

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA
DEPARTAMENTO DE QUÍMICA ANALÍTICA

ANÁLISE QUANTITATIVA POR TÉCNICAS DE REFLECTÂNCIA E
FOTOACÚSTICA NO INFRAVERMELHO MÉDIO COM TRANSFORMADA
DE FOURIER E UTILIZAÇÃO DE CALIBRAÇÃO MULTIVARIADA.

Melissa Umata Lucato

Tese de Doutorado

Orientador: Celso Ulysses Davanzo

CAMPINAS

2005

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE QUÍMICA DA
UNICAMP**

L933a

Lucato, Melissa Umata.

Análise quantitativa por técnicas de reflectância e fotoacústica no infravermelho médio com transformada de Fourier e utilização de calibração multivariada / Melissa Umata Lucato. -- Campinas, SP: [s.n], 2005.

Orientador: Celso Ulysses Davanzo.

Tese – Universidade Estadual de Campinas, Instituto de Química.

1. Infravermelho. 2. Combustíveis. 3. PET.
4. Calibração multivariada. I. Davanzo, Celso Ulysses.
II. Instituto de Química. III. Título.

Título em inglês: Quantitative analysis by reflectance and photoacoustic techniques in mid-infrared with Fourier transform and multivariate calibration.

Palavras-chave em inglês: Infrared, Fuels, PET, Multivariate calibration.

Área de concentração: Química Analítica.

Titulação: Doutora em Química Analítica.

Banca examinadora: Patrício Guilherme Peralta Zamora, Noemi Nagata, Célio Pasquini, Roy Edward Bruns.

Data de defesa: 20/04/2005.

DEDICATÓRIA

“ Quanto mais eu sinta, quanto mais eu sinta como várias pessoas,
Quanto mais personalidades eu tiver,
Quanto mais intensamente, estridentemente as tiver,
Quanto mais simultaneamente sentir com todas elas,
Quanto mais unificadamente diverso, dispersadamente atento,
Estiver, sentir, viver, for,
Mais possuirei a existência total do universo,
Mais completo serei pelo espaço inteiro fora.
Mais análogo serei a Deus, seja ele quem for,
Porque, seja ele quem for, com certeza que é Tudo,
E fora d'Ele há só Ele, e Tudo para Ele é pouco.”

AFINAL, de Álvaro Campos (Fernando Pessoa)

Este trabalho é dedicado aos meus avós Nilton e Elza Lucato,
símbolo de amor à vida, trabalho, dedicação e esperança.

E dedico também, de forma especial ao meu amor José Renato,
cuja força e determinação foram imprescindíveis para
a realização deste trabalho.

AGRADECIMENTOS

- ❖ Em especial ao meu orientador Celso Davanzo pela dedicação, paciência e amizade;
- ❖ Aos meus tios Júlia e Haruo e família, que me acolheram em Campinas;
- ❖ Aos amigos de grupo: Cláudio Trasferetti, Ricardo Rosseto e Marco Ferrão;
- ❖ A amiga Iara Messerschmidt pelo incentivo e amizade;
- ❖ Aos amigos em Campinas: Sílvia Lucas, Débora Biloti, Rita de Cássia, Milene, Edeilza, Ana Paula, Kesley, Neife e Robson;
- ❖ Aos professores: Patrício Zamora e Noemi Nagata.
- ❖ As três Márcias, pela imensa contribuição a este trabalho: Márcia Spinacé, Márcia Breitzkreitz e Márcia C.Q. Antonelli.
- ❖ A todos os funcionários e professores do IQ da UNICAMP;
- ❖ Aos funcionários da Central Analítica da UNICAMP;
- ❖ Ao meu namorado José Renato e família, pelo apoio e amor;
- ❖ As famílias Umata e Lucato, em especial a minha mãe Helena e ao meu irmão João Paulo;
- ❖ Aos amigos: Kátia, Betânia, Analúcia, Jeane, Mariana, Joseane, Alexandre Emmel, Alcindo, Marcelo Lima, Orlando, Alessandra, Joseane, Angelita, Paula, Douglas, Maurício, Paulo Janissek, Patrícia Murta, Marcelo Barros, Dartagnan entre outros;
- ❖ CNPq, CTPETRO, FINEP.
- ❖ A todos que contribuíram de alguma forma para este trabalho;

CURRICULUM VITAE

Melissa Umata Lucato

A-DADOS PESSOAIS

Data de nascimento: 20/03/73

Local de nascimento: Araraquara- São Paulo

RG: 5 195 711-3 (PR)

CPF: 963355569 87

Carteira de Trabalho: 91812 série: 00043 PR

Título Eleitoral: 517629106/98 da 176 Zona Eleitoral de Curitiba Seção 0056

B-EDUCAÇÃO

Doutorado em Ciências, área Química Analítica

Universidade Estadual Campinas UNICAMP (Campinas – SP)

Título: “Análise Quantitativa por Técnicas de Reflectância e Fotoacústica no Infravermelho com Transformada de Fourier com a utilização de Calibração Multivariada”

Mestrado em Ciência e Engenharia de Materiais

Defesa da Dissertação: 19/02/2000

Universidade Federal do Paraná UFPR (Curitiba- PR)

Título: “Sistema para Análise de Subprodutos do Gás SF₆ por Espectroscopia no Infravermelho com Transformada de Fourier (FTIR)”

Graduação em Química

Período: 1992-1996

Bacharelado em Química

Universidade Federal do Paraná (Curitiba-PR)

C-EXPERIÊNCIA DIDÁTICA

Professora Colaboradora (Disciplinas de Química)

Departamento de Ciências UNICENTRO (Campus Irati)

Horas/aula: 14h semanais

Início: junho de 2004

Programa de Estágio Docência nível II: Disciplina Química Analítica

Instituto de Química UNICAMP

Duração: 72 horas aula

Período: 2º Semestre de 2001.

D-TRABALHOS APRESENTADOS EM CONGRESSOS

“Determinação de parâmetros de qualidade de óleo diesel empregando espectroscopia no infravermelho próximo e médio”.

27ª Reunião anual da sociedade Brasileira de Química (Poços de Caldas- MG)

Período: 30 de maio a 2 de junho de 2004

“Determinação da densidade e índice de cetano do óleo diesel B (metropolitano) por FTIR-ATR e PLS

26ª Reunião Anual da Sociedade Brasileira de Química- Poços de Caldas- MG

Período: 26 a 39 de maio de 2003

“Determinação da Viscosidade Intrínseca e Grupos Carboxílicos Terminais de PET Reciclado por FTIR-PAS E PLS.

XV Congresso Brasileiro de Engenharia e Ciência de Materiais, Natal-RN

Período: 9 a 13 de novembro de 2002.

“Planejamento e Calibração Multivariada de Misturas Ternárias de Açúcares em dados de FTIR-ATR.”

24^a Reunião Anual da Sociedade Brasileira de Química- Poços de Caldas- MG

Período: 28 a 31 de maio de 2001.

E-PUBLICAÇÕES

ARTIGO SUBMETIDO

“Intrinsic viscosity determination of poly(ethylene terephthalate) using a multivariate calibration method’

RESUMO

O presente trabalho teve como principal objetivo a elaboração de metodologias analíticas que fornecessem informações diretas referentes aos problemas analisados, utilizando simultaneamente métodos não destrutivos de obtenção de espectros (FTIR) e calibração multivariada (PLS) com seleção de variáveis (Algoritmo Genético e Algoritmo das Projeções Sucessivas). Para efetivar tal objetivo foram estudados dois diferentes sistemas analíticos:

I-Análise dos combustíveis comerciais (diesel, gasolina e álcool etílico automotivo): Foram analisadas propriedades de amostras comerciais dos combustíveis como: densidade, tipos de hidrocarbonetos, curva de destilação, propriedades de octanagem, entre vários outros, a partir de espectros de FTIR-ATR. Os valores de referência foram obtidos a partir de normas ASTM, que foram correlacionados com resultados da calibração multivariada através do método PLS e MLR, com e sem seleção de variáveis. Os erros obtidos apresentaram valores inferiores aos da reprodutibilidade do método padrão para a maioria das propriedades. Além disso, os resultados entre os modelos propostos foram comparados estatisticamente pelo teste F.

II -Estudo das propriedades de degradação do PET , Poli(tereftalato de etileno): Foram determinadas a viscosidade intrínseca $[\eta]$ e a concentração de grupos carboxílicos terminais (CEG) através da técnica de refletância difusa e técnica fotoacústica no infravermelho com calibração multivariada (PLS). Os modelos foram comparados estatisticamente pelo teste-F.

Os resultados apontaram um método alternativo às laboriosas análises convencionais que empregam reagentes tóxicos, gerando resíduos.

ABSTRACT

This research project was mainly intended to elaborate analytical methodologies that supplied referring direct information to the analyzed problems, using simultaneously non destructive methods of attainment of spectroscopy (FTIR) and multivaried calibration (PLS) with variable selection (Genetic Algorithm and Successive Projections Algorithm). To accomplish this objective, two different analytical systems had been studied:

I- *Analysis of commercial fuels (diesel, gasoline and automotive etilic alcohol)*: - Properties of commercial samples of fuels had been analyzed as: density, types of hydrocarbons, curve of distillation, octane number properties, among several others, from spectra of FTIR. The foundation values had been gotten from ASTM norms, that had been correlated with results of the multivariate calibration through PLS and MLR methods, using or not variable selection. The obtained errors had presented inferior values to the ones for reproductibilidade of the standard method for the majority of the properties. Moreover, the results among the considered models were statistically compared by the F test .

II- *Study of the Properties of PET's degradation, Poly(ethylene therphthalate)*: - Intrinsic viscosity $[\eta]$ and the concentration of terminal carboxylic groups had been determined (CEG) by the technique of diffuse reflectance and photoacoustic technique in the infrared region with multivaried calibration (PLS). The models were statistically compared by the F test.

The results point to an alternative method for the laborious conventional analyses that use toxic reagents and generates residues.

ÍNDICE

1-INTRODUÇÃO E OBJETIVOS.....	01
2-TÉCNICAS ESPECTROSCÓPICAS NO INFRAVERMELHO.....	04
2.1-Introdução	04
2.2- Reflectância Total Atenuada (ATR)	05
2.3- Reflectância Difusa (DRIFTS).....	10
2.4- Espectroscopia Fotoacústica (PAS)	14
3- APLICAÇÕES EM SISTEMAS ANALÍTICOS	19
3.1- Introdução	19
3.2- APLICAÇÃO I: Análise de Combustíveis Comerciais	19
3.2.1- Características gerais da gasolina	19
3.2.2- Especificações técnicas para a gasolina	21
3.2.3- O significado e a importância das especificações para a gasolina.....	22
3.2.4- Características gerais do diesel	26
3.2.5- Especificações técnicas para o diesel	27
3.2.6- O significado e a importância das especificações para o diesel	28
3.2.7- Características gerais do álcool automotivo	33
3.2.8- Especificações técnicas para o álcool automotivo.....	33
3.2.9-O significado e a importância das especificações para o álcool automotivo	33
3.3- Análise espectroscópica dos combustíveis.....	34
3.4- APLICAÇÃO II: Estudo das Propriedades de Degradação do PET	38
3.4.1- O PET e Processos de Degradação	38

3.4.2- Propriedades Indicadoras da Degradação do PET e sua Análise Convencional.....	40
3.5- Análise da Degradação do PET por Técnicas Espectroscópicas	41
4- PARTE EXPERIMENTAL	45
4.1- Determinação dos Parâmetros dos Combustíveis.....	45
4.1.1- Obtenção dos Valores de Referência	45
4.1.2- Obtenção dos Espectros FTIR dos Combustíveis	47
4.2- Determinação das Propriedades de Degradação do PET.....	47
4.2.1- Obtenção das Amostras de PET	47
4.2.2- Determinação da Concentração de Grupos Carboxílicos Terminais (CEG).....	48
4.2.3- Determinação da Viscosidade Intrínseca $[\eta]$	50
4.2.4- Obtenção dos Espectros FTIR das amostras de PET.....	51
5- ANÁLISE MULTIVARIADA.....	53
5.1- Introdução	53
5.2- Pré-tratamento de Dados.....	54
5.2.1- Centrar na Média	55
5.2.2- Diferenciação e Alisamento	55
5.3- Análise por Componentes Principais (PCA)	57
5.4-Calibração Multivariada.....	61
5.4.1- Regressão Linear Múltipla (MLR).....	62
5.4.2- Mínimos Quadrados Parciais (PLS).....	64
5.5- Validação dos Modelos.....	66
5.5.1- Escolha do Número de Variáveis Latentes ou Fatores	66

5.5.2- Avaliação dos Erros de Previsão	69
5.5.3- Detecção de Amostras Anômalas.....	70
6- SELEÇÃO DE VARIÁVEIS	72
6.1- Introdução	72
6.2- Algoritmo Genético	73
6.3-Algoritmo das Projeções Sucessivas	79
7-RESULTADOS E DISCUSSÕES	83
7.1- Resultados da Aplicação I: Análise de Combustíveis	83
7.1.1- Espectros FTIR-ATR dos Combustíveis.....	83
7.1.2- Análise Exploratória através do Método PCA	86
7.1.3- Modelos de Calibração para as Propriedades dos Combustíveis	93
7.1.3.1- Modelos PLS e MLR para as propriedades do diesel.....	97
7.1.3.2- Modelos PLS e MLR para as propriedades da gasolina.....	128
7.1.3.3- Modelos PLS e MLR para as propriedades do álcool automotivo	176
7.1.4- Coeficientes lineares e angulares para os Resultados de Previsão dos Combustíveis.....	187
7.1.5- Comparação entre os modelos de calibração através do Teste-F	191
7.1.5-1 Comparação entre os modelos de calibração do diesel	191
7.1.5-2 Comparação entre os modelos de calibração da gasolina.....	195
7.1.5-3 Comparação entre os modelos de calibração do álcool	200
7.1.6- Atribuição dos Números de Onda Seleccionados pelos Métodos AG e APS	203

7.2- Resultados da Aplicação II: Estudo das Propriedades de Degradação do PET	216
7.2.1-Análise de Espectros DRIFTS e Fotoacústica.....	220
7.2.3- Construção dos Modelos PLS	220
7.2.3-1 Modelos PLS para a determinação de CEG	220
7.2.3-2 Modelos PLS para a determinação de $[\eta]$	228
7.2.4-Atribuição das Regiões Espectrais	234
 8-CONCLUSÕES	 236
9-REFERÊNCIAS BIBLIOGRÁFICAS	239
APÊNDICE A	251
APÊNDICE B	253

LISTA DE ABREVIATURAS

ABNT - Associação Brasileira de Normas Técnicas

AEAC - Álcool Etílico Anidro Combustível

AEHC - Álcool Etílico Hidratado Combustível

AG - Algoritmo Genético

ANP - Agência Nacional do Petróleo

APS - Algoritmo das Projeções Sucessivas

ASTM- American Society for Testing and Materials

ATR-(Attenuated Total Reflectance)- Reflectância Total Atenuada

BTEX - benzeno, tolueno, etilbenzeno e xilenos

CEG - (Carboxilic End Group)- Grupos Carboxílicos Terminais

DRIFTS- (Diffuse Reflectance Infrared Fourier Transform Spectroscopy) - Reflectância Difusa

FTIR-(Fourier Transform Infrared Spectroscopy)- Espectroscopia no Infravermelho com Transformada de Fourier

INPM - Instituto Nacional de Pesos e Medidas

MID - (Mid Infrared Spectroscopy)- Espectroscopia no Infravermelho Médio

MLR - (Multiple Linear Regression)- Regressão Linear Múltipla

MON - (Motor Octane Number)

NBR - Normas Brasileiras

NIPALS: (Nonlinear Iterative Partial Least Squares)

NIR - (Near Infrared Spectroscopy)- Espectroscopia no Infravermelho Próximo

PAS-(Photoacoustic Spectroscopy)- Espectroscopia Fotoaústica

PCA - (Principal Component Analysis)- Análise dos Componentes Principais

PLS - (Partial Least Squares)- Mínimos Quadrados Parciais

PONA - parafinas, oleofinas, naftênicos e aromáticos

PRESS- (Predictive Residual Errors Sum of Square)- Soma dos Quadrados dos Erros de Previsão

RMSEP- (Root Mean Square Error Prediction)- Raiz Quadrada dos Erros Médios de Previsão

RON – (Research Octane Number)

1-INTRODUÇÃO E OBJETIVOS

O presente trabalho beneficia-se da versatilidade e agilidade da utilização da espectroscopia no infravermelho médio com Transformada de Fourier (FTIR-MID) em sistemas analíticos. A aplicabilidade desta técnica está relacionada ao crescente aprimoramento tecnológico de acessórios que são uma alternativa aos métodos tradicionais de transmissão, permitindo uma metodologia de análise com pouca ou nenhuma preparação de amostra^[1-2].

Os métodos tradicionais de análise por FTIR (transmissão) necessitam de um processo prévio de preparação de amostras como a confecção de discos prensados de KBr contendo o analito, dissolução da amostra em solventes orgânicos para a obtenção de filmes e diluição prévia de líquidos muito absorventes. Além disso, a maioria dos processos de preparação de amostra envolve maior tempo de análise, possíveis erros de preparação, riscos de alteração da amostra além de apresentar caráter destrutivo.

Métodos que utilizam a região do infravermelho médio, mas que não envolvem processos de transmissão têm sido utilizados como técnicas de reflectância e fotoacústica com o auxílio de acessórios específicos. Estes acessórios são uma boa alternativa aos métodos tradicionais de transmissão, pois otimizam o tempo de análise e podem ser empregados com sucesso em análises qualitativas e quantitativas^[2].

Com o advento da *Quimiometria*, as técnicas alternativas de FTIR tornaram-se mais eficientes e precisas, uma vez que os métodos quimiométricos possibilitam a análise de misturas complexas (multicomponentes), cujas bandas espectrais apresentam-se sobrepostas ou com presença de interferentes. A aplicação da calibração multivariada resulta na otimização do processo de análise e fornece

modelos de previsão com boa capacidade preditiva ^[3-4]. Cabe ainda ressaltar que estes métodos de calibração podem ser otimizados com o processo de seleção de variáveis, no qual um subconjunto de números de onda com informações realmente pertinentes ao sistema analisado é selecionado e as informações nulas ou redundantes são descartadas. Deste modo, o processo de regressão pode ser efetuado com maior rapidez e precisão.

Em vista das considerações acima mencionadas, o presente trabalho tem como principal objetivo a elaboração de metodologias analíticas eficazes e rápidas que forneçam informações diretas referentes ao problema analisado, utilizando simultaneamente métodos não destrutivos de obtenção de espectros e calibração multivariada com seleção de variáveis.

Para efetivar tal objetivo foram estudados dois diferentes sistemas analíticos, considerando-se o tipo de amostra:

I Análise dos combustíveis comerciais (diesel, gasolina e álcool etílico automotivo).

II- Estudo das propriedades de degradação do PET Poli(tereftalato de etileno).

Em relação à aplicação I, pode-se afirmar que combustíveis como gasolina, diesel e álcool precisam ter uma qualidade mínima, de acordo com sua aplicação, para que a energia fornecida por estes seja aproveitada da melhor forma possível, obedecendo às exigências ambientais. Adicionalmente, as propriedades físico-químicas dos combustíveis devem estar dentro das especificações técnicas ^[5]. No Brasil, os parâmetros dos combustíveis devem obedecer à legislação estabelecida (determinada pela agência Nacional do Petróleo, ANP). Os parâmetros que avaliam a qualidade dos combustíveis são determinados segundo procedimentos descritos por normas ASTM (American Society for Testing and Materials) e ABNT (Associação Brasileira de Normas Técnicas). Estes métodos de análise são

laboriosos, caros e demorados. Contudo, a aplicação de métodos espectroscópicos combinados à calibração multivariada pode otimizar o estudo da qualidade da gasolina, álcool e diesel.

Neste sentido, amostras de combustíveis comerciais foram estudadas através da determinação de seus parâmetros com a utilização da técnica de reflectância total atenuada (ATR), através do acessório para amostras líquidas *circle cell*[®], sem qualquer diluição prévia. Para tal, utilizou-se modelos de previsão para as propriedades dos combustíveis elaborados a partir do método PLS 1 (Partial Least Squares) com seleção de variáveis através dos métodos AG (Algoritmo Genético) e APS (Algoritmo das Projeções Sucessivas).

Considerando-se a aplicação II, foi observada que a determinação do grau de degradação do PET, Poli(tereftalato de etileno) é de extrema importância para a indústria de reciclagem, pois assim é possível direcionar sua aplicação^[6]. A concentração de *grupos carboxílicos terminais* (CEG) e *viscosidade intrínseca* $[\eta]$ são propriedades diretamente relacionadas com o grau de degradação do material. Os métodos químicos tradicionais utilizados para a determinação de CEG e $[\eta]$ consomem tempo, geram resíduos tóxicos e necessitam de reagentes caros.

O presente trabalho apresenta a determinação de CEG e $[\eta]$ através da técnica de reflectância difusa “DRIFTS” (Diffuse Reflectance Infrared Fourier Transform Spectroscopy) e espectroscopia fotoacústica PAS (Photoacoustic Spectroscopy) no infravermelho médio, sem a necessidade da confecção de filmes ou discos de KBr e com aplicação de calibração multivariada. Para tal utilizou-se a técnica de calibração multivariada PLS. Deste modo a metodologia de análise proposta pode ser facilmente implementada em análises de rotina.

2-TÉCNICAS ESPECTROSCÓPICAS NO INFRAVERMELHO

2.1 - INTRODUÇÃO

Espectroscopia é o estudo da interação da radiação eletromagnética com a matéria. As moléculas podem interagir com os campos elétrico e magnético da luz e absorver uma certa quantidade de energia suficiente para promover esta molécula de um nível discreto de energia a outro nível de energia superior (transição energética). A energia relacionada com *transições vibracionais* é da ordem de 1,2 a 60 kJ mol⁻¹ (0,3 a 15 kcal mol⁻¹) e a faixa do espectro eletromagnético em números de onda encontra-se a 100-5000 cm⁻¹, correspondendo à região do infravermelho. A Espectroscopia no Infravermelho faz parte de um sistema geral que é a Espectroscopia Vibracional.^[7]

Os espectros na região do infravermelho médio correspondem à faixa espectral de 4000 a 400 cm⁻¹, na qual observa-se principalmente transições fundamentais dos estados vibracionais de energia. Em geral este tipo de espectro produz picos bem definidos que podem ser atribuídos diretamente a espécies individuais e grupos funcionais, o que faz da técnica uma importante ferramenta qualitativa.

O aprimoramento na instrumentação trouxe como consequência a melhora da resolução dos espectros e na sensibilidade e rapidez da técnica, tornado-a apta para a aplicação em análises quantitativas. Adicionalmente, o desenvolvimento de acessórios que utilizam técnicas de reflexão e fotoacústica facilitou o processo de preparação de amostras, tornado a análise ainda mais ágil.

Dentre as diversas técnicas relacionadas com a espectroscopia no infravermelho, foram exploradas neste trabalho as técnicas de Reflexão Difusa (DRIFTS), Reflexão Total Atenuada (ATR), e Espectroscopia Fotoacústica (PAS).

2.2 – REFLECTÂNCIA TOTAL ATENUADA

A espectroscopia de ATR é indicada para materiais muito absorventes (em relação à radiação) ou amostras muito espessas, cuja análise por técnicas de transmissão é inviável devido à saturação do sinal. É muito aplicada no estudo de superfícies e pode ser empregada em análises quantitativas com pouco ou nenhum preparo de amostras^[8].

A reflexão interna ou refletância total atenuada está relacionada com a reflexão interna de um elemento ótico (cristal ou elemento de ATR) que apresenta um índice de refração elevado^[8]. O fenômeno ocorre quando a radiação entra em um cristal de ATR e é refletida totalmente. Esta reflexão origina uma onda evanescente que se estende até a interface da amostra que está em contato direto com a superfície do elemento de ATR. As propriedades do campo evanescente são ilustradas na figura 2.1:

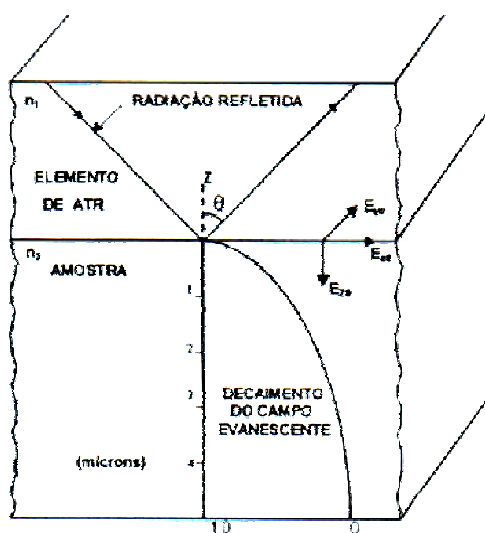


Figura 2.1: Elemento de ATR em contato com amostra^[8].

A radiação infravermelha propaga-se no meio 1 com índice de refração n_1 e sofre reflexão interna na interface com o meio 2 de índice de refração n_2 quando o ângulo de incidência θ excede o ângulo crítico θ_C . O ângulo de incidência θ é medido a partir da normal e o ângulo crítico é definido a partir da relação:

$$\theta_C = \sin^{-1} n_{21} \quad (2.1)$$

Onde $n_{21} = n_2/n_1$.

Ao penetrar o campo evanescente a intensidade da radiação decai exponencialmente com a distância do elemento de ATR/interface da amostra. A propagação do campo evanescente no meio 2 ocorre em determinadas condições:

- 1- A intensidade do campo evanescente no meio 2 não é zero, mas existe um componente normal instantâneo de fluxo de energia no meio 2, cuja média no tempo é zero. Deste modo, não há perda de energia e a propagação da radiação no meio 2 é totalmente refletida no interior do elemento de ATR ;
- 2- O campo evanescente é uma onda não transversa e tem componentes em todas as direções espaciais (E_x0 , E_y0 , E_z0);
- 3- O campo evanescente é restrito à vizinhança da superfície do meio 2 e tem sua intensidade diminuída em função da distância do meio 2 normal à superfície (ao longo do Eixo z, Figura 2.1);

A profundidade de penetração é definida como a distância na direção z em que ocorre a diminuição pela metade do campo evanescente. Para isto, tem-se a equação 2.2:

$$D_p = \frac{\lambda}{2\pi n_1 (\sin^2 \theta - (n_2/n_1)^2)^{1/2}} \quad (2.2)$$

Na qual D_p é a profundidade de penetração; λ é o comprimento de onda; n_2 e n_1 são respectivamente os índices de refração da amostra e do elemento ATR; θ é o ângulo de incidência.

A equação 2.2 mostra a dependência entre a profundidade de penetração e o comprimento de onda. Como consequência, o espectro resultante tem as intensidades das bandas aumentadas com o acréscimo do comprimento de onda (diminuição do número de onda). No entanto, correções destas intensidades podem ser efetuadas matematicamente com auxílio de ferramentas disponíveis no próprio programa de aquisição de espectros.

Vários tipos de materiais são utilizados como elementos de ATR. A escolha do elemento a ser utilizado está relacionada com faixa espectral, índice de refração, robustez e preço. Inicialmente o elemento de ATR mais utilizado era um material chamado KRS-5, constituído de uma mistura eutética de iodeto/brometo de tálio. Este elemento possui uma excelente faixa espectral de uso, no entanto é bastante tóxico, parcialmente solúvel em água e sua superfície apresenta certa plasticidade. Devido a estas desvantagens, o KRS-5 tem sido substituído por materiais mais modernos conforme os exemplos encontrados na tabela 2.1^[8]:

Tabela 2.1: Exemplos de elemento de ATR e respectivos índices de refração:

<i>Elemento ATR</i>	<i>Faixa espectral (cm^{-1})</i>	<i>material</i>	<i>Índice de refração n 1000 cm^{-1}</i>
ZnSe	20000 a 650	seleneto de zinco	2,4
AMTIR	11000 a 625	vidro de GeAsSe	2,5
Si	8300 a 660	silício	3,4
Ge	5500 a 600	germânio	4,0

Os acessórios disponíveis no mercado para a espectroscopia de ATR possuem diferentes configurações, com elemento de ATR com configuração horizontal ou na forma cilíndrica ou bote (barco) e sondas com fibras óticas, que podem ser utilizados na análise não somente de sólidos estendidos, mas também para pastas, materiais particulados e líquidos.

As vantagens da utilização da técnica de ATR em relação às técnicas de transmissão são claramente evidenciadas quando a análise é relacionada com amostras líquidas. Uma das principais dificuldades da análise por transmissão é a análise de amostras aquosas ou outros líquidos muito absorventes. Por serem fortemente absorventes, as bandas características da água podem sobrepor bandas relacionadas com o analito de interesse, além de ocasionar a saturação do sinal. Existem no mercado acessórios da técnica de transmitância com celas de espaçamento reduzido ($6\mu\text{m}$ de espessura) para minimizar o efeito da saturação, no entanto a linearidade dos detectores e a reprodutibilidade da análise ficam comprometidas^[9].

Em contrapartida, a técnica de ATR possui um acessório especialmente desenvolvido para análise de líquidos, a CIRCLE CELL[®], cujo elemento de ATR é cilíndrico.

Devido à pequena profundidade de penetração da técnica de ATR, os efeitos da saturação são diminuídos. Adicionalmente, a geometria cilíndrica do acessório permite um contato efetivo entre amostras e elemento de ATR. Deste modo, o caminho ótico é fixo e bastante reproduzível. Além disso, contrariamente às celas de transmissão, é pouco sensível às variações de temperatura e pressão.

A figura 2.2 mostra exemplos de acessórios de ATR com geometria horizontal e cilíndrica:

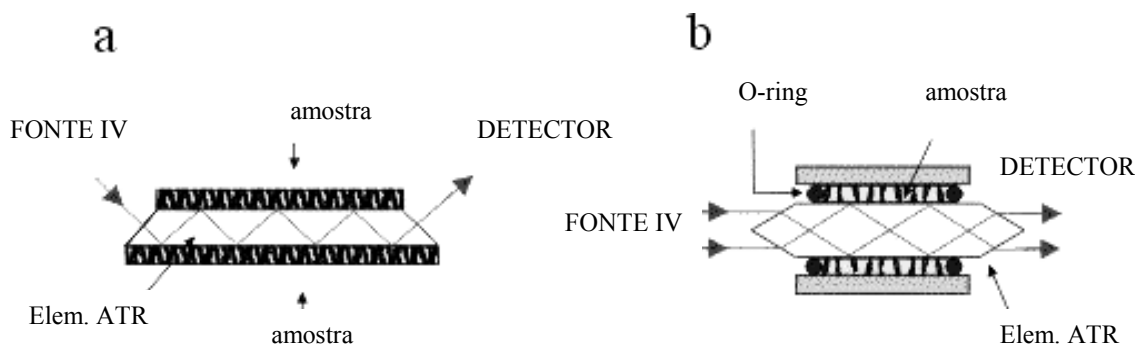


Figura 2.2: Esquema de acessório de ATR. (a) Geometria horizontal. (b) Geometria cilíndrica (Circle Cell)^[9].

Por ser de fácil manuseio e não exigir a diluição da amostra, o acessório circle cell tem sido aplicado em uma ampla diversidade de sistemas analíticos de interesse industrial como análise de açúcares em refrigerante e sucos de frutas^[10] e análise de metanol e metil- terc-butil éter em amostras de gasolinas comerciais^[11].

A técnica também tem sido empregada em análises clínicas: como a determinação de uréia, creatinina, ácido úrico, albumina e glicose em amostras de urina^[12].

Além das aplicações em análises quantitativas, a técnica tem sido utilizada para a determinação de constantes óticas^[13].

2.3 - REFLECTÂNCIA DIFUSA (DRIFTS)

A técnica de reflexão difusa é geralmente empregada na análise de amostras com grande área superficial (particulados ou fragmentados na forma de pós), sem a necessidade de uma preparação laboriosa de amostras, com a possibilidade de analisar materiais pouco ou não reflexivos (opacos). O fenômeno da reflexão não é observado estritamente na superfície da amostra, uma vez que ocorrem interações entre a matriz e a radiação incidente ^[2]. Quando a radiação infravermelha incide sobre a amostra, uma parte é refletida prontamente na sua superfície (reflexão especular): o ângulo de incidência e o ângulo refletido pela superfície são equivalentes. Uma segunda parte da radiação incidente, penetra a superfície da amostra e além de ser parcialmente absorvida, sofre múltiplos espalhamentos nas interfaces das partículas individuais. Em seguida o feixe retorna à superfície em ângulos aleatórios, caracterizando deste modo, o processo da reflexão difusa. Em uma situação *ideal*^[14], ao penetrar a superfície do material, a radiação interage com o interior da amostra, com uma profundidade de 1 a 3 mm, com *um mínimo de quantidade do componente especular*, deixando a superfície em vários ângulos diferentes, conforme representado na figura 2.3:

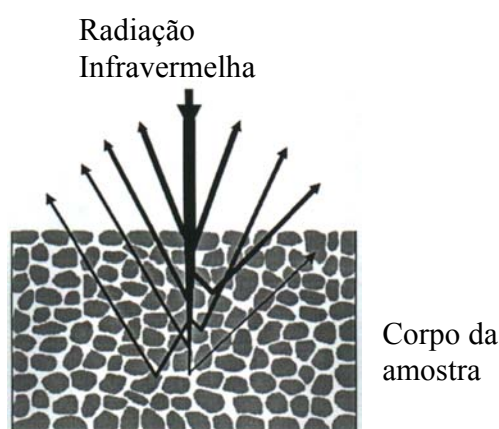


Figura 2.3: Representação da reflexão difusa em um sistema ideal^[14].

No entanto, em situações rotineiras a reflexão difusa ocorre em sistemas não ideais, simultaneamente com a reflexão especular, na qual a radiação refletida contém tipicamente um componente com maior índice de refração, produzindo distorções nos espectros de reflexão difusa. Esta distorção é indesejável, pois pode comprometer a interpretação do espectro, além de afetar a linearidade em análises quantitativas.

A mistura dos efeitos reflectivos em um espectro DRIFTS, é influenciada pela natureza da amostra e o modo como a amostra foi preparada. Amostras altamente absorventes, como os compostos inorgânicos, tendem a restringir a propagação da radiação através da amostra. Como consequência as reflexões acontecem mais próximo da superfície, ocasionando distorções. Em contrapartida, alguns compostos orgânicos fracamente ou moderadamente absorventes não apresentam espectros com grande distorção devido à reflexão especular^[14].

A radiação incidente pode percorrer diferentes caminhos no interior da amostra, com diferente número de espalhamento, com variação da profundidade de penetração. Como consequência, o caminho ótico não é bem definido e varia como uma função da absortividade da amostra. Em outras palavras, nas regiões de baixa absorção, a profundidade de penetração é maior que nas regiões de alta absortividade. Como resultado, além da distorção espectral, o sinal produzido não é uma função linear em relação à concentração da amostra.

A profundidade de penetração da radiação depende de vários fatores para uma dada matriz e concentração do analito, o tamanho da partícula da matriz e do analito governam a profundidade de penetração efetiva (ou o caminho ótico efetivo)^[15]. Deste modo o tamanho das partículas influenciam diretamente a linearidade da análise por DRIFTS. O efeito do tamanho e morfologia das partículas tem sido discutido na literatura^[16].

Várias aproximações para a linearização das informações obtidas pela reflexão difusa foram propostas e a mais aplicada atualmente é a relação de Kubelka-Munk, cuja aplicação está baseada em algumas considerações^[17]:

- 1) Uma camada é composta de partículas, aleatoriamente distribuídas e absorventes que espalham a radiação de maneira uniforme;
- 2) A dimensão das partículas é muito menor que a espessura da camada;
- 3) A direção da radiação incidente e refletida é perpendicular à superfície da camada; assume-se que a radiação é monocromática e que a espessura é infinita;
- 4) A amostra deve estar diluída em uma matriz não absorvente.

A forma mais conhecida da relação Kubelka-Munk é representada pela equação 2.3:

$$f(R_{\infty}) = \frac{(1-R_{\infty})^2}{2 R_{\infty}} = \frac{k}{s} \quad (2.3)$$

Na qual

$f(R_{\infty})$ é a função da equação, denominada função de Kubelka-Munk

R_{∞} é a reflexão da amostra para profundidade infinita

k é o coeficiente de absorção

s é o coeficiente de espalhamento da amostra.

Apesar de ser largamente difundida, esta equação é aplicada com sucesso apenas em situações muito restritas.

A técnica de reflexão difusa no infravermelho médio têm sido aplicada em análises quantitativas de diversos tipos de materiais e sistemas, com resultados satisfatórios. A determinação do *grau de esterificação de pectinas* (polissacarídeos utilizados na indústria alimentícia) foi realizada por Gnanasambandam e

colaboradores^[18]. Os dados foram tratados com métodos univariados e os resultados apresentaram erros relativos de até 7% para pectinas com alto grau de esterificação.

Guauch e colaboradores^[19-20] investigaram a potencialidade de DRIFTS na determinação de *poluentes* (amônio, hidrogenofosfato e os cátions Cu(II), Co(II) e Ni (II) e Fe (III)). Esta técnica de refletância apresentou boa sensibilidade, pois os limites de detecção foram obtidos na faixa de 10^{-4} mol/L. Foram aplicadas correções de Kubelka-Munk, no entanto os autores não utilizaram métodos quimiométricos e os erros foram obtidos na faixa de 10 a 20%.

Embora alguns métodos de calibração univariada tenham sido aplicados, recentemente, grande parte das análises quantitativas multicomponentes tem exigido a utilização de métodos multivariados, uma vez que as bandas características dos vários componentes são sobrepostas. Este fato foi observado por Ferraz e colaboradores^[21] que analisaram a *biodegradação da madeira* ocasionada por fungos. Os autores desenvolveram uma metodologia por DRIFTS combinada ao tratamento quimiométrico, garantindo resultados rápidos com erros de previsão na faixa de 0,3 a 0,8%.

Park e co-autores^[22] também enfatizaram o uso da quimiometria na *determinação de FeO em minérios de ferro* através de DRIFTS. Os autores compararam resultados obtidos em termos de RMSEC (raiz quadrada média dos erros de calibração) e RMSEV (raiz quadrada média dos erros de validação) observados para modelos de calibração construídos a partir de dados submetidos à correção do sinal: MSC (correção multiplicativa do espalhamento), OSC (correção ortogonal do sinal) e para dados sem correção. Foi verificado que a aplicação do tratamento dados foi eficiente na diminuição de erros. Vários algoritmos para calibração multivariada foram aplicados para estudo comparativo.

O teor de *matéria orgânica em solos*^[23] foi determinado através de DRIFTS no infravermelho médio. A metodologia desenvolvida forneceu uma alternativa à análise química convencional que emprega reagentes tóxicos. Os espectros foram obtidos em unidades de Kubelka-Munk (K/S), refletância (R) e $\log(1/R)$. Posteriormente foram submetidos a vários tipos de pré-processamento: MSC, alisamento, derivação e normalização. Foi utilizado o método PLS para calibração e foram observados melhores resultados para dados que sofreram alisamento, diferenciação e normalização.

2.4 – ESPECTROSCOPIA FOTOACÚSTICA

A técnica de espectroscopia fotoacústica permite a análise de sólidos opacos, de qualquer espessura, fragmentados ou contínuos, praticamente sem nenhum preparo de amostras. Adicionalmente, a magnitude do sinal fotoacústico é linearmente dependente da concentração do analito, permitindo análises quantitativas.

O sinal fotoacústico é gerado quando a radiação infravermelha absorvida pela amostra é convertida em calor. O calor é difundido para a superfície da amostra que está em contato direto com um gás inerte (nitrogênio seco, argônio ou hélio). Subseqüentemente ocorre a expansão térmica do gás produzindo os sinais acústicos, que são captados por microfones^[24-25]. A figura 2.4 mostra o esquema da geração do sinal.

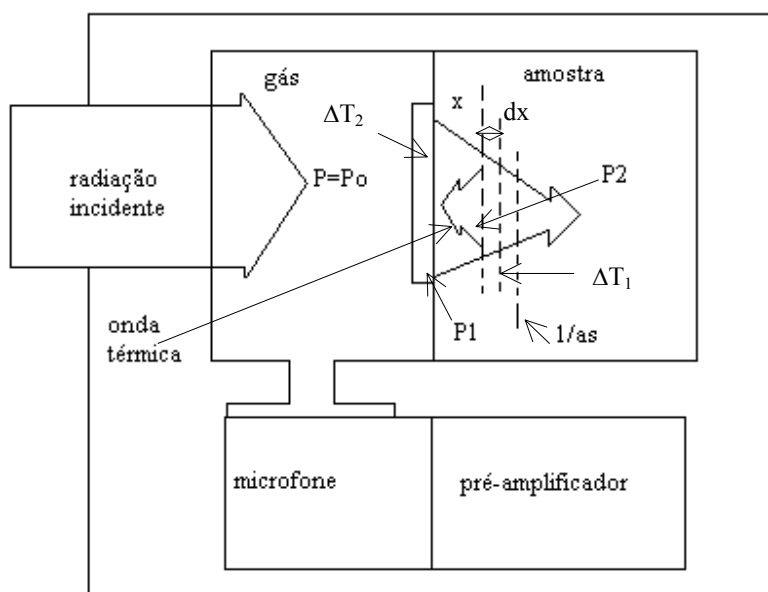


Figura 2.4: Geração do sinal fotoacústico^[24].

Inicialmente a radiação incide sobre a amostra com uma intensidade P_0 . A superfície da amostra reflete parte da radiação, então a intensidade torna-se $(1-R)P_0$. A absorção da radiação causa uma redução adicional de P_0 à profundidade x na amostra, daí: $(1-R)P_0 e^{-\alpha x}$, na qual α é o *coeficiente de absorção*.

A quantidade de energia que pode ser convertida em calor na camada dx é proporcional a $(1-R)I_0 \alpha e^{-\alpha x}$, resultando em uma onda térmica de amplitude ΔT_1 . Cada camada da amostra que oscila na temperatura é uma fonte de ondas térmicas que têm origem no interior da amostra e se propagam até a superfície que está em contato com o gás, local onde o sinal fotoacústico é gerado. A amplitude das ondas térmicas decai durante a propagação com um coeficiente de decaimento ou coeficiente $as = (\pi f/D)^{1/2}$, no qual D é a *difusibilidade térmica* da amostra e f é a

frequência da radiação. Com isso o valor na superfície de ΔT_2 passa a ser proporcional a $(I-R) P_0 \alpha e^{-(\alpha + as) x} dx$.

É importante ressaltar que os coeficientes de decaimento e absorção (as e α) têm um importante papel na geração do sinal fotoacústico: quando $\alpha \ll as$, o sinal aumenta linearmente com α porque a absorção de radiação infravermelha é intensificada. No entanto há um limite para este comportamento, e o sistema deverá perder linearidade em fortes absorções (saturação).

Variando-se a velocidade do espelho (step-scan), estudos do *perfil de profundidade* (L) podem ser realizados, uma vez que:

$$L = (D/\pi f)^{1/2} \quad (2.4)$$

na qual f é a frequência de modulação.

Pode-se observar, portanto, que o perfil de profundidade da radiação é variável com determinado número de onda. Além disso, a técnica PAS está relacionada com as propriedades óticas e térmicas da amostra. A fundamentação teórica envolvendo estas propriedades é complexa e extensa, mas foi integralmente descrita por Rosencwaig e Gersho^[26].

O principal atrativo desta técnica reside no fato de que não são necessárias condições especiais de preparação de amostra. As amostras não precisam ser diluídas em KBr. Entretanto, o *tamanho da partícula* e a *morfologia da superfície* podem em alguns casos afetar a qualidade dos espectros PAS.

O efeito do tamanho da partícula de sólidos sobre os espectros PAS foi investigado por alguns autores^[27-28] que observaram que a intensidade dos picos aumenta com a diminuição do tamanho das partículas e a razão entre os picos não é constante para partículas de diferentes diâmetros. Adicionalmente o formato dos picos é prejudicado com o aumento do tamanho das partículas, devido à influência de propriedades térmicas do material.

O efeito da superfície da amostra sobre os espectros PAS foi examinado por Vidrine ^[29] em amostras de plásticos nitrílicos com diversos tipos de superfície: pós, “pellets”, superfície rugosa e superfície polida. As amostras constituídas de pós finos produzem o melhor espectro porque a transferência de calor do sólido para o gás é favorecida com o aumento da razão área superficial/volume.

A técnica PAS é muito apropriada na análise espectroscópica de superfície, além de transpor dificuldades experimentais comuns em análise no infravermelho como sensibilidade a água ou ar, insolubilidade, infusibilidade e opacidade da amostra. No entanto, em vista da intensa aplicação das técnicas de refletância difusa e ATR, a técnica de espectroscopia fotoacústica tem sido pouco aplicada a análises quantitativas. Alguns fatores contribuem para este fato como à forte dependência da linearidade do método com as propriedades térmicas e óticas da amostra (saturação) e à obtenção de espectros muito ruidosos.

Todavia, a técnica PAS pode ser aplicada com resultados satisfatórios em alguns tipos de análise quantitativa como na determinação de espécies adsorvidas, demonstrado por Pekel e Guven^[30] que investigaram a adsorção de soluções de poli(n-butil metacrilato) em alumina. Através da aplicação de PAS, os autores puderam avaliar a quantidade do polímero adsorvido.

O método PAS foi recentemente utilizado na determinação de proteínas (manchas de sangue) em tecidos (algodão, poliéster)^[31]. Hemoglobina e sangue foram determinados por PAS e micro-ATR através das bandas de amida.

Conforme já discutido anteriormente, uma das limitações da técnica de fotoacústica é a dependência do perfil de profundidade (L) com os números de onda (ν), como pode ser observado na relação $[L = 1/as = (D/\pi\nu\nu)^{1/2}]$. Como resultado para números de onda (cm^{-1}) altos, o perfil de profundidade diminui. No

entanto, a técnica pode corrigir este problema, uma vez que velocidade do espelho (v) pode ser alterada. Deste modo ocorre a modulação da radiação.

Uma das aplicações da técnica fotoacústica no infravermelho médio consiste no estudo de perfis de profundidade, uma vez que maiores profundidades podem ser investigadas (até 15 ou 20 μm contra 2 μm da técnica ATR)^[32]. A técnica também pode ser aplicada em estudos sobre morfologia de polímeros^[33].

3- APLICAÇÕES EM SISTEMAS ANALÍTICOS

3.1- INTRODUÇÃO

O presente trabalho foi baseado na apresentação de duas propostas de análise:

I - Análise espectroscópica dos parâmetros de três diferentes tipos de combustíveis comerciais: gasolina, diesel e álcool automotivo no infravermelho médio, através da técnica de ATR, com o acessório circle-cell e calibração multivariada (PLS) com seleção de variáveis.

II - Análise espectroscópica das propriedades de degradação (*concentração de grupos carboxílicos terminais e da viscosidade intrínseca*) do PET Poli(etileno tereftalato) no infravermelho médio, através das técnicas de Reflexão difusa (DRIFTS) e fotoacústica (PAS) com calibração multivariada (PLS).

Os sistemas estudados apresentam uma série de características particulares quanto à sua análise convencional, análise espectroscópica, legislação, natureza físico-química que serão mencionadas nos próximos itens.

3.2- APLICAÇÃO I: Análise de Combustíveis Comerciais (diesel, gasolina e álcool automotivo).

3.2.1.- Características gerais da gasolina

A gasolina é um combustível constituído basicamente por uma mistura complexa de hidrocarbonetos. Esses hidrocarbonetos são, em geral, formados por moléculas de menor cadeia carbônica (normalmente de 4 a 12 átomos de carbono). Além dos hidrocarbonetos, a gasolina contém compostos oxigenados, compostos de enxofre, compostos de nitrogênio e compostos metálicos, todos eles em baixas

concentrações. A faixa de destilação da gasolina automotiva varia de 30 a 220°C [34].

Os hidrocarbonetos componentes da gasolina são membros das séries parafínica, naftênica e aromática, e suas proporções relativas dependem do petróleo e processos de refino utilizados. Atualmente, a gasolina que sai das refinarias, dotadas de vários processos de refinação, são constituídas de misturas criteriosamente balanceadas desses hidrocarbonetos, visando a atender aos requisitos de desempenho nos motores.

De acordo com a ANP (Agência Nacional do Petróleo) as gasolinas automotivas são classificadas em dois tipos A e C: a primeira é produzida no país, importada ou formulada pelos agentes econômicos autorizados para cada caso, isenta de componentes oxigenados e que atende ao regulamento técnico. A tipo C é aquela constituída de gasolina do tipo A e álcool etílico anidro combustível, nas proporções e especificações definidas pela legislação em vigor^[35].

As gasolinas A (sem álcool anidro) e C (com álcool anidro) podem ser classificadas ainda como gasolina *comum*, gasolina *comum aditivada* e gasolina *Premium*, sendo a última de superior octanagem às demais. Quanto à utilização de aditivos, a gasolina pode ser aditivada ou não aditivada. A gasolina aditivada contém um aditivo constituído de um detergente/dispersante, que tem a função a remoção da goma depositada no sistema mecânico do veículo^[36].

As propriedades da gasolina dependem basicamente^[36]:

- das características físico-químicas do petróleo processado;
- dos processos de refino e das respectivas condições operacionais praticadas;
- das proporções utilizadas de cada fração na formulação da mistura;

- da necessidade de otimizar a produção global da refinaria para atender as demandas dos diversos produtos;
- dos aditivos antioxidantes, anti-detonantes, detergentes, melhoradores de lubricidade, etc.

3.2.2- Especificações técnicas para a gasolina

A qualidade mínima dos diversos tipos de combustíveis utilizados no Brasil é garantida por especificações técnicas legalmente estabelecidas. A qualidade dos combustíveis deve garantir o desempenho energético satisfatório para o bom funcionamento dos motores, respeitando as leis de proteção ambiental. Para tal, é necessário estabelecer parâmetros de qualidade que são designados sob forma destas especificações técnicas. Estas especificações estabelecem a uniformidade de cada combustível, que deve estar com suas propriedades físico-químicas em conformidade com elas^[5].

É preciso que a qualidade dos combustíveis esteja assegurada desde a sua saída da refinaria (ou indústria alcooleira) até os postos de abastecimento. No Brasil, esta tarefa realizada pela ANP ou Agência Nacional do Petróleo.

De acordo com a ANP, o regulamento técnico estabelece que as especificações para a gasolina, álcool automotivo e diesel devem ter suas características físico-químicas determinadas em ensaios pré-definidos por Normas Brasileiras (NBR) e Métodos Brasileiros (MB) da Associação Brasileira de Normas Técnicas (ABNT) ou de normas da American Society for Testing and Materials (ASTM)^[35].

Os valores de cada parâmetro medido também são especificados pelo regulamento técnico da ANP. A tabela 3.1 mostra as especificações da ANP para os parâmetros estudados neste trabalho para a gasolina:

Tabela 3.1: Especificações e normas para gasolina (ANP- Portaria nº 309, 27/12/01).

PROPRIEDADE	ESPECIFICAÇÕES GASOLINA (ANP- Portaria nº 309, 27/12/01)						
	unidade	Limites Gas. Comum		Limites Gas. Premium		Norma ABNT	Norma ASTM
		Tipo A	Tipo C	Tipo A	Tipo C		
Densidade a 20°C	Kg/m³	-	-	-	-	NBR 7148 ^[37] NBR 14065 ^[38]	D 1298 ^[39] D 4052 ^[40]
T10%	°C	65,0	65,0	65,0	65,0	NBR 9619 ^[41]	D 86 ^[42]
T50 %	°C	120,0	80,0	120,0	80,0		
T90%	°C	190,0	190,0	190,0	190,0		
PFE, máx.	°C	220,0	220,0	220,0	220,0		
Resíduo, máx.	% vol.	2,0	2,0	2,0	2,0		
MON, mín.		80,0	82,0	-	-	MB 457 ^[43]	D 2700 ^[44]
IAD, mín.	-	-	87,0	-	91,0	MB 457	D 2699 ^[45] D2700
PI (100 °C)	min	-	360	-	360	NBR 14478 ^[46]	D 525 ^[47]
Álcool etílico anidro	%					NBR 13992 ^[48]	
Benzeno, máx.	% vol	1,2	1,0	1,9	1,5	-	D 3606 ^[49] D 5443 ^[50] D 6277 ^[51]
Oleofínicos, máx.	%vol	38	30	38	30	MB 424 ^[52]	D 1319 ^[53]
Aromáticos, máx.	% vol.	57	45	57	45	MB 424	D 1319

3.2.3- O significado e a importância das especificações para a gasolina^[54]

A *massa específica ou densidade* é definida como a massa de um líquido por unidade de volume a uma dada temperatura de referência. Trata-se de propriedade física fundamental que pode ser usada em conjunto com outras propriedades para caracterizar tanto as frações leves quanto as frações pesadas de petróleo e produtos de petróleo. Esta propriedade é um importante indicador da qualidade para os

combustíveis automotivos, no que se refere à estocagem, manuseio e combustão. Pode ser determinada pelo uso de um densímetro digital.

A *curva de destilação* relaciona as temperaturas em que a gasolina é evaporada nas frações de 10%, 50%, 90% (°C) e PFE (ponto final de evaporação). A destilação tem como objetivo avaliar as características de volatilidade da gasolina. Se a evaporação é insuficiente, a partida e a aceleração durante o aquecimento podem ser afetadas. Já a evaporação em excesso pode, em dias quentes, formar vapores na bomba de gasolina, tornando irregular ou cortando a entrada de combustível. Para obter a volatilidade adequada os componentes da gasolina (hidrocarbonetos leves, médios e pesados) devem estar presentes em proporções corretas. Esta adequação é verificada através da curva de destilação. Os componentes leves vaporizam-se a temperaturas menores, favorecendo a partida do motor, mas sua presença deve ser limitada. A proporção de componentes médios influi no tempo de aquecimento do motor. Os componentes pesados são os mais energéticos, mas sua vaporização e queima são mais difíceis.

A curva de destilação é feita tomando-se 100 mL de amostra que é colocada em um balão de vidro e em seguida é submetido a aquecimento para destilação em condições controladas no destilador automático. Com esse aquecimento, o produto se vaporiza sendo, então, condensado e recolhido em uma proveta de vidro. Após essa operação, as temperaturas anotadas são corrigidas levando-se em conta as perdas que ocorrem por evaporação de pequena parte do produto e a pressão barométrica. Esse teste, além de ser usado no controle da produção da gasolina, pode ser utilizado para identificar a ocorrência de contaminação por derivados mais pesados como o óleo diesel, óleo lubrificante, querosene etc.

As *propriedades de detonação ou octanagem* são representadas pelos parâmetros MON (motor octane number), RON (research octane number) e IAD (índice antidetonante).

Nos motores a gasolina, o combustível é misturado com o ar e comprimido nos cilindros antes da queima (por centelha). Quanto maior a compressão, sem que a mistura detone, maior é a potência do motor. O método MON se refere à situação de carga total em alta rotação (como exigir a potência máxima para uma ultrapassagem em alta velocidade). O teste é feito em motores especiais (motores CFR- Cooperative Fuel Research;), monocilíndricos de razão de compressão variável, equipados com a instrumentação necessária e montados numa base estacionária.

O método RON se refere à situação de carga total em baixa rotação (como acelerar ao máximo um carro parado, quando o sinal abre). O teste é feito em motores semelhantes àqueles utilizados para o teste da octanagem MON.

O índice antidetonante IAD é a média aritmética entre MON e RON. Quando se trata de definir a octanagem requerida pelos motores alguns países como os EUA e o Brasil, em se tratando da gasolina Premium, adotam ao invés do numero de octanagem MON ou RON, o índice antidetonante (IAD) como representativo do desempenho antidetonante do combustível. O que ocorre é que dependendo do projeto do motor do veículo e das condições em que ele opera, o desempenho antidetonante do combustível pode ser melhor representada, em alguns casos pela octanagem MON em outras pela octanagem RON. Com o índice antidetonante (IAD), estima-se o desempenho antidetonante do combustível para um universo mais amplo de veículos o que o coloca em vantagem em relação a octanagem MON ou RON, separadamente.

A *octanagem* de uma gasolina é definida pela comparação com uma mistura contendo iso-octano (ao qual é creditado um número de octanagem igual a 100) presente em uma mistura com n-heptano (numero de octanagem igual a zero). Exemplificando, uma gasolina terá uma octanagem igual a 80 se, durante o teste, apresentar a mesma resistência à detonação apresentada por uma mistura que contém 80% em volume de iso-octano e 20% em volume de n-heptano. A avaliação da octanagem da gasolina é justificada pela necessidade de garantir que o produto atenda às exigências dos motores no tempo de compressão e início da expansão (quando ocorrem aumento de pressão e temperatura) sem entrar em auto-ignição.

O *período de indução* (PI) é realizado através de um equipamento Petrotest (ITRON 16) que fornece uma indicação da tendência da gasolina de formar goma durante estocagem.

A lenta oxidação da gasolina na presença de ar forma produtos como gomas, que se depositam nas superfícies dos compartimentos nos quais o combustível entra em contato dentro do motor, ocasionando o mau funcionamento. Deste modo, um limite máximo de formação de goma é desejado (5mg para 100 mL de gasolina). O período de indução é especificado para controlar a formação de goma e permitir que o combustível chegue ao consumidor com um teor aceitável da mesma, evitando danos ao motor.

A amostra é oxidada no interior de uma bomba de aço inoxidável carregada com oxigênio a 100 psi e aquecida em banho de água. Determina-se então, o tempo necessário para a queda especificada da pressão (gasolina). Basicamente o equipamento utilizado consiste em um banho de oxidação, sistema de monitoramento da pressão e bomba de oxidação.

O *teor de álcool* é a quantidade de álcool etílico anidro combustível (AEAC) misturada à gasolina e que deve estar em conformidade com a especificação em

vigor. O percentual alcoólico era de 22%, quando começou a ser adicionado obrigatoriamente à gasolina. No entanto, este percentual pode ser moderadamente modificado, em função da flutuação na produção de álcool, mas normalmente se encontra na faixa de 20 a 25%. Por causa da calibração dos motores, a mistura do etanol anidro à gasolina não pode variar muito. Deste modo, o monitoramento deste parâmetro torna-se muito importante na avaliação da qualidade da gasolina. O teste é rotineiramente feito pelo “ método da proveta”, no qual é utilizada uma proveta de 100 mL e uma solução aquosa de NaCl é adicionada à amostra, observando-se a variação de volume das duas fases.

O controle do tamanho e da estrutura das moléculas de hidrocarbonetos que constituem a gasolina é relevante para sua qualidade. Os hidrocarbonetos podem ser *aromáticos, olefinicos e saturados*. Os aromáticos têm alta octanagem, mas são tóxicos. Deste modo, a presença de aromáticos, principalmente o *benzeno* é controlada. As oleofinas apresentam bom aproveitamento energético, mas participam da formação de ozônio na baixa atmosfera que é um gás de alta toxicidade. Por esta razão de natureza ambiental, a taxa de olefinicos também é controlada no Brasil^[5].

3.2.4- Características gerais do diesel

O diesel é um combustível que corresponde aos destilados intermediários do petróleo (faixa de destilação de 170 a 320°C) sendo constituído basicamente por hidrocarbonetos classificados em parafínicos, oleofínicos, naftênicos e aromáticos. A distribuição dos hidrocarbonetos nestas quatro categorias é muito importante, pois ajudam a definir as propriedades do combustível. Alguma impureza também pode estar presente em menores quantidades como o nitrogênio e enxofre^[34].

O diesel é o principal produto do fracionamento do petróleo no Brasil. Cerca de 35% do petróleo bruto correspondem à fração do diesel, que possui características de ignição adequadas a motores de combustão interna por compressão. Estes motores são encontrados em caminhões, ônibus, máquinas estacionárias, pequenos barcos, etc.^[5]

A ANP classifica o diesel em duas categorias ^[35]:

- Óleo Diesel Automotivo Metropolitano (tipo D): Produzido no país, importado ou formulado pelos agentes econômicos autorizados para cada caso conforme características constantes no Regulamento Técnico para comercialização nos municípios (em geral municípios com alta densidade populacional) estabelecidos pelo Ministério do Meio Ambiente.

- Óleo Diesel Automotivo Interior (tipo B): Produzido no país, importado ou formulado pelos agentes econômicos autorizados para cada caso conforme características constantes no Regulamento Técnico para comercialização nos demais municípios do país.

Basicamente, os tipos B e D diferem no teor de enxofre permitido de acordo com o regulamento técnico, no qual a quantidade de enxofre para o diesel tipo B (interior) tem um limite superior ao permitido para o diesel tipo D (metropolitano).

Parte do óleo diesel recebe uma aditivação após sair da refinaria. Os aditivos podem apresentar propriedades desemulsificante, antiespuma, detergente e inibidor de corrosão.

3.2.5- Especificações técnicas para o diesel

O combustível adequado para motores a diesel deve apresentar qualidade suficiente para dar um máximo de rendimento energético, estando em conformidade com leis ambientais. A qualidade do diesel também está relacionada

com a minimização de problemas de depósito e desgastes das peças mecânicas e filtros, além de evitar produção excessiva de fumaça. O bom funcionamento do motor com relação à partida e aquecimento adequado também deve ser assegurado por um combustível de qualidade.

Todas estas características são asseguradas pelas especificações do diesel brasileiro que são praticamente as mesmas adotadas internacionalmente.

Na tabela 3.2 estão dispostos os limites permitidos para os parâmetros do diesel estudados nesse trabalho, bem como sua respectiva metodologia regulamentar de análise:

Tabela 3.2 Especificações e normas para o diesel (ANP- Portaria nº 310, 27/12/01).

PROPRIEDADE	ESPECIFICAÇÕES DIESEL (ANP- Portaria nº 310, 27/12/01)				
	<i>unidade</i>	<i>Limites Tipo(B)</i>	<i>Limites Tipo (D)</i>	<i>Norma ABNT</i>	<i>Norma ASTM</i>
<i>Densidade a 20°C</i>	Kg/m ³	820 a 880	820 a 865	NBR 7148 ^[37] NBR 14065 ^[38]	D 1298 ^[39] D 4052 ^[40]
<i>T50%</i>	°C	245,0 – 310,0		NBR 9619 ^[41]	D 86 ^[42]
<i>T85%</i>	°C	370,0	360,0		
<i>Índice de Cetano</i>	—	45 (mínimo)			D 4737-90 ^[55]
<i>Enxofre</i>	% massa	0,35	0,20	NBR 14533 ^[56]	D 1552 ^[57] D 2622 ^[58] D 4294 ^[59]
<i>Viscosidade a 40°C</i>	(mm ² /s)	2,5 a 5,5		NBR 10441 ^[60]	D 445 ^[61]

3.2.6- O significado e a importância das especificações para o diesel^[54]

A *densidade* do diesel permite calibrar os motores de modo que o volume correto de combustível seja injetado na câmara de combustão. Valores altos de densidade levam a uma significativa variação da massa do combustível injetada, impossibilitando uma mistura correta de ar/combustível, aumentando, deste modo,

a emissão de poluentes e material particulado. Valores baixos de densidade reduzem o desempenho, causando perda de potência e de dirigibilidade pela formação de uma mistura pobre, ocasionando o aumento do consumo do combustível^[5].

A capacidade de evaporação do diesel, denominada *volatilidade* está relacionada com a *curva de destilação*, representada, entre outras, pelas temperaturas de destilação a 50% e a 85% do produto. A primeira (T50%) é a temperatura na qual 50% do volume do diesel é destilado. A Análise convencional é realizada destilando-se 100mL da amostra e anotando-se a temperatura que corresponde à fração de 50% de destilado. Esta temperatura corresponde às frações leves e médias do diesel, responsáveis pelo bom desempenho do motor quando este já se encontra em plena carga de funcionamento e nas retomadas de velocidade.

A temperatura T85% corresponde às frações pesadas do combustível. O controle da fração pesada está relacionado com a minimização da formação de depósitos no interior do motor. A análise da temperatura T85% também controla as emissões gasosas dos hidrocarbonetos não queimados, fumaça e óxido de nitrogênio. Adicionalmente, o maior teor de frações pesadas dificulta a vaporização do combustível.

O *número de cetano* mede a qualidade de ignição do diesel e tem influência na partida do motor e seu funcionamento sob carga. Nos motores a diesel, o ar que entra no cilindro é comprimido até atingir as condições ideais para a ignição espontânea do combustível injetado. A ignição deve sofrer, no entanto, um pequeno atraso para que o diesel injetado se aqueça, se vaporize e se misture com o ar. Quanto menor este retardo, maior a qualidade de ignição do diesel. Quando o retardo é excessivo, a quantidade de diesel que permanece na câmara sem queimar no tempo certo é elevada, ocasionando uma liberação de energia demasiada quando

a combustão acontece, provocando esforços anormais sobre o pistão, causando danos mecânicos e perda de potência.

A determinação do número de cetano é feita pela comparação do desempenho do diesel e o desempenho do n-hexadecano, produto parafínico comercializado como *cetano*, o qual é atribuído um número de cetano igual a 100. A um produto aromático é atribuído um número de cetano igual a zero. Isto acontece porque combustíveis com alto teor de parafinas apresentam alto número de cetano, enquanto que o alto teor de aromáticos possuem baixo número de cetano. A determinação desta propriedade requer o uso de um motor teste padrão (motor CFR) operando sob condições também padronizadas.

O *índice de cetano* está correlacionado com o número de cetano, e, portanto, também está ligado à qualidade de ignição. Pela sua praticidade, o índice de cetano pode substituir o número de cetano. É calculado através de uma equação de quatro variáveis determinada pela ASTM (D4737-90) e proporciona a média estimada do número de cetano de combustíveis destilados a partir de sua densidade e temperatura de destilação. A equação 3.1 representa o cálculo do índice de cetano, determinado pela norma ASTM:

$$\begin{aligned}
CCI = & 45,2 \\
& + (0,0892) (T_{10N}) \\
& + [0,131 + (0,901) (B)] [T_{50N}] \\
& + [0,0523 - (0,420)(B)] [T_{90N}] \\
& + [0,00049][(T_{10N})^2 - (T_{90N})^2] \\
& + (107)(B) + (60) (B)^2
\end{aligned} \tag{3.1}$$

Onde:

CCI= índice de Cetano calculado pela equação das 4 variáveis

D = Densidade a 15°C

DN = D-0,85

B = $[e^{(-3,5)(DN)}] - 1$

T₁₀ = 10% da temperatura de destilação

T_{10N} = T₁₀ - 215

T₅₀ = 50% da temperatura de destilação

T_{50N} = T₅₀ - 260

T₉₀ = 90% da temperatura de destilação

T_{90N} = T₉₀ - 310

A *viscosidade* é uma medida da resistência do combustível ao escoamento. Está relacionada com a propriedade de fluidez do diesel, lubrificação e nebulização (atomização). Valores de viscosidade abaixo da faixa podem levar ao desgaste das partes autolubrificantes do sistema de injeção, vazamento na bomba e danos ao pistão. Viscosidade excessiva pode ocasionar o aumento do trabalho da bomba do combustível, além de proporcionar má nebulização do combustível com conseqüente combustão incompleta, aumento da emissão de fumaça e material particulado. O teste é realizado com o auxílio de um viscosímetro de tubo capilar de vidro com temperatura controlada. Esta propriedade é conhecida como viscosidade cinemática.

O *teor de enxofre* representa o controle da quantidade de compostos sulfurados existentes no combustível. A presença de compostos sulfurados afeta a durabilidade do motor e provoca a emissão de gases poluentes. A ação corrosiva e a toxidez de seus compostos são devido à formação de SO₂ e SO₃ durante a

combustão e à formação de ácido sulfúrico na presença de água. A análise pode ser feita através da fluorescência de Raio-X.

3.2.7- Características gerais do álcool automotivo

O álcool, ao contrário da gasolina, é uma substância pura (etanol), embora seja encontrado nos postos como sendo uma mistura de 95% de etanol e 5% de água, em volume. É uma molécula cuja fórmula é C_2H_5OH . O álcool é líquido à temperatura ambiente devido às interações moleculares. É um combustível de fonte renovável, sendo obtido no Brasil, a partir da fermentação dos açúcares obtidos da cana de açúcar^[62].

O etanol é um excelente combustível automotivo, pois apresenta octanagem superior ao da gasolina. Por ser menos volátil que esta, há diminuição das emissões evaporativas na atmosfera, conseqüentemente há uma redução da contaminação do ar. Adicionalmente, não possui enxofre como impureza; logo, a combustão do álcool não lança óxidos de enxofre na atmosfera.

É um combustível que não forma gomas, portanto é considerado mais "limpo" que a gasolina. Tem a desvantagem de ser mais corrosivo no estado líquido que a gasolina, o que demanda um tratamento anticorrosivo nos metais que têm contato com o álcool em sua fase líquida, normalmente através de um revestimento com um metal, como o níquel, usado para revestir os carburadores. Adicionalmente tem um poder calorífico inferior ao da gasolina.

Em 1975 iniciou-se no Brasil O Programa Nacional do Álcool (Proálcool) que visava à substituição gradativa da gasolina pelo álcool, mas em 1979 passaram a ser fabricados veículos já adaptados a este combustível. Em 1982, em função da produção elevada e do interesse do uso do etanol, o governo determinou a mistura obrigatória de álcool anidro à gasolina^[5].

3.2.8- Especificações técnicas para o álcool automotivo

Conforme estipulado pela ANP, os álcoois etílicos combustíveis classificam-se em:

I – Álcool Etílico Anidro Combustível (AEAC), destinado aos distribuidores para a mistura com a gasolina A (sem álcool) para a formulação da gasolina C (com álcool);

II- Álcool Etílico Hidratado Combustível (AEHC) é destinado como combustível em motores de combustão interna de ignição por centelha.

Dentre às diferentes propriedades designadas para a especificação do álcool automotivo no Brasil, destacaram-se neste trabalho as propriedades indicadas na tabela 3.3:

Tabela 3.3-Especificações e normas para o álcool automotivo (ANP- Portaria nº 2, 16/02/02).

PROPRIEDADE	ESPECIFICAÇÕES ÁLCOOL COMBUSTÍVEL (ANP- Portaria nº 2, 16/01/02)			
	<i>unidade</i>	limites	<i>Norma ABNT</i>	<i>Norma ASTM</i>
<i>Densidade a 20°C</i>	Kg/m ³	807,6 a 811,0	NBR 5992 ^[63]	D 4052 ^[40]
<i>Teor Alcoólico</i>	°INPM	92,6 a 93,8	NBR 5992 ^[63]	-

3.2.9- O significado e a importância das especificações para o álcool automotivo:

A determinação da densidade do álcool automotivo tem a mesma importância para os demais combustíveis, uma vez que flutuações nos valores de densidade podem acarretar consumo irregular de combustível. A análise padrão é realizada a 20°C com o auxílio de um densímetro digital.

O teor alcoólico é uma medida da quantidade de álcool puro presente na mistura de álcool e água do AEHC. É estimado na unidade de °INPM (Instituto Nacional de Pesos e Medidas) que expressa a quantidade em gramas de álcool absoluto contida em 100 gramas de mistura hidro-alcoólica. Esta propriedade é determinada a partir de valores tabelados das densidades e das temperaturas de aferição do densímetro.

3.3- Análise Espectroscópica das Propriedades dos Combustíveis:

A habilidade da análise de combustíveis por espectroscopia com métodos quimiométricos pode ser comprovada pelas suas vantagens sobre os métodos convencionais como custo, praticidade, rapidez, entre outras. Andrade *et al* ^[67-68] exemplificou tais vantagens na determinação das propriedades de octanagem, comparando o método ASTM com os métodos espectroscópicos, dos quais o primeiro requer altos custos, manutenção freqüente, maior quantidade de amostra (1 litro contra 20mL do método espectroscópico), e maior tempo de análise (35 minutos contra 5 minutos do método espectroscópico). Adicionalmente, foi citado que o método convencional gera resíduos e necessita de reagentes específicos e caros, além do processo de automação ser bastante complicado, enquanto que na análise por infravermelho a automação é mais simples.

A espectroscopia no *infravermelho próximo* (NIR) tem sido aplicada em combinação com calibração multivariada na análise da gasolina, sobretudo na determinação de hidrocarbonetos alifáticos, aromáticos e olefínicos^[68] e propriedades de octanagem^[70-71]. A Espectroscopia Raman também foi aplicada com êxito na determinação das propriedades de octanagem e pressão de vapor de gasolinas comerciais^[72].

Barzilai *et al*^[73] utilizou NIR com automação on-line, por meio de fibras óticas na determinação de temperaturas de destilação, propriedades de octanagem (MON e RON). Os resultados obtidos através das técnicas de espectroscopia no infravermelho próximo (NIR) e Raman são bastante satisfatórios, fato comprovado por sua ampla utilização no estudo das propriedades dos combustíveis. Adicionalmente, estas técnicas têm a vantagem de serem facilmente adaptadas para análises on-line. No entanto, os espectros obtidos no infravermelho médio mostram picos mais definidos, adequados para a análise qualitativa, quando comparados aos espectros no infravermelho próximo. Além disso, estudos comparativos demonstraram que a calibração realizada com os dados no infravermelho médio apresentou menores erros de previsão, segundo Cooper e co-autores^[74], cujo trabalho empregou a análise de gasolinas comerciais por técnicas de Raman, infravermelho próximo e médio com acessório de ATR. Foram determinados teores de benzeno, tolueno, etilbenzeno e xilenos (BTEX). Simultaneamente, foi utilizado o método PLS (Partial Least Squares) para efetuar a calibração multivariada. Foram obtidos erros de previsão na faixa de 0,070 a 0,356, (dependendo do hidrocarboneto analisado) para os resultados obtidos com o infravermelho médio, que corresponde aos melhores resultados, quando comparados às demais técnicas.

Também na determinação de BTEX, Diehl e colaboradores^[75] utilizaram a espectroscopia no *infravermelho médio combinada à técnica de cromatografia gasosa* (CG/FTIR) para estimar a concentração destes hidrocarbonetos.

As *concentrações de benzeno e tolueno* foram determinadas por FTIR na fase gasosa^[76]. Com este procedimento, os autores afirmaram que a diluição prévia das amostras de gasolina em hexano torna-se desnecessária, otimizando o processo de análise.

Para a vaporização das amostras, foi necessário pequeno volume de amostra (1 μ L), fato que minimiza o impacto ambiental devido à baixa emissão de gases. No entanto, devido a esta pequena quantidade, os autores tiveram alguns problemas com a reprodutibilidade do método, devido à sua baixa sensibilidade.

Teores de *parafinas, olefínicos, naftênicos e aromáticos* (PONA)^[77] foram determinados através de técnicas de transmissão no infravermelho médio em amostras de nafta e gasolinas. Utilizou-se o método PLS na calibração multivariada com bons coeficientes de correlação (maiores que 0,96).

Compostos oxigenados como metanol e metil tert-butil éter que são aditivos utilizados em alguns países foram determinados por Garcia^[11] e co-autores, utilizando-se a *circle cell* e calibração multivariada (PLS). O etanol e outros álcoois adicionados como aditivos foram determinados por FTIR com técnicas de transmissão^[78], com a utilização do método PLS.

As *propriedades de octanagem* foram estudadas em trabalhos que empregaram a técnica no infravermelho médio^[79-81]: (MON) foi determinado por um sistema combinado de FTIR com técnica de transmissão e cromatografia gasosa^[79]; O número de octanagem RON determinado por técnicas de transmissão com celas de KBr com caminhos óticos de 0,05 e 0,025 mm^[80]; Andrade e colaboradores^[81] usaram os métodos PLS, PLS-polinomial e redes neurais com algoritmo genético para produzir modelos de calibração para MON.

A técnica de ATR com acessório de geometria horizontal^[82] foi utilizada para determinar várias propriedades da gasolina, como densidade, pressão de vapor, propriedades de destilação, tipos de hidrocarbonetos, propriedades de octanagem e teores de enxofre e benzeno. Para tal, foi aplicado o método PLS, cujos resultados apresentaram bons coeficientes de correlação e os valores obtidos para os erros de

previsão apresentaram-se abaixo dos valores de reprodutibilidade das técnicas ASTM para muitas das propriedades estimadas.

Apesar da ampla aplicação de FTIR-MID no estudo dos parâmetros da gasolina, a utilização da circle-cell tem sido pouco explorada, embora apresente vantagens como facilidade de adaptação on-line, o que é muito desejável em análises em grande escala.

Para o *diesel*, poucos trabalhos estão disponíveis em literatura, considerando a espectroscopia no infravermelho médio. Adicionalmente, o acessório de ATR mais utilizado possui geometria horizontal e não o uso da “circle cell”.

Parâmetros de densidade, viscosidade cinemática, curva de destilação, número de cetano, entre outras se utilizando ATR com acessório horizontal^[83-84] na região do infravermelho médio e próximo, com calibração multivariada (PLS).

A determinação do teor de enxofre tem sido verificada apenas através de estudos por infravermelho próximo^[85]. Os autores utilizaram métodos de calibração como PCR, PLS e MLR, sendo que o último exigiu seleção prévia de variáveis.

No caso da análise espectroscópica dos parâmetros do *álcool automotivo*, pode-se afirmar que muito pouco tem sido investigado. Uma das razões é o emprego restrito deste combustível, cuja utilização é mais efetiva no Brasil. Deste modo, existem poucos trabalhos disponíveis na literatura, dos quais destacam-se apenas o uso do *infravermelho próximo*^[86-87] na determinação do teor de água em álcool automotivo com erros relativos de 0,07%, através do método PLS.

Mendes *et al*^[88] utilizaram o infravermelho próximo e espectroscopia Raman para determinar o teor alcoólico e teor de água em álcool automotivo e em bebidas alcoólicas. Os autores concluíram que os modelos de calibração PLS-NIR apresentaram melhor desempenho que a calibração realizada com a técnica Raman,

cujos erros relativos (% m/m) para NIR foram de 0,15% para teor alcoólico e 0,16% para teor de água, enquanto que os erros relativos de Raman foram de 0,39% e 0,61% respectivamente.

Os parâmetros de densidade e teor alcoólico foram determinados também por NIR, com modelos PLS em recente trabalho^[89]. Os autores obtiveram erros relativos de 0,02 a 0,03% para densidade e 0,04 a 0,07% para teor alcoólico.

3.4- APLICAÇÃO II: Estudo das Propriedades de Degradação do PET Poli(tereftalato de etileno)

3.4.1- PET e processos de degradação^[90]

O PET é um polímero obtido pela reação de policondensação, a partir do ácido tereftálico ou tereftalato de dimetila e etileno glicol . A estrutura do PET é mostrada na Figura 3.1:

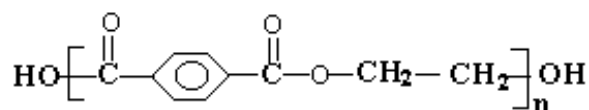


Figura 3.1: Estrutura do PET^[90].

É um poliéster termoplástico e tem como características a leveza, resistência mecânica e transparência, ideais para a produção de embalagens de refrigerantes e de outros produtos, como artigos de limpeza e comestíveis em geral.

A reciclagem das embalagens PET é uma conduta vantajosa, pois além de desviar resíduos plástico dos aterros, utiliza apenas 30 % da energia necessária para

a produção de PET virgem, uma vez que a polimerização do material virgem necessita de aquecimento a alta temperatura sob vácuo. Além disso, a reciclagem evita a incineração do lixo contendo PET, pois sua queima libera gases residuais como monóxido e dióxido de carbono, acetaldeído, benzoato de vinila e ácido benzóico .

O PET pós-consumo reciclado no Brasil pode ser reutilizado na produção de fibras para a fabricação de cordas, lâminas para termoformação e moldagem a vácuo, garrafas de detergentes, carpetes e enchimentos de travesseiros. É possível também reprocessar o polímero para a obtenção de resinas alquídicas usadas na produção de tintas.

Para reutilizar o PET pós-consumo, é necessário que se conheça o grau de degradação sofrido pelo material, pois processos degradativos podem comprometer suas propriedades de interesse. Deste modo, a aplicação específica do material reciclado torna-se adequada. Os processos degradativos relacionados ao PET são *a degradação mecânica, térmica e degradação química*. A *degradação mecânica* tem início no processamento do polímero, uma vez que este é submetido a processos de cisalhamento por atrito ou estiramento, resultando na quebra da cadeia polimérica. A *degradação térmica* ocorre durante o processamento, a secagem ou modificação com extensores de cadeia. Nestes processos, o material é submetido a temperaturas muito elevadas que traz como consequência a despolimerização e cisão das cadeias poliméricas, com liberação de monômeros e oligômeros, aumento da concentração de grupos ácidos, formação de ésteres cíclicos e diminuição da massa molar.

A degradação química pode ser induzida pela presença de ácidos, bases e solventes. As reações químicas ocorrentes são hidrólise, alcoólise, degradação oxidativa e acidólise. O PET é particularmente vulnerável à hidrólise acima de

100°C, portanto remoção de água e outros contaminantes é essencial na etapa de processamento. A hidrólise é uma reação autocatalítica, decorrente da variação do grau de cristalinidade e da hidrofilicidade. A hidrólise do PET é um processo de cisão simples que ocorre de forma aleatória formando moléculas de baixa *massa molar* com *grupos carboxílicos ou hidroxílicos nas extremidades* (Figura 3.2).

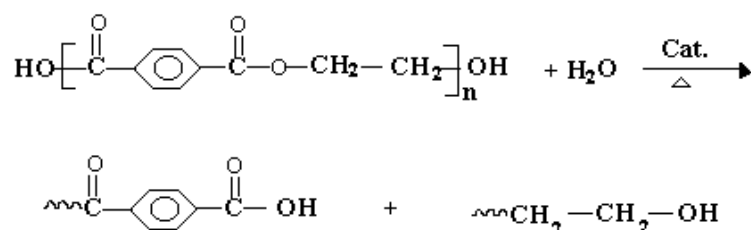


Figura 3.2: Formação de grupos carboxílicos e hidroxílicos terminais^[90].

3.4.2- Propriedades Indicadoras da Degradação do PET e sua análise convencional

Conforme foi descrito no item anterior a redução da massa molar e o surgimento de grupos carboxílicos terminais são características de processos degradativos do PET.

A viscosidade intrínseca $[\eta]$ dos polímeros está diretamente relacionada com sua massa molar média. Esta relação pode ser observada se o polímero é linear e não apresenta muita dispersão de massa molecular como é caso do PET. A análise tradicional de $[\eta]$ é descrita pela norma ASTM D 4603-01^[91], cujo procedimento consome tempo demasiado, utiliza reagentes tóxicos como fenol e tetracloreto de carbono com geração de grande quantidade de resíduo.

A concentração dos grupos carboxílicos terminais (CEG) em amostras de PET pode ser determinada, por via úmida, através do método Pohl^[92], no qual as amostras são dissolvidas em álcool benzílico bi-distilado a alta temperatura (190°C) com subsequente titulação com solução alcoólica de NaOH. A metodologia é complexa, criteriosa e envolve o uso de outros reagentes tóxicos como o clorofórmio e o metanol.

3.5- Análise da degradação do PET por Técnicas Espectroscópicas

Em vista das dificuldades impostas pela análise convencional das propriedades do PET e de outros polímeros comerciais, métodos alternativos têm sido estudados através de técnicas espectroscópicas, sobretudo a espectroscopia no infravermelho médio, da qual destacam-se as técnicas de transmissão, técnicas de reflexão e espectroscopia fotoacústica.

A técnica de *reflectância difusa* com um acessório especial foi utilizada na região do infravermelho médio e próximo na análise de polímeros naturais e sintéticos, entre eles o PET. O autor^[93] demonstrou através das técnicas utilizadas, que a natureza química dos materiais, o tamanho das partículas e a orientação das fibras influenciam as formas das bandas nos espectros e seus valores de reflectância. Adicionalmente foi empregado o método PCA (Análise dos Componentes Principais) para ilustrar o efeito destes parâmetros nos espectros.

Sammon e colaboradores^[94] estudaram a degradação de filmes finos de PET tratados com água pura e solução de KOH a 1% de concentração. As mudanças nas estruturas química e morfológica foram monitoradas pela técnica combinada de *reflexão-absorção (transreflectância)*, na qual analisa-se um filme fino da amostra depositado em um substrato (metálico) que serve como superfície refletora.

Alterações na intensidade das bandas de deformação das ligações de C-H, O-H e C=O foram verificadas e atribuídas às espécies formadas durante o processo de degradação. Também foi possível detectar através dos espectros, mudanças na cristalinidade do material.

Holland e Hay^[95] realizaram um estudo da degradação térmica do PET comercial através de *análise térmica combinada a espectroscopia no infravermelho* com transformada de Fourier (TA-FTIR). Esta técnica foi usada para medir a cinética de degradação polimérica e monitorar as mudanças que ocorreram na estrutura do polímero, durante a degradação térmica.

A determinação da *concentração dos grupos terminais* do PET através da Espectroscopia no Infravermelho teve início através do trabalho de Patterson e Ward^[96]. Os autores desenvolveram um método para estimar a concentração dos grupos O-H e COOH, simultaneamente do espectro de transmissão de um filme de PET. A estimativa da concentração destes grupos foi baseada no número de hidrogênios terminais substituídos por deutério. Isto é possível através da aquisição de um espectro de referência de uma amostra de polímero, na qual os hidrogênios dos grupos terminais foram quantitativamente substituídos, sem afetar os prótons da cadeia principal. Deste modo, a amostra deuterada define o espectro de referência, do qual são subtraídos os espectros a serem analisados.

Zhang e colaboradores^[97] utilizaram a técnica de Patterson e Ward para a determinação da concentração de grupos terminais e a viscosidade Poli(etileno-naftaleno-2,6- dicarboxilato) e do PET. Os autores deste trabalho também determinaram a viscosidade intrínseca através da sua relação com o peso molecular médio.

A *espectroscopia no infravermelho próximo* (NIR) foi empregada na estimativa de grupos carboxílicos terminais (CEG) e viscosidade relativa para

amostras de PET na forma de *chips* e PET fundido^[98-99]. Para tal foi utilizada a calibração multivariada com método PLS, da qual os autores obtiveram bons resultados (erros de previsão de 1,77 para CEG e 0,006 viscosidade relativa). No entanto a faixa de valores destas propriedades foi restrita: para a viscosidade relativa, as amostras variaram de 1,50 a 1,60 e para CEG variaram de 20 a 70 meq/Kg.

Em recente trabalho^[100], foram determinadas simultaneamente as concentrações de grupos carboxílicos e hidroxílicos terminais em filmes de PET, sintetizados em laboratório. Os autores utilizaram o método da substituição de hidrogênios terminais por deutério, no entanto com algumas modificações, como a preparação de um padrão que consiste em amostras de PET preparadas com bis(β -hidroxietiltereftalato), o (BHET), utilizado para determinar a relação entre os picos dos espectros no infravermelho, obtidos com *técnicas de transmissão* dos grupos hidroxílicos e carboxílicos. O estudo possibilitou a identificação e atribuição dos picos dos dois diferentes grupamentos. Adicionalmente, foram obtidos erros relativos entre o método de análise proposto e o método padrão de uma faixa de 3,3, a 12,0%.

Estudos preliminares da determinação de CEG e da viscosidade intrínseca por DRIFTS no infravermelho médio foram realizados^[90]. No entanto, o presente trabalho mostra um novo estudo a respeito da preparação das amostras e também a aplicação de técnica de fotoacústica com calibração multivariada.

A *espectroscopia fotoacústica no infravermelho* foi utilizada na análise de materiais poliméricos, entre eles o PET: Krishnann^[101] mostrou em seu trabalho algumas aplicações da técnica PAS no estudo do perfil de profundidade de filmes de PET.

Jasse^[102] enfatizou o uso de *PAS* na análise de polímeros sintéticos, entre os quais, filmes de PET, plástico nitrílico, poliestireno, entre outros.

Mudanças estruturais no PET foram investigadas em embalagens para alimentos, expostas a radiação ultravioleta. Para tal, foi aplicada a técnica de *PAS* em folhas PET com 225 µm de espessura. Além da versatilidade e sensibilidade do método no estudo da degradação ocasionada pela exposição à radiação UV, também puderam ser investigadas as propriedades térmicas do material e sua relação com o processo degradativo.^[103]

A técnica *PAS* também tem sido utilizada no estudo da degradação térmica de polímeros^[104].

Um estudo comparativo envolvendo técnicas de análise por espectroscopia no infravermelho através de técnicas de transmissão, reflexão e fotoacústica na análise da poliamida-6 reciclada e irradiada com feixes de elétrons na presença de oxigênio foi apresentado por Evora e colaboradores, no qual foi constatado que as mesmas alterações espectrométricas foram verificadas por todas as técnicas^[105].

4 - PARTE EXPERIMENTAL

4.1- DETERMINAÇÃO DOS PARÂMETROS DOS COMBUSTÍVEIS

4.1.1- Obtenção dos Valores de Referência

Todos os valores de referência dos parâmetros estudados de amostras comerciais de combustíveis automotivos foram obtidos na Central Analítica da UNICAMP, onde a maioria das análises foi baseada nas normas ASTM, mostrados no capítulo anterior.

No total foram analisadas 700 amostras de gasolinas (tipo C) comum, aditivada e Premium, que foram distribuídas em dois grupos: 483 amostras de *simples monitoramento* (sem suspeita de adulteração) e 217 amostras de “fiscalização” com *suspeita de adulteração*.

Também foram analisadas 100 amostras de álcool automotivo (AEAC) e 131 amostras de diesel com e sem aditivos e dos tipos D (metropolitano) e B (interior).

Foram analisados os seguintes parâmetros para os combustíveis:

GASOLINA SEM SUSPEITA DE ADULTERAÇÃO (MONITORAMENTO):

- *Massa Específica a 20°C (kg/m³)*: em densímetro digital (DMA 4500);
- *Curva de Destilação*: em destilador automático (NDI 440). Através desta análise foram determinadas as temperaturas de destilação *T10%*, *T50%*, *T90%*, *PFE* (ponto final de evaporação e *teor de resíduos (%)*).
- *Propriedades de Detonação*: em analisador IROX 2000 (Graber). Trata-se de um método padrão secundário, com análise no infravermelho médio. As propriedades analisadas foram *MON* (motor octane number), *RON* (research octane number) e *IAD* (índice antidetonante).

- *Tipos de hidrocarbonetos: (%) de benzeno, aromático e, saturados* determinados em analisador IROX 2000 (Graber)

GASOLINA COM SUSPEITA DE ADULTERAÇÃO (FISCALIZAÇÃO):

Foram avaliados alguns parâmetros similares aos da gasolina de monitoramento: curva de destilação (T10%, T50 %, T90 e PFE); MON, IAD, AEAC e benzeno.

O Período de indução (PI) foi medido através de um equipamento Petrotest (ITRON 16). Esta propriedade foi medida apenas para as gasolinas de fiscalização.

ÓLEO DIESEL:

- *massa específica* a 20°C (kg/m³) (mesmo equipamento da gasolina)
- *curva de destilação* com percentagens de diesel (%) nas temperaturas, T50 e T85°C nos mesmos equipamentos da gasolina.
- *índice de cetano* (através da equação ASTM que emprega valores de densidade e temperatura de destilação a 50%)
- *teor de enxofre* (%) através de Fluorescência de Raio X em um equipamento Spectro Titan.
- *viscosidade a 40°C* foi realizada através de um tubo especial para medidas de viscosidade (Cannon-Feske). Utilizou-se um termômetro (Koehler) e um banho termostatizado.

ÁLCOOL COMBÚSTIVEL (AEAC)

- *massa específica e teor alcoólico* (%) em densímetro digital (DMA 4500):

4.1.2- Obtenção dos Espectros FTIR dos combustíveis

As amostras de combustíveis foram submetidas à análise espectroscópica no infravermelho médio na faixa de 4000 a 650 cm^{-1} . Para tal, utilizou-se a técnica de ATR (acessório circle cell[®] com capacidade para 10 mL de amostra), com elemento de ZnSe. A aquisição dos espectros foi realizada à temperatura ambiente (22 a 25° C).

Os espectros de referência (background) foram adquiridos diretamente da cela (circle cell) vazia e limpa. Um novo espectro de referência foi coletado a cada cinco amostras, para garantir a qualidade do espectro em relação à variação de CO₂ e vapores de água presentes no ambiente de análise e eventual variação na linha base. Todos os espectros foram obtidos com resolução de 4 cm^{-1} , com 128 aquisições (varreduras), em um espectrômetro Bomem MB-100, equipado com detector DTGS (sulfato de triglicina deuterada).

4.2- DETERMINAÇÃO DAS PROPRIEDADES DE DEGRADAÇÃO DO PET

4.2.1- Obtenção das Amostras de PET

Foram utilizados “pellets” de PET virgem (RHOPET[®] S80) fornecidos pela Rhodia Ster S.A. Brasil. O material foi seco a 160°C por 6h em estufa convencional antes de cada processamento. Em seguida o material foi extrudado em uma extrusora mono-rosca Worterx (L/D =30, D=32 mm). Nesta etapa utilizou-se uma rosca do tipo Davis Standar Barrier Screw. O perfil de temperatura utilizado foi 220, 260, 270, 275 e 280°C e rotação de 102 rpm. Foi obtido um conjunto amostras proveniente do reprocessamento por cinco vezes de quatro amostras de PET virgem.

4.2.2- Determinação da concentração dos grupos carboxílicos terminais (CEG):

Os “pellets” processados no item 4.2.1 foram moídos em moinho de facas (Thompson), peneirados (1 mm) e secos a 160°C por 4 h em estufa a vácuo (marca Tecnal, modelo EDGCON 5P). Em seguida a determinação de CEG foi efetuada com base no método Pohl^[92]. Este método é uma micro-titulação que utiliza como titulante uma solução padronizada de NaOH 0,1 mol/L em álcool benzílico/metanol.

Várias etapas foram efetuadas para a realização do método Pohl, das quais destacam-se:

- *Destilação dos Solventes*: O processo de destilação foi utilizado para a purificação dos solventes utilizados.

A destilação do álcool benzílico foi efetuada com pressão reduzida, devido ao seu alto ponto de ebulição (204°C). O objetivo deste processo é eliminar impurezas e traços de ácidos carboxílicos provenientes de processos oxidativos. Para garantir que não houvesse nova oxidação após a destilação, o álcool destilado foi armazenado sob atmosfera inerte (com gás argônio). O álcool benzílico foi utilizado até um dia após sua destilação. Este intervalo de tempo não pode ultrapassar o prazo de 3 dias.

Outro solvente utilizado no processo é o clorofórmio, que foi adicionado a cada amostra analisada. Este solvente foi purificado com a adição de cloreto de cálcio e posterior destilação e pode ser utilizado até uma semana após sua purificação.

- *Preparação do Titulante*: Foi utilizada uma solução de NaOH 0,1M como titulante. Esta solução foi preparada com a dissolução de NaOH em álcool benzílico e 10 mL de metanol. A padronização foi realizada com biftalato de

potássio e solução alcoólica de fenolftaléina 1% como indicador. Foi utilizado um banho de ultrassom para garantir a total dissolução da base.

- *Micro-Titulação*: O processo teve início com a preparação do titulado que consiste na secagem prévia de aproximadamente de 0,100 g a 0,200 g de PET a 160°C por 4h em estufa a vácuo. Em seguida aqueceu-se 10 mL de álcool benzílico a 190°C em um bécker. Rapidamente é adicionada a amostra de PET anteriormente pesada. Aqueceu-se esta mistura por exatamente 1min e 50s. Após este intervalo de tempo, o bécker contendo a amostra foi resfriado em um recipiente contendo água a 25°C por 10 segundos. Adicionou-se 10 mL de clorofórmio para evitar a emulsificação da amostra. Utilizou-se vermelho de fenol 0,1% em etanol como indicador. Em seguida realizou-se a titulação com a solução alcoólica de NaOH 0,1M, com o auxílio de uma micro-bureta de 2 mL ($\pm 0,005$). Todo o processo de análise não deve ultrapassar 3 minutos. Deste modo, todas as etapas foram cronometradas.

Para a verificação da pureza do solvente e para efeito de cálculo, que considera ácidos residuais, uma titulação em branco foi efetuada, de maneira que o titulado branco foi constituído de álcool benzílico e clorofórmio. O volume de NaOH não pode exceder 15µL para o branco, caso contrário, estes solventes não têm mais condições de serem utilizados devido à formação de ácidos residuais.

Após efetuada a titulação em triplicata das amostras de PET, a concentração de grupos carboxílicos terminais foi calculada conforme a expressão 4.1:

$$CEG = \frac{(V_{\text{amostra}} - V_{\text{branco}}) \times 10^{-6} \times \text{mol.L}^{-1}_{\text{NaOH}}}{\text{Amostra PET (g)}} \quad (4.1)$$

4.2.3- Determinação da viscosidade intrínseca $[\eta]$

Esta propriedade foi determinada em um ensaio, descrito pela norma ASTM D 4603^[91]. O método é baseado na determinação do tempo de fluxo de uma solução do polímero de concentração conhecida e o tempo de fluxo do solvente puro em um viscosímetro capilar, sob temperatura constante.

O procedimento foi realizado com a dissolução da amostra em uma solução de fenol / 1,1,2,2,-tetracloroetano (60, 40%) a uma temperatura de 30°C.

Inicialmente, o PET foi moído e em seguida foi submetido à secagem a 160°C por 4h em estufa a vácuo (marca Tecnal, modelo EDGCON 5P). Utilizou-se entre 0,2475g a 0,2525g da amostra seca que foi posteriormente dissolvida na solução à temperatura de 110°C e agitação constante. A viscosidade foi determinada, colocando-se a solução em um viscosímetro (Cannon Ubbelohde), em banho termostaticado a 30°C por 15 min.

A viscosidade intrínseca foi calculada com base na viscosidade relativa, através da equação 4.2:

$$[\eta] = 0,25 (\eta_r - 1 + 3 \ln \eta_r) / C \quad (4.2)$$

Onde:

η_r é a viscosidade relativa = t/t_o

t é o tempo de fluxo médio da solução polimérica (s)

t_o é o tempo de fluxo médio do solvente (s)

C é a concentração da solução polimérica

4.2.4- Obtenção dos Espectros FTIR das amostras de PET

Os experimentos **DRIFTS** (acessório JASCO) foram realizados em dois espectrômetros distintos: Bomem DA-8, utilizando-se um detector MCT (HgCdTe) resfriado com N₂ e Bomem MB-100 com um detector menos sensível DTGS (sulfato de triglicina deuterada). Os espectros de referência foram coletados com o porta-amostras do acessório vazio.

Todos os espectros DRIFTS foram medidos na faixa de 4000 a 500 cm⁻¹ com resolução de 4 cm⁻¹ e um média de 256 aquisições. As *amostras secas* do PET moído foram submetidas a um processo de secagem sob vácuo a 160°C por 4h, transferidas a um dessecador e manuseadas em uma “glove box” com N₂ seco. As amostras foram rapidamente transferidas para o compartimento de amostras do Bomem DA-8 e os espectros foram adquiridos sob vácuo.

As amostras secas não foram analisadas no Bomem MB-100, pois este não dispõe de sistema de análise em vácuo.

As *amostras não secas* do mesmo material foram guardadas em um dessecador com um recipiente com água destilada durante 24 horas para que o material absorvesse umidade e pudesse ser comparado ao material seco. Foram obtidos espectros das *amostras não secas* no Bomem DA-8 e no Bomem-100.

Os espectros **PAS** (acessório MTEC 300) foram adquiridos a partir das *amostras não secas* no equipamento Bomem MB-100, na faixa de 4000 a 500 cm⁻¹, com resolução de 8 cm⁻¹ e 256 aquisições com purga de gás hélio. Como espectros de referência foram utilizados espectros de carvão ativo prensado.

A figura 4.1 resume o experimento realizado:

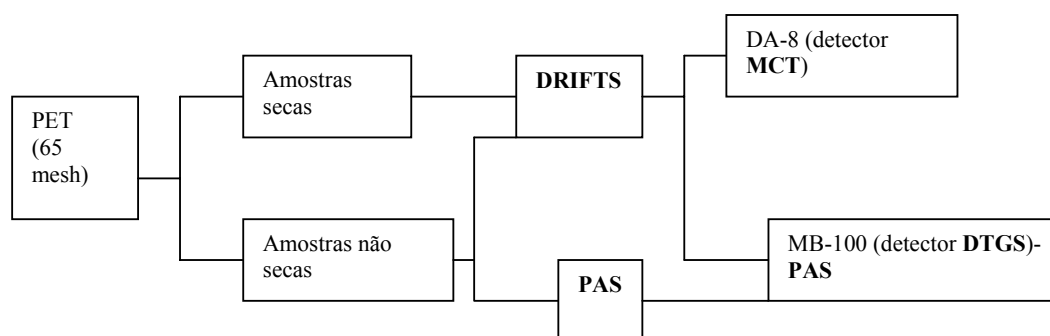


Figura 4.1: Esquema dos experimentos DRIFTS e PAS.

5-ANÁLISE MULTIVARIADA

5.1- INTRODUÇÃO

Os resultados das análises quantitativas podem apresentar alguma dificuldade para serem integralmente compreendidos quando obtidos por métodos instrumentais como a espectroscopia no infravermelho. As dificuldades estão relacionadas na interpretação de espectros e calibração de misturas que têm interferentes que podem absorver em regiões espectrais similares as do analito, assim como na análise de misturas complexas, cujos espectros também podem apresentar bandas sobrepostas, dificultando sua interpretação. Além disso, outros problemas como ruídos espectrais e não-linearidades também podem prejudicar a análise quantitativa.

No entanto, com o advento da quimiometria, o aperfeiçoamento de microcomputadores e programas computacionais específicos, os problemas de modelagem de dados obtidos na espectroscopia no infravermelho foram, em grande parte, superados e a aplicação da técnica em análises quantitativas atualmente proporciona resultados mais precisos.

Para que o tratamento quimiométrico possa ser aplicado, é necessário que os dados do sistema analisado estejam organizado em matrizes. Os espectros individuais das amostras são dispostos em um matriz $\mathbf{X}$ ($n \times m$), na qual as linhas correspondem às amostras e as colunas são aos valores de medidas de intensidade em cada número de onda, denominadas de variáveis *independentes* ^[4]. A disposição dos dados em uma matriz $\mathbf{X}$ pode ser observada na figura 5.1:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}$$

Figura 5.1: Matriz $\mathbf{X}$ ($m \times n$) de variáveis independentes, onde cada linha representa uma amostra e $x_{11}, x_{12}, \dots, x_{1m}$ são as absorvâncias registradas, uma para cada número ou comprimento de onda^[4].

Os valores das concentrações dos analitos ou das propriedades de interesse a serem estimadas são dispostos em uma matriz $\mathbf{Y}$, cujas variáveis são denominadas *dependentes*.

A notação de matrizes e vetores também é um aspecto muito importante na designação dos dados^[3].

- As matrizes são indicadas por letras maiúsculas, em negrito.
- Os vetores são designados por letras minúsculas, em negrito.
- Os valores de escalares não são escritos em negrito.
- O símbolo (T) é designado para representar matrizes ou vetores transpostos.

Os dados obtidos neste trabalho foram submetidos a processos de pré-tratamento e calibração multivariada, com a utilização do programa computacional Unscrambler, versão 7.1 (CAMO).

5.2- PRÉ-TRATAMENTO DE DADOS

Medidas em diferentes unidades e variáveis com diferentes variâncias podem afetar a distribuição e uniformidade dos dados espectrais. Este fato pode ocasionar

dificuldade na análise dos dados, interferindo na extração de informações relevantes, bem como na sua interpretação^[4]. Deste modo, há a necessidade de tratar os dados a serem analisados, seja por meio de pré-processamento ou transformações dos dados originais.

Os espectros adquiridos neste trabalho sofreram algum tipo de operação matemática antes da calibração propriamente dita, objetivando a melhora dos resultados finais.

5.2.1- CENTRAR NA MÉDIA

Existem alguns métodos de pré-processamento que são aplicados nas colunas das matrizes. O tipo mais simples é uma operação que consiste em centrar os dados na média. Este procedimento pode ser aplicado nos dados referentes às propriedades dos analitos e nos dados espectrais. A operação é realizada com a subtração do valor de cada elemento da coluna da matriz pela média de todos os elementos desta coluna^[106].

Este tipo de pré-processamento é o mais indicado para dados espectrais.

5.2.2- DIFERENCIAÇÃO E ALISAMENTO

A diferenciação é um método de transformação de dados que visa corrigir problemas como a variação de linha base, sobreposição de bandas, além de acentuar as informações contidas nos espectros. O método de alisamento reduz os ruídos espectrais.^[106]

O processo de alisamento pode ser efetuado através dos métodos da *média móvel* e do *alisamento polinomial*. O método da média móvel substitui um conjunto de valores por seu valor médio. Como exemplo (figura 5.2), considera-se o

conjunto dos cinco primeiros valores de intensidade do espectro, x_i . A média para os cinco primeiros valores, $i = 1 \dots 5$, fornece um valor médio para o terceiro ponto, x'_3 , no espectro alisado, que corresponde ao ponto central deste grupo. O processo continua iterativamente para os próximos cinco pontos para encontrar x'_4 a partir dos dados originais x_2, x_3, x_4, x_5 e x_6 [107]. Cada grupo é formado, movendo-se um ponto acima do primeiro valor. Desta forma, obtém-se o valor médio para todos os pontos, com exceção dos valores anteriores ao ponto central do primeiro grupo e dos valores posteriores ao ponto central do último grupo, que permanecem os mesmos valores.

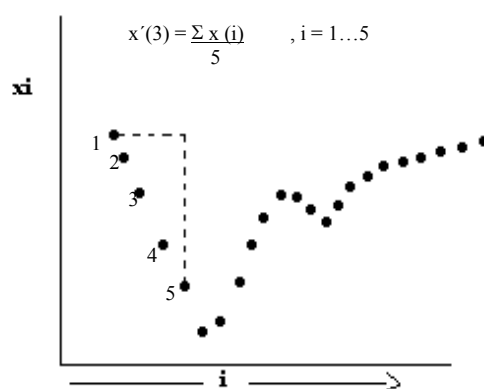


Figura 5.2: Alisamento com média móvel de 5 pontos. Cada novo ponto no espectro alisado é formado pela média de um conjunto de 5 pontos do espectro original [107].

O grau de alisamento é definido pelo número de pontos escolhido, ou seja, da largura da “janela de alisamento”.

O processo matemático da implementação da técnica da média móvel é chamado *convolução*. O espectro resultante, x' (como um vetor), é o resultado da convolução do vetor espectro original, x , com uma função filtro de convolução, w , ou seja:

$$x' = w \otimes x \quad (5.1)$$

Para um processo simples de média móvel com intervalo de 5 pontos, $w = [1,1,1,1,1]$.

No entanto, um outro processo de alisamento com base na média móvel especial foi desenvolvido por Savitzky e Golay^[108]. Neste processo, o filtro de convolução é derivado do método dos mínimos quadrados, empregando um polinômio de grau variável. O alisamento polinomial visa substituir os valores originais por um valor calculado a partir do método de mínimos quadrados polinomial. Desta maneira é feito o ajuste para todos os pontos, e o espectro original apresenta uma redução nos ruídos. Após este procedimento pode-se aplicar as operações de derivação.

Neste trabalho usou-se para os dados dos combustíveis, derivada de 1ª ordem com alisamento Savitzky-Golay com janela de 5 pontos e polinômio de grau variável de 2ª ordem, cujos resultados serão comentados posteriormente.

5.3-ANÁLISE POR COMPONENTES PRINCIPAIS (PCA- Principal Component Analysis)

A Análise de Componentes Principais (PCA) é a base de muitos métodos usados na calibração multivariada^[109]. Além disso, é comumente utilizada como análise exploratória dos dados.

O método PCA pode ser aplicado nas variáveis da matriz $\mathbf{X}$ que apresentam alto grau de colinearidade. Como consequência, informações redundantes e pequena variabilidade de ruídos podem ser removidos^[3].

O propósito da Análise de Componentes Principais é expressar as principais informações das variáveis da matriz $\mathbf{X}$ em um número menor de variáveis, denominadas componentes principais ou fatores^[3].

Através do método PCA, pode-se detectar padrões ou agrupamento de amostras com características semelhantes; estabelecer quais variáveis são mais relevantes para o modelo; identificar variáveis correlacionadas ou redundantes e observar amostras anômalas.

Em termos geométricos, a PCA é um método de projeção que envolve a rotação e transformação dos dados originais (espalhados em um determinado espaço de muitas dimensões), que após a aplicação do método estarão dispostos em novos eixos, fatores ou variáveis latentes, e estarão alinhados na direção de um máximo de variância. Estes novos eixos ou componentes principais são ortogonais entre si. As novas variáveis ou novas coordenadas são resultados da combinação linear das variáveis originais. Deste modo, o método reduz a dimensionalidade dos dados originais^[107].

Os cálculos das componentes principais estão relacionados com a determinação de *autovetores* e *autovalores*, que são extraídos de $\mathbf{X}^T\mathbf{X}$. Os autovalores mostram o quanto de variabilidade cada fator remove da matriz $\mathbf{X}$ e são denotados por $(\tau_1, \tau_2, \dots, \tau_3)$. Os vetores loadings (linhas de $\mathbf{P}$) são os autovetores do produto da matriz $\mathbf{X}^T\mathbf{X}$, que tem τ com autovalores. Isto significa que todo $\mathbf{p}$ satisfaz a relação mostrada na equação 5.2^[3]:

$$\mathbf{X}^T\mathbf{X} \mathbf{p}_a = \mathbf{p}_a \tau_a \quad (5.2)$$

Os autovetores representam os vetores que explicam a máxima variância, que darão a direção dos novos eixos ou fatores. Deste modo, cada componente principal tem um autovetor e um autovalor em particular.

As componentes principais são determinadas com base no critério de máxima variância. Cada subsequente componente principal descreve um máximo de variância, que não é modelada para formar a componente anterior. De acordo com esta afirmação, a primeira componente principal contém mais informação que a segunda componente, que por sua vez, contém mais informação que a terceira, etc.

As componentes principais podem fornecer a percentagem de variância contida em cada fator. Todas as componentes juntas em um modelo fornecem 100% de variância e a maior percentagem é explicada pela primeira componente principal^[110].

O método matemático mais simples para calcular a PCA é o método iterativo NIPALS (nonlinear iterative partial least squares), mas existem outros métodos mais eficazes, baseados na diagonalização de matrizes, como a decomposição do valor singular.

O princípio da PCA é a aproximação da matriz original $\mathbf{X}$ em um produto de duas matrizes menores: a matriz de *loadings* $\mathbf{L}$ e a matriz de *scores* $\mathbf{T}$ ^[110]

$$\mathbf{X} = \mathbf{T} \mathbf{P}^T \quad (5.3)$$

$$\begin{array}{c}
 \begin{array}{|c|} \hline \mathbf{X} \\ \hline \end{array}
 \end{array}
 \begin{array}{c}
 \mathbf{P} \\
 \\
 n
 \end{array}
 =
 \begin{array}{c}
 \begin{array}{|c|} \hline \mathbf{T} \\ \hline \end{array}
 \end{array}
 \begin{array}{c}
 d \\
 \\
 n
 \end{array}
 \begin{array}{c}
 \begin{array}{|c|} \hline \mathbf{P}^T \\ \hline \end{array}
 \end{array}
 \begin{array}{c}
 p \\
 \\
 d
 \end{array}$$

Figura 5.3: Matriz $\mathbf{X}$ de dados originais, com n linhas (objetos) e p colunas (características); $\mathbf{T}$ é a matriz de scores, com n linhas e d colunas (número de componentes principais); $\mathbf{L}$ é a matriz de loadings, com d colunas e p linhas; o $\mathbf{T}$ sobrescrito é a transposta de uma matriz^[110].

As colunas em $\mathbf{T}$ são os vetores *scores* e as linhas em $\mathbf{P}$ são chamadas de vetores *loadings* (ambos os vetores são ortogonais). Em termos geométricos, os *scores* são as novas coordenadas para o novo sistema de eixos e contém informações a respeito das amostras. Os *loadings* fornecem a direção dos eixos ou componentes principais e estão relacionados com informações sobre as variáveis.

A equação 5.3 pode ser escrita como o produto de vetores $\mathbf{t}_i$ e $\mathbf{p}^{[111]}$:

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1^T + \mathbf{t}_2\mathbf{p}_2^T + \dots + \mathbf{t}_a\mathbf{p}_a^T \quad (5.4)$$

No método NIPALS, $\mathbf{t}_1$ e $\mathbf{p}_1^T$ são calculados a partir de $\mathbf{X}$ e o produto $\mathbf{t}_1\mathbf{p}_1^T$ é subtraído de $\mathbf{X}$ pra se obter a matriz residual $\mathbf{E}_1$. Esta matriz de resíduos é então, utilizada para calcular $\mathbf{t}_2$ e $\mathbf{p}_2^T$, conforme mostrado na equação 5.5:

$$\mathbf{E}_1 = \mathbf{X} - \mathbf{t}_1\mathbf{p}_1^T \quad \mathbf{E}_2 = \mathbf{X} - \mathbf{t}_2\mathbf{p}_2^T \quad \dots \quad (5.5)$$

Portanto, o algoritmo NIPALS não calcula todos os valores de *scores* e *loadings* de uma só vez, pois são calculados em um processo iterativo, par a par. (Maiores detalhes a respeito do algoritmo NIPALS, encontram-se no apêndice A)

5.4- CALIBRAÇÃO MULTIVARIADA

Para conhecer a relação matemática entre as matrizes $\mathbf{X}$ e $\mathbf{Y}$ e fazer uso desta para prever as propriedades ou concentrações das amostras, é necessário construir um modelo de previsão e isto é possível através de um processo matemático que normalmente é realizado em duas etapas. Na primeira etapa, as características das informações do método ou instrumento de análise são investigadas e é estabelecido um modelo matemático que representa tal comportamento. Em outras palavras, é construído um modelo que representa a relação $\mathbf{Y} = f(\mathbf{X})$ entre os dois blocos de variáveis (dependentes $\mathbf{Y}$ e independentes $\mathbf{X}$). Este primeiro passo é a fase da *calibração* (ou modelagem), na qual são determinados parâmetros matemáticos do modelo denominados coeficientes de regressão relacionando linearmente os dois blocos de variáveis .

Na segunda etapa, as informações das variáveis independentes e os coeficientes de regressão são utilizados para prever os valores das variáveis dependentes. Esta fase é denominada *previsão* (ou validação)^[4, 111].

A calibração multivariada utiliza simultaneamente várias medidas ($x, x_1, x_2, \dots, x_k$), que formam o conjunto de calibração utilizado para quantificar ou prever as concentrações ou propriedades de interesse (y). Já a calibração univariada utiliza-se apenas de um único valor de x para prever y .

Neste trabalho, foram utilizados dois métodos de calibração multivariada, o PLS e o MLR, os quais será feita uma breve explanação nos próximos tópicos.

5.4.1- REGRESSÃO LINEAR MÚLTIPLA (MLR-Multiple Linear Regression)

O modelo de Regressão Linear Múltipla busca designar um único aspecto da variabilidade dos dados **Y** (propriedades ou concentrações) em relação a cada variável **X** (medidas espectroscópicas) empregada ^[3]. Em outras palavras, busca-se uma relação linear entre as variáveis x e y , determinando-se a solução para o coeficiente **b**.

Este processo pode ser representado matematicamente como^[111]:

$$y = b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_mx_m + e \quad (5.6)$$

Onde y representa as variáveis dependentes; x representa as variáveis independentes; m é o número de variáveis independentes e b é o coeficiente de regressão; e é o erro ou resíduo.

Para n amostras, as variáveis y podem ser escritas como um vetor coluna **y** e as variáveis x podem ser escritas sob forma uma matriz **X**. Deste modo, a expressão assume sua forma mais conhecida, representada pela equação 5.7:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e} \quad (5.7)$$

De acordo com o número de variáveis independentes m e o número de amostras n , há três situações diferentes para se encontrar a solução de **b**:

(1) $m > n$. Há mais variáveis que amostras. Neste caso há um número infinito de soluções para **b**, o que não é desejável;

(2) $m = n$. O número de amostras e variáveis é igual. Esta condição dificilmente é encontrada em situações práticas, entretanto há uma única solução para $\mathbf{b}$, representada pela equação 5.8:

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\mathbf{b} = 0 \quad (5.8)$$

(3) $m < n$. Há mais amostras que variáveis. Não há uma solução exata para $\mathbf{b}$, mas existe uma solução que busca minimizar o comprimento do vetor residual $\mathbf{e}$ através da equação 5.8:

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\mathbf{b} \quad (5.9)$$

A resolução para a equação 5.8 é conhecida como método dos mínimos quadrados, cuja solução é representada pela equação 5.9:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (5.10)$$

Um problema freqüente associado ao método MLR, é que o inverso de $\mathbf{X}^T \mathbf{X}$ pode não existir ou a determinação de $\mathbf{b}$ pode ser muito difícil, originando grandes erros. Este fato é verificado quando existe alta colinearidade em $\mathbf{X}$. Portanto, a técnica exige que o número de variáveis independentes seja reduzido por algum método de seleção de variáveis.

Apesar deste inconveniente, existem vários métodos simples para a seleção de variáveis e, além disso, o método MLR tem a vantagem de ser matematicamente muito simples.

Neste trabalho, o método MLR foi utilizado para construir modelos de calibração para as propriedades dos combustíveis comerciais.

5.4.2- MÍNIMOS QUADRADOS PARCIAIS (PLS – Partial Least Squares)

O método PLS é um método de regressão que fornece modelos que relacionam os blocos de variáveis $\mathbf{X}$ e $\mathbf{Y}$. Deste modo, as informações das medidas espectrais e das concentrações ou propriedades são utilizadas simultaneamente na fase da calibração.

É baseado no método PCA e utiliza as componentes principais para modelar os diferentes analitos e interferências nos dados de $\mathbf{X}$.

A regressão através do método PLS pode solucionar problemas de colinearidade, com satisfatória habilidade de previsão^[3]. Por este motivo esta técnica é adequadamente aplicada em dados espectrais^[112]. As informações das variáveis são comprimidas, tornando-se mais estáveis e como consequência, os modelos são mais fáceis de serem interpretados e os ruídos espectrais podem ser mantidos fora do modelo, na forma de resíduos^[3].

A construção dos modelos PLS baseia-se nas propriedades do algoritmo NIPALS. O modelo simplificado consiste em uma regressão entre os *scores* das matrizes $\mathbf{X}$ e $\mathbf{Y}$. O modelo PLS é considerado primeiramente como uma relação externa entre as matrizes $\mathbf{X}$ e $\mathbf{Y}$ individualmente e subsequenteemente como uma relação interna que relaciona as duas matrizes ($\mathbf{X}$ e $\mathbf{Y}$)^[111].

A relação externa para $\mathbf{X}$ pode ser expressa como a somatória das novas matrizes, originadas da decomposição de $\mathbf{X}$:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} = \sum \mathbf{tp}^T + \mathbf{E} \quad (5.11)$$

E a relação externa de $\mathbf{Y}$ segue o mesmo caminho:

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} = \sum \mathbf{u}\mathbf{q}^T + \mathbf{F} \quad (5.12)$$

Onde: $\mathbf{T}$ e $\mathbf{U}$ são matrizes relacionadas com os *scores* e $\mathbf{P}^T$ e $\mathbf{Q}^T$ estão relacionadas com os *loadings*. $\mathbf{E}$ e $\mathbf{F}$ são as matrizes de resíduos.

A relação interna está baseada na correlação entre os *scores* de $\mathbf{X}$ e $\mathbf{Y}$, obtendo-se a expressão descrita por 5.13:

$$\mathbf{U} = \mathbf{b}\mathbf{T} + \mathbf{e} \quad (5.13)$$

Na qual, $\mathbf{b}$ é o coeficiente de regressão e $\mathbf{e}$ é o vetor de resíduos.

Deste modo, $\mathbf{Y}$ pode ser calculado a partir de $\mathbf{U}$, representado pela equação 5.14 :

$$\mathbf{Y} = \mathbf{b}\mathbf{T}\mathbf{Q}^T + \mathbf{F} \quad (5.14)$$

O número de matrizes em que $\mathbf{X}$ é decomposta é o número de variáveis latentes, componentes principais ou fatores utilizados para modelar as informações do modelo. Se as variáveis latentes para $\mathbf{X}$ e $\mathbf{Y}$ forem calculadas separadamente, pode haver uma relação não satisfatória. Para atenuar este problema e melhorar a relação linear entre $\mathbf{t}$ e $\mathbf{u}$, as variáveis latentes podem ser levemente rotacionadas. No método PLS $\mathbf{u}$ e $\mathbf{t}$ não são calculados independentemente.

Apesar de apresentar um modelo matemático complexo, o método PLS é atualmente muito empregado em análises quantitativas, fornecendo bons resultados.

O método PLS foi utilizado neste trabalho na confecção de modelos de previsão para os parâmetros dos combustíveis e para estimar as propriedades de degradação do PET.

5.5- VALIDAÇÃO DOS MODELOS

A validação de um resultado descreve, em termos gerais, a capacidade que o procedimento analítico adotado possui de fornecer resultados reprodutíveis e seguros para serem utilizados na análise em questão^[110]. Deste modo, a capacidade preditiva de um modelo de calibração deve ser testada. O modelo pode ser validado através da previsão das concentrações dos analitos ou propriedades de algumas amostras que, de preferência, não foram usadas para a elaboração de tal modelo.

Além dos erros obtidos entre os valores de reais e os valores previstos pelo modelo testado, o número e fatores ou variáveis latentes e a presença de amostras anômalas também devem ser considerados^[4].

5.5.1- Escolha do Número de Variáveis Latentes ou Fatores

A determinação do número variáveis latentes é um fator relevante para a qualidade do modelo elaborado quando se utiliza uma técnica de calibração baseada no método PCA, como é o caso do PLS. Existe para cada modelo um número ótimo de fatores, com informações suficientes para o modelamento, sem causar o efeito de superajustamento ou “overfitting”, causado principalmente por um número excessivo de fatores no modelo. O superajustamento é indesejável, pois pode incluir informações irrelevantes como ruídos e interferências e também pode apresentar baixa habilidade de previsão. Por outro lado, a inclusão de um número muito baixo de fatores pode deixar de modelar muitas informações importantes ocasionando o efeito de “underfitting” ou ajustamento inferior.

O número ideal de fatores ou variáveis latentes é dado pelo menor erro de previsão possível, conforme indicado pela flecha na figura 5.4:

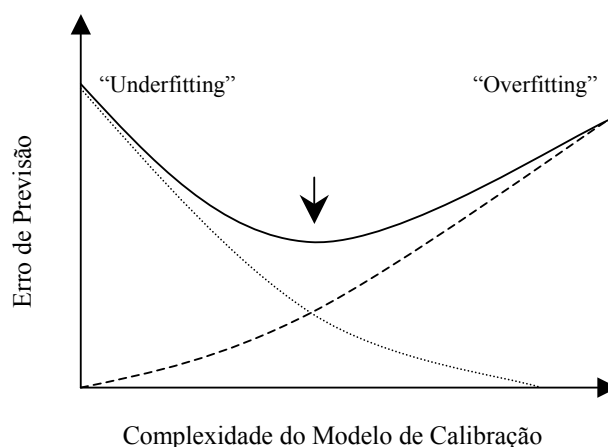


Figura 5.4: Validação do modelo: ilustração do erro de previsão como função da complexidade do modelo de calibração^[3].

Para a avaliação do número ideal de fatores ou variáveis latentes (VL), usa-se como critério a soma dos quadrados dos erros de previsão PRESS (Predictive Residual Errors Sum of Square), dada pela equação 5.15:

$$\text{PRESS} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5.15)$$

Onde y_i é o valor de referência e $\hat{y}_i$ é o valor previsto.

A magnitude de PRESS de cada modelo pode ser obtida através da **validação cruzada**, que é um procedimento utilizado para a escolha do número de fatores. Trata-se de um método de validação interna, cujo procedimento simplificado é descrito a seguir^[4]:

- Uma amostra do conjunto de calibração é removida
- Constrói-se um novo modelo sem a amostra removida

- Usa-se o modelo para fazer a previsão desta amostra.
- Repete-se o procedimento até que todas as amostras, uma por vez, tenham sido excluídas.
- Calcula-se o valor de PRESS

Deste modo, o procedimento é integralmente repetido, variando-se o número de variáveis latentes e comparando-se os valores de PRESS.

A validação cruzada é indicada quando se tem um pequeno número de amostras, caso contrário, usa-se a validação por **conjunto externo**, indicado quando se tem grande número de amostras e deste modo, pode-se dispor de um conjunto de amostras que não estão incluídas no modelo de calibração e que serão utilizadas na etapa da previsão. Esta metodologia é um tipo de validação externa.

No presente trabalho, os dados referentes aos combustíveis foram validados pela validação por conjunto externo, cujo procedimento será mais detalhado na discussão dos resultados. Este tipo de validação foi empregado devido ao grande número de amostras disponíveis. Já os dados referentes à calibração das propriedades do PET foram submetidos à validação cruzada devido ao número reduzido de amostras.

5.5.2- Avaliação dos erros de previsão em amostras teste

A habilidade de previsão de um modelo de calibração pode ser avaliada aplicando-se este modelo para a previsão de amostras teste não incluídas na etapa de calibração (validação interna ou externa). O valor previsto é então comparado ao o valor de referência através do RMSEP (Raiz Quadrada dos Erros Médios de Previsão), dado pela equação 5.16:

$$\text{RMSEP} = \frac{1}{n-1} \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5.16)$$

Onde y_i é o valor de referência e $\hat{y}_i$ é o valor previsto e n é o número de amostras.

Os modelos que fornecem menores valores de RMSEP são selecionados e considerados mais aptos para a previsão de novas amostras.

Uma outra maneira utilizada para se avaliar a habilidade de previsão de um modelo é através do *erro relativo*, dado pela equação 5.17:

$$\text{Erro (\%)} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i| / y_i}{n} \times 100 \quad (5.17)$$

Onde y_i é o valor de referência e $\hat{y}_i$ é o valor previsto e n é o número de amostras.

5.5.3- Detecção de Amostras Anômalas (outliers)

Existem alguns objetos que não se ajustam ao modelo devido a fenômenos inevitáveis que podem acompanhar uma análise quantitativa como erros de operadores, ruídos aleatórios, alterações nas respostas de equipamentos, características peculiares, etc^[3]. A influência destas amostras pode comprometer o desempenho da calibração multivariada. Estas amostras são conhecidas como anômalas ou *outliers*.

A detecção e eliminação de amostras anômalas é um fato que contribui para a qualidade do modelo. Mas é preciso ter critério no descarte de amostras, pois algumas podem conter informações pertinentes ao sistema estudado. Não só amostras irregulares podem ser consideradas *outliers*, variáveis (y) também podem influenciar o modelo.

A identificação de *outliers* em modelos bilineares (PLS, por exemplo) pode ser efetuada observando-se a *influência (leverage)* e os *resíduos* de **X** e **Y**^[3].

A *leverage* é uma medida relacionada com a posição das variáveis independentes (x) em relação uma das outras, ou ao centróide do conjunto de dados e pode ser calculada segundo a equação 5.18^[3]:

$$h_i = 1/l + \sum_{a=1}^k \frac{t_{ia}^2}{\mathbf{t}_a^T \mathbf{t}_a} \quad (5.18)$$

Onde h_i é a *leverage* para o objeto i; $\mathbf{t}_a^T \mathbf{t}_a = \tau_a$ é o autovalor de $\mathbf{X}^T \mathbf{X}$, t_{ia} é o fator e l é o número de objetos.

Formalmente, a *leverage* pode ser interpretada com uma distância de Mahalanobis e indica como uma única variável independente ($\mathbf{x}_i$) está distante do restante das variáveis.

Um valor pequeno de *leverage* significa que o objeto influencia fracamente o modelo de calibração. Se a amostra está distante do centróide do conjunto (alto valor de *leverage*) pode se tratar de uma amostra com concentrações extremas ou valor extremo de um interferente. Por outro lado, pode se tratar de uma variável particularmente informativa^[3].

Para que variáveis informativas não sejam erroneamente descartadas, é necessário avaliar também os resíduos das concentrações $\mathbf{Y}$. Amostras que não se ajustam ao modelo tendem a apresentar altos valores de resíduos.

A influência de cada amostra em particular é obtida através dos Resíduos de Student ou Resíduos Studentizados (r_{ij}) e estes valores podem ser obtidos através de várias técnicas, mas o procedimento mais comum é dividir os resíduos $\mathbf{Y}$ (f_i) por uma estimativa do desvio padrão dos resíduos da regressão, conforme mostra a equação 5.19:

$$r_{ij} = \hat{f}_i / s (1-h_i)^{1/2} \quad (5.19)$$

A detecção de *ouliers* pode ser efetuada com sucesso através da análise do gráfico de resíduos studentizados versus *leverage*. Amostras que apresentam altos valores de resíduos e baixos valores de *leverage* podem ter algum erro no valor da sua concentração. No entanto, amostras com alto resíduo e alta *leverage* devem ser eliminadas.

6- SELEÇÃO DE VARIÁVEIS

6.1- INTRODUÇÃO

As Técnicas Espectroscópicas geralmente originam um grande número de variáveis, tornando o processo de regressão muito demorado. Além disso, algumas variáveis contêm informações que não são pertinentes ao analito (ruídos) ou podem ser colineares (informações redundantes). Adicionalmente, existem regiões espectrais, nas quais as variações das concentrações de um determinado composto não causam variações nos valores de absorbância. Outros problemas do uso de um grande número de variáveis estão relacionados com regiões não lineares, nas quais há desvios da lei de Lambert-Beer devido aos valores muito altos de absorbância. Adicionalmente, alguns métodos de regressão requerem um número de variáveis inferior ao número de amostras como é o caso da Regressão Linear Múltipla^[88].

Deste modo, a habilidade de previsão do modelo de calibração é afetada pelo excesso de variáveis, podendo ter seu desempenho diminuído. Em contrapartida, a escolha das regiões espectrais ou de variáveis específicas pode melhorar a eficiência dos modelos de calibração. Os benefícios da seleção da região espectral adequada relacionam-se com a estabilidade do modelo de calibração em relação à colinearidade, como também na interpretação das relações entre modelo e composição da amostra^[113].

A escolha da região espectral na Espectroscopia no Infravermelho envolve a seleção de um comprimento de onda ou de um subconjunto de comprimentos de onda que resultem em modelos de calibração com menores erros de previsão possíveis.

Existem vários procedimentos para a seleção de regiões espectrais. Alguns destes métodos baseiam-se em critérios objetivos para avaliar desempenho de

determinado conjunto de comprimentos de onda, como a avaliação direta da raiz quadrada dos erros médios de previsão (RMSEP). Outros métodos envolvem a utilização de algoritmos para pesquisa e seleção de regiões espectrais, dentre os quais podem ser citados o Algoritmo Genético e o Algoritmo das Projeções Sucessivas.

Neste trabalho, as regiões espectrais selecionadas utilizadas nos modelos de calibração das propriedades dos combustíveis comerciais foram obtidas a partir dos seguintes métodos:

- Comparação das regiões similares nos gráficos de “*loadings*”, coeficientes de regressão e espectros originais (método PLS):
- Algoritmo Genético (AG)
- Algoritmo das Projeções Sucessivas (APS)

6.3- ALGORITMO GENÉTICO (AG)

Algoritmo genético é um método de otimização em larga escala através do processamento de informações. É inspirado em processos biológicos evolutivos e seu método operacional baseia-se na simulação matemática da Teoria da Evolução das Espécies de Charles R. Darwin (1858), cuja idéia central está fundamentada em um mecanismo de seleção natural, no qual os indivíduos mais adaptados ao meio têm maior chance de sobreviver e constituir um maior número de descendentes, comparativamente aos indivíduos menos adaptados. Conseqüentemente, o processo originaria mudanças nas características e propriedades das populações dos indivíduos, resultantes da evolução biológica. O resultado do processo são as adaptações das espécies em diferentes ambientes^[114, 115].

A análise do processo de seleção natural inspirou pioneiramente John H. Holland na década de sessenta na Universidade de Michigan, a empregar técnicas de otimização em sistemas complexos, que posteriormente receberiam o nome de algoritmo genético.

A aplicação dos algoritmos genéticos em química foi inicialmente explorada por Lucasius e Kateman^[116] na seleção de comprimentos de onda em análises de sistemas multicomponentes. A partir daí, vários pesquisadores têm utilizado este algoritmo na seleção de variáveis de dados espectrofotométricos.

Leardi e Noord^[117] utilizaram o algoritmo genético na seleção de variáveis na construção de modelos de previsão (MLR e PLS) para o teor de hidroxilas de polímeros (poliéteres e polióis) de composição variada. Para tal, os autores selecionaram comprimentos de onda de espectros no infravermelho próximo.

Em dados no infravermelho médio, o algoritmo genético foi aplicado na seleção de variáveis para a determinação simultânea de vários tipos de açúcares, como a glicose, maltose e frutose, mostrando a eficiência do método na análise de multicomponentes, sobretudo quanto à robustez dos modelos de calibração propostos^[118].

Na análise de combustíveis, o AG tem sido aplicado na seleção de comprimentos no infravermelho próximo, com o objetivo de produzir modelos de calibração PLS para a octanagem de amostras comerciais de gasolina^[119]. Com o processo de seleção de variáveis, os autores conseguiram reduzir o número de componentes principais de 9 para 2 e ainda houve diminuição do valor de RMSEP de 0,9 para 0,3 para a propriedade de octanagem RON.

O algoritmo genético básico ou simples envolve as seguintes etapas: Codificação das Variáveis, Criação da População Inicial, Avaliação das Respostas (aptidão), Cruzamento ou Reprodução Sexual e Mutação.

- *Codificação das Variáveis:*

No algoritmo genético as informações ou parâmetros do sistema a ser analisado são codificados de maneira análoga aos genes nos cromossomos. As informações ou parâmetros são representados por números binários, conforme mostras a figura 6.2:

1 1 0 1 0	0 0 1 1 0 1 1	0 0 1 0	0 1 1 0 1 0 0 1 0
x_i	x_j	x_k	x_l

Figura 6.1: Exemplo de uma representação binária de um indivíduo codificado em algoritmo genético^[115].

No exemplo da figura 6.1, cada uma das variáveis x_i , x_j , x_k e x_l representam um gene no cromossomo (indivíduo). Cada gene pode ser atribuído a um determinado parâmetro de um sistema químico a ser otimizado, como por exemplo, temperatura, catalisador, agitação, tempo, etc. Nos casos de seleção de números ou comprimentos de onda, cada um dos genes representa um determinado comprimento de onda. Desta forma fica ilustrada tal similaridade com o processo evolutivo biológico.

- *Criação da População Inicial*

A população inicial de cromossomos ou indivíduos é gerada aleatoriamente, evitando-se uma influência tendenciosa do meio externo. Em alguns casos, para minimizar o tempo de processamento computacional, podem ser introduzidas informações adicionais conhecidas.

A figura 6.2 mostra a um exemplo de geração aleatória de 4 indivíduos ou cromossomos:

Cromosso 1	1 0 0 0 1 1	0 0 1 0 0 0	1 0 1 0 0 1	1 1 1 0 1 0
Cromosso 2	1 0 0 0 1 0	0 0 1 0 0 1	1 0 1 0 1 1	1 1 1 0 0 0
Cromosso 3	1 0 0 0 0 0	0 0 1 0 1 1	1 0 1 0 0 1	1 1 1 0 1 0
Cromosso 4	1 0 0 0 1 1	0 0 1 0 1 0	1 0 1 0 1 0	1 1 1 0 1 1

Figura 6.2: Conjunto de indivíduos representando uma população inicial.

- Avaliação da Resposta (aptidão)

A avaliação da resposta ou aptidão indica a habilidade de cada indivíduo de produzir a melhor resposta ou menor erro entre o valor esperado e o valor previsto. O valor associado ao desempenho de cada indivíduo deve ser encontrado nesta etapa. O término do processo pode ser definido por um número máximo de gerações.

Em alguns algoritmos, a resposta determinará a probabilidade com que um determinado cromossomo irá cruzar, a fim de se estabelecer uma convergência da solução. Em alguns casos são implementados dispositivos pra redefinir o número limite de gerações.

- Cruzamento ou Reprodução Sexual

Este processo determina como ocorre o cruzamento ou recombinação, isto é, a troca de genes entre indivíduos para gerar uma nova população. A partir da população inicial, forma-se uma nova população através da recombinação aleatória dos genes dos cromossomos ou indivíduos. No processo de cruzamento, os pares de cromossomos e os pontos de ruptura também são escolhidos aleatoriamente, seguido da troca de material genético.

Durante o processo de cruzamento, as características dominantes são transmitidas para as gerações futuras. Deste modo, há uma convergência para uma situação otimizada. Em outras palavras, o conjunto de variáveis que apresenta um maior número de informações relevantes para a modelagem do sistema deverá permanecer no modelo através de funções de seleção.

-Mutação

Em uma pequena parcela da população as informações contidas dentro genes podem ser alteradas (alterações no código genético) de uma forma favorável à otimização do sistema. Este processo é conhecido como mutação e ajuda a solucionar o problema do confinamento a ótimos locais.

Na linguagem binária, a mutação pode ser representada pela troca do bit 1 pelo bit 0 ou vice-versa em um dos genes de um determinado cromossomo, conforme exemplificado na figura 6.3:

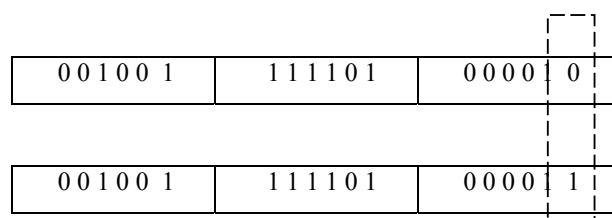


Figura 6.3: Representação do processo de mutação no algoritmo genético: o bit 0 é substituído pelo bit 1, gerando um novo cromossomo.

As principais etapas para o algoritmo genético básico são sumarizadas na figura 6.4:

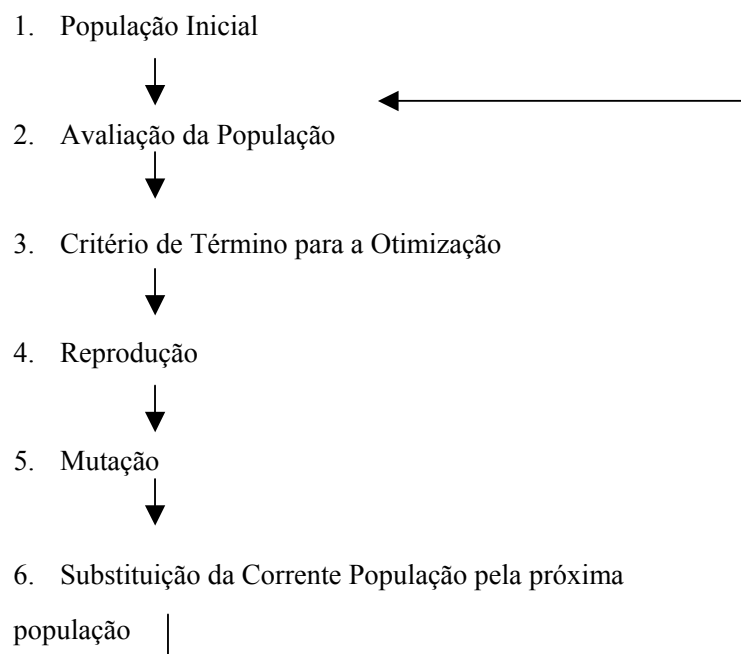


Figura 6.4: Fluxograma mostrando as principais etapas do algoritmo genético básico^[114].

As etapas 2 a 6 são repetidas até que um critério de parada seja obedecido. Normalmente este critério está relacionado com um número máximo de gerações ou até que um erro mínimo seja alcançado.

No caso de seleção de variáveis em análise espectrofotométricas, o algoritmo genético pode indicar quais são as variáveis do espectro de intensidades (absorbância) que melhor representam o analito de interesse. Particularmente, nos espectros da espectroscopia no infravermelho, são selecionados os números de onda (eixo x) que realmente representam o composto analisado.

Conforme anteriormente mencionado, os números de onda do espectro correspondem aos genes do cromossomo, que por sua vez, é representado pelo espectro total. Deste modo, o cromossomo terá o mesmo número de variáveis do espectro original. Cada gene pode assumir um valor binário de 0 (não selecionado) ou 1 (selecionado). Desta forma fica estabelecido quando o número de onda será selecionado ou não até que o critério de parada seja alcançado.

O número máximo de números de onda selecionados pode ser determinado pelas condições de contorno, que podem ser pré-determinadas dependendo do algoritmo.

As condições de contorno também definem outras características desejadas na aplicação do algoritmo genético como o número de indivíduos da população inicial, taxas de cruzamento e mutações e número máximo de gerações.

No presente trabalho foram utilizadas populações de 200 a 300 indivíduos (dependendo do modelo), número de gerações de 100, taxa de cruzamento de 60 % e taxa de mutação de 10% para os modelos de calibração das propriedades dos combustíveis. Foi utilizado para este procedimento o ambiente computacional Matlab versão 6.1. O algoritmo foi desenvolvido por Roberto Kawakami Harroup Galvão (ITA), baseado no modelo original do Algoritmo Genético.

6.2- ALGORITMO DAS PROJEÇÕES SUCESSIVAS (APS)

O Algoritmo das Projeções Sucessivas é um método de natureza determinística de seleção de variáveis, cuja metodologia utiliza operações simples de projeções de vetores e tem o objetivo de minimizar a colinearidade das variáveis originais. Os resultados são bastante reprodutíveis, robustos e o APS pode ser empregado com sucesso e rapidez na seleção de comprimentos de onda para a

calibração multivariada. O método seleciona os comprimentos de onda que produzem os menores valores de RMSEP na previsão do parâmetro de interesse através de um modelo MLR^[120].

A descrição formal do APS^[120] é mostrada no Apêndice B. Informalmente, o método pode ser descrito como um processo de seleção iterativo que opera nas repostas instrumentais da matriz $\mathbf{X}$, cujas linhas são as amostras de calibração e as colunas são as variáveis espectrais. O processo tem início com a escolha aleatória de um vetor coluna $\mathbf{x}_0$. Subseqüentemente, o APS determina qual das colunas (vetores) remanescentes contém a maior projeção em um subespaço (plano) S_0 ortogonal a $\mathbf{x}_0$. Esta coluna, denominada $\mathbf{x}_1$ pode ser considerada aquela que contém a maior quantidade de informações, excluindo-se $\mathbf{x}_0$. A escolha do subespaço a cada iteração é feita, selecionando-se apenas variáveis não colineares. Na próxima iteração, o algoritmo restringe a análise ao subespaço S_0 , tomando-se $\mathbf{x}_1$ como a nova coluna ou vetor de referência, e o processo se repete. A cada iteração, um vetor de comprimento de onda é incluído. O melhor vetor e o critério de parada (número de variáveis a ser selecionado) podem ser otimizados de modo que a capacidade de previsão do modelo seja acentuada.

Cada conjunto de variáveis selecionadas é utilizado para construir um modelo de calibração MLR que será aplicado em um conjunto de amostras de validação para a previsão da propriedade de interesse. Deste modo calcula-se o RMSEP.

A figura 6.5 mostra um exemplo da aplicação do APS, ilustrando a primeira iteração, tomando $\mathbf{x}_3$ como vetor inicial e obtendo-se o vetor de maior projeção $\mathbf{x}_1$:

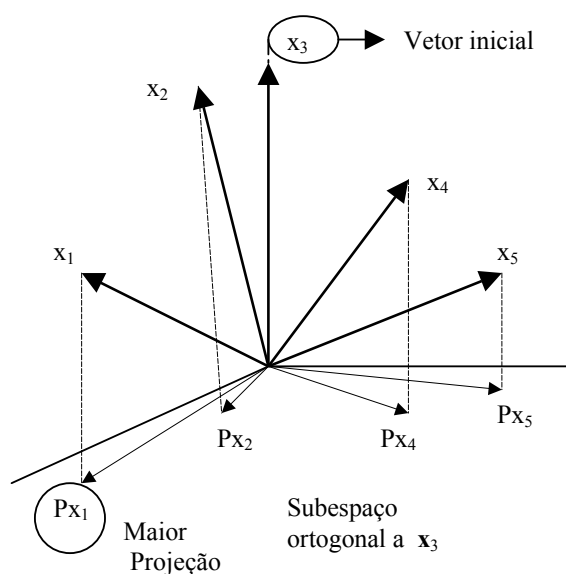


Figura 6.5: Exemplo da aplicação do APS^[120].

O APS não modifica os vetores dos dados originais, uma vez que as projeções são utilizadas somente com propósito de seleção.

O Algoritmo das Projeções Sucessivas é um método de seleção recente e tem sido usado com sucesso para seleção de variáveis na análise espectroscópica de multicomponentes em dados de espectrometria UV-VIS, na seleção de comprimentos de onda^[120], em dados de espectrometria de emissão de plasma, na determinação de vários metais (Mn, Mo, Cr, Ni e Fe)^[121].

O método também foi empregado na espectroscopia no infravermelho, particularmente no infravermelho próximo para a determinação do teor de enxofre em amostras de diesel^[85]. Os resultados mostraram que os valores de RMSEP obtidos a partir dos modelos MLR , com variáveis selecionadas por APS foram inferiores aos valores de RMSEP obtidos por MLR e PLS, com seleção efetuada por algoritmo genético.

Recentemente, o método APS tem sido aplicado também para a seleção de amostras representativas ao conjunto de calibração^[123], demonstrando que o método não é restrito somente à seleção de variáveis.

Neste trabalho o APS foi aplicado para a seleção de números de onda dos dados referentes à análise de combustíveis. Para tal utilizou-se o programa computacional Matlab versão 6.1. O algoritmo foi desenvolvido por Roberto Kawakami Harroup Galvão (ITA).

7. RESULTADOS E DISCUSSÕES

7.1- RESULTADOS DA APLICAÇÃO I: ANÁLISE DE COMBUSTÍVEIS COMERCIAIS

7.1.1- Espectros FTIR-ATR dos Combustíveis

Os espectros com Correção de Linha Base e em Primeira Derivada de algumas amostras de *diesel* obtidas através da técnica de ATR são mostrados na figura 7.1.1:

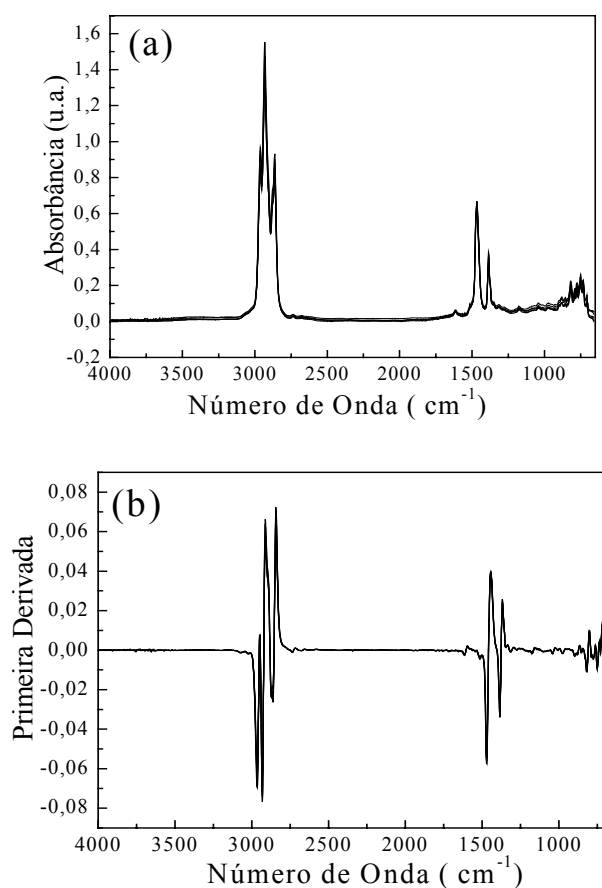


Figura 7.1.1: Espectros ATR no infravermelho médio (4000 a 650 cm^{-1}) para amostras de diesel: (a) Correção de linha base. (b) Primeira Derivada após alisamento Savitzky-Golay.

Os espectros originais apresentaram pequena variação de linha base. Para corrigir este problema, dois tipos de tratamento foram aplicados em modelos distintos: modelos com correção de linha base e modelos com derivada de 1ª ordem com alisamento Savitzky-Golay com janela de 5 pontos e polinômio de 2ª ordem. Deste modo, foram construídos modelos de dois tipos. O mesmo procedimento foi adotado para os espectros da gasolina e álcool, mostrados nas figuras 7.1.2 e 7.1.3, respectivamente:

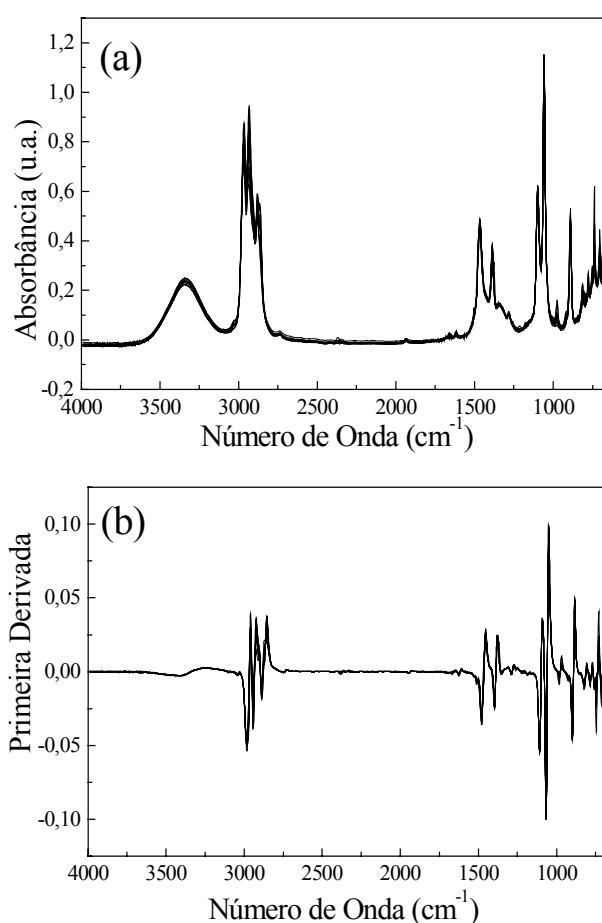


Figura 7.1.2: Espectros ATR no infravermelho médio (4000 a 650 cm^{-1}) para amostras de gasolina: (a) Espectros com correção de linha base. (b) Primeira Derivada após alisamento Savitzky-Golay.

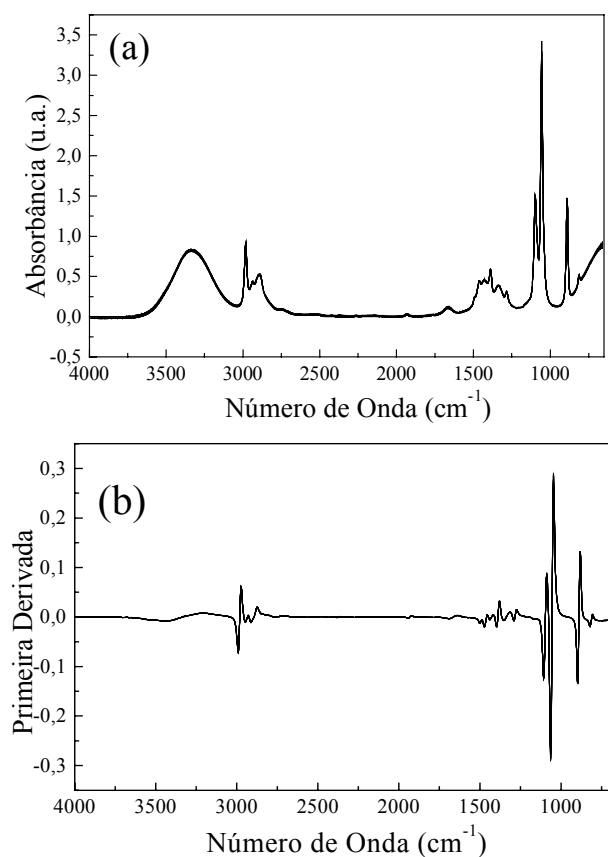


Figura 7.1.3: Espectros ATR no infravermelho médio (4000 a 650 cm^{-1}) para amostras de álcool automotivo: (a) Espectros com correção de linha base. (b) Primeira Derivada com alisamento Savitzky-Golay.

A aplicação da primeira derivada foi efetuada com o objetivo de minimizar os efeitos da variação de linha de base e acentuar os sinais mais fracos para auxiliar no desempenho dos modelos de calibração. O processo de alisamento Savitzky-Golay foi empregado para atenuar os efeitos dos ruídos, que por sua vez são amplificadas com a aplicação da primeira derivada.

Para efeitos comparativos, a correção de linha base foi empregada em outros modelos de calibração, sem aplicação da primeira derivada. O desempenho de tais modelos, quanto capacidade preditiva serão mostrados e discutidos oportunamente.

7.1.2- Análise Exploratória Através do Método PCA

O método PCA foi aplicado inicialmente para observar se há distinção (separação em classes) dos tipos de combustíveis estudados, uma vez que amostras de combustíveis aditivados ou *não conformes* (com propriedades fora dos limites estabelecidos pela ANP) foram estudadas simultaneamente.

A figura 7.1.4 mostra o gráfico de *scores* da primeira e segunda componentes principais, obtido para as amostras de diesel :

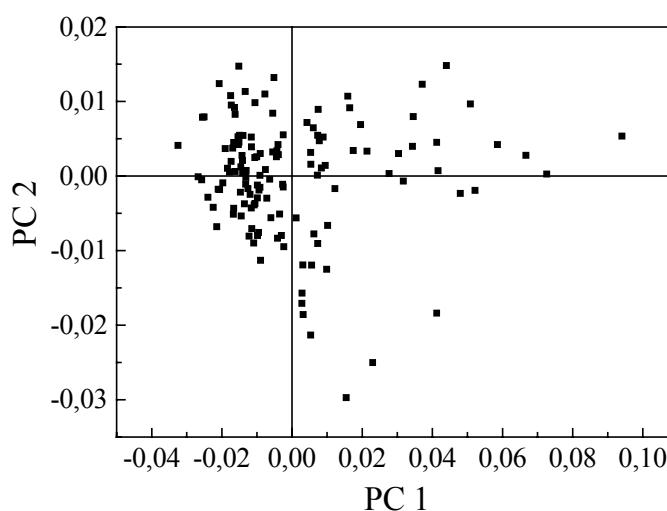


Figura 7.1.4: Gráfico de *scores* para amostras de diesel obtido pela análise PCA.

Por meio do gráfico de *scores* foi possível observar que não há distinção de classes entre os tipos de diesel estudados. Deste modo, os modelos de calibração desenvolvidos neste trabalho contém todos os tipos de amostras de diesel, com exceção dos modelos construídos para a previsão do número de cetano, do qual foram excluídas as amostras aditivadas.

A tabela 7.1.1 mostra os valores de variância explicada para as dez primeiras componentes principais:

Tabela 7.1.1: Porcentagem de variância obtida por PCA para amostras de diesel.

Componente Principal	(%) Variância desta PC	(%) Variância Total
1	80,34	80,34
2	9,81	90,15
3	3,68	93,83
4	2,10	95,93
5	0,99	96,92
6	0,60	97,52
7	0,55	98,07
8	0,37	98,44
9	0,30	98,74
10	0,23	98,97

Para a realização da análise exploratória das amostras da gasolina, tomou-se inicialmente 200 amostras de gasolina do tipo C comum, aditivada e Premium, excluindo-se amostras com suspeita de adulteração e amostras não conformes, para facilitar a distinção das classes, uma vez que as amostras adulteradas, em alguns casos, comportam-se como amostras anômalas. O gráfico de *scores* para o grupo de 200 amostras de gasolinas, sem suspeita de adulteração é mostrado na figura 7.1.5:

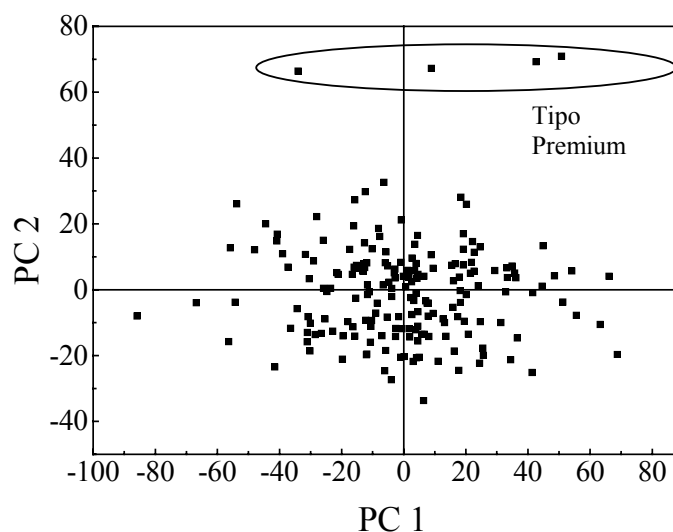


Figura 7.1.5: Gráfico de *scores* para amostras de gasolina sem suspeita de adulteração obtido pela análise PCA. As amostras dentro do círculo correspondem à gasolina do tipo Premium.

O gráfico de *scores* das duas primeiras componentes principais das amostras de gasolina mostra a separação das amostras de gasolina do tipo Premium. Deste modo, estas amostras não foram incluídas nos modelos de previsão para as propriedades da gasolina.

A figura 7.1.6 mostra o gráfico de *scores* para amostras de gasolina, com a inclusão de 200 amostras com suspeita de adulteração no conjunto original das 200 amostras sem suspeita de adulteração:

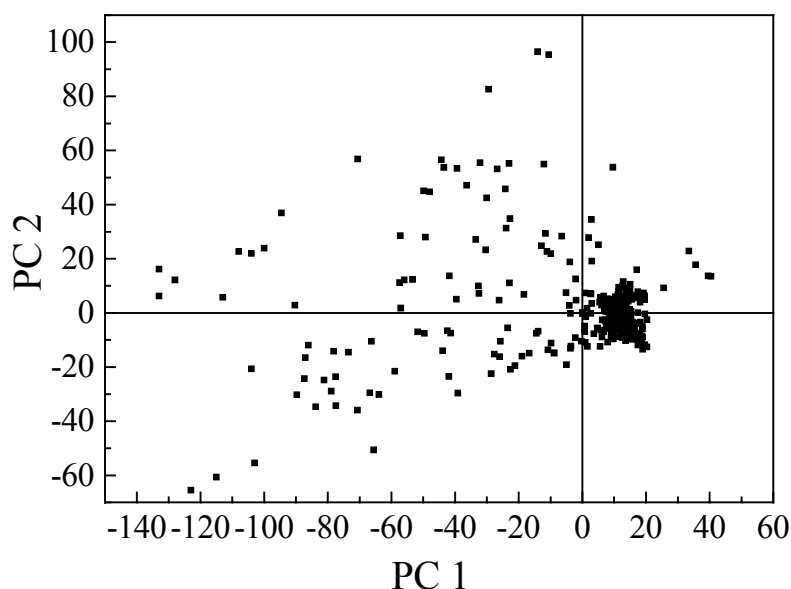


Figura 7.1.6: Gráfico de *scores* obtido pela análise PCA para amostras de gasolina com inclusão das amostras com suspeita de adulteração.

Através do gráfico de *scores* (figura 7.1.6), observou-se que algumas amostras com suspeita de adulteração e amostras não-conformes encontram-se mais espalhadas e distantes do centróide do conjunto de amostras conformes. No entanto, algumas amostras não conformes ou com suspeita de adulteração foram localizadas muito próximas às amostras conformes. Portanto, a eliminação ou inclusão destas amostras dos modelos de previsão foi realizada com mais cautela, que será discutido oportunamente.

A tabela 7.1.2 mostra a porcentagem de variância explicada pelas dez primeiras componentes principais para os conjuntos com e sem a inclusão de amostras não conformes:

Tabela 7.1.2: Porcentagem de variância obtida por PCA para amostras de gasolina.

Componente Principal	Amostras Conformes		Inclusão das amostras não conformes	
	(%) Variância desta PC	(%) Variância Total	(%) Variância desta PC	(%) Variância Total
1	45,52	45,52	56,22	56,22
2	18,4	63,92	16,96	73,18
3	14,17	78,09	12,49	85,67
4	7,06	85,15	6,60	92,27
5	3,22	88,37	2,28	94,55
6	2,99	91,36	1,77	96,33
7	2,68	94,04	1,30	97,63
8	1,24	95,28	0,47	98,10
9	1,09	96,37	0,38	98,48
10	0,84	97,21	0,25	98,73

A análise de componentes principais para as 103 amostras de álcool automotivo do tipo AEAC (álcool etílico anidro combustível) é representada pelo gráfico de *scores* mostrados na figura 7.1.7:

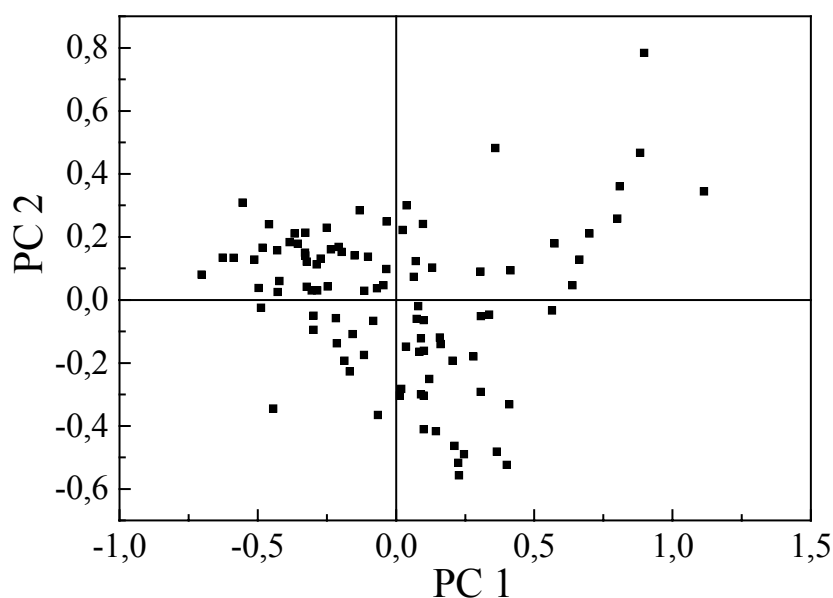


Figura 7.1.7: Gráfico de *scores* obtido pela análise PCA para amostras de álcool automotivo.

Através do gráfico de *scores* das duas primeiras componentes principais, pode-se observar que não houve distinção em classes entre amostras conformes e não-conformes. Portanto todas as amostras foram incluídas nos modelos de previsão para as propriedades deste combustível.

A percentagem de variância explicada para as dez primeiras componentes principais para a análise exploratória do álcool é mostrada na tabela 7.1.3:

Tabela 7.1.3: Porcentagem de variância obtida por PCA para amostras de diesel.

Componente Principal	(%) Variância desta PC	(%) Variância Total
1	53,93	53,93
2	23,16	77,09
3	6,28	83,87
4	5,53	89,40
5	4,20	93,60
6	2,44	96,04
7	1,67	96,92
8	1,09	97,71
9	0,54	98,25
10	0,41	98,66

7.1.3- Modelos para a Previsão das Propriedades dos Combustíveis

Foram elaborados modelos de calibração através dos métodos PLS e MLR com seleção de variáveis (AG e APS) para cada parâmetro estudado das amostras comerciais dos combustíveis (gasolina, diesel e álcool automotivo), conforme mostrado na Tabela 7.1.4:

Tabela 7.1.4: Características dos modelos utilizados para a previsão das propriedades dos combustíveis comerciais:

Modelos PLS	Pré-tratamento	Seleção de variáveis
<i>PLS (A)</i>	Correção linha base	-
<i>PLS (B)</i>	1ª Derivada com alisamento Savitzky Golay	-
<i>PLS (C)</i>	1ª Derivada com alisamento Savitzky Golay	Algoritmo Genético
Modelos MLR	Pré-tratamento	Seleção de variáveis
<i>MLR (A)</i>	1ª Derivada com alisamento Savitzky Golay	Algoritmo Genético
<i>MLR (B)</i>	1ª Derivada com alisamento Savitzky Golay	Algoritmo das Projeções Sucessivas

Deste modo, foram construídos cinco modelos para cada parâmetro dos combustíveis, com os dados centrados na média e validação por conjunto externo. Os modelos *PLS(A)* e *PLS(B)* não foram submetidos aos algoritmos de seleção de variáveis AG e APS, mas as regiões sem informação (coincidentes com a linha base foram descartadas) nos modelos elaborados. Os modelos *PLS (C)* foram construídos a partir de números de onda selecionados pelo algoritmo genético. Estes mesmos números de onda selecionados foram empregados na elaboração dos modelos *MLR (A)*. Os modelos *MLR (B)* foram construídos a partir dos números de onda selecionados pelo Algoritmo das Projeções Sucessivas (APS).

Com o objetivo de facilitar a comparação da habilidade de previsão dos modelos, estes, foram construídos e validados com os mesmos conjuntos de amostras para cada propriedade prevista, sendo que as amostras foram pré-selecionadas da seguinte maneira:

1. *Identificação de amostras anômalas*: inicialmente foi construído um modelo PLS com todas as amostras para identificar amostras anômalas. Este modelo PLS com validação cruzada foi aplicado somente para este fim. As amostras anômalas foram identificadas e eliminadas mediante a observação dos gráficos de resíduos dos dados espectrais X e dos dados de referência Y em função da *leverage*. Amostras com valores altos de *leverage* e resíduos influenciam negativamente no modelo, interferindo no processo de calibração. Deste modo, a exclusão das amostras foi efetuada com a eliminando-se amostras com valores extremos de *leverage* e resíduos. As amostras com altos valores de *leverage*, mas com pequeno valor residual podem conter informações pertinentes ao modelo de calibração, portanto sua remoção foi estudada com mais cuidado. As amostras dispostas grosseiramente fora da reta de regressão também foram consideradas na identificação de *outliers*.

Em alguns casos, as amostras anômalas podem ser visualizadas no gráfico de *scores*, pois se apresentam isoladas do conjunto de amostras.

A exclusão de amostras anômalas em relação a cada parâmetro de cada combustível foi efetuada separadamente, uma vez que estas podem estar associadas com os valores de referência de cada propriedade. A figura 7.1.8 ilustra um exemplo de identificação de possíveis amostras anômalas na previsão da densidade do diesel:

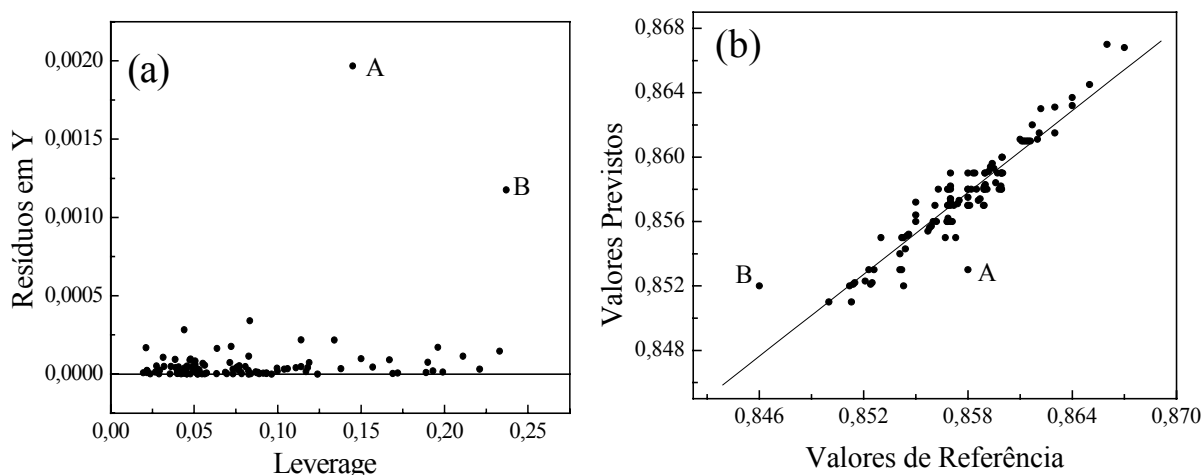


Figura 7.1.8: Identificação de amostras anômalas: (a) Resíduos em Y em função da *leverage*. (b) Reta de regressão.

A amostra “A”, de acordo com a figura 7.1.8(a), apresenta um valor residual muito alto, indicando um possível erro em sua medida de referência. A amostra “B” tem simultaneamente valores altos de resíduo e *leverage*, indicando que deve ser excluída do conjunto de calibração.

As amostras inicialmente indicadas como anômalas foram excluídas do modelo e uma nova calibração foi feita. Subseqüentemente, os valores de RMSEP obtidos antes e depois da exclusão de tais amostras foram comparados para avaliar tal eliminação.

2- Determinação das séries de calibração, validação interna e validação externa: As amostras foram subdivididas em 3 séries ou conjuntos: calibração, validação interna (teste) e validação externa (previsão). A seleção de amostras para o conjunto de calibração é um fator determinante, uma vez que esta série deve abranger todos os valores (baixos, médios e altos) do parâmetro de interesse, tornando-se o mais representativo possível para refletir a habilidade de previsão do

modelo em toda faixa de valores dos parâmetros estudados. A seleção foi efetuada de maneira que houvesse um número parecido de amostras do início ao final de cada faixa de valores do parâmetro estudado. Naturalmente existe um maior número de amostras em torno de um valor médio. Não foram construídos modelos separadamente para os diferentes tipos de diesel e gasolina e as amostras com suspeitas de adulteração foram incluídas em todos os conjuntos. Para todos os parâmetros de todos os combustíveis foi aplicado o mesmo procedimento.

Na etapa da calibração foram construídos modelos PLS com validação externa, utilizando-se os conjuntos de calibração e validação interna (teste). Estes modelos foram utilizados na previsão das propriedades das amostras de validação externa. Deste modo, as amostras do conjunto de validação externo não entraram na etapa da construção do modelo.

Foram considerados para avaliação da habilidade de previsão, os valores de RMSEP e coeficientes de correlação obtidos para o conjunto de validação externa, cujos resultados da previsão das propriedades dos combustíveis serão comentados oportunamente.

7.1.3.1- Modelos PLS e MLR para as propriedades do Diesel

-Densidade

A faixa de valores de densidade para as amostras de diesel analisada foi de 0,8489 a 0,8626 g/cm³, cujos valores foram distribuídos nos conjuntos de calibração (44 amostras), validação interna ou teste (31 amostras) e validação externa ou previsão (53 amostras).

Para os modelos PLS (com correção de linha base e primeira derivada e com variáveis selecionadas por AG) o número de variáveis latentes foi selecionado de acordo com o menor valor de RMSEV, conforme mostrado na figura 7.1.9:

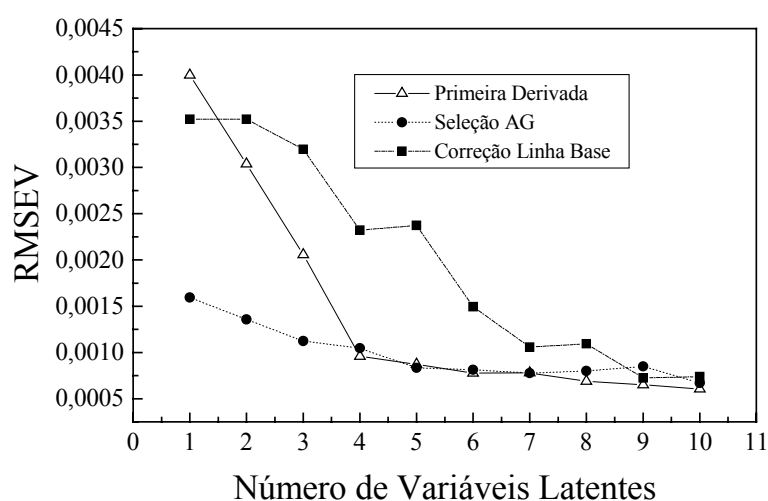


Figura 7.1.9: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para densidade do diesel com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

Para o modelo PLS com correção de linha base, foi possível observar que o número de nove variáveis latentes corresponde ao menor RMSEV. Já para o modelo com primeira derivada, notou-se que o valor de RMSEV não diminui

intensamente a partir de seis variáveis latentes, sendo que a inclusão de uma sétima variável pode ocasionar a superestimação do modelo. O mesmo acontece para o modelo com seleção AG, no qual a inclusão da sexta variável latente não contribui para a diminuição de RMSEV.

Deste modo, determinou-se que o modelo com correção de linha base pode ser descrito com 9 variáveis latentes e o modelo com primeira derivada com 6 variáveis latentes, indicando que a aplicação da derivada pode diminuir o número de fatores usados na calibração. A seleção AG permitiu que o modelo confeccionado a partir das 31 variáveis selecionadas utilizasse apenas 5 variáveis latentes.

O desempenho dos modelos foi avaliado por comparação direta dos valores de RMSEP obtidos entre os valores previstos pelos modelos PLS e MLR de previsão para o conjunto de validação externa e os valores de referência.

A Figura 7.1.10 mostra os gráficos de valores previstos pelos modelos PLS e valores de referência (obtidos de acordo método padrão pra determinação da densidade a 20° C do diesel).

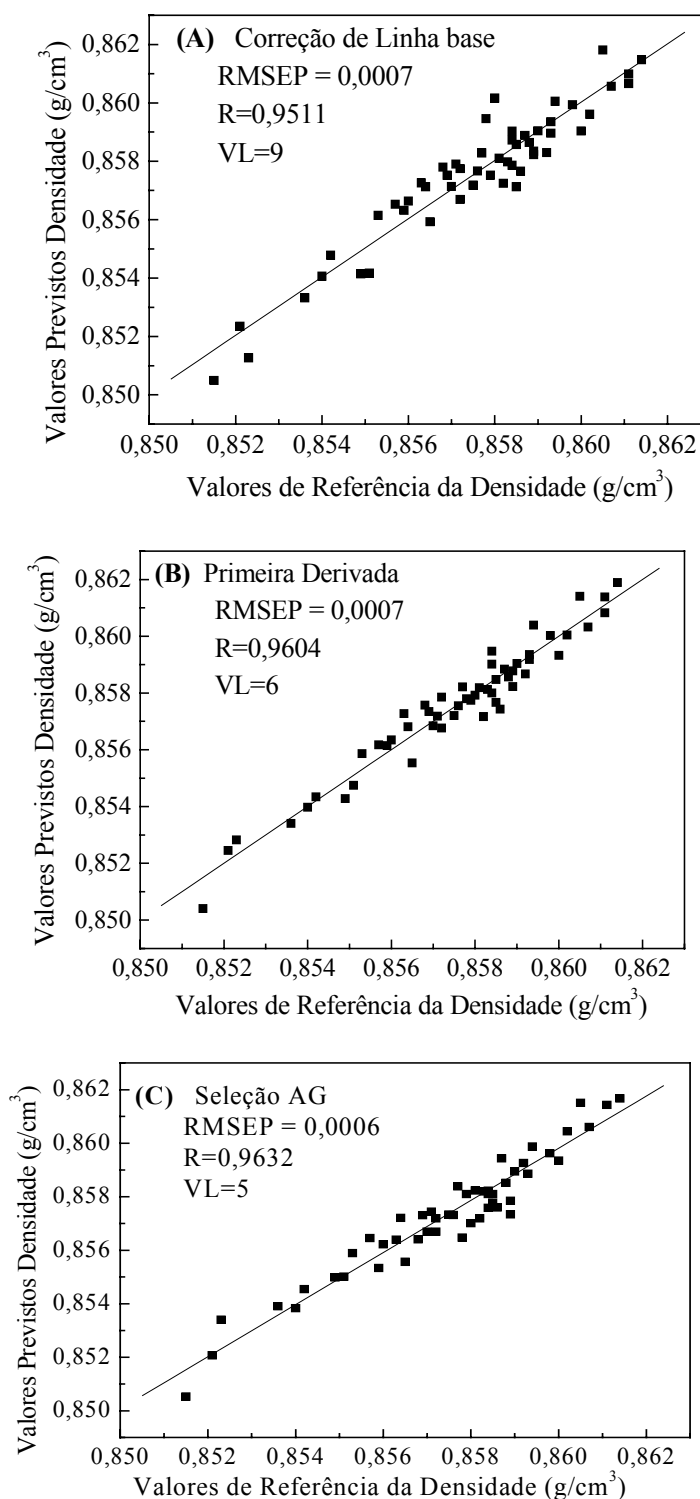


Figura 7.1.10: Comparação entre valores de referência e valores previstos por modelos PLS para a densidade (diesel): (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

Os valores de RMSEP obtidos através dos modelos PLS mostram-se muito próximos. No entanto o menor RMSEP foi observado para o modelo elaborado a partir de 31 números de onda pelo método AG, cujo número de variáveis latentes foi inferior aos demais modelos.

Para o modelo PLS com correção de linha base e primeira derivada foram utilizados 912 e 902 números de onda, respectivamente.

As 31 variáveis utilizadas na elaboração do modelo PLS 7.1.10(C) foram utilizadas na elaboração de um modelo MLR, do qual foi obtido um valor de RMSEP em relação ao mesmo conjunto de previsão empregado em todos os modelos. Um segundo modelo MLR foi elaborado a partir de 43 variáveis selecionadas a partir do método APS que fornece um gráfico de RMSEP em função do número de variáveis independentes, mostrados na figura 7.1.11:

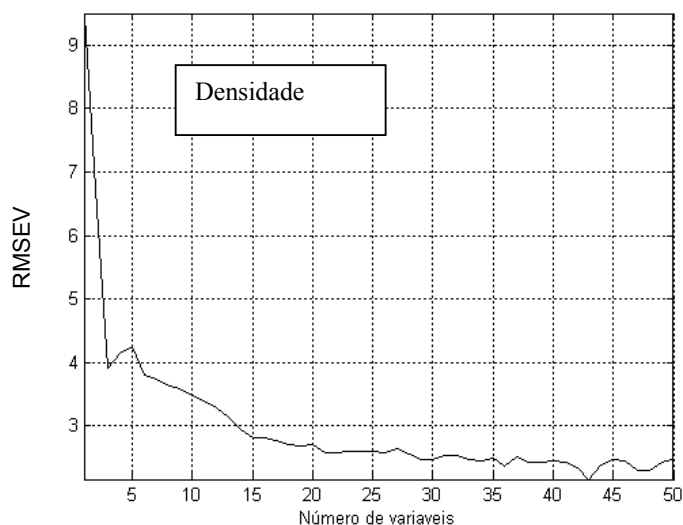


Figura 7.1.11: Resultado da seleção APS para a densidade do diesel. O menor valor de RMSEP corresponde ao número de 43 variáveis selecionadas

O método APS foi efetuado utilizando-se como condições de contorno um número máximo de 50 variáveis a serem selecionadas e um número mínimo de 1 variável. O algoritmo foi efetuado no conjunto de calibração e teste. As variáveis

resultantes (43 números de onda) foram aplicadas no método MLR para as amostras idênticas ao conjunto de previsão do método PLS. Os gráficos dos valores previstos em função dos valores de referência são mostrados na figura 7.1.12:

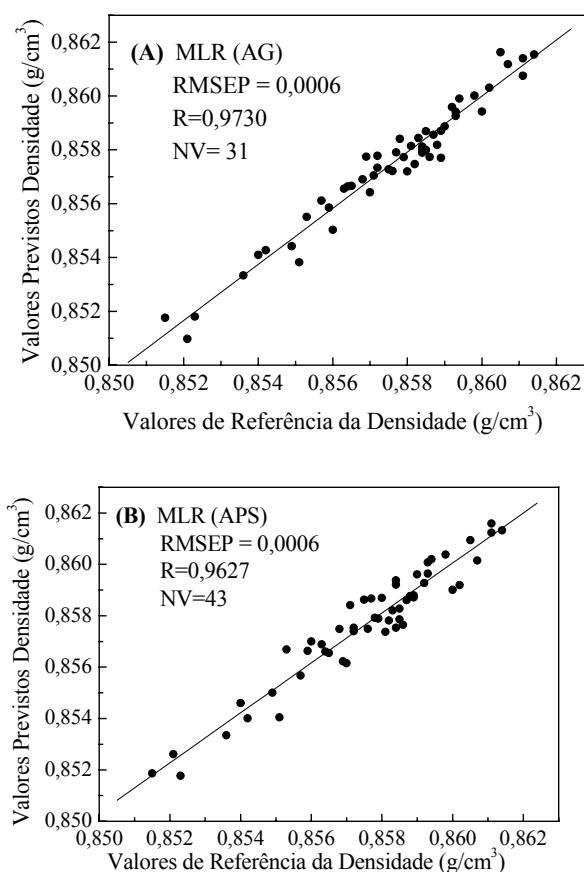


Figura 7.1.12: Comparação entre valores de referência e valores previstos por modelos MLR para o parâmetro de densidade (diesel). (A) variáveis selecionadas por AG e (B) variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

O método AG selecionou um número menor de variáveis que o método APS, embora os resultados obtidos sejam praticamente idênticos.

Os valores de RMSEP obtidos através dos métodos empregados neste trabalho (0,0006 e 0,0007 g/cm³) são menores que os erros de previsão observados na literatura: Fodor e colaboradores^[84], utilizando um acessório de ATR com geometria horizontal, obtiveram erros de previsão de 0,0009 g/cm³ através do

método PLS com 19 fatores e Yang^[123] e co-autores obtiveram erros de 0,002 g/cm³ para a densidade determinada por métodos CG/MS (cromatografia gasosa e espectrometria de massa), com aplicação de métodos de regressão com redes neurais.

Os resultados para a previsão da densidade do diesel estão resumidos na tabela 7.1.5:

Tabela 7.1.5: Resultados para a Previsão da densidade do diesel através de modelos de Calibração Multivariada

DENSIDADE (DIESEL)				
Faixa de valores: 0,8489-0,8626 g/cm ³				
Reprodutibilidade do Método Padrão: 0,0012 g/cm ³				
Modelo	VL	RMSEP	R	NV
PLS (linha base)	9	0,0007	0,9511	912
PLS (1ª Derivada)	6	0,0007	0,9604	902
PLS (AG)	5	0,0006	0,9632	31
MLR (AG)	-	0,0006	0,9730	31
MLR (APS)	-	0,0006	0,9627	43

VL: variáveis latentes; NV: número de variáveis independentes; R: coeficiente de correlação.

Todos os valores de RMSEP obtidos encontram-se abaixo da reprodutibilidade do método padrão (0,0012 g/cm³) e este fato foi utilizado como critério de validação, embora, a rigor, o RMSEP expressa erro e a reprodutibilidade expressa desvio (precisão).

Os métodos de regressão efetuados a partir das variáveis selecionadas (AG e APS) proporcionaram resultados muito próximos dos obtidos sem seleção de variáveis. Portanto um menor número de variáveis independentes pode ser

utilizado para obter o mesmo desempenho dos outros modelos quanto à capacidade preditiva.

- Temperaturas de destilação (T50% e T85%)

Os modelos de calibração PLS e MLR elaborados para as temperaturas de destilação (°C) a 50% e a 85% de destilado foram efetuados através do mesmo procedimento empregado na calibração da densidade. No entanto, foram utilizados diferentes conjuntos de calibração para cada temperatura: na construção dos modelos de T50%, utilizou-se 50 amostras de calibração, 30 amostras de teste e 34 amostras de previsão. Para a determinação de T85% utilizou-se 45 amostras de calibração, 26 amostras de teste e 31 amostras de previsão. O número de amostras difere porque as propriedades têm diferentes *outliers*, ou seja, uma amostra mal modelada em uma propriedade pode não apresentar o mesmo comportamento para a modelagem de outra propriedade e assim pode ser incluída no conjunto de calibração.

As faixas de temperaturas de referência foram de 265,7 a 295, 9 °C para T50% e 331,9 a 360,0 °C para T 85%.

A figura 7.1.13 mostra o número de variáveis latentes selecionados nos modelos PLS para a previsão de T50% e T 85%:

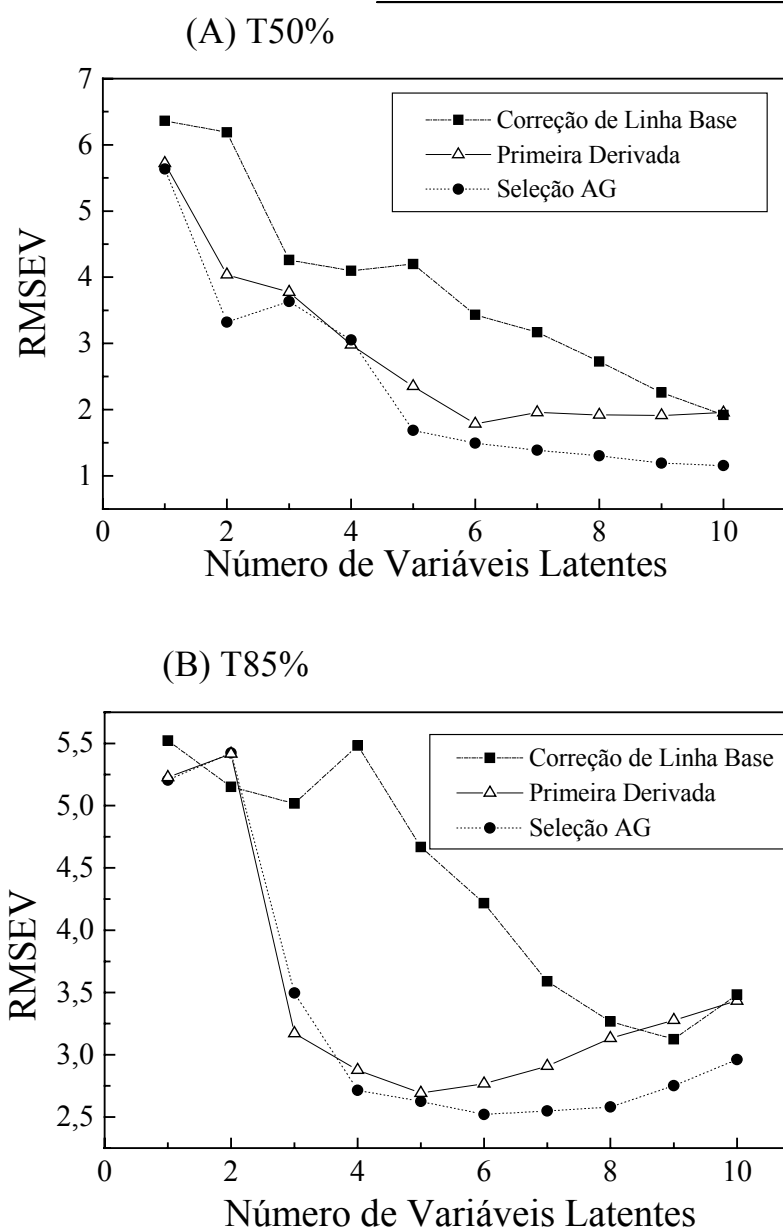


Figura 7.1.13: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para (A) T50% e (B) T85% do diesel com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

Em relação à temperatura de T50%, utilizou-se modelos PLS com 10 variáveis latentes para o modelo com correção de linha base, 6 variáveis latentes para modelos com 1ª Derivada e a partir de variáveis independentes selecionadas pelo método AG. Fodor e colaboradores^[84] utilizaram modelos PLS com alto número de fatores ou variáveis latentes (16 e 17 fatores) para a previsão de T50% e

atribuíram este fato à grande complexidade da composição de um combustível derivado do petróleo como o diesel. Deste modo o uso de 10 fatores ou menos pode ser insuficiente para modelar determinadas propriedades como a densidade, temperaturas de destilação, entre outras.

Neste trabalho, o número máximo de fatores permitidos foi estabelecido em 10 (dez) para a modelagem de qualquer propriedade dos combustíveis analisados. Em relação à temperatura de T85%, os modelos PLS foram descritos com 9 variáveis latentes para o modelo com correção de linha base, 5 para o modelo com primeira derivada e 6 para o modelo efetuado a partir da seleção AG. Para ambas as temperaturas, houve redução do número de variáveis latentes com o emprego da primeira derivada e com o uso dos números de onda selecionados a partir do método AG.

Os resultados da aplicação dos modelos PLS na previsão de T50% e T85% para as amostras do conjunto de previsão, são mostrados nas figuras 7.1.14 e 7.1.15:

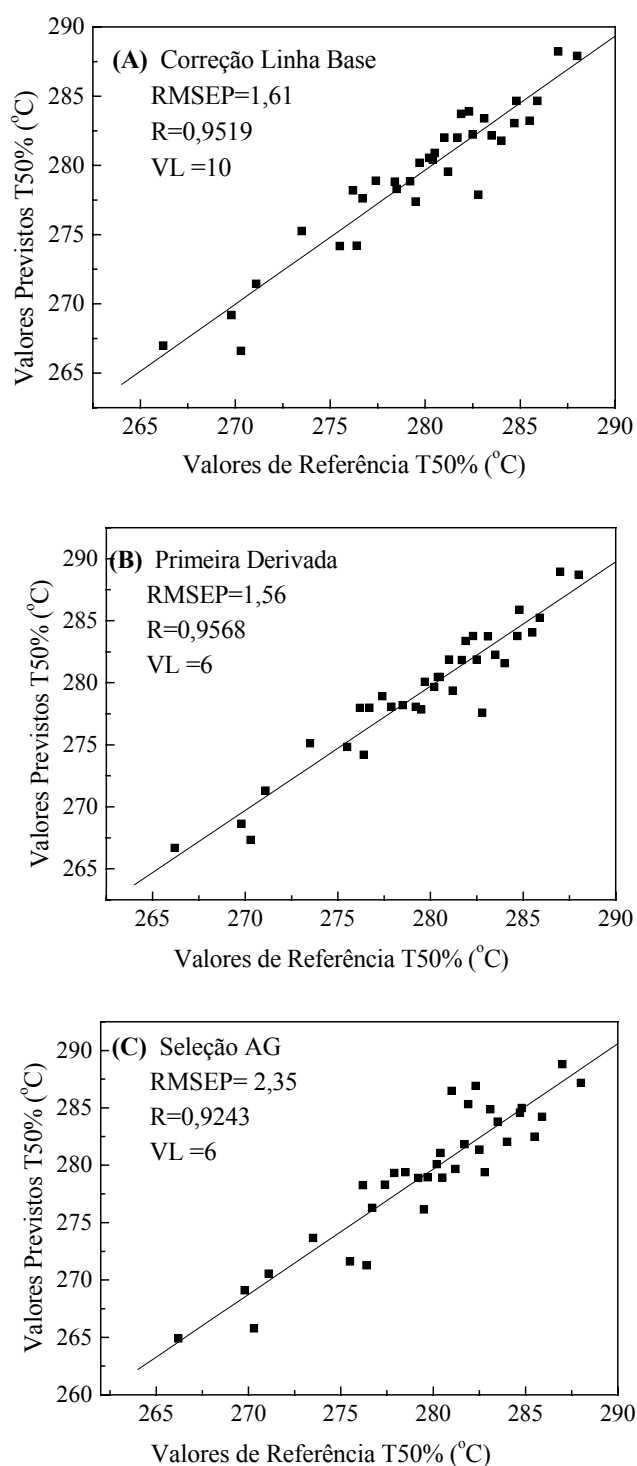


Figura 7.1.14: Comparação entre valores de referência e valores previstos por modelos PLS para o parâmetro T50%. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

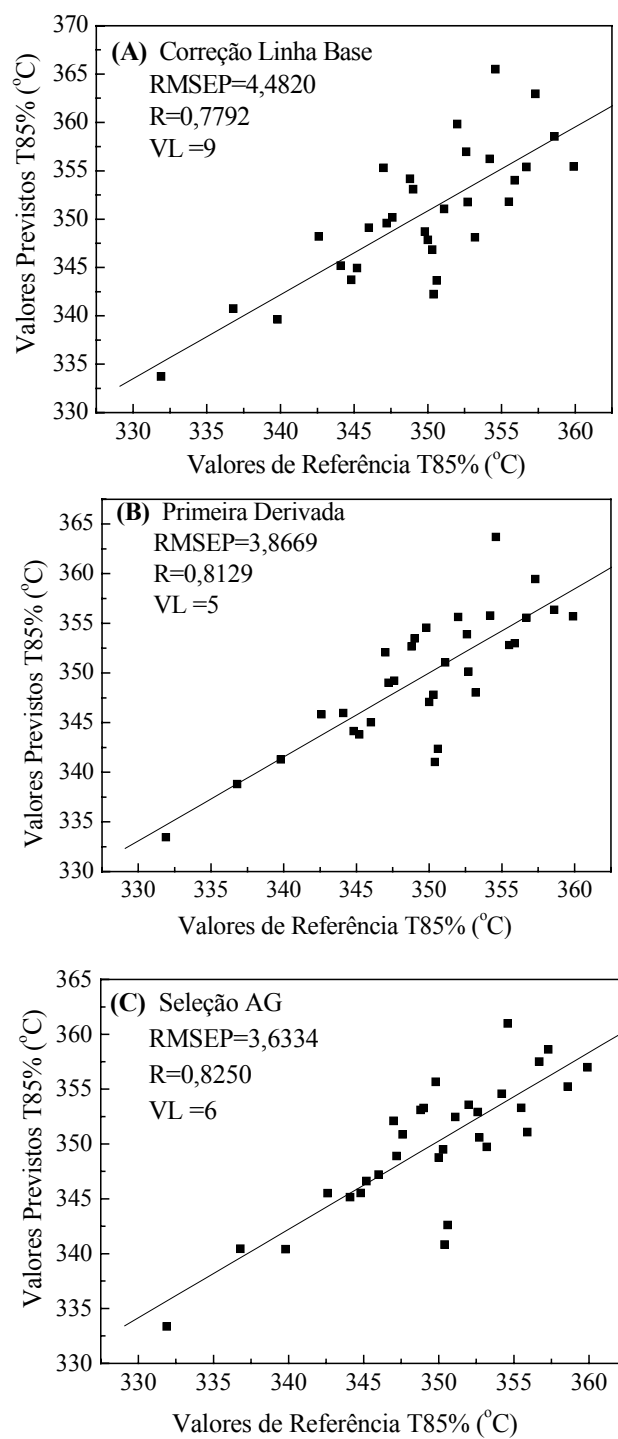


Figura 7.1.15: Comparação entre valores de referência e valores previstos por modelos PLS para o parâmetro T85%. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

O modelo PLS com seleção de variáveis AG produziu o maior valor de RMSEP para a previsão de T50%, embora tenha utilizado um número menor de variáveis latentes. Em relação aos resultados obtidos para a previsão de T85%, o maior valor de RMSEP está relacionado com o modelo com correção de linha base.

O método de referência ASTM apresenta altos valores de repetibilidade para a determinação das temperaturas de destilação (acima de 3°C) e esta falta de precisão pode ser refletida na estimativa FTIR nos valores de RMSEP e nos coeficientes de correlação (R), que para T85% mostraram-se abaixo de 0,85^[84]. Além disso, o valor de RMSEC (Raiz Quadrada dos Erros Médios de Calibração) para o modelo que apresentou o valor mais alto de RMSEP (4,4820°C, figura 7.1.15(A)), corresponde a 2,1117 °C, com coeficiente de correlação de 0,9469. Isto significa que pode ter havido um superajustamento do modelo de calibração, visto que em uma situação ideal, os valores de RMSEC e RMSEP devem ser bem próximos. Os valores de RMSEC para os modelos com 1ª derivada e com seleção AG, correspondem a 2,7277 e 2,7863 °C , respectivamente.

Modelos MLR foram construídos para a previsão das temperaturas de destilação a partir das variáveis selecionadas pelos métodos AG, que selecionou para T50%, 18 variáveis e para T85% , 27 variáveis já empregadas anteriormente nos modelos PLS.

O método APS selecionou 20 variáveis ou números de onda para T50% e 17 variáveis para T85%. O número de variáveis em função dos valores de RMSEV obtidos a través da seleção APS para T50% são mostrados na figura 7.1.16:

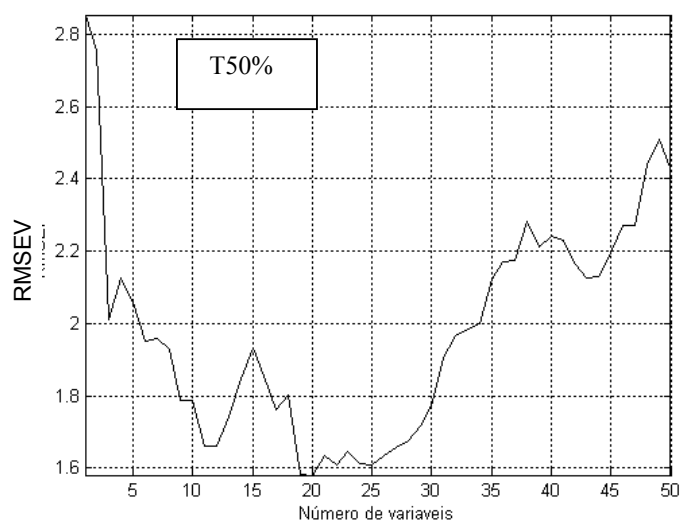


Figura 7.1.16: Resultado da seleção APS para a temperatura a 50% de destilado do diesel (T50%). O menor valor de RMSEV corresponde ao número de 20 variáveis selecionadas

O resultado da seleção através do método APS para T85% é mostrado na figura 7.1.17:

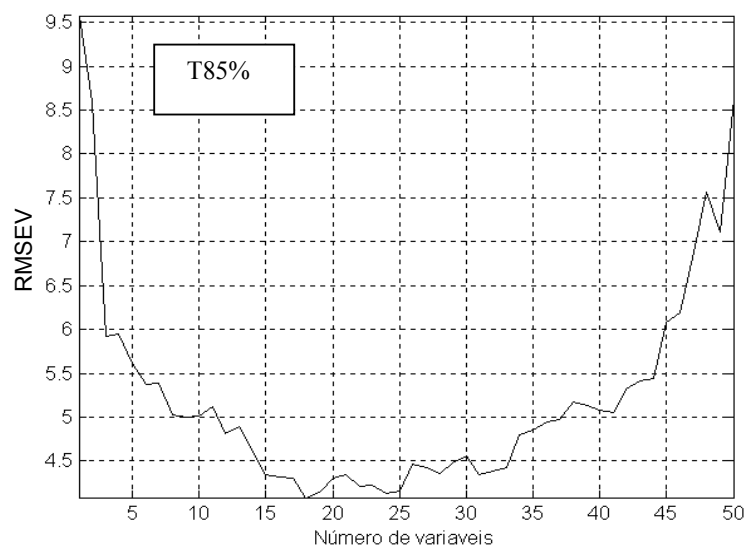


Figura 7.1.17: Resultado da seleção APS para a temperatura a 85% de destilado do diesel (T85%). O menor valor de RMSEV corresponde ao número de 17 variáveis selecionadas.

O resultado dos modelos MLR aplicados nas amostras do conjunto de previsão são mostrados na figura 7.1.18 para T50% e na figura 7.1.19 para a previsão de T85%.

Os valores de RMSEP obtidos pelos modelos MLR (APS e AG) mostraram-se muito próximos dos valores de RMSEP obtidos a través do método PLS para ambas propriedades de destilação.

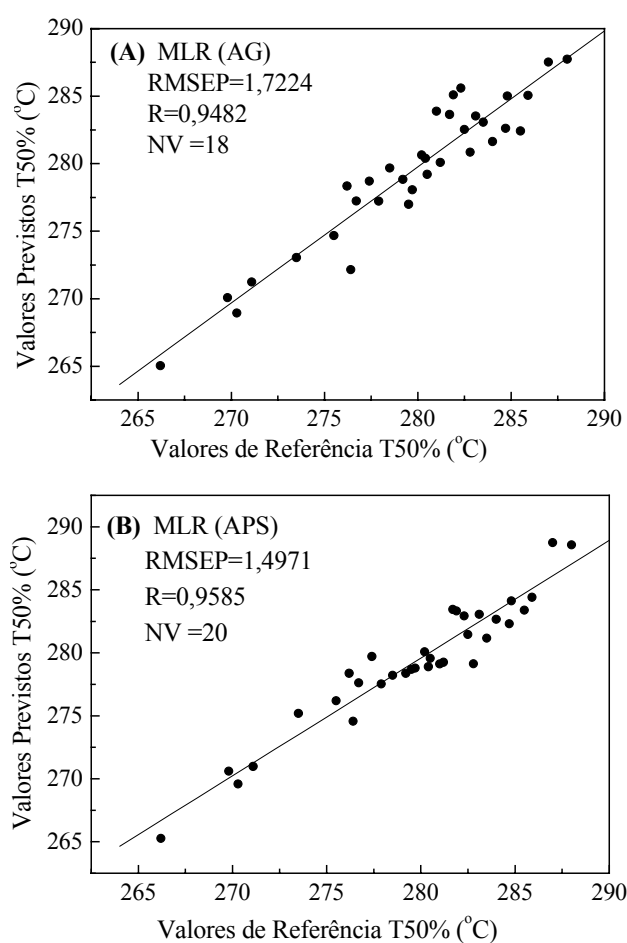


Figura 7.1.18: Comparação entre valores de referência e valores previstos para por modelos MLR para o parâmetro T50% do diesel. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

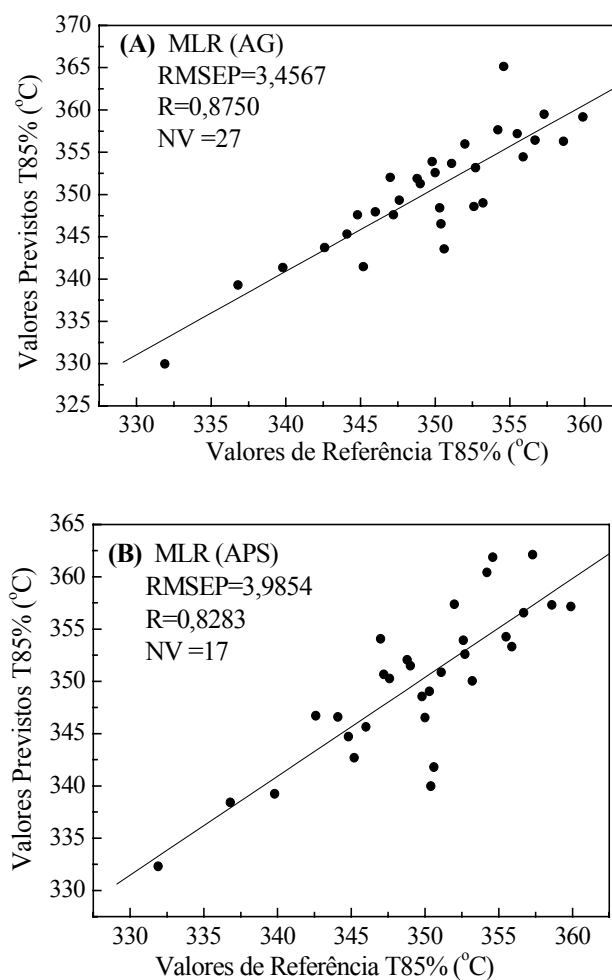


Figura 7.1.19: Comparação entre valores de referência e valores previstos para por modelos MLR para o parâmetro T85% (diesel). (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

Os resultados da previsão das temperaturas de destilação do diesel T50% e T85% estão dispostos na tabela 7.1.6

Tabela 7.1.6: Resultados para a Previsão das temperaturas de destilação T50% e T85% do Diesel através da Calibração Multivariada.

T50%(DIESEL)	Modelo	VL	RMSEP	R	NV
- Faixa : 264,7-295,9 °C - Reprodutibilidade Método Padrão: 9,7 °C	PLS (linha base)	10	1,6148	0,9519	912
	PLS (1ª Derivada)	6	1,5630	0,9568	902
	PLS (AG)	6	2,3556	0,9243	18
	MLR (AG)	-	1,7224	0,9585	18
	MLR (APS)	-	1,4971	0,9482	20
T85%(DIESEL)					
- Faixa : 331,9-359,9 °C - Reprodutibilidade Método Padrão: 10,6 °C	PLS (linha base)	9	4,4820	0,7792	912
	PLS (1ª Derivada)	5	3,8669	0,8129	902
	PLS (AG)	6	3,6334	0,8250	27
	MLR (AG)	-	3,4567	0,8750	27
	MLR (APS)	-	3,9854	0,8283	17

VL: variáveis latentes; NV: número de variáveis independentes; R : coeficiente de correlação

Os valores de RMSEP obtidos em todos os modelos encontram-se abaixo dos valores de reprodutibilidade do método ASTM que correspondem a 9,7°C para T50% e 10,6°C para T 85%.

Os menores valores de RMSEP obtido na previsão das temperaturas de destilação correspondem aos resultados obtidos a partir dos métodos MLR com seleção APS para T50% e com seleção AG para T85%.

Fodor e colaboradores ^[84] obtiveram através do método PLS em dados de FTIR-ATR, um valor de RMSEP de 3°C para T50%.

- Índice de Cetano

Os modelos PLS e MLR elaborados para a previsão do índice de cetano do diesel utilizaram conjuntos de 48 amostras de calibração, 23 amostras de teste e 34 amostras de previsão. Foram excluídas as amostras aditivadas.

A faixa de valores desta propriedade corresponde a 39,9 a 49,3.

O número de variáveis utilizadas nos modelos PLS, segundo o critério de menor RMSEV pode ser visualizado na figura 7.1.20:

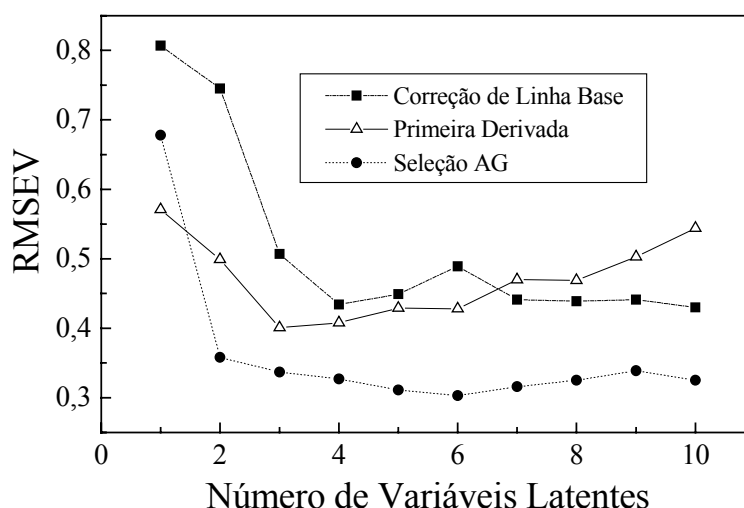


Figura 7.1.20: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para o índice de cetano do diesel com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

O modelo PLS com correção de linha base foi construído com 4 variáveis latentes; o modelo com aplicação da primeira derivada, com 3 variáveis latentes e o modelo com variáveis selecionadas pelo algoritmo genético, com 5 variáveis latentes.

O resultado da previsão do índice de cetano através dos modelos PLS são mostrados na figura 7.1.21:

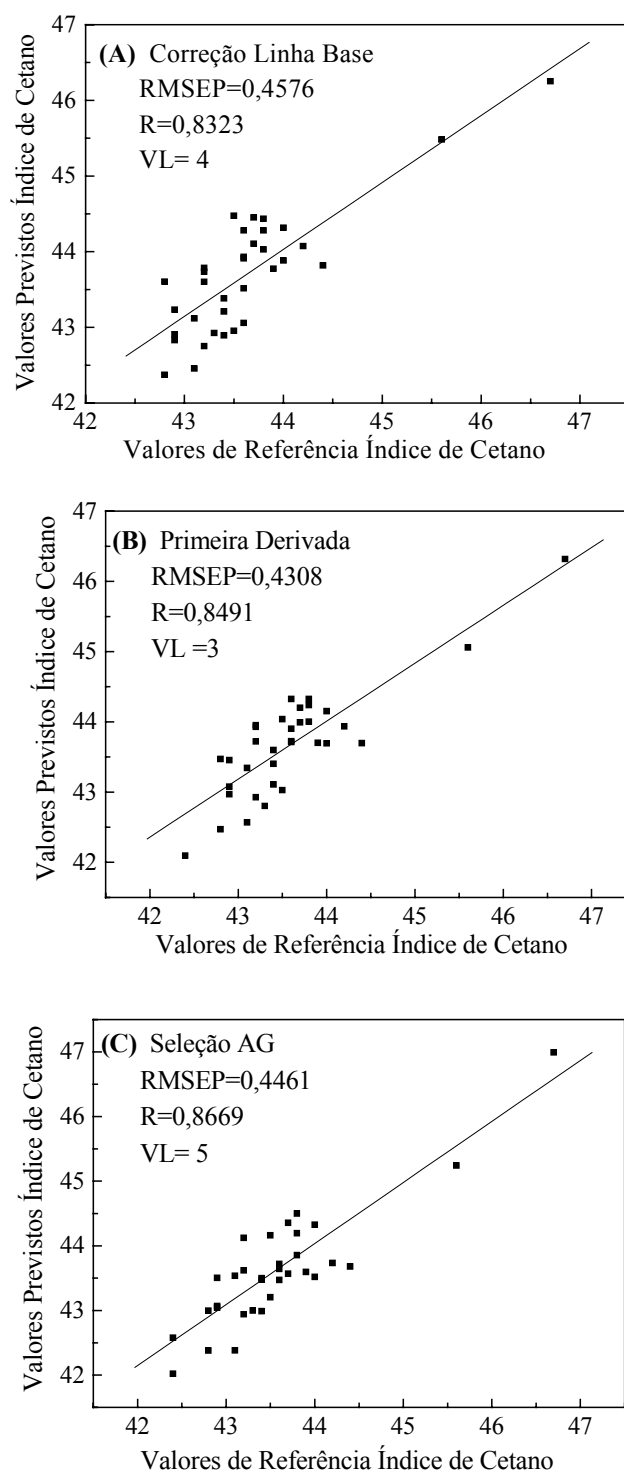


Figura 7.1.21: Comparação entre valores de referência e valores previstos para por modelos MLR para o parâmetro índice de cetano (diesel). (A) Variáveis selecionadas por AG e Variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

Os valores de RMSEP mostraram-se muito próximos, segundo os resultados obtidos pelo método PLS. O menor valor de RMSEP foi obtido pelo modelo de previsão com aplicação da primeira derivada.

Fodor e Kohl^[83] obtiveram um erro padrão de previsão de 0,943 para o índice de cetano, determinado a partir de modelos PLS de dados FTIR-ATR, com cela de ATR horizontal. No entanto, os autores analisaram uma faixa mais ampla de valores desta propriedade (37-58) e obtiveram altos coeficientes de correlação (acima de 0,95).

Através da seleção de variáveis através dos métodos AG e APS foram selecionadas 26 e 9 números de onda, respectivamente. Os 26 números de onda do método AG foram aplicados no método PLS (figura 7.1.21) e o número de variáveis selecionadas pelo método APS em função de RMSEV é mostrado na figura 7.1.22:

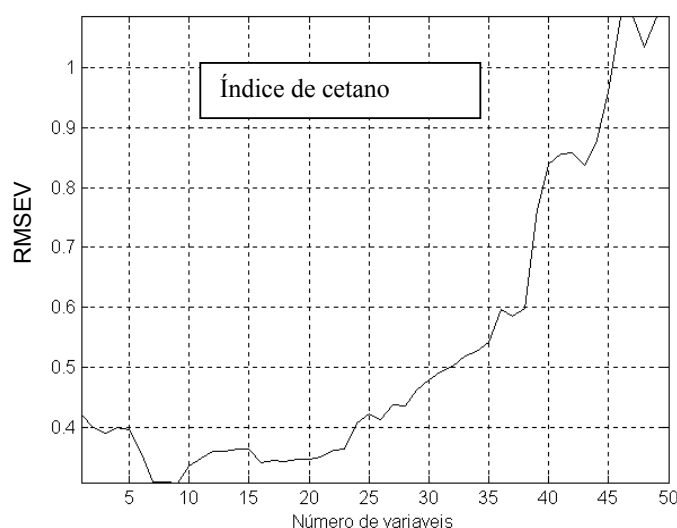


Figura 7.1.22: Resultado da seleção APS para o índice de cetano do diesel
O menor valor de RMSEV corresponde ao número de 9 variáveis selecionadas.

O resultado da previsão dos modelos MLR, construídos a partir dos números de onda selecionados (APS e AG) são mostrados na figura 7.1.23:

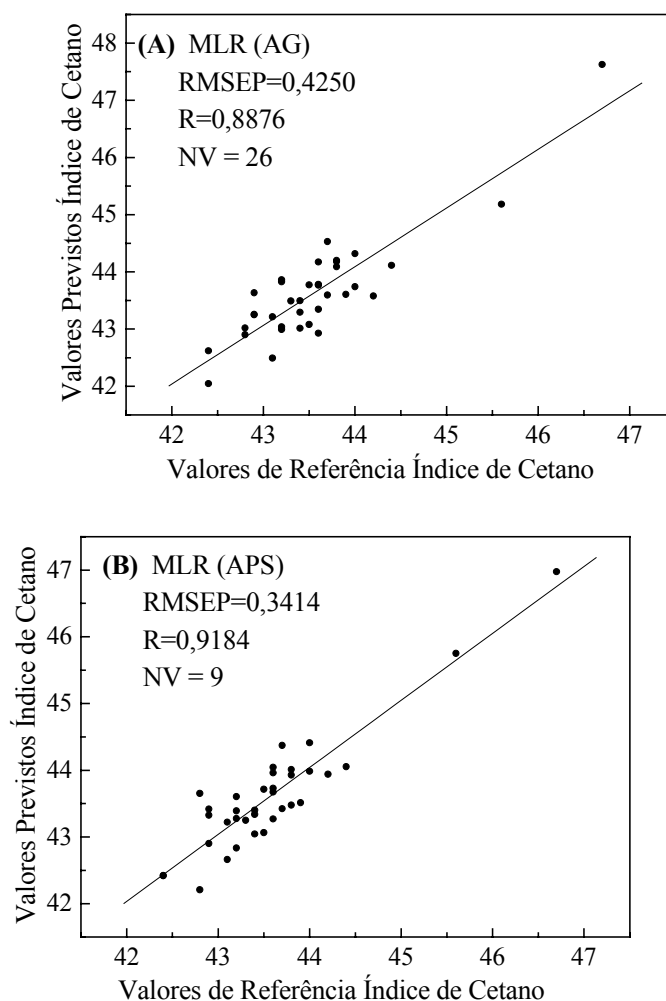


Figura 7.1.23: Comparação entre valores de referência e valores previstos para por modelos MLR para o parâmetro índice de cetano. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

O método APS selecionou um número bem menor de variáveis que o método AG, com maior coeficiente de correlação. Adicionalmente o valor de RMSEP foi inferior ao valor fornecido pelo modelo MLR (AG).

Todos os resultados fornecidos pelos modelos de previsão do índice de cetano do diesel são mostrados na tabela 7.1.7:

Tabela 7.1.7: Resultados para a Previsão do índice de Cetano do Diesel através da Calibração Multivariada.

ÍNDICE DE CETANO (DIESEL)				
Faixa de Valores: 39,9- 49,3				
Reprodutibilidade do Método Padrão: 2,0				
Modelo	VL	RMSEP	R	NV
PLS (linha base)	4	0,4576	0,8376	912
PLS (1ª Derivada)	3	0,4308	0,8491	902
PLS (AG)	5	0,4461	0,8669	26
MLR (AG)	-	0,4250	0,8876	26
MLR (APS)	-	0,3414	0,9184	9

VL: variáveis latentes; NV: número de variáveis independentes;
R: coeficiente de correlação

Os resultados mostraram que a aplicação dos métodos de seleção de variáveis podem melhorar a capacidade preditiva dos modelos propostos para a previsão do índice de cetano, uma vez que os menores valores de RMSEP foram obtidos quando os números de onda selecionados foram empregados. Adicionalmente observou-se aumento do coeficiente de correlação.

Todos os valores de RMSEP fornecidos pelos modelos propostos neste trabalho encontram-se abaixo da reprodutibilidade do método padrão (2,0).

- Teor de Enxofre (%)

A faixa da concentração de enxofre nas amostras de diesel analisadas neste trabalho foi restrita e corresponde a 0,10 a 0,34 (% em massa). Adicionalmente muitas amostras empregadas nos conjuntos de calibração, validação interna e

externa apresentam teor de enxofre semelhante. O conjunto de calibração consiste em 36 amostras de calibração, 24 amostras de teste e 35 amostras de previsão.

O número de variáveis latentes empregadas é mostrado na figura 7.1.24, através do gráfico de RMSEV em função do número de variáveis latentes:

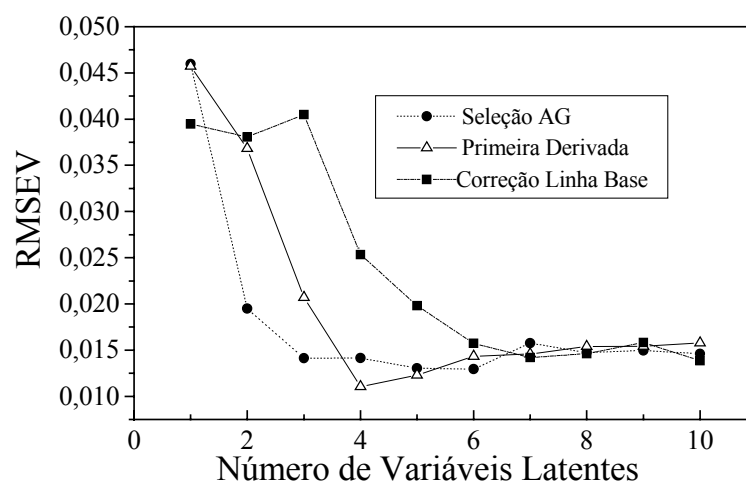


Figura 7.1.24: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para o teor de enxofre do diesel com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

O modelo PLS com correção de linha base foi descrito com sete variáveis latentes. Observou-se através da figura 7.1.24 que a inclusão de mais variáveis latentes não reduz intensamente os valores de RMSEV. O modelo PLS, com primeira derivada pôde ser descrito com um número menor de variáveis latentes (4) e o modelo com variáveis selecionadas pelo algoritmo genético foi descrito com três variáveis latentes.

O resultado da previsão através dos modelos PLS para o teor de enxofre é mostrado na figura 7.1.25:

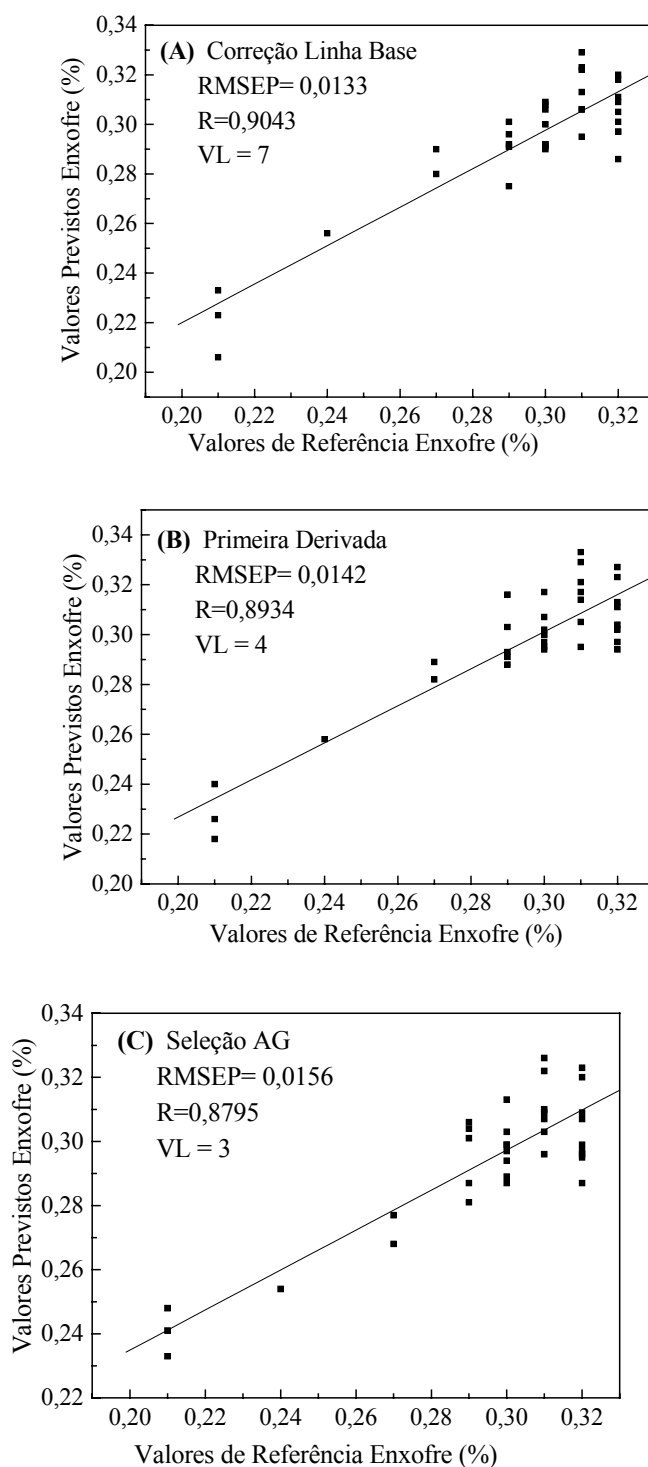


Figura 7.1.25: Comparação entre valores de referência e valores previstos por modelos PLS para o teor de enxofre (diesel): (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

O menor valor de RMSEP (0,013%) e o melhor coeficiente de correlação (0,9043) foram obtidos com o modelo com correção de linha base, embora este modelo seja descrito com um número superior de variáveis latentes, quando comparado aos demais.

O modelo PLS (C) foi construído a partir de 23 números de onda selecionados através do algoritmo genético. O APS selecionou nove números de onda, cujo gráfico de RMSEP em função das variáveis selecionadas é mostrado na figura 7.1.26:

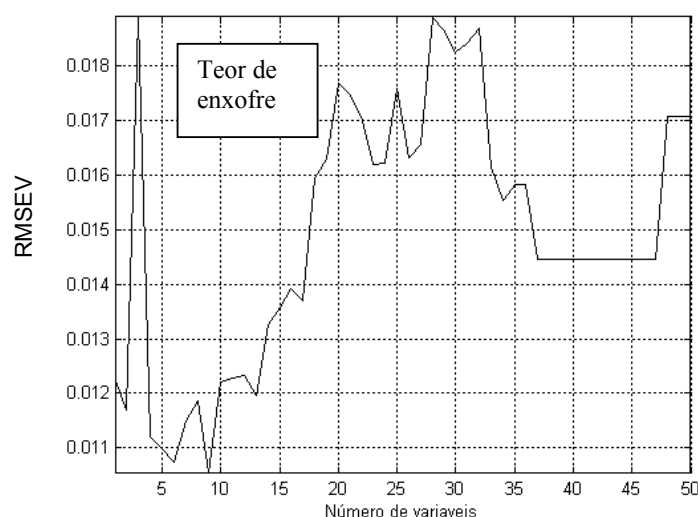


Figura 7.1.26: Resultado da seleção APS para o teor de enxofre (%) do diesel. O menor valor de RMSEP corresponde ao número de 9 variáveis selecionadas.

Os números de onda selecionados pelos métodos AG e APS foram aplicados nos modelos MLR, cuja previsão para o conjunto de validação externa é mostrada na figura 7.1.27:

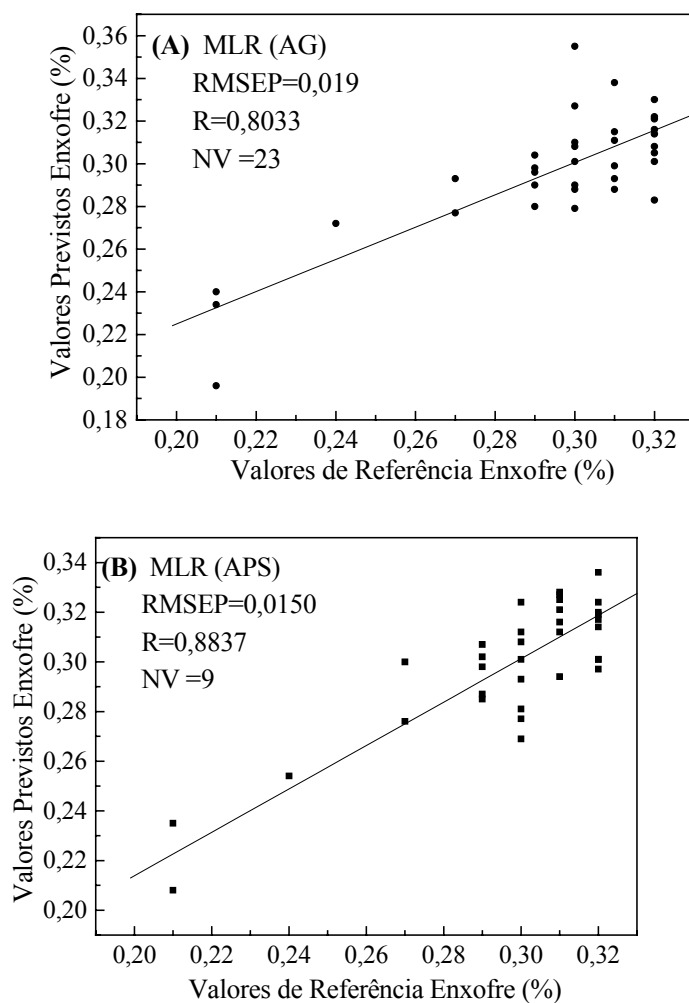


Figura 7.1.27: Comparação entre valores de referência e valores previstos para por modelos MLR para o teor de enxofre. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

O modelo MLR (APS) apresentou melhor capacidade preditiva que o modelo MLR (AG), comparando-se os valores de RMSEP e do coeficiente de correlação. Além disso, o modelo MLR(APS) utilizou um número menor de variáveis independentes.

Em recente trabalho, Breitzkreitz e co-autores^[85] utilizaram os métodos AG e APS para determinar o teor de enxofre em amostras de diesel comercial, através de dados obtidos com espectroscopia no infravermelho próximo. Através dos métodos

PLS (AG) e MLR(AG) com 19 comprimentos de onda e MLR(APS) com nove comprimentos de onda selecionados, os autores obtiveram RMSEP abaixo de 0,03% e coeficientes de correlação superiores a 0,912.

Apesar dos resultados obtidos na região do infravermelho médio apresentarem menores valores de RMSEP para a determinação do teor de enxofre, a espectroscopia no infravermelho próximo possui instrumentação mais simples, em geral e a adaptação para análises on-line é mais facilitada.

A tabela 7.1.8 reúne os principais resultados obtidos através dos modelos de calibração propostos para a determinação do teor de enxofre:

Tabela 7.1.8: Resultados para a Previsão do teor de enxofre do Diesel através da Calibração Multivariada.

TEOR DE ENXOFRE (DIESEL)				
Faixa de Concentração: 0,10-0,34 (%)				
Reprodutibilidade do Método Padrão: 0,05(%)				
Modelo	VL	RMSEP	R	NV
PLS (linha base)	7	0,0133	0,9043	912
PLS (1ª Derivada)	4	0,0142	0,8934	902
PLS (AG)	3	0,0156	0,8795	23
MLR (AG)	-	0,0190	0,8033	23
MLR (APS)	-	0,0150	0,8837	9

VL: variáveis latentes; NV: número de variáveis independentes;
R: coeficiente de correlação.

O menor valor de RMSEP obtido está relacionado com o modelo PLS com correção de linha base. No entanto todos resultados encontram-se abaixo da reprodutibilidade do método padrão ASTM 4294, cujo valor é de 0,05%.

- Viscosidade Cinemática a 40°C

A faixa de valores da viscosidade do diesel analisadas neste trabalho corresponde a 2,797 a 4,123 mm²/s. Foram empregadas 37 amostras no conjunto de calibração, 16 amostras de teste e 27 amostras de validação.

O número de variáveis latentes selecionadas para os modelos PLS são mostradas na figura 7.1.28:

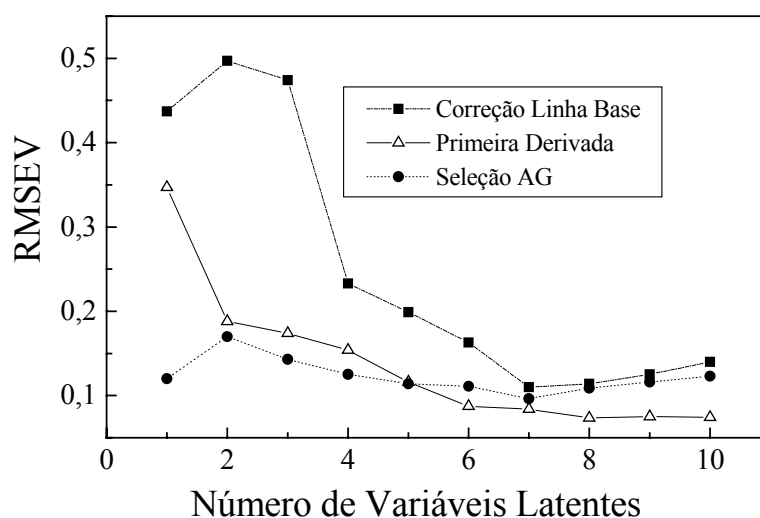


Figura 7.1.28: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para viscosidade do diesel com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

Os modelos PLS com correção de linha base e com seleção AG foram descritos com 7 variáveis latentes e o modelo com primeira derivada foi descrito com 6 variáveis latentes.

O resultado da aplicação dos modelos PLS para o conjunto de previsão é mostrado na figura 7.1.29:

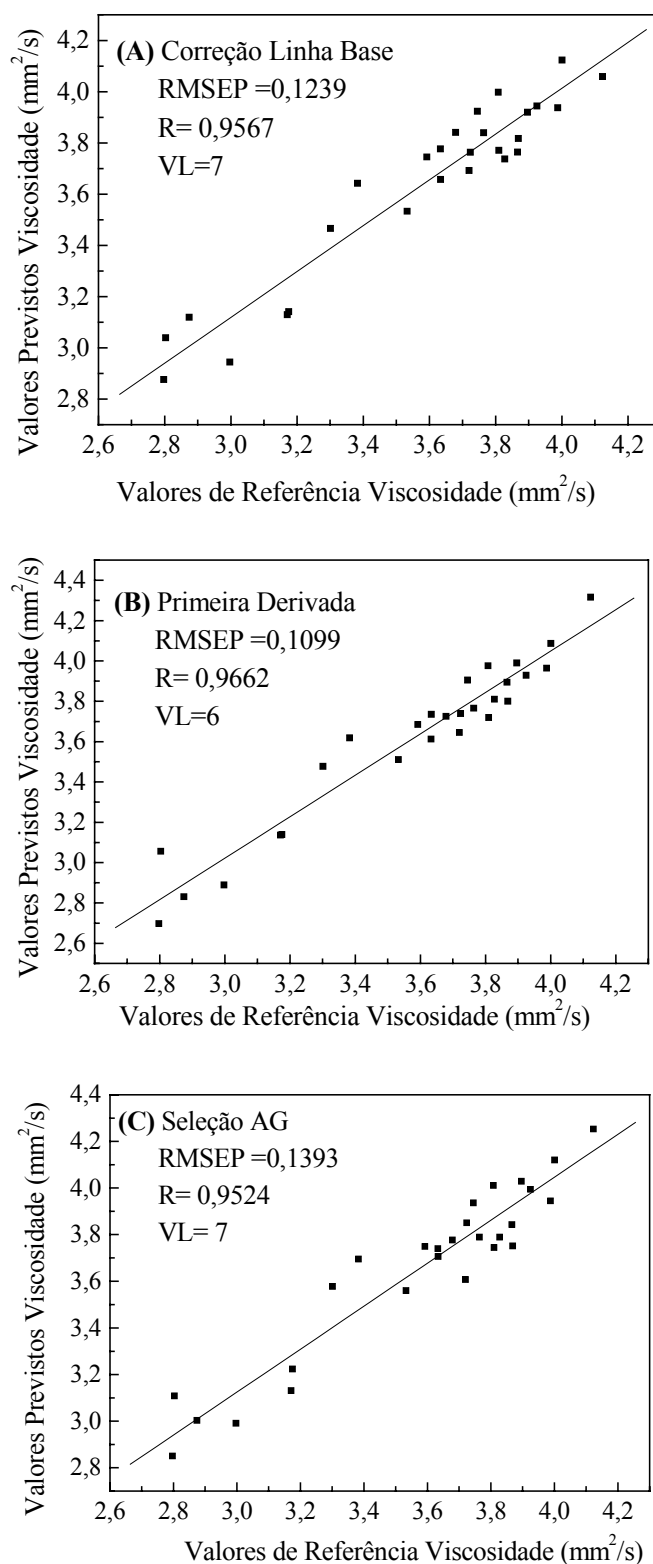


Figura 7.1.29: Comparação entre valores de referência e valores previstos por modelos PLS para viscosidade do diesel. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

O menor valor de RMSEP foi obtido através do modelo PLS com primeira derivada, no qual foi utilizado um menor número de variáveis latentes (6) em relação aos demais.

O modelo PLS (AG) foi construído a partir de 28 números de onda selecionados e o número de variáveis independentes selecionadas pelo método APS é mostrado na figura 7.1.30, que corresponde ao gráfico de RMSEV em função das variáveis escolhidas. Foram selecionadas 20 variáveis independentes através do método APS.

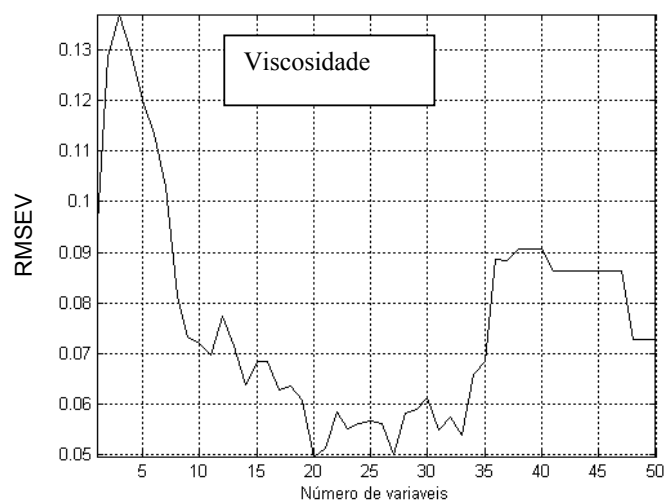


Figura 7.1.30: Resultado da seleção APS para a viscosidade do diesel. O menor valor de RMSEV corresponde ao número de 20 variáveis selecionadas.

Os números de onda selecionados foram aplicados em modelos MLR, cujo resultado é mostrado na figura 7.1.31:

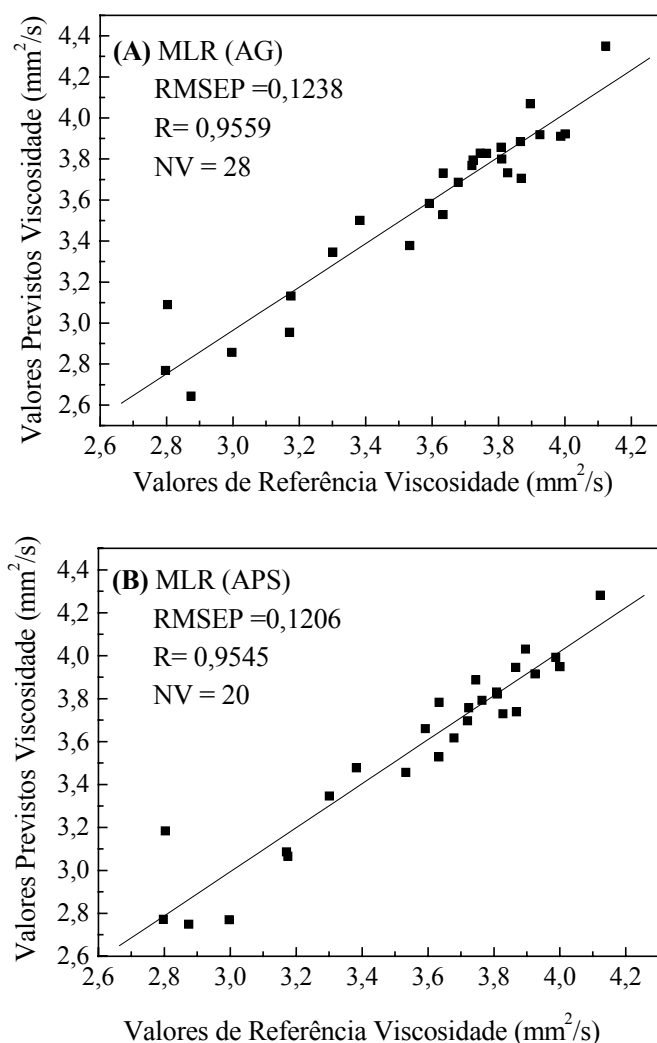


Figura 7.1.31: Comparação entre valores de referência e valores previstos para por modelos MLR para o teor de enxofre. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

Os resultados da previsão da viscosidade através dos modelos MLR mostraram-se muito próximos, indicando capacidade preditiva semelhantes.

A tabela 7.1.9 mostra os resultados obtidos para a previsão da viscosidade do diesel:

Tabela 7.1.9: Resultados para a Previsão da Viscosidade do Diesel através da Calibração Multivariada.

VISCOSIDADE (DIESEL)				
Faixa de Valores: 2,797 a 4,123 mm ² /s				
Reprodutibilidade do Método Padrão: 0,12 (mm ² /s)				
Modelo	VL	RMSEP	R	NV
PLS (linha base)	7	0,1239	0,9567	912
PLS (1ª Derivada)	6	0,1099	0,9662	902
PLS (AG)	7	0,1393	0,9524	28
MLR (AG)	-	0,1238	0,9559	28
MLR (APS)	-	0,1206	0,9545	20

VL: variáveis latentes; NV: número de variáveis independentes;
R: coeficiente de correlação.

A maioria dos valores de RMSEP foram próximos à reprodutibilidade do método padrão, portanto os modelos podem ser aplicados para estimar a viscosidade de amostras de diesel. Os menores valores de RMSEP foram obtidos através dos modelos PLS (1ª Derivada) e MLR(APS), sendo que o modelo PLS (1ª Derivada) forneceu RMSEP menor que a reprodutibilidade do método padrão.

Fodor e colaboradores ^[83-84] obtiveram valores de RMSEP entre 0,09 e 1,33 mm²/s para as amostras de diesel na região do infravermelho médio. No entanto, foram usadas 19 variáveis latentes nos modelos PLS.

7.1.3.2- Modelos PLS e MLR para as propriedades da Gasolina

- Densidade a 20°C

Os modelos de calibração através dos métodos MLR e PLS desenvolvidos com 100 amostras de calibração, 89 amostras de validação interna e 83 amostras de validação externa.

A faixa de valores de densidade dos valores de referência corresponde a 0,7485 a 0,7850 g/cm³.

A figura 7.1.32 mostra o gráfico de RMSEV em função do número de variáveis latentes, utilizadas nos modelos PLS:

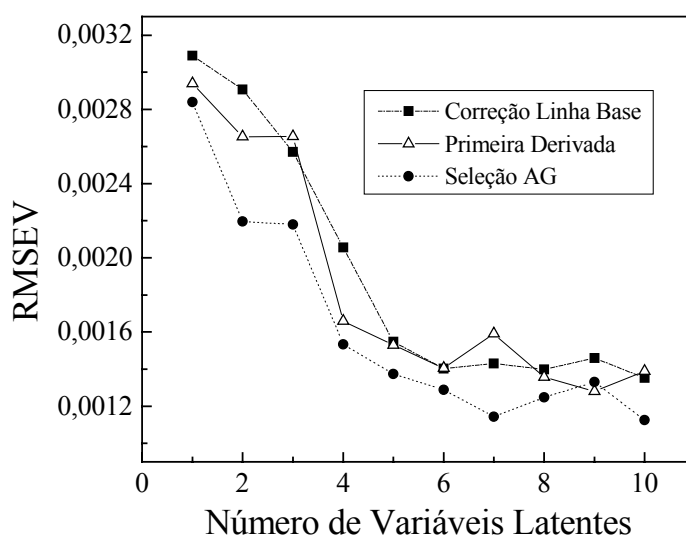


Figura 7.1.32: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para densidade da gasolina com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

Os modelos PLS com correção de linha base e com primeira derivada foram descritos com seis variáveis latentes, enquanto que o modelo construído com números de onda selecionados pelo algoritmo genético utilizou sete variáveis latentes.

O resultado da aplicação dos modelos PLS no conjunto de previsão é mostrado na figura 7.1.33, que representa a relação entre os valores previstos e valores reais:

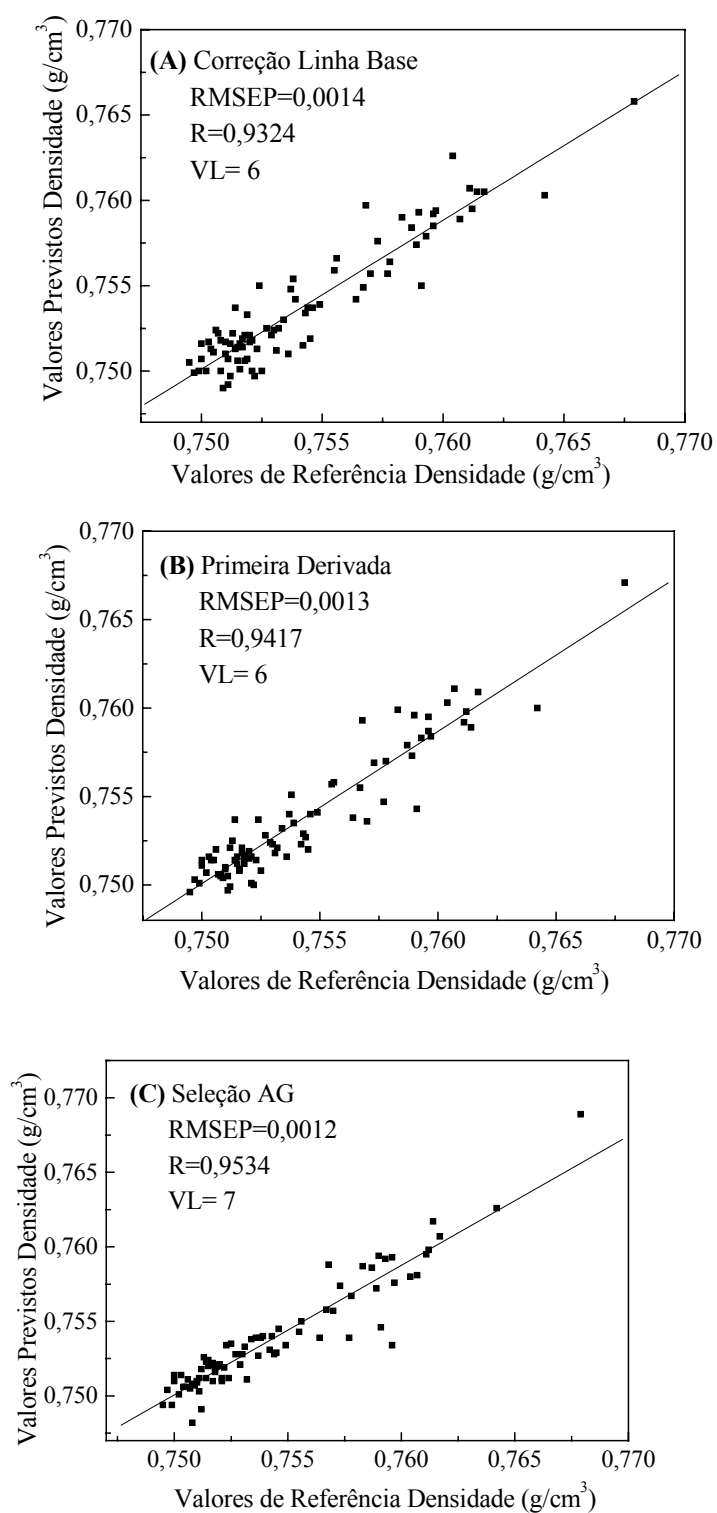


Figura 7.1.33: Comparação entre valores de referência e valores previstos por modelos PLS para a densidade (gasolina): (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

O modelo PLS com correção de linha base foi construído a partir de 1240 números de onda, enquanto que o modelo PLS (1ª derivada) foi construído a partir de 718 números de onda, pois foram excluídas regiões coincidentes com a linha base.

Foi observado que os menores valores de RMSEP foi obtido através do modelo PLS (AG). No entanto todos os valores obtidos foram muito próximos. Devido à complexidade da composição química da gasolina, o número de fatores permaneceu entre seis e sete, uma vez que a aplicação da primeira derivada não reduziu o número de variáveis latentes.

O modelo PLS(AG) foi construído a partir da seleção de 39 números de onda. O APS selecionou 19 números de onda, cujo gráfico de RMSEV em função dos números de variáveis a serem selecionadas são mostradas na figura 7.1.34:

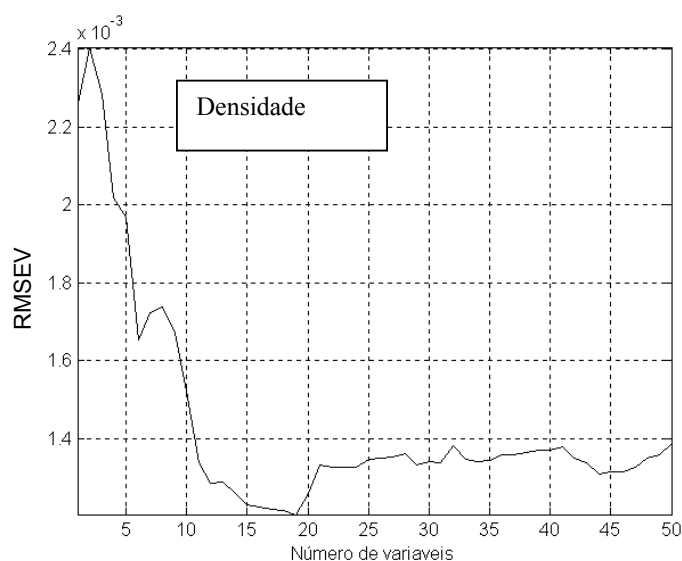


Figura 7.1.34: Resultado da seleção APS para a densidade da gasolina.

O menor valor de RMSEV corresponde ao número de 19 variáveis selecionadas.

O resultado para a determinação da densidade da gasolina para as amostras de validação externa a partir de modelos MLR construídos a partir da seleção AG e APS é mostrado figura 7.1.35:

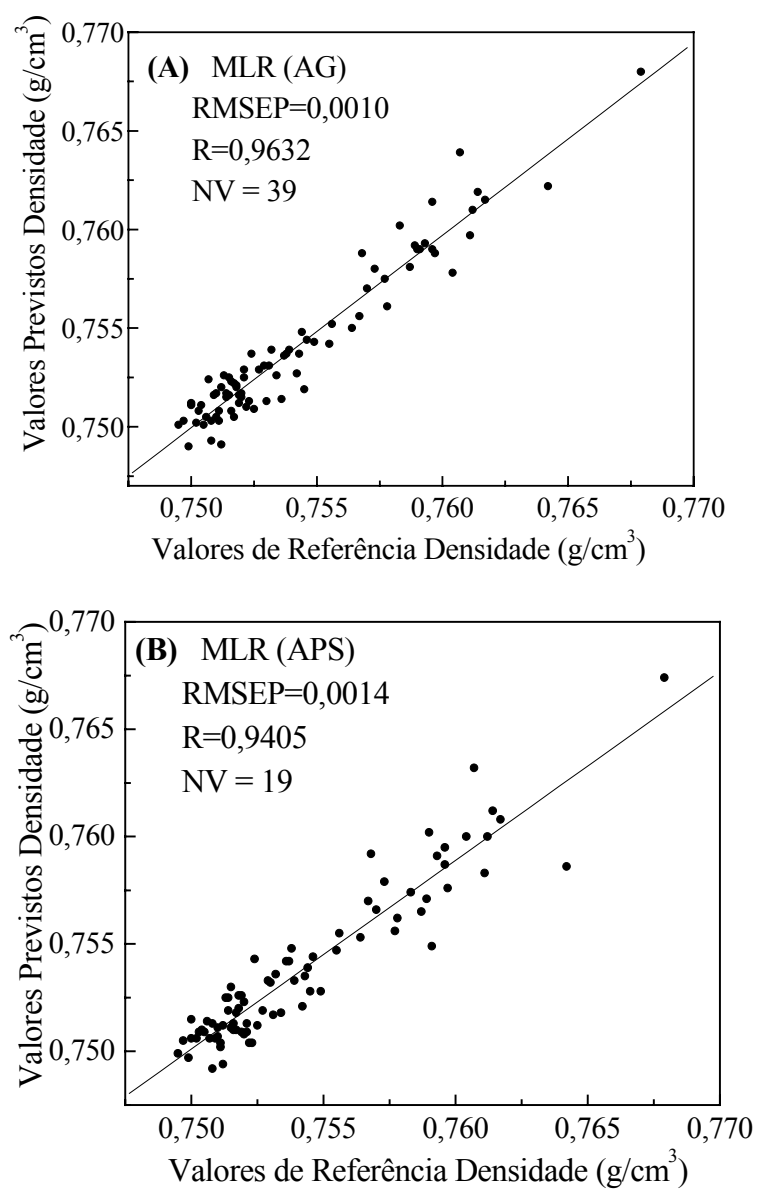


Figura 7.1.35: Comparação entre valores de referência e valores previstos por modelos MLR para o parâmetro de densidade (gasolina). (A) variáveis selecionadas por AG e (B) variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

O modelo MLR (AG) proporcionou o menor valor de RMSEP para a previsão da densidade, com a utilização de 20 variáveis independentes a mais que o modelo MLR (APS).

Tabela 7.1.10: Resultados para a Previsão da densidade da gasolina através de modelos de Calibração Multivariada

DENSIDADE (GASOLINA)				
Faixa de valores: 0,7485 a 0,7850 g/cm ³				
Reprodutibilidade do Método Padrão: 0,0005g/cm ³				
Modelo	VL	RMSEP	R	NV
PLS (linha base)	6	0,0014	0,9324	1240
PLS (1ª Derivada)	6	0,0013	0,9417	718
PLS (AG)	7	0,0012	0,9534	39
MLR (AG)	-	0,0010	0,9623	39
MLR (APS)	-	0,0014	0,9405	19

VL: variáveis latentes; NV: número de variáveis independentes; R: coeficiente de correlação.

De um modo geral, todos os modelos propostos para a determinação da densidade da gasolina apresentaram bons coeficientes de correlação (acima de 0,93). No entanto os valores de RMSEP encontram-se acima do valor de reprodutibilidade do método padrão. Uma hipótese a ser considerada para explicar tal fato é que devido à grande volatilidade das frações leves da gasolina, alterações na composição química podem ter ocorrido durante a análise padrão ou durante obtenção do espectro. Também podem ocorrer perdas quando as amostras estão estocadas. Adicionalmente, todas as análises padrão e espectroscópica devem ser realizadas com rigoroso controle de temperatura. Estes fatores devem influenciar na densidade da gasolina, acarretando erros sistemáticos.

- Temperaturas de destilação (T10% e T90% e PFE)

Foram construídos modelos de previsão das temperaturas a 10% e a 90% de destilado, cujas faixas de valores analisadas correspondem a 48,8 a 69,7°C e 143,9 a 215,5°C, respectivamente. Também foram elaborados modelos para a previsão do ponto final de evaporação (PFE), cuja faixa de temperatura corresponde a 202,2-248,7°C.

Para a modelagem da temperatura de destilação T10% foram empregadas 136 amostras de calibração, 116 amostra de validação interna (teste) e 83 amostras de validação externa (previsão). Para T90%, utilizou-se 125 amostras de calibração, 83 amostras de teste e 69 amostras de previsão. Os modelos para a previsão de PFE utilizaram 85 amostras de calibração, 59 amostras de teste e 45 amostras de previsão.

A figura 7.1.36 mostra o gráfico do número de variáveis latentes em função de RMSEV utilizadas nos modelos PLS:

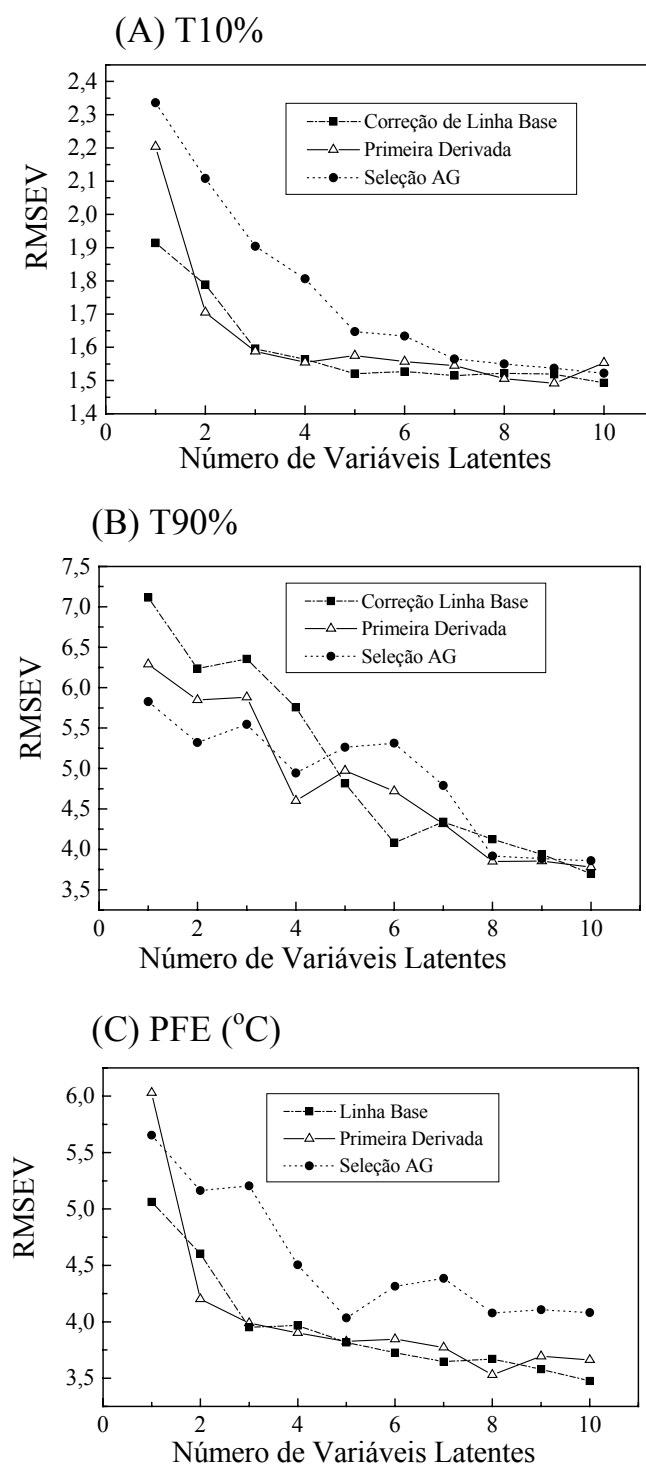


Figura 7.1.36: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para (A) T10%;(B) T90% e (C) PFE da gasolina com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

Os modelos PLS construídos para a seleção de T10% foram descritos com 3 variáveis latentes para os modelos PLS com correção de linha base e em primeira derivada. O modelo PLS com variáveis selecionadas pelo método AG foi construído com 5 variáveis latentes.

Em relação aos modelos PLS construídos para a previsão de T90%, foi utilizado um *maior número* de variáveis latentes, nos quais foram empregadas 6 variáveis latentes para o modelo PLS com correção de linha base, 8 variáveis latentes para o modelo com aplicação da primeira derivada e com seleção de variáveis através do algoritmo genético.

Os modelos PLS para a previsão de PFE foram descritos com 3 variáveis latentes para o modelos com correção de linha base e em primeira derivada e 5 variáveis latentes para o modelo construído com variáveis selecionadas pelo método AG.

Observou-se que a aplicação da primeira derivada não reduziu o número de variáveis latentes para as temperaturas de destilação.

Os resultados da aplicação dos modelos PLS para a previsão de T10%, T90% e PFE são mostrados nas figuras 7.1.37, 7.1.38 e 7.1.39, respectivamente:

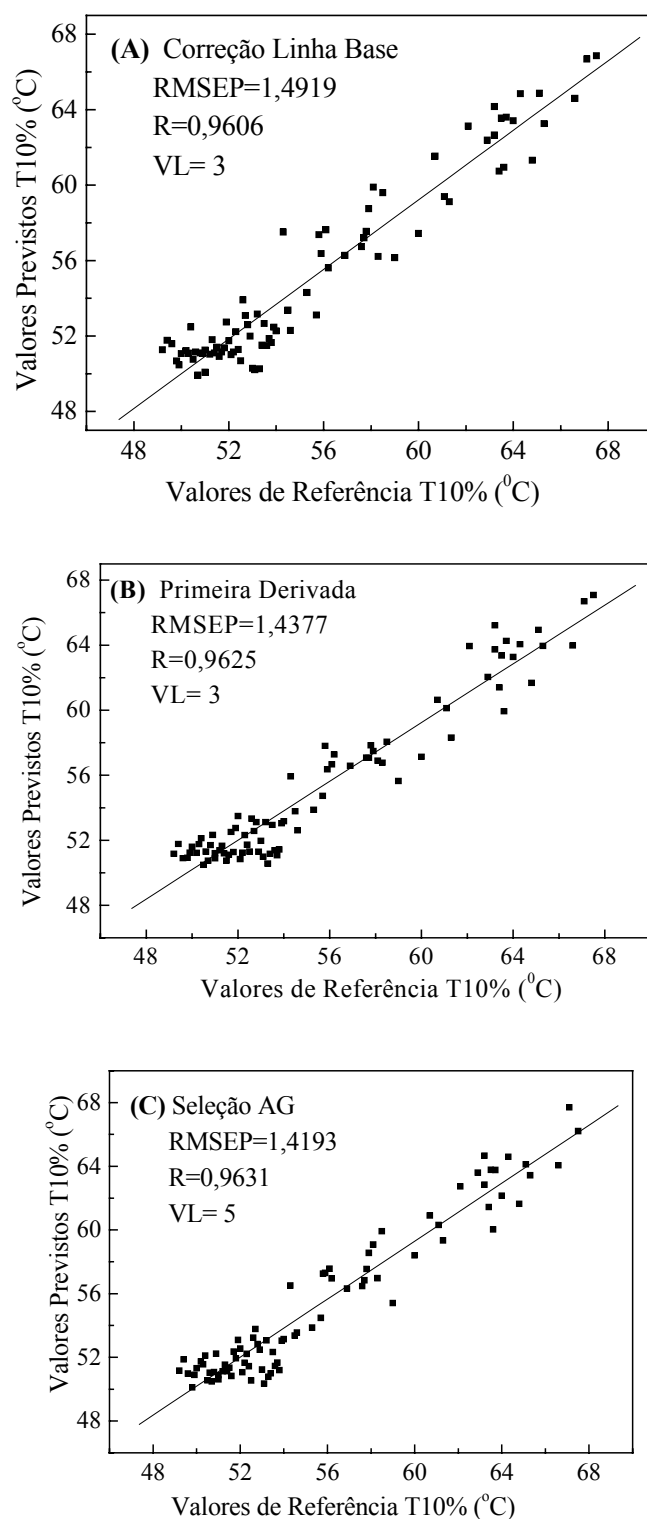


Figura 7.1.37: Comparação entre valores de referência e valores previstos por modelos PLS para temperatura a 10% de destilado da gasolina: (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

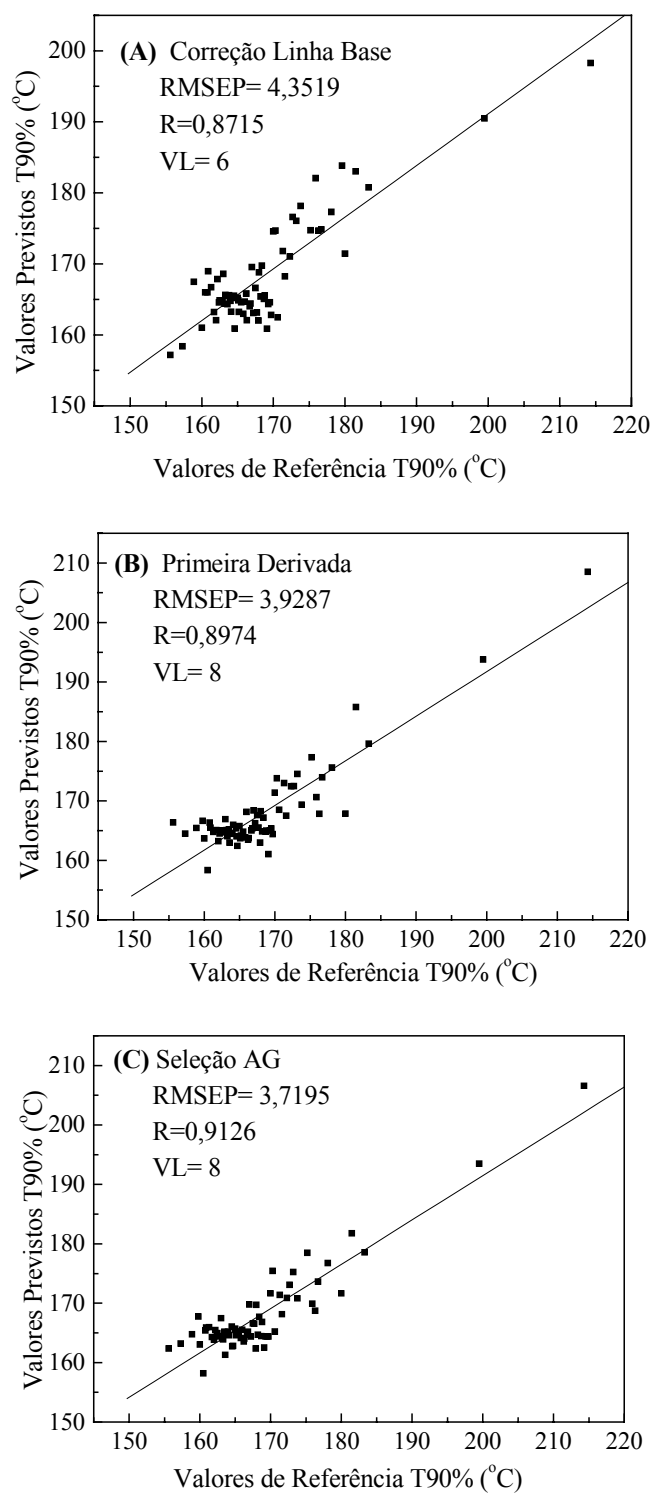


Figura 7.1.38: Comparação entre valores de referência e valores previstos por modelos PLS para temperatura a 90% de destilado da gasolina: (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

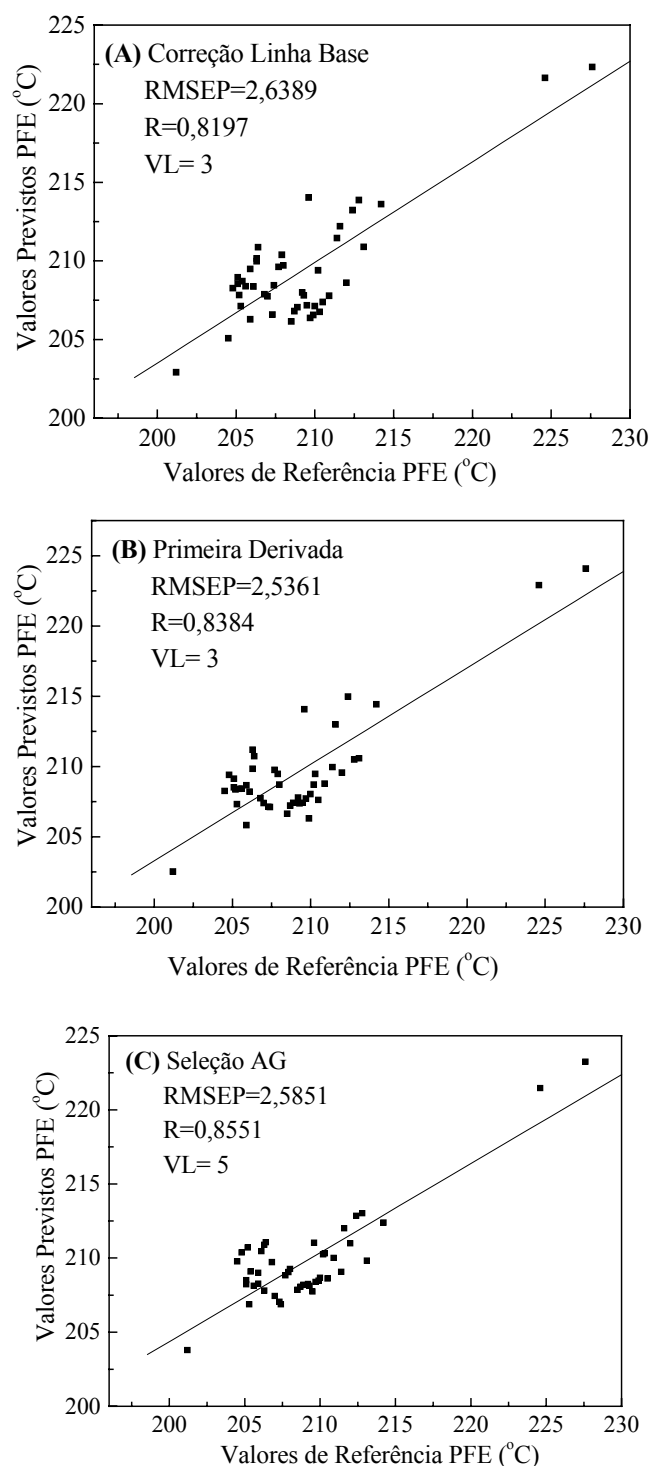


Figura 7.1.39: Comparação entre valores de referência e valores previstos por modelos PLS para o ponto final de evaporação da gasolina: (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

Como pode ser observado nos resultados mostrados na figura 7.1.37 e 7.1.38, o modelo PLS que forneceu o melhor resultado para a previsão de T10% e T90% corresponde ao modelo elaborado com variáveis selecionadas pelo método AG (com 32 e 37 números de onda selecionados para T10% e T90%), segundo o critério de menor RMSEP obtido.

Para a previsão de PFE através dos modelos PLS, observou-se que o modelo que forneceu o menor valor de RMSEP foi o modelo com aplicação da primeira derivada. No entanto, todos os valores de RMSEP obtidos através dos modelos PLS mostraram-se muito próximos para esta propriedade.

Os coeficientes de correlação entre os valores previstos e os valores reais não foram muito altos, sobretudo para T90% e PFE (menores que 0,89). Este resultado pode ser atribuído aos altos valores de repetibilidade da técnica padrão (acima de 6°C), semelhante ao que ocorre na determinação das temperaturas de destilação do diesel. Os valores de RMSEC de T90% e PFE para os modelos que forneceram maiores valores de RMSEP (modelos com correção de linha base) correspondem a 4,2119 °C e 2,4984°C respectivamente. Portanto os valores de RMSEC e RMSEP mostraram-se próximos para estas duas propriedades.

Barzilai *et al*^[73] determinou a temperatura de destilação T10% a partir de dados obtidos por espectroscopia no infravermelho próximo de modo on-line com fibras óticas. Obtiveram para esta propriedade um erro padrão de previsão de 4,09°C através do método PLS. Neste trabalho, foram obtidos RMSEP menores que 2°C para T10%. No entanto, os autores alegaram problemas de flutuação intensa na linha base, devido ao uso de fibras óticas.

Para a aplicação do método MLR foram selecionados números de onda por AG e APS. O método AG selecionou 32, 37 e 33 variáveis para T10%, T90% e PFE, respectivamente.

Os gráficos de variáveis selecionadas em função de RMSEV obtidos pelo método de seleção APS são mostrado na figura 7.1.40 para as propriedades de destilação da gasolina:

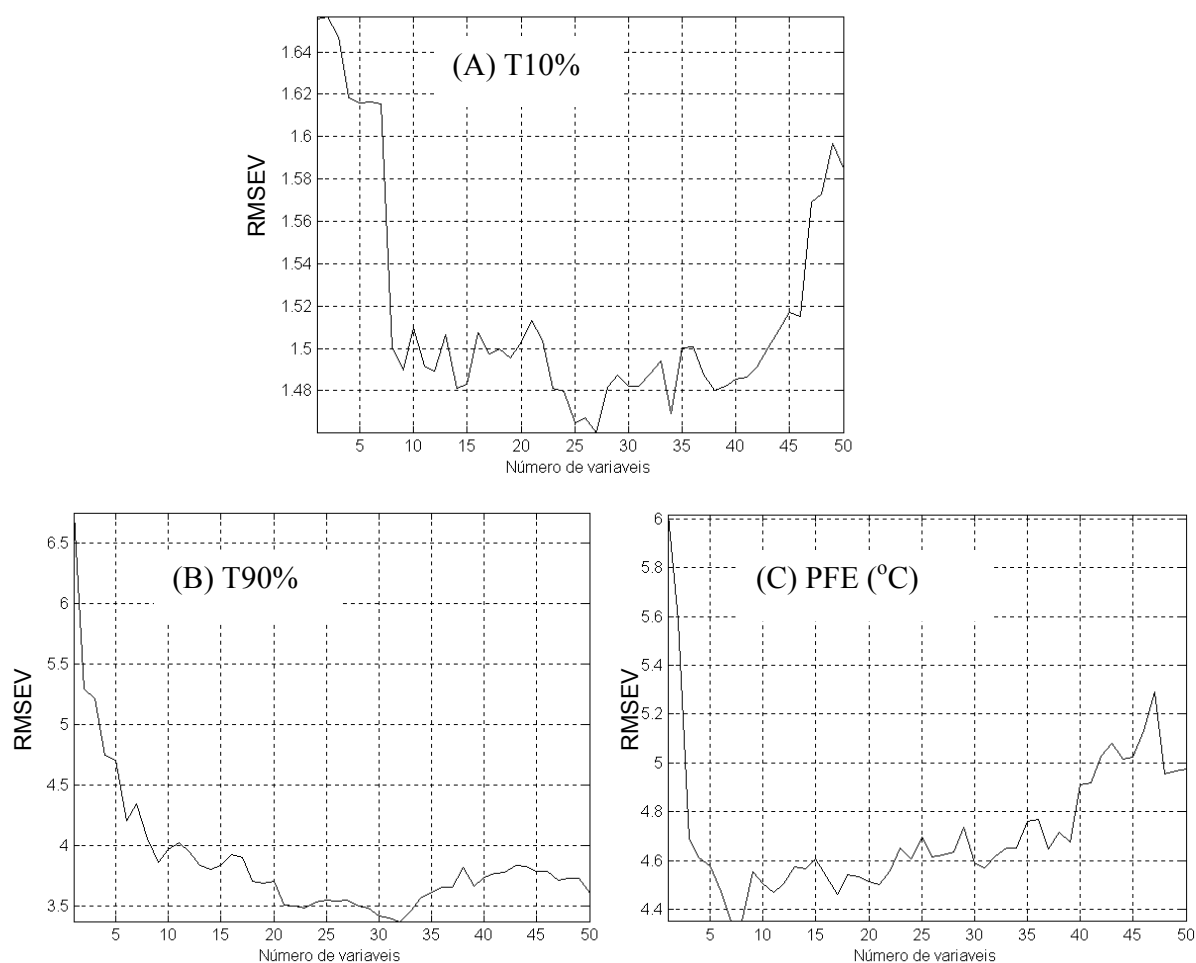


Figura 7.1.40: Resultado da seleção APS. (A) T10% com 27 variáveis selecionadas. (B) T90% com 32 variáveis selecionadas e (C) PFE com 8 variáveis selecionadas através do critério de menor RMSEV.

O método APS selecionou um número de variáveis menor que o método AG em relação às propriedades de destilação da gasolina.

As variáveis selecionadas foram empregadas no método MLR, cujos resultados são mostrados nas figuras 7.1.41, 7.1.42 e 7.1.43:

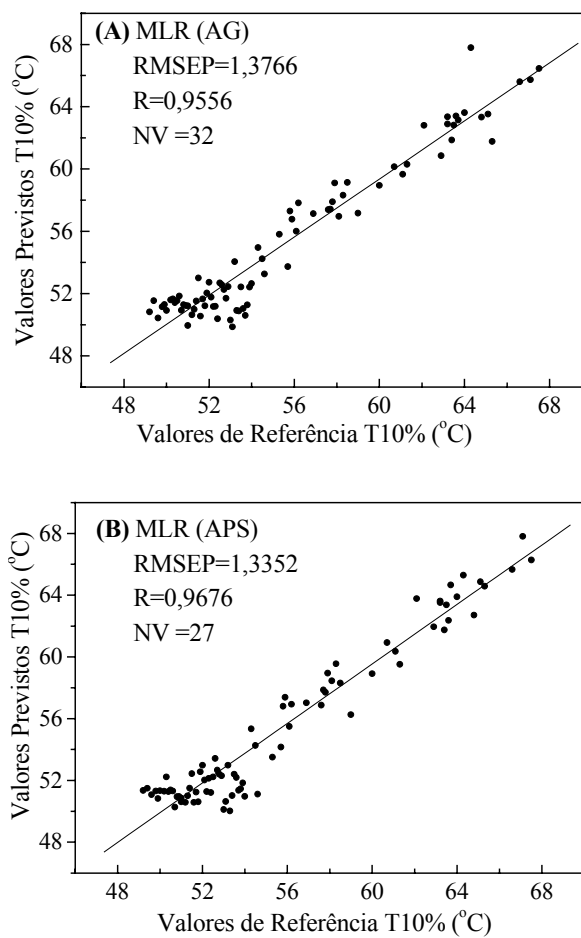


Figura 7.1.41: Comparação entre valores de referência e valores previstos por modelos MLR para a temperatura a 10% de destilado (gasolina). (A) variáveis selecionadas por AG e (B) variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

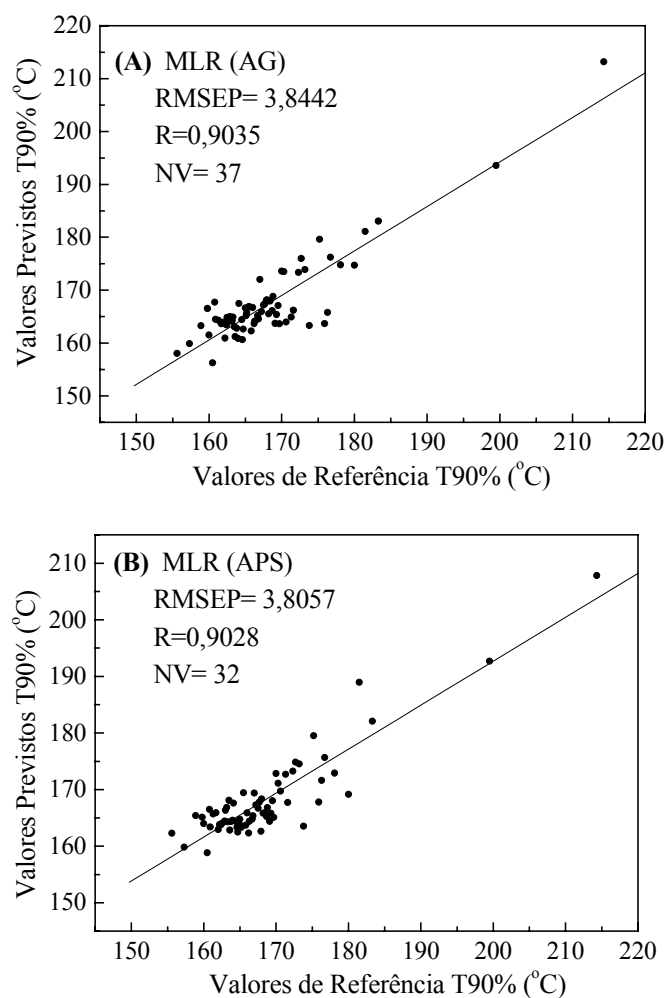


Figura 7.1.42: Comparação entre valores de referência e valores previstos por modelos MLR para a temperatura de 90% de destilado (gasolina). (A) variáveis selecionadas por AG e (B) variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

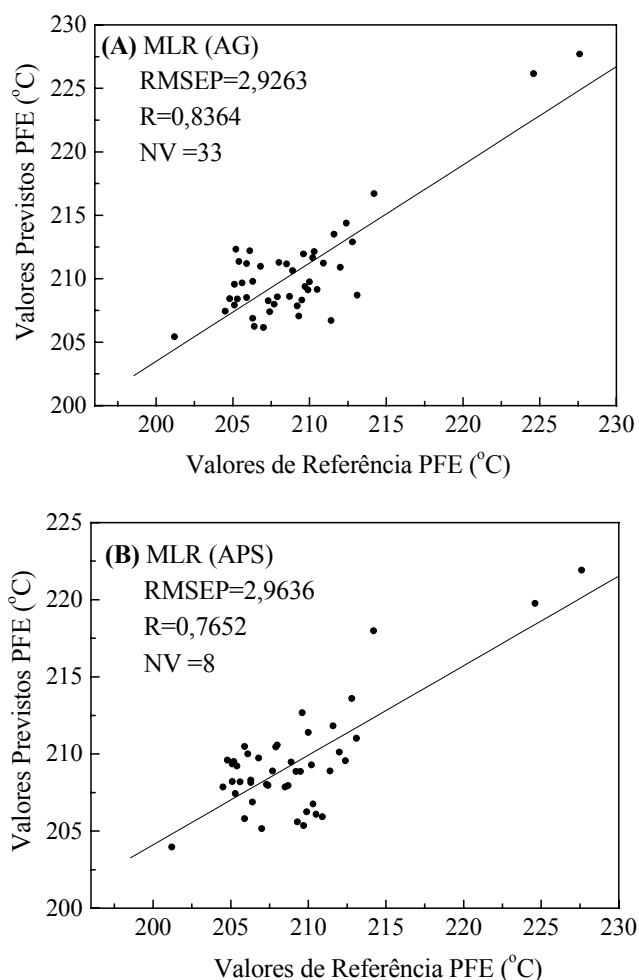


Figura 7.1.43: Comparação entre valores de referência e valores previstos por modelos MLR o ponto final de evaporação (gasolina).(A) variáveis selecionadas por AG e (B) variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

De um modo geral, em relação à previsão das propriedades de destilação da gasolina T10% e T90%, através do método MLR, os menores valores de RMSEP foram obtidos com os modelos construídos com variáveis selecionadas pelo método APS. Em relação à previsão de PFE, o modelo MLR(AG) forneceu menor valor de RMSEP. No entanto o modelo MLR(AG) empregou 33 variáveis contra 8 variáveis usadas no método MLR (APS).

A tabela 7.1.11 mostra um sumário com os resultados obtidos para a previsão das propriedades de destilação da gasolina:

Tabela 7.1.11: Resultados para a Previsão das propriedades de destilação da gasolina através de modelos de Calibração Multivariada

T10% (Gasolina) - Faixa: 48,8-69,7 °C - Reprodutibilidade Método Padrão: 6,8 °C	Modelo	VL	RMSEP	R	NV
	PLS (linha base)	3	1,4919	0,9606	1240
	PLS (1ª Derivada)	3	1,4377	0,9625	718
	PLS (AG)	5	1,4193	0,9631	32
	MLR (AG)	-	1,3766	0,9551	32
	MLR (APS)	-	1,3352	0,9676	27
T90% (Gasolina) - Faixa: 143,9-215,5 °C - Reprodutibilidade do Método Padrão: 8,4 °C	Modelo	VL	RMSEP	R	NV
	PLS (linha base)	6	4,3519	0,8715	1240
	PLS (1ª Derivada)	8	3,9287	0,8974	718
	PLS (AG)	8	3,7195	0,8942	37
	MLR (AG)	-	3,8442	0,9035	37
	MLR (APS)	-	3,8057	0,9028	32
PFE (Gasolina) - Faixa: 202,2-248,7 °C - Reprodutibilidade do Método Padrão: 10,5 °C	Modelo	VL	RMSEP	R	NV
	PLS (linha base)	3	2,6389	0,8197	1240
	PLS (1ª Derivada)	3	2,5369	0,8384	718
	PLS (AG)	5	2,5851	0,8551	33
	MLR (AG)	-	2,9636	0,7652	33
	MLR (APS)	-	2,9263	0,8364	8

VL: variáveis latentes; NV: número de variáveis independentes; R: coeficiente de correlação.

Comparando-se os valores de RMSEP obtidos através dos métodos PLS e MLR, observa-se que os melhores resultados são alcançados através do método MLR com variáveis selecionadas através do método APS para as temperaturas T10% e T90%.

Para PFE, o melhor resultado foi alcançado através do modelo PLS com primeira derivada. No entanto, utilizou-se um número bem superior de variáveis independentes neste modelo. Todos os valores de RMSEP obtidos estiveram abaixo dos valores de reprodutibilidade do método padrão ASTM.

- Propriedades de Octanagem (MON, RON e IAD)

Os modelos de calibração construídos através dos métodos PLS e MLR utilizaram para a propriedade MON (motor octane number): 82 amostras de calibração, 66 amostras de teste e 43 amostras de previsão.

Para RON (research octane number) foram empregadas 73 amostras de calibração, 59 amostras de teste e 53 amostras de previsão.

Para IAD (índice antidetonante), utilizou-se 67 amostras de calibração, 59 amostras de teste e 61 amostras de previsão.

Os números de variáveis latentes escolhidos, segundo o critério de menor RMSEV

para os modelos PLS são mostrados na figura 7.1.44:

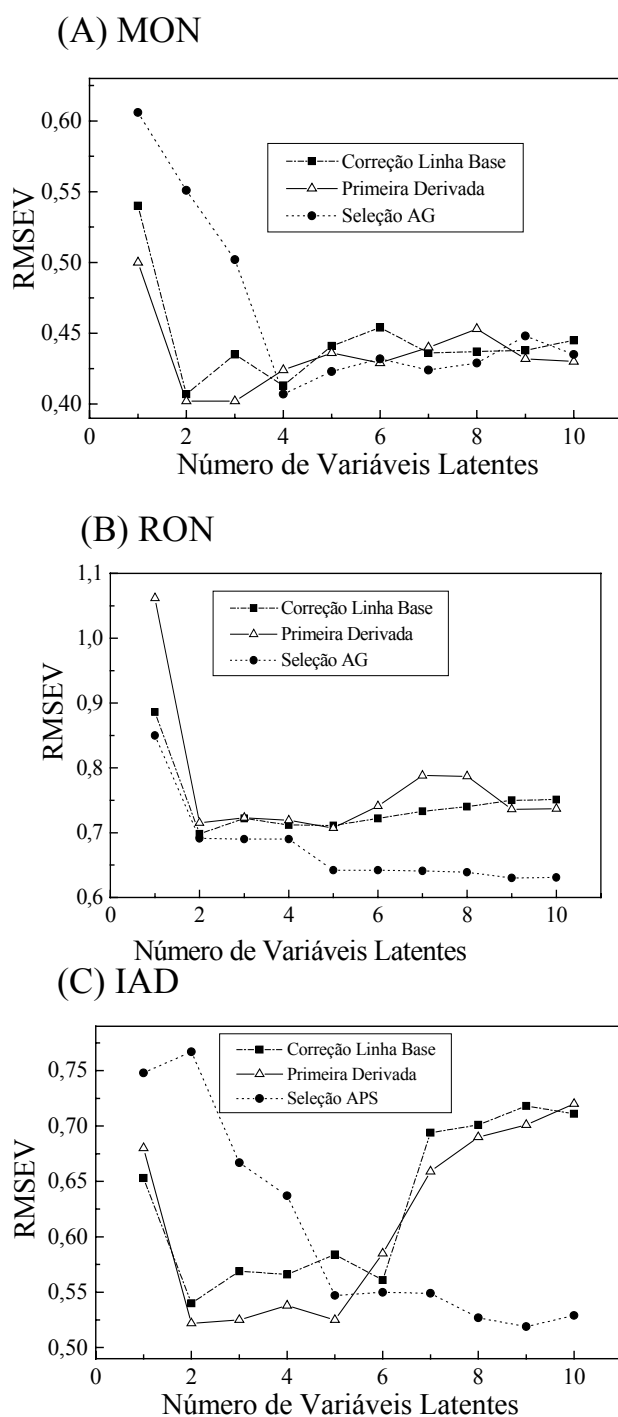


Figura 7.1.44: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para (A) MON, (B) RON e (C) IAD (gasolina) com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

Em relação à propriedade MON, os modelos PLS com correção de linha base e em primeira derivada foram descritos com 2 variáveis latentes. O modelo PLS(AG) foi descrito com 4 variáveis latentes.

Para a propriedade RON, todos os modelos PLS foram descritos com 2 variáveis latentes.

Os modelos PLS construídos para a previsão do índice ani-detonante IAD empregaram 2 variáveis latentes para os modelos PLS com correção de linha base e em primeira derivada e 5 variáveis latentes para o modelo PLS (AG).

Observou-se nos gráficos mostrados na 7.1.44 que o uso de mais variáveis latentes podem acarretar superajustamento (overfitting).

As faixas das propriedades medidas correspondem a 77,9-82,8 para MON; 89,0- 97,5 para RON e 83,2 e 90,1 para IAD.

O resultado da previsão dos modelos PLS é mostrado nas figuras 7.1.45; 7.1.46 e 7.1.47:

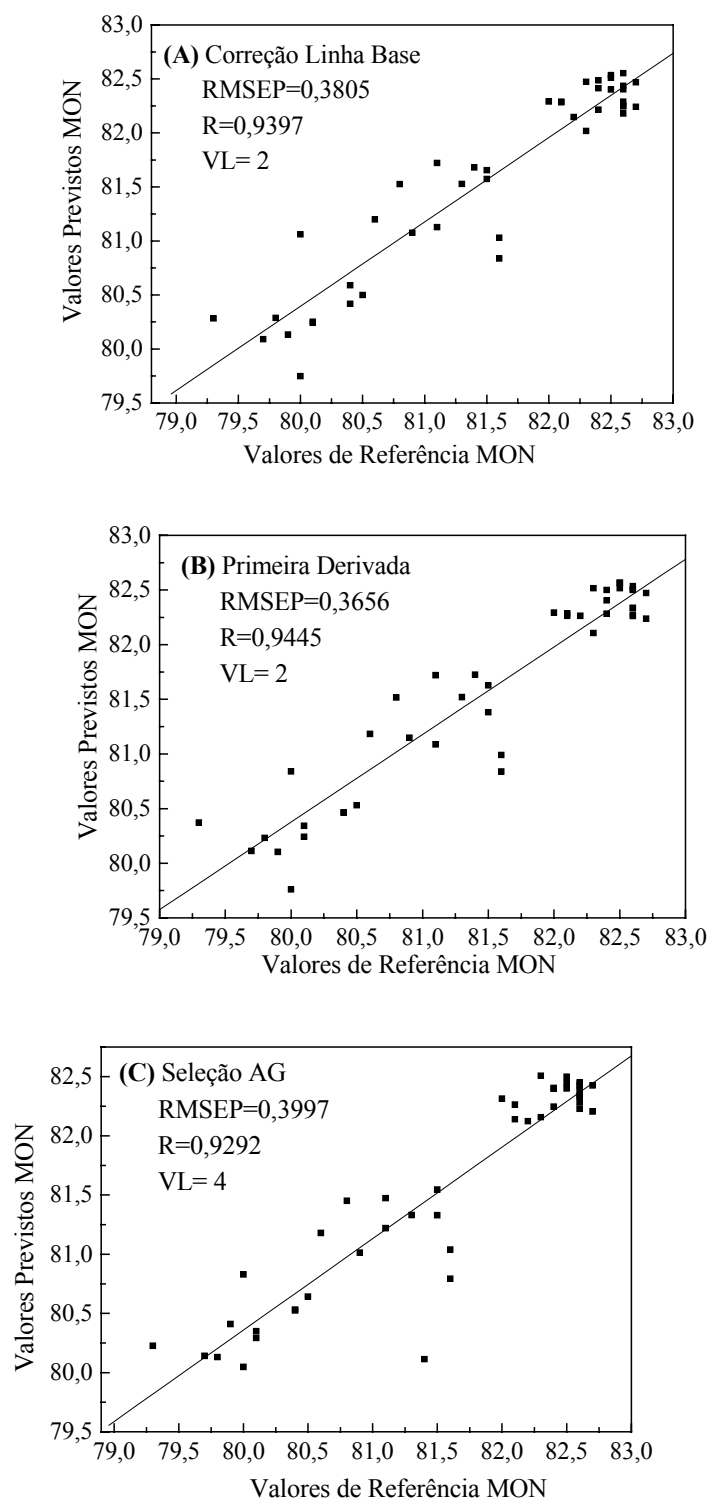


Figura 7.1.45: Comparação entre valores de referência e valores previstos por modelos PLS o parâmetro MON. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

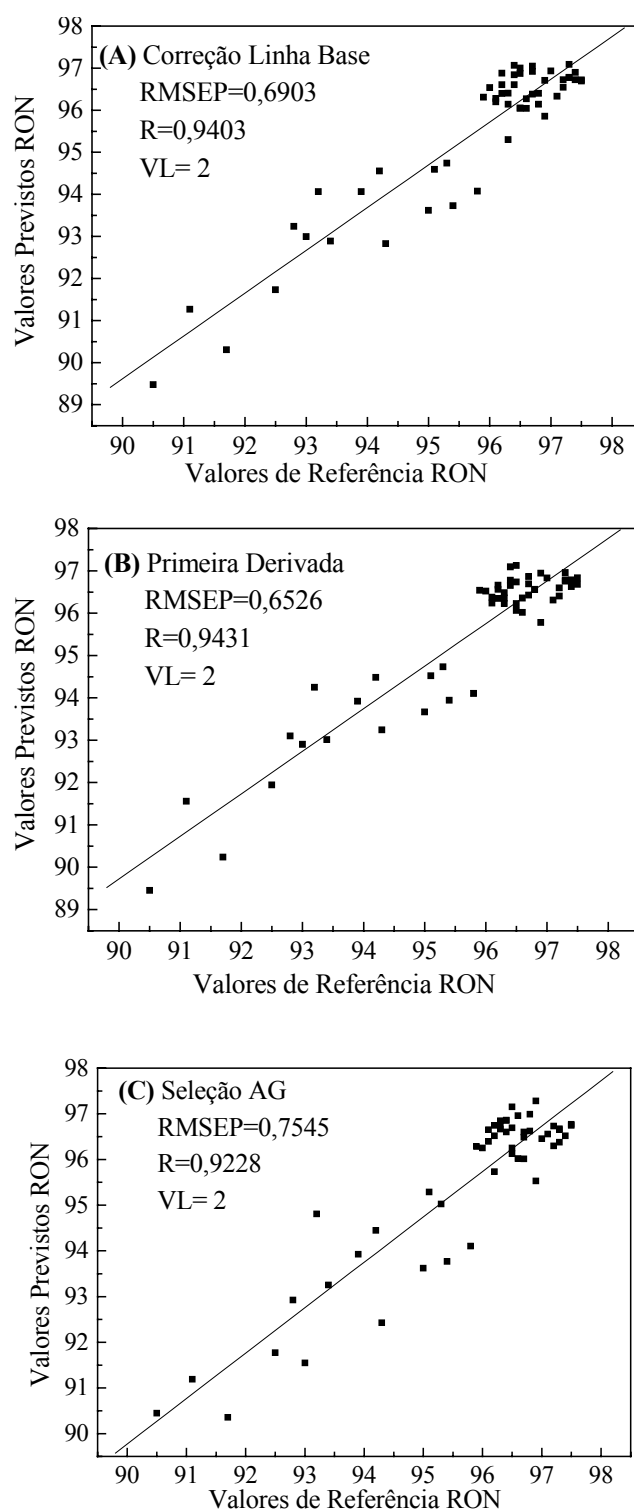


Figura 7.1.46: Comparação entre valores de referência e valores previstos por modelos PLS o parâmetro RON. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

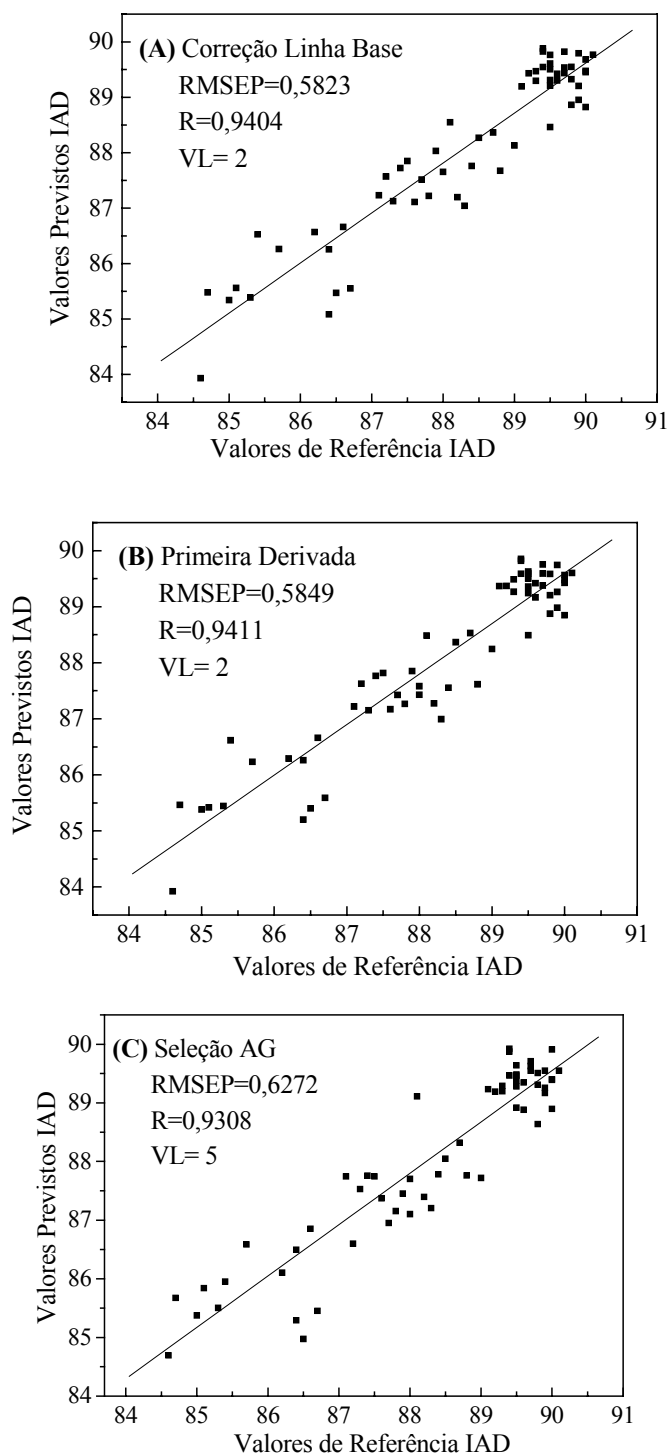


Figura 7.1.47: Comparação entre valores de referência e valores previstos por modelos PLS o parâmetro IAD. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

Os resultados mostraram bons coeficientes de correlação entre os valores reais e os valores previstos para todas as propriedades de detonação.

Em relação à previsão de MON e RON, os menores valores de RMSEP foram obtidos com os modelos PLS com primeira derivada. Para a previsão de IAD, o menor valor de RMSEP foi obtido através do modelo PLS com correção de linha base.

Fodor *et al* ^[82] obtiveram valores de RMSEP de 0,306 a 0,370 para MON; 0,487 a 0,495 para MON e 0,354 a 0,361 para IAD. Para tal, foi utilizado a espectroscopia no infravermelho médio com acessório de ATR horizontal. Os valores de RMSEP são inferiores aos obtidos neste trabalho. No entanto o acessório ATR com geometria horizontal apresenta praticidade inferior ao acessório circle-cell quanto à limpeza e facilidade de adaptação on-line.

Andrade e colaboradores^[67] obtiveram erro padrão de previsão de 0,45 para ambas propriedades MON e RON. Os resultados foram obtidos a partir de técnicas de transmissão no infravermelho médio, com cela de amostra de NaCl.

Os modelos PLS(AG) foram construídos a partir de 27 variáveis selecionadas para a previsão de MON; 18 variáveis selecionadas para RON e 22 variáveis independentes selecionadas para IAD.

As variáveis selecionadas pelo método AG foram empregadas em modelos MLR, assim como as variáveis selecionadas pelo método APS, cujos gráficos de RMSEV em função do número de variáveis são mostrados na figura 7.1.48:

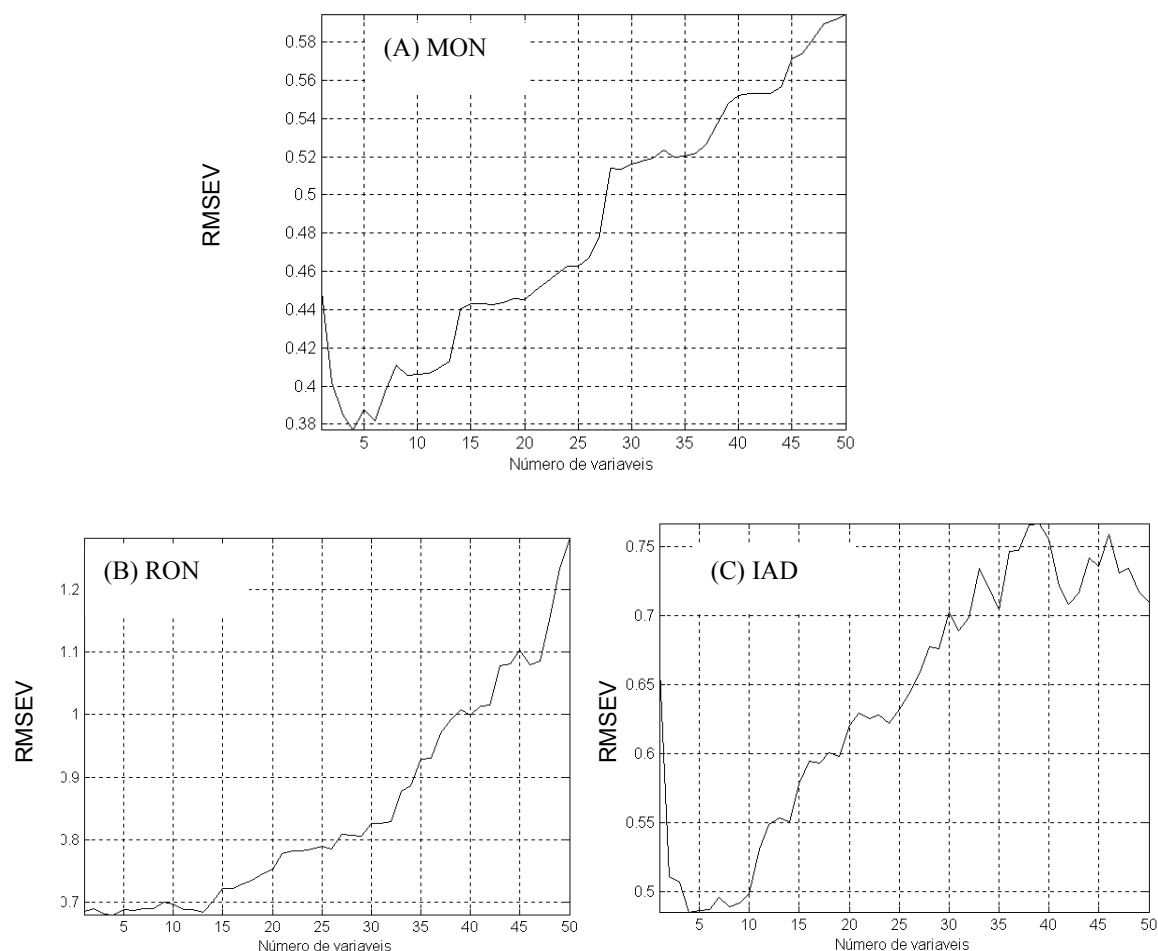


Figura 7.1.48: Resultado da seleção APS. (A) MON. (B) RON e (C) IAD. Todas com 4 variáveis independentes selecionadas através do critério de menor RMSEV.

O método APS selecionou um menor número de variáveis que o método AG. Para todas as propriedades de detonação (MON, RON e IAD) foram selecionadas 4 variáveis independentes.

O resultado da aplicação dos modelos MLR(AG) e MLR(APS) são mostrados nas figuras 7.1.49; 7.1.50 e 7.1.51 para as propriedades MON, RON e IAD, respectivamente:

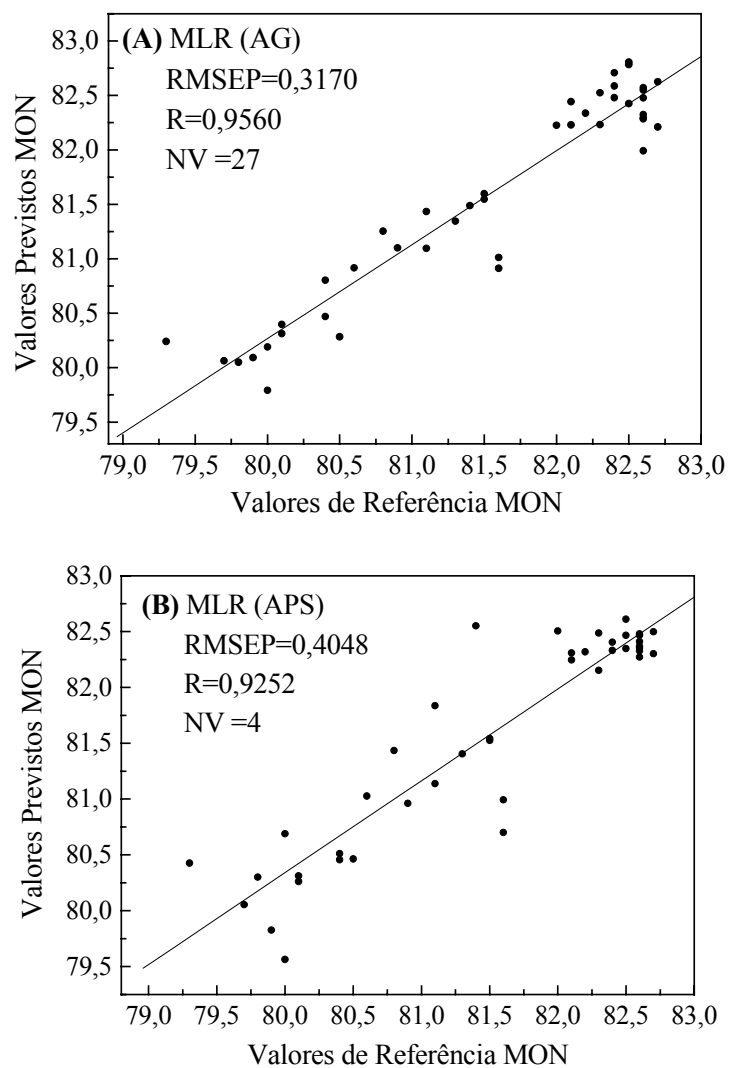


Figura 7.1.49: Comparação entre valores de referência e valores previstos para por modelos MLR para o parâmetro MON da gasolina. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

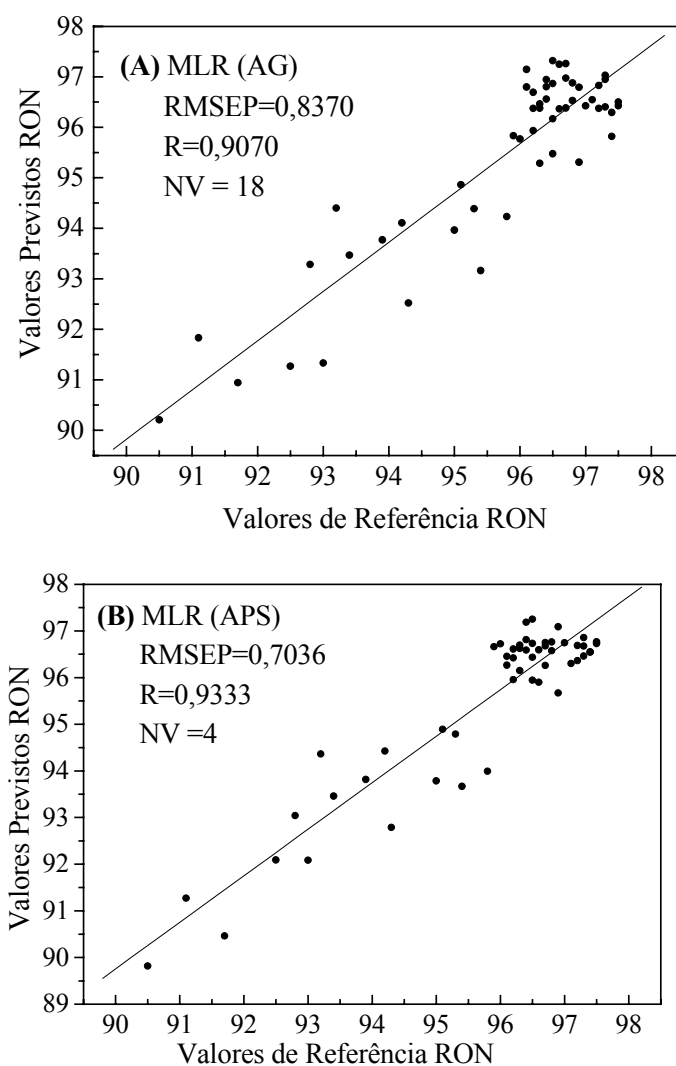


Figura 7.1.50: Comparação entre valores de referência e valores previstos para por modelos MLR para o parâmetro RON da gasolina. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

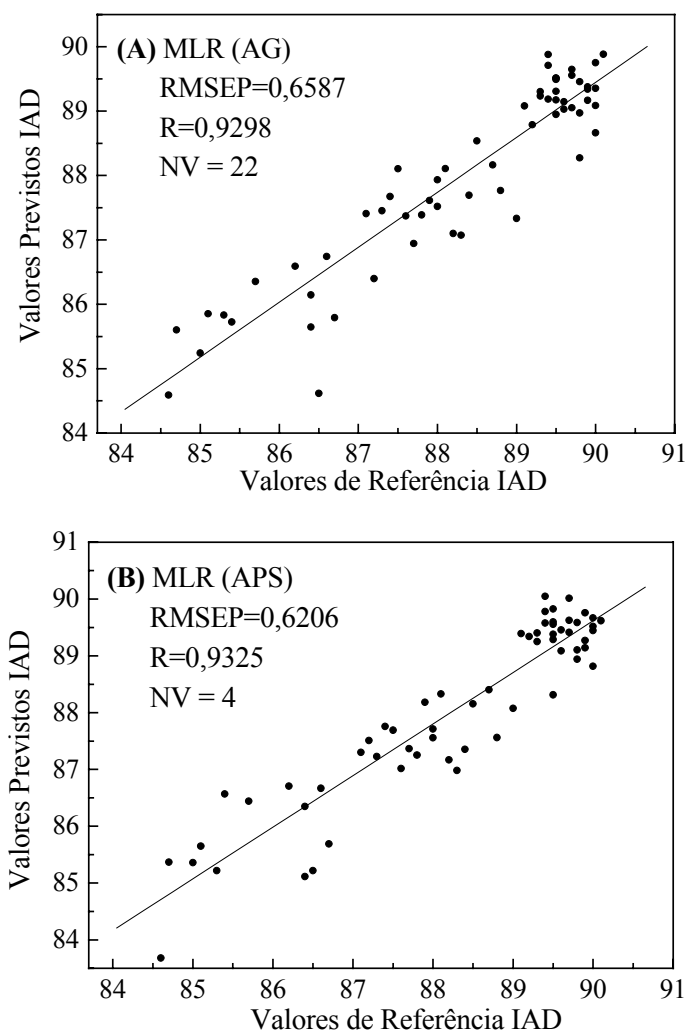


Figura 7.1.51: Comparação entre valores de referência e valores previstos para por modelos MLR para o parâmetro IAD da gasolina. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

O menor valor obtido dentre os modelos MLR para MON, corresponde ao modelo que empregou variáveis selecionadas pelo método AG. Para RON e IAD os menores valores de RMSEP foram obtidos com os modelos MLR(APS), utilizando-se um número inferior de variáveis independentes.

A tabela 7.1.12 mostra os resultados obtidos em todos os modelos de calibração construídos neste trabalho:

Tabela 7.1.12: Resultados para a Previsão das propriedades de detonação da gasolina através de modelos de Calibração Multivariada

MON (gasolina)	Modelo	VL	RMSEP	R	NV
- Faixa de Valores: 77,9-82,8 - Reprodutibilidade do Método Padrão: 0,9	PLS (linha base)	2	0,3805	0,9397	1240
	PLS (1ª Derivada)	2	0,3656	0,9445	718
	PLS (AG)	4	0,3997	0,9292	27
	MLR (AG)	-	0,3170	0,9560	27
	MLR (APS)	-	0,4048	0,9252	4
RON (gasolina)	Modelo	VL	RMSEP	R	NV
- Faixa de Valores: 89,0- 97,5 - Reprodutibilidade do Método Padrão: 0,7	PLS (linha base)	2	0,6903	0,9403	1240
	PLS (1ª Derivada)	2	0,6526	0,9431	718
	PLS (AG)	2	0,7545	0,9228	18
	MLR (AG)	-	0,8370	0,9070	18
	MLR (APS)	-	0,7036	0,9333	4
IAD (gasolina)	Modelo	VL	RMSEP	R	NV
- Faixa de Valores: 83,2-90,1 - Reprodutibilidade do Método Padrão: 0,9	PLS (linha base)	2	0,5823	0,9404	1240
	PLS (1ª Derivada)	2	0,5849	0,9411	718
	PLS (AG)	5	0,6272	0,9308	22
	MLR (AG)	-	0,6587	0,9298	22
	MLR (APS)	-	0,6206	0,9325	4

VL: variáveis latentes; NV: número de variáveis independentes; R: coeficiente de correlação.

Os valores de RMSEP obtidos nos modelos propostos para a previsão de MON e IAD estiveram abaixo da reprodutibilidade do método padrão (0,9). Alguns valores de RMSEP obtidos na previsão de RON estiveram um pouco acima da reprodutibilidade do método padrão (0,7).

- Período de Indução

O parâmetro período de indução foi medido na seguinte faixa 34,2 a 61,7 minutos apenas para as gasolinas com suspeita de adulteração (fiscalização).

Os modelos de calibração PLS e MLR foram construídos a partir de 71 amostras de calibração, 46 amostras de teste e 56 amostras de previsão.

A figura 7.1.52 mostra o número de variáveis latentes selecionadas para os modelos PLS construídos para a previsão de PI:

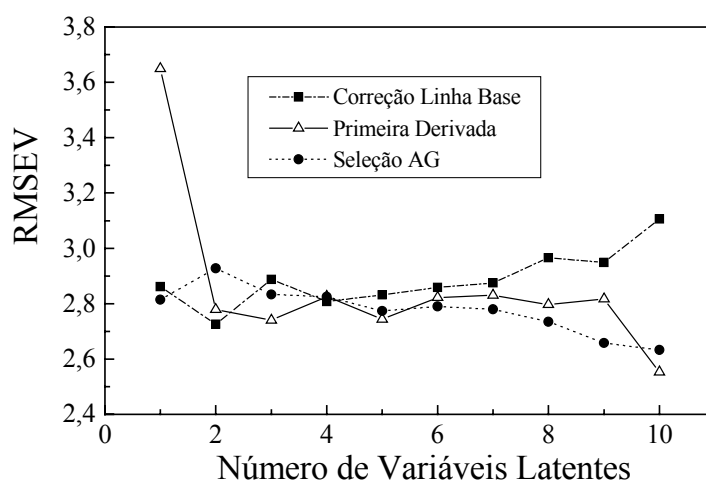


Figura 7.1.52: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para o período de indução da gasolina com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

Os modelos PLS com correção de linha base e PLS com primeira derivada foram descritos com duas variáveis latentes e o modelo PLS com seleção de variáveis através do método AG foi descrito com uma variável latente.

Os resultados da previsão de PI através dos modelos PLS são mostrados na figura 7.1.53:

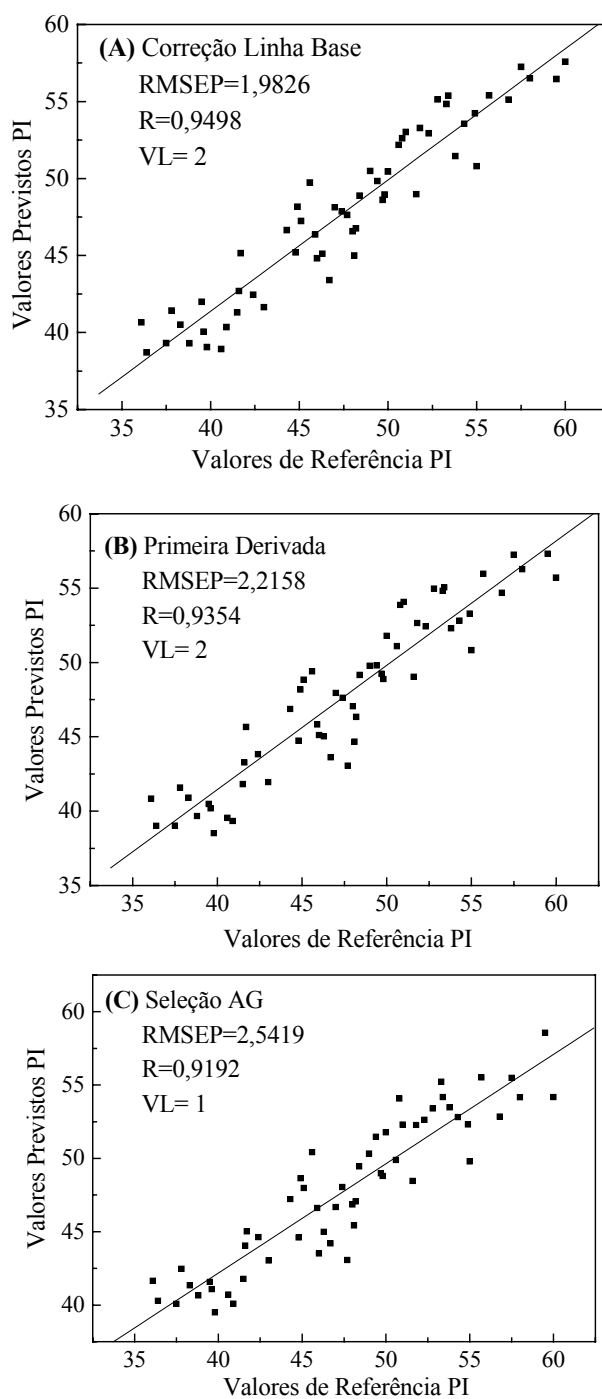
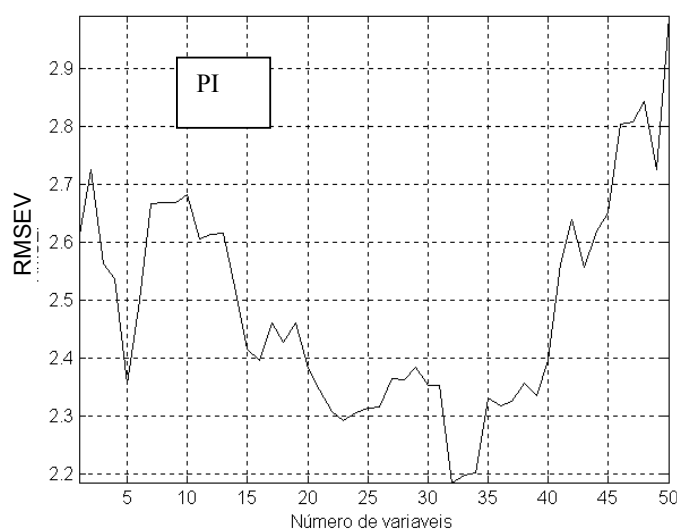


Figura 7.1.53: Comparação entre valores de referência e valores previstos por modelos PLS o parâmetro PI. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

O modelo PLS com correção de linha base forneceu o menor valor de RMSEP e o maior coeficiente de regressão, embora o modelo PLS(AG) tenha sido descrito com uma variável latente a menos.

O modelo PLS (AG) foi construído a partir de 35 variáveis selecionadas, contra 1240 variáveis de PLS (correção de linha base) e 718 variáveis de PLS com primeira derivada.

O método APS selecionou um número de 32 de variáveis, cujo gráfico de RMSEP em função das variáveis selecionadas é mostrado na figura 7.1.54:



7.1.54: Resultado da seleção APS. Foram selecionadas 32 variáveis independentes através do critério de menor RMSEP.

As variáveis independentes selecionadas pelos métodos AG e APS foram aplicadas na construção de modelos MLR e os resultados são mostrados na figura 7.1.55:

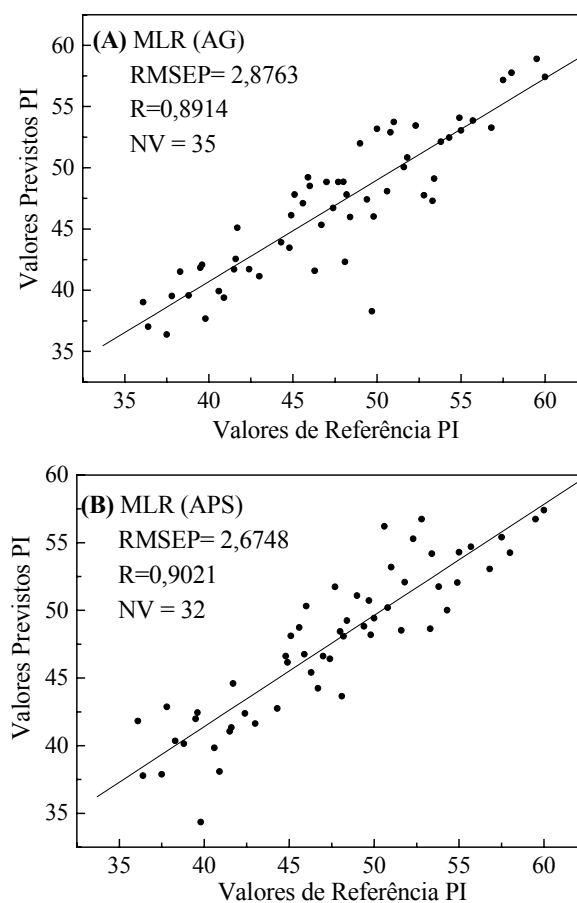


Figura 7.1.55: Comparação entre valores de referência e valores previstos para por modelos MLR para o parâmetro PI da gasolina. (A) Variáveis selecionadas por AG e (B) variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

O modelo construído a partir das variáveis selecionadas pelo método APS forneceu menor valor de RMSEP, quando comparado aos resultados do modelo PLS(AG), que empregou um número maior de variáveis.

Os resultados para a previsão de PI estão resumidos na tabela 7.1.13:

Tabela 7.1.13: Resultados para a previsão do período de indução da gasolina através de modelos de Calibração Multivariada

PI (Gasolina)	Modelo	VL	RMSEP	R	NV
- Faixa : 34,2 – 61,7 - Reprodutibilidade do Método Padrão: 10%	PLS (linha base)	2	1,9826	0,9498	1240
	PLS (1ª Derivada)	2	2,2158	0,9354	718
	PLS (AG)	1	2,5419	0,9192	35
	MLR (AG)	-	2,8763	0,8914	35
	MLR (APS)	-	2,6748	0,9021	32

VL: variáveis latentes; NV: número de variáveis independentes; R: coeficiente de correlação.

O menor valor de RMSEP foi obtido com o modelo PLS com correção de linha base, que empregou o maior número de variáveis. Todos os valores obtidos pelos outros modelos estiveram abaixo da reprodutibilidade do método padrão, cerca de 3,4.

Não foram encontrados dados na literatura para comparação de resultados.

- Tipos de Hidrocarbonetos

Foram construídos modelos de calibração PLS e MLR para a determinação do teor (% em volume) de benzeno, hidrocarbonetos aromáticos e hidrocarbonetos saturados.

Para tal, Utilizou-se 41 amostras de calibração, 40 amostras de teste e 34 amostras de previsão para os modelos de previsão do teor de benzeno. Para o teor de aromáticos, foram empregadas 41 amostras de calibração, 43 amostras de teste e 42 amostras de previsão. Os modelos para a previsão do teor de saturados foram

construídos a partir de 48 amostras de calibração, 55 amostras de teste e 37 amostras de previsão.

A faixa de concentração de cada tipo de hidrocarboneto analisada foi de 0,10 a 1,2 % de benzeno; 8,0 a 21,0% de aromáticos e 32,0 a 60,0% de hidrocarbonetos saturados.

A figura 7.1.56 mostra o número de variáveis latentes selecionadas para os modelos PLS:

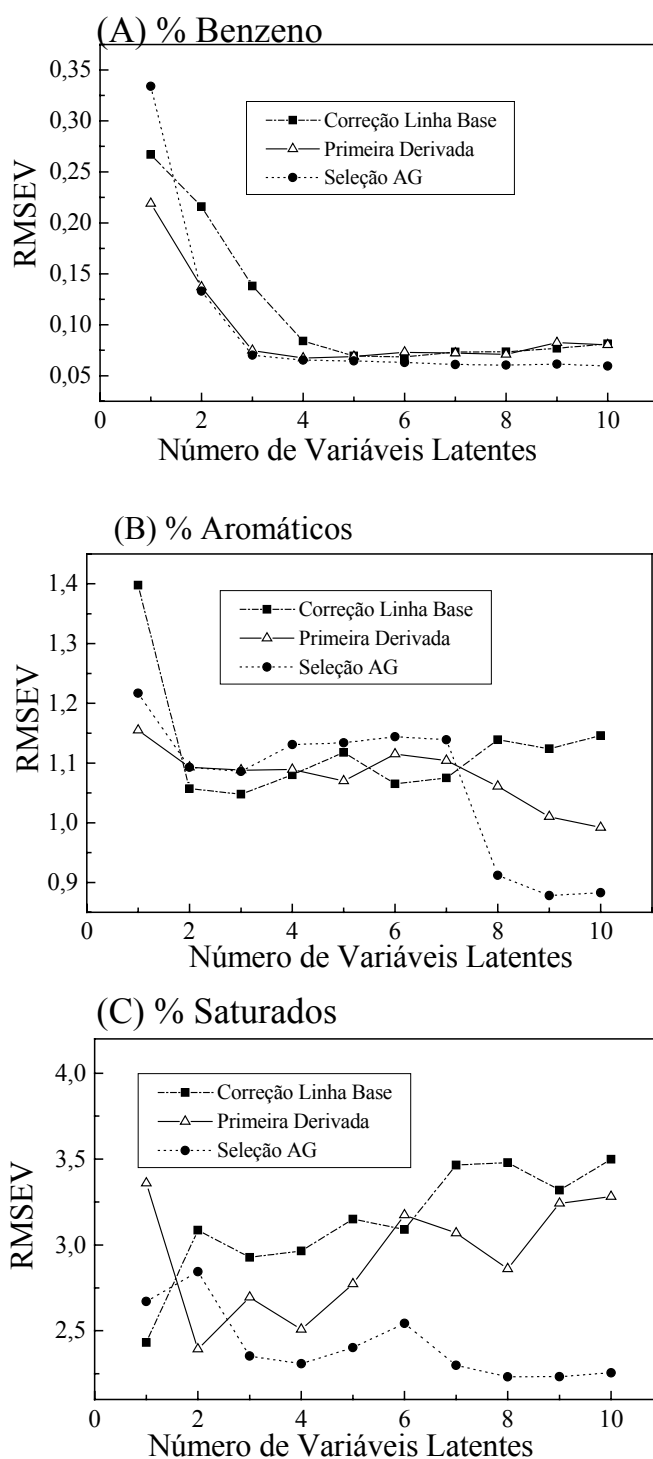


Figura 7.1.56: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para os teores de (A) Benzeno, (B) Aromáticos e (C) Saturados (gasolina) com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

Para a previsão do teor de benzeno, foi utilizado 4 variáveis latentes no modelo PLS com correção de linha base e 3 variáveis latentes para o modelo PLS (AG).

Em relação à previsão do teor de aromáticos, os modelos PLS com correção de linha base e PLS com primeira derivada foram descritos com 2 variáveis latentes e o modelo PLS (AG) foi descrito com 3 variáveis latentes.

Para a previsão do teor de saturados foi empregada 1 variável latente para o modelo com correção de linha base, duas variáveis latentes para PLS(primeira derivada) e 3 variáveis latentes para o modelo PLS com seleção AG.

Os resultados da previsão através dos modelos PLS são mostrados nas figuras 7.1.57; 7.1.58 e 7.1.59. O modelo PLS(AG) para determinação do teor de benzeno utilizou 18 números de onda. Para a determinação do teor de aromáticos e saturados foram empregados 23 números de onda.

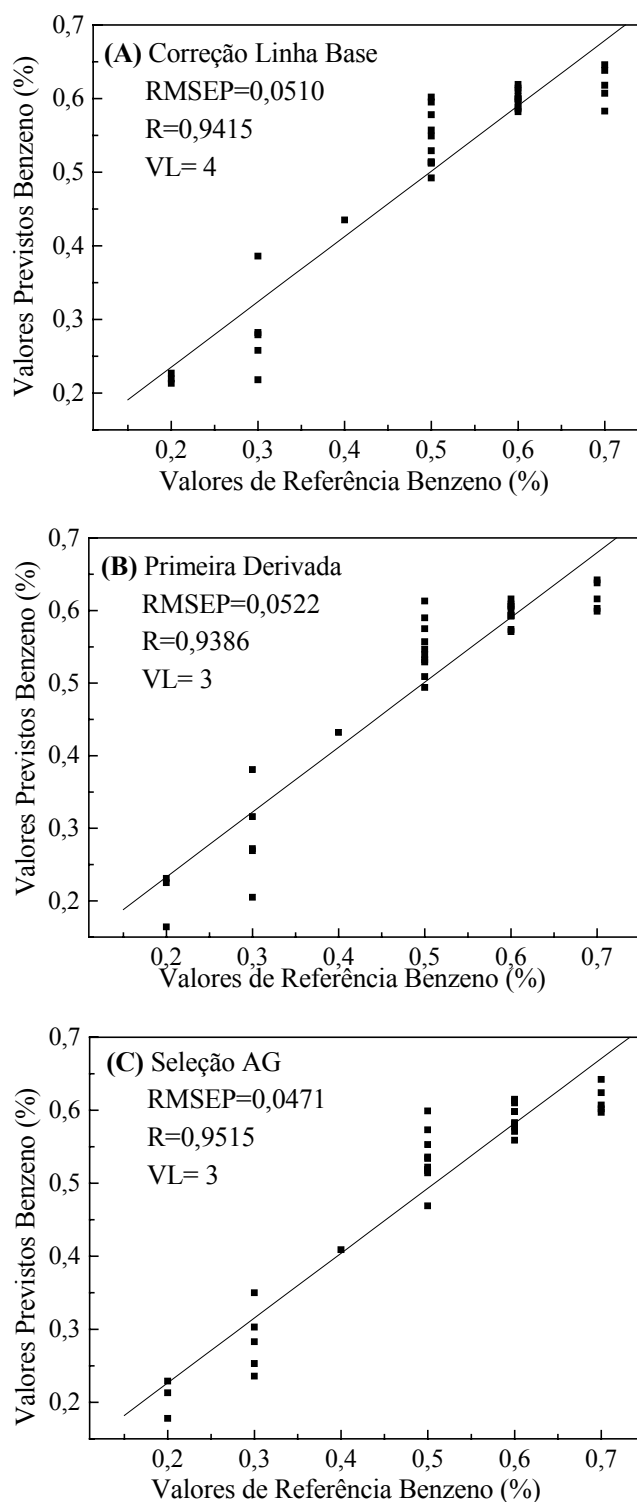


Figura 7.1.57: Comparação entre valores de referência e valores previstos por modelos PLS o teor de benzeno da gasolina. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

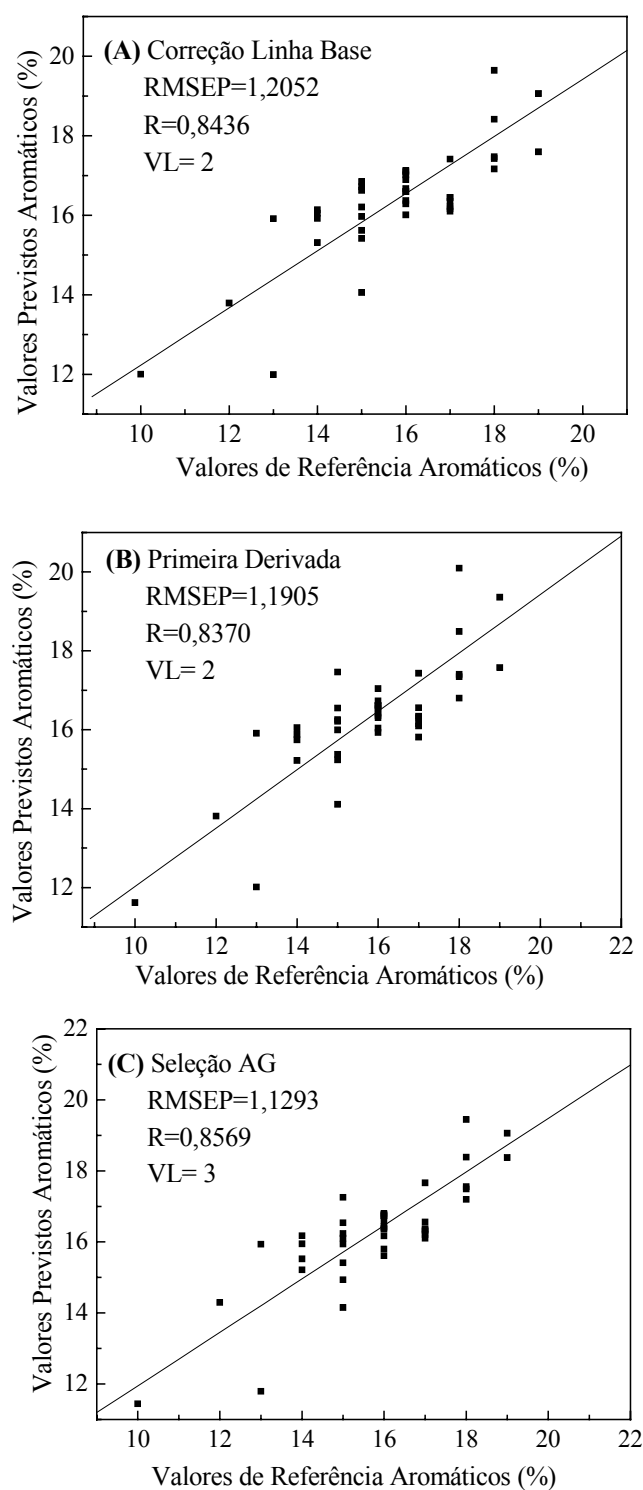


Figura 7.1.58: Comparação entre valores de referência e valores previstos por modelos PLS o teor de aromáticos da gasolina. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

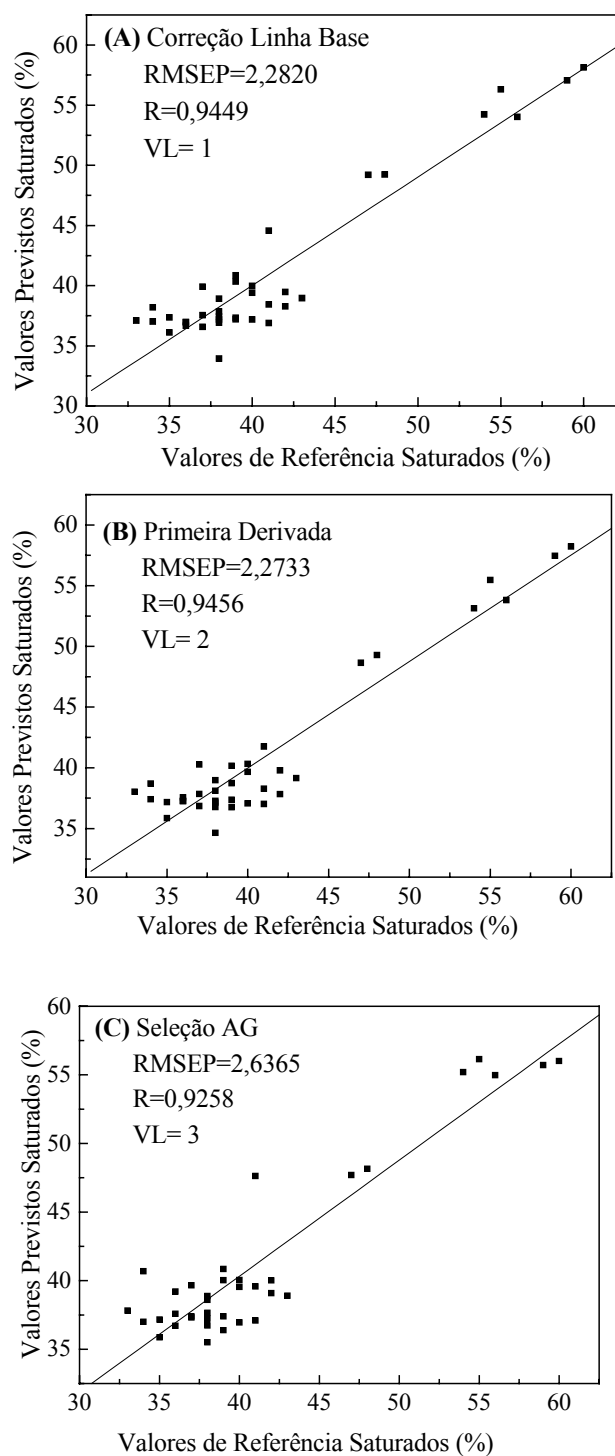


Figura 7.1.59: Comparação entre valores de referência e valores previstos por modelos PLS o teor de saturados da gasolina. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

Em relação à determinação do teor de benzeno, observou-se que o menor valor de RMSEP foi obtido através do modelo PLS com seleção de variáveis(AG).

O menor valor de RMSEP dentre os resultados obtidos para a previsão do teor de aromáticos foi obtido também através do modelo PLS(AG).

Já em relação à previsão do teor de hidrocarbonetos saturados, o menor valor de RMSEP corresponde ao modelo PLS com primeira derivada.

Cooper e colaboradores^[74] obtiveram erro padrão de previsão de 0,08% para o teor de benzeno, através de dados obtidos por espectroscopia no infravermelho médio com ATR, 0,190% através de dados obtidos no infravermelho próximo e 0,101 % através da espectroscopia Raman. Os resultados do presente trabalho mostraram valores de RMSEP abaixo de 0,053%. Adicionalmente foram utilizados de 3 a 4 variáveis latentes, contra 9 variáveis latentes utilizadas por Cooper *et al.*

Fodor e co-autores^[82] obtiveram um valor de erro padrão de previsão igual a 0,104% para o benzeno utilizando-se espectroscopia no infravermelho médio e ATR.

Portanto os resultados para a previsão do teor de benzeno no presente trabalho são compatíveis com principais dados da literatura.

Para o teor de aromáticos, o menor valor de RMSEP obtido por modelo PLS no presente trabalho corresponde a 1,2293%. Fodor *et al*^[82] obtiveram um erro de previsão de 1,912 para o mesmo tipo de hidrocarbonetos.

Fodor e colaboradores^[82] obtiveram um erro padrão de previsão de 2,008% para o teor de saturados. 2,2733% do valor de RMSEP obtido no presente trabalho.

Para a construção dos modelos MLR foram selecionadas variáveis através dos métodos APS e AG. O Número de variáveis selecionadas em função de RMSEP são os resultados obtidos através da seleção APS e são mostrados na figura 7.1.60:

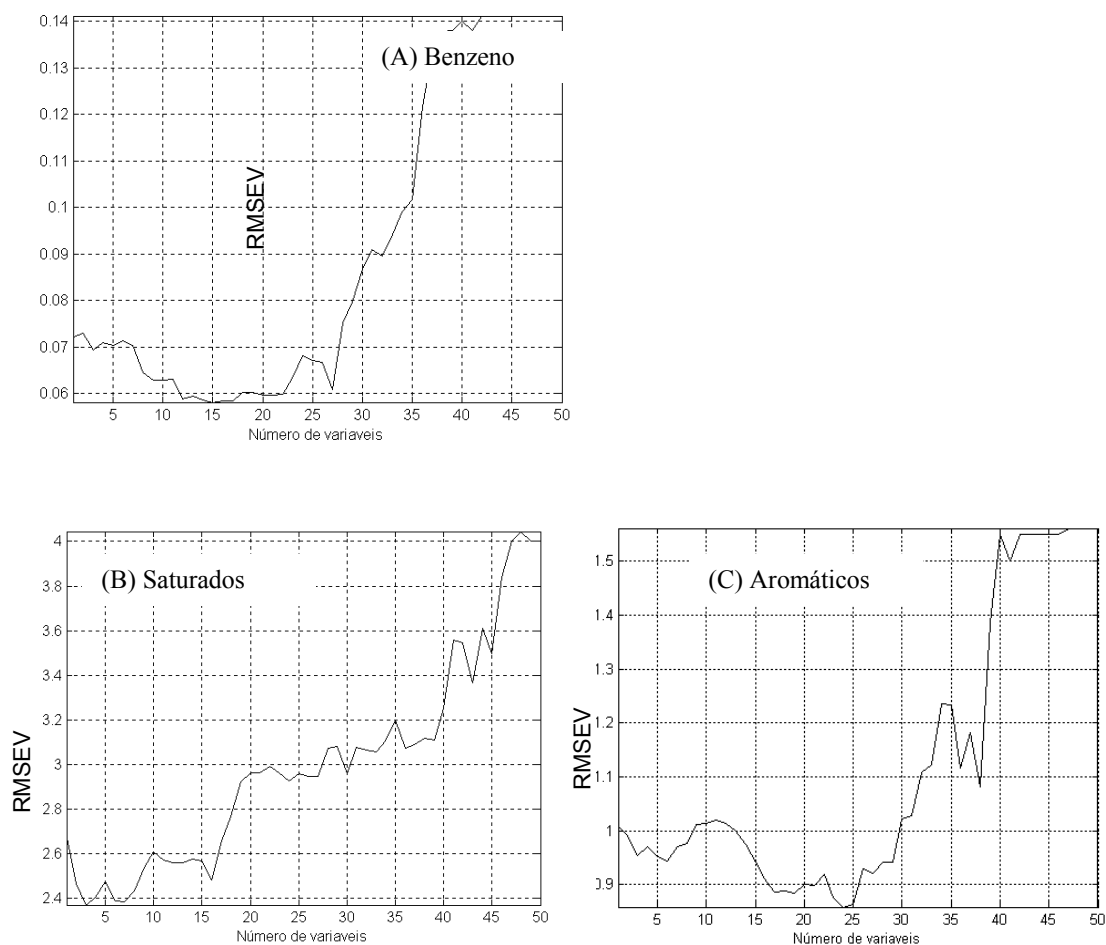


Figura 7.1.60: Resultado da seleção APS. (A) Benzeno (%) com 15 variáveis selecionadas. (B) Saturados (%) com 3 variáveis selecionadas e (C) Aromáticos (%) com 24 variáveis selecionadas através do critério de menor RMSEV.

O método MLR foi empregado, utilizando-se variáveis selecionadas pelos métodos AG e APS, cujos resultados são mostrados nas figuras 7.1.61; 7.1.62 e 7.1.63:

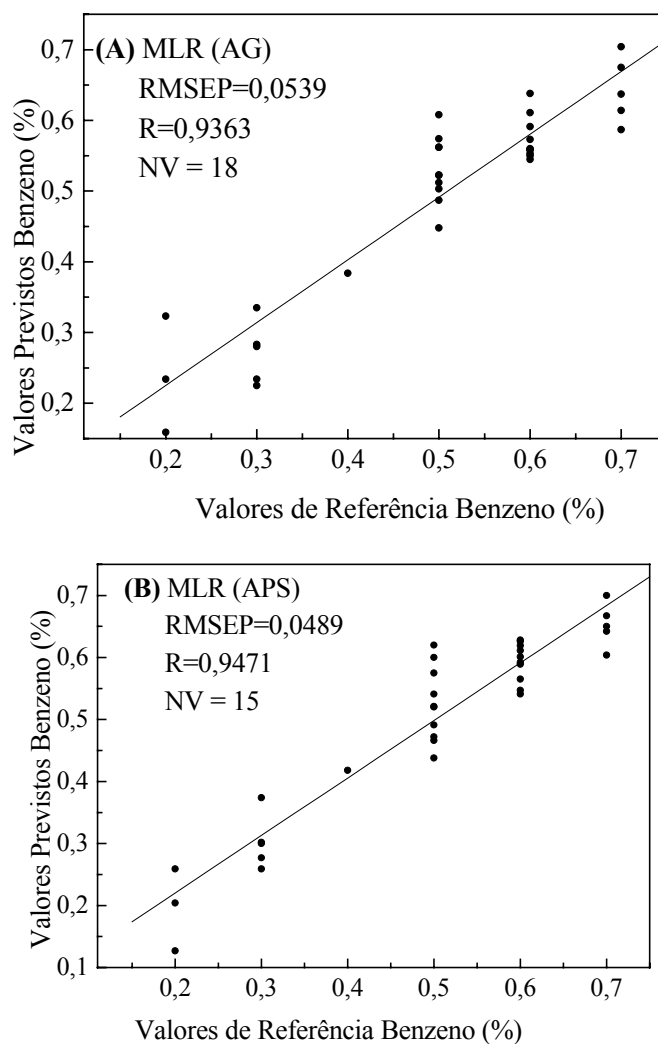


Figura 7.1.61: Comparação entre valores de referência e valores previstos para por modelos MLR para o teor de benzeno da gasolina. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

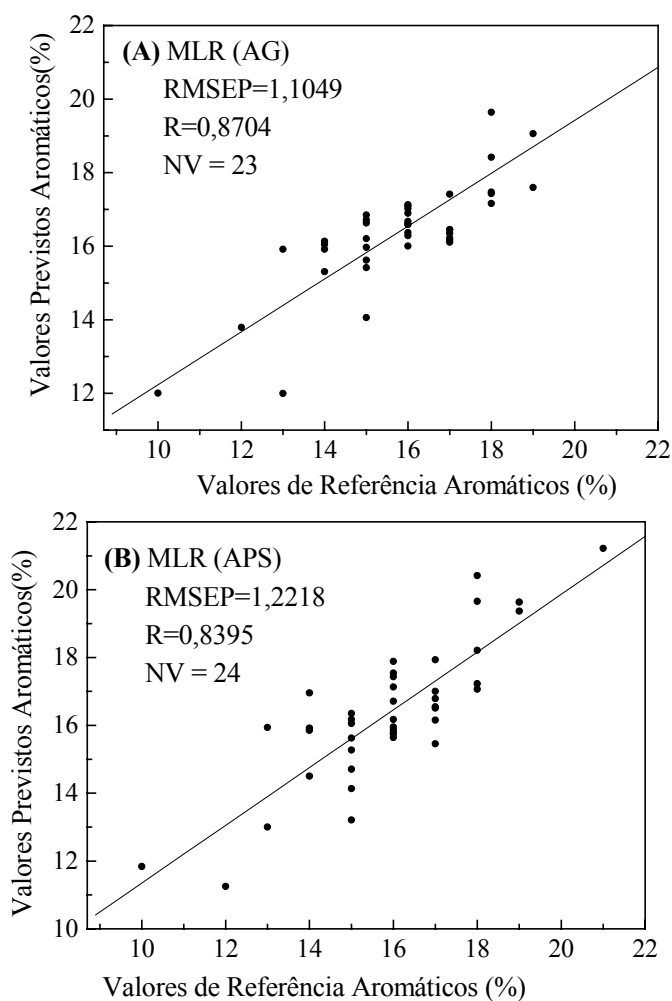


Figura 7.1.62: Comparação entre valores de referência e valores previstos para por modelos MLR para o teor de aromáticos da gasolina. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

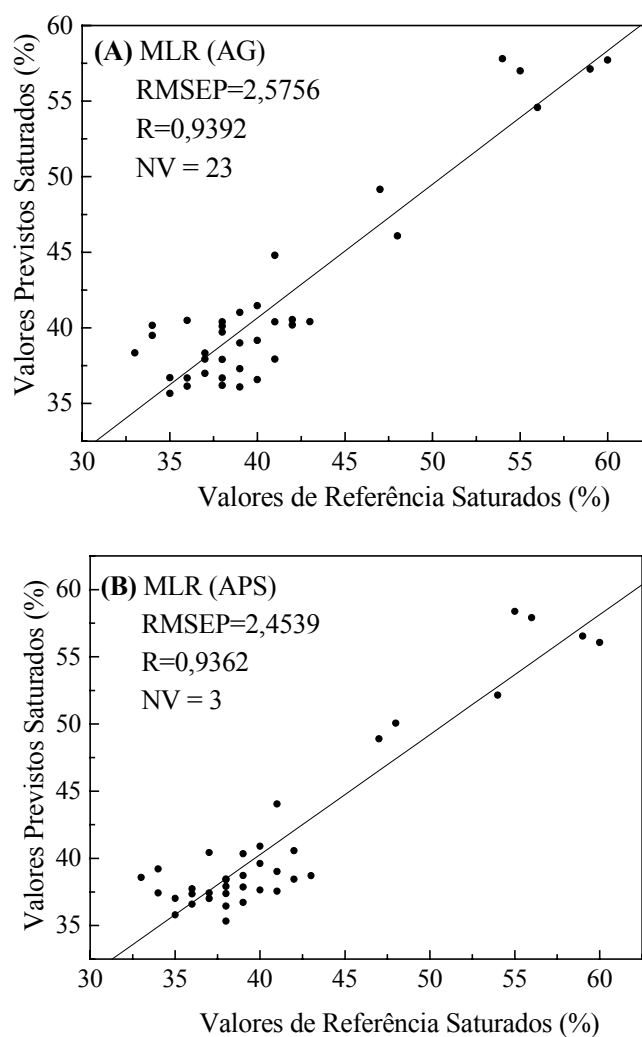


Figura 7.1.63: Comparação entre valores de referência e valores previstos para por modelos MLR para o teor de saturados da gasolina. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

Os modelos MLR(AG) forneceram menor valor de RMSEP para a previsão do teor de aromáticos. Os modelos MLR(APS) forneceram melhores resultados para a previsão dos teores de benzeno e saturados.

Com exceção do teor de aromáticos, o método APS selecionou um menor número de variáveis para a previsão da concentração de hidrocarbonetos na gasolina.

Os resultados fornecidos pelos métodos de calibração para os hidrocarbonetos da gasolina são mostrados na tabela 7.1.14:

Tabela 7.1.14: Resultados para a Previsão dos hidrocarbonetos da gasolina através de modelos de Calibração Multivariada

Benzeno - Faixa : 0,1 – 1,2 (%) - Reprodutibilidade do Método Padrão: 0,28%	Modelo	VL	RMSEP	R	NV
	PLS (linha base)	4	0,0510	0,9415	1240
	PLS (1ª Derivada)	3	0,0522	0,9386	718
	PLS (AG)	3	0,0471	0,9515	18
	MLR (AG)	-	0,0539	0,9363	18
	MLR (APS)	-	0,0489	0,9721	15
Aromáticos - Faixa : 8,0 – 22,0 (%) - Reprodutibilidade do Método Padrão: 3,7%	Modelo	VL	RMSEP	R	NV
	PLS (linha base)	2	1,2052	0,8436	1240
	PLS (1ª Derivada)	2	1,1902	0,8370	718
	PLS (AG)	3	1,1293	0,8569	23
	MLR (AG)	-	1,1044	0,8704	23
	MLR (APS)	-	1,2218	0,8395	24
Saturados - Faixa : 32,0-60,0 (%) - Reprodutibilidade do Método Padrão: 4,2%	Modelo	VL	RMSEP	R	NV
	PLS (linha base)	1	2,2820	0,9449	1240
	PLS (1ª Derivada)	2	2,2733	0,9456	718
	PLS (AG)	3	2,6365	0,9258	23
	MLR (AG)	-	2,5756	0,9322	23
	MLR (APS)	-	2,4539	0,9362	3

VL: variáveis latentes; NV: número de variáveis independentes; R: coeficiente de correlação.

Todos os valores de RMSEP obtidos através dos modelos propostos estiveram abaixo da reprodutibilidade do método padrão.

7.1.3.3- Modelos PLS e MLR para as propriedades do Álcool Automotivo

-Densidade a 20°C

A faixa de densidade das amostras de álcool automotivo analisada neste trabalho corresponde a 808,6 a 811,4 kg/m³.

Os modelos PLS e MLR foram construídos com 33 amostras de calibração, 22 amostras de teste e 31 amostras de previsão.

O número de variáveis latentes selecionadas nos modelos PLS está representado na figura 7.1.64, que mostra os valores de RMSEV em função das variáveis latentes selecionadas:

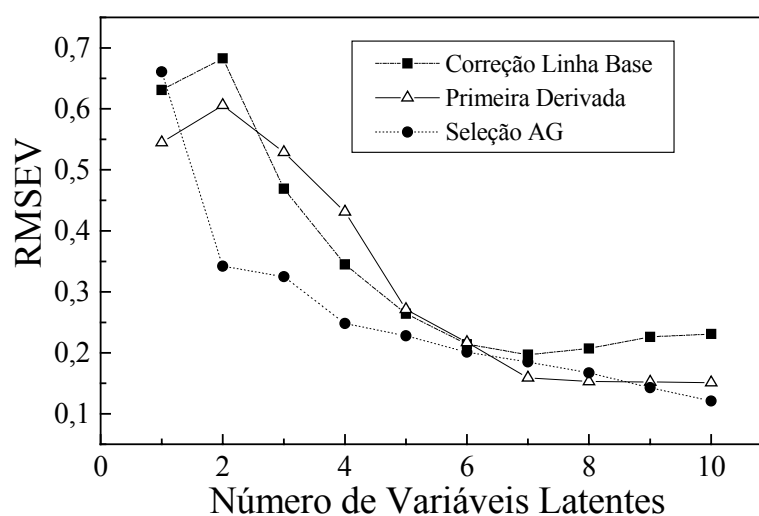


Figura 7.1.64: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para a densidade do álcool com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

Os modelos PLS com correção de linha base e PLS em primeira derivada utilizaram 7 variáveis latentes, enquanto que o modelo PLS(AG) foi descrito com 8 variáveis latentes.

O resultados da aplicação dos modelos PLS para a determinação da densidade do álcool automotivo é mostrado na figura 7.1.65:

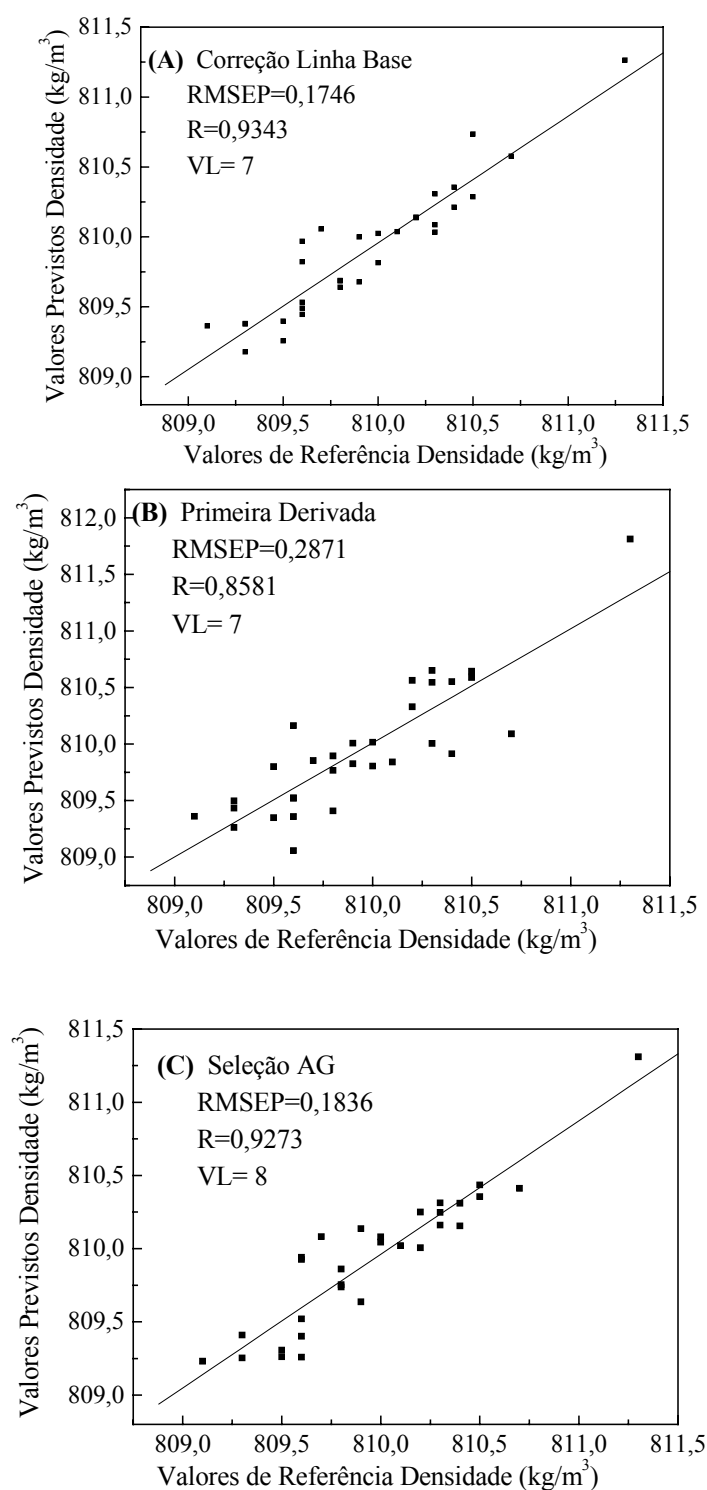
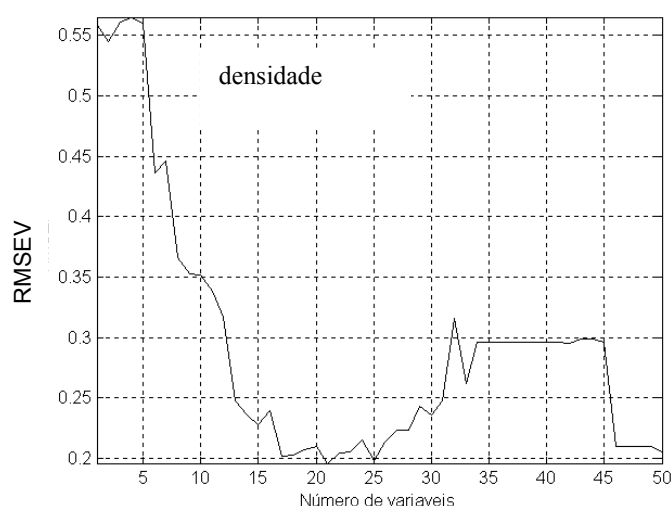


Figura 7.1.65: Comparação entre valores de referência e valores previstos por modelos PLS o parâmetro densidade do álcool. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

O menor valor RMSEP observado corresponde ao resultado fornecido pelo modelo PLS com primeira derivada, embora o valor de RMSEP fornecido pelo modelo PLS(AG) mostrou-se muito próximo a este valor, com a utilização de 27 variáveis independentes.

Através do método APS foram selecionadas 21 variáveis independentes, cujo resultado da seleção é mostrado no gráfico de RMSEV em função das variáveis selecionadas, mostradas na figura 7.1.66:



7.1.66: Resultado da seleção APS para a densidade do álcool. Foram selecionadas 21 variáveis independentes através do critério de menor RMSEV.

O método APS selecionou 21 variáveis contra 27 variáveis selecionadas pelo método AG.

As variáveis selecionadas pelos métodos AG e APS foram utilizadas nos modelos MLR, cujos resultados são mostrados na figura 7.1.67:

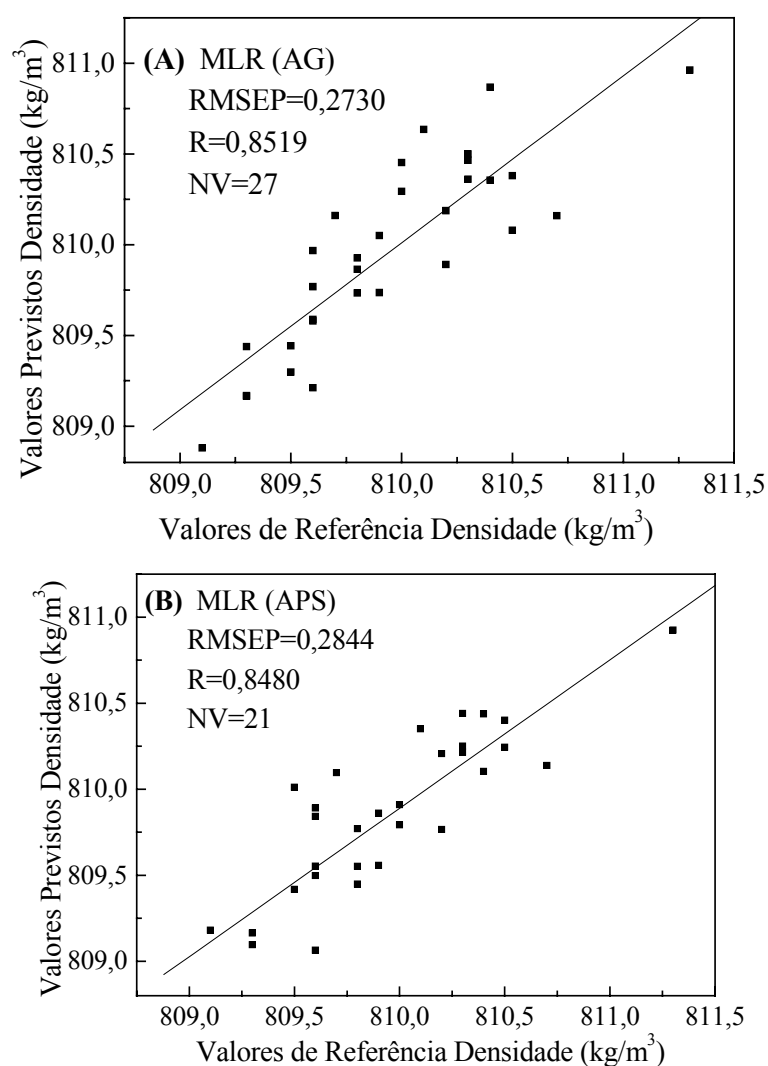


Figura 7.1.67: Comparação entre valores de referência e valores previstos para por modelos MLR para o parâmetro PI da gasolina. (A) Variáveis selecionadas por AG e variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

Os resultados fornecidos pelos modelos construídos através dos métodos MLR(AG) e MLR(APS) mostraram-se muito próximos, com capacidade preditiva semelhante, adicionalmente os número de variáveis selecionadas pelos dois método foram muito próximos.

A tabela 7.1.15 resume os resultados obtidos para a previsão da densidade do álcool automotivo:

Tabela 7.1.15: Resultados para a previsão da densidade do álcool automotivo através de modelos de Calibração Multivariada

Densidade (álcool)	Modelo	VL	RMSEP	R	NV
- Faixa : 808,6 a 811,4 kg/m ³ - Reprodutibilidade do Método Padrão: 1%	PLS (linha base)	7	0,1746	0,9343	1138
	PLS (1ª Derivada)	7	0,2871	0,8372	1128
	PLS (AG)	8	0,1836	0,9273	27
	MLR (AG)	-	0,2730	0,8519	27
	MLR (APS)	-	0,2844	0,841	21

VL: variáveis latentes; NV: número de variáveis independentes; R: coeficiente de correlação.

Os valores obtidos neste trabalho estiveram abaixo da reprodutibilidade do método padrão, que é de 1%, pouco mais de 8 kg/m³. Adicionalmente, estes valores são compatíveis aos encontrados na literatura (0,19 a 0,25 kg/m³), obtidos por Oliveira e colaboradores ^[89] com espectros adquiridos no infravermelho próximo.

- Teor Alcoólico

A faixa de valores estudada para a determinação do teor alcoólico foi de 92,4 a 93,4%.

Os modelos construídos através dos métodos PLS e MLR utilizaram 29 amostras de calibração, 20 amostras de teste e 26 amostras de previsão.

Os números de variáveis latentes selecionadas para a construção dos modelos PLS estão representados na figura 7.1.68 que mostra o gráfico de RMSEV em função do número de variáveis latentes:

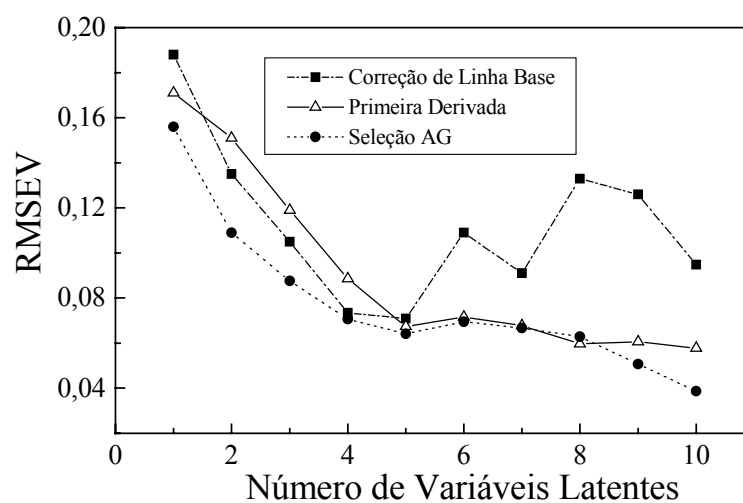


Figura 7.1.68: RMSEV em função dos números de variáveis latentes para os modelos de previsão PLS para teor alcoólico do álcool com correção de linha base, primeira derivada e com variáveis selecionadas por Algoritmo Genético.

O modelo PLS com correção de linha base foi descrito com 4 variáveis latentes, enquanto que os modelos PLS com primeira derivada e PLS(AG) foram descritos com 5 variáveis.

O resultado da previsão dos modelos PLS para a determinação do teor alcoólico é mostrado na figura 7.1.69:

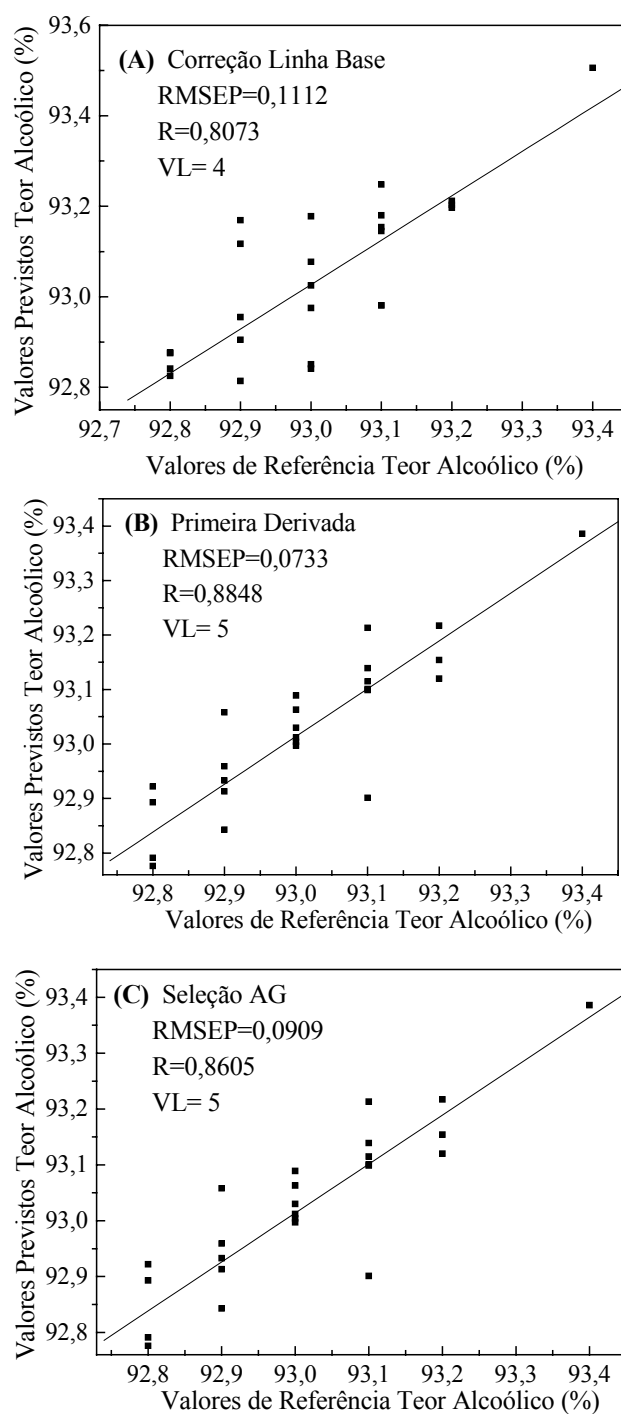
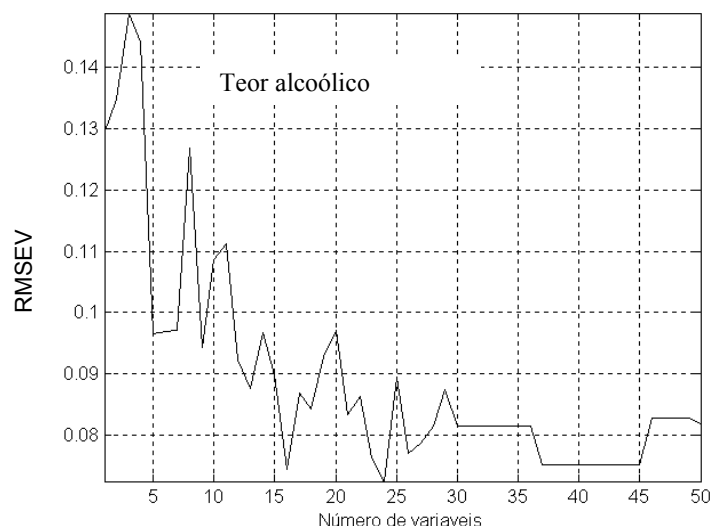


Figura 7.1.69: Comparação entre valores de referência e valores previstos por modelos PLS o parâmetro teor alcoólico. (A) Correção de linha Base, (B) 1ª Derivada e (C) Seleção Algoritmo Genético. VL é o número de variáveis latentes; R é o coeficiente de correlação.

O menor valor de RMSEP obtido corresponde a 0,0733%, fornecido pelo modelo PLS com primeira derivada. O modelo PLS (AG) forneceu erro muito próximo a este resultado (0,0909%) com o uso de 27 números de onda selecionados.

O método APS selecionou um número menor de variáveis (24), cujo gráfico de RMSEV em função do número de variáveis selecionadas é mostrado na figura 7.1.70:



7.1.70: Resultado da seleção APS. Foram selecionadas 24 variáveis independentes através do critério de menor RMSEV.

As variáveis selecionadas pelos métodos AG e APS foram empregadas nos modelos MLR, cujos resultados da previsão são mostrados na figura 7.1.71:

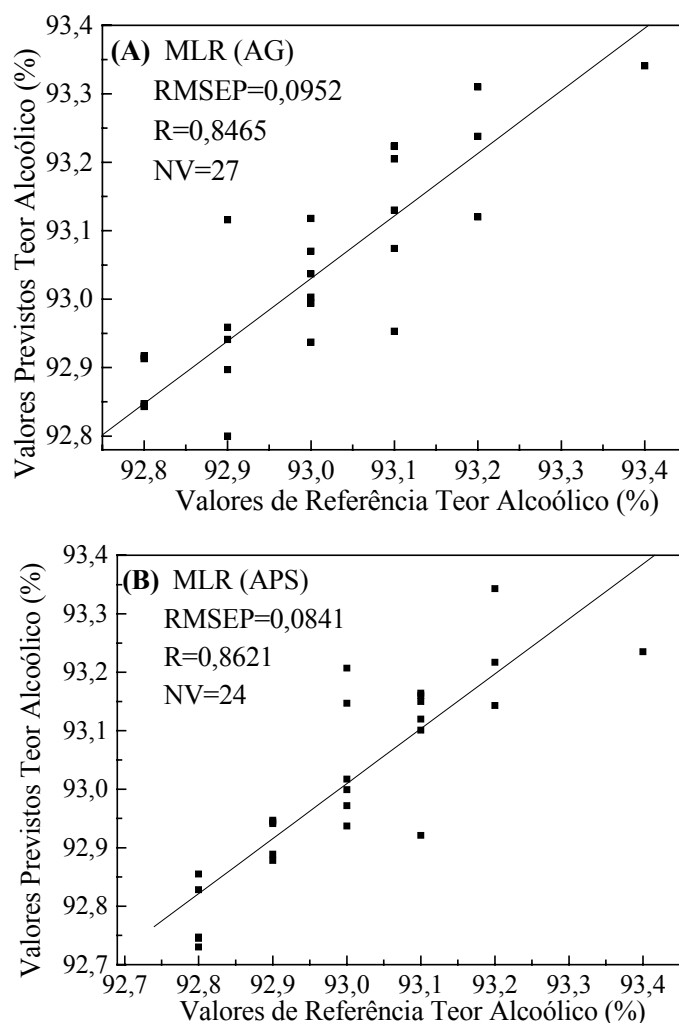


Figura 7.1.71: Comparação entre valores de referência e valores previstos para por modelos MLR para o teor alcoólico. (A) Variáveis selecionadas por AG e (B) variáveis selecionadas por APS. “NV” indica o número de variáveis selecionadas.

O modelo MLR(APS) produziu menor valor de RMSEP para a previsão do teor alcoólico, quando comparado ao valor obtido por MLR(AG), embora muito próximos.

A tabela 7.1.16 reúne os principais resultados obtidos para a previsão do teor alcoólico de amostras de álcool automotivo.

Tabela 7.1.16: Resultados para a previsão o teor alcoólico do álcool automotivo através de modelos de Calibração Multivariada.

Teor alcoólico (álcool) - Faixa : 92,4-93,4 - Reprodutibilidade do Método Padrão: 1%	Modelo	VL	RMSEP	R	NV
	PLS (linha base)	4	0,1112	0,8073	1138
	PLS (1ª Derivada)	5	0,0733	0,8848	1128
	PLS (AG)	5	0,0909	0,8605	27
	MLR (AG)	-	0,0925	0,8465	27
	MLR (APS)	-	0,0841	0,8621	24

VL: variáveis latentes; NV: número de variáveis independentes; R: coeficiente de correlação.

O menor valor de RMSEP obtido corresponde a 0,0733% e foi obtido através do modelo PLS com primeira derivada. Os valores de RMSEP obtidos por modelos MLR mostraram-se próximos a este valor.

O valor de RMSEP obtido por Oliveira *et al*^[89] encontra-se na faixa de 0,0399 a 0,0673% para dados referentes ao infravermelho próximo.

Todos os valores obtidos para a previsão de teor alcoólico do álcool automotivo, encontram-se abaixo da reprodutibilidade do método padrão, cerca de 0,9%.

7.1.4- Coeficientes Angulares e Lineares para os resultados de Previsão dos Combustíveis

Uma boa indicação da concordância entre os valores previstos e os valores medidos pelos métodos de referência pode ser indicada pelas características dos coeficientes da reta do gráfico no qual estão dispostos estes valores.

O coeficiente angular (α) expressa a inclinação da reta deve tender a “um”. O coeficiente linear é o ponto de intersecção no eixo Y e deve tender a zero, deste modo, a reta passa pela origem dos valores, em uma situação ideal.

Deste modo, foram analisados os coeficientes angulares e lineares das retas obtidas pelos gráficos dos valores de referência em função dos valores previstos pelos modelos PLS e MLR propostos neste trabalho.

As tabelas 7.1.17, 7.1.18 e 7.1.19 mostram, respectivamente, os valores de coeficientes lineares (intercepto) e angulares (inclinação) para os resultados obtidos na previsão dos parâmetros do diesel, álcool automotivo e gasolina:

Tabela 7.1.17: Coeficientes (linear e angular) estimados para parâmetros do diesel :

DIESEL											
Parâmetro	Faixa (conc.)	PLS (cor.linha base)		PLS (1ª Derivada)		PLS (AG)		MLR (APS)		MLR(AG)	
		Coef.linear	α	Coef.linear	α	Coef.linear	α	Coef.linear	α	Coef.linear	α
<i>Densidade</i>	0,8489-0,8626 g/cm ³	0,0012	0,9985	0,0331	0,9611	0,0262	0,9692	0,0250	0,9708	-0,0310	1,0371
<i>T50%</i>	264,7-295,9 °C	8,3334	0,9690	-0,8991	1,0022	-26,128	1,0921	18,002	0,9585	-2,3490	1,0070
<i>T85%</i>	331,9-359,9 °C	73,299	0,7908	53,782	0,8463	68,115	0,8061	15,710	0,9565	5,9367	0,9852
<i>I.Cetano</i>	39,9- 49,3	5,0811	0,8852	7,6982	0,8243	2,4888	0,9443	-0,1427	1,0040	-1,0434	1,0257
<i>Enxofre</i>	0,10-0,34 %	0,0648	0,7760	0,0769	0,7422	0,1098	0,6247	0,0390	0,8735	0,0745	0,7532
<i>Viscosidade</i>	2,797-4,123 mm ² /s	0,4357	0,8944	-0,056	1,026	0,3596	0,9216	-0,0802	1,024	-0,2058	1,0567

α é a inclinação da reta (coeficiente angular) e coef. linear é o ponto de interseção no eixo Y

Tabela 7.1.18: Coeficientes (linear e angular) estimados para parâmetros do álcool automotivo:

ÁLCOOL AUTOMOTIVO											
Parâmetro	Faixa (conc.)	PLS (cor.linha base)		PLS (1ª Derivada)		PLS (AG)		MLR (APS)		MLR(AG)	
		Coef.linear	α	Coef.linear	α	Coef.linear	α	Coef.linear	α	Coef.linear	α
<i>Massa Específica</i>	808,6-811,4 kg/m ³	76,274	0,9057	-6,4768	1,008	70,545	0,9129	111,23	0,8625	64,658	0,9201
<i>Teor alcoólico</i>	92,4-93,4 %	1,8218	0,9806	11,4765	0,8767	-14,620	1,0188	5,6234	0,9396	7,9969	0,9143

α é a inclinação da reta (coeficiente angular) e coef. linear é o ponto de interseção no eixo Y

Tabela 7.1.19: Coeficientes (linear e angular) estimados para parâmetros da gasolina :

GASOLINA											
Parâmetro	Faixa (conc.)	PLS (cor.linha base)		PLS (1ª Derivada)		PLS (AG)		MLR (APS)		MLR(AG)	
		Coef.linear	α	Coef.linear	α	Coef.linear	α	Coef.linear	α	Coef.linear	α
<i>Densidade</i>	0,7485-0,7850 g/m ³	0,0953	0,8730	0,1034	0,9622	0,0529	0,9293	0,0927	0,8765	0,020	0,9728
<i>T10%</i>	48,80-69,70 °C	3,9909	0,9008	4,9080	0,8769	4,6442	0,9108	1,7430	0,9633	3,3937	0,9328
<i>T90%</i>	143,90-215,50 °C	45,5982	0,7278	41,6231	0,7506	42,25	0,7460	37,370	0,7765	25,96	0,8413
<i>PFE</i>	202,20-248,70 °C	75,5982	0,6401	66,005	0,6864	18,590	0,6005	88,020	0,5804	48,680	0,7740
<i>MON</i>	77,90-82,80	17,954	0,7805	16,250	0,8015	18,590	0,7720	14,540	0,8224	11,149	0,8639
<i>RON</i>	89,00-97,50	-1,8393	1,0162	-0,3280	1,0056	0,3163	0,9939	-0,0108	0,9974	2,028	0,9754
<i>IAD</i>	83,20-90,1	8,3961	0,9024	8,4701	0,9014	10,704	0,8761	7,8752	0,9082	12,638	0,8534
<i>PI</i>	34,20-61,70 min	7,2684	0,8527	8,0238	0,8356	12,399	0,7460	8,5310	0,8219	7,4978	0,8300
<i>Benzeno</i>	0,10-1,20 %	0,0575	0,8878	0,05357	0,8951	0,04857	0,8892	0,0348	0,9272	0,0477	0,8878
<i>Aromáticos</i>	8,0- 21,0 %	5,0380	0,7192	4,6225	0,7403	4,4145	0,7531	2,8286	0,8518	2,2687	0,8747
<i>Saturados</i>	32,0-60,0 %	3,9909	0,9008	4,9080	0,8769	60,4869	0,8458	4,5667	0,8927	5,3019	0,8838

α é a inclinação da reta (coeficiente angular) e coef. linear é o ponto de interseção no eixo Y

Algumas propriedades como o PFE apresentaram valores muito ruins de coeficientes angulares e lineares. No entanto este fato está relacionado com as próprias limitações das técnicas de referência.

De um modo geral, foi observado que os modelos construídos com o método MLR apresentaram melhores coeficientes linear e angular quando comparados aos resultados obtidos a partir dos modelos PLS para os parâmetros de T85%, índice de cetano e teor de enxofre para o diesel; densidade, T10%, T90%, PFE, MON, teores de benzeno, aromáticos e saturados para a gasolina e massa específica para o álcool automotivo. Os parâmetros apresentaram coeficientes lineares e angulares condizentes com os obtidos através dos modelos PLS.

Este fato demonstra que a maioria dos modelos MLR estiveram menos susceptíveis a erros sistemáticos que os modelos PLS e portanto, podem produzir resultados mais confiáveis.

7.1.5- Comparação entre os modelos de calibração através do Teste-F

O principal critério utilizado neste trabalho para determinar qual modelo apresenta melhor desempenho na previsão da propriedade de interesse está relacionado com o menor valor de RMSEP. No entanto, os valores de RMSEP obtidos na previsão das propriedades dos combustíveis apresentaram-se muito próximos, então uma comparação mais rigorosa entre os modelos de calibração propostos tornou-se necessária.

Para efetuar tal comparação, foi utilizado o teste-F^[124]. Este procedimento visa averiguar se as diferenças obtidas entre os resultados obtidos para o conjunto de previsão dos diversos modelos são estatisticamente significativas. O método compara dois valores de variância e é capaz de detectar diferenças sistemáticas em uma estimativa conjunta.

O teste-F é realizado segundo a equação 7.1, na qual p_i e p_j representam os graus de liberdade (números de amostras do conjunto de previsão menos um):

$$F(p_i, p_j) = \left(\frac{RMSEP_i}{RMSEP_j} \right)^2 \quad (7.1)$$

Os valores de $F(i,j)$ são comparados com um valor de F crítico tabelado com um nível de confiança de 95%. Quando estes valores forem maiores que F crítico, os modelos comparados são considerados estatisticamente diferentes.

7.1.5-1 Comparação entre modelos de calibração das propriedades do diesel:

O teste-F foi aplicado neste trabalho para detectar diferenças significativas entre os modelos propostos para previsão das propriedades do diesel, gasolina e álcool automotivo.

O teste-F foi aplicado da seguinte maneira:

- na comparação entre os modelos PLS;
- na comparação entre os modelos MLR(AG) e MLR (APS);
- na comparação entre o menor RMSEP dos modelos PLS e os RMSEP fornecidos pelos modelos MLR(AG) e MLR(APS)

A tabela 7.1.20 mostra um sumário dos valores de RMSEP obtidos para o conjunto de previsão das propriedades do diesel. Também estão listados o número de amostras empregado nas séries de previsão de cada propriedade:

Tabela 7.1.20: Valores de RMSEP obtidos pelos métodos de calibração PLS e MLR para o conjunto de previsão das propriedades do diesel:

Parâmetro	VALORES DE RMSEP (DIESEL)					
	* N° de Amostras	PLS (cor.linha base)	PLS (1ª Derivada)	PLS (AG)	MLR(AG)	MLR (APS)
<i>Densidade</i>	53	0,0007 (9)	0,0007 (6)	0,0006 (5)	0,0006 [31]	0,0007[43]
<i>T50%</i>	34	1,6148 (10)	1,5630 (6)	2,3556 (6)	1,7224 [18]	1,4971 [20]
<i>T85%</i>	31	4,4820 (9)	3,8669 (5)	3,6334 (6)	3,4567 [27]	3,9329 [17]
<i>I.Cetano</i>	34	0,4576 (4)	0,4748 (3)	0,4935 (5)	0,4250 [26]	0,3414 [9]
<i>Enxofre</i>	35	0,0130 (7)	0,0134 (4)	0,0156 (3)	0,0190 [23]	0,0150 [9]
<i>Viscosidade</i>	27	0,1239 (7)	0,1099 (6)	0,1393 (7)	0,1238 [28]	0,1206 [20]

() Variáveis Latentes; [] números de onda selecionados e * número de amostras de previsão

Com exceção do teor de enxofre e da viscosidade, os menores valores de RMSEP foram obtidos através dos modelos MLR com seleção de variáveis. No entanto, o Teste-F deve afirmar quais resultados são estatisticamente deferentes.

Através dos valores de RMSEP aplicados na equação 7.1 e os respectivos graus de liberdade, foram determinados valores de $F(p_i, p_j)$ que foram comparados aos valores críticos (F_c), obtidos em uma tabela de distribuição F com um nível de 95% de confiança.

A tabela 7.1.21 mostra os resultado da aplicação do Teste-F para a comparação entre os modelos construídos através do método PLS: com correção de linha base, primeira derivada e modelo construído com variáveis selecionadas pelo método AG. Os valores acima de F_c de cada propriedade foram assinalados em **negrito**:

Tabela 7.1.21: Teste-F para modelos PLS para a previsão das propriedades do diesel:

TESTE-F (modelos PLS)			
Teste F	Densidade ($F_c = 1,69$)		
PLS (linha base)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (1ª derivada)	-	1	1,36
	1	-	1,36
T50% ($F_c = 1,84$)			
PLS (linha base)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (1ª derivada)	-	1,07	2,12
	1,07	-	2,27
T85% ($F_c = 1,84$)			
PLS (linha base)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (1ª derivada)	-	1,34	1,52
	1,34	-	1,13
Índice de Cetano ($F_c = 1,84$)			
PLS (linha base)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (1ª derivada)	-	1,07	1,16
	1,07	-	1,08
Teor de Enxofre ($F_c = 1,84$)			
PLS (linha base)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (1ª derivada)	-	1,06	1,20
	1,06	-	1,35
Viscosidade ($F_c = 1,90$)			
PLS (linha base)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (1ª derivada)	-	1,27	1,26
	1,27	-	1,61

De acordo com os resultados obtidos, apenas a comparação entre o RMSEP obtido através do modelo PLS(AG) e os modelos PLS(linha base) e PLS (1ª derivada) em relação à propriedade T50% apresentaram um valor de F superior ao F crítico, sendo considerados significativamente diferentes. O valor de RMSEP obtido por PLS (AG) é maior que os demais.

A tabela 7.1.22 mostra a aplicação do teste-F para a comparação entre os modelos construídos através do método MLR com variáveis selecionadas pelo método AG e pelo método APS:

Tabela 7.1.22: Aplicação do Teste-F entre os modelos MLR(AG) e MLR(APS).

TESTE-F (MODELOS MLR)						
Teste F	Densid. (Fc= 1,69)	T 50% (Fc=1,84)	T85% (Fc = 1,84)	I. Cetano (Fc = 1,84)	Enxofre (Fc = 1,84)	Viscos. (Fc=1,90)
MLR(APS) e MLR(AG)	1,36	1,32	1,29	1,56	1,60	1,05

Para a maioria das propriedades (T50%, índice de cetano, teor de enxofre e viscosidade), os modelos MLR(APS) forneceram menores valores de RMSEP quando comparados aos modelos MLR(AG). Além disso, o método APS selecionou números menores de variáveis independentes. No entanto, baseado nos resultados fornecidos pelo teste-F, é possível concluir que não há diferenças significativas entre os modelos MLR propostos, uma vez que nenhum valor de $F(p_i, p_j)$ calculado esteve acima de F crítico.

Os modelos MLR foram comparados com o modelo PLS que apresentou menor valor de RMSEP para cada propriedade do diesel. O resultado da aplicação do Teste-F para a comparação entre o modelo PLS de menor RMSEP e os modelos MLR(AG) e MLR(APS) é mostrado na tabela 7.1.23.

Tabela 7.1.23: Aplicação do Teste-F entre o modelo PLS de menor RMSEP e modelos MLR(AG) e MLR(APS).

Teste F	Densid. (Fc= 1,69)	T 50% (Fc=1,84)	T85% (Fc = 1,84)	I. Cetano (Fc = 1,84)	Enxofre (Fc =1, 84)	Viscosid. (Fc=1,90)
	PLS (AG)	PLS (1ª Der.)	PLS (AG)	PLS (lin. base)	PLS (lin.base)	PLS (1ª Der.)
MLR (APS)	1,36	1,09	1,17	1,80	1,33	1,20
MLR (AG)	1	1,21	1,10	1,16	2,13	1,27

Observou-se que apenas os resultados obtidos pelos modelos PLS e MLR (AG) para a previsão do enxofre foram considerados significativamente diferentes. O valor de RMSEP fornecido pelo modelo MLR (AG) foi superior aos demais.

7.1.5-2 Comparação entre modelos de calibração das propriedades da gasolina:

A tabela 7.1.24 mostra todos os valores de RMSEP obtidos pelos modelos propostos para a previsão das propriedades da gasolina e os números de amostras de previsão.

Tabela 7.1.24: Valores de RMSEP obtidos pelos métodos de calibração PLS e MLR para o conjunto de previsão das propriedades do diesel:

Parâmetro	VALORES DE RMSEP (GASOLINA)					
	* N° de Amostras	PLS (cor.linha base)	PLS (1ª Derivada)	PLS (AG)	MLR (AG)	MLR (APS)
<i>Densidade</i>	83	0,0014(6)	0,0013(6)	0,0012(7)	0,0011 [39]	0,0014 [19]
<i>T10%</i>	83	1,4919 (3)	1,4377(3)	1,4193(5)	1,3766 [32]	1,3352 [27]
<i>T90%</i>	69	4,3519(6)	3,9287 (8)	3,7195(8)	3,8442 [37]	3,8057 [32]
<i>PFE</i>	45	2,6389 (3)	2,5369 (3)	2,5851 (5)	2,9636 [33]	2,9263 [8]
<i>MON</i>	43	0,3805 (2)	0,3656 (2)	0,3997 (4)	0,3170 [27]	0,4048 [4]
<i>RON</i>	53	0,6903 (2)	0,6526 (2)	0,7545 (2)	0,8370 [18]	0,7036[4]
<i>IAD</i>	61	0,5823 (2)	0,5849 (2)	0,6272 (5)	0,6587 [22]	0,6206 [4]
<i>PI</i>	56	1,9826 (2)	2,2158 (2)	2,5419 (1)	2,8763 [35]	2,6748 [32]
<i>Benzeno</i>	34	0,0510 (4)	0,0522 (4)	0,0471 (3)	0,0539 [18]	0,0489 [15]
<i>Aromáticos</i>	42	1,2052 (2)	1,1902 (2)	1,1293 (3)	1,1044 [23]	1,2218 [24]
<i>Saturados</i>	37	2,2820 (1)	2,2733 (2)	2,6365 (3)	2,5756 [23]	2,4539 [3]

() Variáveis Latentes; [] números de onda selecionados e * número de amostras de previsão

Os menores valores de RMSEP para as propriedades T90%, PFE, RON, IAD, PI, teores de benzeno e saturados foram obtidos através do método PLS. Para a densidade, T50%, MON e teor de aromáticos, os menores valores de RMSEP foram obtidos através do método MLR com seleção de variáveis. No entanto, a previsão de várias propriedades apresentaram resultados muito próximos.

Inicialmente, foram comparados os valores de RMSEP obtidos entre os modelos PLS com correção de linha base, com aplicação da primeira derivada e com seleção AG, cujo resultado da aplicação do Teste-F é mostrado na tabela 7.1.25:

Tabela 7.1.25: Teste-F para modelos PLS para a previsão das propriedades da gasolina:

TESTE-F (modelos PLS)			
Dens. (Fc=1,53)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,15	1,36
PLS (1ª derivada)	1,15	-	1,17
T10% (Fc=1,53)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,08	1,10
PLS (1ª derivada)	1,08	-	1,03
T90% (Fc= 1,53)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,23	1,37
PLS (1ª derivada)	1,23	-	1,12
PFE (Fc=1,69)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,08	1,04
PLS (1ª derivada)	1,08	-	1,03
MON (Fc=1,69)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,08	1,26
PLS (1ª derivada)	1,08	-	1,20
RON (Fc= 1,53)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,12	1,20
PLS (1ª derivada)	1,12	-	1,37
IAD (Fc=1,53)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,01	1,16
PLS (1ª derivada)	1,01	-	1,15
PI (Fc=1,53)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,25	1,64
PLS (1ª derivada)	1,25	-	1,31
benzeno (Fc=1,84)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,05	1,17
PLS (1ª derivada)	1,05	-	1,23
Arom. (Fc=1,69)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,03	1,14
PLS (1ª derivada)	1,03	-	1,11
saturados (Fc=1,69)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	1,01	1,33
PLS (1ª derivada)	1,01	-	1,34

Entre os modelos PLS, apenas foram observadas diferenças significativas em relação aos modelos PLS (linha base) e PLS (AG) na previsão da propriedade

período de indução (PI).O modelo PLS (AG) forneceu um valor de RMSEP superior aos demais modelos PLS:

Tabela 7.1.26: Aplicação do Teste-F entre o modelo PLS de menor RMSEP e modelos MLR(AG) e MLR(APS) para gasolina.

TESTE-F (MODELOS MLR)											
	Dens.	T10%	T90%	PFE	MON	RON	IAD	PI	Benz	Aro.	Sat.
F crítico	1,53	1,53	1,53	1,69	1,69	1,53	1,53	1,53	1,84	1,69	1,69
MLR (APS) e MLR (AG)	1,19	1,06	1,02	1,02	1,63	1,41	1,13	1,16	1,21	1,22	1,10

Nenhuma diferença significativa foi observada entre os modelos MLR construídos para determinar as propriedades da gasolina. No entanto os modelos MLR (APS) forneceram menores (mas não significativos) valores de RMSEP para a previsão da maioria das propriedades, quando comparados aos modelos MLR(AG). Além disso, o método APS selecionou um menor número de variáveis independentes para a maioria das propriedades da gasolina.

O melhor resultado dentre os modelos construídos a partir do método PLS (modelo que forneceu menor RMSEP) foi comparado, através do teste-F, com os modelos obtidos através do método MLR, os resultados são mostrados na tabela 7.1.27:

Tabela 7.1.27: Aplicação do Teste-F entre o modelo PLS de menor RMSEP e modelos MLR(AG) e MLR(APS) para a gasolina.

TESTE-F (modelos MLR e PLS)		
Dens. (Fc=1,53)	MLR (AG)	MLR (APS)
PLS (AG)	1,19	1,36
T10% (Fc= 1,53)	MLR (AG)	MLR (APS)
PLS (AG)	1,06	1,13
T90% (Fc=1,53)	MLR (AG)	MLR (APS)
PLS (AG)	1,07	1,05
PFE (Fc=1,69)	MLR (AG)	MLR (APS)
PLS (1ª Derivada)	1,36	1,33
MON (Fc=1,69)	MLR (AG)	MLR (APS)
PLS (1ª Derivada)	1,33	1,22
RON (Fc=1,53)	MLR (AG)	MLR (APS)
PLS (linha base)	1,64	1,16
IAD (Fc=1,53)	MLR (AG)	MLR (APS)
PLS (linha base)	1,28	1,13
PI (Fc=1,53)	MLR (AG)	MLR (APS)
PLS (linha base)	2,10	1,82
benzeno (Fc= 1,84)	MLR (AG)	MLR (APS)
PLS (linha base)	1,12	1,09
Arom. (Fc=1,69)	MLR (AG)	MLR (APS)
PLS (AG)	1,05	1,17
saturados (Fc=1,69)	MLR (AG)	MLR (APS)
PLS (1ª Derivada)	1,28	1,17

De acordo com os resultados da aplicação do teste-F, somente os modelos construídos para a previsão de RON e PI mostraram diferenças significativas. O modelo MLR(AG) construído para determinar RON apresentou um valor de RMSEP superior aos demais. Em relação à previsão de PI, os modelos MLR(AG) e

MLR(APS) apresentaram valores superiores de RMSEP. No entanto para a maioria das propriedades, os métodos MLR com seleção de variáveis não apresentou diferenças significativas, com vantagem se tratar de um modelo matemático mais simples.

7.1.5-3 Comparação entre modelos de calibração das propriedades do álcool automotivo:

A tabela 7.1.28 mostra valores de RMSEP obtidos através dos modelos PLS e MLR para a previsão as propriedades do álcool automotivo.

Tabela 7.1.28: Valores de RMSEP obtidos pelos métodos de calibração PLS e MLR para o conjunto de previsão das propriedades do álcool automotivo:

Parâmetro	VALORES DE RMSEP (ÁLCOOL AUTOMOTIVO)					
	* N° de Amostras	PLS (cor.linha base)	PLS (1ª Derivada)	PLS (AG)	MLR(AG)	MLR (APS)
<i>Densidade</i>	31	0,1746 (7)	0,2871(7)	0,1836(8)	0,2730 [27]	0,2844 [21]
<i>Teor Alcoólico</i>	26	0,1112 (4)	0,0733(5)	0,0909 (5)	0,0925 [27]	0,0841 [24]

() Variáveis Latentes; [] números de onda selecionados e * número de amostras de previsão

Os modelos PLS apresentaram valores de RMSEP com diferenças um pouco maiores, quando comparados aos resultados obtidos para os outros combustíveis. Através da aplicação do teste-F foi averiguado a significância destas diferenças, cujo resultado é mostrado na tabela 7.1.29:

Tabela 7.1.29: Teste-F para modelos PLS para a previsão das propriedades do álcool automotivo:

TESTE-F (modelos PLS)			
Densidade (Fc=1,84)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	2,70	1,10
PLS (1ª derivada)	2,70	-	2,44
Teor Alcoólico (Fc=1,96)	PLS (linha base)	PLS (1ª derivada)	PLS (seleção AG)
PLS (linha base)	-	2,30	1,50
PLS (1ª derivada)	2,30	-	1,54

Em relação à propriedade de densidade, foi observado que há diferenças significativas entre os modelos PLS (linha base) e PLS (1ª Derivada) e entre os modelos PLS (1ª derivada) e PLS (AG). O modelo PLS (1ª derivada) apresenta um valor de RMSEP (0,2871) superior aos demais valores de RMSEP fornecidos pelos outros modelos PLS.

A tabela 7.1.30 mostra os resultados da aplicação do teste-F na comparação entre os modelos MLR com seleção de variáveis construídos para a previsão das propriedades do álcool automotivo.

Tabela 7.130: Aplicação do Teste-F entre os modelos MLR(AG) e MLR(APS) para o álcool.

TESTE-F (modelos MLR)		
	Densidade (Fc= 1,84)	Teor Alcoólico (Fc=1,96)
MLR (APS) e MLR (AG)	1,08	1,21

Os valores de F obtidos na comparação entre os modelos MLR (AG) e APS não foram superiores ao F crítico tabelado, portando pode-se dizer que não foram observadas diferenças significativas entre ambos.

A tabela 7.1.31 mostra os resultados da aplicação do Teste-F entre o modelos PLS que forneceu menor valor de RMSEP e os modelos MLR(APS) e MLR(AG);

Tabela 7.1.31: Aplicação do Teste-F entre o modelo PLS de menor RMSEP e modelos MLR(AG) e MLR(APS).

TESTE-F (modelos MLR e PLS)		
Densidade (Fc=1,84)	MLR (AG)	MLR (APS)
PLS (linha base)	2,44	2,65
Teor Alcoólico (Fc= 1,96)	MLR (AG)	MLR (APS)
PLS (1ª derivada)	1,59	1,32

Foram observadas diferenças significativas para os resultados obtidos na previsão da densidade do álcool automotivo. Neste caso, os modelos MLR(AG) e MLR(APS) forneceram valores de RMSEP superiores aos modelos PLS.

Para os modelos construídos para a previsão do teor alcoólico não apresentaram diferenças significativas para ambos os métodos (PLS e MLR).

7.1.6- Atribuição dos Números de Onda Seleccionados pelos Métodos AG e APS

Em vista dos vários tipos de hidrocarbonetos, impurezas e aditivos presentes nos combustíveis, a atribuição de todos os números de onda seleccionados não é um trabalho muito simples. Além das bandas fundamentais os espectros apresentam bandas de harmônicas, combinações e sobreposições de vários modos vibracionais.

Felizmente, vários números de onda seleccionados correspondem às principais absorções de hidrocarbonetos.

A figura 7.1.72 mostra um esquema com as principais bandas relacionadas com os números de onda escolhidos pelos métodos AG e APS^[125-126].

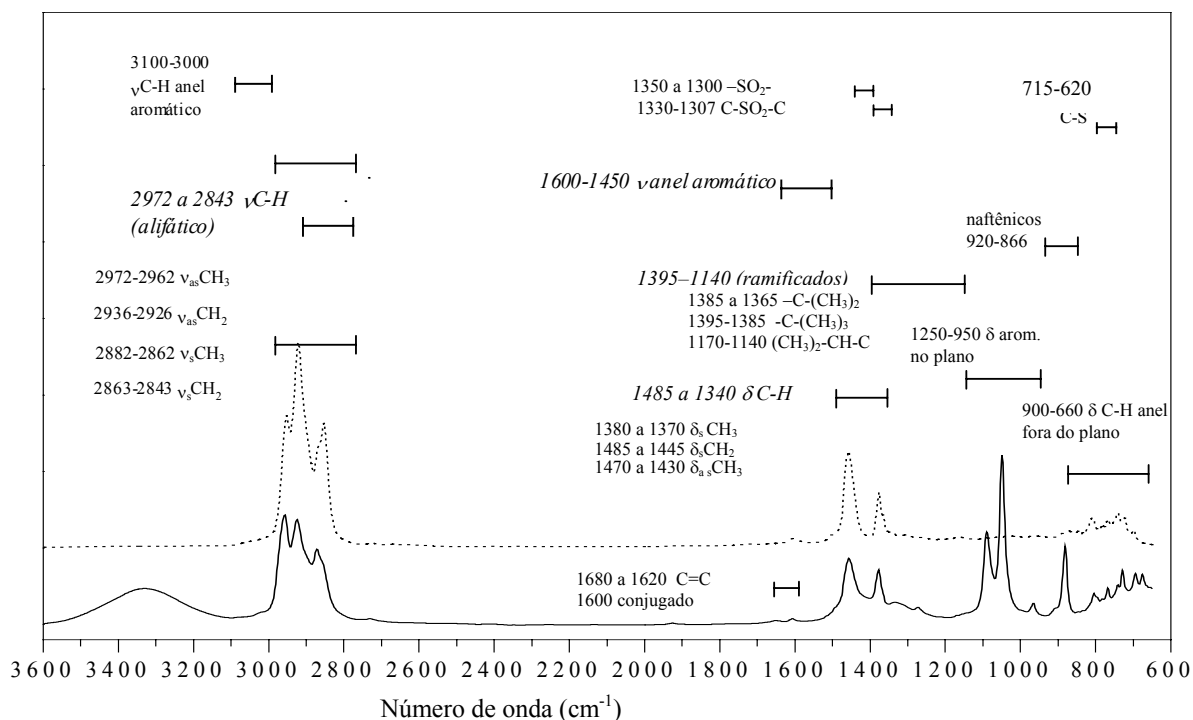


Fig.7.1.72: Esquema de atribuição de bandas no infravermelho médio. (----) espectro de diesel e (—) espectro gasolina. O símbolo ν corresponde à deformação no eixo da ligação (estiramento) e δ à deformação angular.

7.1.5-1 Números de onda selecionados para as propriedades do diesel

Em relação aos parâmetros do **diesel**, com exceção da densidade e T50%, o método AG selecionou um maior número de ondas quando comparado ao método APS. No entanto, os dois métodos selecionaram números de onda muito próximos ou semelhantes, como pode ser observado na tabela 7.1.32:

Tabela 7.1.32: Variáveis (números de onda) selecionados para a o diesel:

Parâmetro (DIESEL)	APS		AG	
	n ^o	Variáveis Selecionadas (cm ⁻¹)	n ^o	Variáveis Selecionadas (cm ⁻¹)
<i>Densidade</i>	43	660 667 682,7 694 705,9 719 734,8 746,6 756 790 796 804,2 889,4 1059 1132 1193 1207 1367 1440 1448 1460 1629 1656 1701 1714 1763 2831 2848 2858 2869 2877 2896 2906 2912 2918 2922 2929 2931 2935 2945 2952 2966 2976	31	744,4 746,4 756 763,7 790 821,6 833,2 858 862 873,6 945 991 1049 1078 1236 1348 1353 1357 1377 1464 1531 1589 1633 1743 2808 2812 2837 2949 2976 3107 3122
<i>T50%</i>	20	659,6 688,5 734,8 742,5 752,1 773,3 802,3 1352 1386 1438 1448 2844 2854 2906 2912 2924 2931 2941 2952 2964	18	690,4 752,1 788,8 804,2 810 873,6 877,5 883 891 1303 1389 1392 1404 1431 1531 2856 2947 2977
<i>T85%</i>	17	659,6 692,3 734,8 756,0 790,7 1462 1450 1560 2845 2855 2960 2908 2926 2943 2933 2953 2964	27	676,9 684,6 719,3 730,9 784,9 831,2 908,4 921,9 1371 1412 1433 1456 1468 1487 1496 1518 1763 2773 2779 2812 2920 2940 2950 2962 2964 2982 2995
<i>I. Cetano</i>	9	660 758 817,7 2746 2914 2926 2937 2951 2962	26	669,2 721,3 756 783 817,7 819,6 829,3 864 9 18 919,9 935,4 1298 1315 1394 1411 1417 1450 1531 1549 1562 2785 2947 2957 2998 2001
<i>Enxofre</i>	9	660 694 802 1467 2746 2920 2931 2952 2964	23	673,1 711,6 740,6 752,1 756 810 860 903 1313 1355 1431 1463 2766 2783 2804 2837 2847 2912 2914 2922 3003 3022 3024
<i>Viscosidade</i>	20	660 687 719 734 744,5 758 790 1446 1468 1500 2846 2856 2906 2912 2916 2926 2937 2945 2960 2970	28	660 661 706 758 792,6 804,2 806,2 825,4 854,9 921,9 960 1299 1325 1359 1429 1450 1481 1579 1607 2879 2883 2927 2931 2945 2962 3056 3099 3110

De um modo geral, os números de onda selecionados estão relacionados aos principais modos vibracionais dos hidrocarbonetos. A faixa de 3000 - 3100 cm^{-1} é referente ao estiramento axial de C-H de aromáticos ou oleofinas.

A faixa de 2976 a 2906 cm^{-1} representa os estiramentos assimétricos dos grupamentos CH_3 e CH_2 e 2896 a 2831 cm^{-1} corresponde aos estiramentos simétricos (CH_3 e CH_2). A faixa de 1640 a 1670 cm^{-1} é característica de ligações C=C de oleofinas.

As ligações C=C do esqueleto do anel aromático absorvem na faixa de 1500 a 1600 cm^{-1} .

Os números de onda entre 1460 a 1365 correspondem à deformação angular no plano de CH_3 e CH_2 ; e de 1350 a 1150 cm^{-1} correspondem deformação angular simétrica fora do plano de CH_2 . A região de 1365 a 1385 cm^{-1} é atribuída a estiramentos C-H de hidrocarbonetos ramificados.

A faixa de 1059 a 1078 cm^{-1} refere-se a deformação angular do C-H do anel aromático no plano.

Os números de onda abaixo de 889,4 (890) cm^{-1} podem ser atribuídos às vibrações fora do plano de C-H do anel aromático e 719-722 cm^{-1} pode estar relacionado com a deformação angular assimétrica de CH_2 parafínico, observada com mais intensidade nos hidrocarbonetos lineares com cadeias longas.

Em relação aos números de onda selecionados para a **densidade** do diesel, a faixa de 2976-2831 se refere aos modos vibracionais de estiramento da ligação C-H. Os números de onda entre 1464 a 1348 cm^{-1} (seleção AG) e 1460 a 1367 cm^{-1} (seleção APS) correspondem aos modos de deformação angular de C-H. A faixa selecionada entre 1207 a 1132 cm^{-1} (APS) e 1236 cm^{-1} (AG) são absorções referentes aos hidrocarbonetos ramificados. Entre 1236 a 945 cm^{-1} existem números de onda atribuídos a deformação angular C-H de anel no plano. Abaixo de 900 cm^{-1} para os números de onda selecionados pelos dois métodos, correspondem

às principais bandas de C-H de aromáticos fora do plano, que são em geral muito intensas. Os números de onda selecionados próximos a 720 cm^{-1} - 750 cm^{-1} também podem ser atribuídos à deformação angular $(-\text{CH}_2)_4$.

Para o parâmetro **T50%**, as faixas de 2964 a 2844 cm^{-1} (APS) e 2947 a 2856 cm^{-1} (AG) estão relacionadas aos modos de deformação axial de C-H. Números de onda que correspondem às deformações angulares de C-H de parafinas foram atribuídos as faixas de 1448 a 1352 cm^{-1} (APS) e 1431 a 1389 cm^{-1} . O número de onda de 1531 cm^{-1} (AG), provavelmente corresponde a ruído espectral.

Abaixo de 802 cm^{-1} (APS) e 891 cm^{-1} (AG) foram selecionados números de onda que correspondem às deformações angulares fora do plano de C-H do anel aromático.

Para **T85%**, o método APS selecionou nove variáveis que correspondem ao estiramento C-H de parafinas (2964 - 2845). Estes números de onda são praticamente coincidentes com os números de onda selecionados para T50% com método APS. Os números de onda 1450 e 1462 cm^{-1} , selecionados por APS correspondem aos modos de deformação angular CH_3 e CH_2 , assim como a faixa de 1496 a 1371 cm^{-1} selecionada pelo método AG. O número de 921 cm^{-1} (AG) pode estar relacionado com absorções de naftênicos, porém trata-se de uma absorção muito fraca. Abaixo de 831 cm^{-1} , os números de onda são atribuídos à deformação angular C-H fora do plano de aromáticos.

Para o **índice de cetano**, o método de APS selecionou 9 contra 26 números de onda do método AG. No entanto, foi verificado que algumas variáveis não correspondem aos modos dos principais grupamentos como 935 e 1298 cm^{-1} e a faixa de 1531 a 1562 cm^{-1} .

Entre 2962 a 2919 cm^{-1} (APS) e 2998 a 2947 cm^{-1} (AG) estão relacionadas com estiramentos de CH_3 e CH_2 parafínico, o que é mais concordante com a

propriedade analisada, uma vez que esta é dependente da concentração de cetano e outros hidrocarbonetos de cadeia linear. Os números de onda selecionados por AG abaixo de 900 cm^{-1} (com exceção de 721 cm^{-1}) podem estar relacionados com vibrações de aromáticos, que também não apresentam relação direta com a propriedade medida.

Para o **teor de enxofre**, as variáveis 694 cm^{-1} (APS) e 673 cm^{-1} (AG) estão relacionadas com a vibração C-S que ocorre entre 715 a 620 cm^{-1} . O número de onda de 1313 cm^{-1} (AG) pode ser atribuído como C-SO₂-C. As ligações S-O-C relacionam-se com os números de onda 802 , 810 e 860 cm^{-1} .

Em relação à **viscosidade**, a grande maioria das variáveis selecionadas correspondem aos estiramentos C-H em 2970 a 2846 cm^{-1}), deformações angulares (1450 a 1359 cm^{-1} e 719 cm^{-1}). No entanto os números de onda 1500 cm^{-1} (AG) e 1579 , 921 e 960 cm^{-1} (APS) estão associados a ruídos. Abaixo de 854 cm^{-1} , as variáveis estão atribuídas aos modos de deformação C-H de anel aromático fora do plano.

7.1.5-2 Números de onda selecionados para as propriedades da gasolina

Tabela 7.1.33: Variáveis (números de onda) selecionados para a gasolina.

Parâmetro (Gasolina)	APS		AG	
	n ^o	Variáveis Selecionadas (cm ⁻¹)	n ^o	Variáveis Selecionadas (cm ⁻¹)
<i>Densidade</i>	19	660 673 686,6 703,9 717,5 729 736,7 759,9 802,3 958 1045,3 1055 1273 1473 2924 2935 2957 2964 2979	39	690,4 746,4 750 768 817,7 839 848,6 854,4 910,3 925,7 948,9 1113 1150 1216 1225 1310 1317 1344 1462 1475 1483 1498 1502 1571 1610 1620 1637 1650 1654 1670 2819 2843 2879 2894 2918 2949 2991 3047 3060
<i>T10%</i>	27	660 665,4 673 684,6 703,9 707,7 727,1 736,7 759,9 790,7 800,4 873,7 918 968 1041,4 1047 1054 1273 1448 1458 2912 2921 2931 2943 2954 2964 2975	32	704,7 719,3 823,4 853,4 869,8 877,2 887,2 950,8 968,2 997 1030 1296 1340 1392 1440 1567 1504 1527 1573,8 1602 1645 2815 28534 2860 2901 2909 1912 2955 2947 2960 3066 3105
<i>T90%</i>	32	660 673,2 675 684,6 694 703,9 707,7 729 736,7 759,9 790,7 801 873,7 918 958 1047 1056 1095 1128 1207 1358 1437 1440 1462 1548 2910 2920 2931 2941 2952 2970 2976	37	673,2 671,1 678,9 684,6 690,4 702,0 707,8 769,5 794,6 821,6 820,7 948,9 973,9 985 1054 1080 1178 1181 1209 1236 1243 1307,6 1315,3 1357,7 1392,5 1460 1461 1462 1558 1593 1620 1695 2779 2850 2883 2933 3078
<i>PFE</i>	8	684,6 736 1057 1460 1677 2931 2955 2964	33	675,0 713,6 784,9 817,7 856,3 858,2 869,8 954,6 966,2 1089,7 1151,4 1234,3 1269 1278 1394 1456 1460 1470 1479 1539 1653 1674 1681 2808 2814 2850 2891 2904 2933 2951 2953 2972 3089
<i>MON</i>	4	736 1460 2931 2958,5	27	692,3 916,1 1060 1076 1124 1132 1178 1209 1243 1262 1288 1296 1309 1336 1392 1425 1504 1605 1618 2792 2802 2806 2894 2898 2964 2991 3035
<i>RON</i>	4	736 1056,9 2941 2960	18	752,1 761,8 810 848,6 860 869,8 898,7 918 920 968,2 1072 1174 1178 1284 1307 1313 2772 2935
<i>IAD</i>	4	682,7 736 2877 2931	22	767,6 837 840 952 1132 1161 1193 1236 1348 1440 1454 1517,8 2831 1581 1679 1683 2908 2912 2945 2947 3039
<i>PI</i>	32	660 665,4 671,1 684,6 694,3 703,9 707,7 729 736 759,9 773,9 783 798,4 893,6 901 958,5 1039 1047 1056 1093 1207 1357 1430 1442 1460 2844 2914 2923 2933 2948 2958 2981	35	7001,1 707,8 757,9 763,7 779,1 896,8 900,7 945 952 960 1035 1113 1145 1381 1394 1396 1441 1445 1487 1504 1549 1579 1587 1614 1627 1658 1660 2793 2866 2868 2876 2928 2932 2943 3018
<i>Benzeno</i>	15	665,4 684 703,4 715 727,1 736 754,9 800,4 958,5 989,4 1056,9 2920 2931 2953 2960	18	669,2 687,7 618,2 1061 1092 1157 1257 1263 1269 1440 1460 2765 2771 2824 2908 2914 2955 3002
<i>aromáticos</i>	24	660 667 673,2 675,3 688,5 703,9 717,5 736,7 618,2 775,3 815,8 837 873,7 918 1039 1047 1093 1449 1458 2925 2957 2966 2933 2979	23	698,1 894,9 987,4 1008,7 1091,6 1144 1157 1169 1429 1554 1601 1606 1632 2798 2810 2835 2889 2920 3022 3045 3064 3070 3107
<i>Saturados</i>	3	1458 734,8 2960,5	23	711,6 792,6 810,0 845,4 835,1 867,9 921,9 923,9 923,8 1014,4 1060 1282 1313 1335 1410 1421 1450 1498 1524 1602 2920 2968 2990 3085

A tabela 7.1.33 mostra os números de onda selecionados para as propriedades da gasolina. De um modo geral, o método APS selecionou um menor número de variáveis que o método AG.

A maioria dos números de onda selecionados, assim como no caso do diesel, correspondem às principais absorções características dos hidrocarbonetos.

O método APS selecionou 19 números de onda para **densidade**, contra 39 números de onda selecionados pelo método AG. Em relação aos números de onda selecionados pelo método APS, quatro são atribuídos aos estiramentos axiais C-H de parafinas (2964 a 2924 cm^{-1}); nove foram atribuídos à deformação angular fora do plano de aromáticos (802 - 660 cm^{-1}). O método AG também selecionou números referentes a C-H de parafinas (2819 a 2991 cm^{-1}); Os números de onda na região de 1600 cm^{-1} são atribuídos a vibrações do esqueleto do anel aromático.

A faixa de 1344 a 1475 cm^{-1} está relacionada com C-H (deformação angular) de CH_3 e CH_2 . Números de onda referentes à deformação angular de C-H de aromáticos fora do plano foram selecionados na faixa de (854 a 690 cm^{-1}) e dois números relacionados ao estiramento C-H de aromáticos: 3047 e 3060 cm^{-1} .

Em relação a **T10%** foram selecionados 27 números de onda método APS e 32 para o método AG. Em relação ao estiramento C-H de parafinas, foram selecionados sete números de onda por APS e AG; em relação aos aromáticos (deformação angular fora do plano) foram verificados 11 números de onda selecionados por APS e cinco selecionados por AG. O método AG selecionou ainda, números de onda na faixa de 1567 a 1573 cm^{-1} , atribuídos a modos de vibração do anel aromático, bem como ao estiramento C-H do anel na faixa de 3105 a 3066 cm^{-1} .

Para **T90%**, através do método APS foram selecionados 32 números de onda (APS) e 37 números de onda (AG). Os números de onda (seleção APS) na faixa de

2976 a 2910 cm^{-1} e 2850 a 2933 cm^{-1} (AG) foram atribuídos C-H de parafinas. Abaixo de 873 cm^{-1} , os números de onda foram atribuídos de C-H de aromáticos fora do plano.

A faixa de 1392 a 1178 cm^{-1} corresponde a absorções de hidrocarbonetos ramificados. Observou-se que para esta propriedade foram selecionados números de onda relacionados as absorções de C-O do álcool presente na gasolina (1000 a 1100 cm^{-1}).

Foram selecionados oito (APS) e 33 (AG) números de onda selecionados para **PFE**. Dos oito números de onda (APS) tem-se 2964 a 2932 referentes ao estiramento assimétrico de CH_2 e CH_3 de parafinas e 684,6 e 736 cm^{-1} relacionados com deformação angular C-H de aromáticos fora do plano. O número de onda 1460 corresponde deformação angular de C-H (CH_3 e CH_2). O número de onda 1677 cm^{-1} é atribuído a ligação C=C do anel.

O método AG selecionou números de onda referentes ao estiramento C-H de parafinas (2972 a 2808 cm^{-1}). Foram selecionados números de onda referentes aos modos de deformação angular CH_3 e CH_2 (1479 a 1394 cm^{-1}).

Também foram selecionados números de onda referentes a hidrocarbonetos ramificados (1278 a 1151 cm^{-1}). Abaixo de 900 cm^{-1} os números de onda correspondem à deformação angular C-H de aromáticos fora do plano.

Em relação ao parâmetro **MON** foram selecionados quatro números de onda (APS) e 27 (AG). É importante ressaltar que a octanagem da gasolina está associada principalmente a presença de hidrocarbonetos aromáticos e ramificados.

Os números de onda (APS) que correspondem aos estiramentos C-H são 2931 e 2958 cm^{-1} . A absorção em 1460 cm^{-1} está associada à vibração do esqueleto do anel aromático. E o número de onda 736 cm^{-1} corresponde à deformação angular de C-H fora do plano do anel aromático.

Para os números de onda selecionados por AG, quatro estão relacionados ao estiramento C-H (2921 a 2894 cm^{-1}); cinco às absorções dos ramificados (1124 a 1243 cm^{-1}) e (1392 a 1309 cm^{-1}) à deformação angular (CH_3 e CH_2). O número de onda 3035 cm^{-1} é atribuído a deformação axial de C-H de aromáticos. Os números de onda 1605 e 1618 relacionam-se a C=C com conjugação. O número de 692 cm^{-1} é atribuído à deformação angular de C-H fora do plano do anel aromático.

A seleção para MON mostrou que vários números de onda estão relacionados com aromáticos e ramificados, indicando que contém sentido físico aplicar tais números de onda na modelagem desta propriedade.

Para **RON** o método APS selecionou quatro números de onda, dos quais dois estão relacionados ao estiramento C-H (2960 e 2941 cm^{-1}) e um a deformação angular de C-H de aromáticos fora do plano (736 cm^{-1}).

O método AG selecionou um número de onda relacionado ao estiramento de C-H (2935 cm^{-1}). A faixa de 898 a 752 cm^{-1} foi atribuída a deformação angular C-H de aromáticos fora do plano e 1174 e 1178 cm^{-1} foram relacionados a hidrocarbonetos ramificados. Em relação ao parâmetro **IAD** foram selecionados quatro números de onda (APS), dos quais dois foram atribuídos ao estiramento C-H (2877 e 2931 cm^{-1}) e dois foram atribuídos à deformação angular C-H de aromáticos fora do plano (736 e 682 cm^{-1}). É interessante ressaltar que o número de onda 736 cm^{-1} foi selecionado por APS nas três propriedades de octanagem (MON, RON e IAD).

O método AG selecionou números de onda relacionados ao estiramento C-H de (2947 a 2908 cm^{-1}); atribuídos a aromáticos fora do plano e quatro relacionados a ramificados.

Em relação ao parâmetro **PI**, foram selecionados 32 números de onda (APS), dos quais sete foram atribuídos ao estiramento C-H de parafinas (2844 a 2981 cm^{-1}); quatro a deformação angular CH_3 e CH_2 (1460 a 1357 cm^{-1}); treze são

correspondentes a C-H de aromáticos fora do plano (660 a 798 cm^{-1}) O método AG selecionou 35 números de onda, dos quais cinco foram atribuídos ao C-H de parafinas (2943 a 2866 cm^{-1}); cinco foram atribuídos ao estiramento do anel (1487 – 1587 cm^{-1}) e correspondem a C=C ($1627, 1658$ e 1660 cm^{-1}).

Para o teor de **benzeno** foram selecionados 15 números de onda (APS) e 18 (AG). Para APS seis números de onda foram atribuídos à deformação angular de C-H de aromáticos fora do plano (665 a 755), esta região é característica do benzeno; três a deformação angular de aromáticos no plano (958 a 1056 cm^{-1}); quatro ao estiramento C-H (2920 a 2960 cm^{-1}). Para o método AG dois números de onda estão relacionados com a deformação angular de aromáticos no plano (669 a 687 cm^{-1}); quatro são referentes a CH_3 e CH_2 de (2824 - 2955 cm^{-1}). Acima de 3000 cm^{-1} os números de onda são atribuídos ao estiramento de C-H de aromáticos.

Em relação ao teor de **aromáticos** foram selecionados 24 números de onda (APS) e 23 (AG). Em relação ao método APS foram selecionados números de onda referentes ao estiramento C-H (2979 a 2925 cm^{-1}). Os números de onda referentes a deformação angular C-H de aromáticos fora do plano correspondem a faixa de (873 a 660 cm^{-1}); os números de onda relacionados a deformação angular de C-H de aromático no plano foram observados na faixa de 1093 - 918 cm^{-1} .

Em relação aos números de onda obtidos por AG, três números de onda foram atribuídos a C-H (2835 , 2889 e 2920 cm^{-1}). Foram atribuídos ao estiramento C-H de aromático os números de onda selecionados entre 3022 a 3107 cm^{-1} . Os números de onda referentes a deformação angular de aromáticos fora do plano (C-H) foram relacionados a 894 e 698 cm^{-1} .

A faixa entre 1008 e 1169 cm^{-1} corresponde aos modos de deformação angular no plano de C-H de aromáticos. Os números de onda de 1601 e 1554 estão associados a vibrações do esqueleto.

Para o teor de **saturados** foram selecionados três números de onda através do método APS : 2960 cm^{-1} referente ao estiramento C-H de parafinas; 1458 cm^{-1} referente à deformação angular CH_3 e 734 referente a deformação angular de $-(\text{CH}_2)_4$. Em relação aos números de onda selecionados pelo método AG, estes foram atribuídos ao estiramento C-H de parafinas e aos modos de deformação angular CH_3 e CH_2 .

7.1.5-3 Números de onda selecionados para as propriedades do álcool automotivo

A figura 7.1.73 mostra um esquema com as principais bandas de absorção do álcool automotivo anidro:

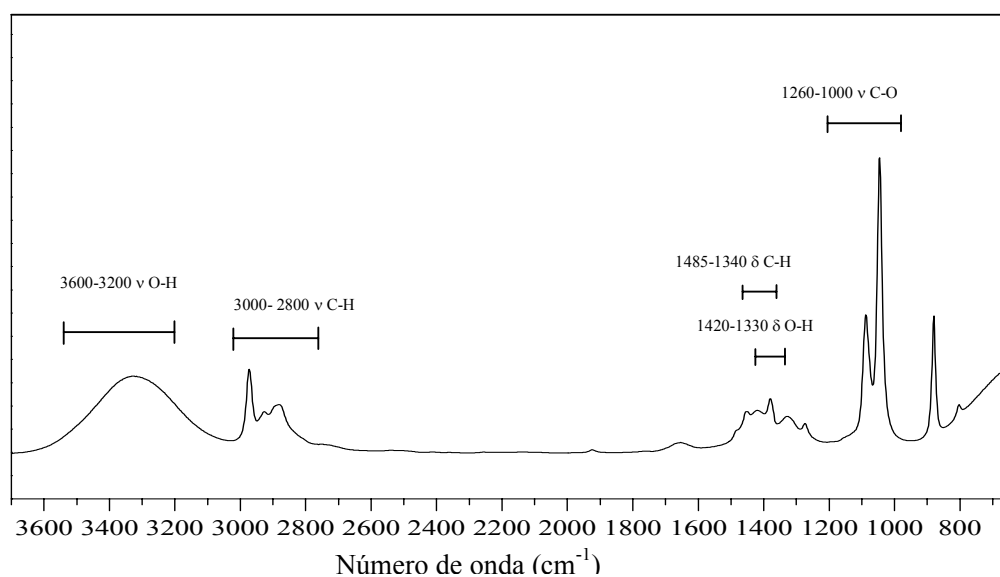


Fig. 7.1.73: Esquema de atribuição de bandas no infravermelho médio para espectro do álcool automotivo. O símbolo v corresponde ao estiramento e δ à deformação angular.

Em relação aos espectros do álcool foram selecionados números de onda relacionados a deformação angular O-H fora do plano em 650 a 780 cm^{-1} ; deformação angular no plano de O-H em 1330 e 1420 cm^{-1} ; estiramento de C-O entre 1000-1260 cm^{-1} e 3200 a 3600 cm^{-1} para estiramento de O-H associado, conforme mostrado na tabela 7.1.34:

Tabela 7.1.34: Variáveis (números de onda) selecionados para o álcool automotivo.

Parâmetro (ÁLCOOL)	APS							AG								
	nº	Variáveis Selecionadas (cm ⁻¹)							nº	Variáveis Selecionadas (cm ⁻¹)						
<i>Massa Específica</i>	21	650	652	658	653	661,5	667,3	731	27	707,8	740,6	763,7	835,1	865,9	995	
		881,4	1035	1039	1041	1043	1045			1018	1151	1483	1485	1512	1629	1680
		1049	1047	1051	1084	1464	2926			1693	1803	2785	2787	2790	2817	2841
		3018	3394							2852	3018	3167	3481	3535	3557	3668
<i>Teor Alcoólico</i>	24	660	663,4	665,4	669,2	673,1	676,9		27	721,3	725,1	804,9	877,5	952,7	1068	
		690,4	700,1	869,8	881,4	1028	1033			1172	1263	1325	1356	1423	1562	1635
		1037	1045	1053	1056	1059	1063			1807	2682	2887	2891	2968	3089	3101
		1084	1095	1755	2972	3427	1124			3173	3211	3330	3546	3566	3583	3653

Em relação aos parâmetros do álcool, foi verificado que para a massa específica foram selecionados 21 números de onda através do método APS e 27 números de onda através do método AG. Dos números de onda selecionados por APS, sete podem estar relacionados com deformação angular O-H fora do plano (650 a 731 cm^{-1}); nove foram atribuídos ao estiramento C-O (1035 a 1084 cm^{-1}); um número de onda foi atribuído à deformação angular no plano de O-H (1464 cm^{-1}) e um número de onda ao estiramento O-H associado (3394 cm^{-1}). Dos números de onda selecionados por AG quatro correspondem à deformação angular O-H fora do plano (707 a 763 cm^{-1}); dois a estiramento C-O (1018 e 1151 cm^{-1}) e quatro ao estiramento O-H associado (3360 a 3481 cm^{-1}).

Para o teor alcoólico foram verificados 24 números de onda selecionados por APS e 27 selecionados por AG. O método APS selecionou onze números de onda atribuídos ao estiramento C-O (1028 a 1124 cm^{-1}). Oito são referentes à

deformação angular O-H fora do plano (660 a 700 cm^{-1}) e um (3427 cm^{-1}) relacionado ao estiramento O-H associado. O método AG selecionou dois números de onda relacionados à deformação angular O-H fora do plano (721 e 725 cm^{-1}). Três são atribuídos ao estiramento C-O (1068 a 1263 cm^{-1}); dois números de onda (1325 e 1536 cm^{-1}) à deformação angular no plano de O-H e seis números de onda relacionados com estiramento de O-H associado (3653 a 3211 cm^{-1}).

7.2-RESULTADOS DA APLICAÇÃO II: ESTUDO DAS PROPRIEDADES DE DEGRADAÇÃO DO PET:

7.2.1- Análise de Espectros DRIFTS e Fotoacústica (PAS) do PET

Foram obtidos três tipos de espectros: espectros DRIFTS das amostras submetidas a um processo de secagem; espectros DRIFTS de amostras sem processo de secagem e espectros de fotoacústica (PAS) das amostras não secas.

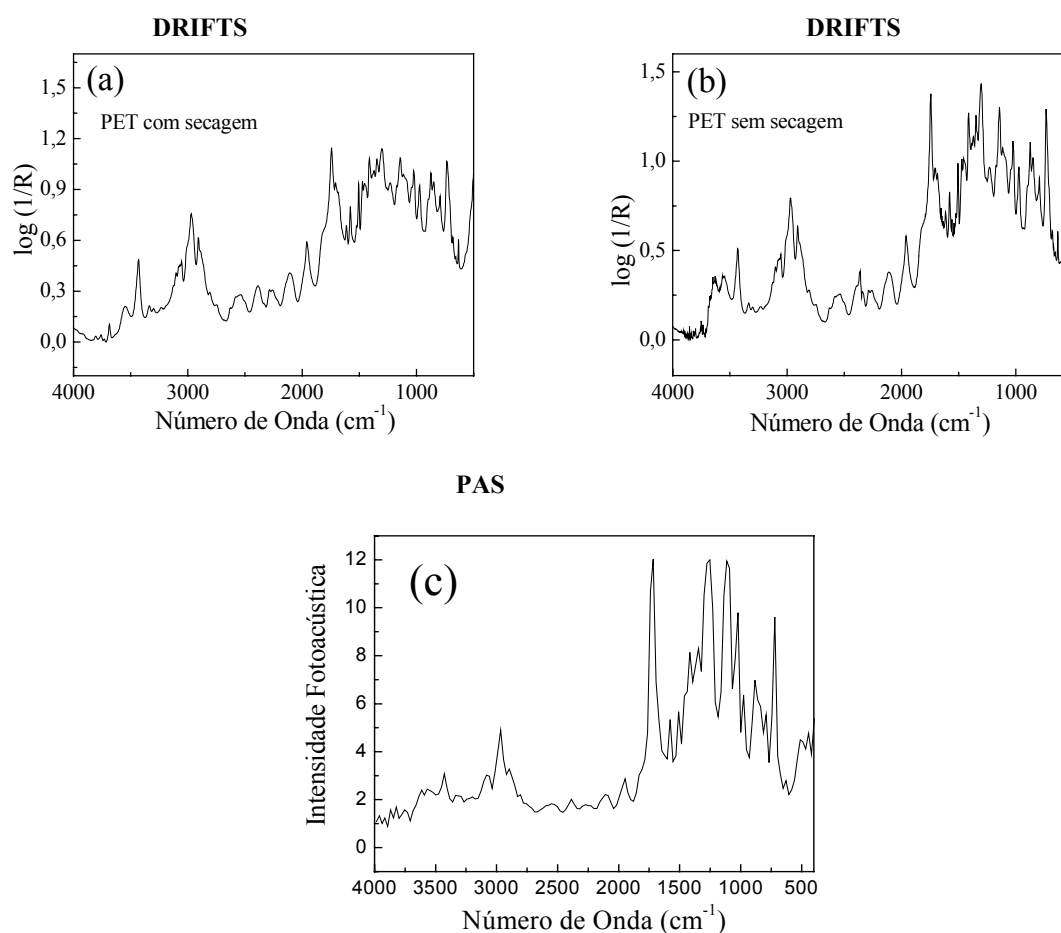


Figura 7.2.1- Espectros de amostras de PET no infravermelho médio. (A) DRIFTS de amostras secas. (b) DRIFTS de amostras não secas. (c) Fotoacústica de Amostras não secas.

Os espectros DRIFTS foram examinados antes da aplicação da calibração multivariada para detectar distorções ocasionadas pela reflexão especular, que pode ocorrer em espectros de reflexão difusa.

Com o objetivo de identificar as distorções causadas pela reflexão especular, os espectros DRIFTS foram comparados com outras técnicas de FTIR, que não apresentam interferência da componente especular com as técnicas de transmissão e ATR. Deste modo foi possível identificar as regiões distorcidas e eliminá-las antes da aplicação da calibração multivariada.

As regiões com distorção podem interferir negativamente no processo de calibração, acarretando erros, pois muitas vezes apresentam relações não-lineares com as concentrações a serem determinadas.

Para viabilizar tal comparação, os espectros foram divididos em duas regiões:

- 4000 a 1800 cm^{-1} , comparados com um espectro de transmissão do filme do material.
- 1800 a 500 cm^{-1} , comparados a um espectro de ATR do material. Utilizou-se esta técnica (ATR com geometria horizontal) nesta faixa porque nesta região existem fortes absorções, exigindo filmes finos que são difíceis de serem obtidos.

Os espectros obtidos através de DRIFTS das amostras com e sem processo de secagem em comparação com espectros de absorbância e ATR são mostrados na figura 7.2.2:

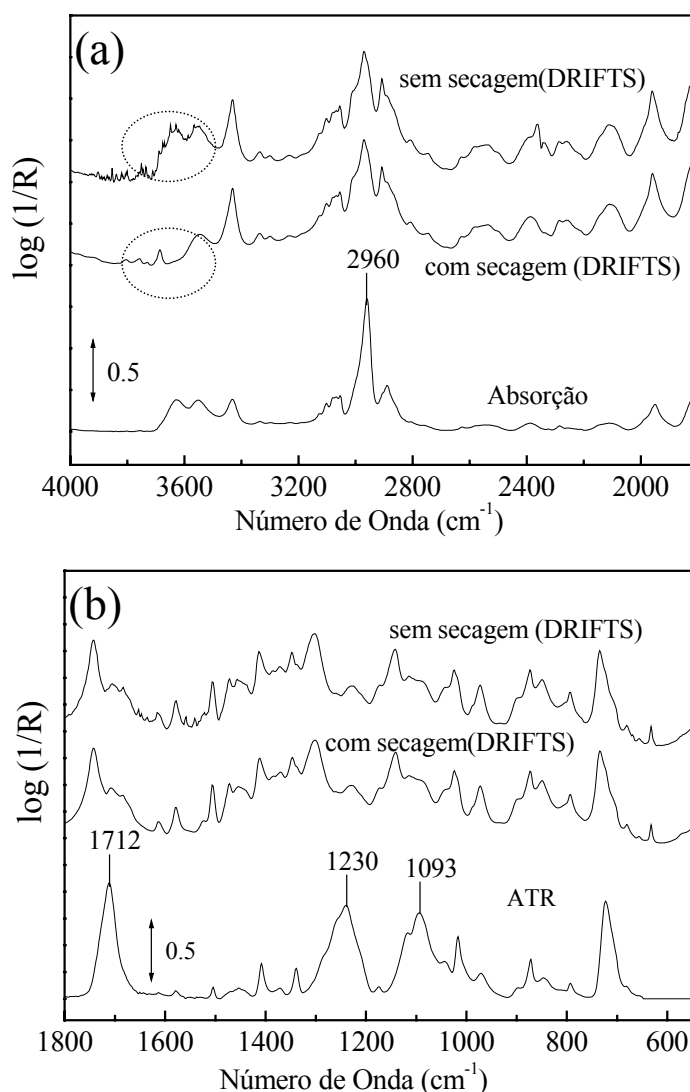


Fig. 7.2.2: Espectros FTIR do 1º ciclo de extrusão: (a) DRIFTS (com e sem secagem) e absorbância em 4000-1800 cm^{-1} e (b) DRIFTS (com e sem secagem) e ATR em 1800-500 cm^{-1} . As bandas assinaladas apresentam distorção. As regiões circuladas correspondem às bandas de absorção da água.

A comparação dos espectros da figura 7.2.2(a) mostra uma distorção na região de 2960 cm^{-1} . Esta banda apresenta correspondência a 2970 cm^{-1} nos espectros DRIFTS (um deslocamento de 10 cm^{-1} em relação ao espectro de absorção). Além disso, a razão entre as intensidades das bandas também é diferente que para a absorção.

Este fato indica uma provável distorção devido à reflexão especular. Os espectros de amostras secas e não secas também apresentam uma diferença na região de 3630 cm^{-1} devido às absorções da água (regiões circuladas).

A figura 7.2.2(b) mostra a região de 1800 a 500 cm^{-1} dos espectros DRIFTS de amostras com e sem secagem e o espectro ATR (em baixo) com devida correção. Podem ser observadas três bandas a 1712 , 1230 e 1093 cm^{-1} (figura 7.2.2(b)) características dos grupamentos éster presentes no PET. Os “máximos” destas bandas no espectro ATR não coincidem com os espectros DRIFTS, sugerindo distorção por reflexão especular. Não foram observadas diferenças significantes entre os espectros com e sem secagem nesta região (1800 a 500 cm^{-1}). A tabela 7.2.1 mostra os deslocamento das regiões espectrais distorcidas dos espectros DRIFTS em relação aos espectros de absorbância e ATR.

Tabela 7.2.1: Deslocamento de bandas do espectro DRIFTS em relação aos espectros de Absorção e ATR.

<i>Espectros de referência (absorbância e ATR)</i>	<i>Espectros DRIFTS</i>	<i>Deslocamento</i>
2960 cm^{-1}	2970 cm^{-1}	10 cm^{-1}
1750 cm^{-1}	1712 cm^{-1}	28 cm^{-1}
1300 cm^{-1}	1230 cm^{-1}	70 cm^{-1}
1150 cm^{-1}	1093 cm^{-1}	57 cm^{-1}

Observou-se as bandas distorcidas apresentaram deslocamento de até 70 cm^{-1} . Deste modo, a inclusão ou não destas regiões no modelos de calibração foi determinada com maior critério.

7.2.3- Construção dos modelos PLS

Foram construídos modelos PLS distintos para espectros PAS e DRIFTS com amostras secas e não secas (separadamente) a partir dos dados centrados na média e com correção de linha base. Foi aplicado o método de validação cruzada devido ao pequeno número de amostras (20 amostras).

Como critério para seleção do melhor modelo para previsão da concentração de grupos carboxílicos terminais CEG e viscosidade intrínseca $[\eta]$, utilizou-se os valores de RMSEP. Também foram observados os coeficientes de correlação R e o erro percentual médio (dado pela equação 5.17).

O número de variáveis latentes usadas na construção dos modelos PLS foi determinado através do menor valor de RMSEP.

7.2.3.1- Modelos PLS para determinação de CEG (grupos carboxílicos terminais)

Os modelos PLS foram construídos em relação a sub-regiões espectrais que foram definidas não só pela presença das distorções, mas também com a observação dos coeficientes de regressão (figura 7.2.3) que indicam a relevância de cada região no processo de calibração.

Para a determinação de CEG foram construídos modelos em relação a 3 sub-regiões :

Em relação ao primeiro modelo, considerou-se toda região espectral [figura 7.2.3(a)]. A faixa espectral total corresponde a 3360 cm^{-1} a 500 cm^{-1} .

No entanto, devido as fortes distorções, a região abaixo de 1830 cm^{-1} foi excluída. Então foi considerada uma segunda sub-região de 3660 a 1830cm^{-1} , da qual também foi excluída a região de absorção do CO_2 que corresponde à faixa de 2100 a 2200 cm^{-1} [figura 7.2.3(b)]. Deste modo, o segundo modelo foi construído com as sub-regiões de $1930\text{-}2290\text{ cm}^{-1}$ e $2384\text{-}3660\text{cm}^{-1}$.

O terceiro modelo foi construído a partir das sub-regiões do segundo modelo, excluindo-se ainda a região próxima a 2960 cm^{-1} que também apresenta distorção [figura 7.2.3(c)]. Portanto as sub-regiões utilizadas neste modelo foram: $1930\text{-}2290\text{ cm}^{-1}$, $2384\text{-}2850\text{ cm}^{-1}$ e $3010\text{-}3660\text{cm}^{-1}$.

Cada um dos modelos com diferentes sub-regiões foram aplicados na previsão de CEG dos três tipos de amostras: amostras secas (detector MCT), amostras não secas (detector DTGS) e amostras não secas detector DTGS.

A construção destes modelos teve como objetivo o estudo da influência do processo de secagem, das regiões selecionadas e do tipo de detector no desempenho dos modelos de calibração.

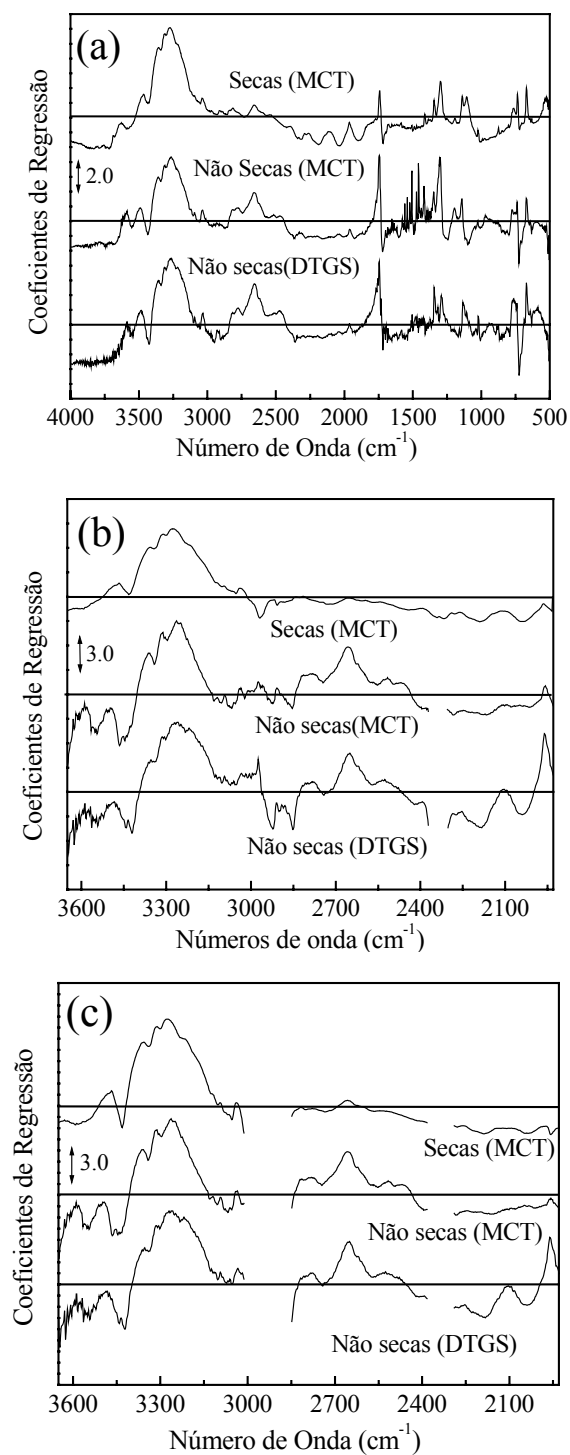


Figura 7.2.3: Coeficientes de Regressão para CEG nas regiões de: (a) 4000 a 500cm⁻¹; (b) 1930-2290, 2384-3660 cm⁻¹ (c) 1930-2290, 2384-2850, 3010-3660 cm⁻¹.

A modelagem através dos espectros de fotoacústica foi feita com a utilização de toda região espectral. Os espectros PAS apresentaram-se muito ruidosos, por isso foi aplicado o alisamento Savitzky-Golay com polinômio de 2ª ordem e 5 janelas antes da etapa de calibração. Deste modo foram construído 10 modelos PLS 1 para a previsão da concentração de grupos carboxílicos terminais na faixa de 27 a 165 $\mu\text{Eq/g}$. Nove destes modelos foram construídos, utilizando-se as diferentes sub-regiões, a partir dos espectros DRIFTS (de amostras secas e não secas e detectores MCT e DTGS) e um modelo foi construído a partir dos espectros de fotoacústica (detector DTGS). A tabela 7.2.2 mostra os resultados da previsão da concentração de CEG de amostras de PET submetidas e não submetidas ao processo de secagem, cujos espectros foram adquiridos através das técnicas DRIFTS e fotoacústica (PAS), utilizando-se os detectores MCT e DTGS:

Tabela 7.2.2: Resultados dos modelos PLS aplicados para a previsão de CEG de amostras de PET

Concentração de Grupos Carboxílicos Terminais (CEG)							
	Amostras	Detec.	R	VL	RMSEP	Erro rel. (%)	Regiões (cm^{-1})
D R I F T S	secas	MCT	0,9649	5	11,33	14,74	Espectro total
	secas	MCT	0,9630	3	11,63	15,15	1930-2290, 2384-3660
	secas	MCT	0,9665	3	11,08	14,31	1930-2290, 2384-2850, 3010-3650
	Não secas	MCT	0,9906	4	5,95	7,43	Espectro total
	Não secas	MCT	0,9800	3	8,46	13,48	1930-2290, 2384-3660
	Não secas	MCT	0,9835	3	7,81	11,97	1930-2290, 2384-2850, 3010-3650
	Não secas	DTGS	0,9876	5	6,76	11,27	Espectro total
	Não secas	DTGS	0,9933	4	4,98	6,85	1930-2290, 2384-3660
	Não secas	DTGS	0,9907	4	5,84	7,74	1930-2290, 2384-2850, 3010-3650
PAS	Não secas	PAS	0,9483	5	11,96	19,37	Espectro total

R é o coeficiente de correlação e VL é o número de variáveis latentes

Foi observado que houve um decréscimo do número de variáveis latentes com a exclusão das regiões distorcidas, sem alterações significativas nos valores de RMSEP, R e erro relativo (%). No entanto para as amostras submetidas à secagem (detector MCT), os valores de RMSEP foram mais altos. O modelo relacionado com a técnica PAS também apresentou alto valor de RMSEP e erro relativo percentual.

Os valores assinalados em **negrito** correspondem aos resultados obtidos com os modelos de maior confiabilidade (apesar de não representarem o menor RMSEP obtido), pois excluem totalmente as regiões distorcidas.

A figura 7.2.4 mostra o número de variáveis independentes utilizadas na construção dos modelos de previsão de CEG com todas as regiões distorcidas excluídas (resultados em **negrito** da tabela 7.2.2):

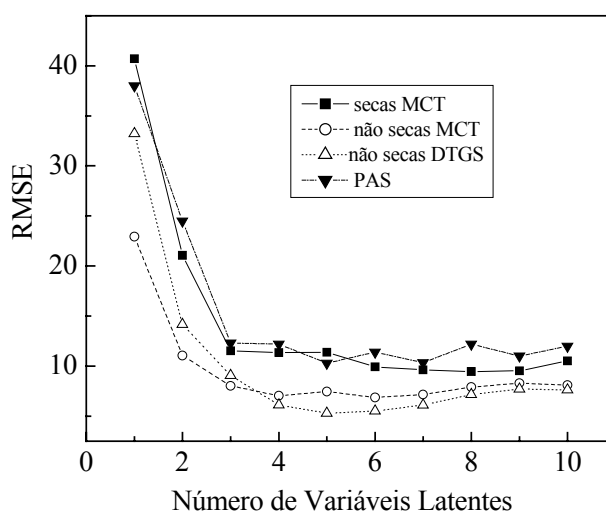


Figura 7.2.4: Número de variáveis latentes selecionadas nos modelos PLS utilizados na previsão da concentração de CEG para o modelo com regiões distorcidas excluídas (DRIFTS) e com toda região espectral (PAS).

Os gráficos dos valores previstos pelos modelos PLS em função dos valores reais para o modelo com todas as regiões distorcidas excluídas são mostrados na figura 7.2.5:

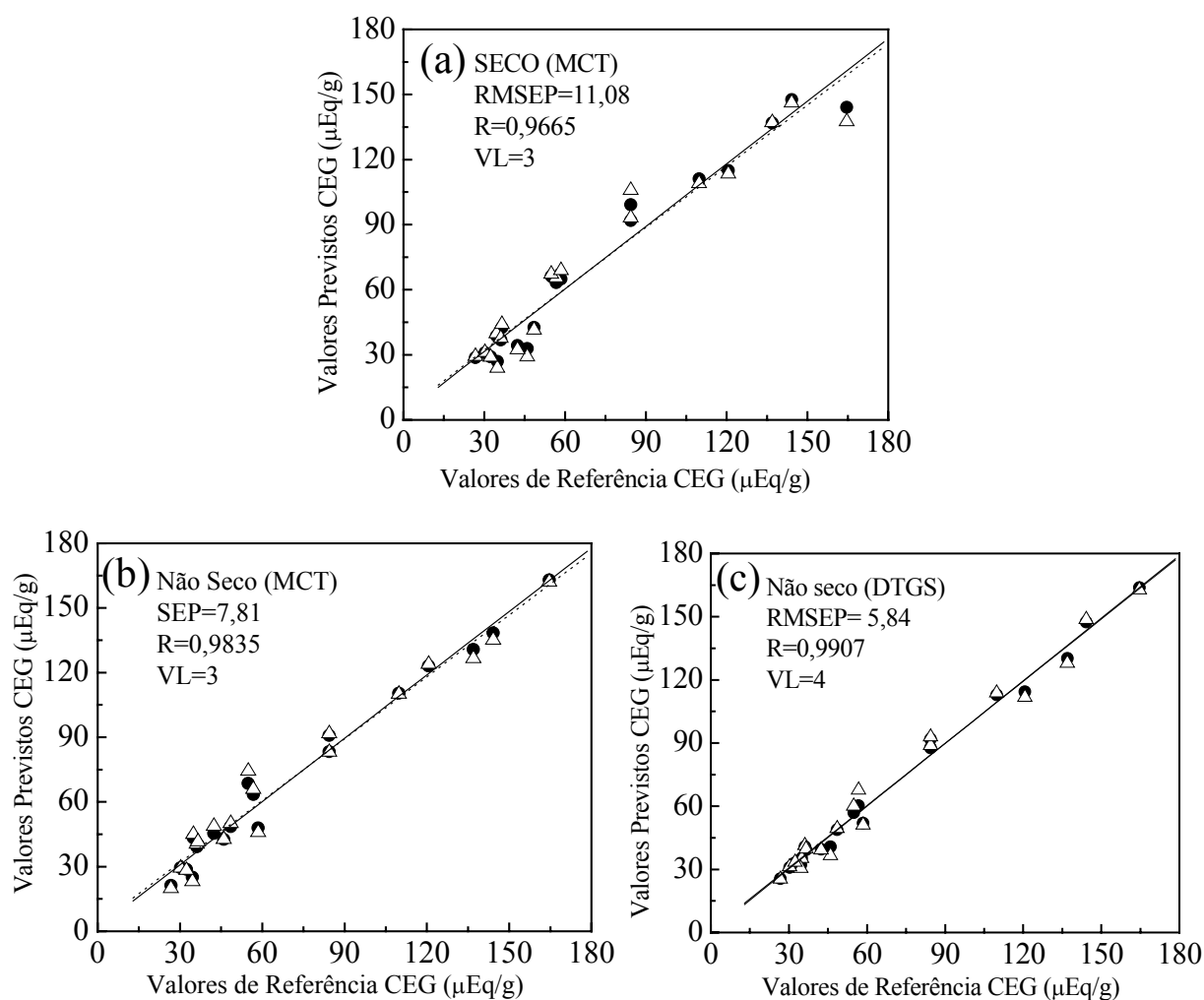


Figura 7.2.5: Valores previstos em função de valores de referência para a previsão da concentração de CEG referente aos modelos construídos com as sub-regiões de 1930-2290, 2384-2850, 3010-3660 cm^{-1} . (●) Amostras de calibração. (Δ) Amostras de validação.

O gráfico de valores de referência em função dos valores previstos para CEG através do modelo construído a partir dos espectros de fotoacústica são mostrados na figura 7.2.6:

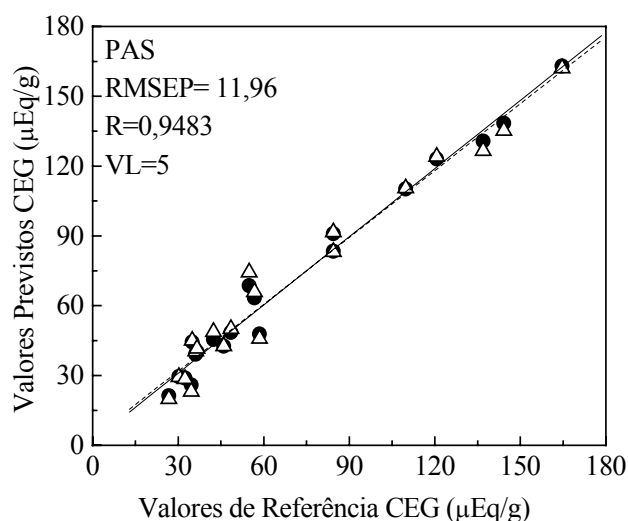


Figura 7.2.6: Valores previstos (PAS) em função de valores de referência para a previsão da concentração de CEG referente ao modelo construído com as sub-regiões de 1930-2290, 2384-2850, 3010-3660 cm^{-1} . (●) Amostras de calibração. (Δ) Amostras de validação.

Para verificar se houve diferenças significativas entre os modelos, consideram-se as diferentes técnicas, detectores e processos de secagem, foi aplicado o Teste-F com 95% de confiança para os valores de RMSEP da tabela 7.2.2 (assinalados em **negrito**) do conjunto de amostras secas (MCT), não secas (MCT), não secas (DTGS) e PAS (DTGS).

Os resultados encontram-se na tabela 7.2.3:

Tabela 7.2.3. Teste-F para modelos de previsão de CEG.

Teste-F ($F_c=2,16$)	Não secas (MCT)	Não secas (DTGS)	Não secas (PAS)
Am. Secas (MCT)	2,01	<i>3,60</i>	1,07
Não secas (MCT)	-	1,78	<i>2,34</i>
Não secas (DTGS)	1,78	-	<i>4,19</i>

Os valores assinalados em *itálico* estão acima do valor tabelado ($F_{\text{crítico}}=2,16$), indicando que há diferenças significativas entre os resultados obtidos

com amostras secas (MCT) e amostras não secas (DTGS) e amostras não secas e PAS.

Para as amostras não secas, os resultados mostram que o uso de diferentes detectores não influenciou na previsão de CEG. Os resultados obtidos para PAS apresentaram diferenças significativas e o maior valor de RMSEP para CEG. Este fato pode ser atribuído a grande quantidade de ruídos nos espectros PAS, apesar das operações de alisamento.

A diferença significativa entre os modelos construídos para amostras secas (MCT) e não secas (DTGS) pode ser atribuída ao uso de diferentes detectores neste caso em específico, pois os resultados indicaram que o processo de secagem não influenciou os outros resultados (amostras secas e não secas, medidas com o mesmo detector), mas há uma tendência do detector MCT fornecer maiores valores de RMSEP.

Os maiores valores de RMSEP foram verificados para as baixas concentrações de CEG. Este fato pode ser justificado por limitações da técnica de referência (método Pohl) que utiliza micro titulação com reagentes de alta pureza, com etapas rigorosamente cronometradas e que não são eficazes para medidas quantitativas de concentrações muito baixas dos grupos carboxílicos terminais.

7.2.3.2- Modelos PLS para determinação de $[\eta]$ (viscosidade intrínseca)

A faixa de valores de viscosidade intrínseca das amostras analisadas corresponde a 0,346 a 0,780 dLg⁻¹.

Para a viscosidade intrínseca, também foram considerada três sub-regiões (espectros DRIFTS) na construção dos modelos PLS.

Além da faixa espectral completa [figura 7.2.7(a)], foram construídos modelos a partir das sub-regiões de 3660-2260cm⁻¹[figura 7.2.7(b)], excluindo-se a região que apresenta maior distorção (abaixo de 2260cm⁻¹).

O terceiro modelo considerou a faixa de 3660-2260cm⁻¹ sem a região distorcida na faixa de 2960 cm⁻¹ [figura 7.2.7(c)],

Os gráficos de coeficientes de regressão em função dos números de onda são mostrados na Fig 7.2.7:

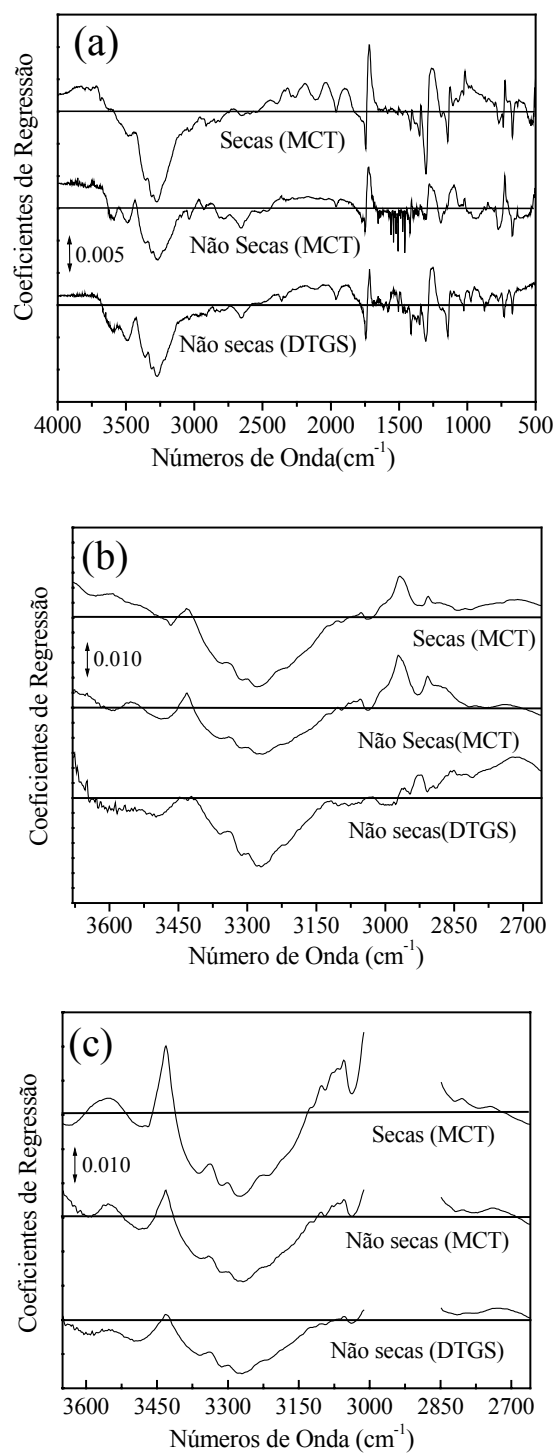


Figura 7.2.7: Coeficientes de Regressão para $[\eta]$ nas regiões de: (a) 4000 a 500cm⁻¹; (b) 3660-2660 cm⁻¹ (c) 3660-3010, 2850-2660 cm⁻¹.

Foi construído um modelo PLS a partir dos espectros PAS para a previsão de $[\eta]$, utilizando toda a região espectral (4000 a 500 cm^{-1}).

Os resultados para a regressão dos modelos com espectros DRIFTS e PAS estão expressos na tabela 7.2.4:

Tabela 7.2.4: Calibração Multivariada (PLS) para $[\eta]$ para espectros DRIFTS e PAS:

Viscosidade Intrínseca $[\eta]$							
	Amostras	Detec.	R	VL	RMSEP	Erro rel. (%)	Regiões (cm^{-1})
D R I F T S	secas	MCT	0,9687	4	0,033	5,095	Espectro total
	secas	MCT	0,9595	3	0,037	5,172	2660-3650
	secas	MCT	0,9580	3	0,038	5,054	2660- 2850, 3010-3660
	Não secas	MCT	0,9725	4	0,032	4,840	Espectro total
	Não secas	MCT	0,9565	3	0,038	5,517	2660-3650
	Não secas	MCT	0,9592	3	0,037	4,859	2660- 2850, 3010-3660
	Não secas	DTGS	0,9600	4	0,037	5,714	Espectro total
	Não secas	DTGS	0,9660	3	0,034	4,975	2660-3650
	Não secas	DTGS	0,9674	3	0,034	5,178	2660- 2850, 3010-3660
PAS	Não secas	PAS	0,9674	4	0,028	4,728	Espectro total

R é o coeficiente de correlação e VL é o número de variáveis latentes

De maneira semelhante aos resultados mostrados para a previsão de CEG, o melhor modelo está associado à eliminação das regiões distorcidas. Os resultados destes modelos estão assinalados em negrito na tabela 7.2.4.

A figura 7.2.8 mostra o número de variáveis independentes utilizadas na construção dos modelos PLS de previsão de com todas as regiões distorcidas excluídas (resultados em negrito da tabela 7.2.2):

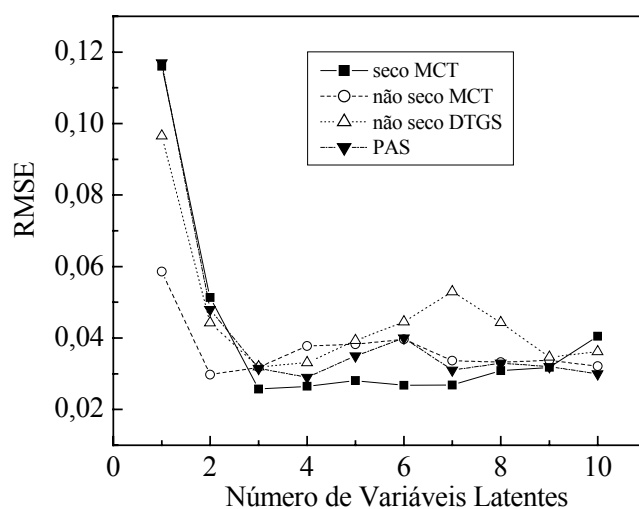


Figura 7.2.8: Número de variáveis latentes selecionadas nos modelos PLS utilizados na previsão da $[\eta]$ para o modelo com regiões distorcidas excluídas (DRIFTS) e com toda região espectral (PAS).

Os gráficos dos valores previstos pelos modelos PLS para $[\eta]$ (considerando-se amostras secas, e não secas e diferentes detectores) em função dos valores reais para o modelo com todas as regiões distorcidas excluídas são mostrados na figura 7.2.9:

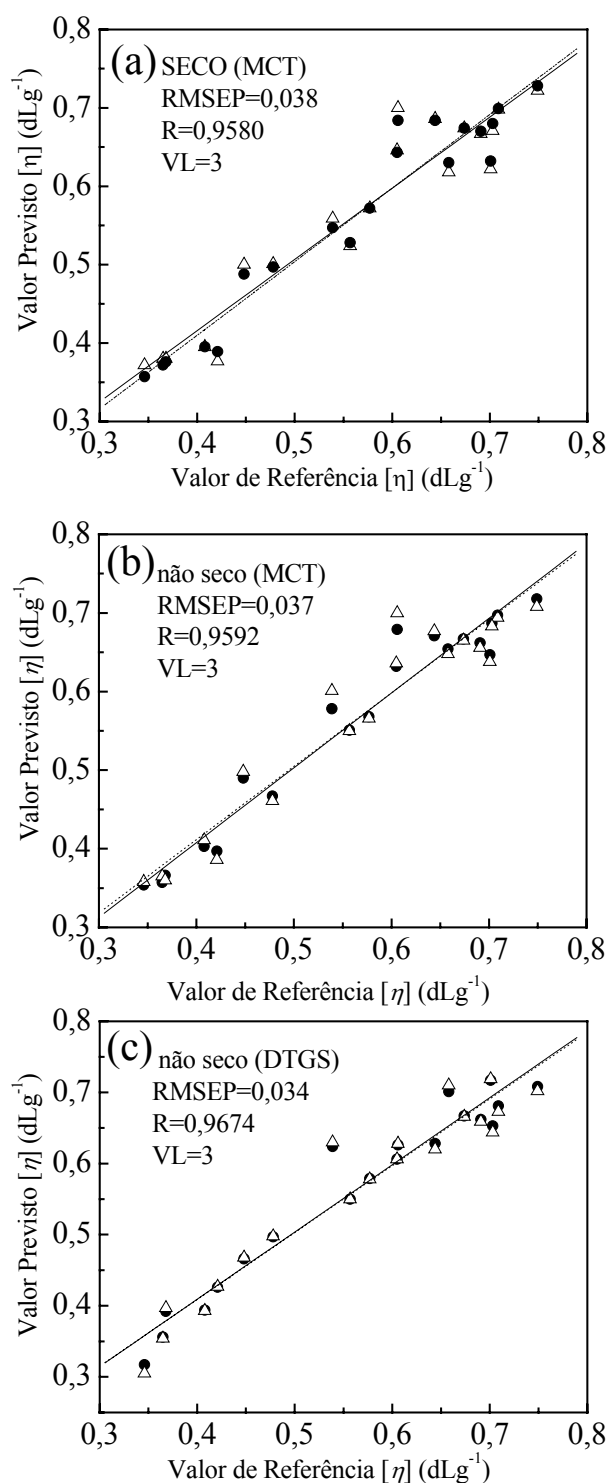


Figura 7.2.9: Valores previstos em função de valores de referência para a previsão de $[\eta]$ referente aos modelos construídos com as sub-regiões de 2660- 2850 e 3010-3660 cm^{-1} . (●) Amostras de calibração.(Δ) Amostras de validação.

O gráfico de valores de referência em função dos valores previstos para $[\eta]$ através do modelo construído a partir dos espectros de fotoacústica são mostrados na figura 7.2.10:

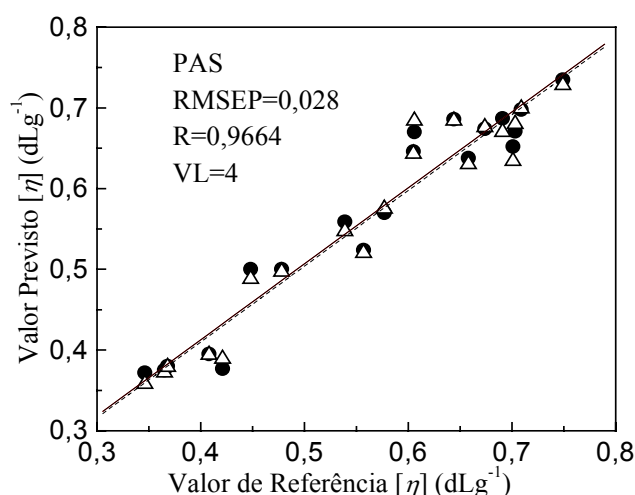


Figura 7.2.10: Valores previstos (PAS) em função de valores de referência para a previsão da concentração de $[\eta]$ referente ao modelo construído com as sub-regiões de 2660- 2850 e 3010-3660 cm^{-1} . (●) Amostras de calibração.(Δ) Amostras de validação.

Os resultados para a previsão de $[\eta]$ mostram valores muito próximos de RMSEP, o que indica que não houve diferenças significativas entre os resultados dos modelos, independente do processo de secagem, detector utilizado ou técnica espectroscópica (DRIFTS ou PAS). Esta afirmação foi confirmada através da aplicação do Teste-F (95% de confiança) para os valores assinados em negrito da tabela 7.2.4, com resultados mostrados na tabela 7.2.5:

Teste-F ($F_c=2,16$)	Não secas (MCT)	Não secas (DTGS)	Não secas (PAS)
Am. Secas (MCT)	1,05	1,24	1,84
Não secas (MCT)	-	1,18	1,75
Não secas (DTGS)	1,18	-	1,47

Os valores de F da tabela 7.2.5 estiveram abaixo do valor de F crítico, por isso os modelos foram considerados estatisticamente semelhantes.

Os modelos apresentaram erros médios relativos de 7,5% a 14% para DRIFTS e 19% para PAS, considerando que para os maiores valores de CEG, o erro relativo cai para 4%. O desvio padrão para o método de referência foi estimado em 1%.

Para $[\eta]$ o erro relativo médio encontram-se na faixa de 4 a 5%. O desvio padrão do método de referência é de 3%.

7.2.4- Atribuição das sub-regiões espectrais

A atribuição de todas as bandas relacionadas ao processo de calibração não é uma tarefa fácil, uma vez que está presente não somente bandas de absorções de transições fundamentais, mas também bandas oriundas de bandas harmônicas e combinações, além de sobreposições de outras bandas relacionadas com a cadeia carbônica das moléculas de PET.

No entanto, algumas atribuições podem ser explicadas. Na faixa de 3700 a 3500 cm^{-1} aproximadamente ocorrem as absorções do estiramento --OH (3547 cm^{-1}) dos grupos hidroxílicos terminais^[100].

Para os espectros das amostras não secas este modo apresenta-se parcialmente sobreposto com absorções da água, cujas bandas são bastante alargadas e têm início em 3756 cm^{-1} e 3657 cm^{-1} (estiramentos --OH anti-simétrico e simétrico). O pico de 3431 cm^{-1} corresponde à segunda harmônica da carbonila (C=O), cuja transição fundamental (primeira harmônica) ocorre em aproximadamente 1715 cm^{-1} .

A faixa de 3229-3335 cm^{-1} corresponde a uma região onde são encontradas absorções de grande importância para a construção dos modelos. Nesta faixa

encontra-se o pico de 3271 cm^{-1} correspondente ao grupo carboxílico terminal, que está sobreposto a outras absorções da cadeia carbônica do PET. Através do gráfico de regressão foi verificado que o valor máximo dos coeficientes correspondem exatamente a absorção de 3271 cm^{-1} , que está positivamente correlacionada com a concentração de CEG e a relação inversa é verificada para $[\eta]$ que está negativamente correlacionada. Isto é atribuído ao fato de que à medida em que ocorre o processo degradativo, a viscosidade intrínseca diminui e a concentração de grupos carboxílicos terminais aumenta.

É importante ressaltar que a região de 3650 a 3100 cm^{-1} foi essencial para o processo de calibração. Esta faixa espectral correspondente às absorções dos grupos carboxílicos terminais ($-\text{COOH}$), dos grupos hidroxílicos terminais ($-\text{OH}$) e das moléculas de água (H_2O).

Na região de 1990 a 2660 cm^{-1} estão compreendidas absorções correspondentes às bandas harmônicas e combinações das fortes absorções fundamentais utilizadas na modelagem.

8- CONCLUSÕES E PROPOSTAS

O presente trabalho reafirma que a substituição dos métodos químicos por FTIR é bastante oportuna e requer apenas um curto tempo de análise. Adicionalmente, o equipamento é de fácil operação e o método não gera resíduo químico.

A aplicação da calibração multivariada em dados de FTIR potencializa a espectroscopia no infravermelho médio, facilitando a resolução dos problemas quantitativos.

Quanto aos combustíveis, pode-se afirmar que o método utilizado é bastante aplicável na estimativa de propriedades que qualificam o combustível. Os valores de RMSEP mostraram-se inferiores aos da reprodutibilidade do método padrão para os parâmetros do diesel como densidade, temperaturas de destilação, índice de cetano, teor de enxofre e viscosidade. Em relação aos parâmetros da gasolina, foram obtidos valores de RMSEP abaixo da reprodutibilidade dos métodos padrão para as temperaturas de destilação, PFE, MON, IAD, PI e teores de benzeno, aromáticos e saturados. Em relação aos parâmetros massa específica e teor alcoólico do álcool automotivo, pode-se afirmar que ambos os valores de RMSEP estiveram abaixo da reprodutibilidade do método padrão.

O uso dos métodos de seleção de variáveis (APS e AG) forneceu bons resultados com a vantagem de utilizar um método matemático mais simples (MLR), com gráficos (valores previstos vs valores reais) que apresentam coeficientes lineares (intercepto) e angulares melhores que os gráficos PLS. Além disso, a seleção de variáveis pode minimizar o efeito das sobreposições de picos e de não linearidade dos dados, proporcionando a construção de modelos mais robustos. O método APS, na maioria dos casos, selecionou um número menor de

variáveis quando comparado ao método AG, com a vantagem de ser de natureza determinística e selecionar números de onda que contenham o máximo de informações.

A técnica empregada (FTIR-ATR) com o acessório circle cell[®] mostrou-se bastante adequada, pois possibilitou a análise direta do combustível, sem a necessidade de diluição de amostras. Adicionalmente, a circle cell[®] é de fácil amostragem e limpeza. Este acessório também pode ser adaptado para análises “on-line”.

Através do trabalho realizado, foi possível aplicar a análise espectroscópica no infravermelho médio, no estudo das propriedades de degradação do PET, minimizando as dificuldades impostas pelos métodos de referência, que são laboriosos e geram resíduos tóxicos.

Foi aplicado o método PLS para previsão de CEG e $[\eta]$ com boa correlação com os métodos de referência. Este método é bastante aplicável, desde que sejam tomados devidos cuidados com a seleção das regiões espectrais. O procedimento requer um pequeno preparo de amostras (moagem) com pouca manipulação dos dados. No entanto foi verificado que o método produz melhores resultados para concentrações de CEG mais altas.

O uso de um detector menos sensível (DTGS) não compromete a eficiência da calibração. A etapa de secagem não influenciou os resultados obtidos para $[\eta]$, indicando que esta propriedade pode ser analisada de maneira direta e simples. Este tipo de análise requer pouco investimento e pode ser adaptada com facilidade em análises de rotina.

Como continuidade deste trabalho, pode-se citar algumas propostas como a aplicação de seleção de amostras na construção de modelos de previsão para as propriedades dos combustíveis e adaptação do acessório circle-cell para as análises

on-line. A metodologia também poderia ser aplicada na avaliação das propriedades de vários tipos de óleos lubrificantes e outros combustíveis.

Propriedades de degradação de outros tipos de polímeros como PVC ou polietileno também poderiam ser avaliadas pela metodologia proposta neste trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] SCOTTER, C.N.G.; Non-destructive Spectroscopic Techniques for the Measurement of Food Quality, *Trends in Food Science & Technology*, **8**, p.285-292, 1997.
- [2] COATES, J.; Vibrational Spectroscopy: Instrumentation for Infrared and Raman Spectroscopy, *Applied Spectroscopy Reviews*, **33**, p. 267- 425, 1998.
- [3] MATERNS, H.; NAES, T. *Multivariate Calibration*. New York, John Wiley & Sons, p.419, 1989.
- [4] FERREIRA, M.C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, P. L. O.; Quimiometria I: Calibração Multivariada, um Tutorial, *Química Nova*, **22**, p.724-731, 1999.
- [5] Combustíveis de Qualidade, uma busca constante, *Ciência Hoje*, **28**, p.1-7, 2000.
- [6] SAMMARCO, C.; DELFINI, L.; Mercado Brasileiro tem Potencial para Ampliar a Reciclagem de Plásticos, *Plástico Industrial*, **3**, p. 106-119, 1999.
- [7] HARRIS, D.C.; BERTOLUCCI, M.D. *Symmetry and Spectroscopy, an Introduction to Vibrational and Eletronic Spectroscopy*. New York, Dover Publications, 1978.
- [8] MIRABELLA, F.M.; Internal Reflection Spectroscopy, *Applied Spectroscopy*, **21**, p.45-178, 1985.
- [9] GRDADOLNIK,J.; An Attenuated Total Reflection Infrared Spectroscopy of Water Solutions, *The Internet Journal of Vibration Spectroscopy [www.ijvs.com]*, **6**, p.12, 2002.
- [10] GUARRIGUES, S.; RAMBLA, F. J.; GUARDIA, M. L.; Comparative Study of Reflectance Cells for PLS- Determination of Sugars in Soft Drinks, *Fresenius Journal of Analytical Chemistry*, **362**, p.137-140, 1998.

- [11] GARCIA, F.X.; LIMA, L.; MEDINA, J.C.; Determination of Methanol and Methyl tert-Butyl Ether in Gasoline by Infrared Spectroscopy Using the Circle[®] cell and Multivariate Calibration, *Applied Spectroscopy*, **47**, p.1036-1039, 1993.
- [12] HEISE, H.M.; VOIGT, G.; LAMPEN, P.; KÜPPER, L.; RUDOLOF, S.; WERNER, G.; Multivariate Calibration for the determination of Analytes in Urine Using Mid-Infrared Attenuated Total Reflection Spectroscopy, *Applied Spectroscopy*, **55**, p.434-443, 2001.
- [13] BERTIE, J.E.; ZANG, S. L.; MANJI, R.; Infrared Intensities of Liquids X: Accuracy of Current Methods of Obtaining Optical Constants from Multiple Attenuated Total Reflection Measurements Using the Circle^{*} Cell, *Applied Spectroscopy*, **46**, p.1660-1665, 1992.
- [14] WETZEL, D.L.; Near-infrared Reflectance Analysis, Sleeper Among Spectroscopy Techniques, *Analytical Chemistry*, **55**, p. 1165-A- 1176A, 1983.
- [15] TEVRUCHT, M.L.E.; GRIFFITHS, P.R.; **Quantitative Investigation of Matrices for Diffuse Reflectance Infrared Fourier –Transform Spectrometry**, *Talanta*, **38**, p.839-849, 1991.
- [16] OLINGER, J.M.; GRIFFITHS, P.R.; Effects of Sample Dilution and Particle-Size Morphology on Diffuse Reflection Spectra of Carbohydrate Systems in the Near-Infrared and Mid-Infrared. Par I: Single Analytes, *Applied Spectroscopy*, **47**, p.687-694, 1993.
- [17] WENDLANT, W. M.; HECHT, H.G.; Reflectance Spectroscopy, New York, Interscience, 1996.
- [18] GNANASAMBANDAM, R.; PROCTOR, A.; Determination of Pectin Degree of Sterification by Diffuse Reflectance Fourier Transform Infrared Spectroscopy, *Food Chemistry*, **68**, p.327-332, 2000.
- [19] GHAUCH, A.; RIMA, J.; CHAREF, A.; SUPTIL, J.; FACHINGER, C.; MARTIN-BOUYER, M.; Quantitative Measurements of Ammonium, Hydrogenophosphate and Cu(II) by Diffuse Reflectance Spectrometry, *Talanta*, **48**, p.385-392, 1999.

- [20] GHAUCH, A. Use of Diffuse Reflectance Spectrometry in Spot Test Reaction for Quantitative Determination of Cations in Water, *Chemosphere*, **40**, p.1327-1333, 2000.
- [21] FERRAZ, A.; BAEZA, J.; RODRIGUEZ, J.; FEER, J.; Estimating The Chemical Composition of Biodegraded Pine and Eucalyptus Wood by DRIFT Spectroscopy and Multivariate Analysis, *Bioresource Technology*, **74**, p.201-212, 2000.
- [22] PARK, K.S.; LEE, H.; JUN, C.H.; PARK, K.H.; JUNG, J.W.; KIM, S.B.; Rapid Determination of Feo Content in Sinter Ores Using DRIFT Spectra and Multivariate Calibrations, *Chemometrics and Intelligent Laboratory Systems*, **51**, p.163-173, 2000.
- [23] MASSERSCHMIDT, I; CUELBAS, C. J.; POPPI, R.J.; ANDRADE, J.C.; ABREU, C.A.; DAVANZO, C.U.; Determination of Organic Matter in Soils by FTIR/Diffuse Reflectance and Multivariate Calibration, *Journal of Chemometrics*, **13**, p.265-273, 1999.
- [24] McCLELLAND, J.F.; Photoacoustic Spectroscopy, *Analytical Chemistry*, **55**, p.89A-105A, 1983.
- [25] COLEMAN, P.B.; Pratical Sampling Techniques for Infrared Analysis. Ed. CRC Press, 1993.
- [26] ROSENCWAIG, A.; GERSHO, A.; Theory of Photoacoustic Effect with Solids *Journal of Applied Physics*, **47**, p.64-69, 1976.
- [27] ROCKLEY, N.L.; WOODARD, M.K.; ROCKLEY M.G.; The effect of Particule Size on FT-IR-PAS Spectra, *Applied Spectroscopy*, **38**, p.329-334, 1984.
- [28] KINNEY, J.B.; STALEY, R.H.; Photo-Acoustic Cell For Fourier-Transform Infrared Spectrometry Of Surface Species, *Analytical Chemistry*, **55**, p.343-348, 1983.
- [29] VIDRINE, D.W.; Photoacoustic Fourier Transform Infrared Spectroscopy of Solid Samples, *Applied Spectroscopy*, **34**, p.314-319, 1980.
- [30] PEKEL, N.; GÜVEN, O.; Fourier Transform Infrared- Photoacoustic Spectroscopy of Poly (N-butyl methacrylate) Adsorbed from solution on Alumina, *Journal of Applied Polymer Science*, **69**, p.1669-1674, 1998.

- [31] DALEN, G.; Protein in Cloths: Evaluations of Analytical Techniques, *Applied Spectroscopy*, **54**, p.1350-1356, 2000.
- [32] NIU, B.; URBAN, M. W.; Recent Advances in Stratification and Film Formation of Latex Films; Attenuated Total Reflection and Step-Scan Photoacoustic FTIR Spectroscopy Studies, *Journal Applied Polymer Science*, **70**, p.1321-1348, 1997.
- [33] LUDWIG B.W.; URBAN, M. W.; Rheo-photoacoustic FTIR and morphology studies of molecular weight-strain dependence in poly(vinylidene), *Polymer*, **39**, p.5899-5912, 1999.
- [34] <http://www.petropar.com/>, consultado em 20/03/2002
- [35] <http://anp.gov.br>, consultado em 16/05/2002
- [36] DAUZACKER, E. R.; PALOMBO, E.; O Impacto das Futuras Especificações na Formulação da Gasolina Brasileira, *Boletim Técnico Petrobrás*, **46**, p.256 -269, 2003.
- [37] Norma Brasileira Método NBR 7148 . Petróleo e produtos de petróleo – Determinação da massa específica, densidade relativa e °API- Método do densímetro.
- [38] Norma Brasileira Método NBR 14065. Destilados de Petróleo e óleos viscosos- Determinação da massa específica e da densidade relativa pelo densímetro digital.
- [39] American Society for Testing and Materials (ASTM), Method D 1298. Density, Relative Density (Specific Gravity), of API Gravity of Crude Petroleum and Liquid Petroleum Products by Hydrometer Method.
- [40] American Society for Testing and Materials (ASTM) Method D 4052 . Density and Relative Density of Liquids by Digital Density Meter.
- [41] Norma Brasileira Método NBR 9619 Produtos de Petróleo - Determinação das Propriedades de Destilação.
- [42] American Society for Testing and Materials (ASTM) Method D 86 Distillation of Petroleum Products.

- [43] Método Brasileiro MB 457 Combustível – Determinação das características antidetonantes- índice de octano- Método Motor.
- [44] American Society for Testing and Materials (ASTM) Method D 2700 Knock Characteristics of Motor and Aviation Fuels by the Motor Method.
- [45] American Society for Testing and Materials (ASTM) Method D 2699 Knock Characteristics of Motor and Aviation Fuels by the Research Method.
- [46] Norma Brasileira Método NBR 14478 Gasolina – Determinação da estabilidade à oxidação pelo método do período de indução.
- [47] American Society for Testing and Materials (ASTM) Method D525 Oxidation Stability of Gasoline (Induction Period Method)
- [48] Norma Brasileira Método NBR 13992 Gasolina Automotiva – Determinação do Teor Alcoólico de álcool etílico anidro combustível (AEAC)
- [49] American Society for Testing and Materials (ASTM) Method D 3606 Benzene and Toluene in Finished Motor an Aviation Gasoline by Gas Chromatography
- [50] American Society for Testing and Materials (ASTM) Method D 5443 Paraffin, Naphthene and Aromatic Hydrocarbon Type Analysis in Petroleum Distillates through 200°C by Multi-Dimensional Gas Chromatography.
- [51] American Society for Testing and Materials (ASTM) Method D 6277 Determination of Benzene in Spark-Ignition Engine Fuels Using Mid-Infrared Spectroscopy.
- [52] Norma Brasileira Método NBR 424 Produtos Líquidos de Petróleo- determinação dos Tipos de Hidrocarbonetos pelo Indicador de Absorção por Fluorescência
- [53] American Society for Testing and Materials (ASTM) Method D 1319 Hydrocarbons Types in Liquid Petroleum Products by Fluorescence Indicator Adsorption.
- [54] <http://www.demec.ufmg.br>, consultado em 14/05/2003.
- [55] American Society for Testing and Materials (ASTM) Method D 4737-90 Standard Test Method for Calculated Cetane Index by Four Variable Equation.

- [56] Norma Brasileira (ABNT) Método NBR 14533 Produtos de Petróleo – Determinação do enxofre por espectrometria de fluorescência de Raios (Energia Dispersiva)
- [57] American Society for Testing and Materials (ASTM) Method D1552 Sulphur in Petroleum Products (High- Temperature Method)
- [58] American Society for Testing and Materials (ASTM) Method D 2622 Sulphur in Petroleum Products by X-Ray Spectrometry
- [59] American Society for Testing and Materials (ASTM) Method D 4294 Sulphur in Petroleum Products by X-Ray Fluorescence Spectroscopy
- [60] Norma Brasileira (ABNT) Método NBR 10441- Produtos de petróleo- Líquidos, transparentes e Opacos- Determinação da viscosidade cinemática e cálculos da viscosidade dinâmica.
- [61] American Society for Testing and Materials (ASTM) Method D 445 Kinematic Viscosity of Transparent and Opaque Liquids (and the Calculation of Dynamic Viscosity).
- [62] <http://www.ipiranga.com.br>, consultado em 24/08/2003.
- [63] Norma Brasileira (ABNT) Método NBR 5992 – Determinação da Massa Específica e do Teor Alcoólico do álcool Etílico e suas misturas com água .
- [64] Norma Brasileira (ABNT) Método NBR 10891- Álcool Etílico- Determinação do pH.
- [65] Norma Brasileira (ABNT) Método NBR 10547- Álcool Etílico- Determinação da Condutividade Elétrica.
- [66] Society for Testing and Materials (ASTM) Method D1125 Electrical Conductivity and Resistivity of water.
- [67] ANDRADE, J.M.; MUNIATEGUI, S.; LÓPEZ, P.; PRADA, D. Costs, Laboratory Safety, Productivity and Faster Research Octane Number and Motor Octane Number Determinations in Industrial Chemistry Laboratories, *Analyst*, 120, p. 249-253, 1995.

- [68] ANDRADE, J.M.; GARCIA, M.V.; LOPEZ-MAHIA, P.; PRADA, D. A Review of the Main Factors Influencing the FT-IR-PLS Abilities Exemplified with Petrochemical Qualimetric Applications, *Talanta*, **44**, 2167-2184.
- [69] KELLY, J.J.; CALLIS, J.B. Nondestructive Analytical Procedure for Simultaneous Estimation of the Major Classes of Hydrocarbon Constituents of Finished Gasolines, *Analytical Chemistry*, **62**, p.1444-1451, 1990.
- [70] KELLY, J.J.; BARLOW, C.H.; JINGUJI, T.M.; CALLIS J.B.; Prediction of Gasoline Numbers from Near-Infrared Spectral Feature in the Range 660-1215 nm, *Analytical Chemistry*, **61**, p.313-320, 1989.
- [71] KALIVAS, J.H.; Two data sets of near infrared spectra, *Chemometrics and Intelligent Laboratory Systems*, **37**, p. 255-259, 1997.
- [72] FLECHER, P.E.; WELCH, W.T.; ALBIN, S.; COOPER, J.B.; Determination of octane numbers and Reid Vapor Pressure in Commercial Gasoline Using Dispersive Fiber-Optic Raman Spectroscopy, *Spectrochimica Acta Part A*, **53**, p.199-206, 1997.
- [73] BARZILAI, I.L.; SELA, I.; BULATOV, V.; ZILBERMAN, I.; SCHECHTER, I. On-line Remote Prediction, Properties by Combined Optical Methods, *Analytica Chimica Acta*, **339**, p.193-199, 1997.
- [74] COOPER, J.B.; WISE, K.L.; WELCH, W.; SUMNER, M.B.; WILT, B.K.; BLEDSOE, R.R.; Comparison of Near-IR, Raman and MID-IR Spectroscopies for the Determination of BTEX in Petroleum Fuels, *Applied Spectroscopy*, **51**, p.1613-1620, 1997.
- [75] DIEHL, J.W.; FINKBEINER, J.W.; DISANZO, F.P.; Determination of Benzene, Toluene, Ethylbenzene, and Xylenes in Gasolines by Gas Chromatography/Deuterium Isotope Dilution Fourier Transform Infrared Spectroscopy, *Analytical Chemistry*, **65**, p.2493-2496, 1993.
- [76] ANREUS, E.L.; GARRIGUES, S.; LA GUARDIA, M.; Vapour generation-Fourier transform infrared spectrometric determination of benzene, toluene and methyl tert.-butyl ether in gasolines, *Analytica Chimica Acta*, **333**, p.157-165, 1996.
- [77] IOB, A.; ALI, M.A.; TAWABINI, B.S.; ABBAS, N.M.; Hydrocarbon group (PONA) analysis of reformat by FT-IR spectroscopy, *Fuel*, **75**, p.1060-1064, 1996.

- [78] IOB, A.; BUENAFE, R.; ABBAS, N.M. ; Determination of oxygenates in gasoline by FTIR, *Fuel*, **77**, p.1861-1864, 1998.
- [79] ICHIKAWA, M.; NONAKA, N.; TAKADA, I.; ISHIMORI, S.; Estimation of the Octane Number of Automobile Gasoline by Fourier Transform Infrared Absorption Spectrometry, *Applied Spectroscopy*, **46**, p.966-971, 1992.
- [80] IOB, A.; ALI, M. A.; TAWABINI, B. S.; ANABTAWI, J. A.; ALI, S.A.; AL-FARAYEDHI, A.; Prediction of Reformate Research Octane Number by FT-IR Spectroscopy, *Fuel*, **74**, p.227231, 1995.
- [81] ANDRADE, J.M.; SÁNCHEZ, M.S.; SARABIA, L.A.; Applicability of high-absorbance MIR Spectroscopy in Industrial Quality Control of Reformed Gasolines, *Chemometrics and Intelligent Laboratory Systems*, **46**, p.41-55, 1999.
- [82] FODOR, G. E.; KOHL, K.B. MASON, R.L.; Analysis of Gasolines by FT-IR Spectroscopy, *Analytical Chemistry*, **68**, p.23-30, 1996.
- [83] FODOR, G. E.; KOHL, K.B.; Analysis of Middle Distillate Fuels by Midband Infrared Spectroscopy, *Energy & Fuels*, **7**, p.598-601, 1993.
- [84] FODOR, G. E; MASON, R.A.; HUTZLER, S.A.; Estimation of Middle Distillate Fuel Properties by FT-IR, *Applied Spectroscopy*, **53**, p.1293-1298, 1999.
- [85] BREITKREITZ, M.C.; RAIMUNDO, Jr, I.M.; ROHWEDDER, J.J.R.; PASQUINI, C.; DANTAS FILHO, H.A.; JOSÉ, G.E.; ARAÚJO, M.C.U.; Determination of total sulfur in diesel fuel employing NIR spectroscopy and multivariate calibration, *Analyst*, **128**, p.1204-1207. 2003.
- [86] GUCHARDI, R.; PASQUINI, C.; Evaluation of a Dual Beam Near-infrared Spectroscopy Based on Acoustic-optic Tunable Filters, *Applied Spectroscopy*, **55**, p.454-457, 2001.
- [87] SILVA, H.E.B.; PASQUINI, C.; Dual-Beam Near-Infrared Hadamark Spectrophotometer, *Applied Spectroscopy*, **55**, p.715- 714, 2001.
- [88] MENDES, L.S.; OLIVEIRA, F.C.C.; SUAREZ, P.A.Z.; RUBIM, J.C. Determination of Ethanol in Fuel Ethanol and Beverages by Fourier Transform (FT)- Near Infrared and FT-Raman Spectroscopies, *Analytica Chimica Acta* , **493**, p.219-231, 2003.

- [89] 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. *Química Nova*, **27**, p.218-225, 2004.
- [90] SPINACÉ, M.A.S. Poli(tereftalato de etileno): reprocessamento por extrusão e metodologias de caracterização. Tese de Doutorado. Universidade Estadual de Campinas, 2000.
- [91] ASTM American Society for Testing and Materials. Designation D 4603-01, Standard Test Method for Determining Inherent Viscosity of Poly(ethylene terephthalate) (PET) by glass capillary viscosimeter.
- [92] POHL, H.A.; Determination Of Carboxyl End Groups in a Polyester, Polyethylene Terephthalate, *Analytical Chemistry*, **26**, p. 1614. 1954.
- [93] PAPINI, M.; Analysis Of The Reflectance Of Polymers in the near and Mid-Infrared Regions, *Journal of Quantitative Spectroscopy & Radiative Transfer*, **57**, p.265-274, 1997.
- [94] SAMMON, C.; YARWOOD, J.; EVERALL, N.; An FT-IR Study of the effect of hydrolytic on the Structure of Thin PET Films, *Polymer Degradation and Stability*, **67**, p.149-158, 2000.
- [95] HOLLAND, B.J.; HAY, J.N.; The Thermal Degradation of PET and Analogous Polyesters Measured by Thermal Analysis- Fourier Transform Infrared Spectroscopy, *Polymer*, **43**, p.1835-1846, 2002.
- [96] PATTERSON, D.; WARD, I.M.; The Assignment of the Carbonyl and Hydroxyl Absorptions in the Infra-red Spectrum of Polyethylene Terephthalate, *Transactions Of The Faraday Society*, **57**, p.291-294, 1957.
- [97] ZHANG, H.; RANKIN, A.; WARD, I.M.; Determination of the End-Group Concentration and Molecular Weight of Poly(ethylene naphthalene-2,6-dicarboxylate) Using Infra-red Spectroscopy, *Polymer*, **37**, p. 1079-1085. 1996.
- [98] VANUUM, M.P.B.; LAMMERS, H.; DEKLEIJN, J.P.; Process Analysis: Properties of Poly(Ethylene Terephthalate) Measured By Near Infrared Spectroscopy, 1. At-Line Analysis of Poly(Ethylene Terephthalate) Chips, *Macromolecular Chemistry and Physics*, **196**, p.2023-2028, 1995.

- [99] VANUUM, M.P.B.; LAMMERS, H.; DEKLEIJN, J.P.; Process Analysis: Properties of Poly(ethylene terephthalate) Measured by Near Infrared Spectroscopy, 2. At-line Analysis of Poly(ethylene terephthalate) Melt, *Macromolecular Chemistry and Physics*, **196**, p.2029-2034, 1995.
- [100] AL-ABDULRAZZAK, S.; LOFGREN, E.A.; JABARIN, S.A.; End-Group Determination in Poly(Ethyleneterephthalate) by Infrared Spectroscopy, *Polymer International*, **51**, p. 174-182, 2002.
- [101] KRISHNAN, K. Some Applications of Fourier Transform Infrared Photoacoustic Spectroscopy, *Applied Spectroscopy*, **35**, p.549-557, 1981.
- [102] JASSE, B.; Fourier-Transform Infrared Photoacoustic Spectroscopy of Synthetic Polymers, *Journal of Macromolecular Science-Chemistry*, **A26(1)**, p.43-67, 1989.
- [103] BARBU, A.; BRATU, I.; Structural Investigations of UV-Irradiated Packaging Polymeric Foils, *Journal of Molecular Structure*, **410-411**, p.229-231, 1997.
- [104] QINGDE, S.; JUNJUN, M.; GUIWEN, Z.; MAOSEN, Z.; A Study on the Structure of Degraded PVC by Photoacoustic Spectroscopy, *Journal of Molecular Structure*, **403**, p.231-236, 1997.
- [105] EVORA, M.C.; GONÇALEZ, O.L.; DUTRA, R.C.L., DINIZ, M.F.; WIEBECK, H.; SILVA, L.G.A.; Comparação de Técnicas FTIR de Transmissão, Reflexão e Fotoacústica na Análise de Poliamida-6, Reciclada e Irradiada, *Polímeros: Ciência e Tecnologia*, **12**, p.60-68, 2002.
- [106] THOMAS, E.V.; A Primer on Multivariate Calibration, *Analytical Chemistry*, **66**, p.775A-804A, 1994.
- [107] ADAMS, M.J.; Chemometrics in Analytical Spectroscopy, Cambridge, The Royal Society of Chemistry, 1995.
- [108] SAVITZKY, A.; GOLAY, M.J.E.; Smoothing and Differentiation of Data by Simplified Least Square Producers, *Analytical Chemistry*, **36**, p.1627-1639, 1964.
- [109] WOLD, S.; ESBENSEN, K.; GEALDI, P.; Principal Component Analysis, *Chemometrics and Intelligent Laboratory Systems*, **2**, p. 37-52, 1987.

- [110] OTTO, M.; Chemometrics, Statistics and Computer Application in Analytical Chemistry, New York, Wiley-Vch, 1999.
- [111] GELADI, P.; KOWALSKI, B.R.; Partial Least -Squares Regression: A Tutorial, *Analytica Chimica Acta*, **8**, p.1- 17, 1986.
- [112] HAALAND, D.M.; THOMAS, E.V.; Partial Least -Squares Methods for Spectral Analyses. 1. Relation to Other Quantitative Calibration Methods and the Extraction of Qualitative Information, *Analytical Chemistry*, **60**, p.1193-1202, 1988.
- [113] JIANG, J.; BERRY, R.J.; SIESTER, H.W.; OZAKI, Y.; Wavelength Interval Selection In Multicomponent Spectral Analysis by Moving Window Partial Least-Squares Regression with Applications to Mid-Infrared And Near-Infrared Spectroscopic Data, *Analytical Chemistry*, **74**, p. 3555, 2002.
- [114] COSTA, P.A.F.; POPPI, R.J.; Algoritmo Genético em Química, *Química Nova*, **22**, p. 405-411, 1999.
- [115] LUCASIUS, C.B.; KATEMAN, G.; Understanding and Using Genetic Algorithms Part 1. Concepts, Properties and Context, *Chemometrics and Intelligent Laboratory Systems*, **19**, p.1-33, 1993.
- [116] LUCASIUS, C.B.; KATEMAN, G.; Genetic Algorithms for Large-Scale Optimization in Chemometrics - An Application, *Trends in Analytical Chemistry*, **10**, p.254-261, 1991.
- [117] RIMBAUD, J.D.; MASSART, D.L.; LEARDI, R.; NOORD, O.E. De; Genetic Algorithms as a Tool for Wavelength Selection in Multivariate Calibration, *Analytical Chemistry*, **67**, p. 4295-4301, 1995.
- [118] COSTA, P.A.F.; POPPI, R. J.; Aplicação de Algoritmos Genéticos na Seleção de Variáveis em Espectroscopia no Infravermelho Médio. Determinação Simultânea de Glicose, Maltose e Frutose, *Química Nova*, **25**, p.46-52, 2002.
- [119] GOICOECHEA, H.C.; OLIVIERI, A.C.; Wavelength Selection for Multivariate Calibration Using a Genetic Algorithm: A Novel Initialization Strategy, *Journal Of Chemical Information And Computer Sciences*, **42**, p.1146-1153, 2002.

- [120] ARAÚJO, M.C.U.; SALDANHA, T.C.B.; GALVÃO, R.K.H.; YONEYAMA, T.; CHAME, H.C.; VISANI, V.; The Successive Projections Algorithm for Variable Selection in Spectroscopic Multicomponent Analysis, *Chemometrics and Intelligent Laboratory Systems*, **57**, p.65-73, 2001.
- [121] GALVÃO, R.K.H.; PIMENTEL, M.F.; ARAÚJO, M.C.U.; YONEYAMA, T.; VISANI, V.; Aspects of the Successive Projections Algorithm for Variable Selection in Multivariate Calibration Applied to Plasma Emission Spectrometry, *Analytica Chimica Acta*, **443**, p.107-115, 2001.
- [122] DANTAS, H.A. F.; GALVÃO, R.K.H.; ARAÚJO, M.C.U.; SILVA, E.C.; SALDANHA, T.C.B; JOSÉ, G.E.; PASQUINI, C.; RAIMUNDO, I.M. JR; ROHWEDDER, J.J.R.; A Strategy for Selecting Calibrations Samples for Multivariate Modeling, *Chemometrics and Intelligent Laboratory Systems*, **72**, p.83-91, 2004.
- [123] YANG, H.; RING, Z.; BRIKER, Y.; McLEAN, N.; FRIESEN, W.; FAIRBRIDGE, C.; Neural Network Prediction of the Cetane Number and Density of Diesel Fuel from its Chemical Composition Determined by LC and GC-MS, *Fuel*, **81**, p.65-74, 2002.
- [124] BARROS NETO, B.de; SCARMÍNIO,I.S.; BRUNS, R.E.; Como fazer experimentos : pesquisa e desenvolvimento na ciência e na indústria, Campinas, Editora da Unicamp, 2002.

APÊNDICE A

- Exemplo prático da aplicação do Algoritmo NIPALS

A determinação das componentes principais utilizando-se uma matriz de dados com base no algoritmo NIPALS é realizada da seguinte forma:

$$\mathbf{X} = \begin{bmatrix} 2 & 1 \\ 3 & 2 \\ 4 & 3 \end{bmatrix}$$

Passo 0. Escala-se na média e normaliza-se o comprimento para 1:

$$\mathbf{X} = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Passo 1. Estima-se o vetor *loading* $\mathbf{p}^T$. Usualmente usa-se a primeira linha de $\mathbf{X}$:

$$\mathbf{p}^T = (-1/\sqrt{2} \quad -1/\sqrt{2})$$

Passo 2. Computa-se um novo vetor score $\mathbf{t}$:

$$\mathbf{t} = \mathbf{X} \mathbf{p} = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Compara-se os vetores $\mathbf{t}$ novos e velho. Caso sejam um mesmo (com um certo erro), o processo pode avançar até o passo 5, pois houve convergência senão segue-se o passo 3.

Passo 3. Computa-se novo vetor $\mathbf{p}^T$

$$\mathbf{pT} = \mathbf{t}^T \mathbf{X} = (1 \ 0 \ 1) \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = (-2/\sqrt{2} \quad -2/\sqrt{2})$$

Normaliza-se o comprimento do vetor loading para 1:

$$\mathbf{p}^T = \frac{\mathbf{p}^T}{\|\mathbf{p}^T\|} = (-1/\sqrt{2} \quad -1/\sqrt{2})$$

Passo 4. Repetir o passo 2 se o número de iterações não exceder o valor pré-definido.

Passo 5. Determina-se a matriz de resíduos:

$$\mathbf{E} = \mathbf{X} - \mathbf{t}\mathbf{p}^T = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 \\ -0 \\ -1 \end{bmatrix} (-1/\sqrt{2} \quad -1/\sqrt{2}) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Se o número de componentes principais ao número previamente fixado, pode-se avançar até o passo 7, senão segue-se até o passo 6:

Passo 6. Usar a matriz residual $\mathbf{E}$ como nova matriz $\mathbf{X}$ e calcular os componentes principais adicionais a partir do Passo 1.

Passo 7. Como resultado, a matriz $\mathbf{X}$ é representada por um modelo de componente principal, ou seja:

$$\mathbf{X} = \mathbf{T} \mathbf{P}^T = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} (-1/\sqrt{2} \quad -1/\sqrt{2})$$

APÊNDICE B

- Descrição do Algoritmo das Projeções Sucessivas

Como notação tem-se as letras maiúsculas em negrito para matrizes e operadores lineares; letras minúsculas em negrito para vetores coluna e letras em itálico para escalares; elementos de uma seqüência são escritos em itálico com um índice entre parênteses e o subscrito T indica transposto(^T); o símbolo (‘) indica valor previsto.

Os caracteres empregados na descrição do algoritmo são:

$k(n)$ significa o comprimento de onda selecionado a n (enésima) iteração de APS;

$\mathbf{X}_{cal}$ é a matriz de dimensão ($M_{cal} \times J$) das respostas instrumentais ou variáveis independentes para M_{cal} misturas de calibração e J comprimentos de onda;

$\mathbf{Y}_{cal}$ é a matriz de dimensão ($M_{cal} \times A$) das concentrações de calibração ou variáveis dependentes para A analitos;

$\mathbf{Y}_{test}$ é a matriz de dimensão ($M_{test} \times A$) das concentrações para M_{test} misturas.

O Algoritmo das Projeções Sucessivas tem início com a escolha aleatória de um comprimento de onda, ao qual é incorporado um novo a cada iteração até que o número de comprimento de ondas N seja alcançado. O propósito deste procedimento é escolher variáveis que contenham um mínimo de redundância, minimizando os problemas de colinearidade. Assumindo que N e um primeiro comprimento de onda $k(0)$ sejam conhecidos:

Passo 0. Antes da primeira iteração ($n = 1$), toma-se o vetor coluna $\mathbf{x}_j$ ($j = 1, \dots, J$) de $\mathbf{X}_{cal}$

Passo 1. Toma-se um conjunto de comprimentos de onda S que não foram selecionados ainda. Isto é: $S = \{j \text{ quando } 1 \leq j \leq J\}$ e $j \notin \{k(0), \dots, k(n-1)\}$

Passo 2. Calcula-se a projeção de $\mathbf{x}_j$ em um subespaço (plano) ortogonal a $\mathbf{x}_{k(n-1)}$ como:

$$\mathbf{P}\mathbf{x}_j = \mathbf{x}_j - (\mathbf{x}_j^T \mathbf{x}_{k(n-1)}) \mathbf{x}_{k(n-1)} (\mathbf{x}_{k(n-1)}^T \mathbf{x}_{k(n-1)})^{-1},$$

para todo $j \in S$

Onde $\mathbf{P}$ é o operador de projeção.

Passo 3. Considera-se $k(n) = \arg(\max \|\mathbf{P}\mathbf{x}_j\|, j \in S)$.

Passo 4. Considera-se $\mathbf{x}_j = \mathbf{P}\mathbf{x}_j, j \in S$.

Passo 5. Considera-se $n = n + 1$. Se $n < N$ volta-se ao passo 1.

Finalização: Os comprimentos de onda resultantes são $\{k(n); n=0, \dots, N-1\}$

Após a escolha dos comprimentos de onda resultantes, constrói-se um modelo de calibração MLR. Usa-se tal modelo para a previsão de um conjunto de validação. Calcula-se o RMSEP.