

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

INSTITUTO DE QUÍMICA

DISSERTAÇÃO DE MESTRADO

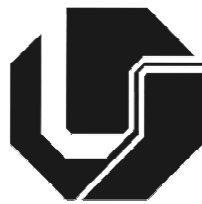
**Aplicação de Espectroscopia no Infravermelho Médio e Análise
Discriminante por Mínimos Quadrados Parciais para Classificação de
Biodieseis e Misturas Biodiesel/Diesel.**

Mestrando: Carlos Alberto Tadeu Zanetti

Orientador: Prof. Dr. Waldomiro Borges Neto

Uberlândia

Janeiro 2014



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

INSTITUTO DE QUÍMICA

DISSERTAÇÃO DE MESTRADO

**Aplicação de Espectroscopia no Infravermelho Médio e Análise
Discriminante por Mínimos Quadrados Parciais para Classificação de
Biodieseis e Misturas Biodiesel/Diesel.**

**Dissertação apresentada ao Programa
de Pós-Graduação em Química da
Universidade Federal de Uberlândia,
como requisito para obtenção de título
de Mestre em Química.**

Mestrando: Carlos Alberto Tadeu Zanetti

Orientador: Prof. Dr. Waldomiro Borges Neto

Área de concentração: Química Analítica

Uberlândia

Janeiro 2014

Dedico esse trabalho em especial
à minha esposa
Érica Regina Tomazinho Zanetti,
ao meu filho ou filha que irá nascer,
Francisco ou Clara
às três mulheres que sempre me incentivaram e
confiaram em mim
minha tia,
Abadia Lourdes Corrêa
minha mãe,
Edna Regina Corrêa Zanetti
minha irmã,
Rita de Cássia Corrêa Zanetti Magalhães.
Ao meu eterno exemplo
Jesus Zanetti †
meu pai.

“É melhor tentar e falhar, que preocupar-se e ver a vida passar. É melhor tentar, ainda que em vão que sentar-se, fazendo nada até o final. Eu prefiro na chuva caminhar, que em dias frios em casa me esconder. Prefiro ser feliz embora louco, que em conformidade viver.”

Martin Luther King



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA



Ata da defesa de DISSERTAÇÃO DE MESTRADO junto ao Programa de Pós-Graduação em Química, do Instituto de Química da Universidade Federal de Uberlândia
DEFESA DE DISSERTAÇÃO DE MESTRADO ACADÊMICO, NÚMERO 189/PPQUI.

DATA: 28/03/2014

DISCENTE: Carlos Alberto Tadeu Zanetti

MATRÍCULA: 11212QMI003

TÍTULO DO TRABALHO: Aplicação de espectroscopia no Infravermelho Médio e análise discriminante por mínimos quadrados parciais para classificação de biodieleses e misturas biodiesel/diesel.

ÁREA DE CONCENTRAÇÃO: Química

LINHA DE PESQUISA: Espectroanalítica Aplicada

PROJETO DE PESQUISA DE VINCULAÇÃO: Aplicação de espectroscopia no Infravermelho e análise discriminante por quadrados mínimos parciais para controle de qualidade de biocombustíveis.

Às quatorze horas do dia vinte e oito de março do ano dois mil e quatorze, no Auditório Prof. Manuel Gonzalo Hernández Terrones, Bloco 5I, no Campus Santa Mônica, reuniu-se a Banca Examinadora composta pelos Professores Doutores Luiz Carlos Moutinho Pataca, da Fundação Centro Tecnológico de Minas Gerais, Sérgio Antônio Lemos de Moraes e Waldomiro Borges Neto, do Instituto de Química da UFU, professor(a) orientador(a) e presidente da mesa. Iniciando os trabalhos, o(a) presidente da mesa apresentou o(a) candidato(a) e a Banca Examinadora, agradeceu a presença do público e discorreu sobre as normas e critérios para a realização desta sessão, baseadas na Norma Interna nº 03/2013/PPQUI. Em seguida, o(a) presidente da mesa concedeu a palavra ao(a) candidato(a) para a exposição do seu trabalho e, em seqüência, aos examinadores, em ordem sucessiva, para arguir o(a) apresentador(a). A duração da apresentação e o tempo de arguição e resposta deram-se conforme as normas do Programa. Ultimada a arguição, desenvolvida dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu os conceitos finais e A provou o(a) candidato(a). Por sugestão da Banca Examinadora, o título do trabalho será mantido como o original

Esta defesa de Dissertação de Mestrado Acadêmico é parte dos requisitos necessários à obtenção do título de Mestre. O competente diploma será expedido após cumprimento do estabelecido nas normas do Programa, legislação e regulamentação internas da UFU. As correções observadas pelos examinadores deverão ser realizadas no prazo máximo de 30 dias. Nada mais havendo a tratar, deu-se por encerrada a sessão às 16 horas e 40 minutos e lavrada a presente ata, que após lida e aprovada, foi assinada pela Banca Examinadora.

Prof. Dr. Luiz Carlos Moutinho Pataca
CETEC

Prof. Dr. Sérgio Antônio Lemos de Moraes
IQUFU

Orientador(a): Prof. Dr. Waldomiro Borges Neto – IQUFU

Universidade Federal de Uberlândia - Avenida João Naves de Ávila, nº 2121, Bairro Santa Mônica - 38.408-144 - Uberlândia - MG

+55 - 34 - 3239-4385

cpqquimica@ufu.br

<http://www.cpqquimica.iq.ufu.br>

AGRADECIMENTOS

- . À **DEUS** pois é o começo, o meio e o fim de todas as coisas, me deu a vida, a família, os amigos, a força e a perseverança para a realização desse trabalho.
- . À minha esposa **Érica** pela paciência e por todos seus cuidados com nossa família, pessoa sem a qual eu não teria começado nem terminado esse trabalho.
- . À minha **Mãe** e à minha **Tia Abadia** pelas orações, pelos estímulos dados aos meus estudos, desde os primeiros livros até o que sou hoje, sem elas não teria chegado até aqui.
- . Ao meu **Pai** (in memórian) pelo seu bom humor, seu exemplo e seu caráter.
- . À minha **IRMÃ** e a meu cunhado **Marcio Paulo Magalhães** por me aturarem muitas vezes, me escutando e me recebendo com muito carinho.
- . Ao Prof. Dr. **Waldomiro Borges Neto**, pela eterna paciência, pelos vários puxões de orelha (que hoje vejo que foram necessários), pelos conhecimentos transmitidos, pela amizade, pela confiança e por sua dedicação que apesar de se esconder como uma pessoa dura tem um enorme coração. A você Dr. Waldomiro, minha eterna gratidão.
- . A todos os professores do departamento de química do Colégio Cenecista Dr. José Ferreira que direta ou indiretamente me ajudaram para a realização desse trabalho, em especial ao meu amigo, irmão e exemplo **Prof. Esp. João Carlos de Lima**.
- . Ao meu sensei **Moises Cardoso** e a toda **família Jiu Jitsu** da academia Judô Oriente. Todos sempre perguntando sobre o andamento do trabalho e não deixando eu “fugir” muito dos treinos, permitindo uma integração entre meu corpo e mente que foi fundamental em muitos momentos difíceis.
- . Aos meus colegas do LQT que nunca esquecerei: **Eloisa** (Elo) a qual eu não tenho palavras para agradecer, **Hery** (Pokemon) pela ajuda e pelos longos finais de semana me ensinando e acompanhando no laboratório, **Lucas** (Maza) pelas várias horas gastas em me ajudar. E ainda ao **Felipe** Bachion, **Letícia** (Abobrón), **Sarmento** (Caramunhão), **Fábio**, **Edvando**, **Zé Eduardo** (Buiatte), pelo companheirismo, amizade, viagens, pelos momentos de tensão, alegria, pelas festas, enfim por terem feito parte de minha vida.
- . À FAPEMIG pelo financiamento do projeto 17.014/11.
- . À TRANSPETRO pelo fornecimento do diesel.
- . À CARAMURU pelo fornecimento de biodieseis.
- . Ao IQUFU por fornecer toda estrutura física e tecnológica para a realização deste trabalho.
- . À UFU pela oportunidade de aprimorar os meus conhecimentos.

“Porque um dia é preciso parar de sonhar, tirar os planos das gavetas e, de algum modo, começar...” (Amyr Klink)

RESUMO

Frente aos estímulos econômicos gerados pelos agronegócios em torno da produção de biocombustíveis e benefícios ao meio ambiente, a produção de biodiesel vem ganhando mercado, fazendo-se necessário o controle de qualidade ao longo de toda sua cadeia de produção e distribuição. No Brasil, essa fiscalização é feita pela Agência Nacional de Petróleo, Gás Natural e Biocombustíveis (ANP) que estabelece parâmetros baseados em normas internacionais como, por exemplo, da American Society for Testing and Materials (ASTM). Na literatura, encontra-se grande quantidade de trabalhos que visam o controle de qualidade na produção de biodiesel e misturas com diesel. Porém, a maioria dessas técnicas demandam grandes quantidades de reagentes, são demoradas e utilizam instrumentos analíticos caros.

Ao encontro desse contexto técnicas espectroscópicas, principalmente no infravermelho apresentam-se como uma alternativa rápida, direta, não destrutiva, limpa e eficiente, podendo ser aplicadas *in situ* no controle de qualidade do biodiesel em todas as etapas da cadeia produtiva e fiscalização. As potencialidades dessas técnicas tornam-se quase ilimitadas quando aliadas ao uso de métodos quimiométricos para extrair informações relevantes de dados complexos tornando-os mais simples de serem aplicados nas rotinas de laboratórios.

Diante da situação apresentada, o presente trabalho visa à utilização de técnicas de espectroscopia no infravermelho médio, aliada a técnica de classificação por análise discriminante por mínimos quadrados parciais (PLS-DA) para o desenvolvimento de método analítico para classificação de amostras de biodieseis de acordo com sua matriz energética em relação aos biodieseis puros (B100) e de suas misturas em diesel com teor de 5% de biodiesel (B5). Foram desenvolvidos também, modelos de classificação para verificação da conformidade quanto ao teor de B5.

O uso da Espectroscopia no Infravermelho Médio aliada ao método de Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA) foi capaz de classificar com eficiência próxima a 100% as amostras de interesse dos modelos desenvolvidos nesse trabalho. Portanto sua aplicação torna-se viáveis para o monitoramento e fiscalização da qualidade destes tipos de combustíveis.

Palavras-chaves: Biodieseis, Quimiometria, PLS-DA.

ABSTRACT

Faced with the economic stimulus generated by agribusiness around the production of biofuels and environmental benefits, biodiesel production has been gaining market making it necessary quality control throughout its chain of production and distribution. In Brazil, such supervision is done by the National Agency of Petroleum, Natural Gas and Biofuels (ANP) establishing parameters based on international standards, eg, American Society for Testing and Materials (ASTM). It is found, in literature, great amount of studies aimed quality control in the production of biodiesel and blends with diesel. However, most of these techniques require large quantities of reagents, used are time consuming and ue expensive analytical instruments.

Within this context, Spectroscopic techniques, especially in the infrared are presented as a rapid, direct, non-destructive, clean and efficient alternative, and can be applied in situ quality control of biodiesel in all stages of the production chain and supervision. The potential of these techniques become almost limitless when combined with the use of chemometric methods to extract relevant information from complex data making them simpler to be applied in routine laboratories.

Given the situation described, the present work ims at the use of spectroscopy in the mid infrared, combined the technique of discriminant analysis classification by partial least squares (PLS-DA) for the development of analytical method for classifying samples according biodiesels with their energy relative to pure biodiesels (B100) and their blends with diesel content of 5% biodiesel (B5). , Classification models for compliance with the contents of B5 were also developed.

The usage of Spectroscopy in the Middle Infrared allied to the method of discriminant analysis by Partial Least Squares (PLS-DA) was able to sort through nearly 100% efficiency the samples of interest of the models developed in this work. So your application becomes feasible for the monitoring and supervision of the quality of these types of fuels.

Keywords: Biodiesels, Chemometrics, PLS-DA.

LISTA DE SIGLAS

ABNT – Associação Brasileira de Normas Técnicas.

ANP – Agência Nacional do Petróleo, Gás Natural e Biocombustíveis

ASTM – American Society for Testing and Materials

ATR – Reflectância Total Atenuada (do inglês, *Attenuated Total Reflectance*)

B100 – Biodiesel puro

B5 – 5% de biodiesel e 95% de diesel em volume

BER – Biodiesel Etílico Residual

BES – Biodiesel Etílico de Soja

BMR – Biodiesel Metílico Residual

BMS – Biodiesel Metílico de Soja

BMS_{Sebo} – Biodiesel Metílico de Sebo

CEN – Comitê Europeu de Normalização (do francês, *Comité Européen de Normalisation*)

MIR – Infravermelho Médio (do inglês, *Middle Infrared*)

NIPALS – Iterativo Não Linear Mínimos Quadrados Parciais (do inglês, *Nonlinear Iterative Partial Least Squares*)

PC – Componente Principal (PC do inglês, *Principal Components*)

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

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

PLS-DA – Análise discriminante por mínimos quadrados parciais (do inglês, *Partial Least Squares for Discriminant Analysis*)

SIMCA – Modelagem Suave Independente Por Analogias de Classes (SIMCA, do inglês, *Soft Independent Modeling of Class Analogy*)

SIMPLS – Valor singular de Decomposição (SIMPLS do inglês, *Singular Value Decomposition*)

VL – Variável Latente

LISTA DE TABELAS

Tabela 1 – Regiões espectrais do infravermelho.....	10
Tabela. 2 – Amostras de calibração e previsão de B100.....	25
Tabela. 3 – Amostras de calibração e previsão de B5.....	26
Tabela. 4 – Amostras de calibração e previsão de BMS com diferentes teores.....	26
Tabela 5 – Valores de estimativa previstos para amostras de interesse.....	35
Tabela 6 – Valores de probabilidade para amostras de interesse.....	36
Tabela 7 – Valores previstos para amostras B100 de BER.....	38
Tabela 8 – Valores de probabilidade para amostras B100 de BER.....	40
Tabela 9 – Valores previstos para amostras B100 de BMR.....	42
Tabela 10 – Valores de probabilidade para amostras B100 de BMR.....	44
Tabela 11 – Valores previstos para amostras B100 de BES.....	46
Tabela 12 – Valores de probabilidade para amostras B100 de BES.....	48
Tabela 13 – Valores previstos para amostras B100 de BMS.....	50
Tabela 14 – Valores de probabilidade para amostras B100 de BMS.....	52
Tabela 15 – Valores previstos para amostras B100 de BMSebo.....	54
Tabela 16 – Valores de probabilidade para amostras B100 de BMSebo.....	56
Tabela 17 – Valores previstos para amostras B5 de BER.....	58
Tabela 18 – Valores de probabilidade para amostras B5 de BER.....	60
Tabela 19 – Valores previstos para amostras B5 de BMR.....	62
Tabela 20 – Valores de probabilidade para amostras B5 de BMR.....	64
Tabela 21 – Valores previstos para amostras B5 de BES.....	66
Tabela 22 – Valores de probabilidade para amostras B5 de BES.....	68
Tabela 23 – Valores previstos para amostras B5 de BMS.....	69
Tabela 24 – Valores de probabilidade para amostras B5 de BMS.....	71
Tabela 25 – Valores previstos para amostras B5 de BMSebo.....	73
Tabela 26 – Valores de probabilidade para amostras B5 de BMSebo.....	75
Tabela 27 – Valores previstos para amostras com 5% de Biodiesel de BMS.....	77
Tabela 28 – Valores de probabilidade para amostras com 5% de Biodiesel de BMS.....	79

LISTA DE FIGURAS

Figura 1 – Demanda energética mundial.....	02
Figura 2 – Produção de Biodiesel em 2010.....	07
Figura 3 – Reação global de transesterificação.....	09
Figura 4 – Espectrofotômetro FT-MIR com acessório de ATR horizontal.....	12
Figura 5 – Acessório de ATR horizontal.....	13
Figura 6 – Decomposição da matriz X nas matrizes de escores (T), pesos (P) e resíduos (E).....	15
Figura 7 – Representação de duas componentes principais.....	16
Figura 8 – Rotação no eixo do PC (Variável Latente).....	17
Figura 9 – Espectros originais de B100 sem correção de linha de base e sem cortes.....	30
Figura 10 – Espectros de B100 após correção da linha de base por baseline e cortes.....	31
Figura 11 – Espectros originais de B5 sem correção de linha de base e sem cortes.....	31
Figura 12 – Espectros de B5 após correção da linha de base por baseline e cortes.....	32
Figura 13 – Espectros originais de BMS com diferentes teores sem correção de linha de base e sem cortes.....	32
Figura 14 – Espectros de BMS com diferentes teores após correção da linha de base por baseline e cortes.....	33
Figura 15 – Gráfico de estimativas de Previsão de amostras geradas pelo modelo PLS-DA. Amostras de calibração: Classe de interesse(▼), Outras classes (✱). Amostras de validação (●).....	34
Figura 16 – Gráfico de Probabilidade de amostras geradas pelo modelo PLS-DA. Amostras de calibração: Classe de interesse(▼), Outras classes (✱). Amostras de validação (●).....	35
Figura 17 – Previsão de amostras de BES-R utilizando PLS-DA. Amostras de calibração: BER (▼), Outros biodieseis (✱). Amostras de validação (●).....	37
Figura 18 – Probabilidade de classificação de amostras de BER utilizando PLS-DA. Amostras de calibração: BER (▼), Outros biodieseis(✱). Amostras de validação (●).....	39
Figura 19 – Previsão de amostras de BMR utilizando PLS-DA. Amostras de calibração: BMR (▼), Outros biodieseis(✱). Amostras de validação (●).....	41
Figura 20 – Probabilidade de classificação de amostras de BMR utilizando PLS-DA. Amostras de calibração: BMR (▼), Outros biodieseis(✱). Amostras de validação (●).....	43

Figura 21 – Previsão de amostras de BES utilizando PLS-DA. Amostras de calibração: BES (▼), Outros biodieseis(*). Amostras de validação (●).....	45
Figura 22 – Probabilidade de classificação de amostras de BES utilizando PLS-DA. Amostras de calibração: BES (▼), Outros biodieseis(*). Amostras de validação (●)....	47
Figura 23 – Previsão de amostras de BMS utilizando PLS-DA. Amostras de calibração: BMS (▼), Outros biodieseis(*). Amostras de validação (●).....	49
Figura 24 – Probabilidade de classificação de amostras de BMS utilizando PLS-DA. Amostras de calibração: BMS (▼), Outros biodieseis(*). Amostras de validação (●).....	51
Figura 25 – Previsão de amostras de BMSebo utilizando PLS-DA. Amostras de calibração: BMSebo (▼), Outros biodieseis(*). Amostras de validação (●).....	53
Figura 26 – Probabilidade de classificação de amostras de BMSebo utilizando PLS-DA. Amostras de calibração: BMSebo (▼), Outros biodieseis(*). Amostras de validação (●).....	55
Figura 27 – Previsão de amostras de B5 preparados com BER utilizando PLS-DA. Amostras de calibração: BER (▼), Outros B5(*). Amostras de validação (●).....	57
Figura 28 – Probabilidade de classificação de B5 preparados com BER utilizando PLS-DA. Amostras de calibração: BER(▼), Outros B5(*). Amostras de validação (●).....	59
Figura 29 – Previsão de amostras de B5 preparados com BMR utilizando PLS-DA. Amostras de calibração: BMR(▼), Outros B5(*). Amostras de validação (●).....	61
Figura 30 – Probabilidade de classificação de B5 preparados com BMR utilizando PLS-DA. Amostras de calibração: BMR(▼), Outros B5(*). Amostras de validação (●).....	63
Figura 31 – Previsão de amostras de B5 preparados com BES utilizando PLS-DA. Amostras de calibração: BES(▼), Outros B5(*). Amostras de validação (●).....	65
Figura 32 – Probabilidade de classificação de B5 preparados com BES utilizando PLS-DA. Amostras de calibração: BES (▼), Outros B5(*). Amostras de validação (●).....	67
Figura 33 – Previsão de amostras de B5 preparados com BMS utilizando PLS-DA. Amostras de calibração: BMS(▼), Outros B5(*). Amostras de validação (●).....	69
Figura 34 – Probabilidade de classificação de B5 preparados com BMS utilizando PLS-DA. Amostras de calibração: BMS(▼), Outros B5(*). Amostras de validação (●).....	70
Figura 35 – Previsão de amostras de B5 preparados com BMSebo utilizando PLS-DA. Amostras de calibração:BMSebo(▼), Outros B5(*). Amostras de validação(●).....	72

Figura 36 – Probabilidade de classificação de B5 preparados com BMSebo utilizando PLS-DA. Amostras de calibração: BMSebo(▼), Outros B5(*). Amostras de validação (●)....	74
Figura 37 – Previsão de amostras com 5% BMS utilizando PLS-DA. Amostras de calibração: B5 de BMS(▼), BX(*). Amostras de validação (●).....	76
Figura 38 – Probabilidade de classificação de amostras com 5% BMS utilizando PLS-DA. Amostras de calibração: B5 de BMS(▼), BX(*). Amostras de validação (●).....	78

Sumário

<i>1. Introdução</i>	1
<i>2. Revisão Bibliográfica</i>	5
2.1. Biodiesel.....	6
2.1.1. Definição e problema.....	6
2.1.2. Produção e matrizes energéticas utilizadas.....	8
2.2. Espectroscopia no infravermelho médio (MIR).	10
2.2.1 Espectroscopia no infravermelho médio (MIR) e suas aplicações..	10
2.2.2. Espectroscopia de Reflectância Total Atenuada	11
2.3. Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA)...	14
2.3.1. Quimiometria.	14
2.3.2. Análise de Componentes Principais (PCA).....	15
2.3.3. Análise por Quadrados Mínimos Parciais (PLS).....	17
2.3.4. Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA)	19
<i>3. Objetivos</i>	21
3.1. Objetivo Geral.....	22
3.2. Objetivos Específicos.....	22
<i>4. Procedimento Experimental</i>	23
4.1. Amostras	24
4.2. Espectros	27
4.3. Construção dos modelos	27
<i>5. Resultados e Discussões</i>	29
5.1. Tratamento dos espectros.....	30

5.2. Modelos PLS-DA.....	33
5.2.1. Classificação de biodieseis B100.....	37
5.2.2. Classificação de biodieseis B5.....	56
5.2.3. Classificação de biodieseis BX quanto ao teor de 5% de Biodiesel em diesel (B5).	75
9. <i>Conclusão</i>	80
10. <i>Referências Bibliográficas</i>	83

1. Introdução

1. Introdução

Ao longo dos anos a necessidade mundial em obter energia vem aumentando, a maior parte dessa energia é derivada de combustíveis fósseis como petróleo, carvão e gás natural. Porém, a reserva limitada de combustíveis fósseis chamou a atenção de pesquisadores para procura de combustíveis alternativos que pudessem ser produzidos a partir de biomassa [1]. A crescente demanda energética mundial indica que é necessário implantar novas fontes energéticas junto a matriz principal [2]. A figura 1 ilustra essa demanda energética até o ano de 2015, mostrando que ao longo dos anos há sempre um aumento na demanda de milhões de barris por dia de óleo equivalente.

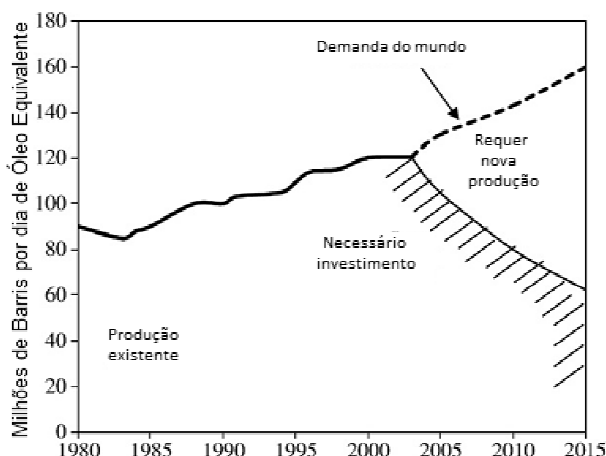


Figura 1 – Demanda energética mundial

Além disso, a instabilidade política dos países produtores de petróleo, elevando o preço do barril e as preocupações ambientais, resultou na necessidade da busca de fontes renováveis de energia. Entre essas fontes renováveis vêm se destacando, o bioetanol, óleos vegetais, gordura animal, energia eólica, e energia solar. [3][4][5].

Países em desenvolvimento, como o Brasil, que se destacam na disponibilidade de terras agricultáveis, clima favorável e disponibilidade de mão de obra, torna a produção de biocombustíveis vantajosa frente a outros países [6][7].

O biodiesel, proveniente de óleos vegetais, é um biocombustível dentre as alternativas existentes, que ficou muito atraente por causa de seus benefícios ambientais e vantagens em relação ao petrodiesel, como baixos níveis de emissão de monóxido de carbono e material particulado durante a combustão, a não liberação de compostos de

enxofre na atmosfera, melhores propriedades combustíveis como o número de cetano e menores emissões dos principais gases do efeito estufa, pois o gás carbônico produzido na sua combustão é quase que totalmente fixado durante o cultivo das oleaginosas [1].

Diante disso, o Brasil tem investido na construção de indústrias de produção de biodiesel, reforçada por políticas públicas, inserindo o biodiesel na matriz energética brasileira, de maneira gradual e progressiva, estimulado pelo Programa Nacional de Produção e Uso de Biodiesel (Lei nº 11.097).

Então, frente ao estabelecimento do mercado de biodiesel e sua crescente expansão, faz-se necessário o controle de qualidade ao longo de toda sua cadeia de produção e distribuição. No Brasil, essa fiscalização é feita pela Agência Nacional de Petróleo, Gás Natural e Biocombustíveis (ANP) que estabelece parâmetros baseados em normas internacionais como, por exemplo, da American Society for Testing and Materials (ASTM) e do Comitê Europeu de Normalização (CEN, do francês *Comité Européen de Normalisation*). Essas técnicas utilizam grandes quantidades de reagentes, além de serem laboriosas e utilizarem uma instrumentação analítica de valor muito elevado [8].

A determinação da concentração de etanol ou metanol em biodiesel é estabelecida pelas normas ABNT NBR 15343 e ASTM D6584, sendo realizada pela técnica de Cromatografia em Fase Gasosa. Por sua vez, a determinação do teor de glicerídeos em biodiesel é estabelecida pela norma ABNT NBR 15908, também usando Cromatografia em Fase Gasosa, [9].

Embora as técnicas cromatográficas sejam técnicas instrumentais bem estabelecidas, capazes de fornecer resultados de análises de biodiesel com precisão e exatidão, apresentam limitações devido à necessidade de preparo de amostras, consumo de reagentes, geração de resíduos e dificuldade quando aplicadas *in situ* para monitoramento da reação de transesterificação [10]. Neste contexto, análises espectroscópicas, especificamente no infravermelho, se apresentam como uma alternativa rápida, direta, não destrutiva, limpa e eficiente, podendo ser aplicadas *in situ* no controle de qualidade do biodiesel em todas as etapas da cadeia produtiva e fiscalização [11]. As potencialidades dessas técnicas tornam-se quase ilimitadas quando aliadas ao uso de métodos quimiométricos para extrair informações relevantes de dados complexos tornando-os mais simples de serem aplicados nas rotinas de laboratórios [12].

Diante da situação exposta, o Laboratório de Biocombustíveis da Universidade Federal de Uberlândia (LABIO-UFU), desde 2007, sob a coordenação do Professor Doutor

Manuel Gonzalo Hernandez-Terrones (*in memoriam*), vem desenvolvendo pesquisas para avaliar o potencial de diversas oleaginosas das regiões do cerrado e Norte de Minas Gerais e processos de produção e caracterização de biodieseis obtidos através da transesterificação via rotas metílica [13] e etílica [14], obedecendo a critérios estabelecidos pela ANP. As matérias-primas testadas no processo de produção de biodiesel foram: pinhão-mansão, pequi, baru, macaúba, girassol, canola, algodão, milho, soja, óleo residual e sebo bovino [15]. Em 2009 o Professor Doutor Waldomiro Borges Neto criou o Laboratório de Quimiometria do Triângulo (LQT) dando início à parceria com o LABIO-UFU em pesquisas visando à otimização multivariada dos parâmetros de produção de biodieseis [16], desenvolvimento de algoritmos matemáticos para determinação de coeficientes de expansão térmica e volumétricos para diferentes biodieseis [17][18], e posteriormente o desenvolvimento e validação de métodos baseados na aplicação de técnicas analíticas aliadas a métodos quimiométricos para controle de qualidade de biodieseis e suas misturas com diesel [19] [20]. Os grupos participam do programa de Pós-Graduação em Química do IQUFU e integram redes de pesquisas em Minas Gerais como: Rede de Pesquisa, Desenvolvimento e Inovação em Biocombustíveis de Minas Gerais; Rede Mineira de Química (RQ-MG) e Rede de Laboratórios de Biocombustíveis de Minas Gerais (RLBio-MG).

Fazendo parte das pesquisas desenvolvidas pelo LQT e inseridas no contexto da aplicação de técnicas analíticas aliadas a métodos quimiométricos para controle de qualidade de biodieseis, o presente trabalho visa à utilização de técnicas de espectroscopia no infravermelho médio (MIR), aliada ao método de classificação por análise discriminante por quadrados mínimos parciais (PLS-DA) para o desenvolvimento de método analítico para; classificação de amostras de biodieseis de acordo com sua matriz energética em relação aos biodieseis puros (B100) e de suas misturas em diesel com teor de 5% de biodiesel (B5) e modelos de classificação para verificação da conformidade quanto ao teor de B5.

2. Revisão Bibliográfica

2.1. Biodiesel

2.1.1. Definição e problema.

A ANP define o biodiesel como sendo um combustível composto de alquil ésteres de ácidos carboxílicos de cadeia longa, produzido a partir da transesterificação e ou/esterificação de matérias graxas, de gorduras de origem vegetal ou animal [9]. É um combustível biodegradável, proveniente de fonte renovável, apresenta uma maior viscosidade, lubricidade e ponto de fulgor, em relação ao biodiesel, melhorando assim, a segurança durante o transporte e o manuseio do produto [21].

Devido à busca de fontes renováveis de combustíveis causadas pelas preocupações com a diminuição das reservas de petróleo, preocupações ambientais e a busca por vantagens econômicas e sociais, a produção e utilização de biodiesel vem recebendo crescente atenção mundial.

A tentativa de diminuir a dependência dos combustíveis derivados do petróleo fez com que vários países começassem a utilizar inúmeras matérias-primas na produção de biodiesel e outras fontes renováveis de energia. O Brasil desde as décadas de 70 e 80, estimulados por planos de governo como o PROÁLCOOL desenvolveu tecnologias de ponta na produção de etanol. Recentemente a utilização de biomassa proveniente da cana de açúcar na produção de bioetanol combustível colocou o Brasil como um dos líderes na tecnologia de produção desse combustível. A consolidação deste mercado potencializou o interesse na produção de biodiesel aproveitando a diversidade nas fontes oleaginosas presentes nos diferentes biomas nacionais [22].

Dessa maneira o biodiesel foi introduzido na matriz energética brasileira em 13 de janeiro 2005, com a Lei nº 11.097, aprovada pelo Congresso Nacional. O biodiesel puro é denominado de B100, enquanto que suas misturas com o diesel de petróleo são genericamente denominadas de BX, onde X é a porcentagem v/v de biodiesel no diesel. A adição de biodiesel em diesel é obrigatória e a quantidade de biodiesel em diesel tem sido aumentada gradualmente, em 2008 essa adição era de 2% (v/v) (2% de biodiesel e 98% de óleo diesel) denominado B2, no ano de 2010 foi implantado o B5 (equivalente a 5% de biodiesel e 95% de óleo diesel) que é utilizada até os dias atuais. Essas misturas estão aprovadas para uso em todo território brasileiro e devem ser produzidas segundo as

especificações técnicas definidas pela Agência Nacional de Petróleo, Biocombustíveis e Gás Natural (ANP) [9]. De acordo com figura 2, o Brasil foi o segundo maior produtor de biodiesel do mundo em 2010, com produção de 2,4 bilhões de litros [23].

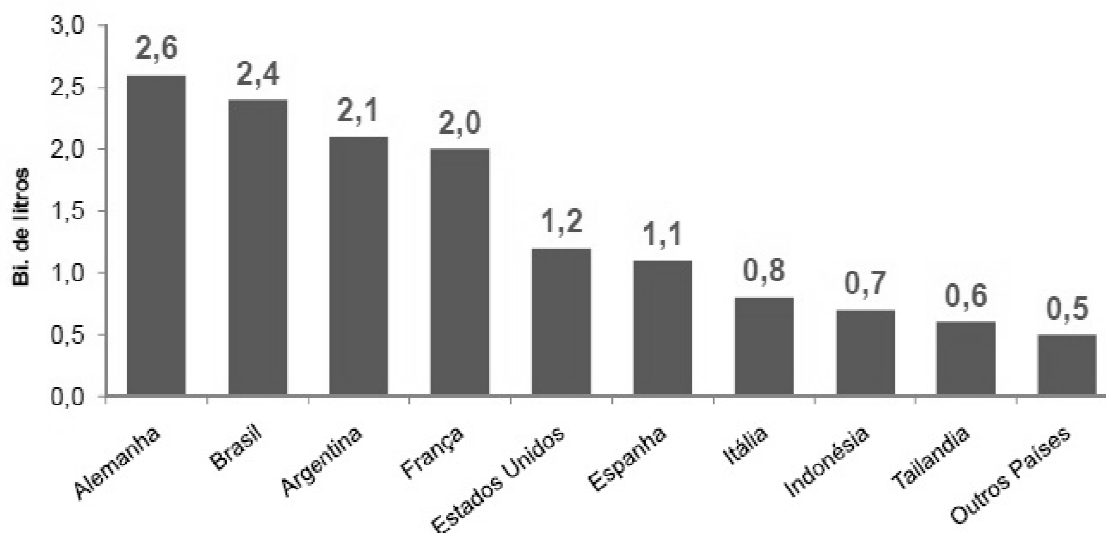


Figura 2 – Produção de Biodiesel em 2010

A principal fonte de matéria prima utilizada para a produção de biodiesel é o óleo de soja. A matéria prima da produção desse óleo, a soja, apresenta grande destaque na economia nacional, pois o Brasil é um dos principais exportadores dessa oleaginosa [24].

Já o uso de óleos vegetais provenientes de frituras, óleo residual, como matéria prima na produção de biodieseis poderia ser mais bem explorada e sua utilização deveria receber incentivos por parte do governo, uma vez que traria inúmeros benefícios para a sociedade, pois reduziria problemas relacionados ao seu descarte, e seu aproveitamento reduziria o custo quanto à matéria-prima, além de aumentar a produção de biodiesel [25].

Devido a sua baixa solubilidade em água, o óleo residual constitui como um fator negativo no que se refere à sua degradação em unidades de tratamento de esgotos por processos biológicos e, quando presentes em mananciais utilizados para abastecimento público, causam problemas no tratamento de água [26]. Pesquisas apontam que cada litro de óleo despejado no esgoto tem capacidade de poluir mais de 25 mil litros de água [27]

O óleo também interfere na passagem de luz na água, retarda o crescimento vegetal e interfere no fluxo de água, além de impedir a transferência do oxigênio para a água o que impede a vida nestes sistemas [28]. Além disso, agrava o efeito estufa, já que o contato da

água poluída pelo óleo ao desembocar no mar gera uma reação química que libera gás metano, um componente muito mais problemático que o gás carbônico [29].

De qualquer forma a utilização de biodiesel apresenta vantagens ambientais quando comparado ao óleo diesel por não conter enxofre e baixos níveis de emissão de material particulado e monóxido de carbono, redução na emissão de gases causadores de efeito estufa e ainda, o gás carbônico emitido na combustão de biodiesel é compensado quase que totalmente durante o cultivo das oleaginosas, que absorvem esse gás do meio ambiente [2].

2.1.2. Produção e matrizes energéticas utilizadas.

Os biodieseis podem ser produzidos a partir de óleos vegetais ou gorduras animais. Entre as matérias primas mais utilizada estão soja, algodão, palma, canola e óleo de fritura; entre as gorduras estão as bovina, suína e frango. A escolha da matéria-prima é fator preponderante nos custos da produção do biodiesel, podendo definir a viabilidade ou não do processo. Portanto, deve-se levar em conta aspectos como: a composição química do óleo extraído, o elevado rendimento energético por unidade de área e a produção abundante de óleos a custos reduzidos, a manutenção da produção de alimentos, priorizando a produção do biodiesel a partir das matérias-primas não alimentares, o baixo custo de produção e alta escala que pode levar a priorizar óleos ou gorduras residuais em relação a óleos refinados ou reciclados [30] [31].

De acordo com a ANP: o Brasil produziu até maio de 2013 1,15 bilhões de litros de Biodieseis e 72,30% da produção de biodiesel no Brasil entre Janeiro e Agosto de 2013 é proveniente do óleo de soja. Em segundo lugar, aparece o sebo bovino com 20,68% seguido do óleo de algodão com 2,83%. Já o óleo de fritura começa a despontar no cenário das matérias primas dos biocombustíveis, visto que não há necessidade de extração e refino desse óleo [32] e, uma estimativa da oferta de óleos residuais de frituras, revela um potencial de oferta no país superior a 30 mil toneladas por ano [33]. Considerando um rendimento de 80% na transesterificação do óleo residual, poder-se-ia atingir uma produção de 24 mil toneladas por ano de biodiesel, mas, atualmente apenas 3,0% do óleo residual descartado é reaproveitado [9]. Os alcoóis mais utilizados no processo de produção são o metanol e o etanol. O metanol apresenta maior reatividade que o etanol diminuindo o tempo e a temperatura para ocorrência da reação, porém é mais tóxico que o

etanol. Já o etanol é obtido de fontes renováveis, com produção bem consolidada no Brasil [34], mas promove uma maior dispersão do glicerol no biodiesel dificultando a separação das fases, necessitando de uma maior razão entre álcool e óleo para aumentar o rendimento da reação.

A produção de biodiesel em escala industrial utiliza uma razão molar óleo: álcool de 1:6, na presença de aproximadamente 0,40% (m/m) de catalisador. O excesso de álcool faz-se necessário para deslocar o equilíbrio estabelecido na reação em direção à formação de biodiesel [32]. A figura 1 mostra de forma generalizada a representação da reação de conversão do triglicerídeo em biodiesel [35]. Essa reação ocorre principalmente pela reação de transesterificação alcalina devido a sua facilidade de execução. O início da reação ocorre com a formação do alcóxido proveniente da base (KOH ou NaOH) e álcool e posteriormente, o ataque nucleofílico do mesmo sobre o sítio carboxílico do éster, conforme descrito no mecanismo representado na figura 3 [36].

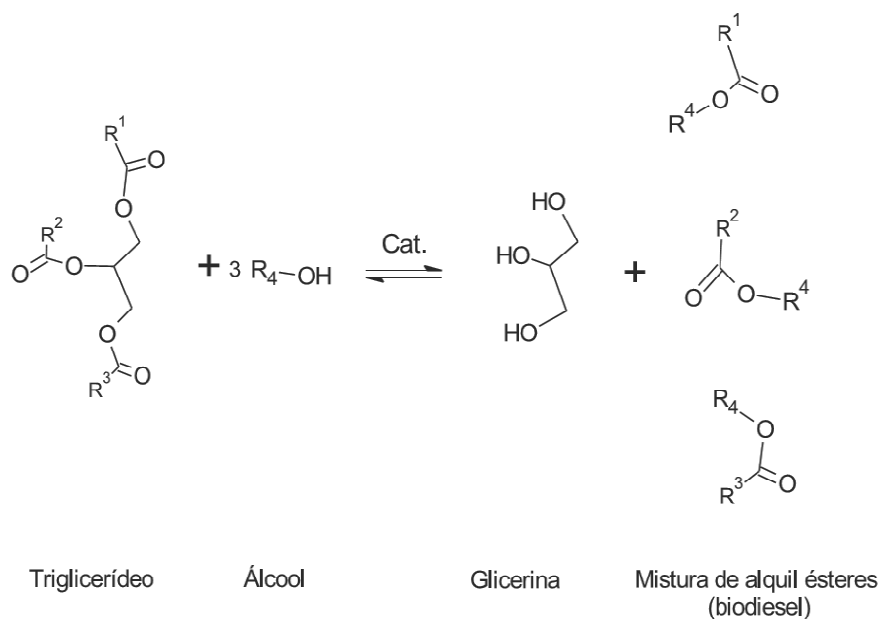


Figura 3 – Reação global de transesterificação.

2.2. Espectroscopia no infravermelho médio (MIR).

2.2.1 Espectroscopia no infravermelho médio (MIR) e suas aplicações.

As análises espectroscópicas são baseadas na medida da quantidade de radiação produzida ou absorvida por moléculas ou espécies químicas de interesse e podem ser classificadas de acordo com a região do espectro eletromagnético envolvida na medida [37]. A região espectral do infravermelho está situada antes da região do visível no espectro eletromagnético e abrange a radiação com números de onda no intervalo de aproximadamente 12800 a 10 cm^{-1} ou comprimento de onda de 780 a 100000 nm. O espectro no infravermelho médio compreende a radiação com números de onda de 4000 a 200 cm^{-1} [37], esses espectros são muito utilizados tanto para análises quantitativas como também para análises qualitativas [38]. A região do infravermelho próximo, de 4.000 a 14.000 cm^{-1} é também usada em determinação quantitativa em rotina de certas espécies como água, dióxido de carbono, enxofre e outros compostos utilizados na indústria e agricultura. Já no infravermelho distante, de 15.000 a 100.000 cm^{-1} , o uso tem sido na determinação de estruturas de espécies inorgânicas e organometálicas, com base em medidas de absorção. A tabela 1 apresenta os valores aproximados para essas regiões [37].

Tabela 1 – Regiões espectrais do infravermelho.

Região	Número de onda (ν), cm^{-1}	comprimento de onda (λ), nm	Frequência (ν), Hz
Próximo (NIR)	12.800 a 4.000	780 a 2.500	$3,8 \times 10^{14}$ a $1,2 \times 10^{14}$
Médio (MIR)	4.000 a 200	2.500 a 5.000	$1,2 \times 10^{14}$ a $6,0 \times 10^{12}$
Distante (FIR)	200 a 10	5.000 a 100.000	$6,0 \times 10^{12}$ a $3,0 \times 10^{11}$

A absorvância em uma frequência particular é característica de um grupo funcional presente no composto químico [39]. Por exemplo, a banda por volta de 1.655 cm^{-1} pode ser atribuída às vibrações de deformação axial da ligação C=C, característica de alquenos e a banda em cerca de 3.010 cm^{-1} às vibrações de deformação axial da ligação C-H da ligação dupla. Observa-se a presença de um pico de absorvância na faixa de 1.700 cm^{-1} a 1.750 cm^{-1} , proveniente da presença de carbonila (C = O) e a presença de bandas na região de

1.000 cm^{-1} a 1.300 cm^{-1} atribuída às vibrações de deformação axial da ligação C-O. Os ésteres metílicos possuem absorção com três bandas na região de 1.165 cm^{-1} a 1.265 cm^{-1} , ao passo que na mesma região para o éster etílico sobressaem duas bandas [37]. Isso oferece oportunidades analíticas quase que ilimitadas para muitas áreas de produção e controle de qualidade em vários setores industriais. Esse fato é consequência da velocidade e da facilidade desse tipo de instrumentação, pois as amostras não necessitam de tratamento prévio, são utilizadas em pequena quantidade, além disso, a técnica é não destrutiva [39].

Discutindo especificamente sobre os espectros de diesel, esses apresentam três regiões espectrais de absorção significativas, que são as regiões: entre 2.840 cm^{-1} e 3.000 cm^{-1} em que se localizam as vibrações de deformação axial das ligações C-H dos grupamentos de metila e metileno; bandas de intensidade intermediária na região de 1.300 cm^{-1} a 1.500 cm^{-1} proveniente das vibrações de deformação angular das ligações C-H do grupo metila e metileno; banda de na região de 720 cm^{-1} proveniente das vibrações de deformação angular assimétrica resultante das deformações C-H do grupamento metileno [37].

2.2.2. Espectroscopia de Reflectância Total Atenuada

A reflexão de radiação pode ser de quatro tipos: especular, difusa, interna e total atenuada. A espectroscopia de Reflectância Total Atenuada (ATR do inglês *Attenuated total reflectance*) baseia-se no fato de que quando um feixe de radiação passa de um meio mais denso (cristal de ATR) para um menos denso (amostra) ocorre uma reflexão. Consiste no contato da amostra com um elemento de reflexão interna construída com um material de alto índice de refração, normalmente o seleneto de zinco (ZnSe). A fração do feixe incidente que é refletida aumenta com o ângulo de incidência ocorrendo a reflexão total. Nesse processo, o feixe se comporta como se penetrasse um pouco no meio menos denso antes que a reflexão ocorra. A profundidade de penetração, que varia de uma fração até vários comprimentos de onda, depende do comprimento de onda da radiação incidente, dos índices de refração dos dois materiais e do ângulo do feixe incidente em relação à interface. A radiação penetrante é chamada onda evanescente. Se o meio menos denso absorve essa onda evanescente, ocorre atenuação do feixe nos comprimentos de onda das bandas de absorção [37].

Os espectros de ATR são semelhantes, mas não iguais aos espectros comuns de absorção. Em geral, enquanto as mesmas bandas são observadas, suas intensidades relativas diferem. As absorvâncias, embora dependam do ângulo de incidência, são independentes da espessura da amostra, uma vez que a radiação penetra apenas alguns micrômetros na mesma.

Uma das maiores vantagens da espectroscopia de ATR é que os espectros de absorção são obtidos rapidamente em uma grande variedade de tipos de amostras com um mínimo de preparação. Assim, linhas, fios, tecidos, fibras, pós, pastas, suspensões, polímeros, borrachas e outros materiais podem ser analisados pressionando as amostras sobre o cristal ATR [37].



Figura 4 – Espectrofotômetro FT-MIR com acessório de ATR horizontal.



Figura 5 – Acessório de ATR horizontal.

Apesar da espectroscopia no infravermelho médio fornecer informações químicas características de cada composto, os espectros podem conter sobreposição de sinais que dificultam a interpretação dos mesmos [7]. As informações contidas nos espectros obtidos são muito semelhantes e complexas impossibilitando a análise direta do analista fazendo necessária a aplicação de métodos quimiométricos que permitem a obtenção de informações qualitativas e quantitativas a partir dos complexos espectros no infravermelho. Devido à complexidade apresentada por suas matrizes, os combustíveis são estudados, frequentemente, por análise de dados multivariados [40], [41].

2.3. Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA)

2.3.1. Quimiometria.

A Quimiometria pode ser definida como a aplicação de métodos matemáticos e estatísticos assim como lógica formal, para o tratamento de dados químicos de forma a extrair uma maior quantidade de informação e otimização dos resultados analíticos [42].

Há algum tempo a obtenção de dados pelos analistas ocorria de forma lenta e a quantidade desses dados analisados era pequena. Com a evolução dos computadores e instrumentos computadorizados, a quantidade de dados obtidos aumentou de forma significativa, a obtenção desses dados tornou-se rápida [7] [42] e a quimiometria vai ao encontro dessa situação como uma ferramenta viável na resolução desses problemas.

Conforme o objetivo de estudo, a quimiometria pode ser dividida em várias áreas como: processamento de sinais analíticos, planejamento e otimização de experimentos, reconhecimento de padrões e classificação de dados, métodos de inteligência artificial, calibração multivariada, dentre outras [42].

Dentre as áreas apresentadas acima, as de maior destaque são: planejamento e otimização de experimentos na qual se busca encontrar quais as variáveis que mais afetam um determinado processo, assim como a interação entre essas variáveis [43]; calibração multivariada, que se busca estabelecer um modelo que relacione uma série de medidas (químicas ou espectrais) realizadas em amostras com uma determinada propriedade (concentração, por exemplo) [43]; reconhecimento de padrões e classificação de dados em que a partir de uma grande quantidade de informações (medidas químicas ou espectrais, por exemplo) sobre uma série de objetos, pretende-se encontrar agrupamentos de amostras (objetos) que são similares entre si e, assim, detectar tendências nos dados [43].

Em geral, estes métodos têm sido aplicados com sucesso na visualização dos dados, na classificação, na resolução de curvas multivariadas e na predição em química analítica, química ambiental, engenharia, investigação médica e na indústria [44] [45].

Nesse trabalho trataremos sobre a Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA, do inglês, *Partial Least Squares for Discriminant Analysis*) que é um método de reconhecimento de padrões e classificação de dados supervisionados (denominação que será discutida adiante) e para melhor entendimento será introduzido o

conceito de Análise de Componentes Principais (PCA, do inglês, *Principal Components Analysis*), que é utilizado como um método de ortogonalização de vetor e Quadrados Mínimos Parciais (PLS, do inglês *Partial Least Squares*) da qual o PLS-DA é uma variante.

2.3.2. Análise de Componentes Principais (PCA)

Na análise de componentes principais ocorre a projeção dos dados originais de grandes dimensões para dimensões menores. Essa projeção é realizada pela PCA em que o objetivo é transformar dados multivariados para que as informações importantes possam ser visualizadas mais facilmente [46] [47]. Nesse caso, a matriz \mathbf{X} é decomposta em duas matrizes menores de escores (\mathbf{T}) e de pesos (\mathbf{P}) e mais uma matriz de resíduos que indica a parte que não é modelada, como mostra a figura 6 e a equação 1.

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (1)$$

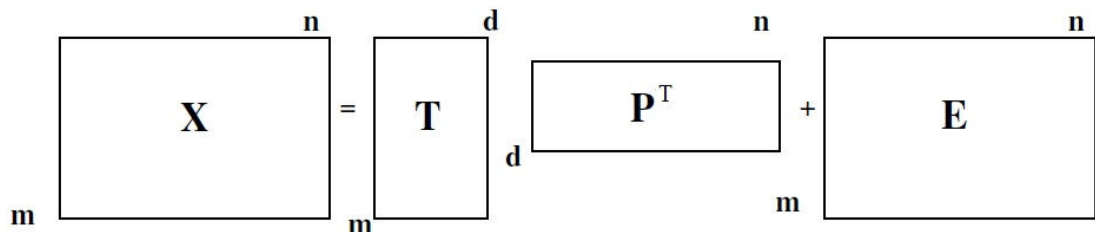


Figura 6 – Decomposição da matriz \mathbf{X} nas matrizes de escores (\mathbf{T}), pesos (\mathbf{P}) e resíduos (\mathbf{E}).

Neste caso, \mathbf{X} é a matriz original constituída de “m” linhas (amostras) e “n” colunas (variáveis); \mathbf{T} é a matriz de escores com “m” linhas e “d” colunas (número de componentes principais); \mathbf{P} é a matriz de pesos com dimensões “d” e “n” em que d é o numero de linhas e “n” o de colunas e \mathbf{E} é a matriz de resíduos com “m” linhas e “n” colunas.

Os escores representam as coordenadas das amostras no sistema de eixos formados pelos componentes principais. Cada componente principal é constituído pela combinação linear das variáveis originais e os coeficientes da combinação são denominados pesos. Matematicamente, os pesos são os cossenos dos ângulos entre as variáveis originais e os

componentes principais (PC), representando, portanto, o quanto cada variável original contribui para uma determinada PC. Já os escores, representam as relações de similaridade entre as amostras. A avaliação dos pesos permite entender quais variáveis mais contribuem para os agrupamentos observados no gráfico dos escores. Através da análise conjunta do gráfico de escores e pesos, é possível verificar quais variáveis são responsáveis pelas diferenças observadas entre as amostras [48].

O número de componentes principais a ser utilizado no modelo PCA é determinado pela porcentagem de variância explicada. Assim, seleciona-se um número de componentes de tal maneira que a maior porcentagem da variação presente no conjunto de dados originais seja capturada. A primeira componente principal (PC1) é traçada no sentido da maior variação no conjunto de dados; a segunda (PC2) é traçada ortogonalmente à primeira, com o intuito de descrever a maior porcentagem da variação não explicada pela PC1 e assim por diante [48]. Ou seja, cada componente principal subsequente descreve um máximo de variância contida nos dados, sendo que a primeira componente principal determina a maior variância dos dados. No segundo componente principal, há menos informação do que no primeiro e, assim, sucessivamente. A figura 7 representa graficamente duas componentes principais no caso de 2 variáveis (x_1 e x_2) em que a maior variância dos dados está sendo explicada pela primeira componente principal (PC₁).

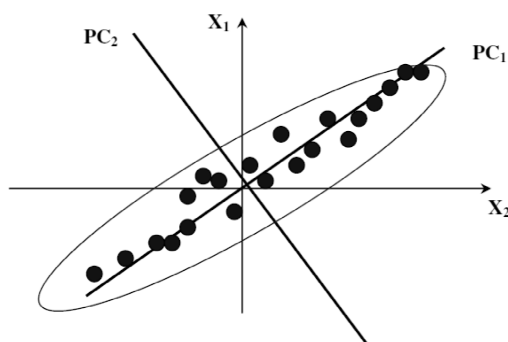


Figura 7 – Representação de duas componentes principais.

Geralmente a variância dos dados está sendo explicada na ordem decrescente, ou seja, $PC_1 > PC_2 > PC_3 > \dots > PC_A$, o que torna possível a visualização dos dados através dos escores de um componente versus outro. Os escores e os pesos podem ser calculados segundo o algoritmo NIPALS (do inglês *Nonlinear Iterative Partial Least Squares*) [47] ou SIMPLS (do inglês *Singular Value Decomposition*) [49].

2.3.3. Análise por Quadrados Mínimos Parciais (PLS)

Este método não requer um conhecimento exato de todos os componentes presentes nas amostras, podendo realizar a previsão de amostras mesmo na presença de interferentes. Porém, estes interferentes também devem estar presentes por ocasião da construção do modelo [47]. No PLS, a informação de **Y** é incorporada, de forma que cada PC do modelo sofre uma pequena modificação para buscar a máxima covariância entre **X** e **Y** e passa a receber a terminologia de Variável Latente (VL) conforme mostrado na Figura 8.

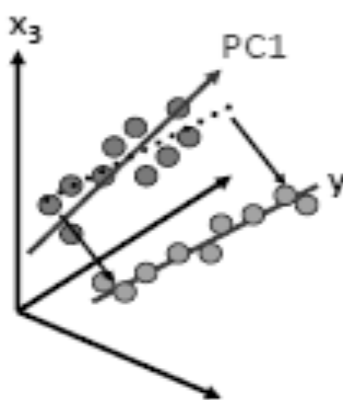


Figura 8 – Rotação no eixo do PC (Variável Latente).

O modelo PLS é obtido através de um processo iterativo, no qual se otimiza ao mesmo tempo a projeção das amostras sobre o(s) peso(s), para a determinação dos escores, e o ajuste por uma função linear dos escores da matriz **X** aos escores da matriz **Y** de modo a minimizar os desvios. Essa otimização simultânea ocasiona pequenas distorções nas direções dos pesos, de modo que, rigorosamente eles perdem a ortogonalidade, levando a pequenas redundâncias de informação. Porém, são essas pequenas redundâncias que otimizam a relação linear entre os escores, e estas distorções da ortogonalidade entre os componentes principais no PLS fazem com que não sejam mais componentes principais (que são ortogonais) e sim variáveis latentes (VL) [42][50].

Y é uma matriz de dimensão $(m \times z)$, em que z é o número de colunas de **Y**. Tem-se a decomposição de ambas as matrizes **X** de dimensão $(m \times n)$ e **Y** em suas matrizes de escores e pesos conforme as Equações 2 e 3:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E}_X \quad (2)$$

$$\mathbf{Y} = \mathbf{UQ}^T + \mathbf{E}_Y \quad (3)$$

onde \mathbf{X} é a matriz de dados (medida instrumental), \mathbf{Y} é a matriz de resposta (concentração, por exemplo), \mathbf{T} e \mathbf{U} são os escores para as duas matrizes de dados, \mathbf{P} e \mathbf{Q} são os respectivos pesos, \mathbf{E}_X e \mathbf{E}_Y são os respectivos resíduos, ou seja, as matrizes que contém a parte não modelada [42][50][51].

Uma relação linear é, então, estabelecida entre os escores de \mathbf{X} e os escores de \mathbf{Y} para cada VL ($h=1,2, \dots, A$) [52].

$$\mathbf{u}_h = \mathbf{b}_h \mathbf{t}_h \quad (4)$$

em que, \mathbf{b}_h é o vetor de coeficientes de regressão do modelo linear para cada VL, obtido através da Equação 5.

$$\mathbf{b}_h = \frac{\mathbf{u}_h^T \mathbf{t}_h}{\mathbf{t}_h^T \mathbf{t}_h} \quad (5)$$

O melhor modelo deverá minimizar as matrizes de resíduos \mathbf{E}_X e \mathbf{E}_Y e ao mesmo tempo, obter uma relação linear entre \mathbf{t} e \mathbf{u} .

Para a determinação do número correto de VL o método mais utilizado é o de validação cruzada, o qual se baseia na habilidade de previsão de um modelo construído por parte de um conjunto de dados seguido pela previsão do restante do conjunto desse dados, que é realizada pelo modelo construído. A validação cruzada pode ser realizada em blocos, ou seja, um número determinado de amostras é deixado de fora no processo de construção do modelo e a seguir essas amostras são previstas pelo modelo construído, ou ainda por um caso conhecido como “leave-one-out” (deixe uma fora por vez), em que uma amostra é deixada de fora no processo de construção do modelo e a seguir essa amostra é prevista pelo modelo construído.

2.3.4. Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA)

A Análise discriminante por mínimos quadrados parciais, PLS-DA, é utilizada na parte da quimiometria de reconhecimento de padrões e classificação de dados, que como descrito anteriormente, consiste em encontrar a similaridade entre amostras a partir de uma grande quantidade de informações. Similaridade essa que não seria possível distinguir apenas pela simples observação dos dados.

Esses métodos de reconhecimento de padrões e classificação de dados podem ser divididos em métodos supervisionados e não supervisionados.

Nos métodos não supervisionados os agrupamentos são feitos sem a necessidade do conhecimento prévio do conjunto de amostras ou das classes. Essas técnicas utilizam apenas medidas de propriedades com intuito de se observar agrupamentos naturais. Um dos exemplos desse método é o próprio PCA, apresentado anteriormente.

Os métodos supervisionados necessitam ter conhecimento inicial sobre as categoriais (ou classes) das amostras que irão ser modeladas, esses métodos utilizam esse conhecimento prévio do conjunto de amostras ou das classes que o compõem para identificar as amostras desconhecidas [53].

Entre os métodos supervisionados podemos citar a Modelagem Suave Independente Por Analogias de Classes (SIMCA, do inglês, *Soft Independent Modeling of Class Analogy*) [54] e o PLS-DA.

O SIMCA, utilizado na classificação de amostras, cujo conjunto de dados apresenta uma grande dimensionalidade, faz uso de componentes principais para localizar os objetos no espaço multidimensional, ou seja, o espaço multidimensional fica delimitado pela construção do modelo PCA para cada categoria de amostras. Uma amostra será classificada como pertencente a uma dada classe previamente modelada se possuir características que o permitam ser inserido no espaço multidimensional do modelo construído [53][55].

O método SIMCA obteve considerável sucesso como ferramenta de classificação [39], devido ao fato de que a grande maioria dos casos a variabilidade entre os grupos é maior que a variabilidade dentro dos grupos. Mas quando variabilidade dentro do grupo é maior que a variabilidade entre grupos o método SIMCA não conseguirá distinguir entre os grupos. Nesses casos, o PLS-DA tem sido uma boa alternativa [39].

A análise discriminante por mínimos quadrados parciais (PLS-DA) é uma variante do PLS utilizada quando a resposta Y é binária. O PLS é usado para encontrar relações fundamentais entre as duas matrizes X e Y , ou seja, trata-se de uma abordagem baseada na variável latente, para modelar as estruturas de covariância nestes dois espaços (X e Y).

Um modelo PLS procura a direção multidimensional no espaço da matriz X que explica a direção de máxima variação multidimensional no espaço da matriz Y . A regressão por mínimos quadrados parciais é particularmente indicada quando a matriz dos preditores (X) tem mais variáveis do que observações, e quando existe multicolinearidade entre os valores de X [56][57][58].

Ou seja, o método PLS-DA é um método supervisionado, desenvolvido a partir de algoritmos para mínimos quadrados parciais (PLS), é utilizado na quimiometria para classificação de amostra quando é necessária a redução de variáveis e não está claro se as diferenças entre grupos irão dominar a variabilidade total das amostras.

As classes, em um modelo PLS-DA, são definidas por conhecimento prévio do conjunto ou através de uma análise exploratória de dados, como por exemplo, PCA. O bloco Y em um modelo PLS-DA informa a classe a qual a amostra pertence, em que 1 é adotado para a classe de interesse que se deseja classificar, outro número para as demais classes e zero para a classes das amostras que se deseja prever.

Para calcular um valor limite que divide as classes evitando falsas classificações das previsões futuras pode se utilizar o teorema Bayesiano [59][60], que estabelece valores limite para definir um limite de confiança entre as classes.

Diferente do PCA, método não supervisionado de classificação, que não consegue prever novas amostras [61], os modelos PLS-DA são capazes de prever a classe de novas amostras.

Dessa forma, os modelos PLS-DA são capazes de classificar amostras de biodieseis e misturas de biodiesel/diesel em conformes e não conformes, quanto a composição, resíduos de processos, matéria prima utilizada em sua produção e outras possíveis adulterações usando espectroscopia MIR.

3.Objetivos

3.1. Objetivo Geral

Desenvolver metodologias analíticas empregando Espectroscopia no Infravermelho Médio aliada ao método de Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA) capazes de classificar amostras de biodieseis (B100) e biodieseis em misturas com diesel, quanto a matriz energética utilizada em sua produção, tanto metílico quanto etílico, provenientes de soja, soja residual e sebo. Além disso, classificar amostras de misturas biodieseis em dieiseis com diferentes teores (BX) em conforme ou não conforme quanto ao teor de B5.

3.2. Objetivos Específicos

Obtenção de espectros MIR de amostras dos biodieseis B100 (metílicos e etílicos) de soja, soja residual e sebo;

Obtenção de espectros MIR de amostras dos biodieseis B5 (metílicos e etílicos) de soja, soja residual e sebo;

Obtenção de espectros MIR de mistura biodieseis metílicos de soja (BMS) em diesel com diferentes faixas de concentração.

Construção dos modelos PLS-DA para classificar as amostras conforme sua matriz energética para B100 e B5.

Construção de modelo PLS-DA para classificar mistura de BMS em diesel com diferentes faixas de concentração em conformes e não conformes.

4. Procedimento Experimental

4.1. Amostras

Os biodieseis utilizados nesse trabalho são: Biodieseis Etílicos Residuais (BER), Biodieseis Metílicos Residuais, Biodieseis Etílicos de Soja (BES), Biodieseis Metílicos de Soja (BMS) e Biodieseis Metílicos de Sebo (BMSebo).

Todos os BMSebo e 20 amostras de BMS foram fornecidos pela usina de Biodiesel Caramuru, em Goiás, Brasil.

Já os BER, BMR, BES e mais algumas amostras de BMS foram produzidos no LQT pelas mestrandas Eloisa Guimarães e Hery Mitsutake sobre orientação e supervisão do Prof. Dr. Waldomiro Borges Neto e Prof. Dr. Douglas Queiroz Santos.

Os biodieseis metílicos e etílicos de soja foram sintetizados a partir de óleo refinado de soja, adquirido no comércio local. Os biodieseis foram produzidos pela reação de transesterificação de óleo soja com os álcoois metílico e etílico na razão de massa 5:1 óleo/álcool na presença de 1% (m/m) de KOH em relação ao óleo utilizado como catalisador. O tempo de reação para as rotas metílica e etílica foi o mesmo, aproximadamente 80 minutos. Após o tempo de reação, a fase glicerol foi removida e o biodiesel lavado com água quente (aproximadamente 80 °C) para retirar impurezas. A etapa de secagem do biodiesel foi realizada empregando evaporador rotativo por 1 hora a 88 rpm e 80 °C.

O biodieseis metílico e etílico residuais foram obtidos a partir de óleos de fritura domésticos, utilizando o mesmo processo descrito acima, diferenciando-se apenas pela filtração realizada antes do processo de síntese.

Para a classificação de biodieseis através da MIR e PLS-DA foi utilizado um total de 100 amostras de B100, separadas aleatoriamente em amostras de calibração (amostras de treinamento) e validação conforme a tabela 2.

Tabela. 2 – Amostras de calibração e previsão de B100

Biodieseis (B100)	Amostras	
	Calibração	Validação
Etílico Residual (BER)	15	5
Metílico Residual (BMR)	15	5
Etílico de Soja (BES)	15	5
Metílico de Soja (BMS)	15	5
Metílico de Sebo (BMSebo)	15	5

Para a classificação das misturas de biodiesel/diesel com teor de 5% de biodieseis (B5), as amostras foram preparadas adicionando 0,15g de biodiesel em 2,85 g de diesel, para obtenção de cada amostra de B5 utilizada. As pesagens foram realizadas em uma balança analítica (Sartorius, modelo BP211D). A solução de B5 foi homogeneizada em agitador Vortex.

Nota-se que na preparação das amostras, descrita acima, foi realizada a adição de biodiesel em diesel em 5% massa/massa devido à facilidade em trabalhar com medições de massa ao invés de volumes, mas como a legislação determina, as misturas de biodiesel em diesel devem ser feitas em 5% volume/volume, com um erro aceitável de 0,5%[62]. Porém, como a densidade aproximada dos biodieseis a 25°C é de 0,87g/cm³[63] e a do diesel é aproximadamente 0,85g/cm³[64], considerando as pequenas variações dessas densidades e cuidados tomados durante as medições, o erro aceitável de 0,5% volume/volume não é ultrapassado. Ou seja, considerando as massas utilizadas de biodieseis em dieseis nas misturas preparadas os teores reais ficam em torno de 5,1% a 5,3% volume/volume.

Então foi utilizado um total de 100 amostras de B5, separadas aleatoriamente em amostras de calibração e validação conforme a tabela 3.

Tabela. 3 – Amostras de calibração e previsão de B5

Biodieseis (B5)	Amostras	
	Calibração	Validação
Etílico Residual (BER)	15	5
Metílico Residual (BMR)	15	5
Etílico de Soja (BES)	15	5
Metílico de Soja (BMS)	15	5
Metílico de Sebo (BMSebo)	15	5

Já para a classificação de misturas de biodieseis metílico de soja em diesel com diferentes teores, foram utilizadas amostras de B100 e B5 de BMS preparadas e utilizadas anteriormente. As amostras de BX de BMS com teores acima e abaixo de 5%, foram preparadas pesando 0,15g de BMS e variando-se a massa de 2,85 g de diesel, de forma a se obter os BMS com teores acima de B5 (B6 a B10) e abaixo de B5 (B4 a B1). A proporção utilizada também foi massa/massa, como realizada para as amostras de B5, situação justificada anteriormente. As pesagens foram realizadas em uma balança analítica (Sartorius, modelo BP211D).

Ao final foram obtidas 65 amostras, separadas aleatoriamente em amostras de calibração e validação conforme a tabela 4.

Tabela. 4 – Amostras de calibração e previsão de BMS com diferentes teores

BMS	Amostras	
	Calibração	Validação
B5	21	11
Diferentes teores (BX) (abaixo de 5%, acima de 5% e Diesel puro)	22	11

4.2. Espectros

Os espectros 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 (ATR) com cristal de ZnSe, resolução de 4 cm^{-1} e 16 varreduras. Os dados dos espectros foram pré-processados utilizando o software Matlab 6.1 (MathWorks Inc.) e PLS Toolbox 3.5 (Eigenvector Research) através de rotina desenvolvida no laboratório. Inicialmente foi calculado o espectro médio para cada amostra a partir dos espectros em quintuplicata. As linhas de base espectrais foram corrigidas empregando o método baseline e as partes mais ruidosas dos espectros foram eliminadas.

Para eliminar a flutuação da linha de base, que influencia os resultados finais, pois os sinais medidos pelo instrumento consistem de sinal verdadeiro e ruído aleatório, foi utilizada a correção da linha de base (baseline) que faz a remoção matemática das fontes de dados indesejáveis. As extremidades dos espectros utilizados, que apresentam ruídos e dados desnecessários à matriz de dados, foram eliminadas. Uma análise exploratória de dados em busca da detecção de amostras anômalas (Outliers) foi realizada utilizando a PCA, as 5 amostras anômalas foram substituídas, por outras amostras da mesma origem e mesma classe.

4.3. Construção dos modelos

Foram construídos, 5 modelos para classificação de biodieseis B100 e 5 modelos para a classificação de biodieseis B5 quanto ao tipo de biodieseis presentes nas amostras. Ou seja, modelos de classificação para os biodieseis puros, que venham a classificar as amostras de previsão de B100 em BER, BMR, BES, BMS e BMSebo, modelos para a classificação de mistura de 5% de biodiesel em diesel, que venham a classificar as amostras de B5 em BER, BMR, BES, BMS e BMSebo.

Foi criado um modelo para classificação de amostras de BMS conforme o seu teor de biodieseis em conformes e não conformes. Ou seja, os modelos classificam as amostras de previsão que apresentam o teor 5%, sendo essas conformes e as demais em não conformes.

Para a classificação de biodiesel B100 e B5 foram utilizadas 75 amostras para calibração e 35 amostras para previsão, como mostra as tabelas 2 e 3, e para a classificação das amostras de BMS quanto ao teor de 5% foram utilizados 43 amostras para calibração e 22 amostras para previsão conforme tabela 4.

Todos os modelos foram construídos utilizando pré-processamento de dados centrados na média e algoritmo de regressão SIMPLS. A matriz **Y** foi construída atribuindo-se valor zero para as amostras de validação, 1 para as amostras de interesse e dois para outras amostras.

Outro ponto importante na construção dos modelos é a determinação do número de variáveis latentes. Como as variáveis originais são substituídas pelas novas variáveis, variáveis latentes, que são usadas para representar o conjunto de dados, é preciso avaliar quantas variáveis latentes são necessárias para representar o conjunto de dados.

Para isso, deve-se analisar a porcentagem de variância capturada, ou seja, a quantidade de dados originais que está sendo explicada em função do número de variáveis latentes. Essas informações são fornecidas pelo programa utilizado. A qual foram analisadas, e a quantidade determinada levando-se em conta que; quanto maior a quantidade de variáveis latentes, menor é o erro na classificação, pois maior a quantidade de informação capturada. Porém, usando uma grande quantidade de variáveis latentes, pode ocorrer a incorporação de dados desnecessários e ruídos ao modelo, comprometendo a classificação.

Antes da construção dos modelos foi realizada uma análise exploratória dos dados utilizando a PCA em busca de *outlier* (amostra com erros) e uma pré classificação para verificar o comportamento dessas amostras. As amostras em que foram constatadas erros foram substituídas por outras da mesma classe, apenas a amostra 55 de BMS foi excluída pra a construção do modelo de classificação de BMSebo.

5. Resultados e Discussões

5.1. Tratamento dos espectros

Os espectros puros, de biodieseis B100, B5 e de BX de BMS, antes do pré-processamento são apresentados respectivamente pelas figuras 9, 11 e 13. Pode-se observar a variação da linha de base e a extremidade ruidosa e sem sinal analítico de interesse.

Após a correção de linha de base, foi fixada a faixa de espectros para a construção do modelo entre 650 a 3082 cm^{-1} para os espectros de B100, 680 a 3250 cm^{-1} para os espectros de B5 e de BX de BMS, como mostrado nas figuras 10, 12 e 14.

É possível observar apenas pequenas diferenças no perfil espectral dessas amostras. Assim, o uso de ferramentas quimiométricas se faz necessário para construir modelos de classificação para diferenciar os tipos de biodieseis.

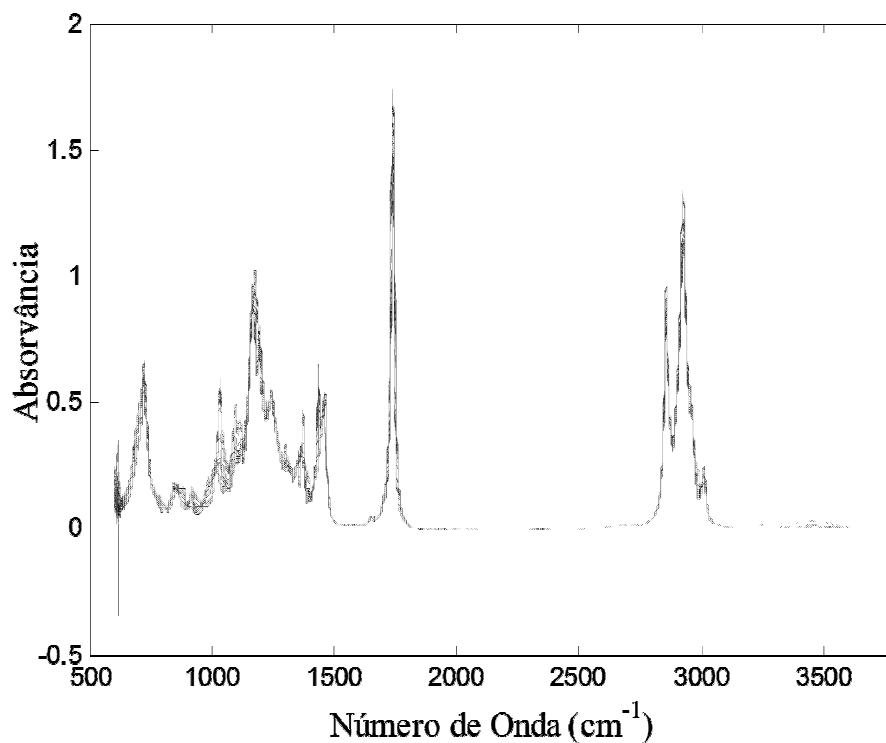


Figura 9 – Espectros originais de B100 sem correção de linha de base e sem cortes.

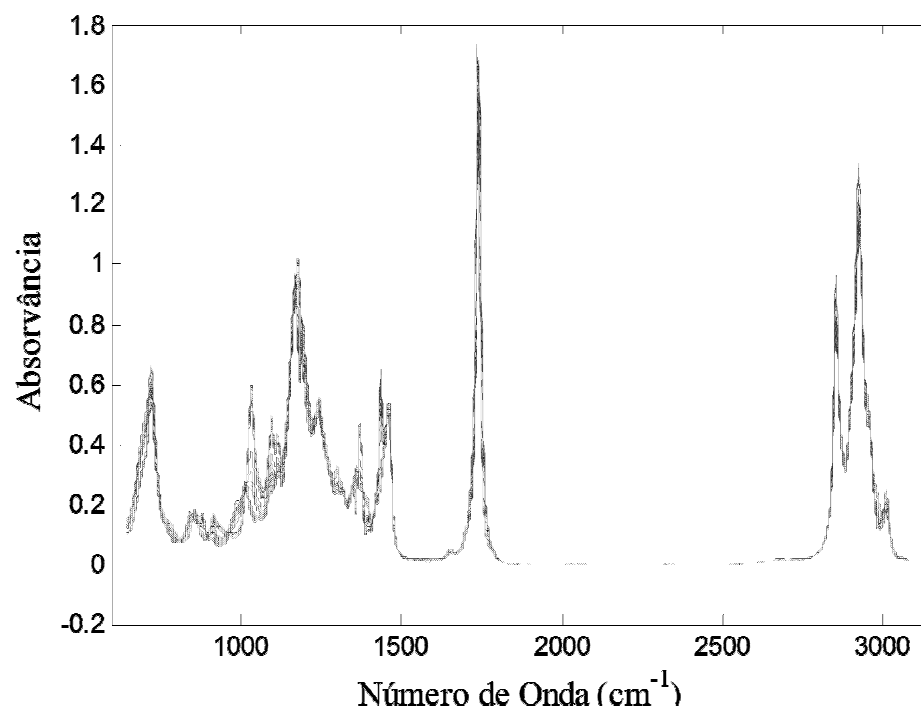


Figura 10 – Espectros de B100 após correção da linha de base por baseline e cortes.

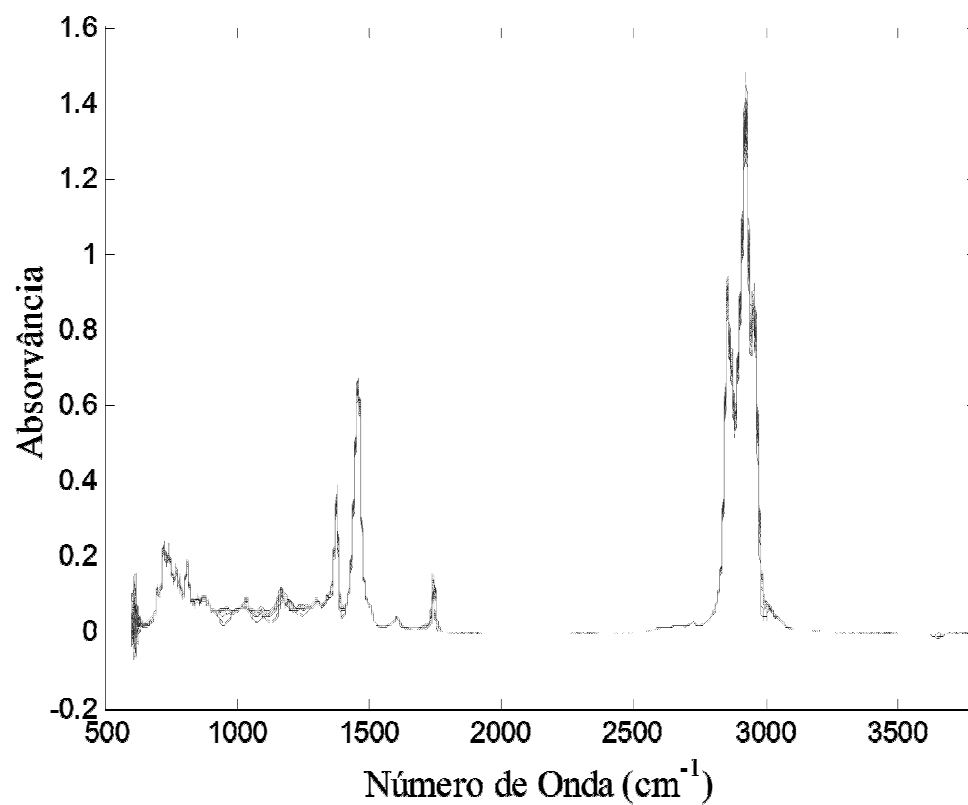


Figura 11 – Espectros originais de B5 sem correção de linha de base e sem cortes.

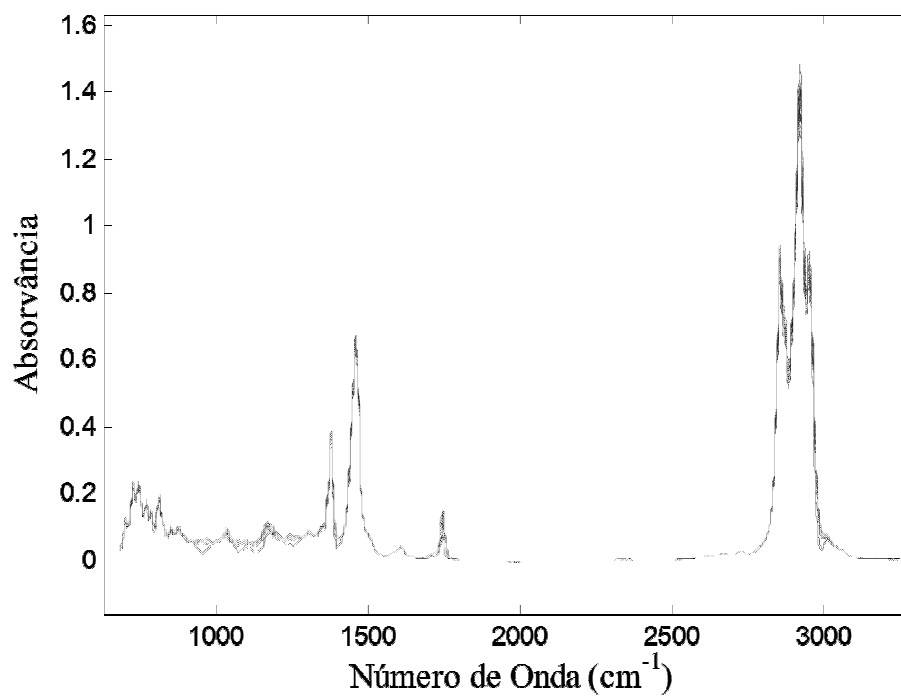


Figura 12 – Espectros de B5 após correção da linha de base por baseline e cortes.

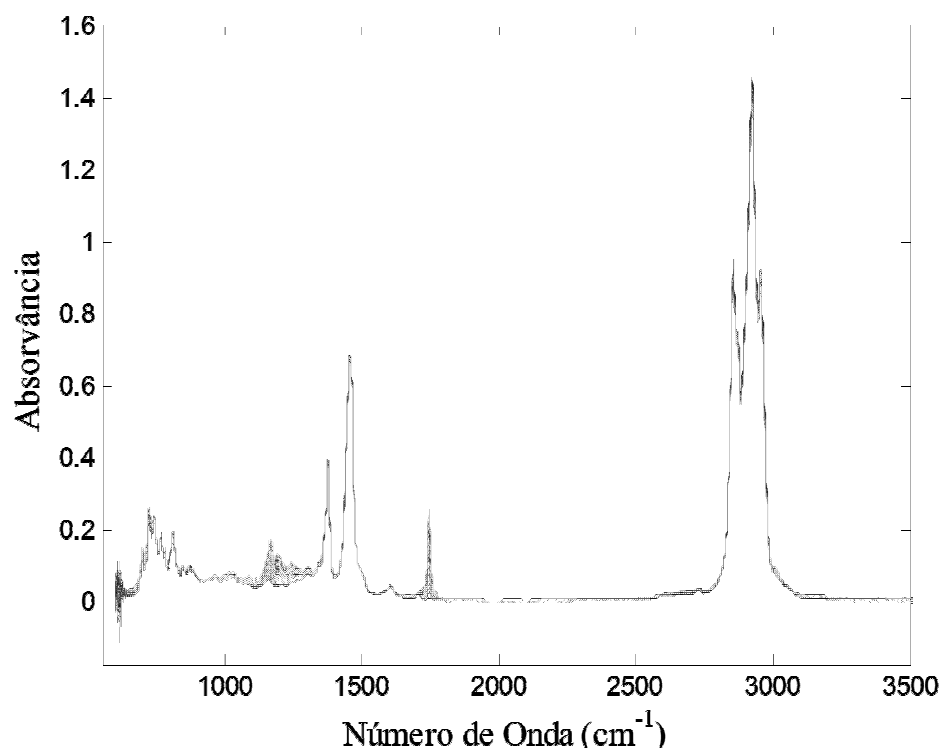


Figura 13 – Espectros originais de BMS com diferentes teores sem correção de linha de base e sem cortes.

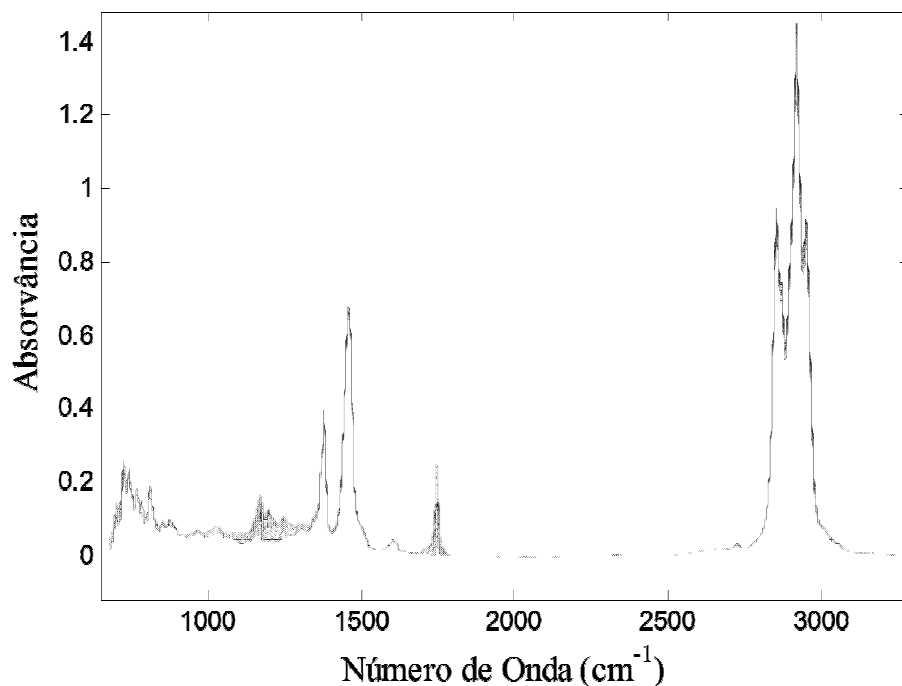


Figura 14 – Espectros de BMS com diferentes teores após correção da linha de base por baseline e cortes.

5.2. Modelos PLS-DA

Os modelos PLS-DA fornecem dados de estimativa e de probabilidade das amostras analisadas pertencerem ou não a classe de interesse. Eles são apresentados na forma de gráficos de estimativa e probabilidade, agrupando as amostras que são semelhantes, permitindo uma separação visual dos grupos [26]. Fornecem também tabelas com os valores das estimativas e probabilidade para essas classificações.

Para estabelecer limite de confiança entre as classes presentes, para os dados de estimativa, foi utilizada uma função presente no PLS Toolbox 3.5 (Eigenvector Research) baseada no teorema Bayesiana [31,32]. Esses limites são as linhas tracejadas presentes nos gráficos de estimativa. Valores estimativa acima desses limites indicam que a amostra pertence à classe modelada de interesse em quanto as amostras que apresentam valores abaixo desses limites pertence a outras classes de amostras, mas de forma alguma a classe modelada de interesse.

A figura 15 apresenta as estimativas de previsão das amostras. Um valor próximo de zero indica a nova amostra não pertence à classe de interesse; um valor de um indica

que a amostra pertence à classe de interesse. Os agrupamentos indicam os conjuntos de amostras semelhantes, e nesse trabalho as amostras representadas em vermelho (▼), em forma de triângulo invertido, representam as amostras dos biodieseis de interesse (Classe 1) utilizadas para calibração todas agrupadas acima do limite de confiança, linha vermelha tracejada. As amostras de cor verde na forma de estrelas (✱) são dos demais biodieseis do conjunto de calibração (classe 2), agrupados na parte de baixo do gráfico. Já as amostras representadas por círculos pretos (●) são o conjunto teste de previsão (amostras de validação), e são excluídas da construção do modelo (classe 0), a sua separação é baseada em suas semelhanças com a classe de calibração, as amostras na parte superior do gráfico pertencem a classe de interesse em vermelho, já as outras, encontradas na parte inferior pertencem a outra classe de biodiesel e não pertencem a classe de interesse.

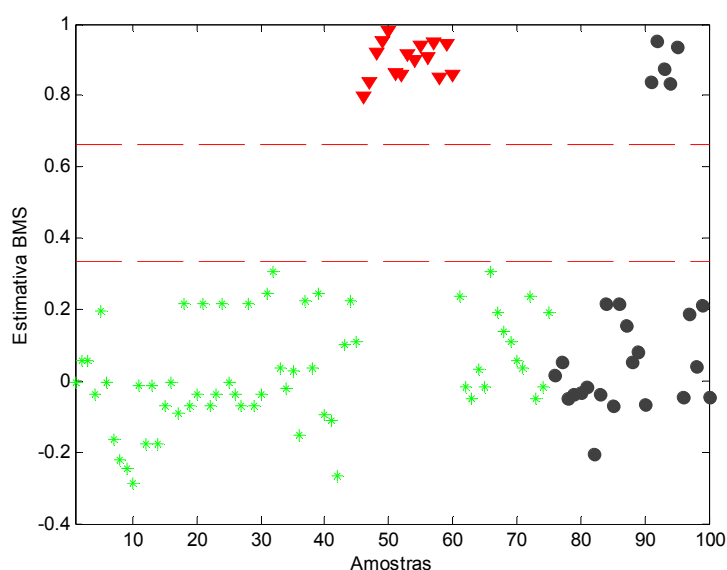


Figura 15 – Gráfico de estimativas de Previsão de amostras geradas pelo modelo PLS-DA. Amostras de calibração: Classe de interesse(▼), Outras classes (✱). Amostras de validação (●).

A tabela 5 quantifica a estimativa de previsão para as amostras de calibração e previsão, mostradas graficamente pela figura 15. Quanto mais esses valores aproximarem de 1 maior a semelhança com as amostras de interesse. Esses valores também podem ser comparados com o valor do limite de confiança que é gerado pelo modelo, representado graficamente pelas linhas vermelhas, quantificando essa separação.

Tabela 5 – Valores de estimativa previstos para amostras de interesse.

Amostras de Calibração		Amostras de Previsão	
Amostras	Estimativa	Amostras	Estimativa
46	0.79661	91	0.83857
47	0.83857	92	0.95396
48	0.91944	93	0.87587
49	0.95356	94	0.83576
50	0.97982	95	0.93751
51	0.86420		
52	0.85797		
53	0.91467		
54	0.89835		
55	0.93989		
56	0.90902		
57	0.94793		
58	0.85111		
59	0.94652		
60	0.86070		

O modelo também calcula a probabilidade de uma amostra pertencente para a classe que está sendo modelada [26], a figura 16, mostra a classificação visual da probabilidade das amostras pertencerem a classe de interesse. Acima, agrupadas em 1, encontra-se as amostras de calibração da classe de interesse, em vermelho, e o conjunto de previsão, em preto, classificados como pertencentes a classe de interesse. Abaixo as amostras de calibração e previsão que não pertencem à classe de interesse.

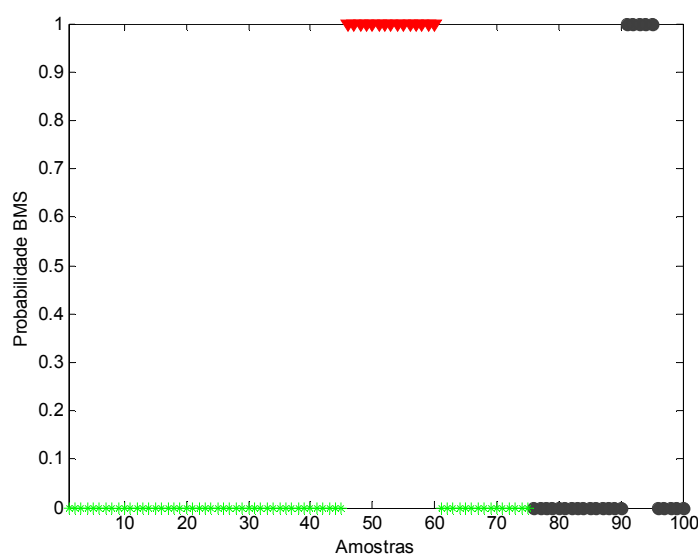


Figura 16 – Gráfico de Probabilidade de amostras geradas pelo modelo PLS-DA. Amostras de calibração: Classe de interesse(▼), Outras classes (*). Amostras de validação (●).

Já a tabela 6 quantifica a probabilidade das amostras de classificação e previsão pertencerem à classe de interesse.

Tabela 6 – Valores de probabilidade para amostras de interesse.

Amostras de Calibração		Amostras de Previsão	
Amostras	Probabilidade	Amostras	Probabilidade
46	99.9997	91	100.000
47	100.000	92	100.000
48	100.000	93	100.000
49	100.000	94	100.000
50	100.000	95	100.000
51	100.000		
52	100.000		
53	100.000		
54	100.000		
55	100.000		
56	100.000		
57	100.000		
58	100.000		
59	100.000		
60	100.000		

A seguir serão apresentados em três partes os resultados e as discussões obtidas através dos modelos que classificam as amostras de; B100 de acordo com sua matriz energética, amostras de B5 de acordo com sua matriz energética e que classificam as amostras de mistura biodiesel/diesel preparadas com diferentes teores (BX) de BMS em conformes e não conformes de acordo com a legislação vigente, ou seja, se apresentam ou não teor de 5% de biodiesel em diesel (B5).

5.2.1. Classificação de biodieseis B100.

5.2.1.1. Biodiesel Etílico Residual (BER)

Para os modelos de classificação de BER, foram utilizados 6 variáveis latentes com variância explicada 99.90% para a matriz de dados e 91,28 para a matriz de resposta .

No gráfico de previsão para BER (figura 17), observa-se o agrupamento das amostras de calibração para BES, em vermelho, e de previsão (amostras de validação), agrupamento de cinco amostras em preto de BER. Todos eles acima do limite de confiança, estabelecido em 0.6080 (representado pela linha tracejada na figura 17), se distinguindo das demais amostras e aproximando-se do valor um. Um valor próximo de um indica que a amostra pertence à classe de interesse, no caso BES. Todas as amostras de BER foram classificadas corretamente.

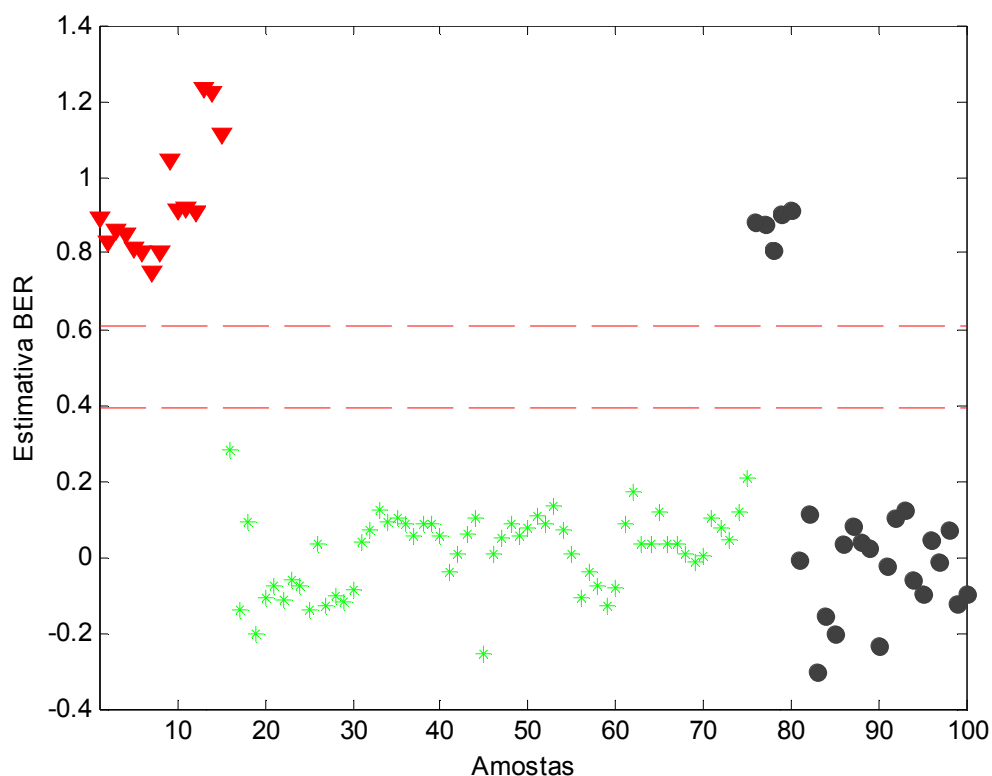


Figura 17 – Previsão de amostras de BER utilizando PLS-DA. Amostras de calibração: BER (▼), Outros biodieseis (*). Amostras de validação (●).

A tabela 7 apresenta os valores de estimativa para essas amostras de BER, onde se quantifica a representação gráfica da figura 17. Nota – se os valores previstos das estimativas das amostras de calibração e previsão próximos a um, indicando que pertencem à classe de interesse, BER. Todos valores de estimativa, estão acima do limite de confiança estabelecido

Tabela 7 – Valores previstos para amostras B100 de BER

Amostras de Calibração		Amostras de Previsão	
BER		BER	
Amostras	Estimativa	Amostras	Estimativa
1	0.89211	76	0.88172
2	0.82688	77	0.87515
3	0.85896	78	0.80799
4	0.85248	79	0.90301
5	0.81472	80	0.91557
6	0.80171		
7	0.74829		
8	0.80358		
9	1.04228		
10	0.91372		
11	0.91990		
12	0.90791		
13	1.23316		
14	1.22333		
15	1.11438		

O gráfico de probabilidade, gerados pelos modelos PLS-DA, das amostras de previsão pertencer à classe BER deixa claro essa classificação. As amostras de BER, tanto de calibração como de previsão estão todos agrupados em um, enquanto as outras amostras agrupadas em zero, como é mostrado na figura 18.

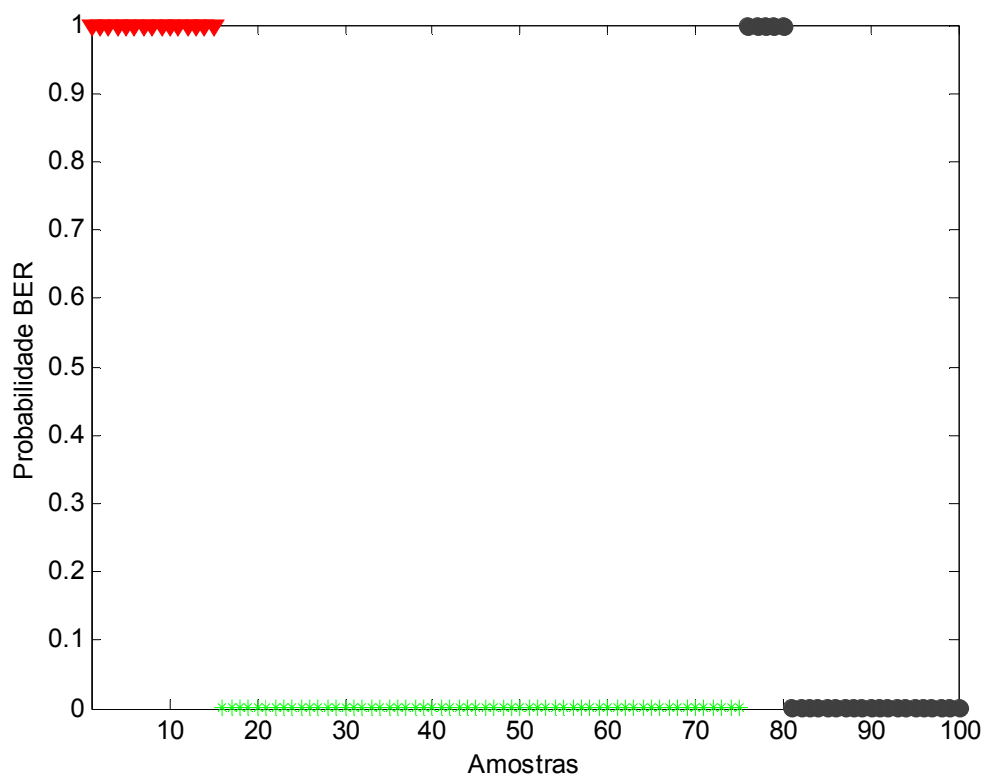


Figura 18 – Probabilidade de classificação de amostras de BER utilizando PLS-DA. Amostras de calibração: BER (▼), Outros biodieseis(*). Amostras de validação (●).

A tabela 8 quantifica a representação visual apresentada pela figura 18. Nesse caso a representação visual é tão clara para a classificação das amostras de previsão que dúvidas quanto a probabilidades de serem ou não de amostras de BER são inexistentes.

Tabela 8 – Valores de probabilidade para amostras B100 de BER

Amostras de Calibração		Amostras de Previsão	
BER		BER	
Amostras	Probabilidade	Amostras	Probabilidade
1	100.000	76	100.000
2	100.000	77	100.000
3	100.000	78	100.000
4	100.000	79	100.000
5	100.000	80	100.000
6	100.000		
7	100.000		
8	100.000		
9	100.000		
10	100.000		
11	100.000		
12	100.000		
13	100.000		
14	100.000		
15	100.000		

O modelo classificou com 100% de acerto as amostras de BER.

5.2.1.2. Biodiesel Metílico Residual (BMR)

Para os modelos de classificação de BMR, foram utilizados 6 variáveis latentes com variância explicada 99.87% para a matriz de dados e 82,88 para a matriz de resposta .

A figura 19 mostra a separação das amostras de calibração de BMR, em vermelho dos demais conjuntos de amostras de calibração, em verde, e a correta classificação das cinco amostras de previsão, em preto, presentes na parte superior da figura, todas elas acima do limite de confiança estabelecido em 0,5134, representado pela linha tracejada em vermelho.

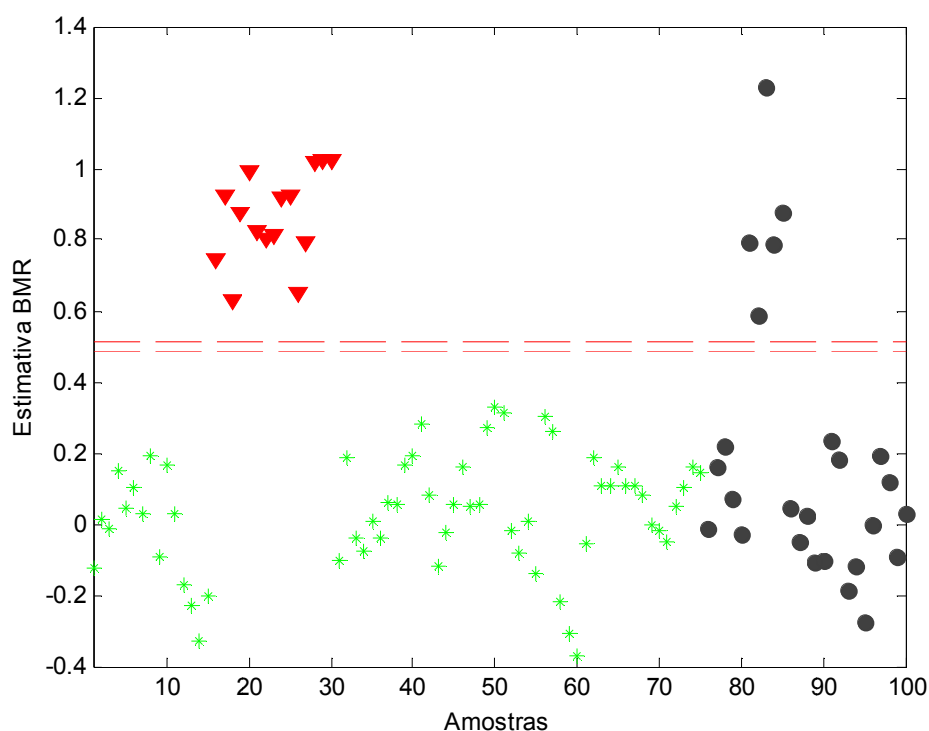


Figura 19 – Previsão de amostras de BMR utilizando PLS-DA. Amostras de calibração: BMR (▼), Outros biodieseis (*). Amostras de validação (●).

A tabela 9 quantifica os valores de estimativa previstos, todos acima do limite de confiança de 0,5134 e aproximando-se de um. Porém se compararmos esses valores com os valores das estimativas de previsão dos modelos de classificação de BER (tabela 7) nota-se uma maior variação desses valores de estimativas. Situação facilmente observada pela comparação dos gráficos de estimativa (figuras 17 e 19).

Isso ocorre por diferenças existentes dentro da própria classe, e não compromete a classificação das amostras, pois não há classificações incorretas. Essas diferenças podem ter origem a partir da matéria prima utilizada, no caso óleo residual, pois podem conter resíduos de diferentes tipos de gordura animal, diferentes tipos de óleos vegetais e diferentes alterações estruturais nos triésteres presentes devido ao diferentes aquecimentos a que foram submetidos.

Possivelmente o modelo PLS-DA seria sensível o suficiente para separar amostras de uma mesma classe de biodiesel quanto a essas diferenças.

Tabela 9 – Valores previstos para amostras B100 de BMR

Amostras de Calibração		Amostras de Previsão	
BMR		BMR	
Amostras	Estimativa	Amostras	Estimativa
16	0.74253	81	0.79359
17	0.92306	82	0.58750
18	0.62928	83	1.23030
19	0.87550	84	0.78593
20	0.99227	85	0.87550
21	0.82121		
22	0.80274		
23	0.81510		
24	0.91645		
25	0.92306		
26	0.64799		
27	0.79255		
28	1.01633		
29	1.02476		
30	1.02304		

Quando é apresentada a probabilidade das amostras pertencerem à classe dos BMR fica evidente a classificação, como é apresentado na figura 20.

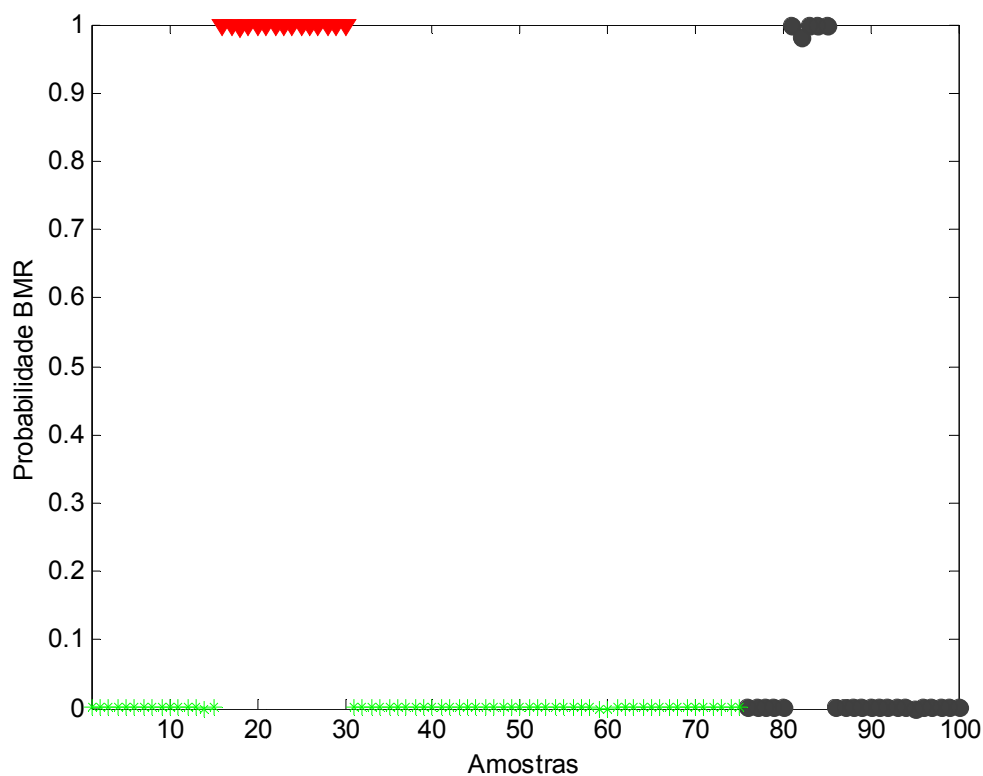


Figura 20 – Probabilidade de classificação de amostras de BMR utilizando PLS-DA. Amostras de calibração: BMR (▼), Outros biodieseis(*). Amostras de validação (●).

Apesar da variação dos valores de probabilidade não serem todos de 100% como ocorrido nos modelos de classificação dos BER (tabela8), a probabilidade das amostras de previsão pertencer a classe de interesse, BMR é superior a 98,2% como mostrado na tabela 10.

O cálculo de probabilidade gerada pelo modelo PLS-DA facilita a visualização e a tomada de decisão das amostras de previsão pertencer ou não a a classe de interesse.

Tabela 10 – Valores de probabilidade para amostras B100 de BMR

Amostras de Calibração		Amostras de Previsão	
BMR		BMR	
Amostras	Probabilidade	Amostras	Probabilidade
16	99.9944	81	99.9991
17	10.0000	82	98.2010
18	99.6347	83	100.000
19	99.9999	84	99.9988
20	100.000	85	99.9999
21	99.9996		
22	99.9993		
23	99.9996		
24	100.000		
25	100.000		
26	99.8255		
27	99.9990		
28	100.000		
29	100.000		
30	100.000		

As amostras de interesse de BMR foram classificadas com 100% de acerto, não ocorrendo nenhuma classificação errônea.

5.2.1.3. Biodiesel Etilico de Soja (BES)

Foram utilizadas 5 variáveis latentes para os modelos de classificação de BES, com variância explicada 99.86% para a matriz de dados e 86,56% para a matriz de resposta .

A figura 21 apresenta a estimativa de previsão para amostras de BES. Ela mostra a separação das amostras de calibração, em vermelho, e de previsão, 5 pontos em preto, todos acima do limite de confiança estabelecido em 0,5137, prevendo corretamente todas as amostras pertencentes a classe do BES.

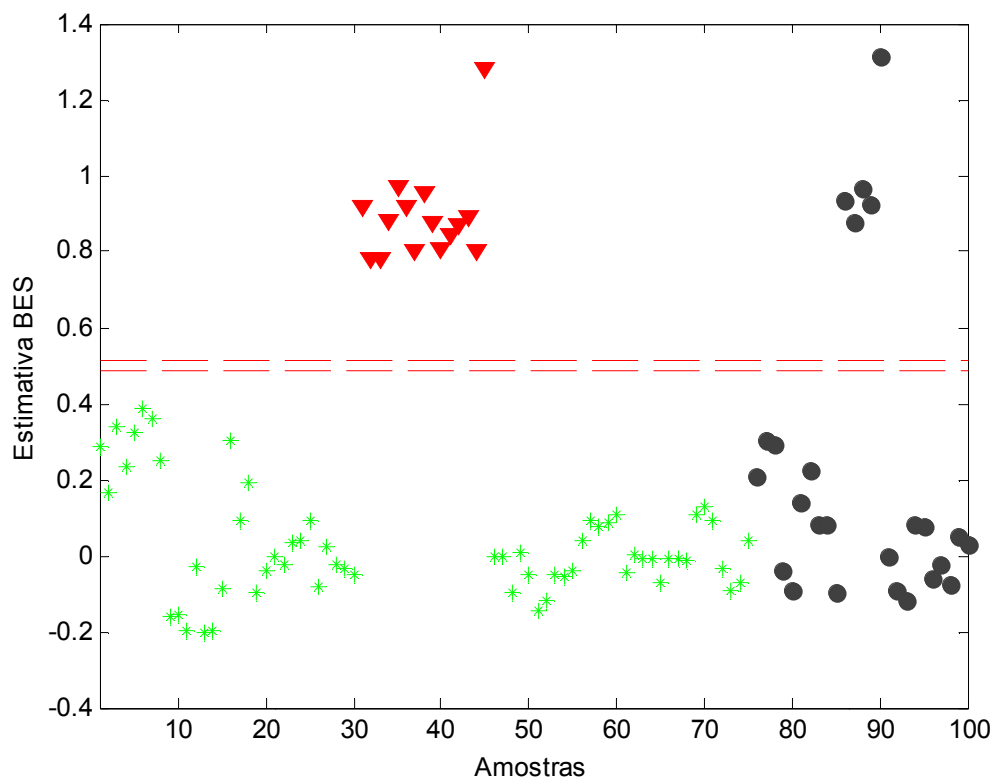


Figura 21 – Previsão de amostras de BES utilizando PLS-DA. Amostras de calibração: BES (▼), Outros biodieseis(*). Amostras de validação (●).

A tabela 11 apresenta os valores estimados para as amostras pertencerem ou não a classe dos BES, nota-se esses valores variando em torno de um, indicando que essas amostras pertencem à classe de interesse, BES.

Tabela 11 – Valores previstos para amostras B100 de BES.

Amostras de Calibração		Amostras de Previsão	
BES		BES	
Amostras	Estimativa	Amostras	Estimativa
31	0.91605	86	0.93543
32	0.78290	87	0.87860
33	0.78061	88	0.96753
34	0.88054	89	0.92398
35	0.96970	90	1.31132
36	0.91733		
37	0.80385		
38	0.95585		
39	0.87809		
40	0.80567		
41	0.84597		
42	0.87181		
43	0.89428		
44	0.80238		
45	1.28194		

Facilitando ainda mais a interpretação dessa classificação, o gráfico da tabela 22, apresenta a probabilidade das amostras pertencerem à classe dos BES, observe a evidente separação entre as amostras de calibração e previsão, agrupadas todas em um, e as demais amostras pertencentes a outras classes de biodieseis, em zero.

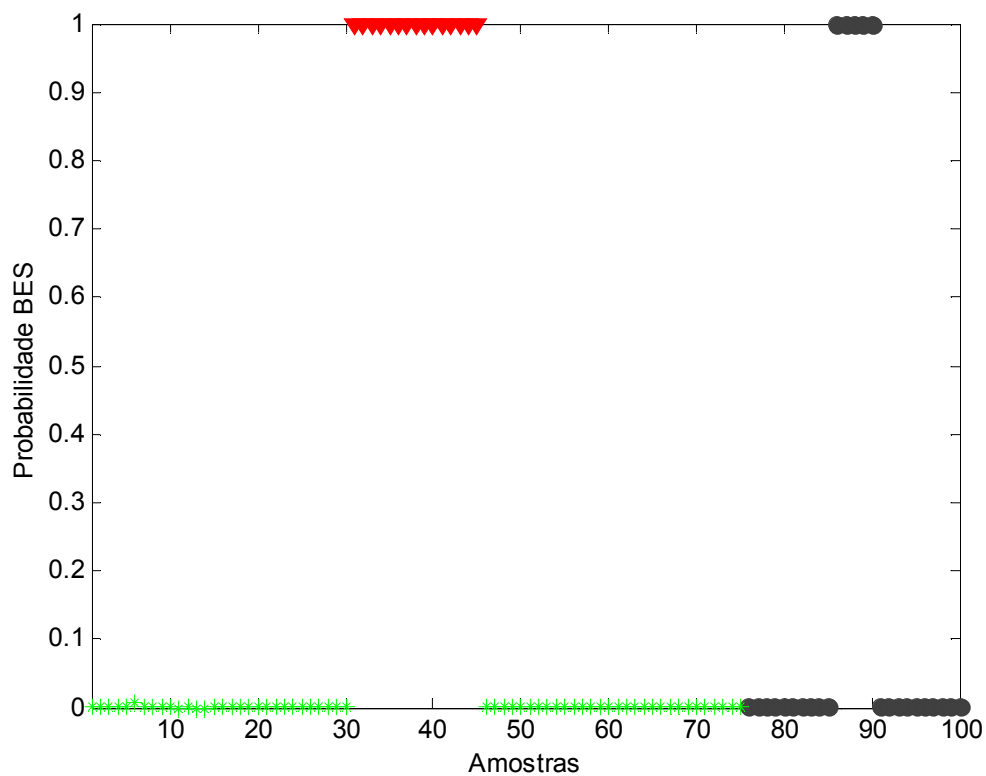


Figura 22 – Probabilidade de classificação de amostras de BES utilizando PLS-DA. Amostras de calibração: BES (▼), Outros biodieseis(*). Amostras de validação (●).

Quantificando as probabilidades apresentadas pela figura 22, a tabela 12 fornece a probabilidade como sendo de 100% para as amostras de previsão, pertencerem à classe dos BES.

Tabela 12 – Valores de probabilidade para amostras B100 de BES

Amostras de Calibração		Amostras de Previsão	
BES		BES	
Amostras	Probabilidade	Amostras	Probabilidade
31	100.000	86	100.000
32	99.9999	87	100.000
33	99.9999	88	100.000
34	100.000	89	100.000
35	100.000	90	100.000
36	100.000		
37	100.000		
38	100.000		
39	100.000		
40	100.000		
41	100.000		
42	100.000		
43	100.000		
44	100.000		
45	100.000		

Todas as amostras de BES foram classificadas com 100% de acerto.

5.2.1.4. Biodiesel Metílico de Soja (BMS)

Foram utilizadas 5 variáveis latentes para os modelos de classificação de BMS, com variância explicada 99.83% para a matriz de dados e 74,48% para a matriz de resposta .

As estimativas de previsão para as amostras de BMS são apresentadas pelo gráfico da figura 23, onde, na parte superior do gráfico, as vermelhas representam as amostras de calibração e cinco amostras em pretos são de previsão, todas elas foram separadas das demais amostras, acima do limite de confiança estabelecido em 0,5658. Porém a amostra de número 97 proveniente de BMSebo, indicada pela seta azul, ficou acima do limite de confiança de 0,4342 estabelecido para as classes não pertencentes ao BMS.

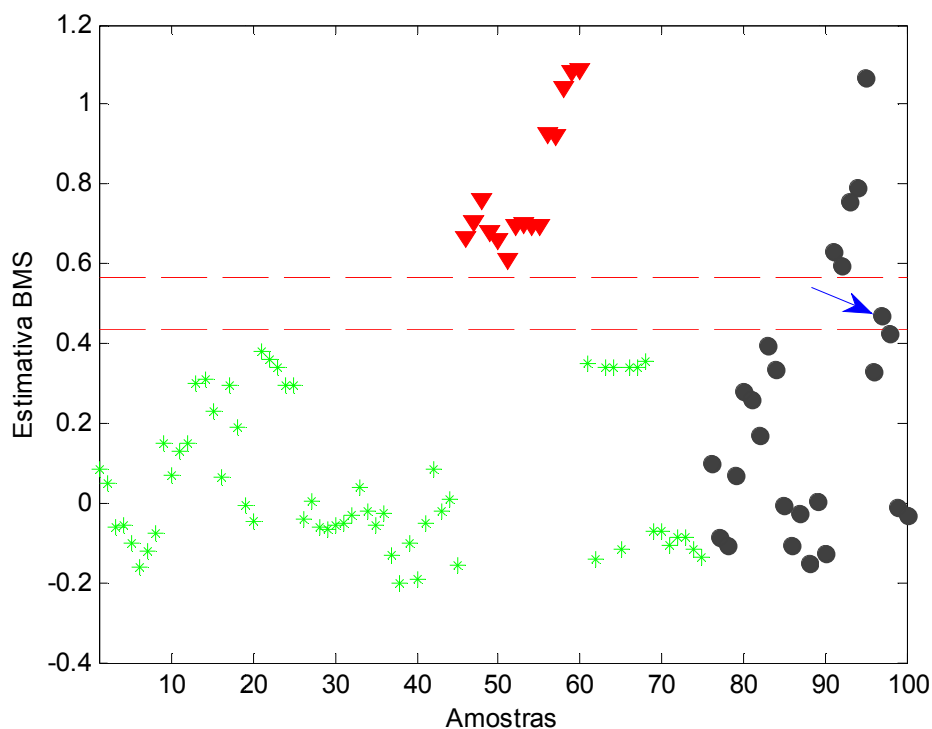


Figura 23 – Previsão de amostras de BMS utilizando PLS-DA. Amostras de calibração: BMS (▼), Outros biodieseis (*). Amostras de validação (●).

A tabela 13 quantifica a classificação visual apresentada pela figura 23, confirmando os valores de estimativa das amostras de calibração e previsão de BMS (amostras de 46 a 60 e 91 a 95) acima do limite de confiança estabelecido (0,5658). Também apresenta o valor de estimativa para a amostra 97, indicada pela seta azul, que se encontra acima do limite de confiança (0,4342) para as outras classes.

O valor de estimativa de previsão da amostra 98 de BMSebo, também é apresentada, apenas para não existir dúvidas quanto a sua localização abaixo do limite de confiança estabelecido para outras classes de biodiesel.

Essa estimativa encontrada para amostra 97 de BMSebo indica que há semelhanças entre ela e a classe de interesse modelada que é de BMS. Uma hipótese levantada foi que algumas amostras de BMS e todas as amostras de BMSebo serem provenientes da empresa Caramuru S/A que adicionam o antioxidante sintético terc-butil-hidroquinona (TBHQ), descartada posteriormente devido a sua pequena concentração nos biodieseis, porém nenhuma dessas situações foram confirmada.

Apesar de estar entre os limites de confiança estabelecido a amostra 97 não é classificada como sendo de BMS, porem poderia causar, a um analista, dúvida quanto a sua classificação.

Tabela 13 – Valores previstos para amostras B100 de BMS

Amostras de Calibração		Amostras de Previsão		
BMS				
Amostras	Estimativa	Amostras	Estimativa	Classe
46	0.66648	91	0.62995	BMS
47	0.70456	92	0.59462	BMS
48	0.76210	93	0.75729	BMS
49	0.68278	94	0.79368	BMS
50	0.66192	95	1.06804	BMS
51	0.61116			
52	0.69802	97	0.47157	BMSebo
53	0.70254	98	0.42275	BMSebo
54	0.69735			
55	0.69489			
56	0.92687			
57	0.92092			
58	1.04221			
59	1.08036			
60	1.08597			

O gráfico de probabilidade, figura 24, agrupa as amostras provenientes de BMS todas proxima a um, classificando-as corretamente como pertencentes a classe de interesse. A a amostras 97 de BMSebo, discutida anteriormente não faz parte do grupo mencionado acima.

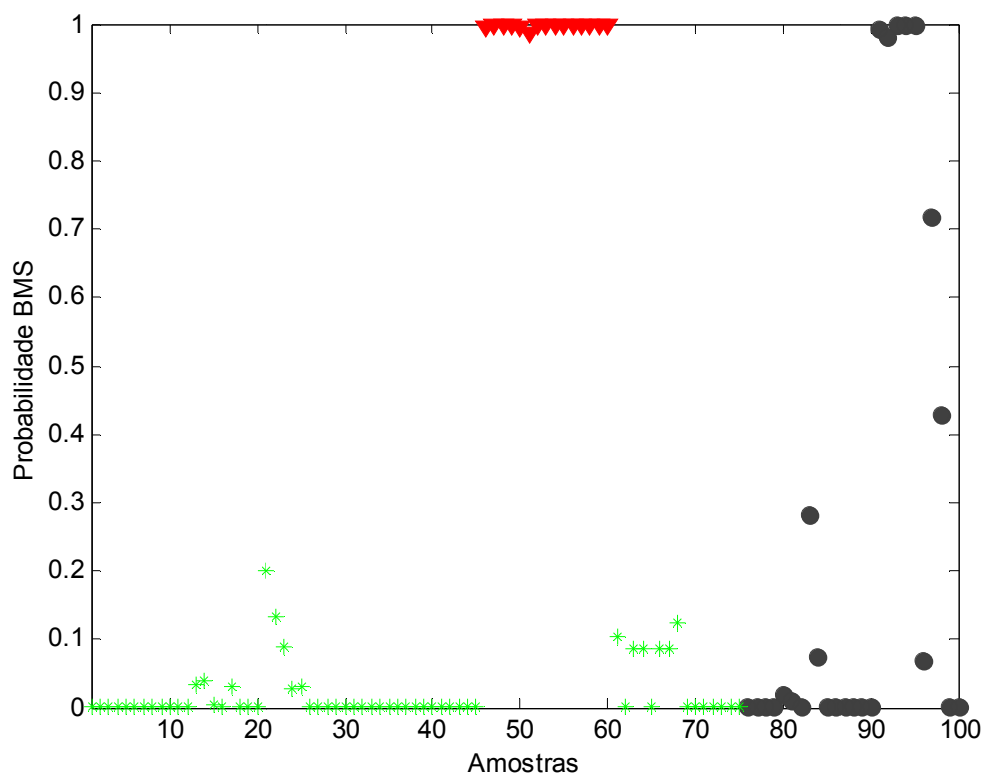


Figura 24 – Probabilidade de classificação de amostras de BMS utilizando PLS-DA. Amostras de calibração: BMS (▼), Outros biodieseis(*). Amostras de validação (●).

A tabela 14 quantifica a probabilidade da figura 24. Todas as amostras de BMS apresentam probabilidade acima 98% de pertencerem a classe de interesse. Já a amostra 97 de BMSebo apresenta esse valor bem abaixo de 98% como mostra a tabela 14.

Os dados de probabilidades apresentados pelo modelo PLS-DA, contribuiu bastante para o esclarecimento da estimativa de classificação da amostra 97.

Tabela 14 – Valores de probabilidade para amostras B100 de BMS

Amostras de Calibração		Amostras de Previsão		
BMS				
Amostras	Probabilidade	Amostras	Probabilidade	Classe
46	99.6542	91	99.1740	BMS
47	99.8599	92	98.0923	BMS
48	99.9639	93	99.9594	BMS
49	99.7652	94	99.9827	BMS
50	99.6182	95	100.000	BMS
51	98.7178			
52	99.8371	97	71.6498	BMSebo
53	99.8535	98	42.9428	BMSebo
54	99.8342			
55	99.8237			
56	99.9992			
57	99.9991			
58	99.9999			
59	100.000			
60	100.000			

A estimativa apresentada para a classificação das amostras de BMS deixou dúvida apenas em uma amostra, porém que não chegou a ser classificada erroneamente. O cálculo de probabilidade gerado pelo modelo elimina essa dúvida possibilitando sua correta classificação. Portanto conclui-se que as amostras de BMS foram classificadas com 100% de acerto.

5.2.1.5. Biodiesel Metílico de Sebo (BMSebo).

Em uma análise dos dados de BMSebo, utilizando PCA e análise prévia do modelo PLS-DA para classificação dessa classe de biodiesel, foi notado o comportamento anômalo de uma amostra de calibração, proveniente de BMS. Essa amostra foi excluída do conjunto de calibração, utilizando-se então um conjunto de 99 amostras, 14 para calibração e 5 para previsão.

Nesse modelo foram utilizadas 7 variáveis latentes, com variância explicada 99.93% para a matriz de dados e 85,34% para a matriz de resposta .

A figura 25 apresenta as amostras de calibração de BMSebo acima do limite de confiança estabelecido em 0,6280, porém das 5 amostras pertencentes ao conjunto de

previsão apenas 3 amostras estão acima do limite de confiança estabelecido. Duas amostras, 96 e 97, não foram classificadas corretamente como pertencentes à classe dos BMSebo. Elas são indicadas pelas setas azuis na figura 25.

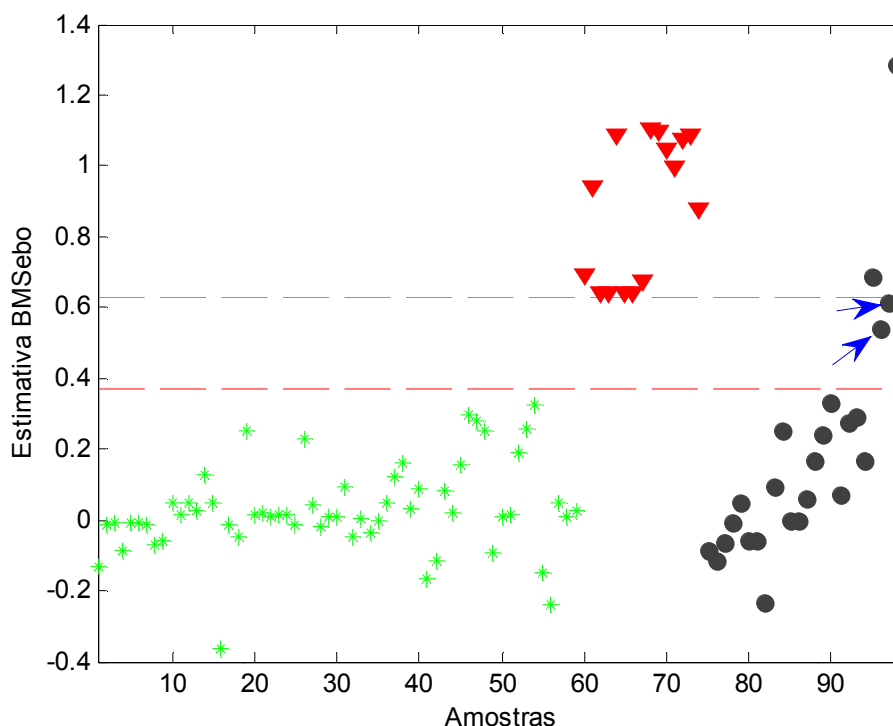


Figura 25 – Previsão de amostras de BMSebo utilizando PLS-DA. Amostras de calibração: BMSebo (▼), Outros biodieseis(*). Amostras de validação (●).

A tabela 15 apresenta quantitativamente as estimativas da tabela 25, nota-se o valor da previsão das amostras 96 e 97 abaixo do limite de confiança estabelecido pelo modelo (0,6280).

Uma hipótese levantada para classificação das duas amostras de BMSebo deve-se ao fato de que algumas amostras de BMS e todas as amostras de BMSebo serem provenientes da empresa Caramuru S/A que adicionam o antioxidante sintético terc-butil-hidroquinona (TBHQ) aumentando a semelhança entre essas amostras de previsão ocasionando a classificação das duas amostras abaixo do limite de confiança. Porém como discutido anteriormente, a hipótese foi descartada devido a pequena concentração desse antioxidante nos biodieseis, entretanto essas situações não foram confirmada.

Tabela 15 – Valores previstos para amostras B100 de BMSebo

Amostras de Calibração		Amostras de Previsão	
BMSebo		BMSebo	
Amostras	Estimativa	Amostras	Estimativa
61	0.93958	95	0.68764
62	0.64227	96	0.54046
63	0.64224	97	0.61531
64	1.08928	98	1.28645
65	0.64225	99	1.18552
66	0.64226		
67	0.67425		
68	1.10652		
69	1.09631		
70	1.04527		
71	0.99413		
72	1.07479		
73	1.08928		
74	0.87576		

A figura 26 apresenta a probabilidade da classificação para as amostras de BMSebo. Em comparação com a estimativa apresentada pela figura 25, o gráfico de probabilidade não apresenta classificações duvidosas para as cinco amostras de previsão de BMSebo. Observe, agrupadas em um, em vermelho as amostras de calibração em preto as amostras de previsão da classe de interesse.

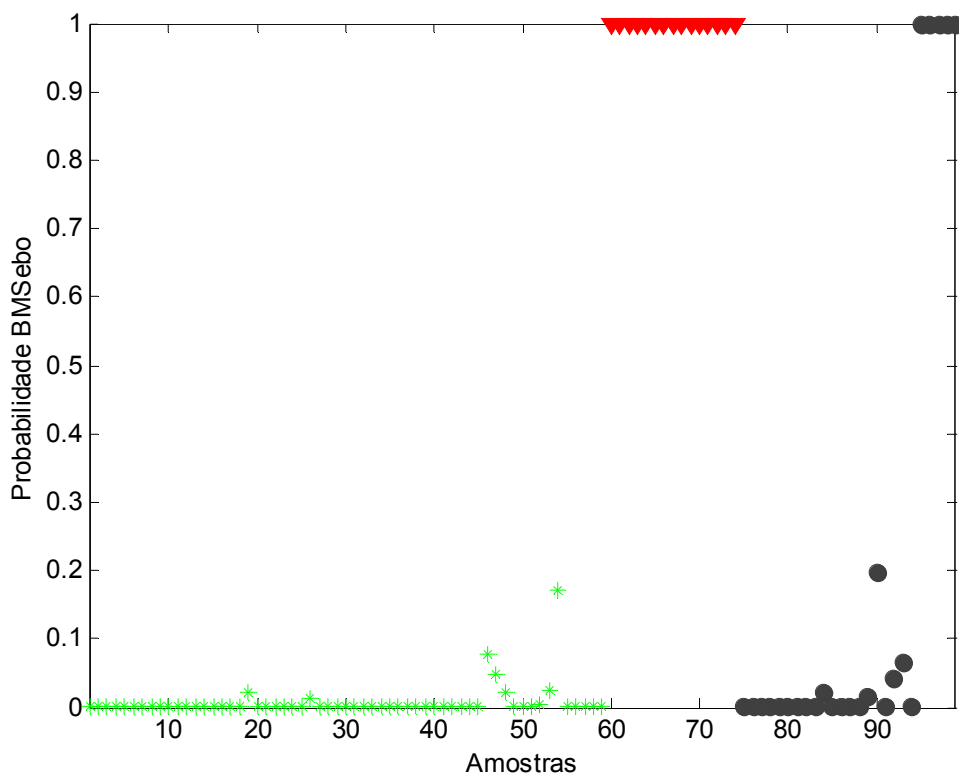


Figura 26 – Probabilidade de classificação de amostras de BMSebo utilizando PLS-DA. Amostras de calibração: BMSebo (▼), Outros biodieseis(*). Amostras de validação (●).

A tabela 16 quantifica os valor de probabilidade das amostras pertencerem ou não a classe dos BMSebo. Todas as amostras de previsão para BMSebo apresentam probabilidade de classificação superior 99.8%.

Comparando com os valores de probabilidade para a classificação para BMS, tabela 14, onde amostras também ultrapassaram o limite de confiança estabelecido, temos na classificação para BMSebo melhores resultados de probabilidade para esse tipo de comportamento.

Tabela 16 – Valores de probabilidade para amostras B100 de BMSebo

Amostras de Calibração		Amostras de Previsão	
BMSebo		BMSebo	
Amostras	Probabilidade	Amostras	Probabilidade
61	100.000	95	99.9997
62	99.9973	96	99.8023
63	99.9973	97	99.9913
64	100.000	98	100.000
65	99.9973	99	100.000
66	99.9973		
67	99.9994		
68	100.000		
69	100.000		
70	100.000		
71	100.000		
72	100.000		
73	100.000		
74	100.000		

Esse modelo apresentou os piores resultados de estimativa em comparação a todos os outros modelos apresentados nesse trabalho.

Porém as duas amostras que ficaram abaixo do limite de confiança, não chegaram a ser agrupadas com outras amostras provenientes de outras classes de biodieseis.

Quando o modelo PLS-DA apresenta o gráfico e os valores de probabilidade para essa classificação ela ocorre com 100% de acerto.

5.2.2. Classificação de biodieseis B5.

5.2.2.1. Biodiesel Etilico Residual (BER)

Foram utilizados 7 variáveis latentes para os modelos de classificação para B5 de BER, com variância explicada 99.34% para a matriz de dados e 76,87% para a matriz de resposta .

As amostras em vermelho agrupadas na parte superior do gráfico apresentadas na figura 27, são as amostras de calibração e as em preto as amostras de previsão. Todas elas encontram-se acima do limite de confiança estabelecido em 0,5667, portanto foram classificadas corretamente como amostras de misturas BER em diesel.

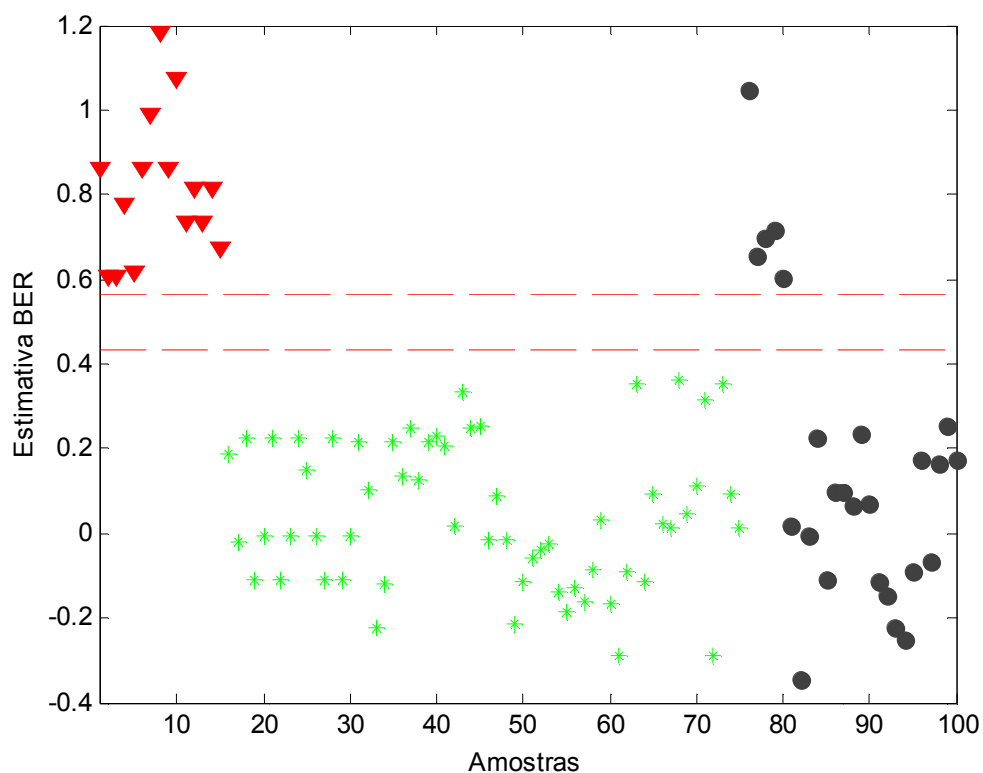


Figura 27 – Previsão de amostras de B5 preparados com BER utilizando PLS-DA. Amostras de calibração: BER (▼), Outros B5(*). Amostras de validação (●).

A tabela 17 quantifica os valores de estimativas para as amostras de misturas de BER em diesel, apresentados graficamente pela figura 27. Nota-se todos os valores de estimativa superior ao limite de confiança estabelecido.

Tabela 17 – Valores previstos para amostras B5 de BER

Amostras de Calibração		Amostras de Previsão	
BER		BER	
Amostras	Estimativa	Amostras	Estimativa
1	0.86457	76	1.04821
2	0.60886	77	0.65304
3	0.60886	78	0.69921
4	0.77773	79	0.71540
5	0.61503	80	0.60194
6	0.86457		
7	0.98827		
8	1.18277		
9	0.86186		
10	1.07369		
11	0.73689		
12	0.81494		
13	0.73689		
14	0.81494		
15	0.67415		

A figura 28 apresenta a probabilidade das amostras pertencerem à classe de interesse, no caso mistura de BER em diesel. Confirmando a separação já apresentada pela figura 27, observa-se as amostras agrupadas em um, estabelecendo alta probabilidade de pertencerem a classe de interesse modelada.

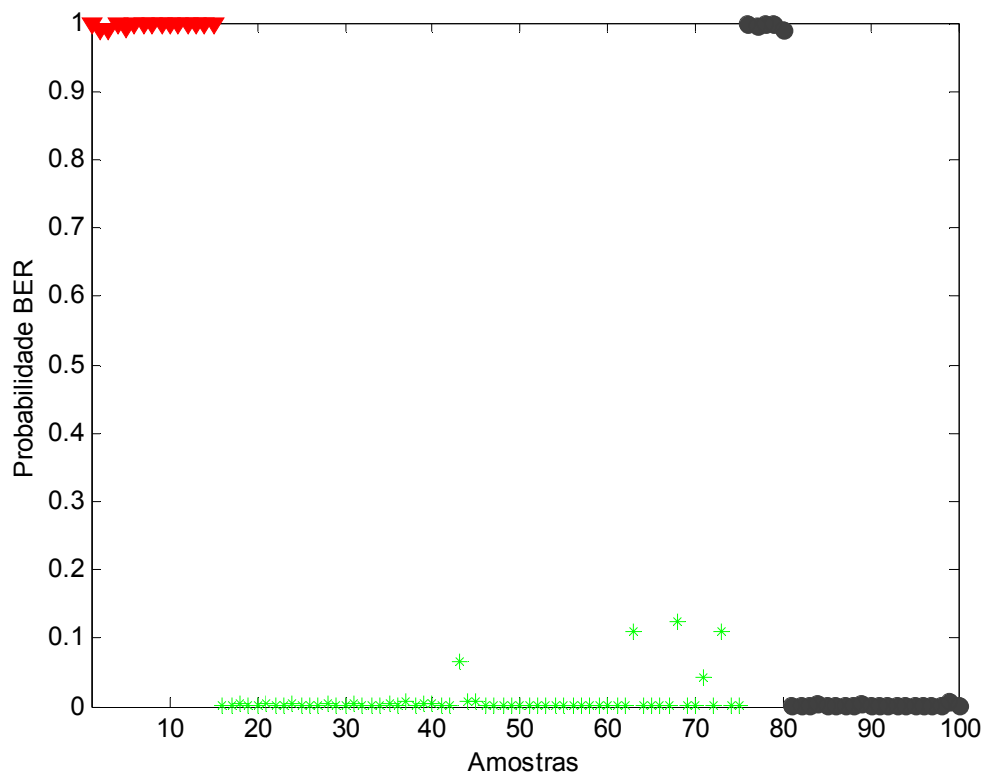


Figura 28 – Probabilidade de classificação de B5 preparados com BER utilizando PLS-DA. Amostras de calibração: BER(▼), Outros B5(*). Amostras de validação (●).

Quantificando os valores apresentados pela figura 27, a tabela 18 mostra os valores de probabilidade dessas amostras pertencerem ou não a classe modela, como sendo superior a 99,7% .

Tabela 18 – Valores de probabilidade para amostras B5 de BER

Amostras de Calibração		Amostras de Previsão	
BER		BER	
Amostras	Probabilidade	Amostras	Probabilidade
1	99.9988	76	100.000
2	99.0500	77	99.7014
3	99.0500	78	99.9104
4	99.9885	79	99.9412
5	99.1782	80	98.8426
6	99.9988		
7	100.000		
8	100.000		
9	99.9987		
10	100.000		
11	99.9664		
12	99.9957		
13	99.9664		
14	99.9957		
15	99.8256		

Todas as amostras B5 de BES foram classificadas com 100% de acerto por esse modelo.

5.2.2.2. Biodiesel Metílico Residual (BMR)

Foram utilizadas 6 variáveis latentes para os modelos de classificação B5 de BER com variância explicada 97.87% para a matriz de dados e 81,69% para a matriz de resposta .

A figura 29 mostra a separação das amostras de BMR em relação às outras classes presentes sendo que em vermelho são as amostras de calibração, e em preto na parte superior do gráfico, amostras de previsão. Todas as amostras da mistura de BMR em diesel encontram-se acima do limite de confiança estabelecido em 0.5893.

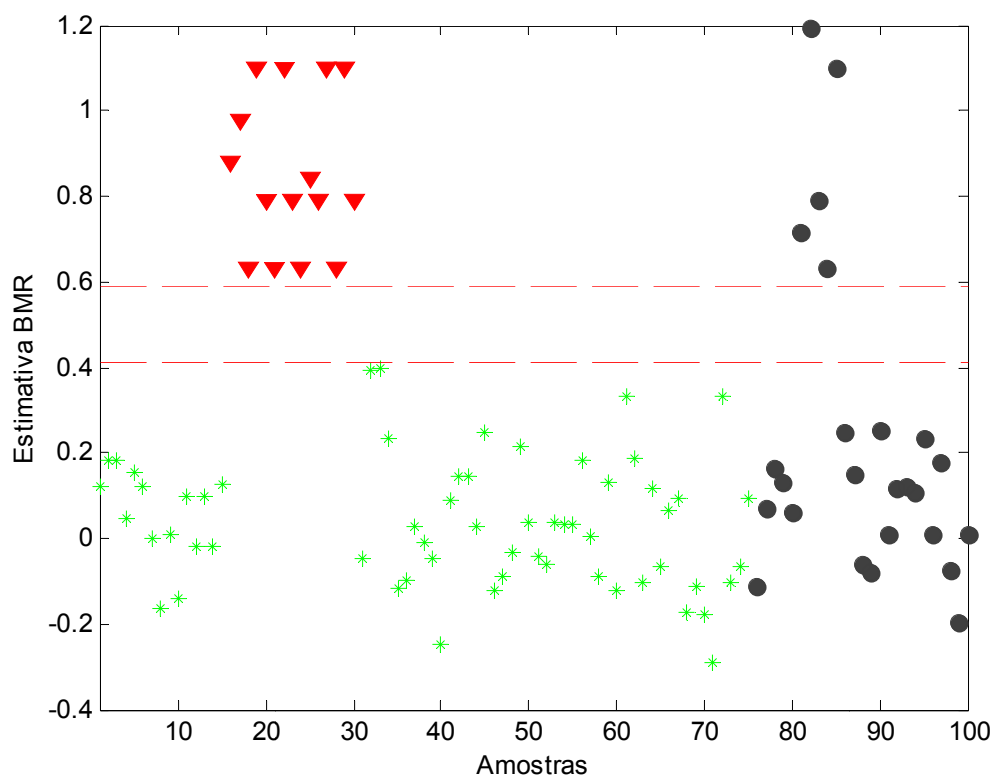


Figura 29 – Previsão de amostras de B5 preparados com BMR utilizando PLS-DA. Amostras de calibração: BMR(▼), Outros B5(*). Amostras de validação (●).

A tabela 19 quantifica os valores de estimativas apresentados pela figura 29, em que todas as amostras apresentam estimativas acima do limite de confiança estabelecido.

Tabela 19 – Valores previstos para amostras B5 de BMR

Amostras de Calibração		Amostras de Previsão	
BMR		BMR	
Amostras	Estimativa	Amostras	Estimativa
16	0.87798	81	0.71387
17	0.98003	82	1.19372
18	0.63228	83	0.79272
19	1.10034	84	0.63228
20	0.79272	85	1.10034
21	0.63229		
22	1.10034		
23	0.79272		
24	0.63227		
25	0.84307		
26	0.79272		
27	1.10034		
28	0.63228		
29	1.10034		
30	0.79272		

A figura 30 apresenta a probabilidade dessas amostras pertencerem a classe da mistura de BMR em diesel. Observe todas as amostras pertencentes a essa mistura agrupadas em um.

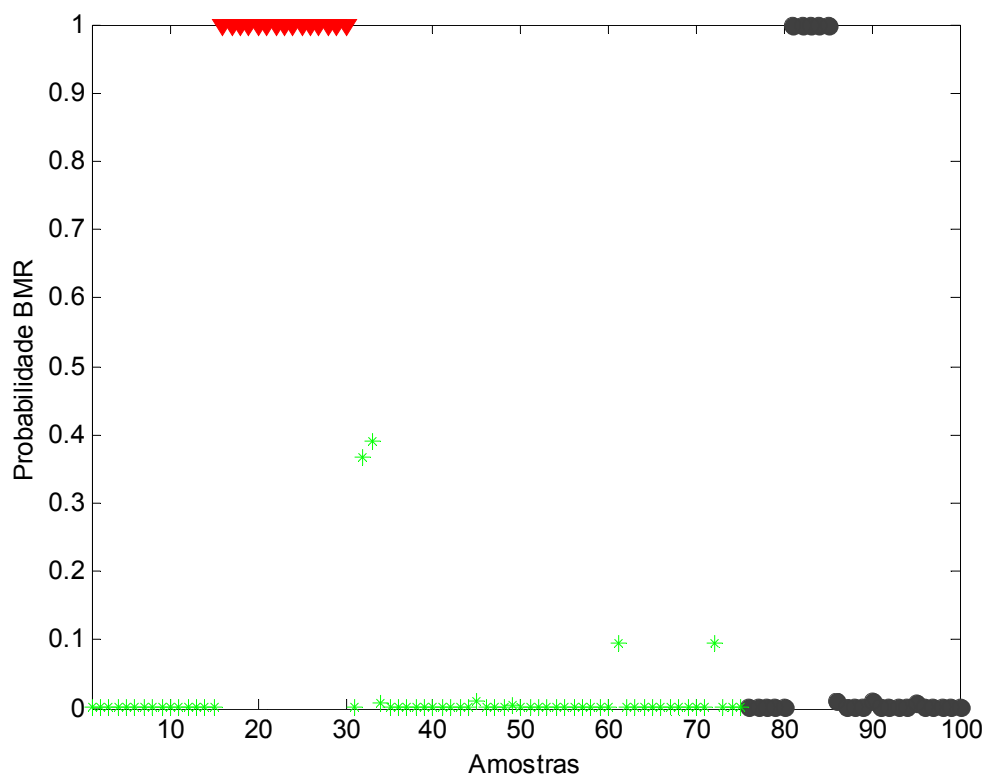


Figura 30 – Probabilidade de classificação de B5 preparados com BMR utilizando PLS-DA. Amostras de calibração: BMR(▼), Outros B5(*). Amostras de validação (●).

Os valores de probabilidade das amostras de previsão pertencerem a classe dos B5 de BMR estão acima de 99,9% como mostra a tabela 20, que apresenta quantitativamente os valores de probabilidade da figura 30.

Tabela 20 – Valores de probabilidade para amostras B5 de BMR

Amostras de Calibração		Amostras de Previsão	
BMR		BMR	
Amostras	Probabilidade	Amostras	Probabilidade
16	100.000	81	99.9939
17	100.000	82	100.000
18	99.9043	83	99.9996
19	100.000	84	99.9044
20	99.9996	85	100.000
21	99.9044		
22	100.000		
23	99.9996		
24	99.9044		
25	99.9999		
26	99.9996		
27	100.000		
28	9.99044		
29	1.00000		
30	9.99996		

A classificação das amostras de B5 de BMR ocorreu com 100% de acerto como mostra os dados apresentados acima.

5.2.2.3. Biodiesel Etilico de soja (BES)

Foram utilizadas 5 variáveis latentes para os modelos de classificação de BES com variância explicada 92.42% para a matriz de dados e 76.52% para a matriz de resposta .

A figura 31 apresenta a estimativa para classificação de amostras B5 de BES. As amostras acima do limite de confiança, com valor de 0,5619, que estão em vermelho são as amostras de calibração e as em preto, amostras de previsão. Todas as amostras B5 de BES forma separadas e então classificadas corretamente.

Porém uma amostra previsão pertencente à classe dos B5 de BER, indicada pela seta em azul na figura 31, ficou acima do primeiro limite de confiança, de 0,4381, estabelecido para separar com segurança as classes que não são a de interesse, ou seja, separar os B5 que não são provenientes de BES.

A situação descrita nessa estimativa ocorre devido a semelhanças encontrada pelo modelo, entre essa amostra B5 de BER e a classe de interesse nessa classificação, que é B5

de BES. Pois além de serem biodiesel etílicos, a variabilidade da composição dos óleos residuais é muito grande. Caso o óleo residual fosse proveniente apenas de soja e descartados após pouca utilização sem várias alterações de temperatura promovendo poucas modificações em sua estrutura, como por exemplos transesterificações, suas semelhanças serão grandes.

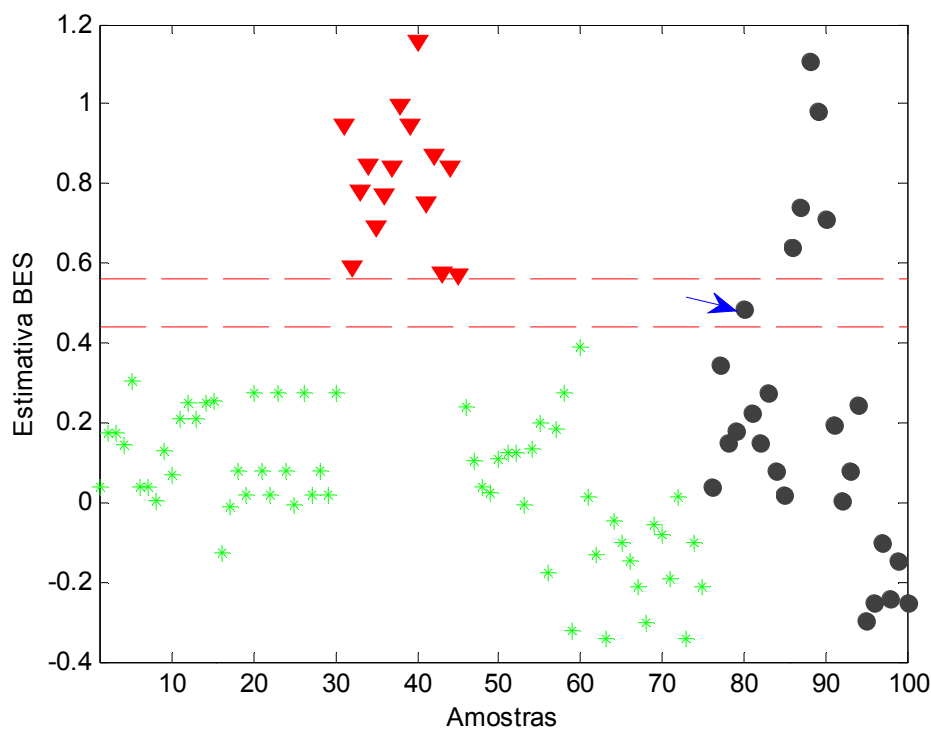


Figura 31 – Previsão de amostras de B5 preparados com BES utilizando PLS-DA. Amostras de calibração: BES(▼), Outros B5(*). Amostras de validação (●).

A tabela 21 apresenta a quantificação da estimativa para as amostras de B5 de BES. Ela deixa clara a classificação dessas amostras, principalmente para três amostras de calibração que ficam próximas ao limite de confiança. Para amostra proveniente de BES o valor de estimativa é 0,4846.

Tabela 21 – Valores previstos para amostras B5 de BES.

Amostras de Calibração		Amostras de Previsão	
BES		BES	
Amostras	Estimativa	Amostras	Estimativa
31	0.94550	86	0.64234
32	0.59286	87	0.74100
33	0.77936	88	1.10659
34	0.84724	89	0.98067
35	0.69313	90	0.71280
36	0.77204		
37	0.84205		
38	0.99505		
39	0.94550		
40	1.15843		
41	0.75344		
42	0.86934		
43	0.57689		
44	0.84205		
45	0.56939		

A figura 32 apresenta as probabilidades das amostras pertencerem ou não a classe do B5 de BES. Todas agrupadas próximo a um, apresenta elevada probabilidade de pertencerem a classe de interesse. A amostra proveniente de BER, indicada pela seta em azul, apesar de não estar agrupadas com as amostras de outras classes de B5, também não apresenta elevada probabilidade de classificação para B5 de BES.

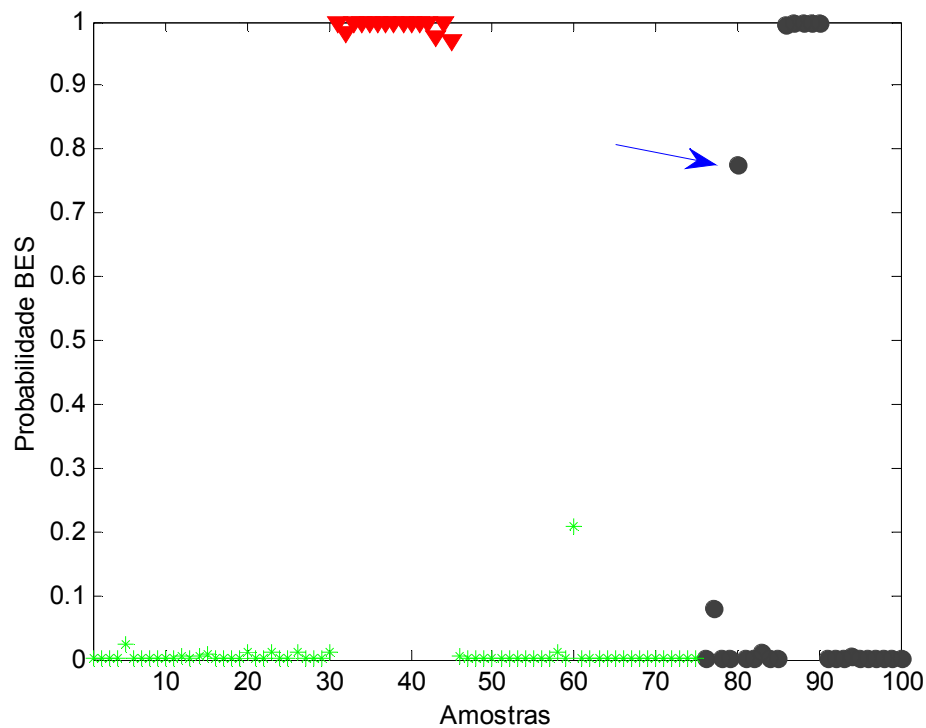


Figura 32 – Probabilidade de classificação de B5 preparados com BES utilizando PLS-DA. Amostras de calibração: BES (▼), Outros B5(*). Amostras de validação (●).

A tabela 22 quantifica esses valores para essa probabilidade. A classificação das amostras de previsão para BES apresenta probabilidades superiores a 99,5% enquanto a probabilidade da amostra de BER é de 77,53%.

Tabela 22 – Valores de probabilidade para amostras B5 de BES

Amostras de Calibração		Amostras de Previsão	
BES		BES	
Amostras	Probabilidade	Amostras	Probabilidade
31	99.9998	86	99.5430
32	98.3557	87	99.9644
33	99.9866	88	100.000
34	99.9976	89	99.9999
35	99.8792	90	99.9262
36	99.9839		
37	99.9973		
38	99.9999		
39	99.9998		
40	100.000		
41	99.9744		
42	99.9986		
43	97.5253		
44	99.9973		
45	97.0069		

A partir do comportamento apresentado pelas amostras de previsão para B5 de BES para as estimativas e probabilidade de classificação, em comparação com as amostras B5 de BER, conclui-se apesar de existirem semelhanças não é possível existir confusão nessa classificação. Portanto pode se afirmar que o modelo apresentou 100% de acerto na classificação de amostras de B5 de BES.

5.2.2.4. Biodiesel Metílico de soja (BMS)

Foram utilizadas 6 variáveis latentes para os modelos de classificação de BMS com variância explicada 98.33% para a matriz de dados e 87.32% para a matriz de resposta .

A figura 33 apresenta a estimativa das amostras pertencerem à classe dos B5 de BMS. As amostras em vermelho, de calibração e em preto, amostras de previsão que se encontram na parte superior do gráfico, acima do limite de confiança estabelecido em 0,6634 são amostras de B5 de BMS. Essa figura apresenta a nítida divisão entre as classes.

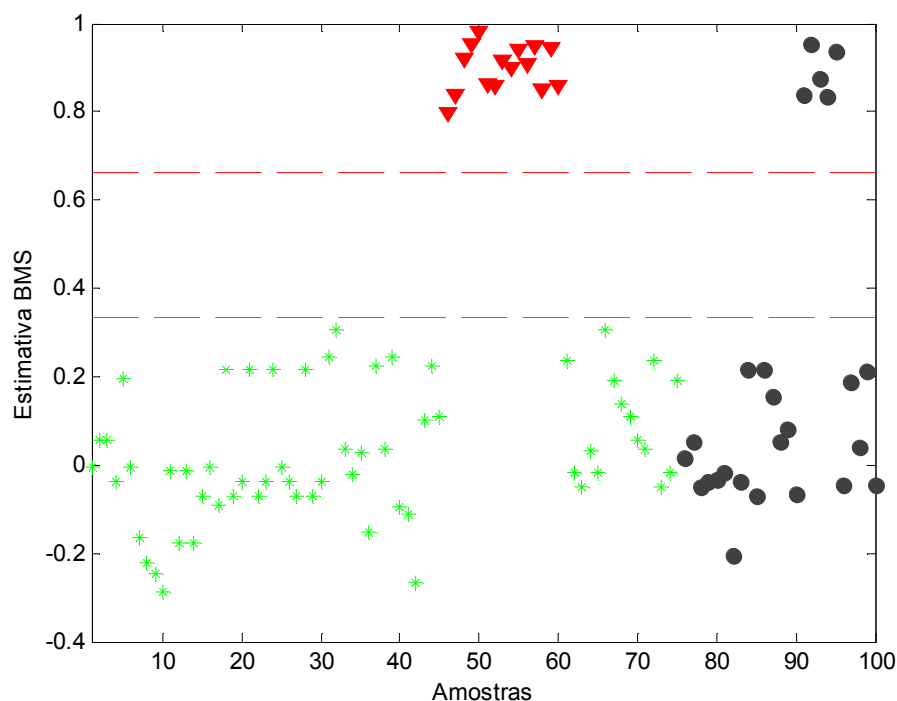


Figura 33 – Previsão de amostras de B5 preparados com BMS utilizando PLS-DA. Amostras de calibração: BMS(▼), Outros B5(*). Amostras de validação (●).

A tabela 23 quantifica os valores previstos para a classificação de amostras B5 de BMS, todos os valores de estimativa encontram-se a cima do limite de confiança.

Tabela 23 – Valores previstos para amostras B5 de BMS

Amostras de Calibração		Amostras de Previsão	
BMS		BMS	
Amostras	Estimativa	Amostras	Estimativa
46	0.79661	91	0.83857
47	0.83857	92	0.95396
48	0.91944	93	0.87587
49	0.95356	94	0.83576
50	0.97982	95	0.93751
51	0.86420		
52	0.85797		
53	0.91467		
54	0.89835		
55	0.93989		
56	0.90902		
57	0.94793		
58	0.85111		
59	0.94652		
60	0.86070		

Já a figura 34 apresenta a probabilidade para a classificação das amostras pertencerem ou não a classe dos B5 de BMS. Não há dúvida que todas as amostras de previsão de B5 de BMS foram classificadas corretamente.

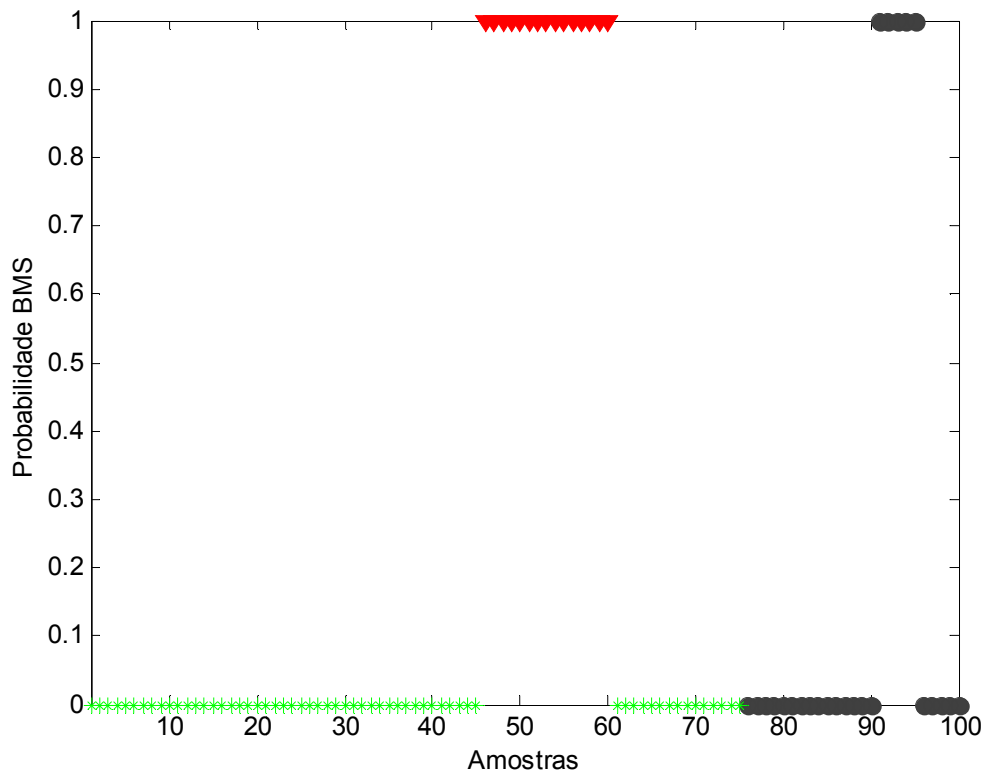


Figura 34 – Probabilidade de classificação de B5 preparados com BMS utilizando PLS-DA. Amostras de calibração: BMS(▼), Outros B5(*). Amostras de validação (●).

Quantificando as observações de probabilidade da figura 34, a tabela 24 apresenta o valor de 100% para todas as amostras de previsão pertencer à classe dos B5 de BMS.

Tabela 24 – Valores de probabilidade para amostras B5 de BMS

Amostras de Calibração		Amostras de Previsão	
BMS		BMS	
Amostras	Probabilidade	Amostras	Probabilidade
46	99.9997	91	100.000
47	100.000	92	100.000
48	100.000	93	100.000
49	100.000	94	100.000
50	100.000	95	100.000
51	100.000		
52	100.000		
53	100.000		
54	100.000		
55	100.000		
56	100.000		
57	100.000		
58	100.000		
59	1.00000		
60	1.00000		

A classificação das amostras de BMS ocorreu com 100% de acerto.

5.2.2.5. Biodiesel Metílico de Sebo (BMSebo)

Foram utilizadas 4 variáveis latentes para os modelos de classificação de BMSebo com variância explicada 86.02% para a matriz de dados e 84.70% para a matriz de resposta .

A figura 35 mostra a separação das amostras de BMSebo em relação às outras classes presentes sendo que em vermelho são as amostras de calibração, e em preto na parte superior do gráfico, amostras de previsão. Todas as amostras B5 de BMSebo em encontram-se acima do limite de confiança estabelecido em 0.5320.

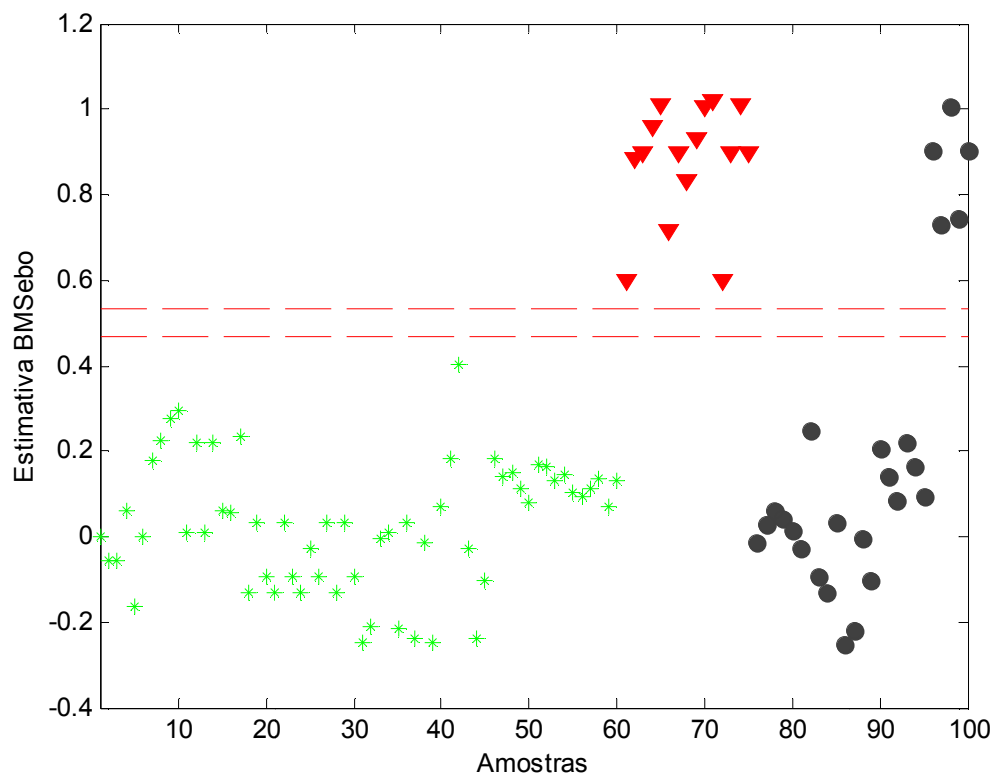


Figura 35 – Previsão de amostras de B5 preparados com BMSebo utilizando PLS-DA. Amostras de calibração: BMSebo(▼), Outros B5(*). Amostras de validação (●).

A tabela 25 quantifica os valores de estimativas apresentados pela figura 35, em que todas as amostras apresentam estimativas acima do limite de confiança estabelecido.

Tabela 25 – Valores previstos para amostras B5 de BMSebo

Amostras de Calibração		Amostras de Previsão	
BMSebo		BMSebo	
Amostras	Estimativa	Amostras	Estimativa
61	0.60114	96	0.90242
62	0.88548	97	0.72829
63	0.89947	98	1.00775
64	0.96090	99	0.74442
65	1.01114	100	0.90242
66	0.71665		
67	0.89700		
68	0.83065		
69	0.92898		
70	1.00381		
71	1.01971		
72	0.60114		
73	0.89947		
74	1.01114		
75	0.89700		

A figura 36 apresenta a probabilidade dessas amostras pertencerem a classe da mistura de BMSebo em diesel. Observe todas as amostras pertencentes a essa mistura agrupadas em um.

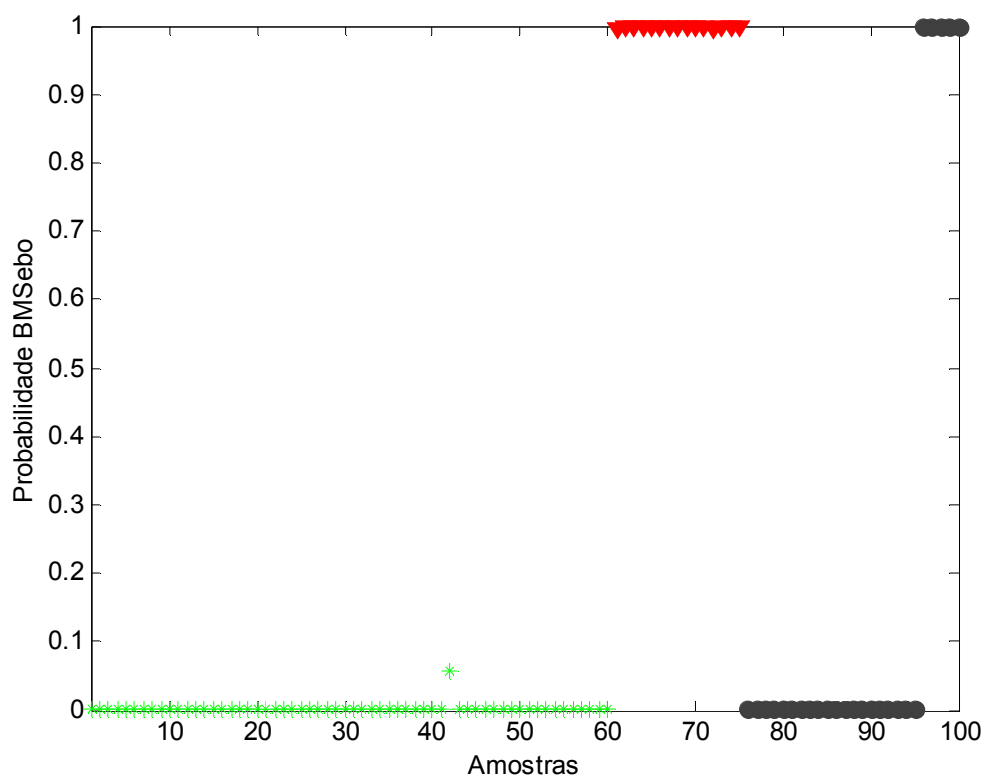


Figura 36 – Probabilidade de classificação de B5 preparados com BMSebo utilizando PLS-DA. Amostras de calibração: BMSebo(▼), Outros B5(*). Amostras de validação (●)

Os valores de probabilidade das amostras de previsão pertencerem a classe dos B5 de BMSebo estão acima de 99,9% como mostra a tabela 26, que apresenta quantitativamente os valores de probabilidade da figura 36.

Tabela 26 – Valores de probabilidade para amostras B5 de BMSebo

Amostras de Calibração		Amostras de Previsão	
BMSebo		BMSebo	
Amostras	Probabilidade	Amostras	Probabilidade
61	99.5850	96	100.000
62	100.000	97	99.9975
63	100.000	98	100.000
64	100.000	99	99.9987
65	100.000	100	100.000
66	99.9960		
67	100.000		
68	100.000		
69	100.000		
70	100.000		
71	100.000		
72	99.5850		
73	100.000		
74	100.000		
75	100.000		

A classificação das amostras de B5 de BMSebo ocorreu com 100% de acerto como mostra os dados apresentados anteriormente.

5.2.3. Classificação de biodieseis BX quanto ao teor de 5% de Biodiesel em diesel (B5).

Para o modelo que classifica as amostras da mistura biodiesel/diesel preparadas com diferentes teores de BMS em conformes e não conformes de acordo com a legislação vigente (ou seja, se as amostras apresentam ou não teor de 5% de biodiesel em diesel), foram utilizadas 8 variáveis latentes, com variância explicada 99.90% para a matriz de dados e 91.69% para a matriz de resposta.

Para esse modelo também são apresentados os valores previstos de estimativa e as probabilidades das amostras de previsão, estarem ou não em conformidade quanto ao teor de 5% de biodiesel metílico de soja em diesel.

A figura 37 apresenta a estimativa de previsão para essa classificação. Acima e em vermelho amostras de calibração de B5 de BMS, as amostras em preto acima dos limites de confiança, estabelecido em 0,5810 são de previsão provenientes de B5 de BMS. As amostras em verde e em preto na parte inferior do gráfico são respectivamente amostras de

calibração e previsão de misturas de BMS com diferentes teores em diesel, conforme tabela 4.

Todas as amostras de B5 de BMS forma classificadas corretamente.

Apenas duas amostras de previsão que não pertencem a classe dos B5 de BMS encontra-se fora do primeiro limite de confiança, estabelecido em 0,4190 para a separação das misturas de BMS em diesel com outros teores de biodiesel. Uma dessas amostras apresenta um teor de BMS inferior a 5% (igual a 4,45%) e a outra amostra, apresenta um teor de BMS superior a 5% (igual a 8,44%). A semelhança de uma das amostras é possivelmente encontrada pelo modelo, pois seu teor aproxima-se de 5%, enquanto a outra provavelmente é ocasionada por compostos presentes no óleo e no diesel.

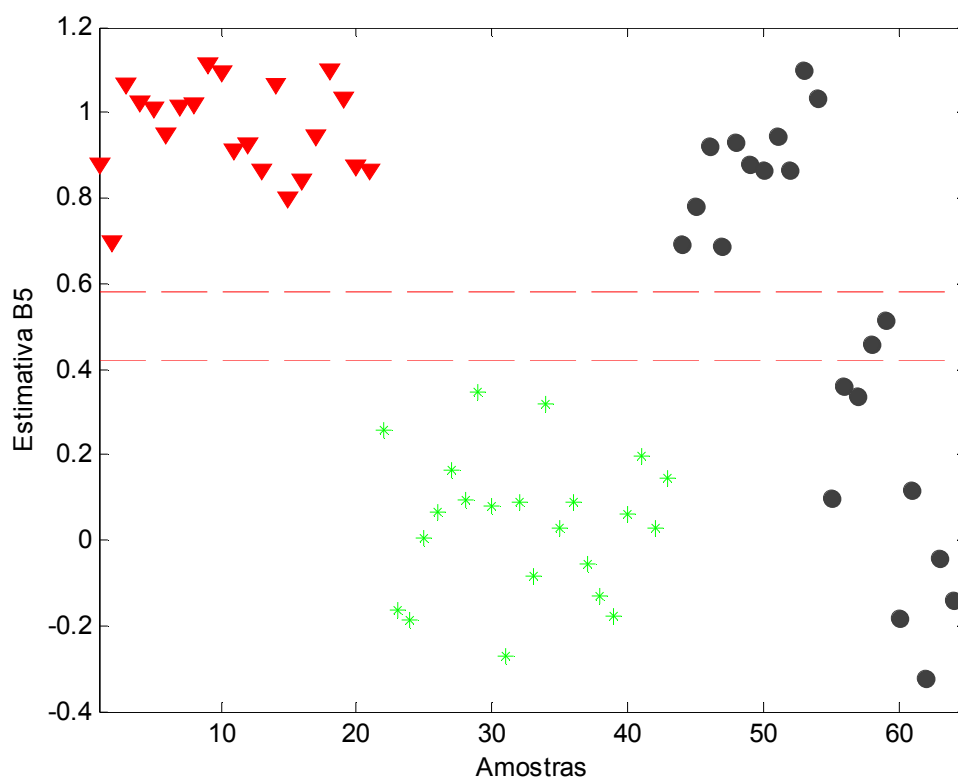


Figura 37 – Previsão de amostras com 5% BMS utilizando PLS-DA. Amostras de calibração: B5 de BMS (▼), BX (✱). Amostras de validação (●).

Os valores de estimativa para essa classificação são apresentados pela tabela 27, todos se encontram acima do limite de confiança, Classificando corretamente as amostras de previsão de B5 de BMS.

Tabela 27 – Valores previstos para amostras com 5% de Biodiesel de BMS.

Amostras de Calibração		Amostras de Previsão	
BMS		BMS	
Amostras	Estimativa	Amostras	Estimativa
1	0.87918	44	0.69292
2	0.69483	45	0.78206
3	1.06562	46	0.91932
4	1.02432	47	0.68865
5	1.01084	48	0.93272
6	0.95056	49	0.88128
7	1.01541	50	0.86755
8	1.01807	51	0.94577
9	1.11537	52	0.86461
10	1.09543	53	1.09777
11	0.91448	54	1.03506
12	0.92422		
13	0.86681		
14	1.06625		
15	0.80106		
16	0.84355		
17	0.94577		
18	1.09777		
19	1.03506		
20	0.87451		
21	0.86755		

Esclarecendo ainda mais essa classificação o gráfico de probabilidade classifica corretamente as amostras de previsão de B5 de BMS. Não levantando dúvidas quanto a a que as amostras pertencem.

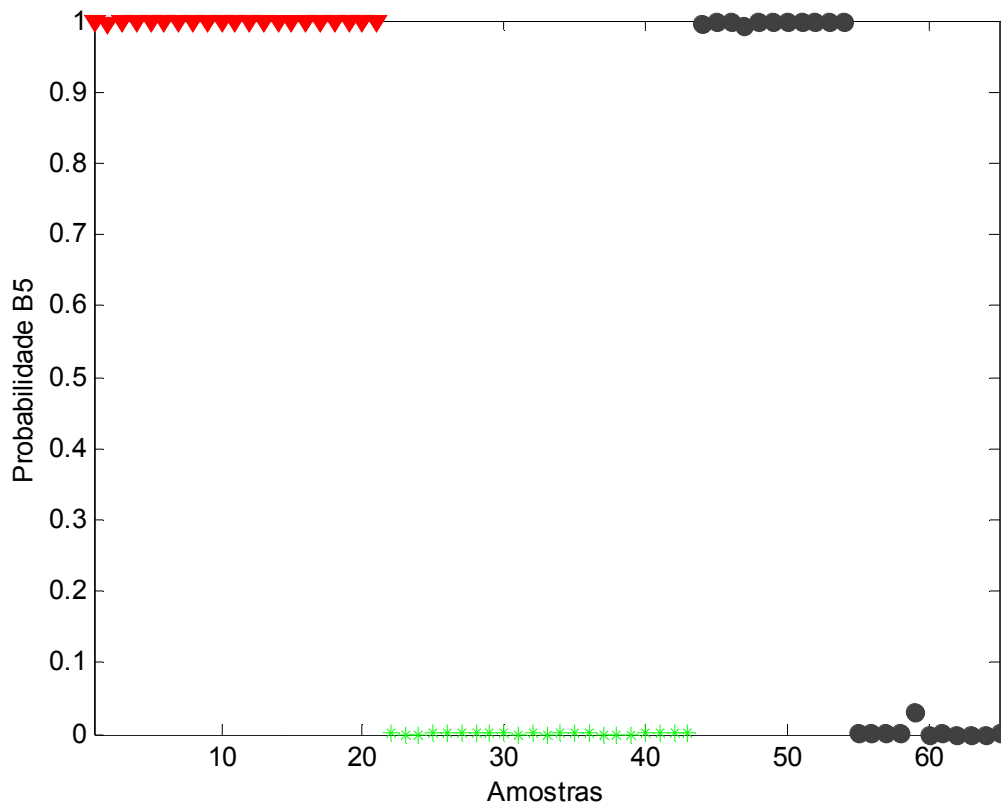


Figura 38 – Probabilidade de classificação de amostras com 5% BMS utilizando PLS-DA. Amostras de calibração: B5 de BMS(▼), BX(*). Amostras de validação (●).

A tabela 28 apresenta a quantificação da probabilidade dessas amostras pertencerem à classe dos B5 de BMS. Essa probabilidade é superior a 99,4% .

Tabela 28 – Valores de probabilidade para amostras com 5% de Biodiesel de BMS.

Amostras de Calibração		Amostras de Previsão	
BMS		BMS	
Amostras	Probabilidade	Amostras	Probabilidade
1	99.9998	44	99.5081
2	99.5513	45	99.9902
3	100.000	46	99.9999
4	100.000	47	99.4120
5	100.000	48	100.000
6	100.000	49	99.9998
7	100.000	50	99.9997
8	100.000	51	100.000
9	100.000	52	99.9996
10	100.000	53	100.000
11	99.9999	54	100.000
12	100.000		
13	99.9997		
14	100.000		
15	99.9953		
16	99.9991		
17	100.000		
18	100.000		
19	100.000		
20	99.9997		
21	99.9997		

O modelo foi capaz de prever com 100% de certeza quais amostras são de misturas de BMS em diesel com teor de 5%. Apesar de duas amostras ultrapassarem o primeiro limite de confiança estabelecido, não ocorrem falsas classificações, mostrando ainda que pequenas diferenças de concentração podem ser detectadas pelo modelo. Em estudos futuros podem ser determinados quais as pequenas diferenças de concentração que esse tipo de modelo consegue capturar.

Os gráficos e valores de probabilidade esclarecem ainda mais essa classificação.

9. Conclusão

9. Conclusão

A espectroscopia no Infravermelho médio (MIR) com técnica de Reflectância Total Atenuada (ATR) mostra-se promissora em comparação a técnicas cromatográficas utilizadas para análise de biocombustíveis, pois se apresenta como uma técnica rápida, não utiliza reagente, não gera resíduos e requer pouca preparação de amostras.

Quando aliada a quimiometria, na necessidade de analisar grandes quantidades de informações, o grande sinergismo entre as duas possibilita ganhos ilimitados nas análises realizadas. Podemos destacar a associação de técnicas espectroscópicas com a técnica de classificação de dados PLS-DA que vem apresentando bom resultado em diversos trabalhos publicados.

Diante dos resultados, podemos concluir que o uso da Espectroscopia no Infravermelho Médio aliado ao método de Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA) foi;

- capaz de classificar, com eficiência de 100% de acerto as amostras de biodieseis (B100), quanto a matriz energética utilizada em sua produção, tanto metílico quanto etílico, provenientes de soja, soja residual e sebo.

- capaz de classificar, com eficiência de 100% de acerto, as amostras de biodieseis em diesel com teor de 5% (B5), quanto a matriz energética utilizada em sua produção, tanto metílico quanto etílico, provenientes de soja, soja residual e sebo.

- capaz classificar as amostras de misturas de biodieseis metílicos de soja em dieseis de diferentes teores (BX) em conforme ou não conforme de acordo com a legislação vigente, ou seja, se apresentam ou não teor de 5% de biodiesel em diesel.

Portanto, o uso destes modelos é uma alternativa promissora para o monitoramento e fiscalização destes tipos de combustíveis quanto ao tipo de óleo ou gordura utilizado em sua produção, quanto ao tipo de álcool utilizado em sua transesterificação e quanto ao teor da mistura biodiesel/diesel determinado em lei. Ainda, dentre suas vantagens, podemos

propor o uso dessa metodologia para utilização em equipamentos portáteis, para análises “*in situ*”, inclusive feitas por pessoal não especializado, pois a interpretação dos resultados obtidos pelos modelos construídos utilizando PLS-DA são de fácil interpretação, facilitando ainda mais o monitoramento e fiscalização sobre os biocombustíveis.

10. Referências Bibliográficas.

10. Referências Bibliográficas.

- [1] PETERSON, C. L.; HUSTRULID, T. Carbon cycle for rapeseed oil biodiesel fuels. *biomass and bioenergy*, v. 14, n. 2, p. 91-101, março 1998.
- [2] CORDEIRO, C. S.; SILVA, F. R. DA; WYPYCH, F.; RAMOS, L. P. Catalisadores heterogêneos para a produção de monoésteres graxos (biodiesel). *Química Nova*, v. 34, p. 477-486, 2011.
- [3] ATADASHI, I.M; AROUA, M. K; AZIZ, A. R. A; SULAIMAN, N. M. N; The effects of water on biodiesel production and refining technologies: A review. *Renewable and Sustainable Energy Reviews*, v.16, p. 3456-3470, 2008.
- [4] BALABIN, R. M; SAFIEVA, R. Z; LOMAKINA, E.I; Analysis of biodiesel density, kinematic viscosity, methanol and water contents using near infrared (NIR) spectroscopy. *Fuel*, v.90, p.2007–2015, 2011.
- [5] BALABIN, R. M; SMIRNOV, S. V. Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel. *Analytica Chimica Acta*, v.692, p.63–72, 2011.
- [6] BALAT, M. Potential alternatives to edible oils for biodiesel production – A review of current work. *Energy Conversion and Management*, v.52, p.1479–1492. 2011.
- [7] FERRÃO, M. F.; VIERA, M. S.; PAZOS, R. E. P.; FACHINI, D.; GERBASE, A. E.; MARDER, L. Simultaneous determination of quality parameters of biodiesel/diesel blends using HATR-FTIR spectra and PLS, iPLS or siPLS regressions. *Fuel*, London, v.90, p.701–706, 2011.
- [8] MONTEIRO, M. R; AMBROZINA, A. R. P; LIÃO, L. M; FERREIRA, A. G. Critical review on analytical methods for biodiesel characterization. *Talanta*, v.77, p.593-605, 2008.

- [9] ÂGENCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS – ANP, Resolução nº14, de 11 de maio de 2012. Regulamenta inserção do biodiesel na matriz energética brasileira. Disponível em: <http://nxt.anp.gov.br>. Acessado: 19/12/ 2013.
- [10] OLIVEIRA, F. C. C.; BRANDÃO, C. R. R.; RAMALHO, H. F.; COSTA, L. A. F.; SUAREZ, P. A. Z.; RUBIM, S. J. C. Adulteration of diesel/biodiesel blends by vegetable oil as determined by Fourier transform (FT) near infrared spectrometry and FT-Raman spectroscopy, *Analytica Chimica Acta*, v. 587, p. 194-199, 2007.
- [11] KILLNER, M. H. M.; ROHWEDDER, J. J. R.; PASQUINI, C. A PLS regression model using NIR spectroscopy for on-line monitoring of the biodiesel production reaction, *Fuel*, v. 90, p. 3268-3273, 2011.
- [12] LAVINE, B.; WORKMAN, J. Chemometrics, *Analytical Chemistry*, v. 80, p. 4519-4531, 2008.
- [13] 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, Julho, 2011.
- [14] LIMA, A. L.; LIMA, A. P.; PORTELA, F.; SANTOS, D.Q.; BORGES NETO, W.; HERNANDEZ-TERRONES, M. G. Parâmetros da transesterificação etílica com óleo de milho para produção de biodiesel. *Eclética Química* (UNESP. Araraquara. Impresso), v. 35, p. 101-106, 2011.
- [15] PEREIRA, N. R. Estudos oleoquímicos e obtenção de ésteres metílicos e etílicos a partir de óleos do dipteryx ala vog (baru). Dissertação.(Mestrado em Química)-Instituto de Química, Universidade Federal de Uberlândia, Uberlândia, 2010.
- [16] SILVA, T. A. R.; NETO, W. B. Estudo da redução da acidez do óleo residual para a produção de biodiesel utilizando o planejamento fatorial fracionado. *Revista Virtual de Química*, v. 5, p. 828-839, 2013.

- [17] SANTOS, D. Q.; VIEIRA, D. F. Determinação de coeficiente de expansão térmica do biodiesel e seus impactos no sistema de medição volumétrico. *Eclética Química*, v. 35, p.107-112, 2010.
- [18] SILVA, T. A. R.; NETO, W. B.; SANTOS D.Q. ; LIMA, A. P. Model for the correction of the specific gravity of biodiesel from residual oil. *Independent Journal of Management & Production*, v. 4, p. 136-147, 2013.
- [19] GONTIJO, L. C.; GUIMARÃES, E.; MITSUTAKE, H.; SANTANA, F. B.; SANTOS, D. Q.; NETO,W. B. Quantification of soybean biodiesels in diesel blends according to ASTM e 1655 using mid-infrared spectroscopy and multivariate calibration, *Fuel*. v.117, p.1111-1114, 2014.
- [20] FONSECA, F.S. Espectrometria no infravermelho médio e calibração multivariada por PLS para quantificação de biodiesel metílico e etílico de soja, milho e girassol em misturas com diesel. Dissertação. (Mestrado em Química)-Instituto de Química, Universidade Federal de Uberlândia, Uberlândia, 2011.
- [21] KNOTHE,G.; K. R. STEIDLEY. A comparison of used cooking oils A very heterogeneous feedstock for biodiesel, *Bioresource Technology*. v.100, p. 5796-5801, 2009.
- [22] Ministério da Agricultura, Pecuária e Abastecimento. Anuário estatístico de agro energia 2012. Secretaria de Produção e Agroenergia. Brasília: MAPA/ACS, 2013. 284 p.
- [23] Biodieselbr. Disponível em: <http://www.biodieselbr.com/energia/demanda-energia.htm>. Acessado: 02/02/ 2014.
- [24] ALVES, J. C.; POPPI, R. J. Biodiesel content determination in diesel fuel blends using near infrared (NIR) espectroscopy and support vector machines (SVM). *Talanta*, v.104, p.155-161, 2013.

[25] SABUDAK, T.; YILDIZ, M. Biodiesel production from waste fryin oils and its quality control, *Waste Management*, v.30, p.799-803, 2010.

[26] ABIOVE Associação Brasileira das Indústrias de Óleos Vegetais. Disponível em: www.abiove.com.br Acessado em: 30/11/2013.

[27] SABESP Companhia de Saneamento básico do Estado de São Paulo. Disponível em: www.sabesp.com.br/ Acessado em: 25/11/2013.

[28] MITTELBACH, M.; TRITTHART, P. Diesel fuel derived from vegetable oils, III. Emission tests using methyl esters of used frying oil. *Journal of the American Oil Chemists' Society*, v. 65, n. 7, p. 1185-1187, julho 1988.

[29] FERRARI, R. A.; OLIVEIRA, V. DA S.; SCABIO, A. Biodiesel de soja - taxa de conversão em ésteres etílicos, caracterização físico-química e consumo em gerador de energia. *Química Nova*, v. 28, n. 1, p. 19-23, 2005.

[30] MA, F.; HANNA, M.A. Biodiesel production, A review. *Bioresource Technology*, v.70, p. 1-15, 1999.

[31] QUINTELLA, C. M.; TEIXEIRA, L. S. G.; KORN, M. G. A. *et al.* Cadeia do biodiesel da bancada à indústria: uma visão geral com prospecção de tarefas e oportunidades para P&D&I. *Química Nova*, v. 32, n. 3, p. 793-808, 2009.

[32] COSTA NETO, P. R.; ROSSI, R. F. S. Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em frituras. *Química Nova*, v. 23, p.531-534, 2000.

[33] HOLANDA, A. Biodiesel e inclusão social. Brasília: Câmara dos Deputados, coordenação de publicações, 2005. disponível em <<http://www2.camara.leg.br/a->

camara/altosestudos/pdf/biodiesel-e-inclusao-social/biodiesel-e-inclusao-social>.

(Acessado 21 de fevereiro, 2014).

[34] LÔBO, I. P.; FERREIRA, S. L. C.; CRUZ, R. S. D.; Biodiesel: Parâmetros de qualidade e métodos analíticos, *Química Nova*, v.32, p.1596-1608, 2009.

[35] LEUNG, D. Y. C.; WU, X.; LEUNG, M. K. H. A review on biodiesel production using catalyzed transesterification. *Applied Energy*, v. 87, p. 1083-1095, 2010.

[36] FERNANDES, D. M. Obtenção e caracterização físico-química dos biodieseis metílico e etílico de algodão. Dissertação. (Mestrado em Química)-Instituto de Química, Universidade Federal de Uberlândia, Uberlândia, Julho, 2011.

[37] SKOOG, D.A.; HOLLER, f. J.; NIEMAN, T. A. *Princípios de análise instrumental*. 5. ed. são Paulo: Bookman, 2002.

[38] BARBOSA, Luiz Cláudio de Almeida. *Espectroscopia no infravermelho: na caracterização de composto orgânicos*. Ed. da UFV, 2007, Viçosa, MG.

[39] PATACA, L. C. M. Análise de mel e própolis utilizando métodos químiométricos de classificação e calibração. Tese de Doutorado. Instituto de Química. Departamento de Química Analítica. Universidade Estadual de Campinas, Campinas, 2006

[40] SOARES, I. P.; RUSSO, R. M. DE O.; PRATES, R. G. D. *et al.* Avaliação da eficiência das técnicas ESI-MS e ATR/FTIR na determinação de adulteração de bx com querosene e óleo. *Química Nova*, v. 34, n. 8, p. 1439-1442, 2011.

[41] SOARES, I. P.; REZENDE, T. F.; SILVA, R. C.; CASTRO, E. V. R.; FORTES, I. C. 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, n. 3, p. 2079-2083, maio 2008.

- [42] VANDEGINSTE, B. G. M.; MASSART, D. L.; BUYDENS, L. M. C.; JING, S.; LEWI, P. J.; SMEYERS-VERBEKE, J. *Handbook of chemometrics and qualimetrics: part B*. Amsterdam: Elsevier, 1998.
- [43] FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, L. O. Quimiometria I: calibração multivariada, um tutorial. *Química Nova*, v.22, p.724–731, 1999.
- [44] LIU S., KOKOT S., WILL G. J. Photochem. Photobio. C: *Photochemistry Reviews*, v. 10, p.159-172, 2009.
- [45] BENNINGHOFF L., CZARNOWSKI D., DENKHAUS E., LEMKEM K. Analysis of human tissues by total reflection X-ray fluorescence. Application of chemometrics for diagnostic cancer recognition. *Spectrochim. Acta B*, v.52, p. 1039-1046, 1997.
- [46] MARTENS, H.; NAES, T. *Multivariate calibration*. New York: Wiley, 1996.
- [47] GELADI, P.; KOWALSKI, B. Partial Least-Squares Regression: A Tutorial. *Analytical Chimica Acta*, v.185, p. 1, 1986.
- [48] SOUZA, A, M; POPPI, R, J. Experimento didático de Quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: um tutorial, parte I. *Química Nova*, v. 35, n. 1, p.223-229, 2012.
- [49] DE JONG, S. SIMPLS - An alternative approach to partial least-squares regression. *Chemometrics Intell. Lab. Syst*, v.18, p. 251-263, 1993.
- [50] BEEBE, K. R.; KOWALSKI, B. R. An Introduction to Multivariate Calibration and Analysis, *Anal. Chem*, v.59, p.1007A-1017A, 1987.

[51] MILLER, J. N.; MILLER, J. C. *Statistics and Chemometrics for Analytical Chemistry*. London: Prentice Hall, 2000.

[52] OTTO, M. *Chemometrics – Statistics and Computer application in Analytical Chemistry*, Weinheim: Wiley, 1999.

[53] BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, B. *Chemometrics- A Practical Guide*. New York: Wiley, 1998.

[54] MASSART, D. L.; VANDEGINSTE, B. G. M.; DEMING, S. N.; MICHOTTE, Y.; KAUFMAN, L. *Chemometrics: a textbook*. In: *Data handling in science and technology*. 1st edition. Amsterdam: Elsevier Science Publisher B.V., 1988. cap. Supervised Pattern Recognition, p. 385-413.

[55] WOLD, S. Pattern recognition by means of disjoint principal component models, *Pattern Recognition*, handbook. Capítulo 33. p. 203-241. 1976.

[56] TENENHAUS M., VINZIA V. E., CHATELINC Y. M., LAUROB C. PLS path modeling. *Comput. Stat. Data An*, v. 48, p.159-205, 2005.

[57] WOLD S., RUHE A., WOLD H., DUNN W. J. The collinearity problem in linear regression: the partial least squares (PLS) approach to generalized inverses. *SIAM, J. Sci. Stat. Comp*, v.5, p.735-743, 1984.

[58] M. R. BARKER, W., Partial least squares for discrimination. *Journal of Chemometrics*, v.17, p.166-173, 2003.

[59] B.M. WISE, N.B. GALLAGHER, R. BRO, J.M. SHAVER, W. WINDIG, R.S. KOCH, *Chemometris Tutorial for PLS_Toolbox and Solo*, Eigenvector Research, Inc., 3905 West Eaglerock Drive, Wenatchee, WA 98801 USA, 2006.

[60] DE ALMEIDA, M.R., CORREA, D.N., ROCHA, W.F.C., SCAFI, F.J.O. & POPPI, R.J. Discrimination between authentic and counterfeit banknotes using Raman spectroscopy and PLS-DA with uncertainty estimation. *Microchemical Journal*, v.109, p.170-177, 2013.

[61] NETO, W. B. Parâmetros de Qualidade de Lubrificantes e óleo de oliva através de espectroscopia vibracional, calibração multivariada e seleção de variáveis. [S.l.]: Universidade Estadual de Campinas, 2005.

[62] ÂGENCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS – ANP, Resolução nº50, de 23 de dezembro de 2013. Regulamenta as especificações do óleo diesel de uso rodoviário, e as obrigações quanto ao controle da qualidade a serem atendidas pelos diversos agentes econômicos que comercializam o produto em todo o território nacional. Disponível em: <http://nxt.anp.gov.br>. Acessado: 19/12/ 2013.

[63] CAVALCANTE, R. M. Predição da densidade de biodiesel proveniente de diferentes matérias-primas. Dissertação. (Mestrado em Tecnologia de Processos Químicos e Bioquímicos), Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010

[64] PETROBRÁS. Densidade e poderes caloríficos superiores. Disponível em: <http://www.investidorpetrobras.com.br/pt/servicos/formulas-de-conversao>. Acessado em : 20/01/2014.