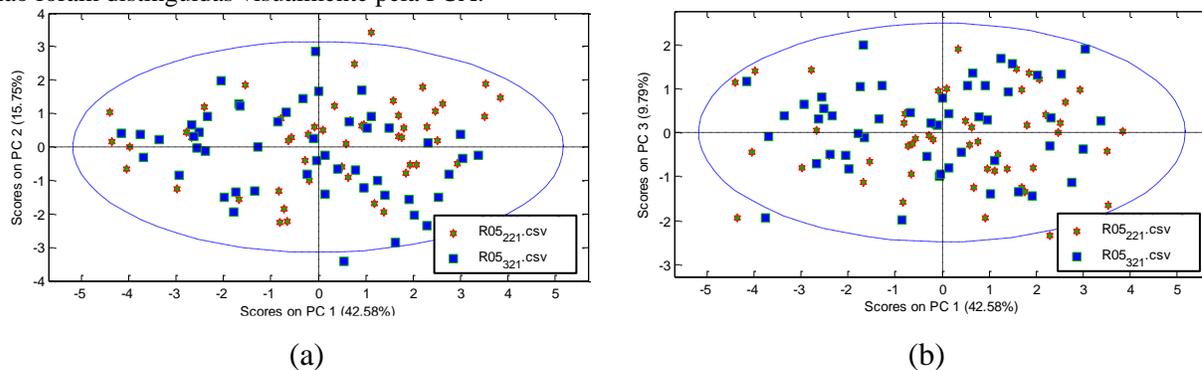


Fonte: O autor.

Os grupos de amostras foram então estudados aos pares a fim de analisar o potencial de separação visual no gráfico dos escores das PCs. Ao serem analisados por meio de PCAs, os 1431 pares de grupos de amostras apresentaram 3 padrões distintos, assim como nos casos anteriores: alguns dos pares não puderam ser distinguidos, alguns puderam ser distinguidos parcialmente e outros puderam ser distinguidos por completo.

A Figura 63 mostra um exemplo de par cujas amostras não puderam ser distinguidas por meio de uma análise visual dos scores da PCA, mesmo ao se analisar mais de 5 PCs, o que aconteceu em cerca de 420 pares, 29,35% dos 1431 pares totais. Nota-se que as amostras de ambos os grupos possuem scores tais que as tornam homogêneas, não sendo possível dizer que as amostras reais de cada grupo possuem origens distintas.

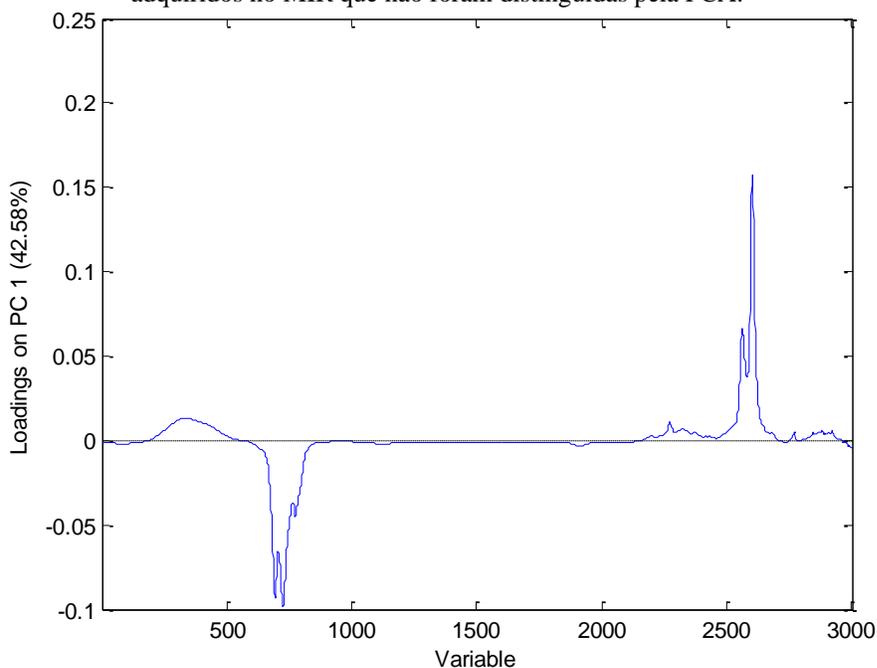
Figura 63 – Scores em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, com espectros adquiridos no MIR, que não foram distinguidas visualmente pela PCA.



Fonte: O autor.

Os *loadings* apresentados na Figura 64 mostram que as variáveis que mais contribuíram para a variabilidade explicada por PC1 foram referentes às regiões entre 3020 e 2802  $\text{cm}^{-1}$  (bandas características de estiramentos das ligações C-H) e entre 1200 e 1100  $\text{cm}^{-1}$  (estiramento C-O).

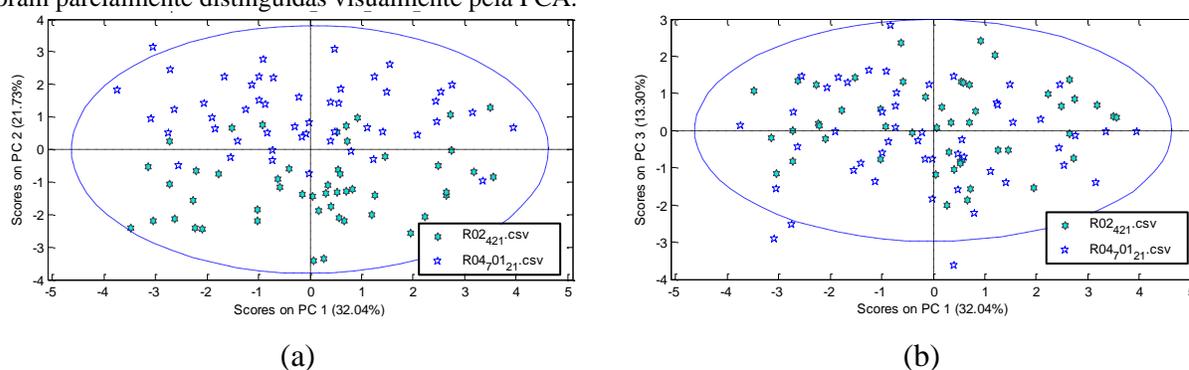
Figura 64 – *Loadings* em PC1 para um par de amostras com espectros adquiridos no MIR que não foram distinguidas pela PCA.



Fonte: O autor.

Quanto ao segundo caso, em que as amostras não foram completamente distinguidas, a Figura 65 mostra os *scores* em PC1xPC2 e PC1xPC3 para um par de amostras que se enquadraram nessa situação. É possível visualizar que os *scores* de PC1xPC2 apresentam os grupos com uma tendência de separação, mas ainda parcialmente misturados. Nestes casos, não é possível concluir que as amostras reais dos dois grupos possuem origens distintas, mas há um indicativo de que podem ser de diferente procedência. Cerca de 780 pares apresentaram este comportamento, o que representa 54,51% do total.

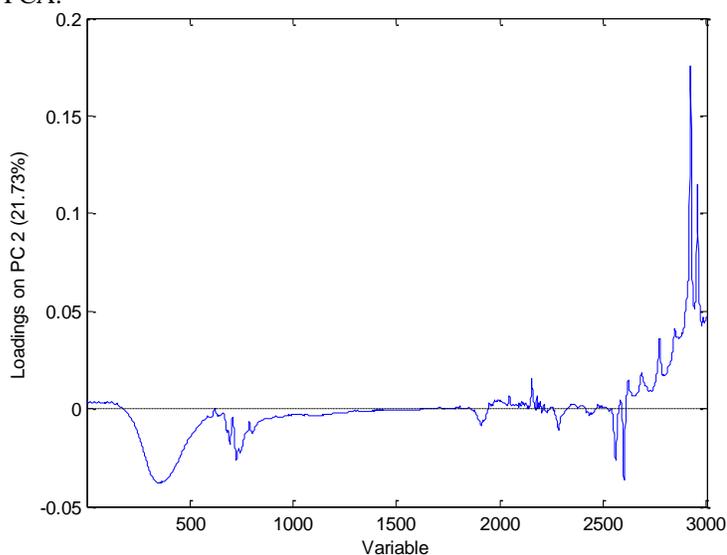
Figura 65 – *Scores* em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras, com espectros adquiridos no MIR, que foram parcialmente distinguidas visualmente pela PCA.



Fonte: O autor.

Os *loadings* apresentados na Figura 66 mostram que em PC2, responsável por explicar a maior variabilidade entre as amostras, as maiores contribuições, se devem às bandas nos menores números de onda, referentes ao estiramento C-O entre 1200 e 1100  $\text{cm}^{-1}$  e também à região entre 3662 e 3101  $\text{cm}^{-1}$ , da banda larga de estiramento da ligação O-H.

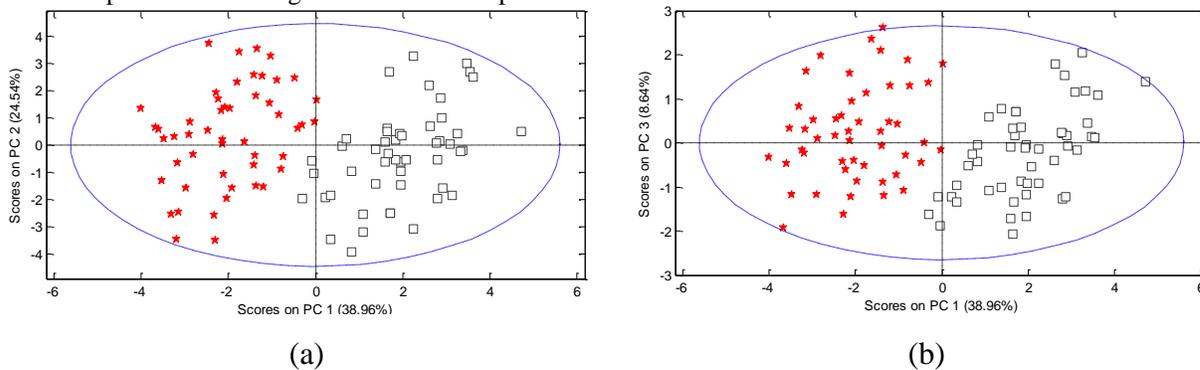
Figura 66 – *Loadings* em PC1 (a) e PC2 (b) para um par de amostras, com espectros adquiridos no MIR, parcialmente distinguidas pela PCA.



Fonte: O autor.

Por fim, foram cerca de 230 pares que se apresentaram completamente separados por meio de análise visual dos *scores* das PCs, o que representa aproximadamente 16,07% de todos os 1431 pares. Assim, os espectros MIR, empregando a estratégia desenvolvida, foram menos eficientes para distinguir pares de amostras que os espectros obtidos na região NIR. A Figura 67 mostra os *scores* em PC1xPC2 e PC1xPC3 para um par de amostras que foram completamente distinguidas pela PCA.

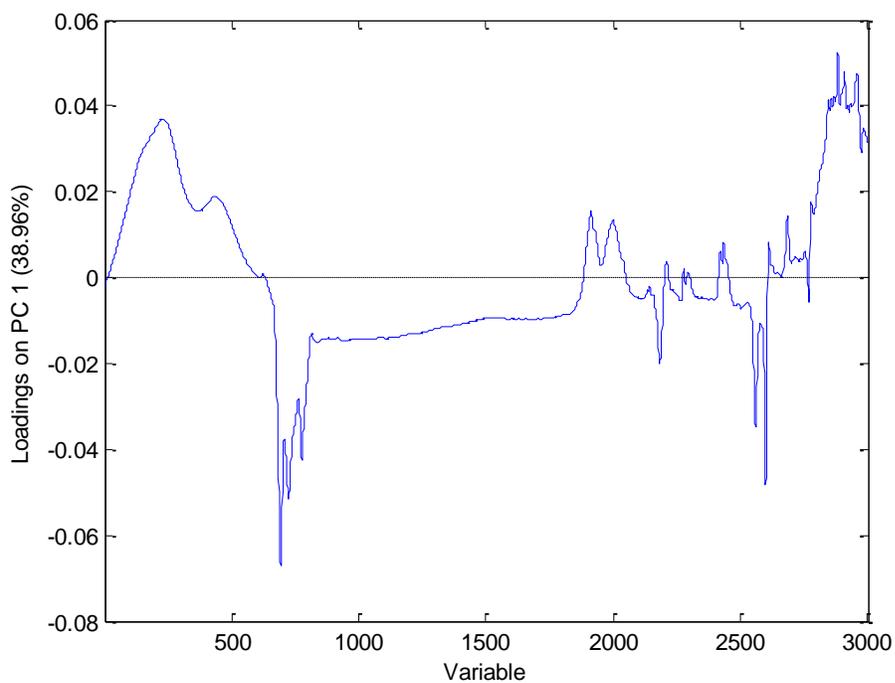
Figura 67 – *Scores* em PC1xPC2 (a) e PC1xPC3 (b) para um par de amostras com espectros adquiridos no MIR que foram completamente distinguidas visualmente pela PCA.



Fonte: O autor.

Para este caso, são apresentados os *loadings* em PC1 na Figura 68, responsável por explicar 38,96% da variabilidade dos dados. Nesta PC, as variáveis que mais explicam esta variabilidade são referentes às bandas características de estiramentos das ligações C-H, na região entre 3020 e 2802  $\text{cm}^{-1}$ , assim como o intervalo entre 3662 e 3101  $\text{cm}^{-1}$ , referente ao estiramento da ligação O-H e estiramento C-O entre 1200 e 1100  $\text{cm}^{-1}$ .

Figura 68 – *Loadings* em PC1 para um par de amostras com espectros adquiridos no MIR completamente distinguidas pela PCA.



Fonte: O autor.

## 5 CONCLUSÕES

Foi desenvolvida uma estratégia para estimar incertezas de medição em dados espectrais, algo não encontrado na literatura. Isso permitiu a geração computacional de amostras virtuais dentro da faixa das incertezas estimadas. Foi possível, então, utilizar tanto as amostras reais quanto as virtuais na tentativa de avaliar o potencial de separação dessas amostras pelos scores da PCA.

A aplicação da metodologia proposta aos dados dos ensaios convencionais da ANP mostrou que foi possível distinguir completamente, através da PCA, 97% dos 1431 pares possíveis de amostras. Os outros pares, 3% do total, foram parcialmente separados pelos *scores* da PCA, apresentando uma alta tendência de separação, o que indica que muito provavelmente as amostras dos pares também possuem origem distintas. Isso, no entanto, não pode ser afirmado com segurança, utilizando-se a metodologia empregada neste trabalho, visto que, nestes casos, as amostras não estão completamente separadas pela análise visual da PCA.

Para as curvas de destilação das gasolinas estudadas, empregando a estratégia proposta, foi possível separar 100% dos pares analisados. Desses, apenas 1 par apresentou as amostras muito próximas na análise visual das PCs, necessitando de 7 PCs para visualizar uma clara separação entre as amostras. Já para os outros 1430 pares, as amostras se apresentaram claramente separadas nas primeiras PCs.

Para os dados espectrais, os pares de amostras apresentaram 3 comportamentos:

- a) **Pares que não puderam ser distinguidos:** 13,9% dos pares, quando se utilizou o equipamento portátil no NIR; 20,54%, quando os espectros foram adquiridos com um equipamento de bancada no NIR; e 29,35%, quando os espectros foram adquiridos no MIR. Nestes casos, não é possível afirmar com segurança quanto à origem destas amostras, mas há um indicativo de que podem possuir a mesma origem.
- b) **Pares que puderam ser distinguidos parcialmente:** 14,67% dos pares, quando os espectros foram adquiridos no NIR com equipamento portátil; 13,98%, quando empregado equipamento de bancada no NIR; e 54,51%, com espectros no MIR. Nestes casos, é possível visualizar uma forte tendência de separação, mas ainda há sobreposição entre amostras.
- c) **Pares que puderam ser distinguidos por completo:** 71,42% dos pares de amostras com espectros adquiridos na região NIR com equipamento portátil; 65,48% dos pares de amostras com espectros obtidos no NIR com equipamento de bancada; e 16,07% dos pares de

amostras com espectros MIR. Para esses pares, as amostras se apresentaram completamente separadas pelos *scores* das PCs, o que indica que elas possuem, de fato, origens distintas.

Desta maneira, conclui-se que o uso desta metodologia foi eficiente na separação de amostras em pares em dados provenientes das análises convencionais normatizadas pela ANP, das curvas de destilação das gasolinas e dos dados espectrais na região do NIR (com o instrumento portátil MicroNIR e com o instrumento de bancada, SPECTRUM 400). A eficiência foi mais reduzida para os dados espectrais obtidos na região do MIR, visto que só foi possível separar por completo apenas 16,07% dos pares, valor baixo ao se comparar com os obtidos para as demais técnicas.

Empregando-se a estratégia proposta, o uso das análises convencionais e das curvas de destilação permitem uma melhor separação das amostras do que quando se utilizam os dados espectrais. No entanto, a instrumentação necessária para essas análises, além de requerer equipamentos específicos e mão-de-obra especializada, não é comumente encontrada nas sedes das polícias científicas. Portanto, o uso de dados de origem espectroscópicas na região do NIR poderiam ser suficientes para aplicação prática por tais instituições, tendo em vista a alta porcentagem de casos em que houve separação completa das amostras em pares, assim como a existência de espectrofotômetros nos locais de trabalho. Apenas nos casos em que não houver separação na PCA, as análises físico-químicas podem ser utilizadas. Além disso, o uso de dados originários de instrumento portátil (MicroNIR) permite a análise de amostras *in situ*, sem a necessidade de deslocamento tanto de equipamentos dos centros de estudos para os locais de crime, assim como transporte das amostras.

## 5.1 PERSPECTIVAS FUTURAS

Para a continuidade deste trabalho e necessidade de abranger o estudo feito para analisar de maneira mais profunda os dados coletados, pretende-se:

- a) Realizar análise supervisionada com SIMCA para todos os conjuntos de dados incluídos neste estudo a fim de aumentar a confiabilidade dos resultados já obtidos e definir limites de classificação mais objetivos;

- b) Realizar análise com fusão de dados, tendo em vista que este trabalho teve foco no estudo individual de cada técnica. Esta comparação por fusão de dados associada a técnicas multivariadas de classificação pode aumentar o nível de separação entre os pares de gasolina, possivelmente chegando a diferenciar pares que não puderam ser separados por algumas das estratégias utilizadas neste trabalho;
- c) Comparar o desempenho de técnicas de classificação multivariada na discriminação de amostras de gasolinas em pares e identificar aquelas com maior capacidade de discriminação das amostras de gasolina;
- d) Aperfeiçoar a metodologia desenvolvida para a geração de espectros virtuais com base na incerteza expandida em cada comprimento de onda, realizando estudos com espectros de outras matrizes diferentes de gasolina;
- e) Procurar melhorar a capacidade de separação observada para os espectros MIR, selecionando regiões espectrais importantes, mas que sejam menos afetadas pelas incertezas.

## REFERÊNCIAS

- ALEME, H. G.; COSTA, L. M.; BARBEIRA, P. J. S. Determination of gasoline origin by distillation curves and multivariate analysis. **Fuel**. v. 87, p. 3664-3668, 2008.
- AL-GHOUTI, M. A.; AL-DEGS, Y. S.; AMER, M. Determination of motor gasoline adulteration using FTIR spectroscopy and multivariate calibration. **Talanta**. v. 76, p. 1105-1112, 2008.
- ALMEIDA, S. Q. **Estudo do efeito da adição de solventes nos parâmetros físico-químicos que caracterizam a qualidade da gasolina automotiva**. 2002. 82 F. Dissertação (Mestrado Profissional em Regulação da Indústria de Energia) – Universidade Salvador, Bahia, 2002.
- AMAT-TOSELLO, S.; DUPUY, N.; KISTER, J. Contribution of external parameter orthogonalisation for calibration transfer in short waves—Near infrared spectroscopy application to gasoline quality. **Anal. Chim. Acta**. v. 642, p. 6-11, 2009.
- ANP. Nota técnica. 2018. Disponível em: < <http://www.anp.gov.br/images/central-de-conteudo/notas-estudos-tecnicos/notas-tecnicas/nota-tecnica-gasolina-formulada-fev2018.pdf> > Acessado em 08/01/2021.
- Associação Brasileira de Normas Técnicas – ABNT NBR 14065 - Destilados de petróleo e óleos viscosos — Determinação da massa específica e da densidade relativa pelo densímetro digital. Rio de Janeiro, 2013.
- Associação Brasileira de Normas Técnicas – ABNT NBR ISO 17025 - Requisitos gerais para a competência de laboratórios de ensaio e calibração. Qualidade que gera confiança. Rio de Janeiro, 2017.
- BALABIN, R. M.; SAFIEVA, R. Z. Gasoline classification by source and type based on near infrared (NIR) spectroscopy data. **Fuel**. v. 87, p. 1096-1101, 2007.
- BALABIN, R. M.; SAFIEVA, R. Z.; LOMAKINA, E. I. Gasoline classification using near infrared (NIR) spectroscopy data: comparison of multivariate techniques. **Analytica Chimica Acta**. v. 671, p. 27-35, 2010.
- BAREA-SEPÚLVEDA, M.; FERREIRO-GONZÁLEZ, M.; CALLE, J. L. P.; BARBERO, G. F.; AYUSO, J.; PALMA, M. Comparison of different processing approaches by SVM and RF on HS-MS eNose and NIR Spectrometry data for the discrimination of gasoline samples. **Microchemical Journal**, 2021.
- BARNETT, I.; KHANG, M. Discrimination of brands of gasoline by using DART-MS and Chemometrics. **Forensic Chemistry**, v. 10, p. 58–66, 2018.
- BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, M. B. **Chemometrics: A Practical Guide**, New York: Wiley-Interscience, 1998.

BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML. **Guide to the expression of uncertainty in measurement**. ISO, 2008.

BIPM, IEC, IFCC, ISO, IUPAC, OIML. **International Vocabulary of Basic and General Terms in Metrology**. ISO, 1993.

BRERETON, R. G. **Chemometrics: Data Analysis for the Laboratory and Chemical Plant**, Chichester: Wiley, 2003.

BUCKMANN, J. H.; SARKIS, J. E. S. O conceito de incerteza aplicado aos processos de medição associados à preparação de uma solução de referência para a calibração. **Química Nova**, v. 25, n. 1, p. 111-116, 2002.

CHASIB, K. F.; KARTHIK, S. An Investigation on the Feasibility of Simulating the Distillation Curves and ASTM Distillation Temperature. **International Journal of Oil, Gas and Coal Engineering**, v. 5, n. 5, p. 80-89, 2017.

CORREIA, R. M.; DOMINGOS, E.; CÁO, V.; ARAUJO, B. R. F.; SENA, S.; PINHEIRO, L. U.; FONTES, A. M.; AQUINO, L. F. M.; FERREIRA, E. C.; FILGUEIRAS, P. R.; ROMÃO, W. Portable near infrared spectroscopy applied to fuel quality control. **Talanta**, v. 176, p. 26-33, 2018.

CORSETTI, S.; ZEHENTBAUER, F. M.; MCGLOIN, D.; KIEFER, J. Characterization of gasoline/ethanol blends by infrared and excess infrared spectroscopy. **Fuel**, v. 141, p. 136-142, 2015.

DE OLIVEIRA, F. S. *et. al.* Screening analysis to detect adulterations in Brazilian gasoline samples using distillation curves. **Fuel**. v. 83, p. 917-923, 2004.

DOBLE, P.; SANDERCOCK, M.; PASQUIER, E.; PETCOZ, P.; ROUX, C.; DAWSON, M. Classification of premium and regular gasoline by gas chromatography/mass spectrometry, principal component analysis and artificial neural networks. **Forensic science international**. v. 132, p. 26-39, 2003.

DO BRASIL, N. I.; ARAÚJO, M. A. S.; SOUSA, E. C. M. **Processamento de Petróleo e Gás: petróleo e seus derivados, processamento primário, processos de refino, petroquímica, meio ambiente**. Rio de Janeiro: LTC, 2012.

ESCALANTE, V. E. J. **Seis-Sigma: Metodología y Técnicas**. Limusa. México, 2010.

EURACHEM/CITAC. **Guide. Quantifying Uncertainty in analytical Measurement**, 2 edition, 2002.

FERNANDES, D. D. S. **Novas estratégias para seleção de variáveis por intervalos em problemas de classificação**. 2016. 139 F. Tese (Doutorado em Química) – Universidade Federal da Paraíba, 2016.

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

FERREIRO-GONZÁLEZ, M.; AYUSO, J.; ÁLVAREZ, J. A.; PLAMA, M.; BARROSO, A. G. Gasoline analysis by headspace mass spectrometry and near infrared spectroscopy. **Fuel**. v. 153, p. 402-407, 2015.

FLUMIGNAN, D. L. *et. al.* Screening Brazilian C gasoline quality: application of the SIMCA chemometric method to gas chromatographic data. **Analytica Chimica Acta**. v. 595, p. 128-135, 2007.

FLUMIGNAN, D. L.; BORALLE, N.; DE OLIVEIRA, J. E. Screening Brazilian commercial gasoline quality by hydrogen nuclear magnetic resonance spectroscopic fingerprintings and pattern-recognition multivariate chemometric analysis. **Talanta**. v. 82, p. 99-105, 2010.

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, p. 53-60, 2008.

GEMPERLINE, P. **Practical Guide to Chemometrics**. Boca Raton, FL: CRC Press: Taylor & Francis Group, 2006.

GULLBERG, R. G. Estimating the measurement uncertainty in forensic blood alcohol analysis. **Journal of Analytical Toxicology**, v. 36, p. 153-161, 2012.

GUTIÉRREZ, P. H.; DE LA VARA, S. **Control Estadístico de la Calidad y Seis Sigma**. Ed. 2 Mc Graw Hill. México, 2009.

INMETRO, Instituto Nacional de Metrologia, Normalização e Qualidade Industrial, VIM; **Vocabulário Internacional de Metrologia: Conceitos Fundamentais e Gerais e Termos Associados** (VIM 2012), 1a ed. Luso-brasileira, Rio de Janeiro, 2012. Acesso em 18/01/2021. Disponível em: < [http://www1.ipq.pt/PT/Metrologia/Documents/VIM\\_IPQ\\_INMETRO\\_2012.pdf](http://www1.ipq.pt/PT/Metrologia/Documents/VIM_IPQ_INMETRO_2012.pdf) >

INSTITUTO NACIONAL DE METROLOGIA, NORMALIZAÇÃO E QUALIDADE INDUSTRIAL, VIM; **Vocabulário Internacional de Metrologia: Conceitos Fundamentais e Gerais e Termos Associados (VIM 2008)**, 1a ed., Rio de Janeiro, 2009.

KHANMOHAMMADI, M.; GARMARUDI, A. B.; GHASEMI, K.; GUARDIA, M. Quality based classification of gasoline samples by ATR-FTIR spectrometry using spectral feature selection with quadratic discriminant analysis. **Fuel**. v. 111, p. 96-102, 2013.

KOWALSKI, B. R. Chemometrics: Views and Propositions. **Journal of Chemical Information and Computer Sciences**. v. 15, n.º.4, p. 201-203, 1975.

LIMA, R. *et. al.* Multivariate analysis of the physico-chemical parameters of "type c" gasoline marketed in the vale do Juruá -Acre. **South American Journal of Basic Education, Technical and Technological**. v. 6, p. 27-37, 2019.

- LI, S.; DAI, L. Classification of gasoline brand and origin by raman spectroscopy and a novel re-weighted LSSVM algorithm. **Fuel**. v. 96, p. 146-152, 2012.
- MALDONADO, M.; BARREIRO, P.; GUTIÉRREZ, R.; VERGARA, G. Mid- infrared uncooled sensor for the identification of pure fuel, aditives and adulterants in gasoline. **Fuel Processing Technology**. v. 171, p. 287-292, 2018.
- MABOOD, F. et al.; Detection and estimation of Super premium 95 gasoline adulteration with Premium 91 gasoline using new NIR spectroscopy combined with multivariate methods. **Fuel**, v. 197, p.388-396, 2017.
- MARINI, F. **Chemometrics in Food Chemistry**. 1st. ed. [s.l.] ELSEVIER, 2013. v. 53.
- MAROTO, A.; BOQUÉ, R.; RIU, J.; RIUS, F. X. Evaluating uncertainty in routine analysis. Trends in Analytical Chemistry. V. 18, issues 9-10, p. 577-584, 1999.
- OLIVEIRA, E. C.; AGUIAR, P. F. Comparação de diferentes abordagens para avaliação da incerteza na cromatografia gasosa do gás natural. **Quím. Nova**, v. 32, n. 6, p. 1655-1660, 2009.
- OLIVEIRA, F. C. C.; SOUZA, A. T. P. C.; DIAS, J. A.; DIAS, S. C. L.; RUBIM, J. C. A escolha da faixa espectral no uso combinado de métodos espectroscópicos e quimiométricos. **Quím. Nova**, v. 27, n. 2, p. 218-225, 2004.
- OTTO, M. **Chemometrics: Statistics and Computer Application in Analytical Chemistry**. 3rd. ed. Leipziger, Germany: Wiley-VCH Verlag GmbH & Co, 2017.
- PAC. OptiDist – Benchmark in Atmospheric Distillation Testing. Manual de Equipamento. Houston. Texas, USA. 2015.
- 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.
- PASQUINI, C. Near infrared spectroscopy: A mature analytical technique with new perspectives - A review. **Analytica Chimica Acta**. v. 1026, p. 8-36, 2018.
- PERRY. **Perry's Chemical Engineers Handbook**. McGraw-Hill Companies, Inc, Seventh Edition, 1999.
- PETRACO, N. D. K.; GIL, M.; PIZZOLA, P. A.; KUBIC, T. A. Statistical Discrimination of Liquid Gasoline Samples from Casework. **Journal of Forensic Sciences**. v. 53, n. 5, p. 1092-1101, 2008.
- PETROBRAS. **Gasolina: Informações técnicas**. PETROBRAS, julho de 2020. Disponível em: <[http://sites.petrobras.com.br/minisite/assistenciatecnica/public/downloads/Manual de Gasolina.pdf](http://sites.petrobras.com.br/minisite/assistenciatecnica/public/downloads/Manual_de_Gasolina.pdf)>. Acessado em 05/01/2021.

PORTARIA MAPA Nº 75, de 05.03.2015. **MAPA – Ministério da Agricultura, Pecuária e Abastecimento**, 06 de março 2015.

RESOLUÇÃO ANP Nº 40, de 28.10.2013 - Dispõe sobre as especificações das gasolinas de uso automotivo. **ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis**, 25 de out. de 2013.

RODRIGUES, L. L. V. **Avaliação do desempenho de motores a combustão interna com ignição por faísca utilizando misturas de gasolina e etanol**. Trabalho de Conclusão de Curso de Engenharia Mecânica, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, 2017.

RUDNEV, V. A.; BOICHENKO, A. P.; KARNOZHITSKIY, P. V. Classification of gasoline by octane number and light gas condensate fractions by origin with using dielectric or gas-chromatographic data and chemometrics tools. **Talanta**. v. 84, p. 963-970, 2011.

SANTOS, F. D. et al. Discrimination of oils and fuels using a portable NIR spectrometer. **Fuel**, v. 283, p. 1-10, 2021.

SILVA, N., C. et al. Classification of Brazilian and foreign gasolines adulterated with alcohol using infrared spectroscopy. **Forensic Science International**, v. 253, p. 33 - 42, 2015.

SILVA, N., C. **Detecção e quantificação de sibutramina em medicamentos naturais e classificação de gasolinas brasileiras e venezuelanas utilizando a espectroscopia no infravermelho**. 2013. 99 F. Dissertação (Mestrado em Química) – Universidade Federal de Pernambuco, 2013.

SILVA, N., C. **Uso da espectroscopia NIR no desenvolvimento de um simulador para gasolina e na transferência de calibração entre instrumentos de bancadas e portátil**. 2017. 124 F. Tese (Doutorado em Engenharia Química) – Universidade Federal de Pernambuco, 2017.

SKOOG, D. A.; HOLLER, F. J.; CROUCH, S. R. **Principles of Instrumental Analysis**. 7 ed., Boston: Cengage Learning, 2017.

SKOOG, D. A.; WEST, D. M.; HOLLER, F. J.; CROUCH, S. R. **Fundamentos de Química Analítica**. Tradução da 8ª edição norte-americana. São Paulo: Thomson Learning, 2007.

SMITH, J.M.; VAN NESS, H.C.; ABBOTT, M.M. **Introdução à Termodinâmica da Engenharia Química**. 7. Ed. Rio de Janeiro: Editora LTC, 2007.

SPEIGHT, J. G. **The Chemistry and Technology of Petroleum**. New York: Marcer Derker, Inc, Thrid Edition, Revised and Expanded, 1999.

TAKESHITA, E. V. **Adulteração de gasolina por adição de solventes: análise dos parâmetros físico-químicos**. Dissertação de mestrado em Engenharia Química, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, 2006.

TOSATO, F. *et. al.* Análise de amostras de combustíveis por fotometria, NIR portátil e RMN de <sup>1</sup>H – uma comparação com os resultados encontrados por técnicas normatizadas. **Química Nova**, v. 43, n. 2, p. 155-167, 2020.

VEMPAPATU, B. P; KANAUIA, P. K. Monitoring petroleum fuel adulteration: A review of analytical methods. **Trends in Analytical Chemistry**, v. 92, p. 1-11, 2017.

VIDAL, O. **A Gasolina, sua Composição e seus Efeitos – Relato de uma Experiência no Ensino Médio. 1999.** 138p. Monografia Especialização em Ensino de Química Experimental para o Segundo Grau – Setor de Ciências Exatas, Universidade Federal do Paraná, Paraná, Curitiba, 1999.

WERNIMONT, G. T.; **Use of Statistics to Develop and Evaluate Analytical Methods**, AOAC: Arlington, 1985.

WORKMAN, J.; WEYER, L. **Practical Guide to Interpretive Near-Infrared Spectroscopy.** [s.l.] CRC Press: Taylor & Francis Group, 2009.

ZHAO, J.; ZHENG, Y.; CHEN, M.; YANG, M.; MA, Q.; ZHANG, Z. Effect of sampling modes on the discrimination of commercial gasolines using direct analysis in real time mass spectrometry. **International Journal of Mass Spectrometry**, v. 463, 2021.

## APÊNDICE A – ROTINA USADA NO MATLAB PARA A GERAÇÃO DE AMOSTRAS VIRTUAIS

Após obtidas as incertezas em cada comprimento de onda e o intervalo das incertezas (o qual é igual a duas vezes o valor da incerteza), assim como realizado pré-processamento dos dados a serem estudados com SNV e deslocamento ao longo do eixo y, foi utilizada a seguinte rotina para obtenção dos espectros virtuais:

```

%% SEÇÃO 01: OBTENÇÃO DA MATRIZ COM AS COMBINAÇÕES DAS AMOSTRAS EM PARES

n = size(DADOS_SEMOUT,1); % é o número de amostras sem os outliers

% Obter a matriz contendo todos os possíveis pares combinados de amostras.
% Essa matriz é facilmente obtida do algoritmo de Kennard-Stone, a matriz D.

[m,dminmax,D] = ks(DADOS_SEMOUT,n);

for i = 1:size(D,1);
    for j = 1:size(D,2);
        if D(i,j) == 0;
            D(i,j) = NaN;
        end
    end
end

for i = 1:factorial(size(D,1))/[factorial(size(D,1)-2)*2];
    [value, index] = min(D(:));
    [row, col] = ind2sub(size(D), index);
    ord(i,1) = row;
    ord(i,2) = col;
    ord(i,3) = value;
    D(row,col) = NaN;
end

clear m D dminmax i j value index row col

%% SEÇÃO 02: GERAÇÃO DAS AMOSTRAS VIRTUAIS

% Dividindo em intervalos

in = 150; % Definir aqui o intervalo

nv = size(NUM_ONDAS,2) % Aqui é obtido o número de variáveis (número de onda
ou comprimento de onda), sendo "NUM_ONDAS" a matriz que armazena estas variáveis

```

```

S = DADOS_SEMOUT(:, 1:in:nv);

% Gerar as amostras virtuais

IR_NOMES_TOTAL = []; % matriz que armazenará todos os nomes das amostras do IR
IR_REGIAO_TOTAL = []; % matriz que armazenará a região de cada amostra do IR
Rdn_total = []; % Matriz que armazenará os valores randômicos gerados entre 0
e 1

tml = 49% Determinar aqui quantidade de amostras simuladas

tm = tml +1; % Número de amostras em cada grupo (uma amostra real + tml virtu-
ais)

var = ((nv-1)/(in))+1; % aqui é o número de variáveis, ou seja, as absorbân-
cias. Portanto, é o número de colunas da matriz

for i=1:n

    Rdn = rand(tm,var);
    Rdn_total = [Rdn_total; Rdn];

    nomes = repmat((IR_NOMES_SEMOUT(i,:)),tm,1);
    R = repmat((IR_REGIAO_SEMOUT(i,:)),tm,1);

    IR_NOMES_TOTAL = [IR_NOMES_TOTAL;nomes]; %matriz com todos os nomes das
amostras
    IR_REGIAO_TOTAL = [IR_REGIAO_TOTAL;R];

end

% Preenchendo a matriz com o linspace

nt = n*tm; % número de amostras totais

Rdn_linspace = zeros(nt, nv); % aqui cria uma matriz de zeros que será preen-
chida com as incertezas, ela contém linhas para todas as amostras e todas as
colunas de variáveis
janela = (1:in:nv);

for i=1:nt
    for j=1:(var-1)
        l = linspace(Rdn_total(i,j), Rdn_total(i,j+1), (in+1));
        Rdn_linspace(i,janela(j):janela(j)+(in)) = l;
    end
end
end

```

```
% GERAR MATRIZ COM CÓPIAS

IR_repmat = []; % essa matriz conterà os espectros originais replicados

for i =1:size(IR_DADOS_SEMOUT,1)

    rep = repmat (IR_DADOS_SEMOUT(i,:),50, 1);
    IR_repmat = [IR_repmat; rep];

end

% CRIAR MATRIZ COM AS INCERTEZAS A SEREM ADICIONADAS

incertezas_aleatorias = repmat(int,nt,1) .* Rdn_linspace;

% SUBTRAIR AS INCERTEZAS

IR_repmat_sub = IR_repmat - repmat(inc, nt,1);

% SOMAR COM AS INCERTEZAS ALEATÓRIAS

IR_repmat_soma = IR_repmat_sub + incertezas_aleatorias;

% SUBSTITUIR OS ESPECTROS NA POSIÇÃO DOS ORIGINAIS PELOS ESPECTROS ORIGINAIS

IR_DADOS_FINAIS = IR_repmat_soma;

for i=1:50:nt
    IR_DADOS_FINAIS(i,:) = IR_repmat(i,:);
end
```

**APÊNDICE B – ESPECTROS VIRTUAIS DE TESTES COM DIFERENTES VALORES PARA INTERVALOS DE VARIAÇÃO LINEAR DO COEFICIENTE RANDÔMICO DA INCERTEZA DOS ESPECTROS COM EQUIPAMENTO PORTÁTIL**

Figura B1 - Espectro virtual na região NIR (equipamento portátil) obtido com intervalos de 2 pontos.

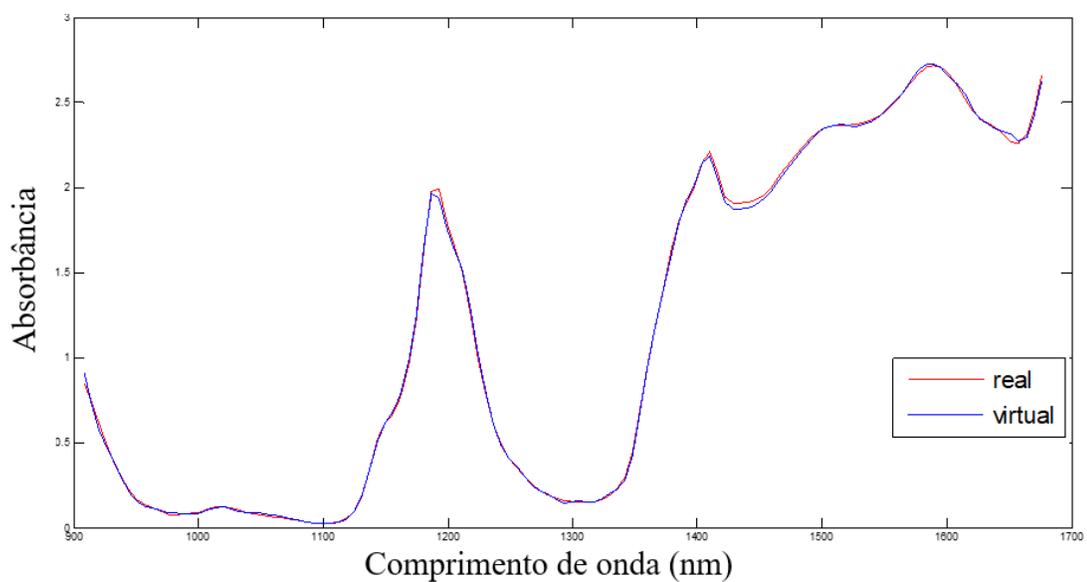
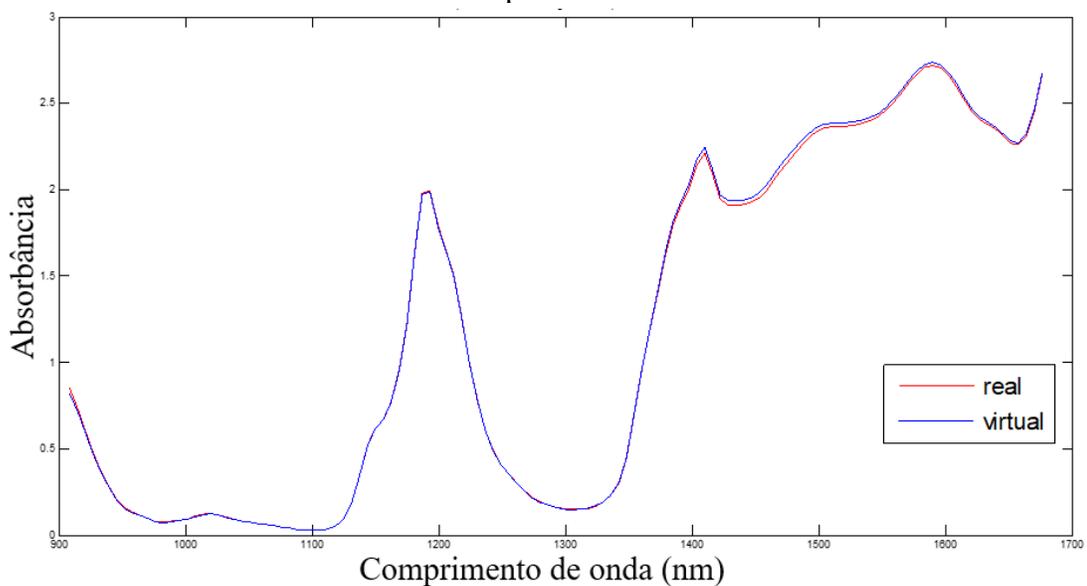


Figura B2 - Espectro virtual na região NIR (equipamento portátil) obtido com intervalos de 31 pontos.



**APÊNDICE C – ESPECTROS VIRTUAIS DE TESTES COM DIFERENTES VALORES PARA INTERVALOS DE VARIAÇÃO LINEAR DO COEFICIENTE RANDÔMICO DA INCERTEZA DOS ESPECTROS NIR COM EQUIPAMENTO DE BANCADA**

Figura C1 - Espectro virtual na região NIR (equipamento de bancada) obtido com intervalos de 50 pontos.

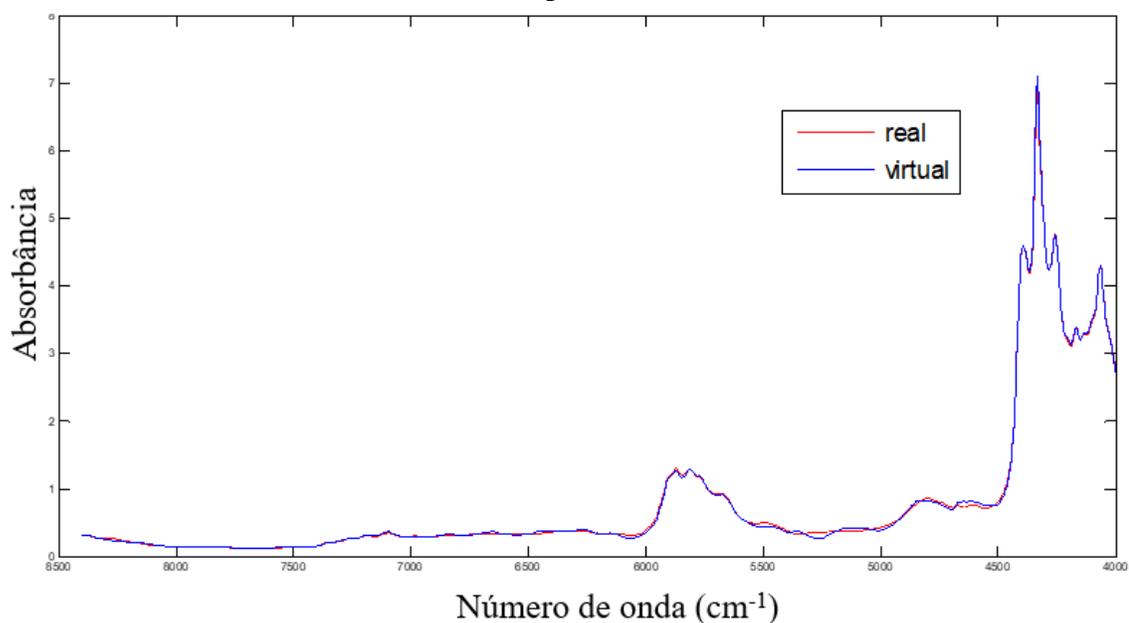


Figura C2 - Espectro virtual na região NIR (equipamento de bancada) obtido com intervalos de 75 pontos.

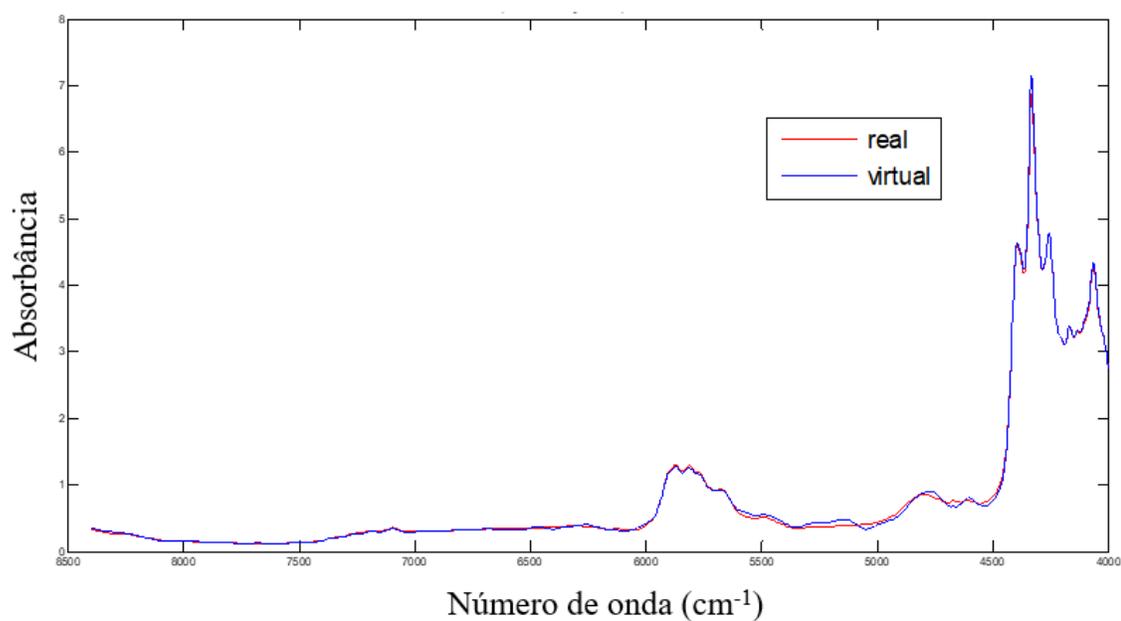


Figura C3 - Espectro virtual na região NIR (equipamento de bancada) obtido com intervalos de 100 pontos.

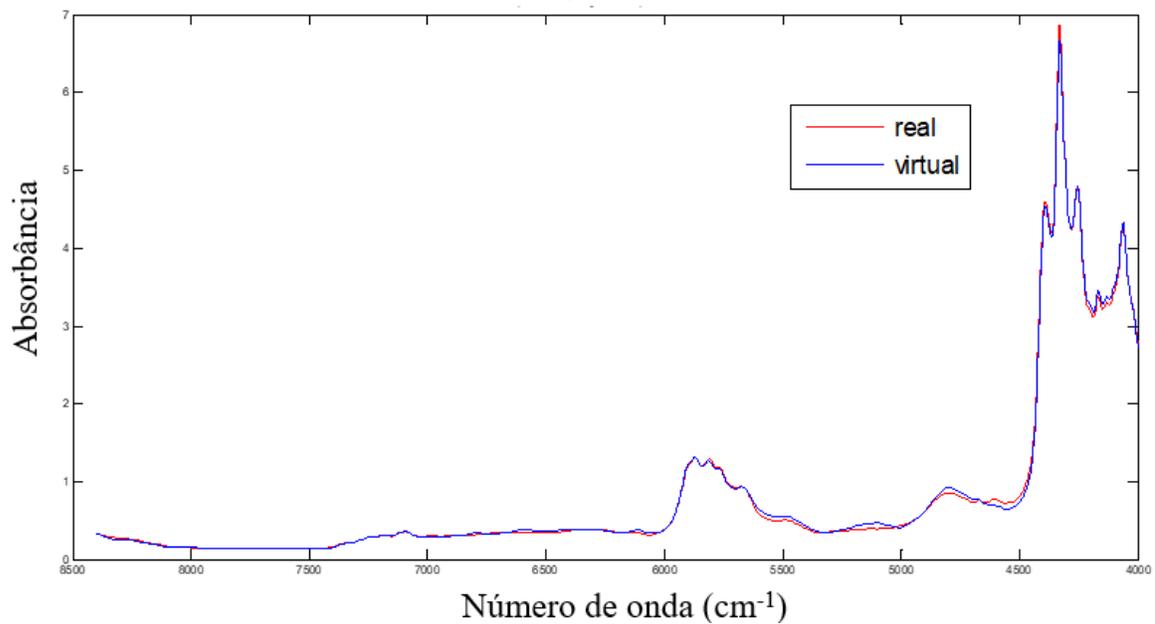
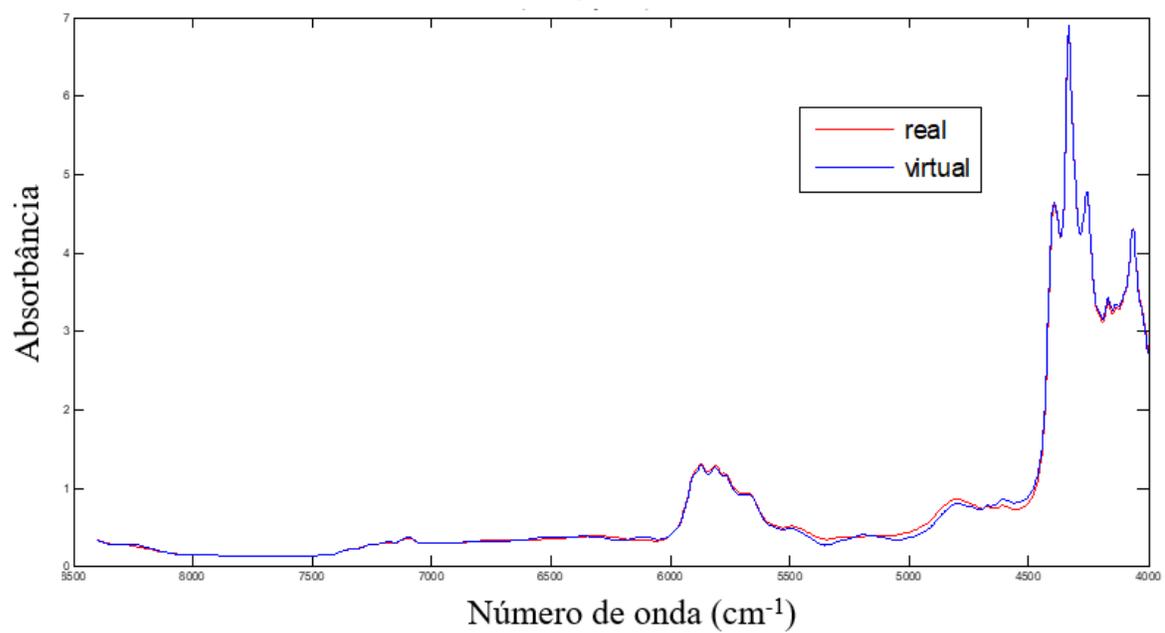


Figura C4 - Espectro virtual na região NIR (equipamento de bancada) obtido com intervalos de 150 pontos.



**APÊNDICE D – ESPECTROS VIRTUAIS DE TESTES COM DIFERENTES VALORES PARA INTERVALOS DE VARIAÇÃO LINEAR DO COEFICIENTE RANDÔMICO DA INCERTEZA DOS ESPECTROS MIR**

Figura D1 - Espectro virtual na região MIR obtido com intervalos de 5 pontos.

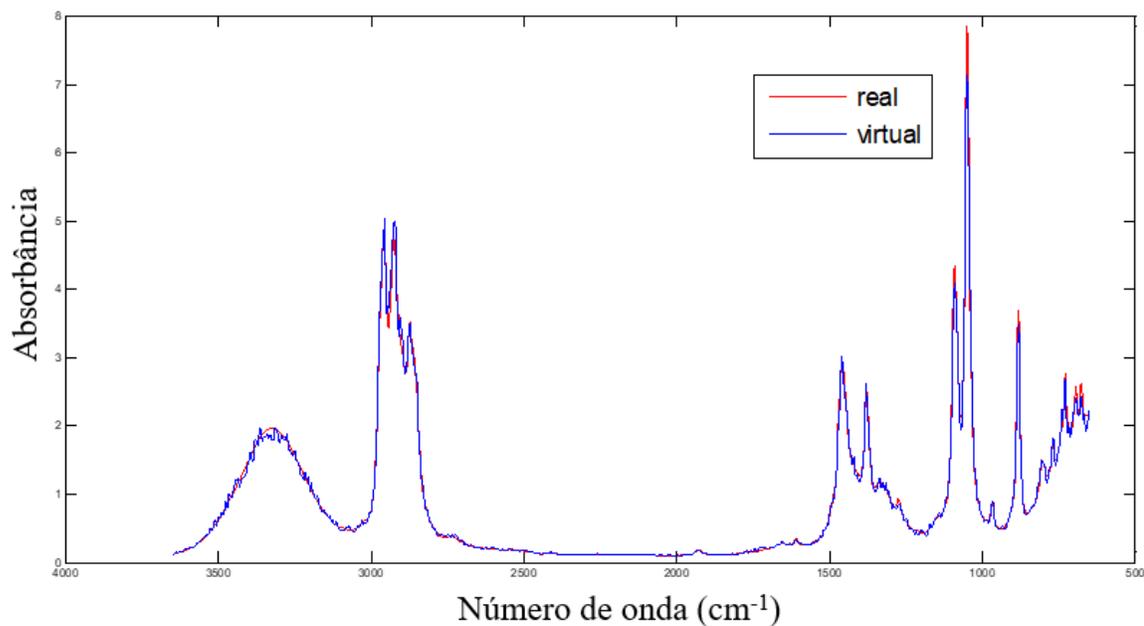


Figura D2 - Espectro virtual na região MIR obtido com intervalos de 25 pontos.

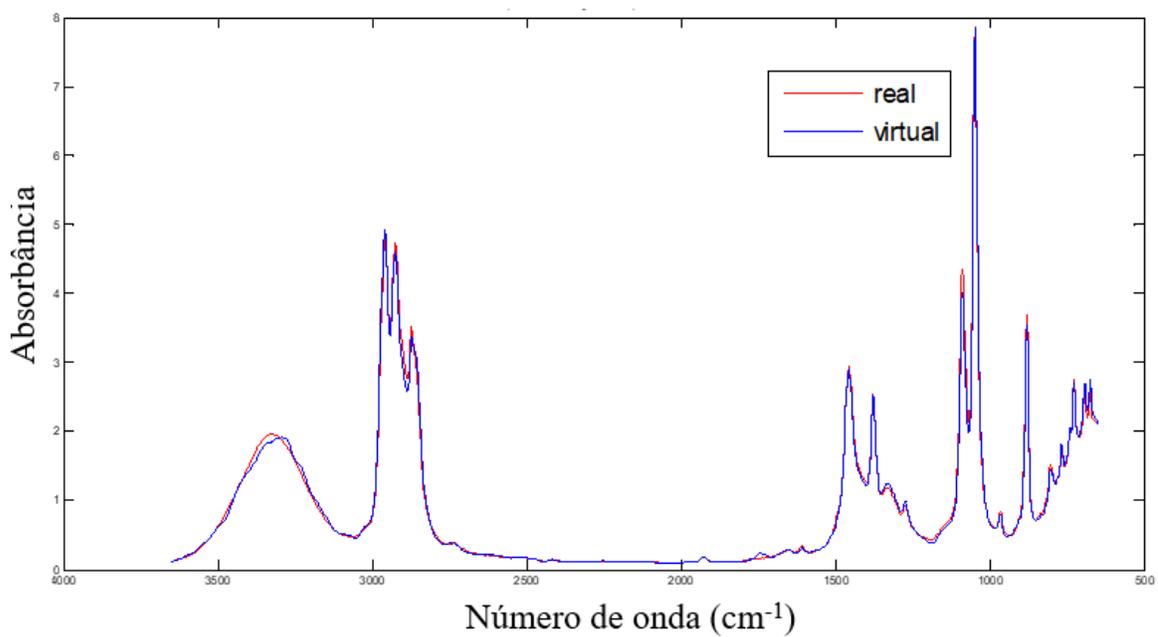


Figura D3 - Espectro virtual na região MIR obtido com intervalos de 50 pontos.

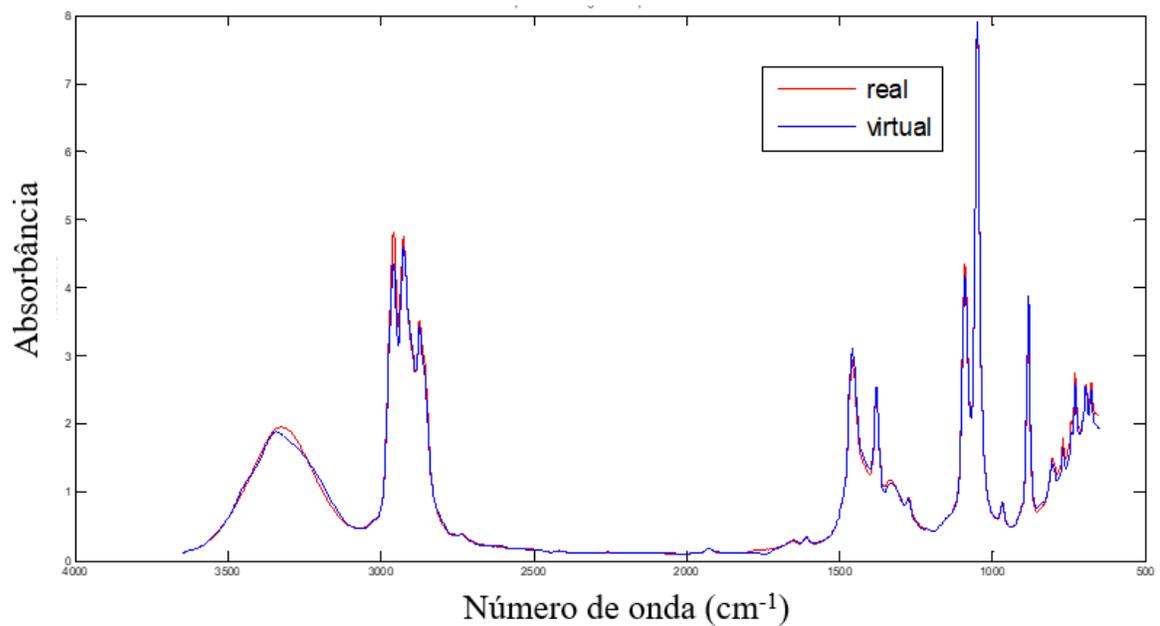


Figura D4 - Espectro virtual na região MIR obtido com intervalos de 67 pontos.

