

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE QUÍMICA

JAMILA BITTENCOURT DE SALES

**ESTUDO DA VIABILIDADE DE IMPLEMENTAÇÃO DE METODOLOGIA PARA
ANÁLISE DE PARÂMETROS DA GASOLINA EMPREGANDO ESPECTROSCOPIA
DE INFRAVERMELHO PRÓXIMO E QUIMIOMETRIA**

Porto Alegre, Janeiro de 2018.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE QUÍMICA

JAMILA BITTENCOURT DE SALES

**ESTUDO DA VIABILIDADE DE IMPLEMENTAÇÃO DE METODOLOGIA PARA
ANÁLISE DE PARÂMETROS DA GASOLINA EMPREGANDO ESPECTROSCOPIA
DE INFRAVERMELHO PRÓXIMO E QUIMIOMETRIA**

Trabalho de conclusão apresentado junto à atividade de ensino “Projeto Tecnológico” do Curso de Química Industrial, como requisito parcial para a obtenção do grau de Químico Industrial pela Universidade Federal do Rio Grande do Sul.

Prof. Dr. Marco Flôres Ferrão
Orientador

Porto Alegre, Janeiro de 2018.

AGRADECIMENTOS

Ao meu professor orientador, Marco Flôres Ferrão, pela paciência, pelos ensinamentos, pela disponibilidade.

À colega Camila Veríssimo, por me transmitir grande parte do seu conhecimento, pelos conselhos e amizade.

A todos os colegas de trabalho que de alguma forma me ajudaram, compartilhando conhecimentos técnicos, emprestando materiais, ou com palavras de conforto e tranquilidade.

Aos amigos que percorreram todo esse caminho na UFRGS junto comigo, ou perto de mim, tornando a trajetória menos difícil.

Ao meu amor, Egon, pelo apoio, companheirismo, cuidado, e por ser presente em todos os momentos.

Aos meus “irmãos”, Giuliano e Karen, por se preocuparem e entenderem as minhas ausências.

Aos meus pais, Angela e Paulo, pelo amor incondicional, pelo suporte, pelas frases de apoio ou consolo, por acreditarem em mim e no meu sucesso, por serem incentivadores de pessoas, por valorizarem a educação.

A Deus, por me proporcionar conhecer, conviver e aprender com todas essas pessoas boas que estão ou estiveram no meu caminho.

“A vida é a arte do encontro”.

(Vinícius de Moraes)

RESUMO

A gasolina é um combustível de composição complexa, constituído por hidrocarbonetos aromáticos, olefínicos e saturados e alguns compostos contendo enxofre e metais. A estrutura química dos hidrocarbonetos da gasolina tem grande influência na sua qualidade, que é assegurada através de diversas análises especificadas pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP). Um dos parâmetros mais importantes no controle de qualidade da gasolina é o número de octanas, que é muito relacionado com a composição, e indica a resistência do combustível à auto-ignição. Com o intuito de garantir o atendimento às especificações de forma mais rápida e precisa, muitos métodos alternativos têm sido estudados, em substituição às análises clássicas, que são caras e trabalhosas. Neste contexto, buscou-se, através deste trabalho, que foi realizado em empresa do setor petroquímico, implementar metodologia para determinação do número de octano motor (MON) e teor de aromáticos e olefinas da gasolina, utilizando espectroscopia no infravermelho próximo e quimiometria. Para tanto, diversos espectros de gasolinas foram coletados, pré-tratados pelo uso de primeira derivada e submetidos à calibração multivariada utilizando o algoritmo de regressão por quadrados mínimos parciais (PLS). O modelo criado para o MON obteve raiz da média quadrática dos resíduos de previsão (RMSEP) igual a 0,2951 e boa correlação com as técnicas tradicionais, principalmente para a faixa de MON de 82,0 e 83,6, devendo ser completado para possível utilização em outras faixas. A regressão para predição de aromáticos apresentou RMSEP igual a 1,0978% e coeficiente de correlação da curva de predição de 0,6030, logo, a metodologia proposta precisa ser reavaliada antes de ser utilizada. Já o modelo para determinação de olefinas, com $R^2=0,9529$ e RMSEP igual a 0,5744%, demonstrou alta correlação com a metodologia de referência, podendo ser uma alternativa para análise de gasolinas.

Palavras-chave: Gasolina. MON. Octanagem. Espectroscopia no Infravermelho Próximo. Quimiometria. Calibração multivariada. Análise por Componentes Principais (PCA). Regressão por Quadrados Mínimos Parciais (PLS).

ABSTRACT

Gasoline is a fuel of complex composition formed by aromatic, olefinic and saturated hydrocarbons and some compounds containing sulfur and metals. The chemical structure of the gasoline's hydrocarbons has a huge influence in her quality, which is ensured through various analyses specified by National Agency of Petroleum, Natural Gas and Biofuels (ANP). One of the most important parameter to control the gasoline's quality is the octane number, which is so related to the composition and indicates the resistance to the fuel self-ignition. With the purpose of guaranteeing the specifications more quickly and accurately, many alternative methods have been studied, in substitution of the classic analyses, which are expensive and laborious. In this context, it was sought through this work, which was carried out in a petrochemical company, implement methodology to determine Motor Octane Number (MON) and aromatic and olefins content of gasoline by Near Infrared (NIR) Spectroscopy and chemometrics. For this purpose several spectra were acquired, pre-treated by first derivative and subjected to multivariate calibration using Partial Least Squares (PLS) algorithm. The model developed for MON got Root Mean Square Error of Prediction (RMSEP) value of 0.2951 and a good correlation with traditional techniques, especially between MON 82.0 and 83.6, should be completed to be used in another range. The regression for aromatic's prediction presented RMSEP as 1.0978% and correlation coefficient of the prediction curve of 0.6030, therefore the proposed methodology needs to be re-evaluated before being used. The model for olefins determination, however, with $R^2 = 0.9529$ and RMSEP as 0.5744%, demonstrated high correlation with the reference methodology, and can be implemented to gasoline analyses.

Key-words: Gasoline. MON. Octane. Near Infrared Spectroscopy. Chemometrics. Multivariate Calibration. Principal Components Analysis (PCA). Partial Least Squares Regression (PLS).

LISTA DE FIGURAS

FIGURA 1: Ciclo de trabalho de um motor ciclo Otto	12
FIGURA 2: Motor CFR utilizado para análise de octanagem da gasolina	13
FIGURA 3: Tipos de vibrações moleculares.....	15
FIGURA 4: Bandas de absorção no infravermelho próximo	16
FIGURA 5: Equipamento FT-NIR utilizado para coleta dos espectros	27
FIGURA 6: Espectros de infravermelho próximo das amostras de gasolina, obtidos na região de 11100 a 3980 cm^{-1}	29
FIGURA 7: Gráfico de <i>scores</i> da PC1 versus a PC2, de todas as amostras.....	30
FIGURA 8: Gráfico de <i>loadings</i> da PC1.....	31
FIGURA 9: Gráfico de <i>scores</i> da PC1 versus a PC2, de todas as amostras, considerando apenas a faixa espectral que contém as variáveis mais importantes.....	32
FIGURA 10: Gráfico de <i>scores</i> da PC1 versus PC2, com todas as amostras, considerando apenas as variáveis mais importantes, e com os dados espectrais tratados pela primeira derivada	34
FIGURA 11: Gráfico dos valores de MON de referência versus valores preditos pelo modelo para amostras de calibração e predição	35
FIGURA 12: Gráfico comparativo dos valores de MON de cada amostra do conjunto de predição obtidos pelo método de referência e pelo modelo	36
FIGURA 13: Gráfico dos teores de aromáticos pelo método de referência versus valores preditos pelo modelo para amostras de calibração e predição.....	38
FIGURA 14: Gráfico comparativo dos valores de aromáticos de cada amostra do conjunto de predição obtidos pelo método de referência e pelo modelo	39
FIGURA 15: Gráfico dos teores de olefinas pelo método de referência versus valores preditos pelo modelo para amostras de calibração e predição	41
FIGURA 16: Gráfico comparativo dos valores de olefinas de cada amostra do conjunto de predição obtidos pelo método de referência e pelo modelo	42

LISTA DE TABELAS

TABELA 1: Regiões espectrais do infravermelho	14
TABELA 2: Figuras de mérito obtidas por PLS após diversos pré-tratamentos dos dados espectrais para MON	33
TABELA 3: Figuras de mérito obtidas por PLS após diversos pré-tratamentos dos dados para teor de aromáticos.....	37
TABELA 4: Figuras de mérito obtidas por PLS após diversos pré-tratamentos dos dados para teor de olefinas.....	40
TABELA 5: Comparativo entre custos e benefícios das metodologias de referência e da técnica proposta	44
TABELA 6: Comparativo dos principais impactos das metodologias de referência e da técnica proposta	45

LISTA DE ABREVIATURAS E SIGLAS

ANP	Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
ASTM	do inglês <i>American Society for Testing and Materials</i>
CFR	do inglês <i>Cooperative Fuel Research</i>
DHA	Análise Detalhada de Hidrocarbonetos, do inglês <i>Detailed Analysis of Hydrocarbons</i>
FAR	Infravermelho Distante, do inglês <i>Far Infrared</i>
FCC	Craqueamento Catalítico Fluidizado, do inglês <i>Fluid Catalytic Cracking</i>
FT-NIR	Infravermelho Próximo com Transformada de Fourier, do inglês <i>Fourier Transform Near Infrared</i>
IAD	Índice Antidetonante
MID	Infravermelho Médio, do inglês <i>Mid Infrared</i>
MON	Número de Octano Motor, do inglês <i>Motor Octane Number</i>
MSC	Correção de espalhamento multiplicativo, do inglês <i>Multiplicative Signal Correction</i>
NIR	Infravermelho Próximo, do inglês <i>Near Infrared</i>
PC	Componente Principal, do inglês <i>Principal Component</i>
PCA	Análise por Componentes Principais, do inglês <i>Principal Component Analysis</i>
PLS	Mínimos Quadrados Parciais, do inglês <i>Partial Least Squares</i>
RON	Número de Octano Research, do inglês <i>Research Octane Number</i>
VL	Variáveis Latentes

SUMÁRIO

1 INTRODUÇÃO	9
1.1 GASOLINA E SUA COMPOSIÇÃO	9
1.2 TIPOS DE GASOLINA	10
1.3 MOTOR CICLO OTTO E OCTANAGEM	11
1.4 ESPECTROSCOPIA DE INFRAVERMELHO	14
1.5 INFRAVERMELHO PRÓXIMO – NIR	16
1.6 TÉCNICAS QUIMIOMÉTRICAS DE ANÁLISE EXPLORATÓRIA	17
1.6.1 Organização dos dados	17
1.6.2 Pré-tratamento dos dados	18
1.6.3 Análise por Componentes Principais – PCA	19
1.6.4 Calibração Multivariada	20
1.6.5 Método dos Quadrados Mínimos Parciais - PLS	22
2 SITUAÇÃO ATUAL	23
3 OBJETIVOS	25
4 PROPOSTA TECNOLÓGICA	26
5 METODOLOGIA	27
6 RESULTADOS E DISCUSSÕES	29
6.1 AVALIAÇÃO DOS ESPECTROS	29
6.2 MON	32
6.3 TEOR DE AROMÁTICOS	36
6.4 TEOR DE OLEFINAS	39
7 AVALIAÇÃO DE CUSTOS	43
8 IMPACTO AMBIENTAL	45
9 CONCLUSÕES	46
REFERÊNCIAS	48

1 INTRODUÇÃO

1.1 GASOLINA E SUA COMPOSIÇÃO

A gasolina é constituída basicamente por diferentes hidrocarbonetos, com diversas estruturas moleculares: alifáticas, naftênicas, aromáticas, saturadas, insaturadas, normais e ramificadas. Não é relevante em sua aplicação cotidiana a proporção de cada uma dessas estruturas¹, no entanto, a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) limita a, no máximo, 25% os compostos olefínicos e 35% os compostos aromáticos, em volume, presentes na gasolina². Ela pode possuir mais de 350 tipos de cadeias diferentes, de 4 a 12 carbonos, e faixa de destilação de 30 a 220 °C³.

As características das cadeias carbônicas que constituirão a gasolina variam conforme o petróleo processado, os processos de refino que a geraram e as diferentes correntes petroquímicas usadas na sua elaboração. A gasolina é composta por uma mistura de naftas oriundas de diferentes etapas nas refinarias: da Destilação Direta, do Craqueamento Catalítico Fluidizado (FCC), da Reforma Catalítica, do Coqueamento Retardado, do Craqueamento Catalítico e da Alquilação^{1,3}, sendo que a maior fonte de gasolina em todas as refinarias é a corrente de nafta produzida por FCC, que possui alto número de octano em função do alto teor de hidrocarbonetos aromáticos presentes. As naftas de destilação direta têm usualmente baixa octanagem, pois são muito parafínicas¹.

A ANP determina que se utilize a norma da *American Society for Testing and Materials* ASTM D1319⁴ para a determinação volumétrica dos grupos de hidrocarbonetos presentes na gasolina, mas também abre a possibilidade de se fazer a análise por cromatografia gasosa, sem definir norma específica².

A ASTM D1319 é uma norma que emprega a técnica de cromatografia líquida, utilizando uma coluna de vidro empacotada com sílica-gel onde os grupos de componentes da gasolina são separados e marcados por indicadores de adsorção por fluorescência, que tornam-se visíveis pela aplicação de uma luz ultravioleta. É um método barato e simples, porém com baixa reprodutibilidade e precisão. Além disso, as informações sobre a composição são bem limitadas⁵. Em substituição, a cromatografia gasosa com detector de ionização de chama tem sido muito utilizada pela indústria petroquímica, pois ela separa, através de uma coluna capilar de alta resolução, os componentes da gasolina quase que individualmente, fornecendo uma Análise Detalhada de Hidrocarbonetos, DHA, como a técnica também é chamada. Existem várias metodologias para esta análise, mas em todas elas

se usa uma coluna longa (50 ou 100 metros de comprimento), e uma rampa de aquecimento no forno do equipamento, de forma que ocorra a separação dos componentes. Além de ser uma análise mais complexa, ela usa hélio como gás de arraste, o que a torna mais cara, e pode levar até 180 minutos para gerar um resultado⁶.

Algumas gasolinas podem receber adição de compostos oxigenados, como álcoois e éteres, para aumentar sua octanagem. No Brasil, devido às características agrícolas do país, com grandes áreas de plantação de cana-de-açúcar, o governo brasileiro, desde a década de 1930, regulamenta, por lei federal, a adição de etanol derivado da cana à gasolina, em porcentagens que variaram muito ao longo dos anos¹. De acordo com a Portaria nº 75 do Ministério da Agricultura, Pecuária e Abastecimento, o percentual atualmente permitido de adição de etanol à gasolina comum é 27%, e nas gasolinas premium, de 25%⁷. Em outros países podem ser utilizados outros compostos oxigenados, como o MTBE (metil, terc-butil éter), que foi proibido devido à contaminação que pode acarretar aos lençóis freáticos em especial^{1,3}. Nas gasolinas de aviação se utiliza o $\text{Pb}(\text{C}_2\text{H}_5)_4$ (chumbo tetraetila), que é proibido para veículos terrestres no Brasil e na Europa por questões toxicológicas¹.

A gasolina que sai das refinarias é composta basicamente por hidrocarbonetos e pequenas quantidades de compostos sulfurados que existem no petróleo, indesejáveis pelo impacto ambiental que produzem¹. A legislação brasileira impõe um limite de 50 mg/kg de enxofre nesta gasolina², sendo que esse composto é removido por hidrogenação das frações, no processo de refino. A ANP também determina limites para outros elementos-traços na especificação da gasolina, como silício, fósforo e chumbo, compostos que têm origem no processamento do petróleo².

1.2 TIPOS DE GASOLINA

Segundo Farah(2012), e Gauto(2016), são produzidos os seguintes tipos de gasolina automotiva no Brasil atualmente:

- Gasolina tipo A comum: gasolina produzida pelas refinarias de petróleo, sem adição de etanol, e que não é comercializada ao público, somente é entregue às companhias distribuidoras. Não tem número de octanas definido, pois não deve ser utilizada pura.
- Gasolina tipo A *Premium*: gasolina formulada com naftas de maior octanagem do que as utilizadas na gasolina comum (maior proporção de naftênicos e aromáticos), portanto possui maior número de octanas, na faixa de 95.

- Gasolina tipo A Podium[®]: gasolina desenvolvida pela Petrobras para atender aos requisitos de qualidade de motores nacionais e importados de altas taxas de compressão e alto desempenho.
- Gasolina tipo C comum: mistura da gasolina tipo A comum com etanol anidro, na proporção legal prevista, feita nas distribuidoras de combustíveis.
- Gasolina tipo C *Premium*: mistura da gasolina tipo A *premium* com etanol anidro, na proporção legal prevista.
- Gasolina tipo C Podium[®]: mistura da gasolina tipo A Podium com etanol anidro, na proporção legal prevista.

Existe ainda a gasolina tipo C aditivada, que é formulada pelas distribuidoras de combustíveis, pela adição de substâncias com propriedades detergentes, antioxidantes e redutoras de fuligem à gasolina tipo C comum. O tipo de aditivo varia conforme a distribuidora de combustível. A octanagem dessa gasolina é 87¹.

Gasolinas especiais, como a de aviação ou a de competição, podem ou não possuir etanol em sua composição, e possuem octanagem bem mais alta que as citadas anteriormente¹.

1.3 MOTOR CICLO OTTO E OCTANAGEM

A gasolina é utilizada como combustível para motores que operam sob o ciclo Otto. A função desses motores é produzir energia mecânica na forma de movimento, a partir de energia térmica liberada na combustão da gasolina^{1,3,8}.

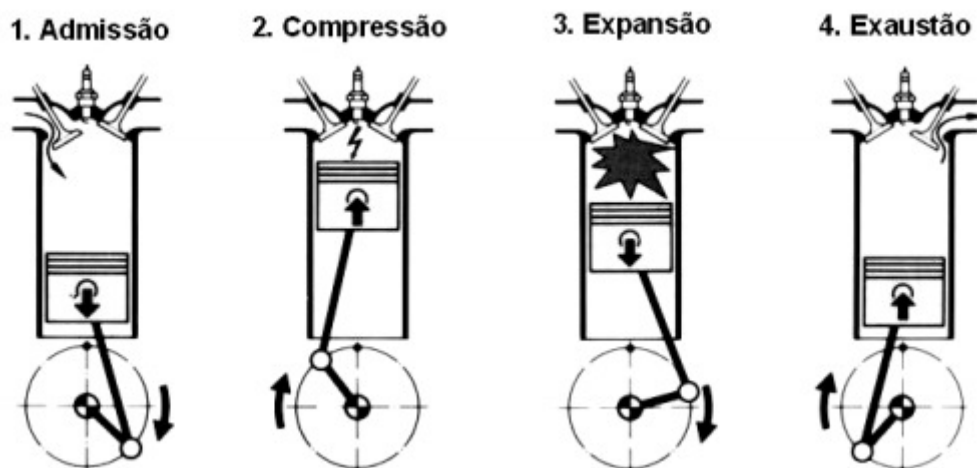
Os motores do ciclo Otto possuem uma câmara de combustão onde ocorre a reação de queima do combustível. Dentro dessa câmara, desliza um pistão que será responsável por transformar a energia liberada na reação química em energia mecânica, através de seu movimento, que é transmitido por um conjunto de eixos e engrenagens às rodas do veículo, levando ao deslocamento deste^{1,3}. Conforme descrito por Gauto(2016), e ilustrado na Figura 1, o processo ocorre basicamente em quatro etapas:

- Admissão: após a mistura com o ar (oxigênio), a gasolina é admitida por meio de uma válvula para dentro do cilindro de reação, enquanto o pistão se move para aumentar o espaço no interior da câmara.
- Compressão: o pistão faz um movimento ascendente, comprimindo a gasolina em um pequeno espaço dentro da câmara.

- **Combustão:** com o auxílio da vela de ignição, é produzida uma centelha que dá origem à reação de combustão. Essa reação gera gases a alta temperatura e pressão que, ao se expandirem, forçam o pistão para baixo provocando o movimento, transformando a energia térmica liberada em energia mecânica.
- **Descarga:** depois de concluída a reação, abre-se a válvula de descarga para que o gás carbônico e o vapor d'água produzidos na reação sejam removidos da câmara, seguindo para o escapamento do motor.

Após a descarga, o processo se reinicia, com a admissão de nova quantidade de mistura ar-gasolina⁸.

FIGURA 1: Ciclo de trabalho de um motor ciclo Otto



Fonte: MILHOR (2002)

Na fase de compressão no motor ciclo Otto, a gasolina só pode entrar em combustão a partir da ignição pela centelha, sem haver combustão espontânea^{3,8}. A autoignição da mistura ar-combustível é chamada de fenômeno de detonação, e quando ela ocorre pode haver oscilações de pressão na câmara que não são absorvidas pelo pistão, causando danos mecânicos ao motor⁸. Surge daí a definição de octanagem ou número de octanas.

A octanagem mede a resistência da gasolina à detonação dentro da câmara de combustão do ciclo Otto, ou seja, indica o quanto a mistura ar-combustível resiste às altas pressões e temperaturas formadas na câmara antes da queima. Quanto maior o número de octanas, maior será a capacidade de o combustível ser comprimido, sob altas temperaturas, na câmara de combustão, sem que ocorra a detonação, o que resulta em um melhor desempenho do veículo, que opera com maiores taxas de compressão¹.

O número de octanagem baseia-se na comparação da gasolina a um padrão composto de isoctano (2,2,4-trimetilpentano) e n-heptano, em determinadas proporções. O método de comparação atribui ao isoctano uma octanagem 100, e ao n-heptano o número de octanas zero^{1,8}. Dessa forma, se tivermos uma gasolina com octanagem 87, significa que ela apresenta resistência à detonação equivalente a uma mistura de 87% de isoctano e 13% de n-heptano. De forma geral, compostos parafínicos possuem baixa octanagem, enquanto os olefínicos, naftênicos e aromáticos possuem octanagem alta^{1,3}.

Dois ensaios podem ser feitos para a determinação do número de octano: o MON (*Motor Octane Number*), segundo a norma ASTM D2700, e o RON (*Research Octane Number*), segundo a norma ASTM D2699. Ambos são feitos num motor padrão (CFR – *Cooperative Fuel Research*) de um único cilindro, em diferentes condições de rotação, temperatura da mistura ar-combustível e avanço da centelha de ignição^{10,11}, que está mostrado na Figura 2. As condições de teste do método de MON são bem mais drásticas que as do método de RON, representando condições de operação mais severas, relativas a maiores rotações do motor. O motor CFR é produzido pela Waukesha Motor Company, o único aceite globalmente para a determinação e certificação de características anti-detonantes de combustíveis¹².

FIGURA 2: Motor CFR utilizado para análise de octanagem da gasolina



A determinação do número de octano é realizada variando-se a taxa de compressão do motor quando se queima a amostra, até que se obtenha uma intensidade sonora padrão. Depois disso, queimam-se misturas de isoctano e n-heptano até que, nas mesmas condições, se alcance a mesma intensidade de detonação. O número de octano da amostra será igual ao percentual de isoctano na mistura padrão que alcançar a mesma detonação do produto³.

A qualidade antidetonante da gasolina também pode ser expressa pelo índice antidetonante (IAD), definido pela média aritmética de RON e MON^{2,8}.

1.4 ESPECTROSCOPIA DE INFRAVERMELHO

A radiação infravermelha corresponde à parte do espectro eletromagnético situada entre as regiões do visível e das micro-ondas¹⁴, compreendendo a radiação com número de onda entre 12800 e 10 cm^{-1} ou com comprimento de onda de 0,78 a 1000 μm ¹⁵.

O espectro infravermelho é subdividido em infravermelho próximo (NIR, do inglês *near infrared*), infravermelho médio (MID) e infravermelho distante (FAR)¹⁵, conforme mostra a Tabela 1. A região do NIR é a mais próxima da região visível, e a FAR é a mais próxima das micro-ondas.

TABELA 1: Regiões espectrais do infravermelho

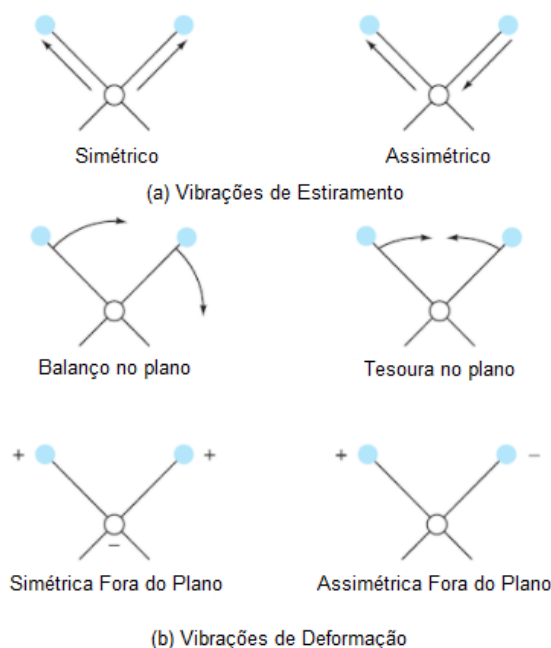
Região	Comprimentos de onda (λ), μm	Números de onda, cm^{-1}	Frequências, Hz
Próximo	0,78 a 2,5	12800 a 4000	$3,8 \times 10^{14}$ a $1,2 \times 10^{14}$
Médio	2,5 a 50	4000 a 200	$1,2 \times 10^{14}$ a $6,0 \times 10^{12}$
Distante	50 a 1000	200 a 10	$6,0 \times 10^{12}$ a $3,0 \times 10^{11}$
Mais usada	2,5 a 15	4000 a 670	$1,2 \times 10^{14}$ a $2,0 \times 10^{13}$

Fonte: HOLLER, F. J. *et al* (2009)

Para absorver a radiação infravermelha, uma molécula deve vibrar, produzindo uma variação no momento de dipolo, que gera um campo que interage com o campo elétrico da radiação¹⁵. Baseando-se na ideia de que as ligações átomo-átomo dentro das moléculas vibram com frequência característica, ao ser emitida uma radiação nesta mesma frequência, ocorre absorção da radiação, e a amplitude da vibração molecular aumenta. Moléculas diatômicas homonucleares tem momento dipolar igual a zero, e portanto, não absorvem a radiação infravermelha¹⁵.

As vibrações moleculares podem ser do tipo estiramento, que alteram a distância entre os átomos, ou de deformação, que modificam o ângulo entre as ligações. A Figura 3 ilustra os tipos de vibrações.

FIGURA 3: Tipos de vibrações moleculares



Fonte: HOLLER, F. J. *et al* (2009)

À temperatura ambiente, a maioria das moléculas está vibrando no seu estado menos energético permitido pela mecânica quântica, ou seja, no nível zero de energia. Quando absorvem a luz de uma determinada frequência, elas são excitadas a um nível de energia vibracional maior¹⁶. Portanto, ao contrário da mecânica clássica, onde os sistemas vibracionais podem ter qualquer energia potencial, os sistemas vibracionais quânticos podem admitir apenas certas energias discretas, o que quer dizer que os níveis de energia vibracionais são quantizados¹⁵. No caso em que a transição energética é do nível 0 de energia para o nível 1, ela é considerada “fundamental”. Quando as transições são do estado fundamental para outros níveis de energia, elas são chamadas de “sobretons”¹⁶. Quando um quantum de energia é absorvido por duas vibrações ao invés de uma, forma-se uma “banda de combinação”. E quando a energia de uma vibração é influenciada por outras vibrações na molécula, ocorrem bandas de “acoplamento vibracional”¹⁵.

Na espectroscopia de infravermelho médio os máximos de absorção são consequência de vibrações fundamentais das ligações covalentes encontradas nos compostos orgânicos¹⁴. Já

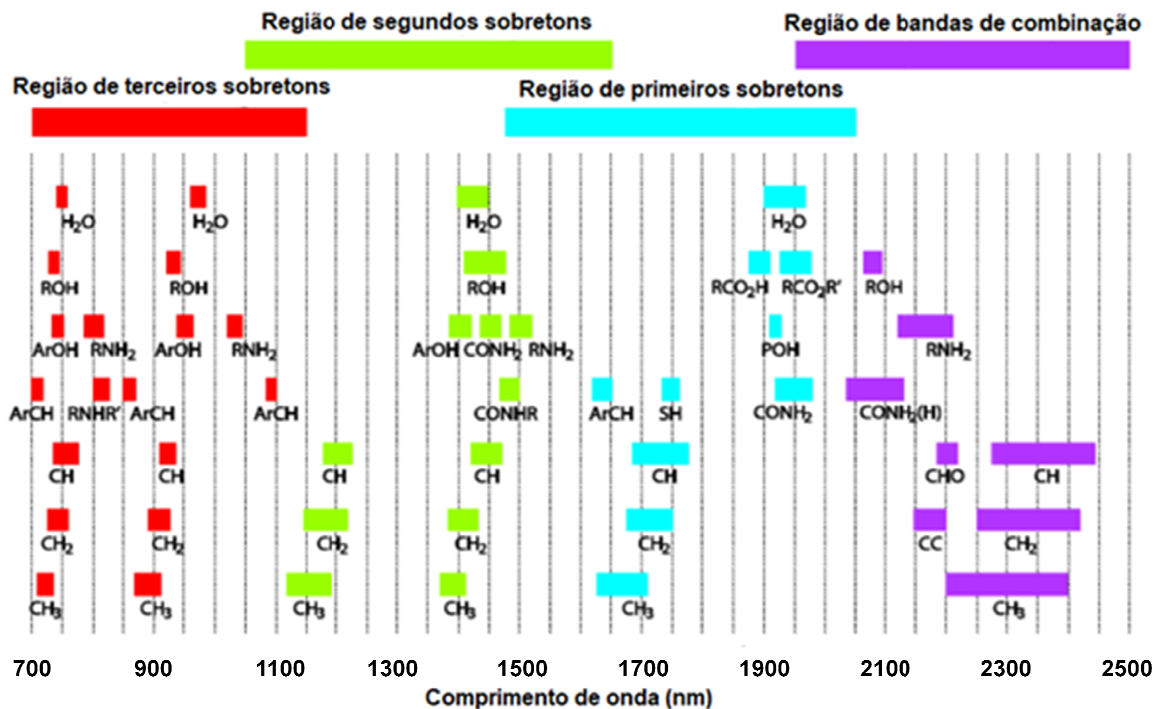
na região do infravermelho distante ocorre a absorção devido às vibrações de ligações entre átomos metálicos e ligantes orgânicos e inorgânicos ¹⁵.

1.5 INFRAVERMELHO PRÓXIMO – NIR

A região espectral do NIR, conforme mostra a Tabela 1, estende-se de 780 nm a 2500 nm (12800 a 4000 cm^{-1}). As principais bandas de absorção nesta região são sobretons ou combinações de vibrações fundamentais que ocorrem na região do MID¹⁵. Somente os grupos funcionais que contém hidrogênio é que possuem absorções relevantes nesta região (CH, NH, OH e SH). As demais absorções dos grupos C=O, C=C, C-F, C-Cl, C-O e C-N são fracas e geralmente não são observáveis no NIR¹⁴.

Conforme mostra a Figura 4, em um espectro de compostos orgânicos as bandas de combinação ocorrem entre 2000 e 2500 nm, as bandas harmônicas de primeira ordem (primeiros sobretons) na faixa de 1500 a 2000 nm, as harmônicas de segunda ordem (segundos sobretons) na faixa de 1100 a 1500 nm e as harmônicas de terceira ordem (terceiros sobretons) na faixa de 700 a 1100 nm.

FIGURA 4: Bandas de absorção no infravermelho próximo



Fonte: NIR SPECTRAL REGION, 2017 (adaptado)

As bandas espectrais no NIR normalmente são largas e sobrepostas¹⁵, porém mais fracas (menores absorvidades) que as bandas de absorção no MID¹⁴. Dificilmente se consegue correlações simples do espectro com características físicas da amostra, sendo geralmente necessárias técnicas de calibração multivariadas¹⁵. Atualmente, os instrumentos NIR encontrados no mercado já vêm com softwares para o desenvolvimento de modelos de calibração¹⁵.

As principais aplicações do NIR são em análises quantitativas de materiais industriais e em controle de processos¹⁵.

1.6 TÉCNICAS QUIMIOMÉTRICAS DE ANÁLISE EXPLORATÓRIA

O grande avanço tecnológico na área de análise instrumental nos últimos anos fez surgir instrumentos que geram até centenas de milhares de dados por análise, dados estes que precisam ser tratados e interpretados para que tenham alguma utilidade química. Neste contexto surgiu a quimiometria: na busca por ferramentas matemáticas e estatísticas que extraíam informações úteis de grandes conjuntos de dados¹⁸.

A análise de dados é uma parte essencial em todo experimento, sendo chamada “univariada” quando somente uma variável é medida sistematicamente para várias amostras, e “multivariada” quando muitas variáveis são analisadas simultaneamente e correlacionadas, gerando quantidade muito maior de informação¹⁹.

A aplicação de métodos matemáticos a um conjunto de dados multivariado, como por exemplo, espectros no infravermelho próximo, permite uma simplificação do mesmo no sentido de comprimir o espaço dimensional a que este está confinado, possibilitando desta forma uma melhor interpretação e visualização²⁰.

Os métodos quimiométricos podem ser aplicados em dados multivariados com os propósitos qualitativos (análise exploratória e reconhecimento de padrões) e quantitativos (calibração)²⁰.

1.6.1 Organização dos dados

Após coletados, os dados devem ser devidamente organizados em um arranjo de linhas e colunas, constituindo uma matriz X , onde as linhas representam as n -amostras e as colunas representam as m -variáveis. No caso de um espectro no NIR, as variáveis correspondem aos comprimentos de onda, e os elementos da matriz (X_{nm}) são as respostas para as variáveis, no

caso, as medidas de intensidade de absorção da amostra em cada comprimento de onda^{18, 19, 20, 21}. Após organizada a matriz, é importante a visualização dos dados através de gráficos, a fim de verificar a distribuição das amostras antes de qualquer tratamento¹⁸.

1.6.2 Pré-tratamento dos dados

A etapa de pré-tratamento dos dados é muito importante para o sucesso da análise multivariada. O objetivo é reduzir as variações indesejáveis dos dados, decorrentes da análise instrumental, que podem dificultar a extração de informações mascarando uma variação de interesse.

Existem diversos métodos de pré-tratamento de dados utilizados em espectroscopia, entre eles podemos citar: técnicas de alisamento (*smoothing*), normalização de dados, aplicação de derivadas, aplicação de correção do espalhamento multiplicativo (MSC).

1.6.2.1 Técnicas de alisamento (*Smoothing*)

O alisamento dos dados é aplicado para redução da componente aleatória dos dados e aumento da razão sinal/ruído¹⁸.

Define-se um pequeno segmento do espectro (uma janela), com um número ímpar de comprimentos de onda (m). A cada intervalo de $2m+1$, as intensidades de cada comprimento de onda são usadas para calcular uma única resposta, o valor no ponto central da janela. Feito isso, a janela se move, percorrendo todo espectro^{18, 22}.

No alisamento pelo método Savitzky-Golay, é utilizada uma média ponderada entre os pontos da janela, em que os pesos são obtidos por um ajuste polinomial, de grau escolhido^{18, 23}.

1.6.2.2 Derivadas

Deslocamentos de linha de base podem ser corrigidos aplicando-se a primeira derivada do espectro. Já para inclinações de linha de base, aplica-se a segunda derivada^{18, 22}.

O método mais utilizado para o cálculo das derivadas é o de Savitzky-Golay, que usa um polinômio de grau n para ajustar o valor no ponto central da janela móvel. Quando se define o tamanho da janela, acaba que se aplica, juntamente com as correções da linha base, um alisamento do espectro^{18, 23}.

1.6.2.3 Normalização

Nessa transformação, todos os valores de uma amostra são divididos por um fator de normalização. O fator geralmente utilizado é o valor médio das variáveis de cada amostra, ou o valor máximo delas. Com este tratamento todas as amostras serão dimensionadas em uma escala predeterminada. No caso da normalização de máximo, todos os valores serão maiores que zero e menores que 1^{18,23}.

1.6.2.4 MSC – Correção do espalhamento multiplicativo

A MSC corrige efeitos de espalhamento aditivos e multiplicativos na absorbância, causados por fenômenos físicos como mudanças no caminho ótico, na sensibilidade do detector, diferenças no tamanho de partículas, rugosidade, opacidade, o que pode causar deslocamento da linha base^{18,22,23}.

Assim como as derivadas, a MSC também corrige deslocamentos da linha base, mas a derivada segunda corrige a inclinação na linha base, o que não ocorre com MSC. Por atuarem de formas diferentes, a MSC e a derivada podem ser aplicadas uma após a outra¹⁸.

1.6.2.5 Outras correções de linha base (*Baseline*)

Deslocamentos de linha de base também podem ser corrigidos subtraindo-se um valor arbitrário de cada espectro. Por exemplo, pode-se usar uma intensidade média relativa apenas à linha de base, pode-se escolher os pontos para definir a nova linha de base, ou ainda pode-se usar o menor valor de cada espectro para corrigir cada um deles^{18,23}.

O ideal seria escolher uma função para cada espectro, o que é praticamente impossível para um conjunto grande de amostras¹⁸.

1.6.3 Análise por Componentes Principais – PCA

Grandes tabelas de dados geralmente contêm grande quantidade de informação que fica parcialmente escondida devido à complexidade dos dados e sua difícil interpretação. A análise de componentes principais é um método de projeção que ajuda a visualizar toda a informação contida nos dados multivariados em um espaço de dimensão menor, tornando mais evidentes as diferenças entre as variáveis e as relações entre as amostras¹⁸. A PCA

representa o conjunto dos dados em um novo sistema de eixos, que são denominados componentes principais (PC)^{18,23}.

As n-variáveis originais geram, através de suas combinações lineares, n-componentes principais, que por sua vez são ortogonais entre si²⁴. A primeira PC (PC1) é definida pela direção que descreve a maior variância dos dados originais, e carrega mais informação; a segunda PC (PC2) tem máxima variância no espaço ortogonal à PC1; as componentes subsequentes são ortogonais às anteriores e obtidas até que se descrevam todas as variáveis restantes^{18,23,24}. Embora a informação estatística presente nas n-variáveis originais seja a mesma das n-componentes principais, é comum que mais que 90% da informação esteja presente em apenas poucas PC^{18,24}.

Matematicamente, na PCA, a matriz inicial \mathbf{X} é decomposta no produto de duas matrizes, uma de *scores* (\mathbf{T}) e uma de *loadings*, ou pesos (\mathbf{L}), mais uma matriz de erros (\mathbf{E}), de tal maneira que:

$$\mathbf{X} = \mathbf{TL}^T + \mathbf{E} \quad (1)$$

Os *scores* expressam as relações entre as amostras, ou seja, as suas coordenadas no sistema de eixos formados pelas componentes principais. Os pesos, ou *loadings*, indicam as relações entre as variáveis. Plotando-se PC1 *versus* PC2, se observa, de forma privilegiada, os pontos no espaço n-dimensional, as diferenças e similaridades entre as amostras. Já um gráfico de *loadings* mostra a contribuição das variáveis originais, ou seja, as variáveis com maior peso na combinação linear das componentes principais mais importantes^{18,22,23,24}. No caso de um espectro de infravermelho, o gráfico de *loadings* ajuda a determinar quais as regiões do espectro que mais impactam nas características das amostras.

1.6.4 Calibração Multivariada

A calibração multivariada, ou regressão, relaciona dois blocos de dados, sendo \mathbf{X} a matriz dos dados experimentais, onde as amostras são representadas nas linhas e as variáveis medidas (comprimentos de onda) são representados nas colunas, e \mathbf{Y} a matriz formada pelas variáveis de interesse (MON, por exemplo). Desta forma, a calibração multivariada relaciona a propriedade de interesse das amostras com as muitas respostas²⁰. O objetivo da calibração é construir um modelo $\mathbf{Y}=\mathbf{f}(\mathbf{X})$, para explicar a variação de \mathbf{Y} a partir da variação de \mathbf{X} ²³.

De acordo com PARREIRA(2003), a calibração pode ser dividida em uma etapa de modelagem, onde se estabelece uma relação entre o sinal medido e a propriedade que se

deseja quantificar das amostras, e outra etapa de validação, onde se confirma que o modelo se aplica às amostras, gerando valores para a propriedade de interesse iguais ou próximos dos experimentais.

Para a etapa de calibração, seleciona-se um grupo de amostras, e os dados dos blocos X e Y são coletados e pré-tratados conforme a necessidade. Com este grupo são construídos diversos modelos, até que o melhor deles é selecionado. Para validar o modelo escolhido, existem duas alternativas¹⁸:

- utilizar um outro conjunto de amostras com dados de X e Y conhecidos, mas que não tenham participado da etapa de calibração, o que é conhecido como validação externa. Os resultados encontrados são então comparados aos valores experimentais, e são calculados os resíduos. O modelo é melhor quanto menores forem os resíduos^{18,20};
- utilizar o método de Validação Cruzada, onde as mesmas amostras são usadas na calibração e na validação. A matriz de dados é dividida em pequenos grupos, e um grupo é retirado da matriz para a construção do modelo. Em seguida ele é testado no modelo e os resíduos obtidos entre os valores reais e estimados são computados. Isso é repetido para todos os pequenos grupos do conjunto de dados. Se o conjunto de amostras não for muito grande, pode-se remover uma amostra de cada vez para a validação cruzada, ao invés de formarem-se grupos^{18,20,23}.

Os parâmetros comumente utilizados para avaliar a qualidade e confiabilidade dos modelos construídos são *Root Mean Square Error of Calibration* (RMSEC, ou raiz da média quadrática dos resíduos de calibração), *Root Mean Square Error of Validation* (RMSECV, ou raiz da média quadrática dos erros de validação) e *Root Mean Square Error of Prediction* (RMSEP, ou raiz da média quadrática dos resíduos de previsão)^{18, 20, 23}. O RMSE (também chamado de resíduo) é a medida da diferença entre valores da propriedade de interesse preditos pelo modelo e os valores experimentais, que estão sendo modelados¹⁸. Pode ser calculado conforme a equação:

$$\text{RMSE} = \sqrt{\sum \frac{(y_{pred} - y_{ref})^2}{N}} \quad (2)$$

Onde y_{ref} são os valores experimentais, e y_{pred} é o valor encontrado pelo modelo.

Após a validação, o modelo construído já pode ser usado para a previsão das propriedades de interesse em amostras onde estes valores são desconhecidos.

1.6.5 Método dos Quadrados Mínimos Parciais - PLS

A regressão por quadrados mínimos parciais (PLS) é um método de compressão de dados, que reduz o espaço multivariado a poucas combinações lineares das variáveis originais, denominadas fatores ou variáveis latentes^{18,25}.

A PLS, em uma única etapa, relaciona a matriz original de dados X (dados espectrais) com a matriz das respostas de interesse Y (MON, teor de aromáticos, teor de olefinas), decompondo-as nas variáveis latentes, e otimiza a decomposição na direção da previsão da propriedade Y, para se obter um modelo de calibração com menor erro de previsão da propriedade de interesse^{18, 25, 26}. Logo, as matrizes X e Y são decompostas em:

$$\mathbf{X} = \mathbf{TL}^T + \mathbf{E} \quad (3)$$

$$\mathbf{Y} = \mathbf{UQ}^U + \mathbf{F} \quad (4)$$

onde T e U são as matrizes de *scores* das matrizes X e Y, L e Q são as matrizes de *loadings* das matrizes X e Y, e E e F são as matrizes de resíduos, contendo os dados que não foram modelados.

O número de variáveis latentes que será utilizado no modelo é determinado durante o processo de validação (ou crossvalidação)²⁰.

2 SITUAÇÃO ATUAL

A ANP é o órgão nacional que especifica as características físico-químicas da gasolina e de todos os derivados de petróleo comercializados no país, garantindo que os combustíveis tenham a qualidade mínima necessária para o desempenho esperado.

A octanagem da gasolina, que é intrinsecamente dependente da sua composição, é um dos principais fatores da qualidade de ignição no motor. A ANP determina a utilização da norma ASTM D2700 para a análise de Número de Octano Motor, e especifica em 82,0 o valor mínimo de MON para a gasolina que contém 26% de álcool etílico anidro adicionado. A ANP especifica também o Índice Antidetonante – IAD, que é a média aritmética entre MON e RON, em um valor mínimo de 87,0 para a mesma proporção gasolina-álcool. O RON, por sua vez, não tem limite determinado pela ANP, mas a metodologia a ser utilizada na sua determinação é a ASTM D2699². As duas normas de octanagem utilizam um motor padrão (CFR), em diferentes condições de análise, para a comparação da queima da gasolina com misturas combustíveis padrões^{10,11}. Além do alto investimento com o equipamento, este necessita condições especiais para sua instalação e pessoal treinado para sua operação. Desta forma, diversos estudos vêm sendo realizados para o desenvolvimento de metodologias alternativas para a determinação da octanagem de maneira mais barata, rápida e prática.

A gasolina é composta por diversos tipos de hidrocarbonetos diferentes, no entanto, a ANP especifica apenas o teor máximo de aromáticos em 35% em volume na gasolina com 26% de álcool etílico anidro adicionado, e o teor máximo de olefinas, em 25%². O método para esta análise pode ser ASTM D1319, que é por cromatografia líquida em coluna de vidro preenchida com sílica-gel, onde os grupos são separados e marcados por um indicador fluorescente⁴, ou pode ser por cromatografia gasosa, sem norma específica. A cromatografia gasosa com detector de ionização de chama, que é normalmente utilizada para este fim, possui uma coluna de cem metros, por onde a amostra é arrastada geralmente pelo gás hélio, e, com a aplicação de um gradiente de temperatura, os componentes são separados quase que individualmente, levando até três horas para finalizar a análise⁶.

Na indústria, todas as correntes que formam a gasolina devem ter suas composições e valores de octanagem determinados para que se faça a mistura de forma otimizada, visando a especificação do produto final. Desta forma, realizam-se análises cromatográficas diárias de diversas amostras, no laboratório, demandando tempo, custo e disponibilidade de pessoal treinado. A octanagem dos intermediários é, sempre que possível, determinada por inferência.

Nenhuma das duas análises costuma ser feita em linha de produção, visto a complexidade das técnicas padronizadas.

A espectroscopia por NIR tem sido utilizada em diversas aplicações industriais, inclusive na indústria petroquímica, para análise de derivados de petróleo, especialmente óleo diesel e gasolina. Ela foi favorecida pelo surgimento de novos equipamentos robustos e de alto desempenho, tanto espectrofotômetros como computadores, bem como métodos matemáticos que permitem calibrações multivariadas de elevada complexidade²⁵. O NIR pode ser utilizado tanto em bancada quanto em linha de produção, e tem diversas vantagens quando comparado às técnicas tradicionais, como o fato de ser uma técnica não destrutiva, prática e rápida. Esta técnica ainda não é uma metodologia oficial da ASTM, portanto não pode ser usada para a certificação de produtos finais. No entanto, para a determinação de parâmetros durante o acompanhamento da produção, ela surge como uma técnica muito útil e promissora.

3 OBJETIVOS

Este trabalho tem como objetivo a obtenção de uma metodologia para predição de alguns parâmetros da gasolina, tais como MON, teor de aromáticos e de olefinas, a partir de seus espectros no infravermelho próximo associados a técnicas quimiométricas de PCA e PLS, e que esta análise apresente resultados comparáveis com as técnicas usuais e padronizadas, com o intuito de auxiliar no acompanhamento da produção da gasolina e no controle da qualidade dos produtos intermediários.

4 PROPOSTA TECNOLÓGICA

As indústrias, principalmente na área petroquímica, cada vez mais sentem as exigências legais de especificações de seus produtos. Para garantir o atendimento ao mercado com qualidade e competitividade, torna-se necessário o desenvolvimento de técnicas analíticas rápidas e precisas. A espectroscopia de infravermelho próximo (NIR), aliada a métodos quimiométricos de análise de dados, tem sido exaustivamente estudada para uso neste contexto, se mostrado muito eficiente para aplicação industrial.

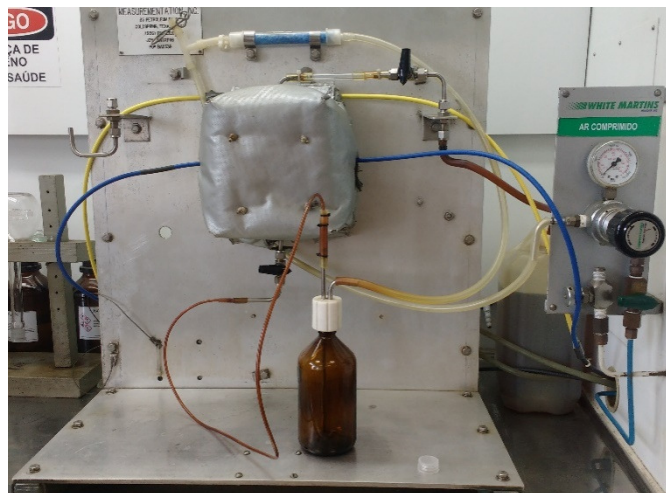
Neste trabalho, busca-se o desenvolvimento de um método rápido e de baixo custo para a determinação de propriedades físico-químicas da gasolina, como sua composição e octanagem. Este método tanto pode ser aplicado em laboratórios, no acompanhamento da produção da gasolina, como pode ser instalado em analisadores em linha, oferecendo resultados instantâneos para a otimização da produção. Desta forma, ao ser investigado o valor de octanagem ou a composição de algum potencial produto ou corrente de mistura, não será necessário empenhar trabalho na análise pelo motor ou aguardar três horas por um cromatograma, pois o método proposto, sem reagentes e em aproximadamente 1 minuto fornecerá resultados confiáveis para serem utilizados no controle da produção.

A capacidade de medir várias propriedades simultânea e rapidamente, sem destruir a amostra e sem gerar resíduos, torna o método ideal para uso nas indústrias petroquímicas, em substituição aos métodos tradicionais, nas análises de rotina e acompanhamento.

5 METODOLOGIA

Para este estudo, 117 amostras de gasolina foram coletadas em frascos de vidro âmbar de 1 litro, de forma que estes fossem preenchidos até 70 a 80% da sua capacidade. Para preservar as características do produto, evitando a perda de frações mais leves e decomposição térmica, os frascos foram mantidos sob refrigeração à temperatura inferior a 4°C até a hora da análise. Durante os 30 minutos anteriores à leitura no espectrofotômetro, as garrafas foram colocadas em banho de gelo. Após este período, transferiu-se, no menor tempo possível, a amostra para outro frasco âmbar, de 500 mL, que possuía em sua tampa as adaptações necessárias para a conexão no equipamento, e procedeu-se a introdução do líquido no sistema de circulação do equipamento, conforme mostrado na Figura 5.

FIGURA 5: Equipamento FT-NIR utilizado para coleta dos espectros



As amostras foram introduzidas no equipamento através da aplicação de uma pressão de ar comprimido no frasco que as continha. Devido às propriedades da gasolina, e a tendência de formação de bolhas, a vazão do ar comprimido precisou ser ajustada para permitir um fluxo constante de produto através do sistema. Após um tempo de circulação, as válvulas eram fechadas, para que a fração de amostra contida na célula estabilizasse à temperatura de 15 °C. Foram testados diferentes períodos de tempo de estabilização da temperatura da amostra na célula, e concluiu-se que com 8 minutos de espera, obtinha-se leituras de espectros mais constantes, de forma que a análise em bancada levou em torno de dez minutos para gerar cada resultado.

Todas as amostras utilizadas neste trabalho tiveram suas propriedades de interesse analisadas, também, pelos métodos tradicionais, ou seja, a composição da gasolina foi feita

por cromatografia pela norma ASTM D6730, e o número de octano motor pela ASTM D2700.

O equipamento utilizado para a aquisição dos espectros foi o FT-NIR NR800 da marca Yokogawa, controlado pelo Software Spectland[®]. Trata-se de um analisador indicado para análises em linha de produção, mas que está instalado na bancada do laboratório para testes e implantações de metodologias que serão posteriormente aplicadas nos processos online. Por ser adequado para uso em locais abertos com condições variáveis, ele possui uma capa térmica impermeável que envolve a célula e as conexões com as fibras óticas, e precisa de algumas adaptações, como um banho de circulação de água com precisão de 0,1°C, para a estabilização da temperatura da célula.

A faixa espectral do NR800 é de 11100 a 3980 cm^{-1} , e a aquisição dos espectros ocorreu com resolução de 8 cm^{-1} , 64 varreduras e em triplicata.

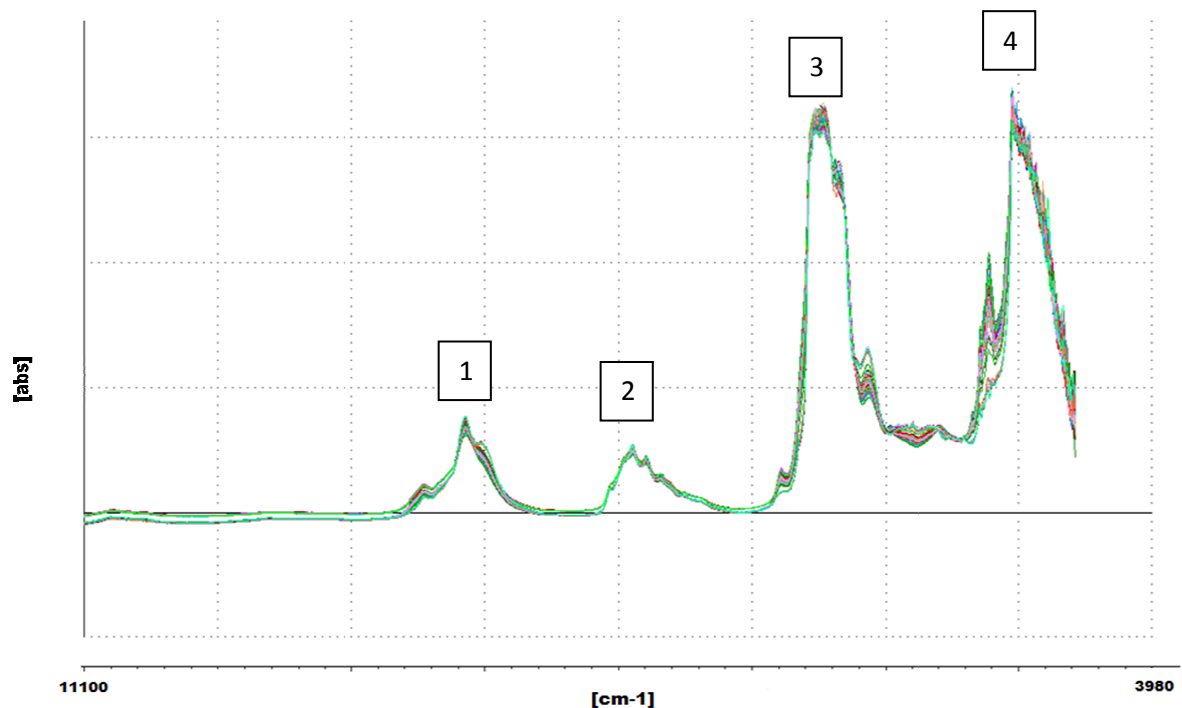
Os dados espectrais das amostras foram tratados pelo software The Unscrambler[®], que é específico para análise de dados multivariados. Estudou-se diversos tipos de pré-tratamentos de dados, como derivadas, normalização, alisamento, MSC e correções de linha de base, e aplicou-se os algoritmos PCA e PLS.

6 RESULTADOS E DISCUSSÕES

6.1 AVALIAÇÃO DOS ESPECTROS

A Figura 6 mostra a sobreposição dos espectros das 117 amostras de gasolina analisadas no espectrofotômetro de infravermelho próximo, sem quaisquer manipulações de dados.

FIGURA 6: Espectros de infravermelho próximo das amostras de gasolina, obtidos na região de 11100 a 3980 cm^{-1}



Fonte: o autor.

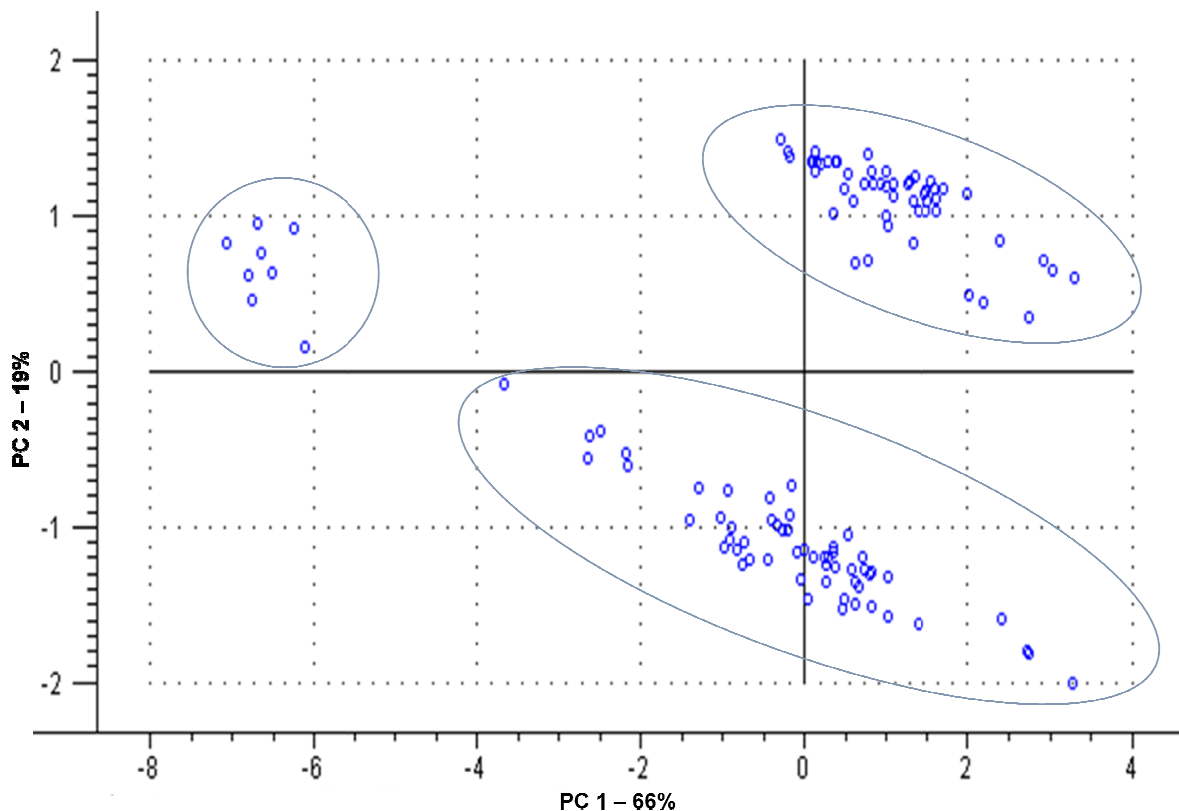
Observa-se na Figura 6:

1. Os espectros são muito semelhantes, com quatro bandas de absorção importantes, características do produto, que estão presentes em todas as amostras, mostrando a similaridade entre elas (de acordo com as regiões do espectro, as bandas 1, 2 e 3 são sobretons, e a 4 é uma banda de combinação);
2. Há uma diferença na linha de base, principalmente no início do espectro, caracterizando dois grupos de amostras. Esta alteração foi devida à abertura e limpeza da célula do equipamento, que ocasionou uma queda da linha de base nas leituras posteriores.

Todos os dados espectrais foram transferidos para o software The Unscrambler® em forma de uma matriz de dados, onde as amostras foram dispostas em linhas e as variáveis (os números de onda) em colunas. Os elementos da matriz são as absorções em cada número de onda. Procedeu-se, então, uma PCA, para melhor interpretação das características das 117 amostras.

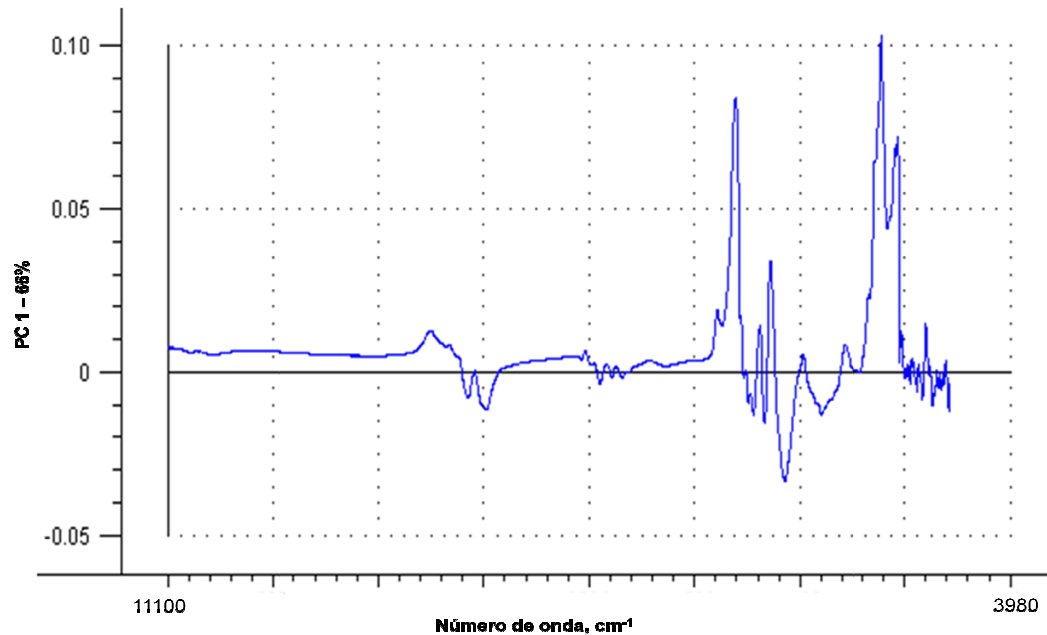
O gráfico de *scores*, apresentado na Figura 7, mostra claramente a separação entre três grupos de amostras. Os dois grupos maiores, separados pela PC2, caracterizam as amostras antes e depois da limpeza da célula, e o terceiro grupo, menor e mais separado do restante pela PC1, são gasolinas com características diferentes das demais. Este gráfico ainda mostra que a primeira PC explica 66% da informação espectral das amostras, e a segunda PC explica 19%.

FIGURA 7: Gráfico de *scores* da PC1 versus a PC2, de todas as amostras



Fonte: o autor.

O gráfico de *loadings*, na Figura 8, apresenta as regiões do espectro com maior variabilidade, ou seja, mostra a importância de cada variável para a construção da PC1.

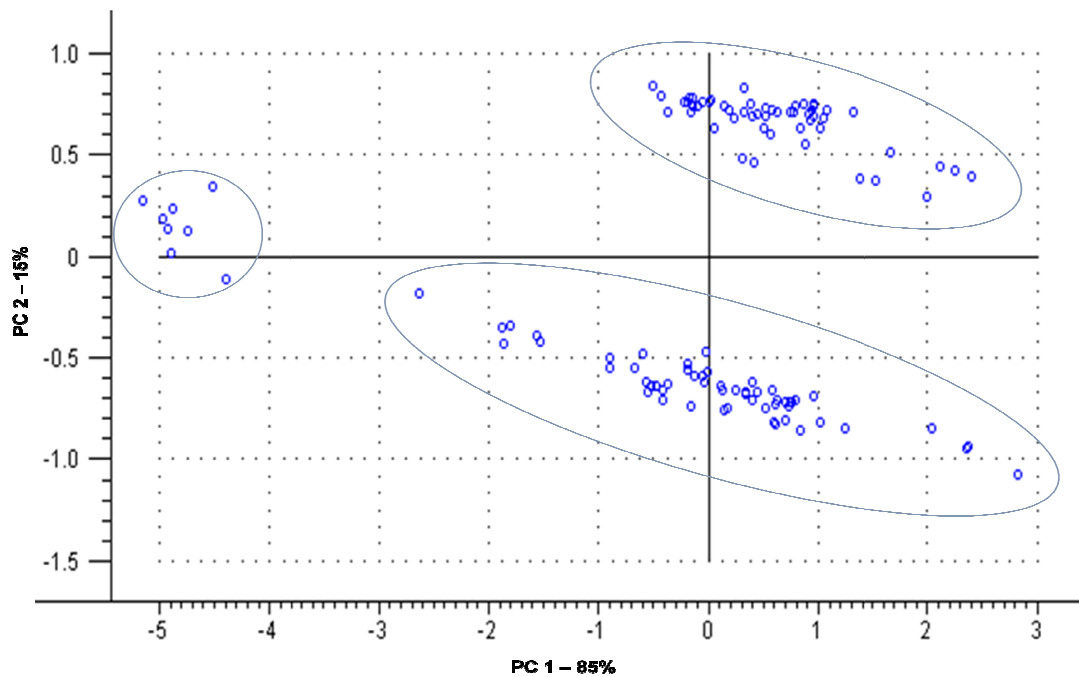
FIGURA 8: Gráfico de *loadings* da PC1

Fonte: o autor.

Baseando-se nesse gráfico e nos espectros originais, se escolheu as variáveis (4 regiões do espectro) que mais influenciam as características das amostras. Procedeu-se então outra PCA, selecionando-se apenas as regiões escolhidas dos espectros de todas as amostras.

O novo gráfico de *scores* gerado desta PCA, mostrado na Figura 9, continuou separando os grupos de amostras de forma muito semelhante ao gráfico da Figura 7, demonstrando que elas devem sofrer algum tipo de tratamento estatístico para corrigir as interferências experimentais durante a aquisição dos espectros. A primeira componente principal, no entanto, passou a explicar 85% das características das amostras, e a segunda, 15%, comprovando que as regiões escolhidas para o estudo foram acertadas, pois reduziram a quantidade de PC necessárias para explicar a relação entre as amostras.

FIGURA 9: Gráfico de *scores* da PC1 versus a PC2, de todas as amostras, considerando apenas a faixa espectral que contém as variáveis mais importantes



Fonte: o autor.

6.2 MON

Após escolhidas as regiões de interesse, diversos pré-tratamentos foram testados, com o intuito de ajustar a linha de base dos dados espectrais para então proceder a calibração. Para avaliar a eficiência dos pré-tratamentos, utilizou-se a ferramenta PLS, correlacionando as variáveis escolhidas com o valor de MON das amostras predito pela análise tradicional.

As regressões dos dados pré-tratados foram feitas utilizando validação-cruzada, e geraram os resultados que se encontram na Tabela 2.

TABELA 2: Figuras de mérito obtidas por PLS após diversos pré-tratamentos dos dados espectrais para MON

PRÉ-TRATAMENTOS	Nº VARIÁVEIS LATENTES	RMSEC	RMSECV	R ² _{cal}	R ² _{val}
Normalização pela média	8	0,2063	0,2519	0,8947	0,8433
Normalização de máximo	8	0,1997	0,241	0,9021	0,8580
Normalização + <i>Baseline</i>	7	0,2179	0,2616	0,8849	0,8345
Alisamento	7	0,2179	0,2603	0,8849	0,8361
Alisamento + <i>Baseline</i>	8	0,1945	0,2401	0,9074	0,8599
Alisamento + 1ª Derivada	8	0,1787	0,2206	0,9233	0,8830
Alisamento + 2ª Derivada	7	0,2063	0,2420	0,8977	0,8593
Baseline	7	0,2079	0,2421	0,8930	0,8552
1ª Derivada	8	0,1721	0,2156	0,9287	0,8883
2ª Derivada	8	0,1727	0,2359	0,9283	0,8669
MSC	8	0,1963	0,231	0,9048	0,8681
<i>Baseline</i> + 1ª Derivada	8	0,1722	0,2248	0,9287	0,8787
MSC + 1ª Derivada	7	0,1854	0,2241	0,9174	0,8793

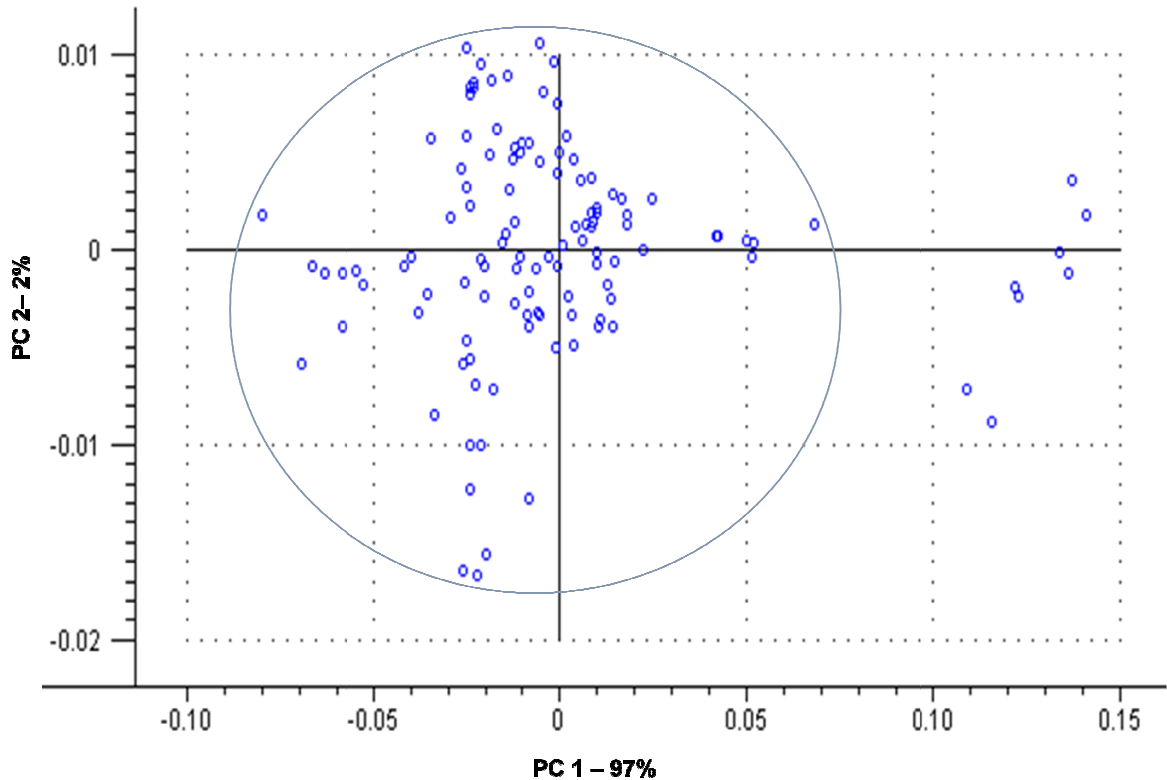
Fonte: o autor.

Em todos os casos, o número de variáveis latentes utilizados no modelo de calibração não variou muito, apenas entre 7 e 8. O RMSEC variou de 0,1721 a 0,2179, e o RMSECV, de 0,2156 a 0,2616. Os menores valores foram encontrados com os dados tratados pela primeira derivada. O coeficiente de correlação da curva de calibração, R²_{cal}, variou de 0,8849 a 0,9287, e o R²_{val}, coeficiente de correlação da curva de validação, de 0,8345 a 0,8883. As maiores correlações também foram encontradas para dados tratados pela primeira derivada.

Após avaliação criteriosa dos resultados apresentados na Tabela 2, optou-se por tratar os espectros com a 1ª derivada para poder então utilizá-los na confecção de um modelo para a determinação do MON da gasolina. Este tratamento foi escolhido em função dos maiores coeficientes de correlação e dos menores RMSEC e RMSEP, comparando com os resultados encontrados após os outros tratamentos.

Para a verificação da decisão, procedeu-se uma PCA com todas as amostras, utilizando as regiões escolhidas do espectro tratadas com derivada primeira. O gráfico de *scores*, desta vez, demonstrou a efetividade do pré-tratamento em corrigir as diferenças entre as linhas de base dos espectros. Além disso, a primeira PC sozinha explicou 97% das propriedades das amostras, conforme mostra a Figura 10.

FIGURA 10: Gráfico de *scores* da PC1 versus PC2, com todas as amostras, considerando apenas as variáveis mais importantes, e com os dados espectrais tratados pela primeira derivada

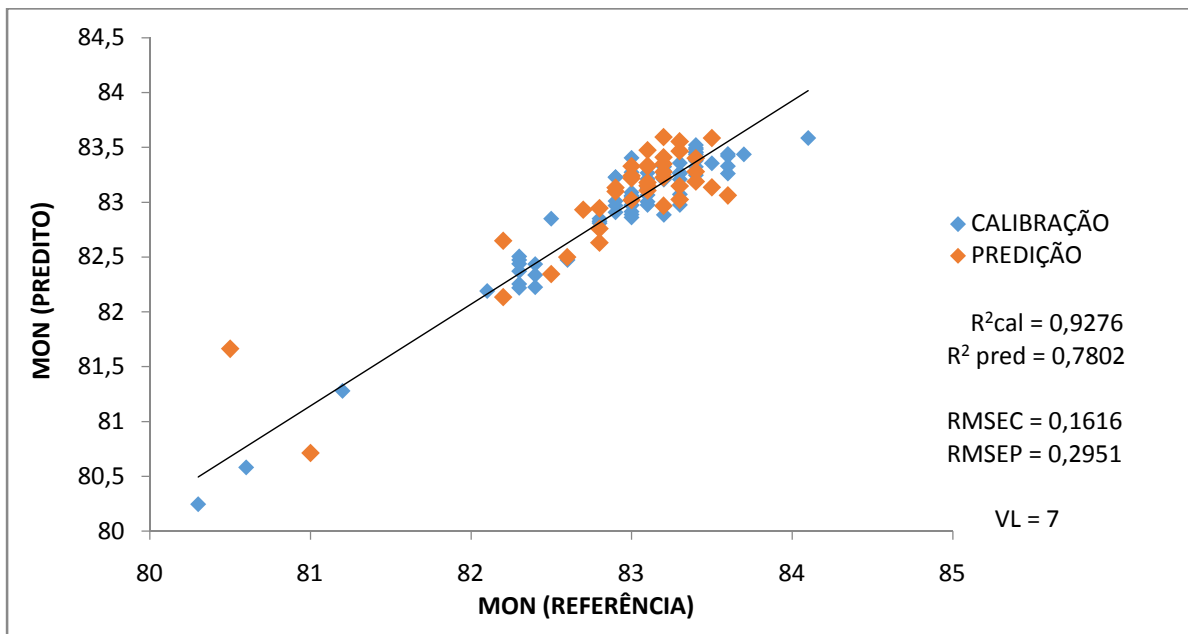


Fonte: o autor.

Após escolhido o melhor pré-tratamento, utilizando todas as amostras, separou-se um grupo de amostras para construção do modelo (calibração) e um grupo menor de amostras para previsão dos valores de MON. A divisão entre as amostras foi feita de forma randômica. As amostras do conjunto de calibração possuem valores de MON entre 80,3 e 83,6, o que abrange toda a variabilidade de octanagem entre as amostras disponíveis.

A Figura 11 mostra um gráfico com os valores obtidos experimentalmente pela metodologia de referência versus os valores previstos pelo modelo das amostras de calibração (em azul) e das amostras de previsão (em laranja). O número de variáveis latentes utilizado na regressão foi 7. Obteve-se um coeficiente de correlação da curva de calibração igual a 0,9276, e da curva das amostras de previsão, igual a 0,7802. O RMSEC encontrado foi 0,1616, e o RMSEP, 0,2951.

FIGURA 11: Gráfico dos valores de MON de referência versus valores preditos pelo modelo para amostras de calibração e predição

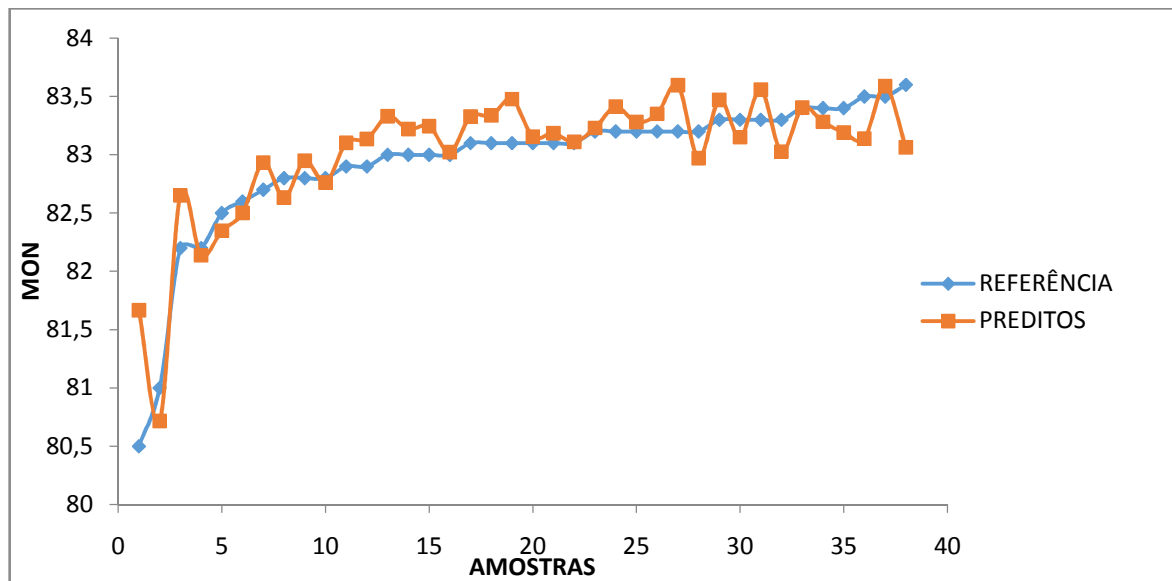


Fonte: o autor.

Como pode ser visto, o conjunto de calibração contém muitas amostras com octanagem entre 82,0 e 83,6, tornando o modelo bem robusto para esta faixa. No entanto, pouquíssimos resultados de MON menores que 82,0 e maiores que 83,7 tornam essas regiões descobertas pelo modelo, aumentando o erro de predição nestes casos.

A Figura 12 mostra um gráfico com os valores encontrados para cada amostra do conjunto de predição (em laranja) plotados sobre os valores de referência de cada uma delas (em azul).

FIGURA 12: Gráfico comparativo dos valores de MON de cada amostra do conjunto de predição obtidos pelo método de referência e pelo modelo



Fonte: o autor.

Podemos novamente perceber que os maiores desvios de predição encontram-se nas regiões de mais alta e de mais baixa octanagem, com a maior diferença entre o valor de referência e o valor predito para a amostra com o menor MON. O conjunto de predição não contém amostras com octanagem próxima a 84, portanto não podemos verificar como seria o erro nesta faixa. Para o restante das amostras verifica-se que o desvio varia sem apresentar tendências. Além disso, um RMSEP de 0,2951 foi considerado muito baixo, levando-se em conta que a reprodutibilidade do método de referência é 0,9.

6.3 TEOR DE AROMÁTICOS

Para correlacionar os espectros da região do infravermelho com o teor de aromáticos, novamente foi realizado um trabalho para verificar o melhor pré-tratamento dos dados espectrais. Aplicou-se, então, diferentes tipos de tratamentos e procedeu-se a regressão pelo método de PLS após cada um deles. Neste caso, apenas foram testados os tratamentos mais efetivos no caso do MON, visto que as técnicas de alisamento e normalização não atuaram na linha de base dos espectros conforme era necessário.

Conforme a Tabela 3, em todos os casos, o número de variáveis latentes utilizados no modelo de calibração foi 1. O RMSEC variou de 1,0134% a 1,0514%, e o RMSECV, de 1,0309% a 1,0727%. Os menores valores foram encontrados com os dados tratados por

baseline e pela primeira derivada. O coeficiente de correlação da curva de calibração, R^2_{cal} , variou de 0,5838 a 0,6648, e o R^2_{val} , coeficiente de correlação da curva de validação, de 0,5668 a 0,6532. As maiores correlações também foram encontradas para dados tratados por *baseline* e pela primeira derivada. Neste caso, como a variação foi mínima, optou-se também por trabalhar com a primeira derivada, pois esta já é utilizada e reconhecida em diversas aplicações com derivados de petróleo.

TABELA 3: Figuras de mérito obtidas por PLS após diversos pré-tratamentos dos dados para teor de aromáticos

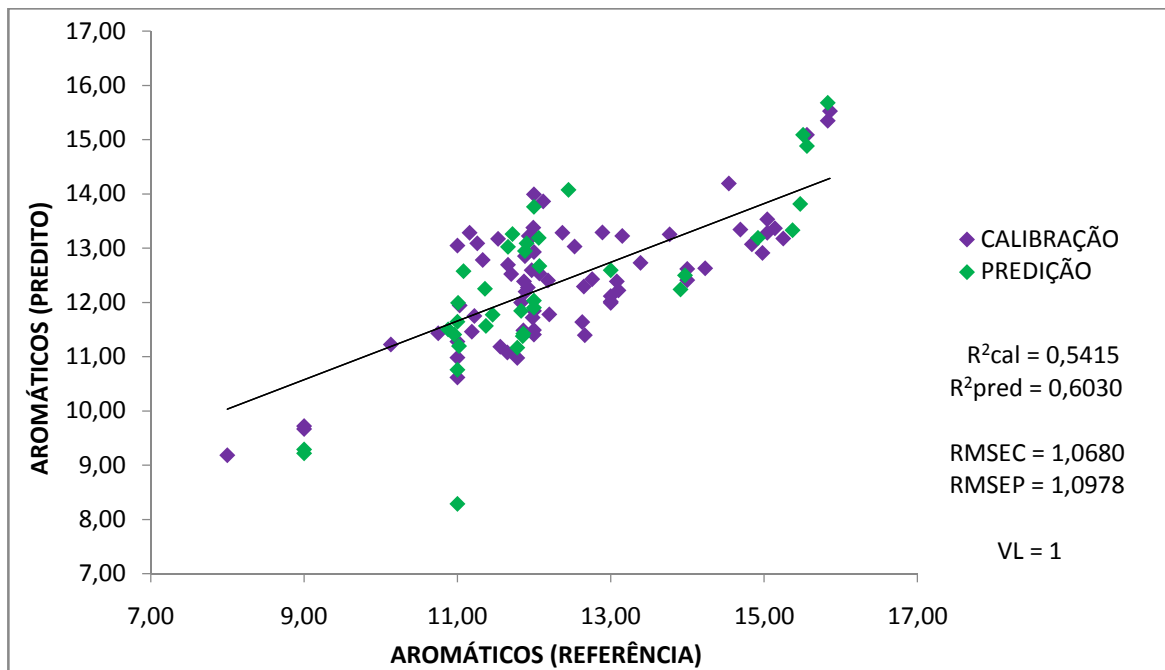
PRÉ-TRATAMENTOS	Nº VARIÁVEIS LATENTES	RMSEC	RMSECV	R^2_{cal}	R^2_{val}
1ª Derivada	1	1,0186	1,0309	0,6614	0,6532
Baseline	1	1,0134	1,0331	0,6648	0,6517
MSC	1	1,0155	1,0379	0,6634	0,6485
MSC + 1ª Derivada	1	1,0457	1,0587	0,6499	0,6411
MSC + Baseline	1	1,0514	1,0727	0,5838	0,5668
2ª Derivada	1	1,0434	1,0658	0,6516	0,6363

Fonte: o autor.

Escolhido o melhor pré-tratamento, separou-se um grupo de amostras para calibração e um grupo menor de amostras para previsão do teor de aromáticos. A divisão entre as amostras foi feita de forma randômica. As amostras do conjunto de calibração possuem teor de aromáticos entre 8,0 e 16,0% em volume, abrangendo toda a variabilidade de aromáticos entre as amostras disponíveis.

A Figura 13 mostra um gráfico com os valores obtidos experimentalmente pela metodologia de referência versus os valores previstos pelo modelo das amostras de calibração (em roxo) e das amostras de predição (em verde).

FIGURA 13: Gráfico dos teores de aromáticos pelo método de referência versus valores preditos pelo modelo para amostras de calibração e predição



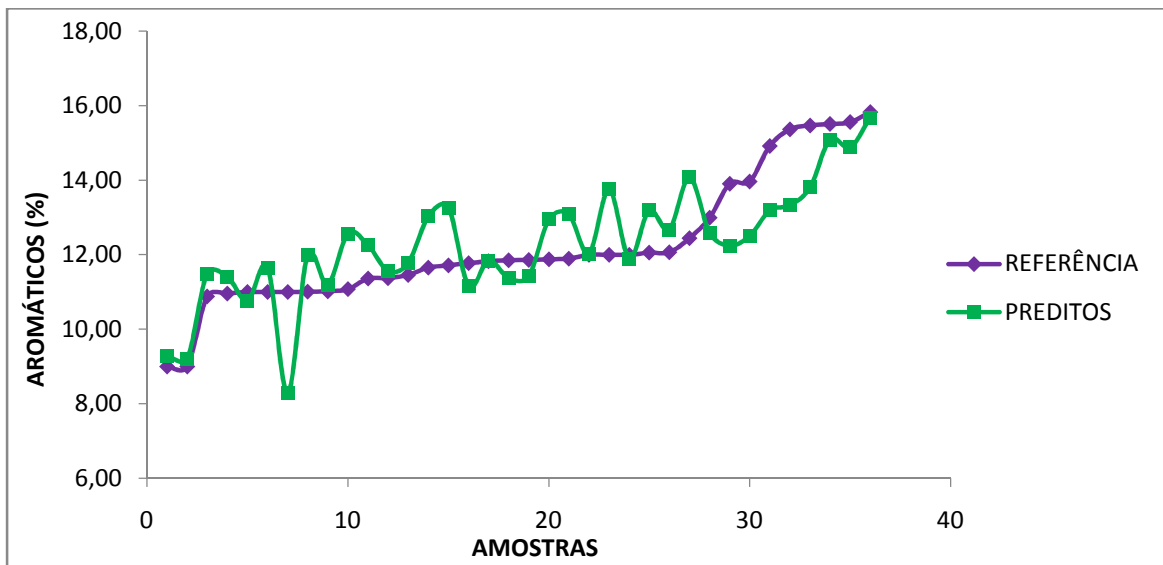
Fonte: o autor.

O número de variáveis latentes utilizado na regressão foi 1. O RMSEC e o RMSEP encontrados foram altos, 1,0690% e 1,0978%, respectivamente. Em contrapartida, os coeficientes de correlação da curva de calibração e de predição ficaram muito baixos, 0,5415 e 0,6030, respectivamente. Sendo o coeficiente de correlação um parâmetro estatístico que representa o grau de associação entre duas variáveis, a baixa correlação encontrada pelo modelo indica que os dados contém informações distintas, ou seja, os resultados pelo método de referência não estão bem relacionados aos do modelo.

Conforme mostra a figura, o conjunto completo de amostras possui poucas representantes com baixo e alto teor de aromáticos, ficando assim, um grupo concentrado de amostras na faixa média do gráfico (11,00 a 14,00%). Ainda assim, mesmo nessa região, percebe-se grandes desvios entre os valores de referência e os encontrados pelo modelo, tanto no conjunto de calibração quanto no de predição.

A Figura 14 mostra o gráfico com os valores encontrados pelo modelo para cada amostra do conjunto de predição (em verde) plotados sobre os valores de referência de cada uma delas (em roxo).

FIGURA 14: Gráfico comparativo dos valores de aromáticos de cada amostra do conjunto de predição obtidos pelo método de referência e pelo modelo



Fonte: o autor.

Podemos perceber que os desvios de predição chegam a ser de até 2,00% do teor de aromáticos, em toda a faixa de concentração estudada. Nota-se uma amostra com um desvio maior. Na parte mais alta da curva, com concentrações maiores de aromáticos, percebe-se que os valores preditos tendem a ser menores que os valores de referência. Enfim, são necessários ajustes no modelo para que seja melhor utilizado nesta faixa de concentração.

6.4 TEOR DE OLEFINAS

Para a modelagem da característica teor de olefinas, novamente foram testados diversos pré-tratamentos dos dados espectrais, aplicando-se na sequência o algoritmo PLS. Os resultados encontram-se na Tabela 4.

As quantidades de variáveis latentes utilizadas nas regressões variou de 2 a 4. Os erros quadráticos de calibração e previsão foram menores e os coeficientes de correlação das curvas foram maiores quando se usou o tratamento *baseline*, mas ainda assim elegeu-se a primeira derivada como pré-tratamento dos espectros, devido ao menor número de variáveis latentes utilizados na PLS.

TABELA 4: Figuras de mérito obtidas por PLS após diversos pré-tratamentos dos dados para teor de olefinas

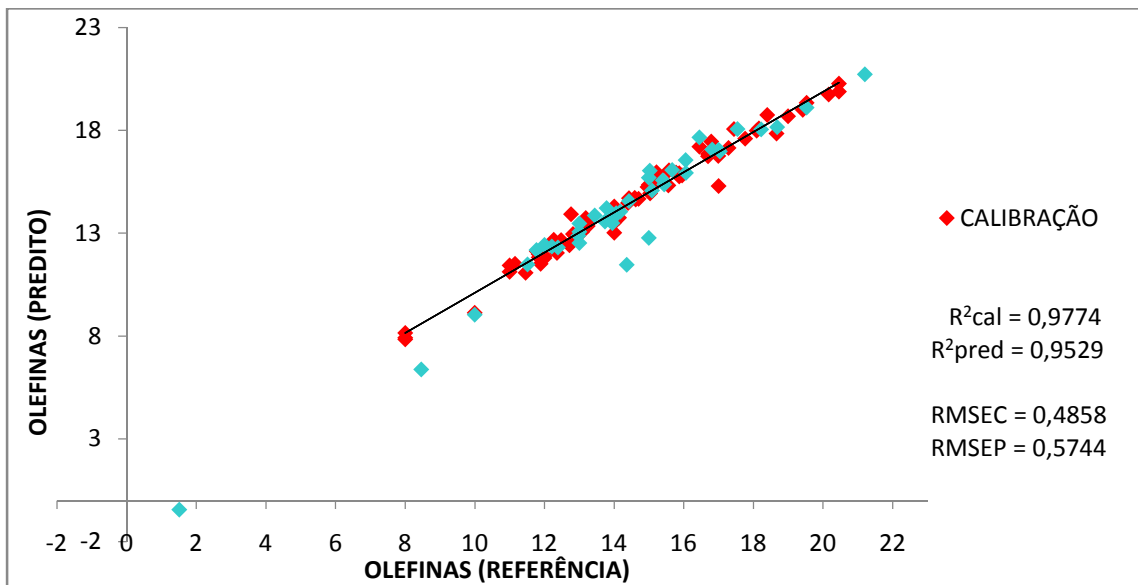
PRÉ- TRATAMENTOS	Nº VARIÁVEIS LATENTES	RMSEC	RMSECV	R²_{cal}	R²_{val}
1ª Derivada	2	0,4051	0,4245	0,9779	0,9757
<i>Baseline</i>	4	0,3683	0,3889	0,9819	0,9797
MSC	3	0,4811	0,5465	0,9702	0,9618
MSC + 1ª Derivada	2	0,4161	0,4411	0,9767	0,9739
MSC + <i>Baseline</i>	4	0,4325	0,4703	0,9759	0,9716
2ª Derivada	3	0,4549	0,4921	0,9734	0,9690

Fonte: o autor.

Após tratar os espectros com a primeira derivada, separou-se o grupo de amostras para calibração e o grupo menor de amostras para previsão dos teores de olefinas. A divisão entre as amostras foi feita de forma randômica. As amostras do conjunto de calibração possuem teores de olefinas entre 8,00 e 20,50% em volume.

No gráfico da Figura 15 estão os valores obtidos experimentalmente pela metodologia de referência versus os valores previstos pelo modelo das amostras de calibração (em vermelho) e das amostras de predição (em azul). O número de variáveis latentes utilizado na regressão foi 2. Obteve-se um coeficiente de correlação da curva de calibração igual a 0,9774, e da curva das amostras de predição, igual a 0,9529. O RMSEC encontrado foi 0,4858%, e o RMSEP, 0,5744%.

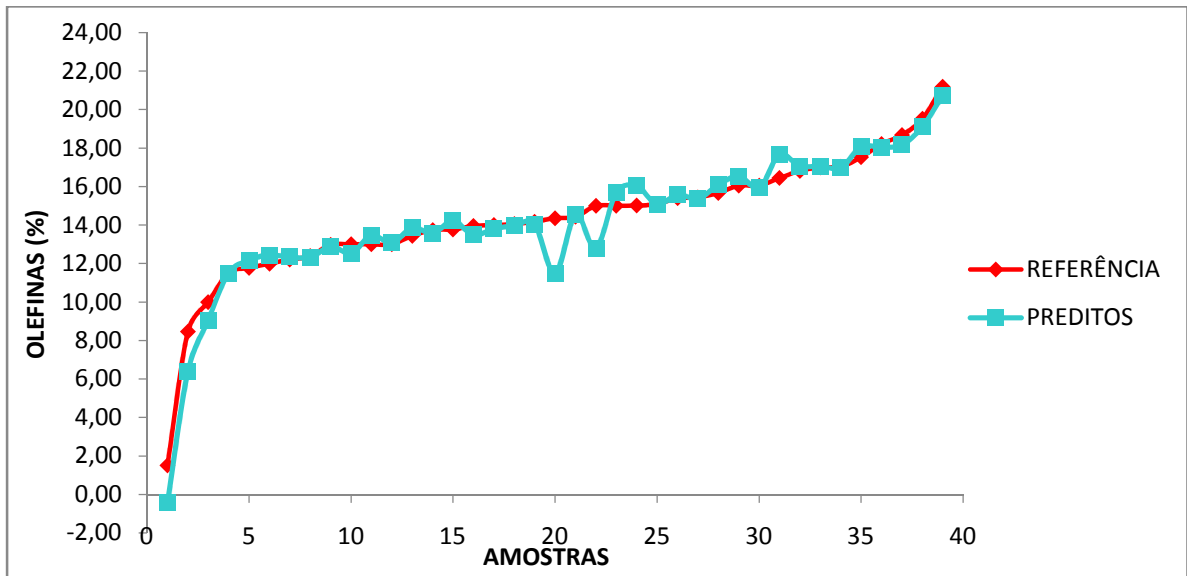
FIGURA 15: Gráfico dos teores de olefinas pelo método de referência versus valores preditos pelo modelo para amostras de calibração e predição



Diferentemente dos outros parâmetros modelados, as amostras do conjunto total possuem teores de olefinas melhor distribuídos, exceto por uma amostra com teor bem abaixo, que ficou no grupo de predição, aumentando o erro para concentrações mais baixas.

Observando a Figura 16 percebemos que os maiores desvios de predição encontram-se na região de mais baixo teor de olefinas, e em duas amostras de concentração intermediária. Essas amostras com comportamento diferenciado devem ser acompanhadas com atenção, pois elas tanto podem ser *outliers* como podem pertencer a outra população de dados, com características diferentes. Estes pontos discrepantes podem acontecer também devido a algum erro no equipamento de referência ou no espectrômetro, que acabaram por fornecer um sinal errôneo. Para o restante das amostras verifica-se que o desvio é baixo e varia sem apresentar tendências.

FIGURA 16: Gráfico comparativo dos valores de olefinas de cada amostra do conjunto de predição obtidos pelo método de referência e pelo modelo



Fonte: o autor.

7 AVALIAÇÃO DE CUSTOS

Nos últimos anos, a técnica de espectroscopia de infravermelho vem sendo largamente utilizada em vários setores da indústria, entre eles o petroquímico, para acompanhamento de produção e otimização de processos.

O método de referência para determinar o número de octano utiliza um motor específico, patenteado, que pode custar até mais de um milhão de reais. Soma-se a isto o gasto com peças e manutenção especializada. Aproximadamente quatro litros de misturas de isoctano e n-heptano de alta pureza são queimados por análise. O resultado analítico demora cerca de trinta minutos para ficar pronto, mas somente no caso de o motor já estar ligado e aquecido antes da análise; caso contrário, deve-se ligar o motor e aguardar cerca de duas horas para que as condições de análise se estabilizem, para aí então iniciar o procedimento com amostra e padrões. A análise inteira deve ser acompanhada por um técnico treinado, e ainda não existe possibilidade de automação.

A determinação da composição da gasolina (teor de aromáticos, olefinas e compostos saturados) por cromatografia gasosa parte da compra de um equipamento de aproximadamente cento e cinquenta mil reais. Os consumíveis do cromatógrafo também são bastante caros, a coluna de separação, principalmente. O gás de arraste utilizado nesta análise geralmente é o hélio de pureza elevada, que tem alto custo devido à dificuldade de obtenção. Além do hélio, são utilizados gases para a queima no detector (hidrogênio e ar sintético). A identificação dos componentes da gasolina é feita em comparação com padrões disponíveis no mercado que, além de caros, muitas vezes não representam a realidade dos produtos analisados, que são específicos para cada indústria. Além disso, para separar todos os componentes (ou quase todos), esta análise demora cerca de 3 horas para então gerar um resultado.

A análise por NIR, apesar de não estar isenta à compra do equipamento, que tem preço na mesma ordem de grandeza de um cromatógrafo, é muito mais simples, rápida e de menor custo para uso na rotina. Os métodos de análise por NIR não necessitam de padrões, apenas de algum solvente para limpeza do equipamento, conforme a aplicação, e em pequena quantidade. Operando o equipamento em bancada, a amostra é introduzida pela aplicação de uma pressão de ar comprimido, que é um gás de processo, produzido na própria indústria, sem custo adicional. Quando o equipamento é utilizado em linha, a amostra é arrastada pela própria pressão da linha, sem necessitar de gás de arraste. O resultado analítico não demora mais do que um minuto para ficar pronto. Ainda mais atrativo é o fato de poder gerar diversos

resultados, para inúmeros parâmetros diferentes, rápida e simultaneamente. Assim, um único analisador pode substituir vários analisadores tradicionais. Outra vantagem do uso do NIR é a possibilidade de análise remota, para o controle de processos das indústrias.

A Tabela 5 apresenta resumidamente os principais custos envolvidos na utilização de cada uma das técnicas de referência e da metodologia proposta neste trabalho.

TABELA 5: Comparativo entre custos e benefícios das metodologias de referência e da técnica proposta

MON - ASTM D2700	COMPOSIÇÃO - ASTM D6730	NIR + QUIMIOMETRIA
Equipamento específico, podendo custar mais de 1 milhão de reais	Cromatógrafo - aproximadamente 150 mil reais	Equipamento - aproximadamente 150 mil reais
Gasto alto com peças - 10% do valor do motor ao ano	Consumíveis caros.	Baixo custo com peças
Consumo alto de reagentes - 4 litros de isoctano/n-heptano por análise	Hélio de alta pureza como gás de arraste + hidrogênio e ar sintético no detector + nitrogênio líquido para baixar temperatura do forno + padrão	Não utiliza padrões, apenas solvente para limpeza. Ar comprimido sem custo.
30 minutos/análise + 2 horas de condicionamento = máximo 44 resultados/dia	3 horas por análise = máximo 8 resultados por dia	1 minuto/análise = 1440 análises/dia + diversos resultados/análise
24 horas de mão de obra	Análise automatizada, porém precisa de supervisão	Possibilidade de análise remota, sem gasto com mão de obra

Fonte: o autor.

8 IMPACTO AMBIENTAL

O MON é determinado através da queima da gasolina no motor, em comparação com a queima de misturas de reagentes-padrão. Logo, trata-se de uma análise destrutiva, com consumo de produto e de solventes. Por se tratar de queima, devemos pensar ainda que há a geração de gases tóxicos, que exigem um bom sistema de exaustão para minimizar a exposição ocupacional do operador, e, assim como nos motores automotivos, um sistema de escape que trate estes gases antes de irem para a atmosfera, sem causar impacto ambiental. A detonação da gasolina no motor é medida através de índices sonoros, logo, a análise expõe o operador a muito ruído, por muito tempo.

Na determinação da composição por cromatografia gasosa, a amostra é queimada no detector de ionização de chama, portanto, deve-se ter exaustão adequada para que os gases não fiquem confinados no ambiente, expondo os operadores ao risco de inalação.

A análise pelo NIR apresenta a característica de ser uma técnica não destrutiva e tem como vantagens a redução do tempo de análise, de custos, da exposição ocupacional e do impacto ambiental quando comparado às metodologias de referência. Quando implantada em linha de produção, não há a menor exposição do operador aos produtos e processos, e o resultado gerado continuamente e on-line facilita o melhor controle de poluição, ganhos de estabilidade e otimização no procedimento das misturas em linha, nas indústrias petroquímicas.

Na Tabela 6 são destacadas algumas características das metodologias de referência e da técnica proposta neste trabalho.

TABELA 6: Comparativo dos principais impactos das metodologias de referência e da técnica proposta

	MON - ASTM D2700	COMPOSIÇÃO - ASTM D6730	NIR + QUIMIOMETRIA
Quantidade de amostra	Elevada	Pequena	Pequena
Resíduo gerado	Elevado	Pequeno	Pequeno
Geração de gases	Sim	Sim	Não
Ruído	Muito elevado	Mínimo	Mínimo
Técnica destrutiva	Sim	Sim	Não

Fonte: o autor.

9 CONCLUSÕES

A espectroscopia de infravermelho próximo aliada às técnicas quimiométricas multivariadas PCA e PLS, possibilitou a construção de modelos matemáticos para os três parâmetros estudados: MON, teor de aromáticos e teor de olefinas. A qualidade dos modelos formados depende de alguns fatores, que em alguns casos mais do que em outros, não foram obedecidos.

Primeiramente, uma boa calibração depende das amostras analisadas, que devem conter todas as variações possíveis encontradas na rotina de análises da indústria. A composição da gasolina depende dos processos de refino utilizados, mas fundamentalmente depende da composição do petróleo destilado. Neste trabalho não foi possível abranger todas as variações de composição possíveis da gasolina, derivadas dos inúmeros tipos de petróleo distintos, portanto, podem haver amostras discrepantes devido à variabilidade do produto.

A calibração de um parâmetro deve abranger toda a faixa de concentração de interesse, com quantidades significativas de amostras em toda o intervalo. Um problema encontrado foi a distribuição não uniforme da quantidade de resultados nas faixas de calibração, aumentando o erro para as amostras que apresentarem valores mais limítrofes do intervalo. No caso do modelo para o MON, a grande maioria das amostras do conjunto analisado possuíam octanagem entre 82,0 e 83,7, tornando o modelo robusto e confiável para este range. Para a determinação de outros valores de octanagem utilizando o modelo, este deve ser alimentado com maior quantidade de amostras de octanagens diferentes. No caso do teor de olefinas, o modelo abrangeu a faixa de 1,00 a 20,00%, mas a maioria das amostras possuía concentração maior que 8,00%, o que pode levar a erros de predição maiores para teores mais baixos de compostos olefínicos.

Para a utilização de um modelo, ele deve abranger todas as condições que uma amostra pode apresentar. Como os modelos foram criados em laboratório, em temperatura constante e controlada, sem vibração, sem interferentes no processo, conclui-se que eles podem ser aplicados nesta condição, porém, para serem utilizados em linha, diversas variações das condições da gasolina no processo devem ser adicionadas aos modelos antes da sua implementação.

Por fim, a aplicabilidade de um modelo depende de uma atualização periódica, para manter a representatividade das amostras, que vão sendo alteradas ao longo do tempo.

Sugere-se, para a continuidade do trabalho, a análise de amostras cujo MON se encontre abaixo de 82,0 e amostras com teor de olefinas abaixo de 8,00% e acima de 20,00%,

possibilitando a previsão de resultados fora de especificação, assim como a revisão do modelo para o teor de aromáticos, que obteve erros muito altos e correlações muito ruins, o que inviabilizou a utilização da regressão criada.

REFERÊNCIAS

- 1 GAUTO, M. (Org.). **Petróleo e gás: princípios de exploração, produção e refino**. Porto Alegre: Bookman, 2016.
- 2 AGÊNCIA NACIONAL DO PETRÓLEO. **Resolução n. 40**, de 25 de outubro de 2013. Regula as especificações das gasolinas de uso automotivo.
- 3 FARAH, M. A. **Petróleo e seus derivados: definição, constituição, aplicação, especificações, características de qualidade**. Rio de Janeiro: LTC, 2012.
- 4 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM D 1319 – 15: Standard Test Method for Hydrocarbon Types in Liquid Petroleum Products by Fluorescent Indicator Adsorption**. West Conshohocken, PA, 2015.
- 5 TAKESHITA, E. V. **Adulteração de gasolina por adição de solventes: análise dos parâmetros físico-químicos**. 2006. 102 f. Dissertação (Mestrado em Engenharia Química) – Universidade Federal de Santa Catarina, Florianópolis, 2006.
- 6 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM D 6730– 1 (2016): Standard Test Method for Determination of Individual Components in Spark Ignition Engine Fuels by 100 Metre Capillary (with Precolumn) High-Resolution Gas Chromatography**. West Conshohocken, PA, 2016.
- 7 BRASIL, Ministério da Agricultura, Pecuária e Abastecimento. **Resolução nº 75**, de 05 de março de 2015. Fixa o percentual obrigatório de adição de etanol anidro combustível à gasolina. Diário Oficial da União, Brasília, DF, 2015.
- 8 BRASIL, N. Í. do, ARAÚJO, M. A. S., SOUZA, E. C. M. de. **Processamento de petróleo e gás: petróleo e seus derivados, processamento primário, processos de refino, petroquímica, meio ambiente**. 2. Ed. São Paulo: Grupo Gen-LTC, 2014.
- 9 MILHOR, C. E. **Sistema de desenvolvimento para controle eletrônico dos motores de combustão interna ciclo Otto**. 2002. 86 f. Dissertação (Mestrado em Engenharia Mecânica) – Universidade de São Paulo, São Carlos, 2002.
- 10 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM D 2700– 17: Standard Test Method for Motor Octane Number of Spark-Ignition Engine Fuel**. West Conshohocken, PA, 2017.

11 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM D 2699– 17:** Standard Test Method for Research Octane Number of Spark-Ignition Engine Fuel. West Conshohocken, PA, 2017.

12 WAUKESHA CFR. Disponível em <<http://www.waukeshacfr.com>>. Acesso em 20 de dezembro de 2017.

13 CFR F2U – Motor Para Determinação do Número de Octanagem. Pensalab. Disponível em <<https://www.pensalab.com.br>>. Acesso em 20 de dezembro de 2017.

14 SIMÃO, C. A. **Compilação de informações sobre a técnica NIR e quimiométrica para previsão de propriedades de combustíveis.** Comunicação Técnica. Centro de Pesquisas e Desenvolvimento Leopoldo A. Miguez de Mello, Rio de Janeiro, 1996.

15 HOLLER, F. J.; SKOOG, D. A.; CROUCH, S. R. **Princípios de Análise Instrumental.** 6. Ed. Porto Alegre: Bookman, 2009.

16 BURNS, D. A.; CIURCZAK, E. W. (Ed.). **Handbook of Near-infrared Analysis.** 6th. Ed. Boca Raton: CRC Press, 2007.

17 NIR SPECTRAL REGION - A guide to near-infrared spectroscopic analysis of industrial manufacturing processes. Metrohm. Disponível em <<https://www.metrohm.com/en/company/news/news-free-monographs>>. Acesso em 10 de dezembro de 2017.

18 FERREIRA, M. M. C. **Quimiometria: Conceitos, Métodos e Aplicações.** Campinas: Ed. Unicamp, 2015.

19 SENA, M. M. de; POPPI, R. J.; FRIGHETTO, R. T.; VALARINI, P. J. Avaliação do uso de métodos quimiométricos em análise de solos. **Química Nova.** São Paulo. v. 23, n. 4, p.547-556, 2000.

20 PARREIRA, T. F. **Utilização de métodos quimiométricos em dados de natureza multivariada.** 2003. 91 f. Dissertação (Mestrado em Química) – Universidade Estadual de Campinas, Campinas, 2003.

21 BORGES FILHO, D. **Aplicação de métodos quimiométricos para a classificação de águas minerais comercializadas em Goiânia-GO.** 2010. 46 f. Dissertação (Mestrado em Química) – Universidade Federal de Goiás, Goiânia, 2010.

22 SOUZA, A. M.; POPPI, R. J. Experimento didático de quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: um tutorial. **Química Nova**. São Paulo. v. 35, n. 1, p. 223-229, 2012.

23 CAMO ASA (1998). **User Manual The Unscrambler® v7.8**. Oslo, 1998.

24 MOITA NETO, J. M.; MOITA, G. C. Uma introdução à análise exploratória de dados multivariados. **Química Nova**. São Paulo, v. 21, n. 4, p. 467-469, 1998.

25 BUENO, A. F. **Caracterização de petróleo por espectroscopia no infravermelho próximo**. 2004. 112 f. Dissertação (Mestrado em Química) – Universidade Estadual de Campinas, Campinas, 2004.

26 MENDES, G.; ALEME, H. G.; BARBEIRA, P. J. S. Determination of octane numbers in gasoline by distillation curves and partial least squares regression. **Fuel**. Belo Horizonte, v. 97, p. 131-136, 2012.