

**UFMG 704**

**D. 435**

**HELGA GABRIELA ALEME**

**UTILIZAÇÃO DE CURVAS DE DESTILAÇÃO E MÉTODOS QUIMIOMÉTRICOS NA  
PREVISÃO DA ORIGEM E NA DETERMINAÇÃO DO TEOR ALCOÓLICO E MASSA  
ESPECÍFICA DA GASOLINA AUTOMOTIVA**

**Dissertação apresentada ao Departamento de Química do  
Instituto de Ciências Exatas da Universidade Federal de Minas  
Gerais, como requisito parcial para obtenção do grau de  
Mestre em Química – Química Analítica.**

**Belo Horizonte**

**2008**

*“Mas, nada, desde cogumelos até as leis científicas,  
pode ser descoberto sem procurar e tentar.”*

Mendeleev, Dmitri

## **AGRADECIMENTOS**

- A Deus pela minha vida e por me permitir realizar tudo que meu coração anseia.
- Ao Prof. Paulo pela amizade, conhecimentos transmitidos, paciência e diversos “puxões de orelha” nos momentos oportunos.
- À Prof. Letícia pela amizade, conhecimentos transmitidos e por me ajudar nos momentos difíceis, com muita calma e sabedoria.
- Aos meus pais, Felipe e Elisabeth, que muitas vezes abdicaram dos seus sonhos para que eu pudesse realizar este sonho: ser mestre. Obrigada por todo apoio e incentivo em toda caminhada.
- Ao meu esposo, Mário, pela paciência, amor, carinho, compreensão e companheirismo.
- Aos funcionários e amigos do LEC-UFMG, em especial Júlio, Natália e Suzy, pelos ensaios executados.
- Ao Prof. Valmir por todo apoio na área de informática e pela amizade.
- À Camila e Itânia pelas discussões relacionadas à Quimiometria.
- Ao GEAPA-UFMG pelo acolhimento e carinho...além de toda amizade que tenho com vocês.
- À Sheila, amiga presente em todos os momentos, sempre apoiando e aconselhando.

- Ao Frank, pela parceria e amizade.
  
- Às amigas Cristiane, Dayse, Gisele, Lucilaine, Milene, Patrícia, Priscilla e Otávia pelas conversas, risadas e companheirismo nos momentos difíceis.
  
- A todos aqueles que de alguma maneira me incentivaram e ajudaram na realização deste trabalho.
  
- À CAPES pela bolsa de estudos a mim concedida durante a execução deste mestrado.
  
- À FAPEMIG e FINEP pelo apoio financeiro.
  
- Ao LEC-UFMG e ANP pelas amostras e ensaios concedidos.

## **RESUMO**

As técnicas quimiométricas *PCA* e *LDA* foram associadas às curvas de destilação (ASTM D86) na determinação da origem de gasolinas comercializadas no estado de Minas Gerais. A *PCA* reduziu o conjunto de dados (com setenta e uma variáveis) a um pequeno número de variáveis mantendo uma elevada variância explicada. Neste modelo, as três primeiras componentes principais representaram 87% do total da variabilidade. Pela primeira componente principal foi possível observar a separação das amostras da *REGAP* e *RLAM* enquanto que pela segunda componente visualizou-se a separação da *RLAM* das refinarias *REDUC*, *REPLAN* e *REVAP*. O modelo *LDA* foi construído utilizando a origem declarada pelos postos revendedores. Nesta técnica, as distâncias entre os grupos foram usadas para determinar a similaridade das amostras. As refinarias *REPLAN/REVAP* apresentaram a menor distância, enquanto que *REDUC/REGAP* a maior distância. Cerca de 80% do conjunto de dados, cuja origem não foi declarada, foram classificadas como pertencentes à *REGAP*, com 95% de acerto. A calibração multivariada *PLS* foi utilizada juntamente com curvas de destilação na determinação do teor alcoólico e da massa específica de gasolinas de diferentes refinarias. Nestas determinações foi realizada uma comparação entre os coeficientes de regressão de todos os algoritmos, após a seleção do número de variáveis latentes. Em ambas as determinações foram observadas, no gráfico das variáveis latentes, a separação do conjunto de amostras em duas faixas de teor alcoólico, mostrando a influência do teor alcoólico na curva de destilação. Os baixos valores de *RMSEC* e *RMSEP* obtidos, associados com a elevada precisão quando comparado às metodologias normalizadas (NBR-13992, ASTM D1298 e D4052) mostraram que a calibração multivariada associada às curvas de destilação foi eficiente para determinar o teor alcoólico e a massa específica em gasolinas de diferentes composições. O método proposto possui baixo custo, reduz o tempo de análise, é fácil de ser implementado e pode substituir a metodologia normalizada usada atualmente.

## ***ABSTRACT***

The chemometrics techniques PCA and LDA were associated with distillation curve (ASTM D86) in determining the origin of gasoline sold in Minas Gerais state. The PCA reduced the number of data (with seventy-one variable) to a small number of variables preserving a high variance explained. In this model, the first three principal components accounted for 87% of the total variability. For the first component was possible to observe the separation of the samples from REGAP and RLAM while the second component is responsible for the separation of RLAM of refineries REDUC, REPLAN and REVAP. The model was built using the LDA according to declared origin made by sealers. In this technique, the distances between the groups were used to determine the similarity of the samples. Refineries REPLAN / REVAP had the lowest distance, while REDUC / REGAP the biggest distance. About 80% of the set data, whose origin has not been declared, were classified as belonging to REGAP, with 95% for chance of correct classification. The multivariate calibration PLS was used with curves of distillation in determining the alcohol content and specific gravity of different gasoline origin. In these determinations, a comparison between the regression coefficients of all algorithms was made after selecting the number of latent variables. In both determinations, the separation of all the samples in two range of alcohol was seen in the graph of latent variables, showing the influence of alcohol content in the distillation curve. The low values of RMSEC and RMSEP obtained, associated with high accuracy when compared to standard methodologies (NBR-13992, ASTM D1298 and D4052) showed that the multivariate calibration associated with distillation curves were efficient to determine the alcohol content and specific gravity of different gasoline compositions. The proposed method is low cost, time reducing, easy to implement, and can be replaced by the usual standard methodology.

# SUMÁRIO

<b>Capítulo 1 – Introdução</b>	<b>1</b>
1.1 Gasolina Automotiva: aspectos gerais	1
1.1.1 Mistura Etanol-Gasolina	4
1.1.2 Qualidade da Gasolina Automotiva Brasileira	5
1.1.3 Não-conformidade e adulteração	7
1.1.4 Atipicidade e origem do combustível	12
1.2 Destilação	14
1.3 Quimiometria	24
1.3.1 Organização e pré-processamento dos dados	24
1.3.2 Análise de Componentes Principais (PCA )	25
1.3.3 Análise Discriminante Linear (LDA)	26
1.3.4. Mínimos Quadrados Parciais (PLS)	27
<b>Capítulo 2 – Objetivos</b>	<b>32</b>
<b>Capítulo 3 – Parte Experimental</b>	<b>33</b>
3.1 Amostras de Gasolina	33
3.2 Reagentes	33
3.3 Equipamentos	33
3.4 Materiais	34
3.5 Qualidade dos Resultados dos Ensaios	34
3.6 Procedimentos Experimentais	34
3.6.1 Ensaio da Destilação (ASTM D86)	34
3.6.2 Ensaio do Teor Alcoólico (NBR 13992)	35
3.6.3 Ensaio da Massa Específica (ASTM D4052 e D1298)	36
3.6.3.1 Ensaio Manual (ASTM D1298)	36
3.6.3.2 Ensaio Automático (ASTM D4052)	37
3.7 Procedimentos de Pré-Processamento e Normalização dos Dados	37
3.8 Testes Estatísticos F e t	38
3.9 Cálculo da Repetibilidade e Reprodutibilidade	39

<b>Capítulo 4 – Resultados e Discussão</b>	40
4.1 Determinação da origem da gasolina comercializada no estado de Minas Gerais	40
4.1.1 Aplicação da PCA	40
4.1.2 Aplicação da LDA	45
4.1.2.1 Classificação de Gasolinas com Origem não Declarada	47
4.2 Determinação do Teor Alcoólico da Gasolina	48
4.2.1 Validação do Método Proposto	62
4.2.2 Verificação da proximidade do comportamento das amostras do conjunto <i>GDE</i> e <i>GPM</i> na determinação do teor alcoólico	63
4.2.3 Avaliação da Repetibilidade e Reprodutibilidade	64
4.3 Determinação da Massa Específica	66
4.3.1 Validação do Método Proposto	73
4.3.2 Avaliação da repetibilidade e reprodutibilidade do método Proposto	74
<b>Capítulo 5 – Conclusões</b>	75
<b>Capítulo 6 – Referências Bibliográficas</b>	77

## ***LISTA DE ABREVIATURAS***

ABNT: Associação Brasileira de Normas Técnicas

ANN: Artificial Neural Network

ANP: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis

ASTM: American Society for Testing and Materials

CENAL: Comissão Executiva Nacional do Álcool

CIDE: Contribuição de Intervenção no Domínio Econômico

CNAL: Conselho Nacional do Álcool

COFINS: Contribuição para o Financiamento da Seguridade Social

FAFEN: Fabrica de Fertilizantes Nitrogenados

FBP: Final Boiling Point

$F_{\text{calc}}$ : F calculado

$F_{\text{tab}}$ : F tabelado

FT-IR: Fourier Transform Infrared Spectroscopy

GDE: Gasolina Dopada com Etanol

GPM: Gasolina Programa de Monitoramento

HCA: Hierarchical Cluster Analysis

IAD: Índice Auto Detonante

IBP: Initial Boiling Point

ICMS: Imposto sobre Circulação de Mercadorias e Prestação de Serviços

LDA: Linear Discriminant Analysis

LEC-UFGM: Laboratório de Ensaio de Combustíveis

LUBNOR: Lubrificantes e Derivados de Petróleo do Nordeste

MLP: Multilayer Perceptron

MON: Motor Octane Number

NBR: Normas Brasileiras

PC: Principal Component

PCA: Principal Component Analysis

PIE: Ponto Inicial de Evaporação

PFE: Ponto Final de Evaporação

PIS: Programa de Integração Social

PLS: Partial Least Squares

PMQC: Programa de Monitoramento da Qualidade dos Combustíveis

PNN: Probabilistic Neural Network

PRESS: Predictive Residual Error Sum of Squares

PROÁLCOOL: Programa Nacional do Álcool

RBC: Rede Brasileira de Calibração

RBPC: Refinaria Presidente Bernardes

RECAP: Refinaria de Capuava

REDUC: Refinaria Duque de Caxias

REFAP: Refinaria Alberto Pasqualini

REGAP: Refinaria Gabriel Passos

REMAN: Refinaria Isac Sabbá

REPAR: Refinaria Presidente Getúlio Vargas

REPLAN: Refinaria de Paulínia

REVAP: Refinaria Henrique Lage

RLAM: Refinaria Lamdulpho Alves

RMPG: Rede Permanente para Geodésia

RMSE: Root Mean Squared Error

RMSEC: Root Mean Squared Error Calibration

RMSEP: Root Mean Squared Error Prediction

RON: Research Octane Number

RPBC: Refinaria Presidente Bernardes

RPI: Refinaria Ipiranga

SIMCA: Soft Independent Modeling of Class Analogy

SIX: Unidade de Industrialização de Xisto

SVM: Support Vector Machine

$t_{\text{calc}}$ : t calculado

$t_{\text{tab}}$ : t tabelado

VE: Variância Explicada

VL: Variável Latente

## **LISTA DE FIGURAS**

- Figura 1** – Refinarias brasileiras e sua localização por estado. 3
- Figura 2** - Volumes de gasolina C comercializados, pelas distribuidoras, em todo território brasileiro (azul) e em Minas Gerais (vermelho) de 2000 a 2007. 5
- Figura 3** - Evolução das não conformidades no Brasil (azul) e em Minas Gerais (vermelho) no período de 2000 a 2007. 8
- Figura 4** - Evolução das não conformidades por ensaio físico-químico em Minas Gerais no período de 2002 a 2007. 9
- Figura 5** – Esquema de composição de preço da gasolina comercializada no estado do Rio de Janeiro. 10
- Figura 6** – Evolução do percentual de amostras de gasolina que apresentaram a presença de marcador no período de 2002 a 2007 da região leste de Minas Gerais. 11
- Figura 7** – Evolução das amostras atípicas no período de 2000 a 2007. 12
- Figura 8** – Evolução do número de amostras com origem desconhecida no período de 2002 a 2007 no estado de Minas Gerais. 13
- Figura 9** - Modelo de destilador manual, segundo a ASTM D86. 16
- Figura 10** – Curva de destilação de gasolinas tipo A com e sem a adição de 25 %(v/v) de etanol anidro. 18
- Figura 11** - Equilíbrio líquido-vapor para sistemas binários de etanol e n-heptano a 1 atm. 19
- Figura 12** – Curvas de destilação de gasolinas (A e B) sem e com a adição e 10% v/v de etanol. 20
- Figura 13** – Curvas de destilação de gasolina e suas misturas com etanol numa proporção volumétrica de 4 e 7 %(v/v). 21

<b>Figura 14</b> – Comportamento da pressão de vapor de acordo com o volume de etanol adicionado na gasolina.	22
<b>Figura 15</b> - Número de variáveis latentes versus valores de <i>PRESS</i> para as dez primeiras variáveis latentes de um modelo PLS.	28
<b>Figura 16</b> - Gráfico de “Leverages” versus Resíduos Studentizados.	29
<b>Figura 17</b> - Gráfico do número de componentes principais versus autovalor	42
<b>Figura 18</b> - Gráfico dos escores das três primeiras componentes principais	42
<b>Figura 19</b> – Gráfico dos escores das duas primeiras componentes principais	43
<b>Figura 20</b> –Gráfico dos escores da primeira e terceira componentes principais	43
<b>Figura 21</b> - Gráfico dos pesos das duas primeiras componentes principais	44
<b>Figura 22</b> – Gráfico percentual recuperado <i>versus</i> temperatura corrigida para uma gasolina com diferentes teores alcoólicos.	48
<b>Figura 23</b> – Gráfico de número de variáveis latentes versus <i>PRESS</i> dos conjuntos <i>GDE</i> e <i>GPM</i> .	50
<b>Figura 24</b> - Gráfico dos pesos das duas primeiras variáveis latentes na determinação do teor alcoólico do conjunto <i>GDE</i> .	52
<b>Figura 25</b> - Gráfico dos pesos das duas primeiras componentes na determinação do teor alcoólico do conjunto <i>GPM</i> .	54
<b>Figura 26</b> – Gráfico das duas primeiras variáveis latentes mostrando a separação das amostras do conjunto <i>GDE</i> em duas faixas de teor alcoólico.	55
<b>Figura 27</b> - Gráfico das duas primeiras variáveis latentes mostrando a separação das amostras do conjunto <i>GPM</i> em duas faixas de teor alcoólico.	56

<b>Figura 28</b> - Valor Real versus Valor Previsto para os modelos normalizados ou não do grupo <i>GDE</i> .	58
<b>Figura 29</b> – Valor Real versus Previsto para os modelos normalizados ou não do grupo <i>GPM</i> .	59
<b>Figura 30</b> – Variação da massa específica para diferentes misturas etanol-gasolina.	66
<b>Figura 31</b> - Gráfico de número de variáveis latentes versus <i>PRESS</i> .	67
<b>Figura 32</b> - Gráfico dos pesos na determinação da massa específica para diferentes normalizações.	69
<b>Figura 33</b> – Gráfico das três primeiras variáveis latentes para a determinação da massa específica com a separação das amostras em dois grupos de teor alcoólico.	70
<b>Figura 34</b> - Valor Real versus Previsto para os modelos construídos com dados não normalizados e normalizados segundo <i>T-IBP</i> e <i>T/FBP</i> .	71

## LISTA DE TABELAS

<b>Tabela 1</b> – Ensaios, métodos e especificações da ANP para a avaliação da qualidade de gasolina.	6
<b>Tabela 2</b> – Parâmetros a serem observados na destilação de acordo com o tipo de amostra discriminado por cada grupo.	15
<b>Tabela 3</b> – Amostras de gasolina utilizadas na determinação da origem.	41
<b>Tabela 4</b> - Número de componentes principais, porcentagem de variância explicada (%VE) e porcentagem de variância explicada acumulada (%VE <sub>acum</sub> ).	41
<b>Tabela 5</b> - Sumário de classificação das amostras usando validação cruzada.	45
<b>Tabela 6</b> - Distância média entre os grupos no modelo LDA com validação cruzada.	46
<b>Tabela 7</b> - Classificação das amostras desconhecidas de acordo com o modelo LDA.	47
<b>Tabela 8</b> – Valores de porcentagem de variância explicada das dez primeiras variáveis latentes no conjunto de calibração na determinação do teor alcoólico nos conjuntos <i>GDE</i> e <i>GPM</i> .	51
<b>Tabela 9</b> - Valores de <i>RMSEC</i> , <i>RMSEP</i> obtidos para diferentes modelos construídos com diferentes conjuntos ( <i>GDE</i> e <i>GPM</i> ) na determinação do teor alcoólico na gasolina a partir das curvas de destilação.	60
<b>Tabela 10</b> - – Valores de $F_{calc}$ (nível de confiança de 95 %) para <i>RMSEC</i> e <i>RMSEP</i> nos conjuntos <i>GDE</i> e <i>GPM</i> , na determinação do teor alcoólico na gasolina a partir das curvas de destilação.	61
<b>Tabela 11</b> - Valores de $t_{calc}$ para a comparação dos resultados obtidos pelos conjuntos <i>GDE</i> e <i>GPM</i> , em relação à quantidade de etanol adicionada e à NBR 13992, respectivamente.	63
<b>Tabela 12</b> – Valores de <i>RMSEP</i> e $t_{calc}$ a 95% de confiança, obtidos para diferentes modelos construídos na determinação do teor alcoólico da gasolina com conjunto de calibração <i>GDE</i> e validação <i>GPM</i> .	64

**Tabela 13** – Valores de repetitividade e reprodutibilidade obtidos para a determinação do teor alcoólico de gasolinas automotivas analisadas pela NBR 13992 e com a metodologia proposta com e sem arredondamento dos resultados obtidos. 65

**Tabela 14** – Valores de porcentagem de variância explicada para as dez primeiras variáveis latentes. 68

**Tabela 15** - Valores de *RMSEC* e *RMSEP* obtidos para diferentes modelos construídos na determinação da massa específica da gasolina. 72

**Tabela 16** - Valores de  $F_{calc}$  para *RMSEC* e *RMSEP* na determinação da massa específica na gasolina a partir das curvas de destilação. 72

**Tabela 17** - Valores de  $t_{calc}$  para na comparação dos resultados obtidos pelo ensaio segundo a norma ASTM D4052 com a metodologia proposta neste trabalho 73

**Tabela 18** – Valores de repetitividade e reprodutibilidade, obtidos para a determinação da massa específica de gasolinas automotivas analisadas pela ASTM D4052, ASTM D1298 e com a metodologia proposta. 74

## 1. Introdução

### 1.1 Gasolina Automotiva: Aspectos Gerais

Gasolina é uma mistura de líquidos, voláteis e inflamáveis derivados de petróleo. Originalmente um subproduto da indústria do petróleo, a gasolina tornou-se o combustível preferido para automóveis devido à sua alta energia de combustão e pela sua capacidade de misturar-se facilmente com o ar em um carburador [1].

A gasolina é uma mistura de hidrocarbonetos que possuem de 4 a 12 átomos de carbono, com pontos de ebulição no intervalo de 30 a 225 °C [2]. Estes hidrocarbonetos são de três tipos: parafínicos (incluindo as cicloparafinas e os ramificados), olefinas e aromáticos. A concentração relativa destes compostos varia consideravelmente dependendo da fonte de petróleo e do processo de refino. As cadeias e comprimentos dos hidrocarbonetos, no intervalo de 4 a 12 átomos de carbono, podem variar com uma distribuição geral consistindo de 4 a 8 % de parafínicos, 2 a 5 % de olefinas, 25 a 40 % de parafinas ramificadas, 3 a 7 % de cicloparafinas e 20 a 50 % de aromáticos [1].

Os principais tipos de gasolina comercializados no Brasil são:

- *Gasolina A Comum*: é a gasolina produzida pelas refinarias de petróleo e entregue diretamente às companhias distribuidoras. Este produto é a base da gasolina disponível nos postos revendedores.
- *Gasolina A Premium*: é elaborada com o objetivo de atender aos veículos (nacionais e importados) de altas taxas de compressão e alto desempenho. Ela possui octanagem maior que a gasolina comum além de um baixo teor de enxofre, o que reduz as emissões de poluentes no meio ambiente. Octanagem é a resistência que a gasolina tem à auto-ignição, o que pode levar à detonação localizada, causando perda de potência e sérios danos ao motor, dependendo de sua intensidade e persistência [3].
- *Gasolina C Comum*: é a gasolina preparada pelas companhias distribuidoras que adicionam etanol anidro à gasolina do tipo A. Esta é a gasolina do tipo C mais simples, que não possuem aditivos em sua composição.
- *Gasolina C Premium*: é a gasolina preparada pelas companhias distribuidoras que adicionam etanol anidro à gasolina do tipo A Premium.
- *Gasolina C Aditivada*: é obtida a partir da Gasolina C Comum, ou Premium, com a adição de um aditivo do tipo detergente-dispersante. Estes aditivos têm a função de manter limpo o sistema de alimentação do combustível contribuindo para minimizar a formação de depósitos no carburador e nos bicos injetores, assim como no coletor e hastes das válvulas de admissão [4]. Esta gasolina recebe um corante que lhe confere uma cor distinta daquela apresentada pela gasolina comum.

Além destes tipos básicos de gasolina existem os seguintes tipos [4]:

- Gasolina Padrão: é especialmente produzida para o uso na indústria automobilística, nos ensaios de avaliação do consumo e das emissões de poluentes, como gases de escapamento e hidrocarbonetos dos veículos por ela produzidos.
- Gasolina de Aviação: apresenta propriedades, requisitos de desempenho e cuidados diferenciados das demais gasolinas para motores de combustão interna e é destinada a aviões de pequeno porte que possuem motores com ignição por centelha. É uma gasolina que apresenta alto índice de desempenho e octanagem, além de possuir em sua composição o aditivo chumbo tetraetila. A gasolina de aviação não deve ser usada em automóveis equipados com conversores catalíticos, pois estes equipamentos são atacados por esse composto a base de chumbo [3].

Até, e durante, a primeira década do século XX, a gasolina era originalmente produzida a partir do petróleo bruto, ou o que poderia ser condensado do gás natural. Entretanto, descobriu-se que frações pesadas de petróleo aquecidas a temperaturas mais altas, sofrem degradação térmica (ou craqueamento) produzindo pequenas moléculas que estavam dentro do intervalo adequado para a gasolina.

Em um primeiro momento, a gasolina obtida pelo craqueamento foi considerada como um produto inferior devido à sua instabilidade de armazenamento, mas como este produto tinha uma alta demanda no mercado, a indústria petrolífera procurou por processos que solucionassem este problema. Alguns métodos utilizados para melhorar a qualidade da gasolina e aumentar a sua oferta são: polimerização, alquilação, isomerização e reformas catalítica e térmica [1].

A polimerização é um processo que converte olefinas gasosas, como propeno e buteno, em moléculas maiores que se enquadram no intervalo da gasolina. Esta conversão pode ser realizada termicamente ou na presença de um catalisador a temperaturas mais baixas.

A alquilação é um processo que combina olefinas e parafinas (como o isobutano) formando isoparafinas de elevada octanagem. Embora a alquilação sem catalisadores seja possível, processos comerciais utilizam cloreto de alumínio, ácido sulfúrico ou fluorídrico como catalisadores quando as reações são feitas em baixas temperaturas, minimizando reações secundárias indesejáveis, tais como a polimerização de olefinas.

A isomerização é um processo que converte hidrocarbonetos de cadeia linear em hidrocarbonetos de cadeia ramificada. Esta conversão é realizada na presença de um catalisador (cloreto de alumínio ativado com ácido clorídrico) que é essencial para inibir reações laterais, como craqueamento e formação de olefinas.

Os processos de reforma térmica e catalítica consistem em se utilizar de calor ou catalisador para reorganizar a estrutura molecular dos hidrocarbonetos. A reforma térmica foi uma evolução natural dos processos mais antigos de craqueamento térmico. Neste processo, o

objetivo é converter óleos pesados em gasolinas de elevada octanagem por meio de aquecimento. A qualidade e quantidade da gasolina (conhecida também como reformada) são dependentes da temperatura. Neste processo quanto maior a temperatura, maior é o número de octanas e menor é o rendimento da gasolina reformada. Por ser menos eficaz e econômico quando comparado ao processo catalítico este tem sido largamente suplantado. O processo de reforma catalítica é feito na presença de um catalisador, como o cloreto de alumínio ativado com ácido clorídrico, sendo este essencial na inibição de reações laterais que podem acontecer durante o processo [1].

No Brasil existem atualmente dezesseis refinarias petrolíferas, como mostradas na Figura 1. Destas, somente treze refinarias são responsáveis pela produção de gasolina, pois as refinarias FAFEN-SE, FAFEN-BA são produtoras de fertilizantes (como amônia, uréia e ácido nítrico) para indústria petroquímica e a LUBNOR, que tem como principais produtos óleo lubrificante, asfalto e gás natural.



**Figura 1** – Refinarias brasileiras e sua localização por estado.

A região sudeste é responsável pelo maior número de refinarias visto que esta possui a maior frota de veículos automotivos.

### 1.1.1 Mistura Etanol-Gasolina

No Brasil, a partir de 1931, o uso de etanol anidro passou a ser obrigatório em toda gasolina comercializada no país, em substituição ao aditivo chumbo tetraetila, inicialmente em uma proporção de 5 %(v/v). Esta porcentagem variou ao longo dos anos devido a alterações promovidas nas diretrizes políticas energéticas e, também, em função dos preços do açúcar e do álcool no mercado externo.

Logo após a primeira crise do petróleo, várias foram as medidas adotadas para reduzir o *déficit* da balança comercial brasileira e, dentre os programas propostos na época, a substituição da gasolina como combustível, em veículos leves, por uma fonte renovável e nacional, como o etanol.

Em 1975, o PROALCOOL foi criado, sendo este um programa de substituição em larga escala dos combustíveis veiculares derivados de petróleo por álcool, financiado pelo governo brasileiro, devido à crise do petróleo de 1973. Um dos objetivos deste programa era evitar o aumento da dependência externa do petróleo.

Após a segunda grande crise do petróleo, em meados de 1979-80, o limite de etanol adicionado foi elevado para o patamar de 22 %(v/v). Nesta época, foram criados organismos como o CNAL (Conselho Nacional do Álcool) e a CENAL (Comissão Executiva Nacional do Álcool) para agilizar a implantação do PROALCOOL [5].

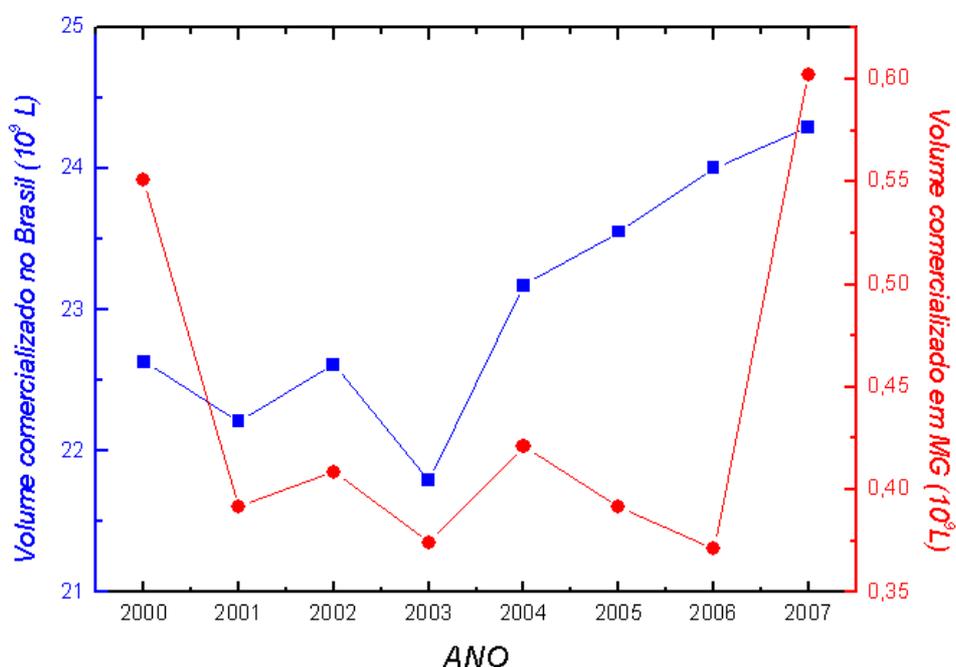
A partir de 1986, o cenário internacional do mercado petrolífero foi alterado com a queda dos preços do barril de petróleo. Com isso, a credibilidade dos programas que forneciam subsídios para a substituição de combustíveis fósseis foi colocada em risco. Um decréscimo no volume de investimentos nos projetos de produção interna de energia foi observado em 1988, devido à escassez de recursos para subsidiar os programas de estímulo aos energéticos alternativos.

Em 1990, em função de uma crise provocada pela redução na safra de cana-de-açúcar aliada ao excesso de carros a álcool em circulação, o teor alcoólico foi reduzido para 13 %(v/v). A situação quanto ao percentual de etanol anidro na gasolina se manteve confusa até 1994, quando o teor de 22 %(v/v) foi oficializado, conforme demanda dos produtores de veículos [6]. Após esta época os teores alcoólicos estabelecidos na legislação foram variados, sendo que atualmente o teor estabelecido é de 25 %(v/v) [7].

### 1.1.2 Qualidade da Gasolina Automotiva Brasileira

Na Figura 2, pode-se visualizar os volumes de gasolina do tipo C comercializados no Brasil e no estado de Minas Gerais [8].

Os volumes comercializados no Brasil entre 2000 e 2003 oscilaram entre 21,8 e 22,6 bilhões de litros de gasolina, sendo que a partir de 2003 observa-se um aumento gradativo nos volumes vendidos, que em 2007 foi de 24,3 bilhões de litros. No Estado de Minas Gerais, em 2007, foram comercializados 601,8 milhões de litros, sendo o segundo maior mercado do país.



**Figura 2** – Volumes de gasolina C comercializados, pelas distribuidoras, em todo território brasileiro (azul) e em Minas Gerais (vermelho) de 2000 a 2007.

A avaliação da qualidade da gasolina é realizada por uma série de ensaios físico-químicos, cujos limites e métodos analíticos são estabelecidos pela Portaria 309 da ANP [7] e são apresentados na Tabela 1.

A seguir é dada uma breve descrição de cada ensaio e seu impacto na qualidade do combustível.

A cor da gasolina determina sua utilização. A gasolina destinada aos veículos automotivos se apresenta geralmente de incolor a amarelada se isenta de corante cuja utilização é permitida no teor máximo de 50 ppm, com exceção da cor azul, restrita à gasolina de aviação e a cor rosa que é reservada a mistura gasolina e metanol.

**Tabela 1** – Ensaios, métodos e especificações da ANP para a avaliação da qualidade de gasolina [7].

Característica	Unidade	Especificação				Método	
		Gasolina Comum		Gasolina Premium		ABNT	ASTM
		Tipo A	Tipo C	Tipo A	Tipo C		
Cor	—	Incolor a amarela				Visual	
Aspecto	—	Límpido e isento de impurezas					
Álcool Etilico Anidro Combustível	%vol	-	25 ± 1	-	25 ± 1	NBR 13992	-
Massa Específica, 20°C	kg/m <sup>3</sup>	Anotar				NBR 7148	D 1298
						NBR 14065	D 4052
Destilação							
10% evaporado, máx.	°C	65,0					
50% evaporado, máx.	°C	120,0	80,0	120,0	80,0	NBR 9619	D 86
90% evaporado, máx.(7)	°C	190,0	190,0	190,0	190,0		
PFE, máx.	°C	220,0					
Resíduo, máx.	%vol	2,0					
Nº de Octano Motor MON, mín.	—	anotar	82,0	—	—	MB 457	D 2700
Índice Antidetonante IAD, mín.	—	anotar	87,0	anotar	91,0	MB 457	D 2699 D 2700 D 3606
Benzeno, máx.	%vol	1,2	1,0	1,9	1,5	—	D 5443 D 6277
Hidrocarbonetos Aromáticos, máx.	%vol	57,0	45,0	57,0	45,0	MB 424	D 1319
Hidrocarbonetos Olefínicos, máx.	%vol	38,0	30,0	38,0	30,0		

O ensaio do aspecto dá uma indicação visual da qualidade e de possível contaminação do produto. Uma amostra de gasolina deve apresentar-se límpida e isenta de materiais em suspensão como água, poeira, ferrugem, etc. Quando esses materiais estão presentes, pode ocorrer a redução da vida útil dos filtros de combustível dos veículos e prejuízo do funcionamento dos motores.

O ensaio do teor alcoólico é realizado com o objetivo de se medir a quantidade de etanol anidro adicionada à gasolina pelas distribuidoras. A adição de álcool etílico provoca o aumento da octanagem e a redução da taxa de emissão de poluentes como o monóxido de carbono [9]. Valores elevados de etanol provocam mau funcionamento do motor.

O ensaio da destilação tem como objetivo avaliar as características de volatilidade da gasolina [10]. É, também, utilizada no controle da produção da gasolina e na identificação de contaminação por derivados de maior massa, como óleo diesel, óleo lubrificante, querosene, etc [4].

A octanagem da gasolina indica sua resistência à compressão sem que haja detonação espontânea, dentro da câmara de combustão do motor, em comparação com uma mistura contendo iso-octano presente em uma mistura com n-heptano [4]. Para a avaliação da octanagem das gasolinas automotivas, encontram-se disponíveis os métodos MON (Motor Octane Number) e RON (Research Octane Number) [11].

O ensaio de teor de aromáticos expressa a quantidade em destes compostos na gasolina. Estes compostos conferem à gasolina uma boa resistência à detonação apresentando, isoladamente, um maior número de octanagem MON e RON do que os demais componentes deste combustível [9]. No entanto são bastante tóxicos para a saúde humana.

O ensaio de teor de olefinas indica a concentração dos hidrocarbonetos com dupla ligação carbono-carbono [4]. Altos teores de olefinas são responsáveis pela instabilidade química da gasolina gerando a formação de goma que pode alterar a qualidade da gasolina podendo provocar sérios danos aos motores [12].

### **1.1.3 Não-Conformidade e Adulteração**

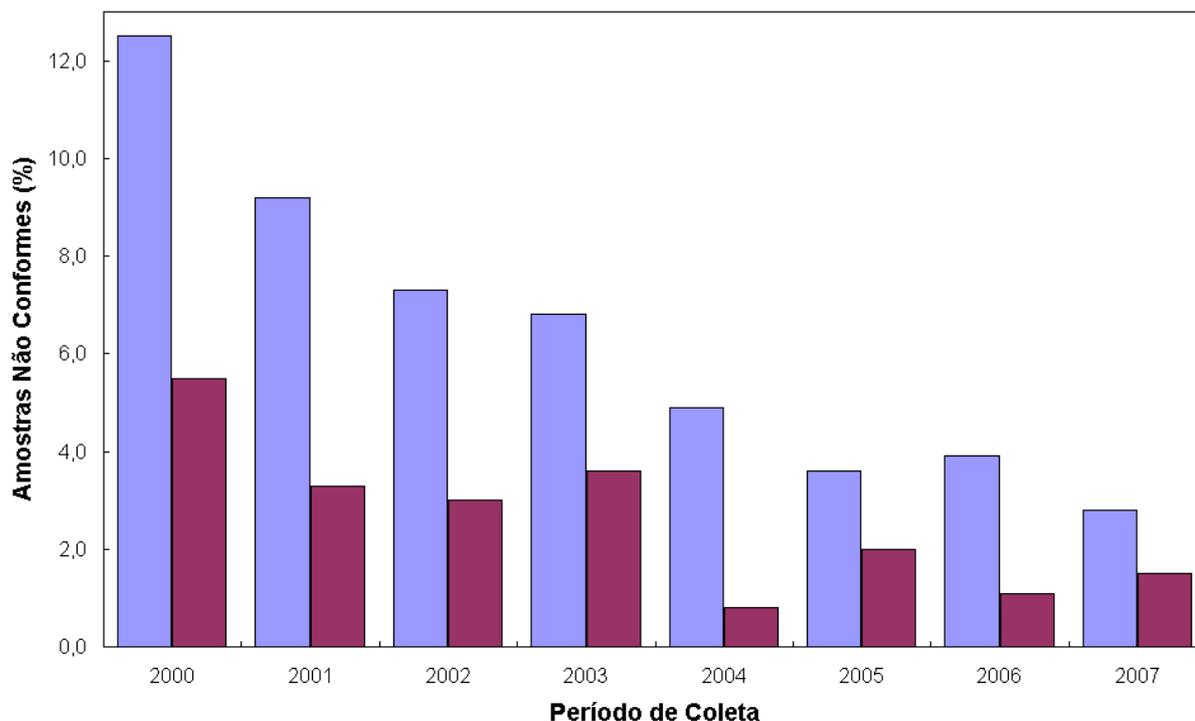
A aprovação da Lei do Petróleo (lei 9478) em 1997 trouxe resultados inesperados com relação à eficiência dos setores petroquímicos e à ampliação de suas atividades. O aumento da concorrência não gerou uma queda nos preços dos combustíveis, tampouco o aumento da qualidade dos produtos oferecidos. Este aumento de concorrência gerou denúncias de fraudes fiscais, sonegações de impostos e a comercialização de combustíveis adulterados, principalmente a gasolina [13].

Para combater a comercialização de combustíveis adulterados, a ANP criou em 1999 o Programa de Monitoramento da Qualidade dos Combustíveis (PMQC) e estabeleceu convênios com várias instituições (como centros de pesquisas e universidades), visando monitorar a qualidade dos combustíveis comercializados. Assim, em abril de 2000 foi fundado o Laboratório de Ensaio de Combustíveis (LEC-UFMG).

O LEC-UFMG é responsável pelo monitoramento da qualidade dos combustíveis comercializados na região leste do estado de Minas Gerais (compreendendo 550 municípios e aproximadamente 2000 postos revendedores), segundo ensaios normatizados e de acordo com as Portarias da ANP [7,14,15].

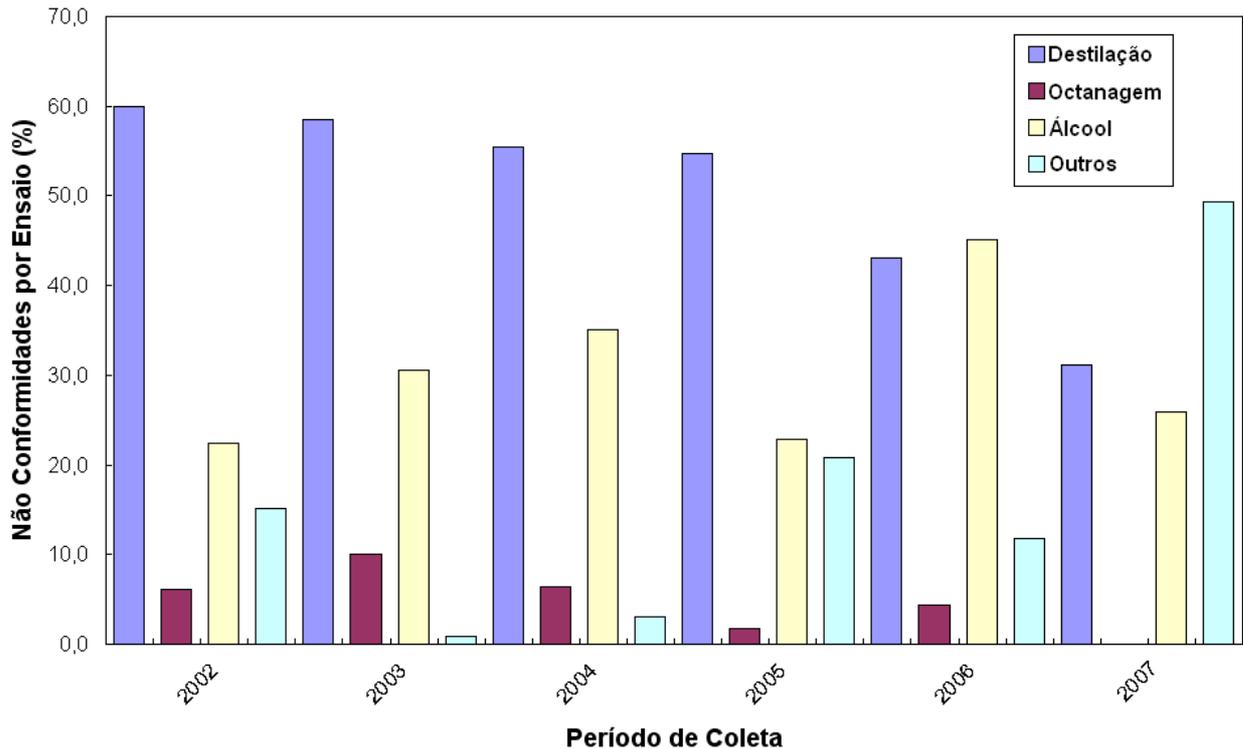
Uma amostra de gasolina é considerada apropriada para o uso, ou seja, “conforme”, quando os resultados dos ensaios estão dentro das especificações estabelecidas pela ANP [7]. Se, um ou mais ensaios estiverem fora das especificações, o combustível é considerado “não conforme”.

A Figura 3 apresenta a evolução do percentual de não conformidades, identificados pelo PMQC (da ANP. Uma queda sistemática das não conformidades pode ser observada, estabilizando em índices em torno de 3 %. Para o Estado de Minas Gerais, ocorreu também uma redução gradativa, com uma estabilização ao redor de 2 %. Estes resultados mostram a importância da atuação da ANP e de seus conveniados no controle da qualidade da gasolina comercializada no Brasil.



**Figura 3** – Evolução das não conformidades no Brasil (azul) e em Minas Gerais (vermelho) no período de 2000 a 2007.

As principais causas de não conformidade na gasolina estão relacionadas ao teor alcoólico, destilação, aspecto, octanagem, entre outros [8]. Na Figura 4 pode-se visualizar a evolução das amostras não conformes por ensaio, onde se pode observar que o ensaio de destilação foi responsável pelo maior percentual de não conformidades no período de 2002 a 2005. A partir de 2006 este percentual foi menor que o do teor alcoólico em 2006, e nos outros ensaios (como aspecto, cor e benzeno) em 2007.



**Figura 4** – Evolução das não conformidades por ensaio físico-químico em Minas Gerais no período de 2002 a 2007.

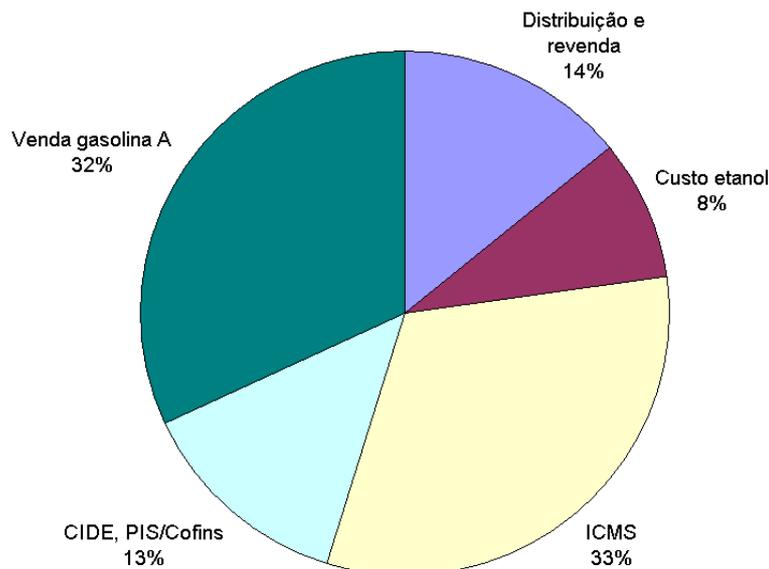
Uma não conformidade no combustível não significa que o mesmo esteja adulterado. Por exemplo, o aparecimento de material particulado ou turvação em uma gasolina que foi mal acondicionada nos postos revendedores. Por outro lado, adulterações criteriosas podem gerar combustíveis que estejam dentro dos limites especificados pela agência reguladora.

A adulteração em combustíveis é a adição de qualquer produto que modifique suas características originais, pois quando o refinador formula a gasolina, este o faz de forma a atender às especificações vigentes e a outras propriedades também necessárias ao perfeito funcionamento do motor [1]. Por isso, a adição de solventes orgânicos, prática comum no mercado brasileiro, pode afetar a qualidade do produto [16]. Os produtos mais comumente adicionados na gasolina são as naftas leves e pesadas, thinner e querosene. A adição de solventes é uma prática comum de adulteração devido à grande diferença de tributação da

gasolina e dos solventes utilizados. Este tipo de adulteração causa um aumento da poluição ambiental, uma diminuição no desempenho do motor, e perdas de receitas fiscais [17].

As refinarias brasileiras são responsáveis por produzir gasolinas do tipo A (comum, aditivada ou de alta octanagem) e encaminhar este produto as distribuidoras. Estas são responsáveis por adicionar a quantidade de etanol anidro determinada pelos órgãos reguladores e vender este produto final para os postos de combustível que finalmente repassam aos consumidores. Nesta cadeia, a adulteração pode acontecer em diferentes etapas, como na distribuidora ou nos postos revendedores.

O preço da gasolina é formado basicamente pela junção de cinco itens: valor comercializado de gasolina A pela refinaria, custo do etanol anidro, os impostos cobrados pela federação (CIDE, PIS/Cofins), e pelo estado (ICMS) e o valor de revenda e distribuição. Um exemplo desta composição de preço pode ser visto na Figura 5 [3].



**Figura 5** – Esquema de composição de preço da gasolina comercializada no estado do Rio de Janeiro.

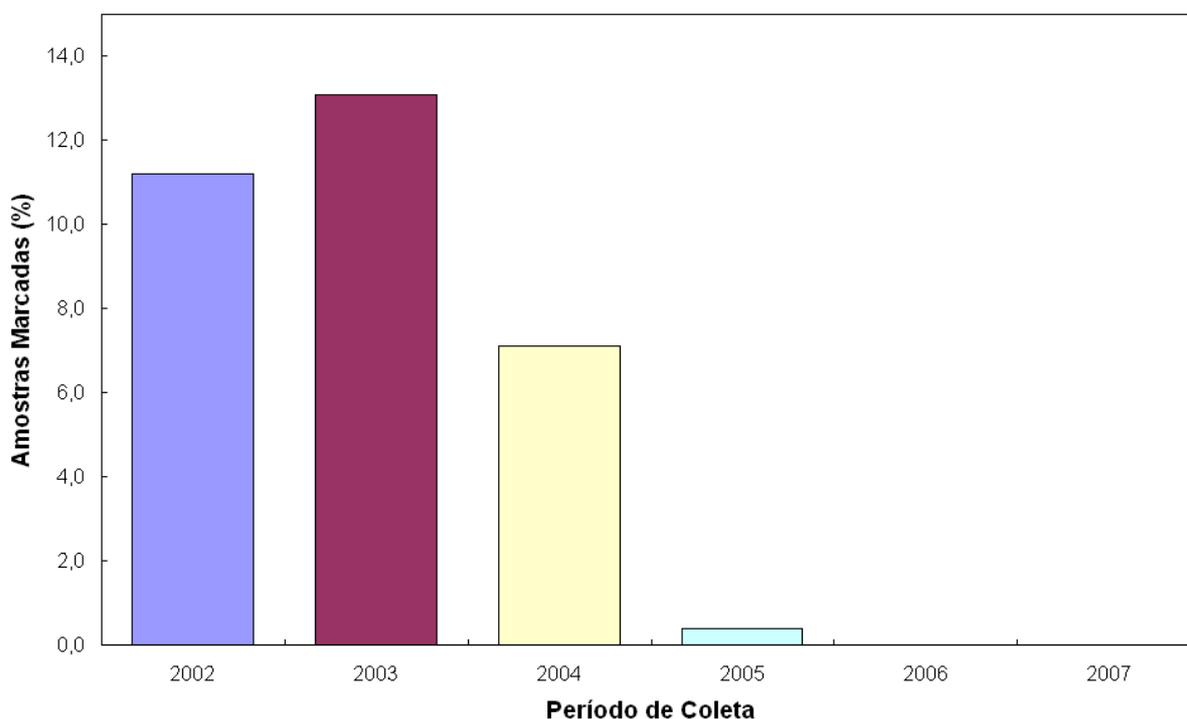
Pela figura observa-se que no estado do Rio de Janeiro, o valor do combustível vendido pela refinaria é próximo do valor de imposto cobrado pelo estado (ICMS). Para este estado, a menor contribuição no preço da gasolina relaciona-se ao custo do etanol adicionado.

Dos cinco fatores apresentados anteriormente, o imposto ICMS é que possui maior contribuição na modificação do preço da gasolina, visto que no Brasil existem diferentes taxas de ICMS. Nos estados de São Paulo, Rio de Janeiro e Minas Gerais os valores são, respectivamente 25, 33 e 25 %. Estas diferentes taxas de circulação e transporte de mercadorias e serviços, para a gasolina e os solventes industriais, podem induzir a adulteração [18-20].

Devido à sua localização geográfica e a disponibilidade de fornecimento de combustível por diferentes refinarias, no estado de Minas Gerais ocorre a comercialização de gasolinas de diferentes estados brasileiros, como São Paulo, Bahia e Rio de Janeiro. Sendo assim, a adulteração e as fraudes fiscais tornam-se um fator lucrativo no mercado de combustíveis.

Como objetivo de coibir as adulterações de gasolina com solventes orgânicos, em 2002 a ANP [21] instituiu o Programa de Marcação de Solventes, que obriga os produtores de solventes a adicionar um marcador químico. A detecção deste marcador na gasolina é uma evidência inequívoca de adulteração. Na Figura 6 pode-se visualizar a evolução das amostras marcadas desde o início do Programa de Marcação de Solventes, na região leste de Minas Gerais.

No período de 2003 a 2005 observou-se uma queda na porcentagem de amostras marcadas, não sendo detectada nenhuma amostra marcada nos anos de 2006 e 2007. Duas possibilidades podem ser atribuídas a este fato: a primeira seria a eficiência da ANP no monitoramento da qualidade e a segunda a não marcação dos solventes pelos produtores como determina a legislação.

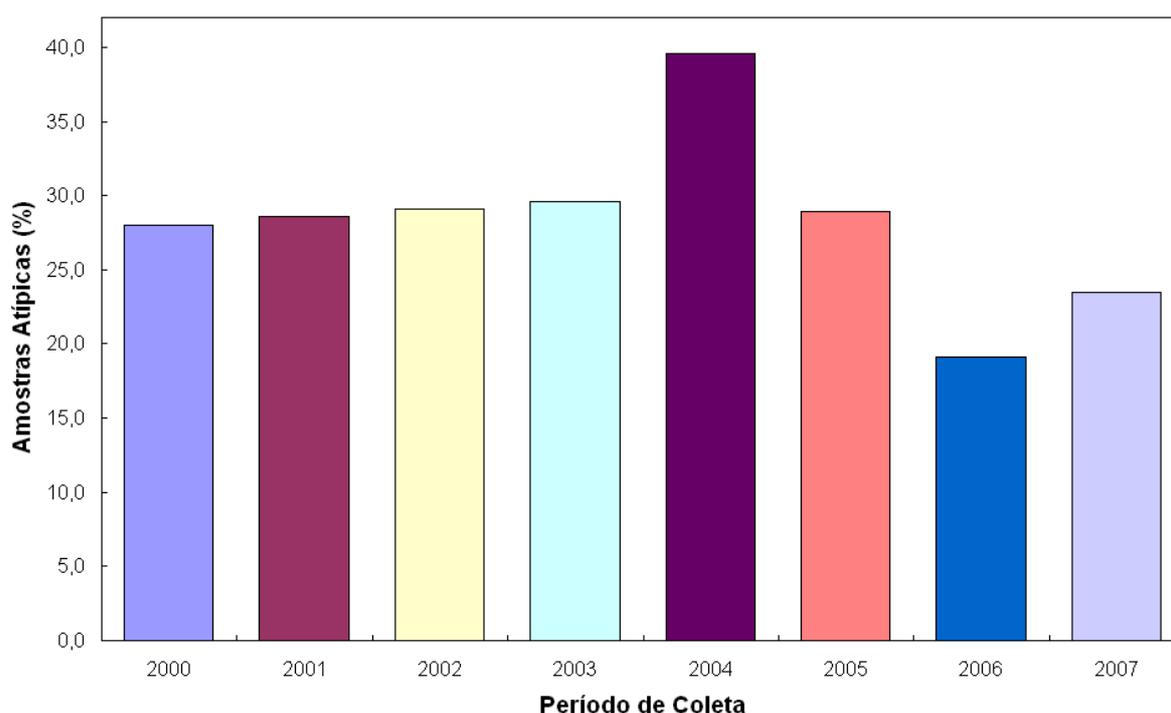


**Figura 6** – Evolução do percentual de amostras de gasolina que apresentaram a presença de marcador no período de 2002 a 2007 da região leste de Minas Gerais.

### 1.1.4 Atipicidade e Origem do Combustível

Uma amostra é considerada atípica quando ela apresenta um perfil diferenciado da maioria das amostras equivalentes, mas continua dentro das especificações, ou seja, conforme. Esta atipicidade pode estar relacionada ao fato da gasolina vir de diferentes fontes (refinaria ou planta petroquímica) ou por uma adição criteriosa de solventes industriais.

Barbeira [22] determinou a composição média da gasolina comercializada em Minas Gerais e detectou amostras com composição atípica utilizando doze parâmetros de ensaios físico-químicos associados a histogramas e curvas de distribuição normal. Com base nesta metodologia a Figura 7 apresenta a evolução das atipicidades observadas nas amostras de gasolina coletadas na região leste do estado de Minas Gerais. Nota-se que o percentual de amostras atípicas é bem mais elevado que os observados para as não-conformidades (Figura 3), sendo de 23,5 % em 2007.



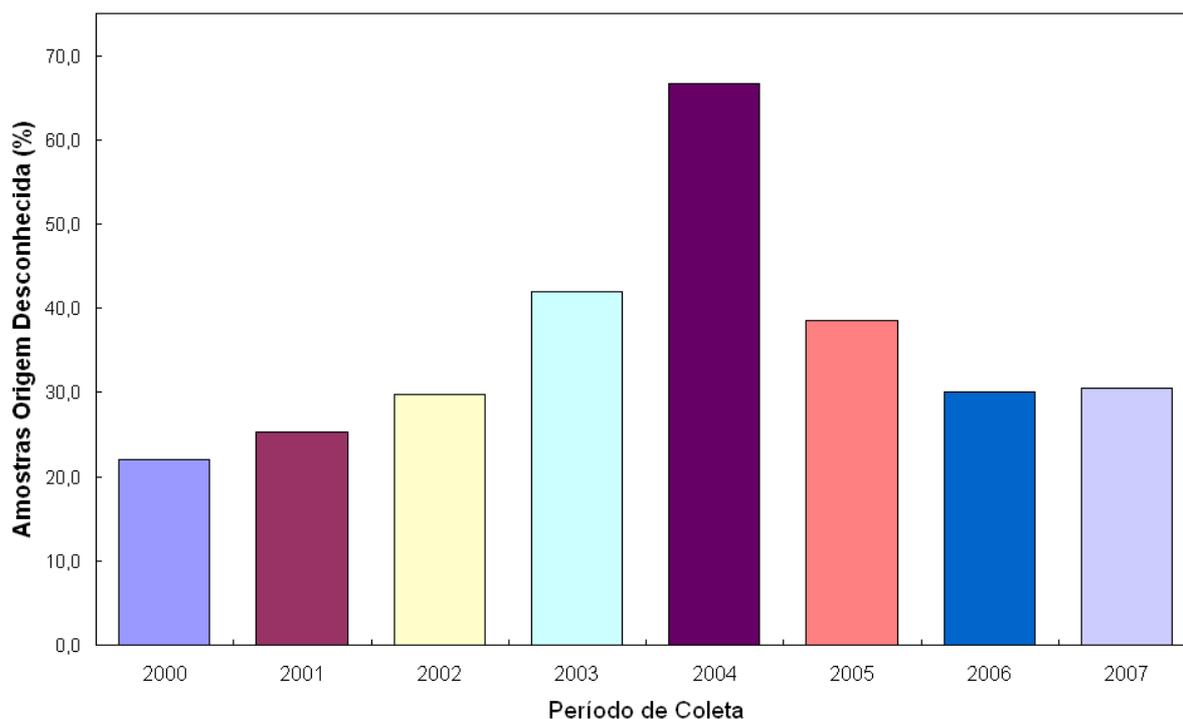
**Figura 7** – Evolução das amostras atípicas no período de 2000 a 2007.

Comparando as Figuras 6 e 7, percebe-se uma discordância entre a porcentagem das amostras marcadas e das atípicas no ano de 2004. Esperava-se que a queda da porcentagem de amostras marcadas fosse compatível a um baixo valor de amostras atípicas, o que não aconteceu, como pode ser visto na Figura 7. Com estes resultados pode-se considerar a possibilidade da adição de solventes industriais não marcados nas gasolinas, formando assim um elevado número de amostras atípicas.

Além das atipicidades, durante a coleta das amostras para o PMQC da ANP, pelo LEC-UFMG, observou-se a presença de amostras que não possuíam uma origem declarada, visto que alguns postos revendedores não apresentam a nota fiscal do combustível adquirido das distribuidoras. Pela Figura 8, pode-se visualizar a evolução das amostras de origem desconhecida, no período de 2002 a 2007.

Ao contrário das não-conformidades, observou-se um aumento gradativo do percentual de amostras com origem desconhecida entre 2002 e 2004, com uma queda gradativa, sendo de 35 % em 2007. Este fato reforça a importância de uma maior rigidez na fiscalização da ANP visto que este fato pode estar encobrindo adulterações e fraudes fiscais.

Como mostrado anteriormente, a maior porcentagem de amostras não conformes está relacionado ao ensaio da destilação. Por isso, um estudo mais aprofundado sobre o assunto se faz necessário.



**Figura 8** – Evolução do número de amostras com origem desconhecida no período de 2002 a 2007 no estado de Minas Gerais.

## 1.2 Destilação

A gasolina é uma mistura de líquidos compostos e por isso não possui uma temperatura de ebulição fixa. Sendo assim, neste caso é necessário utilizar o conceito de curva de destilação. A curva de destilação é um dos mais importantes parâmetros físico-químicos que medem a complexidade de misturas líquidas, sendo este parâmetro relacionado à volatilidade dos componentes da amostra [23].

No Brasil, os ensaios de curva de destilação podem ser realizados seguindo-se as normas ASTM D86 [24] ou NBR 9619 [25], sendo esta baseada na norma ASTM D86.

O método ASTM D86 descreve a destilação à pressão atmosférica de diversos derivados de petróleo e tem por objetivo determinar características de volatilidade, verificando se as proporções leves e pesadas do combustível produzido estão adequadas, buscando um bom desempenho na combustão e na detecção de contaminação com outros produtos.

Para se iniciar o ensaio da destilação (ASTM D86) é necessário que a amostra seja localizada em um dos quatro grupos discriminados na Tabela 2. Isto deverá ser feito baseando-se na composição, pressão de vapor, IBP esperado, FBP esperado, ou combinação de ambos. A gasolina, por exemplo, é considerada pertencente ao Grupo 2 enquanto o óleo diesel é discriminado no Grupo 4. Os arranjos da aparelhagem, a temperatura do condensador, e outras variáveis operacionais são definidas pelo grupo no qual a amostra pertence.

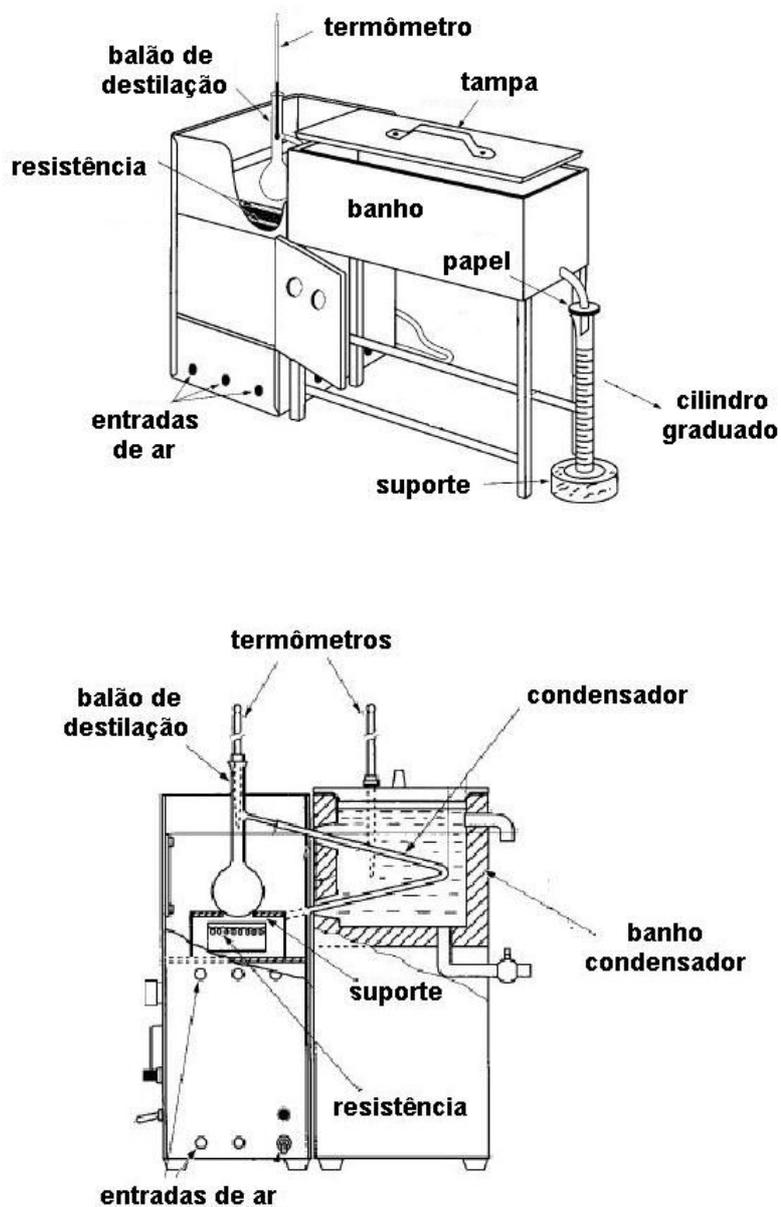
Os componentes básicos de um destilador manual são: balão de destilação, condensador associado a um banho de resfriamento, placa de suporte para o balão, fonte de aquecimento, suporte para o frasco coletor, medidor de temperatura e cilindro para recolher o volume destilado. O sistema é montado conforme mostrado na Figura 9. Além dos componentes básicos, o destilador automático é equipado com outros sistemas para se medir e anotar automaticamente a temperatura e associar o volume recolhido no frasco coletor [24].

Em cada grupo, diferentes condições experimentais, como tipos de balão de destilação, acondicionamento da amostra, temperatura do banho de resfriamento, são empregadas, conforme mostrado na Tabela 2.

O procedimento experimental básico deste ensaio consiste em se transferir cem mililitros da amostra, previamente refrigerada, para um balão de destilação específico, e submetidos a uma rampa de aquecimento, suficiente para manter a taxa de destilação entre 4 a 5 mL por minuto. O vapor destilado é condensado e recolhido em uma proveta refrigerada. O valor do volume do resíduo recolhido e o valor do volume de perda por evaporação também devem ser anotados [24].

**Tabela 2** – Parâmetros a serem observados na destilação de acordo com o tipo de amostra discriminado por cada grupo.

<b>Parâmetro</b>	<b>Grupo 1</b>	<b>Grupo 2</b>	<b>Grupo 3</b>	<b>Grupo 4</b>
Pressão de vapor a 37,8 °C, kPa	≥ 65,5	< 65,5	< 65,5	< 65,5
Ponto inicial de ebulição (IBP), °C	-	-	≤ 100	>100
Ponto final de ebulição (FBP), °C	≤ 250	≤ 250	> 250	> 250
Balão de destilação, mL	125	125	125	125
Termômetro destilação ASTM	7C (7F)	7C (7F)	7C (7F)	8C (8F)
Diâmetro do buraco da placa de suporte, mm	38	38	50	50
Temperatura no início do teste no balão, °C	13-18	13-18	13-18	Não acima da ambiente
Temperatura no início do teste no cilindro coletor e nos 100 mL medidos, °C	13-18	13-18	13-18	temperatura ambiente
Temperatura do cilindro coletor, °C	<10	-	-	-
Temperatura da amostra estocada, °C	<10	<10	Ambiente	Ambiente
Temperatura da amostra após condicionamento prévio para análise, °C	<10	<10	Ambiente ou 9-21	Ambiente ou acima do "pour point"
Temperatura do banho de resfriamento, °C	0-1	0-5	0-5	0-60
Temperatura ao redor do cilindro coletor, °C	13-18	13-18	13-18	± 3
Tempo da primeira aplicação do calor para o ponto inicial de ebulição, min	5-10	5-10	5-10	5-15
Tempo do ponto inicial de ebulição para os 5% recuperados, s	60-100	60-100	-	-
Tempo do ponto inicial de ebulição para os 10% recuperados, min	-	-	-	-
Taxa de condensação dos 5% recuperados até os 5 mL restantes no balão, mL/min	4-5	4-5	4-5	4-5
Tempo de recuperação dos 5 mL restantes até o ponto final, min	5 máx	5 máx	5 máx	5 máx



**Figura 9** – Modelo de destilador manual, segundo a ASTM D86 [24].

Na execução do procedimento experimental do ensaio, as condições experimentais (Tabela 2) devem ser seguidas de acordo com o grupo no qual a amostra pertence.

Após a execução da destilação é necessário corrigir as temperaturas equivalentes a cada volume evaporado para a pressão de 760 mmHg (101,3 kPa). Sem esta correção, a temperatura lida, para uma mesma amostra, entre dois laboratórios em duas localizações diferentes apresentará, em geral, uma diferença menor que 0,1 a 100 °C. Esta correção é realizada aplicando-se a cada temperatura lida a equação de Sydney Young (Equação 1) considerando que a leitura foi realizada em °C [24].

$$C_c = 0,0009 (101,3 - P_k)(273 + t_c) \quad \text{Eq.1}$$

onde,  $C_c$  é a temperatura corrigida a pressão de 1 atm,  $P_k$  é a pressão barométrica prevalecentes no momento e local do ensaio em kPa e  $t_c$  é a temperatura observada.

Além desta correção, é necessária a correção do volume de perda. Esta deve ser realizada por meio de um procedimento aritmético no qual a perda observada de cada percentual evaporado prescrito é deduzida para se obter o correspondente percentual recuperado. Cada temperatura é calculada a partir da Equação 2 [24].

$$T = T_L + \frac{(T_H - T_L)(R - R_L)}{(R_H - R_L)} \quad \text{Eq. 2}$$

onde,

$R$ = percentual recuperado correspondente ao percentual evaporado prescrito,

$R_H$  = percentual recuperado adjacente ao, e maior que  $R$ ,

$R_L$ = percentual recuperado adjacente ao, e menor que  $R$ ,

$T$ = temperatura lida no percentual evaporado prescrito,

$T_H$ = temperatura lida correspondente a  $R_H$  e

$T_L$ = temperatura lida correspondente a  $R_L$ .

Os valores de temperatura corrigidos, juntamente com os percentuais de volume recuperado, podem ser expressos em forma de tabela ou gráfico. Da curva de destilação, os pontos 10, 50, 90 % do volume evaporado e ponto final de evaporação são requeridos no controle de qualidade da ANP. Além destes pontos, o volume do resíduo também é limitado pela agência reguladora [7].

O controle do ponto 10 % do volume da amostra tem como objetivo garantir que a gasolina possua uma quantidade mínima de frações leves que se vaporizem e queimem, com facilidade, na temperatura de partida a frio do motor, facilitando o início do funcionamento do veículo. Neste ponto predominam as frações mais leves do combustível. Pela ANP a amostra, para ser conforme neste item, deve possuir temperaturas menores que 65 °C [7].

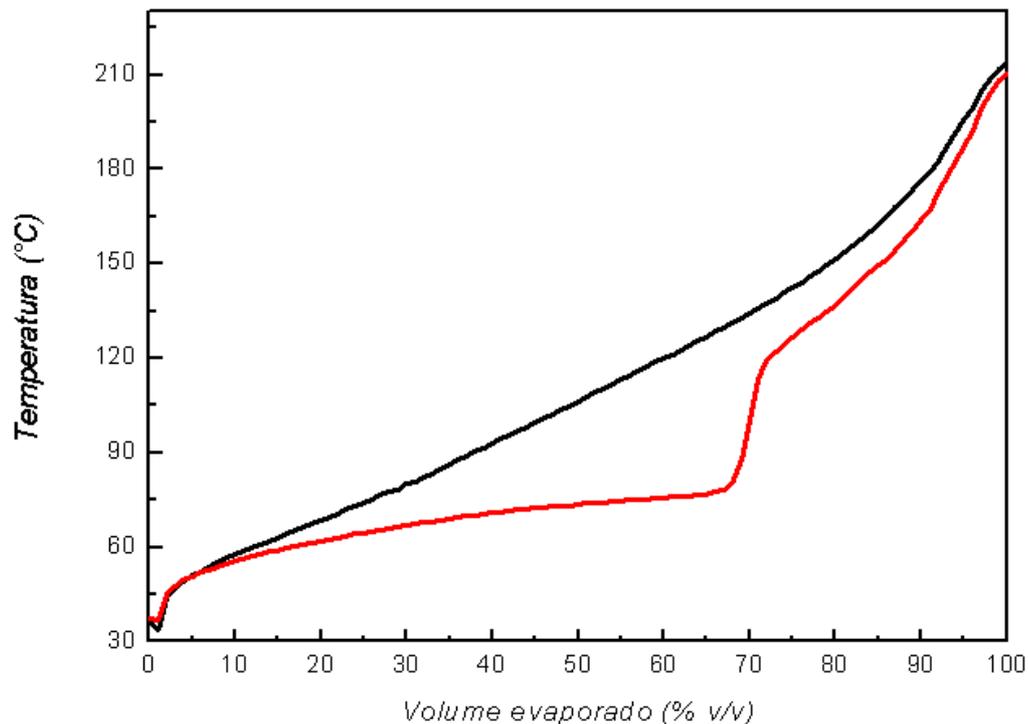
Já o controle do ponto 50 % do volume do produto visa possibilitar a partida fácil do motor, mas sua principal influência se faz sentir no tempo necessário ao seu aquecimento. Essa característica reflete a concentração de componentes de temperatura de destilação intermediária. Para este item, as amostras que possuem temperaturas inferiores a 80 °C são consideradas conformes para a agência reguladora [7].

A limitação do ponto 90 % do volume da amostra tem como objetivo minimizar a formação de depósitos na câmara de combustão e nas velas de ignição. Quanto mais alta for a temperatura necessária para vaporizar as frações mais pesadas, maior será a quantidade de amostra sem queimar que sobrar na câmara de combustão. Amostras com temperaturas inferiores a 190 °C, para este item, são consideradas conformes pela Portaria 309 da ANP [7].

O ponto final de evaporação indica a temperatura mais alta observada na destilação. Este ponto pode ou não coincidir com a vaporização total do produto, já que a parte final da gasolina pode sofrer decomposição térmica, formando assim frações mais leves que reduzem a temperatura na fase final do teste. O controle do ponto final de ebulição da gasolina visa diminuir a formação de depósitos na câmara de combustão e nas velas de ignição [24]. Uma amostra conforme, para este item, deve ter temperaturas menores que 220 °C no ponto final de evaporação.

O resíduo formado na destilação é o conteúdo que sobrou após o término da destilação (ou seja, até o ponto final de ebulição). Um alto volume de resíduo indica que a gasolina possui um elevado teor de frações pesadas advindas do processo de produção ou de contaminação da amostra por produtos mais pesados como, por exemplo, o óleo diesel [4]. Pela Portaria 309 da ANP, tem-se que o resíduo não pode maior que 2,0 mL.

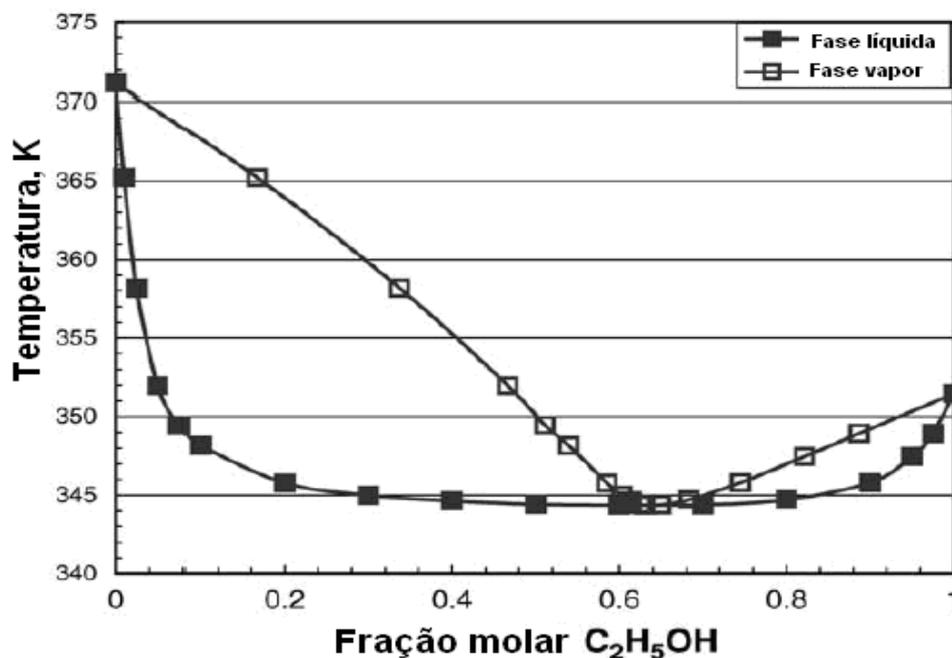
Um exemplo de curva de destilação de uma gasolina isenta de etanol (em preto) e acrescida de 25 % (v/v) de álcool etílico combustível (em vermelho) é mostrado na Figura 10.



**Figura 10** – Curva de destilação de gasolinas tipo A com e sem a adição de 25 % (v/v) de etanol anidro. Sem adição de etanol (preto), com adição de 25 % de etanol (vermelho).

A modificação do comportamento da curva de destilação da gasolina quando há a adição de etanol se deve a formação de um platô na região entre 65 e 75 % de volume evaporado, como visualizado na Figura 10.

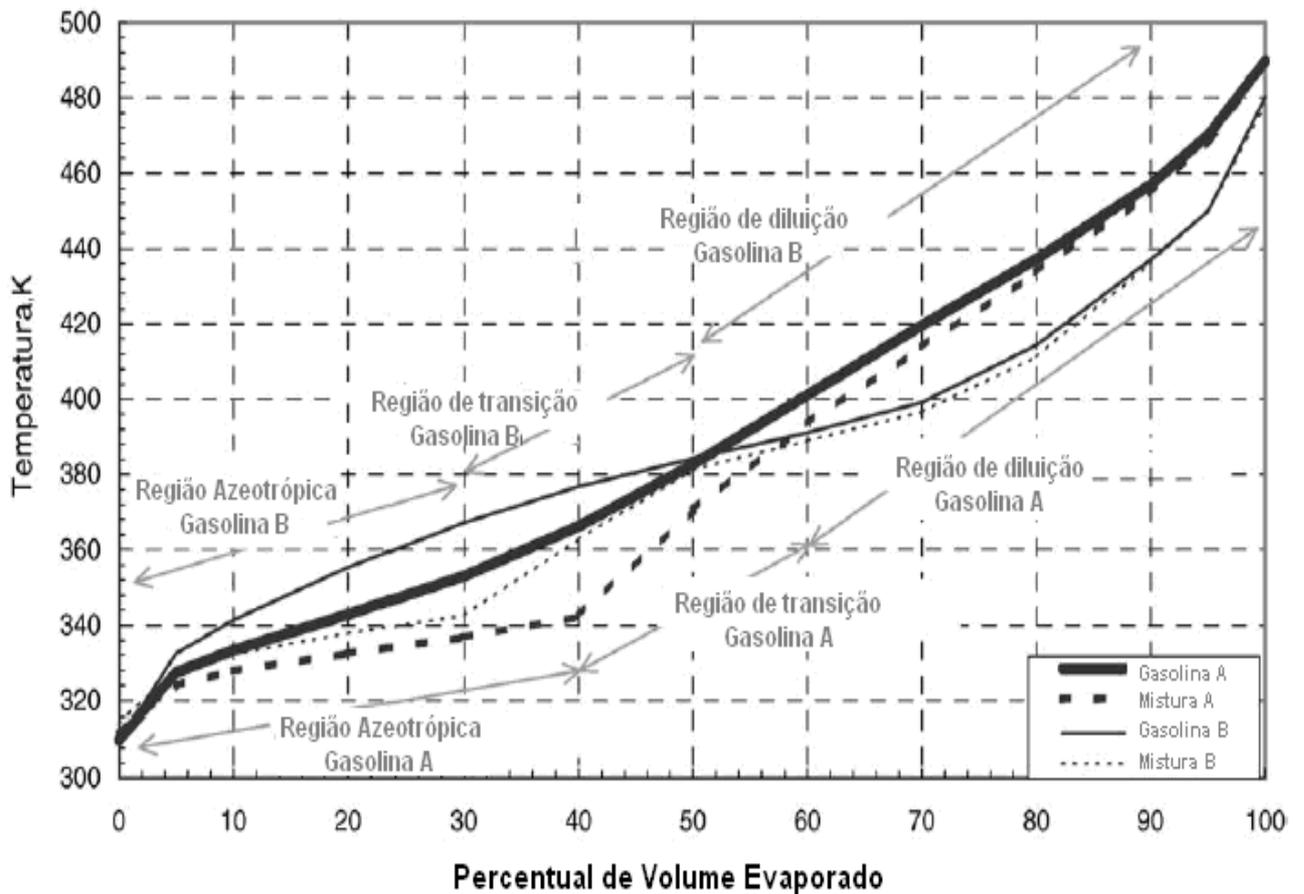
Além da formação do platô pode-se perceber, em geral, a redução da temperatura de ebulição quando se adiciona etanol a gasolina. Este fato está relacionado à formação de misturas azeotrópicas hidrocarboneto-etanol. Misturas de hidrocarbonetos com etanol exibem altos desvios das misturas ideais definidas pela Lei de Raoult, formando misturas azeotrópicas, como mostrado na Figura 11, obtida por French et al [26].



**Figura 11** – Equilíbrio líquido-vapor para sistemas binários de etanol e n-heptano a 1 atm. Figura adaptada de French et al [26].

A mistura gasolina-etanol pode ser modelada similarmente aos sistemas binários hidrocarboneto-etanol com a gasolina, sendo definida como um conjunto de componentes discretos ou, com menos precisão, como um conjunto de pseudo-componentes ou contínua distribuição de hidrocarbonetos [26].

Para cada azeótropo formado na mistura binária hidrocarboneto-etanol haverá uma temperatura no qual esta mistura entrará em ebulição. Assim, para cada tipo de hidrocarboneto haverá a formação de um azeótropo com temperatura de ebulição determinada. Uma vez que os azeótropos formados entraram em ebulição, a mistura de hidrocarbonetos restantes entrará em ebulição, para cada hidrocarboneto, em um específico ponto de ebulição. French et al [26] analisando curvas de destilação definiram três regiões distintas: região azeotrópica, região de transição e região de diluição, como mostrado na Figura 12.



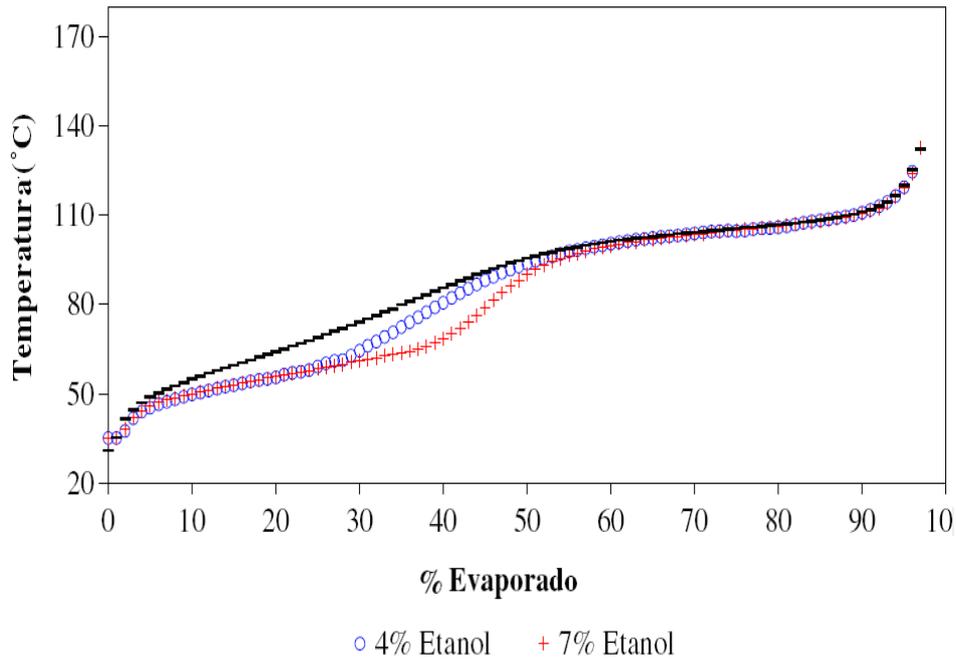
**Figura 12** – Curvas de destilação de gasolinas (A e B) sem e com a adição e 10 %(v/v) de etanol. Figura adaptada de French et al [26].

Pela figura pode-se observar que a região azeotrópica, para ambas as gasolinas, apresentaram temperaturas de ebulição inferiores às observadas nas gasolinas sem adição de etanol. Na região de transição houve a formação de um platô, devido ao etanol adicionado.

Menezes et al [27] observaram em seu estudo que o aumento do teor alcoólico na gasolina provoca uma mudança no comportamento da curva de destilação na região de transição como mostrado pela Figura 13.

Pela figura pode-se perceber que quanto maior é o teor alcoólico da amostra maior é a inclinação do platô formado. A partir dos 55 % de volume evaporado, a adição de etanol não influenciou na destilação, mostrando que a mistura azeotrópica etanol/hidrocarbonetos já foi totalmente evaporada [27].

A adição de etanol na gasolina além de reduzir a temperatura de ebulição das misturas hidrocarboneto-etanol pode também afetar significativamente a volatilidade da gasolina [26-28]. Com isso, algumas propriedades físico-químicas como a pressão de vapor e a entalpia de evaporação são afetadas com a adição de etanol na gasolina.

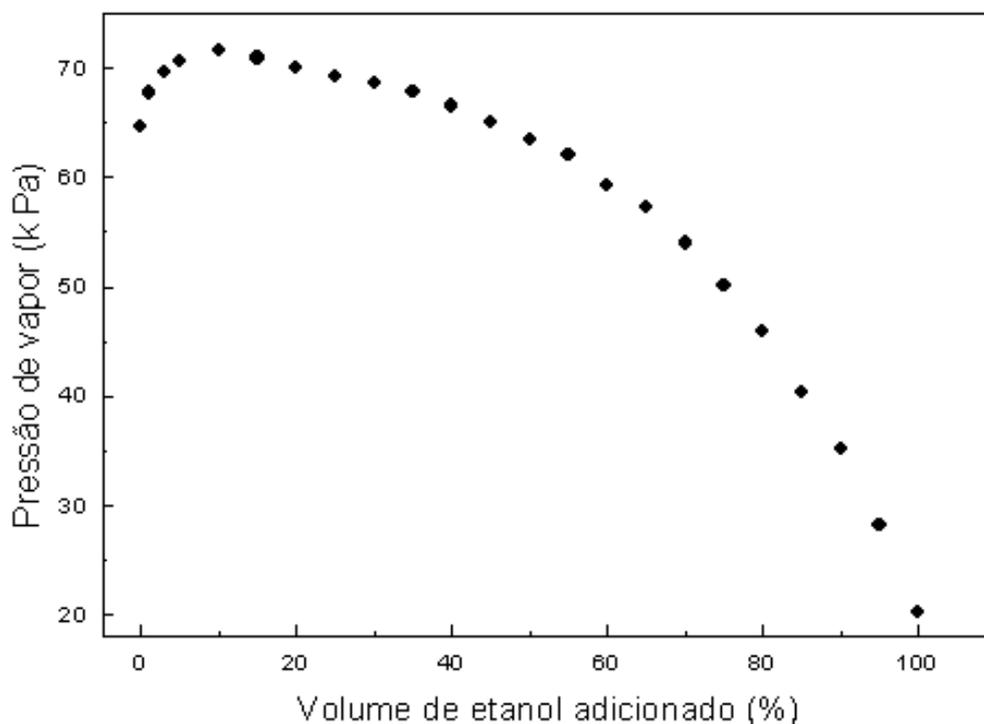


**Figura 13** – Curvas de destilação de gasolina e suas misturas com etanol numa proporção volumétrica de 4 e 7 %(v/v). Figura adaptada de Menezes et al [27].

A adição de etanol em diferentes concentrações causa modificações na pressão de vapor visto que a gasolina possui pressão de vapor aproximadamente três vezes maior que o etanol como mostrado no estudo de Pumphrey et al [29]. Neste estudo, os autores realizaram medições de pressão de vapor para diferentes misturas de etanol-gasolina, verificando-se um declínio sucessivo da pressão de vapor na faixa de 15 a 100 % de etanol, como mostra a Figura 14.

A região de declínio de pressão de vapor entre 15 e 25 %(v/v) de etanol possui um decréscimo em torno de  $0,159 \text{ kPa } ^\circ\text{C}^{-1}$ , enquanto que a região de 25 a 30 %(v/v) apresentou um decréscimo um pouco menos acentuado com uma taxa de  $0,110 \text{ kPa } ^\circ\text{C}^{-1}$ .

Calatuña et al [28] observaram que o aumento da concentração de etanol provoca mudanças na pressão de vapor e na entalpia de vaporização. As gasolinas utilizadas neste trabalho, diferentemente do estudo de Pumphrey et al [29], possuíam diferentes concentrações de hidrocarbonetos aromáticos, parafínicos e olefínicos. O maior aumento observado na entalpia de vaporização das formulações com etanol foi para a gasolina com maior teor de aromáticos. À medida que a concentração de aromáticos diminui o efeito no aumento da entalpia de vaporização, devido à introdução de etanol, diminui. Os autores observaram que após os 5 % em volume de etanol a entalpia de vaporização tende a se estabilizar em um valor fixo de forma similar à pressão de vapor.



**Figura 14** – Comportamento da pressão de vapor de acordo com o volume de etanol adicionado na gasolina. Figura adaptada de Pumphrey et al [29].

Curvas de destilação de gasolina, além de apresentarem um comportamento diferenciado quando se adiciona etanol, elas podem ser utilizadas associadas a técnicas quimiométricas na detecção e avaliação de adulterações em gasolinas com diversos solventes orgânicos [16,30,31] e na previsão da refinaria de origem [13].

Wiedemann e colaboradores [16] utilizaram parâmetros físico-químicos (temperaturas equivalentes a 10, 50, 90 % do volume evaporado e o ponto final de evaporação, MON, RON, IAD, teores de benzeno, hidrocarbonetos aromáticos, alifáticos leves e pesados) juntamente com a cromatografia associadas à análise por aglomeração (HCA) para detecção de adulteração de gasolina com solventes orgânicos. Um conjunto de 22 amostras foi preparado em laboratório com adulterações de solventes do tipo hidrocarbonetos aromáticos, alifáticos leves e pesados e seus parâmetros físico-químicos aplicados ao HCA. O dendograma deste conjunto mostrou a formação de quatro grupos de acordo com o tipo e proporção de solvente utilizado. Os autores concluíram que a adulteração com solventes por adição de hidrocarbonetos alifáticos pesados se mostrou fácil de ser determinada visto que estes não estão presentes naturalmente na gasolina. Já a detecção da adulteração de hidrocarbonetos aromáticos e alifáticos leves se mostrou mais difícil, pois estes componentes estão presentes na composição da gasolina não adulterada.

Curvas de destilação aplicadas ao modelo SIMCA possibilitaram a detecção de adulterações na gasolina brasileira por Oliveira e colaboradores [30]. Um total de 50 amostras de gasolina, sendo 8 adulteradas foi usado neste estudo. Este conjunto foi posteriormente subdividido em um conjunto de previsão composto por 34 amostras e um conjunto teste de 16 amostras sendo que destas, 50 % (ou 8 amostras) eram adulteradas. As porcentagens de 10 a 100 de volume evaporado, em incrementos de 10 %, foram aplicadas das amostras de cada conjunto foi aplicado ao SIMCA. Os autores observaram que a temperatura equivalente a 60 % de volume evaporado foi o parâmetro mais importante na classificação SIMCA. Neste modelo, foi observado um percentual de acerto de 92 % na classificação das 50 amostras bem como a discriminação das amostras classificadas como conformes à legislação. Observou-se também que não foi necessária a utilização de outra técnica para as amostras classificadas como não contaminadas pelo modelo SIMCA. Entretanto, técnicas espectroscópicas ou cromatográficas devem ser aplicadas quando é necessário obter informações sobre o tipo de adulteração observada na gasolina.

As temperaturas equivalentes a 10, 50 e 90 % de volume destilado foram usadas juntamente com outros sete parâmetros físico-químicos na determinação da origem da gasolina comercializada no estado de Minas Gerais [13]. Os parâmetros físico-químicos de um total de 930 amostras de gasolina foram aplicados a análise discriminante na criação de um modelo de calibração para a determinação da origem de amostras de gasolina. Um conjunto teste de 398 amostras foi aplicado posteriormente a este modelo que apresentou um percentual de 97 % de acerto na segregação das amostras, no modelo LDA. Um total de 120 amostras (20 de cada refinaria) escolhidas aleatoriamente do conjunto total de 1380 amostras foi usado na construção de modelos PCA e LDA para a visualização da separação das amostras por refinaria de origem. Com este estudo pode-se observar que com o uso dos parâmetros físico-químicos realizados pelo programa de avaliação da qualidade aliados às técnicas quimiométricas (PCA, LDA e HCA) possibilitaram a visualização e determinação da origem de gasolinas.

Flumignan et al [31] obtiveram um perfil de qualidade da gasolina brasileira através da quantificação de hidrocarbonetos saturados e de etanol utilizando cromatografia gasosa e análise exploratória. Os resultados das análises físico-químicas (incluindo os pontos 10, 50, 90 e PFE da curva de destilação) de 2400 amostras foram colocados em uma planilha e aplicados a análise de aglomerados (HCA), que formou um total de 25 aglomerados. Seis amostras obtidas diretamente de refinarias (Revap, Recap, Replan, RPBC, Reduc e Manguinhos) foram analisadas por cromatografia gasosa pelo método ASTM D6733. Para cada cromatograma obtido, 80 picos representativos de classes de compostos (parafinas, isoparafinas e naftalenos de C3 a C14) foram selecionados para detalhamento na análise estatística. Neste estudo, 13 amostras conformes e 18 não conformes a Portaria 309 da ANP foram utilizadas para a

determinação da quantidade de hidrocarbonetos saturados e de etanol. Foi possível classificar 94 % das amostras que apresentaram o nível de hidrocarbonetos saturados no intervalo entre 34,6 a 46,6 % utilizando HCA e PCA associados à cromatografia gasosa.

Todos estes estudos mostram a grande potencialidade do uso de métodos quimiométricos, em conjunto com diversas técnicas analíticas, na avaliação de diferentes parâmetros físico-químicos da gasolina automotiva.

### **1.3 Quimiometria**

#### **1.3.1 Organização e Pré-Processamento dos Dados**

Para a aplicação dos métodos quimiométricos os dados devem ser organizados na forma de uma matriz de dados  $X$ , chamada de matriz de dados originais. Nesta matriz, as linhas representam as  $n$ -amostras e as colunas as  $m$ -variáveis, sendo cada variável representada por um vetor coluna e cada amostra por um vetor linha [32].

O pré-processamento é uma etapa que consiste em um pré-tratamento dos dados que podem estar expressos em grandezas diferentes ou apresentarem dependências lineares que podem comprometer a análise. Este tratamento prévio dos dados é utilizado para ajustar grandezas em valores equivalentes. Os dois tipos de pré-processamentos mais conhecidos na literatura são a centralização dos dados na média e o autoescalamento [33].

Centrar os dados na média é simplesmente uma adaptação do conjunto de dados para reposicionar o centróide dos dados para a origem do sistema de coordenadas [32]. Para centralizar os dados na média, o valor de cada elemento da variável é obtido subtraindo-se o valor original do valor médio [34].

Autoescalamento é um processo no qual ocorre uma adaptação do conjunto de dados de forma a uniformizar a variância de cada variável [32]. Do ponto de vista da Química Analítica, o autoescalamento mapeia conjunto de dados em um espaço resumido cujos eixos deixaram de ter qualquer significado externas físicas ou químicas [34]. O autoescalamento é feito subtraindo-se cada valor pela média e este resultado é dividido posteriormente pelo desvio padrão.

### 1.3.2 Análise de Componentes Principais (PCA)

A análise de componentes principais é um método não supervisionado de análise exploratória. Na análise exploratória tem-se como objetivo fazer uma avaliação inicial dos dados para encontrar que tipo de informação é possível retirar deles, e assim, definir as diretrizes para um tratamento mais aprofundado. Na PCA, algoritmos são utilizados com o objetivo de reduzir a dimensão dos dados e facilitar a visualização de todo o conjunto.

As componentes principais de um processo são um conjunto particular de combinações lineares das variáveis originais. Geometricamente, as variáveis das componentes principais são os eixos do novo sistema de coordenadas obtido pela rotação do sistema original. Estes novos eixos representam as direções de variabilidade máxima. A informação contida no conjunto completo das  $p$ -componentes principais é exatamente equivalente à informação contida no conjunto completo de todas as variáveis do processo, mas poder-se-á usar bem menos do que as  $p$ -componentes principais para se obter uma descrição satisfatória.

Na PCA, a matriz de dados é decomposta em matrizes de escores e pesos conforme mostrado na Equação 3:

$$X = T L^t \quad \text{Eq.3}$$

onde,  $X$  é a matriz original que contém  $n$  linhas (objetos) e  $p$  colunas (características);  $T$  é a matriz de escores com  $n$  linhas e  $d$  colunas (número de componentes principais);  $L$  é a matriz de pesos com  $d$  colunas e  $p$  linhas; e  $t$  é a transposição da matriz [35].

Os escores descrevem a relação entre as amostras do modelo no espaço e os pesos descrevem a importância de cada variável dentro do modelo [36].

Estudos [13,17,37] têm demonstrado a utilidade de métodos multivariados, como a PCA, aplicados à análise de gasolina para a investigação e detecção da adulteração e origem do combustível.

Para a investigação da adulteração em gasolina, Skrobot et al [17] utilizou a cromatografia gasosa associada às técnicas quimiométricas PCA e LDA (Análise de Discriminantes Linear). Um total de quinze amostras de gasolina consideradas conformes pela Portaria 309 da ANP foram dopadas com quatro tipos de solventes (aguarrás, querosene, solvente de borracha e thinner) no intervalo de 2 a 30 % (v/v). O gráfico das duas primeiras componentes principais demonstrou a separação das amostras dopadas com querosene e thinner enquanto que pelo gráfico PC2xPC4 pode-se visualizar a separação total das amostras dopadas com solvente de borracha e aguarrás. Neste estudo o uso da PCA para a redução do número de variáveis foi necessário, pois no LDA o número de amostras não pode ser menor

que o de variáveis. Um modelo LDA foi construído com quatro componentes principais, visto que elas agregaram um alto percentual de variância explicada do modelo PCA. Com este modelo os autores perceberam que a detecção de adulteração com o solvente thinner pode ser obtida a partir da concentração de 5 % (v/v). Já para os adulterantes solvente de borracha, aguarrás e querosene, pode-se observar a adulteração a partir de 15,10 e 10 % (v/v), respectivamente.

Pereira et al. [37] empregaram métodos quimiométricos (PCA e LDA) para avaliar o comportamento de amostras de gasolina não marcadas e dopadas utilizando análises de infravermelho médio com quatro tipos de solventes (aguarrás, querosene, solvente de borracha e thinner), variando a adição de cada solvente dentro de uma faixa de 1-50 % (v/v) com cada amostra, dando um total de 200 amostras. O modelo quimiométrico apresentou 96,3 % de eficiência na classificação das amostras quanto à adulteração e 93,0 % na identificação de solventes.

### **1.3.3 Análise Discriminante Linear (LDA)**

Análise discriminante é uma técnica estatística apropriada quando a variável dependente é dada por categorias e as variáveis independentes por medidas numéricas. Esta técnica é capaz de lidar com dois ou múltiplos (três ou mais) grupos sendo empregada para classificar elementos de uma amostra ou população [38].

A análise discriminante envolve determinar uma variável estatística, a combinação linear de duas (ou mais variáveis) independentes que discriminarão melhor os grupos definidos *a priori*. A discriminação é conseguida estabelecendo-se os pesos da variável estatística para cada variável para maximizar a variância entre grupos relativa à variância dentro dos grupos.

A análise discriminante é apropriada para testar a hipótese de que as médias de grupo de um conjunto de variáveis independentes para dois ou mais grupos são iguais. Para isso, cada variável independente é multiplicada por seu peso correspondente. O resultado é um escore Z discriminante composto para cada indivíduo da análise. A média do grupo, também chamada de centróide, é calculada a partir da média dos escores discriminantes para todos os indivíduos em um grupo. Os centróides indicam o local mais típico de qualquer indivíduo de um grupo particular. A comparação entre os centróides dos grupos mostra os quão afastados estão os grupos ao longo da dimensão testada [39].

A análise discriminante foi usada em estudos com técnicas como cromatografia [17] e infravermelho médio [37,40], além de parâmetros físico-químicos [13] para avaliação e detecção da adulteração de gasolinas, determinação do tipo ou da refinaria de origem.

Balabin e Safieva [40] tentaram classificar 382 amostras de gasolina de acordo com a origem (refinaria ou processo) e tipo da amostra. Neste estudo, além do LDA outros dois métodos de classificação foram utilizados: SIMCA e MLP (Multilayer Perceptron). Dentre os três métodos acima o que apresentou um menor erro de classificação foi o MLP (12 %) enquanto que o LDA apresentou o maior erro de classificação (20 %). Possíveis razões podem estar relacionadas ao elevado erro de classificação das amostras, tais como o não uso de pré-processamento, de técnicas de classificação mais sofisticadas (SVM, PNN, etc) e a deficiência do número de amostras.

#### 1.3.4. Mínimos Quadrados Parciais (PLS)

O método dos mínimos quadrados, considerado como um processo de calibração multivariada pode ser dividido em duas etapas: modelagem e validação. A etapa de modelagem estabelece uma relação matemática entre as matrizes  $X$  e  $Y$  do conjunto de calibração e a etapa de validação otimiza a relação no sentido de uma melhor descrição do analito(s) de interesse [41]. Uma vez concluída a calibração, o sistema (instrumento físico + modelo matemático) estará apto a ser utilizado para previsão de outras amostras.

O método PLS foi criado por H. Wold e propagado pelo seu filho S. Wold [35]. No método PLS cada variável latente da matriz  $X$  é modificada de modo que a covariância entre ela e o vetor da matriz  $Y$  seja maximizado. Neste método, as matrizes  $X$  e  $Y$  são decompostas em matrizes menores de acordo com as Equações 4 e 5:

$$X = TP^t + E \quad \text{Eq. 4}$$

$$Y = UQ^t + F \quad \text{Eq. 5}$$

onde, em que  $X$ ,  $Y$  são as matrizes que serão decompostas,  $T$  e  $U$  são matrizes de escores contendo linhas ortogonais ( $T$  e  $U$  são matrizes de  $n$  linhas e  $d$  colunas),  $P$  é a matriz de  $d$  linhas e  $p$  colunas da matriz de pesos da matriz  $X$ ,  $E$  é o erro (resíduo) da matriz  $X$ ,  $Q$  é a matriz de  $d$  linhas e  $m$  colunas da matriz de pesos da matriz  $Y$ ,  $F$  é o erro (resíduo) da matriz  $Y$  [33].

Algumas etapas devem ser seguidas para uma boa construção de um conjunto de calibração com alto poder preditivo e baixo sobreajuste:

- 1º) Pré-processamentos dos dados;
- 2º) Escolha do número de variáveis latentes;
- 3º) Verificação da presença de “outliers”;
- 4º) Cálculo do *RMSEC*.

Os pré-processamentos podem ser utilizados nos casos em que os dados possuem grandezas diferentes [39].

O modelo PLS deve ser gerado a partir dos dados de calibração e as respostas do conjunto de previsão devem ser obtidas sempre com o melhor número de fatores, pois a seleção do número de variáveis latentes no PLS é muito importante para uma boa previsão [33,42].

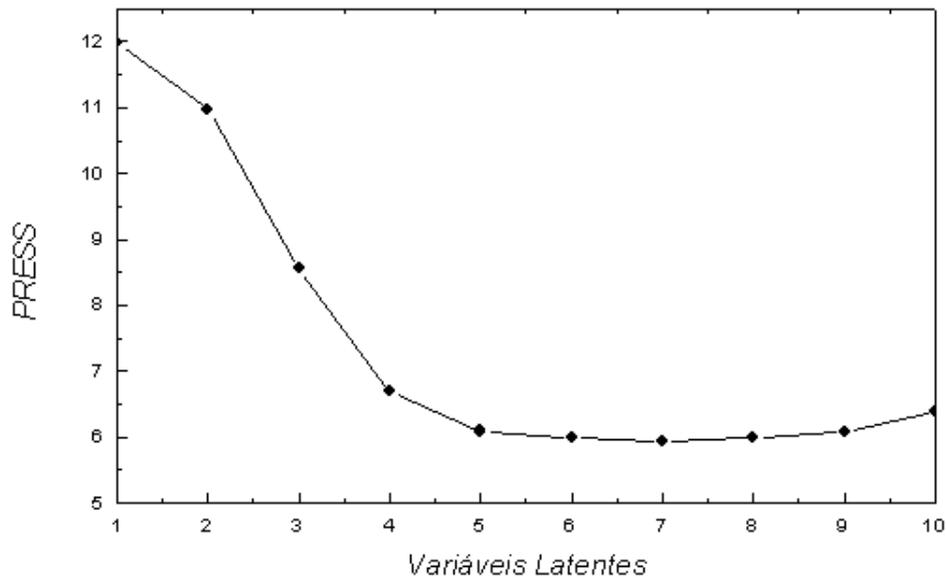
Wakeling and Morris [43] citam em seu estudo que existem somente dois problemas na análise de dados por PLS: o número de dimensões a serem incluídas no modelo e quando este modelo formado se torna significativo. Este estudo mostrou que quanto maior for o número de dimensões incluídas no modelo, maior será o sobreajuste. Um modelo sobreajustado resume bem o conjunto de dados, mas apresenta baixo poder preditivo. Com isso, é necessário um rigoroso teste para a previsão do número de variáveis latentes a serem usadas no modelo PLS [44].

A validação cruzada é um caminho prático para testar a significância preditiva de um modelo. A avaliação dos valores de *PRESS* (*Soma dos Quadrados dos Erros Preditivos Residuais*) obtidos em cada variável latente é feita para se estimar a habilidade preditiva do modelo [44]. Os valores de *PRESS* podem ser calculados a partir da Equação 6 [33]:

$$PRESS = \frac{\sum (Y_{prev} - Y_{real})^2}{N} \quad \text{Eq. 6}$$

onde,  $Y_{prev}$  é o valor calculado pelo modelo e  $Y_{real}$  é o valor verdadeiro, ou obtido por um método de referência e  $N$  é o número de amostras usadas no modelo.

O número correto de variáveis latentes a serem utilizadas no modelo será determinado pelo menor valor de *PRESS* associado às variáveis latentes. A Figura 15 mostra um exemplo de gráfico de variáveis latentes versus *PRESS*. Como se pode observar, os valores de *PRESS* decrescem à medida que o número de variáveis latentes aumenta, até a sétima variável latente. A partir desta variável, os valores de *PRESS* começam a sofrer um ligeiro aumento. Sendo assim, o modelo PLS para este exemplo deve ser construído com sete variáveis latentes.



**Figura 15** – Número de variáveis latentes versus valores de *PRESS* para as dez primeiras variáveis latentes de um modelo PLS.

Após a escolha do número de variáveis latentes para a construção do modelo, deve-se realizar a verificação da presença de “*outliers*”. Esta verificação pode ser feita avaliando-se o gráfico de “*Leverages*” versus Resíduos Studentizados. A “*Leverage*” é a posição das observações das variáveis independentes umas em relação às outras no espaço das  $n$ -variáveis latentes. Um exemplo de gráfico de “*Leverage*” versus Resíduos Studentizados está na Figura 16.

O valor de “*Leverage*” limite (para este exemplo 0,060) pode ser calculado a partir da Equação 7 [45]:

$$\text{Leverage} = 2m / n \quad \text{Eq. 7}$$

onde,  $m$  é o número de variáveis latentes e  $n$  é o de observações.

Analisando as regiões de baixos valores de “*Leverage*”, mas com elevados resíduos pode-se dizer que a amostra está próxima ao centro do conjunto de calibração, mas com um baixo poder preditivo, sendo esta considerada como um “*outlier*” na concentração. Já as amostras concentradas nas regiões de elevado “*leverage*”, mas baixos resíduos, podem demonstrar que o modelo criado não consegue descrever estas amostras, ou que estas possuem baixa similaridade ao restante do conjunto de calibração. Neste caso estas amostras podem ser consideradas como “*outliers*” na curva de destilação. O conjunto de amostras que possui elevados valores de “*leverage*” e com resíduos extremos são consideradas como “*outliers*”. Estes “*outliers*” podem ser eliminados, para que se tenha uma boa habilidade de previsão.

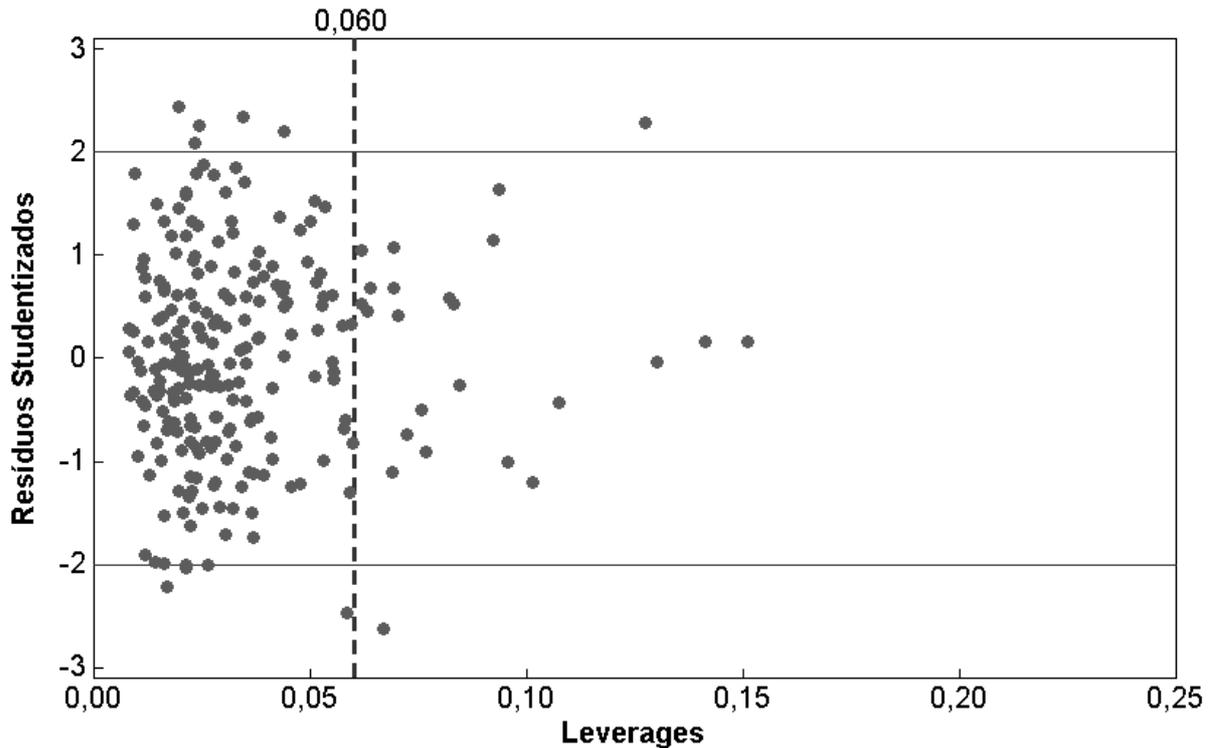


Figura 16 – Gráfico de “Leverages” versus Resíduos Studentizados.

Após a construção do modelo deve-se calcular o valor de *RMSE* (Raiz da Soma dos Erros Quadrados) conforme a Equação 8 [33]:

$$RMSE = \sqrt{\frac{\sum (Y_{prev} - Y_{real})^2}{N}} \quad \text{Eq. 8}$$

O valor de *RMSEC* é o *RMSE* calculado a partir das amostras de calibração sendo esta uma medida de erro na modelagem.

Outra forma de se avaliar a habilidade de previsão de um modelo é a construção de um conjunto de validação externa, também chamado de conjunto de previsão. Este conjunto deve ter aproximadamente 30 % do número de amostras do conjunto de calibração. As variáveis *X* das amostras do conjunto de calibração (para este trabalho, as curvas de destilação) são inseridas na curva de calibração para se obterem os valores de  $Y_{prev}$ . Estes valores de  $Y_{prev}$  são comparados aos valores de  $Y_{real}$  conforme a Equação 6, obtendo-se assim os valores de *RMSEP*.

A calibração multivariada PLS associada a técnicas como infravermelho e cromatografia gasosa foram usadas para prever propriedades de combustível como massa específica [46-48], viscosidade [46,48-50], frações de destilação [46,48,50,51], número de octanas [51,52],

numero de cetano [46,48,51-53], teores de enxofre [54,55], oxigenados [52,56], aromáticos [46,56,57], ponto de ignição [46,49], benzeno [52], MON [47,52,58] e RON [47,52].

Oito parâmetros (ponto de ignição, teor de aromáticos, ponto de congelamento, viscosidade, pontos de destilação (inicial, 10 %, 90 % e final) de querosene foram previstos por Gómez-Carracedo et al [50] aplicando geração de fase vapor, transformada de Fourier no infravermelho médio (FT-MIR) e calibração multivariada PLS. Observou-se neste trabalho que as médias dos erros de previsão foram baixas. Além disso, os valores de repetibilidade e reprodutibilidade foram inferiores aos obtidos a partir dos métodos oficiais, sendo a única exceção relevante a viscosidade.

Oliveira et al [47] determinaram MON e massa específica para gasolina e teor alcoólico e massa específica para o álcool hidratado utilizaram PLS e infravermelho médio. Na determinação do teor alcoólico e massa específica, o método proposto utilizado, comparado ao método do espectro de correlação para a escolha da região espectral, levou à obtenção de modelos PLS com menores valores de *RMSEP*. Para a massa específica foram obtidos os valores de *RMSEP* de 0,0026 e 0,00019 para a gasolina e álcool etílico hidratado respectivamente, enquanto que o valor de *RMSEP* do teor alcoólico no álcool etílico hidratado foi de 0,0399.

Santos et al [48] fizeram um estudo comparativo utilizando técnicas baseadas na espectroscopia de infravermelho (FTIR, FTNIR and FT-Raman) associado ao PLS e ANN com amostras de óleo diesel, para a determinação dos parâmetros massa específica, temperaturas de destilação a 50 e 85 %(v/v) recuperados, viscosidade e teor total de enxofre. Na determinação da massa específica, considerando toda a região espectral, foram obtidos valores de *RMSEP* entre 0,00071 e 0,0029 para as diferentes técnicas, concluindo que a melhor técnica utilizada foi a FTIR-ART1, uma vez que este proporciona uma maior gama espectral, uma melhor relação sinal-ruído e os valores mais baixos de *RMSEP*.

Números de octanas de naftas reformadas cataliticamente foram previstos usando infravermelho médio e PLS por Andrade et al [52] Na construção do modelo de calibração foi observada a presença de dois tipos diferentes de naftas, pelo gráfico das duas primeiras variáveis latentes. Para a nafta de tipo A foram obtidos valores de *RMSEP* para MON e RON de 0,33 e 0,37 enquanto que para a nafta de tipo B 0,35 e 0,42, respectivamente.

## 2. Objetivos

O objetivo principal deste trabalho é a utilização de métodos quimiométricos associados às curvas de destilação, obtidas pela ASTM D86, na avaliação de diferentes parâmetros da gasolina automotiva. Assim, aproveitam-se os resultados obtidos por um ensaio realizado rotineiramente na avaliação da qualidade, para a obtenção de outros parâmetros, diminuindo-se o número de ensaios necessários para a avaliação da sua qualidade e, por consequência, também o tempo de análise e os custos, agilizando o processo analítico.

Desta forma, serão avaliados a origem da gasolina automotiva, bem como o teor alcoólico e a massa específica, importantes parâmetros de qualidade.

A determinação da origem da gasolina automotiva se faz necessária em função do elevado percentual de gasolinas comercializadas com origem não declarada, que podem estar relacionadas a fraudes fiscais, sonegação de impostos e adulteração do combustível.

O teor alcoólico é um dos maiores responsáveis pelas não conformidades da gasolina automotiva no Brasil. Apesar do ensaio ser simples (NBR 13992), são necessários 50 mL de amostra e cerca de 20 minutos para sua realização, que podem limitar o seu uso dependendo da quantidade de amostra disponível.

Apesar de não ter limites estabelecidos pela ANP, a massa específica é um parâmetro muito importante para o bom funcionamento do motor. Além disso, este parâmetro tem se mostrado muito sensível à adulteração pela adição de solventes, sendo um importante indicativo destes procedimentos. Na realização deste ensaio são necessários cerca de um litro de amostra (ASTM D1298), densímetros e termômetros, cuja execução pode ser demorada devido à alta volatilidade da amostra. Para a metodologia ASTM D4052 são necessários apenas 5 mL de amostra e um densímetro automático de elevado custo.

Transformando-se as curvas de destilação (temperatura equivalente em função do percentual do volume evaporado) em uma matriz de dados, serão construídos modelos de calibração que poderão, através da estatística multivariada, determinar a origem da gasolina automotiva ou prever os valores do teor alcoólico e da massa específica, diminuindo o número de ensaios a serem realizados, o volume de amostra necessária, o tempo de análise e seus custos.

### **3. Parte Experimental**

#### **3.1 Amostras de Gasolina**

Para os estudos de origem, determinação de teor alcoólico e massa específica foram usadas gasolinas, comum e aditivada, coletadas pelo LEC-UFMG dentro do PMQC-ANP. As amostras foram armazenadas em frascos de polietileno apropriados e lacrados, e mantidas sob refrigeração (8 a 15 °C) até o momento das análises, para evitar a perda de componentes voláteis.

A origem das amostras foi obtida a partir das informações constantes das notas fiscais de aquisição do produto pelos postos revendedores, apresentadas no ato da coleta. A partir da cidade de origem da distribuidora do combustível, cinco refinarias de origem foram identificadas (REGAP, RLAM, REDUC, REVAP e REPLAN) além de um grande número de amostras com origem desconhecida (amostras cuja origem não pode ser identificada pela falta da nota fiscal de aquisição do produto pelo posto de combustível).

#### **3.2 Reagentes**

No preparo de gasolinas com diferentes teores etanol foi utilizada uma gasolina do tipo A, fornecida pela REGAP, e álcool etílico anidro combustível (AEAC) fornecido pela BR - Distribuidora.

Para a determinação do teor alcoólico das amostras de gasolina foram usadas soluções de cloreto de sódio 10 %(m/v), preparadas utilizando água destilada e cloreto de sódio P.A. da marca Reagen.

#### **3.3 Equipamentos**

As amostras de gasolina foram destiladas em destiladores da marca Herzog modelos HDA 626, 627 e 628, de acordo com a norma ASTM D86 [24].

As medidas de massa específica foram realizadas seguindo-se a norma ASTM D4052 [59], utilizando um densímetro automático Anton Paar, modelo DMA 4500.

### **3.4 Materiais**

Nas destilações das amostras foram utilizados balões de destilação com capacidade para 125 mL (de acordo com a ASTM D86), provetas de vidro de 100 mL com base de metal para medir o volume de amostra e coletar o destilado, termopar tipo PT100 calibrado por organismo credenciado junto à RBC (Rede Brasileira de Calibração), dispositivo centralizador de PT100, placas de suporte do balão de destilação em cerâmica de 38 mm de diâmetro, provetas de 5,0 mL para medir o resíduo, pérolas de vidro e porta gotas.

Na determinação do teor alcoólico (NBR 13992) foram utilizadas provetas de vidro de 100 mL, com tampa de polietileno e resolução de 1 mL, calibradas por organismo credenciado junto à RBC [60].

Na determinação da massa específica pelo método ASTM D4052 [59] utilizou-se seringas de capacidade igual ou superior a 2 mL, adaptáveis ao equipamento, para introdução das amostras. Para o preparo das soluções de gasolina A adicionadas de diferentes quantidades de etanol anidro foram utilizados bureta de 25,00 mL, béqueres de 50 mL e frascos (com tampa) de aproximadamente 20 mL.

Para o preparo das soluções de gasolina do tipo A dopadas com diferentes quantidades de etanol anidro foram usadas buretas de 50,00 mL e provetas de 100,0 mL.

### **3.5 Qualidade dos Resultados dos Ensaios**

Todos os ensaios utilizados neste trabalho são acreditados junto ao INMETRO desde 2006, segundo a norma ISO 17025 [61], garantindo que todos os instrumentos de medição (provetas, medidores de pressão e temperatura, densímetro, etc) são verificados e calibrados periodicamente e que atendem todos os requisitos das normas NBR13992, ASTM D86, D1298 e D4052. A repetibilidade e reprodutibilidade dos ensaios são verificadas periodicamente, segundo a ISO 5725 Parte 2 [62], estando dentro dos limites estabelecidos pelas normas de cada ensaio.

### **3.6 Procedimentos Experimentais**

#### **3.6.1 Ensaio da Destilação (ASTM D86)**

Neste ensaio 100,0 mL da amostra são medidos em uma proveta de vidro e o conteúdo transferido para o balão de destilação apropriado contendo algumas pérolas de vidro, para evitar uma ebulição turbulenta, tomando o cuidado de não derramar nenhum líquido no tubo de

condensação. Um porta-gotas é encaixado na proveta utilizada anteriormente, para que o líquido destilado escoe pela parede interna da proveta não prejudicando a leitura do medidor automático de volume. O conjunto proveta e porta gotas é colocado no compartimento de recolhimento do volume destilado, refrigerado a 15 °C. O balão de destilação é tampado com o centralizador de PT 100 e este sistema é encaixado no destilador, adaptando-se o tubo de condensação do balão ao tubo condensador do equipamento. O tubo de condensação deverá se estender dentro do tubo condensador por uma distância de 25 a 50 mm. O termopar é adaptado ao balão de destilação (que deve estar na posição vertical), tomando-se o cuidado de colocar o bulbo do termopar na saída do tubo de condensação do balão, conforme a norma.

O conjunto de destilação é ajustado e o processo se inicia após o comando do operador no programa de destilação. Após o início do aquecimento, a destilação deve começar entre 5 a 10 minutos. O tempo do ponto inicial de ebulição para os 5 % recuperados deve estar entre 60 a 100 segundos e a taxa de destilação deve ser de 4 a 5 mL min<sup>-1</sup>.

O tempo de recuperação dos 5 mL restantes até o ponto final de ebulição deverá ser de no máximo cinco minutos. Ao final da destilação e do processo de resfriamento, o destilador é desligado e o volume do resíduo contido no balão é medido com o auxílio de uma proveta de 5,0 mL. O volume do resíduo, em mL, é inserido no programa de destilação para que o mesmo faça as correções da pressão atmosférica e do volume do resíduo medido para a obtenção da curva de destilação definitiva.

### **3.6.2 Ensaio do Teor Alcoólico (NBR 13992)**

No ensaio do teor alcoólico, 50,0 mL de gasolina foram medidos em uma proveta limpa e seca de 100,0 mL, com tampa. O volume final da proveta foi atingido com uma solução de NaCl 10 %(m/v), seguido de uma agitação não vigorosa e um repouso de 15 minutos, como descrito na norma NBR 13992 [60]. A leitura do volume final da fase aquosa é realizada após o repouso e expressa em mililitros. Para se calcular o teor de etanol presente na gasolina utiliza-se a Equação 9, no qual A é o volume final da fase aquosa em mililitros.

$$\text{Teor alcoólico (em \%v/v)} = [(A-50) \times 2] + 1 \quad (\text{Eq. 9})$$

quando a diferença (A-50) for inferior a 0,5 mL, deve-se registrar o resultado como menor ou igual a 1 %(v/v) [60].

### **3.6.3 Ensaio da Massa Específica (ASTM D4052 e D1298)**

Para a determinação da massa específica dois métodos podem ser utilizados: o manual (ASTM D1298) e o automático (ASTM D4092).

#### **3.6.3.1 Ensaio Manual (ASTM D1298)**

No ensaio manual, a proveta e o termômetro devem estar a uma temperatura entre aproximadamente 5 °C da temperatura do teste. Transfere-se a amostra para a proveta limpa, sem agitação para evitar a formação de bolhas e minimizar a evaporação dos constituintes mais voláteis da amostra. A proveta deve estar em um local livre de correntes de ar, onde a temperatura média não varie mais que 2 °C durante o tempo do teste.

Insera-se o termômetro e agita-se a amostra com o mesmo, utilizando uma combinação de movimentos verticais e rotacionais para garantir massa específica e temperatura uniformes na proveta. Anota-se a temperatura da amostra com resolução de 0,1 °C e retira-se o termômetro da proveta.

O densímetro é inserido na amostra de modo que ele flutue livremente sem tocar as paredes ou o fundo da proveta. Para líquidos transparentes de baixa viscosidade ou translúcidos, deve-se observar o deslocamento do menisco quando o densímetro é pressionado de 1 a 2 mm abaixo do ponto de equilíbrio e retornado à posição original. Se a posição do menisco variar, deve-se retirar e limpar o densímetro e repetir o processo até que o menisco fique constante. Pressiona-se o densímetro cerca de duas divisões de escala no líquido, liberando com um leve giro de forma que ele não toque as paredes da proveta. Aguarda-se o tempo suficiente para que o densímetro fique em equilíbrio com o líquido. A escala do densímetro deve ser lida até a quarta casa decimal, sendo que para líquidos transparentes a escala é lida no nível da amostra e para líquidos opacos na parte superior do menisco.

Após a leitura da massa específica o densímetro deve ser retirado cuidadosamente da amostra, inserindo-se novamente o termômetro e anotando-se a temperatura com resolução de 0,1 °C. Se esta temperatura diferir da anterior por mais que 0,5 °C devem-se repetir as observações do densímetro e do termômetro até que a temperatura fique estável em 0,5 °C. Se a temperatura estável não for obtida, deve-se colocar a proveta em um banho de temperatura constante e repetir o processo.

Os valores de massa específica obtidos devem ser corrigidos para uma temperatura de 20 °C, segundo uma tabela apropriada [63].

### 3.6.3.2 Ensaio Automático (ASTM D4052)

Na determinação da massa específica pela ASTM D4092, cerca de 0,7 mL de amostra é introduzida em um tubo limpo e seco com o auxílio de uma seringa. Liga-se a iluminação do equipamento e examina-se a amostra no tubo cuidadosamente, assegurando-se que não há bolhas aprisionadas dentro do tubo e que este esteja cheio até um pouco além do ponto de suspensão no lado direito.

Desliga-se a iluminação do equipamento após a introdução da amostra, pois este calor gerado pode afetar a medida da temperatura. Quando o equilíbrio de temperatura estiver estabelecido, anotam-se a massa específica e o valor da temperatura. O valor de massa específica pode ser expresso em  $\text{kg m}^{-3}$  ou  $\text{g mL}^{-1}$  [53].

### 3.7 Procedimentos de Pré-Processamento e Normalização dos Dados

As amostras de gasolina obtidas pelo PMQC-ANP possuem diferentes origens e, conseqüentemente composições [64], que causam pequenas variações de comportamento nas curvas de destilação.

Por isso, nas previsões do teor alcoólico e da massa específica, foram testados dois tipos de normalizações com o intuito de minimizar as diferenças de comportamento das curvas de destilação em função das composições de gasolina de origens distintas.

Em uma das normalizações cada ponto da curva de destilação foi subtraído do ponto inicial de evaporação (*IBP - Initial Boiling Point*), proporcionando curvas de destilação em função da variação de temperatura (*T-IBP*).

Outra normalização testada foi a divisão de cada ponto da curva de destilação pelo ponto final de ebulição (*FBP - Final Boiling Point*), proporcionando valores entre 0 e 1 (*T/FBP*).

Além das normalizações, os pré-processamentos de centralização dos dados na média e autoescalamento foram empregados. Na centralização dos dados, o valor de cada elemento da variável é obtido subtraindo-se o valor original do valor médio. Para o autoescalamento, o valor de cada elemento da variável é obtido subtraindo-se o valor original do valor médio, dividindo-se o resultado anterior pelo desvio padrão.

Na aplicação das técnicas PLS e LDA a validação cruzada do tipo “leave-one-out” foi usada. Esta validação cruzada é feita seguindo os passos a seguir:

- 1 – Calcula-se a matriz de calibração utilizando todas as amostras do conjunto de treinamento, exceto uma;

- 2 – Usa-se a calibração para prever a concentração dos componentes da amostra que foi deixada de fora do conjunto de treinamento;
- 3 – Calcula-se a soma dos erros quadrados (*PRESS*) entre a concentração existente e prevista para a amostra que foi deixada de fora;
- 4 – Retorna-se a amostra excluída para o conjunto de treinamento e retira-se outra amostra do conjunto de treinamento;
- 5 – Calcula-se uma nova calibração para esta nova amostra retirada como feito anteriormente;
- 6 – Retorna-se ao item 2 acima. Adiciona-se o novo valor de *PRESS* calculado no item 3, para o valor de *PRESS* calculado anteriormente. Continua-se este processo até que os valores de *PRESS* de todas as combinações “leave-one-out” sejam computados e somados.

### 3.8 Testes Estatísticos *F* e *t*

O teste *F* é utilizado para comparar a precisão de resultados de dois conjuntos de medidas. Neste caso a comparação é feita a partir dos desvios padrão obtidos em cada conjunto. O valor de *F* calculado ( $F_{calc}$ ) é obtido a partir do quadrado dos desvios padrões como mostrado na Equação 10 [65].

$$F_{calc} = \frac{s_A^2}{s_B^2} \quad \text{Eq. 10}$$

onde,  $s_A^2$  é a variância do conjunto de dados A e  $s_B^2$  é a variância do conjunto de dados B. Neste teste,  $s_A^2$  tem que ser maior que  $s_B^2$ , ou seja, os valores de  $F_{calc}$  sempre serão maiores que 1.

O teste *F* foi aplicado aos valores de *RMSEC* e *RMSEP* nas determinações do teor alcoólico e da massa específica da gasolina automotiva. Isto foi possível, pois os cálculos de *RMSEC* e *RMSEP* (Eq. 8) são semelhantes ao cálculo do desvio padrão.

O teste *t* é usado para comparar dois conjuntos de medidas a fim de se decidir se são ou não diferentes. O nível de probabilidade deve ser escolhido para se concluir se duas medidas diferem uma da outra. O valor de *t* calculado ( $t_{calc}$ ) pode ser calculado pela Equação 11 [66].

$$t_{calc} = \frac{\bar{d}}{\sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}}} \sqrt{n} \quad \text{(Eq. 11)}$$

onde,  $\bar{d}$  é a diferença média entre os métodos a serem comparados,  $d_i$  a diferença entre os métodos para cada amostra e  $n$  o número de pares de dados.

Neste teste, se o valor de  $t_{\text{calc}}$  for maior que o tabelado ( $t_{\text{tab}}$ ) significa que os dois grupos de medidas apresentam resultados diferentes estatisticamente. O valor de  $t_{\text{tab}}$  é calculado de acordo com o nível de significância do teste e o número de graus de liberdade [67].

Na verificação da exatidão do método proposto em comparação aos métodos normatizados o teste  $t$  foi aplicado. Em cada comparação utilizou-se um conjunto de amostras de validação externa (com valores de teor alcoólico e massa específica obtidos pelos métodos normatizados) e os resultados obtidos pela metodologia proposta neste trabalho. Após o cálculo de  $t$  pela Equação 11, os valores de cada modelo foram comparados aos valores de  $t_{\text{tab}}$  a um nível de significância de 95 %.

### **3.9 Cálculo da Repetitividade e Reprodutibilidade**

O cálculo da repetitividade e da reprodutibilidade são importantes ferramentas no controle da qualidade dos resultados de um ensaio. A repetitividade de um ensaio é a diferença entre sucessivos resultados obtidos pelo mesmo operador, no mesmo local, com o mesmo instrumento de medida, sob condições constantes e com o mesmo material teste. Já a reprodutibilidade é diferença entre sucessivos resultados obtidos por diferentes operadores ou diferentes laboratórios ou diferentes equipamentos ou utilizando diferentes materiais teste [59]. Os cálculos da repetitividade e da reprodutibilidade foram realizados segundo a norma ISO/IEC 5725-2 [62].

Para o cálculo da repetitividade e da reprodutibilidade foram utilizadas dez amostras de gasolina. Para cada amostra foram realizados os ensaios físico-químicos por três diferentes analistas, com sete replicadas para cada, resultando em um total de 21 resultados [62].

## **4. Resultados e Discussão**

### **4.1 Determinação da Origem da Gasolina Comercializada no Estado de Minas Gerais**

A determinação da origem da gasolina se faz necessário devido ao elevado número de amostras com origem não declarada. Em 2007, aproximadamente 30 % das amostras coletadas pelo PMQC-ANP não tinham a refinaria de origem declarada [64]. A não declaração da origem pode implicar em amostras adulteradas, ou ainda, a realização de fraudes fiscais. A determinação da origem do combustível ajuda na identificação de amostras desiguais que podem ser associados com adulterações, além de determinar o perfil de combustível, visto que cada refinaria apresenta um perfil diferente devido a processos e/ou tipos de petróleo usados na produção da gasolina.

#### **4.1.1 Aplicação da PCA**

Para aplicação do método PCA, os dados foram autoescalados gerando uma matriz de media zero e variância um. Eles foram projetados em dimensões menores por meio de combinações lineares das variáveis originais gerando as matrizes de escores e pesos. Um total de 2148 gasolinas de diferentes origens declaradas foi utilizado (Tabela 3) na construção do modelo PCA, sendo o maior percentual de amostras relacionado à REGAP e o menor percentual à RLAM.

Para cada componente principal do modelo construído, associou-se uma porcentagem de variância explicada. A soma de cada porcentagem de variância explicada de cada componente resultou na porcentagem de variância acumulada. Na Tabela 4 pode-se visualizar os percentuais de variância explicada de cada componente principal e da variância acumulada das vinte primeiras componentes do modelo PCA. Os resultados mostram que das 71 variáveis originais, somente seis componentes principais são necessárias para explicar 92,2 % da variância total do modelo. Estas 71 variáveis utilizadas compreendem às temperaturas equivalentes entre o ponto inicial de ebulição até o ponto correspondente a 59 % do volume evaporado e entre 85 e 93 % do volume evaporado, além da temperatura do ponto final de ebulição.

**Tabela 3** – Amostras de gasolina utilizadas na determinação da origem.

Origem	Quantidade de amostras	Percentual de amostras
RLAM	17	0,8
REGAP	1890	88,0
REDUC	54	2,5
REPLAN	146	6,8
REVAP	41	1,9

**Tabela 4** – Número de componentes principais, porcentagem de variância explicada (%VE) e porcentagem de variância explicada acumulada (%VE<sub>acum</sub>).

Componente	%VE	% VE <sub>acum</sub>	Componente	%VE	% VE <sub>acum</sub>
1	65,8	65,8	11	0,3	95,1
2	12,3	78,1	12	0,3	95,4
3	9,2	87,3	13	0,3	95,7
4	2,4	89,7	14	0,2	95,9
5	1,3	91,0	15	0,2	96,1
6	1,2	92,2	16	0,2	96,3
7	0,9	93,1	17	0,2	96,5
8	0,7	93,8	18	0,2	96,7
9	0,6	94,4	19	0,1	96,8
10	0,4	94,8	20	0,1	96,9

A Figura 17 apresenta o número de componentes que possuem maior variância explicada e que devem ser utilizadas na construção dos gráficos para a visualização dos agrupamentos das amostras segundo a sua origem. Os resultados mostram que são necessárias somente quatro componentes para a visualização dos agrupamentos.

Na Figura 18, tem-se a visualização dos dados com as três primeiras componentes principais, na Figura 19 observa-se PC1xPC2 e, na Figura 20, PC1xPC3. A Figura 18 foi gerada com o objetivo de se visualizar a separação das amostras de acordo com a refinaria de origem, uma vez que a terceira componente principal apresenta uma variância explicada bastante significativa (9,2 %).

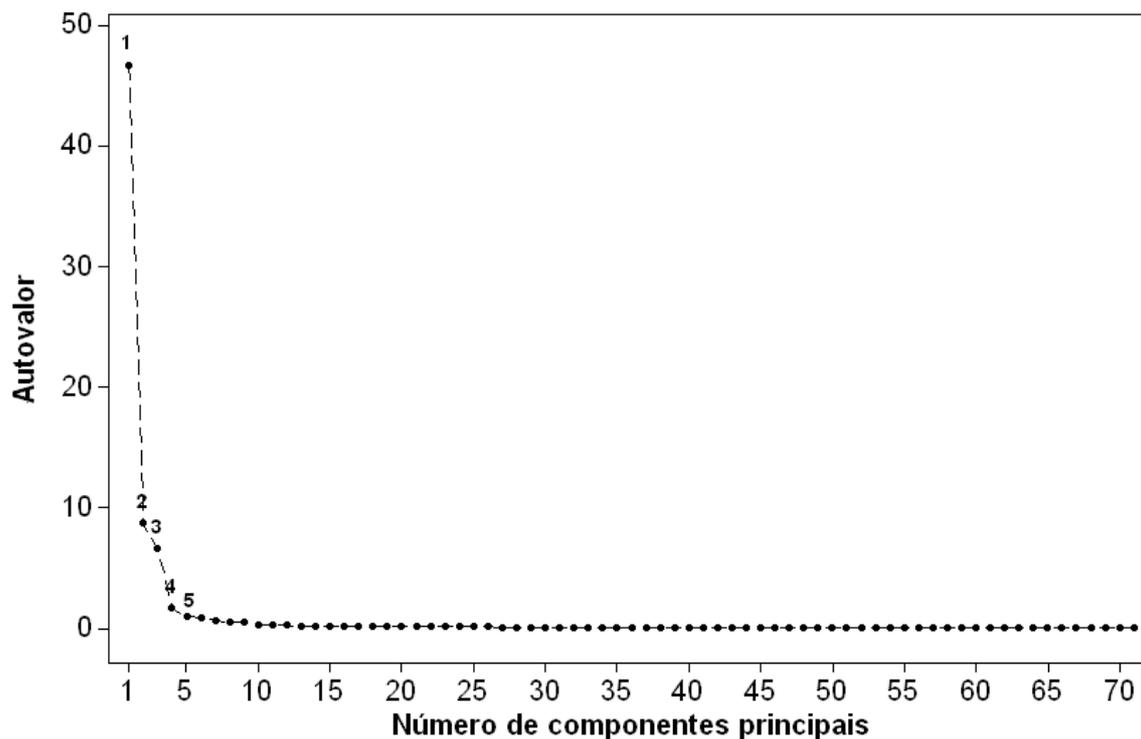


Figura 17 – Gráfico do número de componentes principais versus autovalor.

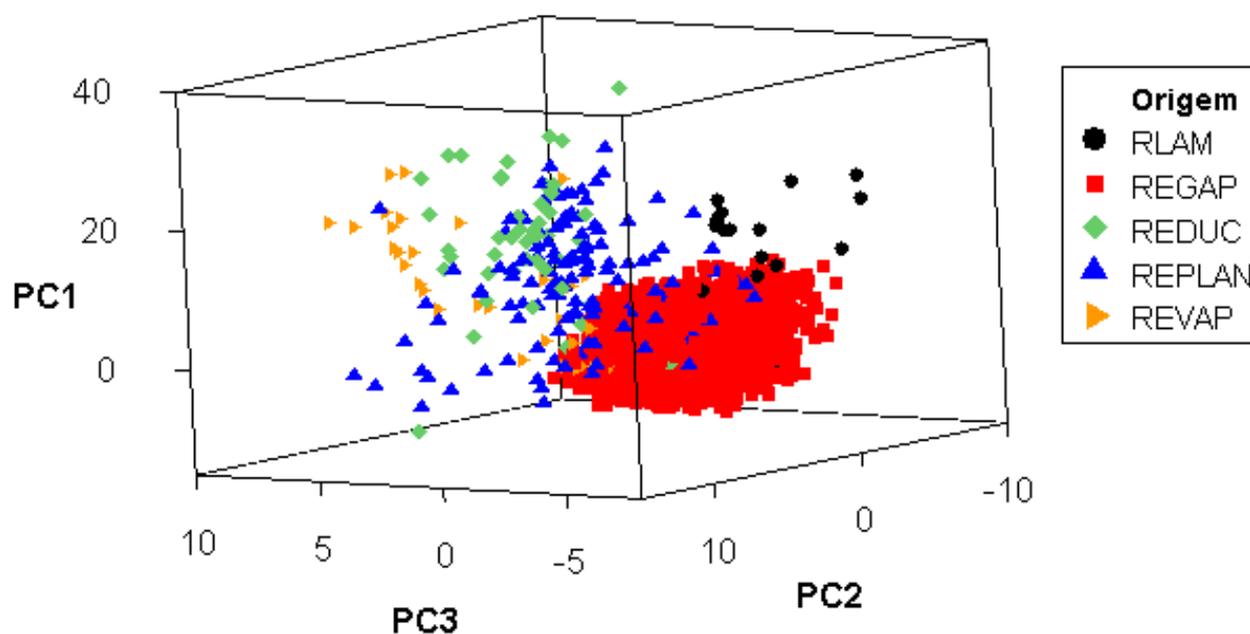
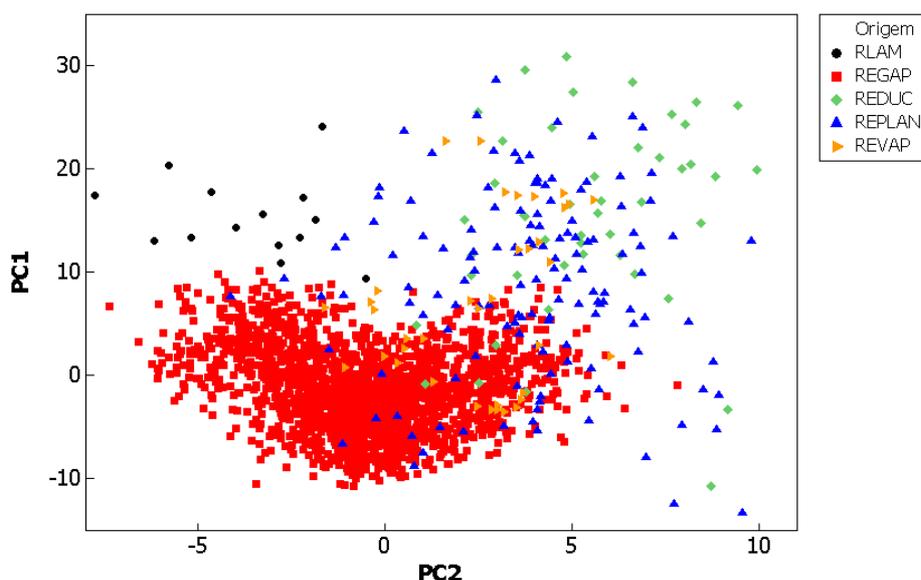
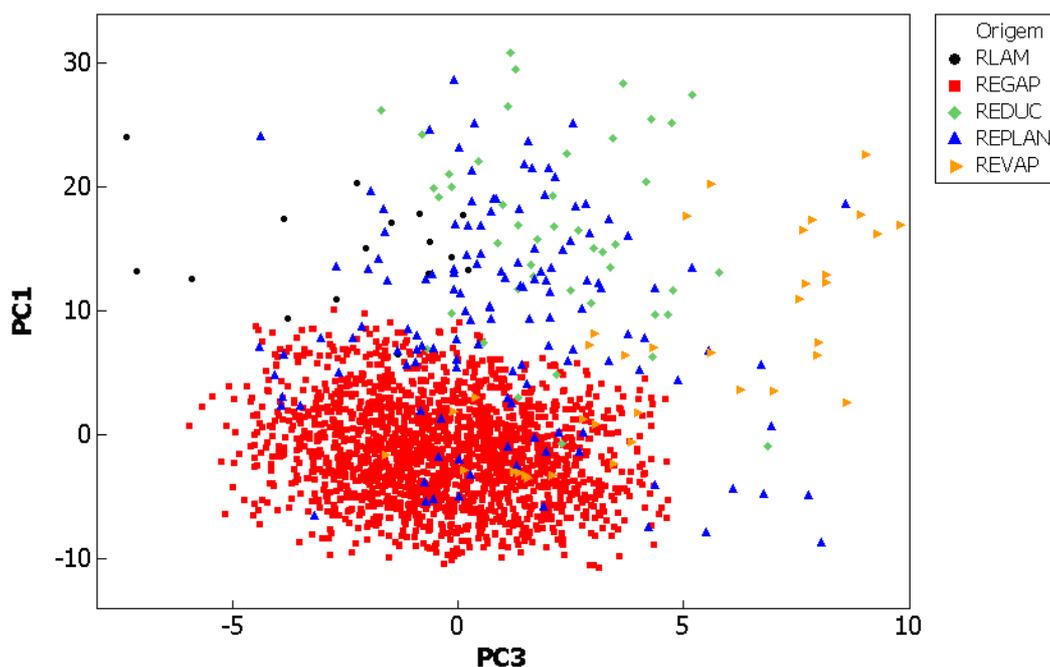


Figura 18 – Gráfico dos escores das três primeiras componentes principais (PC1, PC2 e PC3).

As Figuras 19 e 20 apresentam uma melhor visão das separações das origens por cada componente principal. Na Figura 19, PC1 foi a principal responsável pela separação das refinarias RLAM e REGAP, além da separação das amostras provenientes das refinarias REDUC e REGAP. A segunda componente principal, PC2, separou a refinaria RLAM das refinarias REPLAN, REVAP e REDUC. Na Figura 20, a combinação entre PC1 e PC3 permitiu visualizar a separação entre o conjunto das amostras de REDUC das amostras da refinaria REVAP.



**Figura 19** – Gráfico dos escores das duas primeiras componentes principais (PC1 e PC2).

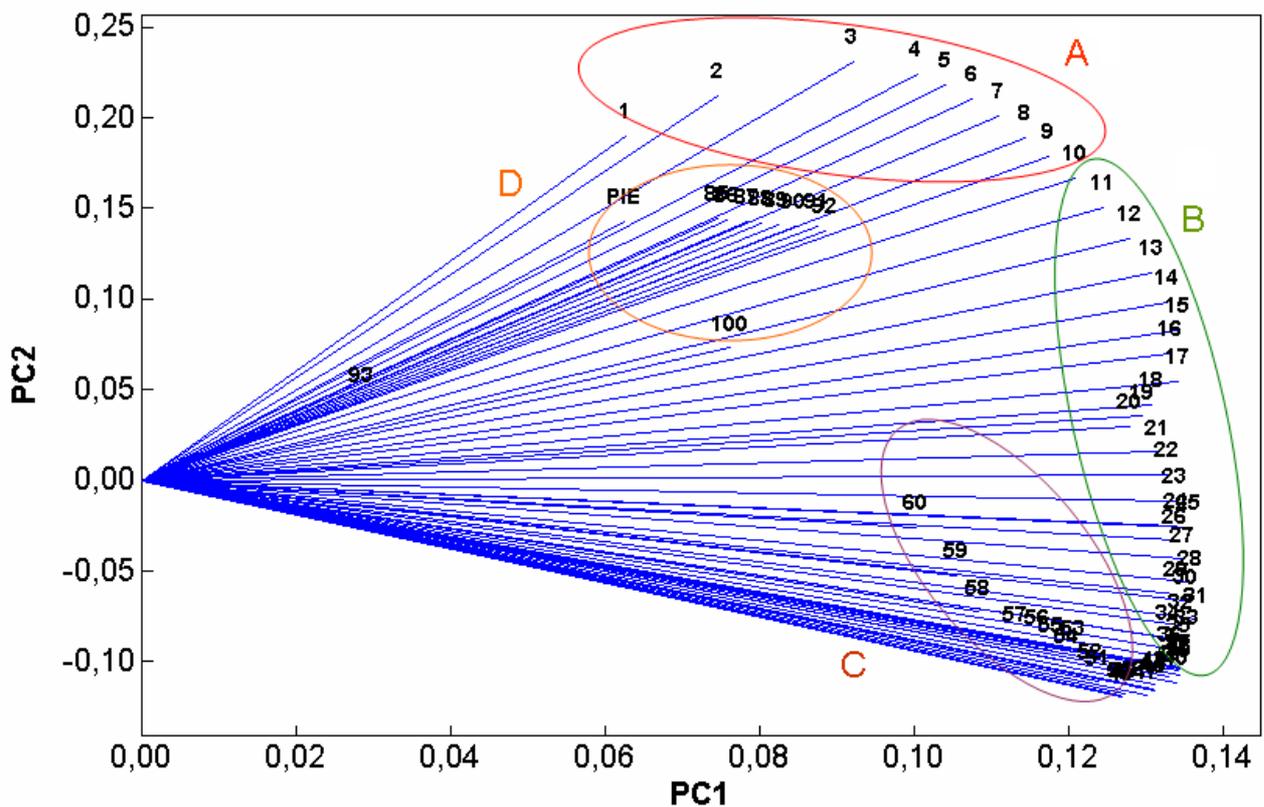


**Figura 20** – Gráfico dos escores da primeira e terceira componentes principais (PC1 e PC3).

As amostras da refinaria REGAP formaram o maior grupo do conjunto de dados, com características bastante semelhantes quando comparada aos demais grupos observados na Figura 18.

O gráfico dos pesos para o modelo PCA foi construído para se observar a importância de cada variável nas PC1 e PC2, visto que estas apresentaram uma maior percentagem de variância explicada (Figura 21).

Na Figura 21, a primeira componente principal mostrou que a região B apresentou a maior importância para a separação das origens, pois PC1 apresentou o maior percentual de variância explicada. Já a região A foi a de maior importância para construção do modelo em PC2. Na região entre 11 a 50 % de volume evaporado ocorre, preferencialmente, a liberação da fração leve da gasolina, podendo estar relacionado aos hidrocarbonetos alifáticos leves (como hexanos) e uma pequena quantidade de hidrocarbonetos olefínicos (como hexenos) [30,68].



**Figura 21** – Gráfico dos pesos das duas primeiras componentes principais.

A) 1 a 10 %, B) 11 a 50 %, C) 50 a 60 % e D) 85 a 92 % de volume evaporado e pontos inicial e final de destilação.

Pela Figura 21 percebe-se que as variáveis 10 e 50 % de volume evaporado que possuem limites estabelecidos pela ANP [7] apresentam elevados pesos em PC1, mostrando que estes parâmetros são importantes não só na avaliação da qualidade do combustível, mas também na determinação de sua origem.

#### 4.1.2 Aplicação da LDA

Para aplicação do método LDA, a refinaria de origem das amostras foi previamente informada. Para a construção do modelo LDA foi necessário utilizar um número de amostras maior que o número de variáveis [13]. Isto se deve a limitação da técnica quimiométrica visto que o modelo não pode ser gerado se o número de amostras for inferior ao de variáveis. O modelo LDA foi obtido a partir do mesmo conjunto amostral do modelo PCA.

O sumário de classificação, na Tabela 5, apresentou o grupo REGAP como o de maior percentual de acerto. Observou-se que um elevado número de amostras da REVAP foi classificado como sendo da refinaria REPLAN ou da REGAP mostrando que um elevado número de amostras pode ter tido a origem informada incorretamente pelo posto revendedor.

**Tabela 5** – Sumário de classificação das amostras usando validação cruzada.

	<i>RLAM</i>	<i>REGAP</i>	<i>REDUC</i>	<i>REPLAN</i>	<i>REVAP</i>
<i>RLAM</i>	15	11	1	0	2
<i>REGAP</i>	2	1848	2	18	3
<i>REDUC</i>	0	5	41	5	3
<i>REPLAN</i>	0	19	5	105	4
<i>REVAP</i>	0	7	5	18	29
<b>Total (N)</b>	17	1890	54	146	41
<b>Acerto (%)</b>	88,2	97,8	75,9	71,9	73,1

As distâncias médias entre os grupos compostos por diferentes refinarias foram calculadas e podem ser visualizadas na Tabela 6.

**Tabela 6** – Distância média entre os grupos no modelo LDA com validação cruzada.

Origens	Distância média
RLAM/REDUC	33,0
RLAM/REGAP	23,8
RLAM/REVAP	33,1
RLAM/REPLAN	35,6
REDUC/REPLAN	29,3
REDUC/REGAP	57,6
REDUC/REVAP	22,5
REGAP/REPLAN	21,9
REGAP/REVAP	26,6
REPLAN/ REVAP	12,1

Os resultados mostram que o par REPLAN/ REVAP apresentou uma menor distância enquanto o par REDUC/REGAP apresentou a maior distância. Isso significa que as amostras provenientes das refinarias REPLAN e REVAP possuem comportamento similar para a curva de destilação, enquanto as amostras da REDUC e da REGAP apresentam comportamento diferenciado.

O percentual de acerto com a validação cruzada, no modelo LDA utilizando a validação cruzada, foi de 95,1 %. Na determinação da origem de gasolina utilizando técnicas quimiométricas e ensaios físico-químicos (massa específica, teor alcoólico, octanagem, teores de hidrocarbonetos olefínicos, parafínicos e aromáticos, além de quatro temperaturas de destilação) realizado por Barbeira et al. [13], observou-se resultado semelhante com emprego de cinco componentes principais para a descrição do modelo.

Estes resultados mostram que o uso das curvas de destilação em conjunto com a análise multivariada pode ser utilizado com sucesso na determinação da origem da gasolina automotiva. A utilização de somente um ensaio para esta análise proporciona redução no tempo de análise, além de menores custos, simplificando o processo analítico, quando comparado ao de Barbeira et al [13].

#### 4.1.2.1 Classificação de Gasolinas com Origem não Declarada

Amostras com origem desconhecida podem encobrir problemas relacionados a adulterações e fraudes fiscais. Em 2004, o aumento das amostras atípicas pode estar relacionado ao aumento do número de amostras com origem desconhecida [64].

A análise discriminante linear foi empregada na determinação das amostras de origem desconhecida, ou seja, amostras em que a origem não foi declarada no momento da coleta do combustível. Desta forma, o modelo construído no item 4.1.2 foi utilizado para classificar um conjunto de 1205 amostras com origem desconhecida. A classificação destas amostras é mostrada na Tabela 7.

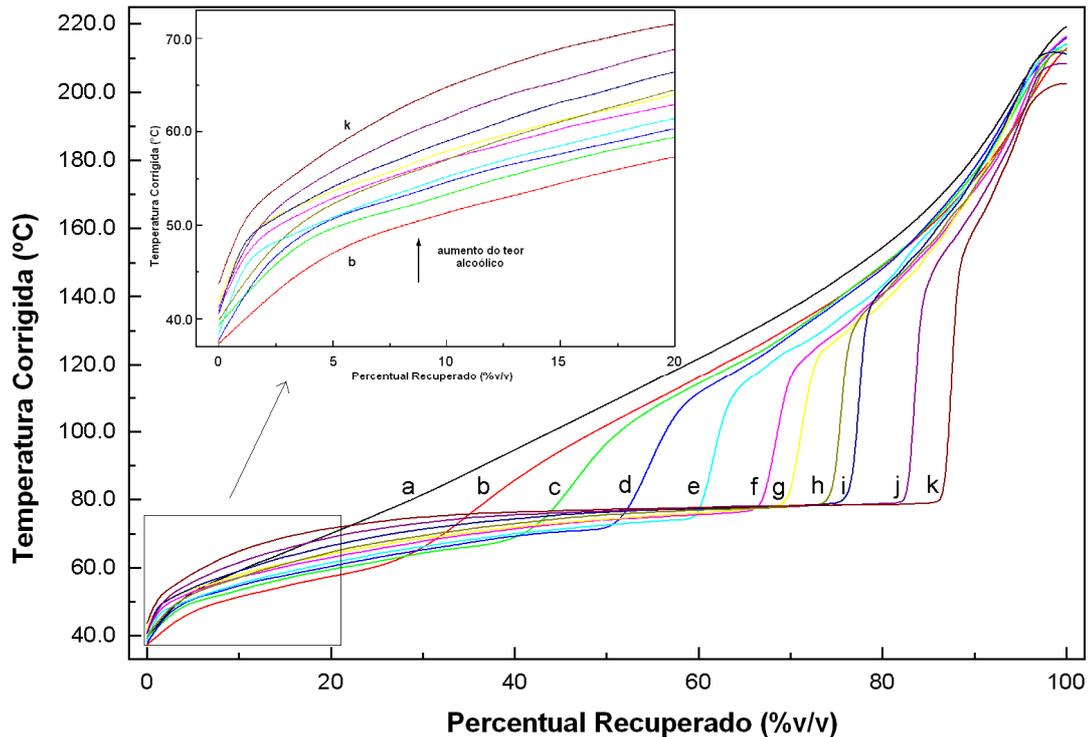
A análise discriminante mostrou que a maioria das amostras desconhecidas tem a REGAP como provável origem, como as amostras utilizadas no modelo. Além disso, os percentuais de origem para estas amostras desconhecidas são distintos daqueles das amostras utilizadas na construção do modelo, como pode ser observado na Tabela 7. Este fato mostra que, a quantidade de amostras dentro de cada conjunto das refinarias não produziu uma classificação tendenciosa das amostras desconhecidas.

**Tabela 7** – Classificação das amostras desconhecidas de acordo com o modelo LDA.

<b>Origem</b>	<b>Amostras com Origem Declarada (%)</b>	<b>Amostras com Origem não Declarada (%)</b>
REGAP	88,0	79,6
REPLAN	6,8	9,5
REVAP	1,9	4,4
REDUC	2,5	3,6
RLAM	0,8	2,9

#### 4.2 Determinação do Teor Alcoólico da Gasolina

Um estudo preliminar da influência do teor alcoólico na curva de destilação da gasolina foi realizado. Curvas de destilação foram obtidas a partir de uma amostra de gasolina do tipo A dopada com diferentes quantidades de etanol no intervalo de 0 a 60 % (v/v). A Figura 22 mostra as alterações na curva de destilação provocadas pela adição de etanol.



**Figura 22** – Gráfico percentual recuperado *versus* temperatura corrigida para uma gasolina com diferentes teores alcoólicos. a) 0; b) 5; c) 10; d) 15; e) 20; f) 25; g) 30; h) 35; i) 40; j) 50 e k) 60 % (v/v).

No início da destilação, na faixa de 0 a 20 % de volume destilado, observa-se que a adição de etanol provoca uma diminuição da temperatura de ebulição devido à formação de azeótropos de mínima na mistura hidrocarbonetos-etanol. A formação de azeótropos nestas misturas se deve ao fato de que misturas de etanol com hidrocarbonetos apresentam grandes desvios na mistura ideal, tal como definido pela lei de Raoult [26]. Percebe-se ainda que, nas curvas de **b** a **k**, a temperatura de ebulição aumenta à medida que o teor alcoólico aumenta. Um platô foi observado nas curvas de destilação devido à adição de etanol na gasolina, uma vez que boa parte do etanol adicionado que não formou azeótropos, sofre ebulição na faixa de 75 a 80 °C. O aumento da concentração de etanol no sistema provoca uma mudança na inclinação e na altura dos platôs, sendo que quanto maior é a quantidade de etanol no sistema maior é a inclinação.

A partir deste estudo, dois conjuntos de amostras foram propostos com objetivo de se determinar o teor alcoólico utilizando curvas de destilação e calibração multivariada PLS. O primeiro conjunto de amostras foi construído a partir de uma gasolina, isenta de álcool, adicionada de concentrações de etanol anidro na faixa de 15 a 30 % (v/v) (Gasolina Dopada com Etanol, *GDE*). O segundo foi composto por amostras coletadas em postos revendedores pelo Programa de Monitoramento da Qualidade da ANP, com concentrações de etanol na faixa de 19 a 28 % (v/v) (Gasolina Programa Monitoramento, *GPM*). Como as amostras coletadas possuem diferentes origens, que afetam diretamente as propriedades do combustível [13,40], procurou-se obter um conjunto de dados representativo que contivesse amostras de diferentes refinarias.

O conjunto *GDE* apresentou a matriz de dados da calibração com 34 amostras enquanto a matriz de previsão foi elaborada com 14 amostras. No conjunto de amostras *GPM* foram utilizadas 94 amostras para a calibração e outras 36 amostras na previsão.

Na previsão do teor alcoólico, para cada conjunto de dados, ou seja, *GDE* e *GPM*, os pré-processamentos e normalizações citados anteriormente foram aplicados além da validação cruzada “*leave-one-out*”.

O teste para prever o número de variáveis a serem utilizadas no modelo foi o do *PRESS*. O número de variáveis latentes que obteve o menor valor de *PRESS* foi o escolhido na construção de cada modelo em cada conjunto amostral [45].

Na Figura 23, para o conjunto *GDE* observou-se que somente duas variáveis foram utilizadas na construção do modelo independente da normalização usada. Já para o conjunto *GPM* foram usadas de quatro a cinco variáveis latentes, dependendo da normalização utilizada.

Os valores de porcentagem de variância explicada para as dez primeiras variáveis latentes de cada grupo, para cada normalização (Tabela 8), mostram que as normalizações empregadas nos conjuntos *GDE* e *GPM* alteraram o comportamento da curva de destilação no modelo PLS com relação aos dados puros.

Para cada variável latente, diferentes valores de porcentagem de variância explicada foram obtidos. No conjunto *GDE*, duas variáveis latentes foram responsáveis por 72,4 % dos dados não normalizados, 77,1 % para a normalização *T-IBP* e 79,0 % para *T/FBP*. Já no conjunto *GPM*, cinco variáveis latentes foram responsáveis por 85,7 % nos dados não normalizados e 90,3 % na normalização *T-IBP*. Neste conjunto quatro variáveis latentes explicaram 93,3 % da variância total do modelo para a normalização *T/FBP*. Observou-se que, para ambos os conjuntos, geralmente a primeira variável latente foi a responsável pela maior porcentagem de variância explicada.

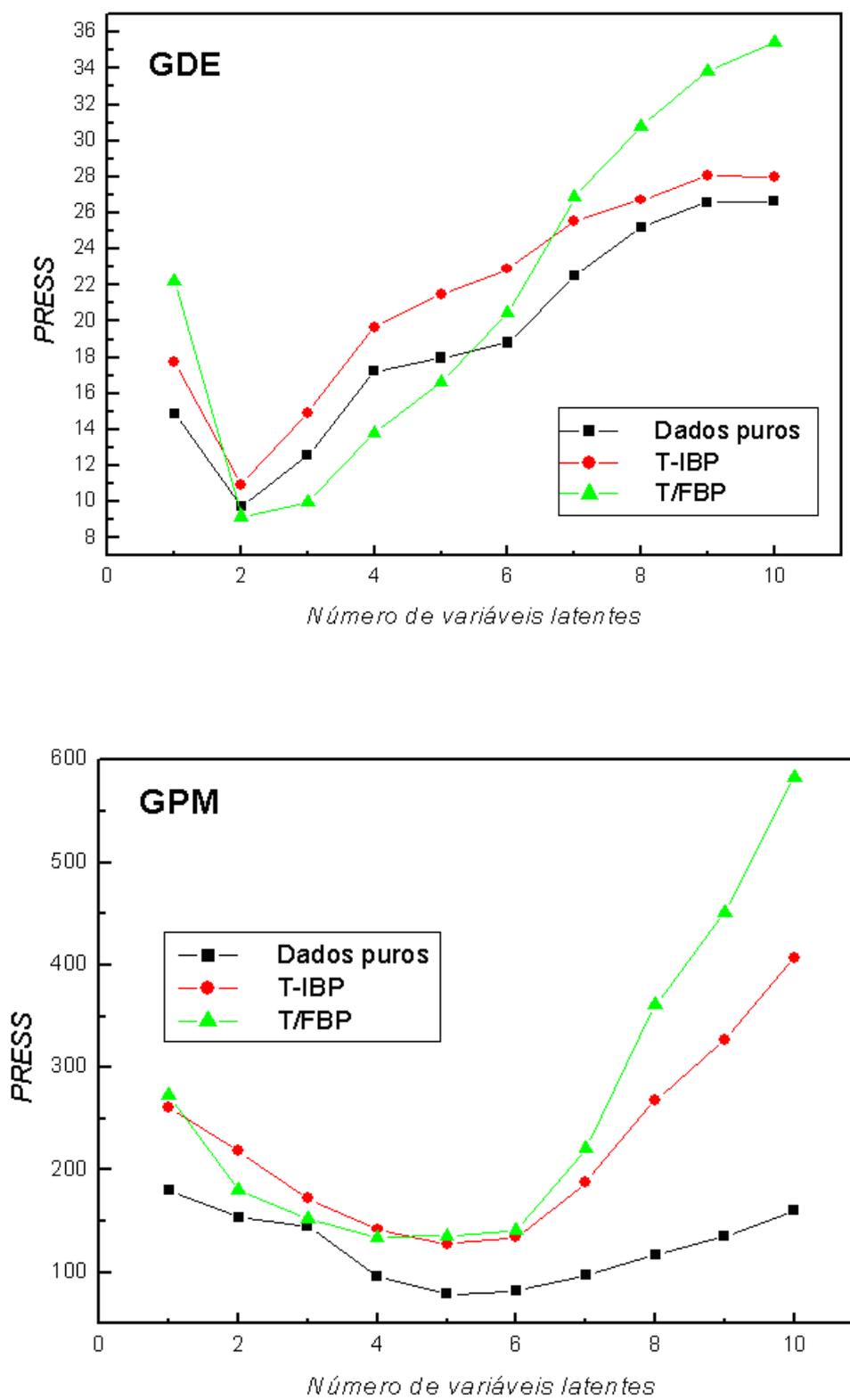


Figura 23 – Gráfico de número de variáveis latentes versus *PRESS* dos conjuntos *GDE* e *GPM*.

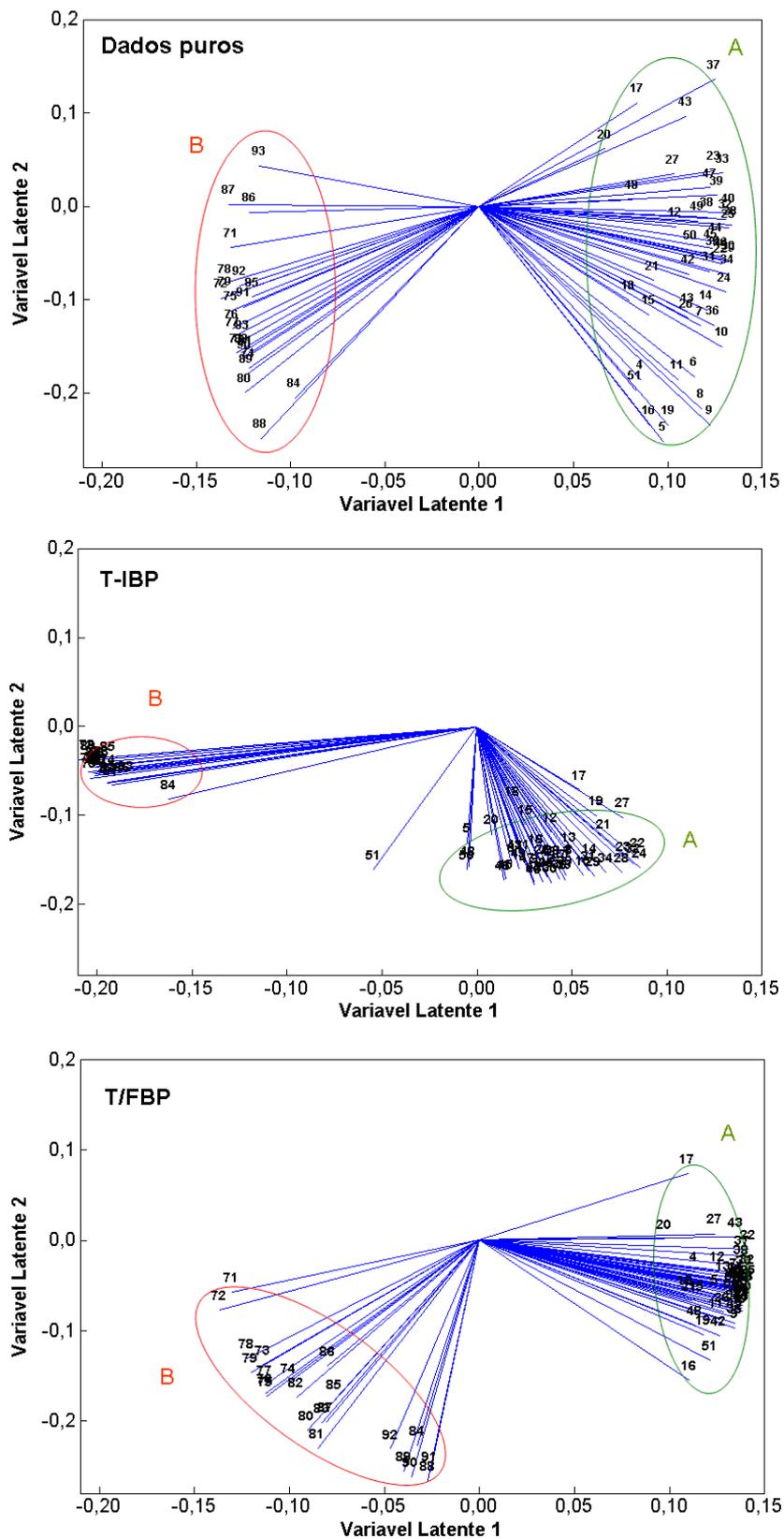
**Tabela 8** – Valores de porcentagem de variância explicada das dez primeiras variáveis latentes no conjunto de calibração na determinação do teor alcoólico nos conjuntos *GDE* e *GPM*.

VL	<i>GDE</i>			<i>GPM</i>		
	Dados	<i>T-IBP</i>	<i>T/FBP</i>	Dados	<i>T-IBP</i>	<i>T/FBP</i>
1	68,5	29,5	66,2	27,4	38,8	49,7
2	3,9	47,6	12,8	20,6	47,3	39,3
3	1,8	2,5	2,9	32,7	2,2	2,7
4	2,6	2,0	1,6	3,6	1,0	1,6
5	3,7	1,6	1,6	1,3	0,9	0,9
6	2,0	1,2	1,1	1,0	1,2	0,5
7	1,3	0,8	0,8	0,8	0,7	0,4
8	1,2	1,0	1,0	0,9	0,5	0,3
9	0,8	1,2	1,2	0,7	0,5	0,4
10	1,4	1,4	1,0	0,8	0,4	0,3

Para cada normalização realizada no conjunto de dados *GDE* obteve-se um gráfico de pesos (Figura 24), mostrando que em cada normalização o comportamento da curva de destilação foi modificado e conseqüentemente diferentes gráficos de pesos foram obtidos. Um comportamento semelhante entre estes gráficos foi a separação dos percentuais de volume evaporado em dois grupos: um de 4 a 51 % (v/v) e outro de 71 a 93 % (v/v).

No gráfico dos dados puros (Figura 24) obteve-se a separação dos percentuais de volume evaporado em forma de "X". A região A está relacionada à parte inicial da curva, entre 4 e 51 % (v/v) e a região B relacionada à parte final da curva de destilação, entre 71 e 93 % (v/v). Os resultados mostram que todas as variáveis utilizadas possuem pesos semelhantes na construção do modelo PLS na primeira variável latente, responsável por 68,5 % da variância explicada (Tabela 8), mostrando-se a mais importante.

A normalização *T-IBP* resultou em uma modificação nos percentuais de variância explicada das variáveis latentes, onde a segunda passa a ter um maior percentual (47,6 %) em relação à primeira (29,5 %), conforme Tabela 8. Comparando com o gráfico dos dados puros pode-se visualizar que a parte relacionada à região A apresenta maior peso na segunda variável latente, enquanto a região B teve seu peso diminuído.



**Figura 24** – Gráfico dos pesos das duas primeiras variáveis latentes na determinação do teor alcoólico do conjunto *GDE*. A) 4 a 51 B) 71 a 93 %(v/v).

Para a normalização *T/FBP* a primeira variável latente apresenta o maior percentual de variância explicada (66,2 %), conforme a Tabela 8. No gráfico dos pesos desta normalização (Figura 24) observa-se que a região B apresenta variáveis com menor peso em VL1 quando comparado aos obtidos com os dados puros, enquanto a região A apresentou peso semelhante.

Os gráficos de pesos para cada normalização no conjunto de dados *GPM* é mostrado na Figura 25. Da mesma forma que para o conjunto *GDE*, houve a divisão dos percentuais de volume evaporado em duas regiões, 4 a 51 e 71 a 93 %(v/v). No entanto, os percentuais de variância explicada para o conjunto de dados *GPM*, foram bastante distintos dos obtidos para o conjunto *GDE*, não apresentando uma variável latente responsável pelo maior percentual da variância explicada.

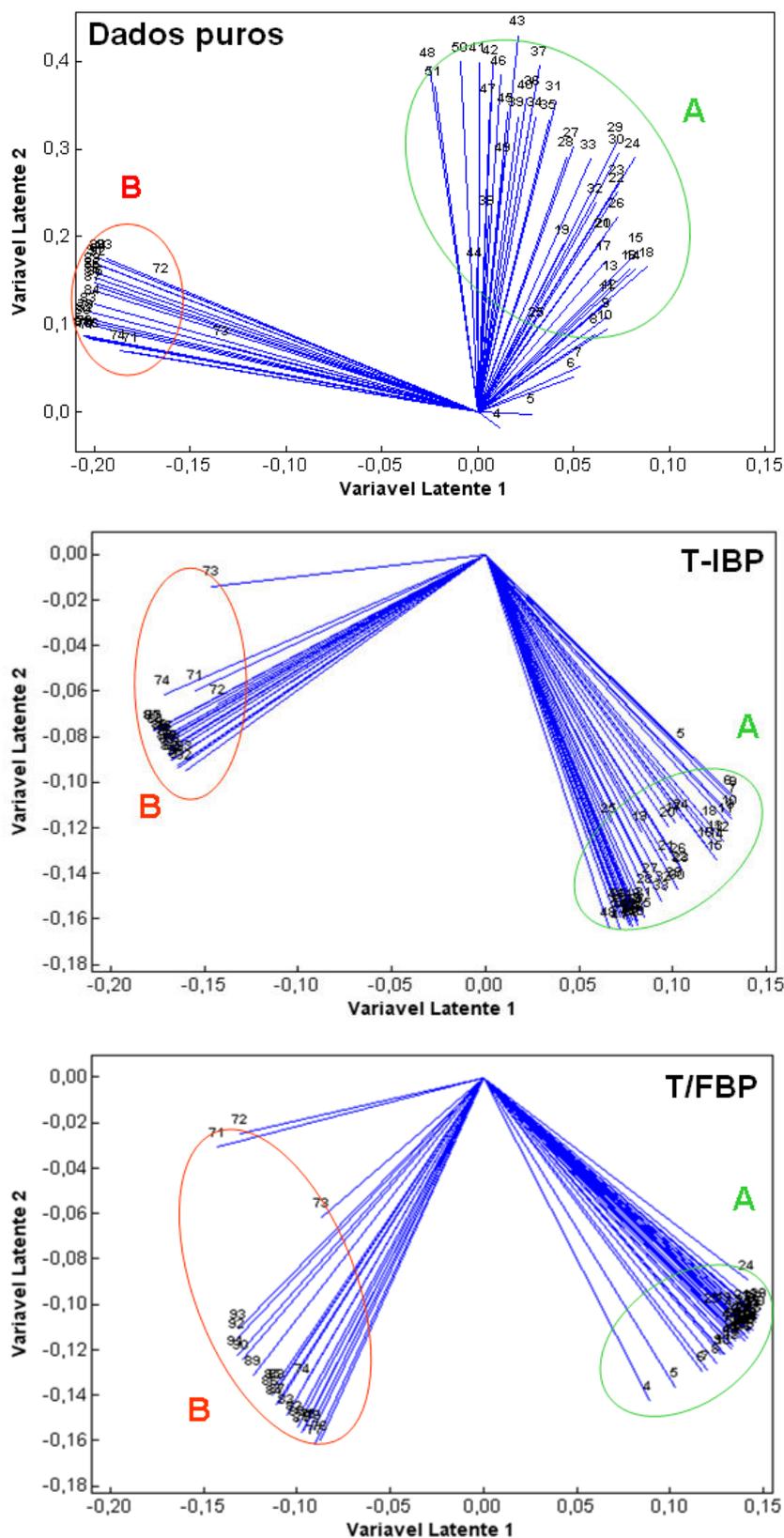
Para os dados puros do conjunto *GPM*, as três primeiras variáveis latentes apresentam percentuais de variância explicada semelhantes (Tabela 8), um comportamento bem diferente do conjunto *GDE*. Esta maior complexidade é resultante das diferenças de composição das gasolinas de origens distintas que compõe o conjunto *GPM*. No gráfico dos pesos (Figura 25), as regiões A e B possuem pesos distintos nas duas primeiras variáveis latentes, diferentemente dos dados puros do conjunto *GDE*.

Para a normalização *T-IBP* as duas primeiras variáveis latentes apresentaram percentuais de variância explicada semelhantes e responsáveis por 86,1 % do total. No gráfico dos pesos (Figura 25) a região A foi a que apresentou maior peso nas duas primeiras variáveis latentes, mostrando-se a mais importante para explicação do modelo.

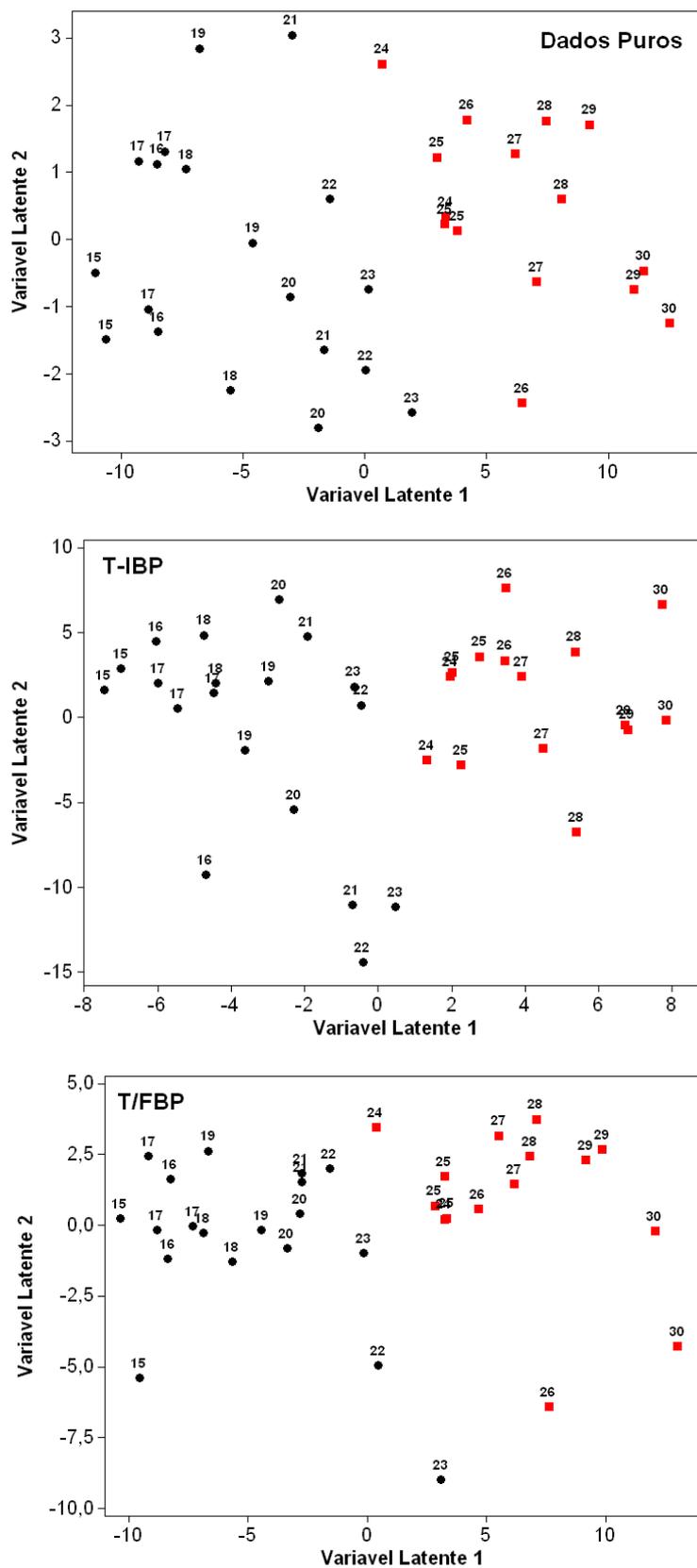
Para os dados normalizados segundo *T/FBP*, as duas primeiras variáveis latentes também apresentaram percentuais de variância explicada semelhantes e são responsáveis por 89,0 % do total. As duas regiões, A e B, apresentaram pesos semelhantes para as duas primeiras variáveis latentes (Figura 25).

É importante salientar que o modelo *GDE* foi construído a partir da adição de diferentes quantidades de etanol em uma mesma gasolina A, enquanto que o conjunto *GPM* foi formado por diversas gasolinas oriundas de diferentes refinarias, formando este um modelo mais robusto em relação ao anterior, refletindo a diversidade de gasolinas comercializadas no Estado de Minas Gerais. Por estes estudos não é possível definir qual das duas regiões (A e B) tem maior importância na explicação dos modelos uma vez que, na maioria dos casos as duas têm pesos equivalentes nas duas primeiras variáveis latentes.

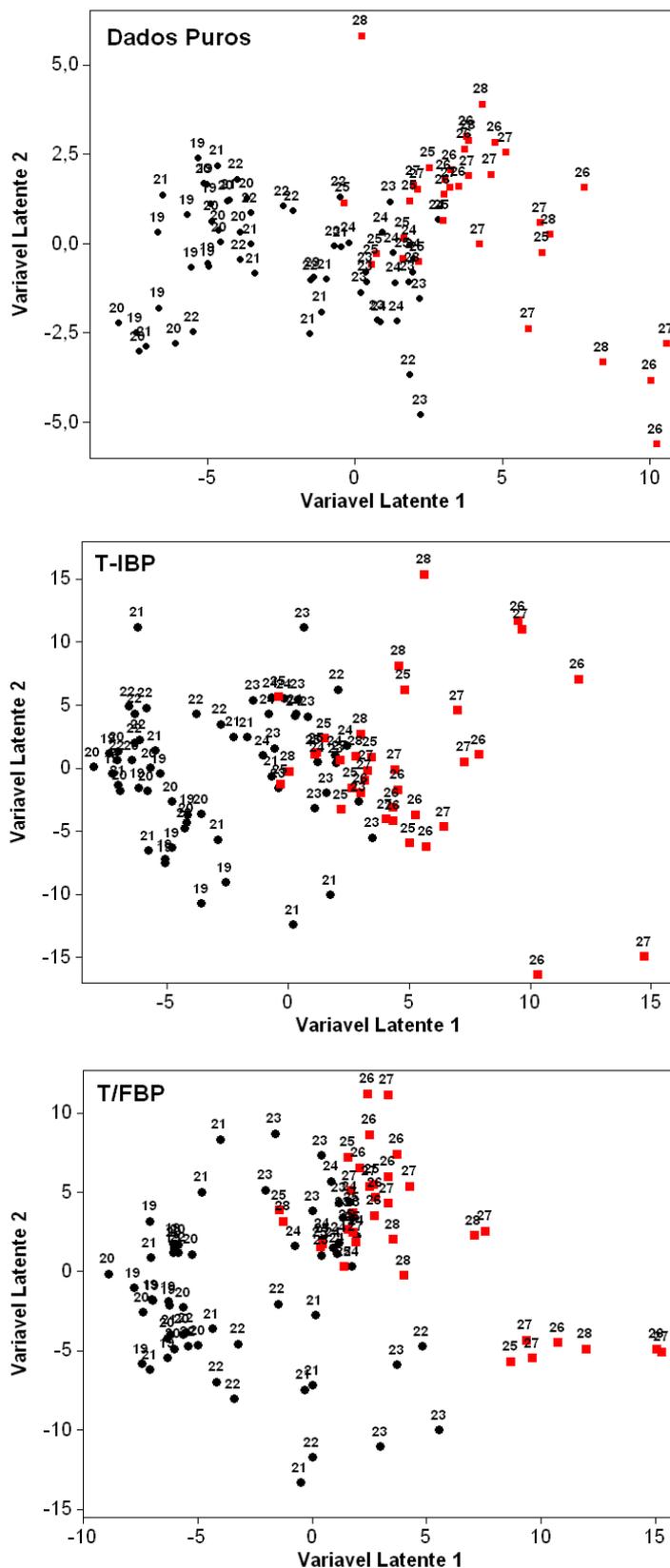
Após a construção dos modelos de calibração percebeu-se que as amostras dos conjuntos *GDE* e *GPM* apresentaram separação de grupos com relação aos diferentes teores alcoólicos nas duas primeiras variáveis latentes (Figuras 26 e 27).



**Figura 25** – Gráfico dos pesos das duas primeiras componentes na determinação do teor alcoólico do conjunto *GPM*. A) 4 a 51 B) 71 a 93 %(v/v).



**Figura 26** – Gráfico das duas primeiras variáveis latentes mostrando a separação das amostras do conjunto *GDE* em duas faixas de teor alcoólico. (●) 15-23 %(v/v); (■) 24-30 %(v/v).



**Figura 27** – Gráfico das duas primeiras variáveis latentes mostrando a separação das amostras do conjunto *GPM* em duas faixas de teor alcoólico. (●) 19-24 % (v/v); (■) 25-28 % (v/v).

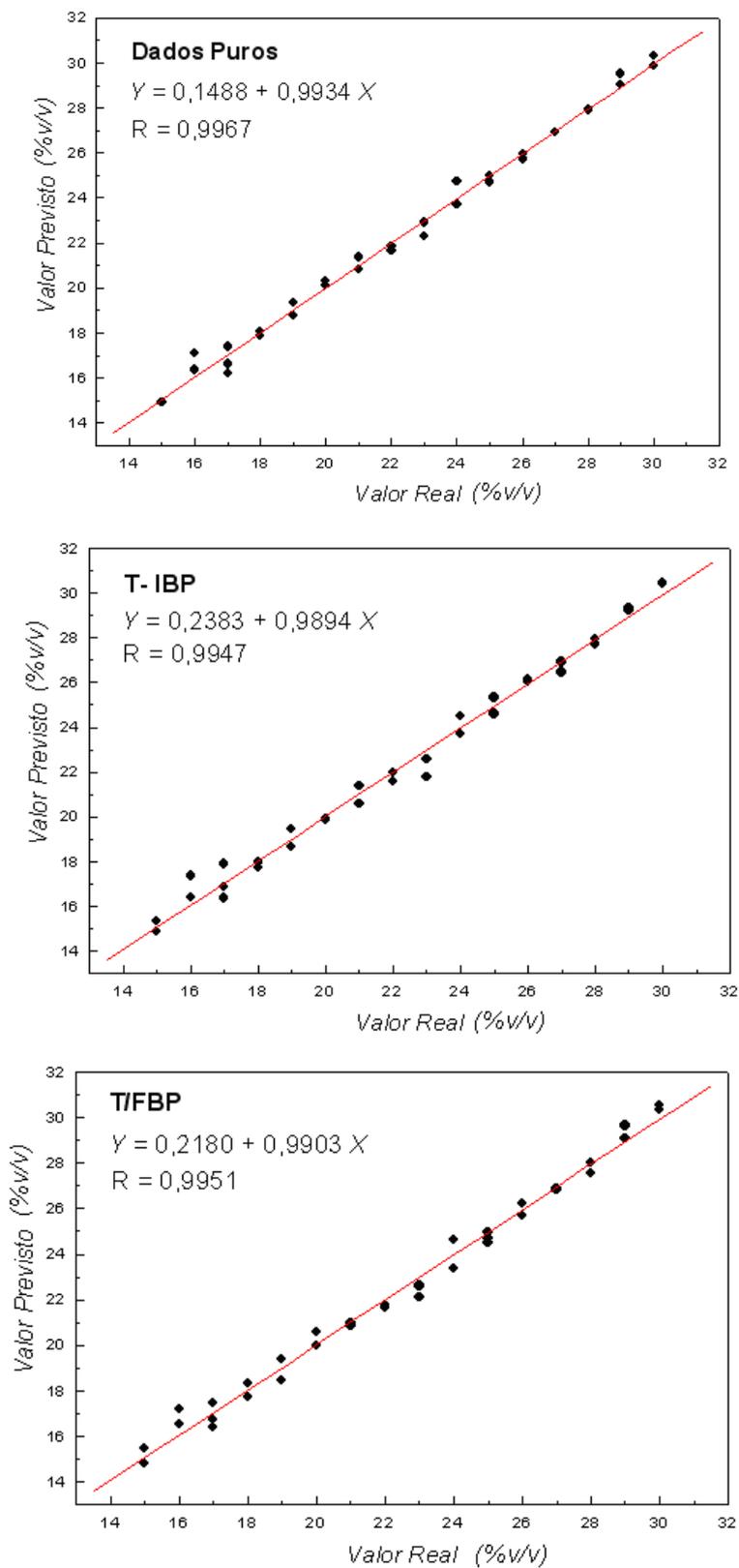
Para o conjunto *GDE* (Figura 26), o conjunto de dados foi separado em dois grupos com teores alcoólicos na faixa de 15 a 23 %(v/v) e 24 a 30 %(v/v), independentemente da normalização empregada. Já para o conjunto *GPM* (Figura 27) a separação foi na faixa de 19 a 24 %(v/v) e 25 a 28 %(v/v), com uma melhor separação para o conjunto dos dados puros. A separação das amostras em dois conjuntos de teor alcoólico pode estar relacionada aos diferentes azeótropos hidrocarboneto-etanol formados, além das variações na pressão de vapor das misturas etanol-gasolina, proporcionadas pelas diferentes quantidades de etanol adicionadas [28,29,69].

A pressão de vapor é uma importante propriedade que fornece indicações do comportamento do combustível sob diferentes condições de operação em motores de ignição por centelha, estando diretamente relacionada com a volatilidade da amostra. Combustíveis com alta pressão de vapor apresentam altas taxas de emissão de compostos voláteis [28].

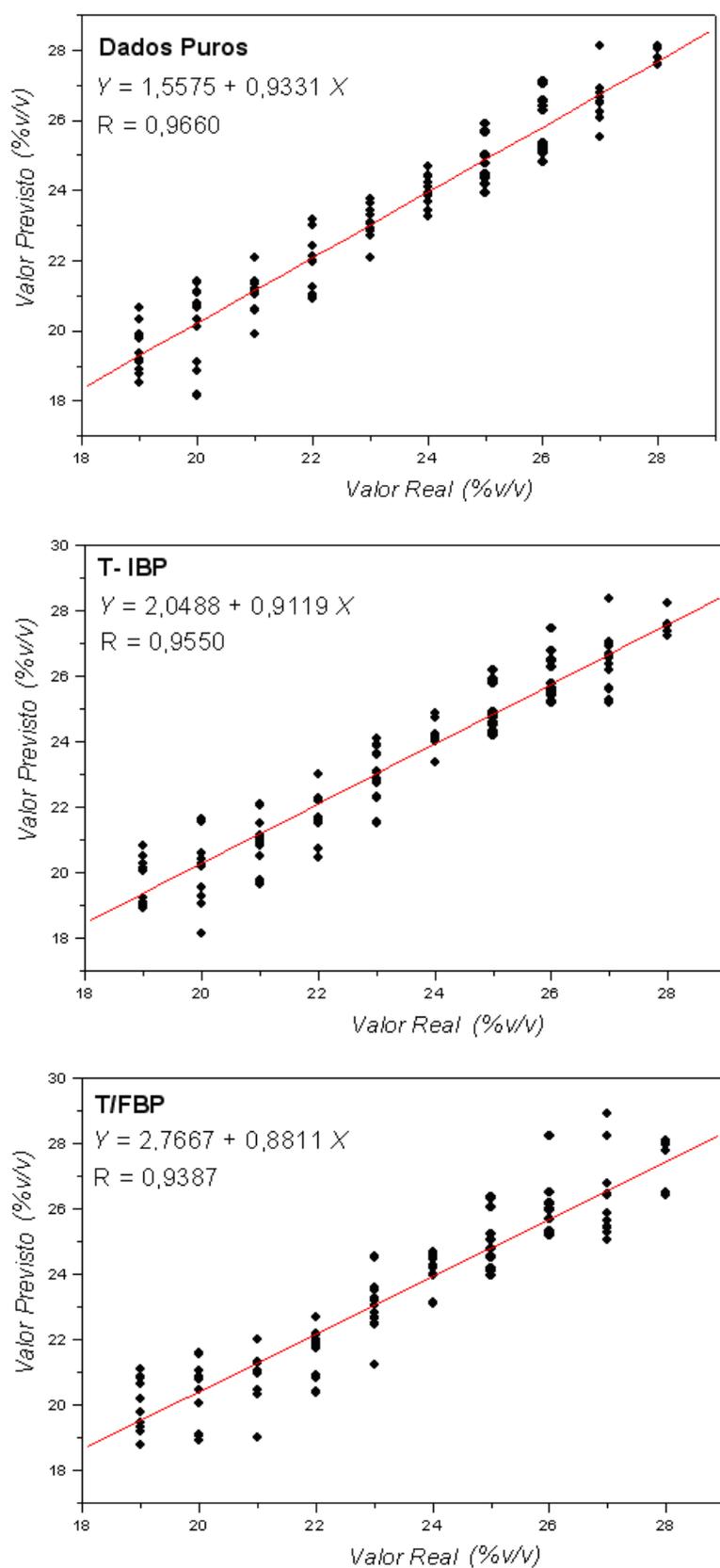
O estudo de Pumphrey et al [29] demonstrou que a adição de etanol causa modificações na pressão de vapor do sistema. Após a realização de diversas medições de pressão de vapor de misturas gasolina-etanol com diferentes concentrações, observaram-se dois comportamentos distintos no sistema, como pode ser visualizado na Figura 14. O primeiro foi o aumento sucessivo da pressão de vapor no intervalo de 0 a 15 %(v/v) e o segundo, um declínio sucessivo a partir de 15 %(v/v) de etanol adicionado. Nesta região de declínio foram observados diferentes decréscimos de pressão de vapor, sendo que na faixa de 15 a 25 %(v/v) observou-se um decréscimo de  $0,159 \text{ kPa } ^\circ\text{C}^{-1}$  e no intervalo de 25 a 30 %(v/v),  $0,110 \text{ kPa } ^\circ\text{C}^{-1}$ . Este resultado mostra o comportamento não linear do etanol em misturas com hidrocarbonetos apolares da gasolina. Com isso, o comportamento de misturas etanol-gasolina não pode ser explicado utilizando a abordagem de comportamento ideal, ou seja, a lei de Raoult não pode ser usada neste caso. Apesar de vários estudos citarem este fato em seus relatos experimentais ou em seus estudos teóricos, este fenômeno necessita de maior exploração, visto que a estrutura de hidrocarbonetos formando azeótropos com etanol na mistura etanol-gasolina ainda é desconhecida [69].

Os valores de teor alcoólico das gasolinas foram previstos em cada modelo com diferentes normalizações e comparados aos valores reais. Os três gráficos, do conjunto *GDE*, apresentaram inclinações e coeficientes de correlação da reta (*R*) bastante semelhantes indicando um bom ajuste (Figura 28).

Já para o conjunto *GPM* (Figura 29), as inclinações obtidas foram menores que as observadas para o conjunto *GDE*, além de apresentarem valores de *R* bem menores, devido à maior dispersão dos pontos. O maior coeficiente de correlação da reta foi obtido para os dados puros, mostrando-se o mais indicado. Estes fatos mostram a grande influência da diversidade da composição das amostras do modelo e a ineficiência das normalizações para minimizá-las.



**Figura 28** – Valor Real versus Valor Previsto para os modelos normalizados ou não do grupo GDE.



**Figura 29** – Valor Real versus Previsto para os modelos normalizados ou não do grupo *GPM*.

Os valores de *RMSEC* e *RMSEP* foram calculados para cada conjunto de amostras a partir dos valores reais e previstos dos modelos de calibração e validação, respectivamente, e são apresentados na Tabela 9. Os valores mostram que o conjunto *GDE* apresentou o menor valor de *RMSEC* para o modelo nos quais os dados não foram normalizados, da mesma forma que o conjunto *GPM*. Para verificar se há alguma diferença estatisticamente significativa entre os valores de *RMSEC* obtidos para as diferentes normalizações, foi realizado o teste F.

Para o conjunto *GDE* o teste F (Tabela 10) mostrou que todos os valores de *RMSEC* não diferem estatisticamente no nível testado, mostrando que qualquer normalização neste caso pode ser usada. Já para o conjunto *GPM*, somente os dados normalizados segundo *T/FBP* apresentaram diferença estatisticamente significativa dos dados não normalizados, a um nível de confiança de 95 %.

A exatidão de cada modelo construído pode ser verificada pelo cálculo do *RMSEP* [53]. De acordo com a Tabela 9, os menores valores de *RMSEP* foram obtidos também para os modelos cujos dados não sofreram pré-processamento ou normalização.

**Tabela 9** – Valores de *RMSEC*, *RMSEP* obtidos para diferentes modelos construídos com diferentes conjuntos (*GDE* e *GPM*) na determinação do teor alcoólico na gasolina a partir das curvas de destilação.

<i>Pré-tratamento e normalização</i>	<i>GDE</i>		<i>GPM</i>	
	<i>RMSEC</i>	<i>RMSEP</i>	<i>RMSEC</i>	<i>RMSEP</i>
<b>Dados puros</b>		0,463		0,838
<b>Dados centrados na média</b>	0,374	0,553	0,792	0,839
<b>Dados autoescalados</b>		0,552		0,953
<b><i>T/FBP</i></b>		0,540		0,967
<b><i>T/FBP</i> centrados na média</b>	0,453	0,610	0,923	0,966
<b><i>T/FBP</i> autoescalados</b>		0,665		1,134
<b><i>T-IBP</i></b>		0,486		1,004
<b><i>T-IBP</i> centrados na média</b>	0,473	0,571	0,954	1,008
<b><i>T-IBP</i> autoescalados</b>		0,784		1,067

A aplicação do teste F nos valores de *RMSEP* (Tabela 10) obtidos mostrou que para o conjunto *GDE*, exceto para o pré-processamento de autoescalamento juntamente com a normalização *T-IBP*, todos os valores de *RMSEP* encontrados foram equivalentes a um nível de confiança de 95 %. No modelo formado pelo conjunto *GPM*, não houve diferença significativa entre os valores encontrados, a um nível de confiança de 95 %, exceto para os dados autoescalados e normalizados segundo *T/FBP*.

Estes resultados mostram, novamente, que as normalizações não apresentaram o efeito desejado, ou seja, de proporcionar modelos com menores erros que os dos dados puros, sejam de calibração ou validação.

**Tabela 10** – Valores de  $F_{calc}$  (nível de confiança de 95 %) para *RMSEC* e *RMSEP* nos conjuntos *GDE* e *GPM*, na determinação do teor alcoólico na gasolina a partir das curvas de destilação.

	<i>GDE</i>		<i>GPM</i>	
	$F_{calc}$ <i>RMSEC</i>	$F_{calc}$ <i>RMSEP</i>	$F_{calc}$ <i>RMSEC</i>	$F_{calc}$ <i>RMSEP</i>
<b>Dados x Dados</b>		1,00 <sup>NS</sup>		1,00 <sup>NS</sup>
<b>Dados centrados na média x Dados</b>	1,00 <sup>NS</sup>	1,43 <sup>NS</sup>	1,00 <sup>NS</sup>	1,01 <sup>NS</sup>
<b>Dados autoescalados x Dados</b>		1,42 <sup>NS</sup>		1,17 <sup>NS</sup>
<b><i>T/FBP</i> x Dados</b>		1,36 <sup>NS</sup>		1,27 <sup>NS</sup>
<b><i>T/FBP</i> centrados na média x Dados</b>	1,47 <sup>NS</sup>	1,73 <sup>NS</sup>	1,61 *	1,29 <sup>NS</sup>
<b><i>T/FBP</i> autoescalados x Dados</b>		2,06 <sup>NS</sup>		1,82 *
<b><i>T-IBP</i> x Dados</b>		1,10 <sup>NS</sup>		1,27 <sup>NS</sup>
<b><i>T-IBP</i> centrados na média x Dados</b>	1,60 <sup>NS</sup>	1,52 <sup>NS</sup>	1,24 <sup>NS</sup>	1,28 <sup>NS</sup>
<b><i>T-IBP</i> autoescalados x Dados</b>		2,87 *		1,60 <sup>NS</sup>
<b><i>Graus de Liberdade</i></b>	33	13	93	34
<b><i>Valores de <math>F_{tab}</math></i></b>	1,79	2,58	1,40	1,78

NS – não significativo; \* - significativo

#### 4.2.1 Validação do Método Proposto

Para o conjunto de amostras *GDE* avaliou-se a exatidão do método proposto, comparando-se os resultados obtidos com os teores alcoólicos obtidos a partir dos volumes de etanol adicionados à gasolina A. Já para o conjunto *GPM* comparou-se o método proposto com os resultados obtidos pela NBR 13992. Os valores de  $t_{calc}$  são apresentados na Tabela 11.

O método da proveta, NBR 13992, considera que os valores de teores alcoólicos devem ser expressos sem casas decimais e que, quando o volume de etanol extraído for menor que 0,5 mL registra-se o resultado como menor ou igual a 1 %(v/v). Já pelo método proposto, a calibração multivariada proporciona resultados com mais de uma casa decimal. A partir destas considerações, para a validação do método proposto, foram utilizados os valores de teores alcoólicos obtidos a partir do conjunto *GPM* com e sem arredondamentos.

O valor de  $t_{tab}$  a 95 % de confiança para o conjunto de amostras do conjunto *GDE* foi de 2,16, enquanto que os valores de  $t_{calc}$  ficaram entre 0,02 e 1,37 (Tabela 11). Desta forma, pode-se concluir que não houve diferença significativa, a um nível de confiança de 95 %, entre todos os modelos PLS do conjunto *GDE*, e as quantidades de etanol adicionadas à gasolina A.

Já o valor de  $t_{tab}$  a 95% de confiança para o conjunto de amostras do conjunto *GPM* foi de 2,03 e o valor de  $t_{calc}$  ficou na faixa de 0,07 a 0,71 para os valores de teor alcoólico não arredondados (uma casa decimal) e 0,16 a 1,00 para os valores arredondados (sem casas decimais).

Com estes resultados pode-se concluir que não houve diferença significativa, a um nível de confiança de 95 %, para todos os modelos PLS, dos conjuntos *GDE* e *GPM*, baseados nas curvas de destilação comparados à quantidade de etanol adicionada e ao método NBR 13992, respectivamente.

Para se verificar qual o modelo mais indicado para a análise do teor alcoólico das amostras de gasolina comercializadas nos postos revendedores, foi estudada o comportamento dos dois conjuntos a um mesmo conjunto de amostras.

**Tabela 11** – Valores de  $t_{calc}$  para a comparação dos resultados obtidos pelos conjuntos *GDE* e *GPM*, em relação à quantidade de etanol adicionada e à NBR 13992, respectivamente.

<b>Pré-tratamento e normalização</b>	<b>GDE</b>	<b>GPM s/ arredondamento</b>	<b>GPM c/ arredondamento</b>
<b>Dados puros</b>	0,02 <sup>NS</sup>	0,07 <sup>NS</sup>	0,18 <sup>NS</sup>
<b>Dados centrados na média</b>	1,36 <sup>NS</sup>	0,71 <sup>NS</sup>	0,53 <sup>NS</sup>
<b>Dados autoescalados</b>	1,37 <sup>NS</sup>	0,66 <sup>NS</sup>	0,53 <sup>NS</sup>
<b>T/FBP</b>	0,36 <sup>NS</sup>	0,11 <sup>NS</sup>	0,16 <sup>NS</sup>
<b>T/FBP centrados na média</b>	1,04 <sup>NS</sup>	0,63 <sup>NS</sup>	1,00 <sup>NS</sup>
<b>T/FBP autoescalados</b>	0,83 <sup>NS</sup>	0,53 <sup>NS</sup>	0,54 <sup>NS</sup>
<b>T-IBP</b>	0,16 <sup>NS</sup>	0,35 <sup>NS</sup>	0,35 <sup>NS</sup>
<b>T-IBP centrados na média</b>	1,24 <sup>NS</sup>	0,63 <sup>NS</sup>	0,63 <sup>NS</sup>
<b>T-IBP autoescalados</b>	0,56 <sup>NS</sup>	0,56 <sup>NS</sup>	0,56 <sup>NS</sup>
<b>Graus de Liberdade</b>	13	34	34
<b>Valores de <math>t_{tab}</math></b>	2,16	2,03	2,03

NS – não significativo; \* - significativo

#### 4.2.2 Verificação da Proximidade do Comportamento das Amostras dos Conjuntos *GDE* e *GPM* na Determinação do Teor Alcoólico

Com o objetivo de verificar a proximidade do comportamento das amostras dos dois conjuntos de gasolinas, associado à curva de destilação, produziu-se um modelo PLS utilizando a curva de calibração do conjunto *GDE* e as amostras de validação do conjunto *GPM*.

Percebe-se que as amostras possuem comportamentos distintos na curva de destilação, devido aos altos valores de *RMSEP* encontrados (Tabela 12), quando comparados aos obtidos anteriormente (Tabela 9). Isto se deve a baixa variabilidade de informação contida nas amostras do conjunto *GDE* quando comparada as do conjunto *GPM* que possuem diferentes origens e, por conseqüência, diferentes composições.

**Tabela 12** – Valores de  $RMSEP$  e  $t_{calc}$  a 95% de confiança, obtidos para diferentes modelos construídos na determinação do teor alcoólico da gasolina com conjunto de calibração  $GDE$  e validação  $GPM$ .

<b>Pré-tratamento e normalização</b>	<b><math>RMSEP</math></b>	<b><math>t_{calc}</math></b>
<b>Dados</b>	5,56	9,86 *
<b>Dados centrados na média</b>	2,96	1,56 <sup>NS</sup>
<b>Dados autoescalados</b>	1,61	3,13 *
<b><math>T/FBP</math></b>	6,70	10,05 *
<b><math>T/FBP</math> centrados na média</b>	3,48	1,31 <sup>NS</sup>
<b><math>T/FBP</math> autoescalados</b>	1,72	2,90 *
<b><math>T-IBP</math></b>	9,53	10,06 *
<b><math>T-IBP</math> centrados na média</b>	4,89	0,92 <sup>NS</sup>
<b><math>T-IBP</math> autoescalados</b>	2,87	1,62 <sup>NS</sup>

$t_{tab\ 95\%} = 2,03$  (n= 94); NS – não significativo; \* - significativo

Além disso, quando o teste  $t$  foi aplicado, somente quatro valores de  $t_{calc}$  se mostraram abaixo do valor de  $t_{tab}$ , indicando que as duas metodologias são significativamente diferentes, para cinco dos casos avaliados, com 95 % de confiança. Isto mostra que os dois conjuntos de gasolinas são distintos e que não podem ser aplicados num mesmo modelo, ressaltando a importância de se utilizar gasolinas de diferentes composições para a construção de um modelo mais próximo do comportamento das gasolinas comercializadas.

Um conjunto de dados na validação de mesma natureza do conjunto observado na calibração também deve ser utilizado, uma vez que as normalizações e pré-processamentos empregados não conseguiram minimizar estas diferenças.

#### **4.2.3 Avaliação da Repetitividade e Reprodutibilidade**

Para a avaliação da repetitividade e reprodutibilidade do método, de acordo com a ISO 5725 Part 2 [62], foram analisadas dez amostras de gasolina provenientes do PMQC-ANP, com diferentes teores alcoólicos. Estas amostras foram analisadas pela NBR 13992 e pelo método proposto, utilizando-se o modelo de calibração nos quais os dados não foram normalizados ou pré-processados, com o conjunto  $GPM$ . Os resultados são apresentados na Tabela 13.

**Tabela 13** – Valores de repetitividade e reprodutibilidade obtidos para a determinação do teor alcoólico de gasolinas automotivas analisadas pela NBR 13992 e com a metodologia proposta com e sem arredondamento dos resultados obtidos.

<b>Método</b>	<b>Repetitividade %(v/v)</b>	<b>Reprodutibilidade %(v/v)</b>
<b>NBR 13992</b>	0,37	0,43
<b>Método Proposto (valor com uma casa decimal)</b>	0,41	0,44
<b>Método Proposto (valor sem casas decimais)</b>	0,50	0,51

A repetitividade e reprodutibilidade do método proposto foram avaliadas de duas formas: utilizando-se os resultados gerados pelo modelo proposto com uma casa decimal e também com os valores de teor alcoólico arredondados (sem casas decimais).

O teste F, aplicado a estes valores, apresentou valores entre 1,05 e 1,83 (para um  $F_{tab}$  de 3,18) mostrando que a repetitividade e reprodutibilidade dos dois métodos são equivalentes, a um nível de confiança de 95 % e nove graus de liberdade, mostrando a eficácia da metodologia proposta.

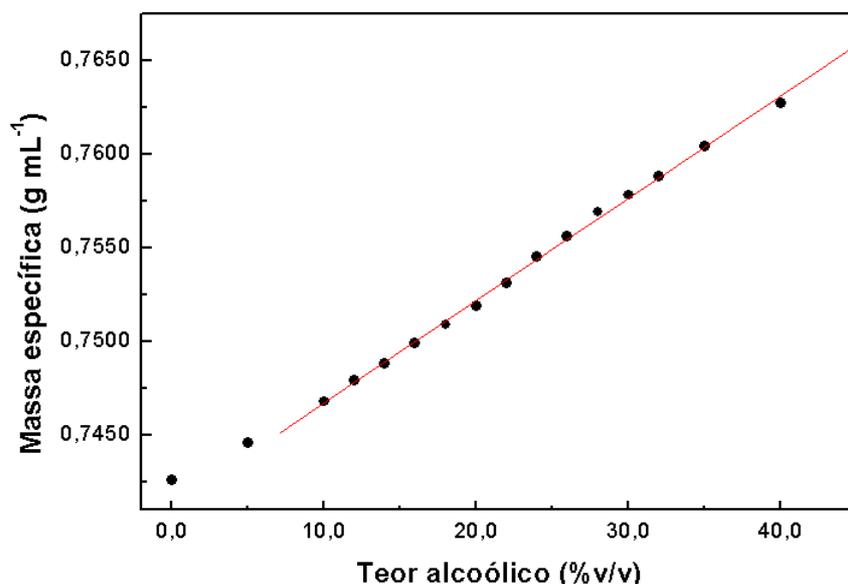
Embora os resultados obtidos com e sem arredondamentos sejam equivalentes aos do método da proveta, os valores com uma casa decimal apresentou valores de repetitividade e reprodutibilidade mais próximos aos do método NBR 13992.

Além de proporcionar resultados equivalentes ao método normalizado, o método proposto proporciona valores com maior número de casas decimais, ou seja, uma maior exatidão, uma vez que na NBR 13992 os valores obtidos devem ser apresentados sem casas decimais [60], devido à resolução da proveta.

Desta forma a metodologia proposta é capaz de prever os valores de teor alcoólico de gasolinas comercializadas em postos revendedores com precisão, exatidão, repetitividade e reprodutibilidade equivalentes ao do método NBR 13992, sem a necessidade de sua realização, diminuindo o número de ensaios a serem realizados e, conseqüentemente, o tempo e os custos de análise.

### 4.3 Determinação da Massa Específica

Para avaliar o comportamento da massa específica em relação ao aumento do teor alcoólico, foi determinada a massa específica, em um densímetro automático, para uma gasolina A adicionada de diferentes quantidades de etanol anidro, na faixa de 0 a 45 %(v/v). Os resultados são apresentados na Figura 30. Esta avaliação foi necessária devido ao comportamento observado anteriormente na determinação do teor alcoólico, onde as amostras se agruparam em dois conjuntos de teores alcoólicos distintos.



**Figura 30** – Variação da massa específica para diferentes misturas etanol-gasolina.

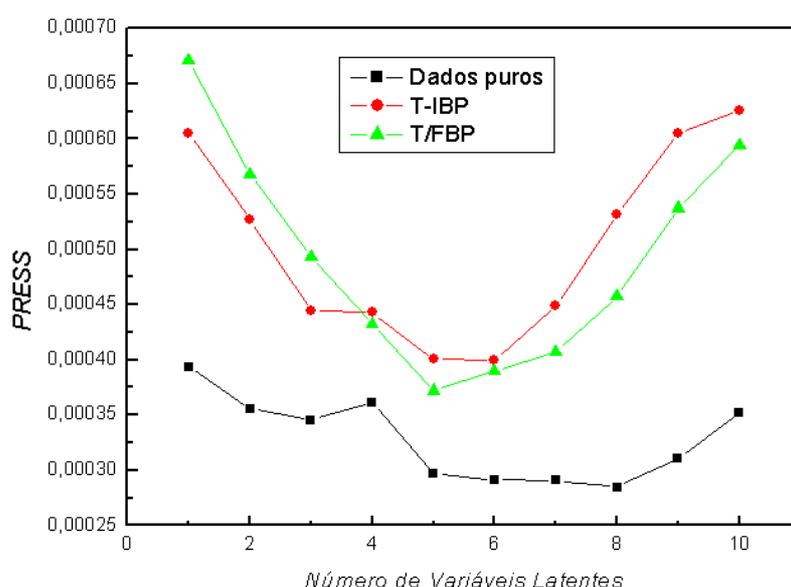
Observou-se uma faixa linear crescente entre 10 e 35 %(v/v) e fora desta faixa, um comportamento não linear. Esta faixa linear contém todo o intervalo de teor alcoólico encontrado para as gasolinas conformes comercializadas nos postos revendedores.

Um conjunto de 95 amostras foi utilizado para a construção do modelo de calibração enquanto 38 amostras foram utilizadas na validação com o objetivo de determinar a massa específica da gasolina utilizando as curvas de destilação e a calibração multivariada. Este conjunto de amostras, do FQCP-ANP, possui diferentes quantidades de etanol, 19 a 28 %(v/v), com origem em cinco diferentes refinarias (REGAP, REDUC, RLAM, REPLAN e REVAP) e massas específicas no intervalo entre 0,74 e 0,76 g mL<sup>-1</sup>.

As amostras foram normalizadas e pré-processadas da mesma maneira realizada na determinação do teor alcoólico, com a criação de nove modelos PLS com diferentes combinações entre normalizações e pré-processamentos, obtendo três modelos de calibração e nove modelos de validação.

O gráfico do número de variáveis latentes versus *PRESS* foi construído para se determinar o número de variáveis latentes necessário para a construção de cada modelo (Figura 31).

Os resultados mostram que oito variáveis latentes podem ser utilizadas na construção da curva de calibração nos dados não normalizados, enquanto cinco e seis variáveis latentes foram usadas nas construções das curvas de calibração dos dados com normalizações *T/IBP* e *T/IBP*, respectivamente.



**Figura 31** – Gráfico de número de variáveis latentes versus *PRESS*.

Após a escolha do número correto de variáveis latentes, foi construído o modelo de calibração para cada conjunto de dados, normalizados ou não, e os valores de porcentagem de variância explicada para as dez primeiras variáveis latentes são apresentados na Tabela 14.

Os resultados mostram que oito variáveis latentes foram responsáveis por 90,4 % da variância explicada nos dados não normalizados, seis por 91,3 % na normalização *T-IBP* e cinco por 94,1 % na normalização *T/IBP*. Observou-se que a primeira variável latente foi destacadamente a responsável pela maior porcentagem de variância explicada somente para os dados puros.

A construção do gráfico dos pesos para as normalizações feitas neste trabalho resultou em comportamento em forma de “V” com duas faixas de porcentagem de volume destilado, 4 a 51 e 71 a 93 %(v/v), como mostrado na Figura 32.

**Tabela 14** – Valores percentuais de variância explicada para as dez primeiras variáveis latentes.

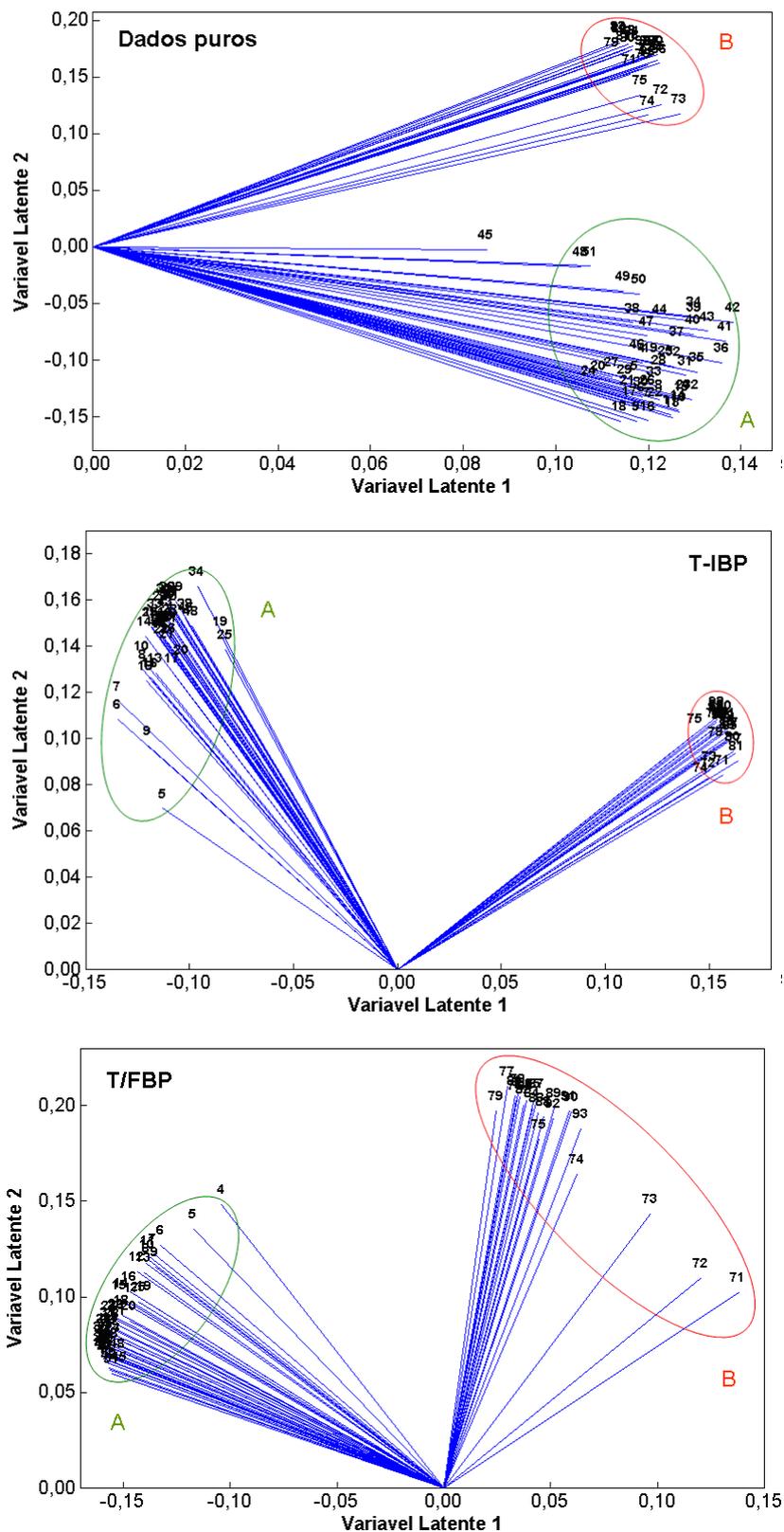
<b>VL</b>	<b>Dados</b>	<b>T-IBP</b>	<b>T/FBP</b>
<b>1</b>	57,8	44,9	54,5
<b>2</b>	21,6	39,9	32,9
<b>3</b>	4,8	3,4	5,3
<b>4</b>	2,3	1,3	0,8
<b>5</b>	1,9	1,3	0,6
<b>6</b>	0,8	0,6	1,1
<b>7</b>	0,6	0,5	0,6
<b>8</b>	0,7	0,4	0,4
<b>9</b>	0,5	0,6	0,3
<b>10</b>	0,5	0,9	0,3

Para os dados puros, as duas primeiras variáveis latentes explicam 79,4 % da variância explicada do modelo, sendo a primeira com o maior percentual de variância explicada, 57,8 %. No gráfico dos pesos dos dados puros pode-se observar que ambas as regiões apresentam pesos equivalentes nas duas primeiras variáveis latentes.

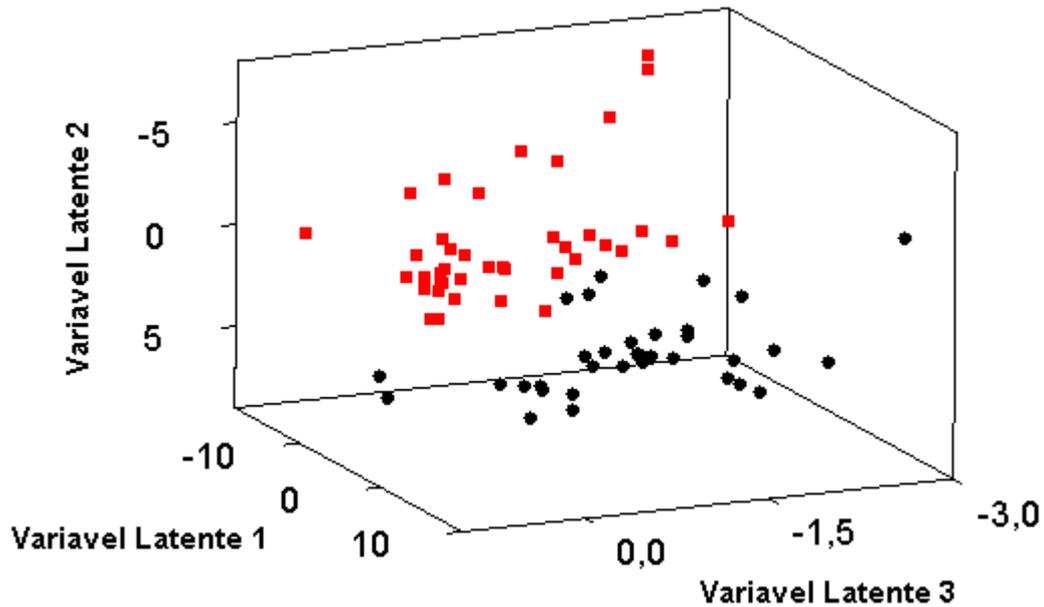
As duas primeiras variáveis latentes, para o modelo com os dados normalizados segundo *T-IBP*, apresentaram semelhantes percentuais de variância explicada (44,9 e 39,9 %). A subtração de cada ponto pelo ponto inicial da curva de destilação (*IBP*) modificou o peso de cada variável, fazendo com que as regiões A e B apresentassem pesos semelhantes na primeira variável latente, da mesma forma que os dados puros, mas a região A apresentou um maior peso na segunda variável latente que a região B.

Para a normalização *T/FBP*, as duas primeiras variáveis latentes explicam 87,4 % da variância explicada do modelo, sendo a primeira com o maior percentual de variância explicada (54,5 %). Pela Figura 32, pode-se notar que a região A apresenta elevado peso na primeira variável latente sendo esta região a de maior importância na construção do modelo, diferentemente que nos modelos anteriores.

Estes modelos, da mesma forma que nos da determinação do teor alcoólico, apresentaram separação por grupos, 19 a 22 e 23 a 28 %(v/v), como mostra a Figura 33. Esta separação se mostrou mais nítida quando se utiliza as três primeiras variáveis latentes, e não duas como no estudo anterior, uma vez que a terceira variável latente destes modelos apresentou um percentual maior que os dos modelos do teor alcoólico.



**Figura 32** – Gráfico dos pesos na determinação da massa específica para diferentes normalizações.



**Figura 33** – Gráfico das três primeiras variáveis latentes para a determinação da massa específica com a separação das amostras em dois grupos de teor alcoólico.

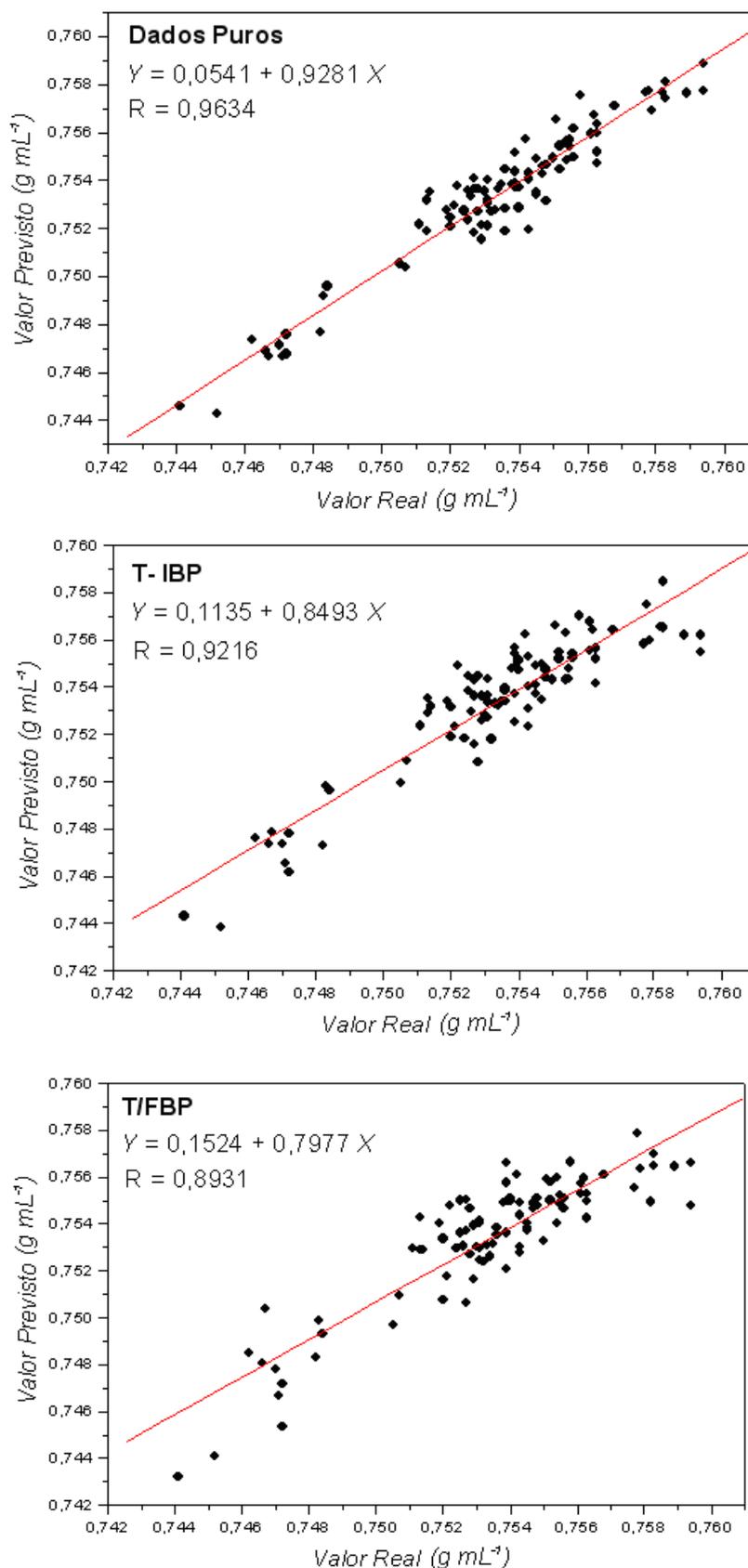
(●) 19 a 22 % (v/v); (■) 23 a 28 % (v/v).

Os valores previstos pelos modelos de calibração foram comparados aos valores reais, obtidos pela ASTM D4052, obtendo-se três retas com diferentes coeficientes de correlação ( $R$ ). Pela Figura 34 percebe-se que o modelo de calibração construído com os dados sem normalização apresentou o maior coeficiente de correlação. Com isto espera-se que o modelo de calibração dos dados não normalizados apresente valores de  $RMSEC$  menores que os modelos em que os dados foram normalizados.

A partir dos valores reais e previstos dos modelos de calibração e validação foram calculados os valores de  $RMSEC$  e  $RMSEP$  (Tabela 15). Os valores dos diferentes modelos de calibração mostraram que os menores valores foram obtidos para os dados que não foram normalizados ou pré-processados, da mesma forma que para a previsão do teor alcoólico.

O teste  $F$ , aplicado para os valores de  $RMSEC$  e  $RMSEP$ , mostrou que as normalizações  $T/FBP$  e  $T-IBP$  proporcionam valores de  $RMSEC$  não somente maiores, mas estatisticamente diferentes, a 95 % de confiança, indicando que o modelo de calibração deve ser construído a partir dos dados não normalizados. O teste  $F$  indicou que os valores de  $RMSEP$  em que os dados não foram pré-processados não diferem estatisticamente no nível de 95 % de confiança.

Estes resultados mostram que as normalizações, mais uma vez, não surtiram o efeito desejado, ou seja, a diminuição dos valores de  $RMSEC$  e  $RMSEP$  dos dados puros.



**Figura 34** – Valor Real versus Previsto para os modelos construídos com dados não normalizados e normalizados segundo *T-IBP* e *T/FBP*.

**Tabela 15** – Valores de *RMSEC* e *RMSEP* obtidos para diferentes modelos construídos na determinação da massa específica da gasolina.

Pré-tratamento e normalização	<i>RMSEC</i>	<i>RMSEP</i>
<b>Dados</b>		0,00111
<b>Dados centrados na média</b>	0,00086	0,00185
<b>Dados autoescalados</b>		0,00241
<b><i>T/FBP</i></b>		0,00134
<b><i>T/FBP</i> centrados na média</b>	0,00144	0,00199
<b><i>T/FBP</i> autoescalados</b>		0,00317
<b><i>T-IBP</i></b>		0,00132
<b><i>T-IBP</i> centrados na média</b>	0,00124	0,00343
<b><i>T-IBP</i> autoescalados</b>		0,00470

**Tabela 16** – Valores de  $F_{calc}$  para *RMSEC* e *RMSEP* na determinação da massa específica na gasolina a partir das curvas de destilação.

	<i>RMSEC</i>	<i>RMSEP</i>
<b>Dados x Dados</b>		1,00*
<b>Dados centrados na média x Dados</b>	1,00*	2,79 <sup>NS</sup>
<b>Dados autoescalados x Dados</b>		4,75 <sup>NS</sup>
<b><i>T/FBP</i> x Dados</b>		1,47*
<b><i>T/FBP</i> centrados na média x Dados</b>	2,82 <sup>NS</sup>	3,25 <sup>NS</sup>
<b><i>T/FBP</i> autoescalados x Dados</b>		8,22 <sup>NS</sup>
<b><i>T-IBP</i> x Dados</b>		1,43*
<b><i>T-IBP</i> centrados na média x Dados</b>	2,10 <sup>NS</sup>	9,66 <sup>NS</sup>
<b><i>T-IBP</i> autoescalados x Dados</b>		18,11 <sup>NS</sup>
<b><i>Graus de Liberdade</i></b>	94	37
<b><i>Valores de <math>F_{tab}</math></i></b>	1,41	1,73

NS – não significativo; \* - significativo

### 4.3.1 Validação do Método Proposto

O teste  $t$  foi realizado para verificar a exatidão do método proposto em comparação ao método do densímetro digital (ASTM D4052) [66]. Com este teste pôde-se notar (Tabela 17) que apenas cinco valores de  $t_{calc}$ , de um total de nove, foram menores que o valor de  $t_{tab}$ . De maneira geral, os modelos cujos dados não foram pré-processados, independentemente da normalização, apresentaram resultados equivalentes aos do método normatizado. Apesar de apresentar maiores valores de  $RMSEC$  e  $RMSEP$ , todos os modelos normalizados  $T-IBP$  proporcionaram resultados idênticos à ASTM D4052, a um nível de confiança de 95 %.

A análise dos valores de  $R$  dos modelos de calibração, juntamente com os resultados dos testes  $t$  e  $F$ , apontaram que o melhor método para se determinar a massa específica é aquele onde os dados das curvas de destilação não são normalizados ou pré-processados, da mesma forma que na determinação do teor alcoólico.

Comparado aos valores descritos na literatura, o valor de  $RMSEP$  para este modelo (0,0011) foi menor que os obtidos por Oliveira et al [47] (0,0026 e 0,0050), usando PLS e FT-NIR, mostrando a grande eficiência da metodologia proposta.

**Tabela 17** – Valores de  $t_{calc}$  para na comparação dos resultados obtidos pelo ensaio segundo a norma ASTM D4052 com a metodologia proposta neste trabalho.

<b><i>Pré-processamento e Normalização</i></b>	<b><math>t_{calc}</math></b>
<b>Dados</b>	1,75
<b>Dados centrados na media</b>	8,65
<b>Dados autoescalados</b>	4,89
<b><i>T/FBP</i></b>	1,79
<b><i>T/FBP</i> centrados na média</b>	7,07
<b><i>T/FBP</i> autoescalados</b>	3,30
<b><i>T-IBP</i></b>	1,13
<b><i>T-IBP</i> centrados na média</b>	0,51
<b><i>T-IBP</i> autoescalados</b>	0,37
<hr/>	
$t_{tab\ 95\%} = 2,03$ (n = 38)	

### 4.3.2 Avaliação da Repetitividade e Reprodutibilidade do Método Proposto

Para a avaliação da repetitividade e reprodutibilidade do método, de acordo com a ISO 5725 Part 2 [62], foram analisadas dez amostras de gasolina provenientes do PMQC-ANP, com diferentes massas específicas. Estas amostras foram analisadas pelas ASTM D4052 (densímetro digital) e D1298 (densímetro de vidro) e pelo método proposto, utilizando-se o modelo de calibração nos quais os dados não foram normalizados ou pré-processados. Os resultados são apresentados na Tabela 18.

Os menores valores de repetitividade e reprodutibilidade foram obtidos para a ASTM D4052, seguido da ASTM D1298 e o método proposto.

O teste  $F$  foi inicialmente aplicado a estes valores em relação ao método ASTM D4052, que apresentou valores entre 4,00 e 29,34 (para um  $F_{tab}$  de 3,18), mostrando que a repetitividade e reprodutibilidade dos métodos ASTM D1298 e o proposto são diferentes da ASTM D4052, a um nível de confiança de 95 % e nove graus de liberdade.

Quando o método proposto é comparado ao da ASTM D1298, obtêm-se os valores de 3,65 e 3,61 para a repetitividade e reprodutibilidade, respectivamente, mostrando que também são diferentes, a um nível de confiança de 95 % e nove graus de liberdade.

Desta forma, o teste  $F$  mostrou que os valores de repetitividade e reprodutibilidade para o método proposto são, além de maiores, diferentes aos obtidos para os métodos normalizados, para um nível de 95 % de confiança, com nove graus de liberdade. Este fato pode estar relacionado aos baixos coeficientes de linearidade das retas do conjunto de calibração, da ordem de 0,96, podendo ainda estar sendo influenciados pelas diferentes composições das amostras de gasolina de diferentes origens. As normalizações testadas não foram capazes de diminuir estes efeitos.

**Tabela 18** – Valores de repetitividade e reprodutibilidade, obtidos para a determinação da massa específica de gasolinas automotivas analisadas pela ASTM D4052, ASTM D1298 e com a metodologia proposta.

<b>Método</b>	<b>Repetitividade (g mL<sup>-1</sup>)</b>	<b>Reprodutibilidade (g mL<sup>-1</sup>)</b>
<b>ASTM D4052</b>	0,00012	0,00020
<b>ASTM D1298</b>	0,00034	0,00040
<b>Método Proposto</b>	0,00065	0,00076

## 5. Conclusões

Os métodos multivariados PCA e LDA aplicados à curva de destilação possibilitaram a separação das amostras de gasolina produzidas em diferentes refinarias, comercializadas dentro do Estado de Minas Gerais.

Aplicando-se a análise de componentes principais pode-se reduzir o número de variáveis do conjunto amostral, mantendo-se uma elevada variância total, apresentando boa eficiência na separação dos grupos de acordo com a refinaria de origem, com apenas quatro componentes principais acumulando aproximadamente 90 % da variabilidade total do sistema. Pela primeira componente observou-se a separação da REGAP e da RLAM enquanto que a PC2 foi responsável por separar a RLAM das refinarias REDUC, REVAP e REPLAN.

Empregando-se a análise discriminante pode-se prever a origem, de forma não tendenciosa, do conjunto de amostras de gasolina com origem desconhecida, com 95,1 % de probabilidade de acerto. As distâncias entre os grupos formados foram geradas de acordo com sua semelhança, sendo que o par REPLAN/REVAP apresentou a menor distância e REDUC/REGAP a maior distância. Estes resultados mostraram que as amostras provenientes da REPLAN possuem composição química próxima as amostras da REVAP enquanto que as amostras da REDUC apresentam composição química diferente da REGAP.

Nas determinações do teor alcoólico e da massa específica observou-se a importância de um estudo preliminar da curva de destilação para a seleção da melhor faixa de trabalho, uma vez que o uso da parte inicial e da parte relativa ao platô formado, referente ao etanol adicionado, produziu altos valores de *RMSEC* e *RMSEP*. Além disso, notou-se a importância de se utilizar gasolinas de diferentes composições para a construção de um modelo mais robusto, que leve em consideração o perfil das gasolinas comercializadas na região de coleta.

O teste F demonstrou que as normalizações e pré-processamentos utilizados não melhoraram significativamente os valores de *RMSEC* e *RMSEP* obtidos nas previsões do teor alcoólico e da massa específica, mostrando que estes não se mostraram adequados.

As curvas de destilação obtidas pela ASTM D86, em conjunto com o método de calibração multivariada PLS, possibilitaram a previsão do teor alcoólico, com o uso de cinco variáveis latentes, na faixa entre 19 e 28 %(v/v), com valores de *RMSEC* e *RMSEP* de 0,792 e 0,838 %(v/v) e valores de repetibilidade e reprodutibilidade, com uma casa decimal, de 0,41 e 0,44 %(v/v), respectivamente. Estes valores de repetibilidade e reprodutibilidade foram estatisticamente equivalentes aos resultados obtidos com a NBR 13992.

A associação de curvas de destilação e calibração multivariada possibilitou, utilizando oito variáveis latentes, a determinação da massa específica da gasolina automotiva, na faixa entre 0,74 e 0,76 mg L<sup>-1</sup>, independentemente da refinaria de origem, com valores de *RMSEC* e *RMSEP* de 0,0009 e 0,0011 mg L<sup>-1</sup>, mostrando-se menores que os descritos na literatura utilizando PLS e FT-NIR [47]. Os valores de repetibilidade e reprodutibilidade obtidos foram de 0,00065 e 0,00076 g mL<sup>-1</sup>, respectivamente mostrando que, além de maiores, são estatisticamente diferentes dos métodos ASTM D4052 e D1298.

A utilização de curvas de destilação associadas às técnicas quimiométricas gerou resultados de boa precisão e exatidão, na análise da origem, teor alcoólico e massa específica da gasolina automotiva. Como a destilação ASTM D86 faz parte dos métodos de rotina na análise da gasolina automotiva, as curvas de destilação ainda podem ser utilizadas na previsão de diferentes parâmetros simultaneamente, diminuindo o número de ensaios a serem realizados, otimizando o processo analítico e diminuindo os custos.

## 6. Referências Bibliográficas

- [1] Speight, J. G. *The Chemistry and Technology of Petroleum*. 4<sup>th</sup> edition. CRC Press: Taylor & Francis Group, 2006. 954p.
- [2] Moreira, L. S.; d'Avila, L. A.; Azevedo, D. A. *Automotive gasoline quality analysis by gas chromatography: Study of adulteration*. *Chromatographia*, 58 (2003) 501-505.
- [3] [www.petrobras.com.br](http://www.petrobras.com.br) , acessado em 09/04/08.
- [4] Gasolina Automotiva. Informativo Técnico da PETROBRAS, REGAP (1996).
- [5] <http://www.biodieselbr.com/proalcool/pro-alcool.htm>, acessado em 26/05/2008.
- [6] <http://ecen.com/eee19/composi.htm>, acessado em 08/05/2008.
- [7] Portaria Agência Nacional do Petróleo nº. 309 de 27/12/2001.
- [8] [www.anp.gov.br](http://www.anp.gov.br), acessado em 15/04/2008.
- [9] Silva,R.; Cataluña, R.; Menezes, E.W.; Samios, D.; Piatnicki, C.M.S. *Effect of additives on the antiknock properties and Reid vapor pressure of gasoline*. *Fuel* 84 (2005) 951-959.
- [10] Côcco, L.C.; Yamamoto, C.I.; von Meien, O.F. *Study of correlations for physicochemical properties of Brazilian gasoline*. *Chemometrics and Intelligent Laboratory Systems* 76 (2005) 55-63.
- [11] ASTM-D2699. Standard test method for Research Octane Number of spark-ignition engine fuel. Washington, DC: American Society for Testing and Materials (1999).
- [12] Pereira,R.C.C.; Pasa, V.M.D. *Effect of mono-olefins and diolefins on the stability of automotive gasoline* . *Fuel* 85 (2006) 1860–1865.
- [13] Barbeira, P.J.S.; Pereira, R.C.C.; Corgozinho, C.N.C. *Identification of gasoline origin by physical and chemical properties and multivariate analysis*. *Energy & Fuels* (2007), 21(4), 2212-2215.
- [14] Portaria Agência Nacional do Petróleo nº. 2 de 16/01/2002, revogada pela Resolução nº 36 de 6/12/2005.
- [15] Portaria Agência Nacional do Petróleo nº. 310 de 27/12/2001, revogada pela Resolução nº 15 de 17/07/2006.
- [16] Wiedemann,L.S.M., d'Avila,L.A., Azevedo,D.A. *Adulteration detection of Brazilian gasoline samples by statistical analysis*. *Fuel* 84 (2005) 467–473.
- [17] Skrobot,V.L., Pereira,R.C.C., Castro,E.V.R., Fortes,I.C.P., Pasa,V.M.D. *Identification of Adulteration of Gasoline Applying Multivariate Data Analysis Techniques HCA and KNN in Chromatographic Data*. *Energy & Fuels* 19 (2005) 2350-2356.
- [18] [www.fazenda.rj.gov.br](http://www.fazenda.rj.gov.br) acessado em 04/06/2007.
- [19] [www.fazenda.mg.gov.br](http://www.fazenda.mg.gov.br) acessado em 04/06/2007.

- [20] Menezes, E.W.de; Cataluña, R.; Samios, D.; da Silva, R. *Addition of an azeotropic ETBE/ethanol mixture in eurosuper-type gasolines*. Fuel 85 (2006) 2567–2577.
- [21] Portaria Agência Nacional do Petróleo 274 de 11/01/2001.
- [22] Barbeira, P.J.S. *Using Statistical tools to detect gasoline adulteration*. Engenharia Térmica, 2 (2002) 48-50.
- [23] Bruno, T.J. *Improvements in the measurement of distillation curves. 1. A composition-explicit approach*. Ind.Eng.Chem.Res. 45 (2006) 4371-4380.
- [24] ASTM-D86. Standard test for distillation of petroleum products at atmospheric pressure. Washington, DC: American Society for Testing and Materials (1987).
- [25] ABNT NBR 9619. Produtos de Petróleo – Determinação da faixa de destilação. Associação Brasileira de Normas Técnicas.
- [26] French, R.; Malone, P. *Phase Equilibra of ethanol fuel blens*. Fluid Phase Equilibra, 228 (2005) 27-40.
- [27] E.W.de Menezes, R. Cataluña, D. Samios, R. da Silva *Addition of an azeotropic ETBE/ethanol mixture in eurosuper-type gasolines*. Fuel 85 (2006) 2567–2577.
- [28] Cataluña, R.; Silva, R. *Desenvolvimento de um equipamento para avaliação do efeito do etanol na pressão de vapor e entalpia de vaporização em gasolinas automotivas*. Química Nova 29 (2006) 580-585.
- [29] Pumphrey, J.A.; Brand, J.I.; Scheller, W.A. *Vapour pressure measurements and predictions for alcohol–gasoline blends*. Fuel 79 (2000) 1405–1411.
- [30] Oliveira, F.S., Teixeira, L.S.G., Araújo, M.C.U., Korn M. *Screening analysis to detect adulterations in Brazilian gasoline samples using distillation curves*. Fuel 83 (2006) 917-923.
- [31] Flumignan, D.L.; Anaia, G.C.; Ferreira, F.O.; Tininis, A.G.; de Oliveira, J.E. *Screening brazilian automotive gasoline quality through quantification of saturated hydrocarbons and anhydrous ethanol by gas chromatography and exploratory data analysis*. Chromatographia 65 (2007) 617-623.
- [32] Kramer, R. *Chemometric Techniques for quantitative analysis*. Marcel Dekker, 1998. 203 p.
- [33] Geladi, P. Kowalsky, B.R. *Partial least-squares regression: a tutorial*. Analytical Chimica Acta, 185 (1986), 1-17.
- [34] Montgomery, D.C. *Introdução ao controle estatístico da qualidade*. Rio de Janeiro: Editora LTC, 2001.
- [35] Otto, M. *Chemometrics: Statistics and computer application in analytical chemistry*. Weinhein, New York, Chichester, Brisbane, Singapore, Toronto: Wiley-VCH, 1999.
- [36] Fernandes, A.P.; Santos, M.C.; Lemos, S.G.; Ferreira, M.M.C.; Nóbrega, J. A. *Pattern recognition applied to mineral characterization of Brazilian coffees and sugar-cane spirits*. Spectrochimica Acta Part B 60 (2005) 717-724.

- [37] Pereira, Rita C. C.; Skrobot, Vinicius L.; Castro, Eustaquio V. R.; Fortes, Isabel C. P.; Pasa, Vanya M. D. *Determination of Gasoline Adulteration by Principal Components Analysis-Linear Discriminant Analysis Applied to FTIR Spectra*. *Energy & Fuels* 20 (2006) 1097-1102.
- [38] Mingoti, S.A. *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Belo Horizonte: Editora UFMG, 2005.
- [39] Hair Jr., J.F.; Anderson, R.E.; Tatham, R.L.; Black, W.C. *Análise Multivariada de dados*. Porto Alegre: Editora Bookman, 2005.
- [40] Balabin, R.M., Safieva, R.Z. *Gasoline classification by source and type based on near infrared (NIR) spectroscopy data*. *Fuel* 87 (2008) 1096-1101.
- [41] Ferreira, M.M.C., Antunes, A.M., Melgo, M.S. et al. *Quimiometria I: calibração multivariada, um tutorial*. *Quím. Nova*, 22 (1999) 724-731.
- [42] Hemmateenejad, B.; Abbaspour, A.; Maghamia, H.; Miri, R.; Panjehshahin, M.R. *Partial least squares-based multivariate spectral calibration method for simultaneous determination of beta-carboline derivatives in Peganum harmala seed extracts*. *Analytica Chimica Acta* 575 (2006) 290–299.
- [43] Wakeling, I. N.; Morris, J.J. *A test of significance for partial least squares regression*. *Journal of Chemometrics* 7 (1993) 291-304.
- [44] Wold, S., Sjöström, M., Eriksson, L. *PLS-regression: a basic tool of chemometrics*. *Chemometrics and Intelligent Laboratory Systems* 58 (2001) 109-130.
- [45] Minitab-14, *Statistical Software for Windows*; Minitab, Inc: State College, PA, 2003.
- [46] Fodor, G. E.; Mason, R. A.; Hutzler, S. A. *Estimation of middle distillate fuel properties by FTIR*. *Applied Spectroscopy* 53 (1999) 1292-1298.
- [47] Oliveira, Flavia C. C.; de Souza, Antonio T. P. C.; Dias, Jose A.; Dias, Silvia C. L.; Rubim, Joel C. *A escolha da faixa espectral no uso combinado de métodos espectroscópicos e quimiométricos*. *Química* 27 (2004) 218-225.
- [48] Vianney, O. S. Jr.; Oliveira, F.C.C.; Lima, D.G.; Petry, A.C.; Garcia, E.; Suarez, P.A.Z.; Rubim, J.C. *A comparative study of diesel analysis by FTIR, FTNIR and FT-Raman spectroscopy using PLS and artificial neural network analysis*. *Analytica Chimica Acta* 547 (2005) 188–196.
- [49] Seinsche, von K.; Luigart, F.; Bartl, P. *Quality control of jet fuels. PLS [Partial least squares] regression analysis of gas chromatographic and IR spectroscopic data*. *Fuel and Energy Abstracts* 38 (1997) 11-11(1).
- [50] Gomez-Carracedo, M.P.; Andrade, J.M.; Calvino, M.; Fernández, E.; Prada, D.; Muniategui, S. *Multivariate prediction of eight kerosene properties employing vapour-phase mid-infrared spectrometry*. *Fuel* 82 (2003) 1211–1218.
- [51] Flumignan, D.L.; Ferreira, F.O.; Tininis, A.G.; de Oliveira, J. E. *Abstracts of Papers, 234th ACS National Meeting, Boston, MA, United States, August 19-23, 2007* (2007), FUEL-011.

- [52] Andrade, J.M.; Muniategui, S.; Prada, D. *Prediction of clean octane numbers of catalytic reformed naphthas using FT-m.i.r. and PLS*. Fuel 76 (1997) 1035-1042.
- [53] Kelly, J. J.; Barlow, C. H.; Jinguji, T. M.; Callis, J. B. *Prediction of gasoline octane numbers from near-infrared spectral features in the range 660-1215 nm*. Anal. Chem. 1989, 61, 313-20.
- [54] Breikreitz, M.C.; Raimundo, I.M., Jr.; Rohwedder, J.J.R.; Pasquini, C.; Dantas Filho, H.A.; Jose, G.E.; Araujo, M.C.U. *Determination of total sulfur in diesel fuel employing NIR spectroscopy and multivariate calibration*. Analyst 128 (2003) 1204-1207.
- [55] Johnson, K.J.; Morris, R.E.; Rose-Pehrsson, S.L. *Evaluating the Predictive Powers of Spectroscopy and Chromatography for Fuel Quality Assessment*. Energy & Fuels 20 (2006) 727-733.
- [56] J.B. Cooper, K.L. Wise, W.T. Welch, R.R. Bledsoe, M.B. Sumner. *Determination of weight percent oxygen in commercial gasoline: a comparison between FT-Raman, FT-IR, and dispersive near-IR spectroscopies*. Appl. Spectrosc. 50 (1996) 917-921.
- [57] J.B. Cooper, K.L. Wise, W.T. Welch, M.B. Sumner, B.K. Wilt, R.R. Bledsoe, *Comparison of Near-IR, Raman, and Mid-IR Spectroscopies for the Determination of BTEX in Petroleum Fuels*. Appl. Spectrosc. 51 (1997) 1613.
- [58] Cooper, J.B.; Bledsoe, R.R.Jr.; Wise, K.L.; Sumner, M.B.; Welch, W.T.; Wilt, B.K. *Process and apparatus for octane numbers and Reid vapor pressure by Raman spectroscopy*. United States Patent 5892228 (1999), 19 pp.
- [59] ASTM-D4052. *Standard test for density and relative density of liquids by digital density meter*. Washington, DC:American Society for Testing and Materials (2002).
- [60] ABNT NBR-13992. *Gasolina Automotiva – Determinação do teor de álcool etílico anidro combustível (AEAC)* (1997).
- [61] ISO/IEC 17025:1999(E). *General requirements for the competence of testing and calibration laboratories*. Genebra, first edition 1999-12-15.
- [62] **ISO/IEC 5725-2** – *Accuracy (trueness and precision) of Measurement Methods and Results – Part 2: Basic Method for the Determination of Repeatability and Reproducibility of a Standard Measurement Method* (2002).
- [63] ASTM-D1298. *Standard Test Method for Density, Relative Density (Specific Gravity), or API Gravity of Crude Petroleum and Liquid Petroleum Products by Hydrometer Method*. American Society for Testing and Materials (2005).
- [64] Relatório Mensal, Programa de Monitoramento da Qualidade dos Combustíveis - ANP, maio, 2008.
- [65] Miller, J.C.; Miller, J.N. *Statistics for Analytical Chemistry*. 3<sup>rd</sup> edition. New York : Ellis Horwood PTR Prentice Hall, 1993.
- [66] Harris, D.C., *Análise Química Quantitativa*, 6<sup>a</sup> ed., LTC, 2005, 876p.

- [67] Magalhães, M. N. ; de Lima, A. C. P. *Noções de probabilidade e estatística*. 5ª edição. São Paulo : EDUSP, 2002.
- [68] Line,D,R, ; CRC - Handbook of Chemistry and Physics, CRC Press Inc.,, 71a edição,1990-1991.
- [69] Balabin, R.M.; Syunyaev, R.Z.; Karpov, S.A.; *Molar enthalpy of vaporization of ethanol-gasoline mixtures and their colloid state*. Fuel 86 (2007) 323–327.