

5º CONGRESSO BRASILEIRO DE PESQUISA E DESENVOLVIMENTO EM PETRÓLEO E GÁS



TÍTULO DO TRABALHO:

Calibração Multivariada Aplicada a Espectros Infravermelhos – Previsão de Propriedades Físico-químicas e Composição Química de Gasolina tipo C.

AUTORES:

Lílian Cristina Côcco, Carlos Itsuo Yamamoto

INSTITUIÇÃO:

Universidade Federal do Paraná – UFPR

Laboratório de Análises de Combustíveis Automotivos – LACAUT_{ets}

Este Trabalho foi preparado para apresentação no 5º Congresso Brasileiro de Pesquisa e Desenvolvimento em Petróleo e Gás- 5º PDPETRO, realizado pela Associação Brasileira de P&D em Petróleo e Gás-ABPG, no período de 15 a 22 de outubro de 2009, em Fortaleza-CE. Esse Trabalho foi selecionado pelo Comitê Científico do evento para apresentação, seguindo as informações contidas no documento submetido pelo(s) autor(es). O conteúdo do Trabalho, como apresentado, não foi revisado pela ABPG. Os organizadores não irão traduzir ou corrigir os textos recebidos. O material conforme, apresentado, não necessariamente reflete as opiniões da Associação Brasileira de P&D em Petróleo e Gás. O(s) autor(es) tem conhecimento e aprovação de que este Trabalho seja publicado nos Anais do 5º PDPETRO.

CALIBRAÇÃO MULTIVARIADA APLICADA A ESPECTROS INFRAVERMELHOS - PREVISÃO DE PROPRIEDADES FÍSICO-QUÍMICAS E COMPOSIÇÃO QUÍMICA DE GASOLINA TIPO C

Abstract

This work describes the attainment of mathematical models, applying multivariate calibration in infrared spectrum with ATR, from 128 gasoline samples with diverse chemical compositions, collected in a period of two and a half years. Infrared spectra had been used to assemble the input matrix for the modeling, whereas the standardized assays and gaseous chromatography had supplied the output matrices. Ninety samples were been used for training and 38 for testing. In order to calibrate chemical composition from chromatography, the techniques of mass spectrometry and chemical ionization were used to identify unknown substances and improve the fitting of the mathematical models. Two hundred and ninety substances were detected and identified, from which 100 were unknown. Six PLS/PCR models were attained to predict some properties as specific mass, Reid vapor pressure, T_{10} , T_{50} , T_{90} and PFE from distillation curve. Another six PLS/PCR models were attained to predict the amount of aromatics, paraffins, isoparaffins, naphtenes, olefins and oxygenates. In general, mathematical models were attained with good training fit, with correlation coefficients higher than 0,975 (T_{10}) and reaching a maximum of 0,998 (naphtenes) and they are able to forecast an average chemical percentage and properties of interest from gasoline, with acceptable prediction errors.

Introdução

Propriedades físico-químicas de combustíveis, em particular da gasolina, são importantes referências de sua qualidade, devendo ser monitoradas a fim de mantê-la, respeitando-se leis ambientais e parâmetros técnicos. Cada propriedade é, em geral, uma complicada função da composição química da gasolina, que depende de características físico-químicas da matéria-prima, bem como das misturas de correntes de diferentes processos de uma refinaria. Por isso, é usual a utilização de correlações matemáticas que forneçam informações relevantes aos profissionais da área de combustíveis automotivos. Estas correlações matemáticas são utilizadas por equipamentos como o espectrômetro de infravermelho portátil para a previsão de propriedades físico-químicas de combustíveis a partir da análise do espectro obtido, reduzindo o custo de implantação de alguns ensaios em laboratórios para controle da qualidade de combustíveis e redução do tempo de obtenção de propriedades físico-químicas desejadas.

Diversas propriedades são previstas através de correlações, tais como densidade, pressão de vapor *Reid* e curva de destilação. A própria composição química também é prevista por correlações, porém as previsões obtidas com as correlações existentes fornecem resultados que dependem muito de calibração e mesmo assim podem apresentar erros de previsão não desejáveis. Algumas correlações internas não estão preparadas para a gasolina nacional, que recebe uma quantidade apreciável de álcool etílico anidro, alterando significativamente a composição química e conseqüentemente suas propriedades. Para corrigir estes desvios de previsão, a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, ANP, implantou um programa de calibração dos analisadores infravermelhos portáteis que realizam o monitoramento, na qual modifica-se o banco de dados destes equipamentos adicionando-se amostras de gasolina com composição química e propriedades conhecidas, melhorando sua resposta. Ainda assim, pode-se melhorar cada vez mais os desvios de previsão, ao aumentar a

abrangência populacional, avaliar outros métodos matemáticos, espectrométricos e configurar novos equipamentos portáteis.

Para a montagem das correlações, utilizam-se dados gerados por qualquer técnica analítica de identificação, tais como espectrometria por infravermelho, ultravioleta ou de massas, espectrometria de ressonância magnética nuclear de hidrogênio (RMN ^1H) ou de carbono 13 (RMN ^{13}C) e cromatografia gasosa ou líquida, como dados de entrada. Os ensaios padrão, como densidade a 20°C e curva de destilação da gasolina, fornecem os dados de saída para posterior aplicação em um método de regressão de dados. O método matemático deve ser capaz de encontrar a relação existente entre as entradas e saídas, obtendo-se desta forma as correlações desejadas. Em resumo, para o desenvolvimento deste tipo de trabalho existe a necessidade da escolha de um método analítico confiável e um tipo de regressão de dados. Pode-se aplicar espectrometria no infravermelho próximo (NIR), pois é uma técnica particularmente popular em aplicações industriais devido a sua rapidez de análise, cujo ensaio não destrutivo necessita de um mínimo de preparação de amostras (Macho e Larrechi, 2002), porém escolheu-se espectrometria no infravermelho médio, devido à sua disponibilidade em laboratório, agregado à interpretação dos grupos funcionais presentes na gasolina.

Na análise de dados multivariados, como os obtidos por espectrometria no infravermelho, existe uma grande dificuldade de identificação do espectro da gasolina. Trabalhos apresentados por Galignani *et al.* (1993a, 1994) identificaram apenas substâncias simples dentro do espectro da gasolina, como exemplo o benzeno (1993a), metil terc-butil éter (MTBE) (1993b) e tolueno (1994). Em razão desta dificuldade de interpretação, um pré-tratamento de dados é indicado, pois poderá melhorar em muito os resultados, desde que se remova ou minimize interferências multiplicativas devido a dispersão, efeitos de partículas, ruído instrumental, mudanças na linha de base, entre outras (Wu *et al.*, 1999). Deve-se observar que uma impressão digital de sucesso envolve amostragem apropriada, técnicas analíticas e estratégias de interpretação de dados (Wang e Fingas, 2003). Desta forma, pode-se encontrar correlações mais aprimoradas para a gasolina nacional e aplicá-las para obtenção de previsões acuradas, com valores próximos às análises experimentais padrão ASTM e NBR, adotadas pela ANP como referências de qualidade para a gasolina brasileira.

A proposta deste trabalho é a aplicação de técnica analítica relativamente rápida de identificação da composição, como exemplo a espectrometria no infravermelho, para redução significativa do tempo de análise para cerca de 5 minutos, viabilizando a aplicação prática de correlações matemáticas, obtidas por PLS ou PCR, para a previsão das propriedades físico-químicas e composição química da gasolina. Ao otimizar o tempo de análise e encontrar nova correlação matemática, o método proposto pode ser aplicado para um monitoramento rápido e eficaz da gasolina nacional.

Metodologia

Foram utilizadas amostras de gasolina comum e aditivada, coletadas em recipiente próprio, de vidro âmbar e armazenadas abaixo de 10°C ou local fresco. A amostragem foi realizada fundamentando-se em resultados prévios obtidos pelos infravermelhos portáteis existentes em laboratório, capazes de fornecer uma composição química aproximada. Também foram escolhidas pelos resultados dos ensaios ASTM, tentando encontrar amostras diferenciadas da média, com a finalidade de abranger um espaço de amostragem amplo. Para obtenção de amostras diferenciadas foi necessário coletar amostras que estavam fora das especificações da ANP. Com isto obteve-se um conjunto de dados de 128 amostras o mais heterogêneo possível, permitindo a construção de modelos robustos. Do total, 22 amostras são gasolinas aditivadas e o restante, gasolina C (comum).

O equipamento utilizado para obter os espectros no infravermelho foi um Excalibur FT, com acessório ATR. A faixa de aquisição de dados foi de 4000 a 400 cm^{-1} (infravermelho médio), com resolução de 4 cm^{-1} , com 32 e 16 *scans* de varredura. A célula escolhida para leitura de amostra foi a de seleneto de zinco. Dados de transmitância foram exportados para utilização em outros programas. O cromatógrafo utilizado para a análise detalhada da composição química da gasolina foi um CG-FID Varian CP-3800. O programa StarDHATM da Varian Inc., versão 5, foi utilizado para quantificar e identificar automaticamente a maioria dos picos cromatográficos presentes na gasolina. O cromatógrafo utilizado para a identificação detalhada da composição química da gasolina foi um CG Varian CP-3800 acoplado a um detector de massas do tipo *ion trap* modelo Saturn 2000. A coluna utilizada foi uma capilar de sílica fundida CP-SIL PONA CB da Chrompack, 0,25 mm de diâmetro interno, 100 m de comprimento e 0,5 mm de filme líquido. A biblioteca NISTTM (versão 2005) foi utilizada como ferramenta adicional na identificação das substâncias presentes na gasolina. Para obtenção dos dados físico-químicos das amostras foram utilizados alguns equipamentos disponibilizados no LACAUT_{ets} (Laboratório de Análises de Combustíveis Automotivos da UFPR). Outros equipamentos utilizados neste estudo foram o densímetro automático modelo DMA 4500 da Antoon Paar que determina a densidade da gasolina a 20 °C em Kg/m^3 , o aparelho de pressão de vapor, modelo Minivap VPS da Grabner Instruments foi utilizado para determinação da PV REID e destiladores automáticos Herzog modelos HDA 628 e 627 forneceram os resultados de curva de destilação.

O PLS *toolbox* do MATLABTM 7.4.0 (R2007a) foi utilizado para gerar os modelos para previsão das variáveis em estudo. O programa permite aplicar diversos tipos de pré-tratamentos matemáticos de dados, além da escolha de algoritmos para regressão. Os pré-tratamentos de dados aplicados aos espectros de infravermelho médio das amostras de gasolina foram: o autoescalamento, aplicação de segunda derivada, alisamento e ortogonalização de sinal. Para as matrizes de dados de saída (composição química e propriedades físico-químicas) aplicou-se o autoescalamento.

Resultados e Discussão

Identificação por Cromatografia Gasosa Acoplada a Espectrometria de Massas

Para obtenção da composição química da gasolina, utilizou-se a técnica de cromatografia gasosa com detector FID (*Flame Ionization Detector*) e um programa específico para quantificação dos compostos conhecidos denominado DHA (*Detailed Hydrocarbon Analysis*). Para identificação de compostos desconhecidos na gasolina, aplicou-se a técnica de cromatografia gasosa com detector de massas (CG-EM). Deste modo os dados de composição química foram calibrados, melhorando a qualidade das matrizes de saída para modelagem matemática. O programa DHA consegue identificar a grande maioria dos compostos mais significativos, porém não estão catalogados em sua biblioteca os compostos nitrogenados e sulfurados, bem como alguns picos olefinicos e aromáticos. A quantidade de desconhecidos em uma amostra com características normais atinge aproximadamente 8% (v/v). Atualmente não é difícil detectar entre 500 a 600 picos em uma amostra de gasolina típica, porém são identificados aproximadamente 200.

Tendo-se identificado muitos picos por comparação com padrões (pentano, hexano, heptano, isoctano, 2,2-dimetilbutano, benzeno e tolueno) e uma mistura de referências da SupelcoTM, quantificou-se uma amostra de gasolina de maneira a confirmar a identificação automática do DHA. Como o DHA utiliza o cromatograma obtido por GC-FID, é importante salientar que este cromatograma apresenta os mesmos picos que a corrida no CG-EM. Iniciou-se então a comparação da mesma amostra injetada no CG-EM, confrontando o espectro de massas com as substâncias identificadas pela biblioteca do DHA (PIANO), observando-se a semelhança entre os cromatogramas e se o pico nomeado correspondia à função química no espectro de massas. Ao encontrar em literatura

as normas ASTM D6729-04 e D6730-01, aproveitou-se também as bibliotecas apresentadas (estavam com um maior número de substâncias possíveis do que o usado pelo DHA). Confirmou-se novamente a maioria das substâncias e outras ainda não identificadas pela ordem de eluição e por espectrometria de massas.

Finalmente identificou-se a função química e o número de carbonos de outras substâncias que não foram identificadas nos passos anteriores, utilizando-se as técnicas auxiliares de espectrometria de massas e de ionização química. Na técnica de ionização química com líquidos, grupos saturados como os parafínicos, isoparafínicos e naftênicos apresentaram fragmentos do tipo $[M-1]^+$, enquanto que, para os compostos insaturados como os aromáticos e olefínicos, os fragmentos gerados foram do tipo $[M+1]^+$. Esta técnica serviu para auxiliar a identificação de grupos que apresentaram o mesmo espectro de massas por impacto de elétrons. Por exemplo, a ionização química permitiu diferenciar os olefínicos dos naftênicos pela formação dos fragmentos $[M+1]^+$ e $[M-1]^+$ respectivamente. Com este procedimento, o teor de desconhecidos baixou de 8% (v/v) para 2,5% (v/v) em média. Em resumo, foram identificadas 290 substâncias por comparação com as normas ASTM, banco de dados do DHA e compostos puros, com apoio complementar dos espectros de massas e ionização química. Salienta-se que diversas amostras de gasolina comum e aditivada foram estudadas por CG-EM e todas apresentaram o mesmo perfil cromatográfico, com as mesmas substâncias, variando-se apenas a quantidade de cada composto. Foram nomeados 100 compostos desconhecidos, sendo que a maioria dos compostos eram aromáticos e olefínicos.

Modelagem Matemática

O objetivo das técnicas PLS (*Partial Least Square*) e PCR (*Principal Component Regression*) é desenvolver um modelo matemático representativo do sistema químico e usá-lo para prever propriedades de amostras testes a partir de seus espectros. São próprias para rotinas, determinações rápidas e não destrutivas da composição química ou outra propriedade de interesse, aplicando-se espectroscopia. São simples de utilizar e aplicam-se ao espectro como um todo; portanto não requerem seleção de picos ou linhas de base. São métodos analíticos robustos e devem ser usados quando existem bandas espectrais sobrepostas ou quando são afetadas por variações de linha de base, ruído espectral, impurezas de amostras e respostas não lineares. Também podem detectar erros operacionais, bem como impurezas ou devido à preparação de amostras, indicando que os resultados podem estar incorretos. O número de dados que podem ser analisados não é fixo, limitado somente pela memória do computador. Um grande número de amostras de calibração pode existir tanto quanto o número de fontes de absorbância presentes na amostra. Tais fontes incluem cada espécie química conhecida, assim como outras fontes de interferência espectral, como interações químicas, óticas e variações na linha de base. Estas técnicas ignoram qualquer dado que se mantenha constante em todos os padrões e amostras. Também compensam automaticamente as fontes de interferência quando estão presentes nas amostras desconhecidas e de calibração. A concentração é uma medida típica, porém outros atributos podem ser utilizados para estudo (George e Willis, 1990).

As ferramentas de regressão PCR e PLS são ferramentas clássicas, muito utilizadas em química analítica. Um princípio importante do PLS é que a modelagem da informação da concentração (matriz **Y**) é tão importante quanto a modelagem da informação da matriz **X**. Na modelagem por PCR, os componentes principais (CPs) são calculados exclusivamente a partir da matriz **X**, e não levam em consideração a matriz **Y**. No PLS, os componentes são obtidos usando ambas as matrizes simultaneamente. Os tratamentos mais convencionais dos mínimos quadrados envolvem encontrar uma variável que maximiza o valor de X^2 , ou o quadrado dos dados experimentais modelados. O PLS encontra uma variável que maximiza o produto dos dados experimentais [**X**] com as concentrações [**Y**] freqüentemente chamado de covariância. Como estas ferramentas são clássicas, não se pretende detalhar matematicamente o que acontece para a inversão das matrizes, já que existem

diversos livros que versam sobre este assunto com riquezas de informação (Brereton, 2000 e 2007; Poppi, 1989). Detalha-se portanto a aplicação e análise de resultados gerados pela modelagem, que é mais importante neste estudo.

Após a calibração da composição química iniciou-se a montagem da matriz de entrada [X], que abrange os espectros no infravermelho das amostras, com 1588 absorbâncias, bem como as matrizes de saída [Y], que referem-se às propriedades físico-químicas da gasolina (massa específica, PV Reid, T_{10} , T_{50} , T_{90} e ponto final de ebulição da curva de destilação) e composição química obtida por cromatografia a gás (teor de aromáticos, parafínicos, isoparafínicos, naftênicos, olefínicos e oxigenados). O conjunto de 128 amostras foi misturado de maneira randômica, para criar o grupo de treinamento, com 90 amostras e outro com 38 amostras para formar o grupo de teste. Desta maneira a representatividade do banco inicial estaria nos subgrupos. Para a obtenção dos modelos matemáticos, aplicou-se diversos pré-tratamentos na matrix **X** de treinamento: autoescalonamento, segunda derivada, alisamento e ortogonalização (Wold *et. al*, 1998 e Brereton, 2007). O ajuste dos dados para o PLS ocorreu para todas as propriedades, com o número de variáveis latentes de 3 a 4 em média, conforme mostrado na tabela 1. A escolha do número de variáveis latentes foi o passo crítico no desenvolvimento de cada modelo. No geral, se um número insuficiente de VL é utilizado, o modelo terá pouca precisão, enquanto que se um número maior do que o necessário for usado, ruído instrumental terá importância significativa na previsão do conjunto de teste. Todas as propriedades físico-químicas e a composição química obtiveram ajuste com as 90 amostras de gasolina do conjunto de treinamento, sendo que em alguns deles houve necessidade de retirada de pontos para melhorar a capacidade preditiva dos modelos. O ajuste ocorreu porque os equipamentos utilizados neste estudo estavam calibrados e o pré-tratamento dos espectros de IV ajudou a minimizar o ruído instrumental do equipamento de infravermelho. Os modelos PLS tiveram capacidade preditiva igual ou superior e, utilizaram menor número de variáveis latentes, do que os modelos PCR.

Tabela 1. Resumo de parâmetros dos modelos PLS e PCR obtidos neste estudo.

Propriedade físico-química ou composição da gasolina	Número de VL/PC	Faixa de utilização (PLS)	RMSEC PLS/PCR	RMSECV PLS/PCR	RMSEP PLS/PCR
Massa específica	3/4	752,0 - 764,0	0,36 / 0,38	3,94 / 3,90	3,50 / 3,48
Pressão de vapor Reid	4/4	17,0 - 65,0	0,65 / 1,31	3,04 / 2,82	2,81 / 2,61
Curva de destilação: T_{10}	6/7	50,0 - 76,0	0,17 / 0,66	2,75 / 2,36	2,18 / 2,03
Curva de destilação: T_{50}	2/4	65,0 - 76,0	0,11 / 0,12	1,36 / 1,38	1,02 / 1,02
Curva de destilação: T_{90}	5/9	100,0 - 195,0	1,59 / 3,72	10,44 / 9,89	9,85 / 9,73
Curva de destilação: T_{PFE}	6/11	160,0 - 320,0	2,86 / 3,38	21,11 / 22,02	17,53 / 18,01
% aromáticos	5/4	1,0 - 20,0	0,25 / 0,58	1,02 / 0,85	1,14 / 1,369
% parafínicos	3/5	8,0 - 23,5	0,21 / 0,21	1,13 / 1,13	1,18 / 1,17
% isoparafínicos	3/4	12,5 - 25,0	0,14 / 0,20	1,48 / 1,42	1,06 / 1,11
% naftênicos	3/3	10,0 - 48,0	0,31 / 0,69	1,08 / 1,14	0,80 / 0,92
% olefínicos	4/4	0,5 - 20,0	0,36 / 0,56	1,26 / 1,26	1,40 / 1,44
% oxigenados	3/3	17,0 - 47,0	0,42 / 0,65	1,66 / 1,61	1,06 / 1,03

Os modelos PLS podem ser inseridos em um equipamento de infravermelho portátil, que execute a leitura do espectro no intervalo de 740,0 a 3800,0 cm^{-1} , e que possua um sensor de fibra ótica com ATR. Os modelos podem fazer parte de um banco de dados interno, no qual poderá ser inserido mais amostras diferenciadas para melhorar a robustez dos mesmos. Leituras de novas

amostras poderão ser realizadas *in situ*, já que o equipamento poderá ser transportado facilmente para estudos em campo, desde que seja desenvolvido um sistema de segurança para transporte.

Tabela 2. Resumo dos desvios do conjunto de teste.

Propriedade	Repetitividade Infravermelho (%)	Reprodutibilidade Equipamento, máx. (%)	Desvio total (IV + EQ)	Média dos desvios de previsão (%)
Massa específica	0,02	0,05	0,1	1,6
Pressão de vapor Reid	1,24	0,33	1,3	3,9
Curva de destilação: T ₁₀	2,22	1,37	2,6	2,8
Curva de destilação: T ₅₀	0,21	0,42	0,5	0,9
Curva de destilação: T ₉₀	0,43	1,6	1,7	3,6
Curva de destilação: T _{PFE}	0,54	0,85	1,0	6,6
% aromáticos	0,58	1,0	1,2	7,1
% parafínicos	2,42	1,0	2,6	4,8
% isoparafínicos	3,26	1,0	3,4	4,4
% naftênicos	1,51	1,0	1,8	2,4
% olefínicos	7,02	1,0	7,1	11,6
% oxigenados	0,66	1,0	1,2	3,3

Na tabela 2 observa-se também que a média dos desvios de previsão também são maiores para as propriedades de PV Reid (3,9%), T₉₀ (3,6%) e ponto final de destilação (6,6%), indicando a dificuldade de previsão pelos modelos obtidos. Essa maior variação é considerada normal comparada às outras propriedades, cuja detecção é mais complexa e sujeita a interferências aleatórias. Observa-se também, que a média dos desvios de previsão para o teor de aromáticos (7,1%) e olefínicos (11,6%) é maior, comparada aos teores dos outros grupos químicos. Nota-se que o próprio erro inserido pelo espectro infravermelho (coluna 4 da tabela 3) já é alto, devido a variações aleatórias provenientes de ruído espectral. Esses erros aleatórios contribuíram para aumentar os valores dos desvios de previsão da composição química.

Conclusões

A técnica de regressão de dados PLS alcançou os melhores ajustes de treinamento para todas as propriedades. Sendo assim, obteve-se modelos matemáticos capazes de identificar com rapidez uma composição química média a partir de espectros no IV, além de prever com pequenos desvios as propriedades de interesse da gasolina. Diversos pré-tratamentos aplicados foram úteis na redução do ruído e um equipamento portátil poderá ser construído, utilizando-se todo o banco de dados gerado. Problemas práticos podem ocorrer com um equipamento portátil, principalmente em relação ao seu transporte, pois o sistema ótico é bastante frágil. Porém este problema poderá ser contornado, escolhendo-se sistemas óticos mais robustos com relação às variações de temperatura e vibrações, os quais ocorrem devido ao deslocamento do equipamento. Outra aplicação interessante seria o de sensor virtual (Zanata, 2005), visando um controle de processo, para manter a qualidade da gasolina dentro das especificações para comercialização. Com o tempo de resposta mais otimizado, o sensor pode ajudar na tomada de decisões para gerenciamento e otimização do processo. A maior dificuldade em gerenciar este tipo de dispositivo dentro de uma unidade industrial, está relacionada com a fragilidade do sensor IV/ATR, que pode sofrer influência de pressão e temperatura, incrustações, riscos ou danos

no cristal de seleneto de zinco. Evidentemente os espectros obtidos com estas fontes de ruído, podem fornecer previsões muito fora da faixa de controle, atrapalhando ou fornecendo previsões errôneas sobre o processo. Para uso industrial, o desenvolvimento de um sensor com proteção, além da melhora na robustez do modelo matemático pode viabilizar a aplicação. A redução do tempo de análise, tanto para o sensor virtual, quanto para o equipamento portátil, é expressiva. O tempo de análise para uma amostra desconhecida passa a ser de 5 minutos, que é o tempo necessário para aquisição de um espectro infravermelho, seguido de previsão de sua composição química e propriedades físico-químicas.

Referências Bibliográficas

AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **ASTM D 6729:** Determination of individual components in spark ignition engine fuels by 100 metre capillary high resolution gas chromatography. 2004.

AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **ASTM D 6730:** Determination of individual components in spark ignition engine fuels by 100-metre capillary high-resolution gas chromatography. 2001.

BRERETON, R. G. **Introduction to multivariate calibration in analytical chemistry.** Analyst. v. 125, p. 2125-2154, October, 2000.

BRERETON, R. G. **Applied chemometrics for scientists.** Hoboken: J. Wiley, 2007, 379 p.

GALLIGNANI, M.; GARRIGUES, S.; de la GUARDIA, M. Direct determination of benzene in gasoline by flow injection-fourier transform infrared spectrometry. **Anal. Chim. Acta.** v. 274, p. 267-274, 1993a.

GALLIGNANI, M.; GARRIGUES, S.; DE LA GUARDIA, M. Flow injection derivative fourier transform infrared determination of methyl tert-butyl ether in gasolines. **Anal. Chim. Acta.** v. 282, p. 543-550, 1993b.

GALLIGNANI, M. *et al.* Comparative study of different approaches for the flow injection-fourier transforms infrared determination of toluene in gasolines. **Talanta.** v. 41, n.5, p. 739-745, 1994.

GEORGE, W. O. and WILLIS, H. A. **Computer methods in UV, visible and IR spectroscopy.** Cambridge: Royal Society of Chemistry, 1990, 216 p.

MACHO, S.; LARRECHI, M.S.. Near infrared spectroscopy and multivariate calibration for the quantitative determination of certain properties in the petrochemical industry. **Trends in Anal. Chem.** v. 21, n.12, p. 799-806, 2002.

POPPI, R. J. **Quantificação de picos cromatográficos superpostos por métodos de calibração multivariada.** 107 f. Dissertação (Mestrado em Química) – Universidade Estadual de Campinas. Campinas, 1989.

WANG, Z.; FINGAS, M. Development of oil hydrocarbon fingerprint and identification techniques. **Marine Poll. Bulletin.** v. 47, p. 423-452, 2003.

WOLD, S. *et al.* Orthogonal signal correction of near infrared spectra. **Chem. Intell. Lab. Systems.** v. 44, p. 175-185, 1998.

WU, W. *et al.* Using contrasts as data pretreatment method in pattern recognition of multivariate data. **Chem. Intell. Lab. Systems.** v.45, p. 39-53, 1999.

ZANATA, D. R. P. **Desenvolvimento de sensor virtual empregando redes neurais para medição da composição em uma coluna de destilação.** 229 f. Dissertação (Mestrado em Engenharia) – Escola Politécnica da Universidade de São Paulo. São Paulo, 2005.