

UNESP – UNIVERSIDADE ESTADUAL PAULISTA
CÂMPUS DE ARARAQUARA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

RAFAEL RODRIGUES HATANAKA

Desenvolvimento, validação, avaliação da incerteza de medição e análise de custos de método baseado em espectroscopia no infravermelho e análise multivariada para previsão de parâmetros de qualidade de amostras de E85

Araraquara

2015

RAFAEL RODRIGUES HATANAKA

Desenvolvimento, validação, avaliação da incerteza de medição e análise de custos de método baseado em espectroscopia no infravermelho e análise multivariada para previsão de parâmetros de qualidade de amostras de E85

Tese apresentada ao Instituto de Química,
Universidade Estadual Paulista, como parte
dos requisitos para obtenção do título de
Doutor em Química

Orientador: Prof. Dr. José Eduardo de Oliveira

Araraquara
2015

FICHA CATALOGRÁFICA

- H361d Hatanaka, Rafael Rodrigues
Desenvolvimento, validação, avaliação da incerteza de medição e análise de custos de método baseado em espectroscopia no infravermelho e análise multivariada para previsão de parâmetros de qualidade de amostras de E85: / Rafael Rodrigues Hatanaka. – Araraquara : [s.n], 2015
150 f. : il.
- Tese (doutorado) – Universidade Estadual Paulista, Instituto de Química
Orientador: José Eduardo de Oliveira
1. Espectroscopia de infravermelho. 2. Incerteza de medição (Estatística). 3. Quimiometria. 4. Análise industrial. 5. Estudos de validação. I. Título.

RAFAEL RODRIGUES HATANAKA

Tese de Doutorado apresentada ao
Instituto de Química, Universidade
Estadual Paulista, como parte dos
requisitos para obtenção do título de
Doutor em Química.

Araraquara, 02 de abril de 2015.



BANCA EXAMINADORA

Prof. Dr. JOSÉ EDUARDO DE OLIVEIRA (Orientador)
Instituto de Química / UNESP / Araraquara – SP



Prof. Dr. EDENIR RODRIGUES PEREIRA FILHO
Universidade Federal de São Carlos / UFSCAR / São Carlos - SP



Prof. Dr. ALBERICO BORGES FERREIRA DA SILVA
Instituto de Química / USP / São Carlos – SP



Prof. Dr. IVANO GEBHARDT ROLF GUTZ
Instituto de Química / USP / São Paulo - SP



Prof. Dr. LEONARDO SENA GOMES TEIXEIRA
Universidade Federal da Bahia / UFBA / Bahia - BA

Curriculum Vitae

Dados pessoais

Nome Rafael Rodrigues Hatanaka

Nascimento 28/06/1986 - Araraquara/SP - Brasil

CPF 833.947.372-72

Formação acadêmica/titulação

2011 Doutorado em Química.
Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, Sao Paulo,
Brasil
Orientador: José Eduardo de Oliveira

2009 - 2010 Mestrado em Química.
Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, São Paulo,
Brasil
Título: Desenvolvimento de método analítico para determinação do teor de
etanol em óleo lubrificante de motores do ciclo Otto usado, Ano de obtenção: 2010
Orientador: Prof. Dr. José Eduardo de Oliveira
Bolsista do(a): Coordenação de Aperfeiçoamento de Pessoal de Nível
Superior

2005 - 2008 Graduação Bacharelado em Química.
Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, São Paulo,
Brasil
Título: Caracterização e Previsão da Qualidade das Gasolinas Comerciais
Brasileiras a partir dos Perfis Cromatográficos obtidos pela ASTM D6729 e Métodos
Quimiométricos de Reconhecimento de Padrões
Orientador: José Eduardo de Oliveira
Bolsista do(a): Fundação para o Desenvolvimento da UNESP

2001 - 2003 Ensino Profissional de nível técnico.

Centro Federal de Educação Tecnológica, CEFET-PA, Brasil

Formação complementar

2009 - 2009 Curso de curta duração em NBR ISO/IEC 17025:2005 Formação de Auditor Interno.

Fundação CERTI, CERTI, Florianópolis, Brasil

2009 - 2009 Curso de curta duração em Formação de Auditores Internos - NBR ISO 14001:200.

Centro das Indústrias do Estado de São Paulo, CIESP, São Paulo, Brasil

2009 - 2009 Curso de curta duração em Incerteza de Medição.

Fundação CERTI, CERTI, Florianópolis, Brasil

2008 - 2008 Curso de curta duração em Implantação da ABNT ISO/IEC 17025:2005.

Fundação CERTI, CERTI, Florianópolis, Brasil

2005 - 2007 Extensão universitária em Química.

Química Júnior Projetos e Consultoria, QJPC, Brasil

Atuação profissional

1. Fundação para o Desenvolvimento da UNESP - FUNDUNESP

Vínculo institucional

2014 - Atual Vínculo: Celetista, Enquadramento funcional: Gerente Técnico, Carga horária: 36. Regime: Integral

2013 - 2014 Enquadramento funcional: Assessor de pesquisa e inovação, Carga horária: 36. Regime: Integral

2. ABNT/CB-164 - Comitê Brasileiro de Tintas

Vínculo institucional

2015 - Atual Vínculo: Voluntário, Enquadramento funcional: Secretário, Carga horária: 1, Regime: Parcial

3. Universidade Estadual Paulista Júlio de Mesquita Filho - UNESP

Vínculo institucional

2011 - 2013 Vínculo: Bolsista, Enquadramento funcional: Bolsista de doutorado, Carga horária: 40. Regime: Integral

2009 - 2010 Vínculo: Bolsista, Enquadramento funcional: Bolsista de mestrado, Carga horária: 40. Regime: Integral

2008 - 2008 Vínculo: Bolsista, Enquadramento funcional: Bolsista de iniciação científica, Carga horária: 20. Regime: parcial

Produção

Produção bibliográfica

Artigos completos publicados em periódicos

1. **HATANAKA, R. R.**, SEQUINEL, R., GUALTIERI, C. E., TERCINI, A. C. B., FLUMIGNAN, D. L., de OLIVEIRA, J. E.

Development and validation of an environmentally friendly attenuated total reflectance in the mid-infrared region method for the determination of ethanol content in used engine lubrication oil. *Talanta (Oxford)*. , v.109, p.191 - 196, 2013.

2. BERGAMASCO, JULIANA, DE ARAUJO, MARCELO V., DE VASCONCELLOS, ADRIANO, LUIZON FILHO, ROBERTO A., **HATANAKA, R. R.**, GIOTTO, MARCUS V., ARANDA, DONATO A.G., NERY, JOSÉ G.

Enzymatic transesterification of soybean oil with ethanol using lipases immobilized on highly crystalline PVA microspheres. *Biomass & Bioenergy*. , v.59, p.218 - 233, 2013.

3. FLUMIGNAN, D. L., SEQUINEL, R., **HATANAKA, R. R.**, BORALLE, N., de OLIVEIRA, J. E.

Carbon Nuclear Magnetic Resonance Spectroscopic Profiles coupled to Partial Least-Squares Multivariate Regression for Prediction of Several Physicochemical Parameters of Brazilian Commercial Gasoline. *Energy & Fuels (Online)*. , v.26, p.5711 - 5718, 2012.

4. FLUMIGNAN, D. L., SEQUINEL, R., **HATANAKA, R. R.**, BORALLE, N., de OLIVEIRA, J. E.

Multivariate calibrations on ¹H NMR profiles for prediction of physicochemical parameters of Brazilian commercial gasoline. *Fuel (Guildford)*. , v.99, p.180 - 187, 2012.

5. SEQUINEL, R., **HATANAKA, R. R.**, GUALTIERI, C. E., FLUMIGNAN, D. L., de OLIVEIRA, J. E., PASSARETTI FILHO, J.

Cromatografia gasosa ultrarrápida: uma visão geral sobre parâmetros, instrumentação e aplicações. *Química Nova (Impresso)*. , v.33, p.2226 - 2232, 2010.

6. **HATANAKA, R. R.**, FLUMIGNAN, D. L., de OLIVEIRA, J. E.

GC Fingerprints Coupled to Pattern-Recognition Multivariate SIMCA Chemometric Analysis for Brazilian Gasoline Quality Studies. *Chromatographia (Wiesbaden)*. , v.70, p.1135 - 1142, 2009.

Capítulos de livros publicados

1. FLUMIGNAN, D. L., de OLIVEIRA, J. E., **HATANAKA, R. R.**, SEQUINEL, R.

Avanços brasileiros no desenvolvimento de normas técnicas analíticas para certificação e controle da qualidade de biodiesel In: BIOENERGIA - Desenvolvimento, Pesquisa e Inovação.1 ed.São Paulo : Editora UNESP, 2012, p. 889-943.

2. SEQUINEL, R., **HATANAKA, R. R.**, FLUMIGNAN, D. L., de OLIVEIRA, J. E.

Overview of Pattern-Recognition Multivariate Algorithms Applied to Petrofuels Quality Control and Origin Determination In: Handbook of Pattern Recognition: Methods and Application.1 ed. Hong Kong: iConcept Press Ltd, 2012, p. 1-26.

Trabalhos publicados em anais de eventos (resumo)

1. MIRANDA, N. T., FERREIRA, P. S., RUFINO, M. P., **HATANAKA, R. R.**, SEQUINEL, R., de OLIVEIRA, J. E., TININIS, A. G., FLUMIGNAN, D. L.

DESENVOLVIMENTO E VALIDAÇÃO DE MÉTODO CROMATOGRÁFICO GASOSO ULTRARRÁPIDO PARA DETERMINAÇÃO DE BTEX EM GASOLINAS COMERCIAIS BRASILEIRAS In: Congresso Latino-Americano de Cromatografia e Técnicas Relacionadas (COLACRO XIV), 2012, Florianópolis.

Livro de Resumos COLACRO XIV. , 2012. p.491 - 491

2. **HATANAKA, R. R.**, SEQUINEL, R., GUALTIERI, C. E., TERCINI, A. C. B., de OLIVEIRA, J. E.

DEVELOPMENT AND VALIDATION OF ANALYTICAL METHOD FOR DETERMINATION OF RESIDUAL ETHANOL IN USED OTTO CYCLE ENGINES LUBRICATING OIL In: XIII Chemometrics in Analytical Chemistry, 2012, Budapeste.

XIII Chemometrics in Analytical Chemistry. , 2012. p.117 - 117

3. SEQUINEL, R., **HATANAKA, R. R.**, GUALTIERI, C. E., CHIARADIA, J. R., PEZZA, L., de OLIVEIRA, J. E.

Determinação de metanol em glicerina proveniente de usinas de biodiesel utilizando a técnica de CG com amostragem por Headspace In: 34 Reunião Anual da Sociedade Brasileira de Química, 2011, Florianópolis.

Anais da 34 Reunião Anual da Sociedade Brasileira de Química. , 2011.

Trabalhos publicados em anais de eventos (resumo expandido)

1. SINZATO, M. M., **HATANAKA, R. R.**, de OLIVEIRA, J. E.

Extração e transesterificação do óleo da semente de *Leucaena* spp. e avaliação da qualidade do biodiesel obtido In: 6º Simpósio Nacional de Biocombustíveis, 2013, Canoas.

6º Simpósio Nacional de Biocombustíveis. , 2013.

2. SINZATO, M. M., **HATANAKA, R. R.**, VIVIANI, V. E., HATANAKA, T., de OLIVEIRA, J. E.

Extração e transesterificação do óleo da semente de *Leucaena* spp. e avaliação da qualidade do biodiesel obtido. In: 5º Congresso da Rede Brasileira de Tecnologia de Biodiesel / 8º Congresso Brasileiro de Plantas Oleaginosas, Óleos, Gorduras e Biodiesel, 2012, Salvador.

5º Congresso da Rede Brasileira de Tecnologia de Biodiesel / 8º Congresso Brasileiro de Plantas Oleaginosas, Óleos, Gorduras e Biodiesel. , 2012. v.1. p.387 - 388

3. CAMARGO, N. L., **HATANAKA, R. R.**, FLUMIGNAN, D. L., de OLIVEIRA, J. E.

AVALIAÇÃO DA QUALIDADE DAS GASOLINAS BRASILEIRAS ATRAVÉS DA APLICAÇÃO DO SIMCA EM PERFIS FAST HRGC In: SIMCRO 2010, Campos do Jordão.

SIMCRO 2010. , 2010.

4. **HATANAKA, R. R.**, GASPARINI, F., SEQUINEL, R., GHANI, Y.A., FLUMIGNAN, D. L., Lima, J.R. de O., de OLIVEIRA, J. E.

Avaliação e adaptação das condições da EN 14103 para quantificação de ésteres em misturas de biodieseis etílicos de soja, babaçu e palma In: III Congresso da Rede Brasileira de Tecnologia de BIODIESEL - RBTB, 2009, Brasília.

III Congresso da Rede Brasileira de Tecnologia de BIODIESEL. , 2009. v.3. p.129 - 130

5. FLUMIGNAN, D. L., **HATANAKA, R. R.**, de OLIVEIRA, J. E.

Caracterização e Previsão da Qualidade de Gasolinas Comerciais Brasileiras a partir dos Perfis Cromatográficos obtidos pela ASTM D6729 e Métodos Quimiométricos de Reconhecimento de Padrões In: 32ª SBQ, 2009, Fortaleza.

32ª Reunião Anual da Sociedade Brasileira de Química. , 2009.

Artigos em jornal de notícias

1. HATANAKA, Rafael Rodrigues

Álcool gel deve ser complemento de higiene das mãos (entrevistado). Tribuna impressa. Araraquara, p.5 - 5, 2009.

2. HATANAKA, Rafael Rodrigues

Químico explica diferença dos alcoóis no combate ao vírus H1N1 (entrevistado). Agência de notícias da Uniara. , 2009.

Produção técnica

Trabalhos técnicos

1. FLUMIGNAN, D. L., Lima, J.R. de O., GASPARINI, F., **HATANAKA, Rafael Rodrigues**, Palombo, A.J.D.A., BARRETOS, D.A., de OLIVEIRA, J. E.

4a. Edição do Programa de Comparação Interlaboratorial em Biodiesel da ANP,
2009

Dedico este trabalho à minha família, e à minha amada esposa Denise.

Agradecimentos

À Deus.

À minha família, pela educação, criação, incentivo e convivência.

À minha esposa pelo apoio, pela paciência e pelo companheirismo, que tornam mais fácil a transposição de qualquer dificuldade em minha vida.

Aos meus amigos.

Ao Cempec e todos os seus colaboradores.

À Fundunesp.

Ao meu orientador, pelos ensinamentos da vida científica, profissional e pessoal.

Ao prof. Paulo Roberto Guimarães Couto que me auxiliou na incerteza de medição.

Ao CNPq pela bolsa durante parte de meu doutorado.

Ao Instituto de Química da Unesp de Araraquara por todos os anos de acolhimento.

“No meio da dificuldade encontra-se a oportunidade”

Albert Einstein

Resumo

O uso do etanol como combustível pode reduzir as emissões dos gases do efeito estufa e a dependência de combustíveis fósseis, já que pode ser obtido de fontes renováveis. Entretanto, este biocombustível pode apresentar problemas de ignição e de dirigibilidade em regiões com clima frio devido à menor pressão de vapor em relação à gasolina. Para evitar tais problemas, pode-se misturar etanol à gasolina em várias proporções (50-85% na Europa e nos Estados Unidos), essas misturas são comercialmente denominadas de E85. Outros parâmetros (como o teor de metanol e de água) também são importantes para garantir a qualidade do combustível e estão previstos em normas internacionais. Sendo assim, o presente trabalho mostra o desenvolvimento, a validação e a estimativa de incerteza de métodos analíticos, utilizando espectroscopia na região do infravermelho (médio e próximo) e regressão multivariada, que permitem quantificar os teores de etanol, metanol, hidrocarbonetos e água em amostras de E85 preparadas em laboratório. A validação se deu através do cálculo das figuras de mérito utilizando o conceito de sinal analítico líquido (NAS). Com isso, foi possível observar que os métodos propostos são capazes de quantificar os teores de etanol, metanol, hidrocarbonetos e água de forma linear, livre de erros sistemáticos, com precisão de 0,2% (m/m), 0,03% (m/m), 0,2% (m/m) e 0,03% (m/m), exatidão de 0,3% (m/m), 0,02% (m/m), 0,2% (m/m) e 0,02% (m/m), respectivamente, e com intervalos de medição cobrindo toda faixa dos parâmetros estabelecidos pelas normas. Além disso, as formas pseudounivariadas dos modelos de regressão multivariada (utilizando o NAS) foram empregadas na estimativa da incerteza de medição dos métodos, permitindo os cálculos de acordo com a abordagem do Guia para Incerteza de Medição e também por simulação de Monte Carlo.

Abstract

The use of ethanol as fuel can reduce emissions of greenhouse gases and dependence on fossil fuels, as it can be obtained from renewable sources. However, this biofuel can present problems of ignition and driving in areas with cold climates due to lower vapor pressure in relation to gasoline. To avoid such problems, ethanol can be mixed with in various proportions (50-85% in Europe and the United States). These mixtures are commercially known as E85. Other parameters (such as methanol and water content) are also important to ensure the quality of the fuel and are fixed by international standards. Thus, the present study shows the development, validation and estimation of the uncertainty of analytical methods using spectroscopy in the infrared region (middle and near) and multivariate regression, which allow quantifying levels of ethanol, methanol, oil and water in E85 samples prepared in laboratory. The validation was made by calculating the figures of merit using the concept of net analytical signal (NAS). Thus, the proposed methods were able to quantify the levels of ethanol, methanol, oil and water linearly, free of systematic errors , with accuracy of 0.2% (m/m), 0.03% (m/m), 0.2% (m/m) and 0.03% (m/m), accuracy of 0.3% (m/m), 0.02% (m/m), 0.2% (m/m) and 0.02% (m/m), respectively, and measuring intervals covering whole range of parameters established by the standards. Furthermore, pseudounivariate forms of multivariate regression models (using NAS) were used to estimate uncertainty of measuring methods, allowing calculations in accordance with the Guide to the Expression of Uncertainty in Measurement and also Monte Carlo simulation.

LISTA DE FIGURAS

Figura 1 – Licenciamento de autoveículos novos no Brasil em 2014 por combustível	35
Figura 2 – Ilustração do ciclo de trabalho de um motor ciclo Otto.	36
Figura 3 – Número de publicações encontradas para as palavras “ethanol” e “fuel” na base de dados Scifinder (22/11/14).....	37
Figura 4 – Diagrama de fases para o sistema etanol, água e gasolina, em volume a 24 °C.....	40
Figura 5 – Representação gráfica do funcionamento do ATR. N é a normal; θ é o ângulo de incidência e d_p (do inglês <i>depth of penetration</i>) é a profundidade de penetração.	45
Figura 6 – Diagrama de energia indicando as transições possíveis para os osciladores harmônicos (A) e anarmônico (B).....	47
(Fonte: ³¹).....	47
Figura 7: Modos de medição empregados na espectroscopia na região do NIR: (a) transmitância; (b) transfletância; (c) refletância difusa; (d) interactância e (e) transmitância através de amostras sólidas.....	48
Figura 8 – Representação esquemática do funcionamento de uma esfera de integração.....	49
Figura 9 – Exemplo ilustrativo de espectros de quatro amostras fictícias, a tabela para comparação dos dois comprimentos de onda (1200 e 3200 cm^{-1}) e o gráfico mostrando a relação inversa entre as absorbâncias.	52
Figura 10: Representação gráfica da decomposição da matriz de dados X pela PCA.....	54

Figura 11 – Gráfico demonstrando as componentes principais PC1 e PC2, bem como a projeção das amostras sobre elas. A linha contínua representa os scores de uma amostra, a linha tracejada representa os <i>loadings</i> e a linha traço dois pontos representa as coordenadas de uma amostra em relação aos eixos originais.....	55
Figura 12 – Representação gráfica da decomposição das matrizes realizada no modelo PLS.....	57
Figura 13 – Representação geométrica da propriedade de ortogonalidade do NAS.....	59
Figura 14 – Exemplo de um diagrama causa e efeito para identificação das fontes de incerteza (X_1 , X_2 e X_3) na estimativa da incerteza de medição do mensurando y	67
Figura 15 – Acessórios utilizados para a leitura das amostras: (A) esfera de integração e (B) cristal de ZnSe para refletância total atenuada.....	79
Figura 16: Acessório original para leitura no modo de medição transfletância para utilização sobre a janela de safira da esfera de integração.....	80
Figura 17 – Acessório de aço inoxidável usinado desenvolvido para leitura no modo de medição transfletância para utilização sobre a janela de safira da esfera de integração.....	80
Figura 18 – Projeções dos acessórios de aço inoxidável usinado utilizados para a coleta dos espectros na região do NIR.	81
Figura 19 – Espectros de várias amostras diferentes de E85 adquiridos na faixa $4000\text{--}650\text{ cm}^{-1}$	84
Figura 20 – Espectros de diferentes amostras de E85 adquiridos na faixa $4000\text{--}1000\text{ cm}^{-1}$	85

Figura 21 – Espectros na região do MIR de cada um dos componentes utilizados na produção das amostras de combustível E85.....	87
Figura 22 – Regiões consideradas para a seleção de variáveis.....	88
Figura 23 – Erros de calibração e validação cruzada em função do número de variáveis latentes para os modelos (A) Etanol; (B) Metanol; (C) HC e (D) Água. As setas indicam o número de variáveis latentes escolhidas.....	90
Figura 24 – Espectros na região do NIR (caminho óptico de 0,2 mm) de cada um dos componentes utilizados na produção das amostras de combustível E85.....	91
Figura 25 – Espectros na região do NIR (caminho óptico de 0,5 mm) de cada um dos componentes utilizados na produção das amostras de combustível E85.....	92
Figura 26 – Erros de calibração em função do número de variáveis latentes para os modelos etanol, metanol; HC e água na região do NIR utilizando o caminho óptico de 0,2 mm. As setas indicam o número de variáveis latentes escolhidas.	94
Figura 27 – Erros de calibração em função do número de variáveis latentes para os modelos etanol, metanol; HC e água na região do NIR utilizando o caminho óptico de 0,5 mm. As setas indicam o número de variáveis latentes escolhidas	95
Figura 28 – Comportamento linear dos valores previstos pelos modelos em função dos valores de referência. Amostras de calibração (o) e amostras de validação (+).....	103
Figura 29 – Elipses de confiança para a inclinação e o intercepto das regressões lineares obtidas através dos valores previstos pelos modelos e os valores de referência.....	104

Figura 30 – Valores de referência em função dos valores de NAS escalar para os quatro modelos.....	105
Figura 31 – Valores dos scores da primeira variável latente em função dos valores de referência. Amostras de calibração (o) e amostras de validação (+).106	
Figura 32 – Comportamento aleatório dos valores dos erros dos modelos em função dos valores de referência. Amostras de calibração (o) e amostras de validação (+).....	107
Figura 32: Diagrama de causa e efeito mostrando as fontes de incerteza na determinação da concentração dos analitos nas amostras.....	114
Figura 33: Fonte de incerteza e suas porcentagens de contribuição para os quatro modelos.....	121
Figura 34 – Histograma representando a PDF resultante da simulação de Monte Carlo para o teor de etanol.....	124
Figura 35 – Histograma representando a PDF resultante da simulação de Monte Carlo para o teor de metanol.....	124
Figura 36 – Histograma representando a PDF resultante da simulação de Monte Carlo para o teor de HC	125
Figura 37 – Histograma representando a PDF resultante da simulação de Monte Carlo para o teor de Água	125
Figura 38 – Gráfico de barras mostrando a porcentagem de contribuição de cada item para o preço total do ensaio de determinação do teor de água por Karl Fishcer	133
Figura 39 – Gráfico de barras mostrando a porcentagem de contribuição de cada item para o preço total do ensaio de determinação do teor de metanol e etanol por cromatografia em fase gasosa	133

Figura 40 – Gráfico de barras mostrando a porcentagem de contribuição de cada item para o preço total do ensaio de determinação do teor de etanol, metanol, HC e água por IV e PLS 134

LISTA DE TABELAS E QUADROS

Tabela 1 – Principais parâmetros de qualidade para o E85 estabelecidos pelas normas internacionais ^{8,9}	38
Tabela 2 – Valores dos teores planejados (obtidos com a função "aleatórioentre") e experimentais para algumas amostras preparadas e utilizadas para a construção dos modelos.	83
Tabela 3 – Regiões selecionadas e os erros de cada modelo na região do MIR.	89
Tabela 4 – Regiões selecionadas e os erros de cada modelo na região do NIR, com utilização do caminho óptico de 0,2 mm.....	92
Tabela 5 – Regiões selecionadas e os erros de cada modelo na região do NIR, com utilização do caminho óptico de 0,5 mm.....	93
Tabela 6 – Valores de RMSECV de todos os modelos, ordenados em ordem crescente.....	96
Tabela 7 – Valores de F calculados para a comparação entre os modelos e valor de F tabelado.....	97
Tabela 8 – Figuras de mérito dos quatro modelos construídos para previsão dos teores de etanol, metanol, HC e água em amostras de E85.	98
Tabela 9 – Replicatas de três amostras de E85 para determinação da precisão em nível de repetibilidade para os quatro modelos.	99
Tabela 10 – Comparação da diferença entre algumas amostras do conjunto de previsão para avaliação da sensibilidade analítica dos modelos Etanol NIR05 e HC NIR05.....	101

Tabela 11 – Comparação da diferença entre algumas amostras do conjunto de previsão para avaliação da sensibilidade analítica dos modelos Metanol MIR e Água NIR05.....	101
Tabela 12 – Parâmetros das regressões lineares para os modelos pseudounivariados.	104
Tabela 13 – Parâmetros das regressões lineares dos valores dos <i>scores</i> da primeira variável latente em função dos valores de referência.....	107
Tabela 14 – Valores das concentrações obtidas em diferentes temperaturas no estudo de robustez.	110
Tabela 15 – Resultados da ANOVA com um fator para avaliar o efeito da variação de temperatura	111
Tabela 16 – Valores das concentrações dos analitos, obtidos através de dez réplicas nos quatro modelos para uma mesma amostra.....	112
Tabela 17 – Valores calculados W e valores críticos W_c da estatística do teste de Shapiro – Wilks para uma amostra quantificada nos quatro modelos.	112
Quadro 1 – Planilha do cálculo de incerteza expandida das medidas das massas das amostras de referência.....	118
Quadro 2 – Componentes de incerteza para a quantificação do etanol	118
Quadro 3 – Componentes de incerteza para a quantificação do metanol	119
Quadro 4 – Componentes de incerteza para a quantificação do HC.....	119
Quadro 5 – Componentes de incerteza para a quantificação da água.....	120
Tabela 18 – Fontes de entrada e respectivas PDFs para estimativa de incerteza do teor de etanol, metanol, HC e água	123

Tabela 19 – Parâmetros estatísticos obtidos através da simulação de Monte Carlo para os teores de etanol, metanol, HC e Água	126
Tabela 20 – Valores de d_{low} , d_{high} e δ para os quatro modelos	127
Tabela 21 – Descrição detalhada dos itens levados em consideração e seus respectivos valores para o cálculo do custo do ensaio de determinação do teor de água por Karl Fisher coulométrico.....	129
Tabela 22 – Descrição detalhada dos itens levados em consideração e seus respectivos valores para o cálculo do custo do ensaio de determinação do teor de etanol e metanol por cromatografia em fase gasosa.....	130
Tabela 23 – Descrição detalhada dos itens levados em consideração e seus respectivos valores para o cálculo do custo do ensaio de determinação do teor de etanol, metanol, HC e água por IV e PLS.....	131
Tabela 24 – Comparação dos custos dos ensaios considerando os equipamentos operando com um quantitativo de 190 amostras mensais e em sua capacidade máxima.....	132

LISTA DE ABREVIATURAS E SIGLAS

ABNT – Associação Brasileira de Normas Técnicas

ANFAVEA – Associação Nacional dos Fabricantes de Veículos Automotores

ANOVA – Análise de Variância (do Inglês *Analysis of Variance*)

ASTM – Formalmente conhecido como *American Society for Testing and Materials*

ATR – Refletância Total Atenuada (do inglês *Attenuated Total Reflectance*)

Bias – Erros Sistemáticos

CD – Custo Direto

Cempecq – Centro Multidisciplinar de Pesquisa em Combustíveis, Biocombustíveis, Petróleo e Derivados

CEN/TS – Comitê Europeu de Normalização (do francês *Comité Européen de Normalisation*)/Especificação Técnica (do inglês *Technical Specification*)

d_p – Profundidade de Penetração (do inglês *Depth of Penetration*)

E10 – Mistura com 10% de etanol e 90% de gasolina

E85 – Mistura com 85% de etanol e 15% de gasolina

EJCR – Região de Confiança Elíptica Conjunta (do inglês *Elliptical Joint Confidence Region*)

EN – Norma Europeia (do inglês *European Standard*)

FFV – Veículos *Flex–Fuel* (do inglês *Flex–Fuel Vehicles*)

FIR – Infravermelho Distante (do inglês *Far Infrared*)

FOM – Figuras de Mérito (do Inglês *Figures of Merit*)

FTIR – Infravermelho por Transformada de Fourier (do inglês *Fourier Transform Infrared*)

HC – Gasolina A

I.R. – Índice de Refração

ICE – Ignição por Centelha Elétrica

ISO GUM – *Guide to the Expression of Uncertainty in Measurement*

IUPAC – União Internacional da Química Pura e Aplicada (do inglês *International Union of Pure and Applied Chemistry*)

IV – Infravermelho

LD – Limite de Detecção

LQ – Limite de Quantificação

M – Número de Interações

MOD – Mão de Obra Direta

MQ – Média dos Quadrados

MSC – Correção do espalhamento Multiplicativo (do inglês *Multiplicative Scatter Correction*)

NAS – Sinal Analítico Líquido (do inglês *Net Analyte Signal*)

NBR – Norma Brasileira

NIR – Infravermelho Próximo (do inglês *Near Infrared*)

NIR02 – Espectros Obtidos com o Espelho de Caminho Óptico de 0,2 mm

NIR05 – Espectros Obtidos com o Espelho de Caminho Óptico de 0,5 mm

PCA – Análise de Componentes Principais (do inglês *Principal Component Analysis*)

PDF – Funções Densidade de Probabilidade (do inglês *Probability Density Functions*)

PLS – Mínimos Quadrados Parciais (do inglês *Partial Least Squares*)

R^2 – Coeficiente de Determinação

RMSEC – Média da Soma dos Quadrados dos Erros da Calibração (do inglês *Root Mean Square Error of Calibration*)

RMSECV – Média da Soma dos Quadrados dos Erros da Calibração Cruzada (do inglês *Root Mean Square Error of Cross-Validation*)

RMSEP – Média da Soma dos Quadrados dos Erros de Previsão (do inglês *Root Mean Square Error of Prediction*)

RSD – Desvio padrão relativo (do inglês *Relative Standard Deviation*)

SDV – Desvio Padrão dos Erros de Validação (do inglês *Standard Deviation of Validation*)

SQ – Soma dos Quadrados

VIM – Vocabulário Internacional de Metrologia

SUMÁRIO

1. INTRODUÇÃO	34
1.1. Notação.....	34
2. REVISÃO DA LITERATURA	35
2.1. Motores de ciclo Otto	35
2.2. E85	36
2.2.1. Etanol	39
2.2.2. Metanol.....	39
2.2.3. Água	39
2.2.4. Hidrocarbonetos	40
2.2.5. Pressão de vapor	40
2.2.6. Acidez e pH	41
2.2.7. Goma.....	41
2.2.8. Cloreto.....	41
2.2.9. Chumbo e Fósforo.....	41
2.2.10. Aparência	42
2.2.11. Cobre.....	42
2.2.12. Enxofre	42
2.3. Os efeitos da adição de etanol na gasolina.....	42

2.4. Espectroscopia na região do infravermelho	43
2.4.1. Espectroscopia na região do MIR.....	44
2.4.2. Espectroscopia na região do NIR	46
2.5. Quimiometria.....	49
2.5.1. Pré-processamento dos dados.....	50
2.5.1.1. Centrado na média	51
2.5.1.2. Correção multiplicativa de sinal	51
2.5.1.3. Derivadas	51
2.5.2. Análise de componentes principais	52
2.5.3. PLS.....	55
2.6. Validação	58
2.6.1. Sinal analítico líquido – NAS	59
2.6.2. Exatidão	60
2.6.3. Precisão	60
2.6.4. Sensibilidade	61
2.6.5. Sensibilidade Analítica	61
2.6.6. Seletividade	61
2.6.7. Linearidade.....	62
2.6.8. Ajuste	62
2.6.9. Razão sinal/ruído.....	62

2.6.10.	Intervalo de medição	63
2.6.11.	Limite de Detecção e Limite de Quantificação	63
2.6.12.	Teste para erros sistemáticos (Bias)	64
2.7.	Incerteza de medição.....	65
2.7.1.	Definição do mensurando.....	66
2.7.2.	Diagrama de causa e efeito.....	66
2.7.3.	Avaliação das incertezas padrão.....	67
2.7.4.	Incertezas correlacionadas e não correlacionadas.....	68
2.7.5.	Cálculo dos coeficientes de sensibilidade	68
2.7.6.	Componentes de incerteza.....	69
2.7.7.	Combinação das componentes	69
2.7.8.	Determinação dos graus de liberdade.....	71
2.7.9.	Determinação do coeficiente de abrangência	71
2.7.10.	Estimativa da incerteza expandida	71
2.8.	Simulação de Monte Carlo	72
2.9.	Gastos de uma empresa	74
2.9.1.	Investimento	74
2.9.2.	Custo	74
2.9.2.1.	Custos diretos.....	75
2.9.2.2.	Custos indiretos de fabricação	75

2.9.2.3.	Custos fixos	75
2.9.2.4.	Custos variáveis	76
2.9.3.	Despesa	76
2.9.4.	Perda	76
2.9.5.	Desperdício	76
3.	OBJETIVOS	77
4.	EXPERIMENTAL	78
4.1.	Reagentes	78
4.2.	Preparação das amostras	78
4.3.	Obtenção dos espectros de infravermelho	78
4.4.	Análise quimiométrica	81
4.5.	Validação	82
4.6.	Estimativa da Incerteza de Medição	82
4.7.	Estimativa de custos	82
5.	RESULTADOS E DISCUSSÃO	83
5.1.	Preparação das amostras	83
5.2.	Espectroscopia na região do infravermelho médio	84
5.3.	Espectroscopia na região do infravermelho próximo	85
5.4.	Regressão multivariada utilizando os espectros na região do MIR	86

5.5. Regressão multivariada utilizando os espectros na região do NIR.....	90
5.6. Seleção dos modelos com base nos erros	95
5.7. Validação	97
5.7.1. Exatidão e precisão	98
5.7.2. Sensibilidade	100
5.7.3. Seletividade	102
5.7.4. Ajuste	102
5.7.5. Linearidade.....	105
5.7.6. Razão sinal/ruído.....	108
5.7.7. Intervalo de medição ou faixa de medição	108
5.7.8. Limites de detecção e quantificação.....	108
5.7.9. Teste para erros sistemáticos (BIAS)	109
5.7.10. Robustez	109
5.8. Incerteza de medição.....	111
5.8.1. Definição do mensurando.....	112
5.8.2. Diagrama de causa e efeito.....	114
5.8.3. Avaliação das incertezas padrão.....	114
5.8.4. Coeficientes de sensibilidade	116
5.8.5. Componentes de incerteza.....	117
5.8.6. Graus de liberdade efetivos.....	121

5.8.7.	Fator de abrangência	121
5.8.8.	Incerteza expandida	122
5.9.	Incerteza de medição pelo método de Monte Carlo.....	122
5.10.	Comparação entre os métodos ISO GUM e simulação de Monte Carlo.....	126
5.11.	Estimativa dos custos dos ensaios	127
5.11.1.	Considerações.....	128
5.11.1.1.	Teor de água por Karl Fischer	128
5.11.1.2.	Teor de etanol e metanol por cromatografia em fase gasosa	129
5.11.1.3.	Teor de etanol, metanol, hidrocarbonetos e água por espectroscopia na região do infravermelho e quimiometria	131
5.11.2.	Avaliação da composição dos custos.....	132
6.	CONCLUSÕES	135
	REFERÊNCIAS.....	137

1. INTRODUÇÃO

Os parâmetros de qualidade do E85 podem variar de acordo com a época do ano e da região de comercialização. Para realizar a caracterização do combustível é necessário um grande número de ensaios normatizados internacionais, os quais muitas vezes são caros e morosos e necessitam que as amostras sejam coletadas e levadas até laboratórios dedicados. Sendo assim, há necessidade de realização de mais pesquisas relacionadas com o desenvolvimento de métodos que forneçam respostas mais rápidas que os métodos convencionais e com confiabilidade compatível.

Diante das considerações acima expostas, a proposta central deste estudo é a utilização de ferramentas quimiométricas para realizar o cruzamento de dados físico-químicos de amostras de E85 com diferentes características a fim de propor métodos analíticos, utilizando espectroscopia no infravermelho (médio e/ou próximo), para obter *fingerprints* das amostras que, aliados a ferramentas quimiométricas, permitam determinar a qualidade, bem como detectar possíveis adulterações e contaminações no combustível estudado.

Além disso, propor uma forma de avaliar a incerteza de medição em métodos que empregam a regressão multivariada por mínimos quadrados parciais (PLS) e comparar os valores das incertezas com os limites dos parâmetros de qualidade para E85 descritos nas Normas internacionais.

1.1. Notação

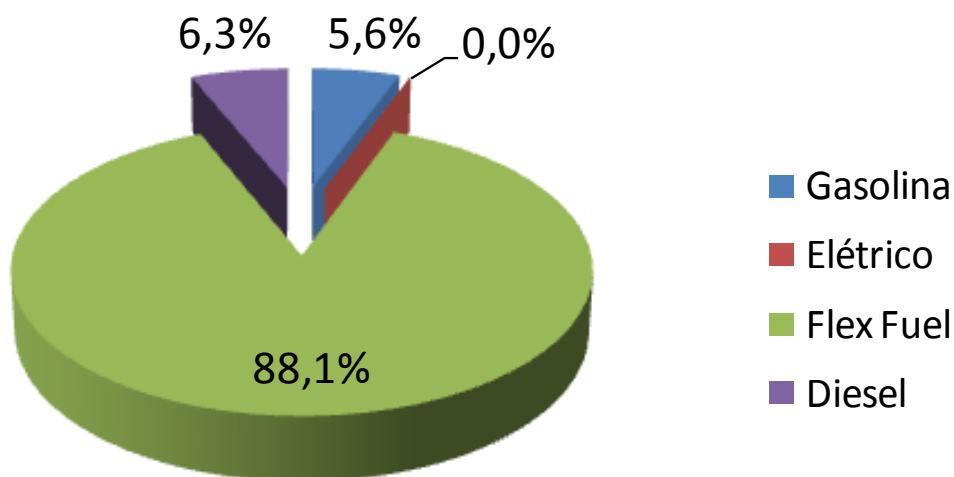
Nesta tese as matrizes são representadas pelas letras em maiúsculo e negrito (**X**), os vetores por letras minúsculas em negrito (**x**) e as grandezas escalares por letras minúsculas ou maiúsculas em itálico ou não (*J* ou *j*, *n* ou *N*).

2. REVISÃO DA LITERATURA

2.1. Motores de ciclo Otto

Inicialmente, será feita uma breve explanação de um motor ciclo Otto (seu funcionamento e as principais partes que o compõe), objetivando facilitar o entendimento, da maioria dos leitores, sobre os problemas causados por combustíveis fora da conformidade nesse tipo de motor.

A fim de demonstrar a importância desse tipo de motor para a indústria automobilística, são ilustrados alguns dados da Associação Nacional dos Fabricantes de Veículos Automotores – ANFAVEA, os quais mostram que a maioria (93,7%, **Figura 1**) dos carros licenciados em 2014 no Brasil são movidos a gasolina ou são do tipo *Flex-Fuel*¹, ou seja, saíram das fábricas equipados com motores ciclo Otto. Na Europa, os motores ciclo Otto propulsionam cerca de 42% dos veículos vendidos em 2013, já nos Estados Unidos esse valor é de aproximadamente 90%².



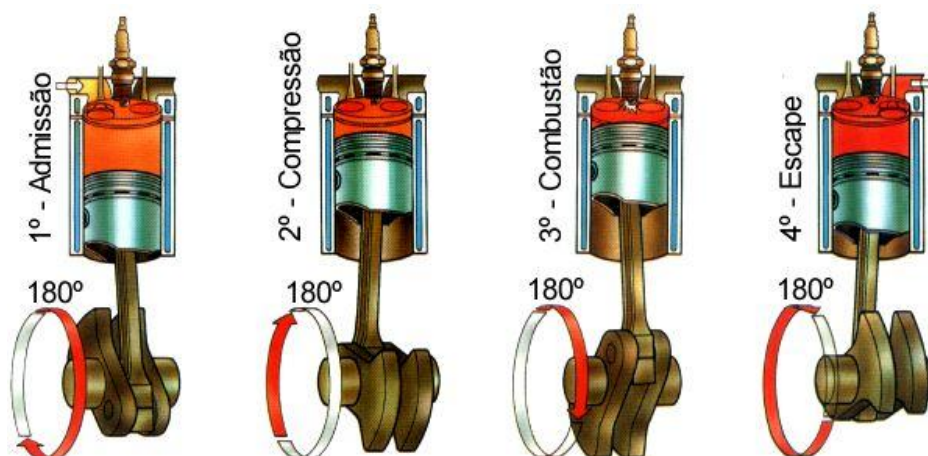
O motor de ignição por centelha elétrica (ICE) foi concebido em 1862, pelo físico francês Alphonse Beau de Rochas, quando descreveu uma sequência de eventos que totalizava quatro tempos (em síntese, é o que ocorre em todo motor ICE atual). Porém,

o motor só foi operacionalizado em 1876 por Nikolaus Otto e se tornou o popular motor Otto somente quando a empresa alemã Otto und Langen passou a fabricá-los ³.

Os quatro tempos do motor são admissão, compressão, expansão e exaustão. Na admissão, com a válvula de admissão aberta e com auxílio do sistema de injeção, ocorre a introdução da mistura ar-combustível na câmara de combustão. Depois, com as válvulas fechadas, o pistão se desloca, utilizando a energia cinética, comprimindo a mistura e a vela de ignição gera uma centelha que promove a reação de combustão.

A combustão gera gases e como as válvulas do sistema estão fechadas, ocorre a elevação da pressão e deslocamento do pistão no sentido contrário ao da compressão (estágio de expansão). Por fim, após a combustão, o pistão volta a se deslocar no mesmo sentido do tempo de compressão mas com a válvula de exaustão aberta, fazendo com que os produtos da reação sejam levados para o sistema de exaustão ³. Os processos descritos acima estão exemplificados na **Figura 2**.

Figura 2 – Ilustração do ciclo de trabalho de um motor ciclo Otto.



(Fonte: www.antonioguilherme.web.br.com/Images/quatrotempos.jpg)

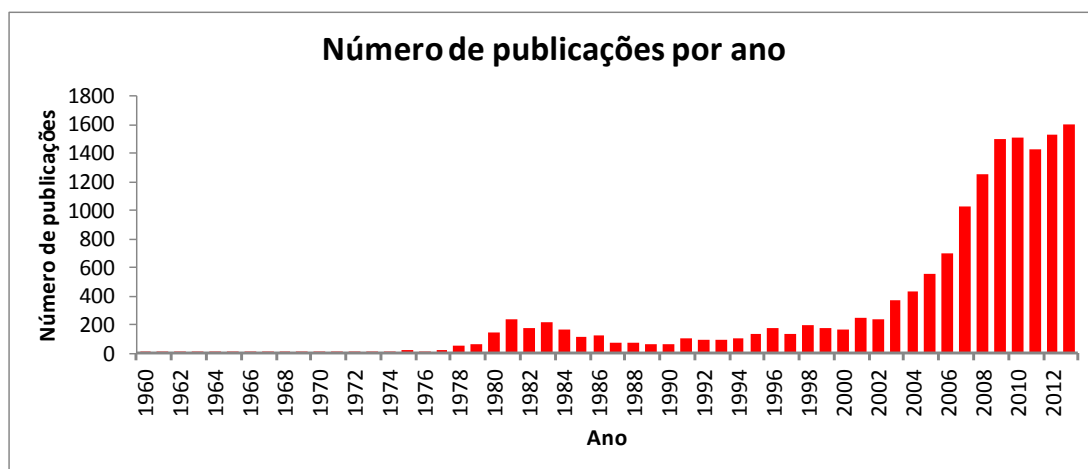
2.2. E85

A crescente demanda e futura escassez de combustíveis fósseis mostram o interesse dos países em apresentar alternativas viáveis para estes combustíveis. Os problemas ambientais associados aos combustíveis de origem fóssil, também são um

estímulo para procura de novas fontes renováveis de energia com menor impacto ambiental. Diante disto houve necessidade de criação de combustíveis alternativos, tecnicamente plausíveis, economicamente competitivos, prontamente disponíveis e aceitáveis do ponto de vista ambiental ^{4,5}.

O etanol vem ganhando grande importância na pesquisa e desenvolvimento (**Figura 3**), uma vez que pode ser produzido a partir de matérias primas renováveis, dentre elas a cana-de-açúcar, o milho e a celulose ^{6,7}. No Brasil, esse biocombustível é utilizado desde os anos 70 em sua forma hidratada (ou a forma anidra misturada na gasolina) e a produção do mesmo é baseada na utilização da cana-de-açúcar como matéria prima. Já nos Estados Unidos, a principal matéria prima é o milho e, assim como na Europa, esse biocombustível é utilizado em misturas (*blends*) com a gasolina.

Figura 3 – Número de publicações encontradas para as palavras “ethanol” e “fuel” na base de dados Scifinder (22/11/14).



(Fonte: Elaborado pelo autor)

A razão entre bioetanol e gasolina em tais misturas pode variar de região para região e também com as diferentes épocas do ano. A nomenclatura depende dessa razão, por exemplo, uma mistura com 10% de etanol e 90% de gasolina é comercialmente conhecida como E10. O comercialmente conhecido E85, o qual possuiria 85% de etanol e 15% de gasolina, na verdade pode variar a concentração de etanol de 51% a 83% (v/v) (nos EUA), de acordo com a ASTM D5798-14 ⁸ e de 50% a 85% (v/v) (na Europa), de acordo com a CEN/TS 15293:2011 ⁹. Essas normas internacionais estabelecem os parâmetros de qualidade do E85.

Tabela 1 – Principais parâmetros de qualidade para o E85 estabelecidos pelas normas internacionais ^{8,9}

Parâmetro	Europa CEN/TS 15293:2011	EUA ATSM D5798-14
	Classes:	Classes:
Teor de Etanol, pressão de vapor	a: 70-85%, 35,0-60,0 kPa b: 70-85%, 50,0-80,0 kPa c: 60-85%, 55,0-80,0 kPa d: 50-85%, mín. 60,0 kPa	1: 51-83%, 38-59 kPa 2: 51-83%, 48-65 kPa 3: 51-83%, 66-83 kPa 4: 51-83%, 66-103 kPa
Aparência	-	Límpido e isento de precipitados e materiais em suspensão (> 21 °C).
Teor de cobre	0,10 mg/kg, máx.	0,07 mg/L, máx.
Densidade	760,0 – 800,0 kg/m ³ a 15 °C	-
Éteres (5 ou mais átomos de Carbono)	11,0% (v/v) máx.	-
Alcoóis superiores (C3-C8)	6,0% (v/v), máx. (C3-C5)	2% (v/v), máx.
Cloreto inorgânico	1,2 mg/kg máx.	1 mg/kg máx.
Teor de metanol	1,0% (v/v) máx.	0,5% (v/v) máx.
Estabilidade oxidativa	360 minutos, mín.	240 minutos, mín. (aplicável somente para os hidrocarbonetos da mistura)
pH _e	-	6,5 – 9,0
Enxofre	10,0 mg/kg, máx.	80 mg/kg, máx.
Teor de água	0,400% (m/m), máx.	1,0% (m/m), máx.
Teor de fósforo	0,15 mg/L, máx.	0,0013 g/L
Acidez total (expresso como ácido acético)	0,005% (m/m)	0,005% (m/m)
Goma	Lavada: 5 mg/100 mL, máx.	Lavada: 5 mg/100 mL, máx. Não lavada: 20 mg/100 mL, máx.
Chumbo	-	0,013 g/L
Condutividade elétrica	1,5 µS/cm	-
Teor de sulfato	4,0 mg/kg, máx.	-

Os limites estabelecidos para os principais parâmetros desse combustível são mostrados na **Tabela 1**. A limitação dos valores possui diferentes objetivos, seja para garantir o bom funcionamento dos sistemas do motor, ou seja, para tentar evitar a emissão de gases além do esperado, entre outros. A seguir, são mostrados os possíveis problemas causados pela não conformidade dos parâmetros.

2.2.1. Etanol

O teor de etanol influencia diretamente na capacidade do sistema eletrônico dos veículos *flex-fuel* (FFV, do inglês *Flex-Fuel Vehicles*) a estabelecer a taxa ótima da mistura ar/combustível. Esse parâmetro também interfere nas propriedades físico-químicas do combustível, como a lubricidade, a tolerância à água e a pressão de vapor⁸.

2.2.2. Metanol

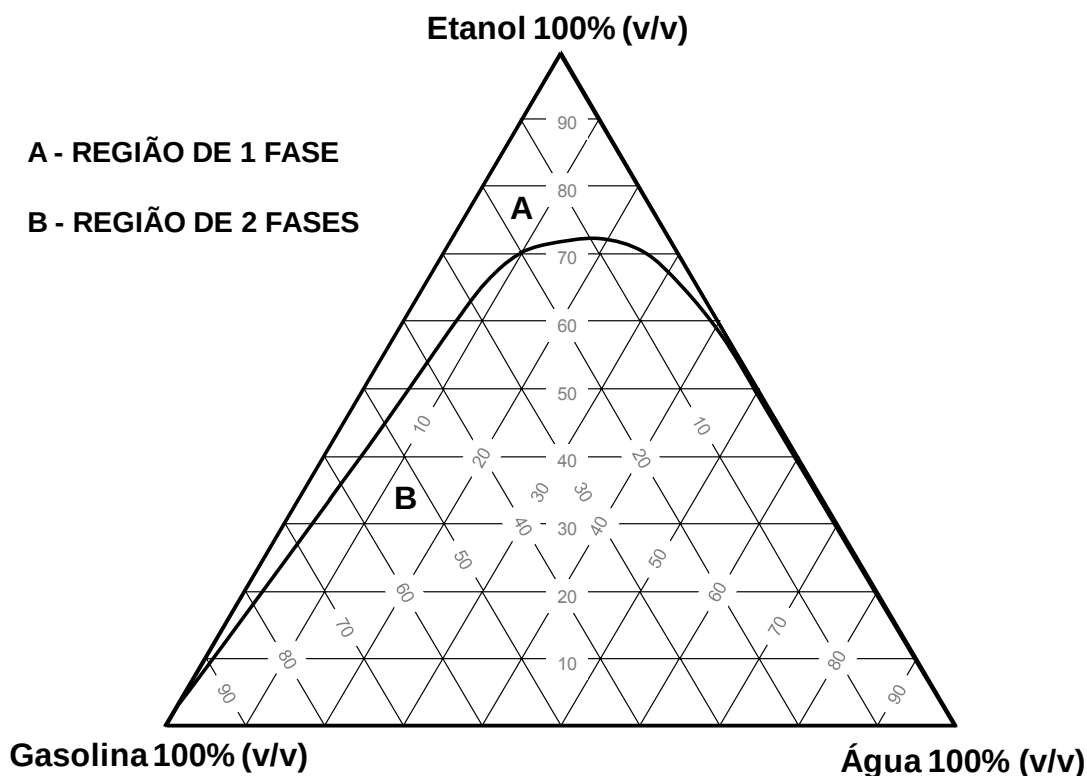
O metanol é tóxico para o ser humano e pode causar problemas de saúde como dores de cabeça, náuseas, vômitos, cegueira e até a morte¹⁰. Além disso, ele pode causar desgaste nos motores e corrosão no sistema de injeção⁸.

2.2.3. Água

Se o teor de água no combustível for maior que os limites estabelecidos, pode ocorrer a descalibração dos sensores dos FFV, além de reduzir o teor energético do combustível, levando a perda de potência e gasto excessivo de combustível. Em altas concentrações, pode ocorrer a separação dos hidrocarbonetos da fase alcoólica, principalmente em regiões de clima frio, acarretando em problemas de partida e de dirigibilidade⁸.

Além disso, a água pode promover a separação de duas fases no tanque: uma fase aquosa de etanol, e uma fase orgânica de gasolina e etanol. Para evitar que ocorra essa separação a mistura deve respeitar os limites da região de uma fase¹¹, como pode ser observado no diagrama da **Figura 4**.

Figura 4 – Diagrama de fases para o sistema etanol, água e gasolina, em volume a 24 °C.



(Fonte: Adaptado de ¹²)

2.2.4. Hidrocarbonetos

O papel dos hidrocarbonetos (gasolina) nas misturas de E85 é elevar a pressão de vapor do combustível. A pressão de vapor é importante para a partida a frio e para a dirigibilidade, principalmente em regiões de clima frio ⁸.

2.2.5. Pressão de vapor

Enquanto a baixa pressão de vapor prejudica a dirigibilidade em climas frios, a pressão de vapor acima dos limites especificados para as diferentes classes podem levar a emissões por evaporação ⁸.

2.2.6. Acidez e pH

Os ácidos, mesmo em concentrações muito baixas, podem causar corrosão a uma vasta gama de metais e ligas metálicas. Como a maior parte das peças do motor são de base metálica, faz-se necessário manter esse parâmetro dentro dos limites especificados para evitar desgaste prematuro dessas peças. Quando o valor do pH dos combustíveis à base de etanol é menor que 6,5, pode ocorrer mau funcionamento nas bombas de combustível devido a formação de filme entre as buchas e o comutador. Podem ocorrer também, problemas nos bicos injetores e nos cilindros em função do desgaste por corrosão. Por outro lado, se o valor de pH for acima de 9,0, podem ocorrer problemas nas partes plásticas das bombas de combustível ⁸.

2.2.7. Goma

O ensaio de goma não lavada mede a quantidade de resíduo após a evaporação de uma alíquota do combustível. Para medir o parâmetro goma lavada, é feito o mesmo procedimento, porém lava-se o resíduo com heptano para remover materiais não voláteis como aditivos e óleos. Teores elevados de goma lavada e goma não lavada no combustível podem causar depósitos no sistema de injeção e emperramento das válvulas de admissão ⁸.

2.2.8. Cloreto

O limite para o teor de cloreto especificado em 1 mg/kg visa evitar que ocorra oxidação das peças metálicas que entram em contato com o combustível, pois o cloreto é corrosivo para muitos metais ⁸.

2.2.9. Chumbo e Fósforo

Fósforo e chumbo no combustível podem provocar desativação do catalisador (que ajuda no controle de emissões de formaldeído e acetaldeído) presente no sistema de exaustão. Por isso a especificação para estes elementos é limitada a 0,013 g/L e 0,0013 g/L, respectivamente ⁸.

2.2.10. Aparência

O combustível deve ser límpido e isento de precipitados e materiais em suspensão. Turbidez ou separação de fase normalmente indicam contaminação⁸.

2.2.11. Cobre

O cobre pode catalisar a reação de oxidação em baixas temperaturas dos hidrocarbonetos presentes no combustível. Portanto, o teor deste metal deve estar abaixo de 0,012 mg/kg. Concentrações acima dessa especificação podem elevar significativamente a taxa de formação de goma⁸.

2.2.12. Enxofre

A limitação no teor de enxofre no E85 visa proteger o motor contra desgaste, deterioração do óleo do motor, corrosão das peças do sistema de exaustão, além da desativação do catalisador⁸.

2.3. Os efeitos da adição de etanol na gasolina

A diferença da natureza química dessas duas fontes de energia faz com que os motores que são abastecidos com eles tenham comportamentos diferentes. Vários trabalhos relatam que a adição de etanol na gasolina reduz as emissões locais dos gases como CO e HC (como benzeno, por exemplo)¹²⁻²⁰ e elevam a emissão de CO₂^{13,14,16,21}, devido à introdução de oxigênio (presente na molécula de etanol). Apesar de a emissão local de CO₂ ser maior, no balanço global ela é menor, pois o etanol provém de uma fonte renovável⁷. Em contrapartida, a emissão de acetaldeído é proporcional ao teor de etanol presente no combustível^{12,18}, principalmente na fase fria do motor²¹. Já a emissão local de NO_x, depende mais das características de operação do motor do que do combustível utilizado^{12,16,21}.

Em relação ao desempenho, os motores passam a consumir maior volume de combustível para realizar a mesma quantidade de trabalho quando alimentados com

etanol²¹ (devido ao menor valor de densidade de energia do etanol em relação à gasolina), por outro lado, ocorre a elevação dos valores do número de octano e do torque^{14,18}.

2.4. Espectroscopia na região do infravermelho

A região do espectro de radiação infravermelha corresponde à parte do espectro eletromagnético situada entre as regiões do visível e das micro-ondas. A região aproximadamente entre 14.000 e 4.000 cm^{-1} é denominada infravermelho próximo (NIR - do inglês *Near Infrared*), a situada entre 4000 e 400 cm^{-1} , denomina-se infravermelho médio (MIR - do inglês *Middle Infrared*) e a região entre 700 e 200 cm^{-1} é denominada infravermelho distante (FIR - do inglês *Far Infrared*)²².

A espectroscopia na região do infravermelho (IV) mede as energias de transição entre estados vibracionais que ocorrem quando uma molécula absorve energia na região do infravermelho. Para que ocorra absorção, a energia do fóton incidente deve ser idêntica à diferença de energia entre dois níveis de energias vibracionais. Além disso, para uma vibração ser ativa é necessário também que o campo elétrico oscilante da radiação incidente possa interagir com a molécula e isso só ocorre se o deslocamento dos átomos na vibração produzir uma mudança no momento de dipolo da molécula.

Os primeiros instrumentos de infravermelho eram do tipo dispersivo, os quais separavam as frequências individuais emitidas de uma fonte de infravermelho pelo uso de prismas ou redes de difração, porém esses instrumentos apresentavam problemas incluindo baixa velocidade de varredura, baixa sensibilidade, entre outros. O método mais moderno em espectroscopia no infravermelho é o Infravermelho por Transformada de Fourier (FTIR – do inglês *Fourier Transform Infrared*). Este método foi desenvolvido para superar as limitações encontradas nos instrumentos por energia dispersiva, principalmente a lentidão no processo de varredura. A solução encontrada foi o emprego de um dispositivo óptico, denominado interferômetro, que produz um único tipo de sinal contendo todas as frequências codificadas dentro dele. O sinal pode ser medido muito rapidamente, geralmente na ordem de um segundo e o tempo base de análise, por amostra, é reduzido para alguns poucos segundos. O sinal resultante é chamado de interferograma e contém informações sobre cada frequência no

infravermelho. A técnica matemática conhecida como Transformada de Fourier “decodifica” as frequências individuais no computador e transforma a informação espectral desejada para análise ²³.

Os espectrômetros FTIR apresentam várias vantagens em relação aos dispersivos ²⁴, tais como: fonte de radiação extremamente estável; rapidez na aquisição dos dados (aproximadamente um segundo por varredura); sensibilidade instrumental; precisão e exatidão das posições das bandas; resolução das bandas ($< 0,1 \text{ cm}^{-1}$). Tais características fazem com que essa técnica ganhe espaço nos laboratórios analíticos e em controle de qualidade de processos industriais. Isto ocorre devido ao baixo custo relativo do ensaio, alta velocidade de aquisição de dados, facilidade de operação, não necessidade de pré-tratamento da amostra (dispensando gasto com reagentes), pequena quantidade de amostra utilizada e alta seletividade, podendo ser usada para determinações qualitativas e quantitativas.

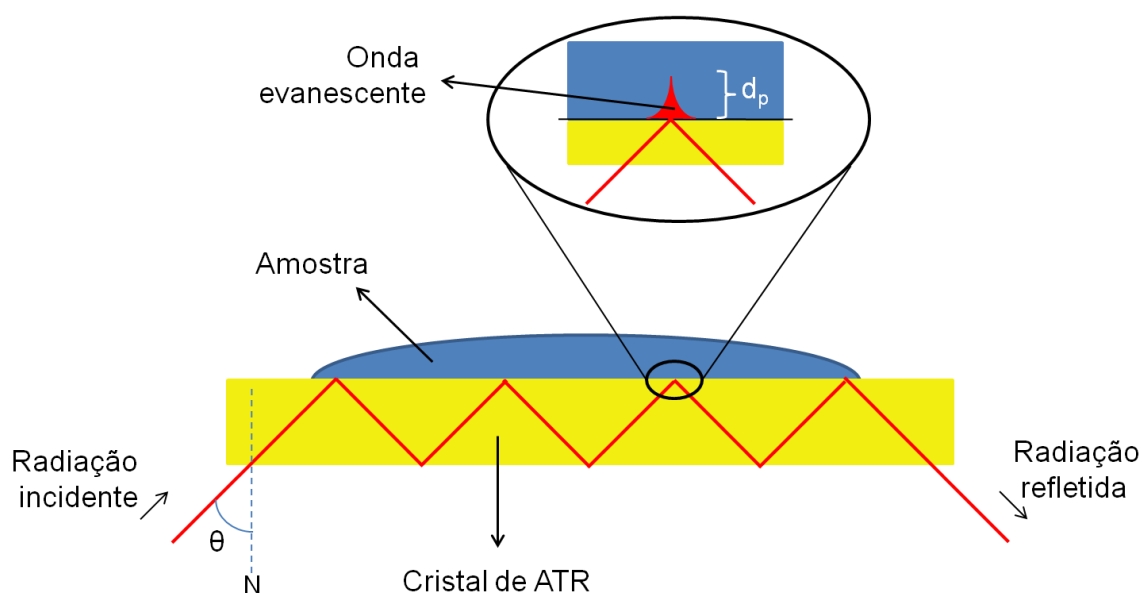
2.4.1. Espectroscopia na região do MIR

Na região do MIR, os diferentes grupos funcionais e suas ligações têm frequências e intensidades de absorção diferentes no infravermelho ^{22,25}, sendo que estas diferenças fazem com que seja possível analisar, por exemplo, etanol (oxigenado) em gasolina (hidrocarbonetos) ²⁶. Além disso, a quantidade de radiação absorvida é diretamente proporcional à concentração da substância e, normalmente, obedece à lei de Lambert-Beer, possibilitando seu uso como método quantitativo, como é o caso das EN 14078:2014 ²⁷ e NBR 15568 ²⁸ para a determinação do teor de biodiesel no diesel.

Nos últimos anos, uma técnica muito rápida para a aquisição dos espectros no infravermelho médio tornou-se disponível para os espectrômetros FTIR possibilitando a análise de um grande número de amostras em pouco tempo. A técnica é conhecida como refletância total atenuada (ATR - do inglês *Attenuated Total Reflectance*). A ATR caracteriza-se pelas reflexões da radiação infravermelha que ocorrem no interior de cristais com alto índice de refração (seleneto de zinco, germânio ou diamante), interagindo apenas com a região da interface da amostra e o cristal (**Figura 5**). No processo de reflexão, o feixe penetra alguns micrômetros no meio com menor índice de refração (amostra) antes que a reflexão ocorra. A radiação penetrante é chamada onda

evanescente. Se o meio menos denso absorve essa onda evanescente, ocorre atenuação do feixe nos comprimentos de onda das bandas de absorção ²⁹.

Figura 5 – Representação gráfica do funcionamento do ATR. N é a normal; θ é o ângulo de incidência e d_p (do inglês *depth of penetration*) é a profundidade de penetração.



(Fonte: Elaborado pelo autor)

O caminho óptico e a profundidade de penetração dependem do número de reflexões, do índice de refração (I.R.) do cristal e do ângulo de incidência como demonstrado pelas equações abaixo ³⁰:

Profundidade de penetração:

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

onde: λ é comprimento de onda da luz incidente, θ é o ângulo de incidência, n_1 e n_2 são os índices de refração do cristal e da amostra, respectivamente.

Caminho óptico efetivo calculado:

$$EPL = Nd_e \quad (2)$$

onde: N é o número de reflexões e d_e é a penetração efetiva, dada por:

$$d_e = \frac{d_{e\perp} + d_{e\parallel}}{2} \quad (3)$$

$$d_{e\perp} = \frac{n_1^2 n_2 \cos \theta}{(n_1^2 - n_2^2)} \cdot \frac{\lambda}{\pi \sqrt{n_1^2 \sin^2 \theta - n_2^2}} \quad (4)$$

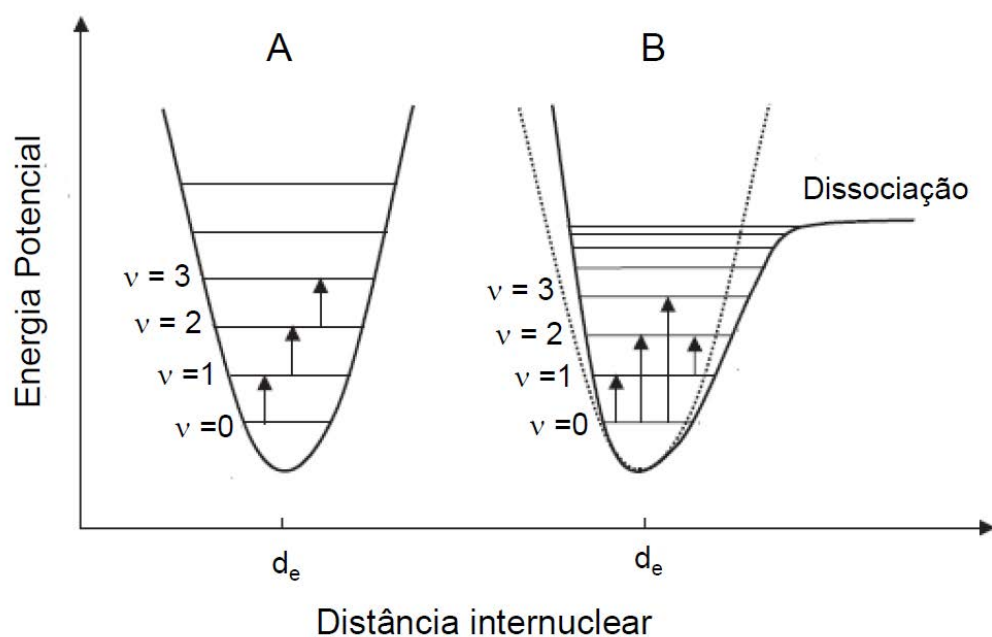
$$d_{e\parallel} = \frac{n_1^2 n_2 \cos \theta}{(n_1^2 - n_2^2)} \cdot \frac{2n_1^2 \sin^2 \theta - n_2^2}{(n_1^2 - n_2^2) \sin^2 \theta - n_2^2} \cdot \frac{\lambda}{\pi \sqrt{n_1^2 \sin^2 \theta - n_2^2}} \quad (5)$$

2.4.2. Espectroscopia na região do NIR

Na região do NIR, além das bandas de combinação, são detectadas as vibrações de primeiro, segundo e terceiro sobretons de bandas fundamentais observadas na região do MIR. Sobretons são as combinações das transições fundamentais ($v = 0$ para $v = 2, 3, 4...$). As principais bandas de absorção na região do NIR são relacionadas aos sobretons e bandas de combinação das transições vibracionais fundamentais dos grupos C-H, N-H, O-H e S-H ^{31,32}. Isso ocorre, pois essas ligações apresentam grande momento de dipolo e alta anarmonicidade. O fenômeno de anarmonicidade elétrica e mecânica ajuda a explicar a observação das transições de sobretons, já que se considerarmos somente a regra de seleção para o MIR ($\Delta v = \pm 1$) e a molécula como um oscilador harmônico, tais transições seriam proibidas.

A vibração das moléculas citadas acima não pode ser descrita por um oscilador com movimento harmônico devido à diferença entre as massas dos átomos. A anarmonicidade mecânica de osciladores pode ser demonstrada pelo diagrama de energia potencial (**Figura 6**). Nos osciladores harmônicos o diagrama de energia é uma parábola perfeita, enquanto que o diagrama referente a um oscilador anarmônico apresenta uma parábola distorcida com espaçamento entre os níveis vibracionais desiguais e contendo energia referente à dissociação da molécula. Neste modelo, transições de $v = 0$ para $v = 2, 3, 4...$ são permitidas e são as de fato observadas no espectro NIR ^{31,32}.

Figura 6 – Diagrama de energia indicando as transições possíveis para os osciladores harmônicos (A) e anarmônico (B).



(Fonte: ³¹)

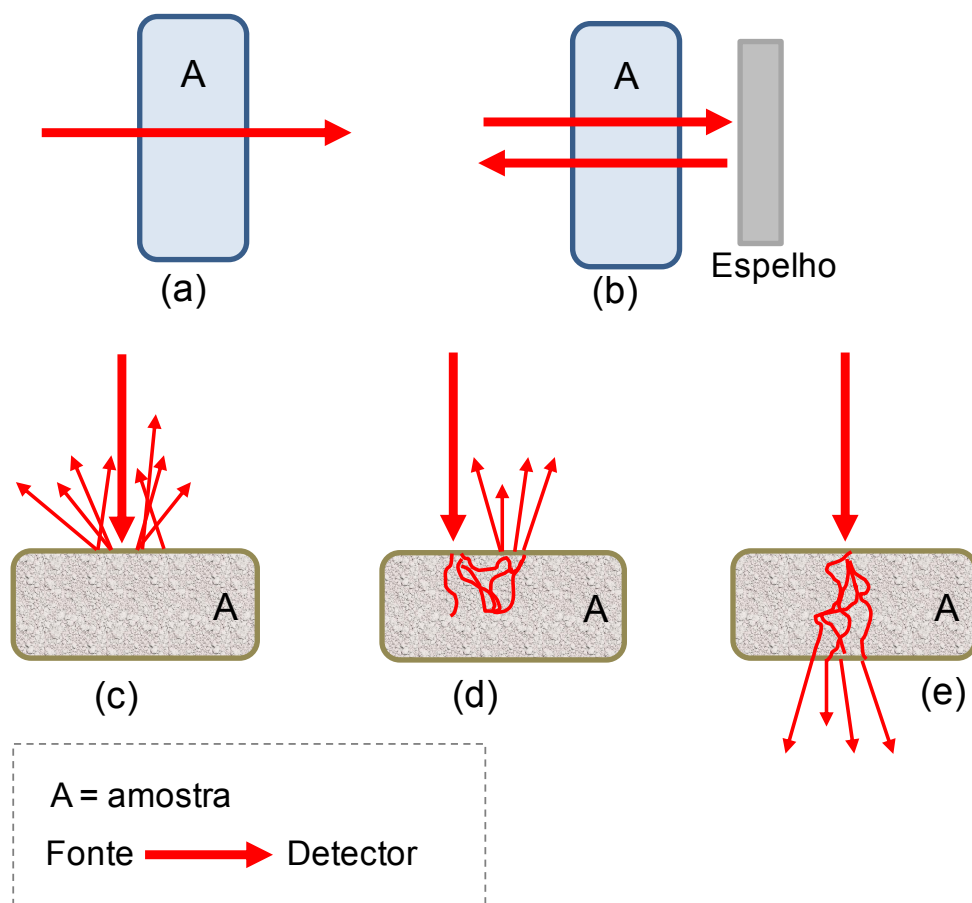
Na espectroscopia NIR, as medidas podem ser realizadas em diferentes modos de medição. A **Figura 7** ilustra os mais comuns ³²:

- ✓ Transmitância: o feixe passa através da amostra em uma cubeta e segue para o detector;
- ✓ Transfletância: o feixe passa pela amostra, é refletido em um espelho e passa novamente pela amostra antes de chegar ao detector;
- ✓ Refletância difusa: o feixe chega à amostra e é refletido de forma difusa, a radiação refletida segue para o detector levando as informações espectrais dessa amostra;
- ✓ Interactância: nesse modo a probabilidade de o feixe incidente interagir com a amostra é maior que na refletância difusa e, conseqüentemente, o

feixe que emerge da amostra contém mais informação sobre os constituintes e reflete melhor a composição real da mesma;

- ✓ Transmitância em amostras sólidas: o caminho óptico efetivo é maior que na transmitância em amostras líquidas devido ao espalhamento interno. Esse modo gera informação que é melhor correlacionada com o conteúdo do interior da amostra do que com o da superfície.

Figura 7: Modos de medição empregados na espectroscopia na região do NIR: (a) transmitância; (b) transfletância; (c) refletância difusa; (d) interactância e (e) transmitância através de amostras sólidas.



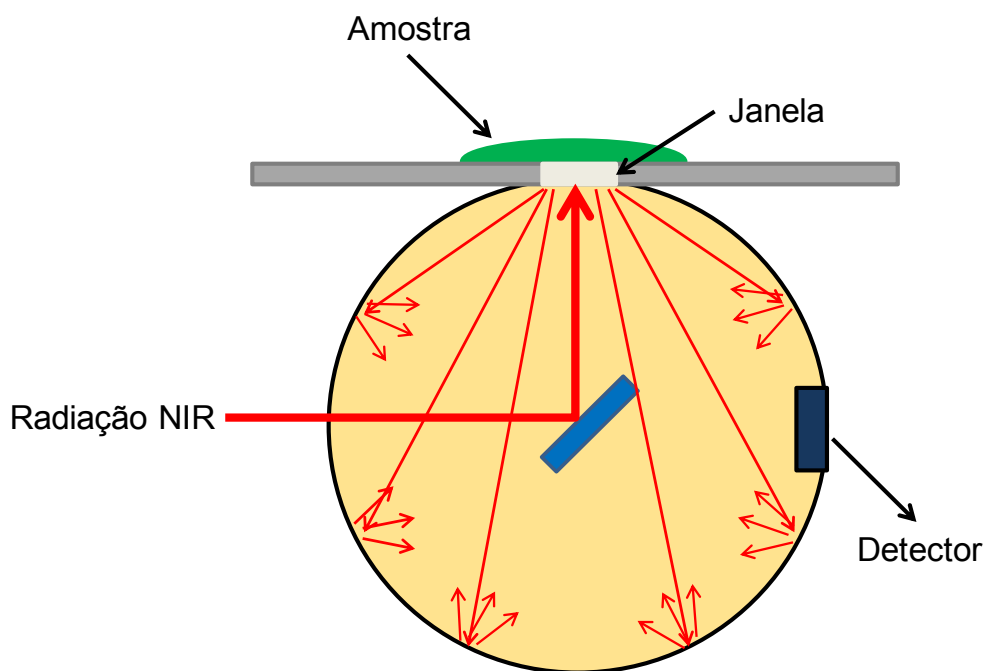
(Fonte: adaptado de ³²)

Um dos acessórios para medidas utilizando o modo de refletância difusa é a esfera de integração. Nesse acessório o feixe de radiação entra na esfera e passa por uma janela de safira até incidir na amostra, interagindo com a mesma. Então, o feixe é refletido de forma difusa e a radiação refletida volta à esfera e é direcionada ao detector após múltiplas reflexões nas paredes da esfera (**Figura 8**). A radiação infravermelha

dispersada é coletada com uma eficiência maior que 95%. Esta técnica apresenta a grande vantagem da facilidade no manuseio de amostras, possibilitando analisar amostras diretamente em frascos pela subtração da contribuição espectral do recipiente

²⁴.

Figura 8 – Representação esquemática do funcionamento de uma esfera de integração.



(Fonte: adaptado de ²⁴)

A espectroscopia na região do MIR e do NIR com transformada de Fourier destaca-se atualmente como uma das mais importantes na análise de *fingerprintings* de misturas complexas. Tal técnica produz uma grande quantidade de informação, muitas vezes complexa e variada, tornando necessário o uso de ferramentas para tratá-las e extrair as informações mais relevantes. Nesse contexto, a quimiometria torna-se indispensável para a análise de dados químicos de natureza multivariada ³³⁻³⁶.

2.5. Quimiometria

No campo da quimiometria, os métodos de reconhecimento de padrões e regressão multivariada podem ser aplicados com diferentes finalidades, entre elas a

análise exploratória de dados, a classificação de amostras e previsão de parâmetros ^{34,37}.

As associações de técnicas espectroscópicas, com ferramentas quimiométricas, são relativamente simples e permitem análises rápidas, precisas, com exatidão comparáveis aos métodos convencionais e fazendo uso de pequenas quantidades de amostra, razões pelas quais têm se mostrado adequadas para obter diversos tipos de informações em matrizes complexas ³⁸⁻⁴⁰.

Tais associações encontram-se amplamente difundidas e foram aplicadas aos mais variados tipos de matrizes, para prever os mais diversos parâmetros físico-químicos, tais como: determinar, qualitativamente, os tipos e as condições de óleos lubrificantes ⁴¹, quantificar contaminantes em óleos lubrificantes ⁴², determinar o teor de biodiesel em diesel ⁴³, além da detecção de adulteração em etanol hidratado combustível ¹⁰.

Há também um estudo mostrando que é possível realizar a quantificação simultânea do teor de etanol e metanol em gasolina (com concentrações dos álcoois entre 0% e 30%, em volume) utilizando espectroscopia na região do infravermelho próximo e regressão multivariada ⁴³. Além disso, a distinção entre amostras de gasolina com etanol e metanol também foi possível empregando NIR e métodos de classificação ⁴⁴. Porém, em nenhum dos estudos citados, observou-se a abordagem da incerteza de medição e dos custos dos ensaios. Ademais, até onde vão nossos conhecimentos, não há estudos para a previsão de parâmetros de qualidade em amostras de E85 utilizando espectroscopia na região do infravermelho e regressão multivariada.

2.5.1. Pré-processamento dos dados

Antes de se iniciar qualquer tipo de análise multivariada é necessária uma análise para verificar a necessidade da realização de uma manipulação matemática prévia (pré-processamento) do conjunto de dados para adequação, ou às vezes até mesmo remoção de possíveis fontes de variação. Por exemplo, muitas vezes as variáveis possuem diferentes dimensões e amplitudes e é necessário um tratamento prévio para expressar cada observação em dimensões e amplitudes equivalentes, sem perda de informações ^{35,46,47}. Existem vários tipos de pré-processamentos que

produzem diferentes transformações nos dados. Entre eles, destacam-se, o centrado na média, a normalização, os Filtros de Savitzky-Golay, a correção do espalhamento multiplicativo (MSC – do inglês *Multiplicative Scatter Correction*) e a variação de padrão normal.

2.5.1.1. Centrado na média

Centrar os dados na média significa subtrair cada valor das variáveis das amostras da média de cada uma das variáveis, para todo o conjunto X . Este processo desloca o eixo do sistema de coordenadas para o centroide dos dados. Do ponto de vista estatístico, centrar na média objetiva prevenir que dados situados distantes do centroide tenham maior influência que os dados que estão próximos a ele ⁴⁸.

2.5.1.2. Correção multiplicativa de sinal

A Correção Multiplicativa de Sinal é usada em aplicações onde existem variações multiplicativas entre os perfis de resposta das amostras. O pré-processamento MSC considera que qualquer espectro da amostra pode ser estimado como um múltiplo de um espectro de referência mais um valor aditivo, sendo muitas vezes o espectro de referência considerado como o espectro médio do conjunto de amostras. A maior parte das variações (multiplicativas) entre os espectros são removidas ⁴⁹.

2.5.1.3. Derivadas

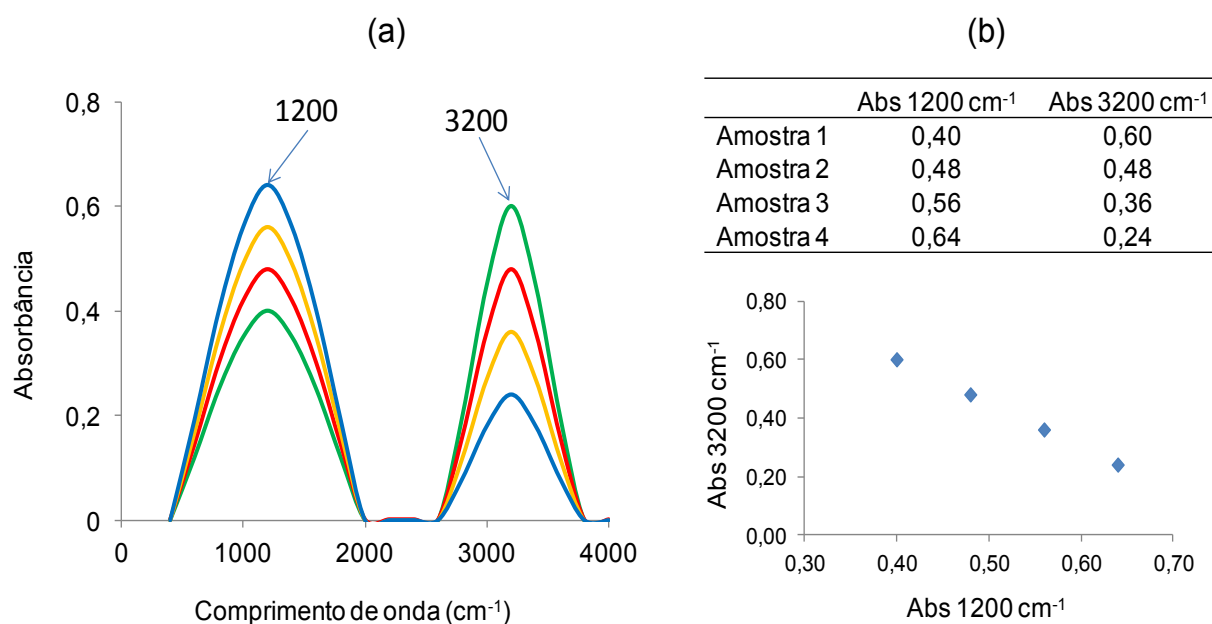
Diferentemente de aplicar MSC ou centrar na média, a derivação dos dados ocorre dentro do vetor de dados de cada amostra e não entre os diferentes vetores do conjunto de dados das amostras. A derivação tem a finalidade de corrigir variações de linha base sem que haja perda de informações quantitativas dos dados originais, pois a derivação de uma função constante é igual à zero. Este método é muito útil em casos onde a linha de base é difícil de ser identificada. Na aplicação da derivada, a escolha do número de pontos do espectro por janela influenciará nos resultados finais do

modelo. Se a escolha não for adequada, poderá haver distorção dos dados e perda de informação quantitativa ⁵⁰.

2.5.2. Análise de componentes principais

Quando queremos observar a relação entre as características analíticas de uma amostra, provavelmente iremos utilizar uma tabela ou um gráfico, como no caso ilustrado na **Figura 9 (a) e (b)**. É possível identificar claramente a relação inversa entre as absorvâncias nos dois comprimentos de onda de quatro amostras fictícias.

Figura 9 – Exemplo ilustrativo de espectros de quatro amostras fictícias, a tabela para comparação dos dois comprimentos de onda (1200 e 3200 cm^{-1}) e o gráfico mostrando a relação inversa entre as absorvâncias.



(Fonte: Elaborado pelo autor)

Vamos supor agora que queremos observar o comportamento de todos os comprimentos de onda dos espectros obtidos de dezenas de amostras diferentes. Tais espectros podem conter centenas de comprimentos de onda, gerando uma tabela muito grande para a observação visual e impossibilitando a construção de um gráfico devido à limitação espacial de três dimensões. Isso pode tornar difícil a identificação de padrões e relações existentes. Portanto, a redução da dimensionalidade dos dados é

necessária para facilitar a sua interpretação e manipulação. Nesse sentido, a Análise de Componentes Principais (PCA - do inglês *Principal Component Analysis*) tem papel fundamental ⁵¹.

A PCA pode ser considerada como a “mãe de todos os métodos em análise multivariada de dados”, pois o seu fundamento é utilizado como base tanto em métodos de reconhecimento de padrão, como em métodos de calibração multivariada ⁵² (como no caso dos mínimos quadrados parciais). Em síntese, a PCA é um método que permite resumir as informações contidas em conjuntos de dados de alta dimensionalidade em algumas poucas dimensões, permitindo encontrar visualmente tendências e similaridades entre as amostras e identificar quais são as variáveis responsáveis por essas características. Sendo assim, a PCA consiste na construção de um novo sistema de eixos ortogonais, chamados de componentes principais, as quais são combinações lineares não correlacionadas das variáveis originais ³³.

Se a matriz **X** é um banco de dados com *I* linhas e *J* colunas, a PCA decompõe a matriz **X** da seguinte forma ⁵³:

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P}^T + \mathbf{E} \quad (6)$$

onde:

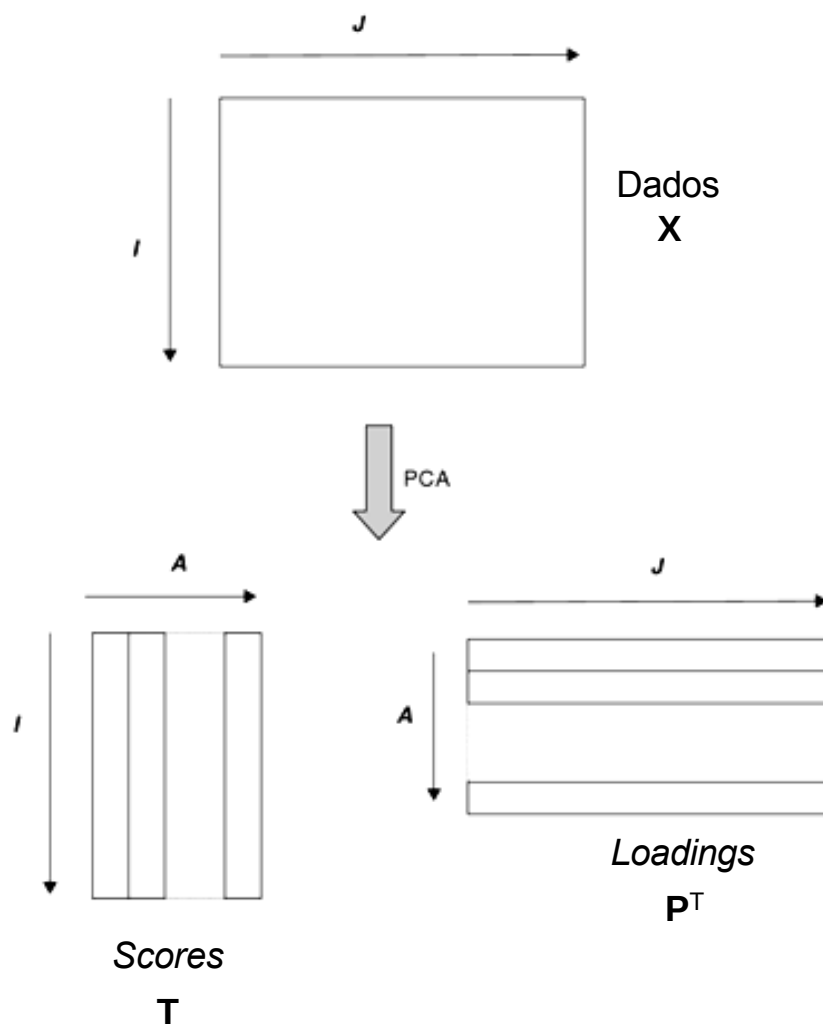
T são os *scores* que possuem tantas linhas quanto a matriz original;

P são os *loadings* que possuem tantas colunas quanto a matriz original. O número de colunas na matriz **T** é igual ao número de linhas na matriz **P**;

E é a matriz de erros.

A matriz de *scores* consiste de uma série de vetores em colunas (**t_a**) e a matriz de *loadings*, uma série de vetores em linhas (**p_a**) (**Figura 10**), cada vetor para cada componente principal. O número máximo de componentes principais (1, 2, 3, ..., A) é o menor valor entre *I* e *J*. Os pares *t* e *p* são ordenados de forma decrescente pela quantidade de variância capturada. Ou seja, a primeira componente principal (representada pelo primeiro par de *scores* e *loadings*) contém a maior quantidade de variação explicada pelo modelo PCA, a segunda contém a segunda maior variação e assim por diante ⁵³.

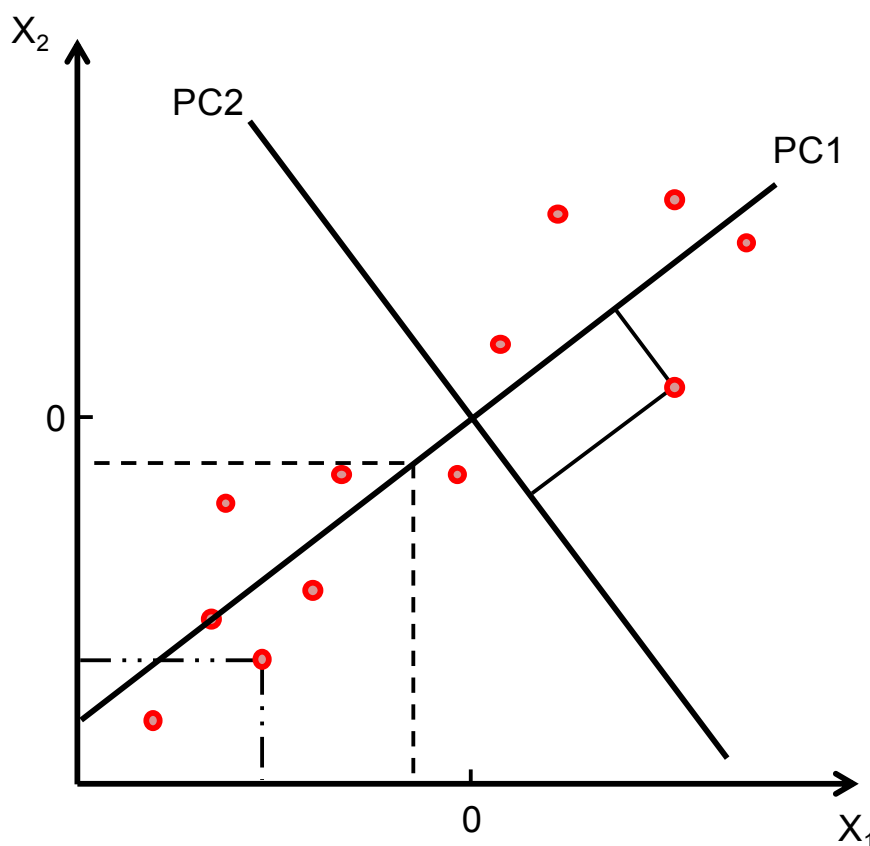
Figura 10: Representação gráfica da decomposição da matriz de dados X pela PCA.



(Fonte: adaptado de ⁵³)

Os *scores* representam as coordenadas de cada uma das amostras em cada componente principal, já os *loadings* fornecem informações sobre a contribuição das variáveis originais para a formação das componentes principais ³¹ (**Figura 11**).

Figura 11 – Gráfico demonstrando as componentes principais PC1 e PC2, bem como a projeção das amostras sobre elas. A linha contínua representa os *scores* de uma amostra, a linha tracejada representa os *loadings* e a linha traço dois pontos representa as coordenadas de uma amostra em relação aos eixos originais.



(Fonte: Adaptado de ³⁵)

2.5.3. PLS

Quando queremos relacionar um sinal analítico com uma propriedade de interesse, por exemplo, a absorvância em um dado comprimento de onda com a concentração de um analito, provavelmente iremos utilizar algum tipo de regressão. É o caso de uma curva analítica na qual é aplicada uma regressão linear. Porém, quando é necessário relacionar mais de um comprimento de onda (uma parte do espectro na região do infravermelho médio, por exemplo) com uma ou mais propriedades de interesse a aplicação de uma regressão linear não é mais possível. Nesses casos a

regressão em mínimos quadrados parciais (PLS – do inglês *Partial Least Squares*) pode ser aplicada.

Existem diversos algoritmos para a realização da PLS, os principais são propostos por Wold and Martens ⁵⁴. Diferentemente da PCA, que modela exclusivamente os dados de uma dada matrix **X**, no modelo PLS **X** e **y** são relacionados da seguinte forma ^{53,55}:

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P}^T + \mathbf{E} \quad (7)$$

$$\mathbf{y} = \mathbf{T} \cdot \mathbf{q}^T + \mathbf{f} \quad (8)$$

onde:

E e **f** são os resíduos de **X** e **y**, respectivamente;

T são os *scores*;

P são os *loadings* de **X**;

q são os *loadings* de **y**.

Essas matrizes são representadas na **Figura 12**. O produto de **T** e **P^T** se aproxima dos dados originais (**X**) e o produto de **T** e **q** dos valores das propriedades de interesse.

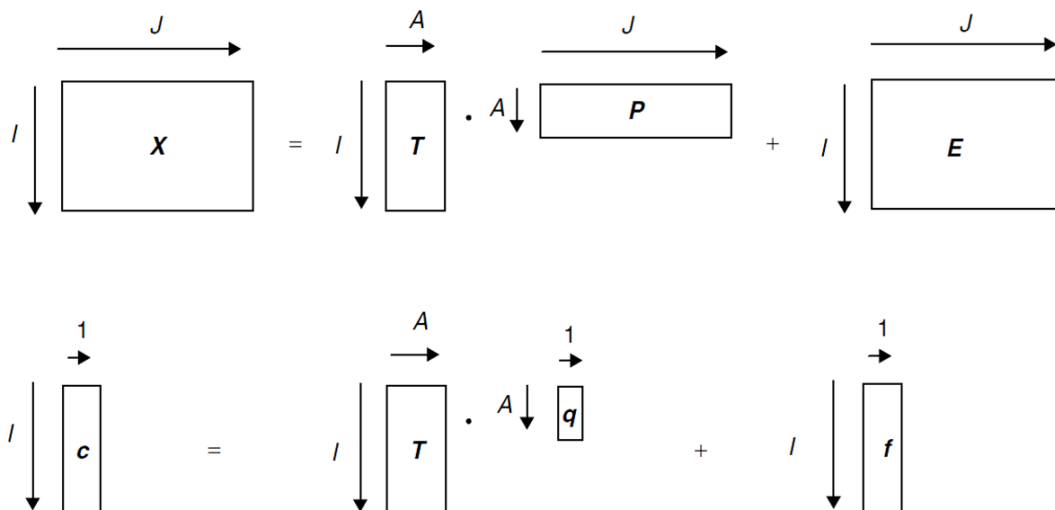
As matrizes **T** e **P** para PLS são diferentes das obtidas na PCA, pois o algoritmo faz mudanças nos valores dos *scores* através de uma série de etapas que são repetidas para cada componente principal, objetivando atingir a maior covariância entre **X** e **y**. Como no cálculo de cada componente principal os valores de concentração dos analitos também são levados em consideração, as componentes principais perdem a total ortogonalidade e passam a ser chamadas de variáveis latentes.

Através do algoritmo PLS também é possível determinar o vetor de regressão (**b**), então se torna possível prever o valor de uma propriedade de interesse (**ŷ**) para novas amostras como se segue ⁵⁵:

$$\hat{\mathbf{y}} = \mathbf{T} \mathbf{q}^T = \mathbf{X} \mathbf{W} (\mathbf{P}^T \mathbf{W})^{-1} \mathbf{q}^T = \mathbf{X} \mathbf{b} \quad (9)$$

onde $\mathbf{W}$ é uma matriz de pesos, determinada pelo algoritmo PLS.

Figura 12 – Representação gráfica da decomposição das matrizes realizada no modelo PLS.



(Fonte: ⁵³)

Nesse processo, é necessário achar o melhor número de variáveis latentes, que normalmente é feito usando um procedimento chamado validação cruzada, no qual o erro mínimo de calibração é determinado. Nesse procedimento, é definida a quantidade de amostras que serão retiradas em cada passo do processo (geralmente uma por vez - *leave-one-out*). Após cada retirada sucessiva destas amostras, as demais serão utilizadas para a construção de um novo modelo de regressão, que será usado para estimar o valor das amostras retiradas. Depois que todas as amostras forem retiradas e a propriedade de interesse estimada, calcula-se a média da soma dos quadrados dos erros (o erro é a diferença entre o valor de referência e o valor previsto pelo modelo) (RMSECV - do inglês *Root Mean Square Error of Cross-Validation*).

Após o desenvolvimento do modelo de calibração, é necessária a avaliação de seu desempenho na predição dos valores de y , através da validação. O objetivo da validação é determinar a reprodutibilidade da calibração e a tendência do modelo comparado com o método de referência ²⁴. Nesta etapa, um conjunto de amostras, diferentes das utilizadas na calibração (conjunto de validação), com valores conhecidos é aplicado no modelo, então, os valores de y previstos são comparados aos valores de referência para avaliação de seu desempenho.

2.6. Validação

Após o desenvolvimento de um novo método analítico, é necessário obter evidências de que ele é adequado para o fim a que se destina nas condições em que será aplicado. Segundo a ABNT NBR ISO IEC 17025:2005 ⁵⁶ "validação é a confirmação por exame e fornecimento de evidência objetiva de que os requisitos específicos para um determinado uso pretendido são atendidos". Existem guias e trabalhos que mostram como realizar o processo de validação tanto para métodos que empregam calibração univariada ⁵⁷⁻⁶¹ (calibração de ordem zero), como para os baseados em calibração multivariada que utilizam um vetor de dados por amostra ⁶²⁻⁶⁵ (primeira ordem).

A ABNT NBR ISO IEC 17025:2005 propõe que a validação seja realizada aplicando uma das seguintes técnicas ou a combinação delas:

- ✓ Calibração com o uso de padrões de referência ou materiais de referência;
- ✓ Comparações com resultados obtidos por outros métodos;
- ✓ Comparações interlaboratoriais;
- ✓ Avaliação sistemática dos fatores que influenciam o resultado;
- ✓ Avaliação da incerteza dos resultados com base no conhecimento científico dos princípios teóricos do método e na experiência prática.

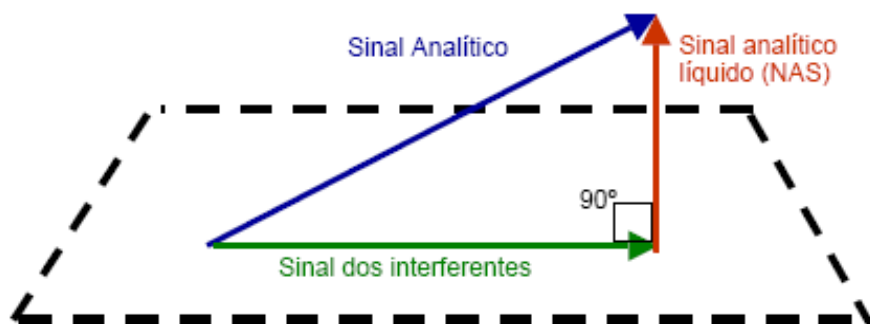
Outra forma de se validar um método é através da determinação das figuras de mérito ⁶¹. O cálculo de algumas figuras de mérito em modelos de calibração multivariada como, por exemplo, exatidão, precisão, ajuste e bias não apresentam maiores dificuldades e são estimados de maneira bastante similar aos métodos de calibração univariada. Entretanto, essa similaridade não é observada para a estimativa de parâmetros como linearidade, sensibilidade, razão sinal/ruído, seletividade e intervalos de abrangência ⁶². Apenas cálculos para modelos de primeira ordem foram descritos nesta revisão.

2.6.1. Sinal analítico líquido – NAS

O sinal analítico líquido (NAS - do inglês *Net Analyte Signal*) é a fração do sinal analítico que é ortogonal aos sinais dos demais compostos presentes na amostra e é de extrema importância na determinação de algumas figuras de mérito para calibrações multivariadas. Os cálculos de seletividade e razão sinal/ruído requerem o cálculo do NAS do analito de interesse. Para métodos de calibração de primeira ordem (como o PLS) obtém-se um vetor de NAS que pode ser representado em sua forma escalar sem perda de informação ⁶².

O método para o cálculo do NAS para modelos multivariados de calibração inversa foi proposto por Lorber ^{62,66}. O NAS é definido, para uma propriedade de interesse k, como sendo a parte do sinal analítico que é ortogonal às contribuições de possíveis interferentes presentes na amostra. Sua propriedade de ortogonalidade pode ser observada pela representação geométrica da **Figura 13**:

Figura 13 – Representação geométrica da propriedade de ortogonalidade do NAS.



(Fonte: ⁶⁴)

O método proposto por Lorber foi corrigido por Ferre, Brown e Rius para possibilitar o cálculo exato do NAS para modelos de calibração construídos a partir dos métodos de regressão baseados no PLS ⁶⁶. O cálculo do NAS é explicado e aplicado em trabalhos da literatura ^{62,64,66-71}. Resumidamente, o cálculo do NAS se dá por operações e propriedades de matrizes para obtenção de uma matriz que pode ser livre da propriedade de interesse (matriz de interferentes) e a partir desta matriz, calcula-se o vetor NAS, que é ortogonal à matriz de interferentes. Uma vez que o vetor NAS é livre

de interferentes, é possível substituí-lo por uma representação escalar, sem perda de informação ⁶⁰.

Com o valor escalar obtido é possível construir uma nova forma de calibração multivariada, em que o modelo pode ser representado em uma forma pseudounivariada. A representação univariada de um modelo de calibração multivariada possibilita avaliar a porção do sinal que participa eficientemente do modelo ^{62,67-71}.

2.6.2. Exatidão

Os valores da raiz quadrada dos erros médios quadráticos de calibração (RMSEC – do inglês *Root Mean Square Error of Calibration*), calibração cruzada (RMSECV) e previsão (RMSEP - do inglês *Root Mean Square Error of Prediction*) podem ser referidos como estimativas da exatidão em trabalhos com calibração multivariada ⁶².

2.6.3. Precisão

Assim como em calibração univariada, pode ser expressa em diferentes níveis ⁶²:

Repetibilidade: segundo a ASTM E1655 ⁶³, são necessárias no mínimo três amostras em concentrações diferentes cobrindo a faixa útil do modelo de calibração e seis replicatas para cada nível de concentração. Então a precisão é calculada:

$$Repetibilidade = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (\hat{y}_{ij} - \hat{\bar{y}}_i)^2}{n(m-1)}} \quad (10)$$

onde, n representa o número de amostras e m o número de replicatas.

Precisão intermediária: a extensão em que a precisão intermediária deve ser determinada depende das circunstâncias em que o método será aplicado, geralmente é calculada a partir de dados obtidos em dias diferentes ou por operadores diferentes.

Reprodutibilidade: é tida com base em estudos interlaboratoriais.

2.6.4. Sensibilidade

Para modelos de calibração multivariada, como PLS, pode ser determinada como ⁶²:

$$SEN = \frac{1}{\|b_k\|} \quad (11)$$

onde, b_k é o vetor dos coeficientes de regressão estimados pelo modelo PLS.

2.6.5. Sensibilidade Analítica

A sensibilidade analítica (γ), não é amplamente abordada em normas ou guias de validação. No entanto, esse parâmetro apresenta a sensibilidade do método em termos da unidade de concentração que é utilizada, sendo definida como a razão entre a sensibilidade e o desvio padrão do sinal de referência (δx) ⁶².

$$\gamma = \frac{SEN}{\|\delta x\|} \quad (12)$$

onde, δx é o desvio padrão do sinal de referência, que pode ser estimado através do desvio padrão do valor de NAS para 15 espectros do branco, por exemplo. O objetivo da utilização de 15 leituras é quantificar a variação do sinal devido à variação instrumental (ruído).

O inverso desse parâmetro, ou seja, γ^{-1} , permite estabelecer a menor diferença de concentração entre amostras, que pode ser distinguida pelo método ⁶².

2.6.6. Seletividade

É a medida do grau de sobreposição entre o sinal da propriedade de interesse e os interferentes presentes na amostra, indicando também, a parte do sinal que é perdida por essa sobreposição. Para modelos de calibração multivariada, a seletividade para a propriedade “k”, SEL_{ki} , é definida como ⁶²:

$$S\hat{E}L_{k,i} = \frac{n\hat{s}_{k,i}}{\|x_{k,i}\|} \quad (13)$$

onde, $n\hat{s}_{k,i}$ é o valor escalar do sinal analítico líquido para a amostra “i” e $x_{k,i}$ representa o vetor de resposta instrumental para a amostra “i”.

2.6.7. Linearidade

A avaliação desta figura de mérito é problemática em calibração multivariada utilizando PLS ou PCR, uma vez que as variáveis são decompostas pelos componentes principais. Assim, uma medida quantitativa para a linearidade não corresponderia a uma tarefa simples, ou mesmo possível. Qualitativamente, os gráficos dos resíduos e de *scores* contra os valores de referência para as amostras de calibração e validação podem indicar se os dados seguem um comportamento linear onde se espera que a distribuição dos resíduos seja aleatória e a dos *scores* linear ⁶².

2.6.8. Ajuste

O ajuste para um modelo não consiste em uma figura de mérito abordada com frequência em guias e normas oficiais. Este parâmetro pode ser estimado a partir da correlação entre os valores previstos pelo modelo e os valores de referência, para a propriedade de interesse. Isso é feito através da reta que melhor se ajusta aos valores de referência e os valores estimados pelo modelo, para as amostras de calibração. Quando o escalar “NAS” é determinado, é possível também determinar o ajuste do modelo através da melhor reta que se ajusta à curva do “NAS” contra a concentração, para as amostras de calibração ⁶².

2.6.9. Razão sinal/ruído

Em calibração multivariada, esta razão indica o quanto da intensidade do NAS da propriedade de interesse é maior do que o desvio padrão do sinal de referência ⁶⁶:

$$S/N_{k,i} = \frac{n\hat{s}_{k,i}}{\delta x} \quad (14)$$

Onde, $n\hat{s}_{k,i}$ é o valor escalar do sinal analítico líquido para a amostra “i” e δx é o desvio padrão do sinal de referência.

2.6.10. Intervalo de medição

É estabelecido determinando a propriedade de interesse em diferentes valores, em um intervalo maior do que se pretende aplicar o método. Através dos resultados obtidos, determina-se a faixa de concentração na qual os resultados apresentam um nível aceitável de incerteza para o método empregado ⁶².

2.6.11. Limite de Detecção e Limite de Quantificação

O limite de detecção (LD) expressa a menor quantidade da espécie de interesse que pode ser detectada e o limite de quantificação (LQ) indica o menor valor que pode ser determinado quantitativamente. Os LD e LQ na calibração multivariada podem ser calculados por ⁷²:

$$LD = \Delta(\alpha, \beta, \nu). RMSEC \sqrt{1 - h_0} \quad (15)$$

onde:

$\Delta(\alpha, \beta, \nu)$ é um parâmetro estatístico que considera a probabilidade de ocorrência de falsos positivos e negativos e o número de graus de liberdade (parâmetro não central da distribuição não central de Student);

$$h_0 = \frac{1}{n} + \sum_{a=1}^A (t_{ia}^2 / \mathbf{t}_a^T \mathbf{t}_a)$$

$$LQ = 100. (RMSEC. \sqrt{1 + h_0}) / RSD(\%) \quad (16)$$

2.6.12. Teste para erros sistemáticos (Bias)

Segundo a IUPAC, erros sistemáticos são calculados pela diferença entre a média da população e o valor verdadeiro e são todas as componentes de erro que não são aleatórias^{62,73}. A norma ASTM E1655⁶³ sugere um teste-t para avaliar se o "bias" é significativo. O "bias" médio pode ser calculado como:

$$bias = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (17)$$

onde, n corresponde ao número de amostras do conjunto de validação externa do conjunto de previsão.

Então, o desvio padrão dos erros de validação (SDV, do inglês *Standard Deviation of Validation*) é obtido por:

$$SDV = \sqrt{\frac{\sum [(y_i - \hat{y}_i) - bias]^2}{n - 1}} \quad (18)$$

Por fim, calcula-se t_{bias} :

$$t_{bias} = \frac{|bias| \sqrt{n}}{SDV} \quad (19)$$

Se o valor de t_{bias} apresentar resultado menor do que o valor de t crítico para n-1 graus de liberdade, onde n é o número de amostras da validação, com 95% de confiança, então, o erro sistemático incluído no modelo pode ser considerado insignificante e desprezado.

A ASTM E1655 também discursa sobre a quantidade mínima de amostras para os conjuntos de calibração e validação. Para o primeiro o número mínimo de amostras deve ser maior que seis vezes o número de variáveis latentes (VL) e no mínimo 24, já para o conjunto de validação, o número de amostras deve ser quatro vezes maior que o número de VL e no mínimo 20.

2.7. Incerteza de medição

Segundo o Vocabulário Internacional de Metrologia (VIM) ⁷⁴, incerteza de medição é um “parâmetro não negativo que caracteriza a dispersão dos valores atribuídos a um mensurando, com base nas informações utilizadas”. Em outras palavras, incerteza de medição é uma faixa de valores na qual o valor medido pode estar estatisticamente contido que, ainda segundo o VIM, pode ser um desvio padrão (incerteza-padrão) ou algum múltiplo dele. A expressão de um resultado de medição somente está completa se além do valor atribuído ao mensurando, também estiver contemplada a incerteza de medição associada a este valor ⁷⁵.

A incerteza de medição compreende componentes provenientes de efeitos sistemáticos, tais como aqueles associados a correções e valores atribuídos a padrões. Algumas das componentes podem ser estimadas por uma avaliação do Tipo A (método de avaliação de incerteza pela análise estatística de séries de observações, que pode ser caracterizado por desvios-padrão). As outras componentes, as quais podem ser estimadas por uma avaliação do Tipo B (método de avaliação de incerteza por outros meios que não a análise estatística de séries de observações), são caracterizadas por desvios-padrão estimados a partir de funções de densidade de probabilidade baseadas na experiência ou em outras informações ^{74,76}.

Para que seja possível a comparação de resultados nas diversas situações, é necessário adotar um procedimento universal para a estimativa da incerteza (como é adotado o sistema internacional, SI, para as unidades) dos resultados de medição. O documento adotado internacionalmente como referência para o estabelecimento da incerteza de um resultado de medição é o “Guide to the Expression of Uncertainty in Measurement” (ISO GUM) ⁷⁶.

A metodologia para a estimativa da incerteza de medição seguindo o ISO GUM ⁷⁴ pode ser resumida através dos seguintes passos principais:

1. Definição do mensurando;
2. Elaboração do diagrama causa e efeito (diagrama de Ishikawa);
3. Estimativas das incertezas das fontes de entrada;
4. Cálculo dos coeficientes de sensibilidade;

5. Cálculo das componentes de incerteza;
6. Combinação das componentes;
7. Cálculo dos graus de liberdade efetivos;
8. Determinação do fator de abrangência;
9. Estimativa da incerteza expandida.

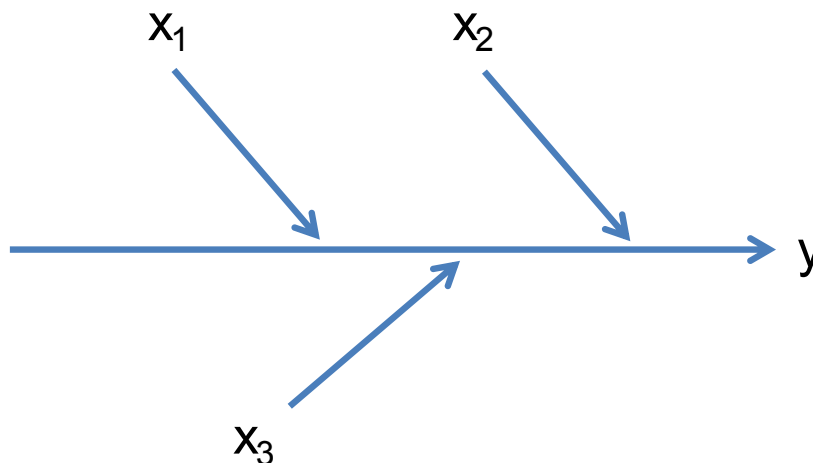
2.7.1. Definição do mensurando

Geralmente, o mensurando y é obtido através de uma relação funcional com as grandezas de entrada x_i , ou seja $y = f(x_1, x_2, \dots, x_n)$. Os valores das grandezas de entrada, bem como suas respectivas incertezas, são determinados diretamente durante a medição ou são obtidos de certificados.

2.7.2. Diagrama de causa e efeito

O objetivo do diagrama causa - efeito é identificar todas as fontes de incerteza que definem a incerteza do mensurando. A **Figura 14** mostra um exemplo desse tipo de diagrama.

Figura 14 – Exemplo de um diagrama causa e efeito para identificação das fontes de incerteza (X_1 , X_2 e X_3) na estimativa da incerteza de medição do mensurando y .



(Fonte: Elaborado pelo autor)

2.7.3. Avaliação das incertezas padrão

A avaliação da incerteza tipo A pode ser realizada pela análise estatística de séries de observações, como no exemplo de medidas em condição de repetibilidade da grandeza de entrada x_i :

$$u(\bar{x}_i) = \frac{s(\bar{x}_i)}{\sqrt{n}} \quad (20)$$

onde:

$s(\bar{x}_i)$ = desvio padrão dos valores individuais do conjunto de repetições;

n = número de medidas realizadas

Já a avaliação do Tipo B, é realizada por um método que não seja a análise estatística de séries de observações. Essa avaliação pode ser feita considerando-se que a variação de x_i tem uma determinada distribuição e um intervalo de dispersão, é o caso da incerteza-padrão devido à resolução de uma indicação digital, pois há uma faixa de sinais de entrada no instrumento, varrendo um intervalo conhecido, que forneceria o mesmo valor (ou seja, se a resolução de uma indicação é 0,1, o valor do

estímulo que produz essa indicação pode estar situado com igual probabilidade em qualquer lugar no intervalo $X-0,1/2$ a $X+0,1/2$) e nesse caso a distribuição é retangular e a estimativa da incerteza padrão é dada por:

$$u(x_i) = \frac{\delta x / 2}{\sqrt{3}} \quad (21)$$

onde:

δx é a resolução da indicação;

Se a distribuição fosse triangular, o divisor da equação acima seria $\sqrt{6}$.

Outro caso ocorre quando a incerteza provém de um certificado de calibração, o qual fornece a incerteza expandida e o fator de abrangência, então a incerteza-padrão é definida por:

$$u(x_i) = \frac{U}{k} \quad (22)$$

onde:

U é a incerteza expandida;

k é o fator de abrangência.

2.7.4. Incertezas correlacionadas e não correlacionadas

Um exemplo ⁷⁵ no qual ocorre uma correlação entre as incertezas das fontes de entrada é no processo de obtenção de um volume de 40 mL, a partir de uma única pipeta com valor nominal de volume igual a 20 mL. Se o volume fosse obtido por duas pipetas de 20 mL as incertezas não apresentariam correlação.

2.7.5. Cálculo dos coeficientes de sensibilidade

O coeficiente de sensibilidade tem a função de estimar a variação do mensurando em relação à variação de uma dada fonte de entrada. Quando se tem uma

relação matemática entre o mensurando y e a fonte de entrada x_i , o coeficiente de sensibilidade pode ser determinado da seguinte forma:

$$c_i = \frac{\partial y}{\partial x_i} \quad (23)$$

Caso não exista uma relação direta entre o mensurando e uma dada fonte de entrada, a determinação do c_i é feita através de experimentos.

2.7.6. Componentes de incerteza

Depois de estimadas as incertezas padrão das fontes de entrada e também os coeficientes de sensibilidade de cada uma delas, é possível avaliar as componentes de incerteza na mesma unidade do mensurando:

$$u_{x_i}(y) = \frac{\partial y}{\partial x_i} u(x_i) = c_{x_i} u(x_i) \quad (24)$$

onde:

$u_{x_i}(y)$ é o componente de incerteza na unidade do mensurado referente à fonte x_i ;

c_{x_i} é o coeficiente de sensibilidade;

$u(x_i)$ é a incerteza-padrão da estimativa de entrada x_i .

2.7.7. Combinação das componentes

Para estimar a incerteza padrão combinada $[u_c(y)]$ para grandezas não correlacionadas, aplica-se a seguinte equação:

$$u_c(y) = \sqrt{\sum_{i=1}^N \left(\frac{\partial y}{\partial x_i} u(x_i) \right)^2} \quad (25)$$

Já para grandezas correlacionadas, a equação é:

$$\begin{aligned}
 u_c^2(y) = & \sum_{i=1}^N \sum_{j=1}^N \frac{\partial y}{\partial x_i} \frac{\partial y}{\partial x_j} u(x_i, x_j) = \sum_{i=1}^N \left[\frac{\partial y}{\partial x_i} \right]^2 u^2(x_i) \\
 & + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial y}{\partial x_i} \frac{\partial y}{\partial x_j} u(x_i, x_j)
 \end{aligned}
 \tag{26}$$

onde:

$u(x_i, x_j) = u(x_j, x_i)$ é a covariância estimada associada com x_i e x_j .

Em termos de coeficientes de correlação, a Equação 26 pode ser reescrita:

$$u_c^2(y) = \sum_{i=1}^N \left[\frac{\partial y}{\partial x_i} \right]^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial y}{\partial x_i} \frac{\partial y}{\partial x_j} u(x_i) u(x_j) r(x_i, x_j)
 \tag{27}$$

onde o coeficiente de correlação pode ser calculado através da seguinte equação:

$$r(x_i, x_j) = \frac{u(x_i, x_j)}{u(x_i) u(x_j)}
 \tag{28}$$

onde $r(x_i, x_j) = r(x_j, x_i)$ e $-1 \leq r(x_i, x_j) \leq +1$. Se as estimativas x_i e x_j são independentes, $r(x_i, x_j) = 0$ e a variação numa delas não implica uma variação esperada na outra. Já para o caso em que todas as estimativas de entrada são correlacionadas, $r(x_i, x_j) = 1$ e a Equação 27 se reduz a:

$$u_c^2(y) = \left[\sum_{i=1}^N \left[\frac{\partial y}{\partial x_i} \right] u(x_i) \right]^2 = \left[\sum_{i=1}^N c_{x_i} u(x_i) \right]^2
 \tag{29}$$

2.7.8. Determinação dos graus de liberdade

Através da chamada fórmula de Welch-Sttterthwaite, é possível obter o número efetivo de graus de liberdade da incerteza padrão combinada:

$$v_{eff} = \frac{u_c^4(y)}{\sum_{i=1}^N \frac{u_i^4(y)}{v_i}} \quad (30)$$

onde:

$u_c(y)$ é a incerteza padrão combinada;

$u_i(y)$ é a incerteza padrão da fonte de entrada, na unidade do mensurando;

v_i é o número de graus de liberdade das fontes de entrada.

O número de graus de liberdade é um número inteiro ⁷⁷. Se o valor calculado de v_{eff} não for um número inteiro, considera-se somente a parte inteira do número. Para uma incerteza padrão do tipo B, considera-se o número de graus de liberdade infinito.

2.7.9. Determinação do coeficiente de abrangência

O fator de abrangência é definido a partir da distribuição t de Student e o mesmo depende da probabilidade de abrangência e dos graus de liberdade efetivos da incerteza padrão combinada do mensurando. É possível observar essa dependência em uma tabela da distribuição t de Student ⁷⁸.

2.7.10. Estimativa da incerteza expandida

O propósito da incerteza expandida (U) é prover um intervalo $y - U$ a $y + U$ com o qual se espera abranger uma extensa fração da distribuição de valores que podem razoavelmente ser atribuídos ao mensurando Y . A incerteza expandida U , para uma determinada probabilidade de abrangência p , é estimada pela seguinte equação:

$$U = k_{(p,\nu)} \cdot u_c(y) \quad (31)$$

onde:

$k_{(p,\nu)}$ é o fator de abrangência. O valor de k é selecionado com base no nível da confiança requerido para o intervalo (tabela da distribuição t).

2.8. Simulação de Monte Carlo

A metodologia para a estimativa de incerteza mostrada no ISO GUM é baseada em algumas suposições, que nem sempre são válidas ⁷⁸:

- ✓ O modelo usado para calcular o mensurando deve possuir uma não linearidade insignificante. Quando o modelo apresenta não linearidade, o truncamento do primeiro termo da série de Taylor pode não ser suficiente para estimar corretamente a incerteza.
- ✓ Validade do teorema do limite central, que afirma que a convolução de um grande número de distribuições resulta em uma distribuição normal. Assim, supõe-se que a distribuição de probabilidade da saída é aproximadamente normal e pode ser representada por uma distribuição t. Na prática, essa distribuição resultante pode não ter um comportamento simétrico ou não tende a uma distribuição normal, o que invalida a abordagem do teorema do limite central.
- ✓ O ISO GUM usa a equação de Welch-Satterthwaite para obter os graus de liberdade efetivos que são necessários para o cálculo da incerteza expandida. Porém, essa equação nem sempre é adequada para todos os casos.

Além disso, a metodologia do ISO GUM pode não ser válida quando uma ou mais fontes de incerteza são muito maiores que outras, ou ainda quando as distribuições das grandezas de entrada não são simétricas.

Nesse sentido, o Suplemento 1 do ISO GUM (ou JCGM 101: 2008) ⁷⁹ fornece orientações básicas para usar a simulação de Monte Carlo para a propagação de

distribuições em metrologia. Esse documento apresenta um método alternativo rápido e robusto para os casos em não se aplicam a abordagem do ISO GUM ou onde ela não é válida.

O Suplemento 1 do ISO GUM indica uma sequência de passos para a estimativa da incerteza:

- a) Definição do mensurando e da fonte de entrada.
- b) Modelagem.
- c) Identificação das funções densidade de probabilidade (PDF – do inglês *Probability Density Functions*), correspondentes a cada fonte de entrada.
- d) Configuração e execução da simulação de Monte Carlo.
- e) Resumo e expressão os resultados.

Os dois primeiros são exatamente os mesmos descritos no ISO GUM. No terceiro passo deve-se selecionar uma PDF que não transmita mais informação do que aquela que é conhecida. Por exemplo, se a única informação disponível sobre uma fonte de entrada são os limites máximo e mínimo, uma PDF uniforme deve ser usada. Feito isso, o número de iterações (M) deve ser selecionado. Por exemplo, quando a probabilidade de abrangência necessária é de 95%, M deve ser de pelo menos maior do que 2×10^5 .

Depois de selecionado o valor de M, o próximo passo é a execução da simulação em si, então a última etapa é resumir e expressar os resultados. Os seguintes parâmetros devem ser relatados como resultados:

- a) A estimativa da grandeza de saída, calculada como a média dos valores gerados na simulação.
- b) A incerteza padrão (desvio padrão destes valores gerados).
- c) A probabilidade de abrangência escolhida (geralmente 95%).
- d) Os valores das extremidades correspondentes ao intervalo de abrangência selecionado.

2.9. Gastos de uma empresa

Nesta seção serão definidos alguns termos aplicados na área de custos de um processo produtivo. No caso desta tese, os processos são os ensaios utilizados para medir os parâmetros de qualidade do combustível E85.

Gasto é aquisição de um bem ou serviço que implica em um desembolso da empresa ⁸⁰. Devido à grande amplitude desse conceito, os gastos podem ser classificados em investimento, custo, despesa, perda e desperdício.

2.9.1. Investimento

São os valores que uma empresa gasta para aquisição de bens que serão estocados até o momento de sua utilização e em bens patrimoniais (máquinas, equipamentos, instalações, etc). É importante salientar que estes últimos sofrem depreciação. Uma característica natural de investimento, do ponto de vista de uma empresa, é que se espera que ele proporcione retorno para a instituição ⁸⁰.

2.9.2. Custo

É todo gasto que representa a aquisição de bens ou serviços que serão utilizados na produção. Esse tipo de gasto se caracteriza por ocorrer somente na atividade produtiva, ou seja, só devem ser incluídos no custo dos produtos ou serviços os insumos necessários à elaboração dos mesmos ⁸⁰. Alguns exemplos de custo são:

- ✓ Reagentes utilizados em um determinado ensaio;
- ✓ Salários e encargos da mão de obra dos profissionais que executam os ensaios;
- ✓ Depreciação dos equipamentos.

Em relação aos produtos fabricados, os custos podem ser classificados em diretos e indiretos. Já em função do volume de produção, podem ser classificados em custos fixos e custos variáveis ⁸⁰.

2.9.2.1. Custos diretos

O custo direto ⁸⁰ (CD), como o próprio nome já diz, está diretamente relacionado com a fabricação. Um exemplo disso são os insumos gastos para a fabricação de uma unidade de produto. Conhecendo-se o preço de aquisição dos insumos gastos em uma unidade, basta multiplicá-lo pela quantidade gasta de cada insumo.

Outro custo direto é a mão de obra direta (MOD), que são os custos com pessoal ligados diretamente à produção de um bem ou serviço. Conhecendo-se os valores pagos aos funcionários por hora de trabalho e o tempo (em horas) que eles gastam na elaboração de cada produto ou serviço, basta multiplicar esses dois valores para obter o custo da mão de obra por unidade.

2.9.2.2. Custos indiretos de fabricação

São todos os custos que necessitam de cálculos para serem distribuídos aos diferentes produtos ou serviços elaborados por uma empresa. Por exemplo, depreciação de equipamentos que são utilizados para fabricação de mais de um produto. Normalmente, os custos fixos são atribuídos aos produtos elaborados por meio de cálculos, pois a maioria dos custos é indireta.

2.9.2.3. Custos fixos

São os custos que se mantêm, independentemente do volume de produção da empresa, dentro de um intervalo relevante. Tal intervalo é uma faixa de quantidades abaixo da qual a empresa não tem interesse em produzir e acima da qual não apresenta capacidade produtiva suficiente. Pode-se citar como exemplos o aluguel, imposto predial e seguros da fábrica.

2.9.2.4. Custos variáveis

Este tipo de custo é aquele que tem seu valor alterado de maneira diretamente proporcional ao volume produzido. Por exemplo, matéria-prima consumida e mão de obra direta.

2.9.3. Despesa

É um gasto que a empresa possui para manter sua estrutura organizacional e visando a obtenção de receita ⁸⁰. São exemplos:

- ✓ Aluguel do escritório central;
- ✓ Salários e encargos do pessoal administrativo;
- ✓ Comissões sobre vendas.

2.9.4. Perda

É o gasto que a empresa fica sujeita quando bens ou serviços são consumidos de maneira anormal às suas atividades, como no caso de inundações, incêndios e greves ⁸⁰.

2.9.5. Desperdício

É "um gasto que a empresa apresenta pelo fato de não ocorrer o aproveitamento normal de todos os seus recursos" ⁸⁰. Por exemplo, produtividade muito abaixo da capacidade.

3. OBJETIVOS

O principal objetivo foi obter um método baseado em espectroscopia na região do infravermelho e regressão multivariada, que fosse mais rápido e mais barato que os métodos oficiais. Além disso, objetivou-se avaliar tal método do ponto de vista metrológico, realizando a validação e estimativa da incerteza de medição. Esse objetivo central pode ser descrito mais detalhadamente como se segue:

- ✓ Produzir em laboratório amostras de E85.
- ✓ Obter os perfis espectroscópicos (nas regiões do infravermelho médio e próximo) das amostras.
- ✓ Correlacionar os perfis espectroscópicos através de ferramentas quimiométricas (regressão multivariada) com os parâmetros de qualidade.
- ✓ Avaliar e validar os modelos quimiométricos obtidos, através de figuras de mérito.
- ✓ Realizar os cálculos da estimativa de incerteza de medição através das diretrizes do ISO GUM e por simulação de Monte Carlo.
- ✓ Efetuar os cálculos dos custos do método proposto e dos oficiais e compará-los.

4. EXPERIMENTAL

4.1. Reagentes

- ✓ Etanol absoluto (99,9%), da marca MERCK, lotes: K43665283 e K45251683405. Validade: 31/07/17 e 31/01/19, respectivamente.
- ✓ Metanol (99,9%), da marca Sigma-Aldrich, lote: SHBC0287V. Validade janeiro/17.
- ✓ Água destilada.
- ✓ Diferentes amostras de gasolina A, cedidas gentilmente pelo laboratório Flumignano (Cuiabá – MT).

4.2. Preparação das amostras

Foram preparadas 80 amostras de E85 a partir da mistura de etanol, diferentes amostras de gasolina A (HC), metanol e água. Os teores variaram, em razão mássica, de forma randômica nos seguintes intervalos: 0-2% para água, 50-90% para etanol, 0-1% para metanol e 0-50% para gasolina. As amostras foram planejadas com o uso da função "aleatórioentre" do aplicativo Microsoft Excel[®] e preparadas em balança analítica (da marca Ohaus, modelo AR2140). Ao final todas tiveram massa final aproximada de 10 g. As amostras foram mantidas em frascos selados para cromatografia (*vial*) até o momento do ensaio.

4.3. Obtenção dos espectros de infravermelho

O modelo do espectrofotômetro utilizado foi um Nicolet 6700 FTIR (da marca Thermo Scientific). Na região do infravermelho médio, utilizou-se um acessório de ATR modelo Smart ARK[™] [Figura 15 - (B)], com um cristal de ZnSe com ângulo de incidência de 45°. Os espectros, na região do MIR, foram coletados no intervalo compreendido entre 4000 cm⁻¹ a 650 cm⁻¹, sendo cada espectro composto pela média de 32 varreduras e com resolução de 4 cm⁻¹, aplicando-se, com o auxílio de uma

seringa e uma agulha (para retirada da amostra do *vial*), o volume de aproximadamente 1 mL de cada amostra.

Figura 15 – Acessórios utilizados para a leitura das amostras: (A) esfera de integração e (B) cristal de ZnSe para refletância total atenuada.

(A)



(B)



Inicialmente, foram realizados alguns testes para aquisição dos espectros das amostras na região do NIR com a utilização de um acessório de transletância (**Figura 16**) que consiste em um espelho de ouro (com uma distância conhecida, que define o caminho óptico) recobrimdo uma bolsa de plástico que contém a amostra. Porém, observou-se que a utilização desse sistema apresentou algumas limitações, como a necessidade da preparação de uma bolsa por amostra e a formação de bolhas na bolsa.

Figura 16: Acessório original para leitura no modo de medição transfletância para utilização sobre a janela de safira da esfera de integração.



Desenvolvemos um novo acessório para que não houvesse a necessidade de preparar (adicionar à bolsa) as amostras. Esse novo acessório consiste em uma peça de aço inoxidável usinada (**Figura 17 e Figura 18**) a qual é colocada sobre uma porção da amostra adicionada anteriormente sobre a janela de safira. Com isso, foi possível manter o mesmo modo de leitura e eliminar o processo de preparação das amostras.

Figura 17 – Acessório de aço inoxidável usinado desenvolvido para leitura no modo de medição transfletância para utilização sobre a janela de safira da esfera de integração.

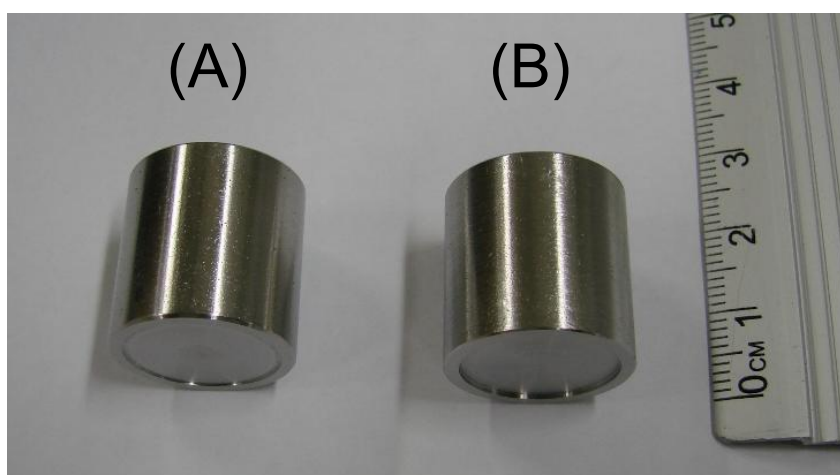
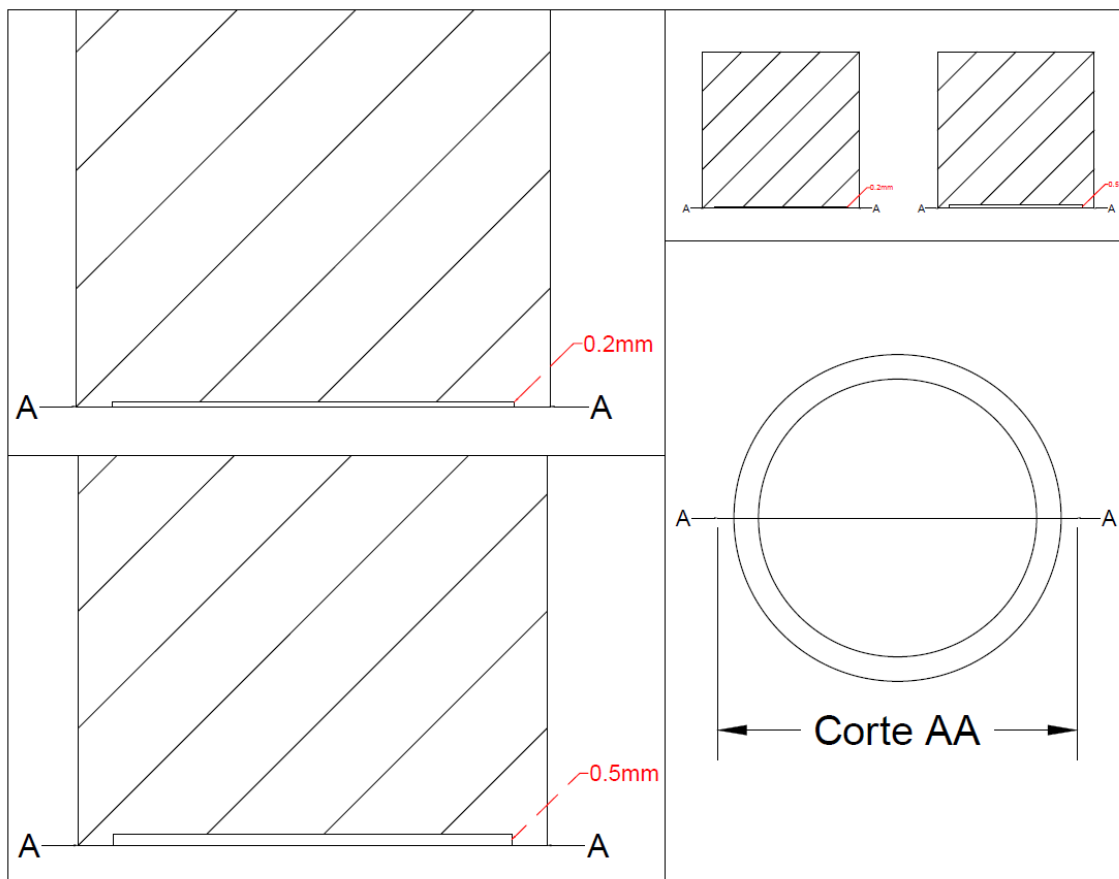


Figura 18 – Projeções dos acessórios de aço inoxidável usinado utilizados para a coleta dos espectros na região do NIR.



Sendo assim, na coleta dos espectros na região do infravermelho próximo, empregou-se o acessório Smart NIR Integrating Sphere [Figura 15 - (B)] com dois caminhos ópticos: 0,2 mm [Figura 17 - (A)] e 0,5 mm [Figura 17 - (B)], na região de 10000 cm^{-1} a 4000 cm^{-1} , aplicando 32 varreduras e com resolução de 4 cm^{-1} .

Para ambos os casos, o espectro do ar foi utilizado como referência e a temperatura do laboratório foi mantida entre $20\text{ }^{\circ}\text{C}$ e $22\text{ }^{\circ}\text{C}$.

4.4. Análise quimiométrica

As análises quimiométricas foram realizadas via Software TQ Analyst 8.0.1.30, da Thermo Scientific (para os testes iniciais), além do Matlab 8.1.0.604 (R2013a), da The Mathworks e PLS_Toolbox 7.3.1, da Eigenvector Research (para a construção e previsão dos modelos).

Objetivando manter a maior variabilidade das amostras no conjunto de dados utilizados para calibrar os modelos, a separação desses conjuntos entre os de calibração (48 amostras) e de previsão (28 amostras) foi realizada através do algoritmo de Kennard-Stone (**ANEXO A**). A somatória das médias quadráticas dos erros da validação cruzada foram obtidos através do processo *leave-one-out*.

4.5. Validação

A validação dos modelos foi feita pela avaliação das figuras de mérito. Para o cálculo das mesmas, foram utilizadas as equações apresentadas na sessão 2.6 Validação, com o auxílio dos algoritmos mostrados nos **ANEXO B** e **ANEXO C**.

4.6. Estimativa da Incerteza de Medição

Os cálculos da estimativa da incerteza de medição, segundo o ISO GUM, foram realizados com o auxílio de planilhas eletrônicas fornecidas pelo Sistema de Gestão da Qualidade do Cempeq (Centro de Monitoramento e Pesquisa da Qualidade da Qualidade de Combustíveis, Biocombustíveis, Petróleo e Derivados, atual Centro Multidisciplinar de Pesquisa em Combustíveis, Biocombustíveis, Petróleo e Derivados). O software utilizado para processar tais planilhas foi o Microsoft Excel® 2007. Já a simulação de Monte Carlo foi realizada no software Crystal Ball.

4.7. Estimativa de custos

As estimativas de custos foram realizadas com auxílio do Microsoft Excel® 2007 e com valores de compra dos insumos para os ensaios nos anos de 2014 e 2015 fornecidos pela Administração do Cempeq. Os valores dos equipamentos também foram fornecidos pela Administração do Cempeq.

5. RESULTADOS E DISCUSSÃO

5.1. Preparação das amostras

As amostras foram planejadas com a função “aleatórioentre” do Microsoft Excel®. Como essa função retorna apenas números inteiros, foi necessário selecionar um intervalo com números com ordem de grandeza maior e depois se efetuar a divisão para se obter os valores na ordem de grandeza desejada, por exemplo o valor de retorno da função foi 52736 g, então ao se dividir por 10000 temos 5,2736 g. Os valores das concentrações planejadas e das concentrações experimentais de algumas amostras são mostradas na **Tabela 2**.

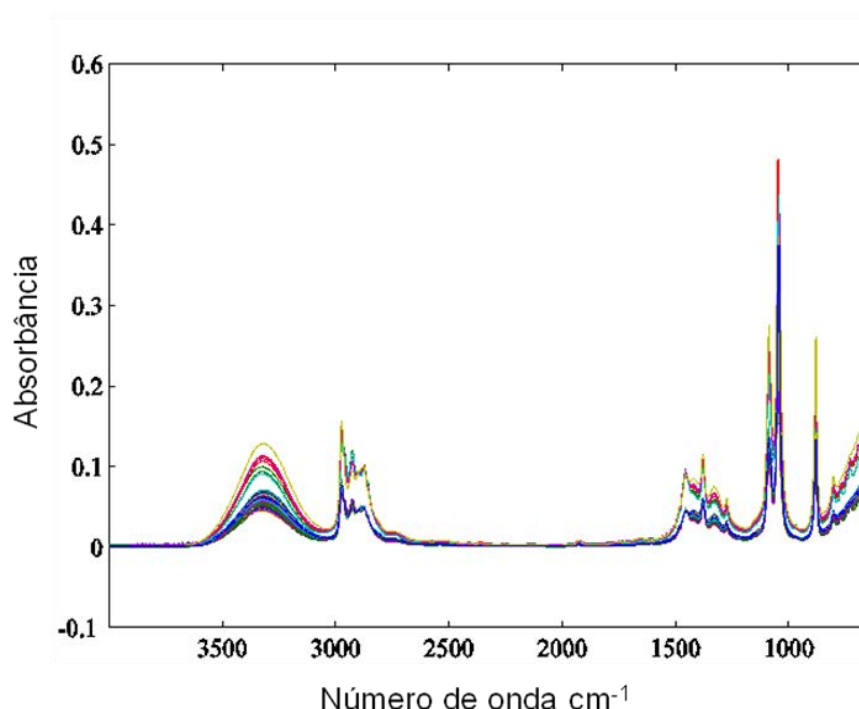
Tabela 2 – Valores dos teores planejados (obtidos com a função "aleatórioentre") e experimentais para algumas amostras preparadas e utilizadas para a construção dos modelos.

Concentração planejada (% m/m)					Concentração experimental (% m/m)			
Nome	Etanol	Metanol	HC	H ₂ O	Etanol	Metanol	HC	H ₂ O
80	93,88	0,60	5,50	0,02	93,89	0,59	5,52	0,01
81	59,20	0,55	40,05	0,20	59,26	0,55	39,98	0,21
82	59,51	0,75	39,55	0,19	58,60	0,77	40,45	0,18
83	95,78	0,88	2,65	0,69	95,54	0,88	2,85	0,72
84	59,78	0,69	39,00	0,53	59,79	0,70	38,94	0,57
85	75,49	0,96	22,80	0,76	75,43	0,97	22,79	0,82
86	82,42	0,03	16,80	0,75	82,44	0,04	16,80	0,73
87	79,94	0,60	18,80	0,66	79,94	0,62	18,77	0,67
88	88,07	0,42	11,10	0,40	87,94	0,43	11,22	0,41
89	84,11	0,77	14,60	0,52	83,94	0,80	14,72	0,54

5.2. Espectroscopia na região do infravermelho médio

A **Figura 19** mostra os espectros de diferentes amostras que foram adquiridos na faixa de $4000\text{--}650\text{ cm}^{-1}$ com o auxílio do acessório de ATR. É possível observar bandas em aproximadamente $3600\text{--}3000\text{ cm}^{-1}$ correspondentes ao estiramento da ligação O-H (presente nas moléculas de água e dos alcoóis). Na região de $3000\text{--}2800\text{ cm}^{-1}$ ocorre o estiramento da ligação C-H devido à contribuição dessa ligação nas moléculas dos alcoóis e da gasolina. Em $1470\text{--}1260\text{ cm}^{-1}$ é possível verificar as deformações das ligações CH_3 e CH_2 e as deformações no plano da ligação O-H, em $1200\text{--}1000\text{ cm}^{-1}$ há o estiramento da ligação C-O (basicamente devido à contribuição das moléculas de etanol e metanol) e na região $900\text{--}800\text{ cm}^{-1}$ é possível observar bandas correspondes ao estiramento das ligações C-C-O⁸¹.

Figura 19 – Espectros de várias amostras diferentes de E85 adquiridos na faixa $4000\text{--}650\text{ cm}^{-1}$.



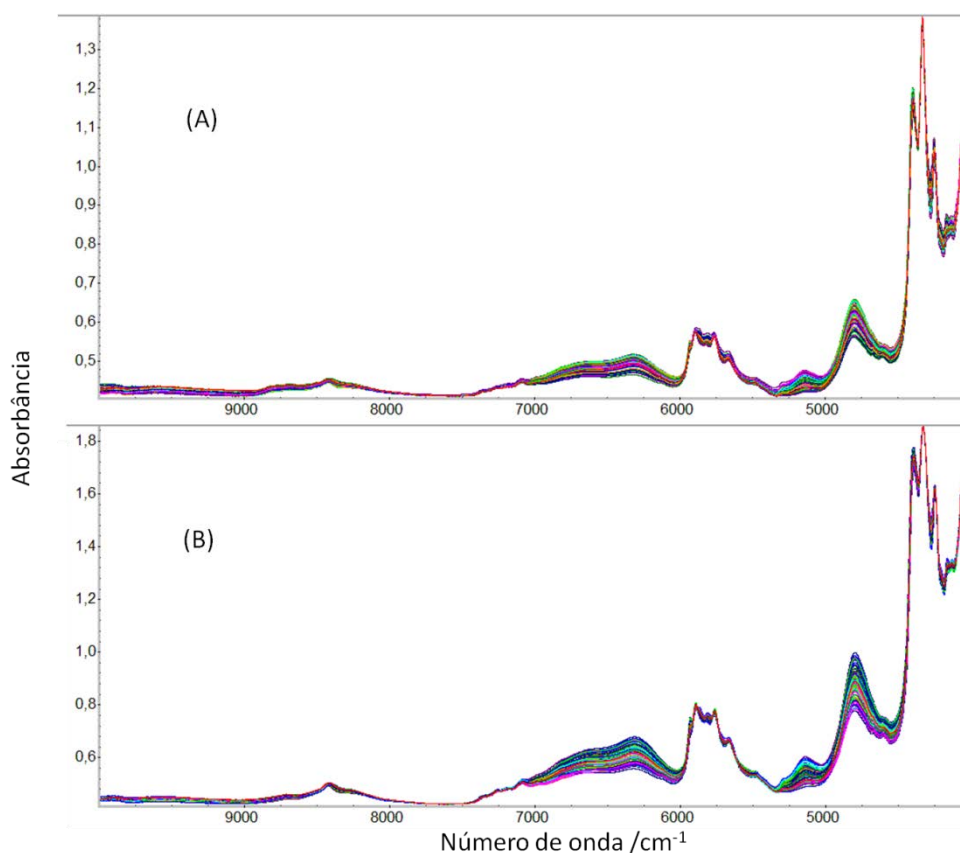
É possível visualizar as diferentes intensidades das bandas devido à variabilidade de composição das amostras e também a existência de regiões do espectro em que não ocorre absorção significativa pelas amostras (aproximadamente $2600\text{--}1700\text{ cm}^{-1}$). Esta informação é importante durante a seleção das variáveis, pois

indica que nessa região não há informações químicas relevantes para a construção dos modelos de regressão multivariada.

5.3. Espectroscopia na região do infravermelho próximo

A **Figura 20** mostra os espectros de diferentes amostras obtidos com a utilização da esfera de integração e do acessório de aço inoxidável confeccionado para este trabalho, com caminhos ópticos de 0,2 mm (A) e 0,5 mm (B). Diferentemente dos espectros obtidos na região do MIR, os obtidos na região do NIR apresentam bandas largas e ocorrem como resultado de sobretons e combinações de vibrações fundamentais⁸². Entretanto, algumas bandas ocorrem com regularidade suficiente para caracterizar grupos moleculares⁸³.

Figura 20 – Espectros de diferentes amostras de E85 adquiridos na faixa 4000–1000 cm^{-1} .



As bandas no intervalo 6000–5000 cm^{-1} e 9090–8200 cm^{-1} são atribuídas respectivamente ao primeiro e segundo harmônicos (o termo em inglês é *overtones*) das vibrações de deformação axial da ligação C-H. Bandas de combinação associadas

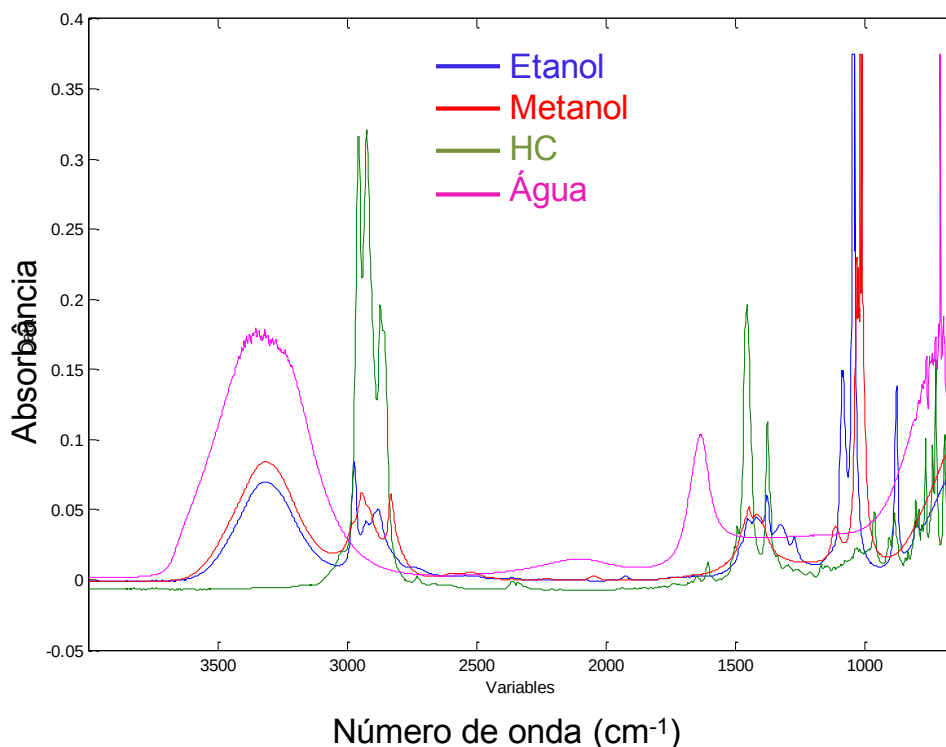
aos grupos C-H podem ser observadas na região $5000\text{--}4000\text{ cm}^{-1}$ ^{10,81}. Também é possível observar, em aproximadamente 6870 cm^{-1} , o primeiro harmônico referente à ligação O-H, presente nas moléculas de álcoois e de água das amostras. Em 5100 cm^{-1} ocorre a banda de combinação da deformação axial e angular da ligação O-H ¹⁰.

É possível observar na **Figura 20 - (A)** que existe uma variação de absorbância mais pronunciada entre as amostras coletadas com caminho óptico de 0,5 mm (região de aproximadamente $7000\text{ a }6000\text{ cm}^{-1}$ e região de aproximadamente $6500\text{ a }5500\text{ cm}^{-1}$) do que as coletadas com caminho óptico de 0,2 mm.

5.4. Regressão multivariada utilizando os espectros na região do MIR

Primeiramente foram construídos modelos para previsão da concentração de etanol (Etanol), metanol (Metanol), gasolina (HC) e água (Água) utilizando todo o espectro, na sequência foram selecionadas algumas regiões e assim novos modelos foram construídos. A seleção dessas regiões (variáveis) se deu, primeiramente, com base na identificação das bandas presentes nos espectros individuais dos componentes utilizados para produção das amostras (**Figura 21**).

Figura 21 – Espectros na região do MIR de cada um dos componentes utilizados na produção das amostras de combustível E85.



Objetivando obter modelos com os menores valores de erros (RMSEC e RMSECV) para a concentração de etanol e metanol, imaginou-se inicialmente que a região referente ao estiramento das ligações C-C-O ($1100\text{--}1000\text{ cm}^{-1}$) conteria grande quantidade de informação útil para os modelos Etanol e Metanol, no entanto quando estes dois modelos foram construídos utilizando apenas essa região do espectro, os valores dos erros foram superiores aos valores de erros ou de LV para o espectro todo (RMSEC = 2,61% (m/m); RMSECV = 3,40% (m/m); LV = 8 para Etanol e RMSEC = 0,07% (m/m); RMSECV = 0,10% (m/m) e LV = 12 para Metanol). Porém, a partir de uma observação da ampliação da faixa espectral $1100\text{--}1000\text{ cm}^{-1}$, foi possível identificar deformações no topo das bandas devido à saturação do detector. Logo, apesar de haver grande quantidade de informação à respeito daquelas moléculas, não foi possível utilizar essa região.

Com a finalidade de facilitar a seleção de variáveis, o espectro foi dividido em algumas regiões (**Figura 22**). Após sucessivos testes com as diferentes regiões e também das combinações entre elas, foram selecionadas as que resultaram em modelos com menores valores de erros (a região D da **Figura 22** não foi utilizada

devido à problemática exposta no parágrafo anterior). Em comparação com o espectro todo, os valores dos erros e o número de variáveis latentes são menores nos modelos em que houve a seleção de variáveis baseada no conhecimento químico, a qual removeu as variáveis que não possuem informações relevantes em relação às propriedades de interesse (concentração das substâncias).

Também foram testados os seguintes pré-processamentos: primeira derivada (Savitzky-Golay) e Correção Multiplicativa de Sinal (MSC, do inglês *Multiplicative Scatter Correction*), centrando e não centrando os dados na média. O pré-processamento matemático que resultou em menores erros para os modelos foi MSC sem centrar os dados na média. A **Tabela 3** mostra as regiões selecionadas, os erros e o número de variáveis latentes para cada modelo.

Figura 22 – Regiões consideradas para a seleção de variáveis

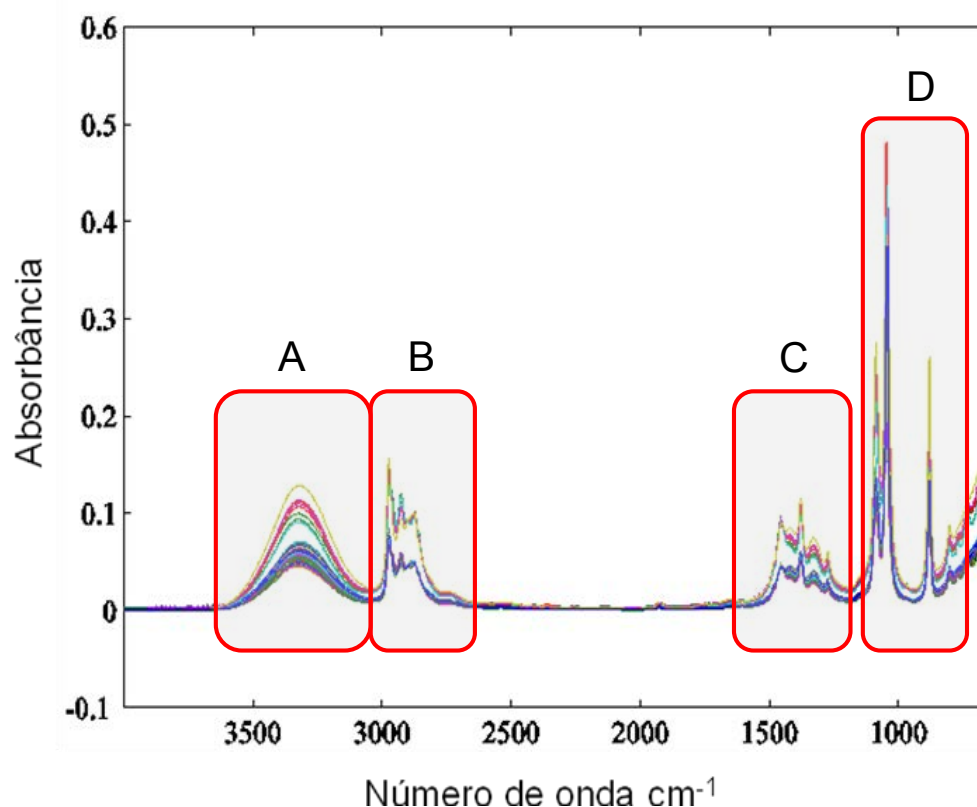


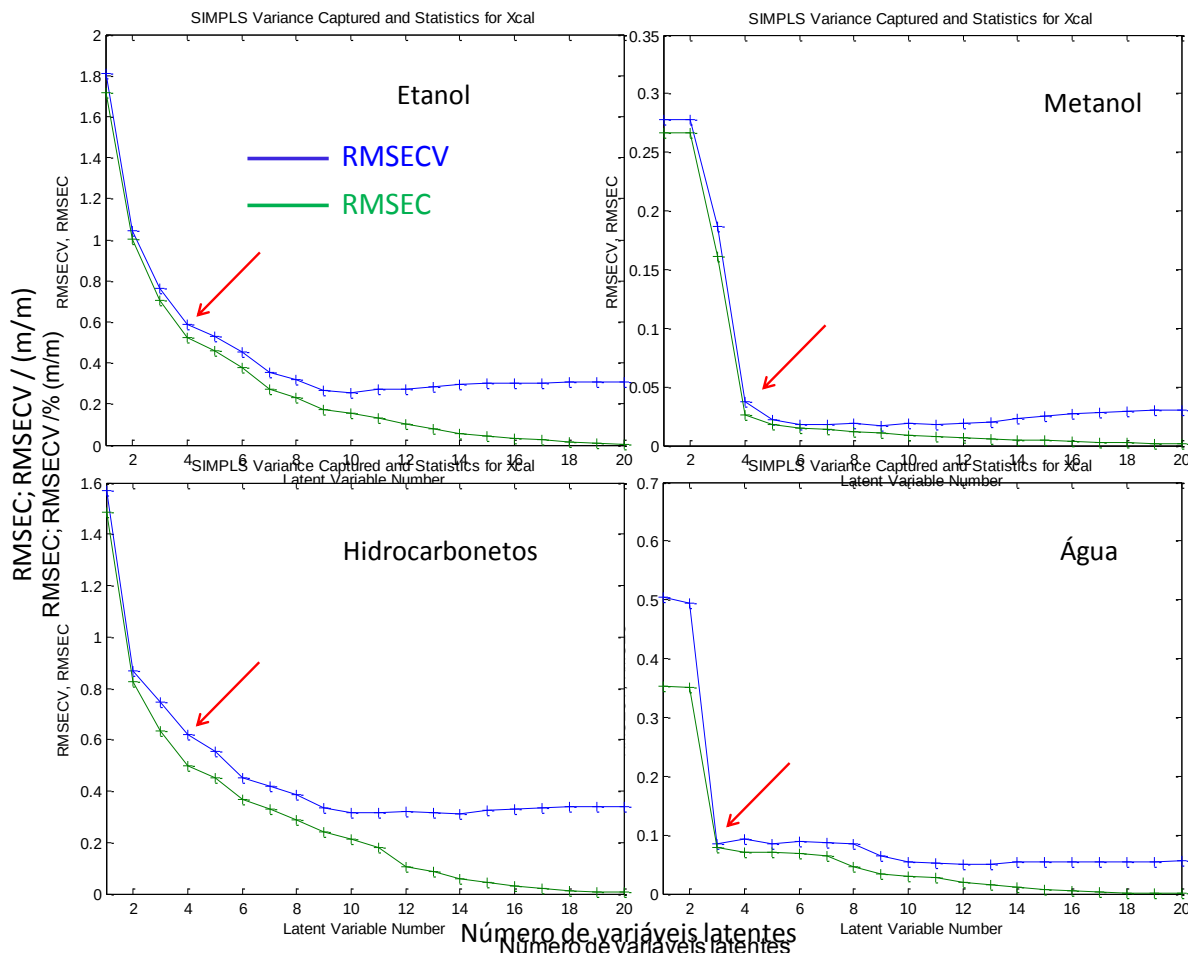
Tabela 3 – Regiões selecionadas e os erros de cada modelo na região do MIR.

Modelo	Região selecionada (cm ⁻¹)	RMSEC % (m/m)	RMSECV % (m/m)	# VL ^a	Todo o espectro		
					RMSEC % (m/m)	RMSECV % (m/m)	# LV ^a
Etanol	2653-3011 1582-1190	0,52	0,59	4	2,05	2,65	4
Metanol	2764-3002	0,02	0,04	4	0,08	0,12	11
HC	2773-3026 1531-1251	0,49	0,62	4	2,05	2,65	4
Água	2614-3640	0,08	0,08	3	0,13	0,16	7

^a Número de variáveis latentes.

A **Figura 23** mostra o comportamento dos valores dos erros de calibração (RMSEC) e os valores de validação cruzada (RMSECV) em função do número de variáveis latentes. O número de VL para cada modelo foi selecionado com base na observação desses gráficos. O critério adotado para essa seleção foi o ponto do gráfico em que a diminuição do número de VL não resulta em um decréscimo significativo nos valores dos erros.

Figura 23 – Erros de calibração e validação cruzada em função do número de variáveis latentes para os modelos (A) Etanol; (B) Metanol; (C) HC e (D) Água. As setas indicam o número de variáveis latentes escolhidas.



5.5. Regressão multivariada utilizando os espectros na região do NIR

Assim como para os espectros na região do MIR, inicialmente foram construídos modelos com todo o espectro e depois com regiões selecionadas com base na variação espectral observada entre as amostras, pois as variações se devem às diferenças de concentração das substâncias que constituem as amostras (**Figura 24** e **Figura 25**), bem como os pré-processamentos. A **Tabela 4** mostra quais foram as regiões selecionadas que resultaram nos menores valores de erros para os modelos construídos com os espectros obtidos com o espelho de caminho óptico de 0,2 mm (NIR02), já a **Tabela 5** para os modelos baseados na região do NIR com caminho óptico de 0,5 mm (NIR05), ambos com pré-processamento MSC.

Figura 24 – Espectros na região do NIR (caminho óptico de 0,2 mm) de cada um dos componentes utilizados na produção das amostras de combustível E85.

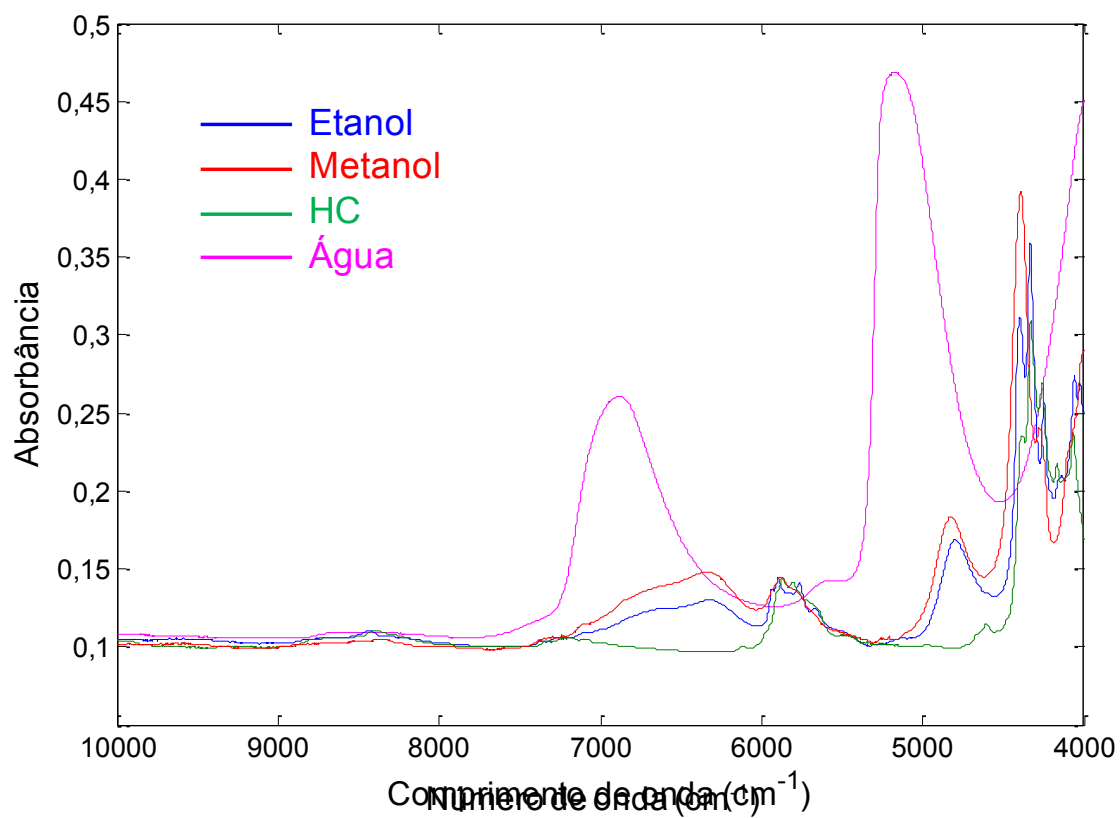


Figura 25 – Espectros na região do NIR (caminho óptico de 0,5 mm) de cada um dos componentes utilizados na produção das amostras de combustível E85.

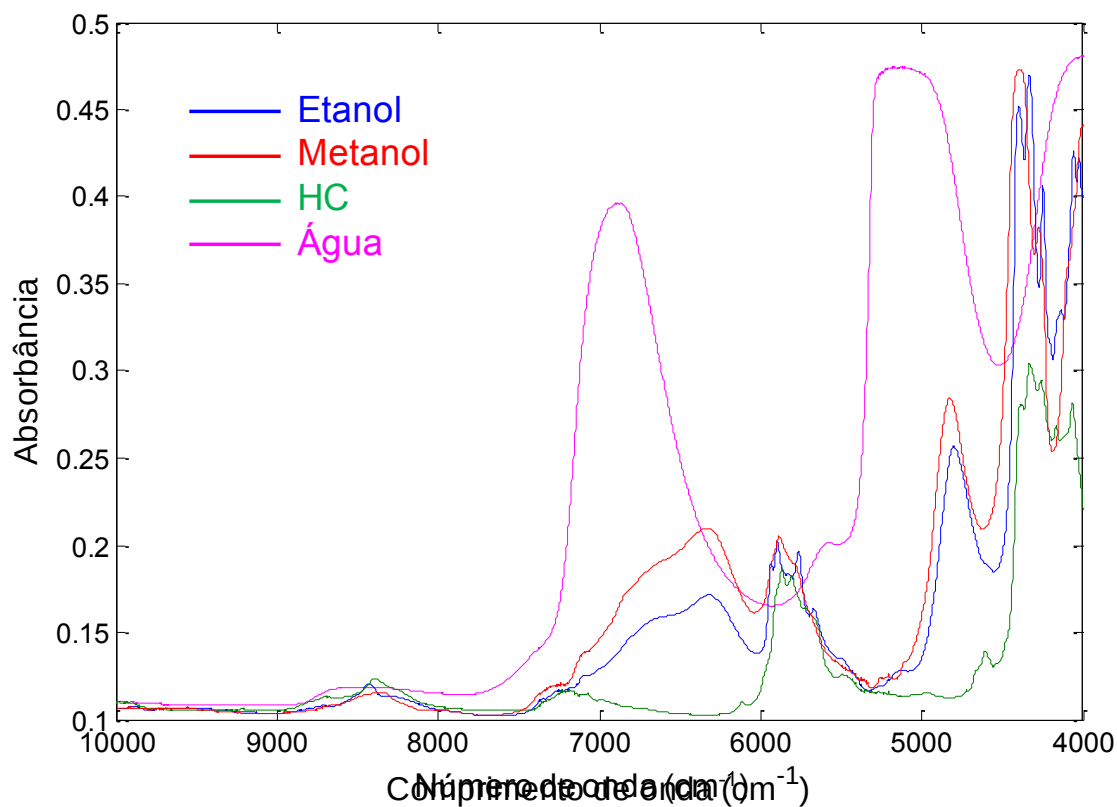


Tabela 4 – Regiões selecionadas e os erros de cada modelo na região do NIR, com utilização do caminho óptico de 0,2 mm.

Modelo	Região selecionada (cm ⁻¹)	RMSEC % (m/m)	RMSECV % (m/m)	# VL ^a	Todo o espectro		
					RMSEC % (m/m)	RMSECV % (m/m)	# LV ^a
Etanol	4000 - 5000	0,27	0,30	4	0,50	0,56	4
Metanol	4000 - 5000	0,05	0,06	5	0,10	0,20	5
HC	4000-5360	0,44	0,50	4	0,46	0,52	4
Água	4000-8900	0,03	0,03	3	0,31	0,34	4

^a Número de variáveis latentes.

Tabela 5 – Regiões seleccionadas e os erros de cada modelo na região do NIR, com utilização do caminho óptico de 0,5 mm.

Modelo	Região seleccionada (cm ⁻¹)	RMSEC % (m/m)	RMSECV % (m/m)	# VL ^a	Todo o espectro		
					RMSEC % (m/m)	RMSECV % (m/m)	# LV ^a
Etanol	4000 - 5330	0,23	0,26	4	0,36	0,44	5
Metanol	4120 - 5000	0,08	0,09	4	0,10	0,14	5
HC	4000 - 5000	0,26	0,27	3	0,23	0,29	5
Água	4000 - 9065	0,02	0,02	4	0,02	0,02	4

^a Número de variáveis latentes.

A **Figura 26** e a **Figura 27** mostram o comportamento dos valores dos erros de calibração (RMSEC) e os valores de validação cruzada (RMSECV) em função do número de variáveis latentes. Assim como nos modelos construídos com os espectros na região do MIR, o número de VL para cada modelo foi seleccionado com base na observação desses gráficos.

Figura 26 – Erros de calibração em função do número de variáveis latentes para os modelos etanol, metanol; HC e água na região do NIR utilizando o caminho óptico de 0,2 mm e as regiões selecionadas mostradas na Tabela 4. As setas indicam o número de variáveis latentes escolhidas.

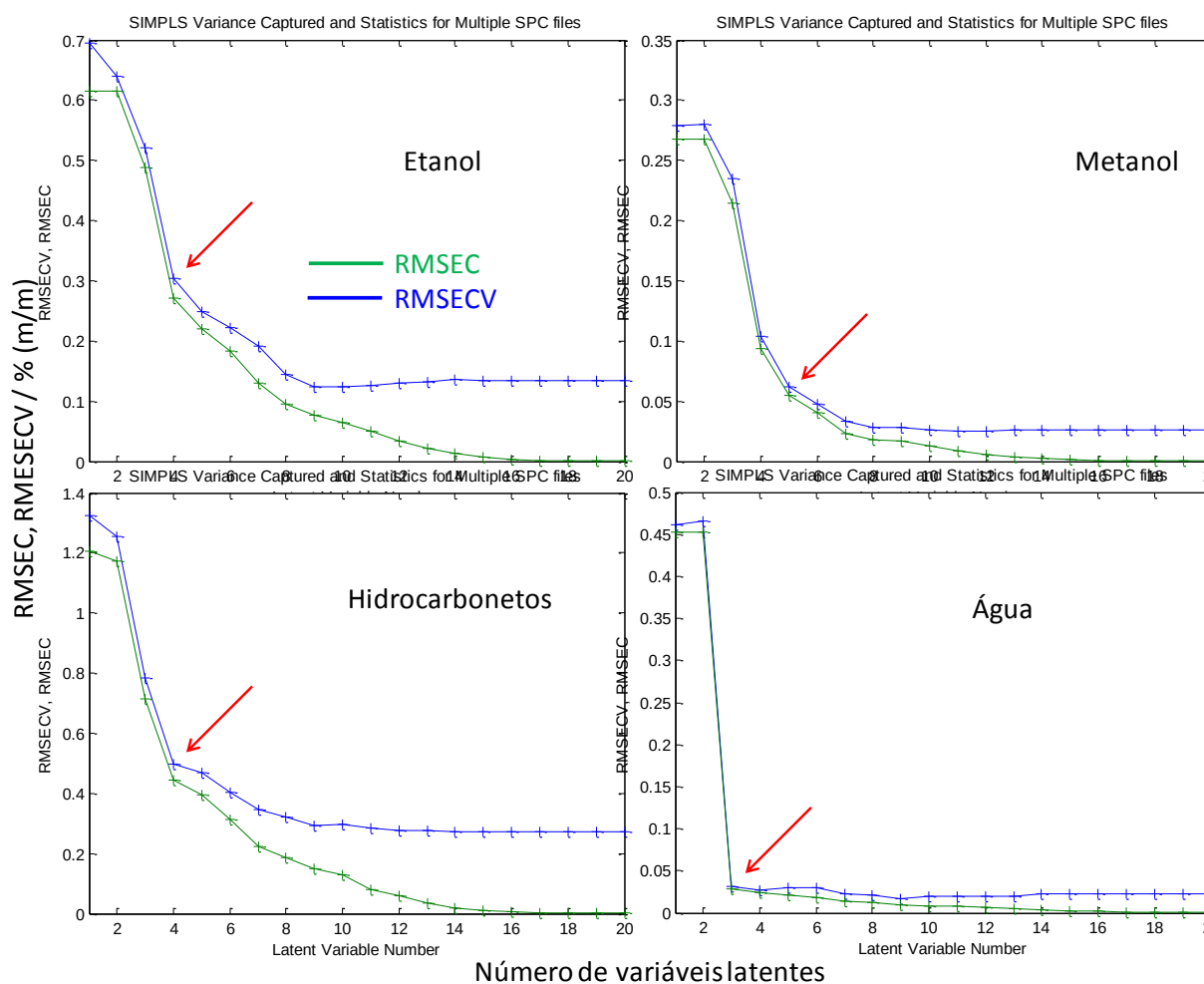
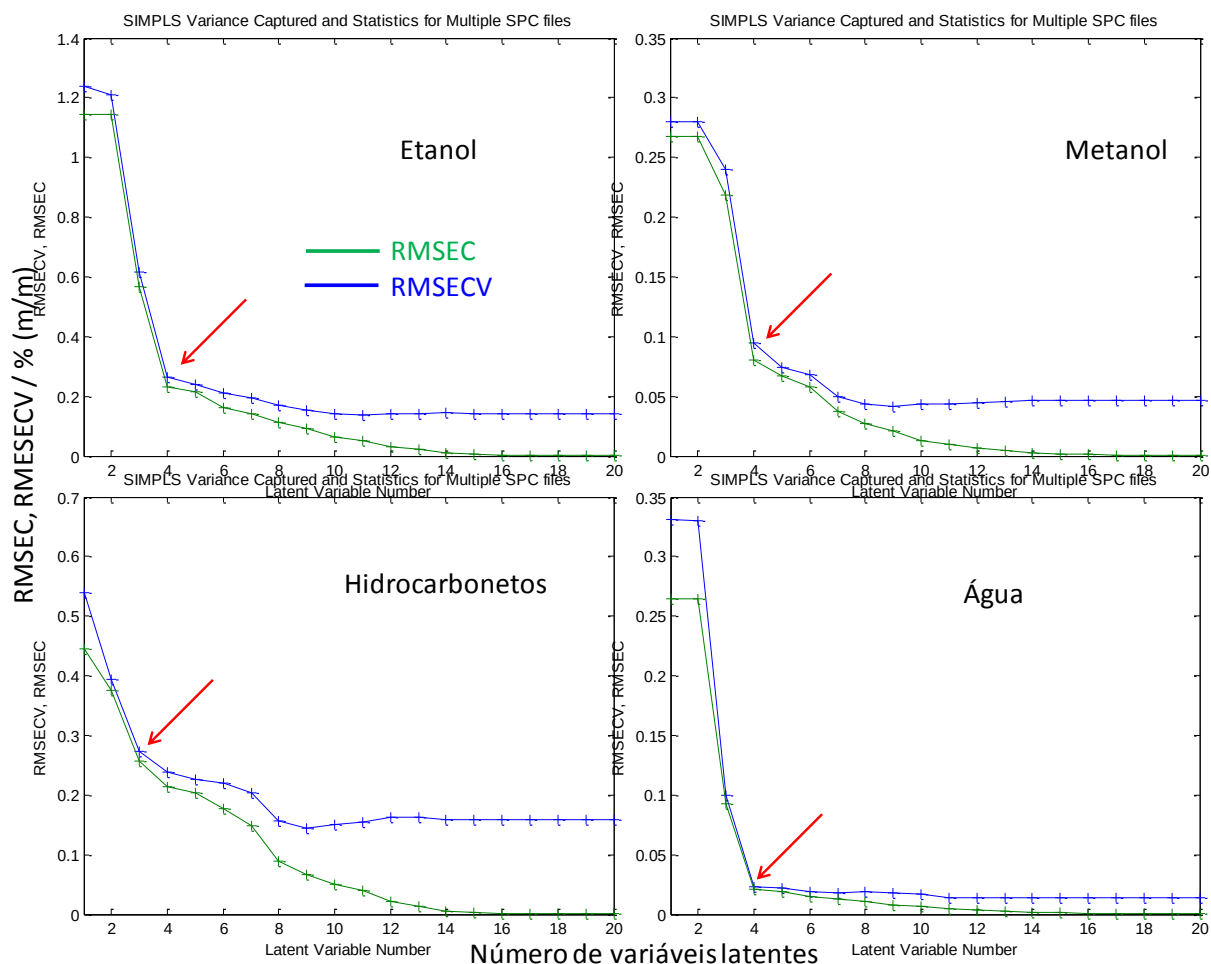


Figura 27 – Erros de calibração em função do número de variáveis latentes para os modelos etanol, metanol; HC e água na região do NIR utilizando o caminho óptico de 0,5 mm e as regiões selecionadas mostradas na Tabela 5. As setas indicam o número de variáveis latentes escolhidas



5.6. Seleção dos modelos com base nos erros

Após a construção dos modelos, era necessário selecionar o com melhor desempenho para cada propriedade de interesse para prosseguir com a validação de cada um deles. Os valores de RMSECV para os modelos são mostrados na **Tabela 6**, ordenados de forma crescente.

Tabela 6 – Valores de RMSECV de todos os modelos, ordenados em ordem crescente.

Modelo	RMSECV (% m/m)	Modelo	RMSECV (% m/m)
Etanol		HC	
NIR05	0,26	NIR05	0,27
NIR02	0,30	NIR02	0,50
MIR	0,59	MIR	0,62
Metanol		Água	
MIR	0,04	NIR05	0,02
NIR02	0,06	NIR02	0,03
NIR05	0,09	MIR	0,08

Após a construção dos modelos, foi necessário estabelecer um critério para a escolha do melhor deles para cada um dos parâmetros analisados. Para tanto, aplicou-se o teste estatístico F nos valores de RMSECV dos modelos, como mostra a Equação 32. Tal teste foi escolhido com base no trabalho publicado por Carneiro et al.⁸⁴.

$$F = \frac{RMSECV_1^2}{RMSECV_2^2} \quad (32)$$

A **Tabela 7** mostra os valores de F calculado para a comparação entre todos os modelos. O valor de F tabelado (graus de liberdade utilizado foi o número de amostras de calibração menos um) é de 1,784, portanto, com base na observação desses valores, é possível afirmar que se pode aceitar a hipótese de que a diferença entre os RMSECV não é significativa ($F_{\text{calculado}} < F_{\text{tabelado}}$), com 95% de confiança, apenas em dois casos: Etanol NIR02 e Etanol NIR05 e HC MIR e HC NIR02.

Tabela 7 – Valores de F calculados para a comparação entre os modelos e valor de F tabelado.

Modelo	Valores de F calculados			F tabelado
	MIR e NIR02	MIR e NIR05	NIR02 e NIR05	
Etanol	3,772	5,231	1,387	1,784
Metanol	2,747	6,337	2,307	
HC	1,551	5,142	3,315	
Água	7,298	13,453	1,844	

Pode-se observar que os melhores modelos para Etanol, HC e Água foram obtidos com a utilização dos espectros coletados na região do NIR com o caminho óptico de 0,5 mm (salvo para Etanol, que não apresenta diferença significativa em comparação com NIR02) e o melhor modelo para Metanol foi o obtido com os espectros na região do MIR. Com base nos valores apresentados anteriormente, prosseguiu-se com a validação dos modelos que apresentaram melhores resultados.

5.7. Validação

Após a seleção dos melhores modelos, seguiu-se com a validação de cada um deles através do cálculo das figuras de mérito (FOM). A **Tabela 8** mostra as principais FOM para cada um dos melhores modelos. Os valores dos erros (RMSEC, RMSECV e RMSEP) dos modelos foram adotados como a exatidão.

Tabela 8 – Figuras de mérito dos quatro modelos construídos para previsão dos teores de etanol, metanol, HC e água em amostras de E85.

FOM	Parâmetro	Etanol NIR05	Metanol MIR	HC NIR05	Água NIR05
Exatidão	RMSEC ^a	0,2	0,02	0,3	0,02
	RMSECV ^a	0,3	0,03	0,3	0,02
	RMSEP ^a	0,3	0,02	0,2	0,02
Precisão	RSD ^c repetibilidade (%)	0,2	0,03	0,2	0,03
Ajuste	Coeficiente de correlação	0,9996	0,9931	0,9995	0,9982
	Inclinação	0,9986	0,9931	0,9993	0,9926
	Intercepto	0,0814	0,0023	0,0219	0,0088
Intervalo de medição ^a		51,6–95,6	0,24–0,95	4,8–46,4	0,22–1,76
Seletividade		0,0086	0,00091	0,015	0,0055
Sensibilidade ^b		0,072	0,00044	0,012	0,060
Sensibilidade analítica (γ) ^a		56,6	18,03	60,3	39,68
γ^{-1} ^b		0,02	0,06	0,02	0,03
Limite de detecção ^a		1,0	0,10	1,1	0,09
Limite de quantificação ^a		-	0,24	-	0,22
Razão Sinal/ruído	Máximo	447	126	239	219
	Mínimo	245	65	29	77
Bias	$t_{\text{calculado}}$	0,745	1,691	0,971	0,338
	$t_{\text{crítico}}$	2,048	2,042	2,042	2,042

^a % (m/m)

^b % (m/m)⁻¹

^c Desvio padrão relativo

A precisão, ao nível de repetibilidade, foi calculada (de acordo com a equação 10) analisando três amostras em diferentes concentrações dos analitos (**Tabela 9**).

Todos os valores dos desvios padrões relativos ficaram abaixo de 0,3%, mostrando boa repetibilidade do método. Quanto à exatidão, os valores relativos entre os maiores valores dos erros médios quadráticos dos modelos e os menores valores do intervalo de medição são: 0,50% para o teor de etanol; 25% para o teor de metanol, 5,6% para o teor de HC e 9,1% para o teor de água.

Tabela 9 – Replicatas de três amostras de E85 para determinação da precisão em nível de repetibilidade para os quatro modelos.

Amostra	Teor de etanol		Teor de Metanol		Teor de HC		Teor de água	
	Média % (m/m)	RSD ^a %	Média % (m/m)	RSD ^a %	Média % (m/m)	RSD ^a %	Média % (m/m)	RSD ^a %
A	53,9		0,50		45,2		0,25	
B	69,8	0,2	0,22	0,03	27,9	0,2	1,79	0,03
C	83,3		0,75		13,5		1,05	

^a Desvio padrão relativo

A repetibilidade para o teor de etanol segundo a ASTM D5501 e a EN1601 é 0,15% (m/m) e 0,2% (m/m), respectivamente. O método proposto para a quantificação de etanol possui um valor de repetibilidade maior ou igual ao estabelecido na norma europeia. Para o metanol, essas mesmas normas estabelecem o limite de repetibilidade igual a 0,02% (m/m) e 0,05% (m/m), respectivamente. O valor de repetibilidade para o método proposto é maior que o limite estabelecido pela norma americana, porém é menor que o estabelecido pela norma europeia. Já para o teor de água, segundo as normas ASTM E1064 e EN15692, o limite de repetibilidade é respectivamente 0,031% e 0,014%. Assim como para a quantificação do etanol, a repetibilidade para a quantificação do teor de água é maior ou igual aos limites estabelecidos nas normas.

5.7.2. Sensibilidade

Sensibilidade pode ser definida como a fração do sinal analítico responsável pelo incremento do valor da concentração dos analitos. Em calibrações multivariadas inversas, como no caso da PLS, a sensibilidade pode ser calculada como o inverso do vetor dos coeficientes de regressão. Por outro lado, a sensibilidade analítica (γ) é uma figura de mérito que permite avaliar de maneira mais simples a sensibilidade de um método analítico. Essa FOM é calculada usando a relação entre a sensibilidade e a razão sinal/ruído. A partir do inverso dessa FOM (γ^{-1}), é possível obter a menor diferença de concentração que o método pode discernir^{65,69,85,86}. Os resultados obtidos mostram que os modelos para a previsão de concentração de etanol e hidrocarbonetos pode diferenciar amostras com 0,02% (m/m) de diferença na concentração, 0,06% (m/m) de diferença para o modelo de metanol e 0,03% (m/m) de diferença para o modelo de previsão do teor de água.

Através da observação dos valores mostrados na **Tabela 10**, é possível afirmar que os modelos Etanol NIR05 e HC NIR05 possuem valores maiores que 0,02% (m/m) para as concentrações entre as amostras, portanto esses modelos foram capazes de diferenciá-las, dentro dos limites de exatidão e precisão. Já os modelos Metanol MIR e Água NIR05, que apresentam uma faixa de medição menor, possuem amostras com diferenças menores que γ^{-1} e, por isso foi possível observar a capacidade dos modelos em diferenciá-las, dentro dos limites da exatidão e precisão (**Tabela 11**).

Tabela 10 – Comparação da diferença entre algumas amostras do conjunto de previsão para avaliação da sensibilidade analítica dos modelos Etanol NIR05 e HC NIR05.

Etanol NIR05				HC NIR05			
Referência	Diferença	Previsto	Diferença	Referência	Diferença	Previsto	Diferença
	% (m/m)				%(m/m)		
53,45	0,33	53,05	0,74	14,89	0,16	14,74	0,27
53,78		53,80		15,05		15,00	
58,49	0,24	58,39	0,26	14,89	-0,16	14,74	-0,27
58,73		58,65		15,05		15,01	
72,01	0,17	71,90	0,27	32,77	-0,68	32,96	-0,38
72,18		72,17		33,45		33,33	
87,51	0,39	87,51	0,12	43,56	-0,61	43,67	-0,57
87,90		87,63		44,17		44,24	

Tabela 11 – Comparação da diferença entre algumas amostras do conjunto de previsão para avaliação da sensibilidade analítica dos modelos Metanol MIR e Água NIR05.

Metanol MIR				Água NIR05			
Referência	Diferença	Previsto	Diferença	Referência	Diferença	Previsto	Diferença
	% (m/m)				%(m/m)		
0,16	0,03	0,21	0,01	0,00	0,09	-0,02	0,10
0,19		0,20		0,09		0,09	
0,30	0,00	0,27	0,03	0,35	0,01	0,36	0,02
0,30		0,30		0,36		0,38	
0,65	0,02	0,62	0,01	0,72	0,02	0,73	0,03
0,67		0,61		0,74		0,76	
0,86	0,03	0,85	0,00	1,43	0,02	1,41	0,03
0,89		0,85		1,45		1,44	

5.7.3. Seletividade

A seletividade, como foi calculada, indica a parte do sinal que se refere exclusivamente ao analito de interesse ^{69,86}. Para exemplificar com os valores mostrados na Tabela 2: para o etanol, 0,86% (valor médio) do sinal foi usado para a calibração do valor do teor desse analito no combustível. Já para os demais analitos: metanol, HC e água, respectivamente 0,09%, 1,50% e 0,55% (valores médios) do sinal foram utilizados por cada modelo na calibração do teor dos mesmos.

5.7.4. Ajuste

O ajuste do modelo foi avaliado utilizando o coeficiente de determinação (R^2) obtido da regressão linear dos pontos formados pelos valores de referência e os valores previstos pelos modelos (**Figura 28**). Todos os valores dos coeficientes foram próximos à unidade (o menor valor foi 0,9931), mostrando a concordância entre os valores de referência e os previstos pelos modelos. Além disso, a **Figura 29** mostra que a região de confiança elíptica conjunta (EJCR – do inglês *Elliptical Joint Confidence Region*) ⁸⁵ contém os valores ideais da inclinação e do intercepto (1 e 0, respectivamente) das retas de regressão linear da **Figura 28**. Avaliou-se essa FOM também através dos modelos na forma pseudounivariada (**Figura 30**), utilizando os valores de referência em função dos valores de NAS. Os parâmetros dessas curvas são mostrados na **Tabela 12**.

Figura 28 – Comportamento linear dos valores previstos pelos modelos em função dos valores de referência. Amostras de calibração (○) e amostras de validação (+).

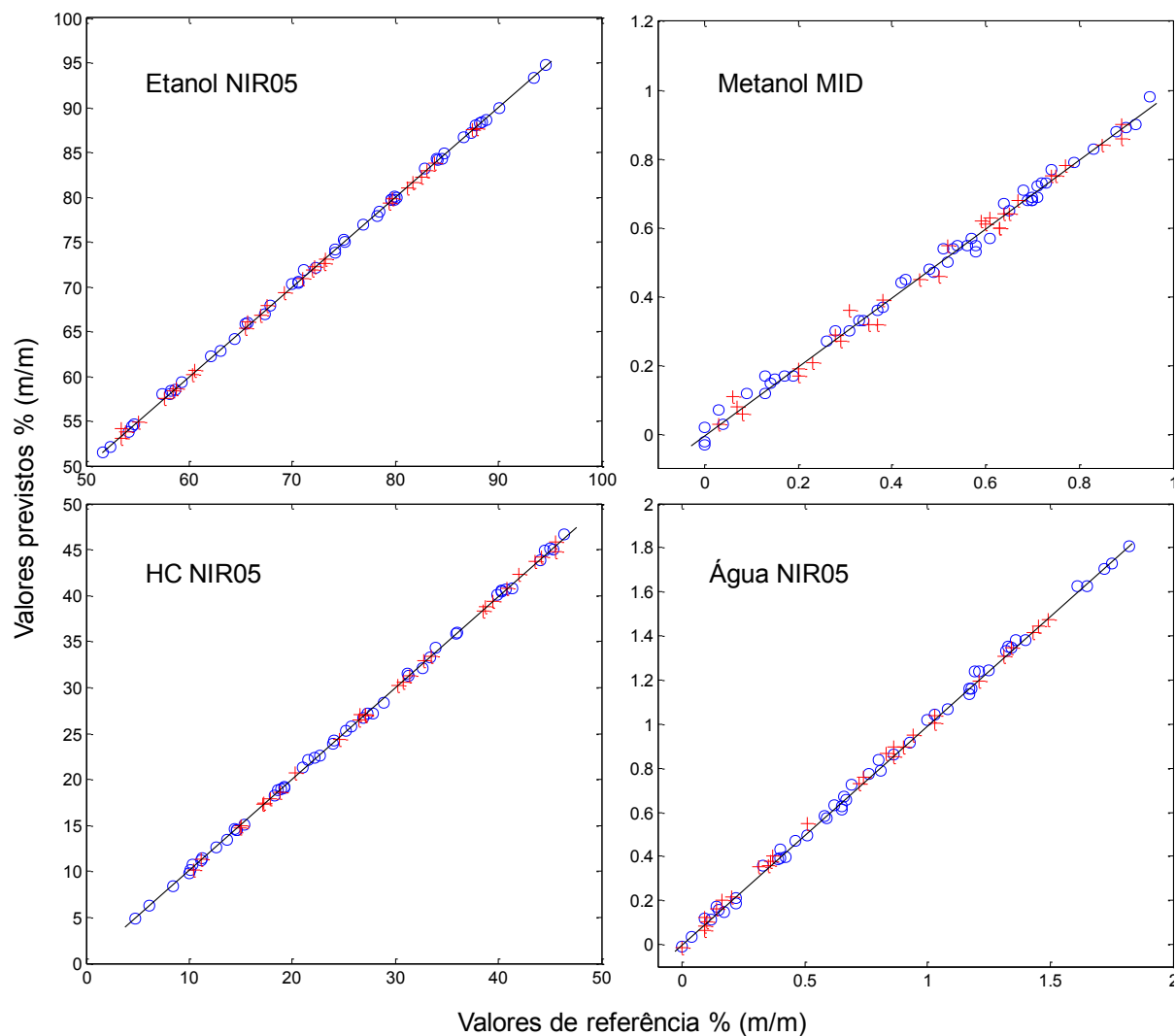


Figura 29 – Elipses de confiança para a inclinação e o intercepto das regressões lineares obtidas através dos valores previstos pelos modelos e os valores de referência.

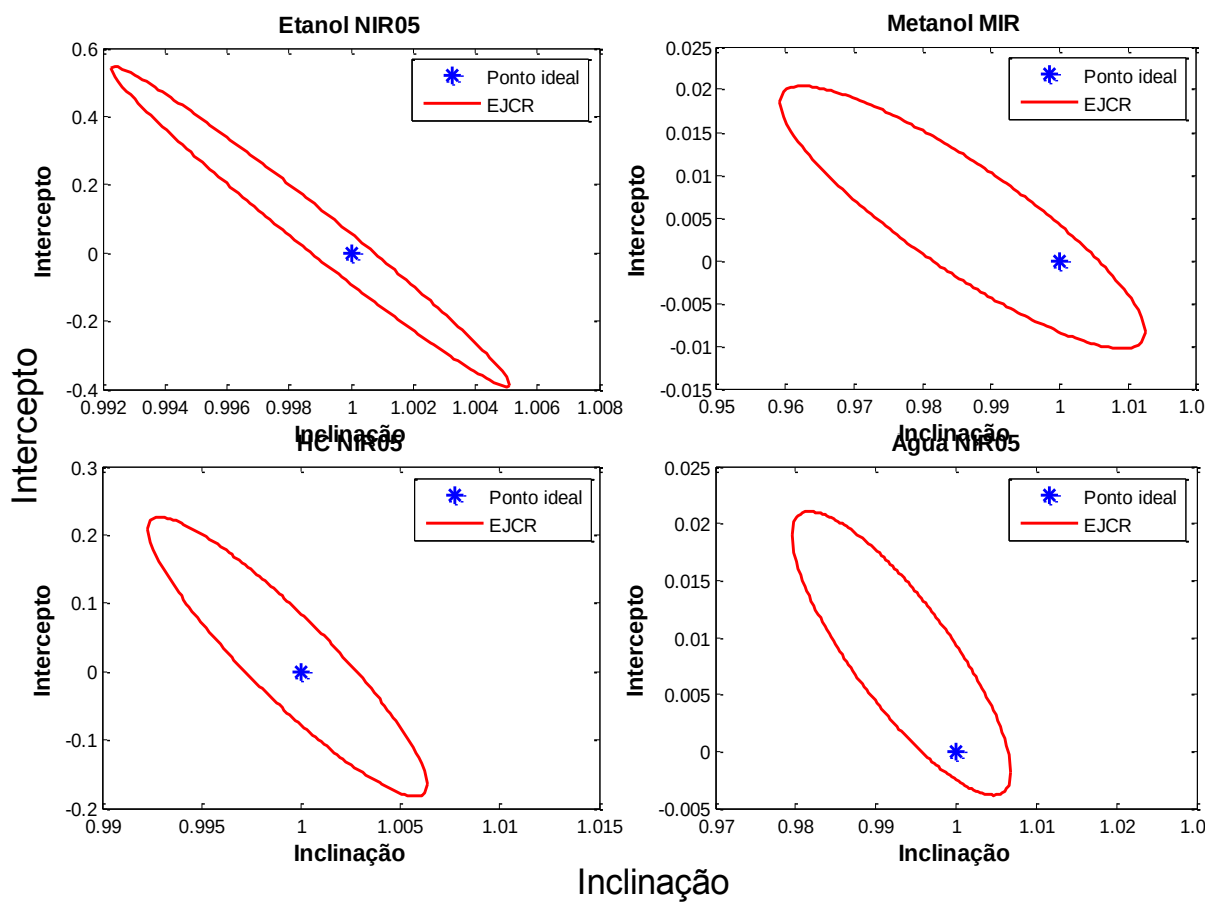
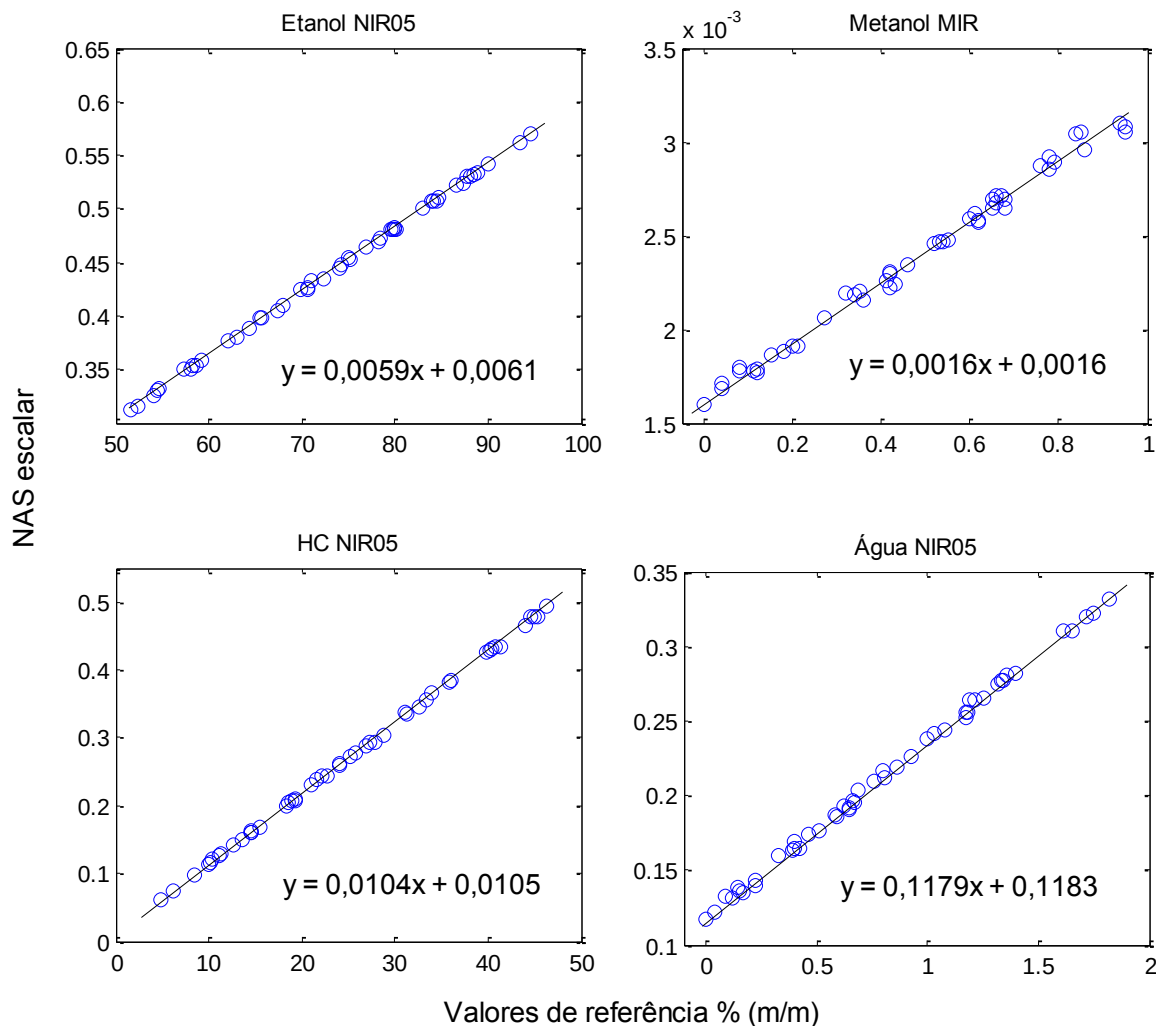


Tabela 12 – Parâmetros das regressões lineares para os modelos pseudounivariados.

	Inclinação	Intercepto	R ²
Etanol NIR05	168,03	-1	0,9996
Metanol MIR	620,99	-1	0,9927
HC NIR05	96,25	-1	0,9995
Água NIR05	8,47	-1	0,9983

Figura 30 – Valores de referência em função dos valores de NAS escalar para os quatro modelos.



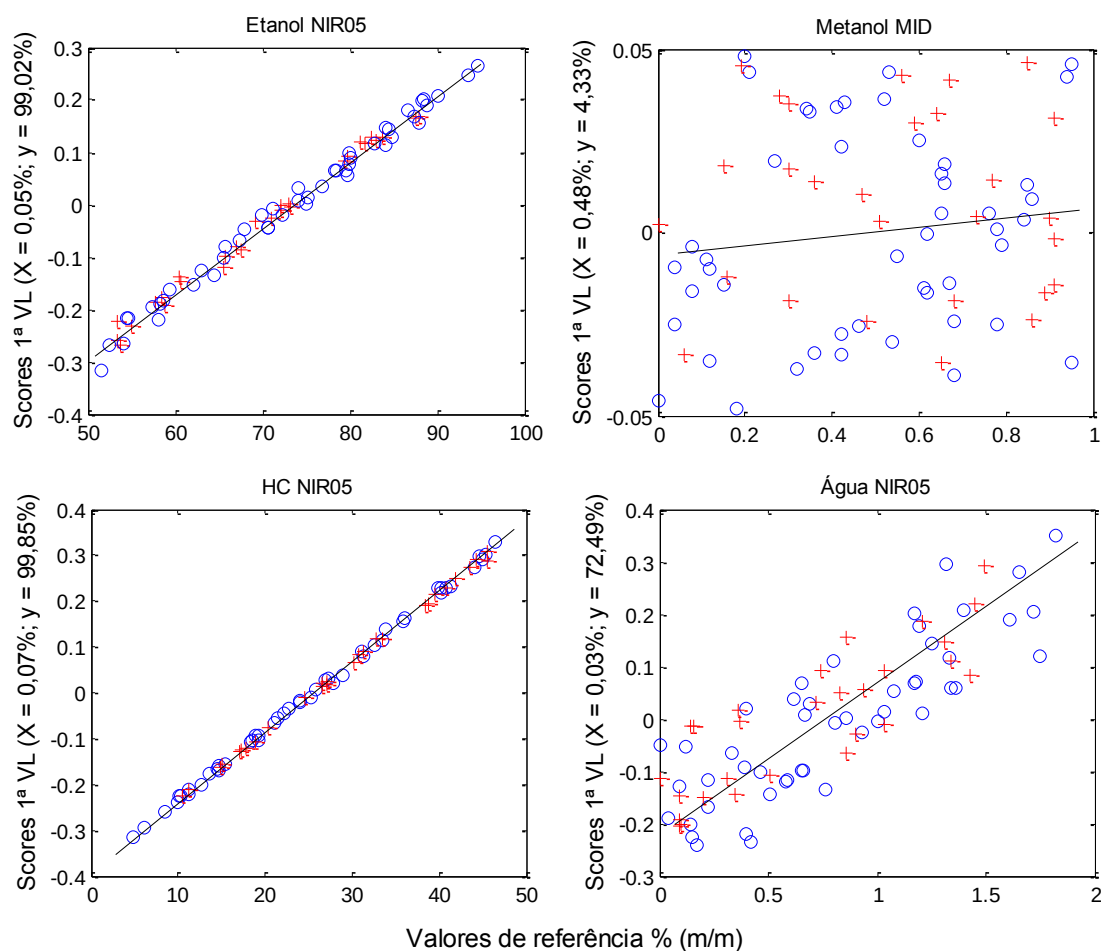
5.7.5. Linearidade

Como os modelos de regressão multivariada não são curvas analíticas tradicionais foi necessário adotar outra forma de avaliação da linearidade, sendo feito de forma qualitativa, ou seja, através da avaliação do comportamento dos valores dos *scores* da primeira variável latente em função dos valores de referência (**Figura 31**). É possível observar a relação linear dos valores dos *scores* em função dos valores de referência para todos os modelos, com exceção do Metanol MIR. Isso não implica na não linearidade desse modelo, pois a primeira variável latente explica apenas 4,33% da variância do vetor $\mathbf{y}$ e o modelo utiliza quatro variáveis latentes. Levando esse fato em

consideração, construiu-se outro gráfico com os valores dos *scores* da terceira variável latente, que explica 92,34% da variância em *y*. A partir desse novo gráfico obtiveram-se os seguintes parâmetros da regressão linear dos pontos: intercepto = -0,0008; inclinação = 0,0017 e coeficiente de determinação = 0,9253.

A linearidade também foi avaliada pela observação da distribuição dos gráficos de resíduos, o qual apresenta um comportamento randômico para os modelos lineares ⁶⁰. A **Figura 32** mostra que todos os modelos apresentam distribuição com característica randômica dos erros.

Figura 31 – Valores dos *scores* da primeira variável latente em função dos valores de referência. Amostras de calibração (○) e amostras de validação (+).

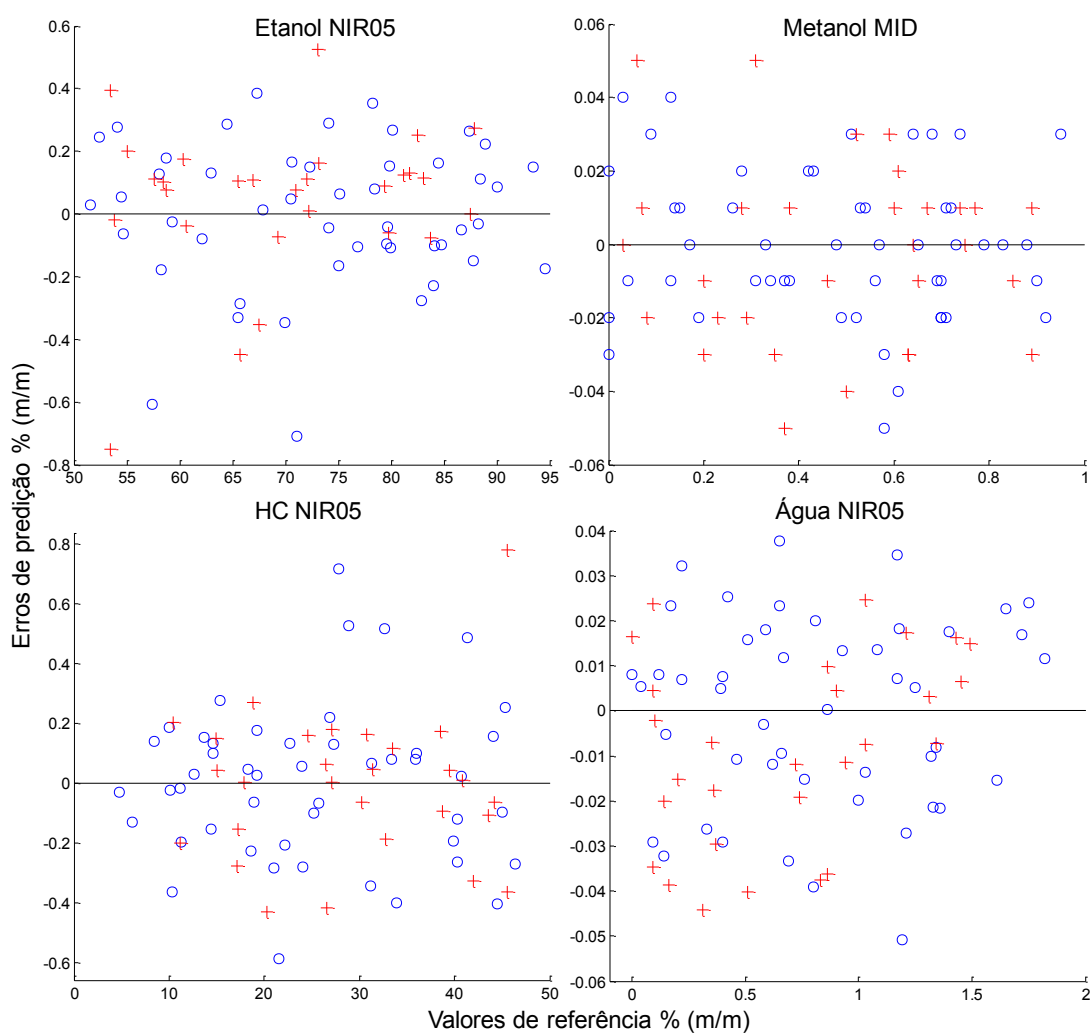


A **Tabela 13** mostra os parâmetros das regressões dos dados apresentados na **Figura 31**.

Tabela 13 – Parâmetros das regressões lineares dos valores dos *scores* da primeira variável latente em função dos valores de referência.

	Inclinação	Intercepto	R ²
Etanol NIR05	0,013	-0,92	0,9902
Metanol MIR	0,013	-0,004	0,0185
HC NIR05	0,015	-0,38	0,9989
Água NIR05	0,24	-0,18	0,711

Figura 32 – Comportamento aleatório dos valores dos erros dos modelos em função dos valores de referência. Amostras de calibração (o) e amostras de validação (+).



5.7.6. Razão sinal/ruído

A razão sinal/ruído, da forma que foi calculada, indica o quanto o valor de NAS escalar está acima do valor do ruído analítico. O modelo que possui o menor valor da razão sinal/ruído é o modelo HC NIR05, o qual é igual a 29, já o que possui o maior valor é o modelo Etanol NIR05, o qual é igual a 447. Portanto, todos os modelos apresentam sinais com pelo menos uma ordem de grandeza maior que o ruído.

5.7.7. Intervalo de medição ou faixa de medição

Apesar de ser possível obter valores numéricos para o LQ para os modelos etanol e hidrocarbonetos e esses serem respectivamente 2,36% e 2,66% (m/m), o menor valor adotado para a faixa de medição desses modelos foi o menor valor da concentração utilizada na calibração dos modelos. Foram adotados tais valores, pois não há evidência experimental que os modelos Etanol e HC sejam lineares até a faixa de concentração próxima ao LQ calculado.

5.7.8. Limites de detecção e quantificação

Os valores de LD foram calculados conforme a Equação 15. De acordo com Felipe-Sotelo et. al.⁷², quando se tem um número elevado de graus de liberdade (maior ou igual a 25), $\alpha = \beta$, $\Delta(\alpha, \beta, v)$ pode ser aproximado de $2 \cdot t_{1-\alpha, v}$. Sendo assim, considerou-se $\alpha = \beta = 5\%$.

Os valores obtidos para os limites de detecção e quantificação mostram que o método é aplicável para a quantificação dos teores propostos nesse trabalho, uma vez que estão abaixo dos limites estabelecidos nas ASTM D5798 e na CEN/TS 15293.

5.7.9. Teste para erros sistemáticos (BIAS)

A FOM Bias foi avaliada de acordo com o teste t, como determina a norma ASTM E1655⁶³. Todos os valores de t_{bias} foram menores que os valores de t crítico, tais resultados mostram que, estatisticamente, não há erros sistemáticos nos modelos, ou seja, os valores previstos e os valores de referência são estatisticamente concordantes.

5.7.10. Robustez

A robustez foi avaliada variando-se a temperatura ambiente de 18 °C a 24 °C, pois esta é a faixa de temperatura em que o equipamento deve ser operado, segundo dados do fabricante. Após a alteração de temperatura aguardou-se pelo menos uma hora para a realização das leituras dos espectros das amostras. Os resultados são mostrados nas quatro primeiras colunas da **Tabela 14**. A partir desses resultados, foi aplicada, com auxílio do Microsoft Excel, a análise de variância (ANOVA)⁸⁸ com um fator para determinar se há diferença entre as médias nas diferentes temperaturas.

Tabela 14 – Valores das concentrações obtidas em diferentes temperaturas no estudo de robustez.

Parâmetro	21 °C	18 °C	24 °C
	54,0	54,2	54,2
Etanol % (m/m)	54,0	53,8	54,4
	54,0	54,3	54,0
	0,58	0,55	0,58
Metanol % (m/m)	0,56	0,57	0,58
	0,56	0,57	0,59
	45,1	45,0	45,1
HC % (m/m)	45,1	45,2	45,0
	45,1	45,0	45,0
	0,27	0,26	0,27
Água % (m/m)	0,25	0,25	0,25
	0,26	0,26	0,27

A **Tabela 15** mostra os valores obtidos pela ANOVA em que todos os valores de F são menores que F crítico. Portanto, aceita-se a hipótese nula que todas as médias são estatisticamente iguais, ou seja, dentro dos limites de variação de temperatura estudados, a temperatura não influencia significativamente (95% de confiança) nos valores de concentração dos analitos.

Tabela 15 – Resultados da ANOVA com um fator para avaliar o efeito da variação de temperatura

Parâmetro	Fonte da variação	SQ	gl	MQ	F	valor-P	F crítico
Etanol	Entre grupos	0,062	2	0,031	0,825	0,482	5,143
	Dentro dos grupos	0,226	6	0,038			
Metanol	Entre grupos	6,74E-4	2	0,00034	2,559	0,1574	
	Dentro dos grupos	7,90E-4	6	0,00013			
HC	Entre grupos	0,013	2	0,0063	1,283	0,344	
	Dentro dos grupos	0,030	6	0,0049			
Água	Entre grupos	4,05E-05	2	2,0E-05	0,240	0,794	
	Dentro dos grupos	5,06E-4	6	8,4E-05			

gl = graus de liberdade

SQ = soma dos quadrados

MQ = média dos quadrados

5.8. Incerteza de medição

Inicialmente, aplicou-se o teste de Shapiro – Wilks para verificar se os resultados dos modelos seguiam a distribuição normal para uma amostra com aproximadamente 84% (m/m) de etanol, 05% (m/m) de metanol, 15% (m/m) de HC e 0,5% (m/m) de água, quantificada 10 vezes (**Tabela 16**). A **Tabela 17** mostra os valores calculados de W do teste de Shapiro – Wilks e seus respectivos valores críticos W_c , para uma probabilidade de 95%. É possível observar que todos os valores de W são maiores do que os valores críticos W_c , sendo assim, podemos inferir que a distribuição dos valores das concentrações previstas pelos quatro modelos seguem a distribuição normal.

Tabela 16 – Valores das concentrações dos analitos, obtidos através de dez réplicas nos quatro modelos para uma mesma amostra.

Réplica	Etanol / % (m/m)	Metanol / % (m/m)	HC / % (m/m)	Água / % (m/m)
01	82,8	0,72	15,8	0,45
02	83,1	0,66	15,5	0,47
03	83,4	0,65	15,3	0,49
04	83,3	0,67	15,4	0,48
05	83,5	0,64	15,2	0,49
06	83,3	0,73	15,3	0,48
07	83,3	0,72	15,3	0,47
08	83,3	0,69	15,3	0,48
09	83,5	0,68	15,2	0,46
10	83,7	0,68	14,9	0,49

Tabela 17 – Valores calculados W e valores críticos W_c da estatística do teste de Shapiro – Wilks para uma amostra quantificada nos quatro modelos.

	Etanol NIR05	Metanol MIR	HC NIR05	Água NIR05
W	0,932	0,983	0,924	0,936
W_c	0,781			

Sendo assim, consideraram-se as distribuições dos modelos como sendo normais e seguiu-se com o cálculo da incerteza de medição para os modelos.

5.8.1. Definição do mensurando

A quantificação dos componentes depende de vários fatores que contribuem para a determinação da incerteza de medição e é fundamental fazer uma análise sobre eles tendo em vista que as equações de cada um deles auxiliarão na determinação da

incerteza de medição. Inicialmente, elaborou-se a equação (Equação 33) do mensurando.

$$y = \mathbf{x}\mathbf{b} \quad (33)$$

onde:

y = concentração de cada componente (etanol, metanol, hidrocarbonetos ou água), em % (m/m);

$\mathbf{x}$ = vetor com o espectro da amostra analisada;

$\mathbf{b}$ = vetor de regressão, obtido pelo PLS.

Para viabilizar os cálculos a partir das definições do ISO GUM ⁷⁶, utilizou-se o modelo na forma pseudounivariada, dessa forma a Equação 33 do mensurando pode ser reescrita:

$$x = \frac{y - b_0}{b_1} (\% \text{ m/m}) \quad (34)$$

onde:

y = valor escalar de NAS para a amostra analisada;

b_0 = coeficiente linear da curva de regressão da forma pseudounivariada de cada um dos modelos (**Figura 30**);

b_1 = coeficiente angular da curva de regressão da forma pseudounivariada de cada um dos modelos (**Figura 30**);

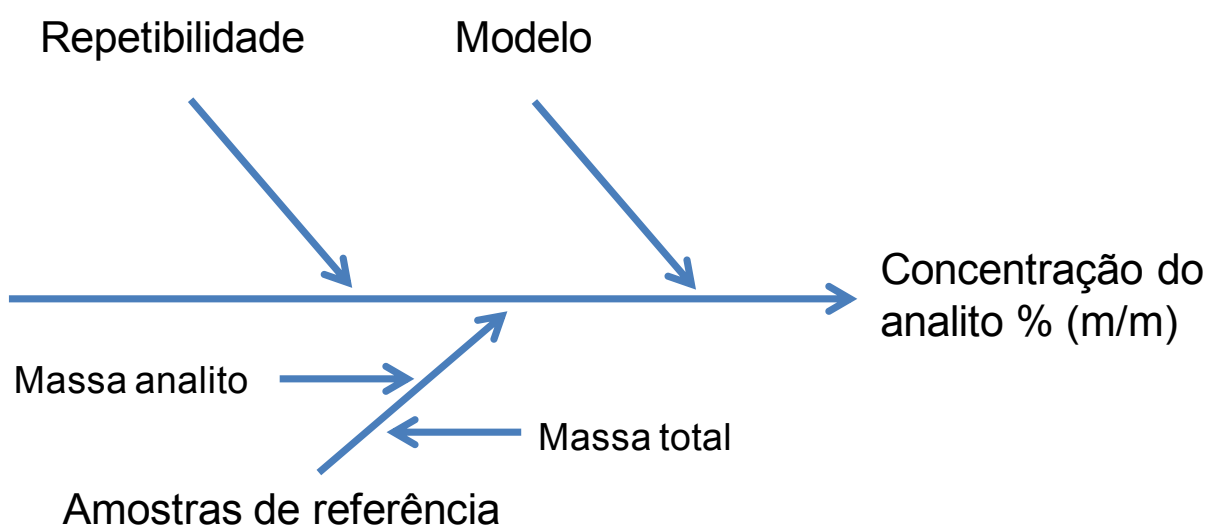
x = concentração de cada componente (etanol, metanol, hidrocarbonetos ou água), em % (m/m).

Os gráficos dos quatro modelos em sua forma pseudounivariada podem ser observados na **Figura 30**.

5.8.2. Diagrama de causa e efeito

O diagrama de causa e efeito (**Figura 32**) para a estimativa da incerteza de medição da concentração dos analitos pelos métodos propostos mostra quais são as fontes levantadas que afetam diretamente o resultado da medição (estão apontando diretamente para o eixo horizontal).

Figura 32: Diagrama de causa e efeito mostrando as fontes de incerteza na determinação da concentração dos analitos nas amostras.



5.8.3. Avaliação das incertezas padrão

A incerteza referente à repetibilidade dos métodos e a incerteza do modelo são do tipo A, já a incerteza referente às amostras de referência são provenientes do certificado de calibração da balança utilizada para o preparo dessas amostras (tipo B).

Para o cálculo da incerteza da curva de regressão linear para o modelo na forma pseudounivariada, considerou-se que para cada valor de x , os valores de y correspondentes são normalmente distribuídos e que cada amostra do conjunto de calibração foi ensaiada uma única vez. Para obtenção da incerteza referente ao modelo, aplicou-se as seguintes equações (35 e 36):

$$u_{curva} = \frac{S_e}{b_1} \cdot \sqrt{\frac{1}{n_t} + \frac{(\bar{y}_o - \bar{y})^2}{b_1^2 \sum (x_i - \bar{x})^2}} \quad (35)$$

$$S_e = \sqrt{\frac{\sum_{i=0}^n (y_i - \hat{y}_i)^2}{n - 2}} \quad (36)$$

onde:

S_e é o erro padrão da estimativa;

b_0 é o coeficiente linear da curva dos modelos na forma pseudounivariada;

b_1 é o coeficiente angular da curva dos modelos na forma pseudounivariada;

n_t é o número de amostras utilizadas na construção da curva;

$\bar{y}_0$ é a média dos valor de NAS escalar para uma dada amostra;

$\bar{y}$ é a média dos valores de NAS escalar das amostras utilizadas para a construção da curva;

$\bar{x}$ é o valor médio dos valores das concentrações de cada analito nas amostras utilizadas para a construção da curva;

x_i é a concentração do analito de interesse na amostras i;

y_i é o valor de NAS escalar para a amostra i;

$\hat{y}_i$ é o valor previsto de NAS escalar para a amostra i.

Aplicando a equação 35 e 36 para valores de NAS escalar correspondentes a 84,2% (m/m) para etanol, 0,43% (m/m) para metanol, 15,4% (m/m) para HC e 0,51% (m/m) para água, obteve-se os seguintes valores:

$$u_{\text{curva.etanol}} = 0,04758 \% \text{ (m/m)};$$

$$u_{\text{curva.metanol}} = 0,00428\% \text{ (m/m)};$$

$$u_{\text{curva.HC}} = 0,04956\% \text{ (m/m)};$$

$$u_{\text{curva.água}} = 0,00373\% \text{ (m/m)}.$$

Para a estimativa da incerteza padrão para a repetibilidade, utilizaram-se os dados da **Tabela 16** e a seguinte equação:

$$u_{repe} = \frac{s}{\sqrt{n}} \% (m/m) \quad (37)$$

onde s é o desvio padrão de medidas realizadas em condição de repetibilidade e n é o número de vezes que essas medidas foram realizadas. Sendo assim:

$$u_{repe. etanol} = 0,07635 \% (m/m);$$

$$u_{repe. metanol} = 0,00980 \% (m/m);$$

$$u_{repe. HC} = 0,07690 \% (m/m);$$

$$u_{repe. água} = 0,00427 \% (m/m).$$

A concentração nos analitos é obtida pela divisão da massa de cada analito pela massa total da amostra, ou seja:

$$teor\ analito \% (m/m) = 100 \frac{m_{analito}}{m_{total}} \quad (38)$$

Para cada medida de massa realizada a balança era tarada, então os valores das fontes de incerteza foram multiplicados por dois e os valores da incerteza padrão das massas são:

$$u_{massa. analito} = 0,00064 \text{ g (Quadro 1);}$$

$u_{massa. total} = 4 \times 0,00064 \text{ g} = 0,00052 \text{ g}$ (como são quatro componentes, aplica-se o multiplicador quatro).

5.8.4. Coeficientes de sensibilidade

O próximo passo foi determinar os coeficientes de sensibilidade. Para repetibilidade e para a curva aplicou-se o coeficiente um, já que cada unidade alterada nessa fonte altera uma unidade na incerteza padrão.

Sabendo que o teor do analito na amostra é dado pela equação 38 e a massa total da amostra preparada foi igual a 10 g, então o coeficiente de sensibilidade adotado para a incerteza da massa do analito para as amostras de referência foi:

$$\frac{\partial teor}{\partial m_{analito}} = \frac{100}{m_{total}} = \frac{100}{10 \text{ g}} = 10 \frac{1}{\text{g}}$$

para as o etanol, mediu-se uma massa de 8,5 g, logo a incerteza da massa total das amostras de referência para esse analito foi:

$$\frac{\partial teor}{\partial m_{total}} = \frac{m_{analito} 100}{m_{total}^2} = \frac{8,5 \text{ g} \cdot 100}{(10 \text{ g})^2} = 8,5 \frac{1}{\text{g}}$$

Para metanol, HC e água as massas foram iguais a 0,5 g, 1,5 g e 0,5 g, então os coeficientes foram respectivamente 0,5 g⁻¹; 1,5 g⁻¹; e 0,5 g⁻¹.

5.8.5. Componentes de incerteza

As componentes são mostradas nos Quadros **2**, **3**, **4**, e **5**. Já a **Figura 33** mostra os gráficos de barras, que apresentam os valores percentuais das fontes que compõem a incerteza de medição dos quatro mensurandos. É possível observar que para Etanol, Metanol e HC as fontes com maior contribuição é a devido à repetibilidade. Isso pode estar relacionado com o fato de a gasolina possuir os componentes mais voláteis do combustível e, portanto, afetar a repetibilidade do método, pois o procedimento de aplicação da amostra no cristal é manual e a mesma pode ficar exposta ao ar por períodos diferentes de tempo em cada ensaio, causando a evaporação dos hidrocarbonetos mais leves e do metanol e alterando suas concentrações. Para diminuir os valores de incerteza, poder-se-ia avaliar a implementação de um sistema automatizado para aplicação das amostras que diminua o efeito da evaporação dos analitos.

Já para Água, não se observa o mesmo comportamento (provavelmente porque a água é menos volátil que os outros analitos). Nesse caso, o maior valor percentual de contribuição foi devido à massa do analito, portanto, para diminuir o valor da incerteza desse mensurando, deve-se diminuir os valores da incerteza que provém da balança.

Quadro 1 – Planilha do cálculo de incerteza expandida das medidas das massas das amostras de referência.

Símbolo	Fontes de incerteza			Distribuição de probabilidade	Divisor	Coeficiente de sensibilidade c_i		Incerteza padrão u_i [mL]	Graus de liberdade n_i	Porcentagem de contribuição
	Nome	Valor (+/-)	Unidade			Valor	Unidade			
$u_{res.Bal.}$	Resolução da balança	0,00010	g	retangular	1,73	1,0000	-	0,00006	infinito	3,26
$u_{cal.Bal.}$	Incerteza de calibração da balança	0,00019	g	normal (k)	2,00	1,0000	-	0,00009	infinito	8,47
$u_{excent.}$	Incerteza da excentricidade da balança	0,00052	g	retangular	1,73	1,0000	-	0,00030	infinito	88,27
Incerteza padrão combinada (u_c)		Distribuição de probabilidade		Graus de liberdade (n_{eff})		Fator de abrangência (k)		Incerteza expandida (U) (Probabilidade @ 95%)		Unidade
0,000320		normal		Infinito		2,00		0,00064		g

Quadro 2 – Componentes de incerteza para a quantificação do etanol

Símbolo	Fontes de incerteza			Distribuição de probabilidade	Divisor	Coeficiente de sensibilidade c_i		Incerteza padrão u_i [% m/m]	Graus de liberdade n_i	Porcentagem de contribuição
	Nome	Valor (+/-)	Unidade			Valor	Unidade			
$u_{repe.etanol}$	Repetibilidade	0,07635	% m/m	normal	1,00	1,00	-	0,07635	9	67,73
u_{curva}	Incerteza do modelo de regressão	0,04758	-	normal	1,00	1,00	-	0,04758	infinito	26,31
$u_{massa.analito}$	Incerteza do analito	0,000639	g	normal	1,00	10,00	1/g	0,00639	infinito	0,47
$u_{massa.total}$	Incerteza da massa total	0,002556	g	normal	1,00	8,50	1/g	0,02173	infinito	5,49
Incerteza padrão combinada (u_c)		Distribuição de probabilidade		Graus de liberdade (n_{eff})		Fator de abrangência (k)		Incerteza expandida (U) (Probabilidade @ 95%)		Unidade
0,093		normal		20		2,09		0,194		% (m/m)

Quadro 3 – Componentes de incerteza para a quantificação do metanol

Símbolo	Fontes de incerteza			Distribuição de probabilidade	Divisor	Coeficiente de sensibilidade c_i		Incerteza padrão u_i [% m/m]	Graus de liberdade n_i	Porcentagem de contribuição
	Nome	Valor (+/-)	Unidade			Valor	Unidade			
$u_{repe.metanol}$	Repetibilidade	0,00980	% m/m	normal	1,00	1	-	0,00980	9	61,2
u_{curva}	Incerteza do modelo de regressão	0,00428	-	normal	1,00	1,00	-	0,00428	infinito	11,7
$u_{massa.analito}$	Incerteza do analito	0,00064	g	normal	1,00	10,00	1/g	0,00639	infinito	26,1
$u_{massa.total}$	Incerteza da massa total	0,00256	g	normal	1,00	0,50	1/g	0,00128	infinito	1,0
Incerteza padrão combinada (u_c)		Distribuição de probabilidade		Graus de liberdade (n_{eff})		Fator de abrangência (k)		Incerteza expandida (U) (Probabilidade @ 95%)		Unidade
0,0125		normal		24		2,06		0,026		% (m/m)

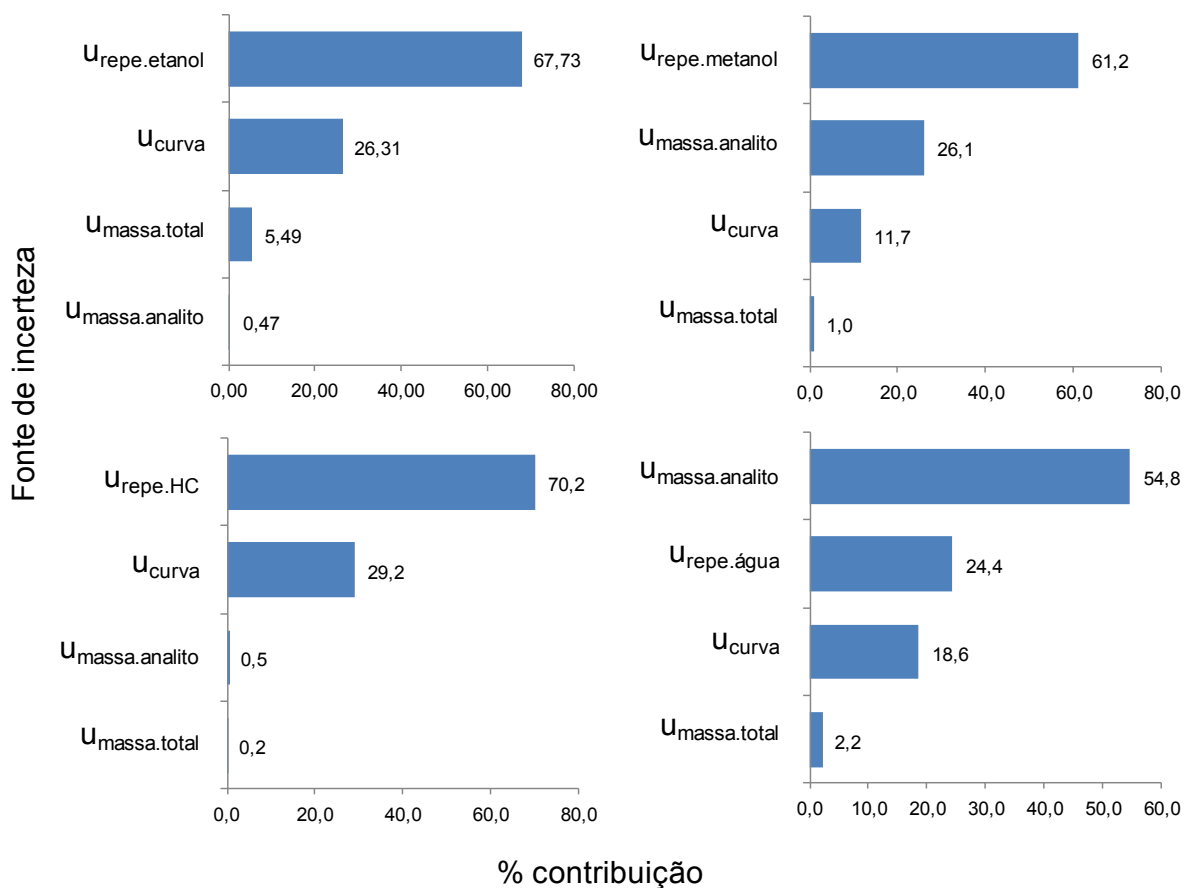
Quadro 4 – Componentes de incerteza para a quantificação do HC

Símbolo	Fontes de incerteza			Distribuição de probabilidade	Divisor	Coeficiente de sensibilidade c_i		Incerteza padrão u_i [% m/m]	Graus de liberdade n_i	Porcentagem de contribuição
	Nome	Valor (+/-)	Unidade			Valor	Unidade			
$u_{repe.HC}$	Repetibilidade	0,07690	% m/m	normal	1,00	1,0000	-	0,07690	9	70,2
u_{curva}	Incerteza do modelo de regressão	0,04956	-	normal	1,00	1,0000	-	0,04956	infinito	29,2
$u_{massa.analito}$	Incerteza do analito	0,00064	g	normal	1,00	10,0000	1/g	0,00639	infinito	0,5
$u_{massa.total}$	Incerteza da massa total	0,00256	g	normal	1,00	1,5000	1/g	0,00383	infinito	0,2
Incerteza padrão combinada (u_c)		Distribuição de probabilidade		Graus de liberdade (n_{eff})		Fator de abrangência (k)		Incerteza expandida (U) (Probabilidade @ 95%)		Unidade
0,09		normal		18		2,10		0,193		% (m/m)

Quadro 5 – Componentes de incerteza para a quantificação da água

Símbolo	Fontes de incerteza			Distribuição de probabilidade	Divisor	Coeficiente de sensibilidade c_i		Incerteza padrão u_i [% m/m]	Graus de liberdade n_i	Porcentagem de contribuição
	Nome	Valor (+/-)	Unidade			Valor	Unidade			
$u_{repe.água}$	Repetibilidade	0,00427	% m/m	normal	1,00	1,0000	-	0,00427	9	24,4
u_{curva}	Incerteza do modelo de regressão	0,00373	-	normal	1,00	1,0000	-	0,00373	infinito	18,6
$u_{massa.analito}$	Incerteza do analito	0,00064	g	normal	1,00	10,0000	1/g	0,00639	infinito	54,8
$u_{massa.total}$	Incerteza da massa total	0,00256	g	normal	1,00	0,5000	1/g	0,00128	infinito	2,2
Incerteza padrão combinada (u_c)		Distribuição de probabilidade		Graus de liberdade (n_{eff})		Fator de abrangência (k)		Incerteza expandida (U) (Probabilidade @ 95%)		Unidade
0,009		normal		151		1,98		0,0171		% (m/m)

Figura 33: Fonte de incerteza e suas porcentagens de contribuição para os quatro modelos.



5.8.6. Graus de liberdade efetivos

Os graus de liberdade efetivos para cada modelo foram obtidos através da equação 30, e seus valores também são mostrados nos quadros 2, 3, 4, e 5.

5.8.7. Fator de abrangência

Os fatores de abrangência, ao grau de 95% de confiança, foram obtidos através da tabela da distribuição t e são mostrados nos quadros 2, 3, 4, e 5.

5.8.8. Incerteza expandida

Por fim, a incerteza expandida é dada pela multiplicação da incerteza combinada u_c pelo fator de abrangência k :

$$U_{etanol} = 0,2\% (m/m)$$

$$U_{metanol} = 0,03\% (m/m)$$

$$U_{HC} = 0,2\% (m/m)$$

$$U_{\acute{a}gua} = 0,02\% (m/m)$$

5.9. Incerteza de medição pelo método de Monte Carlo

Para a simulação de Monte Carlo, foram adotados os valores mostrados na **Tabela 18** para as grandezas de entrada.

Tabela 18 – Fontes de entrada e respectivas PDFs para estimativa de incerteza do teor de etanol, metanol, HC e água

Modelo	Grandeza de entrada	PDF	Parâmetros da PDF
Etanol	Repetibilidade	Normal	Média: 0% (m/m) Desvio: 0,0763% (m/m)
	Massa do analito	Normal	Média: 0% (m/m) Desvio: 0,0064% (m/m)
	Massa total	Normal	Média: 0% (m/m) Desvio: 0,0217% (m/m)
	Coeficiente linear (b_0)	Normal	Média: 0,0061 Desvio: 0,0012
	Coeficiente angular (b_1)	Normal	Média: 0,0059 [% (m/m)] ⁻¹ Desvio: $1,66 \cdot 10^{-5}$ [% (m/m)] ⁻¹
Metanol	Repetibilidade	Normal	Média: 0% (m/m) Desvio: 0,0100% (m/m)
	Massa do analito	Normal	Média: 0% (m/m) Desvio: 0,0064% (m/m)
	Massa total	Normal	Média: 0% (m/m) Desvio: 0,0013% (m/m)
	Coeficiente linear (b_0)	Normal	Média: 0,0016 Desvio: $1,13 \cdot 10^{-5}$
	Coeficiente angular (b_1)	Normal	Média: 0,0015 [% (m/m)] ⁻¹ Desvio: $2,03 \cdot 10^{-5}$ [% (m/m)] ⁻¹
HC	Repetibilidade	Normal	Média: 0% (m/m) Desvio: 0,0769% (m/m)
	Massa do analito	Normal	Média: 0% (m/m) Desvio: 0,0064% (m/m)
	Massa total	Normal	Média: 0% (m/m) Desvio: 0,0038% (m/m)
	Coeficiente linear (b_0)	Normal	Média: 0,0105 Desvio: 0,0009
	Coeficiente angular (b_1)	Normal	Média: 0,0104 [% (m/m)] ⁻¹ Desvio: $3,33 \cdot 10^{-5}$ [% (m/m)] ⁻¹
Água	Repetibilidade	Normal	Média: 0% (m/m) Desvio: 0,0100% (m/m)
	Massa do analito	Normal	Média: 0% (m/m) Desvio: 0,0064% (m/m)
	Massa total	Normal	Média: 0% (m/m) Desvio: 0,0013% (m/m)
	Coeficiente linear (b_0)	Normal	Média: 0,1182 Desvio: 0,0007
	Coeficiente angular (b_1)	Normal	Média: 0,1179 [% (m/m)] ⁻¹ Desvio: 0,0007 [% (m/m)] ⁻¹

Os resultados das simulações, utilizando 2×10^5 iterações, são mostrados nas **Figuras 33, 34 35 e 36 e Tabela 19**.

Figura 34 – Histograma representando a PDF resultante da simulação de Monte Carlo para o teor de etanol

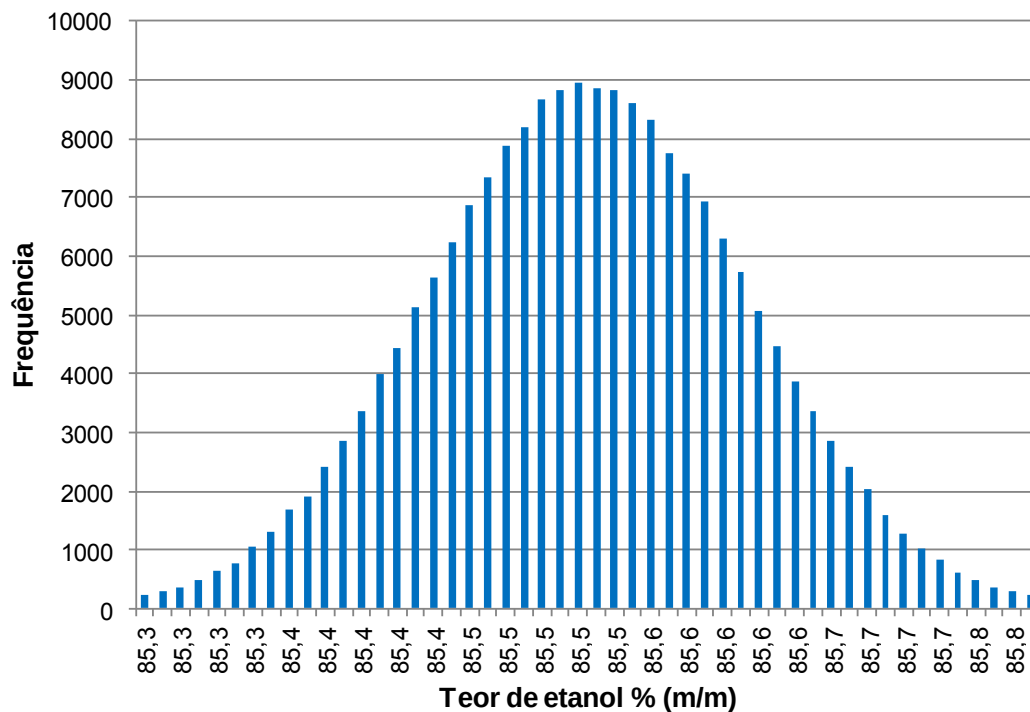


Figura 35 – Histograma representando a PDF resultante da simulação de Monte Carlo para o teor de metanol

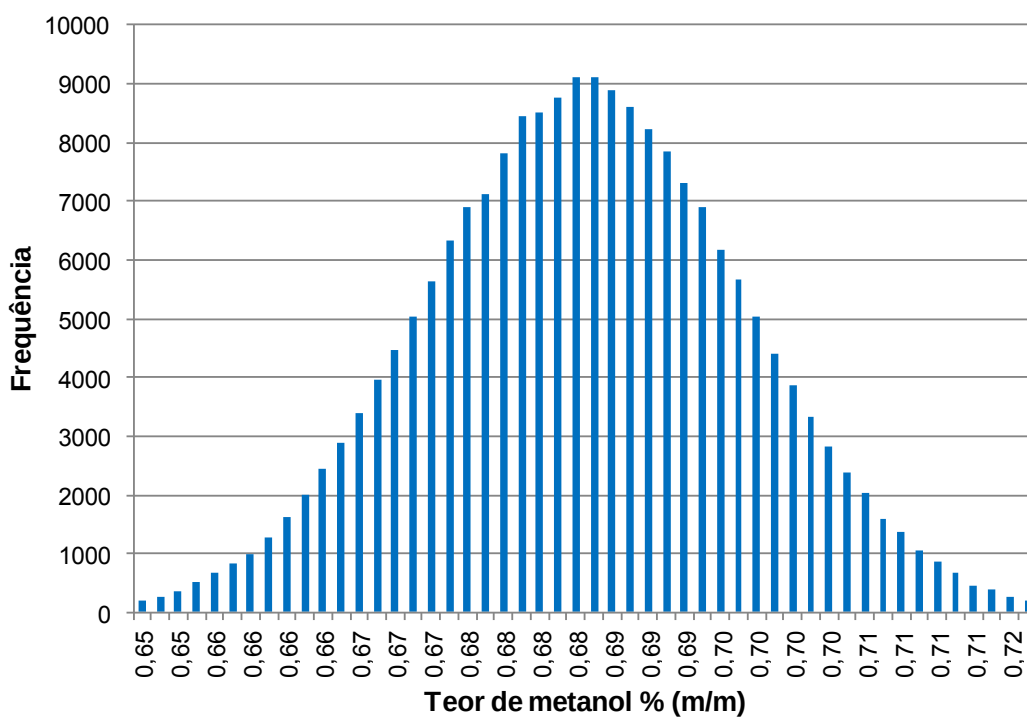


Figura 36 – Histograma representando a PDF resultante da simulação de Monte Carlo para o teor de HC

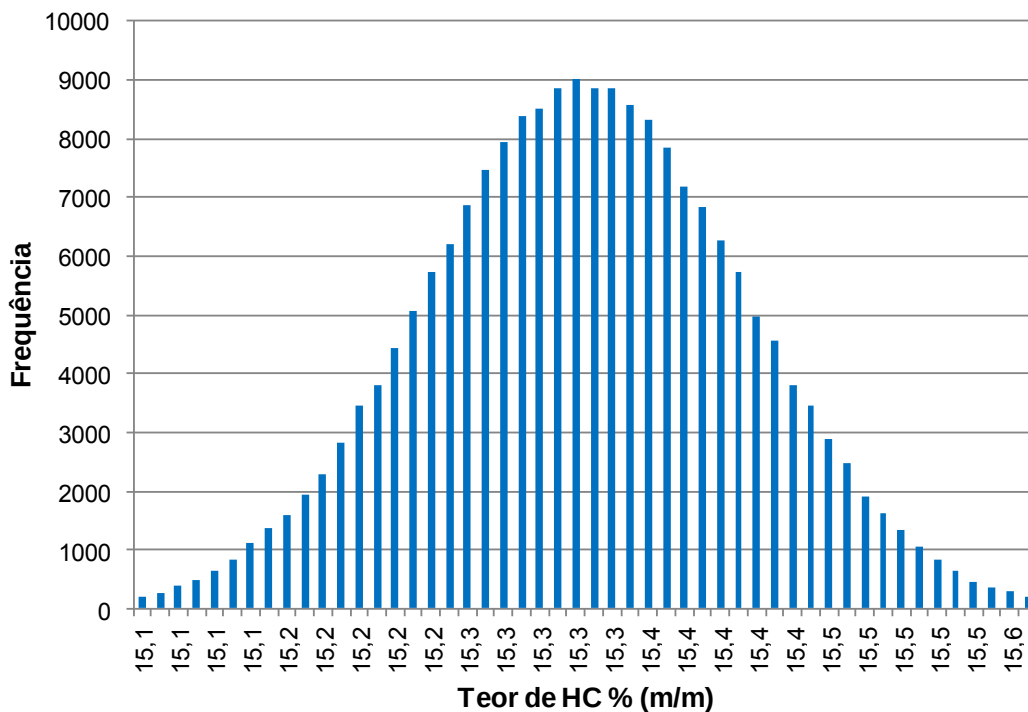


Figura 37 – Histograma representando a PDF resultante da simulação de Monte Carlo para o teor de Água

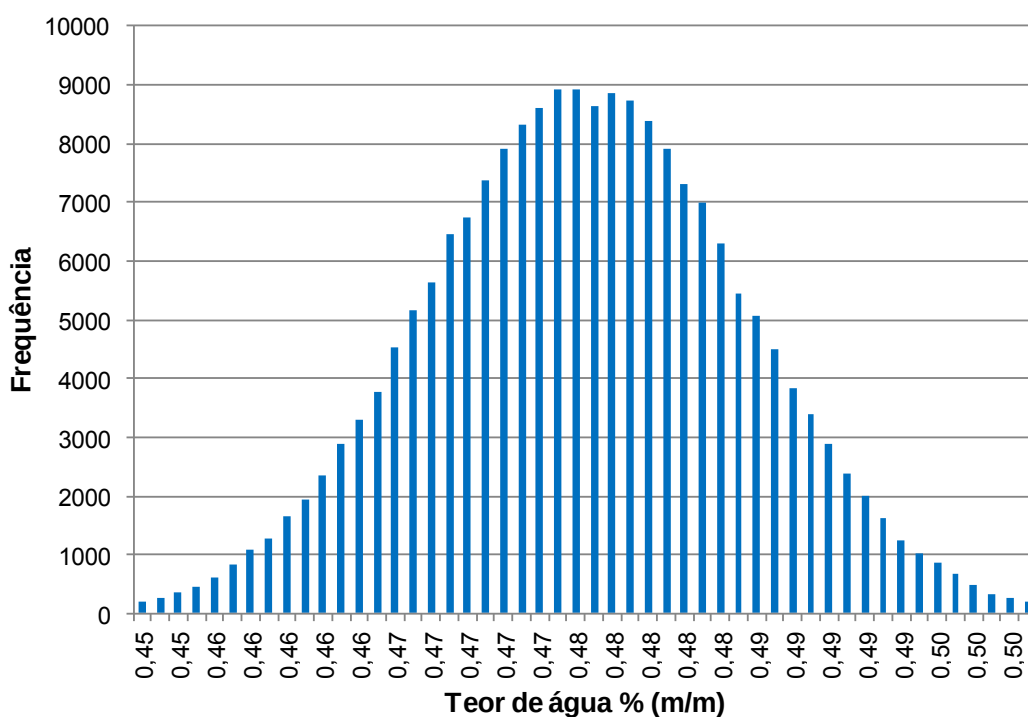


Tabela 19 – Parâmetros estatísticos obtidos através da simulação de Monte Carlo para os teores de etanol, metanol, HC e Água

	Etanol	Metanol	HC	Água
Média (% m/m)	85,5	0,68	15,3	0,47
Desvio padrão (% m/m)	0,09	0,01	0,09	0,009
Menor valor (95%) (% m/m)	85,3	0,66	15,1	0,46
Maior valor (95%) (% m/m)	85,7	0,71	15,5	0,49

5.10. Comparação entre os métodos ISO GUM e simulação de Monte Carlo

O Suplemento 1 do ISO GUM apresenta um procedimento para validar os resultados obtidos pela abordagem dada pelo ISO GUM com os resultados da simulação de Monte Carlo ⁷⁸. Esta técnica baseia-se na comparação dos menores e maiores valores dos pontos finais obtidos por ambos os métodos. Para tanto, deve-se comparar os valores de d_{low} e d_{high} com os valores da tolerância numérica (δ), os quais são calculados da seguinte forma:

$$d_{low} = |y - U(y) - y_{low}| \quad (39)$$

$$d_{high} = |y + U(y) - y_{high}| \quad (40)$$

onde:

y é a estimativa do mensurando;

$U(y)$ é incerteza expandida (ISO GUM);

y_{low} e y_{high} são os limites inferior e superior da simulação de Monte Carlo.

A tolerância numérica de uma incerteza, ou o desvio padrão, pode ser obtida expressando a incerteza padrão como $c \times 10^l$, onde c é um número inteiro com um número de dígitos igual ao número de dígitos significativos da incerteza padrão e l é um número inteiro. Então, a tolerância numérica é expressa como:

$$\delta = \frac{1}{2} \cdot 10^l \quad (41)$$

Se tanto o valor de d_{low} quanto o valor de d_{high} forem menores que o valor de δ , então o cálculo feito segundo o ISO GUM pode ser validado. A **Tabela 20** mostra os valores de d_{low} , d_{high} e δ , obtidos através das equações 39, 40 e 41 para os quatro modelos.

Tabela 20 – Valores de d_{low} , d_{high} e δ para os quatro modelos

Modelo	d_{low}	d_{high}	δ
Etanol	$3,7 \times 10^{-2}$	$3,6 \times 10^{-2}$	5×10^{-2}
Metanol	$3,2 \times 10^{-3}$	$3,2 \times 10^{-3}$	5×10^{-3}
HC	$3,7 \times 10^{-2}$	$3,7 \times 10^{-2}$	5×10^{-2}
Água	4×10^{-4}	4×10^{-4}	5×10^{-4}

A partir dos dados da **Tabela 20** é possível concluir que o método proposto pelo ISO GUM pode ser validado de acordo com o procedimento proposto no Suplemento 1 do ISO GUM para todos os modelos. É importante observar que foram adotados critérios diferentes em cada caso: para Etanol e HC foi adotada a resolução de décimos; para Metanol centésimos e para Água milésimos. Ou seja, na expressão dos resultados, esses mesmos critérios devem ser seguidos.

5.11. Estimativa dos custos dos ensaios

Objetivando fazer uma comparação de custos entre o método desenvolvido e os métodos oficiais, fez-se uma estimativa de custos para os ensaios. Para tanto, foram necessárias algumas considerações.

5.11.1. Considerações

Para a realização dos cálculos, foram considerados apenas os custos diretamente ligados aos ensaios. Custos fixos e indiretos não foram considerados, com exceção dos reagentes utilizados na preparação de curvas analíticas e na calibração dos modelos PLS, as manutenções e consumíveis dos equipamentos utilizados diretamente nos ensaios. São exemplos de custos fixos e indiretos que não foram incluídos: salário do pessoal administrativo, gastos de divulgação e propaganda, gastos com viagens e treinamentos, gastos com pesquisa e desenvolvimento, gastos com obras civis, gastos com limpeza em geral, gastos com acreditação, etc.

Além disso, para o cálculo do custo da hora dos equipamentos, considerou-se que cada equipamento é utilizado apenas em um único ensaio e o tempo para depreciação total de cada equipamento é de cinco anos. Os valores adotados para as manutenções preventivas foram obtidos em orçamentos das empresas fornecedoras, já para as manutenções corretivas, foram valores estimados com base na média dos históricos de manutenções ao longo dos últimos três anos.

Considerou-se um quantitativo de 190 amostras por mês. Este valor corresponde ao número de amostras de etanol hidratado combustível que o Cempeqc ensaia atualmente. Isso é necessário para estimar a quantidade de reagentes que são gastos na preparação das curvas analíticas, ou para construção dos modelos de regressão multivariada (únicos custos indiretos incluídos). Em outro cenário, considerou-se a capacidade máxima de cada técnica, ou seja, o equipamento funcionando sem interrupções durante todo o dia.

5.11.1.1. Teor de água por Karl Fischer

A **Tabela 21** mostra a descrição e os valores dos itens que foram levados em conta na estimativa de custo do ensaio para determinação do teor de água pelo método de Karl Fisher coulométrico segundo a ASTM E1064 ⁸⁸.

Tabela 21 – Descrição detalhada dos itens levados em consideração e seus respectivos valores para o cálculo do custo do ensaio de determinação do teor de água por Karl Fisher coulométrico

Descrição do Custo	Quantidade por ensaio	Unidade	Valor unitário	Valor por ensaio
Titulador coulométrico de Karl Fischer	0,13	h	R\$ 1,72	R\$ 0,23
Manutenção preventiva	0,0004	Un.	R\$ 1.169,74	R\$ 0,51
Manutenção corretiva	0,0004	Un.	R\$ 1.852,00	R\$ 0,81
Reagente Karl Fischer (solução anódica)	0,001	L	R\$ 777,00	R\$ 0,98
Reagente Karl Fischer (solução catódica)	0,02	Ampola	R\$ 333,33	R\$ 7,02
Metanol	0,002	L	R\$ 85,00	R\$ 0,17
Seringa descartável	1	Un.	R\$ 2,00	R\$ 2,00
Agulha de 200 mm x 10 mm	0,01	Un.	R\$ 3,04	R\$ 0,03
Energia elétrica	0,005	kWh	R\$ 0,32	R\$ 0,00
Pessoal direto	0,13	h	R\$ 25,09	R\$ 3,34
TOTAL				R\$ 11,76

5.11.1.2. Teor de etanol e metanol por cromatografia em fase gasosa

A **Tabela 22** mostra a descrição e os valores dos itens que foram levados em conta na estimativa de custo do ensaio para determinação do teor de metanol e etanol em amostras de E85 empregando cromatografia em fase gasosa, segundo a ASTM D5501 ⁸⁹.

Tabela 22 – Descrição detalhada dos itens levados em consideração e seus respectivos valores para o cálculo do custo do ensaio de determinação do teor de etanol e metanol por cromatografia em fase gasosa

Descrição do Custo	Quantidade por ensaio	Unidade	Valor unitário	Valor por ensaio
Cromatógrafo a gás	0,8	h	R\$ 6,50	R\$ 5,40
Manutenção preventiva	0,0004	Un.	R\$ 3.825,00	R\$ 1,68
Manutenção corretiva	0,0004	Un.	R\$ 4.833,33	R\$ 2,12
Gás hélio	0,01	m ³	R\$ 228,78	R\$ 2,29
Gás hidrogênio	0,002	m ³	R\$ 104,72	R\$ 0,16
Ar sintético	0,02	m ³	R\$ 65,43	R\$ 0,98
Gás nitrogênio	0,002	m ³	R\$ 73,76	R\$ 0,11
Septo	0,01	Un.	R\$ 7,00	R\$ 0,07
<i>Liner</i>	0,0004	Un.	R\$ 707,25	R\$ 0,31
Seringa cromatográfica de 5 µL	0,0004	Un.	R\$ 285,71	R\$ 0,13
Coluna cromatográfica	0,0004	Un.	R\$ 2.495,24	R\$ 1,09
<i>Vial</i> de vidro de 2 mL com tampa	1	Un.	R\$ 0,90	R\$ 0,90
Etanol	0,001	L	R\$ 69,20	R\$ 0,10
Metanol	0,00001	L	R\$ 21,25	R\$ 0,00
Isooctano	0,0003	L	R\$ 410,00	R\$ 0,11
Heptano	0,0002	L	R\$ 74,20	R\$ 0,01
Energia elétrica	2,2	kWh	R\$ 0,32	R\$ 0,69
Pessoal direto	0,05	h	R\$ 25,09	R\$ 1,25
TOTAL				R\$ 17,41

5.11.1.3. Teor de etanol, metanol, hidrocarbonetos e água por espectroscopia na região do infravermelho e quimiometria

A **Tabela 23** mostra a descrição e os valores dos itens que foram levados em conta na estimativa de custo do ensaio para determinação do teor de metanol, etanol, HC e água em amostras de E58 empregando o método desenvolvido e validado neste trabalho.

Tabela 23 – Descrição detalhada dos itens levados em consideração e seus respectivos valores para o cálculo do custo do ensaio de determinação do teor de etanol, metanol, HC e água por IV e PLS

Descrição do Custo	Quantidade por ensaio	Unidade	Valor unitário	Valor por ensaio
Espectrofotômetro de infravermelho	0,05	h	R\$ 16,83	R\$ 0,84
Manutenção preventiva	0,0004	Un.	R\$ 2.200,00	R\$ 0,96
Manutenção corretiva	0,0004	Un.	R\$ 3.100,00	R\$ 1,36
Pipeta de Pasteur de 3 ml descartável	1	Un.	R\$ 0,03	R\$ 0,03
Etanol	0,001	L	R\$ 69,20	R\$ 0,08
Metanol	0,0001	L	R\$ 21,25	R\$ 0,00
Água	0,0001	L	R\$ 3,00	R\$ 0,00
Energia elétrica	0,007	kWh	R\$ 0,32	R\$ 0,00
Pessoal direto	0,05	h	R\$ 25,09	R\$ 1,25
TOTAL				R\$ 4,53

A partir dos cálculos dos custos, podemos observar que o ensaio realizado por IV e PLS é o que tem menor custo (R\$ 4,53), enquanto a soma dos ensaios titulométrico e cromatográfico é de R\$ 29,17. Ou seja, o método proposto corresponde a 15,5% do valor da soma dos custos dos métodos oficiais. Isso mostra que o método proposto leva grande vantagem sobre os métodos oficiais em relação ao custo.

Quando se considera que os equipamentos trabalham com suas capacidades máximas de ensaios e aplica-se o mesmo método de cálculo (**Tabela 24**), é possível observar que os custos dos ensaios correspondem a 75% do valor para o método cromatográfico, 53% para o método titulométrico e 48% para o método espectroscópico em relação aos custos desses ensaios quando é considerado um quantitativo de 190 amostras por mês. Nesse mesmo cenário, o método proposto corresponde a 11,2% do valor da soma dos custos dos métodos oficiais, isso mostra que a vantagem econômica do método proposto é maior à medida que se aumenta o número de amostras analisadas, até a capacidade máxima.

Tabela 24 – Comparação dos custos dos ensaios considerando os equipamentos operando com um quantitativo de 190 amostras mensais e em sua capacidade máxima.

Ensaio	Condição de operação	
	190 amostras/mês	Capacidade máxima
Cromatografia em fase gasosa	R\$ 17,41	R\$ 13,07
Titulação coulométrica – Karl Fishcer	R\$ 11,76	R\$ 6,28
Espectrofotômetro de infravermelho	R\$ 4,53	R\$ 2,16

5.11.2. Avaliação da composição dos custos

As **Figuras 37, 38 e 39** mostram, através de gráficos de barras, o quanto cada item contribui percentualmente para a composição do preço de cada ensaio. Nesses gráficos são mostrados apenas os gastos que representam mais de 1% do valor total do ensaio. Podemos observar que para o ensaio de Karl Fisher, 88,1% dos custos são devido ao reagente catódico (59,7%) e ao pessoal (28,5%). Para o ensaio cromatográfico, 66,0% dos custos são devido ao equipamento e suas manutenções e ao gás hélio utilizado como gás de arraste. Já para o ensaio espectroscópico, 79,0% dos custos provêm das manutenções e do pessoal.

Figura 38 – Gráfico de barras mostrando a porcentagem de contribuição de cada item para o preço total do ensaio de determinação do teor de água por Karl Fischer

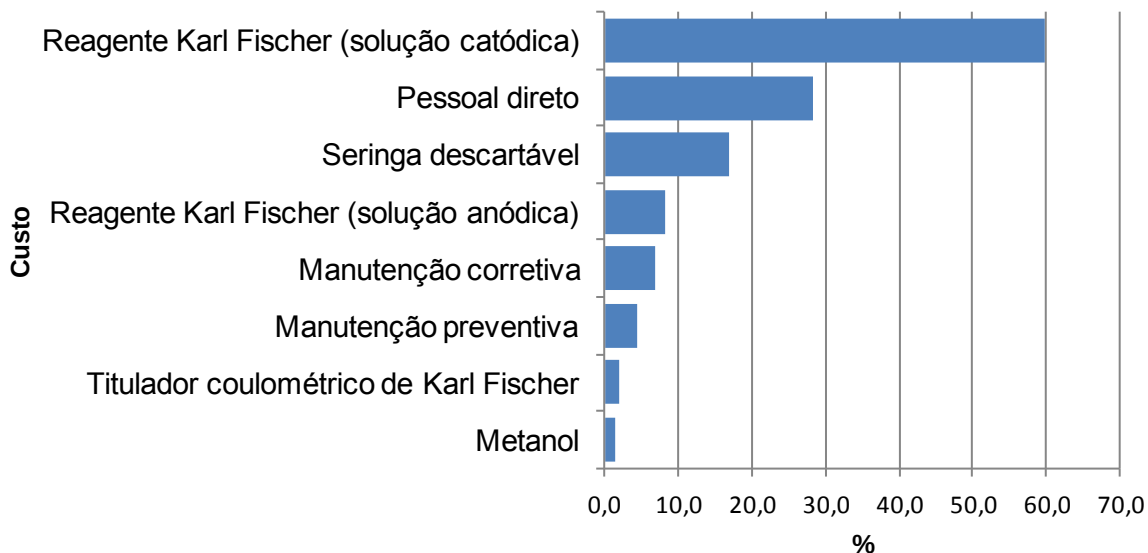


Figura 39 – Gráfico de barras mostrando a porcentagem de contribuição de cada item para o preço total do ensaio de determinação do teor de metanol e etanol por cromatografia em fase gasosa

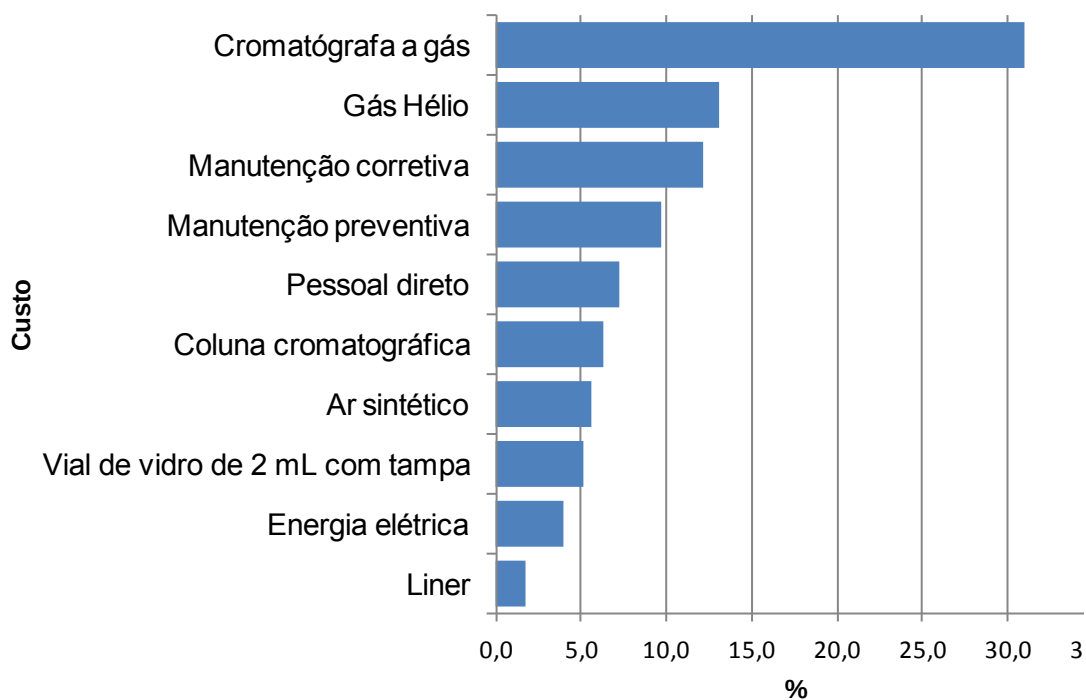
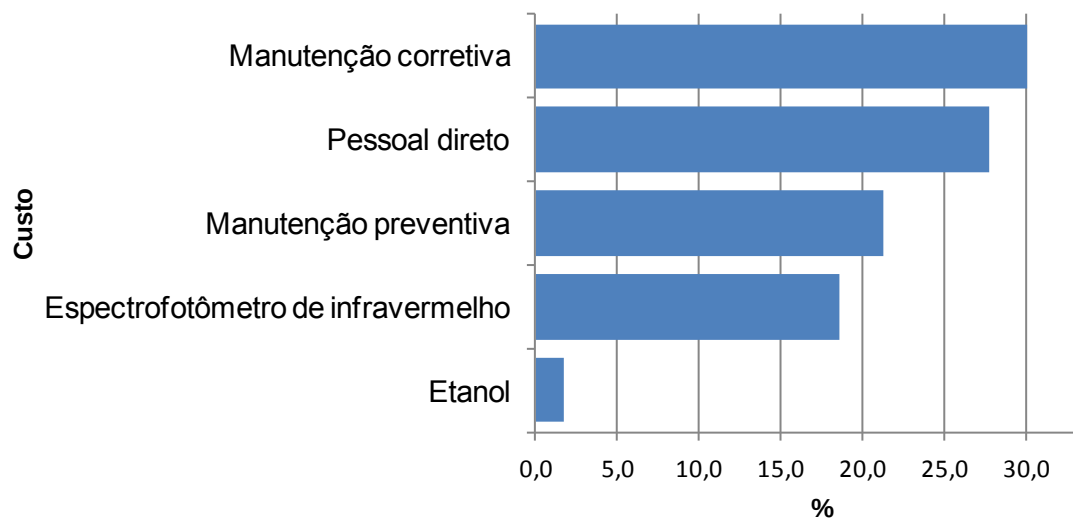


Figura 40 – Gráfico de barras mostrando a porcentagem de contribuição de cada item para o preço total do ensaio de determinação do teor de etanol, metanol, HC e água por IV e PLS



6. CONCLUSÕES

As amostras de E85 foram preparadas objetivando abranger as faixas dos analitos estudados (etanol, metanol, água e hidrocarbonetos) previstos nas normas americana e europeia. A partir dos espectros na região do infravermelho (médio e próximo) dessas amostras, foram construídos modelos de regressão multivariada (PLS) para a previsão dos teores dos analitos estudados.

A seleção de variáveis mostrou-se capaz de remover informações sem relevância para os modelos. Os melhores modelos para a previsão das concentrações de etanol, HC e água, foram os que utilizam os espectros obtidos na região do NIR (caminho óptico de 0,5 mm), já para a previsão da concentração de metanol, foi o que utiliza os espectros no MIR. As figuras de mérito desses modelos mostraram que todos são capazes de quantificar de forma linear, precisa, exata e sem erros sistemáticos os teores dos analitos estudados, mostrando que o método proposto pode ser aplicado no processo de avaliação da qualidade do E85. Além disso, a variação de temperatura entre 18 °C e 24 °C não influencia significativamente os valores de concentração dos analitos.

A utilização da forma pseudounivariada dos modelos permitiu a realização dos cálculos para a estimativa de incerteza dos quatro métodos propostos, tanto pelo método do ISO GUM quanto pela simulação de Monte Carlo. Além disso, os valores das incertezas calculadas seguindo o ISO GUM foram validadas pela simulação de Monte Carlo (com diferentes critérios para cada modelo), ou seja, o método do ISO GUM se mostrou adequado para avaliar as incertezas dos modelos, respeitando a resolução da expressão dos resultados. As incertezas estimadas são compatíveis com os com os teores determinados na ASTM D5798 e na CEN/TS 15293.

A partir das estimativas de custos para os ensaios (desenvolvidos e oficiais), pode-se inferir que os procedimentos envolvidos nas análises das amostras tornam o método aplicável em rotina e com baixo custo em relação aos métodos oficiais.

Logo, conclui-se que o trabalho atingiu os objetivos propostos, gerando assim, métodos que são uma excelente alternativa para a quantificação de etanol, metanol, hidrocarbonetos e água em amostras de E85 e que pode ser aplicado em análises de

rotina. Além disso, este trabalho mostrou que é possível aplicar a abordagem de incerteza de medição aos métodos PLS. A comprovação dessa aplicabilidade permitirá que os laboratórios que seguem a ISO 17025 e os pesquisadores possam gerar mais evidências para a comprovação da validade dos métodos quimiométricos de quantificação. Pois, é necessário que se tenha indicação quantitativa da qualidade do resultado de medição para que se possa comparar esses resultados, seja entre eles mesmos ou com valores de referência fornecidos numa especificação ou numa norma.

Estudos envolvendo espectroscopia na região do infravermelho e métodos quimiométricos poderiam, futuramente, ser desenvolvidos para avaliar sua aplicabilidade para previsão de outros parâmetros, como a pressão de vapor, teores de éteres, estabilidade oxidativa, teor de alcoóis superiores (avaliando a possível interferência que possa existir na quantificação dos outros álcoois) e goma. Além disso, poder-se-ia avaliar o comportamento dos modelos utilizando gasolinas oriundas de diferentes refinarias (ou diferentes processos), visto que a composição dos hidrocarbonetos pode se alterar em função disso.

REFERÊNCIAS

- 1 ASSOCIAÇÃO DOS FABRICANTES DE VEÍCULOS AUTOMOTORES. **Produção, vendas e exportação de autoveículos**. São Paulo, 2014. Disponível em: <<http://www.anfavea.com.br/docs/siteautoveiculos2014.xlsx>>. Acesso em: 1 dez. 2014.
- 2 THE INTERNATIONAL CONUNCIL ON CLEAN TRANSPORTATION. **European Vehicle Market statistics: pocketbook 2014**. Berlin, 2014. Disponível em: <http://www.theicct.org/sites/default/files/publications/EU_pocketbook_2014.pdf>. Acesso em: 2 fev. 2015.
- 3 MILHOR, C. E. **Sistema de desenvolvimento para controle eletrônico dos motores de combustão interna ciclo Otto**. 2002. 89 f. Dissertação (Mestrado em Engenharia Mecânica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2002.
- 4 BAPTISTA, P. et al. Multivariate near infrared spectroscopy models for predicting the methyl esters content in biodiesel. **Analitica Chimica Acta**, v. 607, n. 2, p. 153-159, Jan. 2008.
- 5 DEMIRBAS, A. Importance of biodiesel as transportation fuel. **Energy Policy**, v. 35, n. 9, p. 4661-4670, Sept. 2007.
- 6 JACOBSON, M. Z. Review of solutions to global warming, air pollution, and energy security. **Energy & Environmental Science**, v. 2, n. 2, p. 148-173, Dec. 2009.
- 7 LUO, L.; van der VOET, E.; HUPPES, G. Life cycle assessment and life cycle costing of bioethanol from sugarcane in Brazil. **Renewable & Sustainable Energy Reviews**, v. 13, n. 6/7, p. 1613-1619, Aug./Sept. 2009.
- 8 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM D5798**: standard specification for ethanol fuel blends for flexible-fuel automotive spark-Ignition engines. West Conshohocken, 2014. 10 p.
- 9 EUROPEAN COMMITTEE FOR STANDARDIZATION. **CEN/TS 15293:2011**: automotive fuels: ethanol (E85) automotive fuel: requirements and test methods. Brussels, 2011. 18 p.
- 10 SILVA, A. C. et al. Detection of adulteration in hydrated ethyl alcohol fuel using infrared spectroscopy and supervised pattern recognition methods. **Talanta**, v. 93, p. 129-134, May 2012.
- 11 FERREIRA, S. M.; OLIVEIRA, E. D.; DUARTE, U. Comportamento da gasolina com etanol (E-20) e da gasolina pura após a simulação de um vazamento em colunas de laboratório. **Geologia USP: Série Científica**, v. 4, n. 2, p. 12, Out. 2004.

12 BANCO NACIONAL DE DESENVOLVIMENTO ECONÔMICO E SOCIAL; CENTRO DE GESTÃO E ESTUDOS ESTRATÉGICOS (Org.). **Bioetanol de cana-de-açúcar: energia para o desenvolvimento sustentável**. Rio de Janeiro: BNDES, 2008. 316 p. Disponível em: <<http://www.bioetanoldecana.org/pt/download/bioetanol.pdf>>. Acesso em: 17 set. 2014.

13 NAJAFI, G. et al. Performance and exhaust emissions of a gasoline engine with ethanol blended gasoline fuels using artificial neural network. **Applied Energy**, v. 86, n. 5, p. 630-639, May 2009.

14 AL-HASAN, M. Effect of ethanol–unleaded gasoline blends on engine performance and exhaust emission. **Energy Conversion and Management**, v. 44, n. 9, p. 1547-1561, June 2003.

15 YÜCESU, H. S. et al. Effect of ethanol–gasoline blends on engine performance and exhaust emissions in different compression ratios. **Applied Thermal Engineering**, v. 26, n. 17/18, p. 2272-2278, Dec. 2006.

16 HSIEH, W. D. et al. Engine performance and pollutant emission of an SI engine using ethanol-gasoline blended fuels. **Atmospheric Environment**, v. 36, n. 3, p. 403-410, Jan. 2002.

17 YUKSEL, F.; YUKSEL, B. The use of ethanol-gasoline blend as a fuel in an SI engine. **Renewable Energy**, v. 29, n. 7, p. 1181-1191, June 2004.

18 AGARWAL, A. K. Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines. **Progress in Energy and Combustion Science**, v. 33, n. 3, p. 233-271, June 2007.

19 BALAT, M.; BALAT, H. Recent trends in global production and utilization of bio-ethanol fuel. **Applied Energy**, v. 86, n. 11, p. 2273-2282, Nov. 2009.

20 BOSCH NETO, J. C. **Simulação de emissões de misturas gasolina / etanol em motores de combustão interna**. 2012. 211 f. Tese (Doutorado em Engenharia Mecânica) – Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2012.

21 MELO, T. C. C. de et al. Hydrous ethanol-gasoline blends - combustion and emission investigations on a Flex-Fuel engine. **Fuel**, v. 97, p. 796-804, July 2012.

22 SILVERSTEIN, R. M.; WEBSTER, F. X.; KIEMLE, D. J. **Spectrometric identification of organic compounds**. 7th ed. Hoboken: John Wiley & Sons, 2005. 502 p.

23 CIENFUEGOS, F.; VAITSMAN, D. S. **Análise instrumental**. 4. ed. Rio de Janeiro: Interciência, 2000. 606 p.

24 NESPECA, M. G. **Desenvolvimento de modelos quimiométricos associados à espectroscopia no infravermelho para determinação de parâmetros físico-químicos do óleo diesel**. 2013. 114 f. Dissertação (Mestrado em Química) - Instituto de Química, Universidade Estadual Paulista, Araraquara, 2013.

- 25 PAVIA, D. et al. **Introduction to spectroscopy**. 4th ed. Belmont: Cengage Learning, 2008. 752 p.
- 26 CORSETTI, S. et al. Characterization of gasoline/ethanol blends by infrared and excess infrared spectroscopy. **Fuel**, v. 141, p. 136-142, Feb. 2015.
- 27 EUROPEAN STANDARD. **EN 14078**: liquid petroleum products: determination of fatty acid methyl ester (FAME) content in middle distillates: infrared spectrometry method. Brussels, 2014. 20 p.
- 28 ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 15568**: biodiesel: determinação do teor de biodiesel em óleo diesel por espectroscopia na região do infravermelho médio. Rio de Janeiro, 2008. 13 p.
- 29 SKOOG, D. A.; HOLLER, F. J.; CROUCH, S. R. **Principles of instrumental analysis**. 6th ed. Belmont: Thomson: Brooks/Cole, 2007. 1039 p.
- 30 PIKE TECHNOLOGIES. **ATR**: theory and applications. Madison, 2011. Disponível em: <<http://www.piketech.com/files/pdfs/ATRAN611.pdf>>. Acesso em: 14 jan. 2015.
- 31 BREITKREITZ, M. C. **Desenvolvimento e caracterização de formulações de fármacos pouco solúveis em água empregando espectroscopia de imagem e quimiometria**. 2013. 177 f. (Doutorado em Química) - Instituto de Química, Universidade Estadual de Campinas, Campinas, 2013.
- 32 PASQUINI, C. Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society**, v. 14, n. 2, p. 198-219, Mar./Apr. 2003.
- 33 FERREIRA, M. M. C. et al. Quimiometria I: calibração multivariada, um tutorial. **Química Nova**, v. 22, n. 5, p. 724-731, set./out. 1999.
- 34 FERREIRA, M. M. C. Multivariate QSAR. **Journal of the Brazilian Chemical Society**, v. 13, n. 6, p. 742-753, Nov./Dec. 2002.
- 35 CORREIA, P. R. M.; FERREIRA, M. M. C. Non-supervised pattern recognition methods: exploring chemometrical procedures for evaluating analytical data. **Química Nova**, v. 30, n. 2, p. 481-487, mar./abr. 2007.
- 36 MARTIN, M. J.; PABLOS, F.; GONZALEZ, A. G. Characterization of green coffee varieties according to their metal content. **Analytica Chimica Acta**, v. 358, n. 2, p. 177-183, Jan. 1998.
- 37 BARBEIRA, P. J. S. Using statistical tools to detect gasoline adulteration. **Engenharia Térmica**, v. 1, n. 2, p. 48-50, 2002.
- 38 IOB, A. et al. Hydrocarbon group (PONA) analysis of reformat by FT-ir spectroscopy. **Fuel**, v. 75, n. 9, p. 1060-1064, July 1996.

- 39 MEUSINGER, R. Qualitative and quantitative determination of oxygenates in gasolines using H-1 nuclear magnetic resonance spectroscopy. **Analytica Chimica Acta**, v. 391, n. 3, p. 277-288, June 1999.
- 40 KAISER, C. R. et al. Quality control of gasoline by ¹H NMR: aromatics, olefinics, paraffinics, and oxygenated and benzene contents. **Fuel**, v. 89, n. 1, p. 99-104, Jan. 2010.
- 41 BORIN, A.; POPPI, R. J. Multivariate quality control of lubricating oils using Fourier transform infrared Spectroscopy. **Journal of the Brazilian Chemical Society**, v. 15, n. 4, p. 570-576, July/Aug. 2004.
- 42 BORIN, A.; POPPI, R. J. Application of mid infrared spectroscopy and iPLS for the quantification of contaminants in lubricating oil. **Vibrational Spectroscopy**, v. 37, n. 1, p. 27-32, Jan. 2005.
- 43 FERNADES, H. L. et al. Simultaneous determination of methanol and ethanol in gasoline using NIR spectroscopy: effect of gasoline composition. **Talanta**, v.75, n. 3, p. 804-810, May 2008.
- 44 OUYANG, A.; LIU, J. Classification and determination of alcohol in gasoline using NIR spectroscopy and the successive projections algorithm for variable selection. **Measurement Science & Technology**, v. 24, n. 2, Jan. 2013. doi:10.1088/0957-0233/24/2/025502.
- 45 PIMENTEL, F. M. et al. Determination of biodiesel content when blended with mineral diesel fuel using infrared spectroscopy and multivariate calibration. **Microchemical Journal**, v. 82, n. 2, p. 201-206, Apr. 2006.
- 46 MOITA NETO, J. M.; MOITA, G. C. Uma introdução à análise exploratória de dados multivariados. **Química Nova**, v. 21, n. 4, p. 467-469, jul./ago. 1998.
- 47 SENA, M. M. de et al. Avaliação do uso de métodos quimiométricos em análise de solos. **Química Nova**, v. 23, n. 4, p. 547-556, jul./ago. 2000.
- 48 KRAMER, R. **Chemometric techniques for quantitative analysis**. New York: Marcel Dekker, 1998. 203 p.
- 49 NAES, T. **A user-friendly guide to multivariate calibration and classification**. Chichester: NIR Publications, 2004. 344 p.
- 50 GEMPERLINE, P. (Ed.). **Practical guide to chemometrics**. 2nd ed. Boca Raton: CRC/Taylor & Francis, 2006. 541 p.
- 51 MILLER, J. N.; MILLER, J. C. **Statistics and chemometrics for analytical chemistry**. 4th ed. New York: Prentice Hall, 2000. 271 p.
- 52 VARMUZA, K.; FILZMOSER, P. **Introduction to multivariate statistical analysis in chemometrics**. Boca Raton: CRC Press, 2009. 321 p.

- 53 BRERETON, R. G. **Chemometrics**: data analysis for the laboratory and chemical plant. Hoboken: Wiley, 2003. 489 p.
- 54 WOLD, S.; MARTENS, H.; WOLD, H. The multivariate calibration-problem in chemistry solved by the PLS method. **Lecture Notes in Mathematics**, v. 973, p. 286-293, 1983.
- 55 BRAGA, J. W. B.; SANTOS JUNIOR, A. A. dos; MARTINS, I. S. Determination of viscosity index in lubricant oils by infrared spectroscopy and PLSR. **Fuel**, v. 120, p. 171-178, Mar. 2014.
- 56 ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO/IEC 17025:2005**: requisitos gerais para a competência de laboratórios de ensaio e calibração. Rio de Janeiro, 2005. 28 p.
- 57 INSTITUTO NACIONAL DE METROLOGIA, NORMALIZAÇÃO E QUALIDADE INDUSTRIAL. **Orientação sobre validação de métodos analíticos**: DOQ-CGCRE-008. 2010. Disponível em: <http://www.inmetro.gov.br/Sidoq/Arquivos/CGCRE/DOQ/DOQ-CGCRE-8_03.pdf>. Acesso em: 21 out. 2014.
- 58 RIBANI, M. et al. Validação em métodos cromatográficos e eletroforéticos. **Química Nova**, v. 27, n. 5, p. 771-780, set./out. 2004.
- 59 THOMPSON, M.; ELLISON, S. L. R.; WOOD, R. Harmonized guidelines for single-laboratory validation of methods of analysis - (IUPAC technical report). **Pure and Applied Chemistry**, v. 74, n. 5, p. 835-855, May 2002.
- 60 BRESSOLLE, F.; BROMETPETIT, M.; AUDRAN, M. Validation of liquid chromatographic and gas chromatographic methods - applications to pharmacokinetics. **Journal of Chromatography B: Biomedical Applications**, v. 686, n. 1, p. 3-10, Nov. 1996.
- 61 INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE. **Validation of analytical procedures**: text and methodology Q2 (R1). 2005. Disponível em: <<http://www.ich.org/LOB/media/MEDIA417.pdf>>. Acesso em: 21 out. 2014.
- 62 VALDERRAMA, P.; BRAGA, J. W. B.; POPPI, R. J. Estado da arte de figuras de mérito em calibração multivariada. **Química Nova**, v. 32, n. 5, p. 1278-1287, nov. 2009.
- 63 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM E1655**: standard practices for infrared multivariate quantitative analysis. West Conshohocken, 2012. 29 p.
- 64 RIBEIRO, D. B. M. **Desenvolvimento de metodologias analíticas multivariadas empregando espectroscopia Raman de baixa resolução amplificada por superfície**. 2009. 127 f. Tese (Doutorado em Ciências) – Instituto de Química, Universidade Estadual de Campinas, Campinas, 2009.

- 65 OLIVIERI, A. C. et al. Uncertainty estimation and figures of merit for multivariate calibration. **Pure and Applied Chemistry**, v. 78, n. 3, p. 633-661, Mar. 2006.
- 66 FABER, N. M. Efficient computation of net analyte signal vector in inverse multivariate calibration models. **Analytical Chemistry**, v. 70, n. 23, p. 5108-5110, Dec. 1998.
- 67 ROCHA, W. F. C. et al. Determination and validation of nimesulide in pharmaceutical formulation by near infrared spectroscopy. **Journal of the Brazilian Chemical Society**, v. 21, n. 10, p. 1929-1936, 2010.
- 68 BRÁS, L. G. P. et al. Multiblock PLS as an approach to compare and combine NIR and MIR spectra in calibrations of soybean flour. **Chemometrics and Intelligent Laboratory Systems**, v. 75, n. 1, p. 91-99, Jan. 2005.
- 69 SARRAGUCA, M. C.; LOPES, J. A. The use of net analyte signal (NAS) in near infrared spectroscopy pharmaceutical applications: interpretability and figures of merit. **Analytica Chimica Acta**, v. 642, n. 1/2, p. 179-185, May 2009.
- 70 BLANCO, M. et al. Determination of low analyte concentrations by near-infrared spectroscopy: effect of spectral pretreatments and estimation of multivariate detection limits. **Analytica Chimica Acta**, v. 581, n. 2, p. 318-323, Jan. 2007.
- 71 LIU, L.; ARNOLD, M. A. Selectivity for glucose, glucose-6-phosphate, and pyruvate in ternary mixtures from the multivariate analysis of near-infrared spectra. **Analytical and Bioanalytical Chemistry**, v. 393, n. 2, p. 669-677, Jan. 2009.
- 72 FELIPE-SOTELO, M. et al. Linear PLS regression to cope with interferences of major concomitants in the determination of antimony by ETAAS. **Journal of Analytical Atomic Spectrometry**, v. 21, n. 1, p. 61-68, Jan. 2006.
- 73 VALDERRAMA, P. **Avaliação de figuras de mérito em calibração multivariada na determinação de parâmetros de controle de qualidade em indústria alcooleira por espectroscopia no infravermelho próximo**. 2005. 128 f. Dissertação (Mestrado em Química) - Instituto de Química, Universidade Estadual de Campinas, Campinas, 2005.
- 74 INSTITUTO NACIONAL DE METROLOGIA, NORMALIZAÇÃO E QUALIDADE INDUSTRIAL. **Vocabulário internacional de metrologia: conceitos fundamentais e gerais e termos associados (VIM 2008)**. Rio de Janeiro, 2009. Disponível em: <http://www.inmetro.gov.br/metcientifica/vim/vim_completo.pdf>. Acesso em: 2 nov. 2010.
- 75 INSTITUTO NACIONAL DE METROLOGIA, NORMALIZAÇÃO E QUALIDADE INDUSTRIAL. **NIT-DICLA-021: expressão da incerteza de medição por laboratórios de calibração**. 2013. Rev n° 09. Disponível em: <http://www.inmetro.gov.br/Sidoq/pesquisa_link.asp?seq_tipo_documento=4&cod_uo_numeracao=00587&num_documento=021>. Acesso em: 21 out. 2014.

76 INSTITUTO NACIONAL DE METROLOGIA, NORMALIZAÇÃO E QUALIDADE INDUSTRIAL. **Guia para a expressão de incerteza de medição: GUM 2008:** avaliação de dados de medição. Duque de Caxias, 2012. Disponível em: <http://www.inmetro.gov.br/inovacao/publicacoes/gum_final.pdf>. Acesso em: 21 out. 2014.

77 COUTO, P. R. G. **Estimativa da incerteza da massa específica da gasolina pelo ISO GUM 95 e método de Monte Carlo e seu impacto na transferência de custódia.** 2006. 143 f. Dissertação (Mestrado em Tecnologia de Processos Químicos e Bioquímicos) - Escola de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

78 COUTO, P. R. G.; DAMASCENO, J. C.; OLIVEIRA, S. P. D. Monte Carlo simulations applied to uncertainty in measurement. In: CHAN, V. W. K. (Ed.). **Theory and applications of Monte Carlo simulations.** Rijeka: Intech, 2013. Chap. 2, p. 27-51.

79 JOINT COMMITTEE FOR GUIDES IN METROLOGY. **JCGM 101:2008:** evaluation of measurement data: supplement 1 to the "Guide to the expression of uncertainty in measurement": propagation of distributions using a Monte Carlo method. 2008. Disponível em: <http://www.bipm.org/utls/common/documents/jcgm/JCGM_101_2008_E.pdf>. Acesso em: 21 out. 2014.

80 DUBOIS, A.; KULPA, L.; SOUZA, L. E. de. **Gestão de custos e formação de preço:** conceitos, modelos e instrumentos: abordagem do capital de giro e da margem de competitividade. 3. ed. São Paulo: Atlas, 2009. 272 p.

81 SOCRATES, G. **Infrared and Raman characteristic group frequencies:** tables and charts. 3rd ed. Chichester: Wiley & Sons, 2001. 366 p.

82 BREITKREITZ, M. C. et al. Determination of total sulfur in diesel fuel employing NIR spectroscopy and multivariate calibration. **Analyst**, v. 128, n. 9, p. 1204-1207, Sept. 2003.

83 SIESLER, H. W. et al. (Ed.). **Near-infrared spectroscopy:** principles, instruments, applications. Weinheim: Wiley-VCH, 2002. 361 p.

84 CARNEIRO, H. S. P. et al. Determination of ethanol fuel adulteration by methanol using partial least-squares models based on fourier transform techniques. **Energy Fuels**, v. 22, n. 4, p. 2767-2770, May 2008.

85 SILVA, M. A. et al. Development and analytical validation of a multivariate calibration method for determination of amoxicillin in suspension formulations by near infrared spectroscopy. **Talanta**, v. 89, p. 342-51, Jan. 2012.

86 VALDERRAMA, P. et al. Quantification of chlorpheniramine maleate enantiomers by ultraviolet spectroscopy and chemometric methods. **Analytical and Bioanalytical Chemistry**, v. 397, n. 1, p. 181-188, May 2010.

87 OLIVERI, A. C.; GOICOECHEA, H. C.; INON, F. A. MVC1: an integrated MatLab toolbox for first-order multivariate calibration. **Chemometrics and Intelligent Laboratory Systems**, v. 73, n. 2, p. 189-197, Oct. 2004.

88 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **E1064**: standard test method for water in organic liquids by coulometric Karl Fischer titration. West Conshohocken, 2012. 5 p.

89 AMERICAN SOCIETY FOR TESTING AND MATERIALS. **D5501**: standard test method for determination of ethanol and methanol content in fuels containing greater than 20% ethanol by gas chromatography. West Conshohocken, 2012. 13 p.

Anexo A – Algoritmo Kennard Stone

```
% -----
% Function:
% [model,test]=kenstone(X,k)
% -----
% Aim:
% Uniform subset selection with Kennard and Stone algorithm
% -----
% Input:
% X, matrix (n,p), predictor variables in columns
% k, number of objects to be selected to the model set
% -----
% Output:
% model, vector (k,1), list of objects selected to model set
% test, vector (n-k,1), list of objects selected to test set (optionally)
% -----
% Example:
% X=randn(300,2);
% [model,test]=kenstone(X,20)
% [model]=kenstone(X,20)
% plot(X(test,1),X(test,2),'k. ');hold on;
% plot(X(model,1),X(model,2),'rs','markerfacecolor','r');
% figure(gcf)
% -----
% References:
% [1] R.W. Kennard, L.A. Stone, Computer aided design of experiments,
% Technometrics 11 (1969) 137-148
% [2] M. Daszykowski, B. Walczak, D.L. Massart, Representative subset selection,
% Analytica Chimica Acta 468 (2002) 91-103
% -----
% Written by Michal Daszykowski
% Department of Chemometrics, Institute of Chemistry,
% The University of Silesia
% December 2004
% http://www.chemometria.us.edu.pl

function [model,test]=kenstone(X,k)

[m,n]=size(X);
if k>=m | k<=0
    h=errorlg('Wrongly specified number of objects to be selected to model set.','Error');
    model=[];
    if nargin==2
        test=[];
    end
    waitfor(h)
    return
end

x=[1:size(X,1)]' X;
n=size(x,2);
[i1,ind1]=min(fastdist(mean(x(:,2:n)),x(:,2:n)));
model(1)=x(ind1,1);
x(ind1,:)=[];

[i2,ind2]=max(fastdist(X(model(1),:),x(:,2:n)));
model(2)=x(ind2,1);
x(ind2,:)=[];
```

```

for d=3:k
    [ii,ww]=max(min(fastdist(x(:,2:n),X(model,:))));
    model(d)=x(ww,1);
    x(ww,:)=[];
end

if nargout==2
    test=1:size(X,1);
    test(model)=[];
end

% --->

function D=fastdist(x,y)

% Calculates squared Euclidean distances between two sets of objects

D=((sum(y'.^2))*ones(1,size(x,1)))+(ones(size(y,1),1)*(sum(x'.^2)))-2*(y*x');

```

Fonte: disponível em: <<http://chemometria.us.edu.pl/download/kenstone.M>>.
Acesso em: 01 mar. 2015.

ANEXO B – Algoritmo fmerito

```
function saida=fmerito(Rc,cc,mod);
%-----
% Este programa determina o net analyte signal para uma matriz de espectros
% por pls, tendo como algoritmo base o NIPLS. O grafico de rmsec versus rmsecv
% e determinado por validação cruzada contiguous block.
%
% Sintaxe: saida=fmerito(Rc,cc,mod);
%
% Entradas:
% Rc= matriz de espectros previamente pre-processados.
% cc= vetor de valores de referencia para a propriedade de interesse.
% mod =modelo criado a partir do analysis
% nvl=numero de variaveis latentes utilizadas no modelo pls
%
% Saidas:
%
% nas_c = matriz que conten os vetores nas para as l amostras.
%
% nnas_c = valor escalar do nas(norma euclidiana dos vetores nas)
%
% s = vetor de sensibilidade. Relata a sensibilidade em cada variavel. A sensibilidade
% do modelo e a norma euclidiana de "s".
%
% sel = vetor que conten a seletividade de cada amostra.
%
% sen= sensibilidade do modelo.
%
% H = fator de correção para o calculo do vetor nas para amostras futuras.
% Para um conjunto de amostras desconhecidas(Xv)
%
% constante=escalar utilizado para corrigir os valores de nas
% centrados na media;
%-----
[L1,C1]=size(Rc);
pre=mod.detail.preprocessing{1} (1).keyword;
b=mod.reg;
yest=mod.pred{2};
my=mean(cc);
% calculo do nnas_c
nnas_c=(Rc*b)/norm(b);

if isempty(pre)==0
    Z(1:L1,1)=1;
    Z(1:L1,2)=nnas_c;
    ccc=(cc(:,1)+ones(L1,1))*my;
    cte=pinv(Z)*ccc;
    constante=-cte(1,1)/cte(2,1);
    Nas_c(:,1)=nnas_c(:,1)-constante;
else
    Nas_c=nnas_c;
end

% Fator de correção
H=b/norm(b);

% Sensibilidade(escalar)
```

```

sen=1/norm(b);

% seletividade
for i=1:L1
    sel=nnas_c./norm(Rc(i,:));
end

% Vetor nas para as I amostras de R

nas_c=nnas_c*norm(b)*pinv(b);

% Vetor de sensibilidade
for i=1:L1
    s(i,:)=nas_c(i,:)./yest(i,1);
end
size(s)
s=mean(s);

%saidas
saida.nnas_c=Nas_c;
saida.sel=sel;
saida.sen=sen;
%saida.sen_vetor=s;
saida.H=H;
saida.constante=constante;

```

Fonte: POPPI, R. J. Quimiometria – Figuras de mérito. [mensagem pessoal]. Mensagem recebida por <hatanaka@iq.unesp.br> data de recebimento, 13 out. 2010.

ANEXO C – Algoritmo figmerito

```

function saida = figmerit(x,y,b)
%FIGMERIT Analytical figures of merit for multivariate calibration.
% Calculates analytical figures of merit for PLS and PCR models.
% The inputs are the preprocessed (usually centered and scaled)
% data (x), the preprocessed analyte data (y), and the
% regression vector, (b). Note that for standard PCR and PLS model
% structures that b = model.reg.
% Outputs are the matrix of net analyte signals (nas) for each row
% of (x), the "norm" of the net analyte signal for each row (nnas)
% [with sign], the matrix of sensitivities for each sample (sens),
% and the vector of selectivities for each sample (sel) [which are
% positive regardless of the sign of (nas)].
%
% An earlier version of FIGMERIT estimated a "noise filtered" NAS.
% This is no longer estimated because an improved method for
% calculating the NAS makes it redundant.
%
%Example: given the 7 LV PLS model formed from
% model = pls(x,y,7);
%
%I/O: [nas,nnas,sens,sel] = figmerit(x,y,b);
%I/O: figmerit demo
%
%See also: ANALYSIS, LEVERAG, PLS

%Copyright Eigenvector Research, Inc. 1997-2005
%Licensee shall not re-compile, translate or convert "M-files" contained
% in PLS_Toolbox for use with any software other than MATLAB®, without
% written permission from Eigenvector Research, Inc.
%Barry M. Wise May 30, 1997

% aug 2002, rb, changed to new definition given by the regression vector
% according to Bro, Andersen, "Theory of net analyte signal vectors in
% inverse regression", Journal of Chemometrics
% 4/03 nbg changed the help. removed the redundant b(:)
%

if nargin == 0; x = 'io'; end
varargin{1} = x;
if ischar(varargin{1});
    options = [];
    if nargout==0; evriio(mfilename,varargin{1},options); else; nas =
evriio(mfilename,varargin{1},options); end
    return;
end

warning off backtrace

if nargout == 5
    warning(' In earlier versions of FIGMERIT, a fifth output was possible but is redundant (=nas)
now')
end
if min(size(b))>1
    error(' Third input must be a regression vector for one response')
end

[mx,nx] = size(x);

```

```

nas    = zeros(mx,nx);
nnas   = zeros(mx,1);
sel    = zeros(mx,1);
b      = b(:);

nas    = ((x*b)*inv(b'*b))*b';
sens   = nas;
nnas   = x*b/norm(b);
for i=1:mx
    sel(i) = abs(nnas(i)/norm(x(i,:)));
    sens(i,:) = sens(i,+)/y(i);
end
%Modificado: saídas das variáveis juntas.

saida2.nas=nas;
saida2.nnas=nnas;
saida2.sens=sens;
saida2.sel=sel;

```

Fonte: EIGENVECTOR RESEARCH, INC. PLS_Toolbox. 7.3.1. [S.I.]:
Eigenvector Research, Inc., 2013. Conjunto de programas. 1 CD-ROM.