

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

INSTITUTO DE QUÍMICA

CURSO DE GRADUAÇÃO EM QUÍMICA INDUSTRIAL

LUCAS GUSTAVO DA COSTA

**QUANTIFICAÇÃO DOS TEORES DE BIODIESEL DE CRAMBE E
MAFURRA EM MISTURAS COM DIESEL UTILIZANDO ESPECTROSCOPIA
MIR E FERRAMENTAS QUIMIOMÉTRICAS**

UBERLÂNDIA

2019

LUCAS GUSTAVO DA COSTA

**QUANTIFICAÇÃO DOS TEORES DE BIODIESEL DE CRAMBE E MAFURRA
EM MISTURAS COM DIESEL UTILIZANDO ESPECTROSCOPIA MIR E
FERRAMENTAS QUIMIOMÉTRICAS**

Trabalho de Conclusão de Curso de graduação
apresentado ao Instituto de Química, da
Universidade Federal de Uberlândia, como
requisito parcial para obtenção do grau de
Bacharel em Química Industrial.

Orientador: Prof. Dr. Waldomiro Borges Neto

UBERLÂNDIA

2019

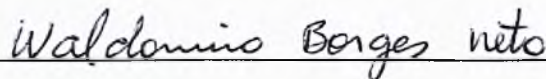
LUCAS GUSTAVO DA COSTA

**QUANTIFICAÇÃO DOS TEORES DE BIODIESEL DE CRAMBE E MAFURRA
EM MISTURAS COM DIESEL UTILIZANDO ESPECTROSCOPIA MIR E
FERRAMENTAS QUIMIOMÉTRICAS**

Trabalho de Conclusão de Curso de graduação
apresentado ao Instituto de Química, da
Universidade Federal de Uberlândia, como
requisito parcial para obtenção do grau de
Bacharel em Química Industrial.


Uberlândia, 04 de julho de 2019.

Banca examinadora:

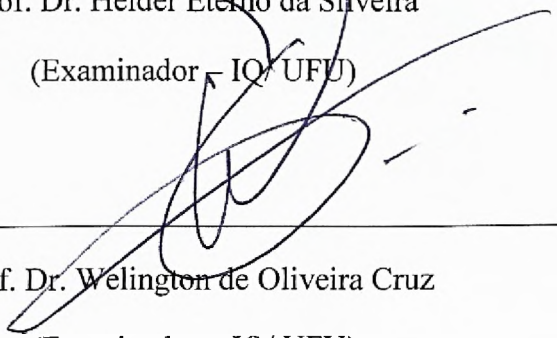


Prof. Dr. Waldomiro Borges Neto

(Orientador – IQ/ UFU)


Prof. Dr. Helder Eterno da Silveira

(Examinador – IQ/ UFU)


Prof. Dr. Wellington de Oliveira Cruz

(Examinador – IQ/ UFU)

AGRADECIMENTOS

Meus agradecimentos à Universidade Federal de Uberlândia e ao Instituto de Química por fornecer toda a estrutura física e tecnológica para o desenvolvimento deste trabalho.

Ao meu orientador, Professor Dr. Waldomiro Borges Neto por me ajudar a encontrar meu lugar dentro da ciência Química.

Aos meus pais pelo apoio em todas as minhas escolhas, respeito e confiança. Agradeço especialmente a minha mãe, por sempre acreditar e me fazer acreditar que sou capaz.

Aos integrantes e parceiros do Laboratório de Quimiometria do Triângulo pelo conhecimento e aprendizado gerados.

Ao Laboratório Multiusuário pela disposição de um dos espectrômetros usados neste trabalho.

À FAPEMIG pelo fomento através da bolsa de iniciação científica.

Agradeço por fim, a todas as pessoas que contribuíram, seja de maneira direta ou não, para a realização deste trabalho e em minha caminhada até aqui.

“[...] It matters not how strait the gate,
How charged with punishments the scroll
I am the master of my fate:
I am the captain of my soul.”

William Ernest Henley

RESUMO

A espectroscopia de absorção no infravermelho médio junto à regressão por quadrados mínimos parciais é o método oficial para monitorar a quantidade de biodiesel em diesel comercial brasileiro. Este método, no entanto, utiliza solventes e um grande número de amostras para a construção das curvas de calibração, o que gera resíduos e aumenta o tempo necessário para a análise. Neste estudo, a fim de desenvolver um método não destrutivo, sendo possível recuperar a amostra depois de sua quantificação, diminuir a quantidade de amostras e fazer uso de uma única curva de calibração, utilizou-se as oleaginosas crambe e mafurra, que não são usadas na alimentação humana, para a produção de biodieseis metílicos e da análise por PLS para determinar os seus teores em misturas com diesel na faixa de concentração de 1,00 a 30,00 (% v/v). O modelo global para o biodiesel metílico de crambe obteve $RMSEC = 0,26$ (% v/v), $RMSECV = 0,35$ (% v/v) e $RMSEP = 0,41$ (% v/v), e o modelo global para o biodiesel metílico de mafurra obteve $RMSEC = 0,47$ (% v/v), $RMSECV = 0,83$ (% v/v) e $RMSEP = 0,69$ (% v/v). Complementarmente, os métodos de seleção de variáveis iPLS, biPLS e siPLS foram empregados nos modelos globais com o intuito de diminuir a faixa espectral necessária para a construção das regressões e melhorar os valores de RMSEP, RMSECV e RMSEC.

Palavras chave: Biocombustíveis; Calibração Multivariada; Seleção de Intervalos; iPLS; biPLS; siPLS.

ABSTRACT

Mid-Infrared absorption spectroscopy associated with the least squares regression is the official method to monitor Brazilian commercial diesel quality. This method, however, uses solvents and a large number of samples for the construction of the calibration curves, which generates waste and increases the time needed for the analysis. In order to develop a non-destructive method, being possible to recover the sample after its quantification, decrease the quantity of samples and make use of a single calibration curve, in this study we used the oilseeds crambe and mafurra, which are not used in human food, for the methyl biodiesel production and PLS analysis for their content determination in mixtures with diesel at a concentration range of 1.00 to 30.00 (% v/v). The global model for crambe methyl biodiesel obtained $RMSEC = 0.26$ (% v/v), $RMSECV = 0.35$ (% v/v) and $RMSEP = 0.41$ (% v/v), and the global model for mafurra methyl biodiesel obtained $RMSEC = 0.47$ (% v/v), $RMSECV = 0.83$ (% v/v) and $RMSEP = 0.69$ (% v/v). Complementary, variables selection methods iPLS, biPLS and siPLS were applied in the global models in order to reduce the spectral range required to regressions construction and to improve the $RMSEP$, $RMSECV$ and $RMSEC$ values.

Keywords: Biofuels; Multivariate Calibration; Selection of Intervals; iPLS; biPLS; siPLS.

LISTA DE FIGURAS

Figura 1 - Estrutura química dos principais ácidos graxos encontrados em óleos vegetais	14
Figura 2 - Representação da reação de transesterificação catalisada por base forte (KOH).....	15
Figura 3 - Representação da reação de hidroesterificação	16
Figura 4 - (a) Plantação de crambe e (b) destaque dos frutos	17
Figura 5 - (a) Árvore da mafurra e (b) destaque de seus frutos.....	18
Figura 6 - Gráfico sobre o volume de óleo diesel comercializado no Brasil em 2018 e primeiro trimestre de 2019.....	19
Figura 7 - Percentual de não conformidades encontradas em óleo diesel comercial brasileiro em abril de 2019	20
Figura 8 - Representação simplificada da análise por ATR.....	22
Figura 9 - Espectros MIR de óleo diesel e biodiesel metílico de crambe puros	23
Figura 10 - Representação das Componentes Principais.....	25
Figura 11 - Representação do vetor NAS, onde: E é o espaço interferente, 1 é o vetor Interferente, 2 é a soma dos espectros das amostras, 3 corresponde ao vetor NAS, 4 é vetor Resíduos e 5 é a soma entre os vetores NAS e Resíduos.....	28
Figura 12 - Espectros de absorbância da mistura biodiesel metílico de crambe-diesel depois de corrigidos	35
Figura 13 - Espectros de absorbância da mistura biodiesel metílico de mafurra-diesel depois de corrigidos	36
Figura 14 - Erros do modelo PLS em relação ao número de VL's.....	37
Figura 15 - Modelo global de regressão por PLS.....	37
Figura 16 - Gráfico iPLS40	39
Figura 17 - Regressão obtida após o uso do iPLS	39
Figura 18 - Faixas selecionadas por siPLS.....	41
Figura 19 - Gráfico para detecção de amostras anômalas	42
Figura 20 - Modelo gerado após o uso do siPLS.....	43
Figura 21 - Erros do modelo PLS em relação ao número de VL's.....	44
Figura 22 - Modelo PLS global.....	45
Figura 23 - Gráfico iPLS5	46
Figura 24 - Regressão obtida após o uso do iPLS	47
Figura 25 - Faixas usadas nos modelos gerados por siPLS.....	49
Figura 26 - Modelo obtido após o uso do siPLS	50
Figura 27 - ERJC para o modelo global da mistura biodiesel metílico de crambe-diesel, onde: (●) representa a posição idealizada do modelo e (*) é a posição real.....	51
Figura 28 - ERJC para o modelo iPLS40 da mistura biodiesel metílico de crambe-diesel, onde: (●) representa a posição idealizada do modelo e (*) é a posição real	53

Figura 29 - ERJC para o modelo global da mistura biodiesel metílico de mafurra-diesel, onde: (●) representa a posição idealizada do modelo e (*) é a posição real.....	54
Figura 30 - ERJC para o modelo siPLS4 da mistura biodiesel metílico de mafurra-diesel, onde: (●) representa a posição idealizada do modelo e (*) é a posição real	55

LISTA DE TABELAS

Tabela 1 - Resultados obtidos por iPLS	38
Tabela 2 - Resultados dos modelos biPLS	40
Tabela 3 - RMSECV para as combinações siPLS	41
Tabela 4 - Valores dos erros para os modelos gerados após uso do siPLS	43
Tabela 5 - Resultados obtidos por iPLS	45
Tabela 6 - Resultados dos modelos biPLS	47
Tabela 7 - RMSECV para as combinações siPLS	48
Tabela 8 - Valores dos erros para os modelos gerados após uso do siPLS	49
Tabela 9 - Figuras de mérito para o modelo global biodiesel metílico de crambe-diesel.....	51
Tabela 10 - Figuras de mérito do modelo iPLS40 de biodiesel metílico de crambe-diesel.....	52
Tabela 11 - Figuras de mérito do modelo global de biodiesel metílico de mafurra-diesel.....	53
Tabela 12 - Figuras de mérito do modelo siPLS4 de biodiesel metílico de mafurra-diesel.....	54

LISTA DE SIGLAS E ABREVIATURAS

ABNT – Associação Brasileira de Normas Técnicas

AGL – Ácidos Graxos Livres

ANP – Agência Nacional do Petróleo, Gás Natural e Biocombustíveis

ASTM – *American Society for Testing and Materials*

B10 – Combustível comercial composto por 90% de óleo diesel e 10% de biodiesel

biPLS – PLS por exclusão de intervalos

EJCR – (do inglês) Elipse Conjunta de Confiança

HATR – (do inglês) Acessório Horizontal de Reflectância Total Atenuada

iPLS – PLS por intervalos

iPLS40 – Análise por iPLS após a divisão do espectro original em 40 faixas

iPLS5 – Análise por iPLS após a divisão do espectro original em 5 faixas

LD – Limite de Detecção

LQ – Limite de Quantificação

MIR – (do inglês) Infravermelho Médio

NAS – (do inglês) vetor Sinal Analítico Líquido

NBR – Norma Técnica Brasileira

PC – (do inglês) vetor Componente Principal

PCA – (do inglês) Análise por Componentes Principais

PLS – (do inglês) Quadrados Mínimos Parciais

PLS1 – Quadrados Mínimos Parciais utilizando um vetor para a correlação com o sinal instrumental.

R – Coeficiente de correlação

RMSE – (do inglês) Erro Quadrático Médio

RMSEC – (do inglês) Erro Quadrático Médio de Calibração

RMSECV – (do inglês) Erro Quadrático Médio de Validação Cruzada

RMSEP – (do inglês) Erro Quadrático Médio de Previsão

SÊN – Sensibilidade

siPLS – PLS por sinergismo de intervalos

siPLS4 – Análise por siPLS para combinações de 4 faixas do espectro original

TAG – Triacilglicerídeos

VL – vetor Variável Latente

SUMÁRIO

1. INTRODUÇÃO	13
1.1. Biodiesel	13
1.2. O crambe.....	17
1.3. A mafurra.....	18
1.4. Mercado	18
1.5. Controle de Qualidade	20
1.6. Espectroscopia no infravermelho médio	21
1.6.1. Espectros MIR do óleo diesel e biodiesel.....	23
1.7. Quimiometria.....	24
1.8. Análise de Componentes Principais (PCA).....	24
1.9. Quadrados Mínimos Parciais (PLS)	25
1.10. Seleção de variáveis.....	26
1.11. Validação	27
1.12. Comparações entre modelos multivariados.....	31
2. OBJETIVOS	32
2.1. Geral	32
2.2. Específicos.....	32
3. MATERIAIS E MÉTODOS	33
4. RESULTADOS E DISCUSSÕES	35
4.1. Mistura biodiesel metílico de crambe em diesel	36
4.2. Mistura biodiesel metílico de mafurra em diesel.....	44
4.3. Outros parâmetros de Validação	50
5. CONCLUSÃO.....	56
6. REFERÊNCIAS	57

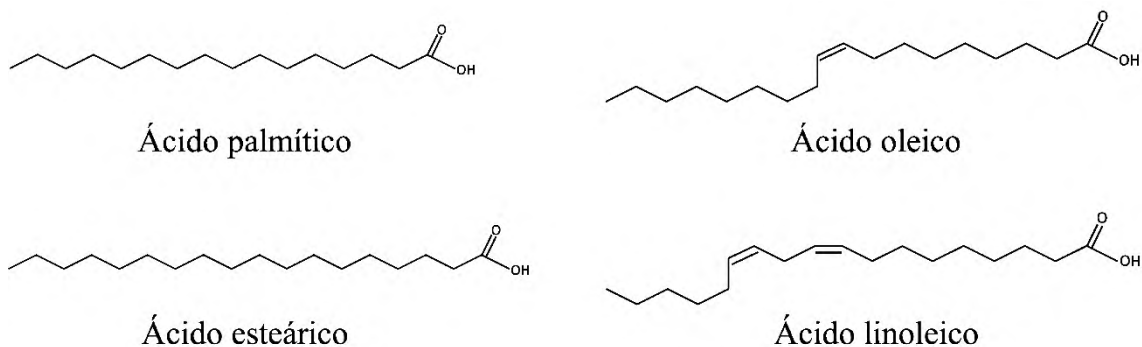
1. INTRODUÇÃO

O aumento exponencial da população mundial e os avanços tecnológicos demandam, cada vez mais, fontes de energia que sejam preferencialmente renováveis e que seus usos como combustíveis não colidam com outros de grande significância, como o uso na alimentação (ANDRADE; MENDONÇA; ROMEIRO, 2018). O desenvolvimento de processos sustentáveis que convertam biomassa em vários produtos com valor agregado contribui de maneira efetiva no aproveitamento de resíduos, geração de renda e redução dos impactos negativos ao meio ambiente (NAZARENO; VIEIRA; OLIVEIRA, 2015).

1.1. Biodiesel

Biodiesel é o combustível feito a partir de óleo vegetal, residual ou gordura animal, e álcool anidro na presença de um catalisador (LIMA et al., 2014). Pensado inicialmente por Diesel e Ford, o processo para sua obtenção somente foi implementado mais de 50 anos depois por Expedito José de Sá Parente (NAZARENO; VIEIRA; OLIVEIRA, 2015).

Os óleos e gorduras são compostos por carbono, hidrogênio e oxigênio que se assemelham estruturalmente quanto à presença de triacilglicerídeos (TAG) de ácidos graxos, mas que se diferenciam pelo tipo e disposição destes ácidos graxos quando combinados com o glicerol. As gorduras (sólidas em temperatura ambiente) são compostas principalmente por triacilglicerídeos de ácidos graxos saturados, enquanto os óleos (líquidos a temperatura ambiente) tem a maior parte formada por triacilglicerídeos de ácidos graxos insaturados. Cerca de 90% dos óleos vegetais produzidos no mundo são compostos de ácidos graxos de cadeia longa, principalmente os ácidos linoleico, esteárico, palmítico e oleico, representados na Figura 1 (SALVADOR et al., 2016).

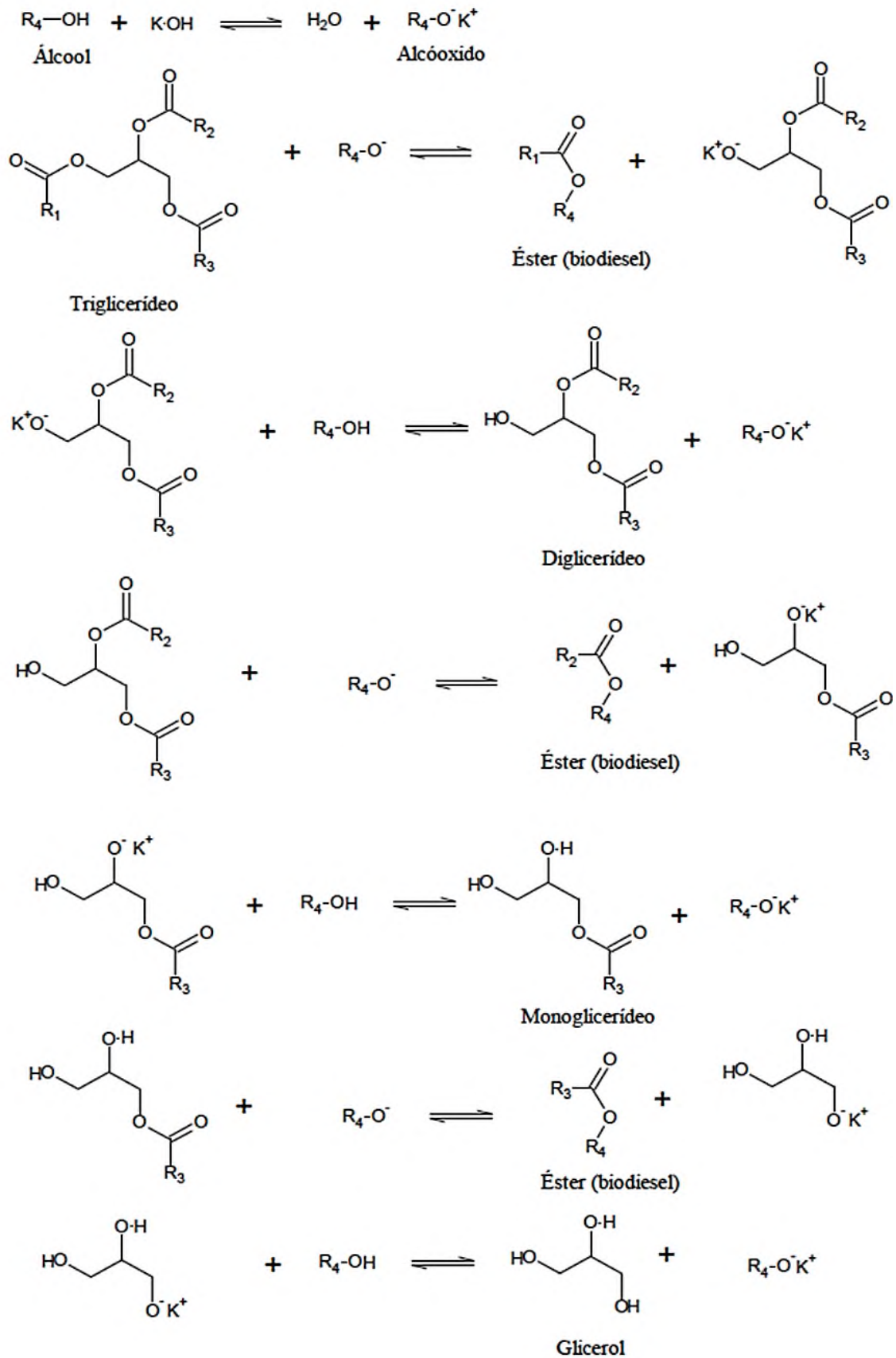
Figura 1 - Estrutura química dos principais ácidos graxos encontrados em óleos vegetais

Fonte: Adaptado de SOARES, 2016.

A transesterificação homogênea via catálise básica, representada na Figura 2, é o processo mais utilizado para a obtenção do biodiesel. Essa reação consiste na interação entre os TAG constituintes dos óleos vegetais/gorduras animais e um álcool de cadeia curta na presença de uma base de Brønsted (receptora de prótons) que atua como catalisador, gerando o éster (chamado de biodiesel) e glicerol. Inicialmente, a base reage com o álcool formando água e o alcóxido correspondente, o alcóxido então ataca o carbono da carbonila do TAG por substituição nucleofílica, gerando um intermediário tetraédrico e um monoéster. O intermediário desprotona o álcool ainda presente no meio reacional e forma um novo acilglicerídeo, que reage novamente com o alcóxido até que reste somente a molécula de glicerol. Como a reação é reversível, se utiliza um excesso de álcool para que o equilíbrio seja deslocado em direção à formação dos produtos (VIEIRA et al., 2017).

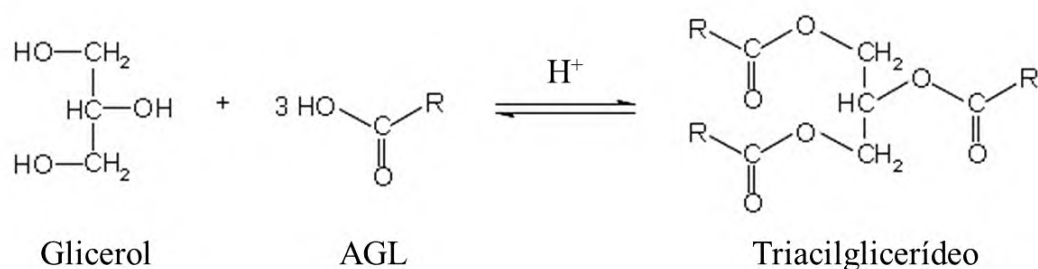
Com o emprego da catálise básica, o uso de reagentes com baixa quantidade de água se faz necessário para evitar a reação paralela de saponificação e o álcool utilizado na síntese pode ser recuperado ao fim do processo e reutilizado após a sua desidratação.

Figura 2 - Representação da reação de transesterificação catalisada por base forte (KOH)



Apesar de a transesterificação ser uma reação simples, muitas variáveis podem interferir no processo, como a relação molar dos reagentes e condição inicial do óleo utilizado. O teor elevado de ácidos graxos livres (AGL) contidos em determinados óleos vegetais pode influenciar qualitativa e quantitativamente na transesterificação, sendo que matérias-primas contendo acidez maior que 0,5% de AGL e umidade maior que 0,25% de água não são adequadas para a reação, fazendo-se necessário o emprego prévio de uma reação de esterificação (ou hidroesterificação), Figura 3, para reduzir o número de AGL presentes no óleo (VIEIRA et al., 2017).

Figura 3 - Representação da reação de hidroesterificação



Fonte: Adaptado de SILVA, 2017.

Dentre as vantagens dos biodieseis, pode-se destacar: i) a ausência de enxofre em sua composição, reduzindo em no mínimo, 15% da emissão de poluentes na atmosfera; ii) o fato de que o volume de gás carbônico liberado durante sua queima é praticamente o mesmo retirado do ar pelas plantações das oleaginosas usadas em sua síntese (LEBID; HENKES, 2015); iii) sua queima resulta em menor quantidade de fuligem, favorecendo uma combustão mais completa, logo, retirando substâncias poliaromáticas da atmosfera com grande potencial carcinogênico (SILVEIRA, 2015). Assim, podemos então enquadrar essa classe de combustíveis na chamada “química verde”, se mostrando uma importante ferramenta a curto e médio prazo, no controle do aquecimento global e outras questões ambientais.

1.2. O crambe

O crambe (*Crambe abyssinica Hochst*), Figura 4, é uma oleaginosa nativa do continente africano que apresenta grande adaptabilidade ao solo e clima brasileiro. É uma planta resistente à geadas e a seca, tem produtividade entre 1000 e 1500 quilos por hectare, sua cultura é considerada de ciclo curto (média de 90 dias), sendo uma boa opção para a safrinha por se tratar de uma cultura de ciclo outono/inverno. Suas sementes possuem cerca de 35% a 60% de óleo (JASPER et al., 2010; SANTOS et al., 2015).

Figura 4 - (a) Plantação de crambe e (b) destaque dos frutos



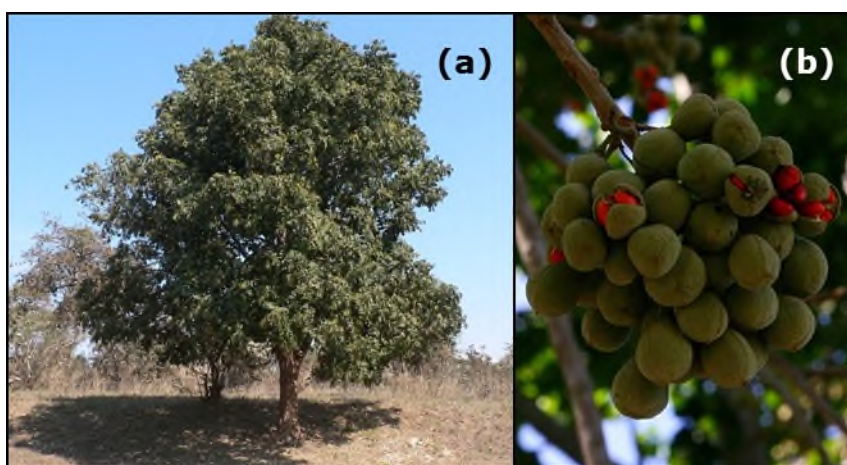
Fonte: FocoRural, 2016.

Para a extração do óleo, tradicionalmente é usada a prensagem e/ou extração com solventes orgânicos, resultando em um óleo com grande teor de ácidos graxos livres, principalmente o ácido úrico (cerca de 60%), que é altamente tóxico e inviabiliza o uso como alimento (LALAS et al., 2012; MELLO et al., 2017; TAVARES et al., 2017), mas com o tratamento adequado através da hidroesterificação ela se torna uma boa substituinte da soja na produção de biodiesel (SANTOS et al., 2015; SILVEIRA, 2015).

1.3. A mafurra

A *Trichilia emética*, popularmente conhecida como mafurra, Figura 5, é uma oleaginosa originária da África Ocidental e Arábia (com o nome de eucandja) que pode atingir cerca de 35 metros de altura e produzir anualmente por volta de 25 quilos de semente verde (MÁQUINA, 2017).

Figura 5 - (a) Árvore da mafurra e (b) destaque de seus frutos



Fonte: PlantZafrica, 2014.

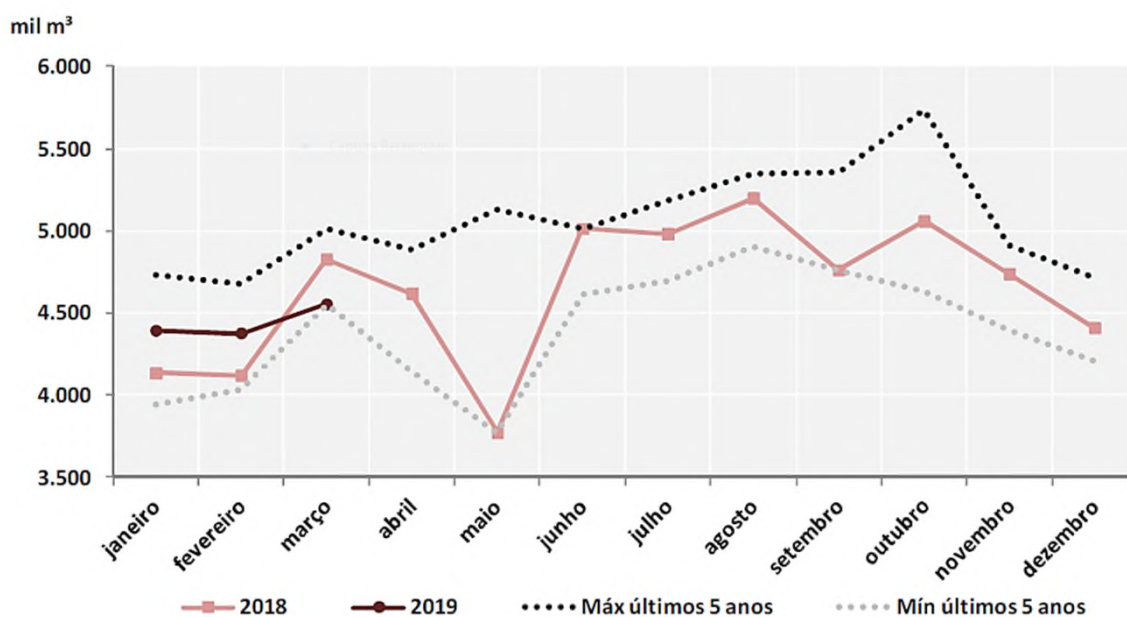
Além de resistente à seca e bem adaptada ao clima subtropical, suas sementes possuem um teor de óleo de 55% a 65%. O sabor amargo de seu óleo inviabiliza o uso na alimentação humana, o que o torna economicamente favorável para a produção de biodiesel (SIMÕES, 2014).

1.4. Mercado

Em 2018, a frota brasileira efetivamente em circulação era de 65,8 milhões de veículos. Destes, 2,0 milhões (3,09%) correspondiam a caminhões e 376,5 mil (0,57%)

a ônibus (EMPRESÔMETRO, 2019), veículos que utilizam prioritariamente o óleo diesel como combustível. No mesmo ano, o volume de diesel combustível comercializado foi de 55,558 bilhões de litros. Em 2019, o acumulado do combustível comercializado até março chega a 13,32 milhões de litros, tendo uma alta de 1,86% em relação ao mesmo período de 2018 como se vê na Figura 6 (ANP, 2019a, b).

Figura 6 - Gráfico sobre o volume de óleo diesel comercializado no Brasil em 2018 e primeiro trimestre de 2019



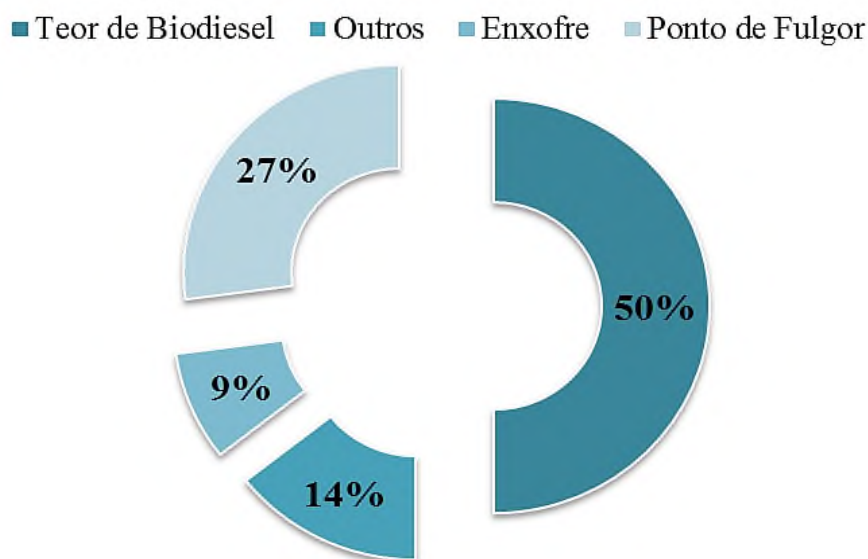
Fonte: ANP, 2019a.

Com a crescente população urbana, que implica em um maior sistema de transporte coletivo, e avanços em pesquisas relacionadas à produção agrícola, sendo esta produção escoada principalmente por estradas, é possível prever que a demanda por esta classe de combustível aumentará (pelo menos em médio prazo) cada vez mais, exigindo assim, o aprimoramento nas suas garantias de qualidade.

1.5. Controle de Qualidade

O diesel combustível vendido no território brasileiro é composto de biodiesel (principalmente feito a partir de soja e gordura animal) e diesel, na proporção de $10,00 \pm 0,50$ (% v/v) biodiesel-diesel (BRASIL, 2016), a essa mistura é dado o nome de B10. Segundo estudo publicado pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), Figura 7, 50% de não conformidades encontradas em B10 comercial estão relacionadas à quantidade de biodiesel na mistura (ANP, 2019b). Logo, se fazem necessários métodos de resposta rápida para o controle de qualidade desse combustível.

Figura 7 - Percentual de não conformidades encontradas em óleo diesel comercial brasileiro em abril de 2019



Fonte: Adaptado de ANP, 2019b.

O alto valor agregado dos biodieseis e a dificuldade de se verificar visualmente misturas biodiesel-diesel não conformes incentivam a adulteração criminosas, ou ainda, os meios de transporte utilizados na distribuição do combustível podem propiciar contaminações por outros produtos petrolíferos. As adulterações e contaminações nas

misturas alteram suas composições, podendo causar danos aos veículos (GUIMARÃES, 2014) e diminuem as vantagens ambientais esperadas.

A norma NBR-15568 é o método oficial empregado na determinação do teor de biodiesel em diesel no Brasil, utilizando a Espectroscopia no Infravermelho Médio e a calibração multivariada por Quadrados Mínimos Parciais (do inglês, PLS). Neste método, duas curvas de calibração são construídas, uma com 160 pontos (amostras), e outra com 45, nas faixas de concentração de 0,00 (% v/v) a 8,00 (% v/v) para a primeira e de 8,00 (% v/v) a 30,00 (% v/v) para a segunda. Há a necessidade de utilizar solventes como o iso-octano e o hexano na preparação das mostras e quando o modelo se mostra inadequado para as predições, com erros elevados, a adição de grupos de 20 amostras é necessária (ABNT, 2008). No entanto, pesquisadores da área acreditam que seja possível desenvolver modelos com a mesma finalidade utilizando uma única curva de calibração com ampla faixa de concentração, menor número de pontos e leitura direta da amostra.

1.6. Espectroscopia no infravermelho médio

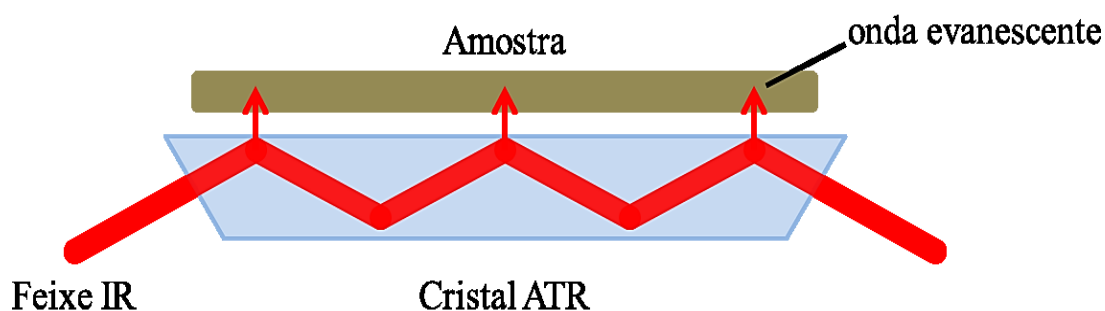
A espectroscopia estuda a interação da radiação eletromagnética com a matéria, sendo um de seus principais objetivos, o estudo dos níveis de energia de átomos ou molécula. Normalmente, as transições eletrônicas estão situadas na região do ultravioleta ou visível, as energias vibracionais na região do infravermelho e as rotacionais na região de microondas (HOLLER; SKOOG; CROUCH, 2009).

O espectrômetro no infravermelho com transformada de Fourier (do inglês, FT-IR) funciona através do método denominado interferômetro de Michelson, que causa um retardo óptico no feixe de radiação eletromagnética, aumentando a sensibilidade, resolução e reprodutibilidade de determinado comprimento de onda. Este instrumento permite que todos os comprimentos de onda da radiação incidam ao mesmo tempo na amostra, resultando em um interferograma que é convertido no espectro através da transformada de Fourier (HAGE; CARR, 2012). A técnica mais antiga para a

obtenção dos espectros no infravermelho é a de transmissão, onde a radiação passa pela amostra, tendo uma parte absorvida e a outra transmitida, podendo ser usada em amostras sólidas, líquidas e gasosas (BARBOSA, 2008). Outra técnica muito utilizada é a medida por refletância, onde a radiação é refletida na superfície da amostra, o tipo de reflexão depende da natureza da superfície da amostra, podendo ser especular, difusa ou refletância total interna (HAGE; CARR, 2012).

Dentre as técnicas de refletância, a Refletância Total Atenuada (do inglês, ATR), exemplificada na Figura 8, se baseia no fato de que mesmo ocorrendo reflexão total na interface de dois meios, a radiação penetra uma determinada distância no meio menos denso (onda evanescente) e pode ser parcialmente absorvida colocando-se uma amostra em contato com o meio mais denso. Esse elemento mais denso é denominado elemento de reflexão interna (BARBOSA, 2008).

Figura 8 - Representação simplificada da análise por ATR



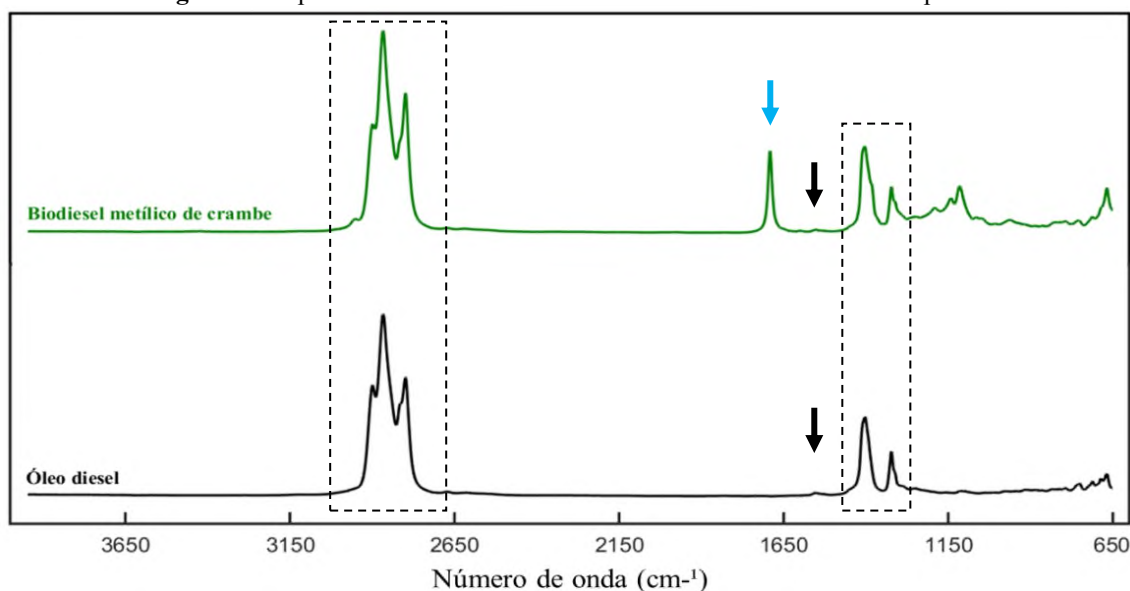
Fonte: Adaptado de BARBOSA, 2008.

A região compreendida no infravermelho médio (do inglês, MIR) abrange a faixa espectral de 4000 a 200 cm^{-1} e é muito utilizada para análises qualitativas e quantitativas, seus espectros apresentam bandas intensas e picos de absorbância bastante altos, o que facilita a interpretação (BARBOSA, 2008).

1.6.1. Espectros MIR do óleo diesel e biodiesel

Os espectros MIR de biodieseis sintetizados a partir de diferentes óleos têm maior diferenciação na região de impressão digital inferior a 1500 cm^{-1} enquanto que a diferença mais proeminente entre os biodieseis e o óleo diesel está na presença da banda de carbonila característica dos ésteres. A Figura 9 mostra os espectros do óleo diesel puro e do biodiesel metílico de crambe.

Figura 9 - Espectros MIR de óleo diesel e biodiesel metílico de crambe puros



Fonte: O Autor.

A presença de bandas na região de $1000\text{ a }1300\text{ cm}^{-1}$ é atribuída às vibrações de deformação axial C–O, as bandas por volta de 1380 cm^{-1} e 1450 cm^{-1} se referem, respectivamente, às deformações simétricas e assimétricas de CH_3 . A banda de baixa intensidade em 1600 cm^{-1} , pouco visível na escala da imagem (indicada pelas setas pretas) é atribuída às deformações axiais da ligação C=C, característica de alcenos. O pico de absorvância na faixa de $1700\text{ a }1750\text{ cm}^{-1}$ é proveniente da carbonila (C=O), e a região entre $2840\text{ e }3000\text{ cm}^{-1}$ se refere às vibrações de deformação axial das ligações C–H dos agrupamentos de metila e metileno (BARBOSA, 2008).

1.7. Quimiometria

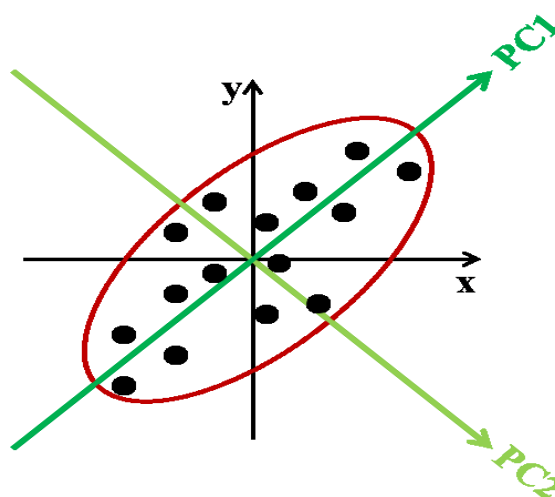
A quimiometria é a parte da química que se utiliza de ferramentas estatísticas, matemáticas e de lógica formal para obter o máximo de informações a partir da análise de dados químicos. Os avanços na instrumentação analítica, gerando dados com muitas variáveis e o grande uso dessas medidas necessita de métodos de estatística multivariada, álgebra matricial e análise numérica para converter a resposta instrumental no dado químico de interesse (BARROS NETO; SCARMINIO, 2006; BRUNS; FAIGLE, 1985), assim, a quimiometria se firma como uma importante parte da química analítica.

1.8. Análise de Componentes Principais (PCA)

Introduzida na química por Malinowski no final dos anos 60, a PCA tem por objetivo transformar dados complexos de tal maneira que a visualização das informações mais importantes seja facilitada (MASSART et al., 1998). Como os dados são projetados em um subespaço dimensional menor, essa análise também é considerada um método de projeção (HERBEGGER, 2007).

Na PCA a matriz de dados originais \mathbf{X} é reconstruída em 2 matrizes menores, a matriz de escores (\mathbf{T}) e a de pesos (\mathbf{P}). A parte não modelada corresponde a uma matriz de resíduos (\mathbf{E}). A diminuição de dimensões é obtida através de novas variáveis ortogonais entre si, as componentes principais (do inglês, \mathbf{PC} 's). As \mathbf{PC} 's são determinadas baseando-se no critério de variância máxima, então a primeira \mathbf{PC} captura mais informação do que a segunda, a segunda mais que a terceira e assim sucessivamente, Figura 10 (MASSART et al., 1998; HERBEGGER, 2007).

Figura 10 - Representação das Componentes Principais



Fonte: Adaptado de BORGES NETO, 2005.

O número de **PC's** usadas no modelo é escolhido pela porcentagem da variância explicada, logo, seleciona-se um número de **PC's** em que a maior quantia da informação dos dados originais seja capturada.

1.9. Quadrados Mínimos Parciais (PLS)

O PLS é o método mais usado para a construção de modelos de calibração multivariada por não ser necessário um conhecimento exato de todos os componentes presentes em uma amostra, realizando a previsão mesmo em presença de interferentes contanto que esses interferentes também estejam presentes na construção do modelo (BEEBE; KOWALSKI, 1987).

No modelo PLS1, uma matriz **X** (correspondente ao sinal instrumental das amostras) e um vetor **y** (correspondente às concentrações do analito de interesse) são incorporados e um processo iterativo aperfeiçoa ao mesmo tempo a projeção das amostras (**X**) sobre os pesos (**y**) determinando assim os escores, que são ajustados por uma função linear minimizando os desvios. Essa otimização simultânea provoca

pequenas mudanças nas direções dos pesos, os fazendo perder a ortogonalidade e levando a redundâncias que evidenciam a relação linear entre eles. Essas distorções das componentes principais no modelo PLS fazem que não sejam mais PC's e sim, variáveis latentes (VL's) que diferencia este modelo do modelo PCA (VANDEGINSTE et al., 1998).

Dois grupos de amostras são usados na construção do modelo: o grupo de calibração, que estabelece a relação entre a resposta instrumental e a propriedade de interesse, criando a equação matemática responsável pela regressão (ou seja, é o modelo em si) e o grupo de previsão, usado para testar o modelo. Para a determinação correta do número de VL's do modelo, se usa o método de validação cruzada, que consiste em usar amostras do grupo de calibração para testar a previsão do modelo antes do grupo de previsão (OTTO, 1999; BORGES NETO, 2005).

1.10. Seleção de variáveis

Algoritmos determinísticos, que localizam regiões no espectro com maior variância entre as amostras, como PLS por intervalos (iPLS), PLS por exclusão de intervalos (biPLS) e PLS por sinergismo de intervalos (siPLS), têm sido utilizados para a seleção das regiões espectrais com maior relevância de informação para a quantificação de determinado analito. Com a aplicação destes algoritmos é possível produzir modelos de regressão com melhor desempenho e menores erros de previsão (BORGES NETO, 2005; GONTIJO, 2016).

O iPLS divide o espectro em subintervalos de igual tamanho e calcula um modelo PLS para cada intervalo. Estes submodelos são comparados com o modelo global, geralmente pelo parâmetro de Erro Quadrático Médio de Validação Cruzada (do inglês, RMSECV), assim as regiões com parâmetros de validação menores que o modelo global são selecionadas para a construção de um novo modelo PLS (NØRGAARD et al., 2000).

No PLS por Exclusão de Intervalos, a região espectral também é dividida em intervalos equidistantes e a região que apresentar maior RMSECV é retirada, sendo calculado um novo modelo com os intervalos restantes, fornecendo nova estimativa RMSECV. Esse processo se repete até que reste apenas um intervalo e os resultados de saída do algoritmo são sumarizados em uma Tabela, chamada Tabela de Exclusão (LEARDI; NØRGAARD, 2014).

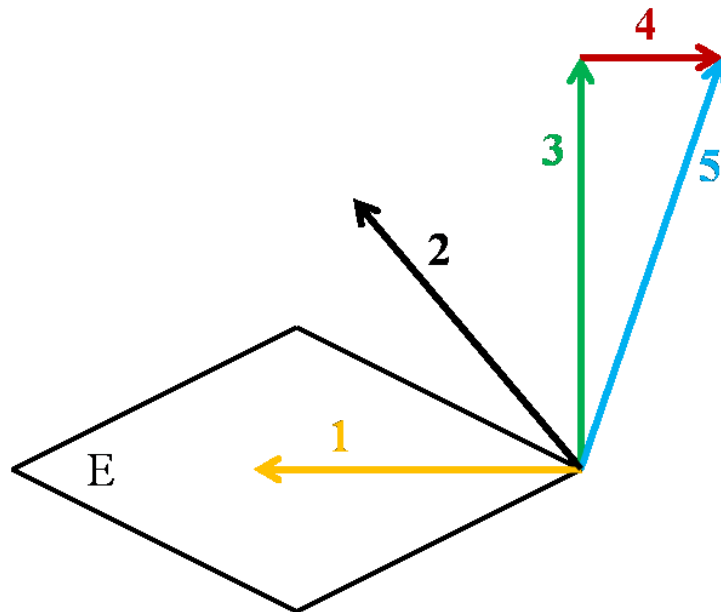
O siPLS divide a matriz espectral em faixas de mesmo tamanho e realiza a combinação de 2, 3 ou 4 intervalos até que se tenha a melhor combinação (que resulte em um menor RMSECV). Escolhida a combinação, um novo modelo PLS é calculado usando essa junção, resultando em um modelo com melhor capacidade preditiva, com menos variáveis e menores erros (MUNCK et al., 2001).

1.11. Validação

A validação é um processo de verificação que certifica se um método desenvolvido apresenta um desempenho adequado nas condições em que ele será aplicado. A validação se dá através da determinação de diversos parâmetros conhecidos como figuras de mérito que em modelos multivariados, recomenda-se serem calculadas através do Sinal Analítico Líquido (do inglês **NAS**) (BRAGA, 2004; VALDERRAMA, 2005; CARVALHO ROCHA, 2007).

O **NAS**, é a parte do sinal analítico ortogonal às contribuições de outros possíveis constituintes da amostra, Figura 11, e com seu uso é possível calcular um valor escalar livre de interferentes, assim, o modelo multivariado pode ser representado em uma forma pseudo-univariada (BRAGA, 2004).

Figura 11 - Representação do vetor NAS, onde: **E** é o espaço interferente, **1** é o vetor Interferente, **2** é a soma dos espectros das amostras, **3** corresponde ao vetor NAS, **4** é vetor Resíduos e **5** é a soma entre os vetores NAS e Resíduos



Fonte: Adaptado de MITSUTAKE, 2015.

Para o cálculo do NAS se usa a matriz de dados reconstruída com A variáveis latentes, $(\hat{\mathbf{X}}_A)$, determinando em seguida a matriz que contém a informação de todas as espécies presentes na amostra, exceto da espécie de interesse K , Equação 1, isso faz com que a matriz $\hat{\mathbf{X}}_{A,-K}$ fique livre de qualquer contribuição da espécie K . Logo, o vetor NAS é obtido conforme a Equação 2:

$$\hat{\mathbf{X}}_{A,-K} = [\mathbf{I} - \hat{\mathbf{y}}_{A,K} \hat{\mathbf{y}}_{A,K}^+] \hat{\mathbf{X}}_A \quad \text{Equação 1}$$

$$\hat{\mathbf{x}}_{A,K}^{\text{nas}} = [\mathbf{I} - \hat{\mathbf{X}}_{A,-K}^T (\hat{\mathbf{X}}_{A,-K}^T)^+] \hat{\mathbf{X}}_A \quad \text{Equação 2}$$

onde $\hat{\mathbf{X}}_{A,-K}$ é uma matriz com todas as informações exceto a espécie de interesse, \mathbf{I} é a matriz identidade, $\hat{\mathbf{y}}_{A,K}$ é o vetor de concentrações da espécie de interesse K estimado com A variáveis latentes, o índice “+” sobrescrito indica a pseudo-inversa do vetor, $\hat{\mathbf{X}}_A$, a matriz de dados reconstruída com A variáveis latentes através do algoritmo SIMPLS e $\hat{\mathbf{x}}_{A,K}^{\text{nas}}$ é o vetor NAS.

Dentre as diversas figuras de mérito existentes, as usadas neste trabalho são: Erros Quadráticos Médios, Sensibilidade, Sensibilidade Analítica, Limites de Detecção e Quantificação e a Elipse de Confiança, descritos em seguida.

O Erro Quadrático Médio (do inglês RMSE) fornece informações sobre o ajuste do modelo em relação aos dados de cada grupo, de calibração (do inglês RMSEC), validação cruzada (do inglês RMSECV) e previsão (do inglês RMSEP) (NØRGAARD et al., 2000), obtido através da Equação 3:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n_g} (y_p - y_r)^2}{n_g}} \quad \text{Equação 3}$$

em que y_p é o valor de concentração da espécie de interesse previsto, y_r o valor real de concentração da espécie de interesse presente na amostra e n_g o número de amostras presentes em cada grupo.

A Sensibilidade ($\hat{S}\hat{E}N$), por definição é a fração do sinal responsável pelo acréscimo de uma unidade de concentração à propriedade de interesse (VALDERRAMA, 2005) e pode ser calculada através da Equação 4:

$$\hat{S}\hat{E}N = \|\mathbf{s}_K^{\text{nas}}\| \quad \text{Equação 4}$$

onde $\mathbf{s}_K^{\text{nas}}$ é o vetor sensibilidade líquida, obtido através da Equação 5:

$$\mathbf{s}_K^{\text{nas}} = \left\| \frac{\hat{\mathbf{x}}_{A,K}^{\text{nas}}}{\mathbf{y}} \right\| \quad \text{Equação 5}$$

em que $\hat{\mathbf{x}}_{A,K}^{\text{nas}}$ é o vetor NAS calculado pela Equação 2 e \mathbf{y} é o vetor concentração.

A Sensibilidade Analítica, Equação 6, é a razão entre a Sensibilidade e o desvio padrão do sinal de referência (BRAGA, 2004),

$$\gamma = \frac{S\hat{E}N}{\|\delta x\|} \quad \text{Equação 6}$$

onde γ é a Sensibilidade Analítica, $S\hat{E}N$ é a Sensibilidade calculada pela Equação 4 e δx é o desvio padrão do sinal estimado pelo NAS para 15 espectros do sinal de referência.

Os Limites de Detecção e Quantificação, Equações 7 e 8, expressam as menores quantidades da espécie de interesse que pode ser detectada e quantificada, respectivamente (BRAGA, 2004; CARVALHO ROCHA, 2007).

$$LD = 3\delta x \frac{1}{S\hat{E}N} \quad \text{Equação 7}$$

$$LQ = 10\delta x \frac{1}{S\hat{E}N} \quad \text{Equação 8}$$

Em que δx é o desvio padrão do sinal de referência e $S\hat{E}N$ é obtido pela Equação 4.

A Elipse Conjunta de Confiança (do inglês EJCR) pode ser definida como o intervalo no qual se pode afirmar com certo grau de confiança ou probabilidade, a presença do valor verdadeiro da propriedade de interesse (BRAGA, 2004; VALDERRAMA, 2005), seu cálculo é feito através da Equação 9:

$$\emptyset_i = \pm t_{(1-\alpha)/2n_{GL}} \sqrt{V(PE)_i} \quad \text{Equação 9}$$

em que $t_{(1-\alpha)/2n_{GL}}$ é o parâmetro estatístico da distribuição t-student com probabilidade $(1 - \alpha)/2n_{GL}$ onde n_{GL} é o número efetivo de graus de liberdade, obtido pela Equação 10 e $V(PE)$ é uma estimativa de variância dos erros de previsão calculados pela Equação 11:

$$n_{GL} = n \left(1 - \sqrt{\frac{RMSEC}{RMSECV}} \right) \quad \text{Equação 10}$$

$$V(PE) = \left(1 + h_i + \frac{1}{n} \right) MSEC_p \quad \text{Equação 11}$$

onde RMSEC e RMSECV são calculados pela Equação 3 e MSEC_p é o Pseudo Erro Quadrático Médio de Calibração, obtido pela Equação 12:

$$MSEC_p = \sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{n_{GL}} \quad \text{Equação 12}$$

1.12. Comparações entre modelos multivariados

Modelos gerados por análises multivariadas podem ser comparados se utilizado dos valores de RMSEP gerados e a distribuição de Fisher–Snedecor (teste F) de acordo com a Equação 13:

$$F_{cal} = \left(\frac{RMSEP_1}{RMSEP_2} \right)^2 \quad \text{(Equação 13)}$$

onde $RMSEP_1$ é maior que $RMSEP_2$. Se o valor de F_{cal} for maior que o valor tabelado da distribuição de Fisher-Snedecor com graus de liberdade igual ao número de amostras de predição e um nível de significância de 5%, não há evidência estatística de que os valores seguem a distribuição normal então, o modelo com $RMSEP_2$ apresenta melhor exatidão (FILGUEIRAS; ALVES; POPPI, 2014).

2. OBJETIVOS

2.1. Geral

Desenvolver um método de quantificação dos teores de biodiesel metílico de crambe e mafurra em misturas com diesel usando espectroscopia MIR e calibração multivariada por PLS na faixa de concentração de 1,00 a 30,00 (% v/v).

2.2. Específicos

- 1 – Desenvolver um método para a quantificação de biodiesel metílico de crambe e de mafurra presentes em misturas com diesel que se diferencie e apresente vantagens em relação ao método oficial;
- 2 – Empregar os métodos de seleção de variáveis iPLS, biPLS e siPLS visando a redução de variáveis utilizadas nos modelos e a melhora em suas capacidades preditivas;
- 3 – Validar os modelos gerados.

3. MATERIAIS E MÉTODOS

Os óleos de crambe e de mafurra usados na síntese do biodiesel passaram por um processo de hidroesterificação na proporção de massas 1:1 (m/m) óleo/álcool metílico mais 0,5 (% m/m) de H₂SO₄ em relação ao volume de álcool, em fervura por 120 minutos sob refluxo.

Em seguida, se separou a fase óleo de cada uma das oleaginosas e a elas, foram adicionado álcool metílico na razão de 30% de suas massas mais 1 (% m/m) de hidróxido de potássio. As misturas foram deixadas em agitação por 60 minutos e depois transferidas para funis de separação para descanso de 24 horas.

Depois do período de descanso, as fases glicerol da catálise básica foram retiradas e os ésteres (biodieseis), lavados com três porções de água destilada a 80°C. Os ésteres foram levados a um evaporador rotativo por 60 minutos, a 80 rpm, pressão de 400 mmHg e 80°C.

As quantidades de reagentes, as condições experimentais e os procedimentos para as catálises ácida e básica seguidos para a síntese dos biodieseis neste trabalho, foram embasados em trabalhos anteriores do Laboratório de Quimiometria do Triângulo, da Universidade Federal de Uberlândia – MG, onde se aperfeiçoou os processos de obtenção de biodieseis através de métodos de otimização multivariada (PORTELA, 2011; 2015).

As amostras foram construídas usando diesel puro, cedido pela TRANSPETRO S/A, localizada em Uberlândia – MG e os biodieseis preparados no Laboratório de Quimiometria do Triângulo, na faixa de 1,00 a 30,00 (% v/v) biodiesel-diesel.

Para a mistura biodiesel metílico de crambe-diesel, foram preparadas 43 amostras para o conjunto denominado calibração e 25 para o conjunto de previsão, totalizando 68 amostras. Para a mistura de biodiesel metílico de mafurra, foram 45 amostras para a calibração e 30 amostras para a previsão, em um total de 75 amostras. A quantidade de amostras de ambos os biodieseis foi determinada seguindo o recomendado pela norma da *American Society for Testing and Materials*, ASTM E1655-05 (2012), que apresenta as Equações 13 e 14, como método de calcular a

quantidade mínima de amostras para a construção de uma curva de calibração usando o PLS, de acordo com a quantidade de **VL**'s que se deseja obter no modelo.

$$C_p = 6(k + 1) \quad \text{Equação 14}$$

$$C_v = 4k \quad \text{Equação 15}$$

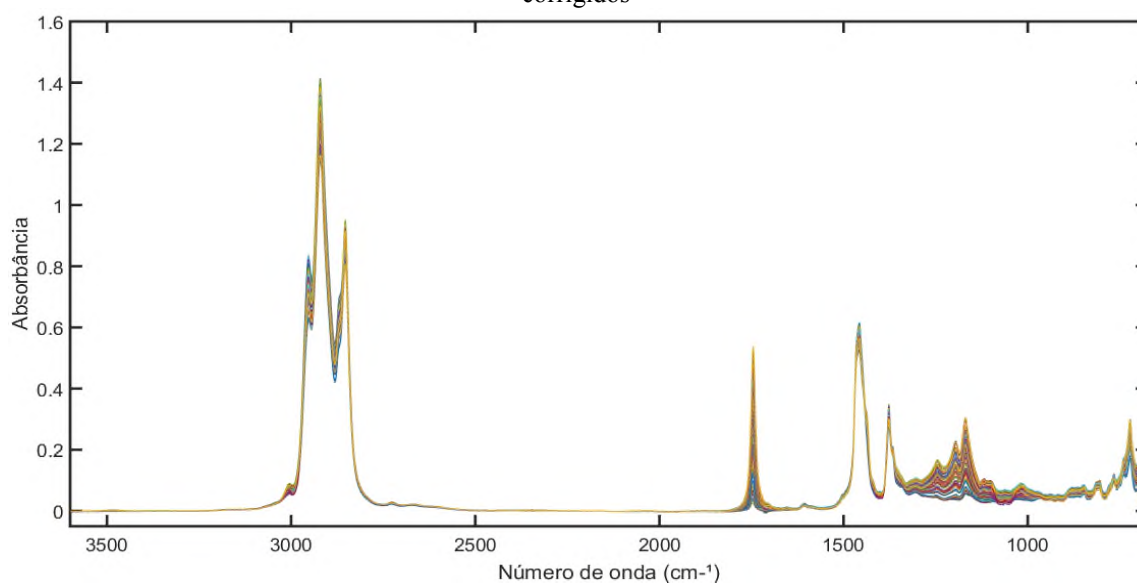
Onde C_p é a quantidade de amostras do conjunto de calibração, k é o número de **VL**'s que se deseja obter no modelo e C_v é a quantidade de amostras do conjunto validação. Quando $k \leq 3$, o grupo de calibração deverá ter no mínimo, 24 amostras.

4. RESULTADOS E DISCUSSÕES

Os espectros de absorvância dos dois conjuntos da mistura biodiesel de crambe-diesel foram obtidos em quintuplicata, na região de 4000 a 600 cm^{-1} , no equipamento PerkinElmer – Spectrum Two, utilizando o Acessório Horizontal de Refletância Total Atenuada (do inglês, HATR) com cristal de ZnSe, resolução de 4 cm^{-1} e 16 varreduras. Já os espectros de absorvância dos conjuntos da mistura biodiesel de mafurra-diesel foram obtidos em quintuplicata, na região de 4500 a 350 cm^{-1} , no equipamento Shimadzu – IRPrestige-21 pertencente ao Laboratório Multiusuário, do Instituto de Química, da Universidade Federal de Uberlândia, utilizando o HATR com cristal de ZnSe, resolução de 16 cm^{-1} e 32 varreduras.

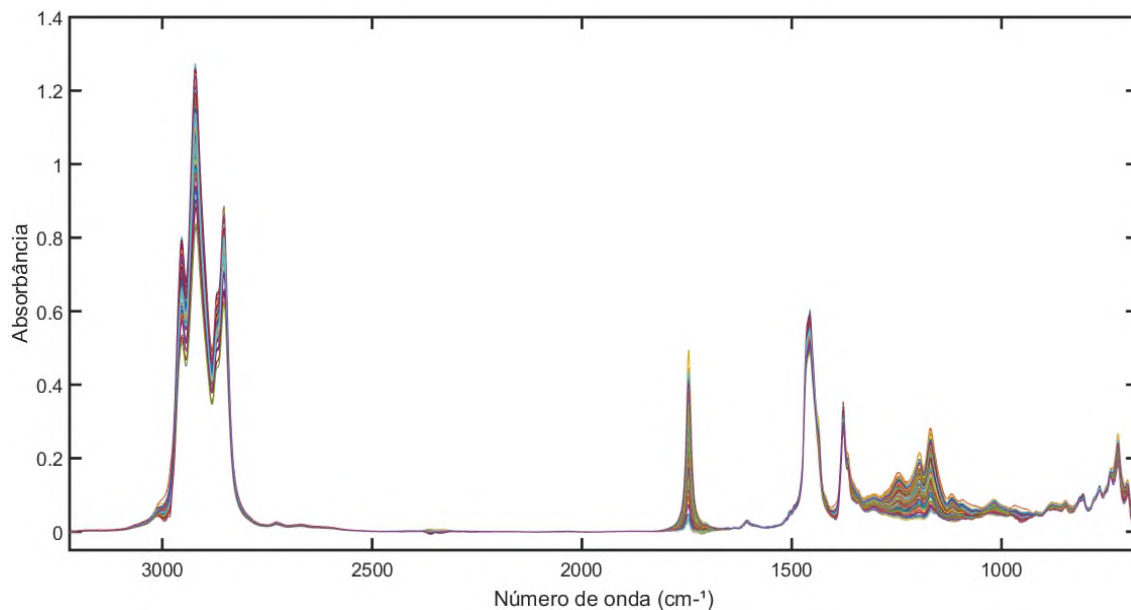
Os dados espectrais das misturas de ambos os biodieseis foram transferidos para ambiente MATLAB, versão R2015a, tiveram suas linhas de base corrigidas, calculada a média das replicatas de cada amostra e as matrizes finais usadas nas regressões foram obtidas depois de um corte nas regiões de interesse espectral ficando com os tamanhos de 68 amostras por 2925 variáveis e de 75 amostras por 2546 variáveis para crambe e mafurra respectivamente, conforme Figuras 12 e 13.

Figura 12 - Espectros de absorvância da mistura biodiesel metílico de crambe-diesel depois de corrigidos



Fonte: O Autor.

Figura 13 - Espectros de absorvância da mistura biodiesel metílico de mafurra-diesel depois de corrigidos



Fonte: O Autor.

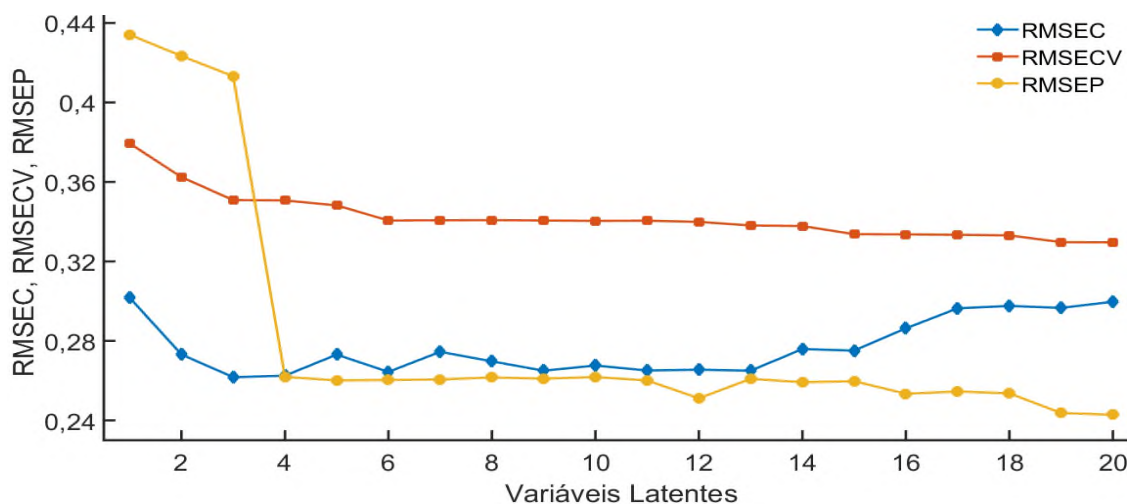
Os modelos globais foram construídos usando o PLS_Toolbox, versão 8.61, tendo como pré-processamento o *mean center* em ambos os blocos (X , amostras, e y , vetor concentração) e *vanetian blinds cross-validation* como método de validação cruzada (12 janelas com 2 amostras por janela).

Os modelos iPLS, biPLS e siPLS foram obtidos através do iPLS_Toolbox para 5, 10, 15, 20, 25, 30, 35 e 40 intervalos nos modelos iPLS, biPLS e siPLS para combinações de 2 e 3 regiões. Para os modelos siPLS com combinações de 4 regiões, os espectros foram divididos em 10, 15, 20, 25, 30, 35 e 40 intervalos. Todos os modelos foram posteriormente recalculados usando o PLS_Toolbox 8.

4.1. Mistura biodiesel metílico de crambe em diesel

A quantidade de variáveis latentes é escolhida de modo a obter o menor valor de RMSECV e a melhor variância explicada, portanto foram escolhidas 3 VL's para a construção do modelo.

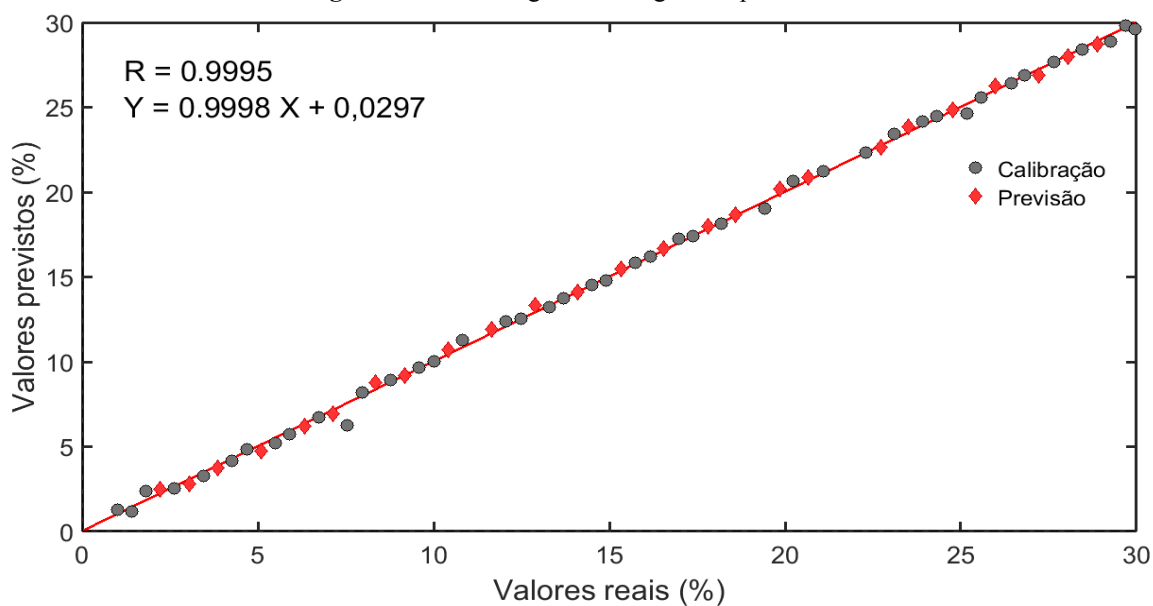
Figura 14 - Erros do modelo PLS em relação ao número de VL's



Fonte: O Autor.

Deste modo, o modelo global, contendo todas as variáveis, Figura 15, obteve um $RMSEC = 0,26$ (% v/v), $RMSECV = 0,35$ (% v/v) e $RMSEP = 0,41$ (% v/v). O valor de $RMSEC$ menor que o $RMSEP$, nos diz que o modelo não possui tendência em suas previsões. Através do coeficiente de correlação (R) é possível perceber ainda, uma boa relação entre os dados de referência e os calculados pelo modelo, já que quanto mais próximo de 1, melhor é a concordância entre os valores medidos e previstos.

Figura 15 - Modelo global de regressão por PLS



Fonte: O Autor.

Os modelos gerados por iPLS, Tabela 1, foram comparados ao modelo global através da estatística F, onde o modelo iPLS40 (negrito) obteve o melhor resultado, já que indicou um alto valor do F calculado.

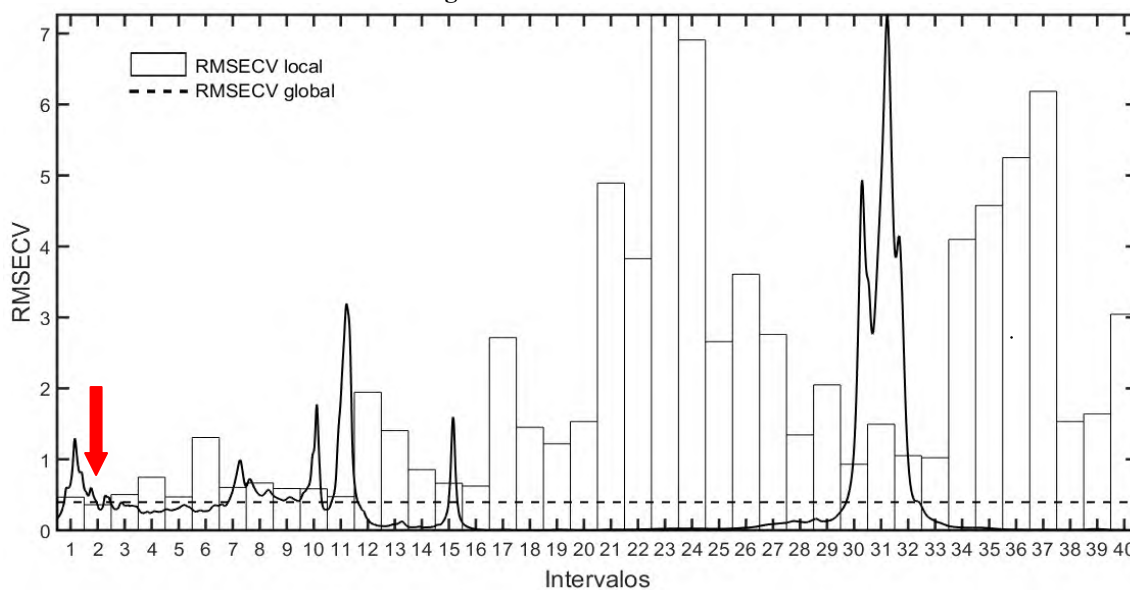
Tabela 1 - Resultados obtidos por iPLS

MODELO iPLS	VARIÁVEIS	FAIXA	VL's	RMSEP (% v/v)	TESTE F*
5	586	1:586	4	0,36	1,30
10	293	1:293	3	0,30	1,90
15	196	1:196	4	0,30	1,90
20	147	1:147	4	0,41	1,01
25	118	1:118	4	0,41	1,01
30	98	1:98	4	0,32	1,67
35	84	1:84	4	0,31	1,84
40	74	75:148	4	0,26	2,58
GLOBAL	2925	1:2925	4	0,41	-----

*Em relação ao modelo global. ($F_{(0,5;25;25)} = 1,95$)

Usando o intervalo 2 (variáveis de 75:148) que corresponde as vibrações relacionadas as deformações angulares de CH₂ de alcenos, como mostrado na Figura 16, foi calculado o novo modelo PLS usando como pré-processamento o *mean center* em ambos os blocos e *vanetian blinds cross-validation* como método de validação cruzada.

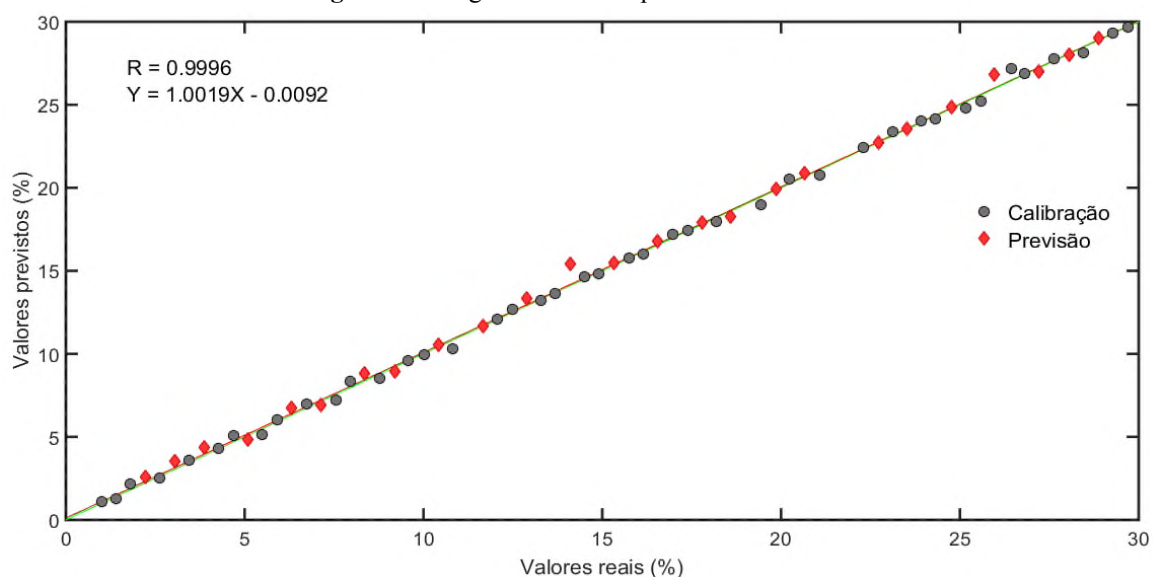
Figura 16 - Gráfico iPLS40



Fonte: O Autor.

A Figura 17 apresenta a regressão obtida após o uso de iPLS, obtendo 4 VL's, RMSEC = 0,23 (% v/v), RMSECV = 0,38 (% v/v), RMSEP = 0,24 (% v/v) e bons valores de coeficientes angular e linear.

Figura 17 - Regressão obtida após o uso do iPLS



Fonte: O Autor.

O biPLS, Tabela 2, foi executado dividindo a matriz **X** na mesma quantidade de faixas que o iPLS, porém não foi capaz de melhorar a exatidão do modelo global. Como os valores de F_{cal} ficaram inferiores ao valor de F tabelado, os modelos são estatisticamente semelhantes então, opta-se pelo modelo global por fazer uso de menos processos matemáticos para sua obtenção.

Tabela 2 - Resultados dos modelos biPLS

MODELO biPLS	VARIÁVEIS	FAIXA	RMSEP (% v/v)	TESTE F*
5	2341	1:586; 1172:2926	0,30	1,90
10	877	1:293; 1759:2050; 2343:2634	0,32	1,67
15	196	1:196	0,33	1,61
20	1465	1:735; 1175:1466; 1613:1904; 2051:2196	0,30	1,90
25	352	1:352	0,38	1,16
30	196	1:196	0,33	1,61
35	837	1:252; 673:756; 1093:1260; 1513:1596; 1931:2179	0,31	1,83
40	1392	1:296; 371:444; 518:590; 664:736; 956:1028; 1175:1320; 1686:2196; 2562:2634; 2854:2926	0,33	1,61
GLOBAL	2925	1:2925	0,41	-----

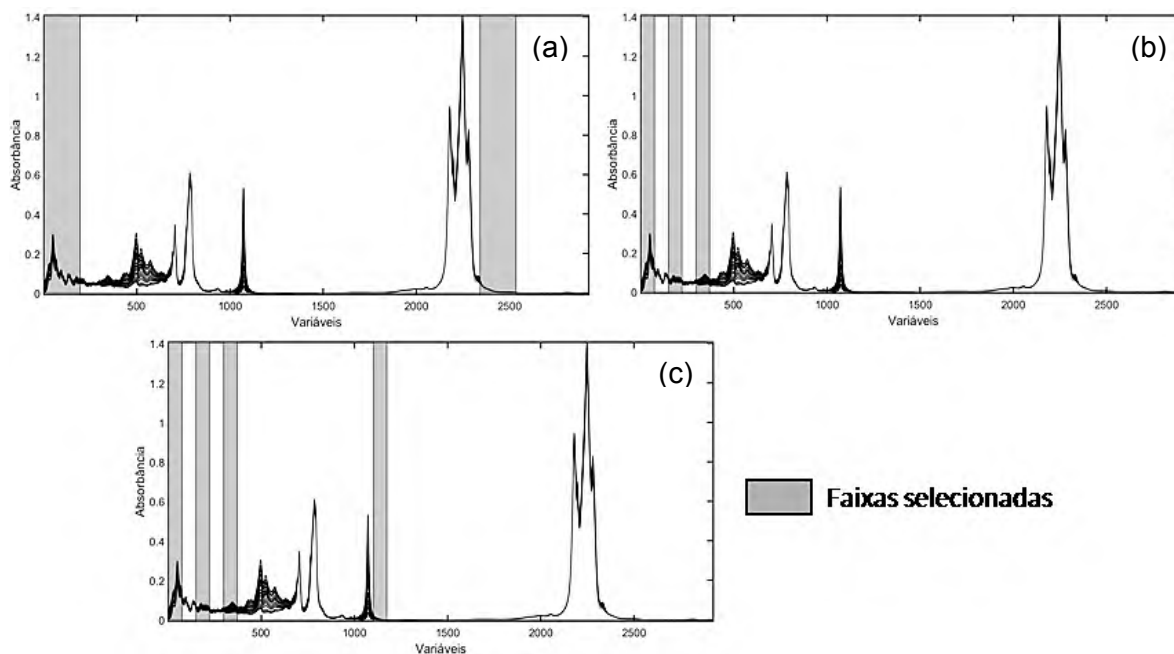
*Em relação ao modelo global. ($F_{(0,5;25;25)} = 1,95$)

No siPLS, foram simulados modelos com 2, 3 e 4 combinações em 5, 10, 15, 20, 25, 30, 35 e 40 divisões do espectro. Os melhores valores de RMSECV para as interações estão destacados em negrito na Tabela 3:

Tabela 3 - RMSECV para as combinações siPLS

Faixas	RMSECV (% v/v)		
	2	3	4
5	0,37	0,37	
10	0,29	0,29	0,28
15	0,27	0,27	0,27
20	0,28	0,28	0,28
25	0,28	0,26	0,26
30	0,28	0,27	0,27
35	0,28	0,25	0,24
40	0,29	0,23	0,23

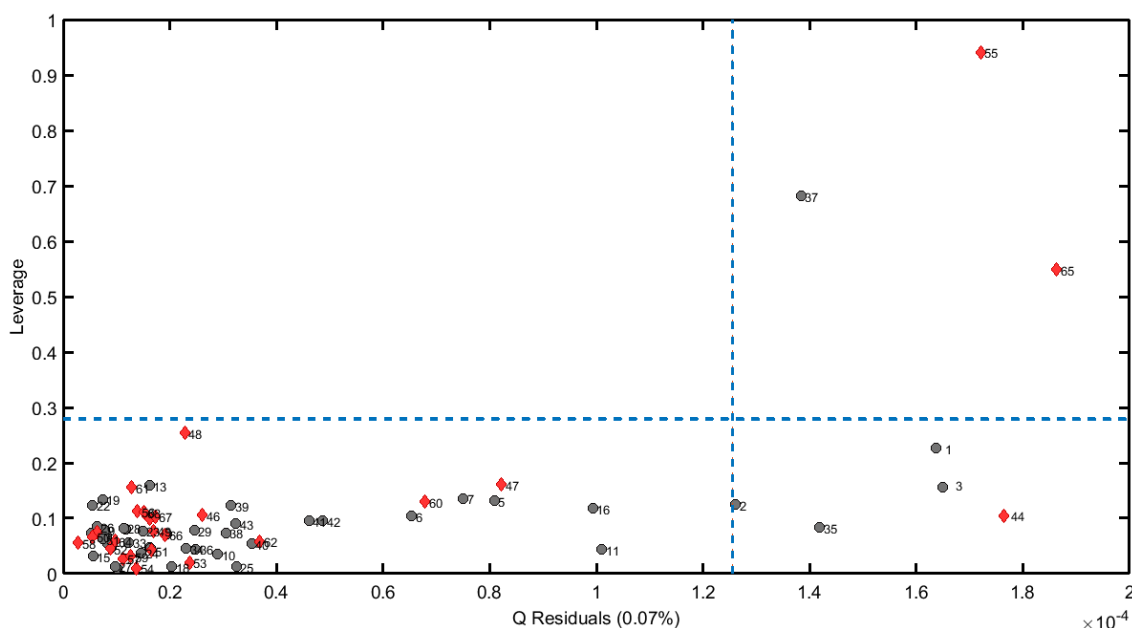
Três modelos, um para cada combinação de faixas mostradas na Figura 18 foram feitos, sendo que: (a) corresponde as faixas utilizadas no modelo com duas combinações, variáveis de 1:196 e 2342:2536; (b) é a combinação de faixas para o modelo de 3 combinações, variáveis de 1:74, 150:222, 297:370; (c) as faixas usadas no modelo de 4 combinações, variáveis de 1:74, 150:222, 297:370, 1102:1174.

Figura 18 - Faixas selecionadas por siPLS

Fonte: O Autor.

Para os modelos com 3 e 4 combinações foi detectado 3 amostras anômalas que aparecem no limite superior da linha tracejada horizontal e ao mesmo tempo, à direita da linha tracejada vertical no gráfico *Leverage versus Q Residuals*, Figura 19. Então, elas foram retiradas e uma amostra do conjunto de calibração foi repassada para o conjunto previsão, a fim de se manter a divisão das amostras em $\frac{2}{3}$ (dois terços) no conjunto de calibração e $\frac{1}{3}$ (um terço) no conjunto de previsão.

Figura 19 - Gráfico para detecção de amostras anômalas



Fonte: O Autor.

A Tabela 4 nos informa os valores de RMSEC, RMSECV, RMSEP e do teste F dos três modelos PLS obtidos, em todos foi usado como pré-processamento o *mean center* em ambos os blocos e *vanetian blinds cross-validation* como método de validação cruzada, obtendo 4 VL's.

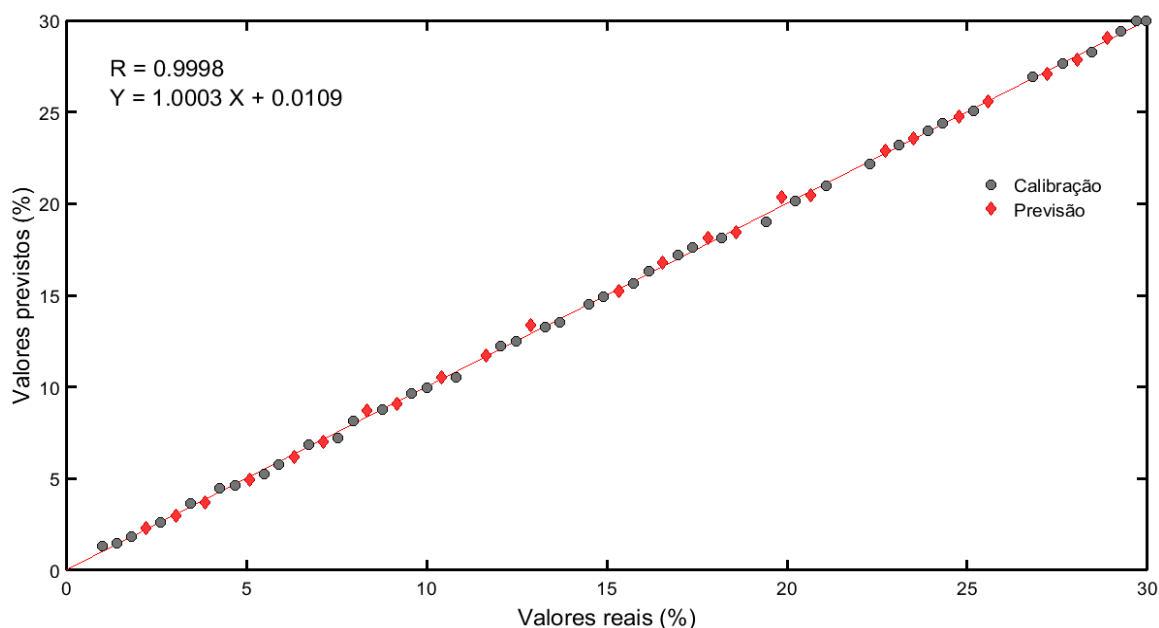
Tabela 4 - Valores dos erros para os modelos gerados após uso do siPLS

Combinações	RMSEC (% v/v)	RMSECV (% v/v)	RMSEP (% v/v)	TESTE F*
2	0,16	0,19	0,30	1,90**
3	0,18	0,24	0,32	1,67
4	0,17	0,22	0,21	3,78
GLOBAL	0,26	0,35	0,41	-----

*Em relação ao modelo global. ($F_{(0,5;25;24)} = 1,92$)

**Em relação ao modelo global. ($F_{(0,5;25;25)} = 1,95$)

Os modelos com 2 e 3 combinações foram descartados após o teste F por não apresentarem melhora na exatidão quando comparados ao modelo global, já que apresentaram valores de F menores que o F tabelado, sobrando então somente o modelo com 4 combinações mostrado na Figura 20:

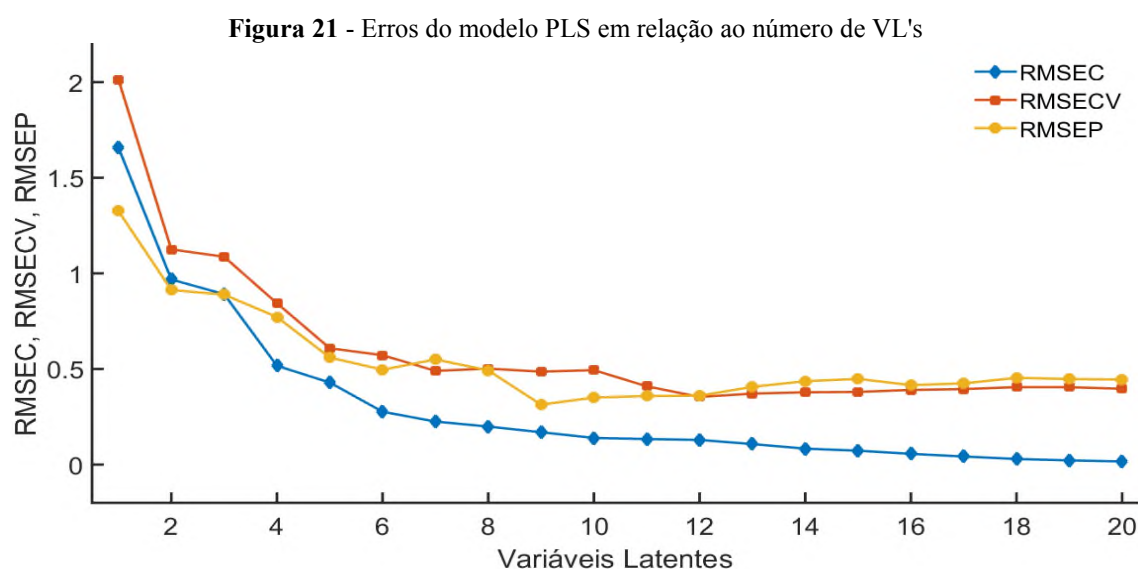
Figura 20 - Modelo gerado após o uso do siPLS

Fonte: O Autor.

É possível notar os bons valores dos coeficientes: angular ($a = 1,0003$), linear ($b = 0,0109$) e de correlação ($R = 0,9998$) conseguidos pelo modelo e que houve uma melhora desses parâmetros quando comparados ao modelo global.

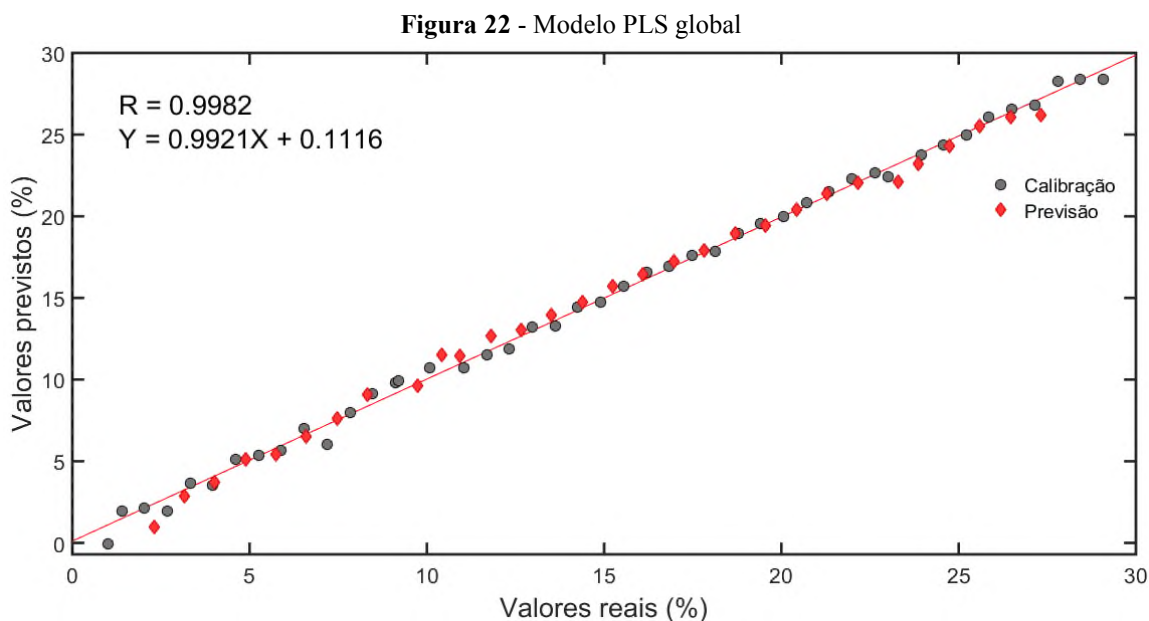
4.2. Mistura biodiesel metílico de mafurra em diesel

Para o modelo contendo misturas de biodiesel metílico de mafurra em diesel foi escolhido 5VL's, já que para esta quantidade ambos os erros são baixos, o RMSEC é menor que o RMSEP conforme mostrado na Figura 21, e a quantidade de amostras nos conjuntos de calibração e previsão ficam dentro do limite mínimo estabelecido pela norma ASTM E1655-05.



Fonte: O Autor.

Então o modelo global representado na Figura 22 obteve $RMSEC = 0,47$ (% v/v), $RMSECV = 0,83$ (% v/v) e $RMSEP = 0,69$ (% v/v) com um $R = 0,9982$ e coeficientes angular e linear satisfatórios.



Fonte: O Autor.

De modo semelhante às análises das misturas biodiesel metílico de crambediesel, os modelos do biodiesel metílico de mafurra gerados por iPLS foram comparados com o modelo global através do teste F. O modelo iPLS5, destacado em negrito na Tabela 5, foi o único a apresentar melhora na capacidade preditiva em relação ao modelo global por obter um F calculado maior que o F tabelado.

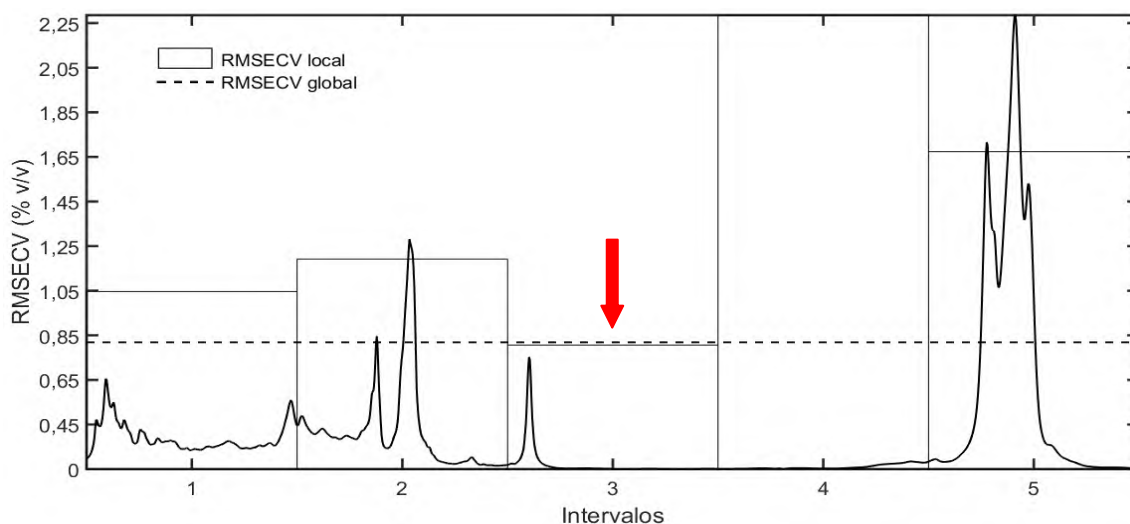
Tabela 5 - Resultados obtidos por iPLS

MODELO iPLS	VARIÁVEIS	FAIXA	VL's	RMSEP (% v/v)	TESTE F*
5	264	530:793	5	0,49	1,93
10	132	134:265	5	0,93	0,55
15	88	530:617	5	0,65	1,13
20	66	530:595	5	0,64	1,17
25	53	531:583	5	0,63	1,18
30	44	530:573	4	0,74	0,86
35	38	533:570	4	0,73	0,88
40	33	530:562	4	0,74	0,86
GLOBAL	2546	1:2546	5	0,69	-----

*Em relação ao modelo global. ($F_{(0,5;30;30)} = 1,84$)

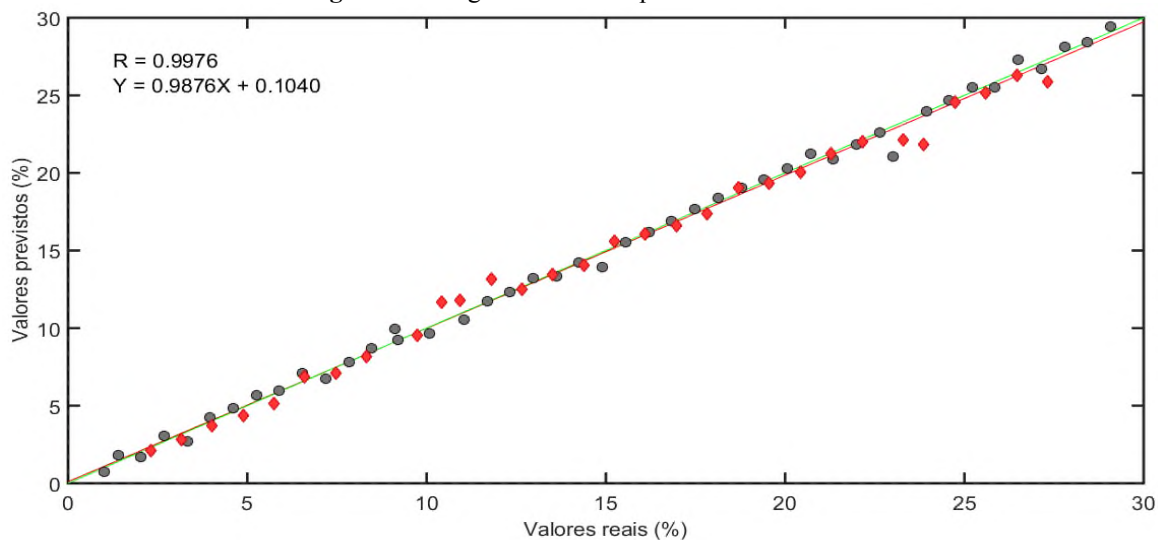
Usando as 263 variáveis do terceiro intervalo do modelo iPLS5 mostrado na Figura 23 e que corresponde à banda atribuída ao estiramento da carbonila (característica de ésteres), foi construído o novo modelo de calibração.

Figura 23 - Gráfico iPLS5



Fonte: O Autor.

O modelo, mostrado na Figura 24 obteve então, $RMSEC = 0,43$ (% v/v), $RMSECV = 0,61$ (% v/v), $RMSEP = 0,49$ (% v/v) e $R = 0,9976$. O menor valor de R deste modelo em relação ao modelo global (diferença de 0,001) pode ser explicado pela maior variância espectral das amostras na faixa selecionada, essa variância por sua vez, contribui para que seja possível quantificar os teores de biodiesel usando somente essas 263 variáveis.

Figura 24 - Regressão obtida após o uso do iPLS

Fonte: O Autor.

O biPLS aplicado para os dados do biodiesel de mafurra, com os dados sumarizados na Tabela 6 não obteve bom desempenho, gerando modelos estatisticamente semelhantes ao modelo global, apresentando valores de F menores que o F tabelado.

Tabela 6 - Resultados dos modelos biPLS

MODELO biPLS	VARIÁVEIS	FAIXA	RMSEP (% v/v)	TESTE F*
5	1452	680:2132	0,53	1,68
10	242	1166:1408	0,52	1,72
15	196	1:196	0,51	1,84
20	161	1166:1327	0,54	1,63
25	938	680:1456; 1554:1635; 1747:1828	0,53	1,68
30	243	924:1005; 1166:1247; 1732:1813	0,56	1,50
35	207	1029:1098; 1166:1235; 1721:1790	0,58	1,42
40	162	518:590;606:656; 799:836	0,52	1,72
GLOBAL	2546	1:2546	0,69	-----

*Em relação ao modelo global. ($F_{(0,5;30;30)} = 1,84$)

A exclusão de faixas espectrais de maneira sistêmica feita pelo biPLS se mostra incapaz (neste caso) de gerar modelos superiores ao modelo gerado com todas as variáveis, mesmo identificando regiões espectrais de alta variância.

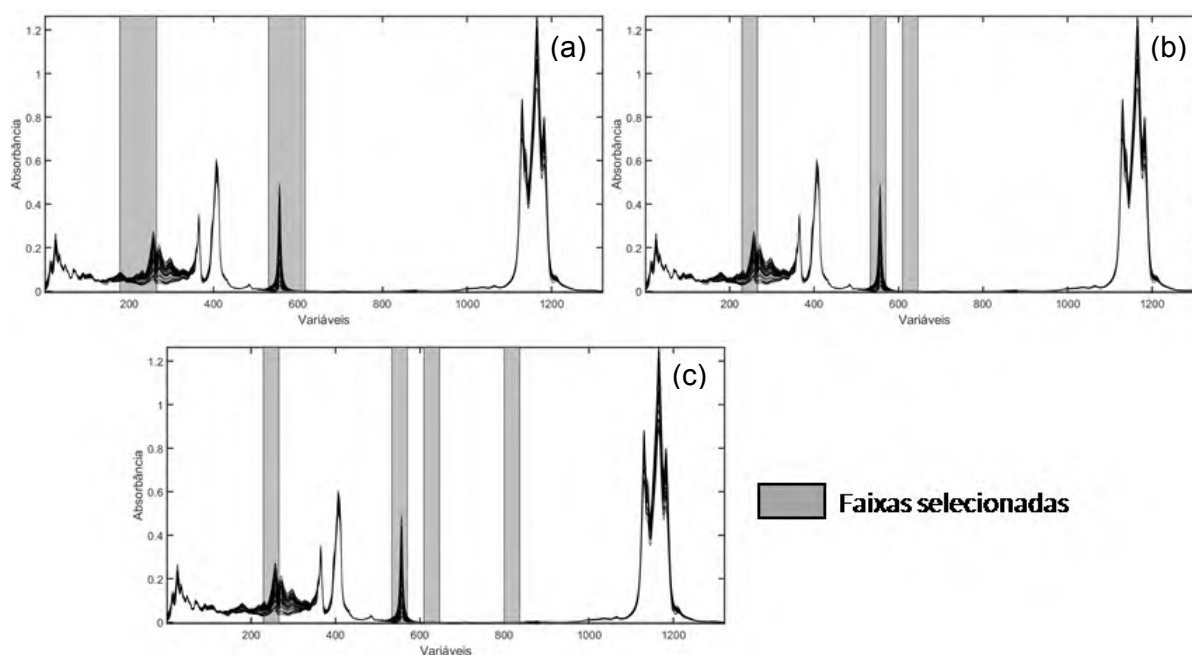
Os modelos simulados por siPLS foram calculados a partir da divisão do espectro em 5, 10, 15, 20, 25, 30, 35 e 40 intervalos, com combinações de 2, 3 e 4 intervalos. Os melhores valores de RMSECV entre as combinações aparecem destacadas em negrito na Tabela 7:

Tabela 7 - RMSECV para as combinações siPLS

Faixas	RMSECV (% v/v)		
	2	3	4
5	0,61	0,60	
10	0,59	0,61	0,61
15	0,46	0,62	0,63
20	0,66	0,53	0,57
25	0,48	0,68	0,82
30	0,49	0,57	0,56
35	0,48	0,46	0,42
40	0,49	0,47	0,56

Três modelos, um para cada combinação de faixas mostradas na Figura 25 foram feitos, sendo que: (a) corresponde as faixas utilizadas no modelo com duas combinações, variáveis de 178:265 e 530:617; (b) é a combinação de faixas para o modelo de 3 combinações, números de variáveis de 229:266, 533:570 e 609:646; (c) as faixas usadas no modelo de 4 combinações, variáveis de 229:266, 533:570, 606:656 e 799:836.

Figura 25 - Faixas usadas nos modelos gerados por siPLS



Fonte: O Autor.

Como os modelos com 3 e 4 combinações apresentaram faixas onde não há informação instrumental (região de linha de base), essas faixas foram retiradas tornando os dois modelos idênticos, porém os diferenciando do modelo com 2 faixas pelo número de variáveis. A Tabela 8 mostra os valores de RMSECV, RMSEC, RMSEP e do teste F dos três modelos PLS, em todos foi usado pré-processamento o *mean center* em ambos os blocos e *vanetian blinds cross-validation* como método de validação cruzada, obtendo 4 VL's.

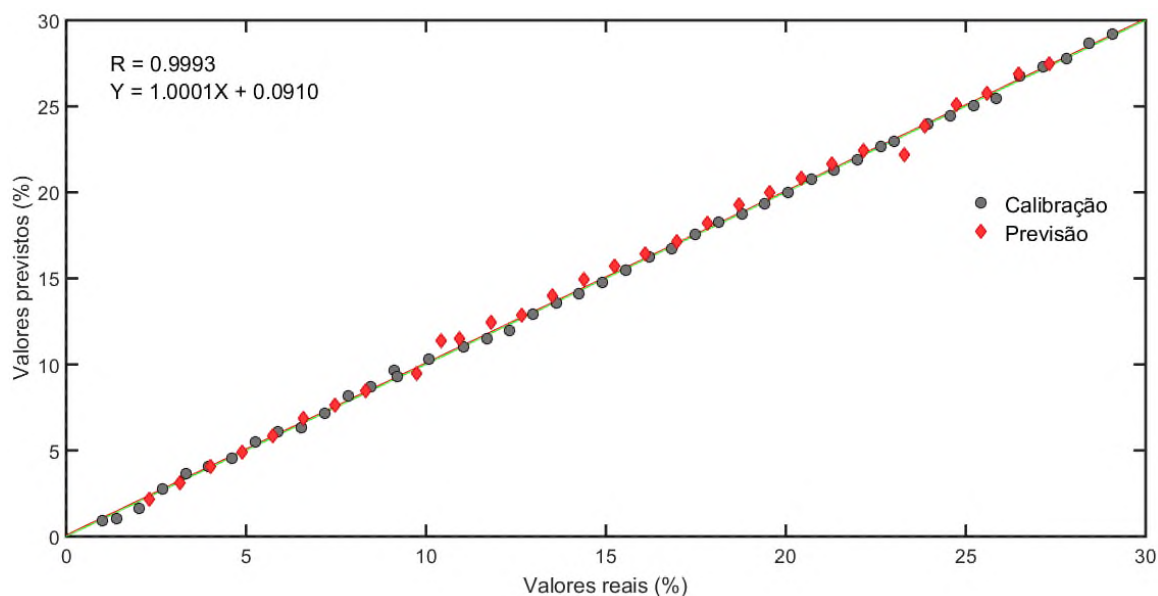
Tabela 8 - Valores dos erros para os modelos gerados após uso do siPLS

Combinações	RMSEC (% v/v)	RMSECV (% v/v)	RMSEP (% v/v)	TESTE F*
2	0,61	0,21	0,57	1,45
3	0,20	0,46	0,42	2,69
4	0,20	0,46	0,42	2,69
GLOBAL	0,47	0,83	0,69	-----

*Em relação ao modelo global. ($F_{(0,5;30;30)} = 1,84$)

A regressão obtida com as 76 variáveis restantes das combinações 3 e 4, é apresentada na Figura 26, onde é possível notar a melhora nos coeficientes angular, linear e de correlação:

Figura 26 - Modelo obtido após o uso do siPLS



Fonte: O Autor.

4.3. Outros parâmetros de Validação

A empregabilidade dos modelos foi avaliada por cinco figuras de mérito além dos RMSE, são elas: Sensibilidade, Limite de Detecção, Limite de Quantificação, Sensibilidade Analítica e Elipse de Confiança (EJCR).

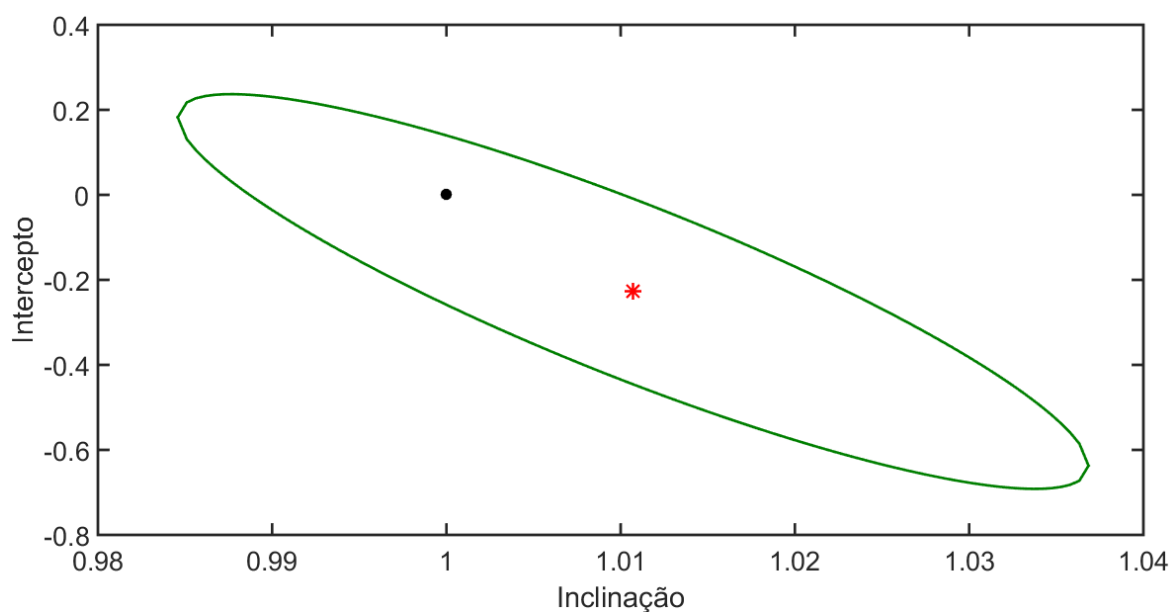
A Tabela 9 apresenta os valores desses parâmetros para o modelo global da mistura de biodiesel metílico de crambe-diesel. Os limites de detecção e quantificação menores que o valor do primeiro ponto da curva (1,00% v/v) confirmam o acerto na escolha da faixa de trabalho para a construção do modelo.

Tabela 9 - Figuras de mérito para o modelo global biodiesel metílico de crambe-diesel

Sensibilidade	0,07
LQ (% v/v)	0,12
LD (% v/v)	0,04
Sensibilidade Analítica	26,98

A elipse de confiança representada na Figura 27, nos mostra a posição do modelo calculado comparada à posição de um modelo ideal em um intervalo de 95% de confiança:

Figura 27 - ERJC para o modelo global da mistura biodiesel metílico de crambe-diesel, onde: (●) representa a posição idealizada do modelo e (*) é a posição real



Fonte: O Autor.

Como o modelo calculado (representado pelo *) está dentro dos limites da elipse, é considerado validado.

Os modelos gerados após o uso do iPLS e siPLS para a mistura de biodiesel metílico de crambe-diesel foram avaliados pelo teste F e não apresentaram diferença estatística entre os seus erros de previsão ($F_{cal} = 1,33$ e $F_{tab} = 1,92$ em relação ao menor RMSEP), logo possuem a mesma capacidade preditiva. Como ambos os modelos passaram por processos matemáticos semelhantes, a discussão sobre se escolher o modelo com menor processamento não se faz relevante, portanto foi escolhido o modelo que usou o menor número de variáveis para a construção da curva de calibração, o iPLS40 com 74 variáveis. A Tabela 10 nos informa os valores das figuras de mérito para este modelo:

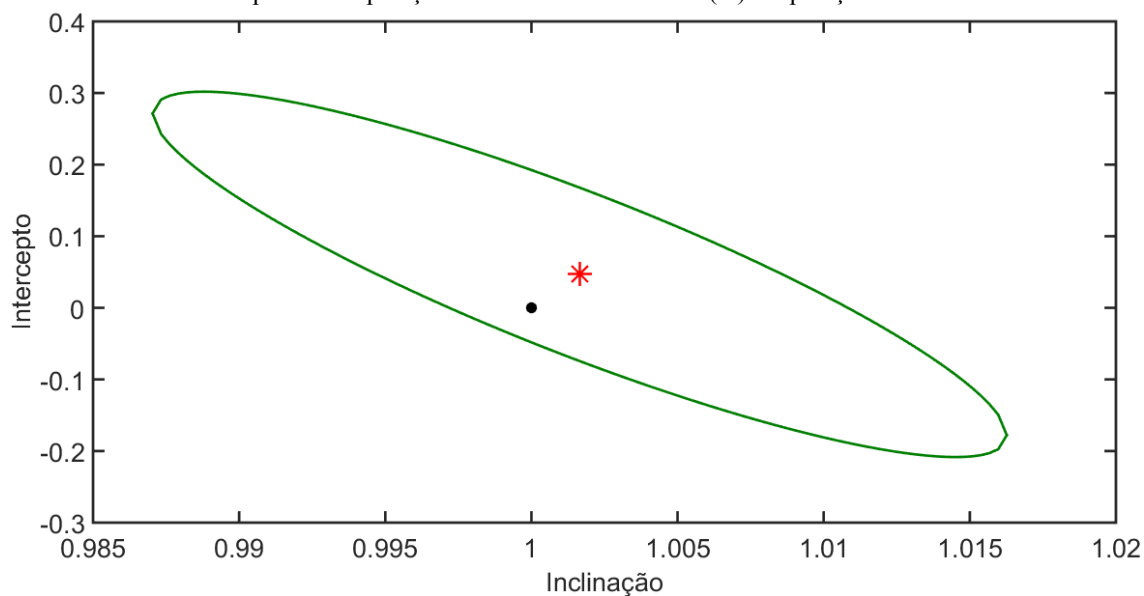
Tabela 10 - Figuras de mérito do modelo iPLS40 de biodiesel metílico de crambe-diesel

Sensibilidade	0,01
LQ (% v/v)	0,25
LD (% v/v)	0,08
Sensibilidade Analítica	30,91

A maior Sensibilidade Analítica do modelo iPLS40 em relação ao modelo global indica seu melhor desempenho, apresentando também limites de LQ e LD menores que o primeiro ponto da curva.

A Figura 28 nos dá sua posição do modelo em relação à posição idealizada dentro de um intervalo de 95% de confiança:

Figura 28 - ERJC para o modelo iPLS40 da mistura biodiesel metílico de crambe-diesel, onde: (●) representa a posição idealizada do modelo e (*) é a posição real



Fonte: O Autor.

Como o modelo (representado pelo *) está dentro da elipse, suas previsões são estatisticamente comprovadas.

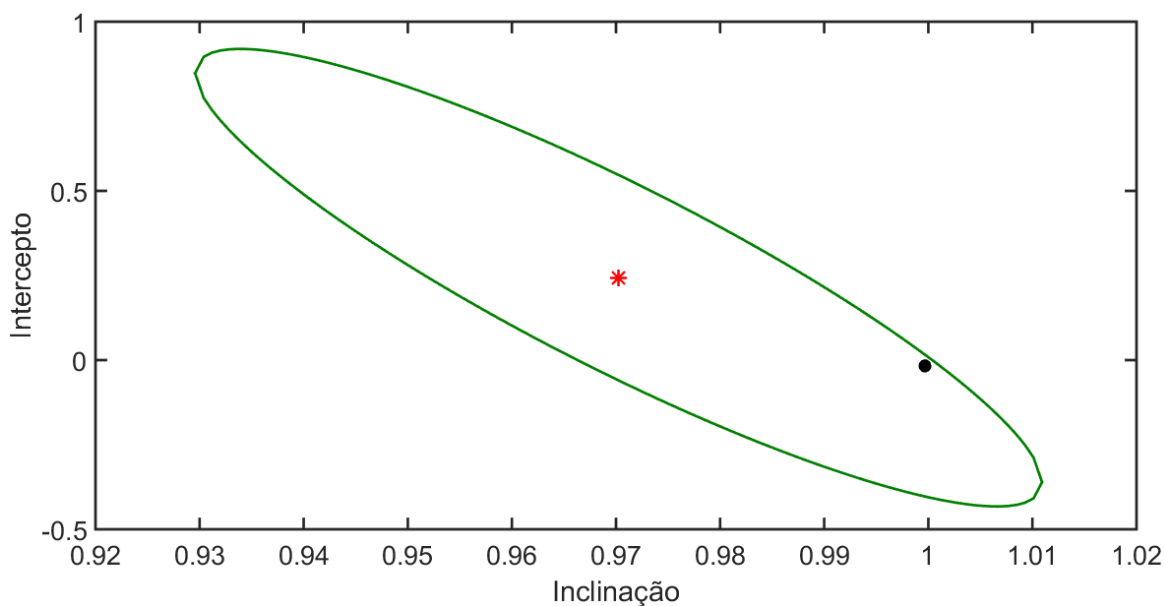
As figuras de mérito para o modelo global da mistura biodiesel metílico de mafurra-diesel com os valores sumarizados na Tabela 11 apresentaram bons resultados, confirmando a assertividade da faixa de trabalho escolhida pelos limites de detecção e quantificação.

Tabela 11 - Figuras de mérito do modelo global de biodiesel metílico de mafurra-diesel

Sensibilidade	0,09
LQ (% v/v)	0,28
LD (% v/v)	0,05
Sensibilidade Analítica	27,85

A EJCR, apresentada na Figura 29, indicou a posição do modelo real próximo aos limites de 95% de confiança, porém como se localiza no interior da elipse, é considerado satisfatório.

Figura 29 - ERJC para o modelo global da mistura biodiesel metílico de mafurra-diesel, onde: (●) representa a posição idealizada do modelo e (*) é a posição real



Fonte: O Autor.

Não havendo diferença estatística entre os modelos gerados pelo iPLS e siPLS ($F_{\text{cal}} = 1,23$ e $F_{\text{tab}} = 1,84$ em relação ao menor RMSEP), para a mistura de biodiesel metílico de mafurra-diesel, foi escolhido novamente aquele que usou menor quantidade de variáveis em sua construção, o siPLS4. Suas figuras de mérito são informadas na Tabela 12:

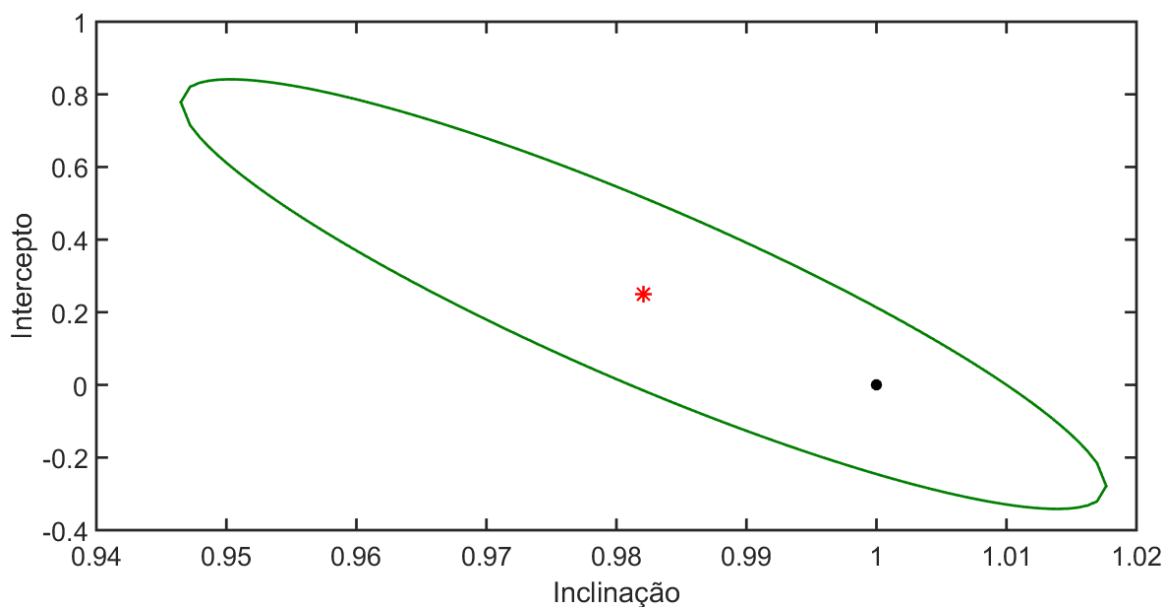
Tabela 12 - Figuras de mérito do modelo siPLS4 de biodiesel metílico de mafurra-diesel

Sensibilidade	0,07
LQ (% v/v)	0,15
LD (% v/v)	0,04
Sensibilidade Analítica	29,05

Com a melhora em todas as figuras de mérito, a diminuição no número de variáveis usadas no modelo de calibração se mostrou muito vantajosa, confirmando a empregabilidade do método.

A Figura 30 apresenta a posição do modelo siPLS4 comparada com a posição idealizada, como o modelo está localizado no interior da elipse, se pode considerá-lo assertivo em suas previsões.

Figura 30 - ERJC para o modelo siPLS4 da mistura biodiesel metílico de mafurra-diesel, onde: (●) representa a posição idealizada do modelo e (✱) é a posição real



Fonte: O Autor.

5. CONCLUSÃO

O método desenvolvido usando a técnica de espectroscopia MIR aliada à análise por Quadrados Mínimos Parciais se mostrou eficiente para a quantificação dos teores de biodiesel metílico de crambe e mafurra em misturas com diesel, na faixa de concentração entre 1,00 a 30,00 (% v/v).

Vantagens como não usar solventes na preparação das amostras, construir uma única curva de calibração contemplando a mesma faixa das duas curvas do método oficial e diminuir significativamente a quantidade pontos usados na construção dos modelos, foram alcançadas. Assim o tempo necessário para a análise do teor dos biodieseis em misturas com diesel se torna menor e não há a geração de resíduos.

Os modelos gerados através dos métodos de seleção de variáveis foram bem sucedidos na diminuição da faixa espectral (com exceção do biPLS) eliminando variáveis não significativas para a análise e melhorando os valores de RMSEC, RMSECV e RMSEP.

Os baixos valores de RMSEC, inferiores a 0,47 (% v/v); RMSECV, menores de 0,83 (% v/v) e RMSEP, inferiores a 0,69 (% v/v) junto com a boa concordância entre os valores previstos e informados ao modelo, com R maiores que 0,9976, reforçam a possível utilização deste procedimento. Os demais parâmetros de validação também obtiveram bons resultados, com valores inferiores a 0,09 para a Sensibilidade, 0,25 (% v/v) para LQ, 0,80 (% v/v) para LD e superiores a 27,85 para a Sensibilidade Analítica.

Os modelos gerados após o uso do iPLS para o biodiesel metílico de crambe e siPLS para o biodiesel metílico de mafurra utilizaram respectivamente, 74 e 76 variáveis para a construção das curvas de calibração. Essa redução nas variáveis além de diminuir a capacidade computacional necessária para o emprego dos modelos viabiliza a construção de detectores para trabalhar na linha de produção, capturando a absorbância dos combustíveis dentro dos dutos e quantificando em tempo real seus teores de biodiesel.

6. REFERÊNCIAS

AMERICAN SOCIETY FOR TESTING AND MATERIALS. **ASTM E-1655-05. Standart Practices for Infrared Multivariate Quantitative Analysis**. Pennsylvania, 2012.

ANDRADE, D.C.; MENDONÇA, T. G.; ROMEIRO, A. R. TENDÊNCIAS DO METABOLISMO DA ECONOMIA BRASILEIRA: Uma análise preliminar à luz da Economia Ecológica. **Revista Iberoamericana de Economía Ecológica**, v. 28, p 66-86, 2018.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR-15568. Biodiesel – Determination of biodiesel content in diesel fuel oil by mid infrared spectroscopy**. Rio de Janeiro, 2008.

BARBOSA, L. C. A. **Espectroscopia no Infravermelho na caracterização de compostos orgânicos**. Viçosa, MG: Editora UFV, 2008.

BARROS NETO, B.; SCARMINIO, I. S.; BRUNS, R. E. 25 anos de quimiometria no Brasil. **Quim. Nova**, v. 29, p. 1401-1406, 2006.

BEEBE, K. R.; KOWALSKI, B. R. An Introduction to Multivariate Calibration and Analysis. **Analytical Chemistry**, v. 59, n. 17, p. 1007A-1017A, 1987.

BORGES NETO, W. **Parâmetros de qualidade de lubrificantes e óleo de oliva através de espectroscopia vibracional, calibração multivariada e seleção de variáveis**. Tese (Doutorado em Química) – Instituto de Química, Universidade Estadual de Campinas, Campinas - SP: 2005.

BRAGA, J. W. B. **Avaliation of Figures of Merit in Multivariate Calibration, applied in a determination of Carbamazepine by Near end Middle Infrared Spectroscopy**. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Estadual de Campinas, Campinas - SP: 2004.

BRASIL. Agência Nacional Do Petróleo, Gás Natural e Biocombustíveis. **ANUÁRIO ESTATÍSTICO BRASILEIRO DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS**, 2018a.

BRASIL. Agência Nacional Do Petróleo, Gás Natural e Biocombustíveis. **Vendas de combustíveis em 2018 ficam estáveis com relação a 2017**. 2018b. Disponível em: <<http://www.anp.gov.br/noticias/5048-vendas-de-combustiveis-em-2018-ficam-estaveis-com-relacao-a-2017>>. Acesso: 10 jun. 19.

BRASIL. Agência Nacional Do Petróleo, Gás Natural e Biocombustíveis. **Síntese Mensal de Comercialização de Combustíveis**, v. 3, 2019a.

BRASIL. Agência Nacional Do Petróleo, Gás Natural e Biocombustíveis. **Boletim de monitoramento da qualidade de biocombustíveis**. Brasília, 2019b.

BRASIL. **Lei nº 13.263, de 23 de março de 2016**. Altera a Lei nº 13.033, de 24 de setembro de 2014, para dispor sobre os percentuais de adição de biodiesel ao óleo diesel comercializado no território nacional. Diário Oficial da União, Brasília. 2016.

BRUNS, R. E.; FAIGLE, J. F. G. QUIMIOMETRIA. **Quim. Nova**, p. 84-98, 1985.

CARVALHO ROCHA, W. F. **Utilization of net analyte signal for validation of models of multivariate calibration by using figure of merit and chart control**. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Estadual de Campinas, Campinas - SP: 2007.

EMPRESÔMETRO - Inteligência de Mercado. **FROTA BRASILEIRA DE VEÍCULOS EM CIRCULAÇÃO, 2019**.

FILGUEIRAS, P. R.; ALVES, J. C. L.; POPPI, R. J. Quantification of animal fat biodiesel in soybean biodiesel and B20 diesel blends using near infrared spectroscopy and synergy interval support vector regression. **Talanta**, v. 119, p. 582-589, 2014.

FOCORURAL. **Crambe pode ser alternativa para o produtor**. 2016, il. color. Disponível em: <http://www.focorural.com/detalhes/e/n/8971/Crambe_pode_ser_alternativa_para_o_produto.html>. Acesso: 17 jun. 19.

GONTIJO, L. C. **Uso de espectrometria no infravermelho médio, calibração multivariada e seleção de variáveis por intervalos na quantificação de biodieseis em misturas com diesel**. Tese (Doutorado em Química) – Instituto de Química, Universidade Federal de Uberlândia, Uberlândia - MG: 2016.

GUIMARÃES, E. **Aplicação de espectroscopia no infravermelho médio e calibração multivariada para a quantificação de adulterantes em biodieséis de óleo de soja e residual e suas misturas com diesel**. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Federal de Uberlândia, Uberlândia - MG: 2014.

HAGE, D. S.; CARR, J. D. **QUÍMICA ANALÍTICA E ANÁLISE QUANTITATIVA**. Tradução de Midori Yamamoto. 1. ed. São Paulo: Pearson Prentice Hall, 2012.

HERBERGER, K. **Chemoinformatics-multivariate mathematical-statistical methods for data evaluation**. In: Vékey, C.; Takes, A.; Vertes, A. (Ed.). *Medical Applications of Mass Spectrometry*, 1 ed. Amsterdam, cap. 7, Ed. Elsevier, 2007.

HOLLER, F. J.; SKOOG, D. A.; CROUCH, S. R. **Princípios de Análise Instrumental**. 6. ed. Porto Alegre: Bookman, 2009.

JASPER, S. P.; BIAGGIONI, M. A. M.; SILVA, P. R. A.; SEKI, A. S.; BUENO, O. C. Análise Energética da cultura do crambe (*Crambe abyssinica* Hochst) produzida em plantio direto. **Engenharia Agrícola. Associação Brasileira de Engenharia Agrícola**, v. 30, n. 3, p. 395-403, 2010. Disponível em: <<http://hdl.handle.net/11449/5007>>. Acesso: 16 jun. 19.

LALAS, S.; GORTZI, O.; ATHANASIADIS, V.; DOURTOGLOU, E.; DOURTOGLOU, V. Full characterization of *Crambe abyssinica* Hochst. seed oil. **Journal of the American Oil Chemists' Society**, vol. 89, p. 2253-2258, 2012.

LEARDI, R.; NØRGAARD, L. Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions. **Journal Chemometrics**, v. 18, p. 486-497, 2004.

LEBID, T.; HENKES, J.A. Óleo de dendê na produção de biodiesel: um estudo de caso das vantagens e desvantagens econômica, ecológica e social da cultura desta oleaginosa para a produção de biodiesel. **R. gest. sust. ambient.**, Florianópolis, v. 4, n. 1, p. 392-415, 2015.

LIMA, L. P.; LOPES, A.; OLIVEIRA, M. C. J.; NEVES, M. C. T.; IAMAGUTI, P. S. BIODIESEL DE DENDÊ EM TRATOR AGRÍCOLA: OPACIDADE DA FUMAÇA

EM FUNÇÃO DAS CONDIÇÕES DE FUNCIONAMENTO DO MOTOR. **Ciência & Tecnologia: Fatec-JB**, v. 6, p. 22-31, 2014.

MÁQUINA, A. D. V. **ESPECTROMETRIA NO INFRAVERMELHO MÉDIO E MÉTODOS QUIMIOMÉTRICOS PLS-DA E PLS: classificação e previsão do teor de biodiesel na mistura de biodiesel/diesel de mafurra, moringa e algodão**. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Federal de Uberlândia, Uberlândia - MG: 2017.

MASSART, D. L.; VANDEGINSTE, B. G. M.; BUYDENS, L. M. C.; DE JONG, S.; LEWI, P. J.; SMEYERS-VERBEKE, J. **Handbook of Chemometrics and Qualimetrics: Part A**. Elsevier: Amsterdam, 1998.

MELLO, B. T. F.; GONCALVES, J. E.; RODRIGUES, G. M.; CARDOZO FILHO, L.; SILVA, C. Hydroesterification of crambe oil (*Crambe abyssinica* H.) under pressurized conditions. **Industrial Crops and Products**, v. 97, p. 110-119, 2017.

MITSUTAKE, H. **Uso de espectroscopia no infravermelho médio e cartas de controle multivariadas no monitoramento da qualidade de biodieseis de soja e fritura residual em misturas com diesel**. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Federal de Uberlândia, Uberlândia - MG: 2015.

MUNCK, L.; PRAM NIELSEN, J.; MØLLER, B.; JACOBSEN, S.; SØNDERGAARD, I.; ENGELSEN, S. B.; NØRGAARD, L.; BRO, R. Exploring the phenotypic expression of a regulatory proteome-altering gene by spectroscopy and Chemometrics. **Analytical Chimica Acta**, v. 446, p. 171-186, 2001.

NAZARENO, J. C.; OLIVEIRA, G.G.; VIEIRA, G.E.G. Aproveitamento do Dendê para Produção de Biodiesel na Região Amazônica Brasileira. **Bioenergia em Revista: Diálogos**, v. 5, p. 1-10, 2015.

NØRGAARD, L. The iToolbox for MATLAB, KVL, Denmark, July 2004. Disponível em: <<http://www.models.life.ku.dk>>. Acesso: 10 jun. 19.

OTTO, M. **Chemometrics – Statistics and Computer application in Analytical Chemistry**, Wiley, 1999.

PLANTZAFRICA. **Trichilia emetica**. 2014, il. color. Disponível em: <<http://pza.sanbi.org/trichilia-emetica>>. Acesso: 17 jun. 19.

PORTELA, F. M. **Efeito da catálise ácida e alcalina na produção e propriedades físico-químicas do biodiesel metílico de pinhão-mansão**. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Federal de Uberlândia, Uberlândia - MG: 2011.

PORTELA, F. M. **Produção de biodieseis etílico e metílico de pinhão-mansão usando diferentes rotas catalíticas**. Tese (Doutorado em Química) – Instituto de Química, Universidade Federal de Uberlândia, Uberlândia - MG: 2015

SALVADOR, P. M.; MARTINS, C. X.; ROMANHOLO FERREIRA, L. F.; AMÉRICO, J. H. P.; TORRES, N. H.; BORTOLETO, G. G. Estudo de diferentes catalisadores para produção de biodiesel: viabilidade dos catalisadores heterogêneos. **bioenergia em revista: diálogos**, n. 2, p. 31-46, 2016.

SANTOS QUEIROZ, D. **Transterificação de triaglicerol de óleos de milho e de soja: Análise quimiométrica do processo e propriedades físico-químicas essenciais do produto, para uso como biodiesel**. Tese (Doutorado em Química) – Instituto de Química, Universidade Federal de Uberlândia, Uberlândia - MG: 2010.

SANTOS, K. A.; BARICCATTI, R. A.; CARDOZO-FILHO, L.; SCHNEIDER, R.; PALÚ, F.; SILVA, C.; SILVA, E. A. J. Extraction of crambe seed oil using subcritical propane: Kinetics, characterization and modeling. **Supercrit. Fluids**, v. 104, p. 54-61, 2015.

SILVA, A. A. **DETERMINAÇÕES COMPARATIVAS DA UTILIZAÇÃO DE HIDRÓXIDO DE POTÁSSIO (KOH) E METÓXIDO DE POTÁSSIO (CH₃OK) PARA OBTENÇÃO DE BODIESEL METÍLICO DO ÓLEO DE SOJA RESIDUAL**. Monografia (Graduação em Química Licenciatura) – Universidade Federal de Uberlândia, Ituiutaba - MG: 2017.

SILVEIRA, E. V. **Otimização do Processo de Trasterificação Metílica via Catálise Alcalina do óleo de crambe (Crambe abyssinicaHochst)**. Dissertação (Mestrado em Química) – Instituto Federal de Educação, Ciência e Tecnologia Goiano, Rio Verde - GO: 2015.

SIMÕES, D. M. B. **Viabilidade de utilização de mafurra (*Trichilia emética*) para a produção de biodiesel**. Dissertação (Mestrado em Engenharia de Energia e Ambiental) – Faculdade de Ciências, Universidade de Lisboa, Lisboa: 2014.

SOARES, K. M. C. **DESENVOLVIMENTO DE NOVA ALTERNATIVA PARA PRODUÇÃO DE BODIESEL VIA CATÁLISE BÁSICA**. Monografia (Graduação em Química Tecnológica) – Universidade Federal de Brasília, Brasília: 2016.

SOUZA, L. M. **Controle de qualidade de óleos de linhaça e ginkgo biloba usando espectrometria no infravermelho médio e ferramentas quimiométricas**. Tese (Doutorado em Química) – Instituto de Química, Universidade Federal de Uberlândia, Uberlândia - MG: 2018.

TAVARES, G. R., MASSA, T. B., GONÇALVES, J. E., SILVA, C., SANTOS, W. D. Assesment of ultrasound-assisted extraction of crambe seed oil for biodiesel synthesis by in situ interesterification. **Renewable Energy**, n. 111, p. 659-665, 2017.

VALDERRAMA, P. **Avaliação de figuras de mérito em calibração multivariada na determinação de parâmetros de controle de qualidade em indústria alcooleira por espectroscopia no infravermelho próximo**. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Estadual de Campinas, Campinas - SP: 2005.

VIEIRA, J. S. C.; SOUSAB, T. L.; ROSASB, L. S.; LIMAC, A. L.; RONCONID, C. M.; MOTA, C. J. A. ESTERIFICAÇÃO E TRANSESTERIFICAÇÃO HOMOGÊNEA DE ÓLEOS VEGETAIS CONTENDO ALTO TEOR DE ÁCIDOS GRAXOS LIVRES. **Quim. Nova**, v. XY, n. 0, p. 1-7, 2017.