

**PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS E PROCESSOS
INDUSTRIAIS – MESTRADO
ÁREA DE CONCENTRAÇÃO EM CONTROLE E OTIMIZAÇÃO DE
PROCESSOS INDUSTRIAIS**

Ricardo Edler Rotta

**DETERMINAÇÃO SIMULTÂNEA DOS PARÂMETROS DE QUALIDADE, PONTO
DE FULGOR, PERCENTAGEM DE BODIESEL, MASSA ESPECÍFICA E
ENXOFRE, NAS BLENDS BODIESEL/DIESEL, EMPREGANDO REFLEXÃO NO
INFRAVERMELHO MÉDIO E REGRESSÃO MULTIVARIADA**

Santa Cruz do Sul, abril de 2010

Ricardo Edler Rotta

DETERMINAÇÃO SIMULTÂNEA DOS PARÂMETROS DE QUALIDADE, PONTO DE FULGOR, PERCENTAGEM DE BIODIESEL, MASSA ESPECÍFICA E ENXOFRE, NAS BLENDS BIODIESEL/DIESEL, EMPREGANDO REFLEXÃO NO INFRAVERMELHO MÉDIO E REGRESSÃO MULTIVARIADA

Dissertação apresentada ao Programa de Pós-Graduação em Sistemas e Processos Industriais - Mestrado, Área de Concentração em Controle e Otimização de Processos Industriais, Universidade de Santa Cruz do Sul, como requisito parcial para obtenção do título de Mestre em Sistemas e Processos Industriais.

Orientador: Prof. Dr. Marco Flôres Ferrão

Santa Cruz do Sul, abril de 2010

Ricardo Edler Rotta

DETERMINAÇÃO SIMULTÂNEA DOS PARÂMETROS DE QUALIDADE, PONTO DE FULGOR, PERCENTAGEM DE BODIESEL, MASSA ESPECÍFICA E ENXOFRE, NAS BLENDS BODIESEL/DIESEL, EMPREGANDO REFLEXÃO NO INFRAVERMELHO MÉDIO E REGRESSÃO MULTIVARIADA

Esta Dissertação foi submetida ao Programa de Pós-Graduação em Sistemas e Processos Industriais – Mestrado – Universidade de Santa Cruz do Sul – UNISC, como requisito parcial para obtenção do título de Mestre em Sistemas e Processos Industriais.

Dr. Marco Flôres Ferrão
Professor Orientador

Dr. Luciano Marder

Dr. Adilson Ben da Costa

Para minha filha Maria Eduarda

AGRADECIMENTOS

A Deus, por iluminar meu caminho;

À minha filha Maria Eduarda, simplesmente por existir;

Aos meus pais, Luiz Antônio Rotta e Helena Audino Edler, por todo o apoio, dedicação e carinho;

Ao meu orientador prof. Dr. Marco Flôres Ferrão, pelas muitas orientações, pelo senso de direção e pela disponibilidade, fatores determinantes para a realização deste trabalho;

À minha namorada, pelo carinho e conforto nos momentos difíceis.

À empresa onde trabalho, à chefia, pares e minha equipe que compreenderam e deram suporte durante os períodos que necessitei estar ausente.

À minha família e amigos, por compreenderem minha ausência nos momentos de confraternização;

Aos professores e secretárias do PPGSPI e ao bolsista Rodrigo Lovato pela colaboração nas análises.

A todos que não foram citados aqui, mas contribuíram para o bom andamento deste trabalho.

“Insanidade é fazer a mesma coisa e esperar resultados diferentes”

Albert Einstein

RESUMO

A busca por novas fontes de recursos renováveis tem gerado várias discussões no âmbito mundial, especialmente relativas ao petróleo que representa 35% da matriz energética no mundo. No Brasil tem-se investido na produção e utilização do Biodiesel, sendo comercializado desde o início de 2008 em mistura com o diesel numa proporção de 2% (B2), sendo esta proporção aumentada, a partir de janeiro de 2010, para 5% (B5). Este trabalho abordou a obtenção de espectros no infravermelho médio com transformada de Fourier, utilizando as técnicas de Transreflectância e Reflexão Total Atenuada, sendo que a segunda apresentou melhores resultados e com a qual os modelos multivariados foram construídos. Para tanto, foram preparadas 45 amostras, que utilizaram dois tipos de diesel, metropolitano e interiorano (que diferem na quantidade de enxofre) para formar blendas com diferentes concentrações de biodiesel, que foram de 8 a 30% (v/v). Para a construção dos modelos, esta dissertação propõe um estudo e o uso do método dos Mínimos Quadrados Parciais (PLS) juntamente com técnicas de seleção de variáveis iPLS e biPLS, para a construção de modelos com maior capacidade de predição, uma vez que utiliza faixas espectrais com maior correlação com a variável estudada. Este trabalho buscou não somente quantificar o biodiesel nas blendas biodiesel/diesel, mas também a quantidade de enxofre, massa específica e ponto de fulgor. Para a quantificação de enxofre o modelo mais robusto foi o que utilizou o algoritmo siPLS com um RMSEP de 16,9 mg/kg e combinou as faixas espectrais de 858 - 960 cm^{-1} e 2732 - 2834 cm^{-1} . Para a predição da quantidade de biodiesel o modelo iPLS foi escolhido, pois alcançou resultados superiores ao siPLS, com um RMSEP de 0,154% (v/v) e utilizou a faixa espectral de 1068 - 1484 cm^{-1} que corresponde ao estiramento assimétrico C-C(O)- dos ésteres do biodiesel. Na determinação da massa específica, o modelo que melhor se ajustou foi o iPLS, por utilizar menor número de variáveis latentes, com um RMSEP de 0,865 kg/m^3 utilizando a faixa espectral de 1274 - 1376 cm^{-1} . Finalmente, para o ponto de fulgor, os modelos iPLS e siPLS alcançaram a mesma performance tanto nas mediadas RMSEP quanto em número de variáveis latentes, sendo escolhido o iPLS com

RMSEP de 1,38°C , que utilizou a faixa espectral 1274 - 1376 cm^{-1} . Todos os modelos que utilizaram o espectro inteiro mostraram capacidades de predição inferiores aos que foram submetidos a procedimentos de seleção de variáveis iPLS ou siPLS, sendo, a seleção de variáveis por métodos determinísticos, uma estratégia promissora para a determinação de propriedades físicas, químicas e físico-químicas.

Palavras-chave: Biodiesel, determinação do ponto de fulgor, espectrometria no infravermelho, regressão por mínimos quadrados parciais, seleção de variáveis.

ABSTRACT

The search for new sources of renewable resources has started various discussions worldwide, especially related to petroleum that represents 35% of all energy resources in the world. Brazil has invested in the production and use of Biodiesel, being commercialized since the beginning of 2008 in mixture with diesel in the proportion of 2% (B2), having this proportion increased, starting in January of 2010, to 5% (B5). This work approaches the use of Fourier transform infrared spectroscopy (FTIR), using transreflectance and total attenuated reflectance (ATR) techniques, being that the second presented better results and it was with it that multivariate regression models were built. Therefore, the 45 samples that utilized metropolitan and countryside diesel (that differed in the quantity of sulfur) were prepared to form blends with different concentrations of biodiesel, which were of 8 to 30% (v/v). For the built of both models, this work proposes a study and the use of the partial least squares (PLS) together with variable selection techniques iPLS and biPLS, for a built of models with greater prediction capacity, which uses spectrum ranges with greater correlation with the studied variable. This work does not only look to quantify the biodiesel in the blends biodiesel/diesel, but also the effect of this quantity on the property quantity of sulfur, specific mass and flash point. For the quantification of sulfur, the more robust model, the algorithm siPLS with a RMSEP of 16.9 mg/kg and combined the spectrum ranges of 858 - 960 cm^{-1} and 2732 - 2834 cm^{-1} was used. For the prediction of biodiesel content the iPLS model was chosen, since it reached results superior to siPLS, with a RMSEP of 0.154% (v/v) and used a spectrum range of 1068 - 1484 cm^{-1} that corresponds to the asymmetric stretching C-C(O) of the esters of biodiesel. In the determination of specific mass the model that most adjusted was the iPLS, for using a fewer number of latent variables, with a RMSEP of 0.865 kg/m^3 when range of 1274 - 1376 cm^{-1} was used. Finally, for the flash point, the iPLS and siPLS models reached the same performance as the RMSEP measures as well as in number of latent variables, being chosen the iPLS with RMSEP of 1.38 $^{\circ}\text{C}$, when the spectral range 1274 - 1376 cm^{-1} was used. All the models that utilized the full spectrum showed prediction capacity inferior to those

submitted to iPLS or siPLS variable selection procedures, being a variable selection for deterministic methods, a promising strategy for the determination of physical, chemical, and physical-chemical properties in biodiesel/diesel blends.

Key words: Biodiesel/diesel blends, flash point determination, infrared spectroscopy, partial least squares regression, variable selection.

LISTA DE ILUSTRAÇÕES

Figura 1 - Consumo no mundo por tipo de combustíveis.	17
Figura 2 - Tipos de deformações que sofrem as ligações quando expostas ao Infravermelho.	22
Figura 3 - Esquema da reflexão interna em um elemento de ATR.....	24
Figura 4 - Representação da propagação da radiação infravermelha através do elemento de ATR.	25
Figura 5 - Esquema que descreve princípio da transreflectância, utilizando o acessório de DRIFTS.	26
Figura 6 - Quadro ilustrativo das etapas envolvidas na realização da pesquisa.	35
Figura 7 - Distribuição das amostras de calibração e de previsão das para as 43 amostras das blendas biodiesel/diesel, variando de 8 a 30 %(v/v).	39
Figura 8 - Três modelos de chapas propostos, com acabamento controlado	41
Figura 9 - Gráfico com os espectros dos três modelos de acessórios (“chapas”).	42
Figura 10 - Dendrograma representativo do agrupamento dos três modelos de chapas, utilizando pré-processamento: normalizados pela variância, distância euclidiana e método de lincagem incremental com transformada SNV	42
Figura 11 - Posicionamento da chapa sobre o acessório de reflexão especular PIKE	43
Figura 12 - Dendrograma representativo da primeira parcela de agrupamentos das triplicadas das quatro chapas modelo A, utilizando pré-processamento: centrado na média, distância euclidiana e método de lincagem flexível com transformada SNV..	44
Figura 13 - Dendrograma representativo da segunda parcela de agrupamentos das triplicadas das quatro chapas modelo A, utilizando pré-processamento: centrado na média, distância euclidiana e método de lincagem flexível com transformada SNV..	45
Figura 14 - Dendrograma representativo da terceira parcela de agrupamentos das triplicadas das quatro chapas modelo A, utilizando pré-processamento: centrado na média, distância euclidiana e método de lincagem flexível com transformada SNV..	45
Figura 15 - Dendrograma representativo da quarta parcela de agrupamentos das triplicadas das quatro chapas modelo A, utilizando pré-processamento: centrado na média, distância euclidiana e método de lincagem flexível com transformada SNV..	46

Figura 16 - Dendrograma representativo da quinta parcela de agrupamentos das triplicadas das quatro chapas modelo A, utilizando pré-processamento: centrado na média, distância euclidiana e método de ligação flexível com transformada SNV..46

Figura 17 - Gráfico com os espectros ATR – FTIR para as quarenta e cinco blendas com faixa concentração variando de 8 a 30,1 %(v/v)..... 47

Figura 18 - Gráfico do RMSECV para os modelos dos 16 intervalos (barras) e modelo global empregando 4 variáveis latentes (linha pontilhada) para a determinação do enxofre nas blendas biodiesel/diesel, com dados centrados na média e correção do espalhamento de luz..... 51

Figura 19 - Gráfico da correlação entre o método de referência e a previsão por iPLS/ATR-FTIR, empregando a faixa espectral 858 - 1064 cm^{-1} (intervalo 15) e 4 variáveis latentes para a determinação do enxofre nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz..... 52

Figura 20 - Gráfico das faixas espectrais utilizadas na construção do modelo siPLS para a determinação do enxofre nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz. 54

Figura 21 - Gráfico da correlação entre o método de referência e a previsão por siPLS/ATR-FTIR, empregando as faixas espectrais 858 - 960 cm^{-1} e 2732-2834 cm^{-1} (intervalos 12 e 30) e 4 variáveis latentes para a determinação do enxofre nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz. 55

Figura 22 - Gráfico do RMSECV para os modelos dos 8 intervalos (barras) e modelo global empregando 1 variáveis latentes (linha pontilhada) para a determinação da % de biodiesel nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz. 59

Figura 23 - Gráfico da correlação entre o método de referência e a previsão por iPLS/ATR-FTIR, empregando a faixa espectral 1068 - 1484 cm^{-1} (intervalo 7) e 1 variável latente para a determinação da % de biodiesel nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz..... 60

Figura 24 - Gráfico das faixas espectrais utilizadas na construção do modelo siPLS para a determinação da % de biodiesel nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz. 62

Figura 25 - Gráfico da correlação entre o método de referência e a previsão por siPLS/ATR-FTIR, empregando as faixas espectrais 856 - 754 cm^{-1} , 962 - 1064 cm^{-1} e 2732 – 2834 cm^{-1} (intervalos 12, 29 e 31) e 4 variáveis latentes para a determinação da % de biodiesel nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz. 63

Figura 26 - Gráfico do RMSECV para os modelos dos 32 intervalos (barras) e modelo global empregando 3 variáveis latentes (linha pontilhada) para a

determinação da massa específica nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.....	67
Figura 27 - Gráfico da correlação entre o método de referência e a previsão por iPLS/ATR-FTIR, empregando a faixa espectral 1274 - 1376 cm^{-1} (intervalo 26) e 3 variáveis latentes para a determinação da massa específica nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.....	68
Figura 28 - Gráfico das faixas espectrais utilizadas na construção do modelo siPLS para a determinação da massa específica nas blendas biodiesel/diesel.	70
Figura 29 - Gráfico da correlação entre o método de referência e a previsão por siPLS/ATR-FTIR, empregando as faixas espectrais 1274 - 1376 cm^{-1} e 1691 - 1793 cm^{-1} e 9 variáveis latentes para a determinação da massa específica nas blendas biodiesel/diesel.....	71
Figura 30 - Gráfico do RMSECV para os modelos dos 32 intervalos (barras) e modelo global empregando 7 variáveis latentes (linha pontilhada) para a determinação do ponto de fulgor nas blendas biodiesel/diesel, com dados autoescalados.....	75
Figura 31 - Gráfico da correlação entre o método de referência e a previsão por iPLS/ATR-FTIR, empregando a faixa espectral 1274 - 1376 cm^{-1} (intervalo 26) e 7 variáveis latentes para a determinação do ponto de fulgor nas blendas biodiesel/diesel, com dados autoescalados.	76
Figura 32 - Gráfico do RMSECV para os modelos dos 32 intervalos (barras) e modelo global empregando 3 variáveis latentes (linha pontilhada) para a determinação do ponto de fulgor nas blendas biodiesel/diesel.	77
Figura 33 - Gráfico da correlação entre o método de referência e a previsão por iPLS/ATR-FTIR, empregando a faixa espectral 1274 - 1376 cm^{-1} (intervalo 26) e 3 variáveis latentes para a determinação do ponto de fulgor nas blendas biodiesel/diesel, com dados centrados na média e correção do espalhamento de luz.....	78
Figura 34 - Gráfico das faixas espectrais utilizadas na construção do modelo siPLS para a determinação do ponto de fulgor nas blendas biodiesel/diesel, com dados autoescalados	80
Figura 35 - Gráfico da correlação entre o método de referência e a previsão por siPLS/ATR-FTIR, empregando as faixas espectrais 858 - 960 cm^{-1} , 1170 - 1376 cm^{-1} (intervalos 26, 27 e 30) com 7 variáveis latentes para a determinação do ponto de fulgor nas blendas biodiesel/diesel, com os dados autoescalados.	81

LISTA DE TABELAS

Tabela 1 – Valores obtidos das propriedades pelo método referência para cada amostra das blendas de biodiesel/diesel.....	37
Tabela 2 – Valores das figuras de mérito para a determinação do enxofre nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR brutos.....	48
Tabela 3 – Valores das figuras de mérito para a determinação do enxofre nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR centrados na média ..	49
Tabela 4 – Valores das figuras de mérito para a determinação do enxofre nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR autoescalados.....	49
Tabela 5 – Valores das figuras de mérito para a determinação do enxofre nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR centrados na média e correção do espalhamento de luz (MSC)	50
Tabela 6 – Valores das figuras de mérito para a determinação do enxofre nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR autoescalados e correção do espalhamento de luz (MSC)	50
Tabela 7 – Valores para os erros de validação cruzada, calibração e previsão para os modelos siPLS, visando a determinação do enxofre nas blendas biodiesel/diesel empregando os dados ATR – FTIR.....	53
Tabela 8 – Valores das figuras de mérito para a determinação da porcentagem de biodiesel nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR brutos..	56
Tabela 9 – Valores das figuras de mérito para a determinação da porcentagem de biodiesel nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR centrados na média.....	56
Tabela 10 – Valores das figuras de mérito para a determinação da porcentagem de biodiesel nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR autoescalados	57
Tabela 11 – Valores das figuras de mérito para a determinação da porcentagem de biodiesel nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR centrados na média e correção do espalhamento de luz (MSC).	58
Tabela 12 – Valores das figuras de mérito para a determinação da porcentagem de biodiesel nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR autoescalados e correção do espalhamento de luz (MSC).	58
Tabela 13 – Valores para os erros de validação cruzada, calibração e previsão para os modelos siPLS, visando a determinação da % de biodiesel nas blendas biodiesel/diesel empregando os dados ATR – FTIR.	61

Tabela 14 – Valores das figuras de mérito para a determinação da massa específica das blendas Biodiesel/Diesel empregando os dados ATR – FTIR brutos	64
Tabela 15 – Valores das figuras de mérito para a determinação da massa específica nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR centrados na média	64
Tabela 16 – Valores das figuras de mérito para a determinação da massa específica nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR autoescalados	65
Tabela 17 – Valores das figuras de mérito para a determinação da massa específica nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR centrados na média e correção do espalhamento de luz (MSC)	65
Tabela 18 – Valores das figuras de mérito para a determinação da massa específica nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR autoescalados e correção do espalhamento de luz (MSC)	66
Tabela 19 – Valores para os erros de validação cruzada, calibração e previsão para os modelos siPLS, visando a determinação da massa específica nas blendas biodiesel/diesel empregando os dados ATR – FTIR	69
Tabela 20 – Valores das figuras de mérito para a determinação do ponto de fulgor nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR brutos	72
Tabela 21 – Valores das figuras de mérito para a determinação do ponto de fulgor nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR centrados na média	72
Tabela 22 – Valores das figuras de mérito para a determinação do ponto de fulgor nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR autoescalados	73
Tabela 23 – Valores das figuras de mérito para a determinação do ponto de fulgor nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR centrados na média e correção do espalhamento de luz (MSC)	73
Tabela 24 – Valores das figuras de mérito para a determinação do ponto de fulgor nas blendas Biodiesel/Diesel empregando os dados ATR – FTIR autoescalados e correção do espalhamento de luz (MSC)	74
Tabela 25 – Valores para os erros de validação cruzada, calibração e previsão para os modelos siPLS, visando a determinação do ponto de fulgor nas blendas biodiesel/diesel empregando os dados ATR – FTIR	79

LISTA DE ABREVIATURAS

ABNT	Associação Brasileira de Normas Técnicas
ANP	Agência Nacional do Petróleo
ASTM	American Society for Testing and Materials
ATR	Attenuated Total Reflectance
CNPE	Conselho Nacional de Política Energética
FT-IR	Fourier Transform Infrared Spectroscopy
HCA	Hierarchical Clustering Algorithms
IEA	International Energy Agency
iPLS	Interval Partial Least Squares
MLR	Multiple Linear Regression
mwPLS	Moving Window Partial Least Squares
NAS	Net Analyte Signal
PCA	Principal Component Analysis
PCR	Principal Components Regression
PLS	Partial Least Squares
RMSE	Root Mean Square Error
RMSEC	Root Mean Square Error of Calibration
RMSECV	Root Mean Square Error of Cross-Validation
RMSEP	Root Mean Square Error of Prediction
siPLS	Synergy Interval Partial Least Squares

SUMÁRIO

INTRODUÇÃO	17
1. APRESENTAÇÃO	17
1.1 Justificativas	18
1.2 Objetivos	19
1.2.1 Objetivo Geral	19
1.2.2 Objetivos Específicos	19
2 FUNDAMENTAÇÃO TEÓRICA.....	20
2.1 Biodiesel.....	20
2.2 Espectroscopia no Infravermelho	21
2.2.1 Reflexão Total Atenuada	23
2.2.2 Transreflectância.....	25
2.3 Técnicas de Análise Multivariada	26
2.3.2 Seleção de Variáveis.....	29
2.4 Recentes aplicações da espectroscopia no infravermelho e análise multivariada na caracterização do biodiesel e suas blendas com diesel.....	32
3 METODOLOGIA.....	34
3.1 Descrição	34
3.2 Etapas	34
4 RESULTADOS	40
4.1 Resultados para o acessório de transreflectância.....	40
4.2 Resultados ATR	47
4.2.1 Enxofre	48
4.2.2 Biodiesel.....	55
4.2.3 Massa específica.....	63
4.2.4 Ponto de fulgor	71
CONCLUSÕES	82
REFERÊNCIAS.....	84

INTRODUÇÃO

1. APRESENTAÇÃO

Ainda que numa escala geológica, de tempo, os combustíveis fósseis continuem sendo gerados, a partir da decomposição de matéria orgânica, não são suficientes para atender a enorme demanda mundial ao longo da escala humana de tempo.

Não só a escassez desses combustíveis, consumido pelo homem há séculos, tem causado sérias transformações nas sociedades contemporâneas, como sua queima para gerar energia é responsável pela emissão de enormes quantidades de gases que estão diretamente relacionados com o aquecimento e com a formação de chuvas ácidas.

Com exceção de hidrelétricas e da energia nuclear, a maior parte de toda a energia consumida no mundo vem do petróleo, do carvão e do gás natural (INTERNATIONAL ENERGY AGENCY – IEA, 2007). A Figura 1 mostra a distribuição do consumo dos recursos energéticos no mundo.

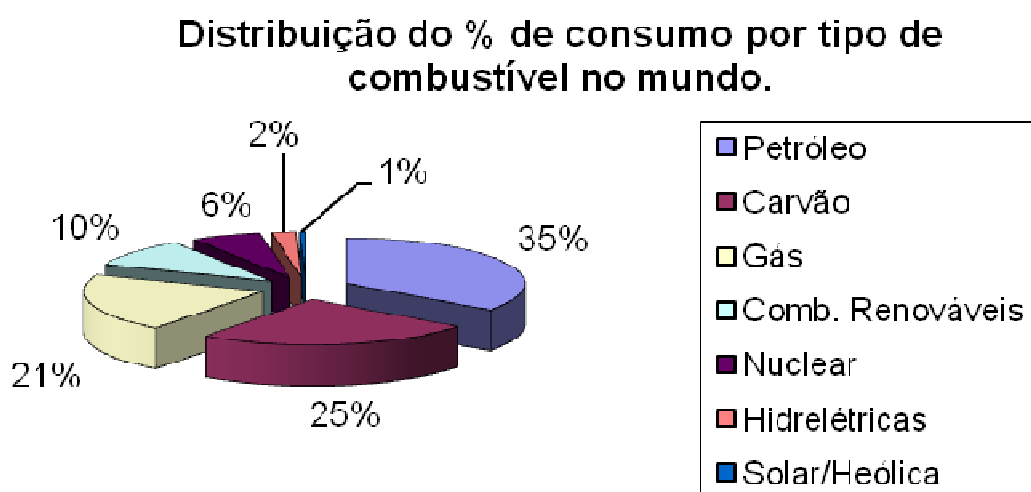


Figura 1 – Consumo no mundo por tipo de combustíveis.
Fonte: INTERNATIONAL ENERGY AGENCY – IEA (2007)

Como essas fontes são limitadas, por serem de origem fóssil, e deverão se esgotar no futuro, a visão de uma fonte alternativa de energia é de vital importância. Neste contexto, o biodiesel é um promissor substituto dos combustíveis fósseis, pois é proveniente de recursos renováveis.

1.1 Justificativas

No Brasil tem-se investido na produção e utilização do biodiesel, sendo comercializado desde o início de 2008 em mistura com o diesel numa proporção de 2%. Esta percentagem veio a aumentar a partir de 2010, por determinação da Resolução CNPE nº6 de 16/09/2009, que definiu a proporção de 5% a partir de janeiro de 2010.

Partindo deste ponto, o controle da porcentagem da mistura é essencial para garantir o atendimento à legislação. Outros pontos a serem citados são as propriedades físicas e químicas deste combustível que podem afetar a funcionabilidade do motor de combustão interna Diesel.

A produção de biocombustíveis alternativos para o óleo diesel, provenientes de óleos vegetais brutos, tem sido tema de diversos estudos (NASCIMENTO, COSTA NETO E MAZUCO, 2001).

Apesar de favorável energeticamente, o uso direto desse produto em motores é problemático devido à sua viscosidade ser maior que a do óleo diesel e baixa volatilidade, o que impede a sua queima completamente. Isto pode acarretar em acúmulo de resíduo não queimado nos bicos injetores dos motores e conseqüente dano ou perda no rendimento. Portanto, visando a reduzir a alta viscosidade dos óleos vegetais, diferentes alternativas têm sido consideradas, sendo, a transesterificação a mais recomendada (MARQUES et al., 2008).

O biodiesel é obtido através da transesterificação que nada mais é do que a separação da glicerina do óleo vegetal. A glicerina torna o óleo mais denso e viscoso. Durante o processo de transesterificação, a glicerina é removida do óleo vegetal, deixando o óleo mais fino e reduzindo a viscosidade.

A Agência Nacional do Petróleo (ANP) monitora algumas propriedades preestabelecidas do biodiesel e, também, sugere alguns métodos analíticos para tal monitoramento (MARQUES et al., 2008). Tais métodos carecem de equipamentos

não comuns em laboratórios e tempos excessivos envolvidos em todo o processo desde a obtenção da amostra e logística de envio para os laboratórios.

Tendo em vista a necessidade de acompanhamento de certas especificações do biodiesel que garantem um bom funcionamento dos motores, surge a necessidade de se utilizar métodos que proporcionem agilidade e acuracidade na obtenção dos resultados.

Neste contexto, o método de transreflectância é pertinente, na medida que se pretende avaliar a possibilidade de utilizar um novo acessório que tornará o processo, além de mais rápido, mais barato.

1.2 Objetivos

1.2.1 Objetivo Geral

Utilizar as ferramentas de espectroscopia no infravermelho e análise multivariada para gerar modelos que determinem a quantidade do biodiesel, enxofre, massa específica e ponto de fulgor das blendas de diesel/biodiesel.

1.2.2 Objetivos Específicos

- Estudar a possibilidade de desenvolver um acessório por transreflectância para a determinação da quantidade de biodiesel nas blendas biodiesel/diesel.
- Utilizar os algoritmos PLS, iPLS e siPLS na identificação das regiões espectrais mais adequadas para a determinação da concentração de biodiesel, da quantidade de enxofre, da massa específica e do ponto de fulgor através dos espectros de infravermelho obtidos para as blendas de biodiesel/diesel.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Biodiesel

O biodiesel é uma fonte alternativa para a redução de poluentes atmosféricos, por ser biodegradável e substituinte dos combustíveis fósseis, já que cada vez mais se pensa na crise ambiental.

A qualidade do biodiesel é afetada por diversos fatores que vão desde a alteração da proporção diesel/biodiesel à adição de outros tipos de óleos que não sofreram o processo de esterificação. Então, para garantir a qualidade destes óleos, ao longo da cadeia de fornecimento, se faz necessário o controle da mistura e de algumas propriedades físicas destes produtos.

Outro fator que afeta a qualidade do biodiesel é a grande variedade de técnicas de produção de ésteres graxos juntamente com a diversidade de oleaginosas com potencial para produzir o óleo vegetal, tornando o processo de obtenção do Biodiesel uma prática com muitas variáveis (MARQUES et al., 2008).

No Brasil, o órgão responsável de estabelecer padrões de comercialização, distribuição, qualidade e fiscalização é a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP).

Entre as especificações de natureza física estabelecidas para a ANP estão as determinações do ponto de fulgor, da massa específica, da porcentagem de biodiesel e da quantidade de enxofre (Resolução ANP 42, 24/11/2004).

O ponto de fulgor é a menor temperatura, sob condições específicas, na qual uma fonte de ignição é capaz de causar combustão do vapor do líquido em análise.

A massa específica é a razão entre a massa de uma quantidade da substância e o volume correspondente, ocupado por esta porção.

A porcentagem de biodiesel é a quantidade, em percentual, de biodiesel puro (B100) misturado na matriz diesel.

A quantidade de enxofre é a quantidade de enxofre contida no diesel e é medida em percentual (%) ou partes por milhão (ppm). Estes valores podem variar conforme legislação local.

Para realizar tais análises, se faz necessária a utilização de um aparato laboratorial considerável e em alguns casos raros, envolvendo tempos longos e consumo de reagentes químicos caros e muitas vezes poluentes.

Neste contexto, o método de espectroscopia no infravermelho pode ser interessante, partindo do ponto que produz uma análise de simples obtenção e de uma forma rápida.

2.2 Espectroscopia no Infravermelho

A espectroscopia no infravermelho é uma ferramenta muito utilizada por sua rapidez na obtenção dos resultados, possibilidade de automação e por não utilizar reagentes tóxicos, utilizados na maioria dos métodos químicos tradicionais.

O objetivo da espectroscopia no infravermelho é a determinação dos grupos funcionais de um dado material. Cada grupo absorve uma frequência característica de radiação na região do infravermelho (PAVIA, LAMPMAN e KRIZ, 1996).

Assim, um gráfico de intensidade de radiação, o espectro de Infravermelho, permite caracterizar os grupos funcionais de um padrão ou de um material desconhecido. A radiação infravermelha na faixa aproximada de 10.000 a 100 cm^{-1} quando absorvida, converte-se em energia de vibração molecular; como resultado, as ligações químicas sofrem deformações angulares e/ou axiais (PAVIA, LAMPMAN e KRIZ, 1996; GEORGE e MCINTYRE, 1994). A Figura 2 mostra alguns dos principais tipos de deformações que podem sofrer as ligações. Os sinais “mais” simbolizam movimentos para fora do plano e os sinais “menos” para dentro.

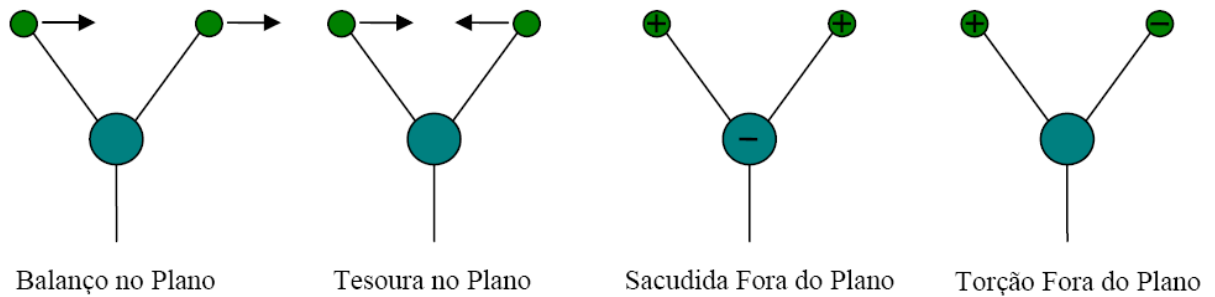


Figura 2. Tipos de deformações que sofrem as ligações quando expostas ao Infravermelho.

Fonte: SKOOG, HOLLER e NIEMAN, 2002

Existem dois tipos de espectrofotômetros no infravermelho. Que são: os Dispersivos (que empregam um monocromador) e os FT-IR (que empregam um interferômetro – largamente usado nos dias de hoje). Os espectrofotômetros no infravermelho com transformada de Fourier (FT-IR), que utilizam um interferômetro, além de mais rápidos superam os do tipo dispersivos pelo fato de não perder muita energia do sinal, face às fendas que o sinal tem que passar. Também não há a necessidade de utilizar fontes duplas.

Os métodos mais utilizados para a obtenção de espectros no infravermelho envolvem transmissão e reflectância, estando disponíveis comercialmente vários tipos de acessórios para amostras sólidas, líquidas e gasosas (BARBOSA, 2007).

Os espectros, na espectroscopia no infravermelho por reflexão, podem ser obtidos, basicamente por três formas: por reflexão difusa, especular ou atenuada (ATR).

A reflexão difusa ocorre em superfícies não totalmente planas, podendo o substrato ser contínuo ou fragmentado (na forma de pó). Neste processo de reflexão, o feixe incidente penetra a superfície da amostra interagindo com a matriz, retornando a superfície da mesma, após absorção parcial e múltiplos espalhamentos (FERRÃO, 2000).

A reflexão especular ocorre na maioria das superfícies que constituem a fronteira entre fases. A radiação refletida depende da natureza do feixe incidente, bem como das propriedades ópticas dos meios em questão. No entanto, para que ela seja observada predominantemente, é necessário que a superfície seja lisa e plana, em toda a extensão da área coberta pelo feixe de luz incidente.

A técnica ATR se caracteriza pelas múltiplas reflexões da radiação infravermelha que ocorrem no interior de cristais, de materiais com alto índice de refração (por exemplo, ZnSe), interagindo apenas com a amostra que estiver superficialmente no cristal (MIRABELLA, 1985).

2.2.1 Reflexão Total Atenuada

A reflexão total atenuada (ATR) é uma técnica de amostragem rápida que requer uma mínima preparação da amostra e, principalmente, que apresenta fácil obtenção de espectros de materiais espessos e fortemente absorventes, difíceis de serem analisados por espectroscopia por transmissão (OLINGER e GRIFFITHS, 1988).

Existem múltiplos elementos de ATR e uma variedade de acessórios desenvolvidos para obter espectros de materiais líquidos, sólidos e viscosos para uma larga faixa de tipos de amostras, incluindo filmes, resíduos, papel, revestimentos sobre papel, pós, tintas, tecidos, espumas, minerais, vidros, etc. Alguns dos mais recentes acessórios permitem o monitoramento de fluxo

A Figura 3 ilustra a reflexão total atenuada através do emprego de um elemento de ATR. Este cristal deve ser composto de um material com alto índice de refração, para que somente uma pequena parte do feixe de radiação incidente seja refletida ao atingir o cristal.

Após o feixe incidente sofrer a difração penetrando no meio n_2 , a geometria do cristal deve promover a reflexão e para tanto $\theta_i < \text{sen}^{-1}(n_3/n_2)$. Esta reflexão interna cria uma onda evanescente a qual se estende na amostra em contato com o cristal. Na região do infravermelho onde a amostra absorve energia, a onda evanescente será atenuada. Num acessório de ATR tem-se um elemento de material adequado, por exemplo, ZnSe, que deverá produzir múltiplas reflexões (FERRÃO, 2000).

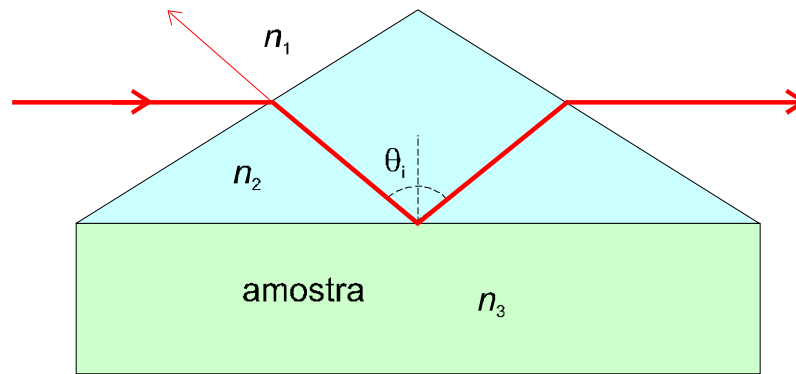


Figura 3 - Esquema da reflexão interna em um elemento de ATR.
Fonte: FERRÃO, 2000.

Os princípios teóricos básicos da espectroscopia de reflexão interna (IRS), foram publicados no início dos anos 60 por Harrick (1960) e Fahrenfort (1961), que descrevem detalhes da teoria e apresentam dados experimentais.

O fenômeno da reflexão interna é observado sob certas condições. Quando a radiação entra em um prisma feito com um índice de refração alto em relação ao meio externo (cristal de ATR) a radiação será refletida totalmente internamente. Para tanto, considere a Figura 4, na qual um feixe de radiação encontra-se no interior do elemento de ATR.

Para um índice de refração tal que $n_2 > n_3$, aplicando-se a lei de Snell o ângulo de refração é imaginário para valores do ângulo de incidência tal que satisfaça a equação 1.

$$\theta_i > \text{sen}^{-1}(n_3/n_2) \quad (1)$$

O ângulo acima do qual esta onda refratada deixa de ser real é denominado de ângulo crítico (θ_c) e é expresso conforme equação 2.

$$\theta_c = \text{sen}^{-1}\left(\frac{n_3}{n_2}\right) \quad (2)$$

Nestas condições o feixe é totalmente refletido na interface n_3/n_2 .

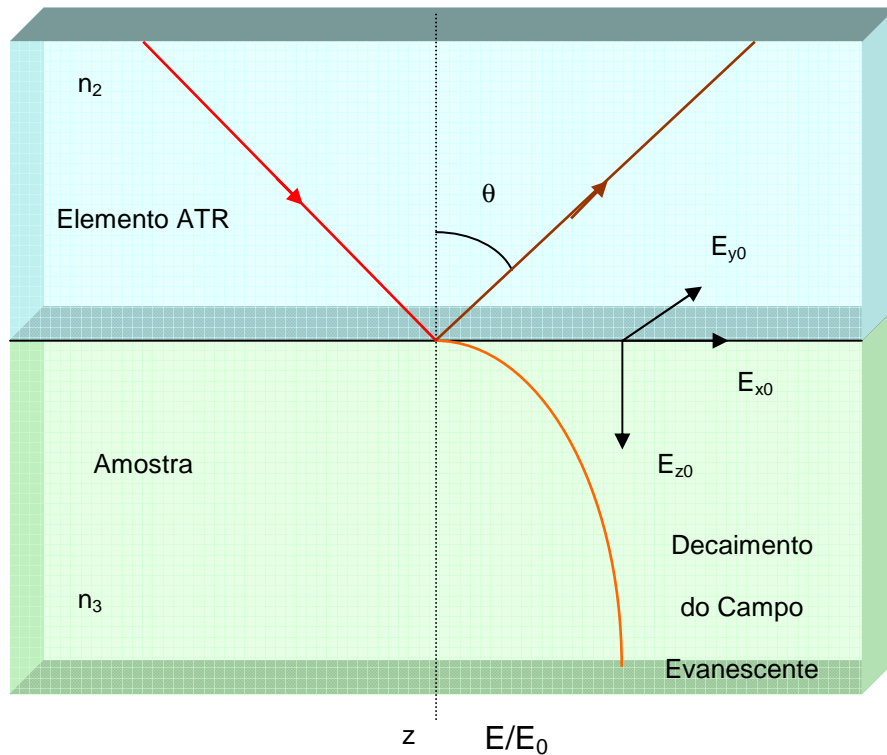


Figura 4 - Representação da propagação da radiação infravermelha através do elemento de ATR.
Fonte: FERRÃO, 2000.

As propriedades do campo evanescente podem ser visualizadas na Figura 4, onde a propagação da radiação infravermelha do meio 2, com índice de refração n_2 , sofre reflexão interna na interface com o meio 3, com índice de refração n_3 , quando o ângulo de incidência excede o ângulo crítico.

2.2.2 Transreflectância

Apesar de existirem comercialmente acessórios que visam otimizar uma determinada técnica de reflexão e minimizar a interferência de outras, um importante fator a ser levado em consideração é a característica da amostra. Por exemplo, se empregarmos um típico acessório de reflexão difusa no infravermelho com transformada de Fourier (DRIFTS) para uma amostra líquida, o fenômeno observado não será a reflexão difusa, mas sim um tipo pouco comum de reflexão que envolve primeiramente a transmissão do feixe no filme líquido, seguido de uma reflexão na superfície do acessório e nova transmissão no retorno da radiação através do filme

líquido, conforme ilustra a Figura 5. No presente trabalho, esta técnica aqui será denominada transreflectância (HELFER et al., 2006).

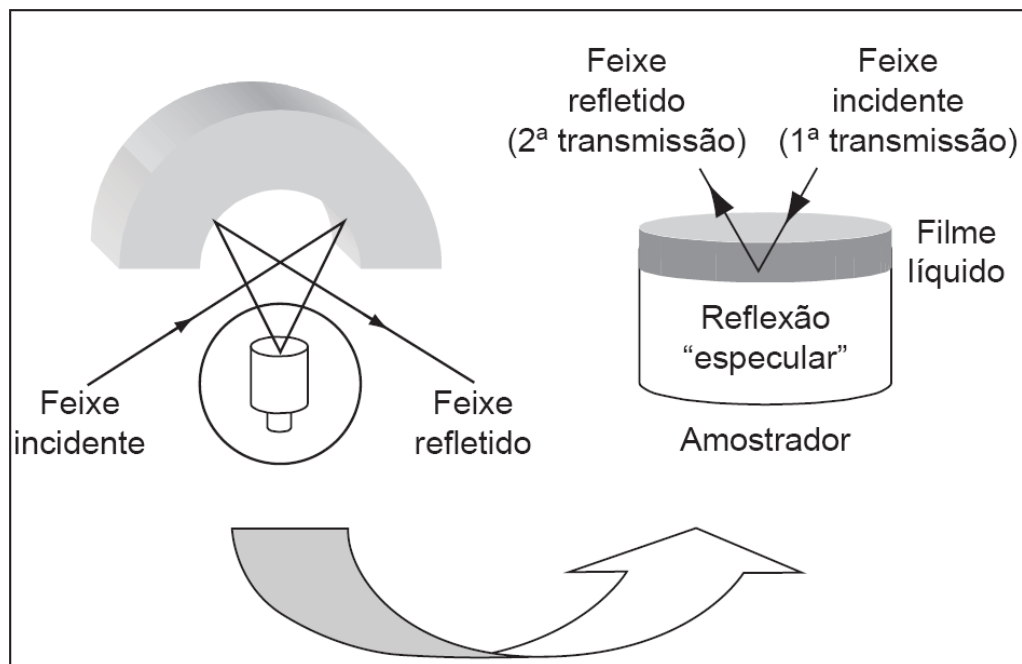


Figura 5 – Esquema que descreve princípio da transreflectância, utilizando o acessório de DRIFTS.

Fonte: HELFER et al., 2006.

O que determina o processo de transreflectância não é o acessório mas sim a maneira com que o raio de infravermelho incide na amostra, reflete no amostrador e passa novamente através da amostra (transmissão + reflexão + transmissão).

2.3 Técnicas de Análise Multivariada

As técnicas quimiométricas para análise multivariada tornaram-se abordagens comuns para análise rápida de amostras complexas de suas assinaturas espectrais e tornam-se mais eficientes com a escolha de regiões espectrais, minimizando, com isso, os erros de previsão. Os benefícios em selecionar-se uma região espectral vão desde a estabilidade do modelo de calibração em relação à colinearidade, como também na interpretação das relações entre o modelo e a composição da amostra (MULLER et al., 2006). Também pode ser entendida como uma área multidisciplinar que aplica ferramentas matemáticas e estatísticas a dados provindos de processos químicos, de natureza multivariada (FERRÃO, 2000).

Dentre as principais técnicas de Análise Multivariada existentes estão:

- Análise por Componentes Principais (PCA);
- Análise Por Agrupamento Hierárquicos (HCA);
- Regressão Linear Múltipla (MLR);
- Regressão por Componentes Principais (PCR);
- Regressão por Mínimos Quadrados Parciais (PLS).

2.3.1 Regressão por Mínimos Quadrados Parciais (PLS)

O PLS foi desenvolvido no início da década de 70 por Herman Wold e foi usada primeiramente na região do infravermelho próximo, onde é difícil designar bandas para componentes particulares. Neste modelo as informações espectrais e as informações de concentrações são usadas ao mesmo tempo para a calibração do modelo (WOLD e SJÖSTRÖM, 1998).

O PLS é uma técnica que começou a ser largamente usada recentemente que generaliza e combina características da Análise de Componentes Principais e da Regressão Múltipla. Sua meta é prever ou analisar dados de variáveis dependentes de um conjunto de dados de variáveis independentes (ABDI, 2007). Diferente dos Modelos de Regressão Múltipla eles podem analisar dados com alto grau de colinearidade (correlacionados), ruídos, alto número de variáveis independentes e, também, gerar modelos para mais de uma variável dependente (WOLD, SJÖSTRÖM E ERIKSSON, 2001).

Esta predição é obtida através da extração das variáveis independentes de um conjunto de fatores ortogonais chamados de variáveis latentes, que possuem um poder de predição maior.

O método consiste na decomposição de uma matriz \mathbf{X} , em termos de várias matrizes unidimensionais \mathbf{M} , acrescidos de um erro (que representa a parte não modelada de \mathbf{X}), conforme a equação 3.

$$\mathbf{X} = \mathbf{M}_1 + \mathbf{M}_2 + \dots + \mathbf{M}_a + \mathbf{E} \quad (3)$$

onde a representa o número de fatores (componentes principais ou variáveis

latentes) selecionados para truncar a igualdade, e **E** representa a matriz de resíduos, relacionadas ao número de componentes escolhido.

As matrizes **M** representam as variáveis latentes e são formadas pelo produto de dois vetores **t** e **p**, escores e os pesos, respectivamente, como mostram as equações 4 e 5.

$$\mathbf{X} = \mathbf{t}_1 \mathbf{p}_1^t + \mathbf{t}_2 \mathbf{p}_2^t + \dots + \mathbf{t}_v \mathbf{p}_a^t + \mathbf{E} \quad (4)$$

$$\mathbf{X} = \mathbf{T}^t + \mathbf{E} \quad (5)$$

t e **p** são vetores coluna, sendo que **t** vai de 1 a n e **p**, que é transposto, de 1 a m.

A dimensionalidade do espaço original é igual ao número de colunas expressas em **X**, ou seja, o número de variáveis originais que é expresso por m. no novo modelo a dimensionalidade é descrita pelo número de matrizes **M_i** necessárias para descrever **X**. Assim, se for possível descrever a matriz **X** através de um número pequeno destas matrizes **M_i** será possível reduzir a dimensionalidade sem perda significativa de informação.

De maneira análoga as variáveis independentes **X** as variáveis **Y**, as dependentes, também são representadas por escores e pesos, como mostra a equação 6.

$$\mathbf{Y} = \mathbf{U}^t + \mathbf{F} \quad (6)$$

A relação entre as matrizes **X** e **Y** pode ser obtida correlacionando-se os escores de cada bloco, com a finalidade de se obter uma relação linear, como mostram as equações 7 e 8.

$$\mathbf{U}_a = \mathbf{b}_a \mathbf{t}_a \quad (7)$$

$$\mathbf{U} = \mathbf{bT} + \mathbf{e} \quad (8)$$

Onde \mathbf{T} é uma matriz resposta para uma série de amostras de calibração, \mathbf{U} é uma matriz contendo as propriedades de todas as amostras e \mathbf{b} é um vetor contendo todos os parâmetros do modelo.

O método PLS busca um ajuste onde as matrizes de ruído \mathbf{E} e \mathbf{F} sejam as menores possíveis e ao mesmo tempo, conseguindo um relação linear entre \mathbf{t} e \mathbf{u} . este ajuste é obtido através de leves mudanças nos escores, buscando a melhor correlação possível.

No PLS a calibração multivariada é realizada utilizando a informação do espectro inteiro para construir um modelo da regressão, relacionado à propriedade de interesse. Por esta razão é chamado de método *full-spectrum* (BORIN e POPPI, 2005).

2.3.2 Seleção de Variáveis

É crescente o número de trabalhos que empregam a espectroscopia no infravermelho em análises quantitativas, principalmente pela utilização de métodos multivariados de análise. Estes métodos utilizam inúmeras informações não relacionadas diretamente com os analitos de interesse, pois utilizam todo o espectro para a calibração do modelo. Devido ao fato de cada variável dependente estar mais relacionada com apenas uma parte do espectro as conclusões podem ser errôneas, devido ao ruído gerado pelo excesso de informações (DI et al., 2010).

Para minimizar este problemas experimentais diversos métodos têm sido utilizados, além do pré-processamento, que são chamados métodos de seleção de variáveis. Isso gera um sub-conjunto com menor número de variáveis (parte do espectro) e que apresentam maior sensibilidade e linearidade com as variáveis dependentes estudadas (COSTA FILHO e POPPI, 2002) .

Para a otimização do modelo PLS podem ser usados algoritmos determinísticos conhecidos como iPLS, biPLS, siPLS e mwPLS para selecionarmos sub-conjuntos de variáveis com maior grau de interesse. Estes métodos geralmente superam o PLS em termos de predição, devido a remoção das partes do espectro que não estão relacionadas com a propriedade estudada, evitando assim ruídos no modelo (PEREIRA et al., 2008).

O algoritmo iPLS (Interval PLS) é uma extensão do PLS e divide o conjunto de dados em um número de intervalos, calcula o modelo PLS para cada intervalo e apresenta os resultados em um gráfico. O método é planejado para dar uma visão geral dos dados e pode ser útil na interpretação de quais sinais do espectro são mais representativos na construção de um bom modelo de calibração (NØRGAARD et al., 2000).

A otimização por biPLS (backward interval PLS), por sua vez, remove regiões espectrais sem relevância, na qual os modelos PLS são inicialmente calculados com cada um dos intervalos deixados de fora. Por exemplo, ao escolher-se 20 intervalos, cada modelo PLS é construído com 19 intervalos, deixando de fora um intervalo por vez. O primeiro intervalo eliminado será aquele que quando deixado de fora resulta no melhor modelo, isto é, aquele que resulta no RMSECV (*Root Mean Square Error of Cross Validation*) mais baixo. Esse procedimento é continuado até que haja um intervalo restante (GUEDES et al., 2007).

O mwPLS (moving window PLS) calcula o iPLS baseado no conceito de uma janela móvel. Para cada variável dependente é calculado um modelo PLS para um dado tamanho de janela (range do espectro). O objetivo deste método é escolher qual é a posição e o tamanho da janela, ao longo do espectro, que retorna o menor RMSECV. Como resultado se tem um gráfico com o RMSECV e o número de variáveis latentes de cada posição da janela, ao longo do espectro (DU et al., 2004).

Finalmente, o siPLS (synergy interval PLS) divide um conjunto de variáveis em um número de intervalos (variáveis de informação) e calcula todas as possibilidades de combinações de dois, três e quatro intervalos. O tempo de processamento pode ser longo dependendo do número de intervalos em que foi dividido o conjunto inicial e o número de elementos à combinar, que foram selecionados (SILVA et al., 2009).

Outros métodos, além dos descritos das derivações do PLS, também podem ser aplicados para o problema de seleção de variáveis, trazendo ganhos importantes na capacidade de predição dos modelos.

Costa Filho e Poppi (2002) aplicaram algoritmos genéticos para a seleção de variáveis, utilizando o infravermelho médio, para a determinação simultânea de misturas de glicose, maltose e frutose. Foram usados modelos PLS e de Regressão

Múltipla Linear com e sem pré-processamento, com erros obtidos em torno de 3% para a porcentagem de açúcar.

Outro trabalho mostrou que a combinação do algoritmo de sucessivas projeções e o algoritmo de eliminação de variáveis não-informativas (UVE – SPA) trouxe redução no erro de previsão do teor de nicotina em uma lâmina de tabaco, de 0,174 para 0,160 % (YE, WANG e MIN, 2008).

Du et al. (2004) desenvolveram algoritmos melhorados, baseados no mwPLS chamados janela móvel de tamanho variável (csmwPLS) e pesquisa de combinação de janela móvel (scmwPLS), conseguindo modelos realmente otimizados. O primeiro varia a posição da janela, no espectro, e o tamanho na busca da região espectral mais relevante. O segundo algoritmo ao mesmo tempo que encontra um tamanho de janela e posição com informações relevantes acha também outras janelas, que combinadas formam um sub-conjunto de janelas de espectro mais representativas. Este último é feito a base de cálculo exaustivo, sendo somente viável com arquiteturas de computadores super potentes.

Xiaobo et al. (2010) apresentaram um artigo de revisão, contemplando os métodos de seleção de variáveis do mais simples aos mais complexos, abordando desde iPLS à algoritmos genéticos, redes neurais, projeções sucessivas, remoção de variáveis não informativas, janela móvel de tamanho variado e outros.

Para a seleção e avaliação dos modelos de regressão multivariada são empregados os erros quadráticos médios de validação cruzada, de calibração e de previsão (COSTA FILHO e POPPI, 2002), calculado segundo a equação 9, mostrada a seguir:

$$RMSE = \sqrt{\frac{\sum (y_p - y_r)^2}{n}} \quad (9)$$

Onde:

y_p = o valor da variável previsto pelo modelo multivariado;

y_r = o valor da variável segundo o método de referência;

n = representa o número de amostras.

Outro critério que também é observado para a seleção do modelo é o número de variáveis latentes empregadas. À medida que aumenta o número de variáveis latentes o modelo passa a apresentar um erro menor, porém não real, pois está modelando junto o ruído. Por isso, dá-se preferência aos modelos que empregam um menor número de variáveis latentes, porém apresentam erros de calibração e previsão baixos e semelhantes.

2.4 Recentes aplicações da espectroscopia no infravermelho e análise multivariada na caracterização do biodiesel e suas blends com diesel

Uma recente aplicação utilizou cartas de controle multivariado junto à espectroscopia do infravermelho próximo para controlar a qualidade do biodiesel. As cartas foram construídas utilizando o Método do Sinal Analítico dos Componentes (NAS – Net Analyte Signal). Este método consiste em separar o sinal do espectro nas componentes NAS, de interação e de ruído. Este método é facilmente adaptável ao monitoramento online de processos de biodiesel e requer somente algumas amostras padrão para calibrar o modelo (POPPI, OLIVEIRA e ROCHA, 2008).

Outro estudo investigou a viabilidade da determinação rápida de contaminantes (gasolina, etileno glicol e água) em lubrificantes, utilizando apenas uma fração do espectro, utilizando um procedimento de calibração multivariada orientado graficamente, chamado iPLS (BORIN e POPPI, 2005).

Ainda, utilizando o infravermelho próximo e calibração multivariada Pereira et al. (2008) determinaram a acidez, o índice de refração e a viscosidade em quatro tipos de óleos vegetais (milho, soja, canola e girassol). Utilizaram os algoritmos iPLS e MLR combinados com o Algoritmo das Projeções Sucessivas.

De maneira semelhante, utilizando métodos analíticos baseados no infravermelho próximo, infravermelho médio e calibração multivariada Lira et al. (2010) proporem um modelo para monitorar a estabilidade de alguns parâmetros de qualidade do biodiesel (estabilidade oxidativa, acidez e quantidade de água). Para a construção dos modelos foram utilizado PLS e a seleção de variáveis otimizada com o Algoritmo das Projeções Sucessivas.

Outro trabalho, que utilizou a espectroscopia no infravermelho com

transformada de Fourier, foi utilizado para calibrar um modelo PLS e identificar adulterações na mistura diesel/biodiesel. No modelo utilizado foram considerados biodiesel obtido a partir de óleos de rícino, algodão e dendê, onde foram adulterados com a adição de óleo de soja. O modelo criado tornou fácil a identificação da quantidade do adulterante, óleo de soja, em cada recurso de biodiesel (FORTES et al., 2008).

3 METODOLOGIA

3.1 Descrição

Este trabalho iniciou com uma pesquisa bibliográfica sobre conceitos e aplicações dos métodos multivariados PLS e iPLS na construção de modelos de regressão, bem como experimentos em laboratório para a obtenção dos espectros, como forma de compreendê-los, para que juntamente com os demais materiais bibliográficos coletados fosse possível desenvolver esta dissertação.

3.2 Etapas

Para a realização deste trabalho foram seguidas as etapas descritas no quadro da Figura 6.



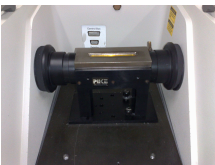
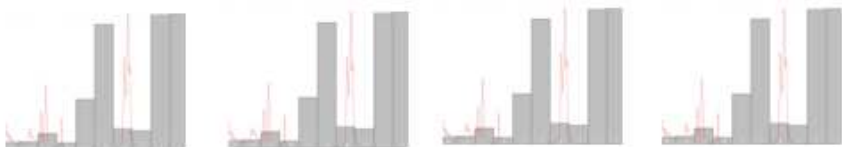
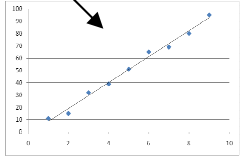
Atividade	Etapa
 <p data-bbox="582 492 925 526">Obtenção das amostras</p>	I
 <p data-bbox="518 795 1013 828">Preparação das blendas (misturas)</p>	II
 <p data-bbox="582 1086 933 1120">Aquisição dos Espectros</p>	III
<p data-bbox="582 1176 925 1209">Tabulação das Matrizes</p>	IV
<p data-bbox="319 1265 1189 1355">Obtenção dos parâmetros físicos, químicos e físico-químicos das blendas</p>	V
 <p data-bbox="478 1579 1021 1612">Construção dos modelos e otimização</p>	VI
 <p data-bbox="790 1792 1061 1825">Escolha do Modelo</p>	VII

Figura 6. Quadro ilustrativo das etapas envolvidas na realização da pesquisa.
 Fonte: Elaborada pelo autor

Etapa I – Obtenção das amostras

O biodiesel utilizado neste trabalho é constituído de ésteres metílicos de óleo de soja, tendo sido doado pela Oleoplan S.A., Veranópolis, Brasil. As amostras de diesel, classificadas como interior e metropolitano, foram fornecidas pela Refinaria Alberto Pasqualini, Canoas, Brasil. Estas amostras diferem basicamente no teor de enxofre.

Etapa II – Preparação das blendas (misturas)

Esta etapa consistiu em preparar misturas de diesel/biodiesel em proporções diferentes, com o objetivo de avaliar o efeito da concentração do biodiesel na mistura.

Para este trabalho, foram utilizadas quarenta e cinco blendas com faixa concentração variando de 8 a 30,1 %(v/v), sendo o incremento, entre cada amostra, de aproximadamente 0,5%. As amostras ímpares foram preparadas com diesel interiorano e as pares, metropolitano. As misturas foram realizadas em balões calibrados, para garantir precisão na medida de concentração de biodiesel/diesel. Todas as formulações foram realizadas no Instituto de Química da UFRGS.

Etapa III – Aquisição dos espectros

Os espectros ATR/FTIR foram obtidos empregando um espectrofotômetro Nicolet Magna 550 FTIR equipado com um acessório de reflexão total atenuada horizontal da Pike com cristal de ZnSe. Todas as amostras das blendas foram analisadas em triplicata na faixa de 650 – 4000 cm^{-1} a temperatura ambiente, sendo realizadas 32 varreduras com resolução de 4 cm^{-1} .

Etapa IV – Tabulação dos dados

Com os sinais espectrais obtidos, foi construída uma matriz de dados com a média das triplicatas dos espectros normalizados de cada amostra. Esta matriz foi utilizada para a construção de modelos de regressão multivariados.

Etapa V – Obtenção dos parâmetros físicos, químicos e físico-químicos das blendas

Todas as blendas formuladas foram caracterizadas utilizando o método determinado pela ANP, sendo realizadas pelo LABCOM, UFRGS.

A massa específica, a 20°C, foi determinada segundo o método ABNT NBR 14065, usando um densímetro digital, modelo Anton Paar DMA 4500. O ponto de

fulgor foi determinado de acordo com o método ABNT NBR 14598, utilizando o equipamento Petrotest (Pensky Martens - ASTM D93), modelo PMA4. O teor de enxofre nas blendas foi determinado seguindo a norma ASTM D 5453 através de um analisador de enxofre total ANTEK, modelo 9000. A Tabela 1 apresenta as amostras das blendas diesel/biodiesel com suas concentrações e respectivos valores das propriedades medidas pelo método referência.

TABELA 1 – VALORES OBTIDOS DAS PROPRIEDADES PELO MÉTODO REFERÊNCIA PARA CADA AMOSTRA DAS BLENDEAS DE BIODIESEL/DIESEL.

Amostra #	1	2	3	4	5	6	7	8	9
% BD (v/v)	8,0	8,5	9,0	9,5	10,0	10,5	11,0	11,5	12,0
Massa específica (kg/m ³)	854,1	862,0	854,5	862,2	854,7	862,4	855,0	862,6	855,3
P. fulgor (C°)	59,5	75,5	59,5	76,5	60,5	76,5	60,5	76,5	60,5
Enxofre(mg/kg)	599,0	401,0	548,0	417,0	544,0	389,0	533,0	383,0	524,0
Amostra #	10	11	12	13	14	15	16	17	18
% BD (v/v)	12,4	13,0	13,5	14,0	14,4	15,0	15,4	16,0	16,5
Massa específica (kg/m ³)	862,8	855,6	862,7	855,8	863,2	856,1	863,4	856,5	863,6
P. fulgor (C°)	78,5	59,5	77,5	60,5	77,5	62,5	77,5	61,5	77,5
Enxofre (mg/kg)	373,0	510,0	382,0	512,0	384,0	502,0	373,0	518,0	368,0
Amostra #	19	20	21	22	23	24	25	26	27
% BD (v/v)	17,1	17,6	18,1	18,6	18,9	19,4	20,1	20,6	21,1
Massa específica (kg/m ³)	856,6	863,9	856,9	864,0	857,1	864,2	857,5	864,4	857,9
P. fulgor (C°)	60,5	77,5	58,5	77,5	60,5	76,5	61,5	77,5	59,5
Enxofre (mg/kg)	492,0	362,0	482,0	375,0	487,0	365,0	480,0	351,0	474,0
Amostra #	28	29	30	31	32	33	34	35	36
% BD (v/v)	21,6	21,9	22,4	22,9	23,4	23,9	24,4	24,9	25,6
Massa específica (kg/m ³)	864,5	858,1	864,8	858,3	865,0	858,6	865,1	859,0	865,5
P. fulgor (C°)	78,5	61,5	77,5	61,5	78,5	62,5	77,5	62,5	77,5
Enxofre (mg/kg)	345,0	483,0	341,0	486,0	354,0	443,0	334,0	450,0	328,0
Amostra #	37	38	39	40	41	42	43	44	45
% BD (v/v)	26,0	26,5	27,1	27,4	28,0	28,5	28,7	29,2	30,1
Massa específica (kg/m ³)	859,3	865,8	859,6	865,7	859,8	866,0	860,0	866,2	860,3
P. fulgor (C°)	63,5	78,5	63,5	79,5	62,5	77,5	62,5	77,5	62,5
Enxofre (mg/kg)	444,0	325,0	432,0	321,0	430,0	312,0	399,0	312,0	428,0

Etapa VI – Construção do modelo e otimização

Nesta etapa foram construídos os modelos de PLS e otimizados com os métodos iPLS, siPLS e biPLS.

Todos os modelos foram construídos, utilizando o MATLAB® versão 7.0 (*TheMath Works*, <http://www.mathworks.com>, EUA), juntamente com o pacote *iToolbox* versão 2 desenvolvido por Lars Norgaard (*iToolbox for MATLAB*, <http://www.models.kvl.dk>, Dinamarca).

Para cada propriedade, foram desenvolvidos modelos PLS dividindo espectro em 2, 4, 8, 16 e 32 partes. Os dados foram utilizados de forma bruta e com os pré-processamentos centrados na média, auto-escalados, centrados na média com correção do espalhamento de luz e auto-escalados com correção do espalhamento de luz.

Para os modelos siPLS foi utilizado o intervalo dividido em 32 partes, combinando 2, 3 e 4 intervalos e utilizaram os mesmos pré-processamentos descritos para o modelo iPLS.

Das quarenta e cinco amostras produzidas, vinte e nove foram selecionadas para a calibração e quatorze para a previsão, sendo que duas amostras, identificadas como 17 e 44, foram rejeitadas por suspeita de contaminação.

A seleção das amostras para a constituição dos conjuntos de calibração e previsão teve dois objetivos principais: construir conjuntos que contenham a informação da matriz de dados utilizados neste estudo e que contenha a maior variabilidade possível em relação à faixa de concentração. Os 43 espectros médios das amostras foram divididos em dois conjuntos, sendo 29 empregados na construção do conjunto de calibração e 14 no conjunto de previsão. A distribuição das concentrações destas amostras para cada uma das substâncias participantes das misturas (biodiesel/diesel) é apresentada na Figura 7, onde se tomou o cuidado de garantir que as amostras com valores extremos de concentração pertençam ao conjunto de calibração.

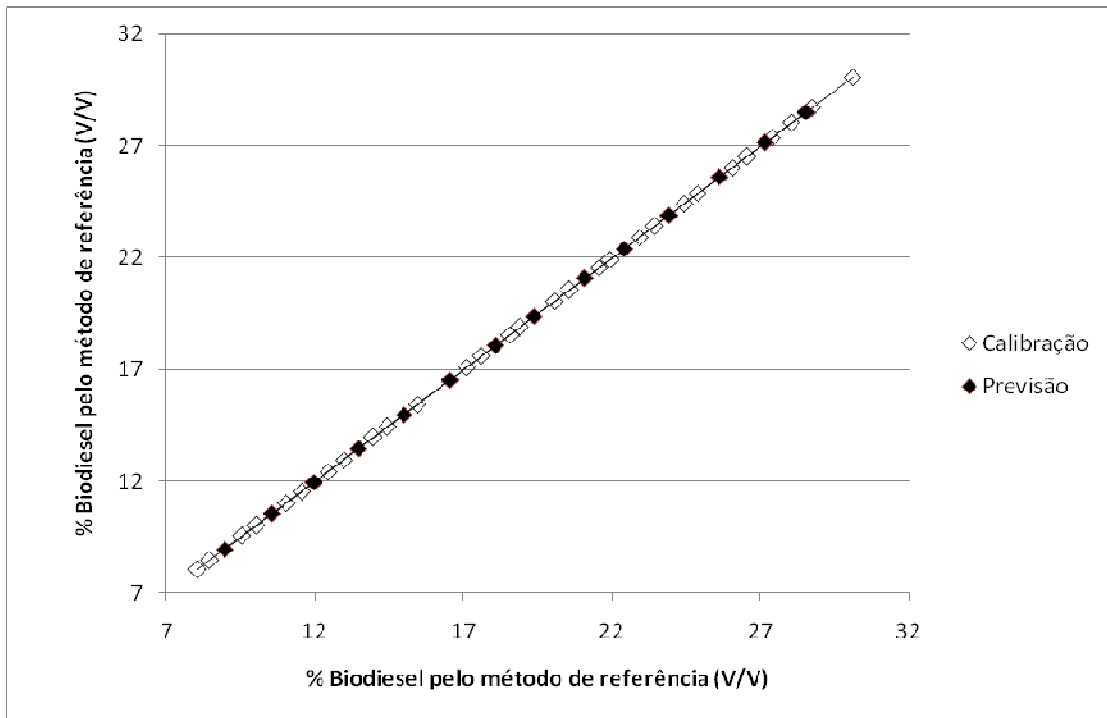


Figura 7 – distribuição das amostras de calibração e de previsão das para as 43 amostras das blends biodiesel/diesel, variando de 8 a 30 % (v/v).
Fonte: Elaborada pelo autor

Etapa VII – Escolha do modelo.

Dentre os modelos desenvolvidos foram selecionados para cada propriedade os modelos que apresentavam, além do RMSECV baixo, também, um erro de calibração e de previsão na mesma ordem de grandeza, para cada propriedade.

4 RESULTADOS

Neste capítulo são apresentados os resultados obtidos para o desenvolvimento do acessório alternativo, utilizando o acessório de Reflexão Especular com o princípio de transreflectância (transmissão+reflexão+transmissão), e os resultados dos espectros obtidos pelo ATR.

Os espectros gerados através do acessório desenvolvido para transreflectância, foram submetidos à análise pelo algoritmo HCA, buscando identificar algum tipo de similaridade que agrupasse as amostras por faixa de concentração.

Já, para os dados gerados pelo ATR, além do HCA, foram aplicados os algoritmos PLS, iPLS e biPLS, sendo na seqüência escolhido o modelo que melhor correlacionou as propriedades quantidade de enxofre, concentração de biodiesel, massa específica e ponto de fulgor das blendas diesel/biodiesel aos espectros obtidos.

4.1 Resultados para o acessório de transreflectância

Primeiramente foram realizados testes com chapas reflexivas construídas a partir de aço inoxidável e alumínio sem acabamento superficial controlado.

Os resultados obtidos demonstraram que surgiram flutuações nos dados, provavelmente devido à reflexão do infravermelho sobre a superfície irregular da chapa e variação da espessura do filme das blendas biodiesel/diesel.

Com isso, foram propostas outras três chapas reflexivas com acabamentos superficiais distintos, forçando as condições de acabamento superficial aos extremos, indo de uma superfície com acabamento controlado e polido, passando por outra com pequenos riscos e a terceira com aplicação de jato de micro-esfera de vidro. A Figura 8 mostra as três chapas desenvolvidas.

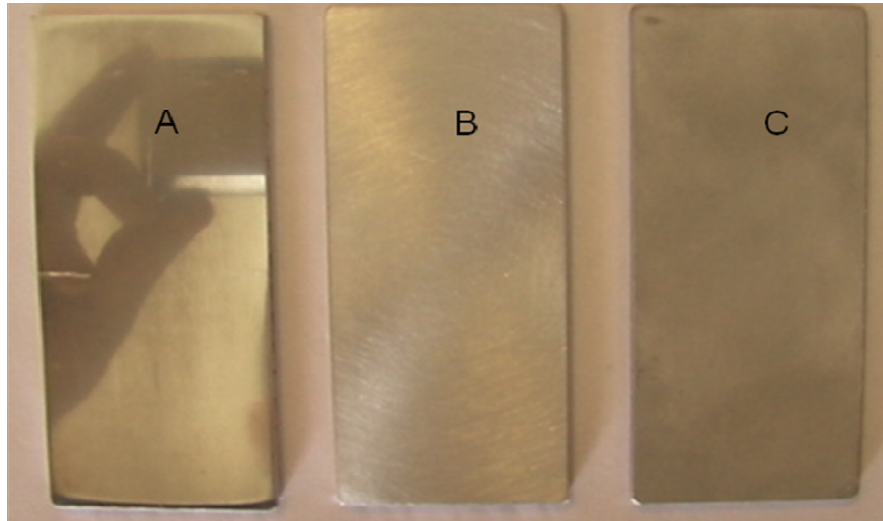


Figura 8 – Três modelos de chapas propostos, com acabamento controlado
Fonte: Registro fotográfico do autor

Na chapa A foi aplicado acabamento com oito tipos de lixa variando de #220 a #2000 (quanto maior o número mais fina é a granulometria do abrasivo) e posteriormente aplicado polimento com massa de polir #2 e massa de polir #1, gerando uma superfície polida. O lixamento foi realizado fixando-se a lixa em uma superfície plana, para corrigir a planicidade do acessório.

Na chapa B foi aplicado lixamento, aos moldes da chapa A porém até a lixa #800, para correção de planicidade e posteriormente foi realizado um lixamento com movimentos circulares com lixa #600.

A chapa C recebeu o mesmo tratamento da chapa B porém ao invés de finalizar com lixa #600 foi aplicado jateamento com micro-esfera de vidro.

As chapas B e C foram desenvolvidas para testar se o acabamento com rugosidade maior seria melhor pra controlar a espessura do filme das blendas diesel/biodiesel, devido ao fato das mesmas se espalharem melhor sobre este tipo de superfície. Já o desenvolvimento da chapa A teve o intuito de se conseguir uma reflexão melhor e mais precisa do feixe do infravermelho, melhorando a reprodutibilidade das informações espectrais das amostras.

A Figura 9 mostra os espectros obtidos para os diferentes modelos de chapa, onde se pode notar uma diferença clara entre os três modelos e a Figura 10 mostra o dendrograma com o resultado do agrupamento realizado entre estes três modelos.

Cabe ressaltar que a um coeficiente de similaridade de 0,962, a chapa A superou as outras duas, formando três agrupamentos por concentração de biodiesel.

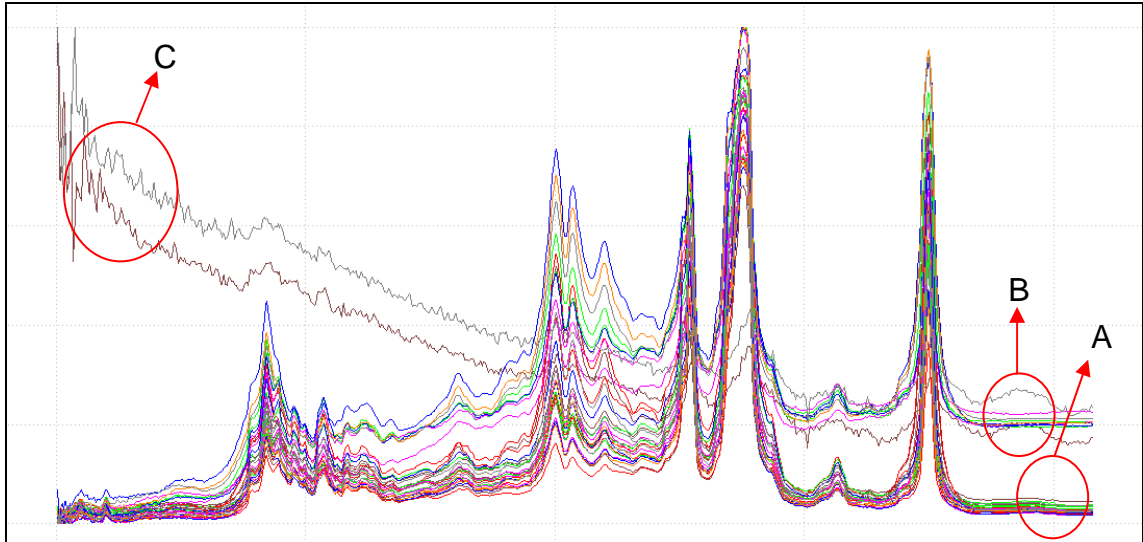


Figura 9 - Gráfico com os espectros dos três modelos de acessórios ("chapas").
Fonte: Elaborada pelo autor.

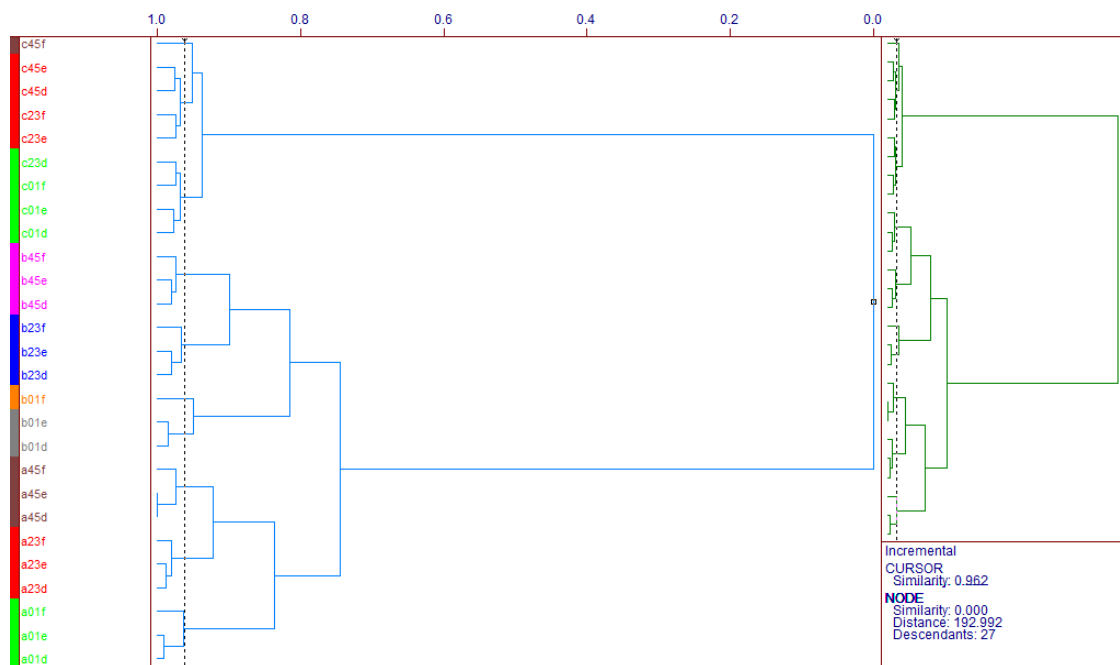


Figura 10 – Dendrograma representativo do agrupamento dos três modelos de chapas, utilizando pré-processamento: normalizados pela variância, distância euclidiana e método de ligação incremental com transformada SNV (dendrograma obtido utilizando o software Pirouette 3.11, <http://www.infometrix.com>, EUA).

Na seqüência foram desenvolvidos vários testes, com quatro chapas modelo A em triplicata com várias faixas de concentração de blendas biodiesel/diesel. Utilizou-se uma configuração diferente no equipamento, reduzindo o número de varreduras, pois verificou-se que quanto mais tempo a amostra permanecia no equipamento maior seria o efeito de formação de gota do filme de óleo na chapa, pois quando a leitura é realizada o filme ficava virado para baixo. No procedimento realizado, a chapa era colocada somente quando o equipamento estivesse pronto para a leitura, reduzindo assim o efeito da formação de gota. A Figura 11 mostra a posição da chapa no acessório PIKE.



Figura 11 – Posicionamento da chapa sobre o acessório de reflexão especular PIKE

Fonte: Registro fotográfico do autor

A seguir, pode ser observado, através de dendrogramas, o nível de agrupamento por similaridade de concentração de biodiesel identificado nos espectros obtidos com quatro protótipos da chapa A, utilizando o algoritmo HCA do software Pirouette.

A nomenclatura das amostras procede da seguinte maneira: ignora-se as duas primeiras letras, a dezena seguinte representa o código crescente representativo da faixa de concentração de biodiesel, a letra representa as quatro chapas modelo A de a até d e o último número é a indicação das triplicatas.

Devido ao grande número de amostras, o dendrograma foi dividido em cinco parcelas para poderem ser visualizadas as nomenclaturas das amostras e os agrupamentos formados a um nível de similaridade de 0,853, que podem ser visualizados da Figura 12 a 16. Do lado direito das cinco figuras pode ser visualizado o dendrograma completo e a parte que esta em ampliação em cor diferenciada.

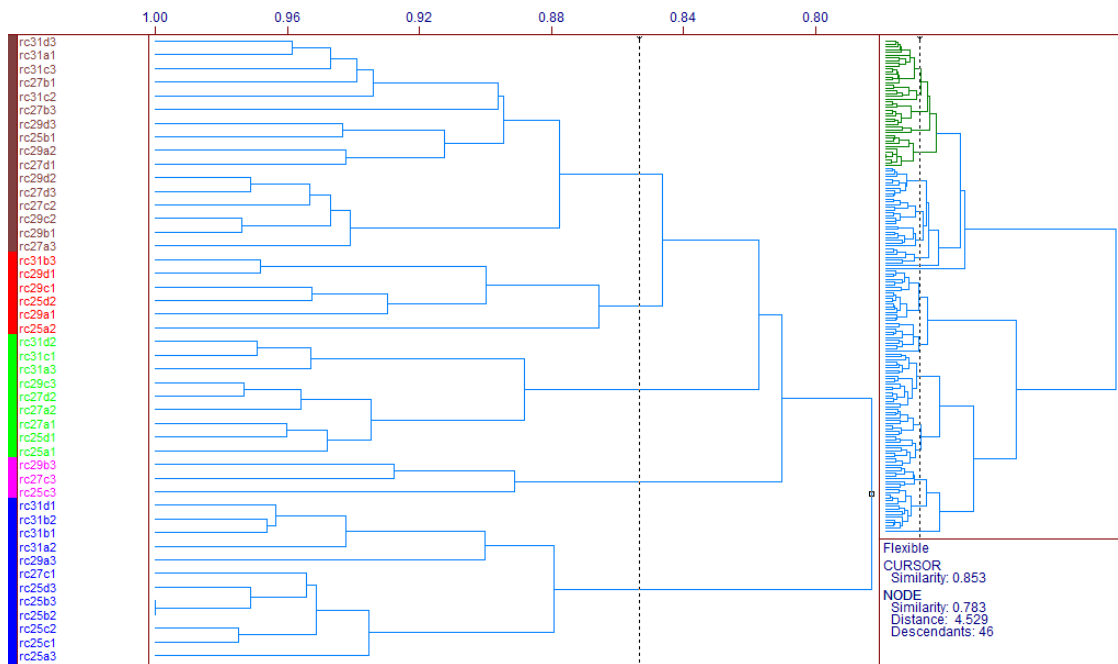


Figura 12 – Dendrograma representativo da primeira parcela de agrupamentos das triplicadas das quatro chapas modelo A, utilizando pré-processamento: centrado na média, distância euclidiana e método de ligação flexível com transformada SNV (dendrograma obtido utilizando o software Pirouette 3.11, <http://www.infometrix.com>, EUA).

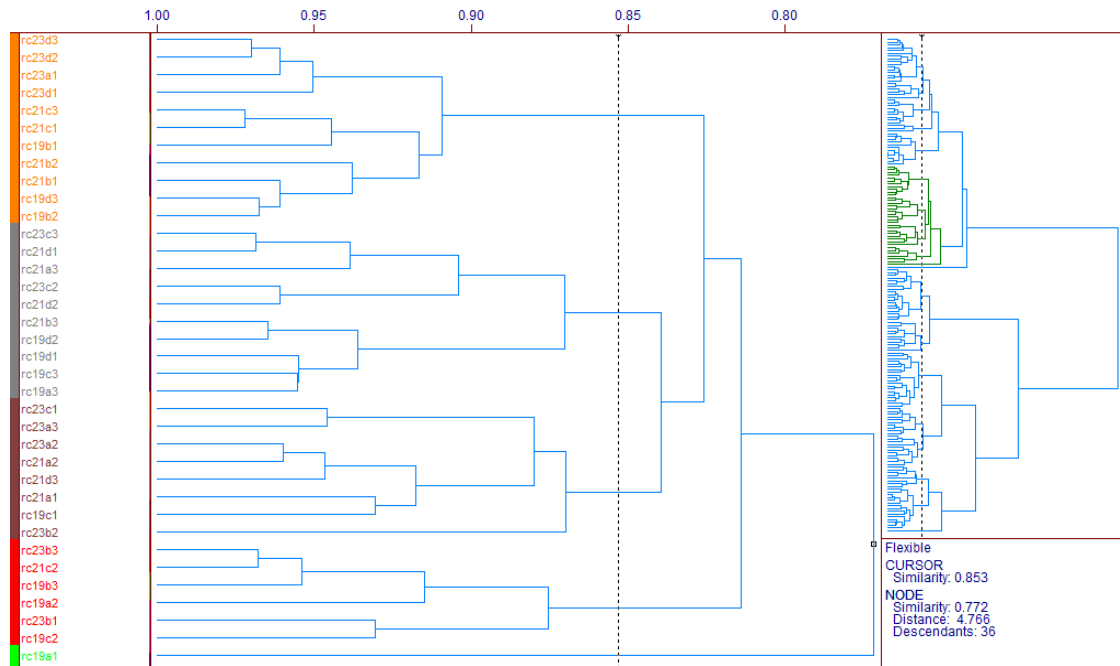


Figura 13 – Dendrograma representativo da segunda parcela de agrupamentos das triplicadas das quatro chapas modelo A, utilizando pré-processamento: centrado na média, distância euclidiana e método de ligação flexível com transformada SNV (dendrograma obtido utilizando o software Pirouette 3.11, <http://www.infometrix.com>, EUA).

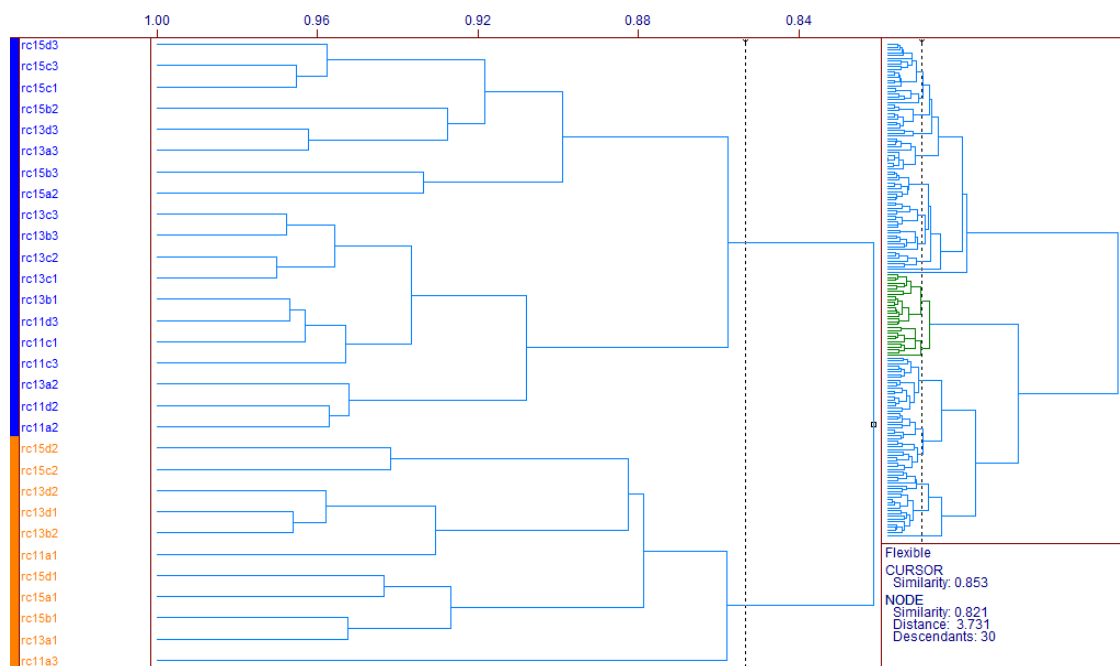


Figura 14 – Dendrograma representativo da terceira parcela de agrupamentos das triplicadas das quatro chapas modelo A, utilizando pré-processamento: centrado na média, distância euclidiana e método de ligação flexível com transformada SNV (dendrograma obtido utilizando o software Pirouette 3.11, <http://www.infometrix.com>, EUA).

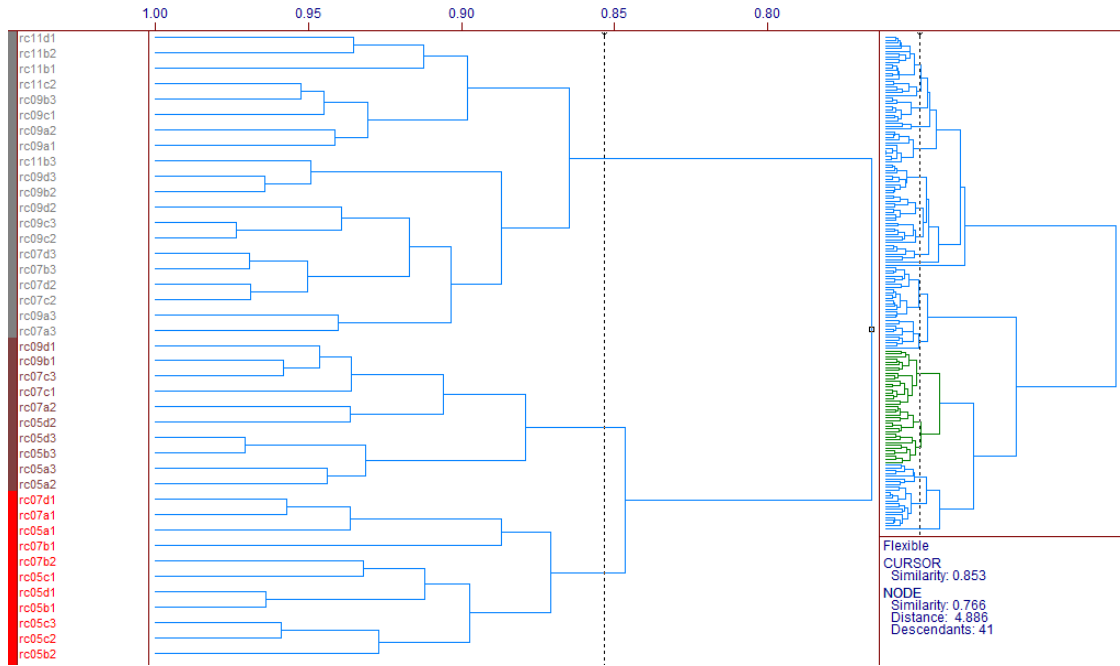


Figura 15 – Dendrograma representativo da quarta parcela de agrupamentos das triplicadas das quatro chapas modelo A, utilizando pré-processamento: centrado na média, distância euclidiana e método de ligação flexível com transformada SNV (dendrograma obtido utilizando o software Pirouette 3.11, <http://www.infometrix.com>, EUA).

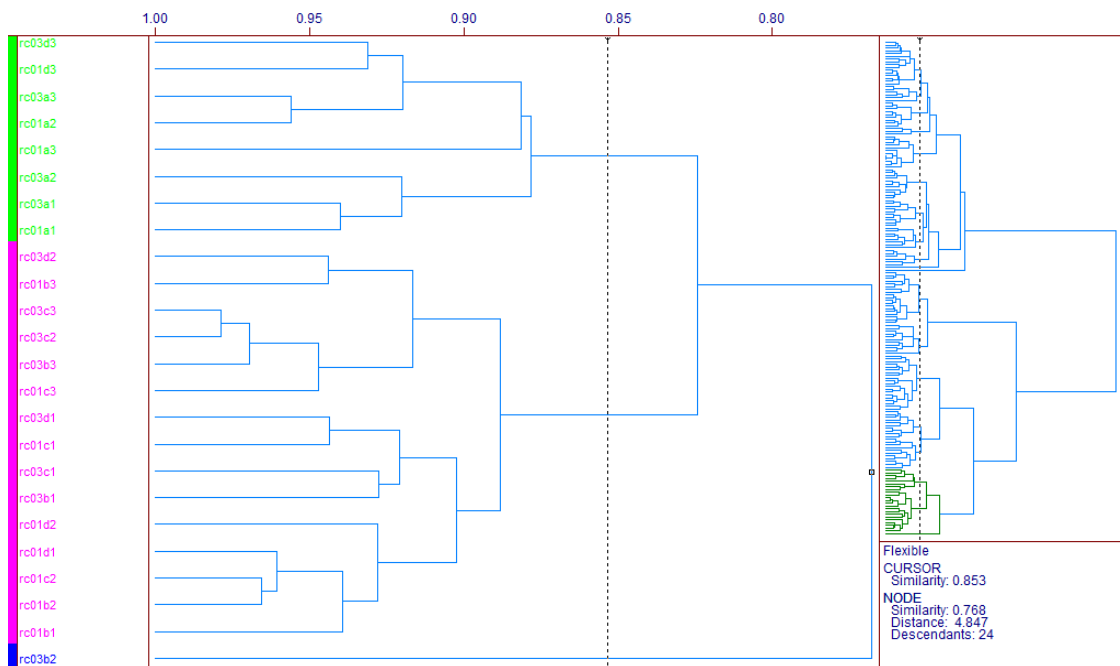


Figura 16 – Dendrograma representativo da quinta parcela de agrupamentos das triplicadas das quatro chapas modelo A, utilizando pré-processamento: centrado na média, distância euclidiana e método de ligação flexível com transformada SNV (dendrograma obtido utilizando o software Pirouette 3.11, <http://www.infometrix.com>, EUA).

Pode-se notar que os grupos são formados por faixas de concentração, ou seja, o método consegue identificar as diferenças nas faixas de concentração, porém devido a ruídos a sensibilidade fica comprometida, necessitando o mesmo ser aprimorado.

4.2 Resultados ATR

Na sequência serão apresentados os principais resultados para as quatro propriedades em estudo (quantidade de enxofre e biodiesel, massa específica e ponto de fulgor), para as blends de biodiesel/diesel empregando os espectros de ATR – FTIR. O conjunto de espectros é apresentado na Figura 17.

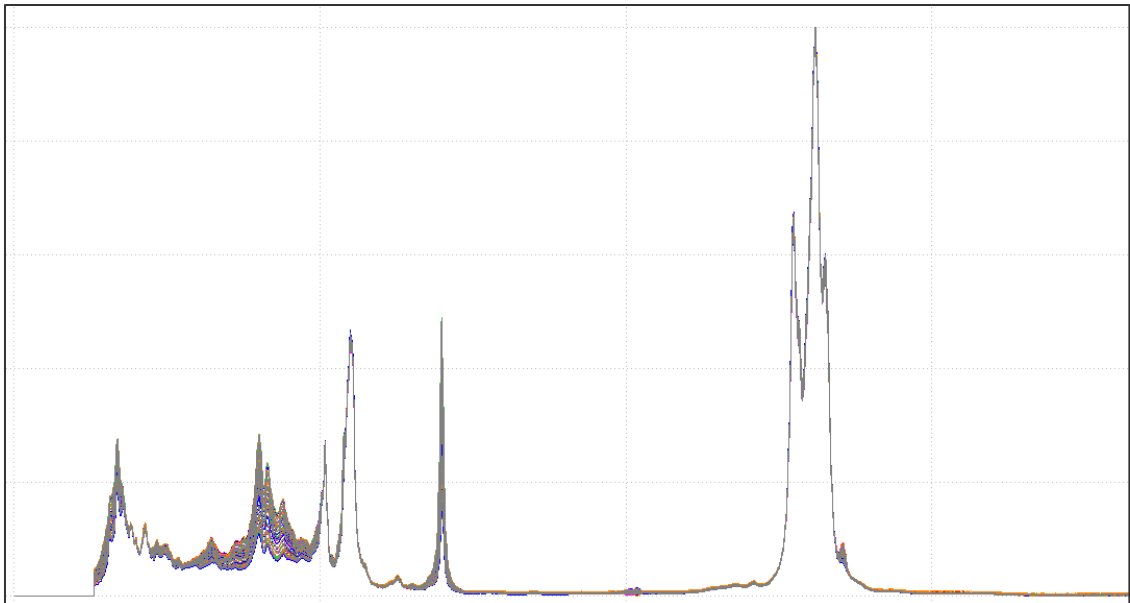


Figura 17 - Gráfico com os espectros ATR – FTIR para as quarenta e cinco blends com faixa concentração variando de 8 a 30,1 %(v/v).

Fonte: Elaborada pelo autor.

Foram desenvolvidos modelos empregando os dados brutos, centrados na média, autoescalados, centrados na média com correção do espalhamento de luz e os dados auto escalados com correção do espalhamento de luz.

Para cada propriedade foram desenvolvidos modelos empregando o espectro inteiro ou através do iPLS, dividindo o espectro em dois, quatro, oito, dezesseis e trinta e dois intervalos. Ainda, foram gerados modelos empregando o

algoritmo siPLS, dividindo o espectro em trinta e duas partes, combinando dois, três ou quatro intervalos.

4.2.1 Teor de enxofre

Os resultados para a quantificação do enxofre, utilizando o algoritmo iPLS e o espectro inteiro (*full spectrum*) são apresentados nas tabelas de 2 a 6, comparativamente ao modelo que emprega o espectro total.

Os modelos visando a quantificação do enxofre, e que empregaram todo o espectro, variaram o seu RMSECV na faixa de 31,6 a 40,3 mg/kg. Deve-se destacar que o melhor RMSECV foi do modelo que empregou os dados autoescalados e correção do espalhamento de luz sendo que o RMSEP ficou na mesma faixa, porém elevado comparativamente ao modelo que empregou os dados centrados na média e correção do espalhamento de luz, que pode ser considerado o melhor modelo quando emprega-se todo o espectro.

TABELA 2 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DO ENXOFRE NAS BLENDAS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR BRUTOS

Modelo	VL^a	Int^b	RMSECV (mg/kg)	R	RMSEC (mg/kg)	R	RMSEP (mg/kg)	R
Full	5	TODOS	33,7	0,8886	19,7	0,9625	30,4	0,9164
i2	4	2	29,3	0,9151	22,7	0,9501	25,0	0,9448
i4	4	4	29,2	0,9159	22,6	0,9503	24,3	0,9466
i8	5	8	28,5	0,9210	18,7	0,9664	12,0	0,9569
i16	6	15	17,9	0,9693	12,3	0,9855	17,8	0,9761
i32	6	30	18,9	0,9658	10,9	0,9887	21,1	0,9756

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 3 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DO ENXOFRE NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR CENTRADOS NA MÉDIA

Modelo	VL^a	Int^b	RMSECV (mg/kg)	R	RMSEC (mg/kg)	R	RMSEP (mg/kg)	R
Full	4	TODOS	34,0	0,8868	19,9	0,9617	30,4	0,9164
i2	5	2	30,3	0,9103	18,0	0,9689	21,1	0,9603
i4	5	4	27,8	0,9258	16,1	0,9750	16,6	0,9746
i8	5	8	28,5	0,9217	15,2	0,9777	15,4	0,9785
i16	4	13	27,9	0,9249	18,3	0,9676	18,2	0,9829
i32	5	26	20,1	0,9614	14,3	0,9803	8,8	0,9934

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 4 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DO ENXOFRE NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR AUTOESCALADOS

Modelo	VL^a	Int^b	RMSECV (mg/kg)	R	RMSEC (mg/kg)	R	RMSEP (mg/kg)	R
Full	2	TODOS	40,3	0,835	31,5	0,9008	38,8	0,8497
i2	4	2	38,3	0,8538	28,2	0,9215	42,3	0,8234
i4	4	3	36,7	0,8642	27,6	0,9248	47,3	0,7737
i8	4	7	29,5	0,9158	19,4	0,9635	11,8	0,9880
i16	5	13	28,7	0,9219	14,8	0,9790	15,2	0,9866
i32	5	30	18,7	0,9665	12,1	0,9860	21,8	0,9647

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 5 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DO ENXOFRE NAS BLENDAIS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR CENTRADOS NA MÉDIA E CORREÇÃO DO ESPALHAMENTO DE LUZ (MSC)

Modelo	VL^a	Int^b	RMSECV (mg/kg)	R	RMSEC (mg/kg)	R	RMSEP (mg/kg)	R
Full	4	TODOS	31,9	0,9007	16,2	0,9748	19,3	0,9661
i2	2	2	31,1	0,9039	26,8	0,9296	34,1	0,8916
i4	2	4	28,1	0,9220	24,1	0,9435	27,2	0,9326
i8	3	8	26,4	0,9321	18,8	0,9660	15,2	0,9788
i16	4	15	18,0	0,9688	14,6	0,9795	16,9	0,9790
i32	4	30	18,6	0,9669	12,3	0,9855	21,6	0,9765

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 6 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DO ENXOFRE NAS BLENDAIS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR AUTOESCALADOS E CORREÇÃO DO ESPALHAMENTO DE LUZ (MSC)

Modelo	VL^a	Int^b	RMSECV (mg/kg)	R	RMSEC (mg/kg)	R	RMSEP (mg/kg)	R
Full	4	TODOS	31,6	0,9014	21,6	0,9549	31,3	0,9076
i2	3	2	32,5	0,8947	24,8	0,9400	33,2	0,8956
i4	6	2	24,1	0,9446	11,9	0,9866	18,9	0,9675
i8	3	8	30,7	0,9070	20,7	0,9585	25,4	0,9409
i16	3	13	30,0	0,9124	21,3	0,956	16,5	0,9753
i32	3	30	22,6	0,9594	18,1	0,9686	23,3	0,9584

^a VL = variáveis latentes

^b Int = intervalo selecionado

Podemos observar na Tabela 5, onde são empregados os dados centrados na média com correção do espalhamento de luz, que os resultados são mais satisfatórios comparativamente aos resultados das demais tabelas. Este modelo utilizou o intervalo 15 que foi o que apresentou o menor erro de validação cruzada com um RMSEC de 14,6 mg/kg e um RMSEP de 16,9 mg/kg, utilizando apenas quatro variáveis latentes.

A Figura 18 mostra um gráfico que relaciona o RMSECV de cada um dos intervalos e o do modelo global.

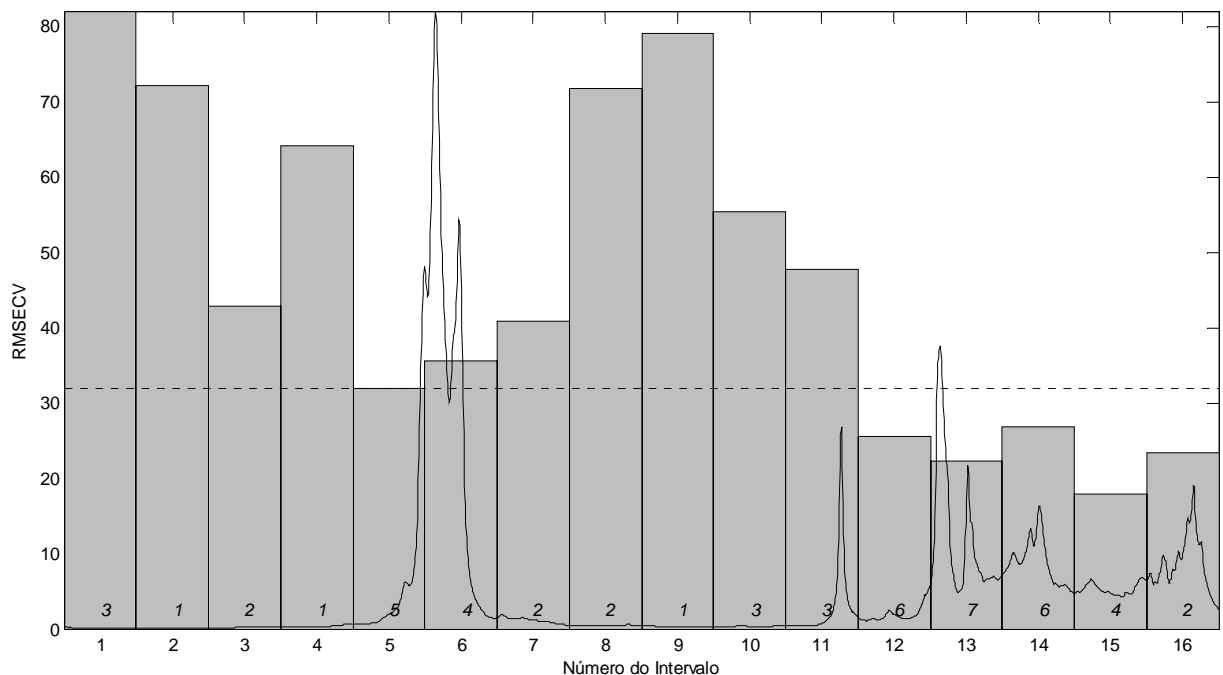


Figura 18 – Gráfico do RMSECV para os modelos dos 16 intervalos (barras) e modelo global empregando 4 variáveis latentes (linha pontilhada) para a determinação do enxofre nas blendas biodiesel/diesel, com dados centrados na média e correção do espalhamento de luz.

Fonte: Elaborada pelo autor

Podemos verificar na Figura 19 o gráfico de correlação entre o método de referência e a previsão feita pelo iPLS ATR – FTIR, empregando a faixa espectral que corresponde a 858 até 1064 cm^{-1} (intervalo 15). O coeficiente de correlação para o modelo de calibração foi de 0,979.

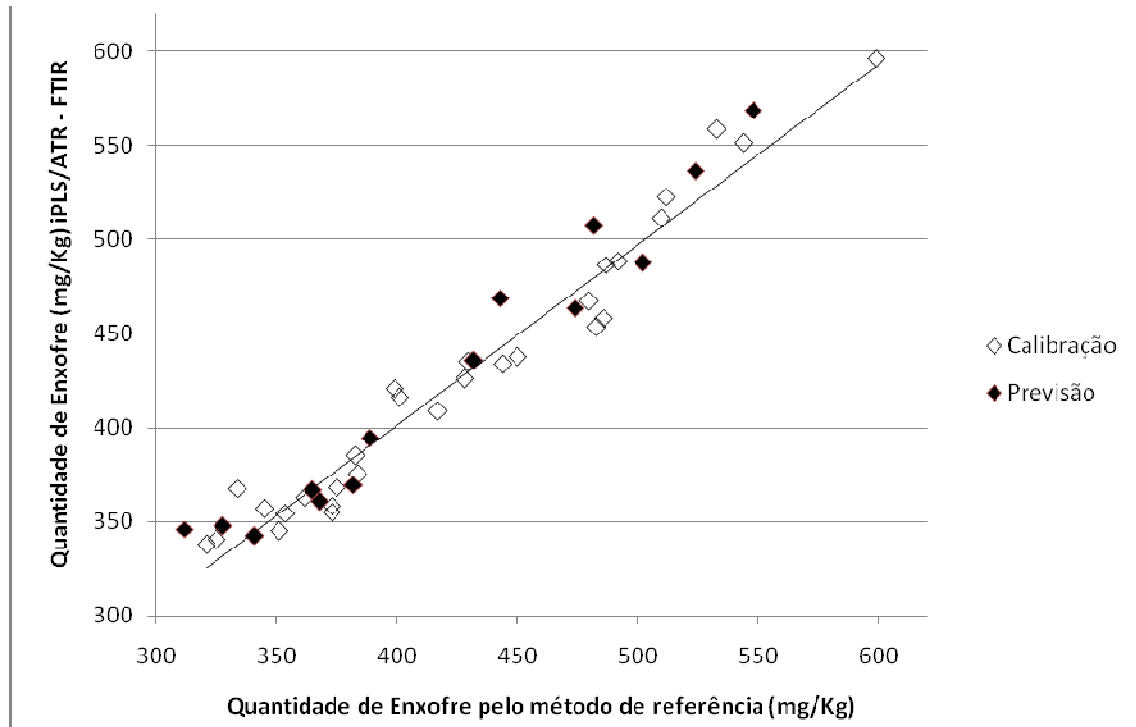


Figura 19 – Gráfico da correlação entre o método de referência e a previsão por iPLS/ATR-FTIR, empregando a faixa espectral 858 - 1064 cm^{-1} (intervalo 15) e 4 variáveis latentes para a determinação do enxofre nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.

Fonte: Elaborada pelo autor

Na Tabela 7 são apresentados os resultados para os modelos siPLS para a determinação do enxofre. Podemos observar que o modelo que combinou dois intervalos (12 e 30) e utilizou os dados centrados na média com correção do espalhamento de luz apresentou uma redução nos erros de forma geral, obtendo-se um RMSEC de 11,1 mg/kg e um RMSEP de 14,8 mg/kg, sendo utilizadas apenas quatro variáveis latentes.

TABELA 7 – VALORES PARA OS ERROS DE VALIDAÇÃO CRUZADA, CALIBRAÇÃO E PREVISÃO PARA OS MODELOS SIPLS, VISANDO A DETERMINAÇÃO DO ENXOFRE NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR

	Nº de intervalos	VL ^a	Int ^b	RMSECV (mg/kg)	RMSEC (mg/kg)	RMSEP (mg/kg)
N	2	6	12 30	16,1	11,9	15,4
	3	7	8 29 30	13,8	8,8	17,1
	4	8	8 20 29 30	14,7	7,4	17,8
M	2	6	12 30	16,9	11,5	14,3
	3	6	12 13 30	17,1	11,5	16,7
	4	6	1 12 30 32	16,8	8,8	16,1
A	2	5	30 32	18,6	12,0	12,0
	3	7	22 30 32	14,2	7,9	17,4
	4	8	22 27 30 32	14,3	6,4	18,1
MM	2	4	12 30	14,1	11,1	14,8
	3	4	12 13 30	15,0	11,3	15,0
	4	5	12 17 19 30	14,3	9,6	15,2
MA	2	5	12 30	13,9	9,0	14,5
	3	5	12 26 31	14,5	10,2	12,6
	4	6	13 25 30 32	13,4	8,3	13,6

^a VL = variáveis latentes

^b Int = intervalos selecionados

Na Figura 20, podemos observar as faixas espectrais utilizadas na construção deste modelo que correspondem a $858 - 960 \text{ cm}^{-1}$ combinada com a faixa de $2732 - 2834 \text{ cm}^{-1}$.

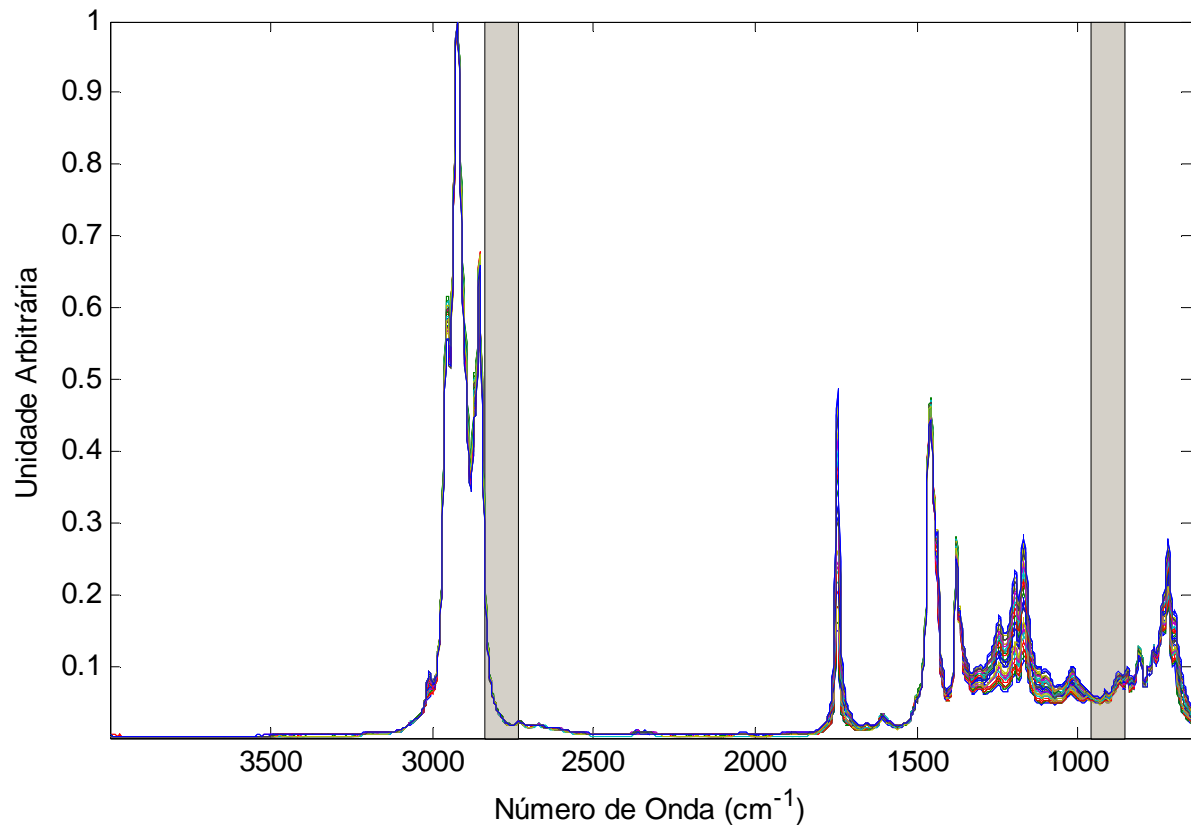


Figura 20 – Gráfico das faixas espectrais utilizadas na construção do modelo siPLS para a determinação do enxofre nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.

Fonte: Elaborada pelo autor

Na Figura 21 é apresentada a correlação entre o método de referência e a previsão feita pelo siPLS ATR – FTIR, onde o coeficiente de correlação foi de 0,9809.

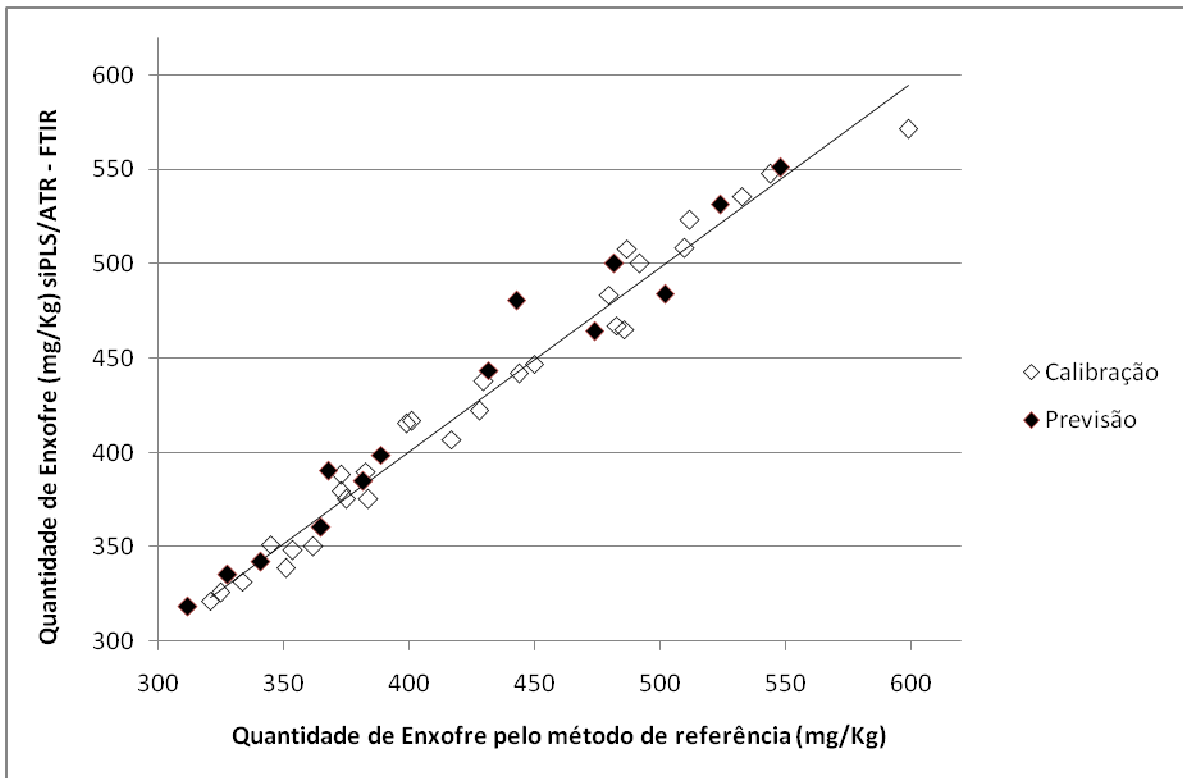


Figura 21 – Gráfico da correlação entre o método de referência e a previsão por siPLS/ATR-FTIR, empregando as faixas espectrais $858 - 960 \text{ cm}^{-1}$ e $2732 - 2834 \text{ cm}^{-1}$ (intervalos 12 e 30) e 4 variáveis latentes para a determinação do enxofre nas blends biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.
Fonte: Elaborada pelo autor

4.2.2 Biodiesel

Para o biodiesel, também, foram criados cinco modelos empregando o espectro total, variando-se por tipo de pre-processamento. A faixa de RMSECV para este modelo foi de 0,388 a 0,468%(v/v). Destaca-se que o melhor modelo, dentre os cinco desenvolvidos, foi aquele que empregou os dados centrados na média e a correção do espalhamento de luz, sendo também o que utilizou o menor número de variáveis latentes, sendo necessária apenas uma para a obtenção deste modelo.

Nas tabelas de 8 a 12 são apresentados os resultados para os modelos iPLS e com o espectro inteiro.

TABELA 8 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DA PORCENTAGEM DE BIODIESEL NAS BLENAS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR BRUTOS

Modelo	VL^a	Int^b	RMSECV %(v/v)	R	RMSEC %(v/v)	R	RMSEP %(v/v)	R
Full	2	TODOS	0,439	0,9978	0,411	0,9980	0,411	0,9982
i2	3	2	0,366	0,9985	0,309	0,9989	0,190	0,9996
i4	3	4	0,360	0,9985	0,302	0,9990	0,225	0,9995
i8	3	7	0,360	0,9985	0,304	0,9989	0,181	0,9998
i16	5	16	0,353	0,9986	0,241	0,9993	0,224	0,9997
i32	4	32	0,366	0,9985	0,276	0,9991	0,172	0,9997

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 9 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DA PORCENTAGEM DE BIODIESEL NAS BLENAS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR CENTRADOS NA MÉDIA

Modelo	VL^a	Int^b	RMSECV %(v/v)	R	RMSEC %(v/v)	R	RMSEP %(v/v)	R
Full	2	TODOS	0,456	0,9976	0,400	0,9982	0,383	0,9984
i2	3	2	0,381	0,9983	0,314	0,9989	0,195	0,9996
i4	3	4	0,371	0,9984	0,305	0,9989	0,219	0,9995
i8	3	7	0,361	0,9985	0,298	0,9990	0,161	0,9998
i16	4	16	0,386	0,9983	0,278	0,9991	0,201	0,9996
i32	3	31	0,354	0,9986	0,311	0,9989	0,346	0,9993

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 10 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DA PORCENTAGEM DE BIODIESEL NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR AUTOESCALADOS

Modelo	VL^a	Int^b	RMSECV %(v/v)	R	RMSEC %(v/v)	R	RMSEP %(v/v)	R
Full	4	TODOS	0,441	0,9978	0,337	0,9987	0,385	0,9986
i2	4	2	0,443	0,9977	0,350	0,9986	0,303	0,9989
i4	4	4	0,407	0,9981	0,304	0,9989	0,146	0,9998
i8	3	7	0,384	0,9983	0,312	0,9989	0,150	0,9998
i16	5	16	0,332	0,9987	0,229	0,9994	0,234	0,9997
i32	3	31	0,343	0,9986	0,301	0,9990	0,338	0,9993

^a VL = variáveis latentes

^b Int = intervalo selecionado

A Tabela 11 mostra que o modelo iPLS que utilizou os dados centrados na média com correção do espalhamento de luz utilizou apenas uma variável latente resultando em RMSEC de 0,295%(v/v) e RMSEP de 0,154%(v/v) e coeficiente de correlação de 0,999 para o modelo de calibração, sendo este o modelo iPLS escolhido.

TABELA 11 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DA PORCENTAGEM DE BIODIESEL NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR CENTRADOS NA MÉDIA E CORREÇÃO DO ESPALHAMENTO DE LUZ (MSC).

Modelo	VL^a	Int^b	RMSECV %(v/v)	R	RMSEC %(v/v)	R	RMSEP %(v/v)	R
Full	1	TODOS	0,388	0,9983	0,362	0,9985	0,306	0,9988
i2	5	2	0,398	0,9982	0,256	0,9992	0,179	0,9996
i4	1	4	0,331	0,9987	0,305	0,9989	0,283	0,9990
i8	1	7	0,321	0,9988	0,295	0,999	0,154	0,9997
i16	2	16	0,695	0,9944	0,595	0,9959	0,511	0,9966
i32	2	10	0,376	0,9984	0,320	0,9988	0,233	0,9993

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 12 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DA PORCENTAGEM DE BIODIESEL NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR AUTOESCALADOS E CORREÇÃO DO ESPALHAMENTO DE LUZ (MSC).

Modelo	VL^a	Int^b	RMSECV %(v/v)	R	RMSEC %(v/v)	R	RMSEP %(v/v)	R
Full	2	TODOS	0,468	0,9975	0,383	0,9983	0,336	0,9986
i2	6	1	0,521	0,9971	0,208	0,9995	0,423	0,9976
i4	2	4	0,377	0,9984	0,302	0,9989	0,227	0,9993
i8	1	7	0,339	0,9987	0,309	0,9989	0,241	0,9993
i16	2	13	0,567	0,9963	0,472	0,9974	0,253	0,9992
i32	3	26	0,478	0,9974	0,393	0,9982	0,434	0,9976

^a VL = variáveis latentes

^b Int = intervalo selecionado

Se observarmos a Figura 22, para qual o espectro foi dividido em 8 partes, o intervalo 7 é aquele que apresenta o menor RMSECV, único com valor abaixo do modelo global, empregando apenas uma variável latente.

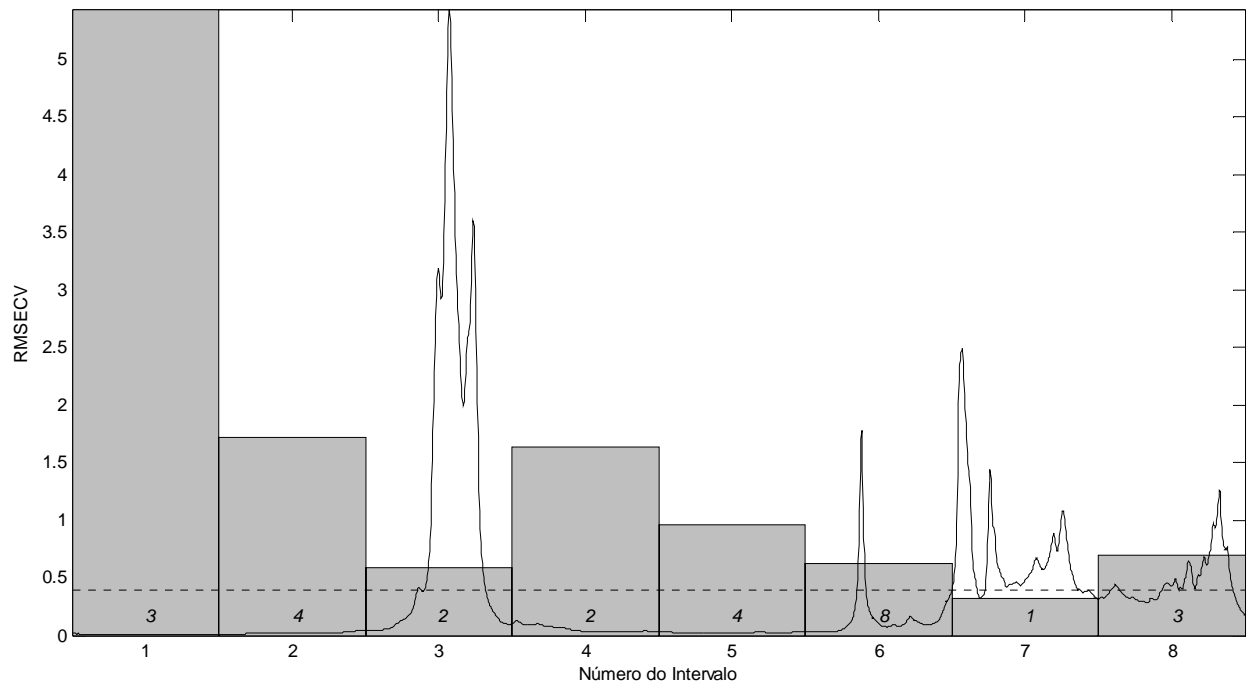


Figura 22 – Gráfico do RMSECV para os modelos dos 8 intervalos (barras) e modelo global empregando 1 variáveis latentes (linha pontilhada) para a determinação da % de biodiesel nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.

Fonte: Elaborada pelo autor

Na Figura 23 é apresentada a curva de correlação entre os valores de referência e os valores previstos pelo modelo iPLS ATR-FTIR.

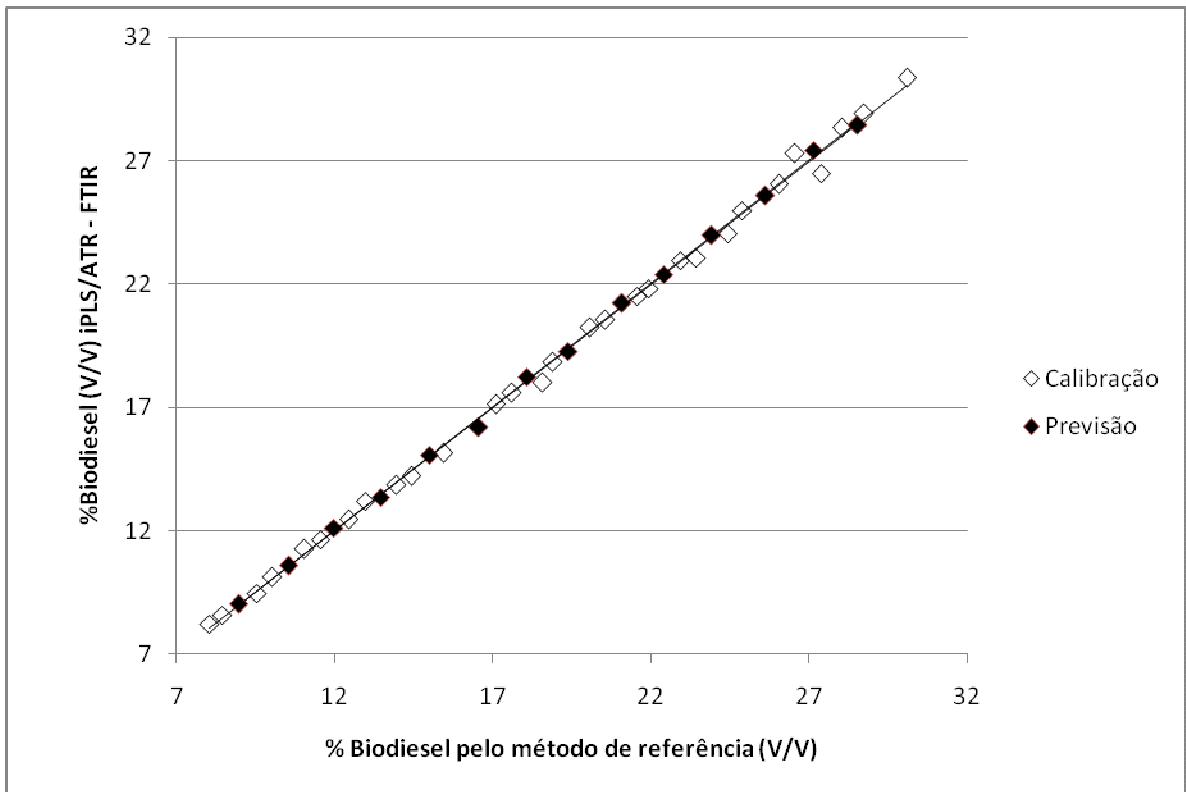


Figura 23 – Gráfico da correlação entre o método de referência e a previsão por iPLS/ATR-FTIR, empregando a faixa espectral $1068 - 1484 \text{ cm}^{-1}$ (intervalo 7) e 1 variável latente para a determinação da % de biodiesel nas blends biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.

Fonte: Elaborada pelo autor

A faixa espectral $1068 - 1484 \text{ cm}^{-1}$ selecionada para a construção do modelo iPLS para a determinação do biodiesel nas blends corresponde a deformação angular simétrica fora do plano das ligações $=\text{C-H}$ de olefinas, bem como, do estiramento assimétrico C-C(O) - dos ésteres do biodiesel (BARBOSA, 2007).

Os modelos desenvolvidos com o algoritmo siPLS são apresentados na Tabela 13.

TABELA 13 – VALORES PARA OS ERROS DE VALIDAÇÃO CRUZADA, CALIBRAÇÃO E PREVISÃO PARA OS MODELOS SIPLS, VISANDO A DETERMINAÇÃO DA % DE BIODIESEL NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR.

	Nº de intervalos	VL ^a	Int ^b	RMSECV %(v/v)	RMSEC %(v/v)	RMSEP %(v/v)
N	2	4	29 31	0,312	0,252	0,263
	3	8	2 29 30	0,247	0,066	0,267
	4	8	2 8 23 31	0,244	0,079	0,295
M	2	4	16 31	0,323	0,259	0,264
	3	8	2 29 30	0,245	0,056	0,290
	4	9	2 18 29 30	0,253	0,050	0,288
A	2	6	19 31	0,308	0,193	0,334
	3	9	21 22 29	0,268	0,069	0,528
	4	10	21 22 29 31	0,232	0,062	0,425
MM	2	2	20 31	0,307	0,273	0,203
	3	4	12 29 31	0,274	0,226	0,243
	4	8	9 10 28 31	0,229	0,079	0,274
MA	2	2	25 31	0,316	0,272	0,184
	3	7	21 23 25	0,252	0,140	0,150
	4	6	5 14 19 31	0,231	0,083	0,306

^a VL = variáveis latentes

^b Int = intervalos selecionados

Comparativamente, o melhor modelo foi o que utilizou os dados centrados na média e correção do espalhamento de luz, combinando 3 intervalos (12, 29 e 31) e empregando 4 variáveis latentes. Entretanto, este modelo não superou o resultado obtido pelo algoritmo iPLS, no qual não só os erros foram menores mas também, principalmente, pelo fato de ter utilizado apenas uma variável latente e ter obtido um coeficiente de correlação próximo de 1,00.

Na Figura 24 são apresentadas as regiões selecionadas para a construção do melhor modelo siPLS e na Figura 25 o gráfico de correlação entre os valores do método de referência e os previstos pelo modelo siPLS ATR-FTIR.

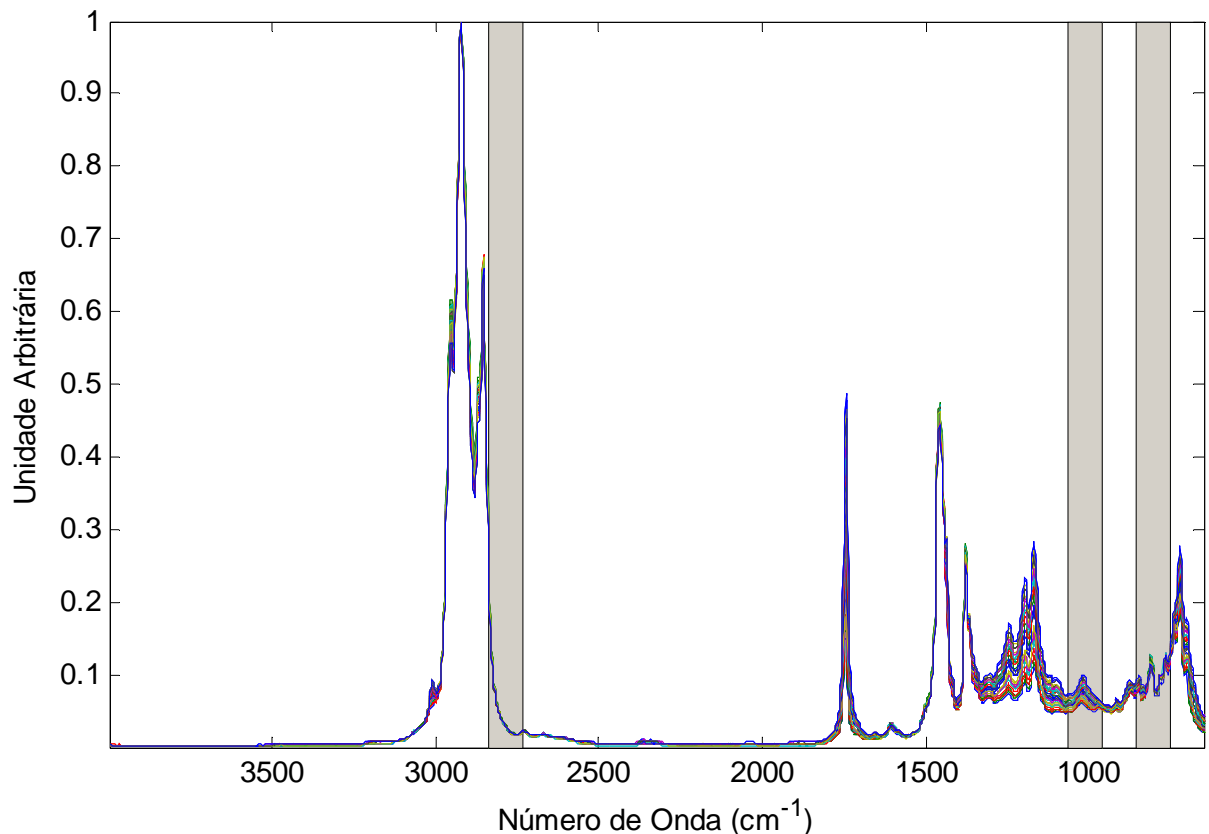


Figura 24 – Gráfico das faixas espectrais utilizadas na construção do modelo siPLS para a determinação da % de biodiesel nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.

Fonte: Elaborada pelo autor

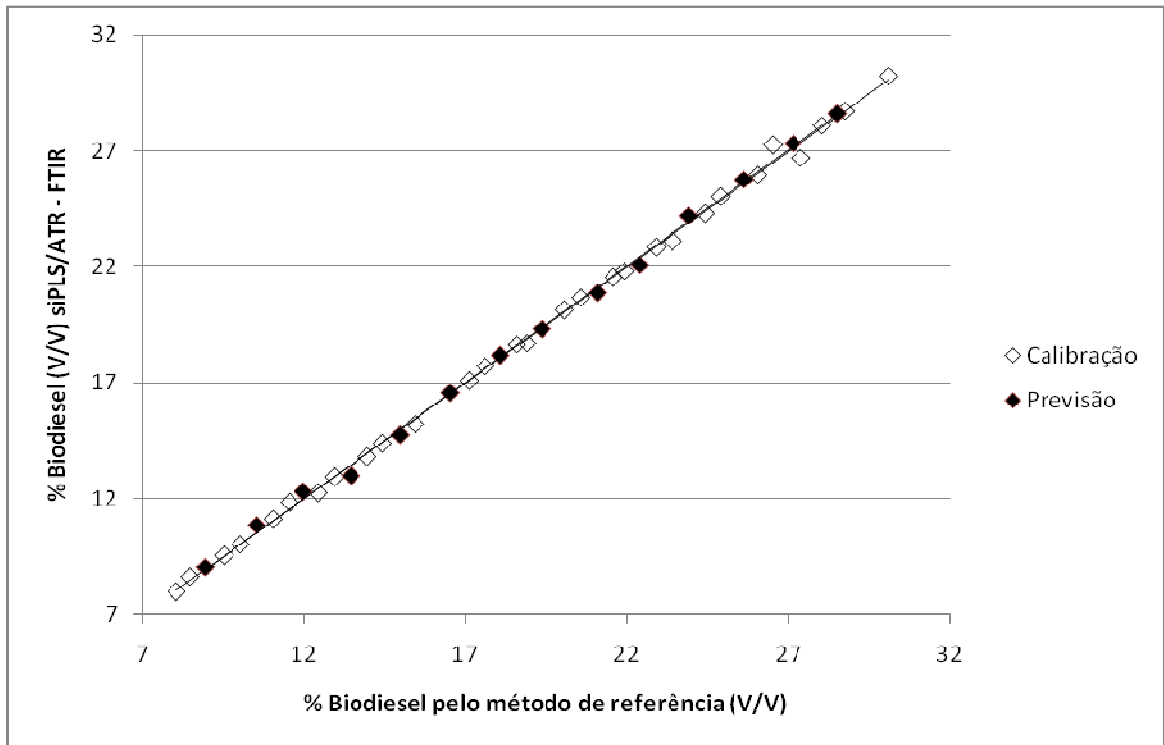


Figura 25 – Gráfico da correlação entre o método de referência e a previsão por siPLS/ATR-FTIR, empregando as faixas espectrais $856 - 754 \text{ cm}^{-1}$, $962 - 1064 \text{ cm}^{-1}$ e $2732 - 2834 \text{ cm}^{-1}$ (intervalos 12, 29 e 31) e 4 variáveis latentes para a determinação da % de biodiesel nas blends biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.

Fonte: Elaborada pelo autor

No caso do modelo siPLS a faixa espectral $962 - 1064 \text{ cm}^{-1}$ também corresponde deformação angular simétrica fora do plano das ligações $=\text{C-H}$ de olefinas, bem como, do estiramento assimétrico C-C(O)-O dos ésteres do biodiesel (BARBOSA, 2007). Já a faixa espectral $2732 - 2834 \text{ cm}^{-1}$ corresponde aos estiramentos simétricos e assimétricos dos CH_2 e CH_3 das cadeias do diesel (BARBOSA, 2007).

4.2.3 Massa específica

Para a quantificação da massa específica das blends estudadas o erro de validação cruzada foi de $1,57$ a $1,91 \text{ kg/m}^3$ para os cinco modelos empregando o espectro total. Cabe destacar que o melhor modelo foi aquele que utilizou os dados centrados na média e correção do espalhamento de luz.

A utilização do algoritmo iPLS, para seleção da região espectral mais

representativa, na construção dos modelos de determinação da massa específica, resultou em modelos de melhor capacidade preditiva, conforme podem ser visualizados nas Tabelas de 14 a 18.

TABELA 14 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DA MASSA ESPECÍFICA DAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR BRUTOS

Modelo	VL^a	Int^b	RMSECV (kg/m³)	R	RMSEC (kg/m³)	R	RMSEP (kg/m³)	R
Full	5	TODOS	1,660	0,8934	0,868	0,9711	1,490	0,9396
i2	4	2	1,450	0,9173	1,090	0,9542	1,260	0,9576
i4	4	4	1,530	0,9083	1,110	0,9527	1,160	0,9633
i8	5	8	1,440	0,9197	0,862	0,9716	0,745	0,9822
i16	6	15	0,862	0,9721	0,472	0,9916	0,673	0,9869
i32	3	31	1,440	0,9180	1,290	0,9354	1,090	0,9728

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 15 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DA MASSA ESPECÍFICA NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR CENTRADOS NA MÉDIA

Modelo	VL^a	Int^b	RMSECV (kg/m³)	R	RMSEC (kg/m³)	R	RMSEP (kg/m³)	R
Full	4	TODOS	1,640	0,8948	0,883	0,9701	1,490	0,9392
i2	5	2	1,460	0,9172	0,788	0,9762	1,140	0,9631
i4	5	4	1,40	0,9257	0,693	0,9817	0,839	0,9786
i8	5	8	1,420	0,9229	0,612	0,9858	0,98	0,9691
i16	4	13	1,430	0,9222	0,812	0,9748	0,743	0,9879
i32	6	26	0,547	0,9888	0,284	0,9969	0,372	0,9957

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 16 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DA MASSA ESPECÍFICA NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR AUTOESCALADOS

Modelo	VL^a	Int^b	RMSECV (kg/m³)	R	RMSEC (kg/m³)	R	RMSEP (kg/m³)	R
Full	2	TODOS	1,910	0,8522	1,520	0,9089	2,090	0,8454
i2	5	1	1,850	0,8638	1,040	0,9579	2,580	0,7532
i4	3	4	1,960	0,8454	1,550	0,9046	2,060	0,8570
i8	4	7	1,530	0,9112	0,926	0,967	0,606	0,9898
i16	3	13	1,500	0,9116	1,220	0,9418	1,180	0,9645
i32	5	30	0,720	0,9804	0,380	0,9945	0,995	0,9694

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 17 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DA MASSA ESPECÍFICA NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR CENTRADOS NA MÉDIA E CORREÇÃO DO ESPALHAMENTO DE LUZ (MSC)

Modelo	VL^a	Int^b	RMSECV (kg/m³)	R	RMSEC (kg/m³)	R	RMSEP (kg/m³)	R
Full	4	TODOS	1,570	0,9039	0,680	0,9824	0,913	0,9751
i2	2	2	1,490	0,9123	1,280	0,9362	1,830	0,8893
i4	2	4	1,390	0,9243	1,160	0,9475	1,430	0,9395
i8	3	8	1,350	0,9288	0,904	0,9686	0,803	0,9800
i16	2	16	1,240	0,9401	1,060	0,9568	0,946	0,9451
i32	3	26	0,881	0,9703	0,674	0,9827	0,865	0,9772

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 18 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DA MASSA ESPECÍFICA NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR AUTOESCALADOS E CORREÇÃO DO ESPALHAMENTO DE LUZ (MSC)

Modelo	VL ^a	Int ^b	RMSECV (kg/m ³)	R	RMSEC (kg/m ³)	R	RMSEP (kg/m ³)	R
Full	4	TODOS	1,630	0,8960	1,000	0,9615	1,700	0,9074
i2	3	2	1,580	0,9021	1,190	0,9450	1,780	0,8981
i4	2	4	1,800	0,8690	1,530	0,9076	1,730	0,9039
i8	3	8	1,540	0,9075	0,987	0,9625	1,510	0,9258
i16	3	13	1,510	0,9129	0,973	0,9636	1,100	0,9658
i32	4	26	0,765	0,9777	0,559	0,9881	0,679	0,9874

^a VL = variáveis latentes

^b Int = intervalo selecionado

Como se pode visualizar na Tabela 17, o melhor modelo iPLS dividiu o espectro em trinta e duas partes, com os dados centrados na média e correção do espalhamento de luz. A Figura 26 mostra o gráfico dos RMSECV, para este modelo escolhido, apresentando três regiões espectrais (intervalos) que apresentam menor erro, aproximadamente metade do erro do modelo global. Estes intervalos são 24, 26 e 30, sendo que o escolhido foi o 26 (correspondente a faixa espectral de 1274 – 1376 cm⁻¹) por utilizar o menor número de variáveis latentes, obtendo-se um erro equivalente aos demais. Para este modelo os valores de RMSEC e RMSEP foram de 0,674 kg/m³ e de 0,865 kg/m³, respectivamente.

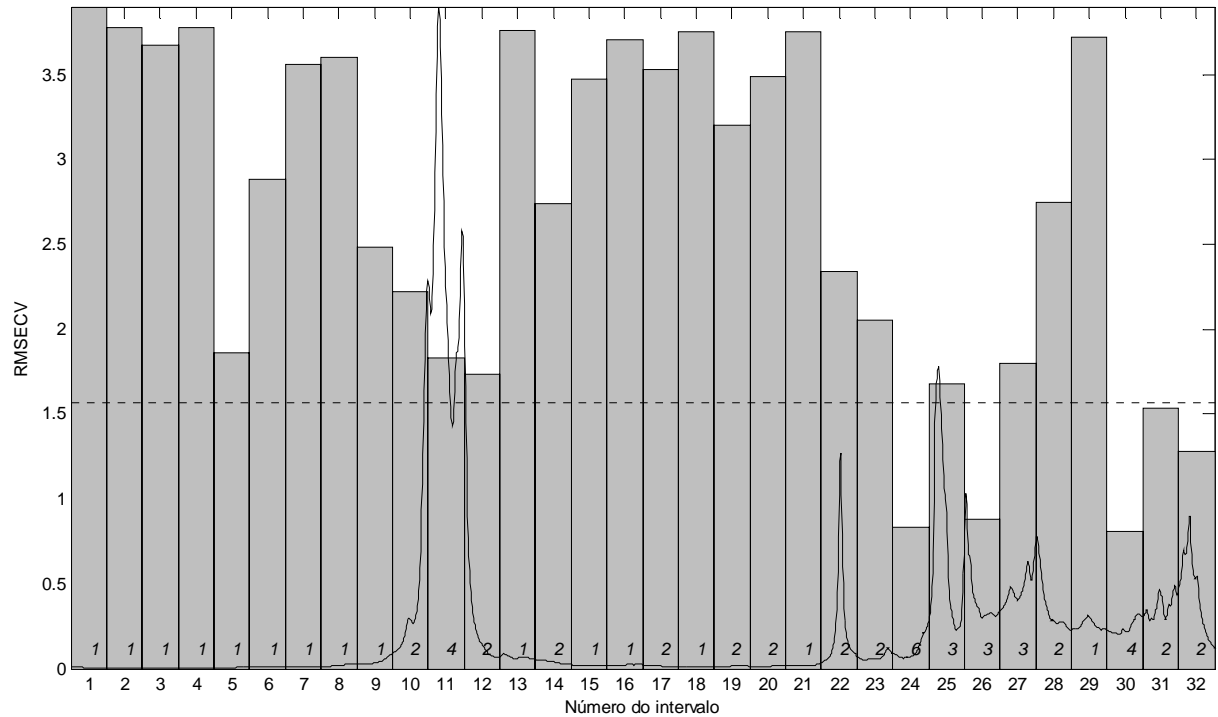


Figura 26 – Gráfico do RMSECV para os modelos dos 32 intervalos (barras) e modelo global empregando 3 variáveis latentes (linha pontilhada) para a determinação da massa específica nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.

Fonte: Elaborada pelo autor

Na Figura 27 é apresentada a curva de correlação entre os valores de referência e os valores previstos pelo modelo iPLS ATR-FTIR, para massa específica.

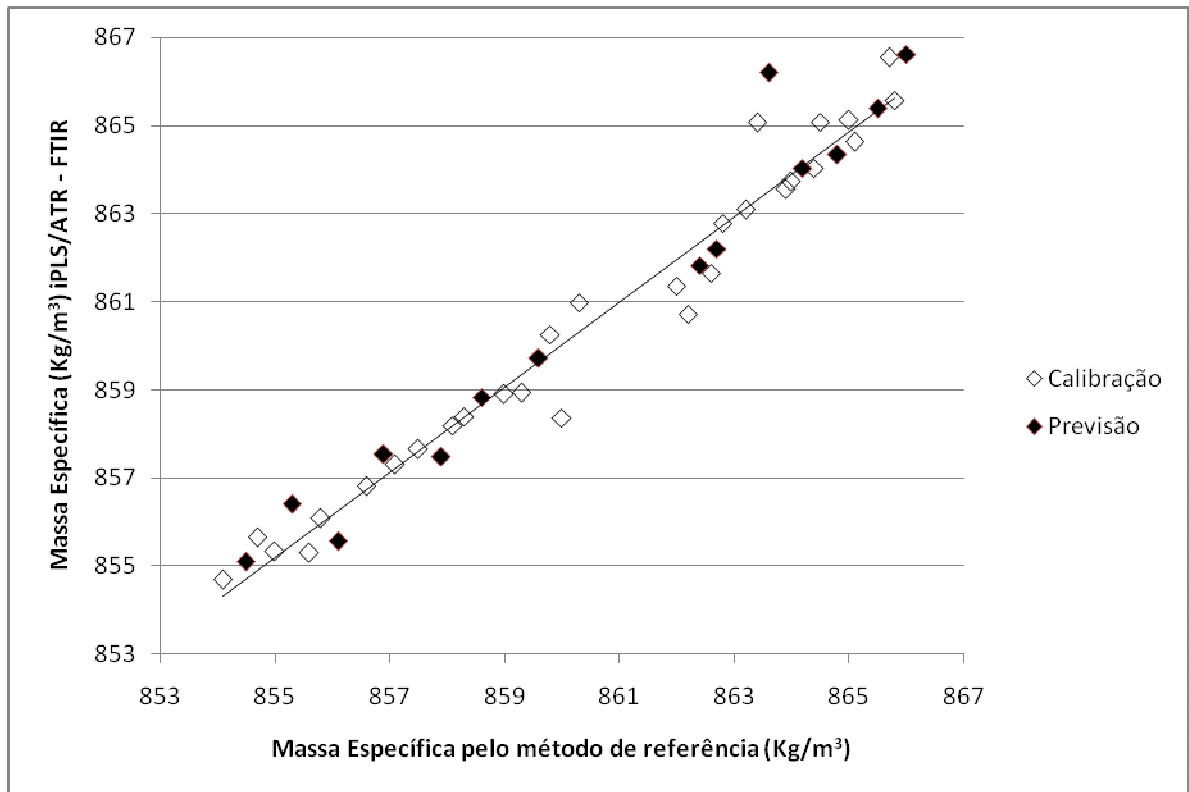


Figura 27 – Gráfico da correlação entre o método de referência e a previsão por iPLS/ATR-FTIR, empregando a faixa espectral $1274 - 1376 \text{ cm}^{-1}$ (intervalo 26) e 3 variáveis latentes para a determinação da massa específica nas blendas biodiesel/diesel, com os dados centrados na média e correção do espalhamento de luz.

Fonte: Elaborada pelo autor

A faixa espectral $1274 - 1376 \text{ cm}^{-1}$ selecionada para a construção do modelo iPLS para a determinação da massa específica das blendas corresponde ao estiramento assimétrico C-C(O)- dos ésteres do biodiesel (BARBOSA, 2007).

Os resultados para os modelos siPLS são apresentados na Tabela 19.

**TABELA 19 – VALORES PARA OS ERROS DE VALIDAÇÃO CRUZADA, CALIBRAÇÃO E PREVISÃO PARA OS
 MODELSO SIPLS, VISANDO A DETERMINAÇÃO DA MASSA ESPECÍFICA NAS BLENDDAS BIODIESEL/DIESEL
 EMPREGANDO OS DADOS ATR – FTIR**

	Nº de intervalos	VL ^a	Int ^b	RMSECV (kg/m ³)	RMSEC (kg/m ³)	RMSEP (kg/m ³)
N	2	9	26 30	0,431	0,159	0,362
	3	10	13 26 30	0,434	0,066	0,359
	4	12	22 26 27 30	0,357	0,050	0,225
M	2	9	26 30	0,408	0,127	0,385
	3	12	14 22 26	0,239	0,064	0,388
	4	12	14 15 22 26	0,227	0,059	0,413
A	2	9	22 26	0,257	0,118	0,358
	3	10	22 26 27	0,288	0,112	0,396
	4	8	22 26 27 30	0,391	0,168	0,436
MM	2	4	12 30	0,656	0,383	0,637
	3	7	9 26 30	0,544	0,223	0,404
	4	12	11 22 23 26	0,392	0,068	0,388
MA	2	4	11 27	0,747	0,575	1,017
	3	9	11 22 24	0,454	0,128	0,536
	4	9	11 17 22 26	0,289	0,137	0,555

^a VL = variáveis latentes

^b Int = intervalos selecionados

Neste caso, o melhor modelo foi aquele que utilizou os dados auto escalados e combinou os intervalos 22 e 26. Apesar deste modelo ter apresentado valores de RMSECV, RMSEC e RMSEP menores que o modelo iPLS este modelo necessitou 9 variáveis latentes, sendo então considerado o modelo iPLS, que usou somente três, o mais adequado para a determinação desta propriedade.

As regiões espectrais selecionadas para este modelo encontram-se na Figura 28 que correspondem as faixas de 1274 - 1376 e 1691 - 1793 cm^{-1} e a curva de correlação entre os valores de massa específica do método de referência e os previstos pelo modelo siPLS ATR-FTIR é apresentada na Figura 29.

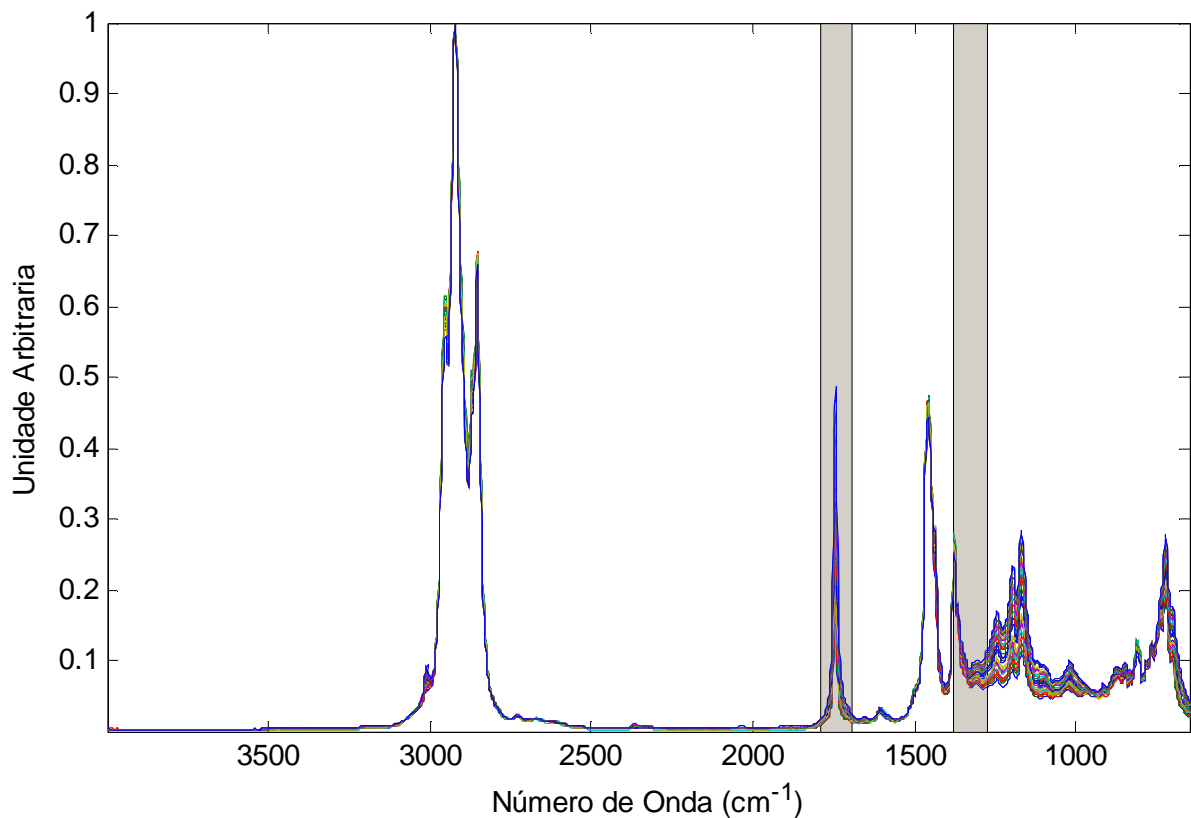


Figura 28 – Gráfico das faixas espectrais utilizadas na construção do modelo siPLS para a determinação da massa específica nas blendas biodiesel/diesel.
Fonte: Elaborada pelo autor

Neste caso além da faixa espectral 1274 - 1376 cm^{-1} , que corresponde ao estiramento assimétrico C-C(O)- dos ésteres do biodiesel, o modelo siPLS

selecionou a faixa espectral 1691 – 1793 cm^{-1} que corresponde ao estiramento do grupo C=O dos ésteres (BARBOSA, 2007).

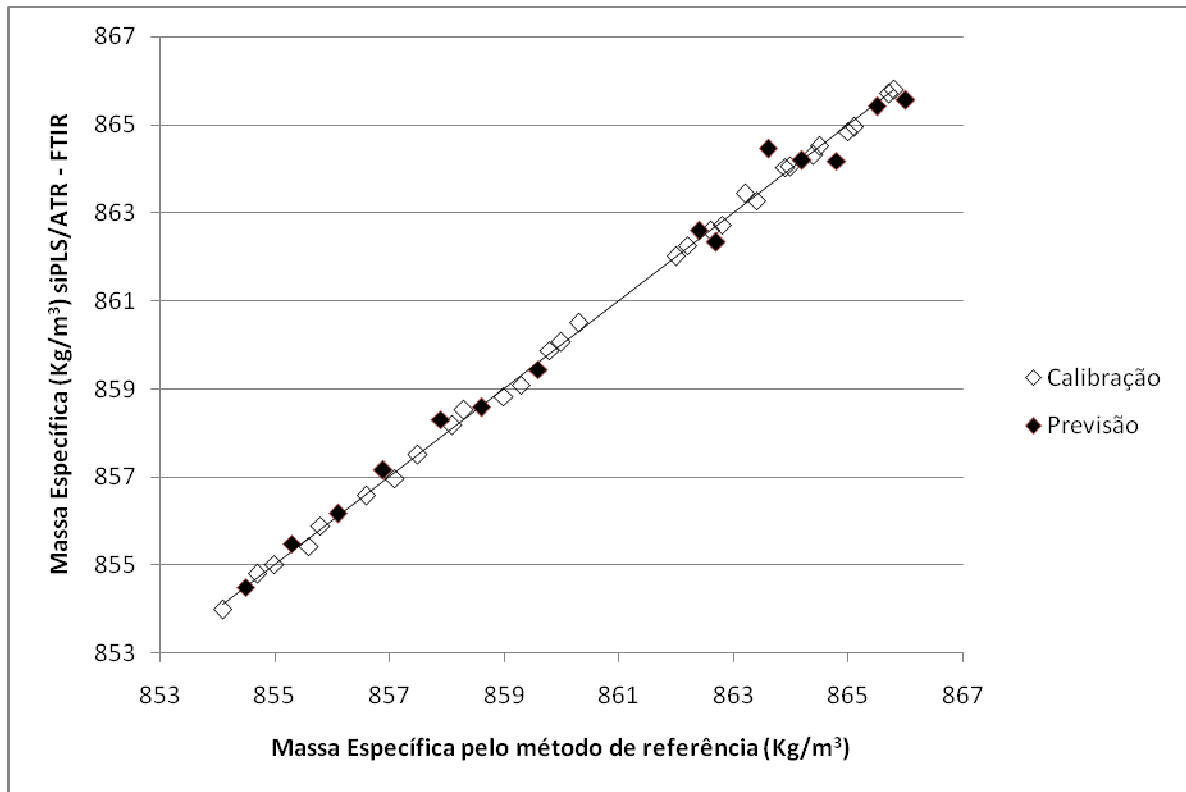


Figura 29 – Gráfico da correlação entre o método de referência e a previsão por siPLS/ATR-FTIR, empregando as faixas espectrais 1274 - 1376 cm^{-1} e 1691 - 1793 cm^{-1} e 9 variáveis latentes para a determinação da massa específica nas blendas biodiesel/diesel.

Fonte: Elaborada pelo autor

4.2.4 Ponto de fulgor

Para a quantificação do ponto de fulgor das blendas estudadas o erro de validação cruzada foi de 4,01 a 4,56 $^{\circ}\text{C}$ para os cinco modelos empregando o espectro total. Destes modelos, o melhor foi o que empregou os dados brutos tendo um RMSEC de 2,18 e um RMSEP de 3,53 $^{\circ}\text{C}$. O coeficiente de correlação do modelo de calibração foi de 0,964.

Os modelos iPLS e espectros completos são apresentados nas tabelas 20 a 24.

TABELA 20 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DO PONTO DE FULGOR NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR BRUTOS

Modelo	VL^a	Int^b	RMSECV (°C)	R	RMSEC (°C)	R	RMSEP (°C)	R
Full	5	TODOS	4,07	0,8730	2,18	0,9643	3,53	0,9186
i2	5	1	4,31	0,8595	2,14	0,9656	3,47	0,9157
i4	4	4	3,72	0,8938	2,67	0,9451	2,93	0,9435
i8	5	8	3,70	0,8960	2,26	0,9615	2,25	0,9623
i16	6	15	1,83	0,9756	1,00	0,9926	1,30	0,9919
i32	6	30	1,85	0,9748	0,97	0,9930	2,26	0,9768

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 21 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DO PONTO DE FULGOR NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR CENTRADOS NA MÉDIA

Modelo	VL^a	Int^b	RMSECV (°C)	R	RMSEC (°C)	R	RMSEP (°C)	R
Full	4	TODOS	4,01	0,8761	2,21	0,9632	3,54	0,9181
i2	5	2	3,69	0,8963	2,05	0,9686	2,91	0,9415
i4	8	3	2,13	0,9664	0,89	0,9941	1,94	0,9751
i8	5	7	3,61	0,9017	1,72	0,9779	1,88	0,9783
i16	6	15	2,45	0,9559	0,91	0,9939	1,31	0,9905
i32	8	26	0,92	0,9737	0,46	0,9984	1,37	0,9869

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 22 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DO PONTO DE FULGOR NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR AUTOESCALADOS

Modelo	VL^a	Int^b	RMSECV (°C)	R	RMSEC (°C)	R	RMSEP (°C)	R
Full	2	TODOS	4,56	0,8344	3,61	0,8981	5,06	0,7904
i2	2	1	4,40	0,8463	3,67	0,8955	5,66	0,7295
i4	3	4	4,53	0,8376	3,61	0,8987	4,83	0,8211
i8	4	7	3,61	0,9029	2,23	0,9625	1,91	0,9747
i16	3	13	3,40	0,9115	2,79	0,9410	2,92	0,9487
i32	7	26	1,12	0,9910	0,60	0,9974	1,38	0,9864

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 23 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DO PONTO DE FULGOR NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR CENTRADOS NA MÉDIA E CORREÇÃO DO ESPALHAMENTO DE LUZ (MSC)

Modelo	VL^a	Int^b	RMSECV (°C)	R	RMSEC (°C)	R	RMSEP (°C)	R
Full	3	TODOS	4,03	0,8732	3,01	0,9310	4,43	0,8539
i2	2	2	3,63	0,8980	3,12	0,9256	4,46	0,8477
i4	2	4	3,40	0,9112	2,83	0,9391	3,54	0,9156
i8	3	8	3,70	0,8953	2,55	0,9510	2,27	0,9624
i16	4	15	2,35	0,9586	1,84	0,9747	1,44	0,9879
i32	3	26	1,89	0,9735	1,42	0,9851	2,51	0,9603

^a VL = variáveis latentes

^b Int = intervalo selecionado

TABELA 24 – VALORES DAS FIGURAS DE MÉRITO PARA A DETERMINAÇÃO DO PONTO DE FULGOR NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR AUTOESCALADOS E CORREÇÃO DO ESPALHAMENTO DE LUZ (MSC)

Modelo	VL ^a	Int ^b	RMSECV (°C)	R	RMSEC (°C)	R	RMSEP (°C)	R
Full	2	TODOS	4,34	0,8508	3,68	0,8947	5,24	0,7829
i2	3	2	4,00	0,8756	3,01	0,9310	4,35	0,8613
i4	2	4	4,25	0,8569	3,58	0,9005	4,12	0,8775
i8	3	8	4,17	0,8651	2,73	0,9437	3,71	0,8972
i16	3	13	3,63	0,9012	2,39	0,9571	3,05	0,9414
i32	4	26	1,62	0,9805	1,14	0,9904	1,95	0,9751

^a VL = variáveis latentes

^b Int = intervalo selecionado

Com relação aos dados autoescalados convém destacar que este utilizou 7 variáveis latentes enquanto que o outro modelo empregou a mesma região espectral (intervalo 26) com os dados centrados na média e correção do espalhamento de luz e apenas 3 variáveis latentes. Cabe salientar que apesar do primeiro ter utilizado mais variáveis latentes se mostrou mais robusto apresentando RMSEC 0,60 °C e RMSEP 1,38 °C (Tabela 22) comparados ao segundo com RMSEC 1,42 °C e RMSEP 2,51 °C (Tabela 23).

Independente do pré-processamento utilizado o intervalo 26 foi o selecionado, correspondendo a faixa espectral 1274 - 1376 cm⁻¹, que está relacionado ao estiramento assimétrico C-C(O)- dos ésteres do biodiesel (BARBOSA, 2007).

A Figura 30 mostra o espectro dividido em 32 partes, com seus respectivos RMSECV, e dados autoescalados.

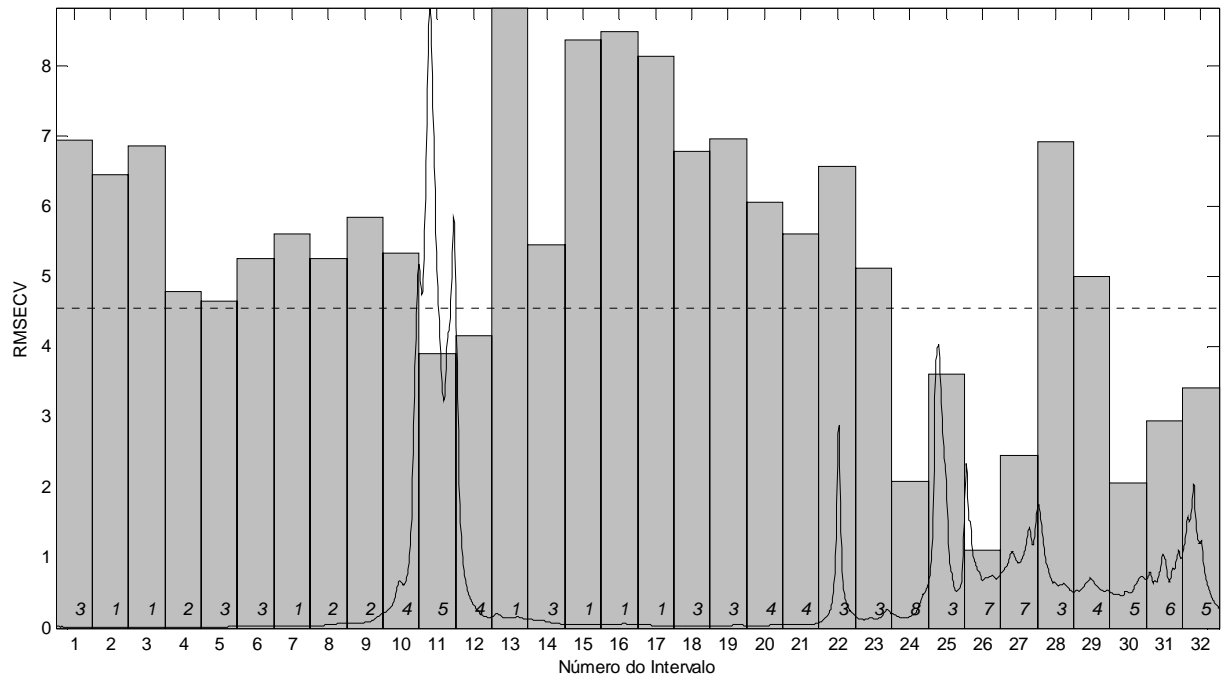


Figura 30 – Gráfico do RMSECV para os modelos dos 32 intervalos (barras) e modelo global empregando 7 variáveis latentes (linha pontilhada) para a determinação do ponto de fulgor nas blendas biodiesel/diesel, com dados autoescalados.

Fonte: Elaborada pelo autor

A curva de correlação entre os valores de referência e os valores previstos pelo modelo iPLS ATR-FTIR pode ser observada na figura 31.

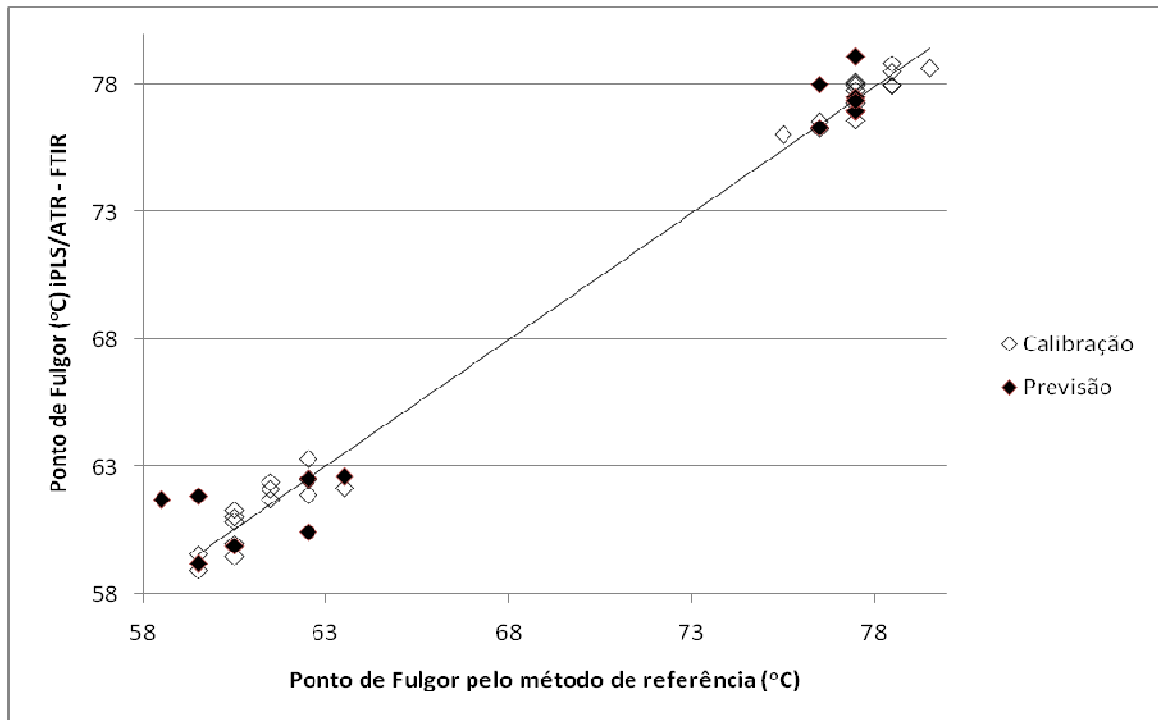


Figura 31 – Gráfico da correlação entre o método de referência e a previsão por iPLS/ATR-FTIR, empregando a faixa espectral 1274 - 1376 cm^{-1} (intervalo 26) e 7 variáveis latentes para a determinação do ponto de fulgor nas blendas biodiesel/diesel, com dados autoescalados.

Fonte: Elaborada pelo autor

A Figura 32 mostra o espectro dividido em 32 partes com os dados centrados na média com correção do espalhamento de luz, para o intervalo 26.

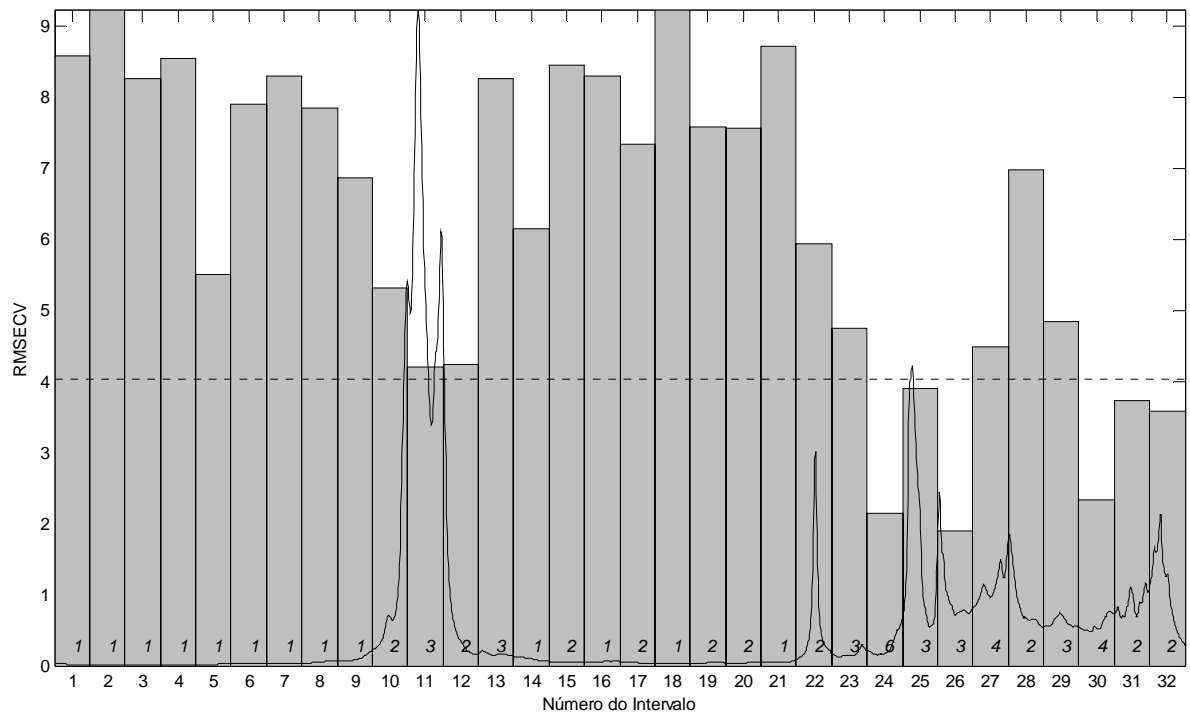


Figura 32 – Gráfico do RMSECV para os modelos dos 32 intervalos (barras) e modelo global empregando 3 variáveis latentes (linha pontilhada) para a determinação do ponto de fulgor nas blendas biodiesel/diesel.

Fonte: Elaborada pelo autor

Na Figura 33 é apresentada a curva de correlação entre os valores de referência e os valores previstos pelo modelo iPLS ATR-FTIR, com os dados centrados na média e correção do espalhamento de luz, para o intervalo 26.

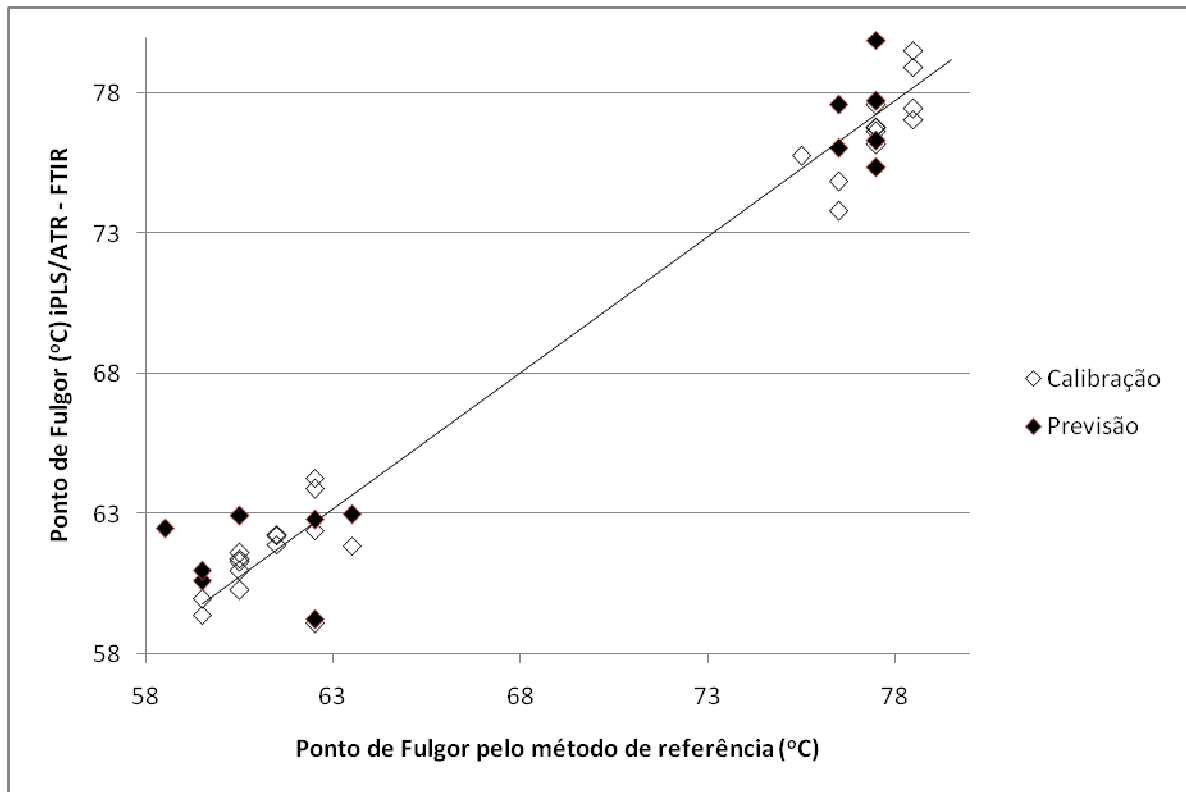


Figura 33 – Gráfico da correlação entre o método de referência e a previsão por iPLS/ATR-FTIR, empregando a faixa espectral 1274 - 1376 cm^{-1} (intervalo 26) e 3 variáveis latentes para a determinação do ponto de fulgor nas blends biodiesel/diesel, com dados centrados na média e correção do espalhamento de luz.

Fonte: Elaborada pelo autor

Com relação a Tabela 25 que apresenta os dados para os modelos siPLS, podemos observar que, de maneira análoga ao iPLS, os dados auto escalados apresentaram melhores resultados, com a seleção de três intervalos combinados (26, 27 e 30) e com utilização, também, de 7 variáveis latentes com RMSEC 0,55 °C e RMSEP 1,28 °C, semelhantes aos encontrados no modelo iPLS selecionado.

Uma possível justificativa para o número maior de variáveis latentes, para modelar este parâmetro físico-químico, seria pela complexidade maior deste em relação aos outros, com efeito que depende de mais variáveis e, ainda, interações entre elas.

TABELA 25 – VALORES PARA OS ERROS DE VALIDAÇÃO CRUZADA, CALIBRAÇÃO E PREVISÃO PARA OS MODELOS SIPLS, VISANDO A DETERMINAÇÃO DO PONTO DE FULGOR NAS BLENDS BIODIESEL/DIESEL EMPREGANDO OS DADOS ATR – FTIR

	Nº de intervalos	VL ^a	Int ^b	RMSECV (°C)	RMSEC (°C)	RMSEP (°C)
N	2	7	21 30	1,54	0,70	2,73
	3	9	19 26 30	1,05	0,46	1,22
	4	9	19 20 26 30	1,05	0,45	1,23
M	2	8	15 26	0,90	0,34	1,30
	3	8	14 15 26	0,82	0,33	1,33
	4	9	14 15 26 30	0,82	0,30	1,27
A	2	7	26 30	1,14	0,58	1,25
	3	7	26 27 30	1,06	0,55	1,28
	4	8	26 27 28 30	1,15	0,57	1,34
MM	2	4	12 30	1,89	1,08	1,62
	3	6	1 12 30	1,46	0,37	1,75
	4	7	11 21 22 28	1,56	0,80	1,37
MA	2	4	11 27	1,79	1,35	2,70
	3	4	11 22 24	1,68	1,25	2,17
	4	9	11 22 24 26	1,03	0,38	1,72

^a VL = variáveis latentes

^b Int = intervalos selecionados

A Figura 34 mostra as faixas espectrais selecionadas (intervalos 26, 27 e 30) para o modelo com dados auto escalados.

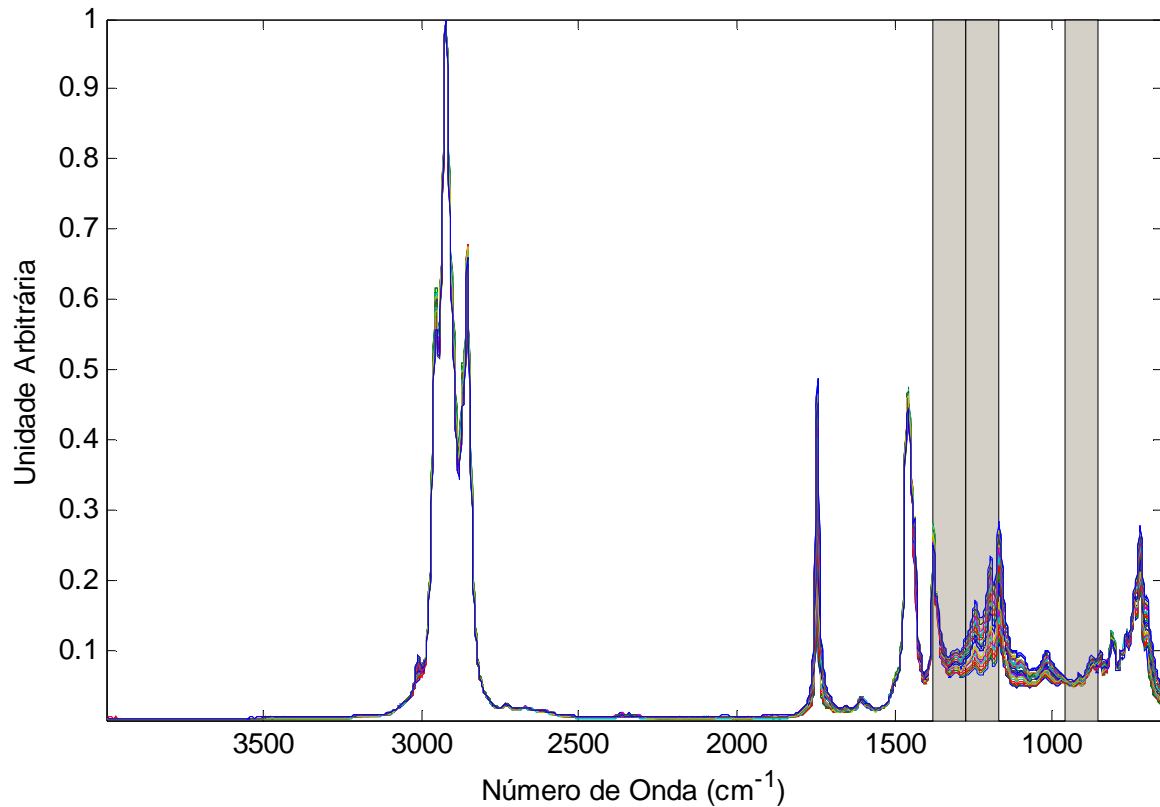


Figura 34 – Gráfico das faixas espectrais utilizadas na construção do modelo siPLS para a determinação do ponto de fulgor nas blends biodiesel/diesel, com dados autoescalados.

Fonte: Elaborada pelo autor

Na Figura 35 é apresentada a curva de correlação entre os valores de referência e os valores previstos pelo modelo siPLS ATR-FTIR.

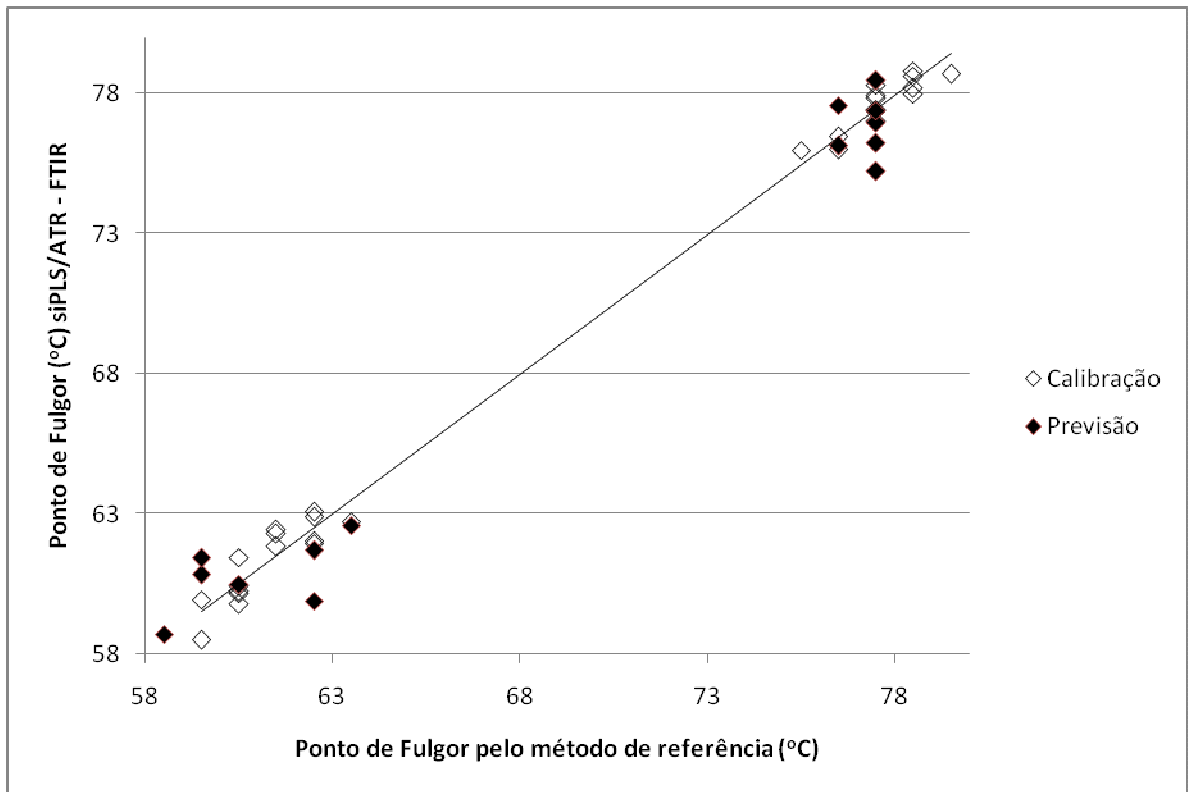


Figura 35 – Gráfico da correlação entre o método de referência e a previsão por siPLS/ATR-FTIR, empregando as faixas espectrais $858 - 960 \text{ cm}^{-1}$, $1170 - 1376 \text{ cm}^{-1}$ (intervalos 26, 27 e 30) com 7 variáveis latentes para a determinação do ponto de fulgor nas blendas biodiesel/diesel, com os dados autoescalados.

Fonte: Elaborada pelo autor

Para o modelo siPLS, visando a determinação do ponto de fulgor das blendas, além faixa espectral $1274 - 1376 \text{ cm}^{-1}$ selecionada para a construção do modelo iPLS que corresponde ao estiramento assimétrico C-C(O)- dos ésteres do biodiesel, foi combinada a faixa espectral $858 - 960 \text{ cm}^{-1}$ que corresponde a deformação angular simétrica fora do plano das ligações =C-H de olefinas (BARBOSA, 2007).

CONCLUSÕES

Com relação ao desenvolvimento do acessório para transreflectância, os resultados obtidos neste estudo não foram satisfatórios, porém a técnica pode ser melhorada na medida em que exista controle na espessura do filme da blenda biodiesel/diesel que é colocada na placa, merecendo maiores estudos para que seja atingido este objetivo.

Com relação a determinação do enxofre, empregando os espectros obtidos por Reflexão Total Atenuada, um modelo com boa capacidade preditiva foi obtido através do iPLS, porém com o siPLS houve uma melhora na capacidade preditiva, sendo o modelo que emprega os dados centrados na média com correção do espalhamento de luz e que combinou as faixas espectrais de $858 - 960 \text{ cm}^{-1}$ e $2732 - 2834 \text{ cm}^{-1}$, foi o que apresentou melhor desempenho para esta propriedade.

Já, para a determinação da porcentagem de biodiesel, o modelo escolhido foi o que utilizou o algoritmo iPLS, que além de utilizar apenas uma variável latente alcançou um coeficiente de correlação próximo de 1,00. Tal modelo iPLS utilizou os dados centrados na média com correção do espalhamento de luz, dividindo o espectro em 8 partes e escolhendo o intervalo 7, que representa a faixa espectral de $1068 - 1484 \text{ cm}^{-1}$ e corresponde ao estiramento assimétrico C-C(O)- dos ésteres do biodiesel.

Para a determinação da massa específica o modelo siPLS retornou menores valores de RMSE de calibração, previsão e validação cruzada de 0,118, 0,358 e 0,257, respectivamente, utilizando os dados auto escalados, porém necessitou de nove variáveis latentes. Já o iPLS atingiu erros de RMSEC e RMSEP de $0,674 \text{ kg/m}^3$ e $0,865 \text{ kg/m}^3$, respectivamente, utilizou os dados centrados na média com correção do espalhamento de luz e foi julgado o melhor por utilizar menor número de variáveis latentes (somente três). A faixa espectral selecionada, para este último foi a de $1274 - 1376 \text{ cm}^{-1}$.

E, finalmente, para determinação do ponto de fulgor, obteve-se resultados semelhantes para os modelos iPLS e siPLS, ambos com os dados autoescalados e utilizando sete variáveis latentes. Pelo fato do modelo iPLS ser de mais simples

construção este foi o escolhido, utilizando o intervalo 26, correspondendo a faixa espectral 1274 - 1376 cm^{-1} .

Todos os modelos que utilizaram o espectro inteiro mostraram capacidades de predição inferiores aos que foram submetidos a procedimentos de seleção de variáveis iPLS ou siPLS.

Com exceção do ponto de fulgor, que utilizou os dados auto escalados, o restante das propriedades utilizou os dados centrados na média com correção de espalhamento de luz, sendo, de forma geral, esta combinação, a mais adequada para a modelagem dos espectros no infravermelho com transformada de Fourier.

REFERÊNCIAS

- ABDI, H. Partial Least Square Regression PLS-Regression. In: Encyclopedia of Measurement and Statistics. Neil Salkind, 2007.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR14065*: Destilados de petróleo e óleos viscosos - Determinação da massa específica e da densidade relativa pelo densímetro digital. Rio de Janeiro, 2006.
- _____. *NBR14598*: Determinação do ponto de fulgor pelo aparelho de vaso fechado Pensky-Martens. Rio de Janeiro, 2007.
- ASTM D93-10A. Standard Test Methods for Flash Point by Pensky-Martens Closed Cup Tester, 18p.
- ASTM D5453-9. Standard Test Method for Determination of Total Sulfur in Light Hydrocarbons, Spark Ignition Engine Fuel, Diesel Engine Fuel, and Engine Oil by Ultraviolet Fluorescence, 10p.
- BARBOSA, Luis Cláudio de Almeida. *Espectroscopia no Infravermelho na caracterização de Compostos Orgânicos*. Editora da UFV: Viçosa, 2007.
- BORIN, A.; POPPI, R. J. Application of Mid infrared spectroscopy and iPLS for the quantification of contaminants in lubricating oil. *Vibrational Spectroscopy*, v. 37, p.27-32, 2005.
- BRASIL. Resolução Nº 42, de 24 de novembro de 2004. Agência Nacional do Petróleo (ANP). Disponível em: <http://nxt.anp.gov.br/nxt/gateway.dll/leg/resolucoes_anp/2008/mar%C3%A7o/ranp%207%20-%202008.xml>. Acesso em: 22 dez. 2008.
- COSTA FILHO, P. A.; POPPI, R. J. Aplicação de algoritmos genéticos na seleção de variáveis em espectroscopia no infravermelho médio. Determinação simultânea de glicose, maltose e frutose. *Quim. Nova*, v. 25, n. 1, p.46-52, 2002.
- DI, W.; Yong, H.; Pengcheng, N.; Fang, C.; Yidan, B. Hybrid variable selection in visible and near-infrared spectral analysis for non-invasive quality determination of grape juice. *Analytica Chimica Acta*, v. 659 p. 229–237, 2010.
- DU, Y.P.; LIANG, Y.Z.; JIANG, J.H.; BERRY, R.J.; OZAKI, Y. Spectral regions selection to improve prediction ability of PLS models by changeable size moving window partial least squares and searching combination moving window partial least squares. *Analytica Chimica Acta*, v. 501, p.183-191, 2004.

FAHRENFORT, J. Attenued total reflection. A new principle for the production of useful infrared reflection spectra of organic compounds. *Spectrochim. Acta*, v. 17, p. 698-709, 1961.

FERRÃO, Marco Flôres. *Aplicação de técnicas espectroscópicas de reflexão no infravermelho no controle de qualidade de farinha de trigo*. 2000. Tese . (Programa de Pós-Graduação em Química – Mestrado e Doutorado) – Universidade Estadual de Campinas – UNICAMP, Campinas, 2000.

FORTES, I. C. P.; CASTRO, E. O. V. R.; SILVA, R. C.; REZENDE, T.F.; SOARES, I. P. Multivariate Calibration by Variable Selection for Blends of Raw Soybean Oil/Biodiesel from Different Sources Using Fourier Transform infrared Spectroscopy (FTIR) Spectra Data. *Energy & Fuels*, v.22, p. 2079-2083, 2008.

GEORGE, Bill; MCINTYRE, Petter. *Infrared Spectroscopy*. John Wiley & Sons, 1994.

GUEDES, H. R.; AZEVEDO, R. G.; FERRÃO, M. F.; CROSSETTI, G. L; DAVANZO, C. U. Optimization regression moldels using biPLS for determination of protein content in commercial wheat flour. In: XIII INTERNATIONAL CONFERENCE ON INDUSTRIAL ENGINEERING AND OPERATIONS MANAGEMENT (ICIEOM) – Foz do Iguaçu, 2007.

HARRICK, N. J. Surface Chemistry from Spectral Analysis of Totally Internally Reflected Radiation. *JPhys. Chem.*, v. 64, p.1110-1114, 1960.

HELPER, G. A.; FERRÃO, M. F.; FERREIRA, C. V.; HERMES, N. Aplicação de métodos de análise multivariada no controle qualitativo de essências alimentícias empregando espectroscopia no infravermelho médio. *Ciênc. Tecnol. Aliment.*, v. 26(4), p.779-786, 2006.

INFOMETRIX Inc. *Pirouette User Guide*. Version 4.0, Bothell, WA, 2008.

International Energy Agency. Disponível em: <<http://www.iea.org/Textbase/stats/graphsearch.asp>>. Acesso em: 16 ago. 2008.

LIRA, L. F. B; ALBUQUERQUE, M. S.; PACHECO, J. G. A.; FONSECA, T. M.; CAVALCANTI, E. H. S.; STRAGEVITCH, L; PIMENTEL, M. F. Infrared spectroscopy and multivariate calibration to monitor stability quality parameters of biodiesel. *Microchem. J.*, p.1-6, 2010.

MARQUES, M. V.; SILVA, C. F. G.; NACIUK, F. F.; FONTOURA, L. A. M. A química, os Processos de Obtenção e as Especificações do Biodiesel. *Revista Analytica*, v. 33 p. 72-87, 2008.

MIRABELLA JR., F. M. Internal Reflection Spectroscopy. *Appl. Spectroscopy*, v. 21, p. 45-178, 1985.

MULLER, V.; FERRÃO, M. F.; FURTADO, J. C.; COSTA, J. M.; CORBELLINI, V. A.; SCROFERNEKER, M. L. Aplicação do biPLS na previsão da atividade lipase de

agentes da cromoblastomicose através de espectroscopia no infravermelho. In: XXVI ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO (ENEGERP) - Fortaleza, 2006.

NASCIMENTO, M.G.; COSTA NETO, P.R.; MAZZUCO, L.M. Biotransformação de óleos e gorduras. *Biocombustíveis: Ciência & Desenvolvimento*, v. 19, p. 28-31, 2001.

NØRGAARD, L. iPLS Toolbox Manual. Disponível em: <www.models.kvl.dk>. Acesso em: 13 set. 2009.

NØRGAARD, L.; *iToolbox Manual. Matlab Version 7.11.0*, 2004.

NØRGAARD, L.; SAUDLAND, A.; WAGNER, J.; NIELSEN, J. P.; MUNCK, L.; ENGELSEN, S. B. Interval partial least-squares regression (i-PLS): A comparative chemometric study with an example from near infrared spectroscopy. *Applied Spectroscopy*, v. 54, p.413–419, 2000.

OLINGER, J. M.; GRIFFITHS, P. R. Quantitative effects of an absorbing matrix on near infrared diffuse reflectance spectra. *Analytical Chemistry*, v. 60, p. 2427-2435, 1988.

PAVIA, Donald L.; LAMPMAN, Gary M.; KRIZ, George S. *Introduction to Spectroscopy*. Harcourt Brace College Publishers, 1996.

PEREIRA, A. F. C.; PONTES, M. J. C.; NETO, F. F. G.; SANTOS, S. R. B.; GALVÃO, R. K. H.; ARAÚJO, M. C. U. NIR spectrometric determination of quality parameters in vegetable oils using iPLS and variable selection. *Food Research International*, v. 41, p.341–348, 2008.

POPPI, R. J.; OLIVEIRA, I. K. DE; ROCHA, W. F. C. Application of near infrared spectroscopy and multivariate control charts for monitoring biodiesel blends. Elsevier, 2008.

SKOOG, D. A.; HOLLER, F. J.; NIEMAN, T. A. *Princípios de Análise Instrumental*. 5 ed. Bookman, 2002.

SILVA, F. E. B.; FERRÃO, M. F.; PARIZOTTO, G.; MÜLLER, E. I.; FLORES, E. M. M. Simultaneous determination of sulphamethoxazole and trimethoprim in powder mixtures by attenuated total reflection-Fourier transform infrared and multivariate calibration. *Journal of Pharmaceutical and Biomedical Analysis*. v. 49 p. 800–805, 2009.

WOLD, S.; SJÖSTRÖM, M. Chemometrics, present and future success. *Chemometrics and Intelligent Laboratory Systems*, v. 44 p. 3-14, 1998.

WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, v. 58 p.109-130, 2001.

XIAOBO Z.; JIEWEN Z.; POVEY, M.J.W.; HOLMES, M., HANPIN, M. Variables selection methods in Near Infrared spectroscopy. *Analytica Chimica Acta*, v. 667 p. 14-32, 2010.

YE, S.; WANG, D.; MIN,S. Successive projections algorithm combined with uninformative variable elimination for spectral variable selection. *Chemometrics and Intelligent Laboratory Systems*, v. 91 p.194–199, 2008.