

**UNIVERSIDADE DE SANTA CRUZ DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS
E PROCESSOS INDUSTRIAIS – MESTRADO
ÁREA DE CONCENTRAÇÃO EM
CONTROLE E OTIMIZAÇÃO DE PROCESSOS**

Marlize Rizzi

**APLICAÇÃO DO MÉTODO ENXAME DE PARTÍCULAS NA OTIMIZAÇÃO DE MODELOS
DE REGRESSÃO MULTIVARIADA APLICADOS NA QUANTIFICAÇÃO DA MISTURA
BODIESEL/ÓLEO VEGETAL/DIESEL**

Santa Cruz do Sul, março de 2008.

Marlize Rizzi

**APLICAÇÃO DO MÉTODO ENXAME DE PARTÍCULAS NA OTIMIZAÇÃO DE MODELOS
DE REGRESSÃO MULTIVARIADA APLICADOS NA QUANTIFICAÇÃO DA MISTURA
BODIESEL/ÓLEO VEGETAL/DIESEL**

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

Orientador: Prof. Dr. João Carlos Furtado

Co-orientador: Prof. Dr. Marco Flôres Ferrão

Santa Cruz do Sul, março de 2008.

AGRADECIMENTOS

Inicialmente agradeço a Deus pela vida, e por estar sempre comigo.

Aos meus pais Avelino e Geni, por me darem carinho, amor, educação e força em todos os momentos de minha vida. Amo vocês.

Ao futuro papai e meu amor Éverton Camargo da Silva pela paciência, compreensão e incentivo. Amo você.

Ao meu filho Gabriel, que está por nascer, que soube colaborar para que juntos pudéssemos concluir esta etapa tão importante em nossas vidas.

Ao meu orientador, prof. Dr. João Carlos Furtado, pela amizade, pelos ensinamentos e a disponibilidade apresentada.

Ao meu co-orientado prof. Dr. Marco Flôres Ferrão pelas muitas orientações, pelo senso de direção, pela disponibilidade e pela oportunidade concedida.

A amiga Márcia Werlang pelas ótimas horas de estudos e a pela sua importante colaboração nesta dissertação.

As secretárias Janaina Ramires Haas e Camila Borges Corrêa pela eficiência e presteza sempre apresentada.

Ao mestre Gustavo Post Sabin, por ter compartilhado comigo seus conhecimentos nesta área.

A CAPES e a FAPERGS pelo financiamento do projeto “casadinhos” e, pela colaboração do Instituto de Química e do Laboratório de Combustíveis da Universidade Federal do Rio Grande do Sul - UFRGS.

A todos aqueles que embora não tenham sido mencionados contribuíram direta ou indiretamente para a realização deste trabalho.

“A imaginação é tudo. É uma prévia das próximas atrações da vida.”

Albert Einstein

RESUMO

A substituição gradativa dos produtos derivados de fontes não renováveis, por produtos de origem renováveis, vem sendo mundialmente discutida, especialmente quanto aos combustíveis. No Brasil, o governo oficializou a mistura biodiesel/diesel nos percentuais obrigatórios de 2% a partir de 2008. Considerando a necessidade de monitoramento da qualidade desta mistura, esta dissertação propõe o estudo e o uso do método enxame de partículas (*Particle Swarm Optimization* – PSO), juntamente com a técnica de regressão multivariada de mínimos quadrados parciais por intervalos (iPLS), visando constituir um algoritmo híbrido capaz de selecionar variáveis mais representativas do constituinte a ser quantificado, buscando gerar modelos de regressão mais robustos. Nesta pesquisa efetuou-se a quantificação de biodiesel, do óleo vegetal e do óleo diesel, utilizando toda a região espectral das misturas, as regiões espectrais sugeridas pela proposta de norma ABNT e as regiões escolhidas pelos algoritmos híbridos. Na quantificação do biodiesel o menor erro de previsão obtido foi através do método híbrido, sendo este erro 61,54% menor que o erro encontrado pelas regiões sugeridas pela proposta de norma ABNT e, 26,72 % menor que o erro encontrado utilizando toda a região espectral das misturas. Já, na quantificação do óleo vegetal o menor erro de previsão, obtido também foi pelo método híbrido, ficando 27,64% menor do que o erro encontrado pelas regiões sugeridas pela proposta de norma ABNT e, 38,62% menor do que erro encontrado utilizando toda a região espectral das misturas. Porém na quantificação do óleo diesel os métodos apresentaram desempenhos semelhantes. Com base nestes resultados, pode-se concluir que a utilização do método enxame de partículas juntamente com o método iPLS, foi capaz de otimizar as soluções, selecionando variáveis do espectro para construção de modelos mais robustos.

Palavras-chave: enxame de partículas, regressão por mínimos quadrados parciais, biodiesel, otimização combinatória, espectroscopia no infravermelho.

ABSTRACT

The gradual substitution of products coming from non-renewable sources by products of renewable origin has been globally discussed lately, especially regarding to the fuels. In Brazil, the government has declared official the biodiesel/diesel mixture in the obligatory percentile of 2% starting from 2008. This dissertation based on the need for monitoring such mixture's quality proposes the study and the use of the particles swarm method (Particle Swarm Optimization - PSO), together with the interval partial least-squares multivariate regression technique (iPLS). It presents as its aims to constitute a hybrid algorithm capable to select more representative variables of the representative to be quantified, seeking to produce more robust regression models. In this research, it was done the biodiesel, the vegetable oil, and the diesel oil quantification, using (a) the whole spectral area of the mixtures, (b) the spectral areas suggested by the norm proposal from the Brazilian Association of technical rules (ABNT), and (c) the areas chosen by the hybrid algorithms. In the biodiesel quantification, the smallest foreseen mistake obtained was through the hybrid method, being this mistake 61.54% smaller than the mistake found by the areas suggested by the norm proposal from ABNT and, 26.72% smaller than the found mistake using the whole spectral area of the mixtures. Yet in the vegetable oil quantification the smallest foreseen mistake, also obtained through the hybrid method, was 27.64% smaller than the mistake found by the areas suggested by the norm proposal from ABNT and, 38.62% smaller than mistake found using the whole spectral area of the mixtures. However, in the diesel oil quantification the methods have presented similar performance. Based on these results, it was possible to conclude that the use of particles swarm method in association with the method iPLS, was capable to optimize the solutions, selecting the spectrum's variables for the construction of more robust models.

Key-words: particles swarm, partial least-squares regression, biodiesel, combinatorial optimization, infrared spectroscopy.

LISTA DE ILUSTRAÇÕES

Figura 1 – Esquema de transesterificação de um triacilglicerídeo.....	19
Figura 2 – Representação da propagação da radiação infravermelho através do elemento de ATR.	22
Figura 3 - Ilustração do método de enxame de partículas.....	26
Figura 4 – Fluxograma das etapas executadas na parte experimental.	34
Figura 5 – Diagrama ternário para as misturas biodiesel/óleo vegetal/diesel.	35
Figura 6 – Distribuição das amostras de calibração e previsão para o biodiesel.....	39
Figura 7 – Distribuição das amostras de calibração e previsão para o óleo diesel.....	39
Figura 8 – Distribuição das amostras de calibração e previsão para o óleo vegetal.....	40
Figura 9 – Representação gráfica de duas partículas.	41
Figura 10 – Exemplo de uma partícula e representação dos seus intervalos selecionados .	42
Figura 11 – Região espectral utilizada na modelagem para a quantificação do biosiesel. ...	46
Figura 12 – Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel utilizando toda a região espectral das misturas.....	47
Figura 13 – Regiões espectrais sugeridas pela proposta de norma ABNT.....	47
Figura 14– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT.....	48
Figura 15 – Exemplo da evolução das partículas para a quantificação do biodiesel, usando o método iPLS-PSO (CD) com 40 intervalos e 100 partículas.	50
Figura 16 – Evolução das três execuções do iPLS-PSO com 20 intervalos e 50 partículas, para quantificação do biodiesel.....	51
Figura 17 – Evolução das três execuções do iPLS-PSO com 40 intervalos e 50 partículas, para quantificação do biodiesel.....	52
Figura 18 – Evolução das três execuções do iPLS-PSO com 60 intervalos e 50 partículas, para quantificação do biodiesel.....	52
Figura 19 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel usando o método iPLS-PSO com 60 intervalos e 100 partículas.	53
Figura 20 – Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 60 intervalos e 100 partículas.	54

Figura 21 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.....	55
Figura 22 – Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.....	55
Figura 23– Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	56
Figura 24– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	57
Figura 25 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.....	58
Figura 26– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.....	58
Figura 27 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO com 40 intervalos e 50 partículas.	60
Figura 28– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 40 intervalos e 50 partículas.	60
Figura 29 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal utilizando toda a região espectral das misturas.	61
Figura 30 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT 62	
Figura 31– Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas, para a quantificação do óleo vegetal.....	64
Figura 32 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas, para a quantificação do óleo vegetal.....	65
Figura 33 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para a quantificação do óleo vegetal.....	65
Figura 34 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO com 40 intervalos e 25 partículas.....	66
Figura 35 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.....	67
Figura 36 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.....	68
Figura 37 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.	68
Figura 38 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	69

Figura 39 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.	70
Figura 40 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.....	71
Figura 41– Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.	71
Figura 42 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.....	72
Figura 43 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.	73
Figura 44 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel utilizando toda a região espectral da mistura.	74
Figura 45 – Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas para a quantificação do óleo diesel.....	76
Figura 46 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas para quantificação do óleo diesel.....	77
Figura 47 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para quantificação do óleo diesel.....	77
Figura 48 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO com 40 intervalos e 25 partículas.	78
Figura 49 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.	79
Figura 50 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.....	80
Figura 51 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.....	80
Figura 52 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	81
Figura 53 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	82
Figura 54 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.....	83
Figura 55 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.....	83
Figura 56 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DC) com 40 intervalos e 25 partículas.....	84

- Figura 57 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 50 partículas..... 85
- Figura 58 – Esquema contendo espectro de todas as misturas, do óleo diesel, do biodiesel e do óleo vegetal. 87
- Figura 59 – Erros de previsão para a quantificação do biodiesel, usando as regiões propostas pelo método iPLS-PSO com 40 intervalos e 50 partículas..... 88
- Figura 60 – Erros de previsão para a quantificação do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT..... 88
- Figura 61 – Valores de RMSECV *versus* RMSEP para diferentes *G_{best}* para quantificação do biodiesel, usando o método iPLS-PSO(CD) com 60 intervalos e 25 partículas. 90
- Figura 62 – Valores de RMSECV *versus* RMSEP para diferentes *G_{best}* para a quantificação do biodiesel, usando o método iPLS-PSO(CD) com 40 intervalos e 100 partículas. 90

LISTA DE TABELAS

Tabela 1- Planejamento experimental das misturas realizado na UNISC	36
Tabela 2– Planejamento experimental das misturas realizado na UFRGS	37
Tabela 3– Parâmetros instrumentais e regiões empregadas na aquisição dos espectros por FT-IR/ATR	38
Tabela 4 – Nomenclatura usada para identificar as implementações	44
Tabela 5– Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do biodiesel.	49
Tabela 6 – Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do óleo vegetal.	63
Tabela 7 – Principais resultados dos modelos, incluindo das cinco implementações da PSO para a quantificação do óleo diesel.	75

LISTA DE ABREVIATURAS

ABNT	Associação Brasileira de Normas Técnicas
AG	Algoritmo Genético
ANP	Agência Nacional de Petróleo, Gás Natural e Biocombustíveis
ATR	<i>Attenuated Total Reflectance</i>
FT-IR	<i>Fourier Transform Infrared Spectroscopy</i>
GC	<i>Gaseous Chromatograph</i>
GPC	<i>Gel Permeation Chromatography</i>
HPLC	<i>High Performance Liquid Chromatography</i>
iPLS	<i>Interval Partial Least-Squares</i>
NIR	<i>Near-infrared Reflectance</i>
PLS	<i>Partial Least-Square</i>
PSO	<i>Particle Swarm Optimization</i>
RMSEC	<i>Root Mean Square Error of Calibration</i>
RMSECV	<i>Root Mean Square Error of Cross Validation</i>
RMSEP	<i>Root Mean Square Error of Prediction</i>
RMSEV	<i>Root Mean Square Error of Validation</i>
TLC	<i>Thin Layer Chromatography</i>
VL's	Variáveis Latentes

LISTA DE SIMBOLOS

i	partícula
x_i	posição da partícula i
n	tamanho do enxame
it	iteração atual
$pbest_i$	melhor posição que a partícula i já obteve durante a busca
$Gbest$	melhor posição encontrada pelas partículas do enxame
c_1, c_2	parâmetros de confiança
v_i	velocidade atual da partícula i
V_{max}	limitador de velocidade
w	fator de inércia
w_{ini}	valor inicial para o fator de inércia
w_{fin}	valor final para o fator de inércia
k	coeficiente de constrição
R	nº total de iterações
$rand()$	função aleatória
$(pbest_i - x_i)$	distância entre a melhor posição encontrada para a partícula i e a posição atual dessa mesma partícula
$(Gbest - x_i)$	distância entre a melhor posição de todas as partículas do enxame e a posição atual da partícula i

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	Justificativa	16
1.2	Objetivos.....	16
1.2.1	Objetivo Geral.....	16
1.2.2	Objetivos Específicos.....	16
1.3	Organização do texto.....	17
2	FUNDAMENTAÇÃO TEÓRICA.....	18
2.1	Biodiesel	18
2.2	Espectroscopia no infravermelho com transformada de Fourier (FT-IR)	20
2.3	Principais técnicas de reflexão no infravermelho.....	21
2.3.1	Reflexão total atenuada	21
2.4	Análise multivariada.....	22
2.5	Regressão por mínimos quadrados parciais (PLS)	23
2.6	Regressão por mínimos quadrados parciais por intervalos (iPLS)	25
3	EXAME DE PARTÍCULAS	26
3.1	Introdução.....	26
3.2	O algoritmo básico de PSO.....	27
3.3	Estrutura do exame de partículas.....	28
3.4	Outros parâmetros do PSO.....	28
3.4.1	Tamanho da população	28
3.4.2	Velocidade Máxima.....	29
3.4.3	Fator de Inércia.....	29
3.4.4	Coeficiente de constrição.....	31
3.5	Outro tipo de representação	31
3.5.1	PSO binário	31
3.6	Algumas aplicações da técnica exame de partículas	32
4	METODOLOGIA	34
4.1	Preparo das misturas	35
4.2	Aquisição dos espectros	37
4.3	Tabulação das matrizes	38
4.4	Modelagem dos dados.....	38
4.4.1	Considerações gerais sobre as implementações da meta-heurística exame de partículas para selecionar variáveis de espectros no infravermelho	40
4.4.2	Estrutura básica do exame de partículas implementado para a otimização dos modelos de regressão	40
5	RESULTADOS	46
5.1	Resultados para a quantificação do biodiesel	46
5.1.1	Modelos para a determinação do biodiesel com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT	46
5.1.2	Resultados gerais e evoluções para a determinação do biodiesel	48
5.1.3	Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO	53
5.1.4	Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CC).....	54

5.1.5 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (DD).....	56
5.1.6 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CD).....	57
5.1.7 Resultados obtidos para a determinação do biodiesel aplicando o iPLS-PSO (DC)...	57
5.1.8 Considerações sobre os modelos para a previsão do biodiesel.....	59
5.2 Resultados para a quantificação do óleo vegetal	61
5.2.1 Modelos para determinação do óleo vegetal com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT.....	61
5.2.2 Resultados gerais e evoluções dos modelos para a previsão do óleo vegetal	62
5.2.3 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO.....	66
5.2.4 Resultados obtido para a determinação do óleo vegetal aplicando o método iPLS-PSO (CC).....	67
5.2.5 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DD)	69
5.2.6 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (CD)	70
5.2.7 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DC)	72
5.2.8 Considerações sobre os modelos para a previsão de óleo vegetal.....	73
5.3 Resultados para a quantificação do óleo diesel	74
5.3.1 Modelo para a determinação do óleo diesel com espectro inteiro.....	74
5.3.2 Resultados gerais e evoluções para a determinação do óleo diesel	75
5.3.3 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO	78
5.3.4 Resultados obtidos para a quantificação do óleo diesel aplicando o método iPLS-PSO (CC).....	79
5.3.5 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DD).....	81
5.3.6 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (CD).....	82
5.3.7 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DC).....	84
5.3.8 Considerações sobre os modelos de óleo diesel	85
5.4 Considerações finais sobre os modelos de biodiesel, óleo vegetal e óleo diesel	86
6 CONCLUSÕES.....	91
7 TRABALHOS FUTUROS.....	92
REFERÊNCIAS	93
ANEXOS.....	98

1 INTRODUÇÃO

Atualmente, a substituição parcial do óleo diesel por combustíveis renováveis está sendo efetuada mundialmente, e os óleos provenientes de fontes vegetais são as principais alternativas para se efetuar esta substituição.

No Brasil tornou-se obrigatório, desde 1º de janeiro de 2008, a adição de 2% de biodiesel (B2) ao diesel; com perspectivas deste biocombustível ser produzido, nos próximos anos, em grande escala para atender a demanda de mercado. Desta forma, como é feito atualmente com outros combustíveis, será necessário monitorar a qualidade desta mistura para garantir que o produto a ser comercializado não lese o consumidor, garanta o perfeito funcionamento dos veículos automotores e contribua realmente para conservação do meio ambiente.

Neste sentido, são crescentes os estudos que buscam desenvolver metodologias alternativas as metodologias tradicionais de quantificação de combustíveis. A técnica de espectroscopia no infravermelho é uma alternativa que está sendo estudada nesta área, devido a possibilidade de realizar análises não destrutivas, mais rápidas, sem gastos de reagentes e gerando informações com qualidade. Associadas a esta técnica são utilizadas ferramentas quimiométricas de regressão para construção de modelos.

Por outro lado, a otimização por enxame de partículas é um método de otimização recente (1995), baseado na analogia do comportamento social dos pássaros. Diversos estudos estão sendo desenvolvidos com esta técnica movidos pela sua simplicidade, robustez e eficiência.

Essa dissertação desenvolveu um método híbrido para quantificar o biodiesel, óleo vegetal e diesel nas misturas. Método este composto pelo método iPLS (*Interval Partial Least-Squares*) otimizado pela meta-heurística enxame de partículas, visando selecionar as variáveis mais representativas para construção de modelos de regressão mais robustos.

1.1 Justificativa

Tendo em vista que o processo de adição do biodiesel (2%) ao diesel está iniciando este ano em nosso país, e as perspectivas são bastante amplas, as indústrias terão que se estruturar de maneira a poder controlar a qualidade bem como evitar a adulteração deste biocombustível. Neste contexto, a espectroscopia no infravermelho associada a ferramentas quimiométricas, pode ser uma solução alternativa aos métodos tradicionais.

O conjunto de dados obtidos neste tipo de misturas biodiesel/diesel apresenta um número extremamente elevado de variáveis, bem como as combinações resultantes possíveis. Com o propósito de selecionar um conjunto de variáveis que possibilitem a construção de modelos com habilidade equivalente na calibração e na previsão, sem a necessidade de avaliar todas as soluções possíveis optou-se por empregar a técnica de otimização chamada de enxame de partículas para efetuar esta seleção.

1.2 Objetivos

1.2.1 Objetivo Geral

O objetivo geral deste trabalho foi verificar a eficiência do método enxame de partículas para selecionar variáveis capazes de otimizar os modelos de análise multivariada, aplicados na quantificação das misturas biodiesel/óleo vegetal/diesel.

1.2.2 Objetivos Específicos

- Estudar métodos de otimização empregados no processo de quantificação de matérias-primas e combustíveis.
- Quantificar misturas de biodiesel, óleo vegetal e diesel, através de modelos de regressão.
- Comparar os resultados obtidos pelo modelo desenvolvido com outros descritos na literatura.

1.3 Organização do texto

O próximo capítulo introduz alguns aspectos importantes para compreensão e desenvolvimento deste trabalho, abrangendo uma visão geral do biodiesel, informações sobre espectroscopia no infravermelho com transformada de Fourier (FT-IR/ATR) e sobre os métodos de regressão que serão utilizados. Já no terceiro capítulo é apresentada a meta-heurística enxame de partículas implementada para otimização dos modelos.

O quarto capítulo descreve a metodologia utilizada na construção desta dissertação iniciando com a elaboração das misturas e a obtenção dos espectros, juntamente com uma descrição detalhada da ferramenta desenvolvida iPLS-PSO e suas variações.

No quinto capítulo são apresentados os resultados obtidos na quantificação das misturas, juntamente com uma análise destes resultados. No sexto capítulo são apresentadas às conclusões deste trabalho e, finalmente no sétimo capítulo são apresentadas sugestões de futuros trabalhos.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Biodiesel

De toda a energia consumida mundialmente, a maior parte vem do petróleo, do carvão e do gás natural, fontes estas limitadas e com previsão de se esgotar no futuro; portando a busca por fontes alternativas de produção de energia é de suma importância. Neste contexto surgem inúmeros estudos, que envolvem a produção de combustíveis alternativos ao óleo diesel, proveniente de óleos vegetais, denominados biocombustíveis (NASCIMENTO, COSTA NETO, MAZZUCO, 2001). Os primeiros estudos envolvendo o uso de óleos vegetais como combustível foram realizados na época do desenvolvimento dos motores a diesel, porém neste período a questão ambiental não era considerada importante. Atualmente, quando as atenções se voltam aos problemas ambientais e a redução da emissão dos gases causadores do efeito estufa, a busca por fontes de produção de combustíveis “limpos” tem sido constante no mundo.

No Brasil, a lei nº 11.097, de 13 de janeiro de 2005, determina percentuais de mistura do biodiesel ao óleo diesel, bem como o monitoramento da inserção deste novo combustível ao mercado, pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. A adição obrigatória do percentual de 2% (B2) iniciou em janeiro deste ano; e a meta é que seja adicionado 5% (B5) a partir de 2013. Segundo esta lei, o biodiesel é um biocombustível derivado de biomassa renovável para uso em motores a combustão interna com ignição por compressão ou, conforme regulamento, para geração de outro tipo de energia, que possa substituir parcial ou totalmente combustíveis de origem fóssil (BRASIL, 2005). Podem ser obtidos através de óleos vegetais ou gorduras animais, novos ou residuais, a partir de processos tecnológicos para conversão de óleos em combustíveis.

Dentre as reações empregadas na produção do biodiesel a partir de óleos vegetais, destaca-se a reação de transesterificação, onde triacilglicerídeos reagem com álcool na presença de uma base ou ácido forte, produzindo uma mistura de ésteres de ácidos graxos e glicerol, conforme ilustrado na Figura 1 (GERIS *et al.*, 2007).

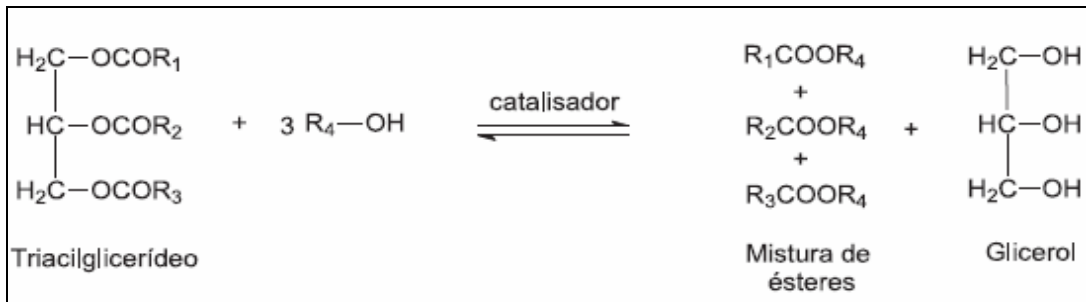


Figura 1 – Esquema de transesterificação de um triacilglicerídeo.
 Fonte: GERIS *et al.*, 2007.

O emprego do biodiesel em substituição total ou parcial ao diesel (pois é um produto miscível) proporciona benefícios ao meio ambiente, pois o CO₂ liberado é de origem de fonte renovável; sociais, pois possibilita a geração de emprego e renda em comunidades agrícolas e; econômicos tendo em vista que o diesel é um dos combustíveis mais consumidos no Brasil, pois a maior parte do transporte de cargas é realizado por meio terrestre (OLIVEIRA, 2006).

Entretanto, um dos problemas dos óleos vegetais é o preço, que tende a ser mais elevado que os de óleos de origem fósseis. Diversos pontos devem ser levados em consideração na hora de calcular o custo do biodiesel: oferta da matéria-prima, mão-de-obra, preço dos produtos agrícolas, infra-estrutura e a competição da matéria-prima com outros mercados. Um outro ponto que não pode ser esquecido são os “subprodutos” gerados durante o processo de transformação (o farelo, a glicerina e o álcool hidratado) que podem fazer com que o preço do biodiesel aumente para compensar os prejuízos de subprodutos, ou reduzido por causa de lucros com derivados do processo (CEPEA, 2006). Nos EUA, as fontes mais comuns para produção do biodiesel são do óleo de soja e o óleo de frituras reciclado, mas o custo ainda elevado de fabricação, e a disponibilidade limitada de algumas matérias primas usadas, fazem com que a aplicação comercial seja ainda limitada (RADICH, 2004).

O Brasil porém é um dos poucos países com tamanha diversidade na produção de biodiesel, a partir de plantas como soja, babaçu, amendoim, mamona, girassol, dendê e palma (GERIS *et al.*, 2007). Estudos estão sendo feitos para produzir, e aperfeiçoar a produção de biodiesel a partir do reaproveitamento de óleos como: óleos de frituras (COSTA NETO *et al.*, 2000; ENCINAR, GONZÁLEZ, RODRÍGUEZ-REINARES, 2005), óleos de oliva usados (DORADO *et al.*, 2004). Entretanto, independente da matéria prima, o produto final deve ser único e padronizado conforme os requisitos definidos pela Agência Nacional de Petróleo, Gás Natural e Biocombustíveis - ANP.

Diversas metodologias para análise do biodiesel estão sendo abordadas, tentando estabelecer um combustível padrão em diversas regiões e países ao redor do mundo. Estão inseridos neste contexto os Estados Unidos, a Europa, o Brasil, a Austrália e o Sul da África. Dentre as principais metodologias que vem sendo utilizadas para a quantificação de derivados, sub produtos e contaminações, destacam-se a cromatografia gasosa (GC), a cromatografia líquida de alta eficiência (HPLC), a cromatografia por permeação gel (GPC) e cromatografia por camada delgada (TLC) (KNOTHE, 2006). Porém estes métodos além de destrutivos, são laboriosos e consomem reagentes químicos.

Levando em consideração os antecedentes históricos e atuais de adulterações de combustíveis no Brasil, associado a necessidade de análise e controle de qualidade desse biocombustível, metodologias alternativas estão sendo propostas como a utilização da espectrofotometria no infravermelho próximo (NIR), infravermelho médio (FT-IR) e Raman (OLIVEIRA, 2006).

Outro trabalho recente desenvolvido por Pimentel *et al.* (2006) foi a utilização da espectroscopia no infravermelho e calibração multivariada na determinação de biodiesel em misturas com diesel mineral.

2.2 Espectroscopia no infravermelho com transformada de Fourier (FT-IR)

O instrumento utilizado atualmente para a obtenção de um espectro no infravermelho é denominado espectrofotômetro no infravermelho com transformada de Fourier (FT-IR). O modo de funcionamento de um equipamento FT-IR é denominado método interferométrico, pois o componente óptico básico deste equipamento é o interferômetro de Michelson (BARBOSA, 2007).

Dentre as vantagens de utilização deste método pode-se mencionar que: apresenta uma maior sensibilidade ou maior razão sinal/ruído, pois utiliza todos os comprimentos da onda ao mesmo tempo durante a obtenção do espectro; possibilita a obtenção de espectros em poucos segundos com maior resolução e maior precisão nas medidas de número de ondas (BARBOSA, 2007).

2.3 Principais técnicas de reflexão no infravermelho

Dentre os principais métodos de reflexão no infravermelho utilizados na aquisição dos dados espectrais mencionam-se: o método por reflexão difusa, o método por reflexão especular (ou externa) e o método de reflexão total atenuada (ATR), sendo este último o que será utilizado neste trabalho.

2.3.1 Reflexão total atenuada

A reflexão total atenuada (ATR – do inglês *attenuated total reflection*) é uma técnica que possibilita a obtenção de dados espectrais de amostras de maneira rápida, além de exigir uma mínima preparação da amostra. Possibilita analisar materiais líquidos, espessos e muito absorventes. Esta técnica caracteriza-se pelas múltiplas reflexões de radiação infravermelha que ocorrem no interior de cristais, interagindo somente com a amostra que estiver superficialmente no cristal (FERRÃO, 2001).

A reflexão na espectroscopia por ATR ocorre quando o feixe de radiação passa de um meio mais denso (cristal de ATR) para um menos denso (amostra), conforme Figura 2. Distorções que ocorram da banda no espectro podem ser ocasionadas pelo tipo de cristal que está sendo utilizado. Dentre os materiais usados para ATR está o ZnSe. Sua faixa de utilização é de 20.000 até 650 cm^{-1} , sendo recomendado para medidas em soluções aquosas. Outro material é o silício (Si) com uma faixa de utilização de 9.000 até 400 cm^{-1} e indicado para amostras altamente espalhativas. Um problema encontrado é a dificuldade de gerar-se boa reprodutibilidade no contato da amostra com o elemento ATR. Isto pode ser observado na variação da intensidade da banda, que é influenciado pelo aumento da pressão e pela área de contato entre o cristal e a amostra. Para medidas quantitativas, deve-se colocar toda a área do cristal em contato com a amostra (FERRÃO, 2001).

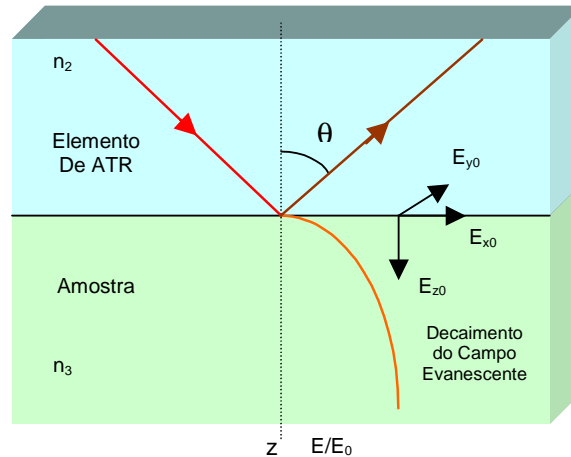


Figura 2 – Representação da propagação da radiação infravermelha através do elemento de ATR.
Fonte: FERRÃO, 2000.

Alguns exemplos de utilização da espectroscopia por reflexão total atenuada: na determinação rápida do índice de cafeína em refrigerantes (PARADKAR, IRUDAYARAJ, 2002) e, na análise de espécies de ácidos gordurosos insaturados em óleos dietéticos (YOSHIDA, YOSHIDA, 2003).

Em função das características das amostras e dos trabalhos que vem sendo realizados foi selecionada a reflexão total atenuada para a aquisição dos espectros das amostras a serem utilizadas neste trabalho.

2.4 Análise multivariada

O desenvolvimento da microeletrônica proporcionou modernos instrumentos de medição, e computadores com elevada capacidade de processamento. Esta união proporcionou um grande avanço nas análises instrumentais, possibilitando a aquisição de um grande número de dados da mesma amostra.

Analisar os dados obtidos é fundamental em todo o experimento realizado. A análise multivariada foi introduzida, ao tratamento dos dados químicos, em virtude da limitação da análise univariada, dando origem a uma área do conhecimento chamada quimiometria (SENA *et al.*, 2000). A análise multivariada refere-se a todos os métodos estatísticos que analisam simultaneamente múltiplas medidas do objeto que está sendo investigado. Pode-se dizer que qualquer análise simultânea de duas ou mais variáveis pode ser considerado

como análise multivariada (HAIR *et al.*, 2005). Nos métodos multivariados, o modelo estatístico considera a correlação entre muitas variáveis simultaneamente, possibilitando obter um número muito maior de informações (SENA *et al.*, 2000).

Quimiometria é uma área multidisciplinar do conhecimento, onde são utilizadas ferramentas matemáticas e estatísticas à análise de dados relativos aos processos químicos. Isto começou a ser amplamente utilizado para analisar amostras, graças a possibilidade de utilizar várias frequências dos espectros. A regressão é uma das mais importantes ferramentas na análise quantitativa. Os procedimentos matemáticos utilizados visam determinar as relações quantitativas existentes entre duas ou mais variáveis (FERRÃO, 2000).

Estas ferramentas de análise multivariada de dados vem sendo amplamente estudadas para aplicação na área industrial, como por exemplo, na análise exploratória na diferenciação química de vegetais (MORGANO *et al.*, 1999); na identificação de diferentes formulações de medicamentos manipulados contendo piroxicam (PARISOTTO *et al.*, 2005), na determinação simultaneamente dos teores de constituintes inorgânicos em sucos e refrigerantes de uva (FERREIRA *et al.*, 2002).

Segundo Gurden, Westerhuis e Smilde (2002), o processo industrial é carente de novas tecnologias que efetuem o monitoramento e o controle de processo em lote. A espectroscopia está se tornando uma opção de se obter em tempo real informações químicas de alta qualidade em intervalos de tempos frequentes em diversas aplicações industriais. Este autor enfatiza o correto manuseio dos dados medidos, a fim de extrair o máximo de informação, visando melhorar a performance e a eficiência do processo.

2.5 Regressão por mínimos quadrados parciais (PLS)

A técnica de regressão por mínimos quadrados parciais (PLS - do inglês *Partial Least Square*) utiliza o espectro por inteiro, ou parte deste, para construção do modelo. A matriz de dados **X**, é decomposta em várias matrizes **M** juntamente com uma matriz de resíduos **E** (KONZEN *et al.*, 2003), conforme Equação 1:

$$\mathbf{X} = \mathbf{M}_1 + \mathbf{M}_2 + \mathbf{M}_3 + \dots + \mathbf{M}_a + \mathbf{E} \quad (1)$$

onde:

a correspondem ao número de variáveis latentes ou componentes principais;

M constituem os componentes principais (ou variáveis latentes), e são compostas por dois vetores, **t** (os escores) e **p** (os pesos), conforme as Equações 2 e 3:

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

$$\mathbf{Y} = \mathbf{UQ}^t + \mathbf{F} \quad (3)$$

onde:

X é a matriz de dados (medida instrumental);

Y é a matriz de concentração;

T e **U** são os escores de **X** e **Y** respectivamente;

P e **Q** são os *loadings* de **X** e **Q**, e

E e **F** são os erros de modelagem de **X** e **Y**.

Para determinar o número de variáveis latentes (VL's) que devem ser empregadas na construção de modelo, não há um método exato, por isso os pesquisadores normalmente utilizam um aumento crescente nas VL's e verificam o comportamento da modelagem (KONZEN *et al.*, 2003).

O modelo construído é considerado eficiente, se for capaz de descrever da melhor maneira a situação real a que é proposto, levando em consideração o maior número possível de variações. Usualmente, os modelos são avaliados a partir da correlação entre a técnica proposta e o método de referência, com base nos valores de correlação e dos erros padrões de calibração (RMSEC) e validação (RMSEV), sendo os erros calculados conforme Equação 4:

$$\text{RMSEC, RMSEV, RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

onde:

y_i e \hat{y}_i corresponde aos valores de referência e estimado respectivamente para a *i*-ésima amostra e, **n** número total de amostras.

Outro erro utilizado para avaliar o modelo é o RMSECV que é obtido quando se utiliza na validação o processo de validação cruzada (*crossvalidation*), que consiste em

retirar uma amostra de cada vez do conjunto de calibração ao longo da validação. A equação para obtenção deste erro é análoga a equação 4, sendo que o i neste caso corresponde a amostra retirada do conjunto e n número total de amostras utilizadas na construção do modelo.

Na área de controle de qualidade e adulterações de combustíveis utilizando o PLS pode-se mencionar alguns trabalhos como: a análise das misturas biodiesel-óleos vegetais (OLIVEIRA, 2006); a análise qualitativa e quantitativa de combustíveis, em veículos leves, álcool e gasolina (SACORAGUE, 2004) e a determinação dos componentes da gasolina (FERREIRA, 2005).

2.6 Regressão por mínimos quadrados parciais por intervalos (iPLS)

Este método é uma extensão da regressão por PLS, onde o espectro é dividido em intervalos equidistantes, e para cada intervalo é construído um modelo PLS. O foco deste método é possibilitar que sejam selecionados regiões do espectro global que possuam informações relevantes, e um novo modelo seja construído a partir destas variáveis selecionadas (NORGAARD *et. al*, 2000).

No método iPLS, o espectro é dividido em n intervalos, e são construídos um modelo para cada intervalo de região espectral, e mais um modelo com o espectro inteiro. Os modelos são comparados utilizando o parâmetro de validação RMSECV/RMSEP (*root mean square error of cross-validation/prediction*), (LEARDI, NORGAARD, 2004). Pode-se propor o seguinte pseudocódigo iPLS:

Dividir os espectros em n intervalos equidistantes

Enquanto o número de intervalos $\leq n$ faça

Constrói o modelo PLS para o n -ésimo intervalo

FimEnquanto

Selecionar o modelo cujo intervalo apresenta o menor RMSECV

Na área de pesquisa de óleos, empregando o método iPLS, menciona-se os trabalhos de quantificação de contaminantes em óleos lubrificantes (BORIN, POPPI, 2005) e, parâmetros de qualidade de lubrificantes e óleo de oliva extra virgem e puro (BORGES NETTO, 2005).

3 EXAME DE PARTÍCULAS

3.1 Introdução

A técnica de otimização por enxame de partículas (PSO – do inglês *Particle Swarm Optimization*), foi criada por Kennedy e Eberhart (1995), a partir da análise comportamental dos pássaros a procura de alimento ou de um local para construção do ninho, conforme ilustrado na Figura 3. Na procura pelo ninho ou alimento, todo o indivíduo (partícula) pode lucrar com as experiências dos membros do grupo (enxame). Propuseram um algoritmo de otimização robusto, simples e de fácil implementação computacional.

A PSO é uma técnica de computação evolutiva, motivada pela simulação do “comportamento social”, que trabalha em uma população potencial de soluções em um espaço de busca, onde, através da cooperação e da competição entre estas soluções é possível encontrar uma solução ótima, quando aplicadas em problemas de otimização (POMEROY, 2003).



Figura 3 - Ilustração do método de enxame de partículas.

Este algoritmo faz uso da teoria dos enxames, onde os pássaros fazem uso de suas experiências e da experiência do bando para localizar o ninho ou alimento. Desta forma a área sobrevoada corresponde ao espaço definido do projeto e encontrar a comida significa encontrar o ótimo ou próximo do ótimo. A possibilidade de encontrarem alimento, baseados na inteligência social, aumenta significativamente, porque um indivíduo aprende com o acerto do outro (PRADO e SARAMAGO, 2005).

3.2 O algoritmo básico de PSO

O algoritmo é baseado em um espaço de busca, onde é inicialmente gerada uma população aleatória de partículas e a estas associadas uma velocidade, ou seja, cada partícula possui uma posição (x_i) e uma velocidade (v_i). A cada interação, cada partícula atualiza sua velocidade, pela Equação 6, e uma nova posição é gerada pela soma de sua posição atual e a nova velocidade, conforme a Equação 5 (ENGELBRECHT, 2005).

$$x_i^{it+1} = x_i^{it} + v_i^{it+1} \quad (5)$$

$$v_i^{it+1} = v_i^{it} + (c_1^{it} * rand()^{it} * (pbest_i^{it} - x_i^{it})) + (c_2^{it} * rand()^{it} * Gbest^{it} - x_i^{it}) \quad (6)$$

A Equação 6, que calcula a velocidade da partícula, é composta também pela função $rand()^{it}$ que gera números aleatórios entre 0 e 1, pelos parâmetros de confiança c_1^{it} e c_2^{it} que indicam quanto a partícula confia em sua experiência (c_1^{it}) e na do enxame (c_2^{it}); e pelos termos de aceleração: $(pbest_i^{it} - x_i^{it})$ que representa a distância entre a melhor posição encontrada para a partícula i e a posição atual dessa mesma partícula, e $(Gbest^{it} - x_i^{it})$, que representa a distância entre a melhor posição de todas as partículas do enxame e a posição da partícula i .

As partículas enquanto percorrem seu espaço de busca geram informações que são armazenadas em vetores, onde:

$X_i = (x_1, x_2, x_3, \dots, x_n)$ corresponde as posições das partículas;

$V_i = (v_1, v_2, v_3, \dots, v_n)$ indica as velocidades das partículas;

$Pbest_i = (pbest_1, pbest_2, pbest_3, \dots, pbest_n)$ corresponde a melhor posição obtida de cada partícula. E n indica o tamanho do enxame.

Armazena-se também, durante o processo de exploração, a melhor posição (*fitness*) já encontrada entre todas as partículas do enxame, que é representado pelo $Gbest$.

3.3 Estrutura do enxame de partículas

O algoritmo de PSO apresenta um conceito simples, de fácil implementação e de eficiência computacional. A seguir um pseudocódigo do PSO (ENGELBRECHT, 2005):

Criar uma população aleatória de partículas, denominada enxame;

Inicializar esta população com posições e velocidades aleatórias;

Repita para cada partícula i

 Calcular a função objetivo (*fitness*);

 Comparar o novo valor resultante da *fitness* da partícula i . Se o valor for melhor que o valor do *pbest* da partícula i , atualize o *pbest* com o novo valor da *fitness*.

 Comparar o novo valor resultante da *fitness* da partícula i com o valor do *Gbest*. Se o valor da *fitness* da partícula i for melhor que o valor do *Gbest*, atualize o valor do *Gbest* com este novo valor;

 Alterar a posição e a velocidade da partícula i ;

Até que a condição de parada seja satisfeita.

3.4 Outros parâmetros do PSO

A versão básica do enxame de partículas foi aperfeiçoada desde sua proposição em 1995. Abaixo são detalhados alguns dos parâmetros importantes relatados na literatura e que contribuíram no aperfeiçoamento deste método.

3.4.1 Tamanho da população

O tamanho da população é determinado em função do problema em que PSO será aplicado. Variam entre 20 e 50 partículas. Soluções ótimas geralmente são obtidas com tamanhos menores de população (MENDES, 2004).

Parsopoulos e Vrahatis (2002) trabalharam com enxames não maiores de 40 partículas. Já Shi e Eberhart (1999) testaram enxames com tamanhos de 20, 40, 60 e 80 partículas.

3.4.2 Velocidade Máxima

Inicialmente, o algoritmo enxame de partículas, não controlava a velocidade das partículas. Para tal foi introduzido um importante parâmetro V_{\max} . Valor muito grande para este parâmetro, facilita a exploração global, enquanto que valor muito pequeno favorece a exploração local. Porém se o valor de V_{\max} for demasiado pequeno, o enxame pode não explorar suficientemente além das regiões localmente boas (ENGELBRECHT, 2005).

O V_{\max} é um parâmetro ajustado pelo usuário conforme as características do problema em questão, normalmente entre [-4, 4]. Assim, se $(v_i) < -4$ então $(v_i) = -4$ e se $(v_i) > 4$ então $(v_i) = 4$ (KENNEDY, EBERHART, SHI, 2001).

3.4.3 Fator de Inércia

Shi e Eberhart (1998a, 1998b) incluíram um novo parâmetro, chamado de fator de inércia ou peso de inércia (*inertia weight*), no algoritmo PSO, representado por (w) conforme visualizado na Equação 7. O papel fundamental do fator de inércia é balancear habilidade de exploração de busca local e global do algoritmo. Observaram que valores maiores para (w) facilita a busca no espaço global, enquanto que valores menores para (w) aumentam a busca no espaço local, fazendo com que as partículas converjam rapidamente, para o ponto ótimo ou não, sem que haja uma boa exploração do espaço de busca.

$$v_i^{it+1} = w * v_i^{it} + (c_1^{it} * \text{rand}()^{it} * (pbest_i^{it} - x_i^{it})) + (c_2^{it} * \text{rand}()^{it} * (Gbest^{it} - x_i^{it})), \quad (7)$$

Uma alternativa interessante para que ocorra um balanço entre a investigação e a exploração é iniciar a busca com um valor alto de (w) e diminuir este valor ao longo das

iterações, como proposto por Shi e Eberhart (1998a; 1998b, 1999). onde propuseram que o valor de (w) deveria diminuir linearmente entre 0,9 e 0,4, aproximadamente. Já Parsopoulos e Vrahatis (2002) utilizaram 1 para o valor inicial de (w), sofrendo uma diminuição gradual até 0.4.

Shi e Eberhart (2001) utilizaram um conjunto de regras Fuzzy para ajustar dinamicamente o valor de (w), mostrando-se ser um método promissor especialmente para problemas em ambientes dinâmicos.

O fator de inércia também pode sofrer uma diminuição linear a cada iteração através da Equação 8 (ENGELBRECHT, 2005).

$$w^{it} = (w_{ini} - w_{fin}) \frac{(R - it)}{R} + w_{fin} \quad (8)$$

onde w_{ini} e w_{fin} é o valor inicial e o valor final para o fator de inércia, respectivamente; R é o número total de iterações.

Como pode ser visto, muitos estudos empregaram w com variação decrescente no seu valor, em contrapartida Zheng. *et al.* (2003) propuseram que o valor de w fosse incrementado linearmente de 0.4 até 0.9 e obtiveram bons resultados.

Já para testar o método enxame de partículas Prado e Saramago (2005) utilizaram o valor da (w) em 0,729, sendo atualizado através da Equação 9, onde f_w é uma constante entre 0 e 1.

$$w_{new} = f_w * w_{old} \quad (9)$$

O fator de inércia também pode sofrer uma redução de maneira não linear. Peram, Veeramachaneni e Mohan (2003) aplicaram a Equação 10 para o cálculo do valor de (w):

$$w^{it+1} = \frac{(w^{it} - 0.4) * (R - it)}{R + 0.4} \quad (10)$$

Recentemente, Chatterjee e Siarry (2006) propuseram uma nova maneira de cálculo de (w) através da variação não-linear, conforme a Equação 11.

$$w^{it} = \left\{ \frac{(R-it)^{nl}}{(R)^{nl}} \right\} * (w_{ini} - w_{fin}) + w_{fin} \quad (11)$$

onde nl representa o índice não-linear.

3.4.4 Coeficiente de restrição

O coeficiente de restrição (K), calculado segundo Equação 12, e incorporado na equação de cálculo da velocidade, conforme Equação 13, visando limitar a velocidade das partículas, e garantir a convergência do PSO (CLERC, 1999; CLERC, KENNEDY, 2002).

$$v_i^{it+1} = K^{it} * [v_i^{it} + (c_1^{it} * rand()^{it} * (pbest_i^{it} - x_i^{it})) + (c_2^{it} * rand()^{it} * Gbest^{it} - x_i^{it})] \quad (12)$$

$$K^{it} = \frac{2}{2 - \varphi - \sqrt{\varphi^2 - 4\varphi}} \quad (13)$$

Sendo que $\varphi = c_1 + c_2$, $\varphi > 4$

3.5 Outro tipo de representação

Originalmente a técnica PSO foi desenvolvida para o universo de variáveis contínuas, mas muitos problemas são definidos num universo de variáveis discretas, onde o domínio é finito. O PSO foi facilmente adaptado para trabalhar com este tipo de variáveis. Este tipo de representação é chamado de PSO binário (ENGELBRECHT, 2005).

3.5.1 PSO binário

Esta representação tenta estender o algoritmo de enxame de partículas para trabalhar em espaço binários. Kennedy e Eberhart (1997) e Kennedy, Eberhart e Shi (2001)

descreveram uma alteração no algoritmo para operar sobre *bits* de caracteres, conforme Equação 14, onde a velocidade é usada como uma probabilidade limiar para determinar se x_i será zero ou um.

$$\text{sigmoid}(v_i^{it}) = \frac{1}{1 + \exp(-v_i^{it})} \quad (14)$$

Kennedy e Spears (1998) comparam o algoritmo binário com diversos tipos de algoritmos genéticos aleatórios. Para tal usaram um gerador aleatório. Os resultados mostram que o enxame de partículas binário foi o único a encontrar o ótimo global em cada uma das experiências, sem considerar as características do problema.

Agrafiotis e Cedeño (2002) usaram o PSO binário para seleção de características, em que (x_i) pode receber apenas 0 e 1 e, representa se a característica da partícula i será ou não selecionada. Neste caso as posições (x_i) são tratadas como uma probabilidade limiar para decidir entre o bit de valores 0 e 1, conforme Equação 15.

$$p_{ij}^{it} = \frac{x_{ij}^{\alpha it}}{\sum_{j=1}^{nx} x_{ij}^{\alpha it}} \quad (15)$$

Sendo que α determina a seleção, e $x_{ij} \in [0, 1]$, onde para valores de $\alpha > 1$ enfatizam indivíduos com alta adaptação, enquanto $\alpha < 1$ aumenta a chance de indivíduos com baixa adaptação serem selecionados.

Na literatura encontram-se pesquisas que relatam o uso do enxame de partículas binário como, por exemplo, na redução e colocação de banco de capacitores (KHALIL *et al.*, 2006).

3.6 Algumas aplicações da técnica enxame de partículas

Apesar de ser uma técnica bastante recente, na área de otimização, muitas pesquisas já foram feitas nas mais diversas áreas, como por exemplo: na análise do tremor humano

usando PSO e redes neurais (EBERHART e HU, 1999), otimização da recarga de combustível em um reator nuclear (MENESES, 2005); em classificações de imagens não supervisionadas (OMRAN, ENGELBRECHT, SALMAN, 2005), problemas complexos de engenharia nuclear (MEDEIROS, 2005) e na otimização de *layouts* (MÜLLER, 2007).

Schwaab (2005), testou algoritmos heurísticos de otimização na estimativa de parâmetros de modelo não lineares, e o método do enxame de partículas, em particular, apresentou resultados superiores aos demais algoritmos heurísticos utilizados neste trabalho, tanto na minimização de funções teste, como na estimativa de parâmetros de problemas reais.

A utilização de heurísticas e meta-heurística na otimização de processos químicos, tem recebido grande atenção por parte de pesquisadores. No entanto, o método de enxame de partículas, ainda é pouco explorado nesta área. Alguns exemplos são: PSO e rede neural na construção de modelos QSAR (WANG *et. al*, 2004) e, treinamentos de redes neurais para modelos QSAR de inibidores de plaquetas (QI SHEN *et al.*, 2006).

Por outro lado, o algoritmo genético (AG) que é uma ferramenta de otimização, apresenta um vasto número de trabalhos utilizando-o com uma forma de selecionar regiões espectrais para construção dos modelos, como por exemplo, na determinação simultânea de glicose, maltose e frutose (COSTA FILHO, POPPI, 2002) e, na otimização de modelos de regressão multivariados empregando métodos de seleção (SABIN, 2007). Já outros trabalhos relatam a utilização da heurística busca tabu na seleção do comprimento de onda (HAGEMAN *et. al*, 2003) e na otimização dos métodos de controle de qualidade de fármacos (KONZEN *et. al*, 2003).

4 METODOLOGIA

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

Para um melhor entendimento, a Figura 4 mostra um fluxograma das etapas que foram executadas na parte experimental, para a composição deste trabalho.

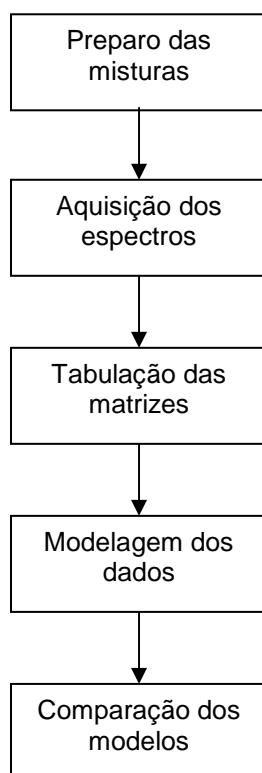


Figura 4 – Fluxograma das etapas executadas na parte experimental.
Fonte: Elaborado pela autora.

4.1 Preparo das misturas

Para a realização das misturas foram utilizados 2 planejamentos experimentais conforme tabelas 1 e 2. As 40 misturas binárias e ternárias conforme Tabela 1 foram realizadas a partir de uma amostra de óleo diesel, fornecida pelo Labcom da Universidade Federal do Rio Grande do Sul – UFRGS, de uma amostra de biodiesel de palma (doador pela Agropalma S/A), uma amostra de biodiesel de soja e uma amostra de óleo de soja (ambas fornecidas pela OleoPlan S.A). Já as 40 misturas binárias e ternárias conforme Tabela 2 foram realizadas a partir de uma segunda amostra de óleo diesel, fornecida pelo Labcom – UFRGS, da mesma amostra de biodiesel de soja e da mesma amostra de óleo de soja (ambas fornecidas pela OleoPlan S.A).

A união dos dois planejamentos resultou em um conjunto total de 80 amostras binárias e ternárias, podendo estas serem visualizadas no diagrama ternário da Figura 5. Cabe ressaltar que na apresentação do diagrama os dois diferentes óleos diesel, bem como os dois diferentes biodiesel utilizados, foram considerados como um único.

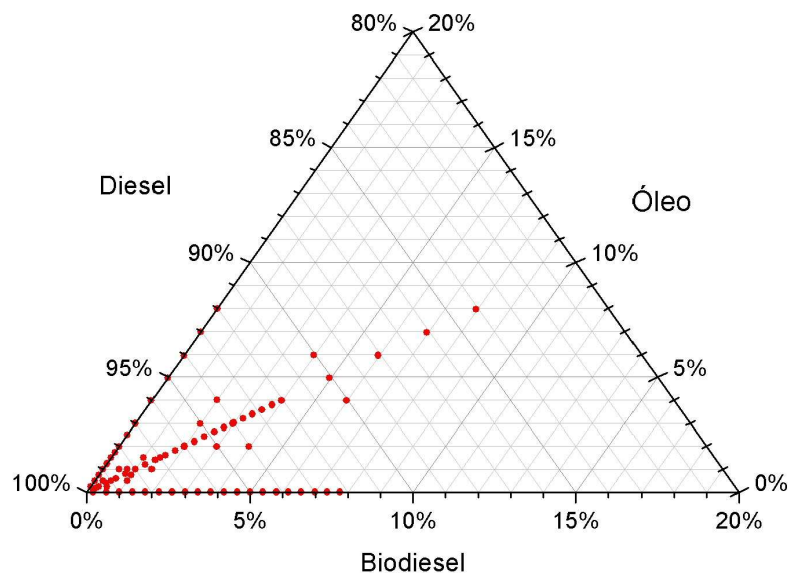


Figura 5 – Diagrama ternário para as misturas biodiesel/óleo vegetal/diesel.
Fonte: elaborado pela autora.

As amostras foram preparadas pela adição de alíquotas de biodiesel e/ou óleo vegetal conforme tabelas 1 e 2, em um balão volumétrico de 25 mL, seguida de adição do

óleo diesel até completar o volume. As soluções foram homogeneizadas e deixadas em repouso pelo tempo mínimo de 24 horas.

Tabela 1 - Planejamento experimental das misturas realizado na UNISC

Biodiesel (% v/v)	Óleo Vegetal (%v/v)	Diesel (%v/v)
0,2	0	99,8
0,2	0,2	99,6
0,6	0	99,4
0,4	0,4	99,2
1	0	99
0,6	0,6	98,8
1,4	0	98,6
0,8	0,8	98,4
1,8	0	98,2
1	1	98
2,2	0	97,8
1,2	1,2	97,6
2,6	0	97,4
1,4	1,4	97,2
3	0	97
1,6	1,6	96,8
3,4	0	96,6
1,8	1,8	96,4
3,8	0	96,2
2	2	96
4,2	0	95,8
2,2	2,2	95,6
4,6	0	95,4
2,4	2,4	95,2
5	0	95
2,6	2,6	94,8
5,4	0	94,6
2,8	2,8	94,4
5,8	0	94,2
3	3	94
6,2	0	93,8
3,2	3,2	93,6
6,6	0	93,4
3,4	3,4	93,2
7	0	93
3,6	3,6	92,8
7,4	0	92,6
3,8	3,8	92,4
7,8	0	92,2
4	4	92

Fonte: elaborado pela autora.

Tabela 2 – Planejamento experimental das misturas realizado na UFRGS

BioDiesel (%v/v)	Óleo Vegetal (%v/v)	Diesel (%v/v)
0,25	0,25	99,5
0,5	0,5	99
1	1	98
1,5	1,5	97
2	2	96
3	3	94
4	4	92
5	5	90
6	6	88
7	7	86
8	8	84
0,25	0,5	99,25
0,5	0,25	99,25
0,5	1	98,5
1	0,5	98,5
0,75	1	98,25
1	0,75	98,25
1	1,5	97,5
1,5	1	97,5
2	3	95
3	2	95
2	4	94
4	2	94
4	6	90
6	4	90
0	0,25	99,75
0	0,5	99,5
0	0,75	99,25
0	1	99
0	1,25	98,75
0	1,5	98,5
0	1,75	98,25
0	2	98
0	2,5	97,5
0	3	97
0	4	96
0	5	95
0	6	94
0	7	93
0	8	92

Fonte: elaborado pela autora.

4.2 Aquisição dos espectros

As amostras foram novamente homogeneizadas antes da aquisição dos espectros, sendo os mesmos obtidos em duplicata utilizando um acessório de reflexão total atenuada

(ATR) com cristal de ZnSe, acoplado a um espectrofotômetro Nicolet Magna 550, na faixa de 650 - 4000 cm^{-1} , conforme parâmetros descritos na Tabela 3. Antes de ser obtido o espectro de cada amostra foi realizado o espectro do branco, sem nenhum composto sobre o cristal.

Tabela 3 – Parâmetros instrumentais e regiões empregadas na aquisição dos espectros por FT-IR/ATR

Parâmetros	Condições
Região espectral	650 - 4000 cm^{-1}
Detector	DTGS
Divisor de feixes	KBr
Varreduras	32
Resolução	4 cm^{-1}
Acessório	Reflexão Total Atenuada - PIKE

Fonte: elaborado pela autora.

4.3 Tabulação das matrizes

Com os sinais espectrais obtidos, foi construída uma matriz de dados com a média das duplicatas dos espectros normalizados de cada amostra. Esta matriz foi utilizada para a construção de modelos de regressão multivariados para a quantificação do biodiesel, óleo vegetal e diesel. Foram implementados testes utilizando o espectro inteiro das misturas, as regiões espectrais propostas pela norma ABNT submetida à consulta popular em fevereiro de 2008 e finalmente com as regiões espectrais selecionadas pelas implementações propostas com o exame de partículas.

4.4 Modelagem dos dados

A seleção das amostras para a constituição dos conjuntos de calibração e previsão teve dois objetivos principais: construir conjuntos que contenham a informação da matriz de dados utilizados neste estudo e que contenha a maior variabilidade possível em relação à faixa de concentração. Os 80 espectros médios das amostras foram divididos em dois conjuntos, sendo 48 empregados na construção do conjunto de calibração e 32 no conjunto

de previsão. A distribuição das concentrações destas amostras para cada uma das substâncias participantes das misturas é apresentada nas Figuras 6, 7 e 8, onde se tomou o cuidado de garantir que as amostras com valores extremos de concentração pertençam ao conjunto de calibração.

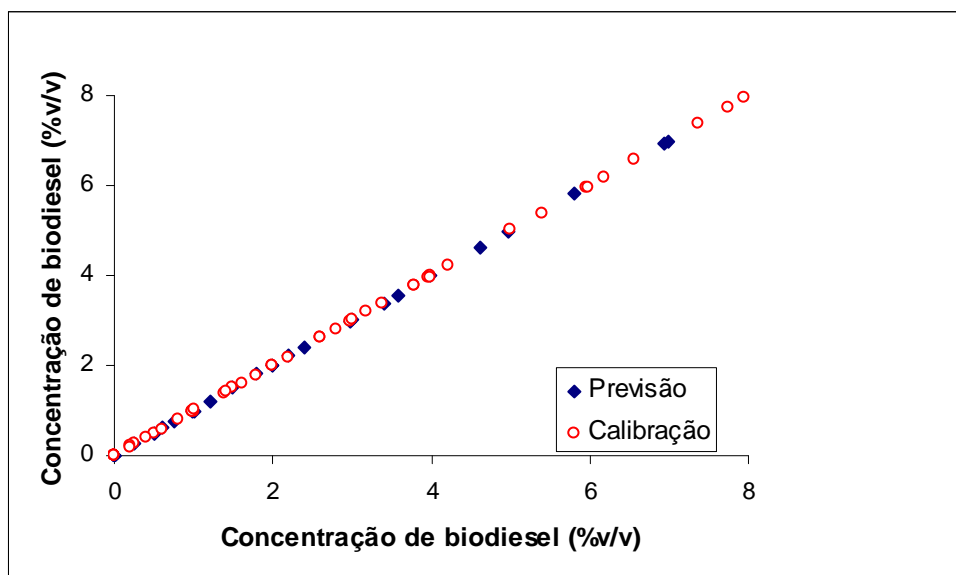


Figura 6 – Distribuição das amostras de calibração e previsão para o biodiesel.
Fonte: elaborado pela autora.

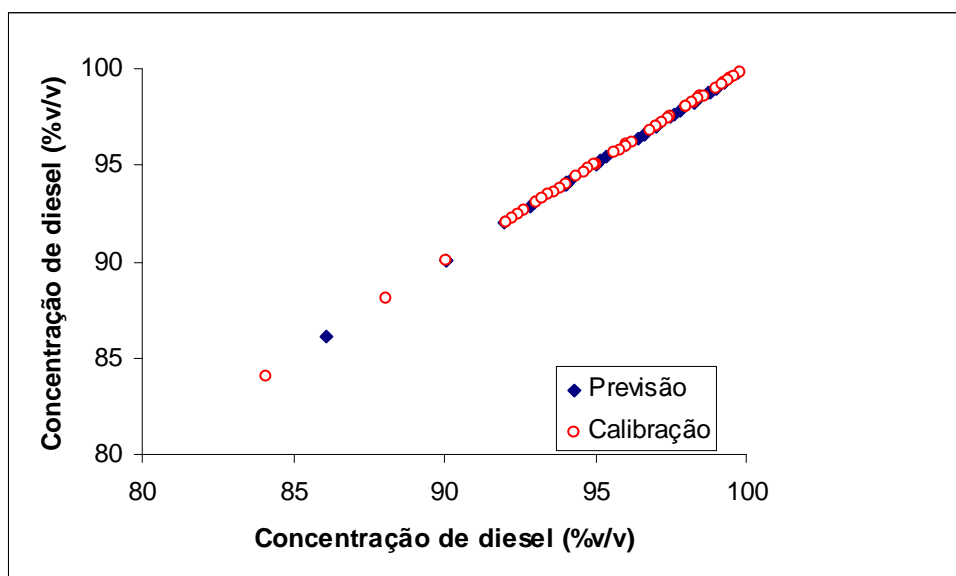


Figura 7 – Distribuição das amostras de calibração e previsão para o óleo diesel.
Fonte: elaborado pela autora.

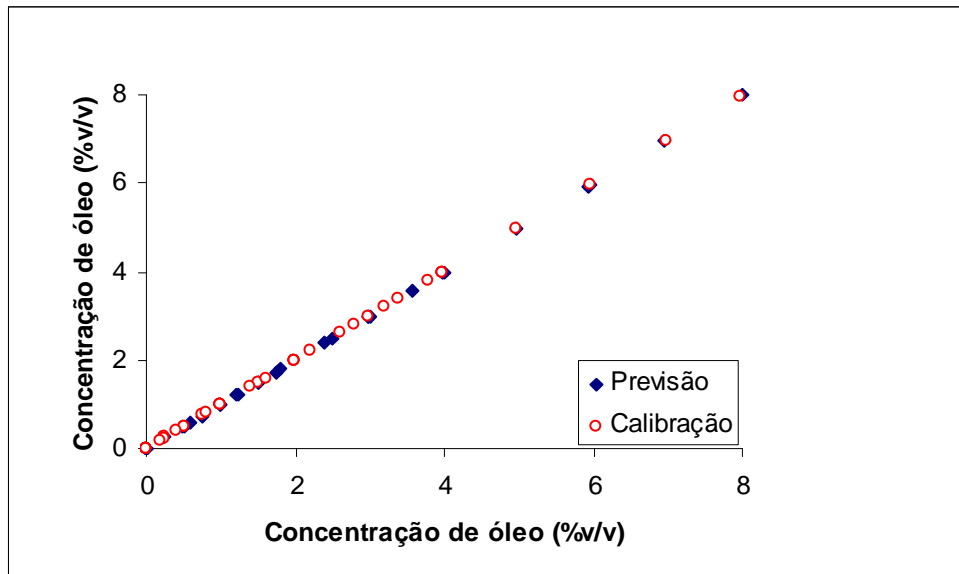


Figura 8 – Distribuição das amostras de calibração e previsão para o óleo vegetal.
Fonte: elaborado pela autora

Todos os modelos construídos neste trabalho, utilizaram o MATLAB[®] versão 6.5 (*The Math Works*, <http://www.mathworks.com>, EUA), juntamente com o pacote *iToolbox* versão 2 desenvolvido por Lars Norgaard (*iToolbox* for MATLAB, <http://www.models.kvl.dk>, Dinamarca).

As descrições das implementações com PSO e os resultados de todas as propostas mencionadas acima, são descritas no decorrer deste trabalho.

4.4.1 Considerações gerais sobre as implementações da meta-heurística enxame de partículas para selecionar variáveis de espectros no infravermelho

O princípio básico do algoritmo enxame de partícula é que as partículas se movimentam no espaço de busca, objetivando encontrar uma solução, orientadas pela sua experiência e pela experiência do enxame.

Como o propósito desta pesquisa é otimizar modelos de regressão multivariados, abaixo são explicados alguns conceitos do enxame de partículas neste contexto e que se mantêm para todas as implementações propostas.

Partícula: é uma possível solução para o problema. A posição da partícula é representada pelo conjunto de comprimentos de onda selecionados, ou seja, a configuração da mesma. A Figura 9 apresenta exemplo de duas partículas.

1	0	1	0	1	1	1	0	0	1
0	0	1	1	1	0	0	0	0	1

Figura 9 – Representação gráfica de duas partículas.
Fonte: elaborado pela autora.

Enxame: um conjunto de soluções (partículas).

Unidade de experiência: corresponde a um intervalo do espectro, podendo este ser ou não utilizado na construção do modelo de regressão.

Fitness: é a qualidade da posição (configuração) da partícula, obtido pelo erro de *cross-validação* (RMSECV) gerado no modelo de regressão.

Tamanho do enxame: corresponde ao número total de partículas utilizadas.

Velocidade: é a magnitude em que a solução corrente é alterada nas “direções” do $Gbest$ e do $pbest_i$.

Fator de Inércia: a intensidade em que a solução corrente é alterada (o percentual transferido do $Gbest$ e do $pbest_i$).

4.4.2 Estrutura básica do enxame de partícula implementado para a otimização dos modelos de regressão

O enxame inicia sua execução criando as partículas, onde cada partícula é composta de um número pré-definido de intervalos, codificadas com 0 e 1 aleatoriamente, e assim, cada partícula possui uma posição (solução inicial) e uma velocidade inicial. Para este problema, cada um dos intervalos da partícula corresponde a uma região do espectro, e os modelos serão construídos somente com os intervalos codificados com 1, conforme Figura 10.

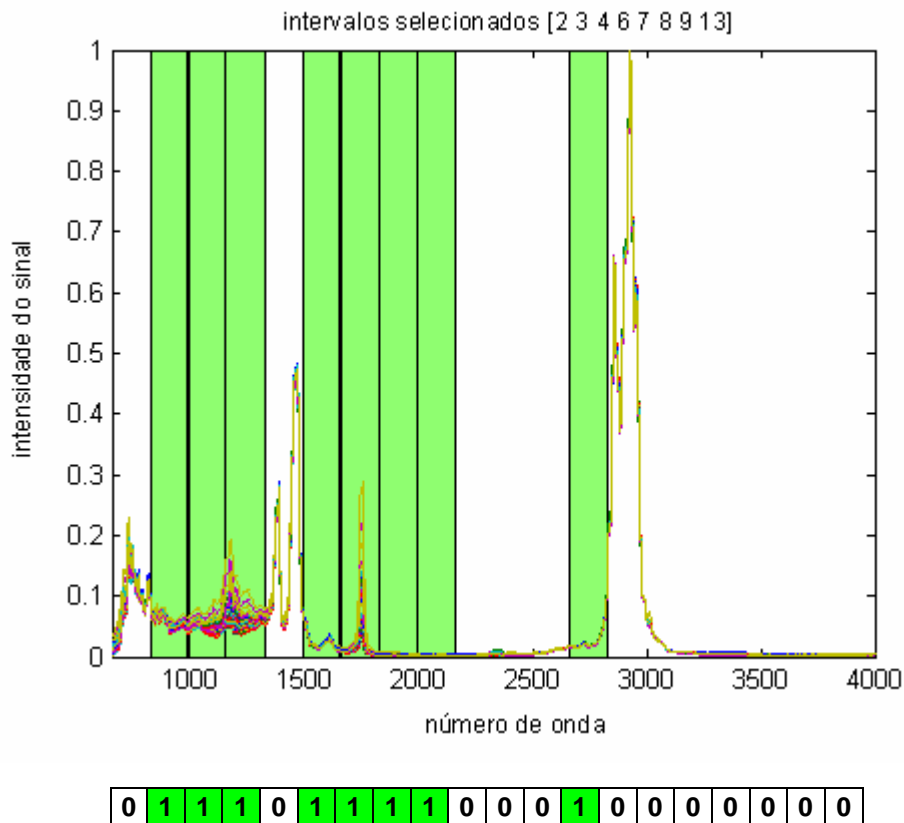


Figura 10 – Exemplo de uma partícula e a representação dos seus intervalos seleccionados. Fonte: elaborado pela autora com base no espectro do infravermelho de amostras de misturas biodiesel/óleo vegetal/diesel.

Quando o enxame é criado a posição inicial de cada partícula i é considerada como a sua melhor posição encontrada, ou seja o $pbest_i$, e a melhor posição entre todas as partículas, passa a ser considerada o $Gbest$ do enxame.

A cada iteração, cada partícula i , se desloca para uma outra posição, com base na soma da sua posição atual e nova velocidade gerada. Nestas implementações a velocidade será calculada através da Equação 7.

Assim, com base no deslocamento da partícula i é gerado um vetor somente com os intervalos seleccionados, e este é passado como parâmetro para a função $plsmodel$ do pacote $iToolBox$, juntamente com um modelo $iPLS$, com o mesmo número de intervalos da partícula i que está sendo avaliada, que gera um modelo de regressão empregando apenas estes intervalos. Este modelo gera um vetor de erros (RMSECV), onde cada valor deste vetor corresponde a utilização de um número diferente de variáveis latentes. O resultado do deslocamento da partícula ($fitness$) será o menor RMSECV, juntamente com o respectivo número de variáveis latentes.

Esta nova posição (valor resultante da *fitness*) da partícula *i* é comparada com o melhor posição que esta partícula *i* obteve até o presente momento ($pbest_i$). Da mesma maneira compara-se a *fitness* da partícula *i* com o valor do *Gbest* do enxame. Neste trabalho em que se pretende minimizar o erro RMSECV, ambos, $pbest_i$ e *Gbest* somente serão atualizados se o valor obtido pela *fitness* da partícula *i* tiver sido menor que a última melhor solução armazenada por esta partícula e pelo enxame, respectivamente.

As partículas terão a oportunidade de evoluir no espaço de busca durante 250 iterações, que é condição de parada.

Através de testes preliminares, foram definidos alguns parâmetros para serem adotados como padrão para todas as implementações, sendo eles:

- a intensidade (fator de inércia) em que a partícula *i* é alterada em direção ao $pbest_i$ (parte cognitiva) será maior do que a intensidade em direção ao *Gbest* (parte social) e que as respectivas intensidades, para ambas direções, serão compostas por valores (percentuais) máximos e mínimos permitidos;
- serão utilizados enxames com tamanho de 25, 50 e 100 partículas;
- a quantidade de intervalos que cada partícula poderá ser constituída serão 20, 40 e 60 intervalos;
- para cada combinação proposta, todas as partículas apresentarão a mesma constituição (20, 40 e 60 intervalos) e, um tamanho fixo do enxame (25, 50 e 100 partículas). Além de que todas as combinações serão executadas em triplicata.

O algoritmo de enxame de partículas implementado busca combinar várias unidades de informação (regiões do espectro), para construção dos modelos de regressão. Nesta dissertação foram implementadas cinco versões do enxame de partículas, onde os resultados obtidos destas soluções, foram comparados com os modelos obtidos utilizando o espectro inteiro das misturas e as sub-regiões indicadas pela proposta da norma ABNT.

Foi criada uma nomenclatura para identificar as versões do enxame de partículas implementado, conforme Tabela 4.

Tabela 4 – Nomenclatura usada para identificar as implementações

Sigla	Definição
iPLS	Método usado para construção do modelo.
PSO	Método de otimização utilizado.
C* ou D**	Sentido do deslocamento da intensidade da partícula i em relação ao G_{best} .
C* ou D**	Sentido do deslocamento da intensidade da partícula i em relação ao $pbest_i$.

* - crescente ** - decrescente

Fonte: elaborado pela autora.

Abaixo são citadas as implementações efetuadas utilizando a PSO.

a) Implementação do Enxame iPLS-PSO

Nesta implementação, manteve-se a estrutura básica do enxame de partículas implementado, mas a intensidade em que a partícula i é alterada (o percentual transferido do G_{best} e do $pbest_i$) foi definida e mantida constantemente durante todas as iterações. Os percentuais utilizados aqui foram 10% do G_{best} e 15 % $pbest_i$.

b) Implementação do enxame iPLS-PSO (CC)

Esta abordagem manteve-se novamente a estrutura básica do enxame de partículas implementada inicialmente, mas a intensidade em que a partícula i é alterada (os percentuais transferidos do G_{best} e do $pbest_i$) foi iniciada com valores mínimos (5%), e no decorrer das iterações sofreram um incremento linear constante até os valores máximos permitidos (15% e 30%, respectivamente).

c) Implementação do Enxame iPLS-PSO (DD)

Com base na estrutura básica do enxame de partículas descrita inicialmente, foi proposta uma nova abordagem em que a intensidade em que a partícula i é alterada (o

percentual transferido do G_{best} e do p_{best_i}) é iniciada com valores máximos (15% e 30%, respectivamente), e no decorrer das iterações sofreram um decremento linear constante até os valores mínimos permitidos de 5%.

c) Implementação do Enxame iPLS-PSO (CD)

Partindo da implementação básica do enxame, nesta quarta proposta, a intensidade em que a partícula i é alterada em relação a parte cognitiva da fórmula da velocidade é definida com um valor mínimo permitido (5%) e a medida que a partícula i se movimenta no espaço de busca este valor sofre um incremento linear constante até o máximo de 15%. Em relação a parte social da fórmula da velocidade, a intensidade em que a partícula i é alterada (percentual transferido do p_{best_i}) é inicialmente definida com um valor máximo permitido (30%), e no decorrer das iterações sofre um decremento linear constante, até atingir o mínimo permitido (5%).

d) Implementação do Enxame iPLS-PSO (DC)

Em contrapartida, esta implementação partiu da proposta básica do enxame, mas a intensidade em que a partícula i é alterada em relação a parte cognitiva da fórmula da velocidade é definida com um valor máximo permitido (15%) e, a medida que a partícula i se movimenta no espaço de busca este valor sofre um decremento linear constante até atingir o mínimo permitido (5%). Em relação a parte social da fórmula da velocidade, a intensidade em que a partícula i é alterada, foi iniciada com um valor mínimo (5%) permitido, e no decorrer das iterações sofre um incremento linear constante, até atingir o máximo permitido (30%).

5 RESULTADOS

5.1 Resultados para a quantificação do biodiesel

Visando quantificar o biodiesel nas misturas, foram desenvolvidos diversos modelos de regressão, e os resultados obtidos estão descritos a seguir. Primeiramente são apresentados os modelos usando o PLS e em seguida, são descritos os modelos obtidos com a otimização através da meta-heurística enxame de partículas combinada ao iPLS.

5.1.1 Modelos para a determinação do biodiesel com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT

Com base no conjunto de dados foram desenvolvidos dois modelos de regressão, ambos utilizando para construção o método PLS. O primeiro modelo, para a quantificação do biodiesel, foi construído utilizando toda região espectral (650 cm^{-1} a 4000 cm^{-1}) conforme Figura 11, e o desempenho do modelo empregando 6 variáveis latentes, é apresentado na Figura 12.

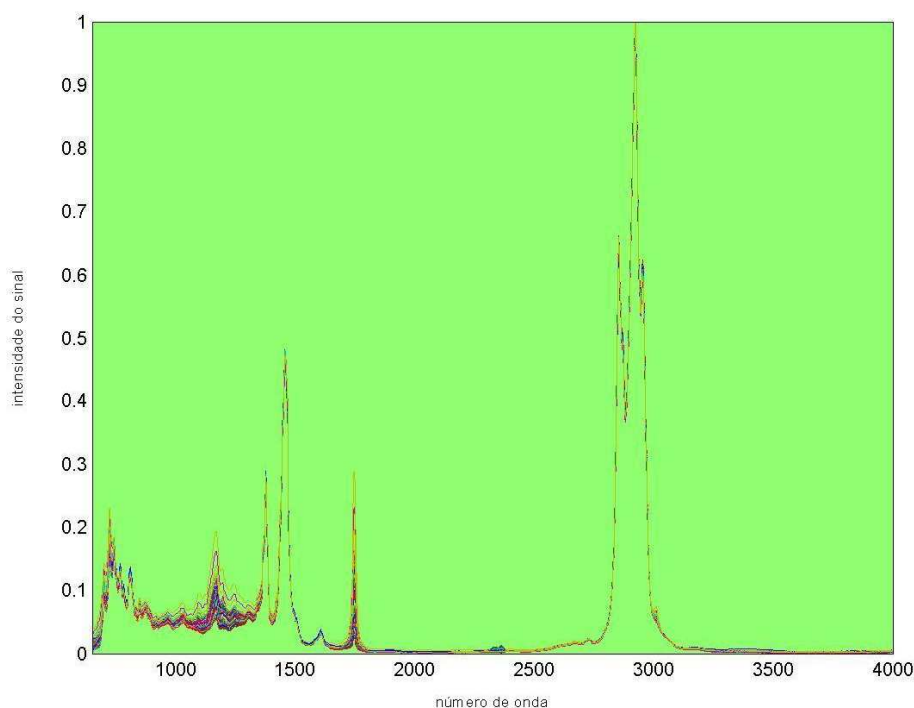


Figura 11 – Região espectral utilizada na modelagem para a quantificação do biodiesel. Fonte: elaborado pela autora.

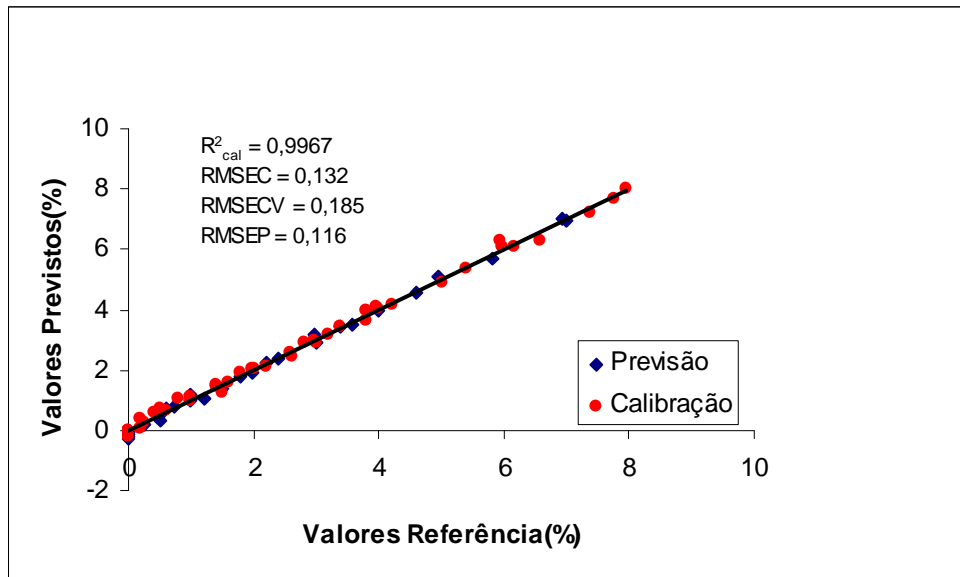


Figura 12 – Valores de referência *versus* previstos para o modelo de previsão do biodiesel utilizando toda a região espectral das misturas.
 Fonte: elaborado pela autora.

O segundo modelo construído para a quantificação do biodiesel, empregou somente as duas regiões espectrais (de 1100 cm^{-1} a 1200 cm^{-1} e de 1735 cm^{-1} a 1750 cm^{-1}) sugeridas pela proposta de norma ABNT. As regiões espectrais, e o desempenho do modelo empregando 5 variáveis latentes são visualizados nas Figuras 13 e 14, respectivamente.

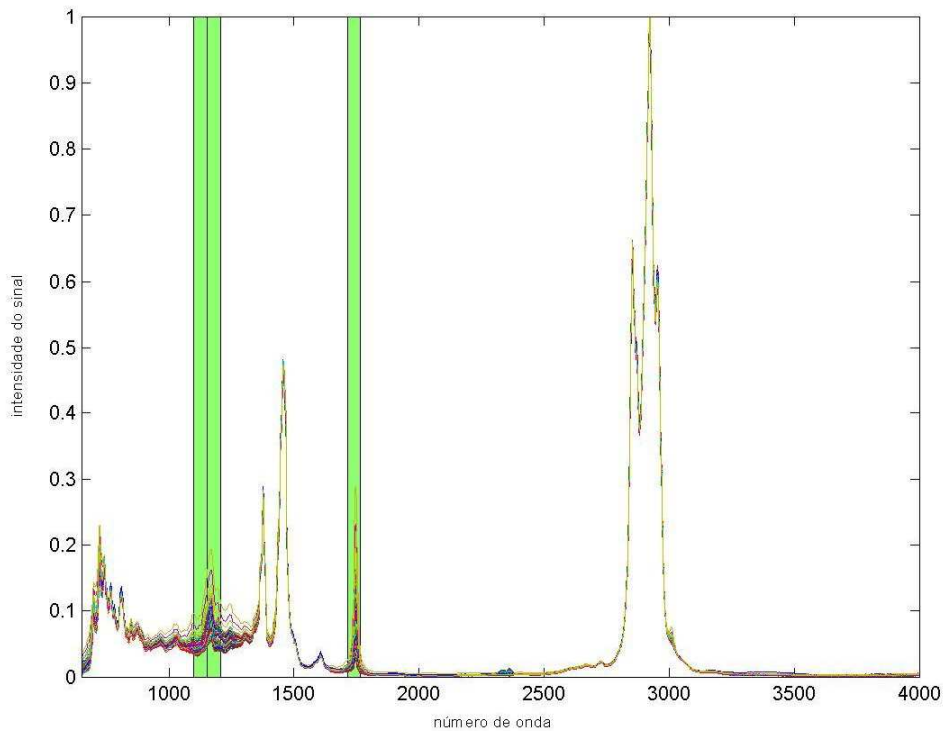


Figura 13 – Regiões espectrais sugeridas pela proposta de norma ABNT.
 Fonte: elaborado pela autora.

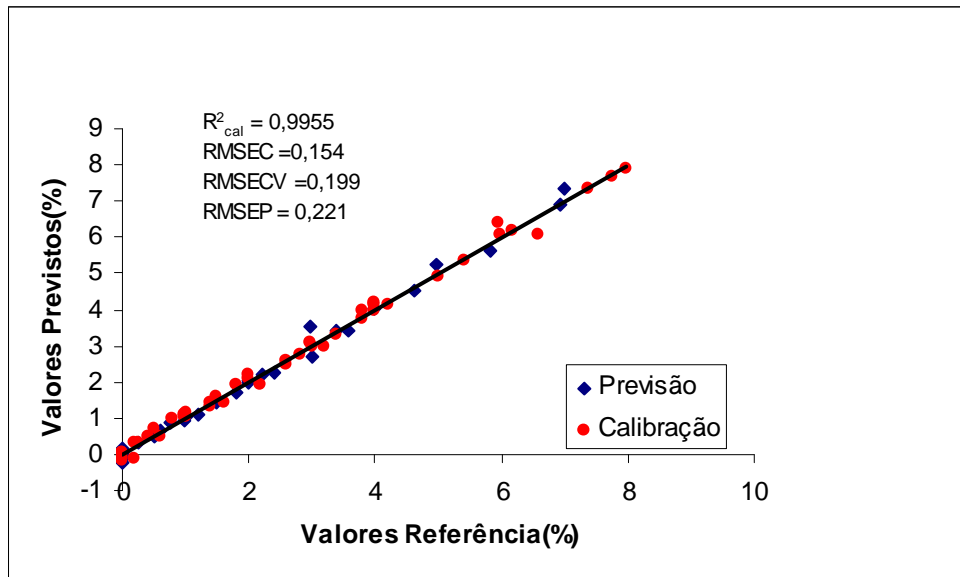


Figura 14– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT.

Fonte: elaborado pela autora.

Podemos observar que o modelo gerado utilizando o espectro inteiro, apresentou uma melhor capacidade preditiva, em relação ao modelo que levou em consideração somente às regiões sugeridas pela proposta de norma ABNT. Um aspecto que justifica este resultado pode ser a presença dos sinais interferentes do óleo nas mesmas regiões espectrais que o biodiesel, além de que outras regiões do espectro podem conter informações importantes para construção de um modelo mais robusto.

Tentando gerar modelos com maior capacidade preditiva para o analito de interesse, sugerimos a implementação de um algoritmo híbrido para a seleção das variáveis (número de onda) para construção dos modelos de regressão multivariados de espectros no infravermelho, que consiste na utilização do método iPLS, otimizado pela meta-heurística chamada enxame de partículas.

5.1.2 Resultados gerais e evoluções para a determinação do biodiesel

Com o propósito descrito anteriormente, foram implementadas cinco versões empregando a otimização por enxame de partículas: iPLS-PSO, iPLS-PSO (CC), iPLS-PSO (DD); iPLS-PSO (CD); iPLS-PSO (DC). Os principais resultados dos modelos, incluindo as implementações com a PSO, para quantificação do biodiesel são apresentados na Tabela 5.

Os resultados completos das implementações do enxame de partículas para quantificação do biodiesel encontram-se no ANEXO A.

Tabela 5– Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do biodiesel.

Modelos	Regiões empregadas	R² cal	RMSEC	RMSECV	RMSEP
espectro inteiro	650 - 4000 cm ⁻¹	0,9967	0,132	0,185	0,116
regiões sugeridas pela proposta de norma ABNT	1101 - 1198 cm ⁻¹ 1736 - 1749 cm ⁻¹	0,9955	0,154	0,199	0,221
iPLS-PSO	874 - 1319 cm ⁻¹ 2240 - 2717 cm ⁻¹ 3055 - 3109 cm ⁻¹ 3446 - 3556 cm ⁻¹	0,9997	0,038	0,095	0,088
iPLS-PSO (CC)	904 - 1327 cm ⁻¹ 2675 - 2756 cm ⁻¹ 3421 - 3502 cm ⁻¹	0,9995	0,052	0,105	0,085
iPLS-PSO (DD)	904 - 1327 cm ⁻¹ 2675 - 2756 cm ⁻¹ 3089 - 3170 cm ⁻¹ 3421 - 3502 cm ⁻¹	0,9996	0,048	0,103	0,086
iPLS-PSO (CD)	904 - 1327 cm ⁻¹ 2675 - 2756 cm ⁻¹ 3421 - 3502 cm ⁻¹	0,9995	0,052	0,105	0,085
iPLS-PSO (DC)	904 - 1327 cm ⁻¹ 2426 - 2756 cm ⁻¹ 3504 - 3585 cm ⁻¹	0,9991	0,070	0,103	0,093

Fonte: elaborado pela autora.

Na Figura 15 é apresentado um exemplo da evolução das partículas, mostrando as mesmas nas suas posições iniciais, bem como após 10 iterações, as posições das mesmas na 100^o iteração e finalmente quando atingem o número máximo de iterações definido para o problema. Pode-se observar que com o passar das iterações as partículas convergiram para uma solução.

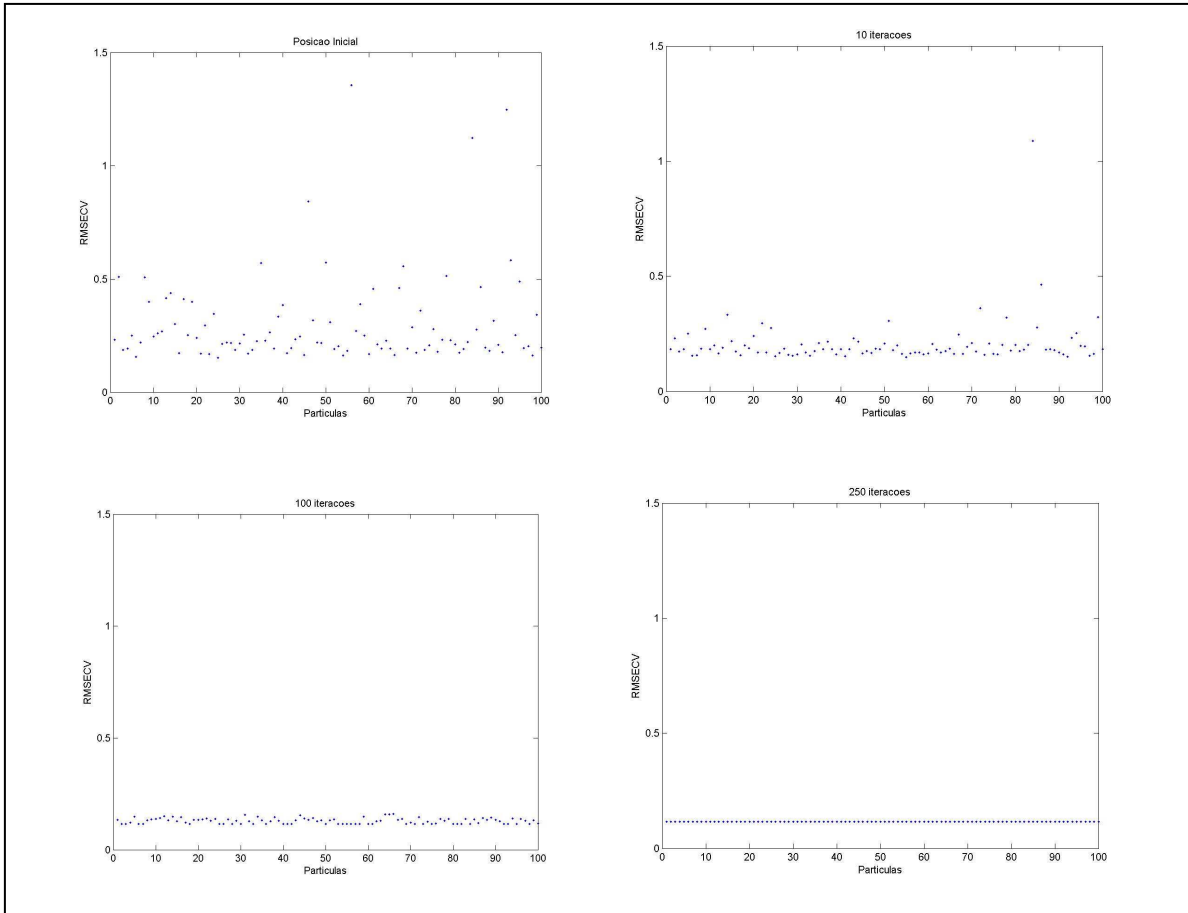


Figura 15 – Exemplo da evolução das partículas para a quantificação do biodiesel, usando o método iPLS-PSO (CD) com 40 intervalos e 100 partículas.

Fonte: elaborado pela autora.

A Figura 16 mostra um comparativo das evoluções para as 3 repetições para a combinação do iPLS-PSO com 20 intervalos e 50 partículas. Observa-se nesta proposta que existe uma tendência de convergência de todos os modelos, sendo necessário em média 44 iterações; mostrando juntamente que o número de iteração definido é suficiente para explorar as melhores combinações entre os 20 intervalos, atingindo um mínimo global.

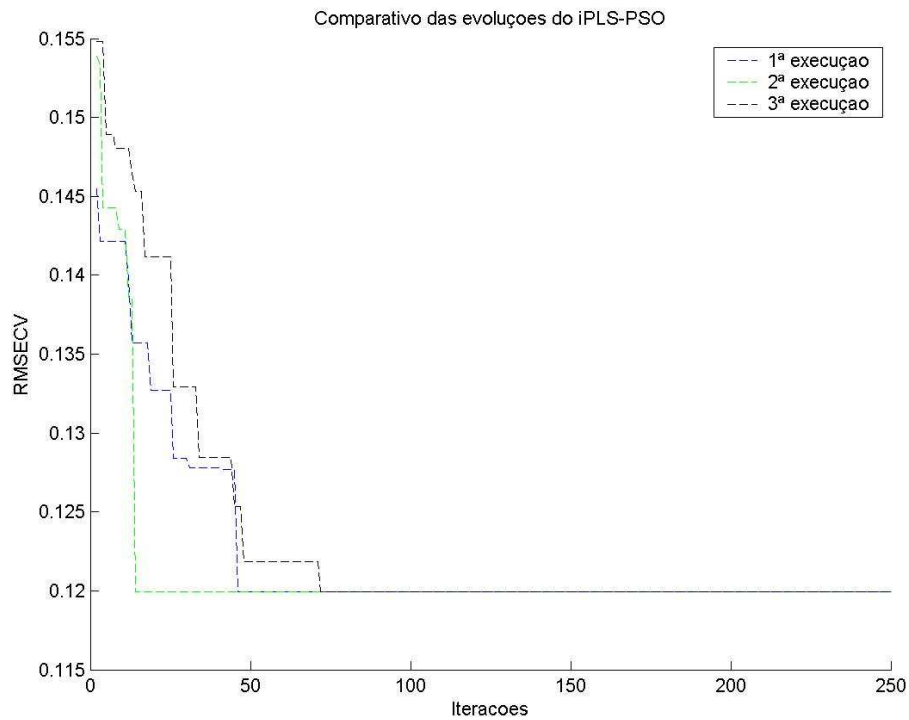


Figura 16 – Evolução das três execuções do iPLS-PSO com 20 intervalos e 50 partículas, para quantificação do biodiesel.
 Fonte: elaborado pela autora.

As Figuras 17 e 18 apresentam um comparativo de 3 evoluções do iPLS-PSO com 40 e 60 intervalos, respectivamente, ambos com 50 partículas. Observa-se um comportamento diferente para ambos, pois nem todos os modelos atingem o mínimo com poucas iterações. Em média para iPLS-PSO com 40 intervalos são necessárias 146 iterações, já para 60 intervalos em média 188. Em virtude disso, todos os resultados apresentados no ANEXO A foram obtidos com 250 iterações, pois um aumento no número de iterações refletiria num aumento considerável de custo computacional (tempo).

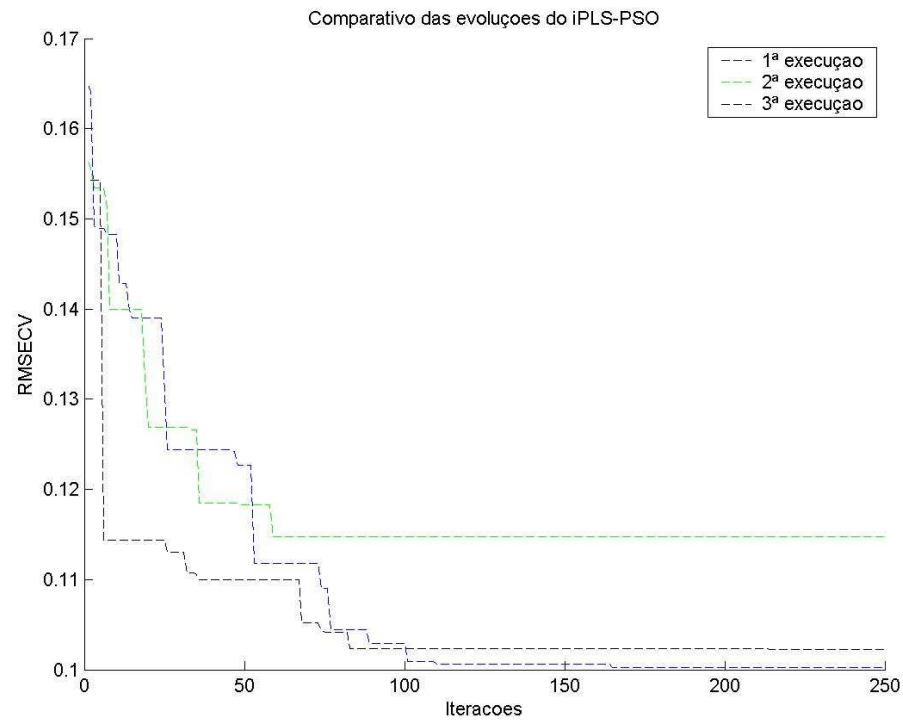


Figura 17 – Evolução das três execuções do iPLS-PSO com 40 intervalos e 50 partículas, para quantificação do biodiesel.
Fonte: elaborado pela autora.

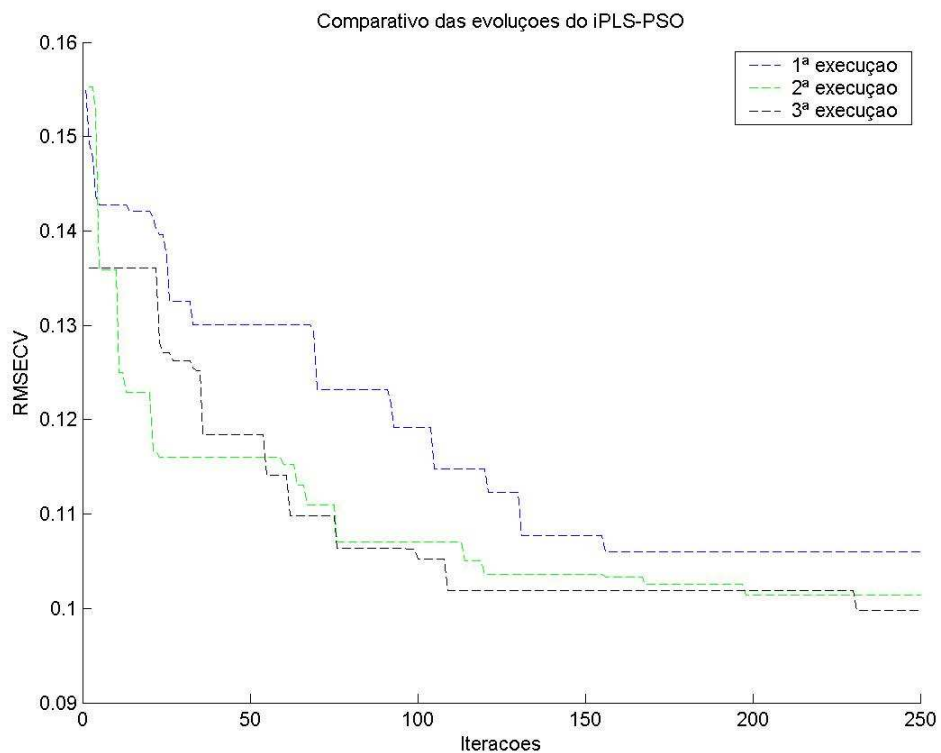


Figura 18 – Evolução das três execuções do iPLS-PSO com 60 intervalos e 50 partículas, para quantificação do biodiesel.
Fonte: elaborado pela autora.

5.1.3 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO

A aplicação da primeira implementação denominada iPLS-PSO no conjunto de dados propostos para a quantificação do biodiesel, selecionou regiões espectrais para construção do modelo que compreendem de 874 a 1319 cm^{-1} com lacunas, de 2240 a 2717 cm^{-1} , de 3055 a 3109 cm^{-1} e de 3446 a 3556 cm^{-1} como podem ser visualizadas na Figura 19. Dentre todos os modelos construídos para este método proposto, o selecionado foi o obtido com 60 intervalos e 100 partículas e o desempenho do modelo empregando 10 variáveis latentes para quantificação do biodiesel, é apresentada na Figura 20.

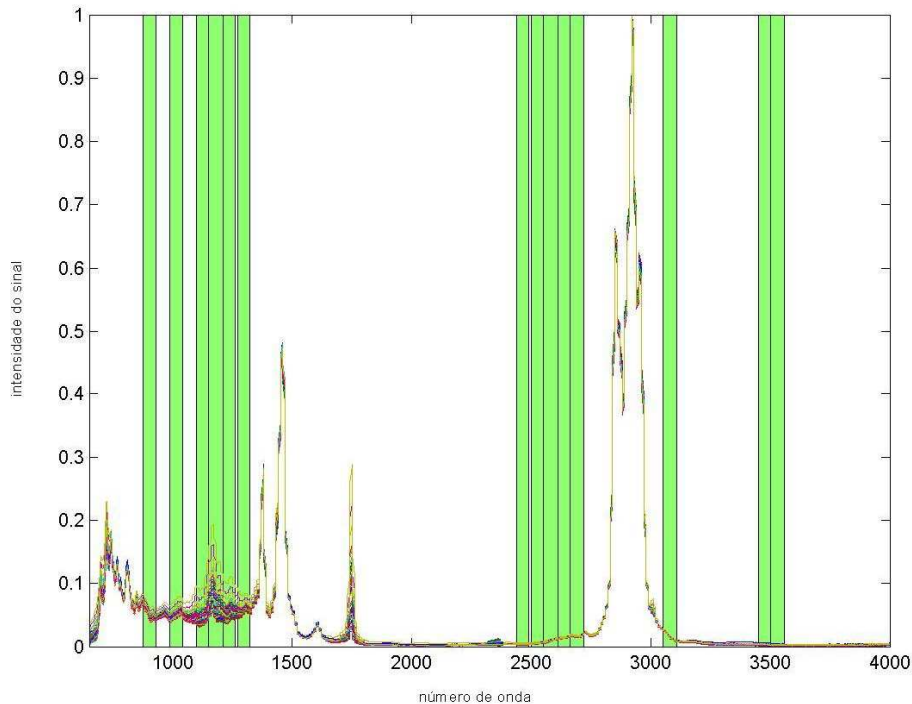


Figura 19 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel usando o método iPLS-PSO com 60 intervalos e 100 partículas.
Fonte: elaborado pela autora.

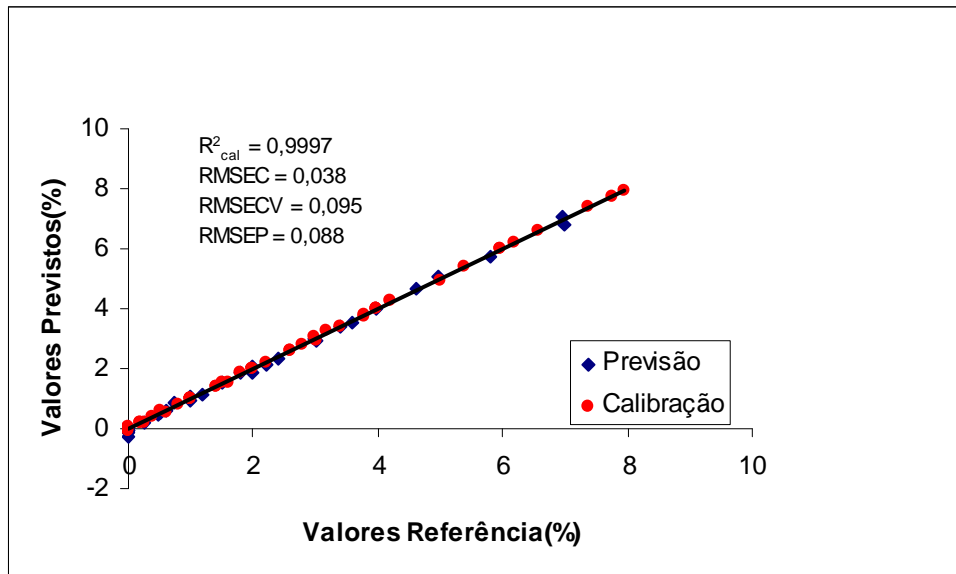


Figura 20 – Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 60 intervalos e 100 partículas.
 Fonte: elaborado pela autora.

5.1.4 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CC)

Já o método iPLS-PSO (CC) selecionou regiões espectrais que compreendem as faixas de 904 a 1327 cm^{-1} , de 2675 a 2756 cm^{-1} e de 3421 a 3502 cm^{-1} , conforme Figura 21. Nesta implementação o modelo selecionado empregou 40 intervalos e 25 partículas e, a sua respectiva curva de calibração utilizando 10 variáveis latentes para a concentração de biodiesel, é apresentada na Figura 22.

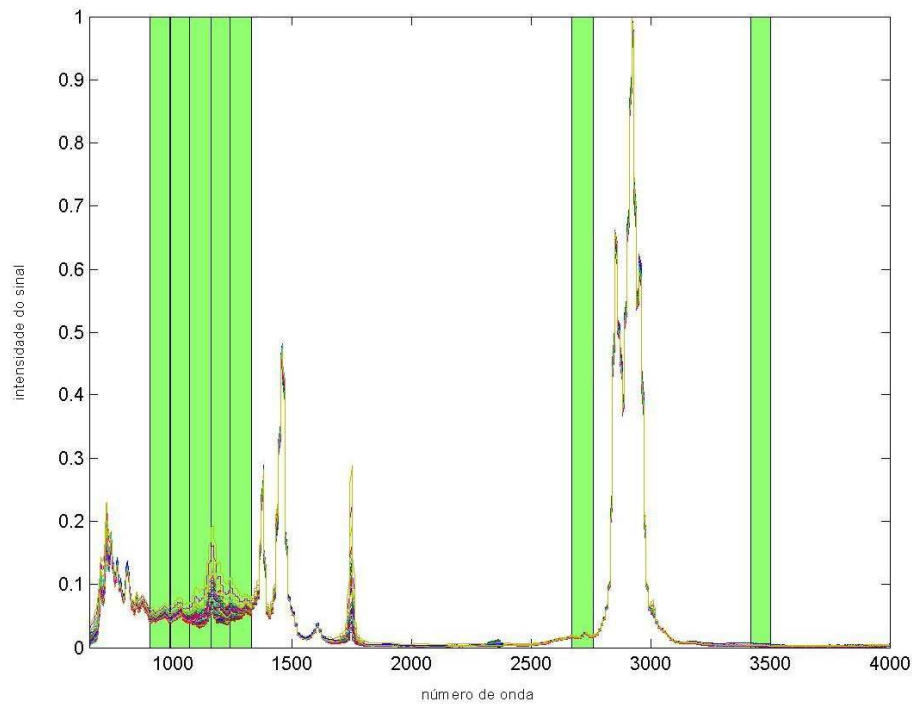


Figura 21 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

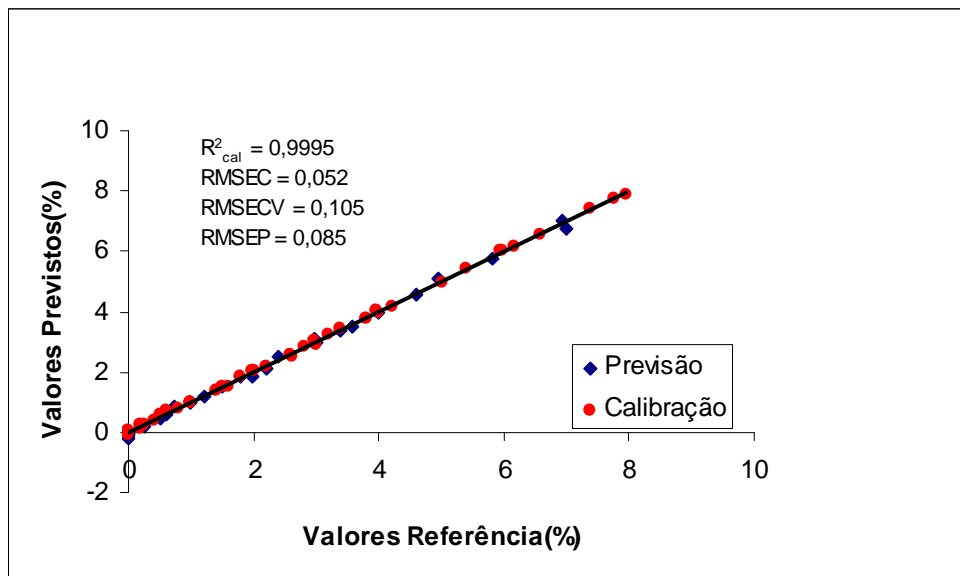


Figura 22 – Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.1.5 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (DD)

No caso da implementação iPLS-PSO (DD) as regiões espectrais selecionadas compreendem as faixas de 904 a 1327 cm^{-1} , de 2675 a 2756 cm^{-1} , de 3089 a 3170 cm^{-1} e de 3421 a 3502 cm^{-1} , conforme Figura 23. Neste caso, o modelo selecionado foi aquele construído com 40 intervalos e 100 partículas, e a curva de calibração, utilizando 10 variáveis latentes, para a concentração de biodiesel é apresentada na Figura 24.

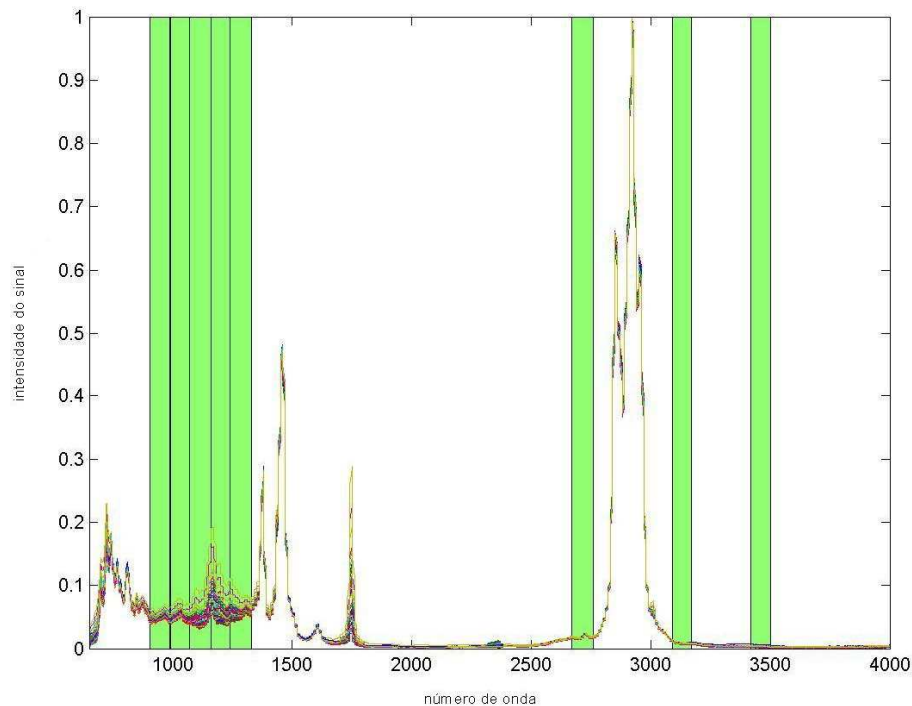


Figura 23– Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

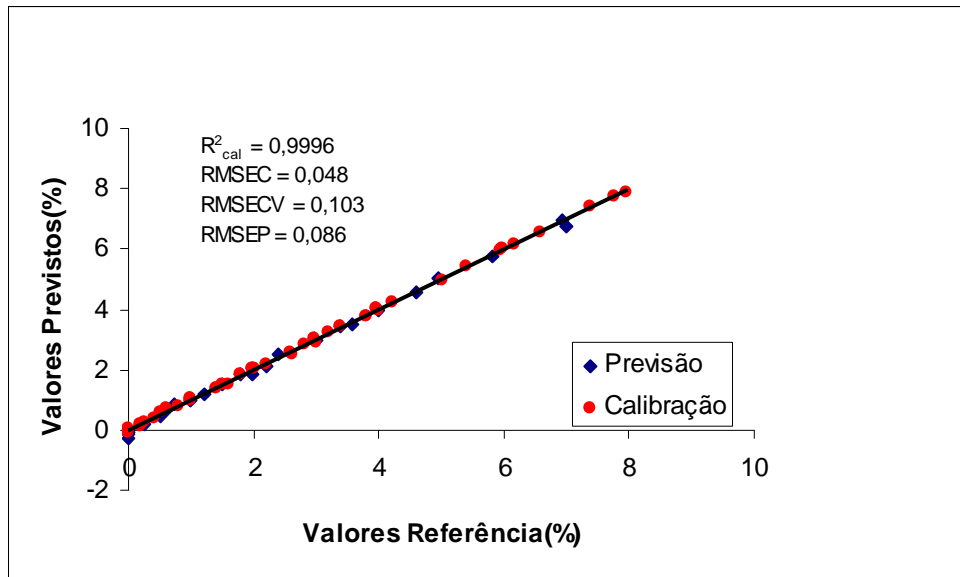


Figura 24– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
 Fonte: elaborado pela autora.

Um aspecto interessante nesta implementação é que conforme ANEXO A, na construção dos modelos com 20 intervalos, as partículas convergiram muito rápido para uma solução, não precisando ser explorado todo o espaço de busca, convergindo em média após 33 iterações.

5.1.6 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CD)

Conforme a Tabela 5, o modelo selecionado pelo método iPLS-PSO (CD) corresponde as mesmas regiões espectrais que a implementação do método iPLS-PSO (CC). As regiões espectrais bem como a curva de calibração pode ser vistas nas Figuras 21 e 22, anteriormente apresentadas.

5.1.7 Resultados obtidos para a determinação do biodiesel aplicando o iPLS-PSO (DC)

Neste método iPLS-PSO (DC) as regiões espectrais selecionadas para construção do modelo compreendem as faixas de 904 a 1327 cm^{-1} com lacunas, de 2426 a 2756 cm^{-1} e,

de 3504 a 3585 cm^{-1} conforme Figura 25. Dentre todos os modelos construídos para este método o modelo selecionado foi o construído com 40 intervalos e 100 partículas, e a curva de calibração utilizando 8 variáveis latentes, para a concentração de biodiesel, é apresentada na Figura 26.

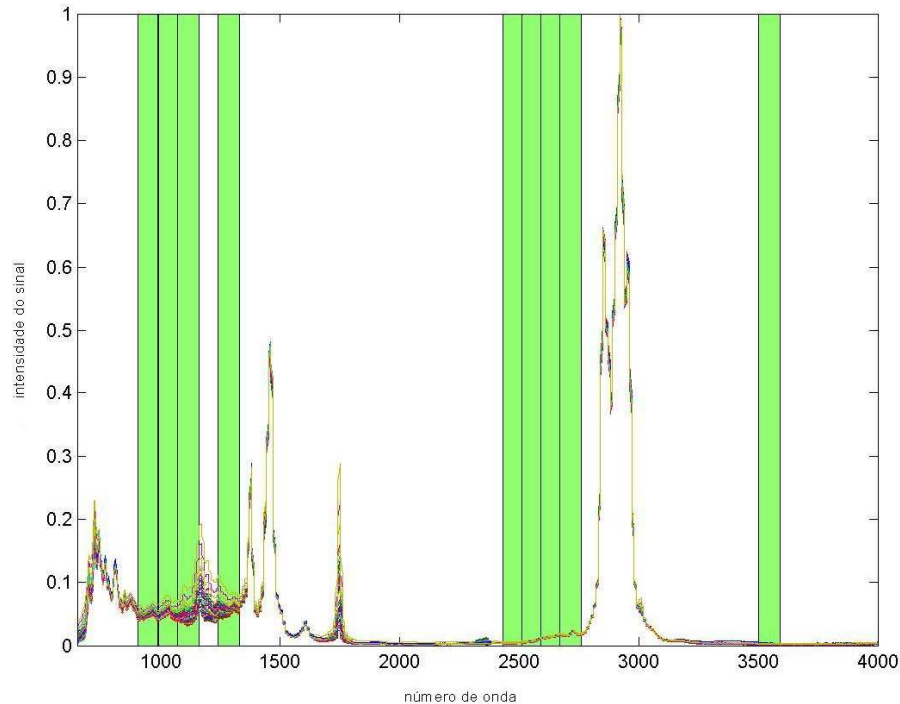


Figura 25 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

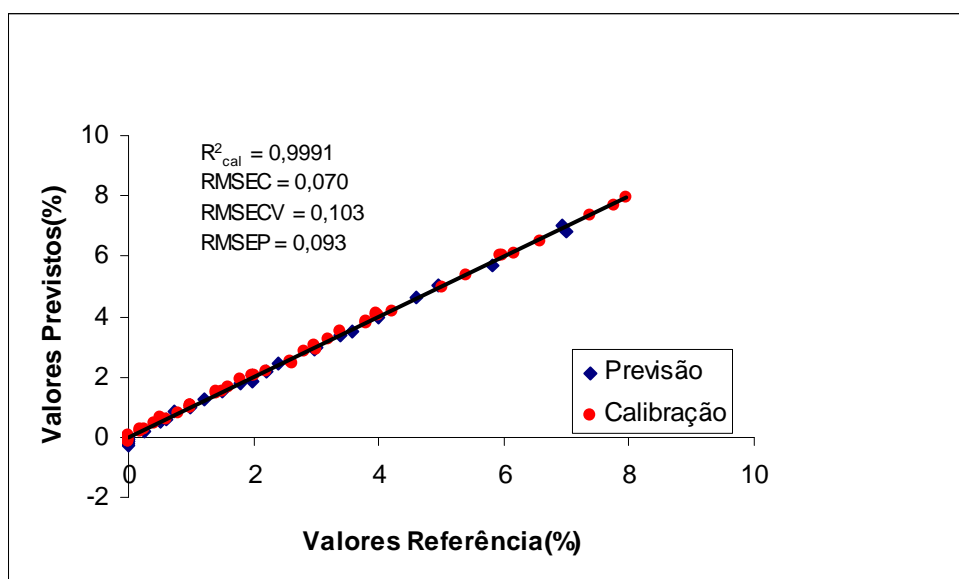


Figura 26– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

5.1.8 Considerações sobre os modelos para a previsão do biodiesel

Conforme evidenciado, todos os modelos construídos para quantificação do biodiesel empregam regiões espectrais semelhantes, e sendo que todos foram capazes de quantificar o analito proposto. Os resultados encontrados empregando as soluções propostas pelo método enxame de partículas, em comparação aos resultados empregando o PLS sobre todo o espectro bem como as regiões espectrais sugeridas pela proposta de norma ABNT, demonstraram que houve uma otimização dos modelos de regressão multivariados.

A região selecionada pelos modelos otimizados contempla em parte a região espectral, sugerida pela proposta de norma, e ampliada principalmente na faixa dos sinais referentes ao C-O. Um aspecto interessante é que em nenhum dos modelos otimizados, selecionou a região da carbonila (C=O), região esta que a proposta de norma preconiza como sendo a principal banda para a quantificação do biodiesel.

Uma hipótese é que a região da carbonila não apresenta informações suficientes para diferenciar a contribuição do sinal do biodiesel e do óleo isoladamente, ou seja, provavelmente as modificações que acontecem do óleo para o biodiesel são pouco pronunciadas nesta região e muito mais pronunciadas na região do sinal correspondente ao C-O.

Com base nestas observações e considerando alguns dos resultados das otimizações propostas pelo método enxame de partículas, é possível obter modelos com boa capacidade preditiva com as informações espectrais das regiões de 904 a 1327 cm^{-1} com lacunas, ilustradas na Figura 27, obtidas pelo método iPLS-PSO em uma de suas soluções. O desempenho deste modelo, com 10 variáveis latentes, pode ser visualizado na Figura 28.

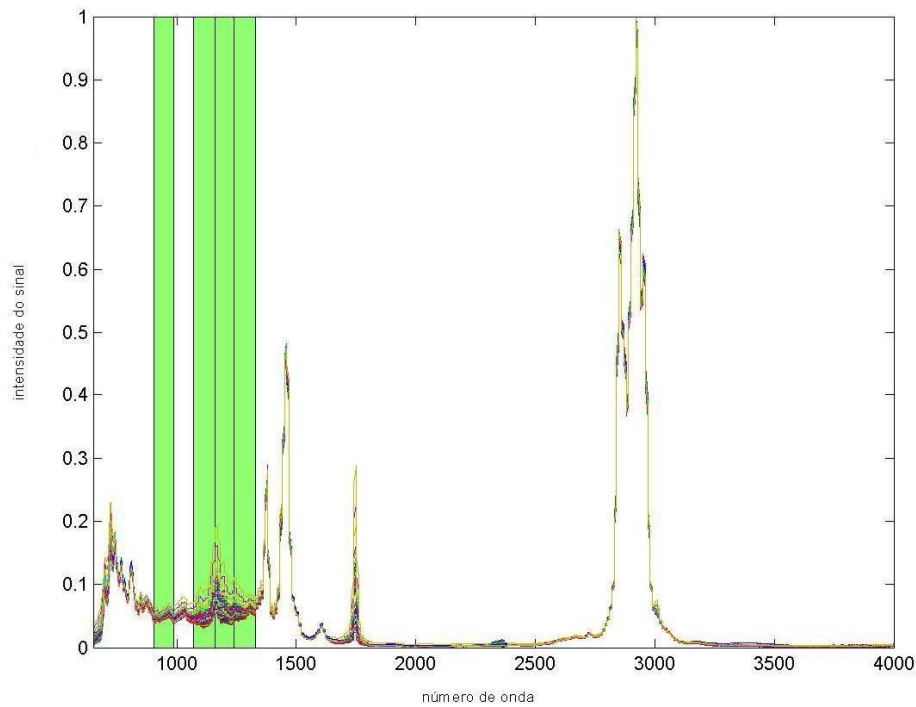


Figura 27 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

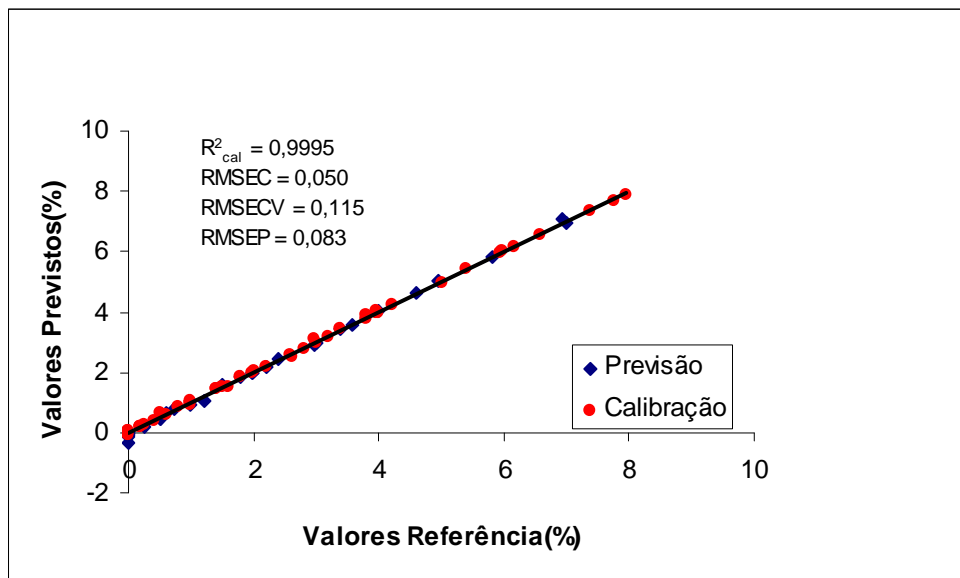


Figura 28– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

5.2 Resultados para a quantificação do óleo vegetal

Neste trabalho também foram desenvolvidos diversos modelos de regressão visando quantificar o óleo vegetal presente nas misturas, com o intuito de verificar a capacidade de detecção de adulterantes nas misturas biodiesel/óleo vegetal/diesel através do uso do FT-IR/ATR. Os principais resultados obtidos estão descritos na seqüência. Primeiramente são apresentados os modelos obtidos sem a otimização através da meta-heurística implementada e em seguida, são descritos os modelos obtidos pela combinação da PSO e do iPLS.

5.2.1 Modelos para determinação do óleo vegetal com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT

Com base no conjunto de dados foram desenvolvidos dois modelos de regressão, ambos utilizando para construção o método PLS. O primeiro modelo, para a quantificação do óleo vegetal foi construído utilizando toda região espectral (650 cm^{-1} a 4000 cm^{-1}) conforme Figura 11, apresentada anteriormente, e o desempenho do modelo é apresentado na Figura 29, empregando 6 variáveis latentes.

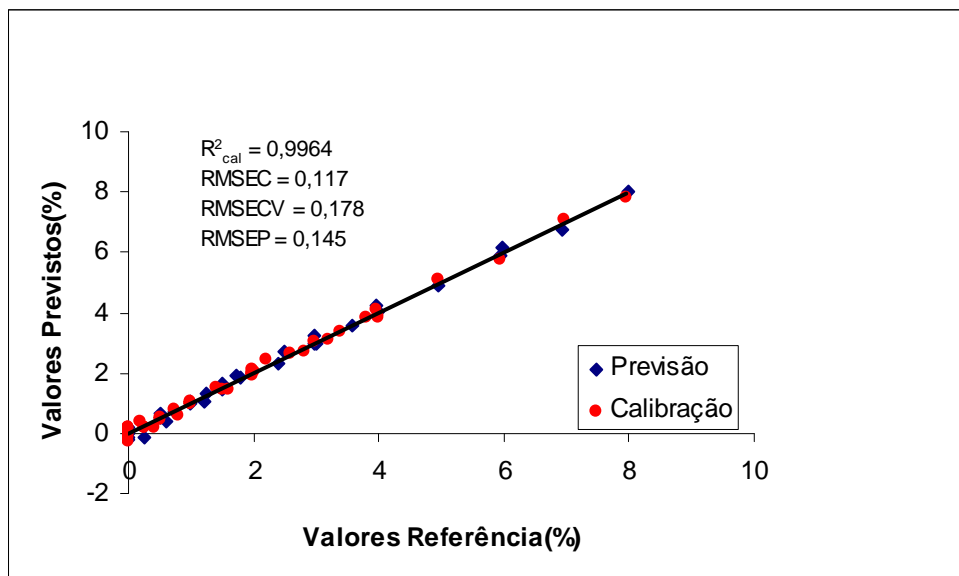


Figura 29 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal utilizando toda a região espectral das misturas.
 Fonte: elaborado pela autora.

O segundo modelo construído para a quantificação do óleo vegetal nas misturas empregou somente as duas regiões espectrais (de 1100 cm^{-1} a 1200 cm^{-1} e de 1735 cm^{-1} a 1750 cm^{-1}) sugeridas pela proposta de norma ABNT, conforme Figura 13. O desempenho do modelo, utilizando 5 variáveis latentes, é apresentado na Figura 30.

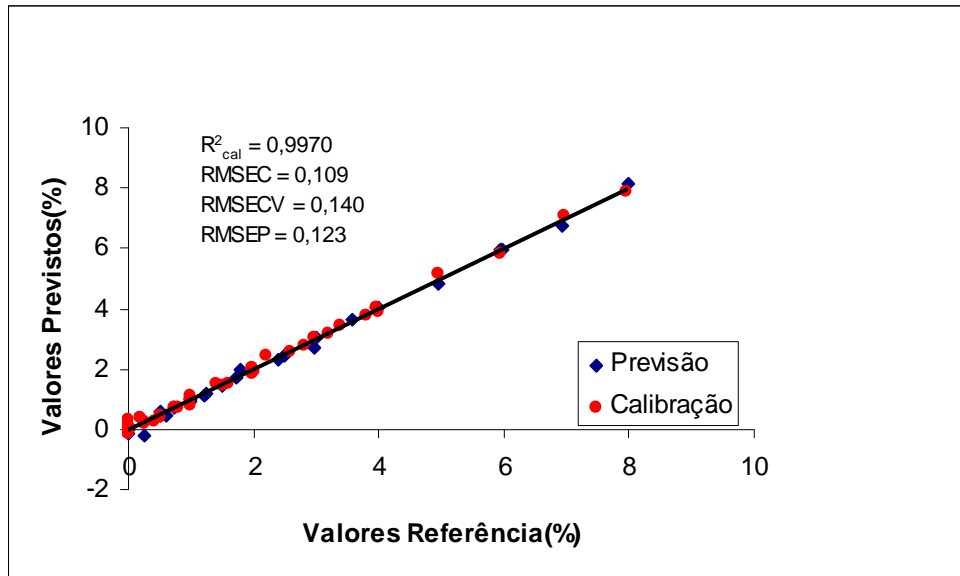


Figura 30 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT. Fonte: elaborado pela autora.

Observando os dois modelos anteriores pode-se verificar que os resultados obtidos pelo modelo construído, empregando somente as regiões espectrais sugerida pela proposta de norma ABNT, apresenta uma melhor capacidade de quantificação do óleo vegetal do que o modelo construído com o mesmo propósito, empregado toda a região do espectro.

Tentando analisar a possibilidade de obter modelos com melhor capacidade preditiva, sugere-se a mesma implementação proposta para construção dos modelos de biodiesel, desta vez com o propósito de otimizar a quantificação do óleo vegetal nas misturas.

5.2.2 Resultados gerais e evoluções dos modelos para a previsão do óleo vegetal

Foram testadas na quantificação do óleo vegetal, as mesmas cinco implementações da otimização por enxame de partículas: iPLS-PSO, iPLS-PSO (CC), iPLS-PSO (DD); iPLS-PSO (CD); iPLS-PSO (DC), descritas anteriormente. Os principais resultados dos modelos, incluindo os obtidos através das implementações com a PSO, são apresentados na Tabela

6. Os resultados completos das implementações do enxame de partículas para a quantificação do óleo vegetal encontram-se no ANEXO B.

Tabela 6 – Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do óleo vegetal.

Modelos	Regiões empregadas	R² cal	RMSEC	RMSECV	RMSEP
espectro inteiro	650 - 4000 cm ⁻¹	0,9964	0,117	0,178	0,145
regiões sugeridas pela proposta de norma ABNT	1101 - 1198 cm ⁻¹ 1736 - 1749 cm ⁻¹	0,9970	0,109	0,140	0,123
iPLS- PSO	904 - 1412 cm ⁻¹ 1583 - 1751 cm ⁻¹ 2509 - 2756 cm ⁻¹ 3504 - 3585 cm ⁻¹	0,9991	0,059	0,083	0,092
iPLS- PSO (CC)	904 - 1412 cm ⁻¹ 1583 - 1751 cm ⁻¹ 1838 - 2006 cm ⁻¹	0,9990	0,060	0,088	0,089
iPLS- PSO (DD)	904 - 1327 cm ⁻¹ 1583 - 1666 cm ⁻¹ 2008 - 2341 cm ⁻¹	0,0995	0,042	0,082	0,089
iPLS- PSO (CD)	804 - 1319 cm ⁻¹ 1992 - 2158 cm ⁻¹ 2495 - 2829 cm ⁻¹ 3334 - 3500 cm ⁻¹	0,9995	0,041	0,081	0,094
iPLS- PSO (DC)	820 - 1751 cm ⁻¹ 2426 - 2756 cm ⁻¹ 3007- 3253 cm ⁻¹	0,9992	0,056	0,085	0,090

Fonte: elaborado pela autora.

Um comparativo das evoluções para 3 execuções do método iPLS-PSO com 20 intervalos e 50 partículas para quantificação do óleo vegetal, é apresentada na Figura 31. Observa-se nesta proposta que existe uma tendência de convergência de todos os modelos, sendo necessário em média de 108 iterações; mostrando juntamente que o número total de iterações definido é suficiente para explorar todas as melhores combinações possíveis entre os 20 intervalos.

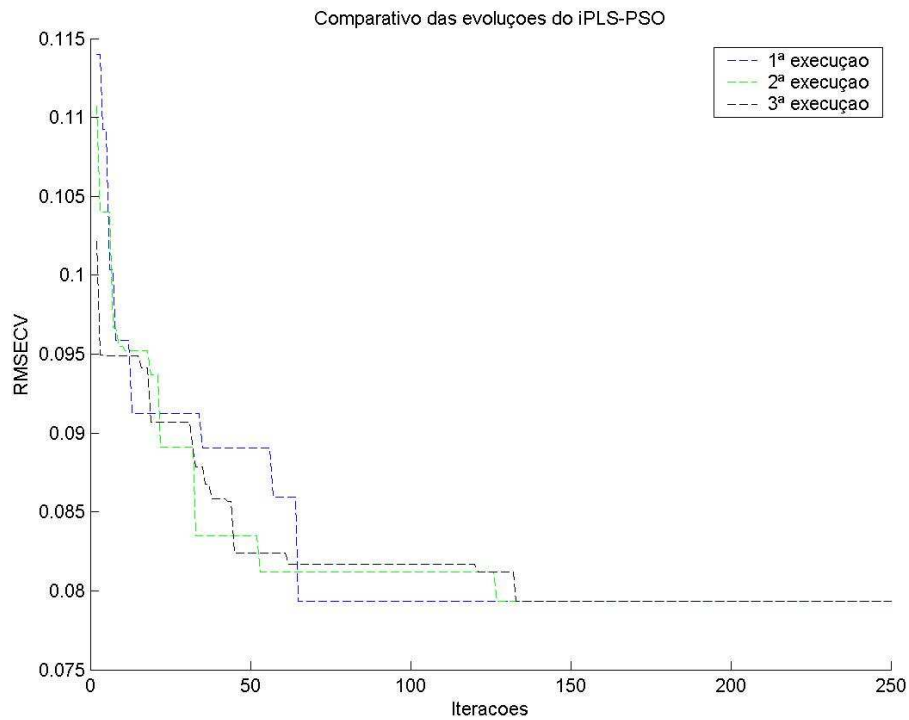


Figura 31– Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas, para a quantificação do óleo vegetal.
 Fonte: elaborado pela autora.

Observa-se um comportamento diferente quando se apresentam as 3 evoluções do iPLS-PSO com 40 e 60 intervalos, ambos com 50 partículas, como pode ser visto nas Figuras 32 e 33. São necessários mais iterações para que os modelos atinjam mínimos locais, em média para iPLS-PSO com 40 intervalos são necessárias 184 iterações, já para 60 intervalos em média 209 iterações.

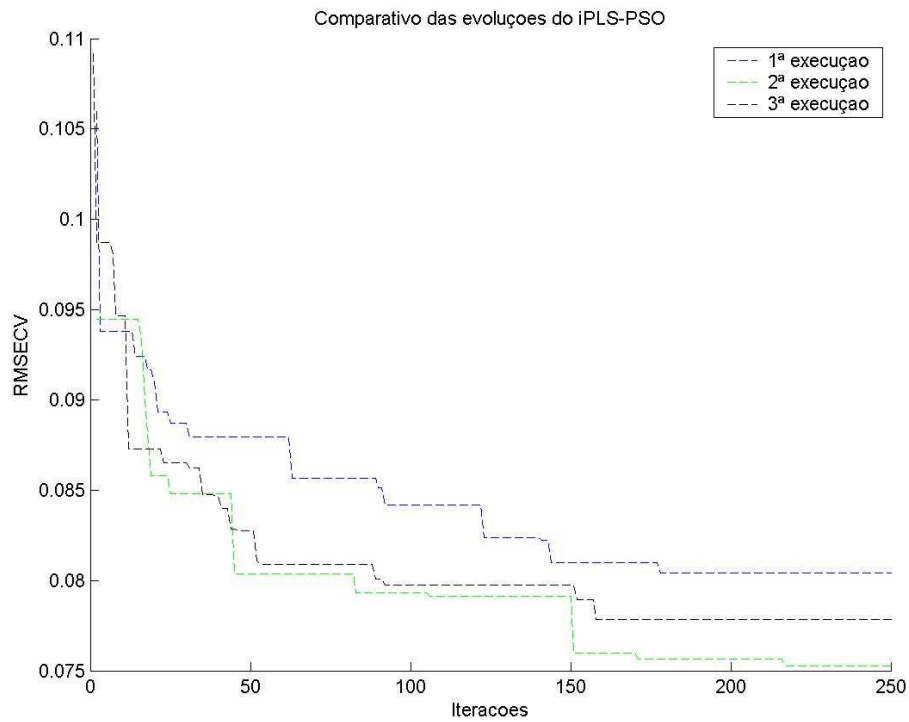


Figura 32 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas, para a quantificação do óleo vegetal.
Fonte: elaborado pela autora.

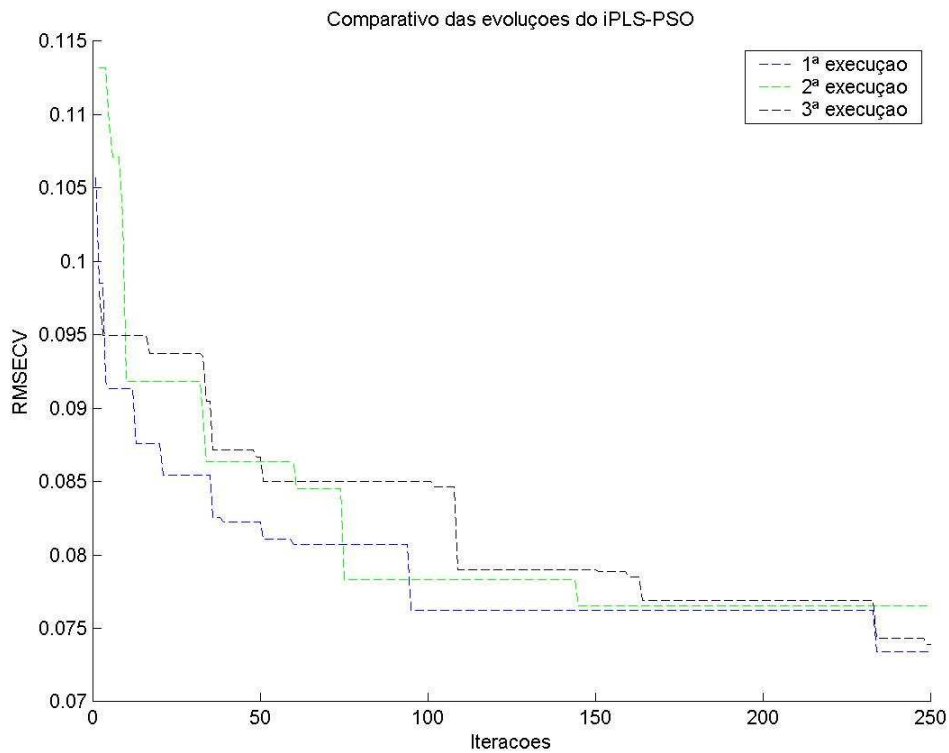


Figura 33 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para a quantificação do óleo vegetal.
Fonte: elaborado pela autora.

5.2.3 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO

Nesta primeira implementação denominada iPLS-PSO aplicada ao conjunto de dados propostos para quantificar o óleo vegetal, as regiões espectrais selecionadas para construção do modelo foram as que compreendem de 904 a 1412 cm^{-1} com lacunas, de 1583 a 1751 cm^{-1} , de 2509 a 2756 cm^{-1} e de 3504 a 3585 cm^{-1} como podem ser visualizadas na Figura 34. Dentre todos os modelos construídos, o selecionado foi o construído com 40 intervalos e 25 partículas e a sua respectiva curva de calibração, com 9 variáveis latentes, para quantificação do óleo vegetal, é apresentada na Figura 35.

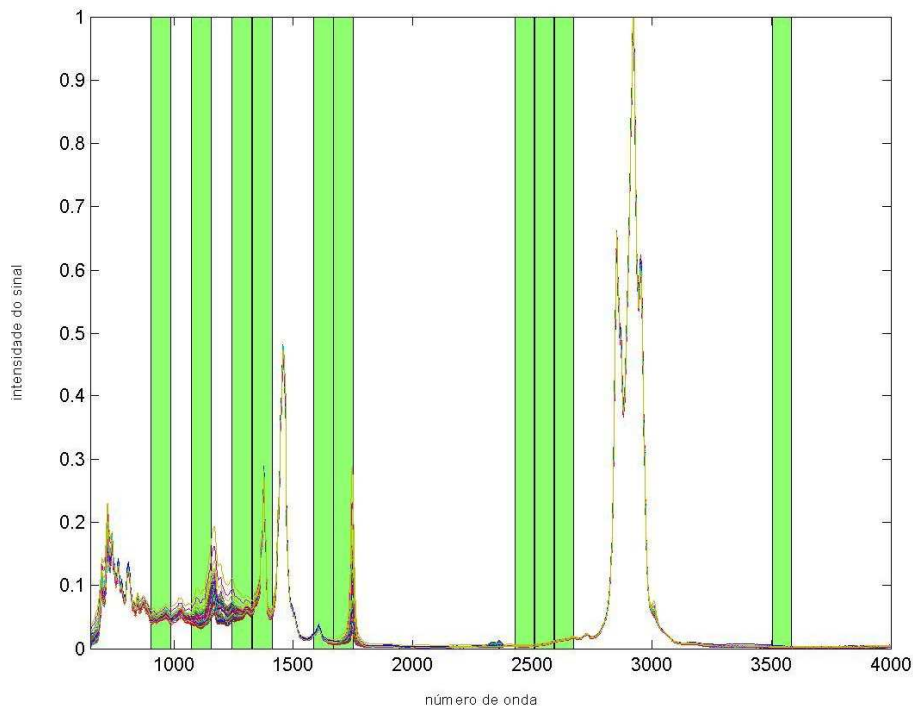


Figura 34 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

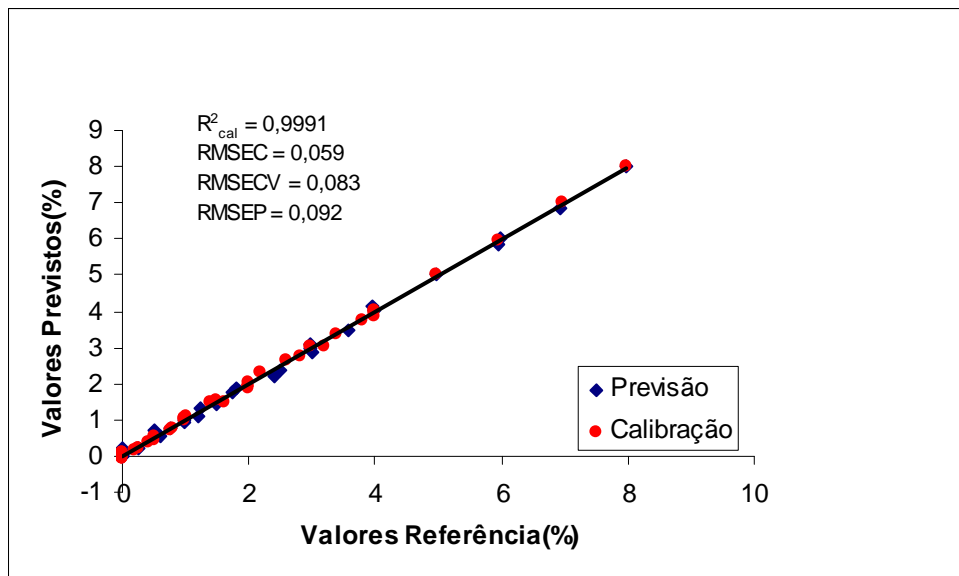


Figura 35 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.
 Fonte: elaborado pela autora.

5.2.4 Resultados obtido para a determinação do óleo vegetal aplicando o método iPLS-PSO (CC)

Na implementação iPLS-PSO (CC) as regiões espectrais selecionadas para construção do modelo de quantificação do óleo vegetal compreendem as faixas de 904 a 1412 cm^{-1} com lacunas, de 1583 a 1666 cm^{-1} e de 2008 a 2341 cm^{-1} , conforme Figura 36. Nesta abordagem, o modelo selecionado empregou 40 intervalos e 25 partículas e a sua respectiva curva de calibração, com 9 variáveis latentes, para a concentração do óleo vegetal, é apresentada na Figura 37.

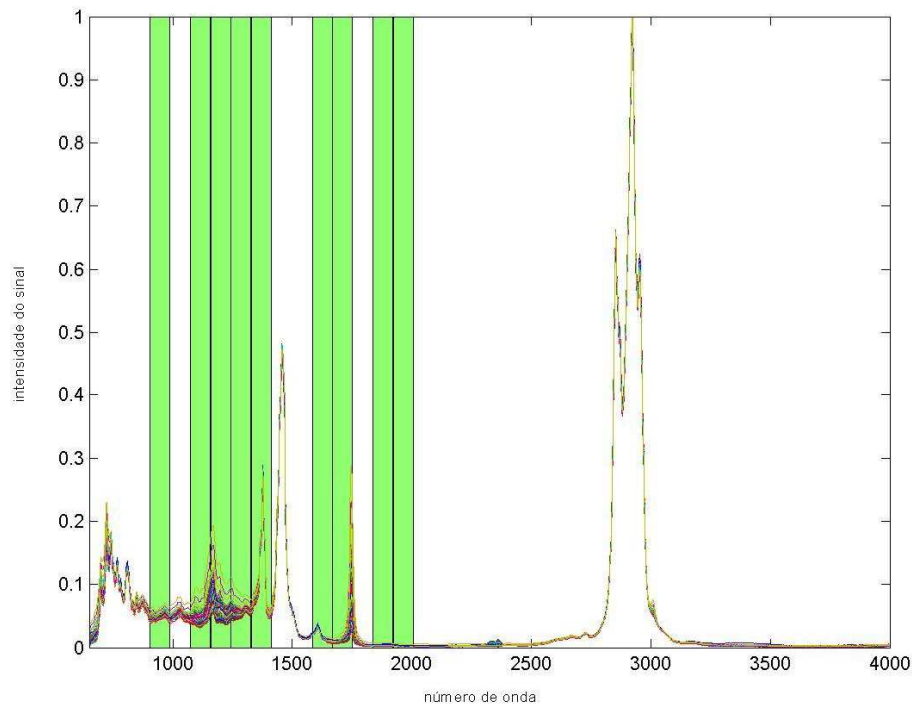


Figura 36 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

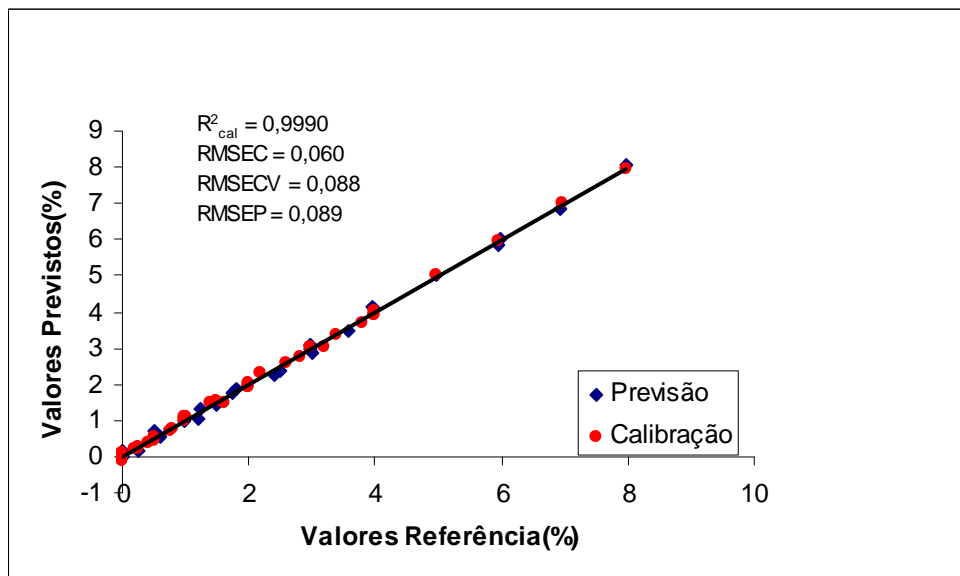


Figura 37 – Valores de referência versus previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.2.5 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DD)

Já o método iPLS-PSO (DD) para a quantificação do óleo vegetal nas misturas, o modelo selecionado foi aquele construído com 40 intervalos e 100 partículas, compreendendo as regiões espectrais selecionadas nas faixas de 904 a 1327 cm^{-1} com lacunas, de 1583 a 1666 cm^{-1} e de 2008 a 2341 cm^{-1} com lacunas, conforme Figura 38, e a curva de calibração, com 10 variáveis latentes, para a concentração do óleo vegetal é apresentada na Figura 39.

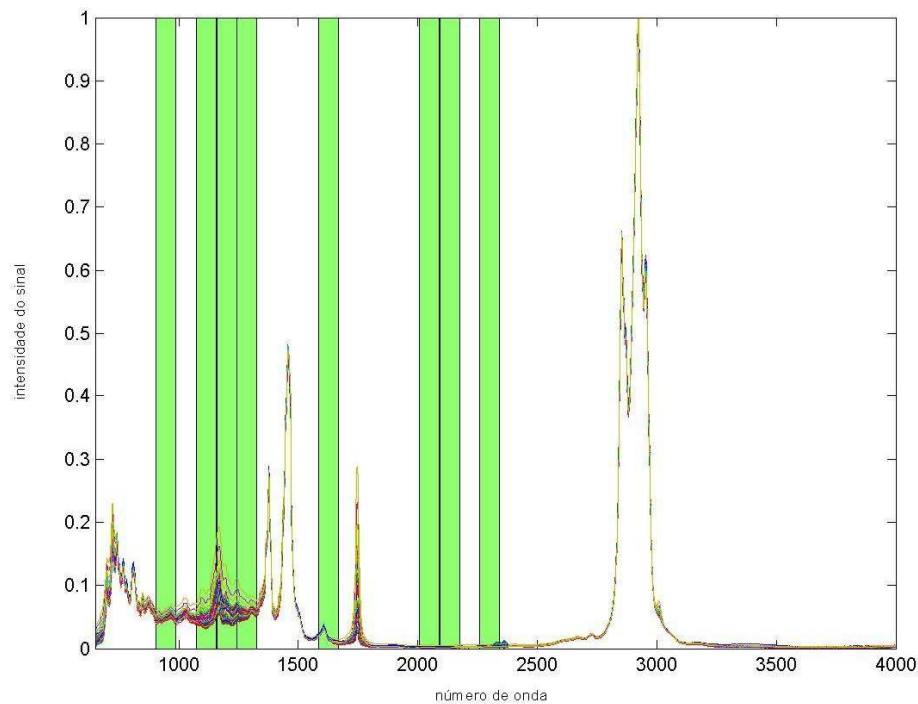


Figura 38 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

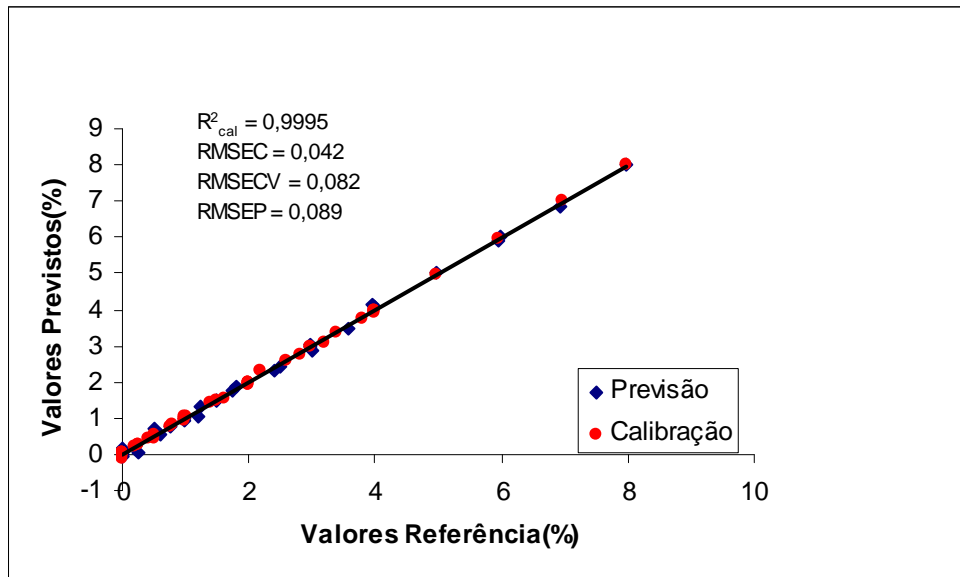


Figura 39 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas. Fonte: elaborado pela autora.

5.2.6 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (CD)

No caso da implementação iPLS-PSO (CD) as regiões espectrais selecionadas compreendem as faixas de 804 a 1319 cm^{-1} , de 1992 a 2158 cm^{-1} , de 2495 a 2829 cm^{-1} e de 3334 a 3500 cm^{-1} , conforme Figura 40. Neste caso, o modelo selecionado, dentre todos os modelos construídos para esta abordagem, foi aquele elaborado com 20 intervalos e 25 partículas, e a curva de calibração, com 10 variáveis latentes, para a concentração do óleo vegetal é apresentada na Figura 41.

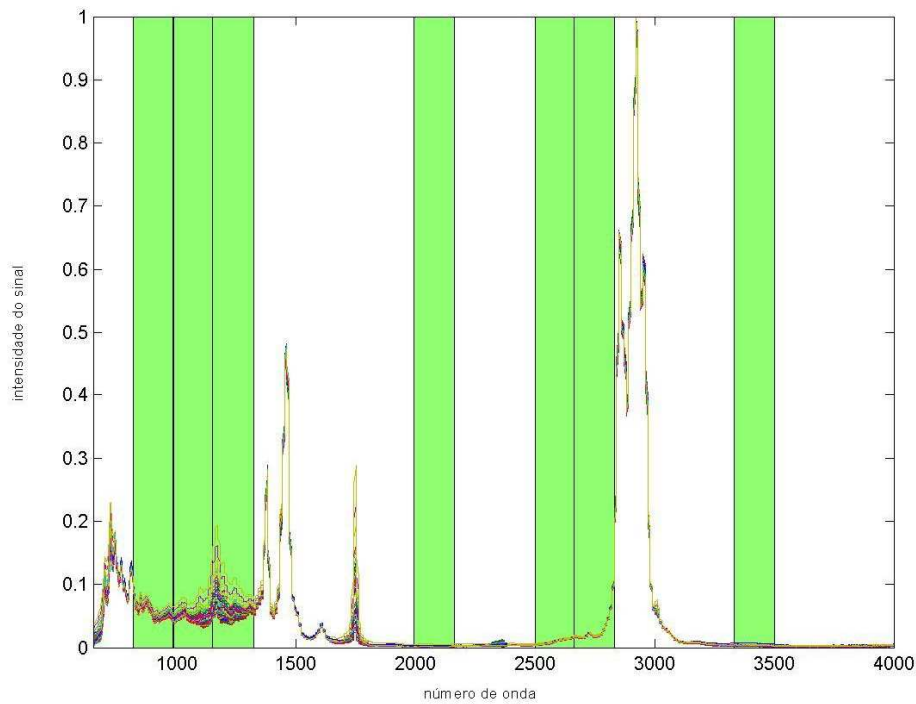


Figura 40 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.
Fonte: elaborado pela autora.

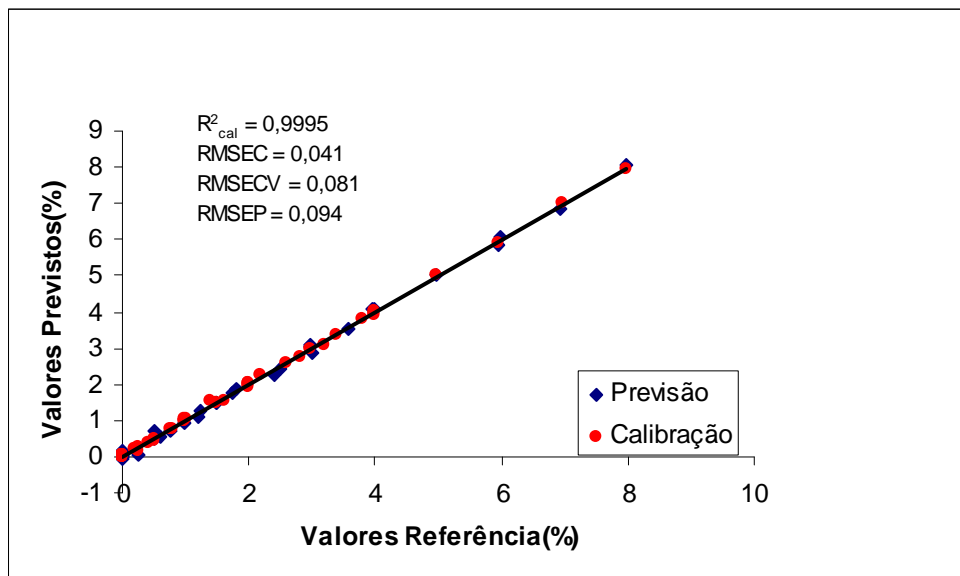


Figura 41– Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.2.7 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DC)

Já a implementação denominada iPLS-PSO (DC) selecionou as seguintes regiões espectrais para construção do modelo de 820 a 1751 cm^{-1} , de 2426 a 2756 cm^{-1} e de 3007 a 3253 cm^{-1} como podem ser visualizadas na Figura 42. Dentre todos os modelos construídos para iPLS-PSO (DC), o selecionado foi aquele construído com 40 intervalos e 100 partículas e a sua respectiva curva de calibração com 10 variáveis latentes, para quantificação do óleo vegetal, é apresentada na Figura 43.

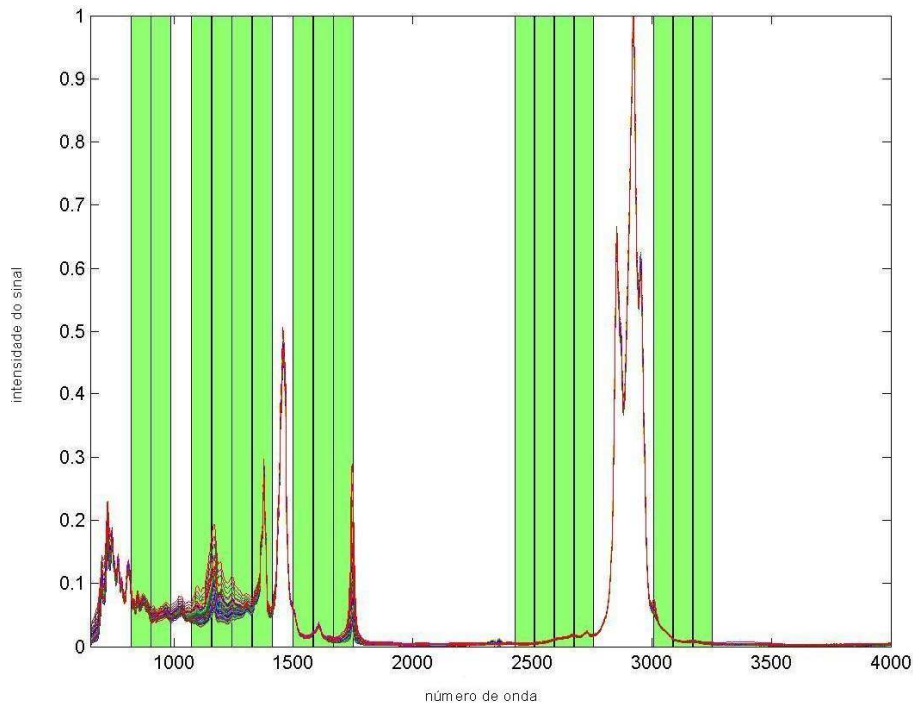


Figura 42 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

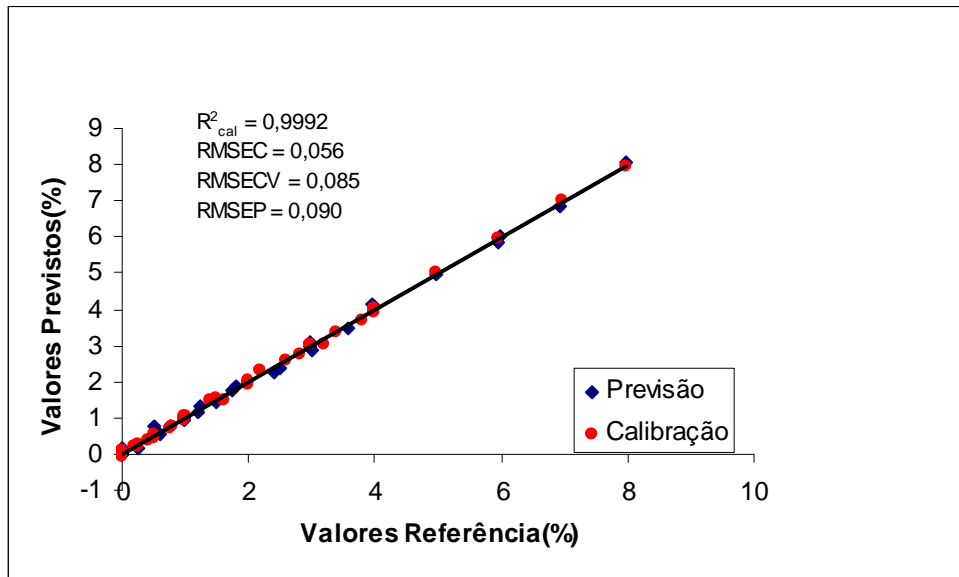


Figura 43 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas. Fonte: elaborado pela autora.

5.2.8 Considerações sobre os modelos para a previsão de óleo vegetal

Analisando os resultados anteriores, podemos observar que todos os modelos construídos para quantificar o óleo vegetal foram capazes de atingir este propósito, apresentando baixos erros de previsão (RMSEP). Cabe ressaltar que os modelos construídos empregando o método PLS com o espectro inteiro das misturas e somente com as regiões sugeridas pela proposta de norma ABNT apresentaram uma capacidade de quantificação do óleo vegetal menor do que os modelos construídos através do método híbrido proposto nesta pesquisa.

Os modelos selecionados através das otimizações aqui propostas apresentaram valores de RMSEP entre 0,089 a 0,094, enquanto que os valores de RMSECV ficaram na amplitude de 0,081 a 0,088.

Do ponto de vista das regiões selecionadas, o modelo para previsão do óleo vegetal além de necessitar da região do C-O, também necessita utilizar as informações referentes ao sinal da carbonila (C=O), conforme as figuras anteriores, que apresentam as regiões selecionadas.

5.3 Resultados para a quantificação do óleo diesel

Finalmente, o outro analito quantificado através de modelos de regressão multivariados nas misturas, neste trabalho foi o óleo diesel. Embora não seja comum quantificar o componente majoritário das misturas, o propósito desta determinação é verificar o comportamento do algoritmo e da técnica FT-IR/ATR. Para isso foram desenvolvidos diversos modelos de regressão. Neste caso, foram construídos modelos usando o PLS com o espectro inteiro das misturas e posteriormente modelos com as implementações propostas com a meta-heurística enxame de partículas em conjunto com o iPLS.

5.3.1 Modelo para a determinação do óleo diesel com espectro inteiro

Para a determinação do óleo diesel também foi desenvolvido um modelo de regressão com o método PLS, utilizando toda a região espectral (650 cm^{-1} a 4000 cm^{-1}) conforme Figura 11 apresentada anteriormente. O desempenho do modelo, com 4 variáveis latentes, é apresentado na Figura 44.

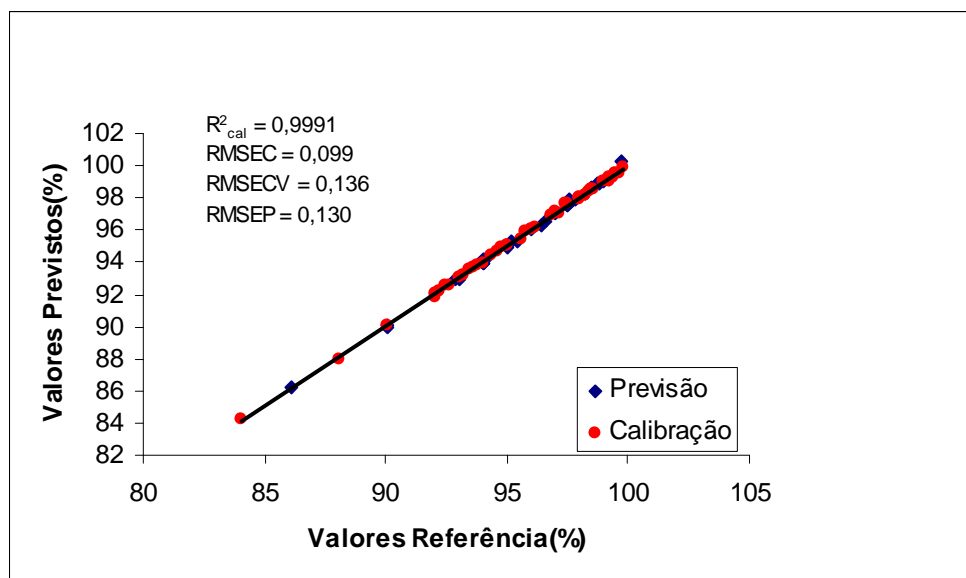


Figura 44 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel utilizando toda a região espectral da mistura.
 Fonte: elaborado pela autora.

5.3.2 Resultados gerais e evoluções para a determinação do óleo diesel

Tentando avaliar se o comportamento das implementações propostas pelo enxame de partículas seria capaz gerar modelos mais robustos para quantificação do óleo diesel, foram testadas as cinco implementações na quantificação deste analito. Os principais resultados dos modelos, incluindo as resultantes das implementações com a PSO são apresentados na Tabela 7.

Tabela 7 – Principais resultados dos modelos, incluindo das cinco implementações da PSO para a quantificação do óleo diesel.

Modelos	Regiões empregadas	R ² cal	RMSEC	RMSECV	RMSEP
espectro inteiro	650 - 4000 cm ⁻¹	0,9991	0,099	0,136	0,130
iPLS- PSO	735 - 903 cm ⁻¹	0,9998	0,048	0,098	0,102
	1159 - 1327 cm ⁻¹				
	3172 - 3253 cm ⁻¹				
iPLS- PSO (CC)	762 - 1207 cm ⁻¹	0,9998	0,047	0,099	0,108
	2719 - 2773 cm ⁻¹				
	3223 - 3332 cm ⁻¹				
	3446 - 3500 cm ⁻¹				
iPLS- PSO (DD)	904 - 1327 cm ⁻¹	0,9997	0,059	0,096	0,112
	2092 - 2175 cm ⁻¹				
	2426 - 2756 cm ⁻¹				
	3089 - 3170 cm ⁻¹				
	3504 - 3585 cm ⁻¹				
	3753 - 3834 cm ⁻¹				
iPLS- PSO (CD)	904 - 1327 cm ⁻¹	0,9999	0,038	0,095	0,108
	1753 - 1836 cm ⁻¹				
	2260 - 2507 cm ⁻¹				
	2675 - 2756 cm ⁻¹				
	3753 - 3834 cm ⁻¹				
iPLS- PSO (DC)	735 - 1327 cm ⁻¹	0,9997	0,055	0,108	0,111

Fonte: elaborado pela autora.

A seguir, com base na implementação iPLS-PSO com 20 intervalos e 50 partículas para quantificação do óleo diesel, é mostrado um comparativo das evoluções para as 3 execuções, conforme a Figura 45. Observa-se nesta proposta que existe uma tendência de convergência de todos os modelos, sendo necessário em média de 107 iterações; comprovando novamente que o número de iterações definido é suficiente para explorar todas as melhores combinações possíveis entre os 20 intervalos.

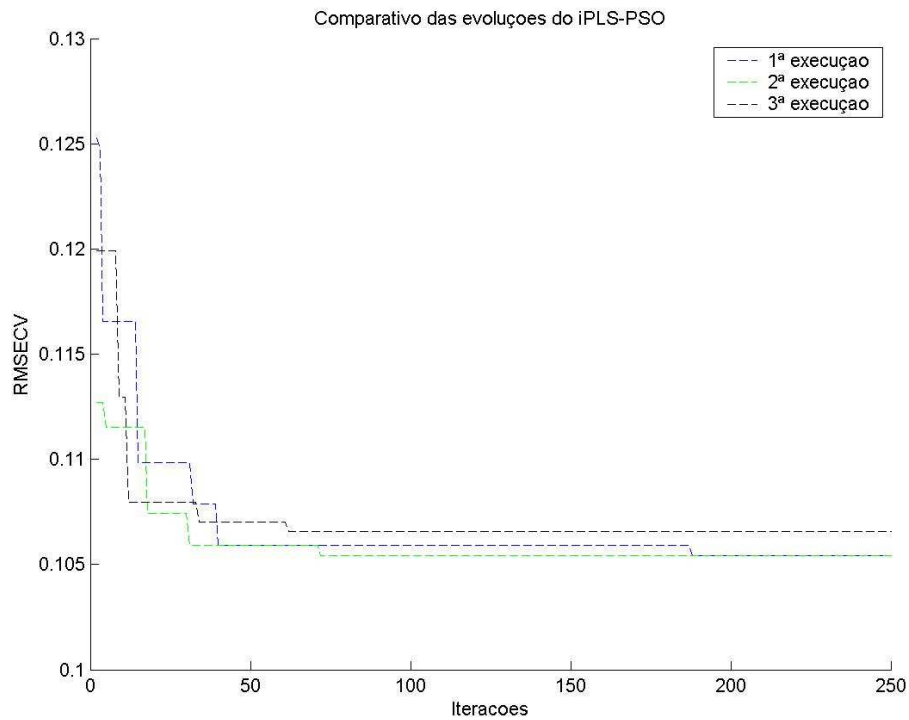


Figura 45 – Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas para a quantificação do óleo diesel.
Fonte: elaborado pela autora.

O comportamento das evoluções para as 3 execuções do método iPLS_PSO com 40 e 60 intervalos, ambos com 50 partículas, é diferente, pois a maioria dos modelos necessita de um número de iterações superior do que o modelo com 20 partículas, para atingirem os mínimos locais. Na evolução iPLS-PSO com 40 intervalos, percebe-se que em média são necessárias 155 iterações, já para 60 intervalos, em média, são necessárias 123 iterações como pode ser visualizada nas Figuras 46 e 47.

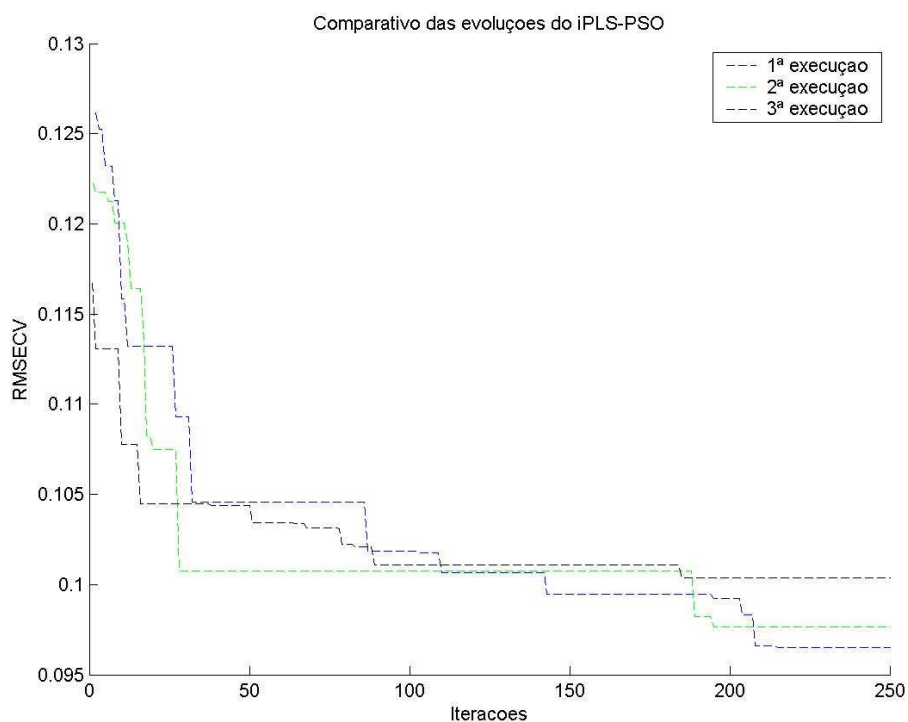


Figura 46 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas para quantificação do óleo diesel.
Fonte: elaborado pela autora.

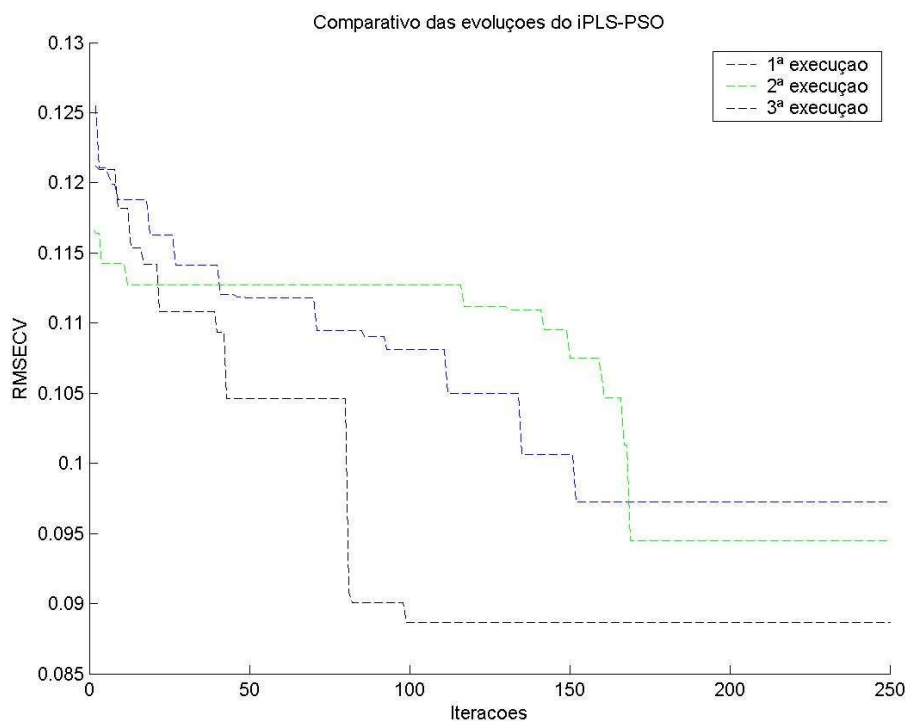


Figura 47 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para quantificação do óleo diesel.
Fonte: elaborado pela autora.

5.3.3 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO

Nesta implementação denominada iPLS-PSO, utilizada para quantificação do óleo diesel no conjunto de dados proposto, selecionou regiões espectrais para construção do modelo que compreendem de 735 a 903 cm^{-1} , de 1159 a 1327 cm^{-1} e de 3172 a 3253 cm^{-1} como podem ser visualizadas na Figura 48. Analisando todos os modelos construídos, o modelo selecionado foi o construído com 40 intervalos e 25 partículas e, a sua respectiva curva de calibração empregando 10 variáveis latentes é apresentada na Figura 49.

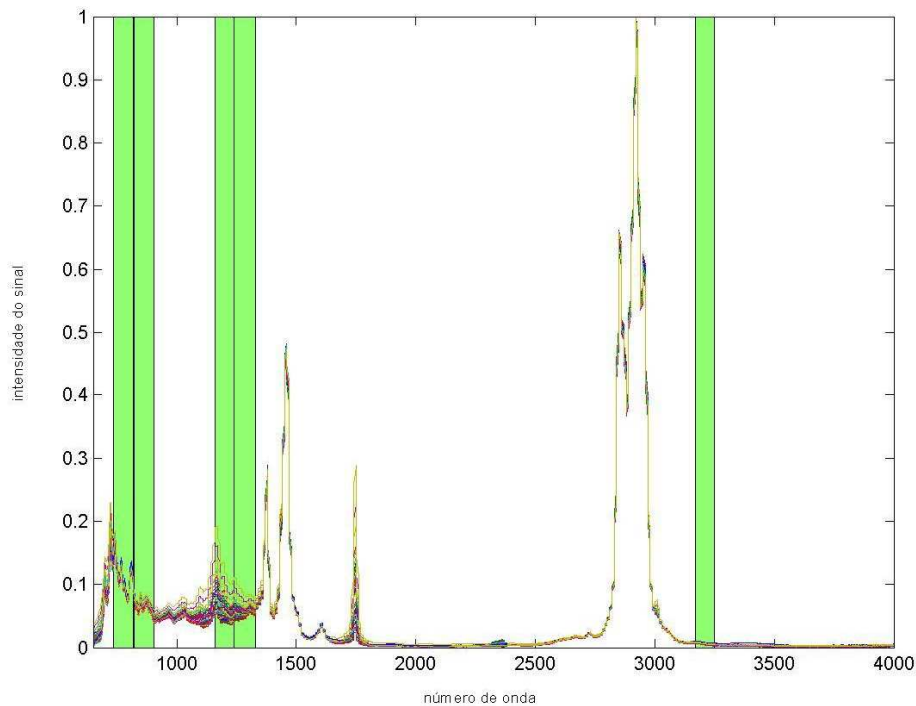


Figura 48 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

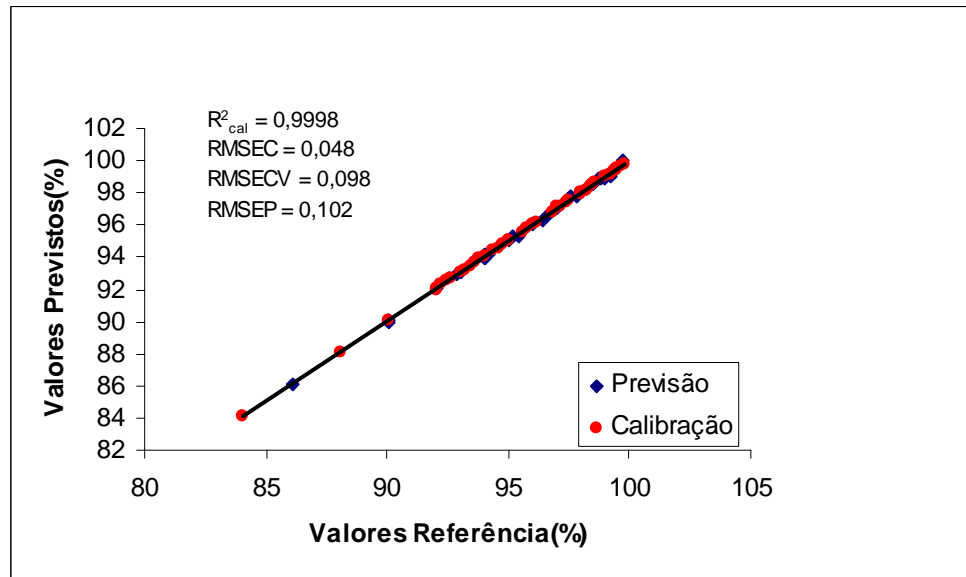


Figura 49 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.
 Fonte: elaborado pela autora.

5.3.4 Resultados obtidos para a quantificação do óleo diesel aplicando o método iPLS-PSO (CC)

Quando aplicado o método proposto iPLS-PSO (CC), para quantificação do óleo diesel, as regiões espectrais selecionadas para construção do modelo compreendem de 762 a 1207 cm^{-1} com lacunas, de 2719 a 2773 cm^{-1} , de 3223 a 3332 cm^{-1} e de 3446 a 3500 cm^{-1} , visualizadas na Figura 50. O modelo selecionado, dentre todos os construídos para este método, foi o proposto com 60 intervalos e 25 partículas e, a sua respectiva curva de calibração empregando 10 variáveis latentes é apresentada na Figura 51.

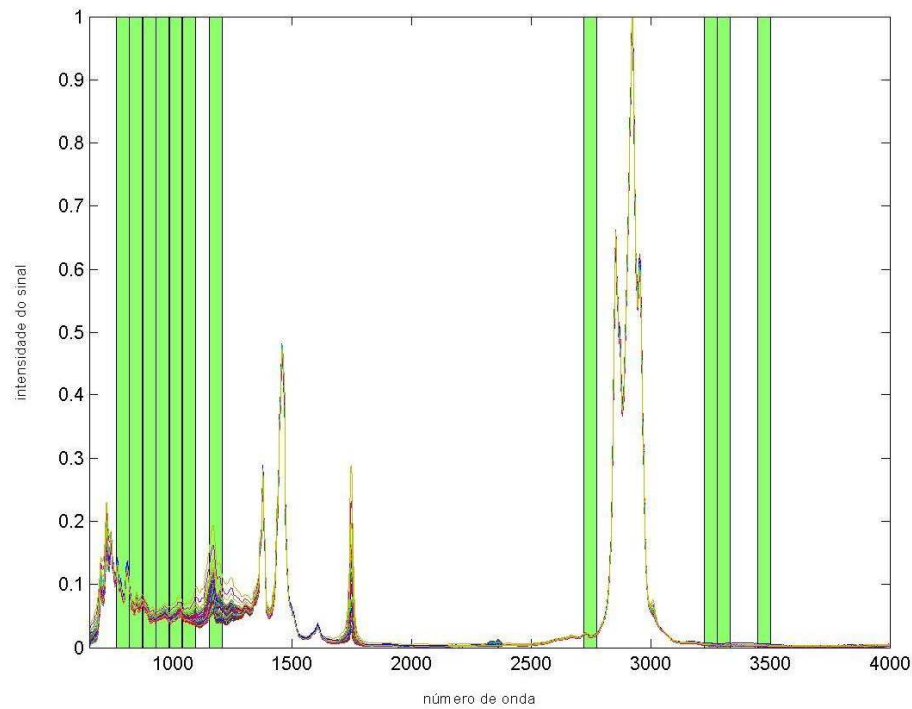


Figura 50 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.
Fonte: elaborado pela autora.

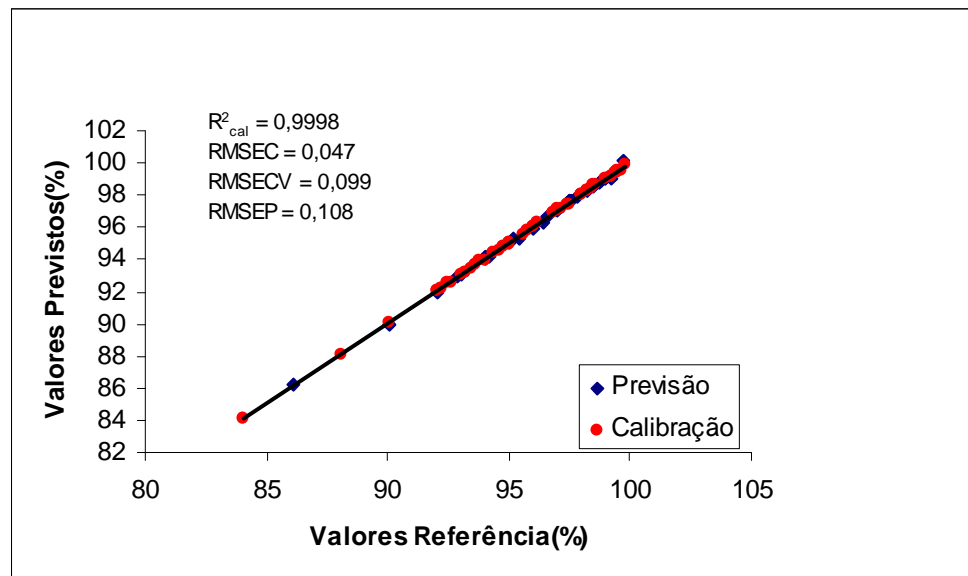


Figura 51 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.3.5 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DD)

Nos testes para quantificação do óleo diesel, aplicando o método iPLS-PSO (DD), foram selecionadas as regiões espectrais que compreendem de 904 a 1327 cm^{-1} com lacunas, de 2092 a 2175 cm^{-1} , de 2426 a 2756 cm^{-1} , de 3089 a 3170 cm^{-1} , de 3504 a 3585 cm^{-1} e de 3753 a 3834 cm^{-1} para a construção do modelo de regressão, visualizadas na Figura 52. O modelo selecionado, dentre todos os construídos para este método, foi o proposto com 40 intervalos e 100 partículas e, a sua respectiva curva de calibração utilizando 8 variáveis latentes é apresentada na Figura 53.

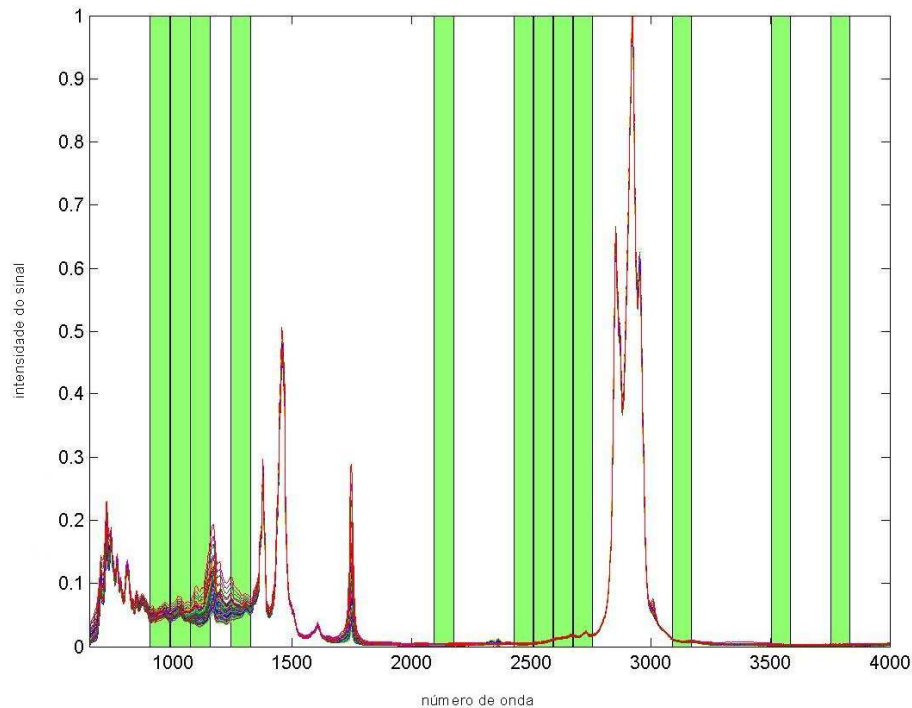


Figura 52 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

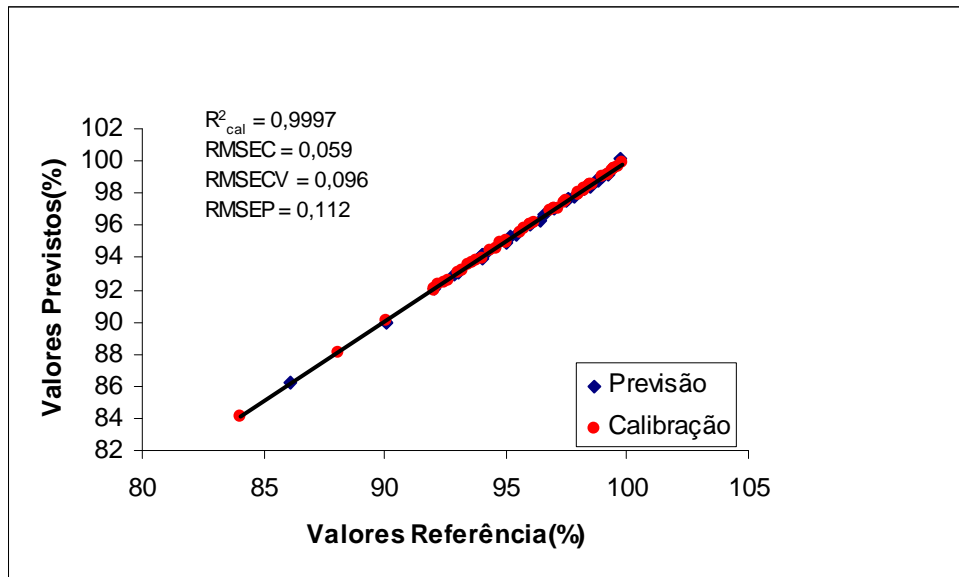


Figura 53 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas. Fonte: elaborado pela autora.

5.3.6 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (CD)

Já a aplicação do método iPLS-PSO (CD) para a quantificação do óleo diesel nas misturas, selecionou as regiões espectrais nas faixas de 904 a 1327 cm^{-1} com lacunas, de 1753 a 1836 cm^{-1} , de 2260 a 2507 cm^{-1} , de 2675 a 2756 cm^{-1} e, de 3753 a 3834 cm^{-1} visualizadas na Figura 54. O modelo selecionado, dentre todos os construídos para este método, foi o proposto com 40 intervalos e 50 partículas e, a sua respectiva curva de calibração utilizando 10 variáveis latentes é apresentada na Figura 55.

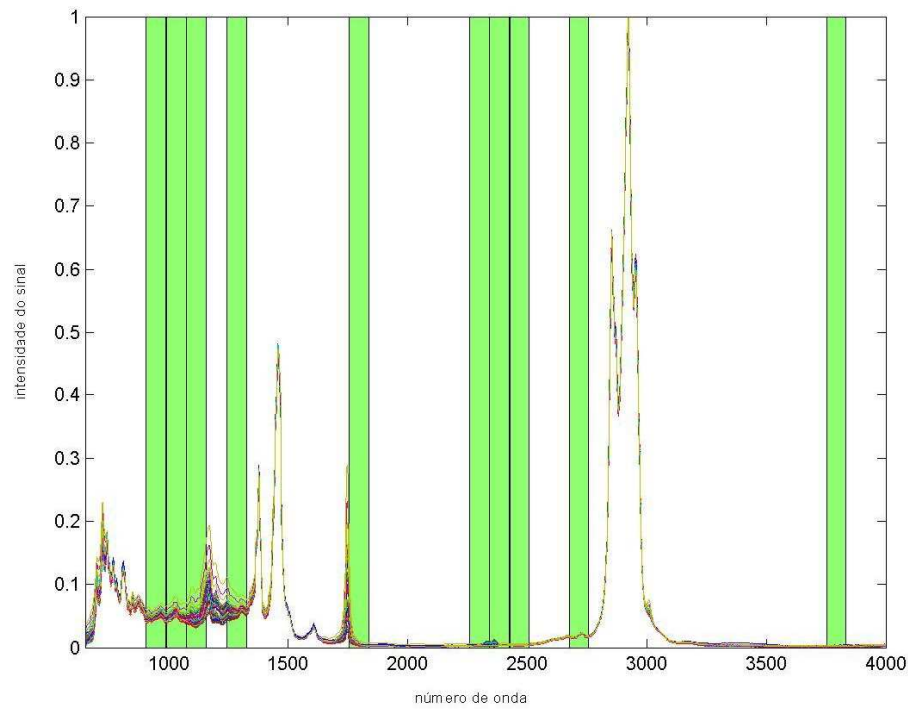


Figura 54 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

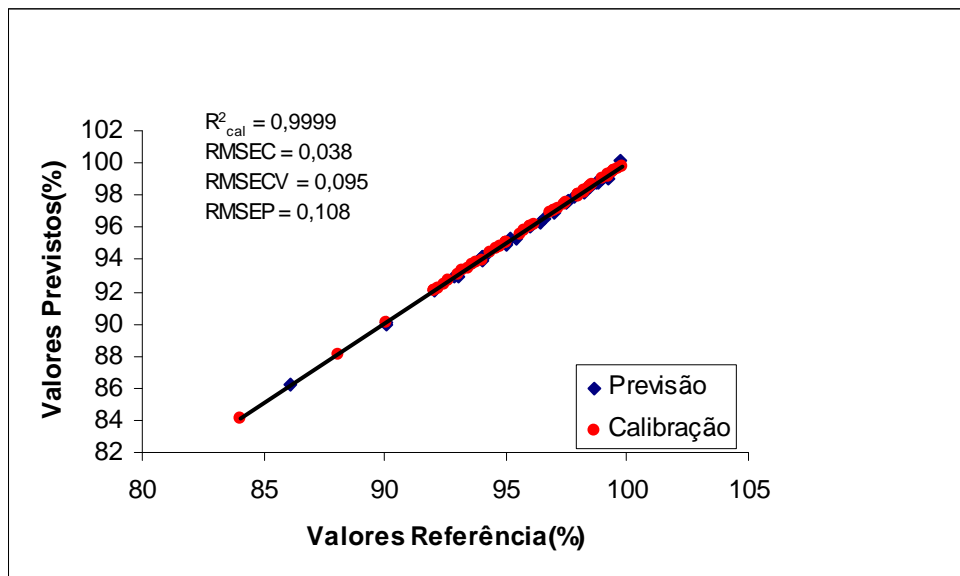


Figura 55 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

5.3.7 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DC)

E finalmente o método iPLS-PSO (DC) foi aplicado para a quantificação do óleo diesel e o modelo selecionado, dentre todos os construídos, foi o que selecionou a região espectral de 735 a 1327 cm^{-1} , visualizada na Figura 56, com 40 intervalos e 25 partículas. A sua respectiva curva de calibração para quantificação do óleo diesel empregando 10 variáveis latentes é apresentada na Figura 57.

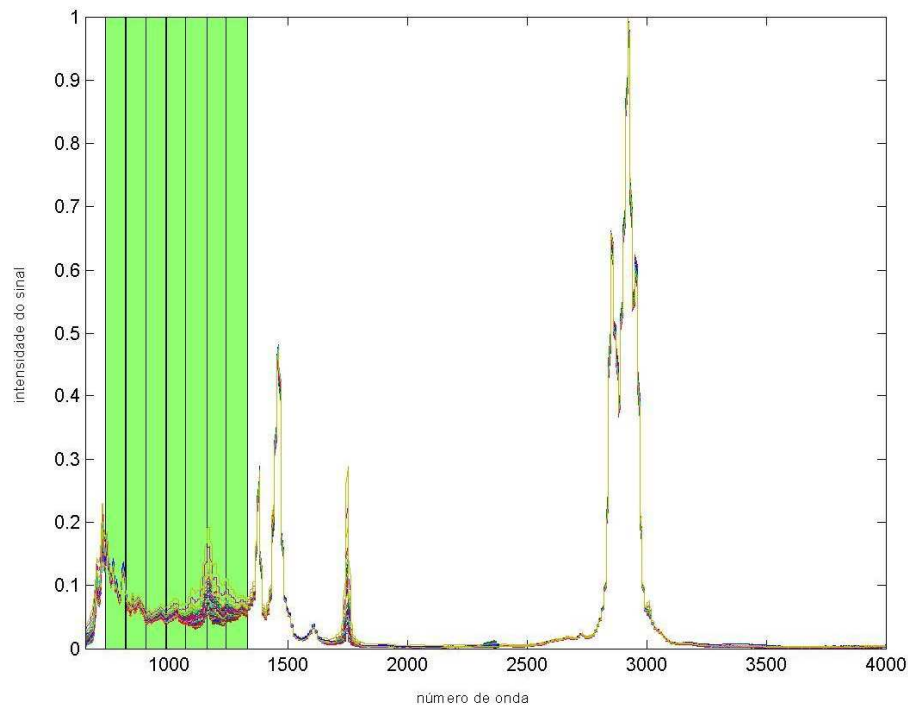


Figura 56 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

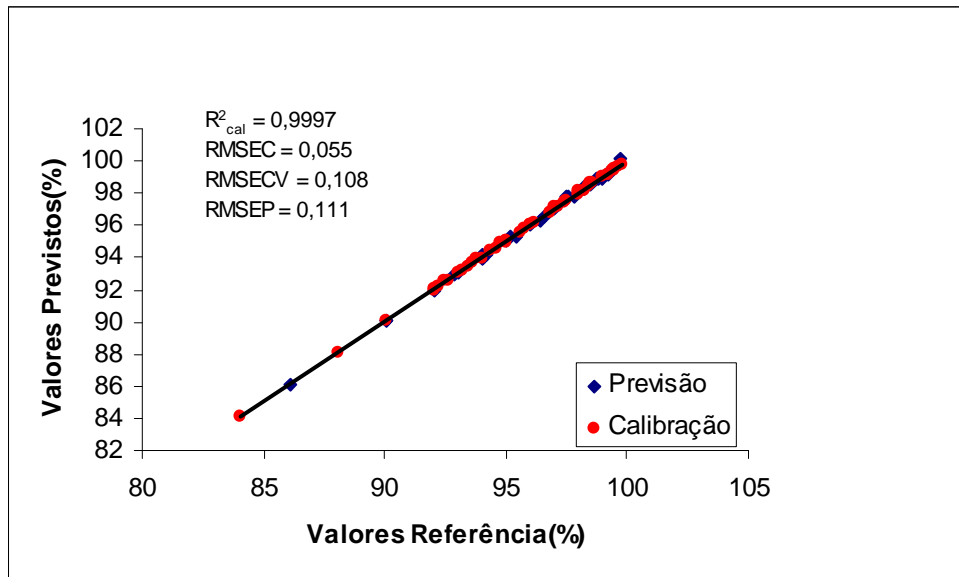


Figura 57 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 50 partículas. Fonte: elaborado pela autora.

5.3.8 Considerações sobre os modelos de óleo diesel

Com base nos resultados mostrados acima, foi possível observar que com baixos erros de previsão (RMSEP) todos os modelos construídos foram capazes de quantificar o óleo diesel. Analisando os resultados obtidos, empregando o PLS sobre o espectro inteiro da mistura e os resultados encontrados empregando as soluções propostas pelo método enxame de partículas, é possível perceber uma melhora na capacidade preditiva dos modelos otimização pela meta-heurística, porém não tão representativa como no caso das determinações de biodiesel e óleo vegetal.

Os modelos selecionados pelas otimizações aqui propostas para a quantificação do óleo diesel, apresentaram valores de RMSEP entre 0,102 a 0,112, enquanto os valores RMSECV ficaram na amplitude de 0,095 a 0,108.

No que diz respeito a comparação dos modelos construídos com as regiões selecionadas pela algoritmo híbrido e com toda região espectral da mistura não houve grandes melhoras, ficando evidenciado que a quantidade de informação capturada pelas variáveis latentes no modelo que emprega toda a região espectral é adequado para a quantificação deste componente da mistura.

Já para os modelos utilizando o algoritmo PSO nenhuma região característica dos hidrocarbonetos (região de 2850-3000 cm^{-1} , que corresponde a absorção de $\nu\text{C-H}$ referente aos estiramentos simétricos e assimétricos das ligações C-H de alcanos; 1465 cm^{-1} que corresponde a absorção de δCH_2 ; e 1375-1450 cm^{-1} , que corresponde a absorção de δCH_3) que constituem o diesel foi selecionada. Observando as regiões selecionadas na construção dos modelos a maior parte dessas correspondem à região do infravermelho médio denominada de impressão digital.

5.4 Considerações finais sobre os modelos de biodiesel, óleo vegetal e óleo diesel

Os resultados obtidos para os modelos visando a determinação, principalmente do biodiesel e da possível presença de óleo vegetal em misturas biodiesel/óleo vegetal/diesel, indicaram que não é possível utilizar uma única região e/ou a combinações das regiões sugeridas pela proposta de norma que garanta a determinação destes dentro do erro especificado pela proposta de norma. Conforme, destacado anteriormente a banda da carbonila (C=O) é insuficiente para a construção de um modelo eficiente e sensível as possíveis adulterações.

Comparando os espectros do biodiesel de soja e do óleo de soja que deu origem a este biodiesel (Figura 58) fica claro que a banda da carbonila é muito semelhante (sinal em 1742 cm^{-1} para o biodiesel e em 1743 cm^{-1} para o óleo) não sendo suficiente para gerar informação de discriminação, mesmo quando aplicada regressão multivariada.

Por outro lado, ratificando os resultados encontrados pela otimização PSO, a região referente ao C-O apresenta modificações bem mais representativas. Se por um lado o principal sinal no espectro do biodiesel encontra-se em 1170 cm^{-1} , por outro lado este mesmo sinal desloca-se para 1160 cm^{-1} para o óleo vegetal. O mesmo é verificado para a faixa de compreende o intervalo entre 1100-1450 cm^{-1} , região para a qual diversos modelos otimizados apresentam soluções tanto para o biodiesel, quanto para diesel com baixos erros dentro da faixa preconizada pela proposta de norma.

Uma outra forma de justificar o maior erro encontrado na previsão do biodiesel através das faixas espectrais sugeridas pela proposta de norma, é pela observação das figuras 59 e 60, onde as barras representam o erro de previsão do biodiesel nas amostras do conjunto de previsão, enquanto que as retas representam a concentração de óleo presentes nesta amostras.

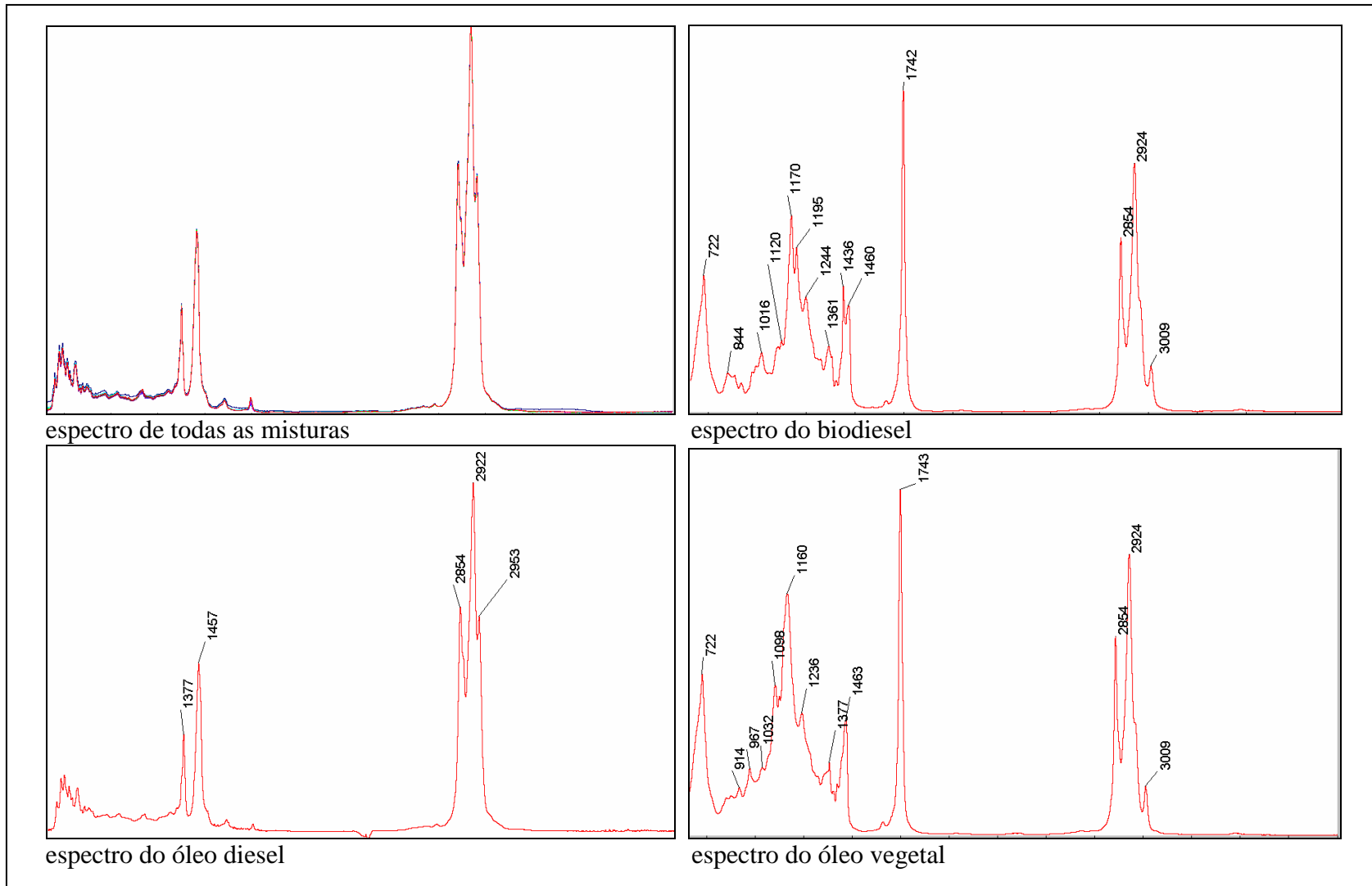


Figura 58 – Esquema contendo espectro de todas as misturas, do óleo diesel, do biodiesel e do óleo vegetal.
 Fonte: elaborado pela autora.

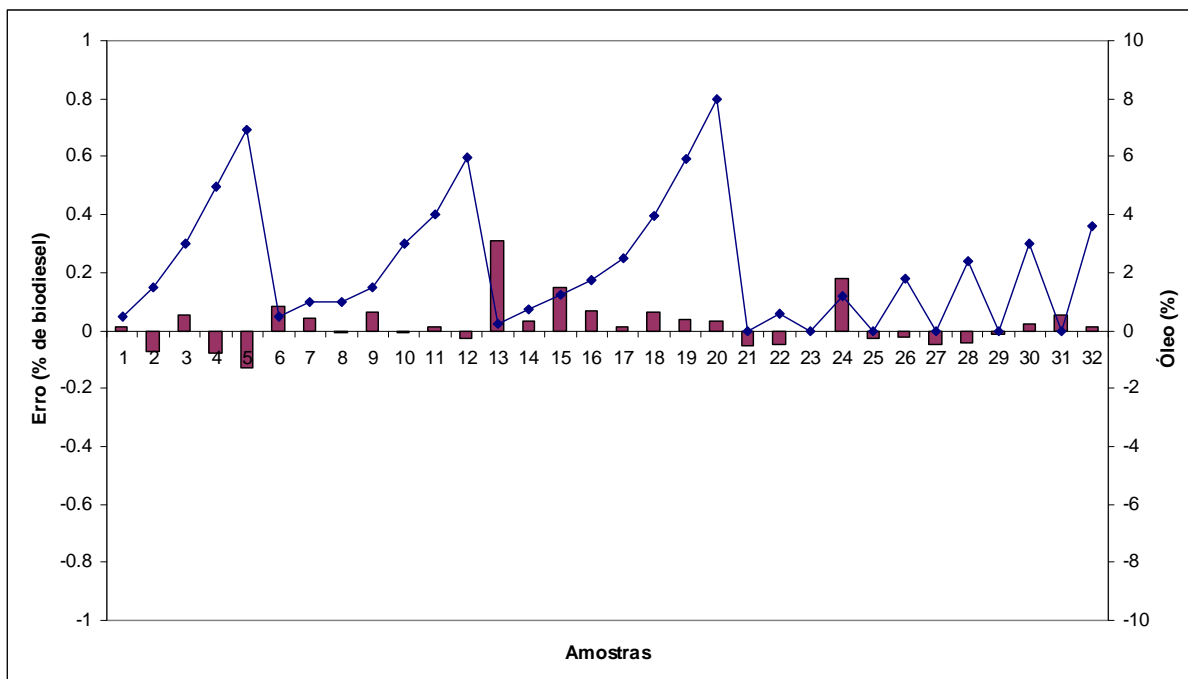


Figura 59 – Erros de previsão para a quantificação do biodiesel, usando as regiões propostas pelo método iPLS-PSO com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

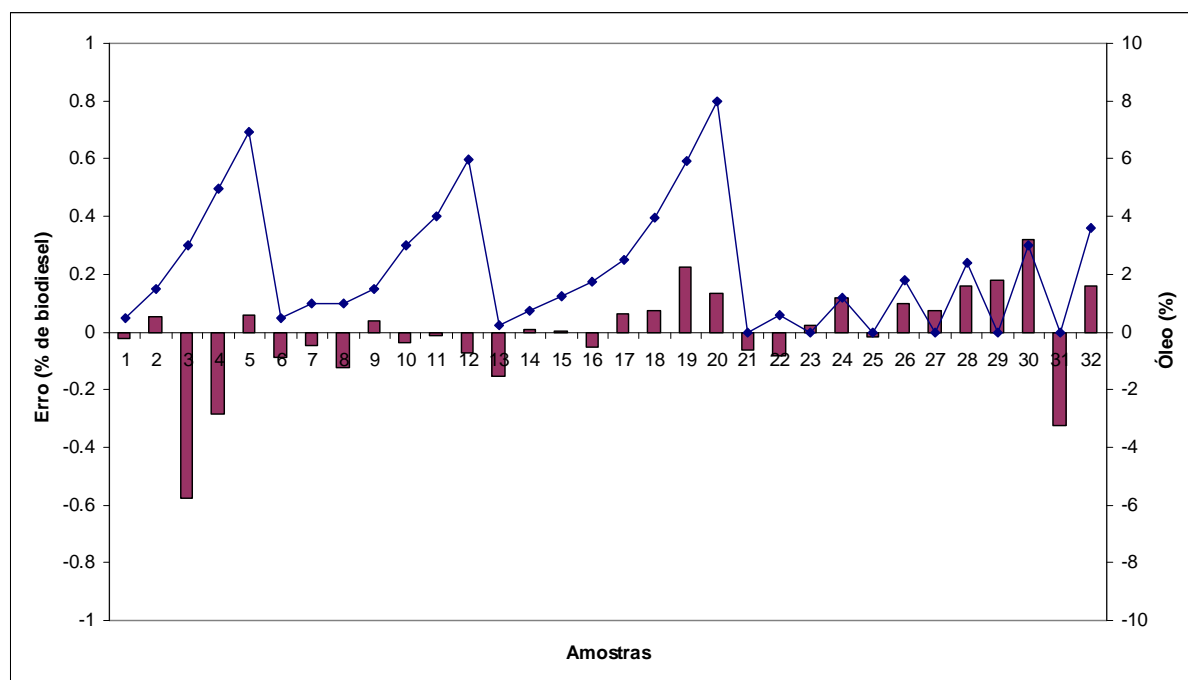


Figura 60 – Erros de previsão para a quantificação do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT.
Fonte: elaborado pela autora.

O modelo empregando somente as regiões sugeridas pela proposta de norma ABNT, de uma forma geral, apresenta maior erro para as amostras que apresentam maior concentração de óleo, sugerindo que contribuição do sinal referente ao óleo esteja afetando a capacidade preditiva deste modelo. Já para o modelo otimizado isto não se verifica.

Em trabalhos empregando regressão multivariada, a exemplo do trabalho de Soares, Rezende e Fortes (2007), foi utilizada uma metodologia para quantificação de misturas biodiesel/diesel empregando o método de seleção de variáveis stepwise e forward apresentando erros de previsão na faixa de 0,65 a 1,40 %, entretanto neste caso não foi possível aplicar sempre o mesmo método de seleção de variáveis para todas as misturas de biodiesel e óleo de soja cru, onde cada grupo de amostra foi estudado separadamente para desenvolver um modelo de calibração adequado. No método proposto nesta dissertação, o RMSEP do melhor modelo de biodiesel, foi de 0,83 % sendo capaz de modelar diferentes tipos de biodiesel (soja e palma) e ainda na presença do óleo de soja, constituindo-se num modelo mais robusto.

Ainda existem na literatura trabalhos empregando análise univariada, Guarieiro *et al.* (2005, 2008), onde não se considera a existência de óleo de soja estando sujeitos, quando da análise de amostras adulteradas, produzir resultados falsos positivos, não sendo estes modelos adequados na análise a campo.

Embora os resultados apresentados nas Tabelas 5, 6 e 7 não indiquem sobre ajuste nos modelos selecionados, em alguns casos o processo de otimização não convergiu de forma satisfatória gerando modelos com RMSECV menor que o RMSEP. A Figura 61 apresentou um exemplo do comportamento desejável para a qual o *G_{best}* evolui de valores maiores RMSECV e RMSEP (pontos a direita) para valores menores de RMSECV e RMSEP (pontos a esquerda). Já na Figura 62 a evolução resulta na diminuição do RMSECV sem que necessariamente o RMSEP tenha o mesmo comportamento.

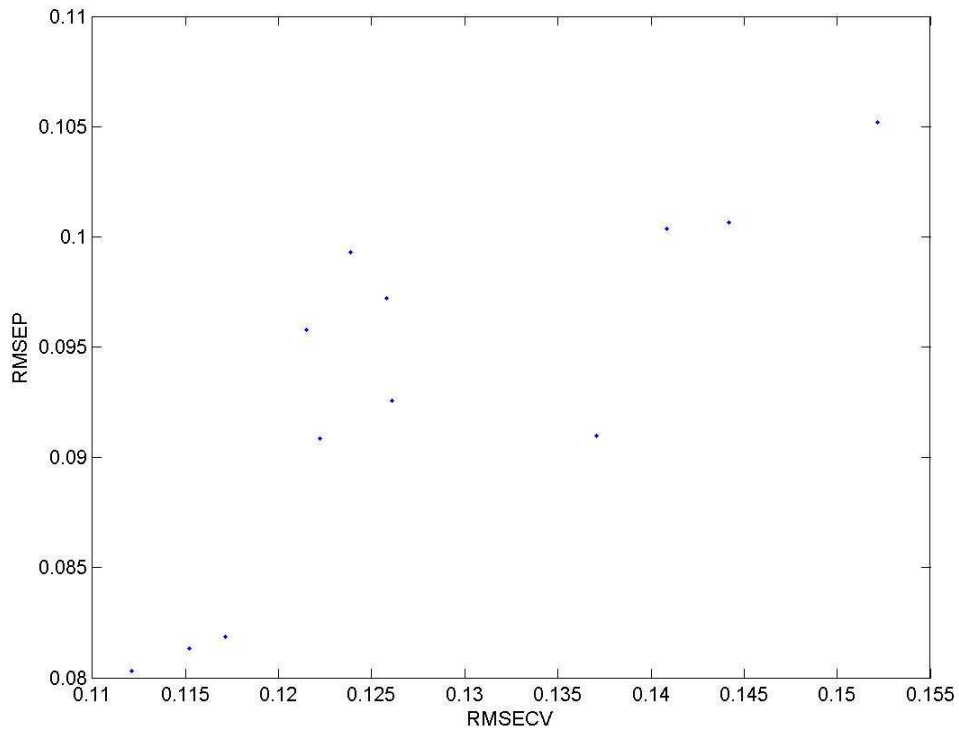


Figura 61 – Valores de RMSECV *versus* RMSEP para diferentes *Gbest* para quantificação do biodiesel, usando o método iPLS-PSO (CD) com 60 intervalos e 25 partículas.

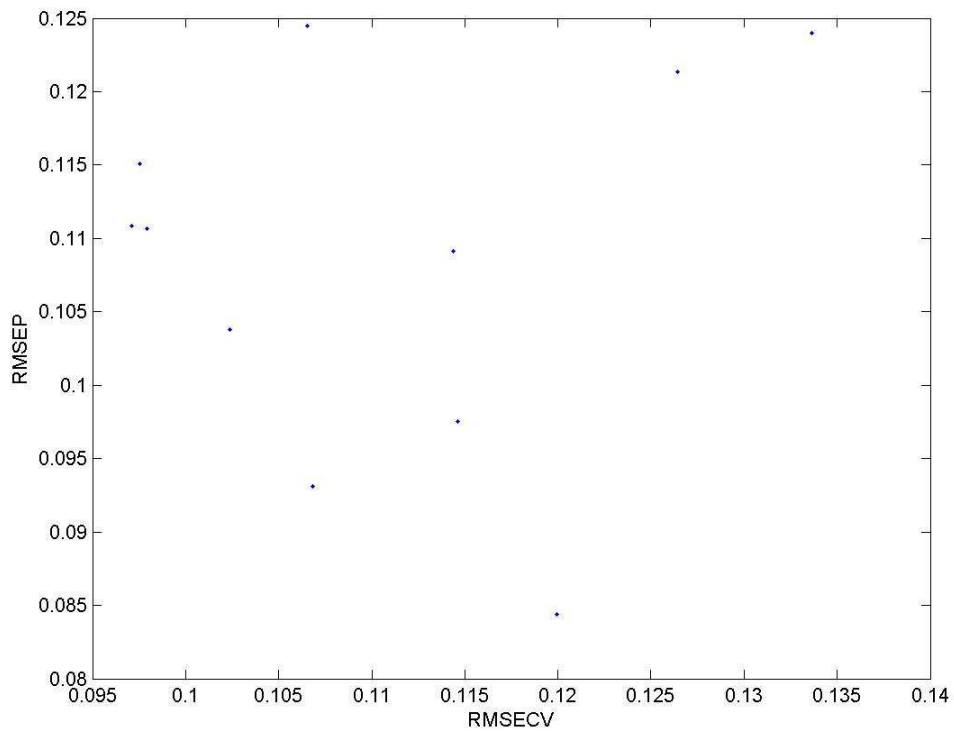


Figura 62 – Valores de RMSECV *versus* RMSEP para diferentes *Gbest* para a quantificação do biodiesel, usando o método iPLS-PSO (CD) com 40 intervalos e 100 partículas.

6 CONCLUSÕES

A presente dissertação mostrou que o método enxame de partículas em conjunto com o método de regressão por mínimos quadrados parciais por intervalos é capaz de desenvolver modelos de previsão de biodiesel, óleo vegetal e óleo diesel através dos espectros de reflexão total atenuada das misturas biodiesel/óleo diesel.

Independente do número de partículas empregados na otimização por enxame de partículas, bem como do número de intervalos para os quais os espectros foram subdivididos, bons modelos com capacidade preditiva maiores ou equivalentes ao emprego do espectro total foram obtidos para qualquer um dos componentes das misturas. Entretanto no que se refere ao modelo para a previsão de biodiesel empregando as regiões sugeridas pela proposta de norma este apresentou baixo desempenho que pode ser atribuído à interferência dos sinais do óleo que ocorrem nas mesmas faixas de absorção do biodiesel, não sendo suficiente a utilização destas faixas espectrais para a construção de modelos multivariados. O comportamento encontrado neste caso é bastante semelhante a interferência quando são construídos modelos univariados a partir de dados espectroscópicos.

As principais regiões selecionadas pelos algoritmos PSO podem ser facilmente atribuídas as principais absorções referentes a composição do biodiesel e do óleo vegetal, destacando-se a região referente ao estiramento C-O para quantificação do biodiesel e do óleo vegetal nas misturas, bem como a região referente a impressão digital para a quantificação do óleo diesel. Nestes modelos a seleção da região de absorção da carbonila (C=O) nem sempre foi necessária na obtenção de modelos com baixos erros de previsão.

Os modelos otimizados a partir dos algoritmos PSO também permitiram a identificação e quantificação da adulteração proveniente de óleos vegetais, enquanto que a proposta de norma é restrita a quantificação do biodiesel no diesel.

Finalmente, a implementação de métodos híbridos, nos quais são empregados conceitos de algoritmos clássicos em combinação com algoritmos heurísticos, permitem a resolução de problemas combinatoriais de forma adequada, somando-se as vantagens de ambas as técnicas para alcançar soluções satisfatórias em tempo real.

7 TRABALHOS FUTUROS

Abaixo se relaciona algum dos futuros trabalhos que podem ser desenvolvidos, com base nesta dissertação:

- ampliar a composição das misturas, utilizando outros tipos de biodiesel como o biodiesel de sebo e de óleo de frituras;
- ampliar a faixa de quantificação do biodiesel até 30 %(v/v), conforme o segundo modelo sugerido pela proposta de norma ABNT;
- comparar os resultados obtidos nesta dissertação com a PSO, com outros algoritmos híbridos ou não, como por exemplo algoritmo genético e busca tabu e;
- empregar algoritmos de compactação de dados, como as wavelets, tentando diminuir a quantidade de variáveis e o tempo de processamento, para construção dos modelos.

REFERÊNCIAS

- AGRAFIOTIS, D. K.; CEDEÑO, W. *Feature selection for structure-activity correlation using binary particle swarms*. Journal of Medicinal Chemistry, v. 45, n. 5, p. 1098-1107, fev. 2002.
- BARBOSA, Luiz Cláudio de Almeida. Espectroscopia no infravermelho na caracterização de compostos orgânicos. 1. ed. Viçosa: Editora Universidade Federal de Viçosa, 2007.
- BORGES NETO, Waldomiro. Parâmetros de qualidade de lubrificantes e óleo de oliva através de espectroscopia vibracional, calibração multivariada e seleção de variáveis. 2005. 130 f. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, 2005.
- BORIN, A.; POPPI, R. J. *Application of mid infrared spectroscopy and iPLS for the quantification of contaminants in lubricating oil*. Vibrational Spectroscopy, v. 37, n. 1, p. 27-32, jan. 2005.
- BRASIL. Lei n. 11.097, de 13 de janeiro de 2005. Disponível em: <http://www.planalto.gov.br/ccivil/_Ato2004-2006/2005/Lei/L11097.htm>. Acesso em: 02 fev. 2008.
- CEPEA. [s.n]. Quanto custa produzir biodiesel. 2006. Disponível em: <<http://cepea.esalq.usp.br/comunicacao/>>. Acesso em 25 ago. 2006.
- CLERC, M. *The swarm and the queen: towards a deterministic and adaptive particle swarm optimization*. In: Proc. Congress on Evolutionary Computation 1999, p. 1951-1957, 1999.
- CLERC, M.; KENNEDY, J. *The particle swarm: explosion, stability, and convergence in a multi-dimensional complex space*. IEEE Transactions on Evolutionary Computation, v. 6, n. 1 p. 58-73, fev. 2002.
- CHATTERJEE, A.; SIARRY, P. *Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization*. Computers and Operations Research, v. 33, n. 3, p. 859-871, mar. 2006.
- COSTA FILHO, P. A. da; POPPI, R. J. Aplicação de algoritmos genéticos na seleção de variáveis em espectroscopia no infravermelho médio. Determinação simultânea de glicose, maltose e frutose. Química Nova, São Paulo, v. 25, n. 1, p. 46-52, jan./fev. 2002.
- COSTA NETO, P. R. *et al.* Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em frituras. Química Nova, São Paulo, v. 23, n. 4, p. 531-537, jul./ago. 2000.
- DORADO, M. P. *et al.* *Kinetic parameters affecting the alkali-catalyzed transesterification process of used olive oil*. Energy e Fuels, v. 18, n. 5, p. 1457-1462, 2004.
- EBERHART, R.C.; HU, X. *Human tremor analysis using particle swarm optimization*. In: Proc. Congress on Evolutionary Computation 1999, p. 1927-1930, 1999.
- ENCINAR, J. M.; GONZÁLEZ, J. F.; RODRÍGUEZ-REINARES, A. *Biodiesel used frying oil. Variables affecting the yields and characteristics of the biodiesel*. Industrial & Engineering Chemistry Research, v. 44, n. 15, p. 5491-5499, 2005.
- ENGELBRECHT, Andries P. *Fundamentals of computational swarm intelligence*. England:

John Wiley & Sons, 2005.

FERRÃO, M. F. Técnicas de reflexão no infravermelho aplicadas na análise de alimentos. Tecno-Lógica., Santa Cruz do Sul, v. 5, n.1, p. 63-85, jan./jun. 2001.

FERRÃO, Marco Florês. Aplicação de técnicas espectroscópica de reflexão no infravermelho no controle de qualidade de farinha de trigo. 2000. 219 f. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, 2000.

FERREIRA, E. C. *et al.* Análise exploratória dos teores de constituintes inorgânicos em sucos e refrigerantes de uva. Eclética Química, v. 27, n. especial, p. 77-90, 2002.

FERREIRA, Marcos Antonio Gonzalez. Uso de análise multivariada para determinar a composição de gasolina a partir da espectroscopia FT-Raman e cromatografia gasosa. 2005. 109 f. Dissertação (Programa de Pós Graduação em Química Inorgânica) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2005.

GERIS, R. *et al.* Biodiesel de Soja – Reação de transesterificação para aulas práticas de química orgânica. Química Nova, São Paulo, v. 30, n. 5, p. 1369-1373, set./out. 2007.

GUARIEIRO, L. L. N. *et al.* Desenvolvimento de metodologia para quantificação das misturas biodiesel:diesel por infravermelho, 2005. Disponível em <www.portalabpg.org.br/PDPetro/3/trabalhos/IBP0250_05.pdf> . Acesso em: 10 dez. 2007.

GUARIEIRO, L. L. N. *et al.* Metodologia analítica para quantificar o teor de biodiesel na mistura biodiesel: diesel utilizando espectroscopia na região do infravermelho. Química Nova, São Paulo, v. 31, n. 2, p. 421-426, 2008.

GURDEN, S. P.; WESTERHUIS, J. A.; SMILDE, A. K. *Monitoring of batch processes using spectroscopy*. American Institute of Chemical Engineers. v. 48, n. 10, p. 2283-2297, out. 2002.

HAIR, J. F. *et al.* Análise Multivariada de Dados. 5. ed. Porto Alegre: Editora Bookman, 2005.

HAGEMAN *et al.* *Wavelength selection with tabu search*. Journal of Chemometrics. v. 17, n. 8-9, p. 427-437, 2003.

KHALIL, T. *et al.* *A Binary particle swarm optimization for optimal placement and sizing of capacitor banks in radial distribution feeders with distorted substation voltages*. In: Conference on Artificial Intelligence and Machine Learning, p. 129-135, 2006.

KENNEDY, J.; EBERHART, R. C. *A discrete binary version of the particle swarm algorithm*. In: Proc. of the 1997 Conference on Systems, Man, and Cybernetics, p. 4104-4109, 1997.

KENNEDY, J.; EBERHART, R. C. *Particle swarm optimization*. In: The 1995 IEEE International Conference on Neural Networks, p. 1942-1948, 1995.

KENNEDY, J.; SPEARS, W. M. *Matching algorithms to problems: an experimental test of the particle swarm and some genetic algorithms on the multimodal problem generator*. In: Proc. of the 1998 International Conference on Evolutionary Computation, p. 78-83, 1998.

KENNEDY, J.; EBERHART, R.C.; SHI, Y. *Swarm intelligence*. San Francisco: Morgan Kaufmann Publishers, 2001.

- KNOTHE Gerhard. *Analyzing biodiesel: standards and other methods*. Journal of the American Oil Chemists Society, v. 83, n. 10, p. 823-833, out. 2006.
- KONZEN, P. H. de A. *et al.* Otimização de métodos de controle de qualidade de fármacos usando algoritmo genético e busca tabu. Pesquisa Operacional, Rio de Janeiro, v. 23, n. 1, p. 189-207, jan./abr. 2003.
- LEARDI, R.; NORGAARD, L. *Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions*. Journal of Chemometrics, v. 18, n. 11, p. 486-497, 2004.
- MÜLLER, Viviane. Otimização de layouts industriais através do método enxame de partículas. 2007. 79 f. Dissertação (Programa de Pós-Graduação em Sistemas e Processos Industriais - Mestrado) - Universidade de Santa Cruz do Sul, Santa Cruz do Sul, 2007.
- MEDEIROS, José Antonio Carlos Canedo. Enxame de partículas como ferramenta de otimização em problemas complexos de engenharia nuclear. 2005. 108 f. Tese (Programa de Pós-Graduação de Engenharia – Doutorado) Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.
- MENDES, Rui. *Population topologies and their influence in particle swarm performance*. 2004. 189 f. Dissertation al fulfillment of the requirements for the degree of Philosophiæ Doctor at the University of Minho, Portugal, 2004.
- MENESES Anderson Alvarenga de Moura. Otimização por enxame de partículas aplicado ao problema combinatório da recarga de um reator nuclear. 2005. 100f. Dissertação (Programa de Pós-Graduação de Engenharia – Mestrado) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.
- MORGANO, M. A. *et al.* Aplicação de análise exploratória na diferenciação de vegetais. *Brazilian Journal of Food Technology*, São Paulo, v.2, n. 1/2, p. 73-79, 1999.
- NASCIMENTO, M. da G.; COSTA NETO, P. R.; MAZZUCO, L. M. Biotransformação de óleos e gorduras. *Biotechnology Ciência & Desenvolvimento*, Brasília, n.19, p. 28-31, mar./abr. 2001.
- NORGAARD, L. *et al.* *Interval partial least-square regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy*. *Applied Spectroscopy*, v. 54, n. 3, p. 413-419, 2000.
- OLIVEIRA, Flávia Carneiro da Cunha. Modelos de calibração multivariada associados à espectroscopia vibracional para análise de misturas diesel-óleos vegetais. 2006. 108 f. Dissertação (Programa de Pós-Graduação do Instituto de Química - Mestrado) - Universidade de Brasília, 2006.
- OMRAN M.G.H.; ENGELBRECHT A. P.; SALMAN, A. 2005. *Dynamic clustering using particle swarm optimization with application in unsupervised image classification*. *Transactions on Engineering, Computing And Technology*, v. 9, p. 199-204, nov. 2005.
- PARADKAR, M. M.; IRUDAYARAJ, J. *Rapid determination of caffeine content in soft drinks using FTIR-ATR spectroscopy*. *Food Chemistry*, v. 78, n. 2. p. 261-266, ago. 2002.
- PARISOTTO, G. *et al.* Análise exploratória aplicada no estudo de medicamentos contendo piroxicam. *Brazilian Journal of Pharmaceutical Sciences*, São Paulo, v. 41, n. 4. p. 499-505, out./dez. 2005.

- PARSOPOULO, K. E.; VRAHATIS, M. N. *Particle swarm optimization method in multiobjective problems*. In: Proc. of the 2002 Symposium on Applied Computing, p. 603-607, 2002.
- PIMENTEL, M. F. *et al.* *Determination of biodiesel content when blended with mineral diesel fuel using infrared spectroscopy and multivariate calibration*. Microchemical Journal, v. 82, n. 2, p. 201-206, abr. 2006.
- PERAM, T., VEERAMACHANENI, K.; MOHAN, C. K. Fitness-distance-ratio based particle swarm optimization. In: Proc. of the 2003 Swarm Intelligence Symposium, p. 174-181, 2003.
- POMEROY, Paul. *An introduction to particle swarm optimization*. 2003 Disponível em: <<http://www.adaptiveview.com/articles/ipsoprnt.html>>. Acesso em: 03 jan. 2006.
- PRADO, J. R. do; SARAMAGO, S. F. .P. Otimização por colônia de partículas. FAMAT em Revista, Uberlândia, n. 4, p. 87-103, abr. 2005.
- QI SHEN *et al.*; *Particle swarm algorithm trained neural network for qsar studies of inhibitors of platelet-derived growth factor receptor phosphorylation*. European journal of pharmaceutical sciences, v. 28, n. 5, p. 369-376, 2006.
- RADICH, Antony. *Biodiesel performance, costs, and use*. 2004. Disponível em: <<http://www.eia.doe.gov/oiaf/analysispaper/biodiesel/>>. Acesso em: 10 de mar. 2008.
- SABIN, Gustavo Post. Otimização de modelos de regressão multivariados empregando métodos de seleção de variáveis. 2007. 111 f. Dissertação (Programa de Pós-Graduação em Sistemas e Processos Industriais - Mestrado) - Universidade de Santa Cruz do Sul, Santa Cruz do Sul, 2007.
- SACORAGUE, Luiz Alexandre. Avaliação de diferentes regiões do espectro do infravermelho próximo na determinação de parâmetros de qualidade de combustíveis empregando ferramentas quimiométricas. 2004. 235 f. Tese (Departamento de Química Analítica - Doutorado) – Universidade Estadual de Campinas, Campinas, 2004.
- SCHWAAD, Marcio. Avaliação de algoritmos heurísticos de otimização em problemas de estimação de parâmetros. 2005. 156 f. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.
- SENA, M. M. de *et al.* Avaliação do uso de métodos quimiométricos em análise de solos. Química Nova, São Paulo, v. 23, n. 4, p. 547-556, jul./ago. 2000.
- SHI, Y.; EBERHART, R. C. *A modified particle swarm optimizer*. In: Proc. of the IEEE International Conference on Evolutionary Computation, p. 69-73, 1998a.
- SHI, Y.; EBERHART, R. C. *Parameter selection in particle swarm optimization*. In: Evolutionary Programming VII. Lecture Notes in Computer Science, p. 591-600, 1998b.
- SHI, Y., EBERHART, R. *Empirical Study of Particle Swarm Optimization*. In: Proc. Congress on Evolutionary Computation, p. 1945-1950, 1999.
- SHI, Y; EBERHART, R. C. *Fuzzy adaptive particle swarm optimization*. In: Proceedings of the IEEE Congress on Evolutionary Computation, p. 101-106, 2001.
- SOARES, I. P.; REZENDE, T. F.; FORTES, I. C. P. Calibração multivariada com seleção de variáveis em amostras de biodiesel adulteradas com óleo de soja cru utilizando dados de

espectroscopia de infravermelho com transformada de fourier. 2007 Disponível em www.biodiesel.gov.br/docs/congresso2007/caracterizacao/10.pdf. Acesso em: 03 de mar. 2008.

WANG, Z. *et al.* *Particle swarm optimization and neural network application for QSAR*. In: Parallel and Distributed Processing Symposium, abr. 2004.

YOSHIDA, S.; YOSHIDA H. *Non destructive analyses of unsaturated fatty acid species in dietary oils by attenuated total reflectance with Fourier transform IR spectroscopy*. Journal Biopolymers, v. 70, n. 4, p. 604-613, nov. 2003.

ZHENG, Y. *et al.* *Empirical study of particle swarm optimizer with an increasing inertia weight*. In: Proceeding of the IEEE Congress on Evolutionary Computation, v. 1, p. 221-226, 2003.

ANEXOS

ANEXO A - Resultados completos das implementações do iPLS-PSO para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	58
	2 ^a	0,120	0,091	63
	3 ^a	0,120	0,091	115
20 intervalos 50 partículas	1 ^a	0,120	0,091	46
	2 ^a	0,120	0,091	14
	3 ^a	0,120	0,091	72
20 intervalos 100 partículas	1 ^a	0,120	0,091	17
	2 ^a	0,120	0,091	56
	3 ^a	0,120	0,091	79
40 intervalos 25 partículas	1 ^a	0,103	0,134	179
	2 ^a	0,101	0,115	241
	3 ^a	0,101	0,120	239
40 intervalos 50 partículas	1 ^a	0,100	0,137	165
	2 ^a	0,115	0,083	59
	3 ^a	0,102	0,120	214
40 intervalos 100 partículas	1 ^a	0,100	0,137	142
	2 ^a	0,100	0,137	205
	3 ^a	0,101	0,107	207
60 intervalos 25 partículas	1 ^a	0,128	0,092	166
	2 ^a	0,114	0,089	238
	3 ^a	0,106	0,090	205
60 intervalos 50 partículas	1 ^a	0,106	0,099	156
	2 ^a	0,101	0,103	198
	3 ^a	0,100	0,091	231
60 intervalos 100 partículas	1 ^a	0,095	0,088	247
	2 ^a	0,107	0,121	168
	3 ^a	0,123	0,089	150

ANEXO A - Resultados completos das implementações do iPLS-PSO (CC) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	61
	2 ^a	0,120	0,091	141
	3 ^a	0,120	0,091	61
20 intervalos 50 partículas	1 ^a	0,120	0,091	94
	2 ^a	0,120	0,091	79
	3 ^a	0,120	0,091	76
20 intervalos 100 partículas	1 ^a	0,120	0,091	90
	2 ^a	0,120	0,091	70
	3 ^a	0,120	0,091	47
40 intervalos 25 partículas	1 ^a	0,103	0,903	152
	2 ^a	0,105	0,085	108
	3 ^a	0,106	0,106	117
40 intervalos 50 partículas	1 ^a	0,101	0,105	172
	2 ^a	0,103	0,093	164
	3 ^a	0,103	0,086	213
40 intervalos 100 partículas	1 ^a	0,106	0,092	159
	2 ^a	0,103	0,116	231
	3 ^a	0,113	0,075	106
60 intervalos 25 partículas	1 ^a	0,099	0,114	169
	2 ^a	0,108	0,097	158
	3 ^a	0,095	0,108	227
60 intervalos 50 partículas	1 ^a	0,100	0,114	147
	2 ^a	0,109	0,097	207
	3 ^a	0,116	0,115	164
60 intervalos 100 partículas	1 ^a	0,101	0,109	227
	2 ^a	0,092	0,106	213
	3 ^a	0,087	0,106	107

ANEXO A - Resultados completos das implementações do iPLS-PSO (DD) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	12
	2 ^a	0,120	0,091	12
	3 ^a	0,120	0,094	23
20 intervalos 50 partículas	1 ^a	0,120	0,091	9
	2 ^a	0,120	0,091	85
	3 ^a	0,120	0,091	44
20 intervalos 100 partículas	1 ^a	0,120	0,091	22
	2 ^a	0,120	0,091	56
	3 ^a	0,120	0,091	33
40 intervalos 25 partículas	1 ^a	0,103	0,101	247
	2 ^a	0,104	0,111	230
	3 ^a	0,101	0,101	231
40 intervalos 50 partículas	1 ^a	0,103	0,917	203
	2 ^a	0,103	0,109	202
	3 ^a	0,108	0,096	160
40 intervalos 100 partículas	1 ^a	0,103	0,086	171
	2 ^a	0,097	0,111	193
	3 ^a	0,105	0,085	138
60 intervalos 25 partículas	1 ^a	0,103	0,101	247
	2 ^a	0,104	0,111	230
	3 ^a	0,101	0,101	231
60 intervalos 50 partículas	1 ^a	0,103	0,917	203
	2 ^a	0,103	0,109	202
	3 ^a	0,108	0,096	160
60 intervalos 100 partículas	1 ^a	0,102	0,091	236
	2 ^a	0,105	0,095	203
	3 ^a	0,087	0,106	333

ANEXO A - Resultados completos das implementações do iPLS-PSO (CD) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	41
	2 ^a	0,120	0,091	76
	3 ^a	0,120	0,091	54
20 intervalos 50 partículas	1 ^a	0,120	0,091	33
	2 ^a	0,120	0,091	105
	3 ^a	0,120	0,091	70
20 intervalos 100 partículas	1 ^a	0,120	0,091	44
	2 ^a	0,120	0,091	33
	3 ^a	0,120	0,091	14
40 intervalos 25 partículas	1 ^a	0,105	0,085	176
	2 ^a	0,113	0,115	242
	3 ^a	0,103	0,091	201
40 intervalos 50 partículas	1 ^a	0,115	0,083	52
	2 ^a	0,115	0,083	85
	3 ^a	0,115	0,083	104
40 intervalos 100 partículas	1 ^a	0,113	0,063	63
	2 ^a	0,097	0,111	169
	3 ^a	0,115	0,083	32
60 intervalos 25 partículas	1 ^a	0,121	0,074	160
	2 ^a	0,124	0,105	163
	3 ^a	0,112	0,080	118
60 intervalos 50 partículas	1 ^a	0,110	0,101	224
	2 ^a	0,108	0,097	106
	3 ^a	0,114	0,094	127
60 intervalos 100 partículas	1 ^a	0,117	0,078	95
	2 ^a	0,095	0,097	193
	3 ^a	0,096	0,102	159

ANEXO A - Resultados completos das implementações do iPLS-PSO (DC) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	115
	2 ^a	0,120	0,091	96
	3 ^a	0,120	0,091	90
20 intervalos 50 partículas	1 ^a	0,120	0,091	77
	2 ^a	0,120	0,091	25
	3 ^a	0,120	0,091	82
20 intervalos 100 partículas	1 ^a	0,120	0,091	9
	2 ^a	0,120	0,091	106
	3 ^a	0,120	0,091	51
40 intervalos 25 partículas	1 ^a	0,112	0,101	51
	2 ^a	0,110	0,092	104
	3 ^a	0,103	0,101	62
40 intervalos 50 partículas	1 ^a	0,116	0,098	155
	2 ^a	0,105	0,103	162
	3 ^a	0,102	0,101	82
40 intervalos 100 partículas	1 ^a	0,105	0,104	59
	2 ^a	0,102	0,108	60
	3 ^a	0,103	0,093	123
60 intervalos 25 partículas	1 ^a	0,126	0,093	163
	2 ^a	0,113	0,101	53
	3 ^a	0,125	0,081	102
60 intervalos 50 partículas	1 ^a	0,112	0,096	51
	2 ^a	0,110	0,147	30
	3 ^a	0,110	0,101	169
60 intervalos 100 partículas	1 ^a	0,107	0,101	185
	2 ^a	0,103	0,116	48
	3 ^a	0,104	0,099	89

ANEXO B - Resultados completos das implementações do iPLS-PSO para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,079	0,095	114
	2 ^a	0,082	0,103	30
	3 ^a	0,082	0,103	48
20 intervalos 50 partículas	1 ^a	0,079	0,095	65
	2 ^a	0,079	0,095	127
	3 ^a	0,079	0,095	133
20 intervalos 100 partículas	1 ^a	0,079	0,095	67
	2 ^a	0,079	0,095	49
	3 ^a	0,079	0,095	80
40 intervalos 25 partículas	1 ^a	0,084	0,098	80
	2 ^a	0,083	0,092	232
	3 ^a	0,081	0,096	238
40 intervalos 50 partículas	1 ^a	0,080	0,010	178
	2 ^a	0,075	0,105	217
	3 ^a	0,078	0,103	158
40 intervalos 100 partículas	1 ^a	0,076	0,102	219
	2 ^a	0,076	0,107	158
	3 ^a	0,080	0,107	231
60 intervalos 25 partículas	1 ^a	0,075	0,116	171
	2 ^a	0,075	0,106	166
	3 ^a	0,076	0,121	240
60 intervalos 50 partículas	1 ^a	0,073	0,101	234
	2 ^a	0,077	0,102	145
	3 ^a	0,074	0,112	249
60 intervalos 100 partículas	1 ^a	0,077	0,110	238
	2 ^a	0,079	0,110	85
	3 ^a	0,076	0,101	166

ANEXO B - Resultados completos das implementações do iPLS-PSO (CC) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,079	0,095	197
	2 ^a	0,082	0,103	77
	3 ^a	0,079	0,095	165
20 intervalos 50 partículas	1 ^a	0,079	0,095	56
	2 ^a	0,079	0,095	49
	3 ^a	0,079	0,095	98
20 intervalos 100 partículas	1 ^a	0,079	0,095	38
	2 ^a	0,079	0,095	49
	3 ^a	0,079	0,095	98
40 intervalos 25 partículas	1 ^a	0,083	0,101	72
	2 ^a	0,088	0,089	167
	3 ^a	0,079	0,105	238
40 intervalos 50 partículas	1 ^a	0,078	0,109	120
	2 ^a	0,076	0,101	174
	3 ^a	0,079	0,109	126
40 intervalos 100 partículas	1 ^a	0,081	0,105	87
	2 ^a	0,075	0,105	201
	3 ^a	0,080	0,094	201
60 intervalos 25 partículas	1 ^a	0,077	0,108	215
	2 ^a	0,069	0,118	201
	3 ^a	0,080	0,106	221
60 intervalos 50 partículas	1 ^a	0,083	0,106	193
	2 ^a	0,080	0,111	231
	3 ^a	0,080	0,102	130
60 intervalos 100 partículas	1 ^a	0,085	0,108	37
	2 ^a	0,077	0,100	144
	3 ^a	0,077	0,114	213

ANEXO B - Resultados completos das implementações do iPLS-PSO (DD) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,079	0,095	205
	2 ^a	0,081	0,096	156
	3 ^a	0,089	0,109	157
20 intervalos 50 partículas	1 ^a	0,079	0,095	91
	2 ^a	0,082	0,103	21
	3 ^a	0,079	0,095	174
20 intervalos 100 partículas	1 ^a	0,079	0,095	100
	2 ^a	0,079	0,095	119
	3 ^a	0,081	0,094	234
40 intervalos 25 partículas	1 ^a	0,070	0,130	186
	2 ^a	0,080	0,101	221
	3 ^a	0,082	0,988	241
40 intervalos 50 partículas	1 ^a	0,074	0,105	241
	2 ^a	0,085	0,099	244
	3 ^a	0,078	0,110	208
40 intervalos 100 partículas	1 ^a	0,075	0,102	156
	2 ^a	0,082	0,089	161
	3 ^a	0,076	0,107	97
60 intervalos 25 partículas	1 ^a	0,080	0,107	227
	2 ^a	0,079	0,101	185
	3 ^a	0,083	0,104	216
60 intervalos 50 partículas	1 ^a	0,079	0,103	246
	2 ^a	0,076	0,103	246
	3 ^a	0,080	0,130	238
60 intervalos 100 partículas	1 ^a	0,074	0,110	249
	2 ^a	0,075	0,100	250
	3 ^a	0,071	0,101	236

ANEXO B - Resultados completos das implementações do iPLS-PSO (CD) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,081	0,096	84
	2 ^a	0,081	0,094	81
	3 ^a	0,082	0,103	105
20 intervalos 50 partículas	1 ^a	0,079	0,095	59
	2 ^a	0,079	0,095	99
	3 ^a	0,082	0,103	69
20 intervalos 100 partículas	1 ^a	0,079	0,095	101
	2 ^a	0,079	0,095	93
	3 ^a	0,079	0,095	90
40 intervalos 25 partículas	1 ^a	0,080	0,099	147
	2 ^a	0,074	0,105	147
	3 ^a	0,082	0,118	250
40 intervalos 50 partículas	1 ^a	0,082	0,102	189
	2 ^a	0,078	0,109	181
	3 ^a	0,078	0,106	222
40 intervalos 100 partículas	1 ^a	0,075	0,110	191
	2 ^a	0,075	0,102	94
	3 ^a	0,070	0,130	95
60 intervalos 25 partículas	1 ^a	0,084	0,106	243
	2 ^a	0,075	0,096	243
	3 ^a	0,070	0,115	199
60 intervalos 50 partículas	1 ^a	0,075	0,103	240
	2 ^a	0,076	0,107	231
	3 ^a	0,085	0,112	56
60 intervalos 100 partículas	1 ^a	0,074	0,105	229
	2 ^a	0,077	0,105	228
	3 ^a	0,075	0,105	229

ANEXO B - Resultados completos das implementações do iPLS-PSO (DC) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,082	0,103	209
	2 ^a	0,085	0,115	49
	3 ^a	0,082	0,103	62
20 intervalos 50 partículas	1 ^a	0,079	0,095	57
	2 ^a	0,079	0,095	24
	3 ^a	0,079	0,095	177
20 intervalos 100 partículas	1 ^a	0,079	0,095	44
	2 ^a	0,079	0,095	37
	3 ^a	0,079	0,095	44
40 intervalos 25 partículas	1 ^a	0,082	0,102	221
	2 ^a	0,087	0,109	47
	3 ^a	0,085	0,122	182
40 intervalos 50 partículas	1 ^a	0,082	0,107	32
	2 ^a	0,076	0,101	206
	3 ^a	0,079	0,100	59
40 intervalos 100 partículas	1 ^a	0,085	0,090	59
	2 ^a	0,085	0,105	209
	3 ^a	0,078	0,102	58
60 intervalos 25 partículas	1 ^a	0,082	0,101	226
	2 ^a	0,085	0,117	41
	3 ^a	0,091	0,116	32
60 intervalos 50 partículas	1 ^a	0,084	0,101	15
	2 ^a	0,082	0,102	225
	3 ^a	0,084	0,110	213
60 intervalos 100 partículas	1 ^a	0,077	0,110	177
	2 ^a	0,080	0,109	38
	3 ^a	0,085	0,113	119

ANEXO C - Resultados completos das implementações do iPLS-PSO para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,107	0,131	103
	2 ^a	0,111	0,139	136
	3 ^a	0,118	0,147	222
20 intervalos 50 partículas	1 ^a	0,105	0,125	188
	2 ^a	0,105	0,125	72
	3 ^a	0,107	0,131	62
20 intervalos 100 partículas	1 ^a	0,107	0,131	35
	2 ^a	0,105	0,125	217
	3 ^a	0,103	0,136	76
40 intervalos 25 partículas	1 ^a	0,116	0,143	85
	2 ^a	0,098	0,102	91
	3 ^a	0,096	0,128	193
40 intervalos 50 partículas	1 ^a	0,097	0,117	215
	2 ^a	0,098	0,127	195
	3 ^a	0,100	0,129	185
40 intervalos 100 partículas	1 ^a	0,097	0,110	171
	2 ^a	0,101	0,125	130
	3 ^a	0,091	0,114	127
60 intervalos 25 partículas	1 ^a	0,103	0,108	250
	2 ^a	0,100	0,131	240
	3 ^a	0,103	0,108	151
60 intervalos 50 partículas	1 ^a	0,097	0,130	152
	2 ^a	0,095	0,140	169
	3 ^a	0,089	0,136	99
60 intervalos 100 partículas	1 ^a	0,107	0,119	187
	2 ^a	0,100	0,120	232
	3 ^a	0,088	0,122	201

ANEXO C - Resultados completos das implementações do iPLS-PSO (CC) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,106	0,129	122
	2 ^a	0,112	0,115	220
	3 ^a	0,106	0,134	69
20 intervalos 50 partículas	1 ^a	0,111	0,139	129
	2 ^a	0,107	0,131	59
	3 ^a	0,112	0,140	75
20 intervalos 100 partículas	1 ^a	0,107	0,131	99
	2 ^a	0,105	0,125	61
	3 ^a	0,101	0,128	96
40 intervalos 25 partículas	1 ^a	0,100	0,112	147
	2 ^a	0,091	0,114	141
	3 ^a	0,116	0,141	99
40 intervalos 50 partículas	1 ^a	0,091	0,114	119
	2 ^a	0,099	0,132	96
	3 ^a	0,096	0,112	232
40 intervalos 100 partículas	1 ^a	0,095	0,147	178
	2 ^a	0,102	0,127	183
	3 ^a	0,097	0,124	223
60 intervalos 25 partículas	1 ^a	0,099	0,142	222
	2 ^a	0,100	0,131	224
	3 ^a	0,099	0,108	133
60 intervalos 50 partículas	1 ^a	0,099	0,127	204
	2 ^a	0,092	0,111	148
	3 ^a	0,091	0,140	227
60 intervalos 100 partículas	1 ^a	0,084	0,140	179
	2 ^a	0,089	0,124	216
	3 ^a	0,087	0,126	204

ANEXO C - Resultados completos das implementações do iPLS-PSO (DD) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,112	0,128	43
	2 ^a	0,111	0,131	3
	3 ^a	0,112	0,122	155
20 intervalos 50 partículas	1 ^a	0,112	0,122	37
	2 ^a	0,101	0,128	62
	3 ^a	0,103	0,136	211
20 intervalos 100 partículas	1 ^a	0,103	0,136	20
	2 ^a	0,103	0,136	45
	3 ^a	0,107	0,131	79
40 intervalos 25 partículas	1 ^a	0,139	0,117	173
	2 ^a	0,101	0,115	49
	3 ^a	0,101	0,114	216
40 intervalos 50 partículas	1 ^a	0,091	0,114	120
	2 ^a	0,102	0,113	81
	3 ^a	0,091	0,114	110
40 intervalos 100 partículas	1 ^a	0,091	0,114	79
	2 ^a	0,096	0,112	241
	3 ^a	0,091	0,114	41
60 intervalos 25 partículas	1 ^a	0,099	0,121	166
	2 ^a	0,109	0,128	60
	3 ^a	0,108	0,133	201
60 intervalos 50 partículas	1 ^a	0,098	0,141	208
	2 ^a	0,090	0,138	245
	3 ^a	0,090	0,118	250
60 intervalos 100 partículas	1 ^a	0,104	0,119	157
	2 ^a	0,086	0,116	131
	3 ^a	0,095	0,130	206

ANEXO C - Resultados completos das implementações do iPLS-PSO (CD) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,131	0,131	80
	2 ^a	0,147	0,147	76
	3 ^a	0,111	0,139	83
20 intervalos 50 partículas	1 ^a	0,101	0,128	108
	2 ^a	0,112	0,122	75
	3 ^a	0,101	0,128	93
20 intervalos 100 partículas	1 ^a	0,111	0,131	48
	2 ^a	0,112	0,122	65
	3 ^a	0,112	0,122	27
40 intervalos 25 partículas	1 ^a	0,091	0,114	80
	2 ^a	0,108	0,111	18
	3 ^a	0,938	0,123	165
40 intervalos 50 partículas	1 ^a	0,091	0,114	74
	2 ^a	0,095	0,108	141
	3 ^a	0,095	0,119	126
40 intervalos 100 partículas	1 ^a	0,087	0,116	175
	2 ^a	0,091	0,114	43
	3 ^a	0,091	0,114	28
60 intervalos 25 partículas	1 ^a	0,100	0,119	185
	2 ^a	0,110	0,136	97
	3 ^a	0,106	0,123	49
60 intervalos 50 partículas	1 ^a	0,105	0,131	135
	2 ^a	0,108	0,134	142
	3 ^a	0,089	0,128	167
60 intervalos 100 partículas	1 ^a	0,106	0,124	82
	2 ^a	0,100	0,109	23
	3 ^a	0,084	0,112	209

ANEXO C - Resultados completos das implementações do iPLS-PSO (DC) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,108	0,131	52
	2 ^a	0,112	0,121	207
	3 ^a	0,111	0,139	159
20 intervalos 50 partículas	1 ^a	0,107	0,131	27
	2 ^a	0,111	0,139	101
	3 ^a	0,101	0,128	135
20 intervalos 100 partículas	1 ^a	0,101	0,127	93
	2 ^a	0,111	0,139	53
	3 ^a	0,105	0,125	53
40 intervalos 25 partículas	1 ^a	0,108	0,111	233
	2 ^a	0,109	0,126	66
	3 ^a	0,116	0,143	81
40 intervalos 50 partículas	1 ^a	0,110	0,116	211
	2 ^a	0,102	0,113	203
	3 ^a	0,096	0,118	51
40 intervalos 100 partículas	1 ^a	0,103	0,121	194
	2 ^a	0,105	0,116	59
	3 ^a	0,986	0,131	35
60 intervalos 25 partículas	1 ^a	0,100	0,131	46
	2 ^a	0,103	0,126	150
	3 ^a	0,109	0,128	106
60 intervalos 50 partículas	1 ^a	0,109	0,128	121
	2 ^a	0,099	0,145	57
	3 ^a	0,101	0,153	192
60 intervalos 100 partículas	1 ^a	0,100	0,121	236
	2 ^a	0,098	0,128	194
	3 ^a	0,108	0,118	231

**UNIVERSIDADE DE SANTA CRUZ DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS
E PROCESSOS INDUSTRIAIS – MESTRADO
ÁREA DE CONCENTRAÇÃO EM
CONTROLE E OTIMIZAÇÃO DE PROCESSOS**

Marlize Rizzi

**APLICAÇÃO DO MÉTODO ENXAME DE PARTÍCULAS NA OTIMIZAÇÃO DE MODELOS
DE REGRESSÃO MULTIVARIADA APLICADOS NA QUANTIFICAÇÃO DA MISTURA
BODIESEL/ÓLEO VEGETAL/DIESEL**

Santa Cruz do Sul, março de 2008.

Marlize Rizzi

**APLICAÇÃO DO MÉTODO ENXAME DE PARTÍCULAS NA OTIMIZAÇÃO DE MODELOS
DE REGRESSÃO MULTIVARIADA APLICADOS NA QUANTIFICAÇÃO DA MISTURA
BODIESEL/ÓLEO VEGETAL/DIESEL**

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

Orientador: Prof. Dr. João Carlos Furtado

Co-orientador: Prof. Dr. Marco Flôres Ferrão

Santa Cruz do Sul, março de 2008.

AGRADECIMENTOS

Inicialmente agradeço a Deus pela vida, e por estar sempre comigo.

Aos meus pais Avelino e Geni, por me darem carinho, amor, educação e força em todos os momentos de minha vida. Amo vocês.

Ao futuro papai e meu amor Éverton Camargo da Silva pela paciência, compreensão e incentivo. Amo você.

Ao meu filho Gabriel, que está por nascer, que soube colaborar para que juntos pudéssemos concluir esta etapa tão importante em nossas vidas.

Ao meu orientador, prof. Dr. João Carlos Furtado, pela amizade, pelos ensinamentos e a disponibilidade apresentada.

Ao meu co-orientado prof. Dr. Marco Flôres Ferrão pelas muitas orientações, pelo senso de direção, pela disponibilidade e pela oportunidade concedida.

A amiga Márcia Werlang pelas ótimas horas de estudos e a pela sua importante colaboração nesta dissertação.

As secretárias Janaina Ramires Haas e Camila Borges Corrêa pela eficiência e presteza sempre apresentada.

Ao mestre Gustavo Post Sabin, por ter compartilhado comigo seus conhecimentos nesta área.

A CAPES e a FAPERGS pelo financiamento do projeto “casadinhos” e, pela colaboração do Instituto de Química e do Laboratório de Combustíveis da Universidade Federal do Rio Grande do Sul - UFRGS.

A todos aqueles que embora não tenham sido mencionados contribuíram direta ou indiretamente para a realização deste trabalho.

“A imaginação é tudo. É uma prévia das próximas atrações da vida.”

Albert Einstein

RESUMO

A substituição gradativa dos produtos derivados de fontes não renováveis, por produtos de origem renováveis, vem sendo mundialmente discutida, especialmente quanto aos combustíveis. No Brasil, o governo oficializou a mistura biodiesel/diesel nos percentuais obrigatórios de 2% a partir de 2008. Considerando a necessidade de monitoramento da qualidade desta mistura, esta dissertação propõe o estudo e o uso do método enxame de partículas (*Particle Swarm Optimization* – PSO), juntamente com a técnica de regressão multivariada de mínimos quadrados parciais por intervalos (iPLS), visando constituir um algoritmo híbrido capaz de selecionar variáveis mais representativas do constituinte a ser quantificado, buscando gerar modelos de regressão mais robustos. Nesta pesquisa efetuou-se a quantificação de biodiesel, do óleo vegetal e do óleo diesel, utilizando toda a região espectral das misturas, as regiões espectrais sugeridas pela proposta de norma ABNT e as regiões escolhidas pelos algoritmos híbridos. Na quantificação do biodiesel o menor erro de previsão obtido foi através do método híbrido, sendo este erro 61,54% menor que o erro encontrado pelas regiões sugeridas pela proposta de norma ABNT e, 26,72 % menor que o erro encontrado utilizando toda a região espectral das misturas. Já, na quantificação do óleo vegetal o menor erro de previsão, obtido também foi pelo método híbrido, ficando 27,64% menor do que o erro encontrado pelas regiões sugeridas pela proposta de norma ABNT e, 38,62% menor do que erro encontrado utilizando toda a região espectral das misturas. Porém na quantificação do óleo diesel os métodos apresentaram desempenhos semelhantes. Com base nestes resultados, pode-se concluir que a utilização do método enxame de partículas juntamente com o método iPLS, foi capaz de otimizar as soluções, selecionando variáveis do espectro para construção de modelos mais robustos.

Palavras-chave: enxame de partículas, regressão por mínimos quadrados parciais, biodiesel, otimização combinatória, espectroscopia no infravermelho.

ABSTRACT

The gradual substitution of products coming from non-renewable sources by products of renewable origin has been globally discussed lately, especially regarding to the fuels. In Brazil, the government has declared official the biodiesel/diesel mixture in the obligatory percentile of 2% starting from 2008. This dissertation based on the need for monitoring such mixture's quality proposes the study and the use of the particles swarm method (Particle Swarm Optimization - PSO), together with the interval partial least-squares multivariate regression technique (iPLS). It presents as its aims to constitute a hybrid algorithm capable to select more representative variables of the representative to be quantified, seeking to produce more robust regression models. In this research, it was done the biodiesel, the vegetable oil, and the diesel oil quantification, using (a) the whole spectral area of the mixtures, (b) the spectral areas suggested by the norm proposal from the Brazilian Association of technical rules (ABNT), and (c) the areas chosen by the hybrid algorithms. In the biodiesel quantification, the smallest foreseen mistake obtained was through the hybrid method, being this mistake 61.54% smaller than the mistake found by the areas suggested by the norm proposal from ABNT and, 26.72% smaller than the found mistake using the whole spectral area of the mixtures. Yet in the vegetable oil quantification the smallest foreseen mistake, also obtained through the hybrid method, was 27.64% smaller than the mistake found by the areas suggested by the norm proposal from ABNT and, 38.62% smaller than mistake found using the whole spectral area of the mixtures. However, in the diesel oil quantification the methods have presented similar performance. Based on these results, it was possible to conclude that the use of particles swarm method in association with the method iPLS, was capable to optimize the solutions, selecting the spectrum's variables for the construction of more robust models.

Key-words: particles swarm, partial least-squares regression, biodiesel, combinatorial optimization, infrared spectroscopy.

LISTA DE ILUSTRAÇÕES

Figura 1 – Esquema de transesterificação de um triacilglicerídeo.....	19
Figura 2 – Representação da propagação da radiação infravermelho através do elemento de ATR.	22
Figura 3 - Ilustração do método de enxame de partículas.....	26
Figura 4 – Fluxograma das etapas executadas na parte experimental.	34
Figura 5 – Diagrama ternário para as misturas biodiesel/óleo vegetal/diesel.	35
Figura 6 – Distribuição das amostras de calibração e previsão para o biodiesel.....	39
Figura 7 – Distribuição das amostras de calibração e previsão para o óleo diesel.....	39
Figura 8 – Distribuição das amostras de calibração e previsão para o óleo vegetal.....	40
Figura 9 – Representação gráfica de duas partículas.	41
Figura 10 – Exemplo de uma partícula e representação dos seus intervalos selecionados .	42
Figura 11 – Região espectral utilizada na modelagem para a quantificação do biosiesel. ...	46
Figura 12 – Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel utilizando toda a região espectral das misturas.....	47
Figura 13 – Regiões espectrais sugeridas pela proposta de norma ABNT.....	47
Figura 14– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT.....	48
Figura 15 – Exemplo da evolução das partículas para a quantificação do biodiesel, usando o método iPLS-PSO (CD) com 40 intervalos e 100 partículas.	50
Figura 16 – Evolução das três execuções do iPLS-PSO com 20 intervalos e 50 partículas, para quantificação do biodiesel.....	51
Figura 17 – Evolução das três execuções do iPLS-PSO com 40 intervalos e 50 partículas, para quantificação do biodiesel.....	52
Figura 18 – Evolução das três execuções do iPLS-PSO com 60 intervalos e 50 partículas, para quantificação do biodiesel.....	52
Figura 19 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel usando o método iPLS-PSO com 60 intervalos e 100 partículas.	53
Figura 20 – Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 60 intervalos e 100 partículas.	54

Figura 21 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.....	55
Figura 22 – Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.....	55
Figura 23– Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	56
Figura 24– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	57
Figura 25 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.....	58
Figura 26– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.....	58
Figura 27 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO com 40 intervalos e 50 partículas.	60
Figura 28– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 40 intervalos e 50 partículas.	60
Figura 29 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal utilizando toda a região espectral das misturas.	61
Figura 30 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT 62	
Figura 31– Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas, para a quantificação do óleo vegetal.....	64
Figura 32 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas, para a quantificação do óleo vegetal.....	65
Figura 33 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para a quantificação do óleo vegetal.....	65
Figura 34 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO com 40 intervalos e 25 partículas.....	66
Figura 35 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.....	67
Figura 36 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.....	68
Figura 37 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.	68
Figura 38 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	69

Figura 39 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.	70
Figura 40 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.....	71
Figura 41– Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.	71
Figura 42 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.....	72
Figura 43 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.	73
Figura 44 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel utilizando toda a região espectral da mistura.	74
Figura 45 – Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas para a quantificação do óleo diesel.....	76
Figura 46 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas para quantificação do óleo diesel.....	77
Figura 47 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para quantificação do óleo diesel.....	77
Figura 48 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO com 40 intervalos e 25 partículas.	78
Figura 49 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.	79
Figura 50 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.....	80
Figura 51 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.....	80
Figura 52 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	81
Figura 53 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	82
Figura 54 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.....	83
Figura 55 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.....	83
Figura 56 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DC) com 40 intervalos e 25 partículas.....	84

- Figura 57 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 50 partículas..... 85
- Figura 58 – Esquema contendo espectro de todas as misturas, do óleo diesel, do biodiesel e do óleo vegetal. 87
- Figura 59 – Erros de previsão para a quantificação do biodiesel, usando as regiões propostas pelo método iPLS-PSO com 40 intervalos e 50 partículas..... 88
- Figura 60 – Erros de previsão para a quantificação do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT..... 88
- Figura 61 – Valores de RMSECV *versus* RMSEP para diferentes *Gbest* para quantificação do biodiesel, usando o método iPLS-PSO(CD) com 60 intervalos e 25 partículas. 90
- Figura 62 – Valores de RMSECV *versus* RMSEP para diferentes *Gbest* para a quantificação do biodiesel, usando o método iPLS-PSO(CD) com 40 intervalos e 100 partículas. 90

LISTA DE TABELAS

Tabela 1- Planejamento experimental das misturas realizado na UNISC	36
Tabela 2– Planejamento experimental das misturas realizado na UFRGS	37
Tabela 3– Parâmetros instrumentais e regiões empregadas na aquisição dos espectros por FT-IR/ATR	38
Tabela 4 – Nomenclatura usada para identificar as implementações	44
Tabela 5– Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do biodiesel.	49
Tabela 6 – Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do óleo vegetal.	63
Tabela 7 – Principais resultados dos modelos, incluindo das cinco implementações da PSO para a quantificação do óleo diesel.	75

LISTA DE ABREVIATURAS

ABNT	Associação Brasileira de Normas Técnicas
AG	Algoritmo Genético
ANP	Agência Nacional de Petróleo, Gás Natural e Biocombustíveis
ATR	<i>Attenuated Total Reflectance</i>
FT-IR	<i>Fourier Transform Infrared Spectroscopy</i>
GC	<i>Gaseous Chromatograph</i>
GPC	<i>Gel Permeation Chromatography</i>
HPLC	<i>High Performance Liquid Chromatography</i>
iPLS	<i>Interval Partial Least-Squares</i>
NIR	<i>Near-infrared Reflectance</i>
PLS	<i>Partial Least-Square</i>
PSO	<i>Particle Swarm Optimization</i>
RMSEC	<i>Root Mean Square Error of Calibration</i>
RMSECV	<i>Root Mean Square Error of Cross Validation</i>
RMSEP	<i>Root Mean Square Error of Prediction</i>
RMSEV	<i>Root Mean Square Error of Validation</i>
TLC	<i>Thin Layer Chromatography</i>
VL's	Variáveis Latentes

LISTA DE SIMBOLOS

i	partícula
x_i	posição da partícula i
n	tamanho do enxame
it	iteração atual
$pbest_i$	melhor posição que a partícula i já obteve durante a busca
$Gbest$	melhor posição encontrada pelas partículas do enxame
c_1, c_2	parâmetros de confiança
v_i	velocidade atual da partícula i
V_{max}	limitador de velocidade
w	fator de inércia
w_{ini}	valor inicial para o fator de inércia
w_{fin}	valor final para o fator de inércia
k	coeficiente de constrição
R	nº total de iterações
$rand()$	função aleatória
$(pbest_i - x_i)$	distância entre a melhor posição encontrada para a partícula i e a posição atual dessa mesma partícula
$(Gbest - x_i)$	distância entre a melhor posição de todas as partículas do enxame e a posição atual da partícula i

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	Justificativa	16
1.2	Objetivos.....	16
1.2.1	Objetivo Geral.....	16
1.2.2	Objetivos Específicos.....	16
1.3	Organização do texto.....	17
2	FUNDAMENTAÇÃO TEÓRICA.....	18
2.1	Biodiesel	18
2.2	Espectroscopia no infravermelho com transformada de Fourier (FT-IR)	20
2.3	Principais técnicas de reflexão no infravermelho.....	21
2.3.1	Reflexão total atenuada	21
2.4	Análise multivariada.....	22
2.5	Regressão por mínimos quadrados parciais (PLS)	23
2.6	Regressão por mínimos quadrados parciais por intervalos (iPLS)	25
3	EXAME DE PARTÍCULAS	26
3.1	Introdução.....	26
3.2	O algoritmo básico de PSO.....	27
3.3	Estrutura do exame de partículas.....	28
3.4	Outros parâmetros do PSO.....	28
3.4.1	Tamanho da população	28
3.4.2	Velocidade Máxima.....	29
3.4.3	Fator de Inércia.....	29
3.4.4	Coeficiente de constrição.....	31
3.5	Outro tipo de representação	31
3.5.1	PSO binário	31
3.6	Algumas aplicações da técnica exame de partículas	32
4	METODOLOGIA	34
4.1	Preparo das misturas	35
4.2	Aquisição dos espectros	37
4.3	Tabulação das matrizes	38
4.4	Modelagem dos dados.....	38
4.4.1	Considerações gerais sobre as implementações da meta-heurística exame de partículas para selecionar variáveis de espectros no infravermelho	40
4.4.2	Estrutura básica do exame de partículas implementado para a otimização dos modelos de regressão	40
5	RESULTADOS	46
5.1	Resultados para a quantificação do biodiesel	46
5.1.1	Modelos para a determinação do biodiesel com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT	46
5.1.2	Resultados gerais e evoluções para a determinação do biodiesel	48
5.1.3	Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO	53
5.1.4	Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CC).....	54

5.1.5 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (DD).....	56
5.1.6 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CD).....	57
5.1.7 Resultados obtidos para a determinação do biodiesel aplicando o iPLS-PSO (DC)...	57
5.1.8 Considerações sobre os modelos para a previsão do biodiesel.....	59
5.2 Resultados para a quantificação do óleo vegetal	61
5.2.1 Modelos para determinação do óleo vegetal com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT.....	61
5.2.2 Resultados gerais e evoluções dos modelos para a previsão do óleo vegetal	62
5.2.3 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO.....	66
5.2.4 Resultados obtido para a determinação do óleo vegetal aplicando o método iPLS-PSO (CC).....	67
5.2.5 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DD)	69
5.2.6 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (CD)	70
5.2.7 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DC)	72
5.2.8 Considerações sobre os modelos para a previsão de óleo vegetal.....	73
5.3 Resultados para a quantificação do óleo diesel	74
5.3.1 Modelo para a determinação do óleo diesel com espectro inteiro.....	74
5.3.2 Resultados gerais e evoluções para a determinação do óleo diesel	75
5.3.3 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO	78
5.3.4 Resultados obtidos para a quantificação do óleo diesel aplicando o método iPLS-PSO (CC).....	79
5.3.5 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DD).....	81
5.3.6 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (CD).....	82
5.3.7 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DC).....	84
5.3.8 Considerações sobre os modelos de óleo diesel	85
5.4 Considerações finais sobre os modelos de biodiesel, óleo vegetal e óleo diesel	86
6 CONCLUSÕES.....	91
7 TRABALHOS FUTUROS.....	92
REFERÊNCIAS	93
ANEXOS.....	98

1 INTRODUÇÃO

Atualmente, a substituição parcial do óleo diesel por combustíveis renováveis está sendo efetuada mundialmente, e os óleos provenientes de fontes vegetais são as principais alternativas para se efetuar esta substituição.

No Brasil tornou-se obrigatório, desde 1º de janeiro de 2008, a adição de 2% de biodiesel (B2) ao diesel; com perspectivas deste biocombustível ser produzido, nos próximos anos, em grande escala para atender a demanda de mercado. Desta forma, como é feito atualmente com outros combustíveis, será necessário monitorar a qualidade desta mistura para garantir que o produto a ser comercializado não lese o consumidor, garanta o perfeito funcionamento dos veículos automotores e contribua realmente para conservação do meio ambiente.

Neste sentido, são crescentes os estudos que buscam desenvolver metodologias alternativas as metodologias tradicionais de quantificação de combustíveis. A técnica de espectroscopia no infravermelho é uma alternativa que está sendo estudada nesta área, devido a possibilidade de realizar análises não destrutivas, mais rápidas, sem gastos de reagentes e gerando informações com qualidade. Associadas a esta técnica são utilizadas ferramentas quimiométricas de regressão para construção de modelos.

Por outro lado, a otimização por enxame de partículas é um método de otimização recente (1995), baseado na analogia do comportamento social dos pássaros. Diversos estudos estão sendo desenvolvidos com esta técnica movidos pela sua simplicidade, robustez e eficiência.

Essa dissertação desenvolveu um método híbrido para quantificar o biodiesel, óleo vegetal e diesel nas misturas. Método este composto pelo método iPLS (*Interval Partial Least-Squares*) otimizado pela meta-heurística enxame de partículas, visando selecionar as variáveis mais representativas para construção de modelos de regressão mais robustos.

1.1 Justificativa

Tendo em vista que o processo de adição do biodiesel (2%) ao diesel está iniciando este ano em nosso país, e as perspectivas são bastante amplas, as indústrias terão que se estruturar de maneira a poder controlar a qualidade bem como evitar a adulteração deste biocombustível. Neste contexto, a espectroscopia no infravermelho associada a ferramentas quimiométricas, pode ser uma solução alternativa aos métodos tradicionais.

O conjunto de dados obtidos neste tipo de misturas biodiesel/diesel apresenta um número extremamente elevado de variáveis, bem como as combinações resultantes possíveis. Com o propósito de selecionar um conjunto de variáveis que possibilitem a construção de modelos com habilidade equivalente na calibração e na previsão, sem a necessidade de avaliar todas as soluções possíveis optou-se por empregar a técnica de otimização chamada de exame de partículas para efetuar esta seleção.

1.2 Objetivos

1.2.1 Objetivo Geral

O objetivo geral deste trabalho foi verificar a eficiência do método exame de partículas para selecionar variáveis capazes de otimizar os modelos de análise multivariada, aplicados na quantificação das misturas biodiesel/óleo vegetal/diesel.

1.2.2 Objetivos Específicos

- Estudar métodos de otimização empregados no processo de quantificação de matérias-primas e combustíveis.
- Quantificar misturas de biodiesel, óleo vegetal e diesel, através de modelos de regressão.
- Comparar os resultados obtidos pelo modelo desenvolvido com outros descritos na literatura.

1.3 Organização do texto

O próximo capítulo introduz alguns aspectos importantes para compreensão e desenvolvimento deste trabalho, abrangendo uma visão geral do biodiesel, informações sobre espectroscopia no infravermelho com transformada de Fourier (FT-IR/ATR) e sobre os métodos de regressão que serão utilizados. Já no terceiro capítulo é apresentada a meta-heurística enxame de partículas implementada para otimização dos modelos.

O quarto capítulo descreve a metodologia utilizada na construção desta dissertação iniciando com a elaboração das misturas e a obtenção dos espectros, juntamente com uma descrição detalhada da ferramenta desenvolvida iPLS-PSO e suas variações.

No quinto capítulo são apresentados os resultados obtidos na quantificação das misturas, juntamente com uma análise destes resultados. No sexto capítulo são apresentadas às conclusões deste trabalho e, finalmente no sétimo capítulo são apresentadas sugestões de futuros trabalhos.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Biodiesel

De toda a energia consumida mundialmente, a maior parte vem do petróleo, do carvão e do gás natural, fontes estas limitadas e com previsão de se esgotar no futuro; portando a busca por fontes alternativas de produção de energia é de suma importância. Neste contexto surgem inúmeros estudos, que envolvem a produção de combustíveis alternativos ao óleo diesel, proveniente de óleos vegetais, denominados biocombustíveis (NASCIMENTO, COSTA NETO, MAZZUCO, 2001). Os primeiros estudos envolvendo o uso de óleos vegetais como combustível foram realizados na época do desenvolvimento dos motores a diesel, porém neste período a questão ambiental não era considerada importante. Atualmente, quando as atenções se voltam aos problemas ambientais e a redução da emissão dos gases causadores do efeito estufa, a busca por fontes de produção de combustíveis “limpos” tem sido constante no mundo.

No Brasil, a lei nº 11.097, de 13 de janeiro de 2005, determina percentuais de mistura do biodiesel ao óleo diesel, bem como o monitoramento da inserção deste novo combustível ao mercado, pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. A adição obrigatória do percentual de 2% (B2) iniciou em janeiro deste ano; e a meta é que seja adicionado 5% (B5) a partir de 2013. Segundo esta lei, o biodiesel é um biocombustível derivado de biomassa renovável para uso em motores a combustão interna com ignição por compressão ou, conforme regulamento, para geração de outro tipo de energia, que possa substituir parcial ou totalmente combustíveis de origem fóssil (BRASIL, 2005). Podem ser obtidos através de óleos vegetais ou gorduras animais, novos ou residuais, a partir de processos tecnológicos para conversão de óleos em combustíveis.

Dentre as reações empregadas na produção do biodiesel a partir de óleos vegetais, destaca-se a reação de transesterificação, onde triacilglicerídeos reagem com álcool na presença de uma base ou ácido forte, produzindo uma mistura de ésteres de ácidos graxos e glicerol, conforme ilustrado na Figura 1 (GERIS *et al.*, 2007).

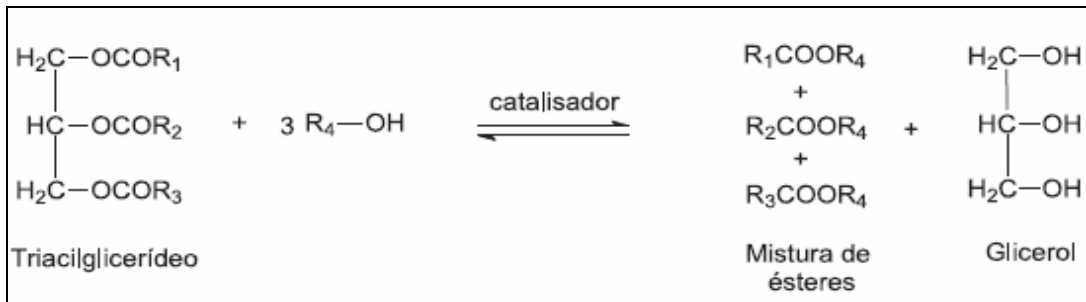


Figura 1 – Esquema de transesterificação de um triacilglicerídeo.
 Fonte: GERIS *et al.*, 2007.

O emprego do biodiesel em substituição total ou parcial ao diesel (pois é um produto miscível) proporciona benefícios ao meio ambiente, pois o CO₂ liberado é de origem de fonte renovável; sociais, pois possibilita a geração de emprego e renda em comunidades agrícolas e; econômicos tendo em vista que o diesel é um dos combustíveis mais consumidos no Brasil, pois a maior parte do transporte de cargas é realizado por meio terrestre (OLIVEIRA, 2006).

Entretanto, um dos problemas dos óleos vegetais é o preço, que tende a ser mais elevado que os de óleos de origem fósseis. Diversos pontos devem ser levados em consideração na hora de calcular o custo do biodiesel: oferta da matéria-prima, mão-de-obra, preço dos produtos agrícolas, infra-estrutura e a competição da matéria-prima com outros mercados. Um outro ponto que não pode ser esquecido são os “subprodutos” gerados durante o processo de transformação (o farelo, a glicerina e o álcool hidratado) que podem fazer com que o preço do biodiesel aumente para compensar os prejuízos de subprodutos, ou reduzido por causa de lucros com derivados do processo (CEPEA, 2006). Nos EUA, as fontes mais comuns para produção do biodiesel são do óleo de soja e o óleo de frituras reciclado, mas o custo ainda elevado de fabricação, e a disponibilidade limitada de algumas matérias primas usadas, fazem com que a aplicação comercial seja ainda limitada (RADICH, 2004).

O Brasil porém é um dos poucos países com tamanha diversidade na produção de biodiesel, a partir de plantas como soja, babaçu, amendoim, mamona, girassol, dendê e palma (GERIS *et al.*, 2007). Estudos estão sendo feitos para produzir, e aperfeiçoar a produção de biodiesel a partir do reaproveitamento de óleos como: óleos de frituras (COSTA NETO *et al.*, 2000; ENCINAR, GONZÁLEZ, RODRÍGUEZ-REINARES, 2005), óleos de oliva usados (DORADO *et al.*, 2004). Entretanto, independente da matéria prima, o produto final deve ser único e padronizado conforme os requisitos definidos pela Agência Nacional de Petróleo, Gás Natural e Biocombustíveis - ANP.

Diversas metodologias para análise do biodiesel estão sendo abordadas, tentando estabelecer um combustível padrão em diversas regiões e países ao redor do mundo. Estão inseridos neste contexto os Estados Unidos, a Europa, o Brasil, a Austrália e o Sul da África. Dentre as principais metodologias que vem sendo utilizadas para a quantificação de derivados, sub produtos e contaminações, destacam-se a cromatografia gasosa (GC), a cromatografia líquida de alta eficiência (HPLC), a cromatografia por permeação gel (GPC) e cromatografia por camada delgada (TLC) (KNOTHE, 2006). Porém estes métodos além de destrutivos, são laboriosos e consomem reagentes químicos.

Levando em consideração os antecedentes históricos e atuais de adulterações de combustíveis no Brasil, associado a necessidade de análise e controle de qualidade desse biocombustível, metodologias alternativas estão sendo propostas como a utilização da espectrofotometria no infravermelho próximo (NIR), infravermelho médio (FT-IR) e Raman (OLIVEIRA, 2006).

Outro trabalho recente desenvolvido por Pimentel *et al.* (2006) foi a utilização da espectroscopia no infravermelho e calibração multivariada na determinação de biodiesel em misturas com diesel mineral.

2.2 Espectroscopia no infravermelho com transformada de Fourier (FT-IR)

O instrumento utilizado atualmente para a obtenção de um espectro no infravermelho é denominado espectrofotômetro no infravermelho com transformada de Fourier (FT-IR). O modo de funcionamento de um equipamento FT-IR é denominado método interferométrico, pois o componente óptico básico deste equipamento é o interferômetro de Michelson (BARBOSA, 2007).

Dentre as vantagens de utilização deste método pode-se mencionar que: apresenta uma maior sensibilidade ou maior razão sinal/ruído, pois utiliza todos os comprimentos da onda ao mesmo tempo durante a obtenção do espectro; possibilita a obtenção de espectros em poucos segundos com maior resolução e maior precisão nas medidas de número de ondas (BARBOSA, 2007).

2.3 Principais técnicas de reflexão no infravermelho

Dentre os principais métodos de reflexão no infravermelho utilizados na aquisição dos dados espectrais mencionam-se: o método por reflexão difusa, o método por reflexão especular (ou externa) e o método de reflexão total atenuada (ATR), sendo este último o que será utilizado neste trabalho.

2.3.1 Reflexão total atenuada

A reflexão total atenuada (ATR – do inglês *attenuated total reflection*) é uma técnica que possibilita a obtenção de dados espectrais de amostras de maneira rápida, além de exigir uma mínima preparação da amostra. Possibilita analisar materiais líquidos, espessos e muito absorventes. Esta técnica caracteriza-se pelas múltiplas reflexões de radiação infravermelha que ocorrem no interior de cristais, interagindo somente com a amostra que estiver superficialmente no cristal (FERRÃO, 2001).

A reflexão na espectroscopia por ATR ocorre quando o feixe de radiação passa de um meio mais denso (cristal de ATR) para um menos denso (amostra), conforme Figura 2. Distorções que ocorram da banda no espectro podem ser ocasionadas pelo tipo de cristal que está sendo utilizado. Dentre os materiais usados para ATR está o ZnSe. Sua faixa de utilização é de 20.000 até 650 cm^{-1} , sendo recomendado para medidas em soluções aquosas. Outro material é o silício (Si) com uma faixa de utilização de 9.000 até 400 cm^{-1} e indicado para amostras altamente espalhativas. Um problema encontrado é a dificuldade de gerar-se boa reprodutibilidade no contato da amostra com o elemento ATR. Isto pode ser observado na variação da intensidade da banda, que é influenciado pelo aumento da pressão e pela área de contato entre o cristal e a amostra. Para medidas quantitativas, deve-se colocar toda a área do cristal em contato com a amostra (FERRÃO, 2001).

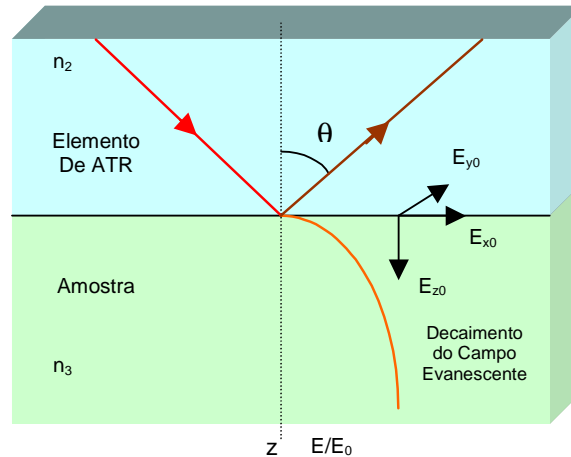


Figura 2 – Representação da propagação da radiação infravermelha através do elemento de ATR.
Fonte: FERRÃO, 2000.

Alguns exemplos de utilização da espectroscopia por reflexão total atenuada: na determinação rápida do índice de cafeína em refrigerantes (PARADKAR, IRUDAYARAJ, 2002) e, na análise de espécies de ácidos gordurosos insaturados em óleos dietéticos (YOSHIDA, YOSHIDA, 2003).

Em função das características das amostras e dos trabalhos que vem sendo realizados foi selecionada a reflexão total atenuada para a aquisição dos espectros das amostras a serem utilizadas neste trabalho.

2.4 Análise multivariada

O desenvolvimento da microeletrônica proporcionou modernos instrumentos de medição, e computadores com elevada capacidade de processamento. Esta união proporcionou um grande avanço nas análises instrumentais, possibilitando a aquisição de um grande número de dados da mesma amostra.

Analisar os dados obtidos é fundamental em todo o experimento realizado. A análise multivariada foi introduzida, ao tratamento dos dados químicos, em virtude da limitação da análise univariada, dando origem a uma área do conhecimento chamada quimiometria (SENA *et al.*, 2000). A análise multivariada refere-se a todos os métodos estatísticos que analisam simultaneamente múltiplas medidas do objeto que está sendo investigado. Pode-se dizer que qualquer análise simultânea de duas ou mais variáveis pode ser considerado

como análise multivariada (HAIR *et al.*, 2005). Nos métodos multivariados, o modelo estatístico considera a correlação entre muitas variáveis simultaneamente, possibilitando obter um número muito maior de informações (SENA *et al.*, 2000).

Quimiometria é uma área multidisciplinar do conhecimento, onde são utilizadas ferramentas matemáticas e estatísticas à análise de dados relativos aos processos químicos. Isto começou a ser amplamente utilizado para analisar amostras, graças a possibilidade de utilizar várias frequências dos espectros. A regressão é uma das mais importantes ferramentas na análise quantitativa. Os procedimentos matemáticos utilizados visam determinar as relações quantitativas existentes entre duas ou mais variáveis (FERRÃO, 2000).

Estas ferramentas de análise multivariada de dados vem sendo amplamente estudadas para aplicação na área industrial, como por exemplo, na análise exploratória na diferenciação química de vegetais (MORGANO *et al.*, 1999); na identificação de diferentes formulações de medicamentos manipulados contendo piroxicam (PARISOTTO *et al.*, 2005), na determinação simultaneamente dos teores de constituintes inorgânicos em sucos e refrigerantes de uva (FERREIRA *et al.*, 2002).

Segundo Gurden, Westerhuis e Smilde (2002), o processo industrial é carente de novas tecnologias que efetuem o monitoramento e o controle de processo em lote. A espectroscopia está se tornando uma opção de se obter em tempo real informações químicas de alta qualidade em intervalos de tempos frequentes em diversas aplicações industriais. Este autor enfatiza o correto manuseio dos dados medidos, a fim de extrair o máximo de informação, visando melhorar a performance e a eficiência do processo.

2.5 Regressão por mínimos quadrados parciais (PLS)

A técnica de regressão por mínimos quadrados parciais (PLS - do inglês *Partial Least Square*) utiliza o espectro por inteiro, ou parte deste, para construção do modelo. A matriz de dados **X**, é decomposta em várias matrizes **M** juntamente com uma matriz de resíduos **E** (KONZEN *et al.*, 2003), conforme Equação 1:

$$\mathbf{X} = \mathbf{M}_1 + \mathbf{M}_2 + \mathbf{M}_3 + \dots + \mathbf{M}_a + \mathbf{E} \quad (1)$$

onde:

a correspondem ao número de variáveis latentes ou componentes principais;

M constituem os componentes principais (ou variáveis latentes), e são compostas por dois vetores, **t** (os escores) e **p** (os pesos), conforme as Equações 2 e 3:

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

$$\mathbf{Y} = \mathbf{UQ}^t + \mathbf{F} \quad (3)$$

onde:

X é a matriz de dados (medida instrumental);

Y é a matriz de concentração;

T e **U** são os escores de **X** e **Y** respectivamente;

P e **Q** são os *loadings* de **X** e **Q**, e

E e **F** são os erros de modelagem de **X** e **Y**.

Para determinar o número de variáveis latentes (VL's) que devem ser empregadas na construção de modelo, não há um método exato, por isso os pesquisadores normalmente utilizam um aumento crescente nas VL's e verificam o comportamento da modelagem (KONZEN *et al.*, 2003).

O modelo construído é considerado eficiente, se for capaz de descrever da melhor maneira a situação real a que é proposto, levando em consideração o maior número possível de variações. Usualmente, os modelos são avaliados a partir da correlação entre a técnica proposta e o método de referência, com base nos valores de correlação e dos erros padrões de calibração (RMSEC) e validação (RMSEV), sendo os erros calculados conforme Equação 4:

$$\text{RMSEC, RMSEV, RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

onde:

y_i e \hat{y}_i corresponde aos valores de referência e estimado respectivamente para a *i*-ésima amostra e, **n** número total de amostras.

Outro erro utilizado para avaliar o modelo é o RMSECV que é obtido quando se utiliza na validação o processo de validação cruzada (*crossvalidation*), que consiste em

retirar uma amostra de cada vez do conjunto de calibração ao longo da validação. A equação para obtenção deste erro é análoga a equação 4, sendo que o i neste caso corresponde a amostra retirada do conjunto e n número total de amostras utilizadas na construção do modelo.

Na área de controle de qualidade e adulterações de combustíveis utilizando o PLS pode-se mencionar alguns trabalhos como: a análise das misturas biodiesel-óleos vegetais (OLIVEIRA, 2006); a análise qualitativa e quantitativa de combustíveis, em veículos leves, álcool e gasolina (SACORAGUE, 2004) e a determinação dos componentes da gasolina (FERREIRA, 2005).

2.6 Regressão por mínimos quadrados parciais por intervalos (iPLS)

Este método é uma extensão da regressão por PLS, onde o espectro é dividido em intervalos equidistantes, e para cada intervalo é construído um modelo PLS. O foco deste método é possibilitar que sejam selecionados regiões do espectro global que possuam informações relevantes, e um novo modelo seja construído a partir destas variáveis selecionadas (NORGAARD *et. al*, 2000).

No método iPLS, o espectro é dividido em n intervalos, e são construídos um modelo para cada intervalo de região espectral, e mais um modelo com o espectro inteiro. Os modelos são comparados utilizando o parâmetro de validação RMSECV/RMSEP (*root mean square error of cross-validation/prediction*), (LEARDI, NORGAARD, 2004). Pode-se propor o seguinte pseudocódigo iPLS:

Dividir os espectros em n intervalos equidistantes

Enquanto o número de intervalos $\leq n$ faça

Constrói o modelo PLS para o n -ésimo intervalo

FimEnquanto

Selecionar o modelo cujo intervalo apresenta o menor RMSECV

Na área de pesquisa de óleos, empregando o método iPLS, menciona-se os trabalhos de quantificação de contaminantes em óleos lubrificantes (BORIN, POPPI, 2005) e, parâmetros de qualidade de lubrificantes e óleo de oliva extra virgem e puro (BORGES NETTO, 2005).

3 EXAME DE PARTÍCULAS

3.1 Introdução

A técnica de otimização por enxame de partículas (PSO – do inglês *Particle Swarm Optimization*), foi criada por Kennedy e Eberhart (1995), a partir da análise comportamental dos pássaros a procura de alimento ou de um local para construção do ninho, conforme ilustrado na Figura 3. Na procura pelo ninho ou alimento, todo o indivíduo (partícula) pode lucrar com as experiências dos membros do grupo (enxame). Propuseram um algoritmo de otimização robusto, simples e de fácil implementação computacional.

A PSO é uma técnica de computação evolutiva, motivada pela simulação do “comportamento social”, que trabalha em uma população potencial de soluções em um espaço de busca, onde, através da cooperação e da competição entre estas soluções é possível encontrar uma solução ótima, quando aplicadas em problemas de otimização (POMEROY, 2003).



Figura 3 - Ilustração do método de enxame de partículas.

Este algoritmo faz uso da teoria dos enxames, onde os pássaros fazem uso de suas experiências e da experiência do bando para localizar o ninho ou alimento. Desta forma a área sobrevoada corresponde ao espaço definido do projeto e encontrar a comida significa encontrar o ótimo ou próximo do ótimo. A possibilidade de encontrarem alimento, baseados na inteligência social, aumenta significativamente, porque um indivíduo aprende com o acerto do outro (PRADO e SARAMAGO, 2005).

3.2 O algoritmo básico de PSO

O algoritmo é baseado em um espaço de busca, onde é inicialmente gerada uma população aleatória de partículas e a estas associadas uma velocidade, ou seja, cada partícula possui uma posição (x_i) e uma velocidade (v_i). A cada interação, cada partícula atualiza sua velocidade, pela Equação 6, e uma nova posição é gerada pela soma de sua posição atual e a nova velocidade, conforme a Equação 5 (ENGELBRECHT, 2005).

$$x_i^{it+1} = x_i^{it} + v_i^{it+1} \quad (5)$$

$$v_i^{it+1} = v_i^{it} + (c_1^{it} * rand()^{it} * (pbest_i^{it} - x_i^{it})) + (c_2^{it} * rand()^{it} * Gbest^{it} - x_i^{it}) \quad (6)$$

A Equação 6, que calcula a velocidade da partícula, é composta também pela função $rand()^{it}$ que gera números aleatórios entre 0 e 1, pelos parâmetros de confiança c_1^{it} e c_2^{it} que indicam quanto a partícula confia em sua experiência (c_1^{it}) e na do enxame (c_2^{it}); e pelos termos de aceleração: $(pbest_i^{it} - x_i^{it})$ que representa a distância entre a melhor posição encontrada para a partícula i e a posição atual dessa mesma partícula, e $(Gbest^{it} - x_i^{it})$, que representa a distância entre a melhor posição de todas as partículas do enxame e a posição da partícula i .

As partículas enquanto percorrem seu espaço de busca geram informações que são armazenadas em vetores, onde:

$X_i = (x_1, x_2, x_3, \dots, x_n)$ corresponde as posições das partículas;

$V_i = (v_1, v_2, v_3, \dots, v_n)$ indica as velocidades das partículas;

$Pbest_i = (pbest_1, pbest_2, pbest_3, \dots, pbest_n)$ corresponde a melhor posição obtida de cada partícula. E n indica o tamanho do enxame.

Armazena-se também, durante o processo de exploração, a melhor posição (*fitness*) já encontrada entre todas as partículas do enxame, que é representado pelo $Gbest$.

3.3 Estrutura do enxame de partículas

O algoritmo de PSO apresenta um conceito simples, de fácil implementação e de eficiência computacional. A seguir um pseudocódigo do PSO (ENGELBRECHT, 2005):

Criar uma população aleatória de partículas, denominada enxame;

Inicializar esta população com posições e velocidades aleatórias;

Repita para cada partícula i

 Calcular a função objetivo (*fitness*);

 Comparar o novo valor resultante da *fitness* da partícula i . Se o valor for melhor que o valor do *pbest* da partícula i , atualize o *pbest* com o novo valor da *fitness*.

 Comparar o novo valor resultante da *fitness* da partícula i com o valor do *Gbest*. Se o valor da *fitness* da partícula i for melhor que o valor do *Gbest*, atualize o valor do *Gbest* com este novo valor;

 Alterar a posição e a velocidade da partícula i ;

Até que a condição de parada seja satisfeita.

3.4 Outros parâmetros do PSO

A versão básica do enxame de partículas foi aperfeiçoada desde sua proposição em 1995. Abaixo são detalhados alguns dos parâmetros importantes relatados na literatura e que contribuíram no aperfeiçoamento deste método.

3.4.1 Tamanho da população

O tamanho da população é determinado em função do problema em que PSO será aplicado. Variam entre 20 e 50 partículas. Soluções ótimas geralmente são obtidas com tamanhos menores de população (MENDES, 2004).

Parsopoulos e Vrahatis (2002) trabalharam com enxames não maiores de 40 partículas. Já Shi e Eberhart (1999) testaram enxames com tamanhos de 20, 40, 60 e 80 partículas.

3.4.2 Velocidade Máxima

Inicialmente, o algoritmo enxame de partículas, não controlava a velocidade das partículas. Para tal foi introduzido um importante parâmetro V_{\max} . Valor muito grande para este parâmetro, facilita a exploração global, enquanto que valor muito pequeno favorece a exploração local. Porém se o valor de V_{\max} for demasiado pequeno, o enxame pode não explorar suficientemente além das regiões localmente boas (ENGELBRECHT, 2005).

O V_{\max} é um parâmetro ajustado pelo usuário conforme as características do problema em questão, normalmente entre [-4, 4]. Assim, se $(v_i) < -4$ então $(v_i) = -4$ e se $(v_i) > 4$ então $(v_i) = 4$ (KENNEDY, EBERHART, SHI, 2001).

3.4.3 Fator de Inércia

Shi e Eberhart (1998a, 1998b) incluíram um novo parâmetro, chamado de fator de inércia ou peso de inércia (*inertia weight*), no algoritmo PSO, representado por (w) conforme visualizado na Equação 7. O papel fundamental do fator de inércia é balancear habilidade de exploração de busca local e global do algoritmo. Observaram que valores maiores para (w) facilita a busca no espaço global, enquanto que valores menores para (w) aumentam a busca no espaço local, fazendo com que as partículas convirjam rapidamente, para o ponto ótimo ou não, sem que haja uma boa exploração do espaço de busca.

$$v_i^{it+1} = w * v_i^{it} + (c_1^{it} * \text{rand}()^{it} * (pbest_i^{it} - x_i^{it})) + (c_2^{it} * \text{rand}()^{it} * (Gbest^{it} - x_i^{it})), \quad (7)$$

Uma alternativa interessante para que ocorra um balanço entre a investigação e a exploração é iniciar a busca com um valor alto de (w) e diminuir este valor ao longo das

iterações, como proposto por Shi e Eberhart (1998a; 1998b, 1999). onde propuseram que o valor de (w) deveria diminuir linearmente entre 0,9 e 0,4, aproximadamente. Já Parsopoulos e Vrahatis (2002) utilizaram 1 para o valor inicial de (w), sofrendo uma diminuição gradual até 0.4.

Shi e Eberhart (2001) utilizaram um conjunto de regras Fuzzy para ajustar dinamicamente o valor de (w), mostrando-se ser um método promissor especialmente para problemas em ambientes dinâmicos.

O fator de inércia também pode sofrer uma diminuição linear a cada iteração através da Equação 8 (ENGELBRECHT, 2005).

$$w^{it} = (w_{ini} - w_{fin}) \frac{(R - it)}{R} + w_{fin} \quad (8)$$

onde w_{ini} e w_{fin} é o valor inicial e o valor final para o fator de inércia, respectivamente; R é o número total de iterações.

Como pode ser visto, muitos estudos empregaram w com variação decrescente no seu valor, em contrapartida Zheng. *et al.* (2003) propuseram que o valor de w fosse incrementado linearmente de 0.4 até 0.9 e obtiveram bons resultados.

Já para testar o método enxame de partículas Prado e Saramago (2005) utilizaram o valor da (w) em 0,729, sendo atualizado através da Equação 9, onde f_w é uma constante entre 0 e 1.

$$w_{new} = f_w * w_{old} \quad (9)$$

O fator de inércia também pode sofrer uma redução de maneira não linear. Peram, Veeramachaneni e Mohan (2003) aplicaram a Equação 10 para o cálculo do valor de (w):

$$w^{it+1} = \frac{(w^{it} - 0.4) * (R - it)}{R + 0.4} \quad (10)$$

Recentemente, Chatterjee e Siarry (2006) propuseram uma nova maneira de cálculo de (w) através da variação não-linear, conforme a Equação 11.

$$w^{it} = \left\{ \frac{(R-it)^{nl}}{(R)^{nl}} \right\} * (w_{ini} - w_{fin}) + w_{fin} \quad (11)$$

onde nl representa o índice não-linear.

3.4.4 Coeficiente de restrição

O coeficiente de restrição (K), calculado segundo Equação 12, e incorporado na equação de cálculo da velocidade, conforme Equação 13, visando limitar a velocidade das partículas, e garantir a convergência do PSO (CLERC, 1999; CLERC, KENNEDY, 2002).

$$v_i^{it+1} = K^{it} * [v_i^{it} + (c_1^{it} * rand()^{it} * (pbest_i^{it} - x_i^{it})) + (c_2^{it} * rand()^{it} * Gbest^{it} - x_i^{it})] \quad (12)$$

$$K^{it} = \frac{2}{2 - \varphi - \sqrt{\varphi^2 - 4\varphi}} \quad (13)$$

Sendo que $\varphi = c_1 + c_2$, $\varphi > 4$

3.5 Outro tipo de representação

Originalmente a técnica PSO foi desenvolvida para o universo de variáveis contínuas, mas muitos problemas são definidos num universo de variáveis discretas, onde o domínio é finito. O PSO foi facilmente adaptado para trabalhar com este tipo de variáveis. Este tipo de representação é chamado de PSO binário (ENGELBRECHT, 2005).

3.5.1 PSO binário

Esta representação tenta estender o algoritmo de enxame de partículas para trabalhar em espaço binários. Kennedy e Eberhart (1997) e Kennedy, Eberhart e Shi (2001)

descreveram uma alteração no algoritmo para operar sobre *bits* de caracteres, conforme Equação 14, onde a velocidade é usada como uma probabilidade limiar para determinar se x_i será zero ou um.

$$\text{sigmoid}(v_i^{it}) = \frac{1}{1 + \exp(-v_i^{it})} \quad (14)$$

Kennedy e Spears (1998) comparam o algoritmo binário com diversos tipos de algoritmos genéticos aleatórios. Para tal usaram um gerador aleatório. Os resultados mostram que o enxame de partículas binário foi o único a encontrar o ótimo global em cada uma das experiências, sem considerar as características do problema.

Agrafiotis e Cedeño (2002) usaram o PSO binário para seleção de características, em que (x_i) pode receber apenas 0 e 1 e, representa se a característica da partícula i será ou não selecionada. Neste caso as posições (x_i) são tratadas como uma probabilidade limiar para decidir entre o bit de valores 0 e 1, conforme Equação 15.

$$p_{ij}^{it} = \frac{x_{ij}^{\alpha it}}{\sum_{j=1}^{nx} x_{ij}^{\alpha it}} \quad (15)$$

Sendo que α determina a seleção, e $x_{ij} \in [0, 1]$, onde para valores de $\alpha > 1$ enfatizam indivíduos com alta adaptação, enquanto $\alpha < 1$ aumenta a chance de indivíduos com baixa adaptação serem selecionados.

Na literatura encontram-se pesquisas que relatam o uso do enxame de partículas binário como, por exemplo, na redução e colocação de banco de capacitores (KHALIL *et al.*, 2006).

3.6 Algumas aplicações da técnica enxame de partículas

Apesar de ser uma técnica bastante recente, na área de otimização, muitas pesquisas já foram feitas nas mais diversas áreas, como por exemplo: na análise do tremor humano

usando PSO e redes neurais (EBERHART e HU, 1999), otimização da recarga de combustível em um reator nuclear (MENESES, 2005); em classificações de imagens não supervisionadas (OMRAN, ENGELBRECHT, SALMAN, 2005), problemas complexos de engenharia nuclear (MEDEIROS, 2005) e na otimização de *layouts* (MÜLLER, 2007).

Schwaab (2005), testou algoritmos heurísticos de otimização na estimativa de parâmetros de modelo não lineares, e o método do enxame de partículas, em particular, apresentou resultados superiores aos demais algoritmos heurísticos utilizados neste trabalho, tanto na minimização de funções teste, como na estimativa de parâmetros de problemas reais.

A utilização de heurísticas e meta-heurística na otimização de processos químicos, tem recebido grande atenção por parte de pesquisadores. No entanto, o método de enxame de partículas, ainda é pouco explorado nesta área. Alguns exemplos são: PSO e rede neural na construção de modelos QSAR (WANG *et. al*, 2004) e, treinamentos de redes neurais para modelos QSAR de inibidores de plaquetas (QI SHEN *et al.*, 2006).

Por outro lado, o algoritmo genético (AG) que é uma ferramenta de otimização, apresenta um vasto número de trabalhos utilizando-o com uma forma de selecionar regiões espectrais para construção dos modelos, como por exemplo, na determinação simultânea de glicose, maltose e frutose (COSTA FILHO, POPPI, 2002) e, na otimização de modelos de regressão multivariados empregando métodos de seleção (SABIN, 2007). Já outros trabalhos relatam a utilização da heurística busca tabu na seleção do comprimento de onda (HAGEMAN *et. al*, 2003) e na otimização dos métodos de controle de qualidade de fármacos (KONZEN *et. al*, 2003).

4 METODOLOGIA

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

Para um melhor entendimento, a Figura 4 mostra um fluxograma das etapas que foram executadas na parte experimental, para a composição deste trabalho.

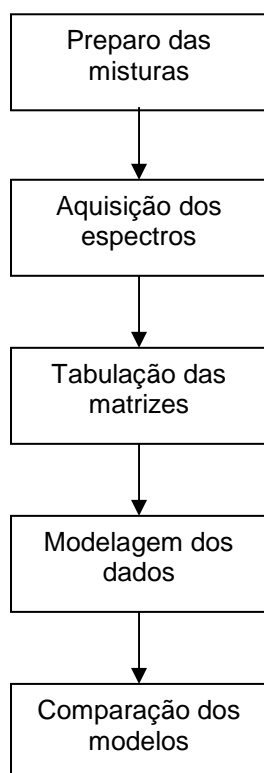


Figura 4 – Fluxograma das etapas executadas na parte experimental.
Fonte: Elaborado pela autora.

4.1 Preparo das misturas

Para a realização das misturas foram utilizados 2 planejamentos experimentais conforme tabelas 1 e 2. As 40 misturas binárias e ternárias conforme Tabela 1 foram realizadas a partir de uma amostra de óleo diesel, fornecida pelo Labcom da Universidade Federal do Rio Grande do Sul – UFRGS, de uma amostra de biodiesel de palma (doador pela Agropalma S/A), uma amostra de biodiesel de soja e uma amostra de óleo de soja (ambas fornecidas pela OleoPlan S.A). Já as 40 misturas binárias e ternárias conforme Tabela 2 foram realizadas a partir de uma segunda amostra de óleo diesel, fornecida pelo Labcom – UFRGS, da mesma amostra de biodiesel de soja e da mesma amostra de óleo de soja (ambas fornecidas pela OleoPlan S.A).

A união dos dois planejamentos resultou em um conjunto total de 80 amostras binárias e ternárias, podendo estas serem visualizadas no diagrama ternário da Figura 5. Cabe ressaltar que na apresentação do diagrama os dois diferentes óleos diesel, bem como os dois diferentes biodiesel utilizados, foram considerados como um único.

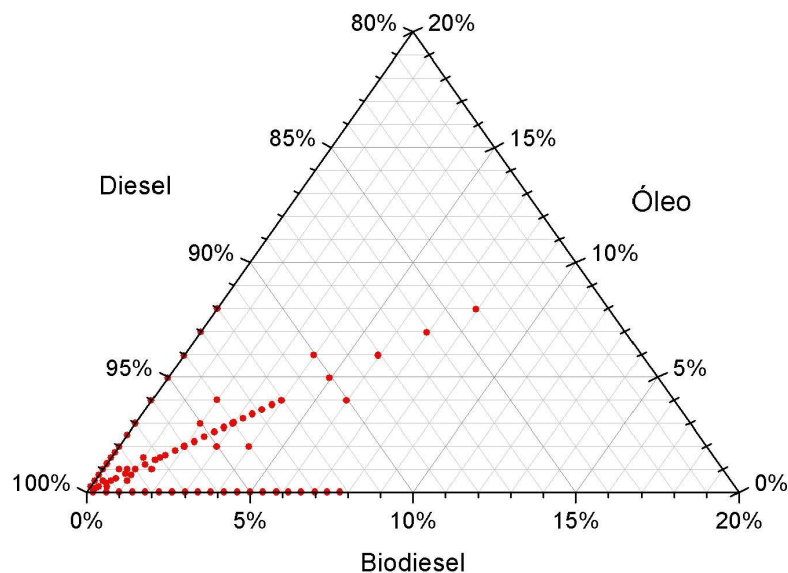


Figura 5 – Diagrama ternário para as misturas biodiesel/óleo vegetal/diesel.
Fonte: elaborado pela autora.

As amostras foram preparadas pela adição de alíquotas de biodiesel e/ou óleo vegetal conforme tabelas 1 e 2, em um balão volumétrico de 25 mL, seguida de adição do

óleo diesel até completar o volume. As soluções foram homogeneizadas e deixadas em repouso pelo tempo mínimo de 24 horas.

Tabela 1 - Planejamento experimental das misturas realizado na UNISC

Biodiesel (% v/v)	Óleo Vegetal (%v/v)	Diesel (%v/v)
0,2	0	99,8
0,2	0,2	99,6
0,6	0	99,4
0,4	0,4	99,2
1	0	99
0,6	0,6	98,8
1,4	0	98,6
0,8	0,8	98,4
1,8	0	98,2
1	1	98
2,2	0	97,8
1,2	1,2	97,6
2,6	0	97,4
1,4	1,4	97,2
3	0	97
1,6	1,6	96,8
3,4	0	96,6
1,8	1,8	96,4
3,8	0	96,2
2	2	96
4,2	0	95,8
2,2	2,2	95,6
4,6	0	95,4
2,4	2,4	95,2
5	0	95
2,6	2,6	94,8
5,4	0	94,6
2,8	2,8	94,4
5,8	0	94,2
3	3	94
6,2	0	93,8
3,2	3,2	93,6
6,6	0	93,4
3,4	3,4	93,2
7	0	93
3,6	3,6	92,8
7,4	0	92,6
3,8	3,8	92,4
7,8	0	92,2
4	4	92

Fonte: elaborado pela autora.

Tabela 2 – Planejamento experimental das misturas realizado na UFRGS

BioDiesel (%v/v)	Óleo Vegetal (%v/v)	Diesel (%v/v)
0,25	0,25	99,5
0,5	0,5	99
1	1	98
1,5	1,5	97
2	2	96
3	3	94
4	4	92
5	5	90
6	6	88
7	7	86
8	8	84
0,25	0,5	99,25
0,5	0,25	99,25
0,5	1	98,5
1	0,5	98,5
0,75	1	98,25
1	0,75	98,25
1	1,5	97,5
1,5	1	97,5
2	3	95
3	2	95
2	4	94
4	2	94
4	6	90
6	4	90
0	0,25	99,75
0	0,5	99,5
0	0,75	99,25
0	1	99
0	1,25	98,75
0	1,5	98,5
0	1,75	98,25
0	2	98
0	2,5	97,5
0	3	97
0	4	96
0	5	95
0	6	94
0	7	93
0	8	92

Fonte: elaborado pela autora.

4.2 Aquisição dos espectros

As amostras foram novamente homogeneizadas antes da aquisição dos espectros, sendo os mesmos obtidos em duplicata utilizando um acessório de reflexão total atenuada

(ATR) com cristal de ZnSe, acoplado a um espectrofotômetro Nicolet Magna 550, na faixa de 650 - 4000 cm^{-1} , conforme parâmetros descritos na Tabela 3. Antes de ser obtido o espectro de cada amostra foi realizado o espectro do branco, sem nenhum composto sobre o cristal.

Tabela 3 – Parâmetros instrumentais e regiões empregadas na aquisição dos espectros por FT-IR/ATR

Parâmetros	Condições
Região espectral	650 - 4000 cm^{-1}
Detector	DTGS
Divisor de feixes	KBr
Varreduras	32
Resolução	4 cm^{-1}
Acessório	Reflexão Total Atenuada - PIKE

Fonte: elaborado pela autora.

4.3 Tabulação das matrizes

Com os sinais espectrais obtidos, foi construída uma matriz de dados com a média das duplicatas dos espectros normalizados de cada amostra. Esta matriz foi utilizada para a construção de modelos de regressão multivariados para a quantificação do biodiesel, óleo vegetal e diesel. Foram implementados testes utilizando o espectro inteiro das misturas, as regiões espectrais propostas pela norma ABNT submetida à consulta popular em fevereiro de 2008 e finalmente com as regiões espectrais selecionadas pelas implementações propostas com o exame de partículas.

4.4 Modelagem dos dados

A seleção das amostras para a constituição dos conjuntos de calibração e previsão teve dois objetivos principais: construir conjuntos que contenham a informação da matriz de dados utilizados neste estudo e que contenha a maior variabilidade possível em relação à faixa de concentração. Os 80 espectros médios das amostras foram divididos em dois conjuntos, sendo 48 empregados na construção do conjunto de calibração e 32 no conjunto

de previsão. A distribuição das concentrações destas amostras para cada uma das substâncias participantes das misturas é apresentada nas Figuras 6, 7 e 8, onde se tomou o cuidado de garantir que as amostras com valores extremos de concentração pertençam ao conjunto de calibração.

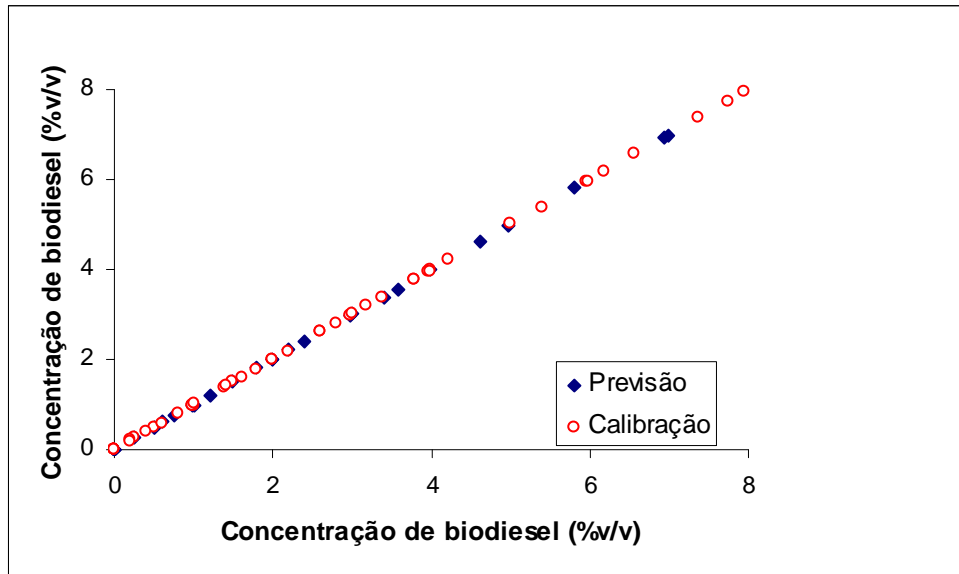


Figura 6 – Distribuição das amostras de calibração e previsão para o biodiesel.
Fonte: elaborado pela autora.

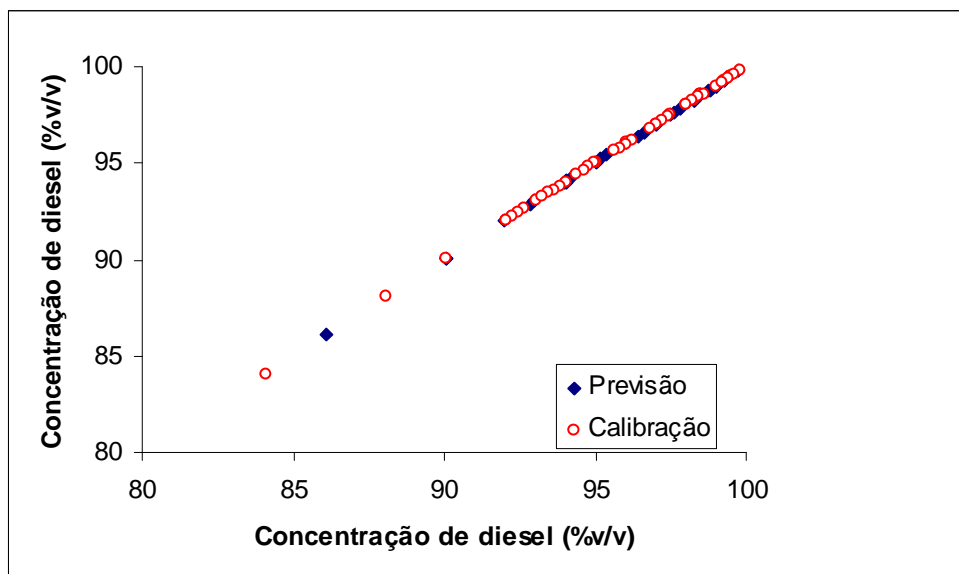


Figura 7 – Distribuição das amostras de calibração e previsão para o óleo diesel.
Fonte: elaborado pela autora.

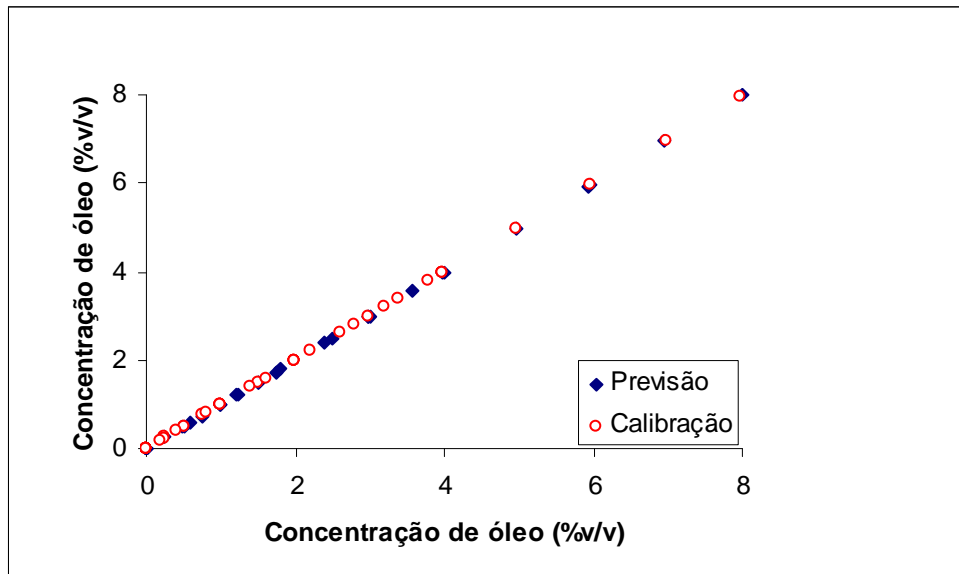


Figura 8 – Distribuição das amostras de calibração e previsão para o óleo vegetal.
Fonte: elaborado pela autora

Todos os modelos construídos neste trabalho, utilizaram o MATLAB® versão 6.5 (*The Math Works*, <http://www.mathworks.com>, EUA), juntamente com o pacote *iToolbox* versão 2 desenvolvido por Lars Norgaard (*iToolbox* for MATLAB, <http://www.models.kvl.dk>, Dinamarca).

As descrições das implementações com PSO e os resultados de todas as propostas mencionadas acima, são descritas no decorrer deste trabalho.

4.4.1 Considerações gerais sobre as implementações da meta-heurística enxame de partículas para selecionar variáveis de espectros no infravermelho

O princípio básico do algoritmo enxame de partícula é que as partículas se movimentam no espaço de busca, objetivando encontrar uma solução, orientadas pela sua experiência e pela experiência do enxame.

Como o propósito desta pesquisa é otimizar modelos de regressão multivariados, abaixo são explicados alguns conceitos do enxame de partículas neste contexto e que se mantêm para todas as implementações propostas.

Partícula: é uma possível solução para o problema. A posição da partícula é representada pelo conjunto de comprimentos de onda selecionados, ou seja, a configuração da mesma. A Figura 9 apresenta exemplo de duas partículas.

1	0	1	0	1	1	1	0	0	1
0	0	1	1	1	0	0	0	0	1

Figura 9 – Representação gráfica de duas partículas.
Fonte: elaborado pela autora.

Enxame: um conjunto de soluções (partículas).

Unidade de experiência: corresponde a um intervalo do espectro, podendo este ser ou não utilizado na construção do modelo de regressão.

Fitness: é a qualidade da posição (configuração) da partícula, obtido pelo erro de *cross-validação* (RMSECV) gerado no modelo de regressão.

Tamanho do enxame: corresponde ao número total de partículas utilizadas.

Velocidade: é a magnitude em que a solução corrente é alterada nas “direções” do $Gbest$ e do $pbest_i$.

Fator de Inércia: a intensidade em que a solução corrente é alterada (o percentual transferido do $Gbest$ e do $pbest_i$).

4.4.2 Estrutura básica do enxame de partícula implementado para a otimização dos modelos de regressão

O enxame inicia sua execução criando as partículas, onde cada partícula é composta de um número pré-definido de intervalos, codificadas com 0 e 1 aleatoriamente, e assim, cada partícula possui uma posição (solução inicial) e uma velocidade inicial. Para este problema, cada um dos intervalos da partícula corresponde a uma região do espectro, e os modelos serão construídos somente com os intervalos codificados com 1, conforme Figura 10.

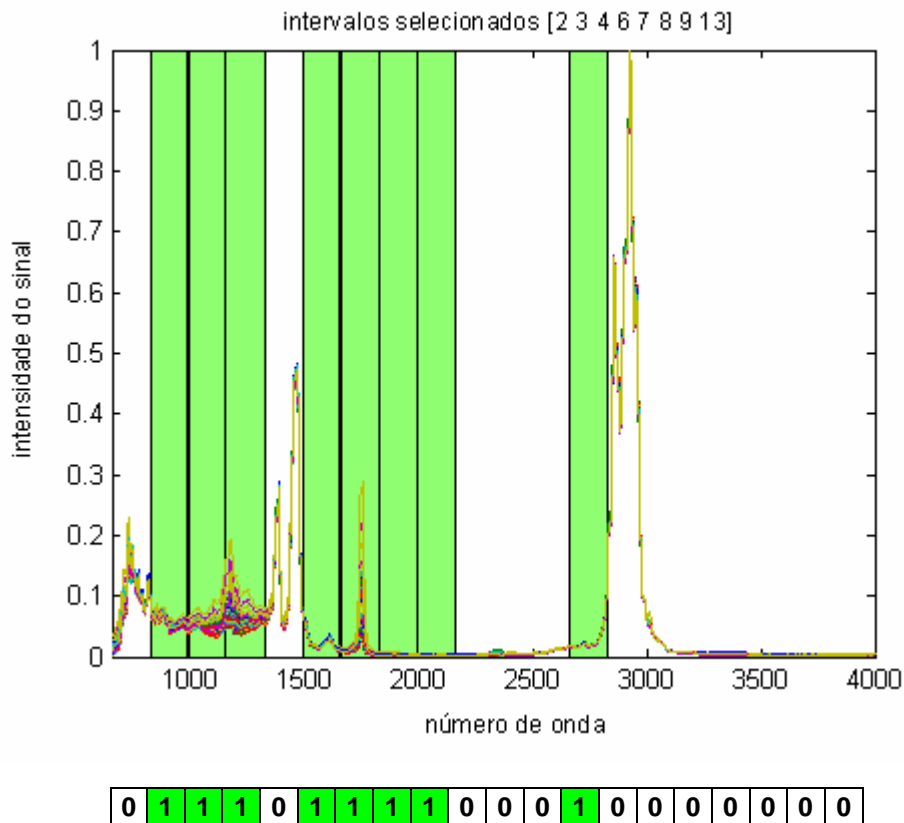


Figura 10 – Exemplo de uma partícula e a representação dos seus intervalos seleccionados. Fonte: elaborado pela autora com base no espectro do infravermelho de amostras de misturas biodiesel/óleo vegetal/diesel.

Quando o enxame é criado a posição inicial de cada partícula i é considerada como a sua melhor posição encontrada, ou seja o $pbest_i$, e a melhor posição entre todas as partículas, passa a ser considerada o $Gbest$ do enxame.

A cada iteração, cada partícula i , se desloca para uma outra posição, com base na soma da sua posição atual e nova velocidade gerada. Nestas implementações a velocidade será calculada através da Equação 7.

Assim, com base no deslocamento da partícula i é gerado um vetor somente com os intervalos seleccionados, e este é passado como parâmetro para a função $plsmodel$ do pacote $iToolBox$, juntamente com um modelo $iPLS$, com o mesmo número de intervalos da partícula i que está sendo avaliada, que gera um modelo de regressão empregando apenas estes intervalos. Este modelo gera um vetor de erros (RMSECV), onde cada valor deste vetor corresponde a utilização de um número diferente de variáveis latentes. O resultado do deslocamento da partícula ($fitness$) será o menor RMSECV, juntamente com o respectivo número de variáveis latentes.

Esta nova posição (valor resultante da *fitness*) da partícula *i* é comparada com o melhor posição que esta partícula *i* obteve até o presente momento ($pbest_i$). Da mesma maneira compara-se a *fitness* da partícula *i* com o valor do *Gbest* do enxame. Neste trabalho em que se pretende minimizar o erro RMSECV, ambos, $pbest_i$ e *Gbest* somente serão atualizados se o valor obtido pela *fitness* da partícula *i* tiver sido menor que a última melhor solução armazenada por esta partícula e pelo enxame, respectivamente.

As partículas terão a oportunidade de evoluir no espaço de busca durante 250 iterações, que é condição de parada.

Através de testes preliminares, foram definidos alguns parâmetros para serem adotados como padrão para todas as implementações, sendo eles:

- a intensidade (fator de inércia) em que a partícula *i* é alterada em direção ao $pbest_i$ (parte cognitiva) será maior do que a intensidade em direção ao *Gbest* (parte social) e que as respectivas intensidades, para ambas direções, serão compostas por valores (percentuais) máximos e mínimos permitidos;
- serão utilizados enxames com tamanho de 25, 50 e 100 partículas;
- a quantidade de intervalos que cada partícula poderá ser constituída serão 20, 40 e 60 intervalos;
- para cada combinação proposta, todas as partículas apresentarão a mesma constituição (20, 40 e 60 intervalos) e, um tamanho fixo do enxame (25, 50 e 100 partículas). Além de que todas as combinações serão executadas em triplicata.

O algoritmo de enxame de partículas implementado busca combinar várias unidades de informação (regiões do espectro), para construção dos modelos de regressão. Nesta dissertação foram implementadas cinco versões do enxame de partículas, onde os resultados obtidos destas soluções, foram comparados com os modelos obtidos utilizando o espectro inteiro das misturas e as sub-regiões indicadas pela proposta da norma ABNT.

Foi criada uma nomenclatura para identificar as versões do enxame de partículas implementado, conforme Tabela 4.

Tabela 4 – Nomenclatura usada para identificar as implementações

Sigla	Definição
iPLS	Método usado para construção do modelo.
PSO	Método de otimização utilizado.
C* ou D**	Sentido do deslocamento da intensidade da partícula i em relação ao G_{best} .
C* ou D**	Sentido do deslocamento da intensidade da partícula i em relação ao $pbest_i$.

* - crescente ** - decrescente

Fonte: elaborado pela autora.

Abaixo são citadas as implementações efetuadas utilizando a PSO.

a) Implementação do Enxame iPLS-PSO

Nesta implementação, manteve-se a estrutura básica do enxame de partículas implementado, mas a intensidade em que a partícula i é alterada (o percentual transferido do G_{best} e do $pbest_i$) foi definida e mantida constantemente durante todas as iterações. Os percentuais utilizados aqui foram 10% do G_{best} e 15 % $pbest_i$.

b) Implementação do enxame iPLS-PSO (CC)

Esta abordagem manteve-se novamente a estrutura básica do enxame de partículas implementada inicialmente, mas a intensidade em que a partícula i é alterada (os percentuais transferidos do G_{best} e do $pbest_i$) foi iniciada com valores mínimos (5%), e no decorrer das iterações sofreram um incremento linear constante até os valores máximos permitidos (15% e 30%, respectivamente).

c) Implementação do Enxame iPLS-PSO (DD)

Com base na estrutura básica do enxame de partículas descrita inicialmente, foi proposta uma nova abordagem em que a intensidade em que a partícula i é alterada (o

percentual transferido do G_{best} e do p_{best_i}) é iniciada com valores máximos (15% e 30%, respectivamente), e no decorrer das iterações sofreram um decréscimo linear constante até os valores mínimos permitidos de 5%.

c) Implementação do Enxame iPLS-PSO (CD)

Partindo da implementação básica do enxame, nesta quarta proposta, a intensidade em que a partícula i é alterada em relação a parte cognitiva da fórmula da velocidade é definida com um valor mínimo permitido (5%) e a medida que a partícula i se movimenta no espaço de busca este valor sofre um incremento linear constante até o máximo de 15%. Em relação a parte social da fórmula da velocidade, a intensidade em que a partícula i é alterada (percentual transferido do p_{best_i}) é inicialmente definida com um valor máximo permitido (30%), e no decorrer das iterações sofre um decréscimo linear constante, até atingir o mínimo permitido (5%).

d) Implementação do Enxame iPLS-PSO (DC)

Em contrapartida, esta implementação partiu da proposta básica do enxame, mas a intensidade em que a partícula i é alterada em relação a parte cognitiva da fórmula da velocidade é definida com um valor máximo permitido (15%) e, a medida que a partícula i se movimenta no espaço de busca este valor sofre um decréscimo linear constante até atingir o mínimo permitido (5%). Em relação a parte social da fórmula da velocidade, a intensidade em que a partícula i é alterada, foi iniciada com um valor mínimo (5%) permitido, e no decorrer das iterações sofre um incremento linear constante, até atingir o máximo permitido (30%).

5 RESULTADOS

5.1 Resultados para a quantificação do biodiesel

Visando quantificar o biodiesel nas misturas, foram desenvolvidos diversos modelos de regressão, e os resultados obtidos estão descritos a seguir. Primeiramente são apresentados os modelos usando o PLS e em seguida, são descritos os modelos obtidos com a otimização através da meta-heurística enxame de partículas combinada ao iPLS.

5.1.1 Modelos para a determinação do biodiesel com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT

Com base no conjunto de dados foram desenvolvidos dois modelos de regressão, ambos utilizando para construção o método PLS. O primeiro modelo, para a quantificação do biodiesel, foi construído utilizando toda região espectral (650 cm^{-1} a 4000 cm^{-1}) conforme Figura 11, e o desempenho do modelo empregando 6 variáveis latentes, é apresentado na Figura 12.

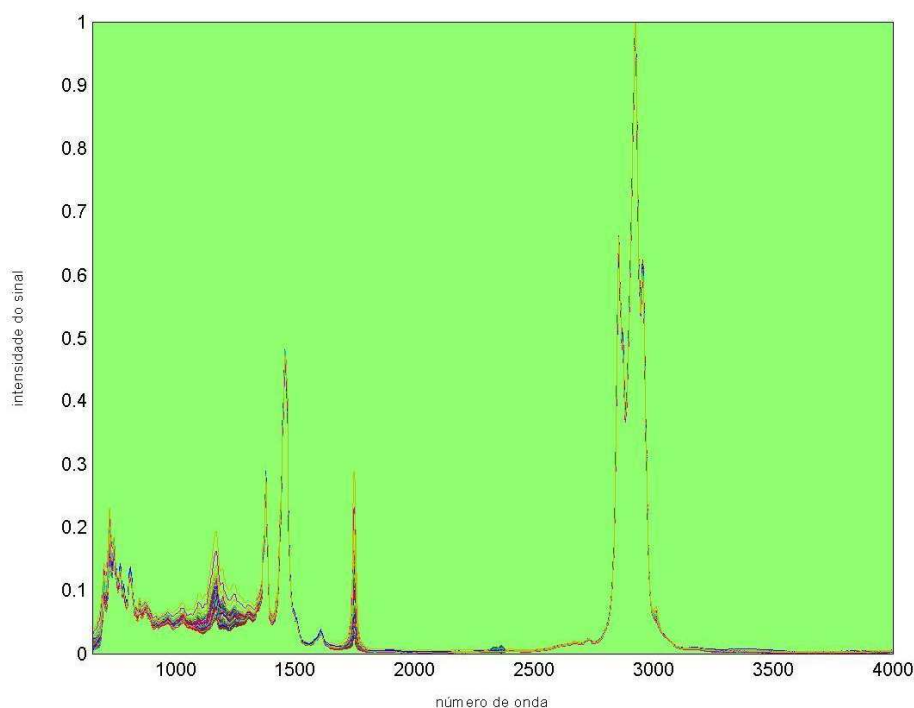


Figura 11 – Região espectral utilizada na modelagem para a quantificação do biodiesel. Fonte: elaborado pela autora.

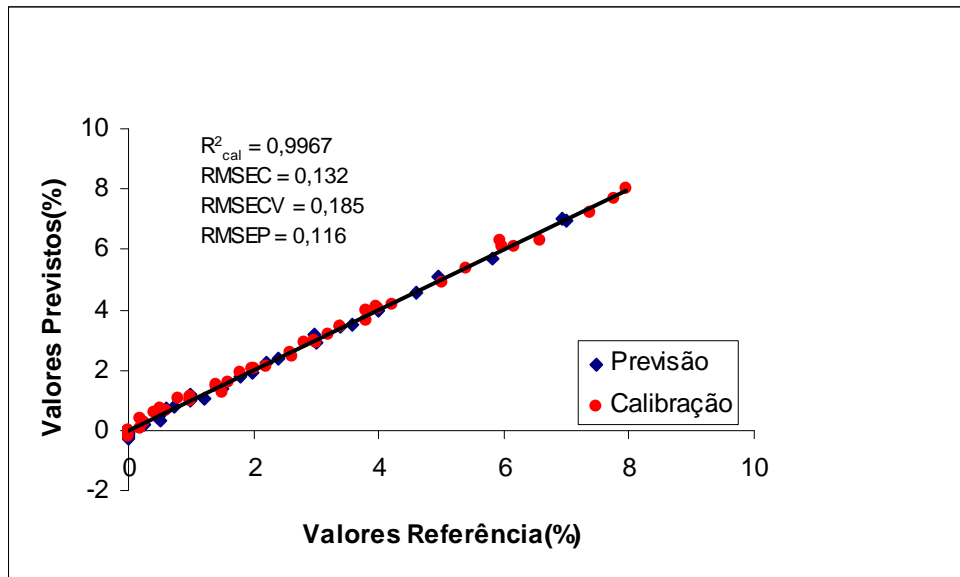


Figura 12 – Valores de referência *versus* previstos para o modelo de previsão do biodiesel utilizando toda a região espectral das misturas.
Fonte: elaborado pela autora.

O segundo modelo construído para a quantificação do biodiesel, empregou somente as duas regiões espectrais (de 1100 cm^{-1} a 1200 cm^{-1} e de 1735 cm^{-1} a 1750 cm^{-1}) sugeridas pela proposta de norma ABNT. As regiões espectrais, e o desempenho do modelo empregando 5 variáveis latentes são visualizados nas Figuras 13 e 14, respectivamente.

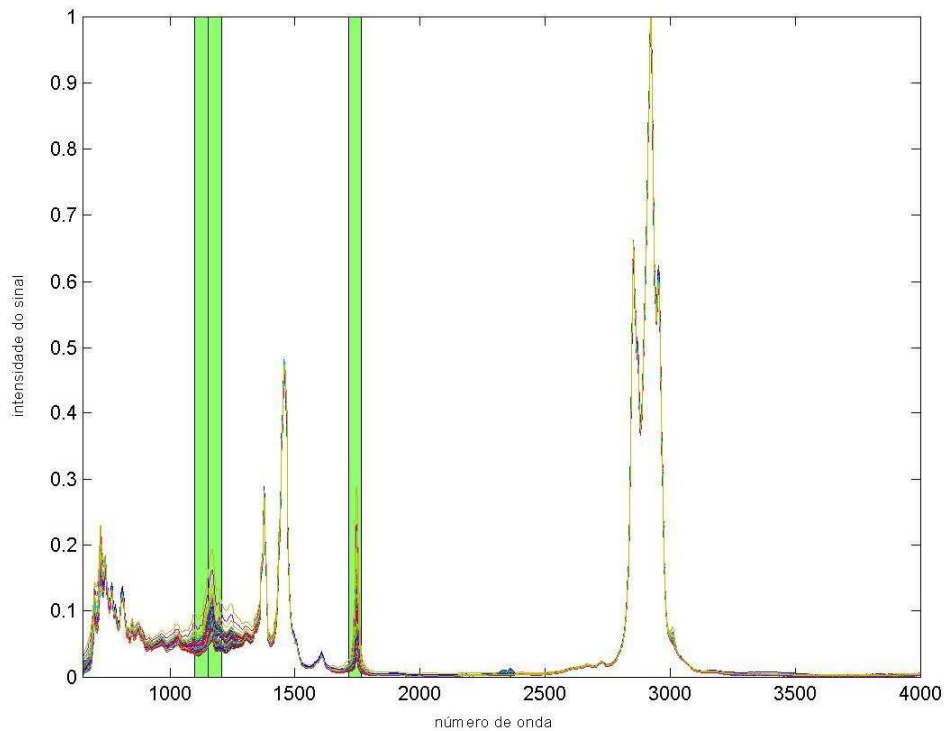


Figura 13 – Regiões espectrais sugeridas pela proposta de norma ABNT.
Fonte: elaborado pela autora.

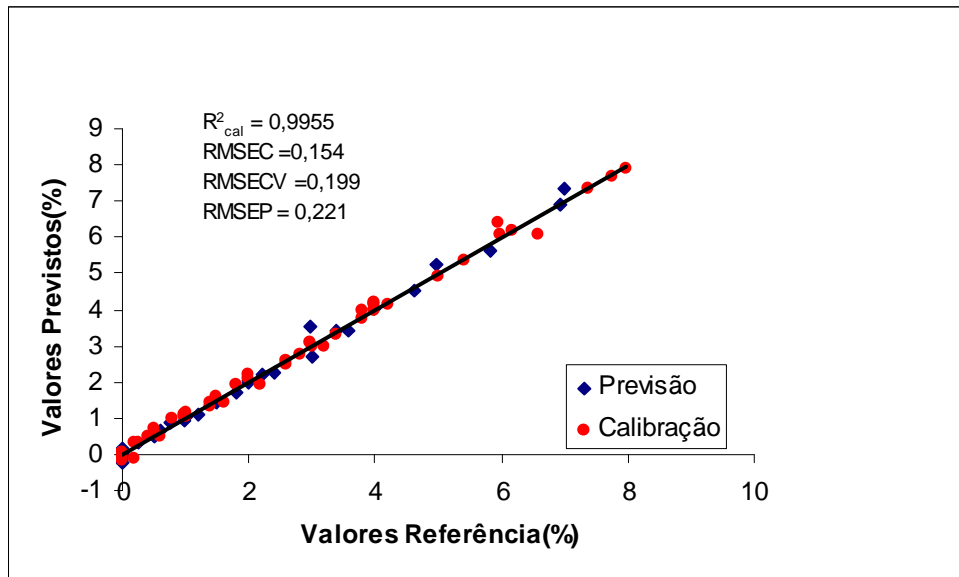


Figura 14– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT.

Fonte: elaborado pela autora.

Podemos observar que o modelo gerado utilizando o espectro inteiro, apresentou uma melhor capacidade preditiva, em relação ao modelo que levou em consideração somente às regiões sugeridas pela proposta de norma ABNT. Um aspecto que justifica este resultado pode ser a presença dos sinais interferentes do óleo nas mesmas regiões espectrais que o biodiesel, além de que outras regiões do espectro podem conter informações importantes para construção de um modelo mais robusto.

Tentando gerar modelos com maior capacidade preditiva para o analito de interesse, sugerimos a implementação de um algoritmo híbrido para a seleção das variáveis (número de onda) para construção dos modelos de regressão multivariados de espectros no infravermelho, que consiste na utilização do método iPLS, otimizado pela meta-heurística chamada enxame de partículas.

5.1.2 Resultados gerais e evoluções para a determinação do biodiesel

Com o propósito descrito anteriormente, foram implementadas cinco versões empregando a otimização por enxame de partículas: iPLS-PSO, iPLS-PSO (CC), iPLS-PSO (DD); iPLS-PSO (CD); iPLS-PSO (DC). Os principais resultados dos modelos, incluindo as implementações com a PSO, para quantificação do biodiesel são apresentados na Tabela 5.

Os resultados completos das implementações do enxame de partículas para quantificação do biodiesel encontram-se no ANEXO A.

Tabela 5– Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do biodiesel.

Modelos	Regiões empregadas	R² cal	RMSEC	RMSECV	RMSEP
espectro inteiro	650 - 4000 cm ⁻¹	0,9967	0,132	0,185	0,116
regiões sugeridas pela proposta de norma ABNT	1101 - 1198 cm ⁻¹ 1736 - 1749 cm ⁻¹	0,9955	0,154	0,199	0,221
iPLS-PSO	874 - 1319 cm ⁻¹ 2240 - 2717 cm ⁻¹ 3055 - 3109 cm ⁻¹ 3446 - 3556 cm ⁻¹	0,9997	0,038	0,095	0,088
iPLS-PSO (CC)	904 - 1327 cm ⁻¹ 2675 - 2756 cm ⁻¹ 3421 - 3502 cm ⁻¹	0,9995	0,052	0,105	0,085
iPLS-PSO (DD)	904 - 1327 cm ⁻¹ 2675 - 2756 cm ⁻¹ 3089 - 3170 cm ⁻¹ 3421 - 3502 cm ⁻¹	0,9996	0,048	0,103	0,086
iPLS-PSO (CD)	904 - 1327 cm ⁻¹ 2675 - 2756 cm ⁻¹ 3421 - 3502 cm ⁻¹	0,9995	0,052	0,105	0,085
iPLS-PSO (DC)	904 - 1327 cm ⁻¹ 2426 - 2756 cm ⁻¹ 3504 - 3585 cm ⁻¹	0,9991	0,070	0,103	0,093

Fonte: elaborado pela autora.

Na Figura 15 é apresentado um exemplo da evolução das partículas, mostrando as mesmas nas suas posições iniciais, bem como após 10 iterações, as posições das mesmas na 100^o iteração e finalmente quando atingem o número máximo de iterações definido para o problema. Pode-se observar que com o passar das iterações as partículas convergiram para uma solução.

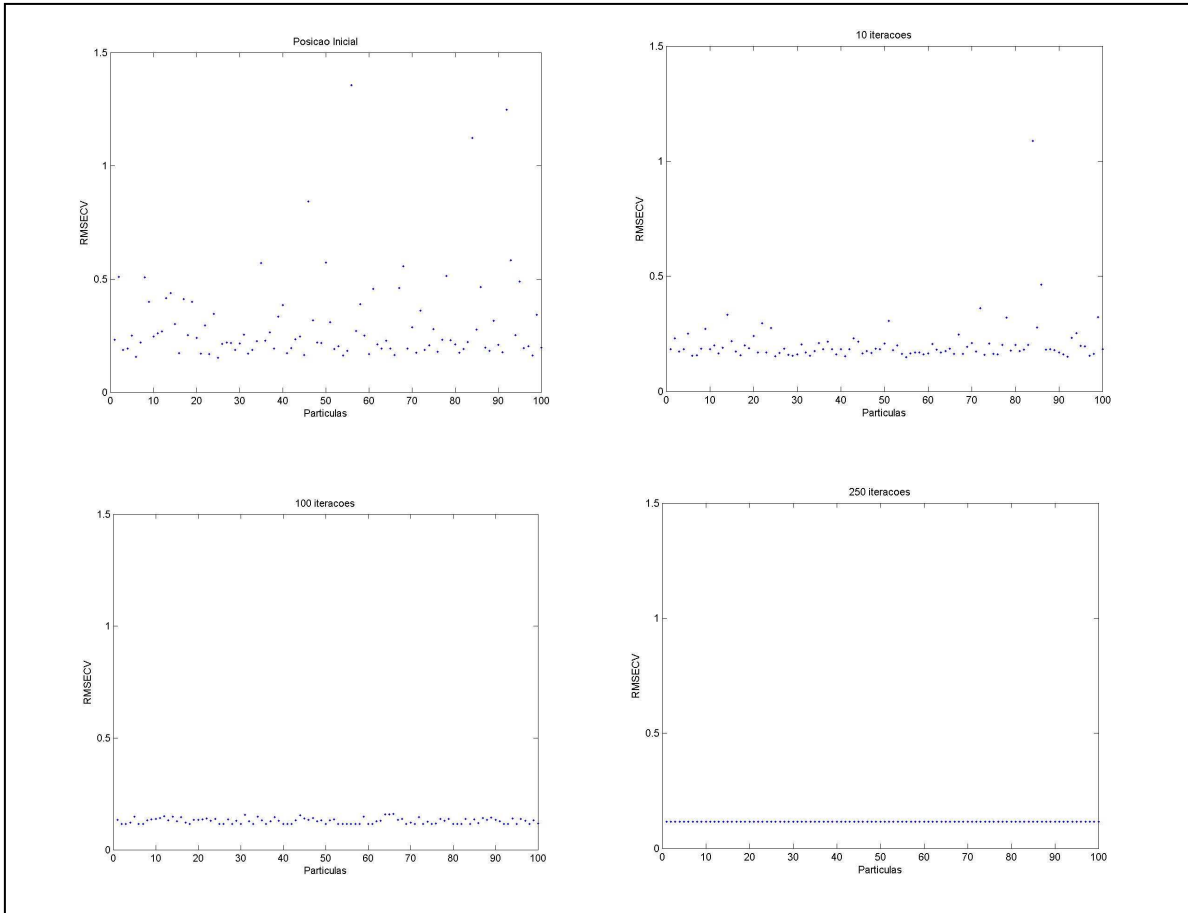


Figura 15 – Exemplo da evolução das partículas para a quantificação do biodiesel, usando o método iPLS-PSO (CD) com 40 intervalos e 100 partículas.

Fonte: elaborado pela autora.

A Figura 16 mostra um comparativo das evoluções para as 3 repetições para a combinação do iPLS-PSO com 20 intervalos e 50 partículas. Observa-se nesta proposta que existe uma tendência de convergência de todos os modelos, sendo necessário em média 44 iterações; mostrando juntamente que o número de iteração definido é suficiente para explorar as melhores combinações entre os 20 intervalos, atingindo um mínimo global.

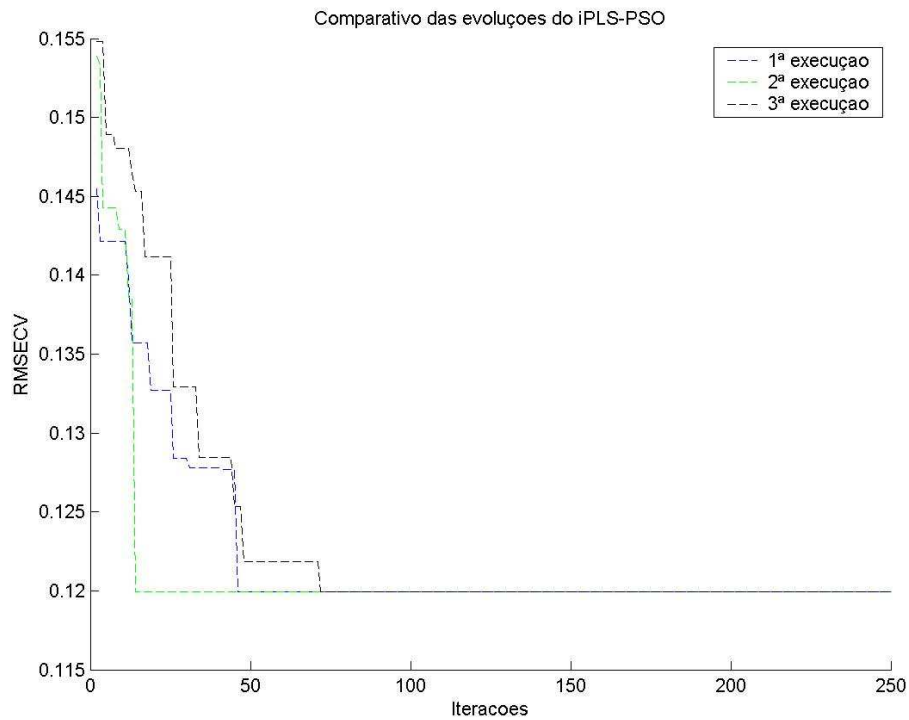


Figura 16 – Evolução das três execuções do iPLS-PSO com 20 intervalos e 50 partículas, para quantificação do biodiesel.
 Fonte: elaborado pela autora.

As Figuras 17 e 18 apresentam um comparativo de 3 evoluções do iPLS-PSO com 40 e 60 intervalos, respectivamente, ambos com 50 partículas. Observa-se um comportamento diferente para ambos, pois nem todos os modelos atingem o mínimo com poucas iterações. Em média para iPLS-PSO com 40 intervalos são necessárias 146 iterações, já para 60 intervalos em média 188. Em virtude disso, todos os resultados apresentados no ANEXO A foram obtidos com 250 iterações, pois um aumento no número de iterações refletiria num aumento considerável de custo computacional (tempo).

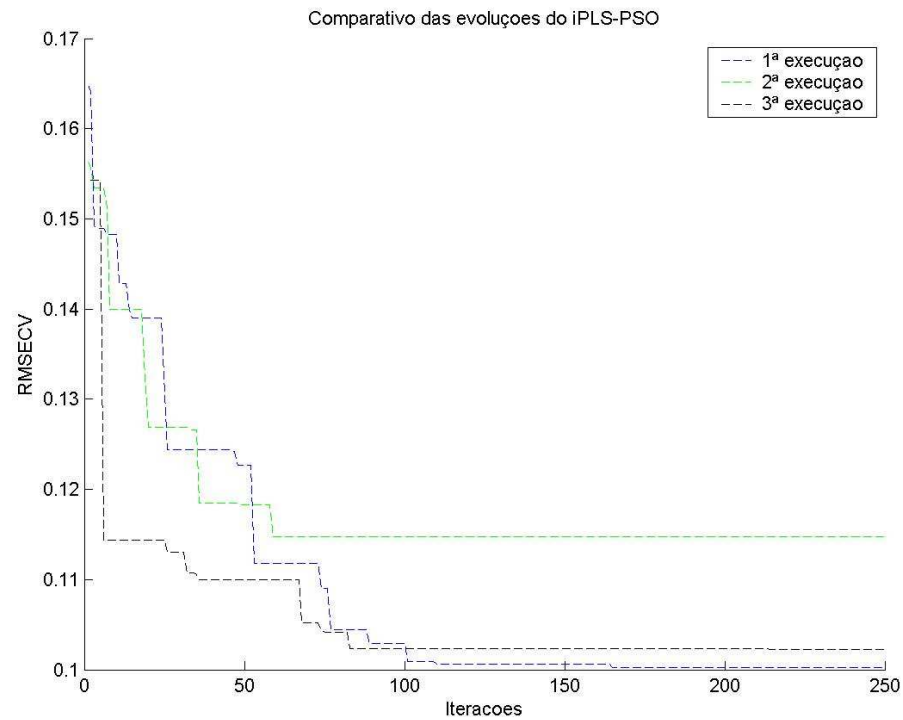


Figura 17 – Evolução das três execuções do iPLS-PSO com 40 intervalos e 50 partículas, para quantificação do biodiesel.
Fonte: elaborado pela autora.

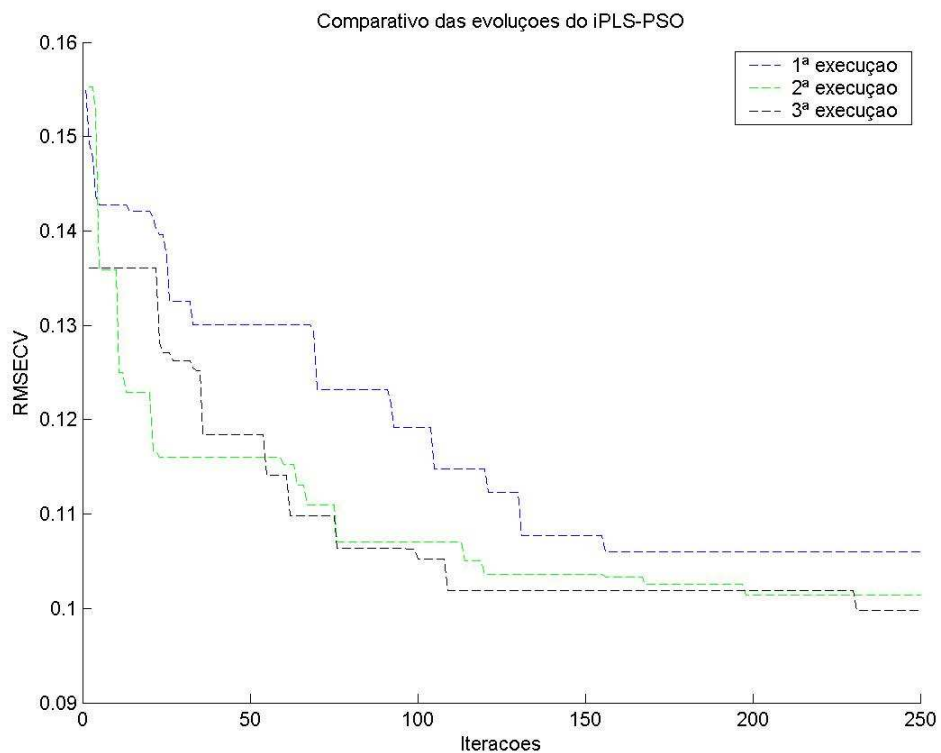


Figura 18 – Evolução das três execuções do iPLS-PSO com 60 intervalos e 50 partículas, para quantificação do biodiesel.
Fonte: elaborado pela autora.

5.1.3 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO

A aplicação da primeira implementação denominada iPLS-PSO no conjunto de dados propostos para a quantificação do biodiesel, selecionou regiões espectrais para construção do modelo que compreendem de 874 a 1319 cm^{-1} com lacunas, de 2240 a 2717 cm^{-1} , de 3055 a 3109 cm^{-1} e de 3446 a 3556 cm^{-1} como podem ser visualizadas na Figura 19. Dentre todos os modelos construídos para este método proposto, o selecionado foi o obtido com 60 intervalos e 100 partículas e o desempenho do modelo empregando 10 variáveis latentes para quantificação do biodiesel, é apresentada na Figura 20.

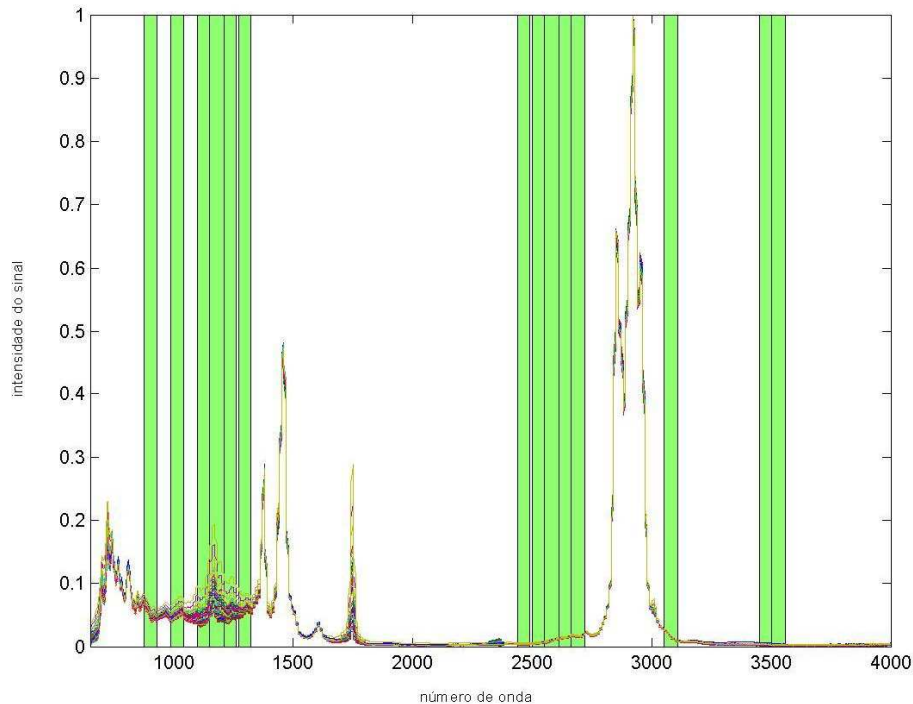


Figura 19 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel usando o método iPLS-PSO com 60 intervalos e 100 partículas.
Fonte: elaborado pela autora.

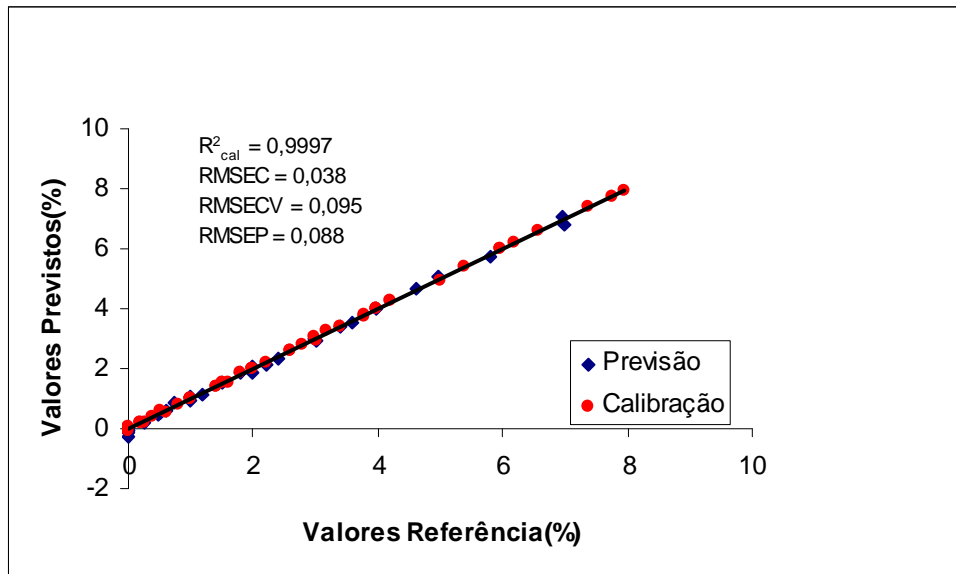


Figura 20 – Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 60 intervalos e 100 partículas.
 Fonte: elaborado pela autora.

5.1.4 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CC)

Já o método iPLS-PSO (CC) selecionou regiões espectrais que compreendem as faixas de 904 a 1327 cm^{-1} , de 2675 a 2756 cm^{-1} e de 3421 a 3502 cm^{-1} , conforme Figura 21. Nesta implementação o modelo selecionado empregou 40 intervalos e 25 partículas e, a sua respectiva curva de calibração utilizando 10 variáveis latentes para a concentração de biodiesel, é apresentada na Figura 22.

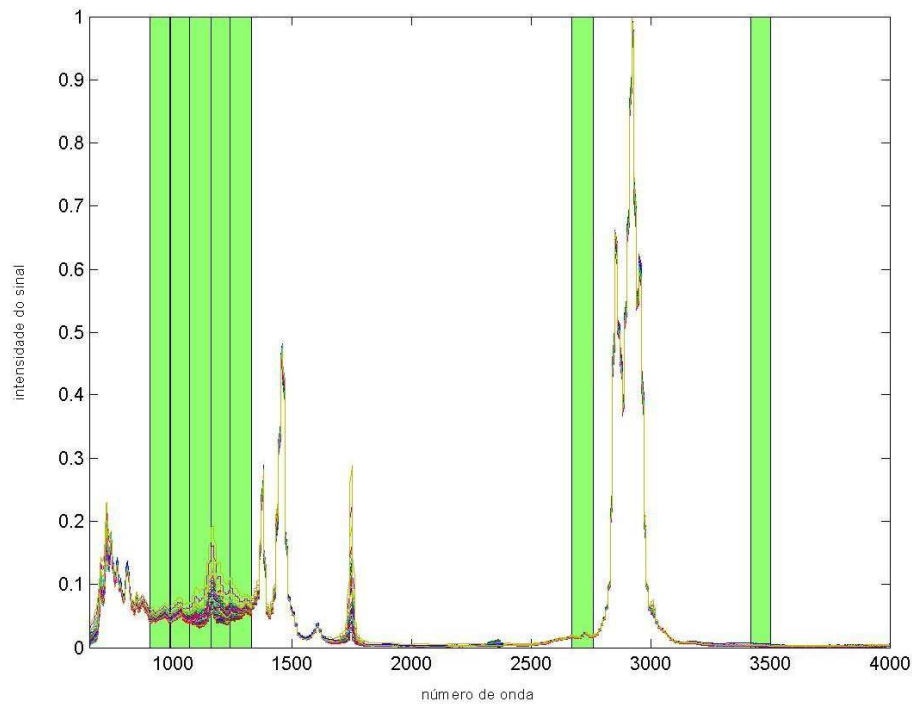


Figura 21 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

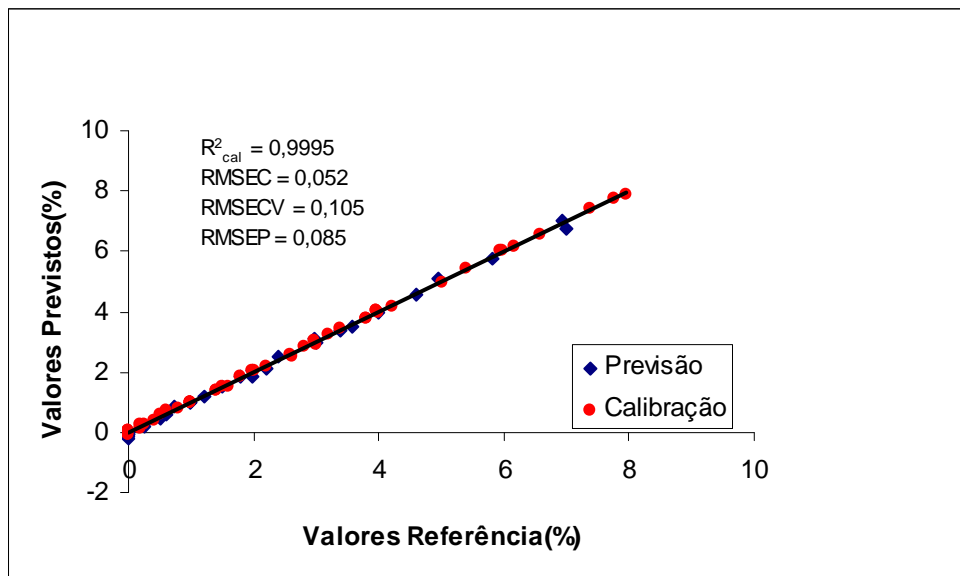


Figura 22 – Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.1.5 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (DD)

No caso da implementação iPLS-PSO (DD) as regiões espectrais selecionadas compreendem as faixas de 904 a 1327 cm^{-1} , de 2675 a 2756 cm^{-1} , de 3089 a 3170 cm^{-1} e de 3421 a 3502 cm^{-1} , conforme Figura 23. Neste caso, o modelo selecionado foi aquele construído com 40 intervalos e 100 partículas, e a curva de calibração, utilizando 10 variáveis latentes, para a concentração de biodiesel é apresentada na Figura 24.

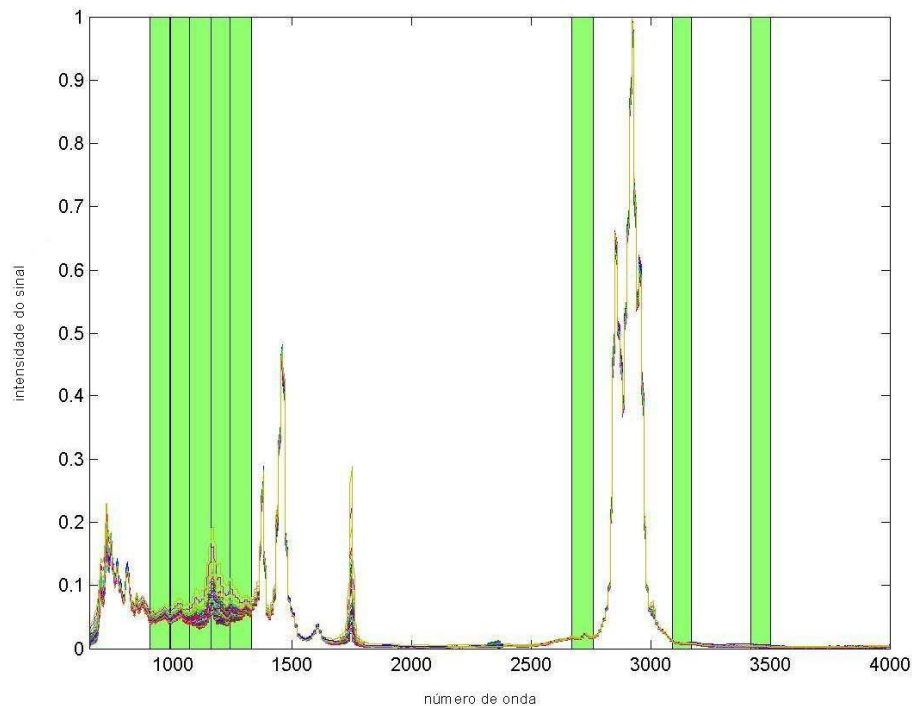


Figura 23– Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

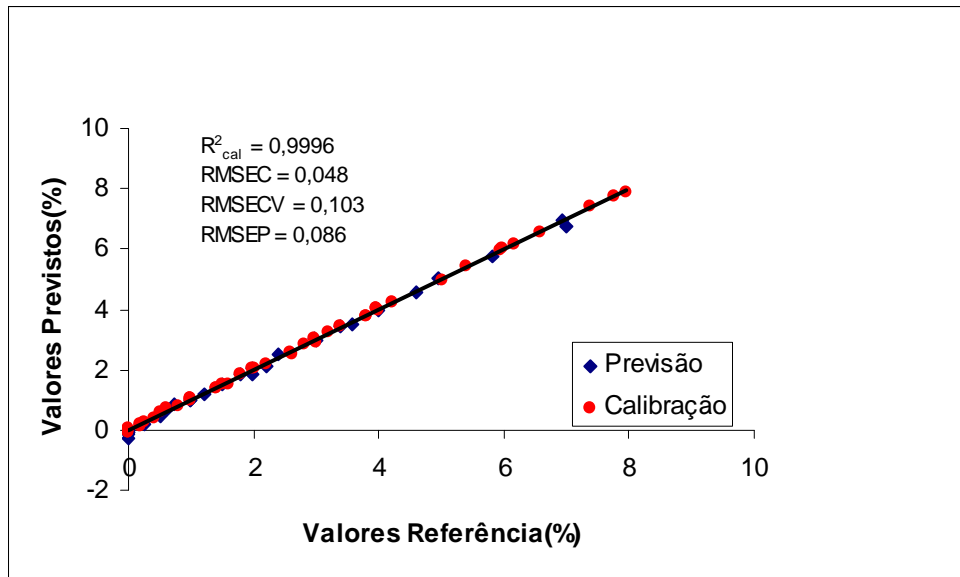


Figura 24– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

Um aspecto interessante nesta implementação é que conforme ANEXO A, na construção dos modelos com 20 intervalos, as partículas convergiram muito rápido para uma solução, não precisando ser explorado todo o espaço de busca, convergindo em média após 33 iterações.

5.1.6 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CD)

Conforme a Tabela 5, o modelo selecionado pelo método iPLS-PSO (CD) corresponde as mesmas regiões espectrais que a implementação do método iPLS-PSO (CC). As regiões espectrais bem como a curva de calibração pode ser vistas nas Figuras 21 e 22, anteriormente apresentadas.

5.1.7 Resultados obtidos para a determinação do biodiesel aplicando o iPLS-PSO (DC)

Neste método iPLS-PSO (DC) as regiões espectrais selecionadas para construção do modelo compreendem as faixas de 904 a 1327 cm^{-1} com lacunas, de 2426 a 2756 cm^{-1} e,

de 3504 a 3585 cm^{-1} conforme Figura 25. Dentre todos os modelos construídos para este método o modelo selecionado foi o construído com 40 intervalos e 100 partículas, e a curva de calibração utilizando 8 variáveis latentes, para a concentração de biodiesel, é apresentada na Figura 26.

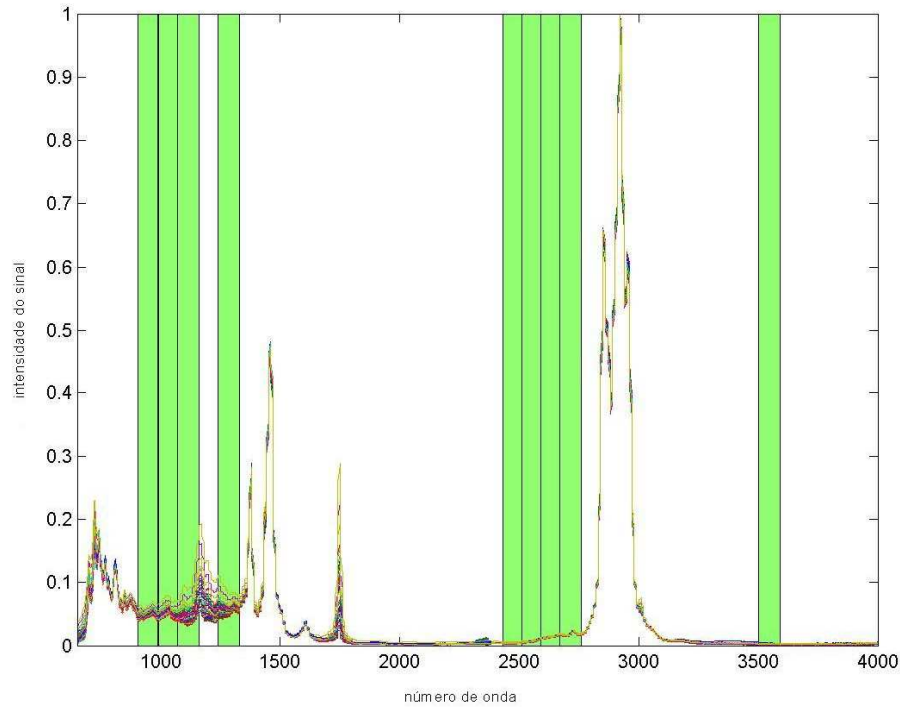


Figura 25 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

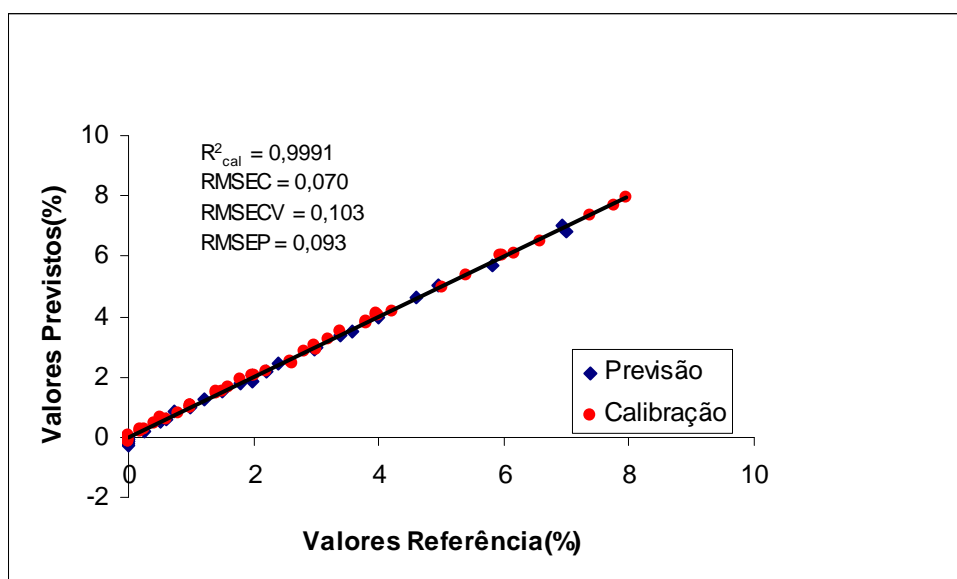


Figura 26– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

5.1.8 Considerações sobre os modelos para a previsão do biodiesel

Conforme evidenciado, todos os modelos construídos para quantificação do biodiesel empregam regiões espectrais semelhantes, e sendo que todos foram capazes de quantificar o analito proposto. Os resultados encontrados empregando as soluções propostas pelo método enxame de partículas, em comparação aos resultados empregando o PLS sobre todo o espectro bem como as regiões espectrais sugeridas pela proposta de norma ABNT, demonstraram que houve uma otimização dos modelos de regressão multivariados.

A região selecionada pelos modelos otimizados contempla em parte a região espectral, sugerida pela proposta de norma, e ampliada principalmente na faixa dos sinais referentes ao C-O. Um aspecto interessante é que em nenhum dos modelos otimizados, selecionou a região da carbonila (C=O), região esta que a proposta de norma preconiza como sendo a principal banda para a quantificação do biodiesel.

Uma hipótese é que a região da carbonila não apresenta informações suficientes para diferenciar a contribuição do sinal do biodiesel e do óleo isoladamente, ou seja, provavelmente as modificações que acontecem do óleo para o biodiesel são pouco pronunciadas nesta região e muito mais pronunciadas na região do sinal correspondente ao C-O.

Com base nestas observações e considerando alguns dos resultados das otimizações propostas pelo método enxame de partículas, é possível obter modelos com boa capacidade preditiva com as informações espectrais das regiões de 904 a 1327 cm^{-1} com lacunas, ilustradas na Figura 27, obtidas pelo método iPLS-PSO em uma de suas soluções. O desempenho deste modelo, com 10 variáveis latentes, pode ser visualizado na Figura 28.

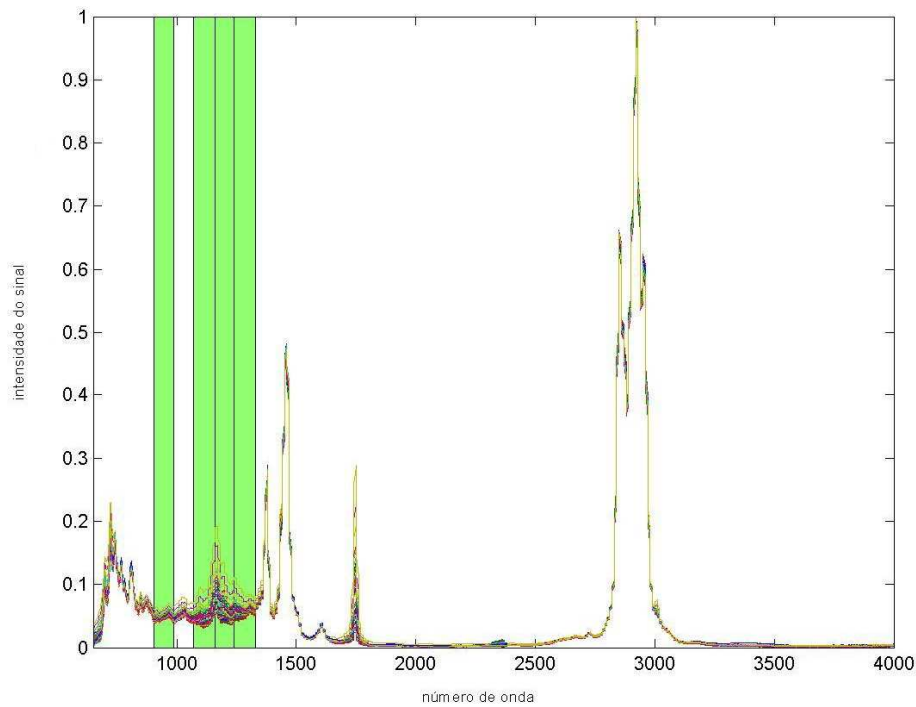


Figura 27 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

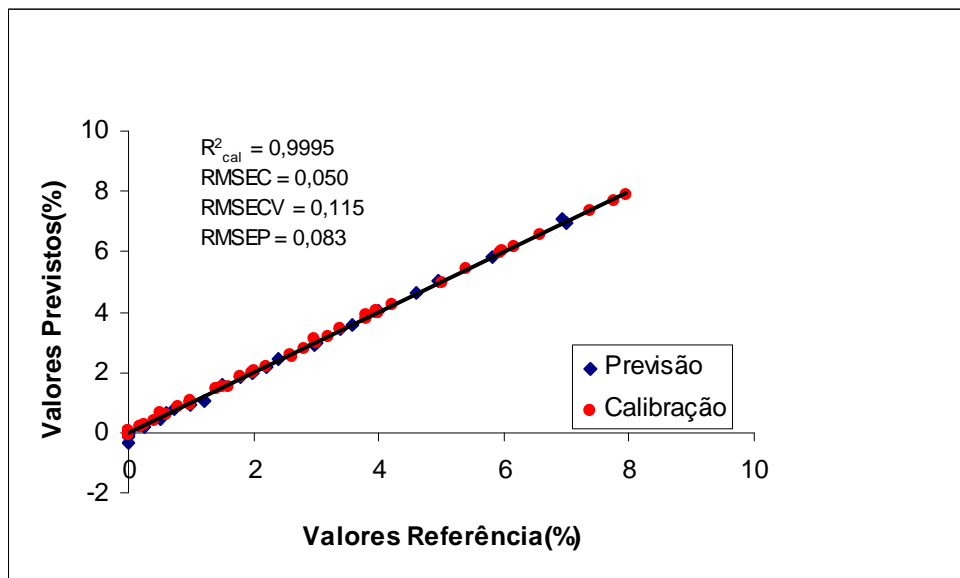


Figura 28– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

5.2 Resultados para a quantificação do óleo vegetal

Neste trabalho também foram desenvolvidos diversos modelos de regressão visando quantificar o óleo vegetal presente nas misturas, com o intuito de verificar a capacidade de detecção de adulterantes nas misturas biodiesel/óleo vegetal/diesel através do uso do FT-IR/ATR. Os principais resultados obtidos estão descritos na seqüência. Primeiramente são apresentados os modelos obtidos sem a otimização através da meta-heurística implementada e em seguida, são descritos os modelos obtidos pela combinação da PSO e do iPLS.

5.2.1 Modelos para determinação do óleo vegetal com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT

Com base no conjunto de dados foram desenvolvidos dois modelos de regressão, ambos utilizando para construção o método PLS. O primeiro modelo, para a quantificação do óleo vegetal foi construído utilizando toda região espectral (650 cm^{-1} a 4000 cm^{-1}) conforme Figura 11, apresentada anteriormente, e o desempenho do modelo é apresentado na Figura 29, empregando 6 variáveis latentes.

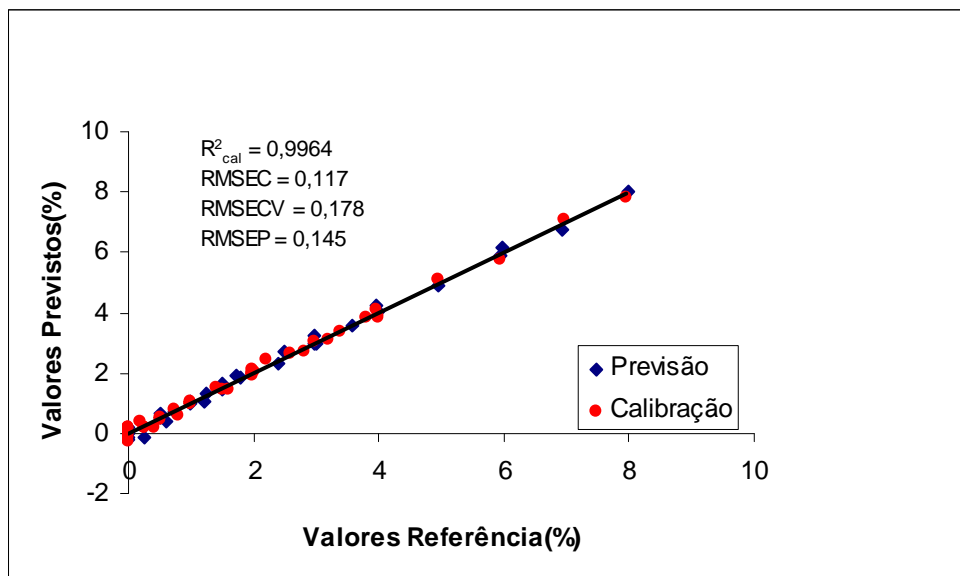


Figura 29 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal utilizando toda a região espectral das misturas.
 Fonte: elaborado pela autora.

O segundo modelo construído para a quantificação do óleo vegetal nas misturas empregou somente as duas regiões espectrais (de 1100 cm^{-1} a 1200 cm^{-1} e de 1735 cm^{-1} a 1750 cm^{-1}) sugeridas pela proposta de norma ABNT, conforme Figura 13. O desempenho do modelo, utilizando 5 variáveis latentes, é apresentado na Figura 30.

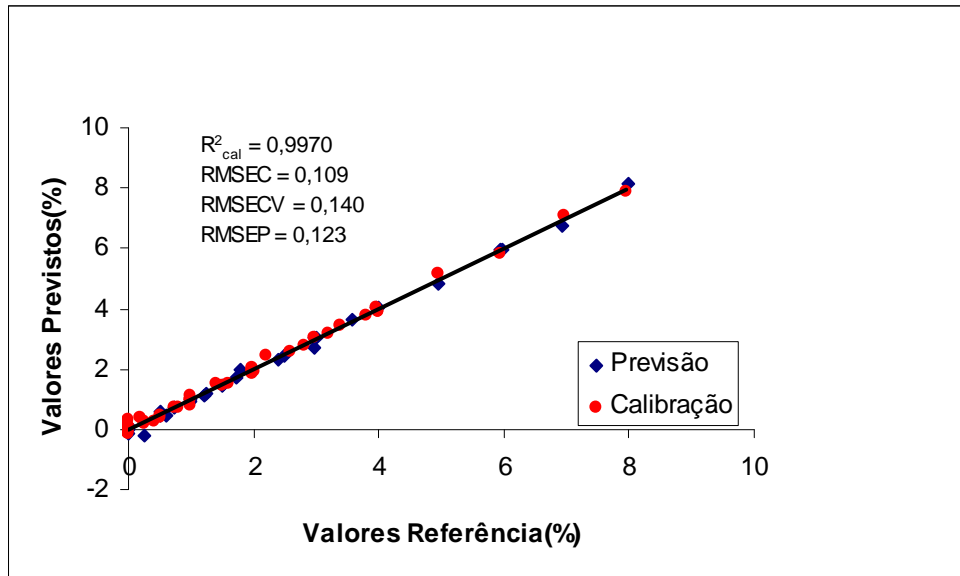


Figura 30 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT. Fonte: elaborado pela autora.

Observando os dois modelos anteriores pode-se verificar que os resultados obtidos pelo modelo construído, empregando somente as regiões espectrais sugerida pela proposta de norma ABNT, apresenta uma melhor capacidade de quantificação do óleo vegetal do que o modelo construído com o mesmo propósito, empregado toda a região do espectro.

Tentando analisar a possibilidade de obter modelos com melhor capacidade preditiva, sugere-se a mesma implementação proposta para construção dos modelos de biodiesel, desta vez com o propósito de otimizar a quantificação do óleo vegetal nas misturas.

5.2.2 Resultados gerais e evoluções dos modelos para a previsão do óleo vegetal

Foram testadas na quantificação do óleo vegetal, as mesmas cinco implementações da otimização por enxame de partículas: iPLS-PSO, iPLS-PSO (CC), iPLS-PSO (DD); iPLS-PSO (CD); iPLS-PSO (DC), descritas anteriormente. Os principais resultados dos modelos, incluindo os obtidos através das implementações com a PSO, são apresentados na Tabela

6. Os resultados completos das implementações do enxame de partículas para a quantificação do óleo vegetal encontram-se no ANEXO B.

Tabela 6 – Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do óleo vegetal.

Modelos	Regiões empregadas	R² cal	RMSEC	RMSECV	RMSEP
espectro inteiro	650 - 4000 cm ⁻¹	0,9964	0,117	0,178	0,145
regiões sugeridas pela proposta de norma ABNT	1101 - 1198 cm ⁻¹ 1736 - 1749 cm ⁻¹	0,9970	0,109	0,140	0,123
iPLS- PSO	904 - 1412 cm ⁻¹ 1583 - 1751 cm ⁻¹ 2509 - 2756 cm ⁻¹ 3504 - 3585 cm ⁻¹	0,9991	0,059	0,083	0,092
iPLS- PSO (CC)	904 - 1412 cm ⁻¹ 1583 - 1751 cm ⁻¹ 1838 - 2006 cm ⁻¹	0,9990	0,060	0,088	0,089
iPLS- PSO (DD)	904 - 1327 cm ⁻¹ 1583 - 1666 cm ⁻¹ 2008 - 2341 cm ⁻¹	0,0995	0,042	0,082	0,089
iPLS- PSO (CD)	804 - 1319 cm ⁻¹ 1992 - 2158 cm ⁻¹ 2495 - 2829 cm ⁻¹ 3334 - 3500 cm ⁻¹	0,9995	0,041	0,081	0,094
iPLS- PSO (DC)	820 - 1751 cm ⁻¹ 2426 - 2756 cm ⁻¹ 3007- 3253 cm ⁻¹	0,9992	0,056	0,085	0,090

Fonte: elaborado pela autora.

Um comparativo das evoluções para 3 execuções do método iPLS-PSO com 20 intervalos e 50 partículas para quantificação do óleo vegetal, é apresentada na Figura 31. Observa-se nesta proposta que existe uma tendência de convergência de todos os modelos, sendo necessário em média de 108 iterações; mostrando juntamente que o número total de iterações definido é suficiente para explorar todas as melhores combinações possíveis entre os 20 intervalos.

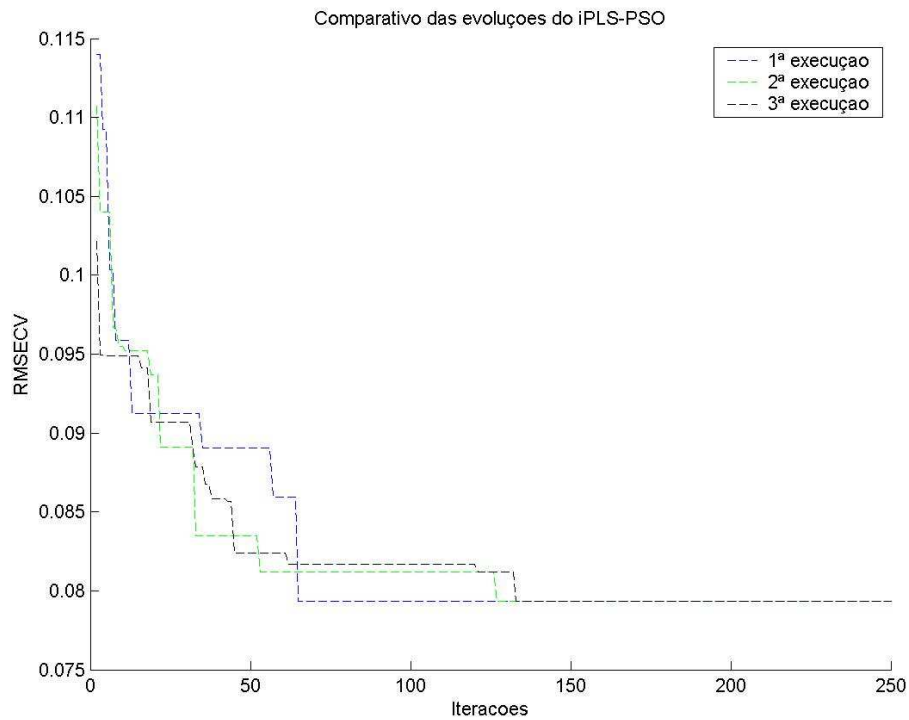


Figura 31– Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas, para a quantificação do óleo vegetal.
 Fonte: elaborado pela autora.

Observa-se um comportamento diferente quando se apresentam as 3 evoluções do iPLS-PSO com 40 e 60 intervalos, ambos com 50 partículas, como pode ser visto nas Figuras 32 e 33. São necessários mais iterações para que os modelos atinjam mínimos locais, em média para iPLS-PSO com 40 intervalos são necessárias 184 iterações, já para 60 intervalos em média 209 iterações.

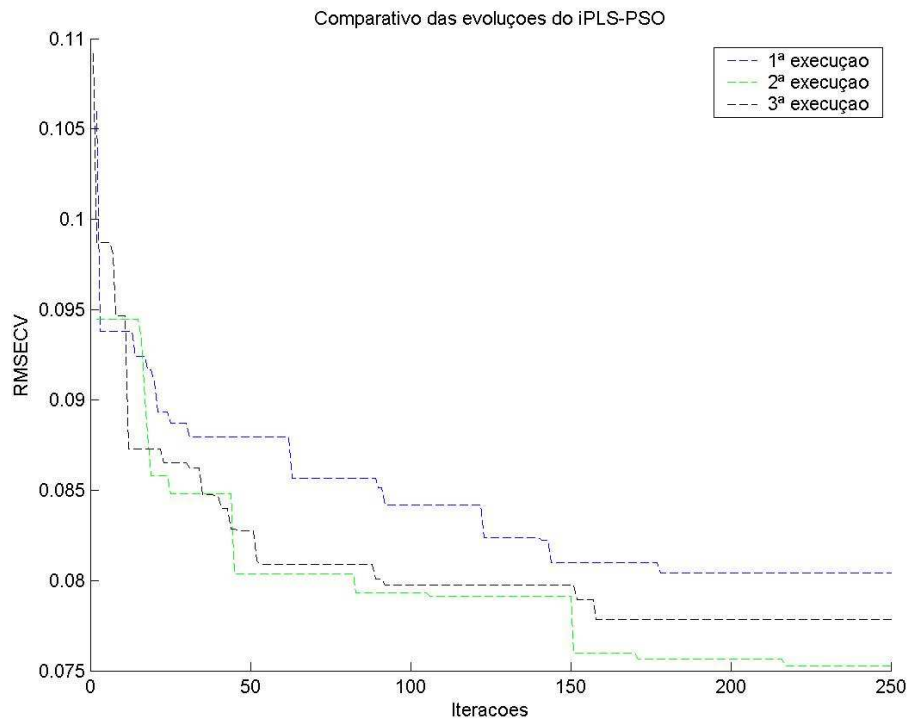


Figura 32 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas, para a quantificação do óleo vegetal.
Fonte: elaborado pela autora.

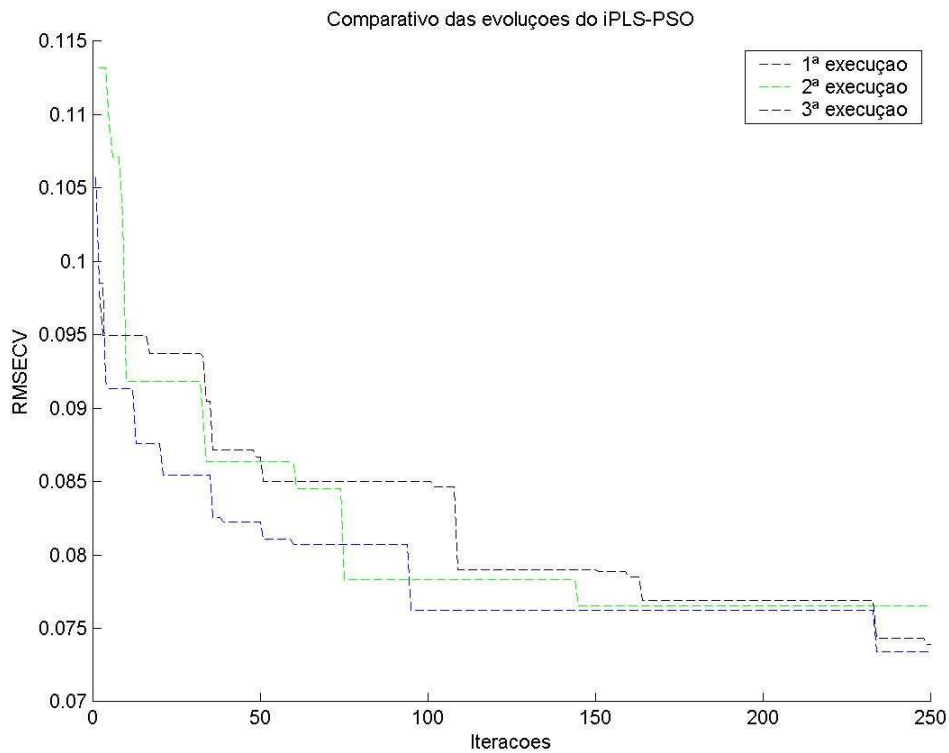


Figura 33 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para a quantificação do óleo vegetal.
Fonte: elaborado pela autora.

5.2.3 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO

Nesta primeira implementação denominada iPLS-PSO aplicada ao conjunto de dados propostos para quantificar o óleo vegetal, as regiões espectrais selecionadas para construção do modelo foram as que compreendem de 904 a 1412 cm^{-1} com lacunas, de 1583 a 1751 cm^{-1} , de 2509 a 2756 cm^{-1} e de 3504 a 3585 cm^{-1} como podem ser visualizadas na Figura 34. Dentre todos os modelos construídos, o selecionado foi o construído com 40 intervalos e 25 partículas e a sua respectiva curva de calibração, com 9 variáveis latentes, para quantificação do óleo vegetal, é apresentada na Figura 35.

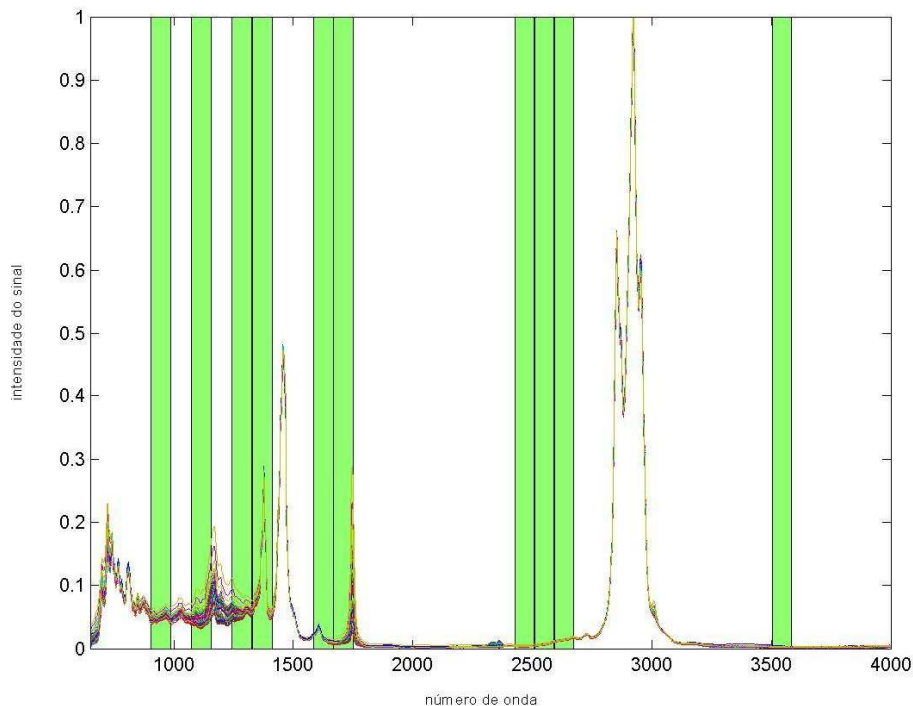


Figura 34 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

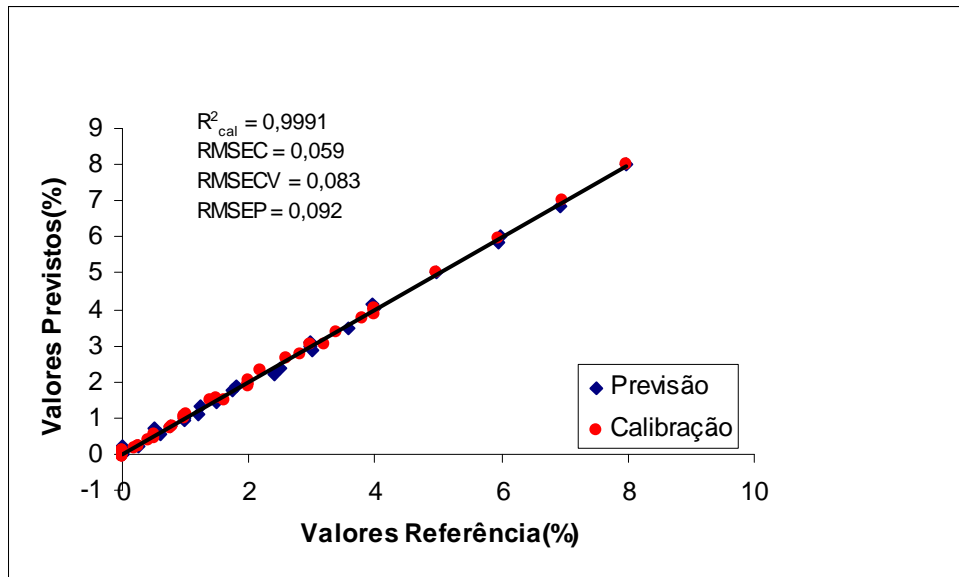


Figura 35 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.
 Fonte: elaborado pela autora.

5.2.4 Resultados obtido para a determinação do óleo vegetal aplicando o método iPLS-PSO (CC)

Na implementação iPLS-PSO (CC) as regiões espectrais selecionadas para construção do modelo de quantificação do óleo vegetal compreendem as faixas de 904 a 1412 cm^{-1} com lacunas, de 1583 a 1666 cm^{-1} e de 2008 a 2341 cm^{-1} , conforme Figura 36. Nesta abordagem, o modelo selecionado empregou 40 intervalos e 25 partículas e a sua respectiva curva de calibração, com 9 variáveis latentes, para a concentração do óleo vegetal, é apresentada na Figura 37.

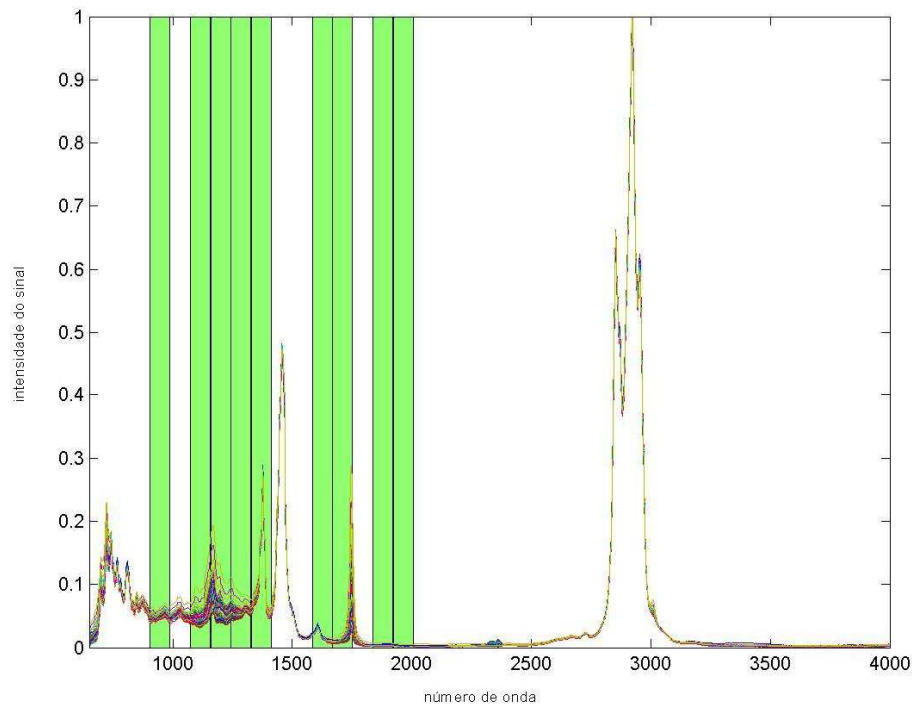


Figura 36 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

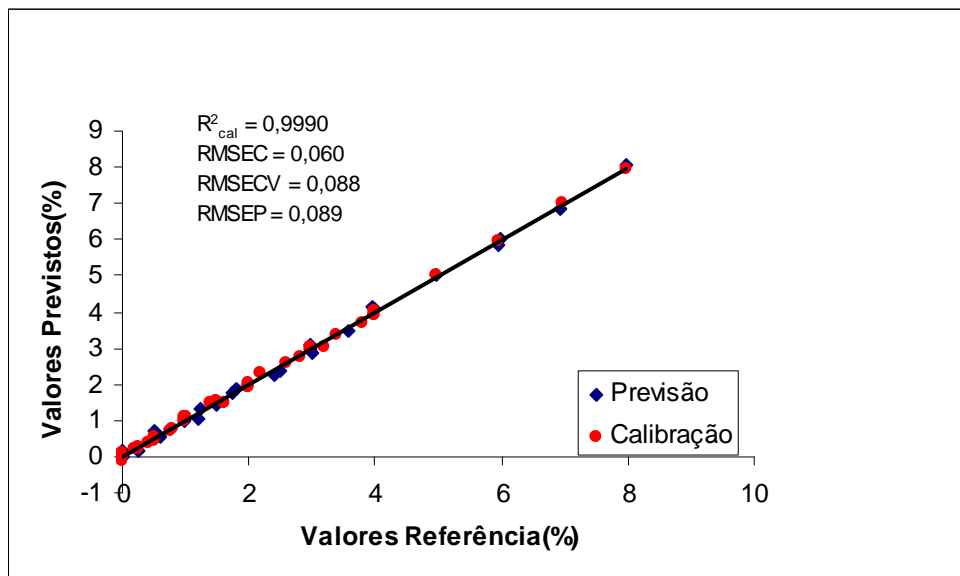


Figura 37 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.2.5 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DD)

Já o método iPLS-PSO (DD) para a quantificação do óleo vegetal nas misturas, o modelo selecionado foi aquele construído com 40 intervalos e 100 partículas, compreendendo as regiões espectrais selecionadas nas faixas de 904 a 1327 cm^{-1} com lacunas, de 1583 a 1666 cm^{-1} e de 2008 a 2341 cm^{-1} com lacunas, conforme Figura 38, e a curva de calibração, com 10 variáveis latentes, para a concentração do óleo vegetal é apresentada na Figura 39.

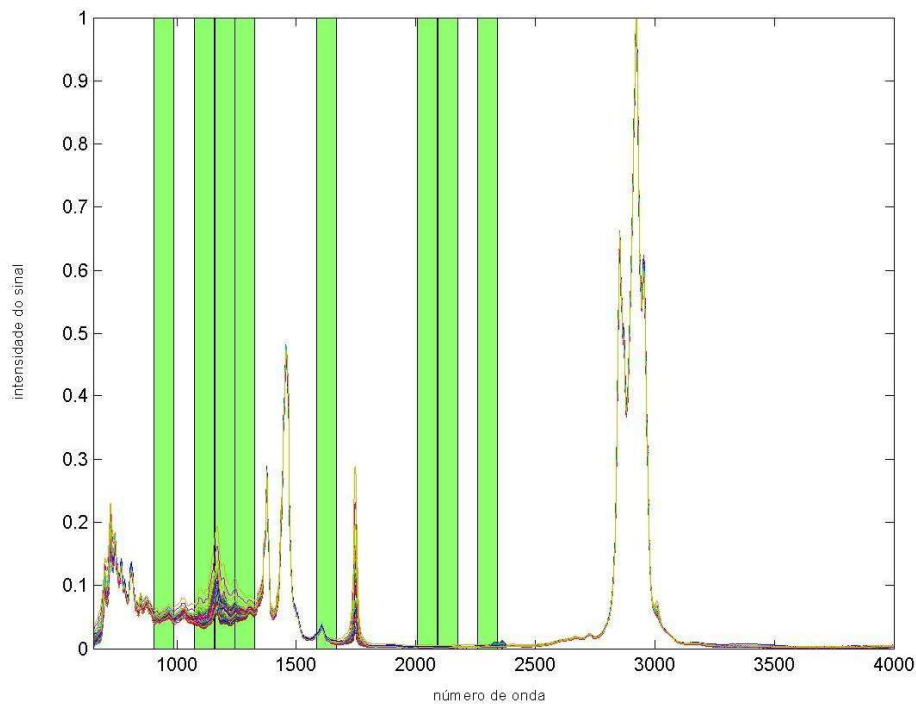


Figura 38 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

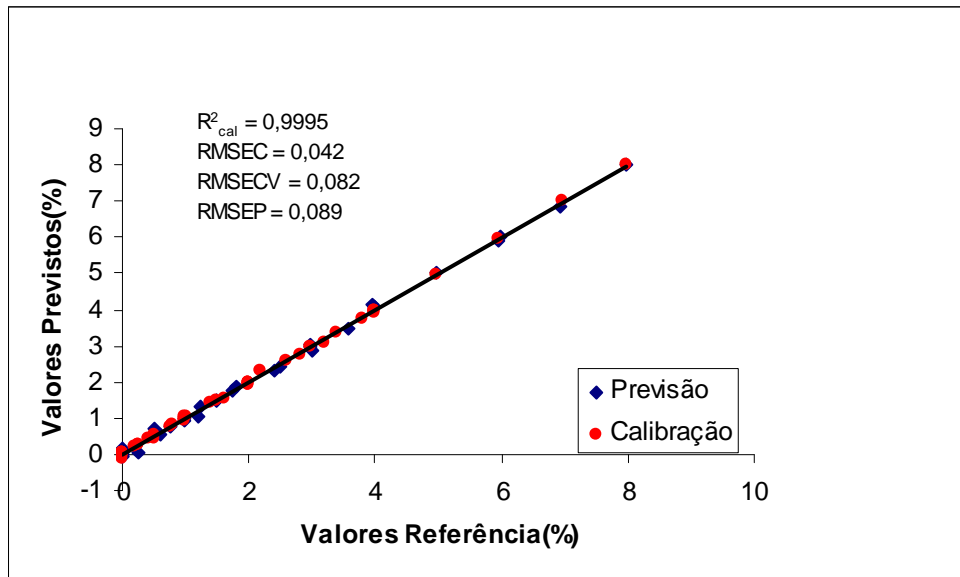


Figura 39 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas. Fonte: elaborado pela autora.

5.2.6 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (CD)

No caso da implementação iPLS-PSO (CD) as regiões espectrais selecionadas compreendem as faixas de 804 a 1319 cm^{-1} , de 1992 a 2158 cm^{-1} , de 2495 a 2829 cm^{-1} e de 3334 a 3500 cm^{-1} , conforme Figura 40. Neste caso, o modelo selecionado, dentre todos os modelos construídos para esta abordagem, foi aquele elaborado com 20 intervalos e 25 partículas, e a curva de calibração, com 10 variáveis latentes, para a concentração do óleo vegetal é apresentada na Figura 41.

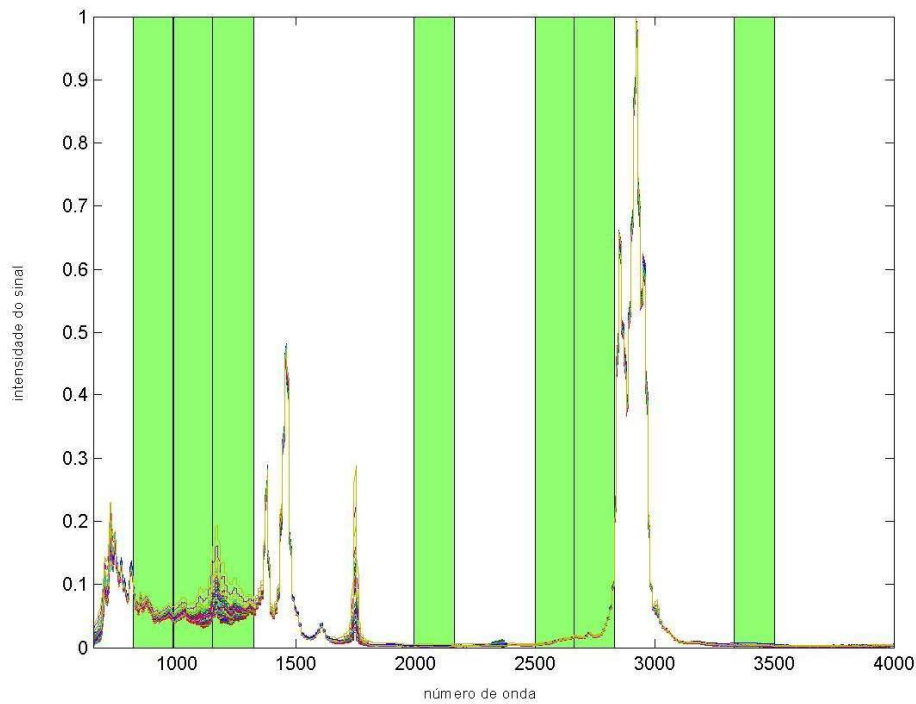


Figura 40 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.
Fonte: elaborado pela autora.

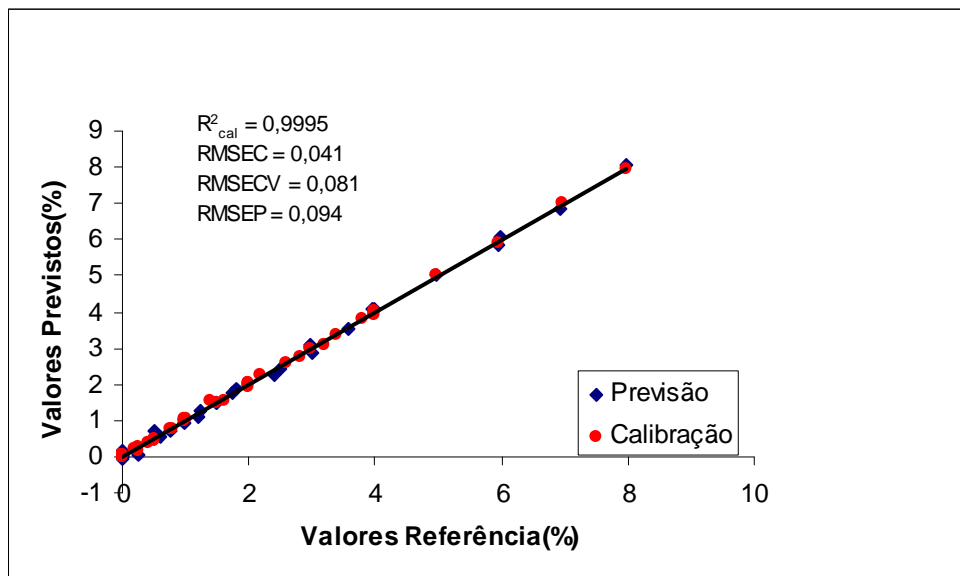


Figura 41– Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.2.7 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DC)

Já a implementação denominada iPLS-PSO (DC) selecionou as seguintes regiões espectrais para construção do modelo de 820 a 1751 cm^{-1} , de 2426 a 2756 cm^{-1} e de 3007 a 3253 cm^{-1} como podem ser visualizadas na Figura 42. Dentre todos os modelos construídos para iPLS-PSO (DC), o selecionado foi aquele construído com 40 intervalos e 100 partículas e a sua respectiva curva de calibração com 10 variáveis latentes, para quantificação do óleo vegetal, é apresentada na Figura 43.

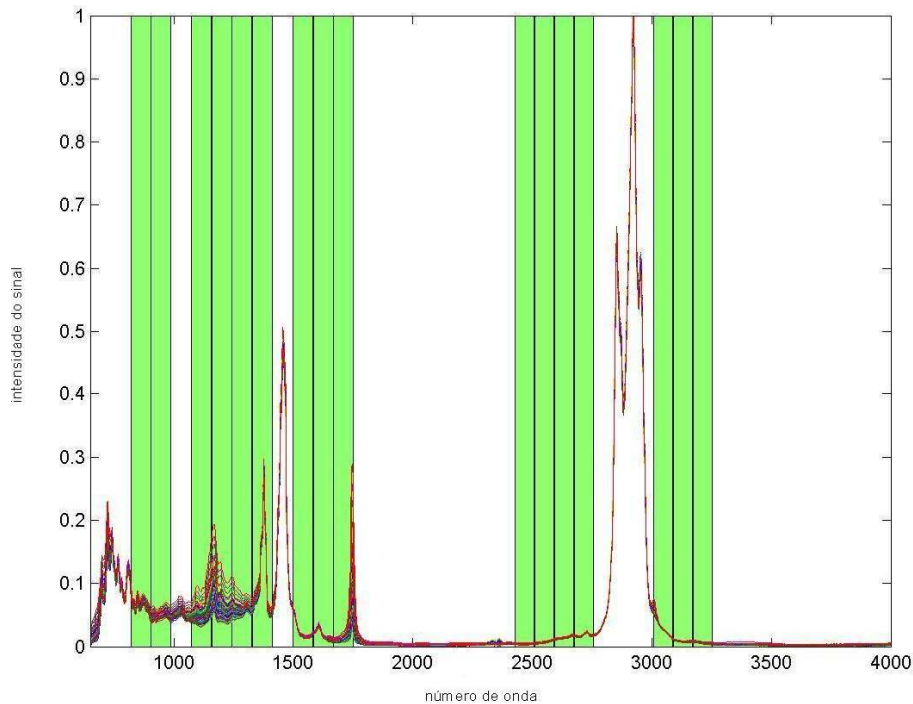


Figura 42 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

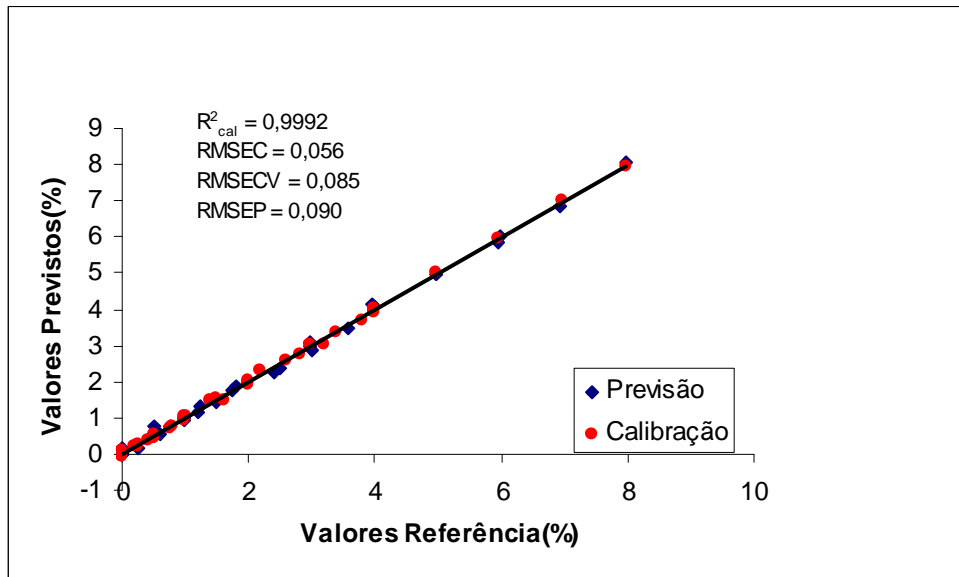


Figura 43 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas. Fonte: elaborado pela autora.

5.2.8 Considerações sobre os modelos para a previsão de óleo vegetal

Analisando os resultados anteriores, podemos observar que todos os modelos construídos para quantificar o óleo vegetal foram capazes de atingir este propósito, apresentando baixos erros de previsão (RMSEP). Cabe ressaltar que os modelos construídos empregando o método PLS com o espectro inteiro das misturas e somente com as regiões sugeridas pela proposta de norma ABNT apresentaram uma capacidade de quantificação do óleo vegetal menor do que os modelos construídos através do método híbrido proposto nesta pesquisa.

Os modelos selecionados através das otimizações aqui propostas apresentaram valores de RMSEP entre 0,089 a 0,094, enquanto que os valores de RMSECV ficaram na amplitude de 0,081 a 0,088.

Do ponto de vista das regiões selecionadas, o modelo para previsão do óleo vegetal além de necessitar da região do C-O, também necessita utilizar as informações referentes ao sinal da carbonila (C=O), conforme as figuras anteriores, que apresentam as regiões selecionadas.

5.3 Resultados para a quantificação do óleo diesel

Finalmente, o outro analito quantificado através de modelos de regressão multivariados nas misturas, neste trabalho foi o óleo diesel. Embora não seja comum quantificar o componente majoritário das misturas, o propósito desta determinação é verificar o comportamento do algoritmo e da técnica FT-IR/ATR. Para isso foram desenvolvidos diversos modelos de regressão. Neste caso, foram construídos modelos usando o PLS com o espectro inteiro das misturas e posteriormente modelos com as implementações propostas com a meta-heurística enxame de partículas em conjunto com o iPLS.

5.3.1 Modelo para a determinação do óleo diesel com espectro inteiro

Para a determinação do óleo diesel também foi desenvolvido um modelo de regressão com o método PLS, utilizando toda a região espectral (650 cm^{-1} a 4000 cm^{-1}) conforme Figura 11 apresentada anteriormente. O desempenho do modelo, com 4 variáveis latentes, é apresentado na Figura 44.

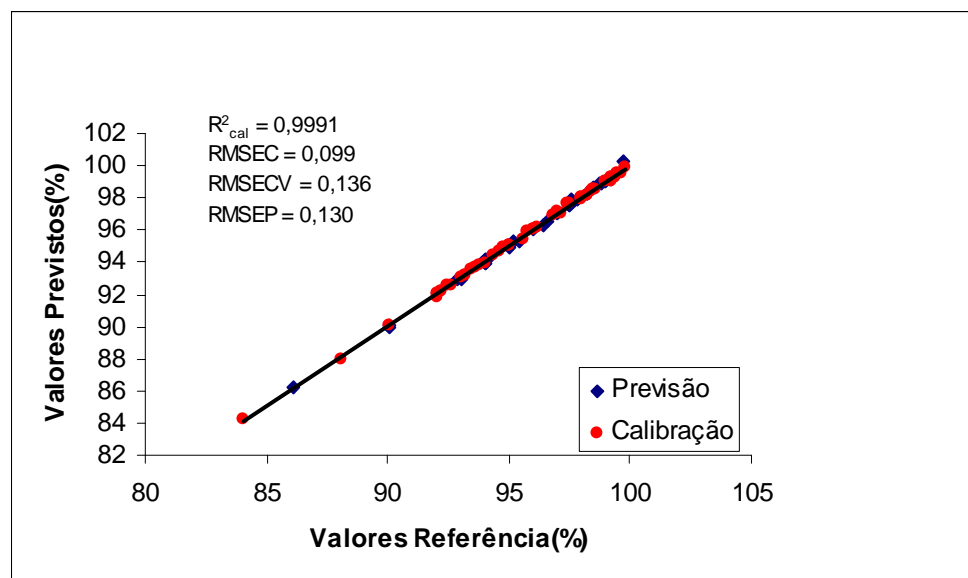


Figura 44 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel utilizando toda a região espectral da mistura.
Fonte: elaborado pela autora.

5.3.2 Resultados gerais e evoluções para a determinação do óleo diesel

Tentando avaliar se o comportamento das implementações propostas pelo enxame de partículas seria capaz gerar modelos mais robustos para quantificação do óleo diesel, foram testadas as cinco implementações na quantificação deste analito. Os principais resultados dos modelos, incluindo as resultantes das implementações com a PSO são apresentados na Tabela 7.

Tabela 7 – Principais resultados dos modelos, incluindo das cinco implementações da PSO para a quantificação do óleo diesel.

Modelos	Regiões empregadas	R ² cal	RMSEC	RMSECV	RMSEP
espectro inteiro	650 - 4000 cm ⁻¹	0,9991	0,099	0,136	0,130
iPLS- PSO	735 - 903 cm ⁻¹	0,9998	0,048	0,098	0,102
	1159 - 1327 cm ⁻¹				
	3172 - 3253 cm ⁻¹				
iPLS- PSO (CC)	762 - 1207 cm ⁻¹	0,9998	0,047	0,099	0,108
	2719 - 2773 cm ⁻¹				
	3223 - 3332 cm ⁻¹				
	3446 - 3500 cm ⁻¹				
iPLS- PSO (DD)	904 - 1327 cm ⁻¹	0,9997	0,059	0,096	0,112
	2092 - 2175 cm ⁻¹				
	2426 - 2756 cm ⁻¹				
	3089 - 3170 cm ⁻¹				
	3504 - 3585 cm ⁻¹				
	3753 - 3834 cm ⁻¹				
iPLS- PSO (CD)	904 - 1327 cm ⁻¹	0,9999	0,038	0,095	0,108
	1753 - 1836 cm ⁻¹				
	2260 - 2507 cm ⁻¹				
	2675 - 2756 cm ⁻¹				
	3753 - 3834 cm ⁻¹				
iPLS- PSO (DC)	735 - 1327 cm ⁻¹	0,9997	0,055	0,108	0,111

Fonte: elaborado pela autora.

A seguir, com base na implementação iPLS-PSO com 20 intervalos e 50 partículas para quantificação do óleo diesel, é mostrado um comparativo das evoluções para as 3 execuções, conforme a Figura 45. Observa-se nesta proposta que existe uma tendência de convergência de todos os modelos, sendo necessário em média de 107 iterações; comprovando novamente que o número de iterações definido é suficiente para explorar todas as melhores combinações possíveis entre os 20 intervalos.

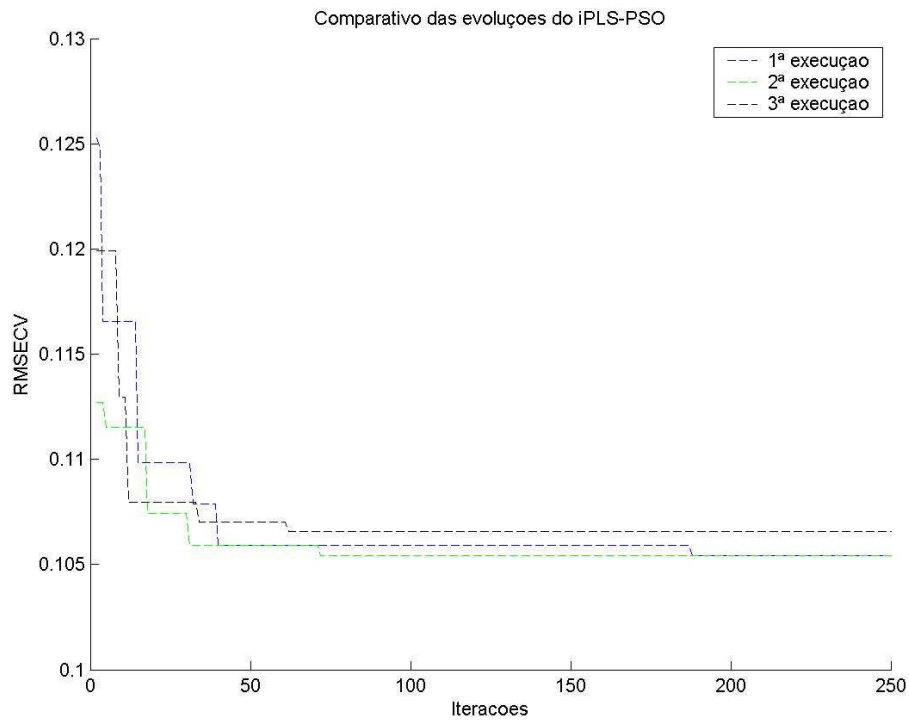


Figura 45 – Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas para a quantificação do óleo diesel.
Fonte: elaborado pela autora.

O comportamento das evoluções para as 3 execuções do método iPLS_PSO com 40 e 60 intervalos, ambos com 50 partículas, é diferente, pois a maioria dos modelos necessita de um número de iterações superior do que o modelo com 20 partículas, para atingirem os mínimos locais. Na evolução iPLS-PSO com 40 intervalos, percebe-se que em média são necessárias 155 iterações, já para 60 intervalos, em média, são necessárias 123 iterações como pode ser visualizada nas Figuras 46 e 47.

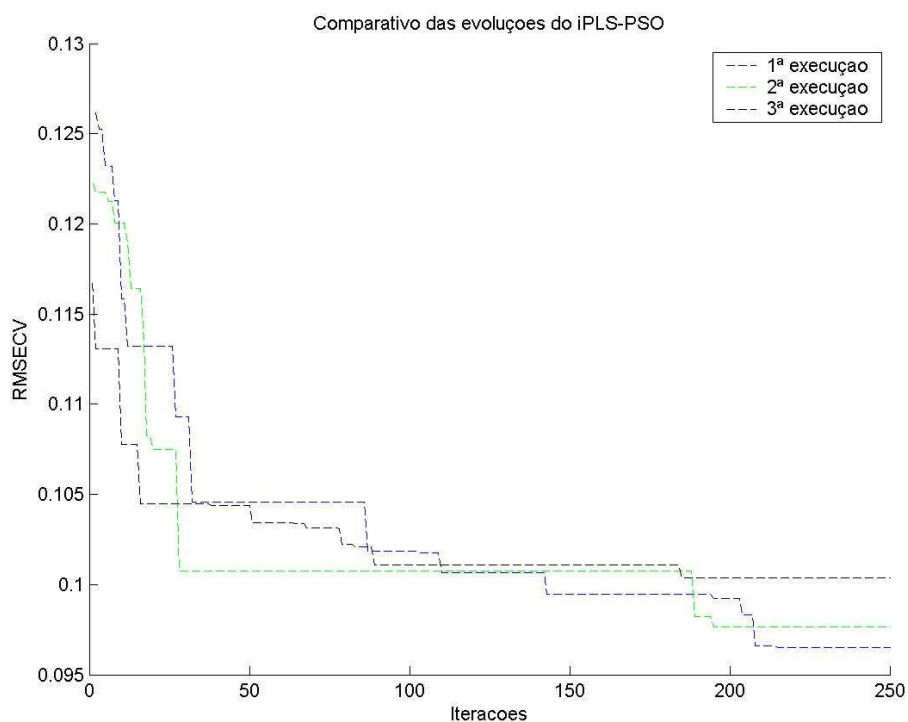


Figura 46 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas para quantificação do óleo diesel.
Fonte: elaborado pela autora.

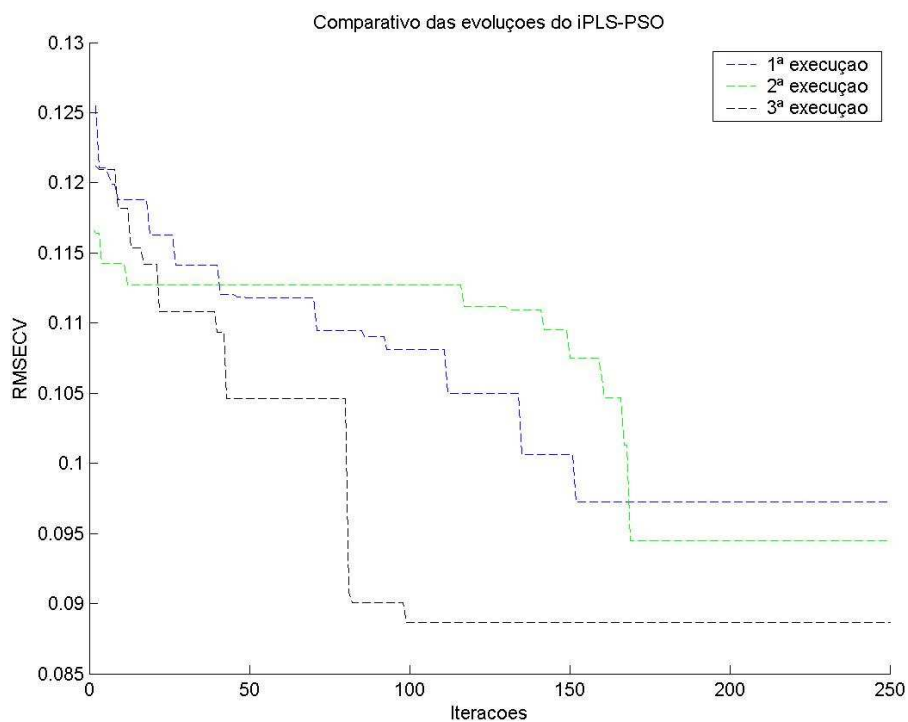


Figura 47 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para quantificação do óleo diesel.
Fonte: elaborado pela autora.

5.3.3 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO

Nesta implementação denominada iPLS-PSO, utilizada para quantificação do óleo diesel no conjunto de dados proposto, selecionou regiões espectrais para construção do modelo que compreendem de 735 a 903 cm^{-1} , de 1159 a 1327 cm^{-1} e de 3172 a 3253 cm^{-1} como podem ser visualizadas na Figura 48. Analisando todos os modelos construídos, o modelo selecionado foi o construído com 40 intervalos e 25 partículas e, a sua respectiva curva de calibração empregando 10 variáveis latentes é apresentada na Figura 49.

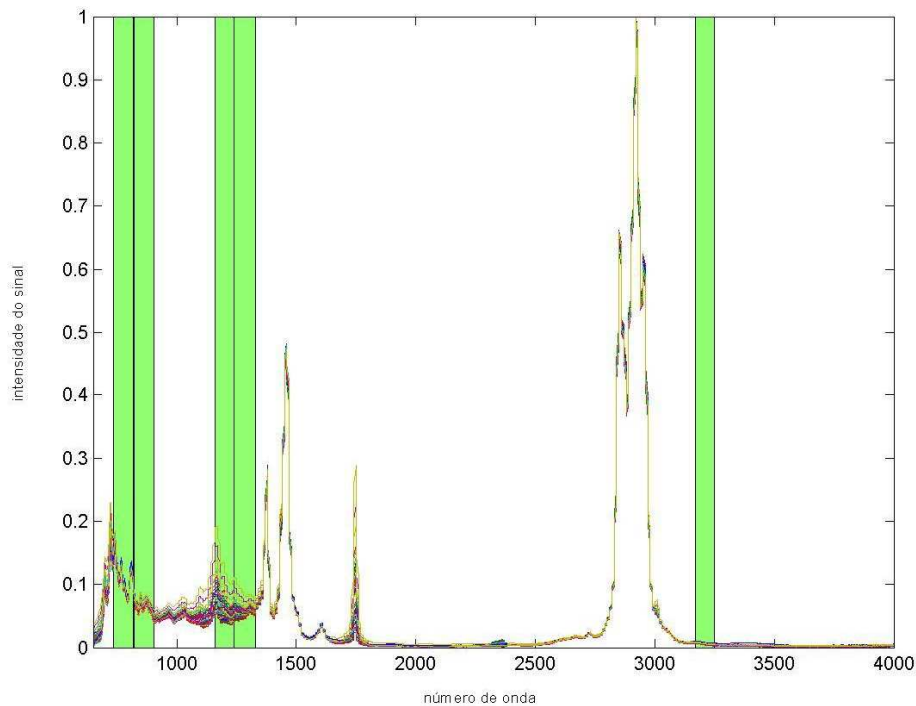


Figura 48 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

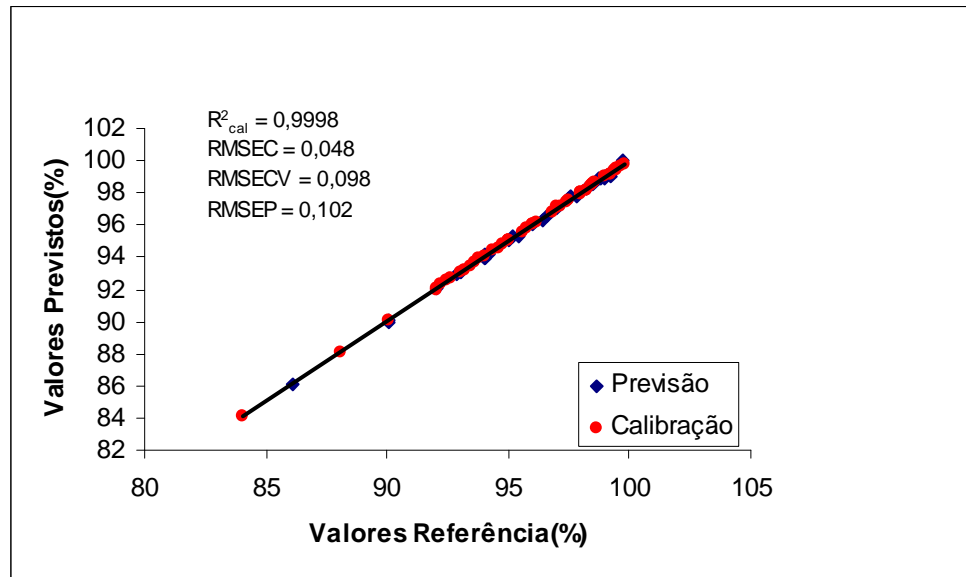


Figura 49 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO com 40 intervalos e 25 partículas. Fonte: elaborado pela autora.

5.3.4 Resultados obtidos para a quantificação do óleo diesel aplicando o método iPLS-PSO (CC)

Quando aplicado o método proposto iPLS-PSO (CC), para quantificação do óleo diesel, as regiões espectrais selecionadas para construção do modelo compreendem de 762 a 1207 cm^{-1} com lacunas, de 2719 a 2773 cm^{-1} , de 3223 a 3332 cm^{-1} e de 3446 a 3500 cm^{-1} , visualizadas na Figura 50. O modelo selecionado, dentre todos os construídos para este método, foi o proposto com 60 intervalos e 25 partículas e, a sua respectiva curva de calibração empregando 10 variáveis latentes é apresentada na Figura 51.

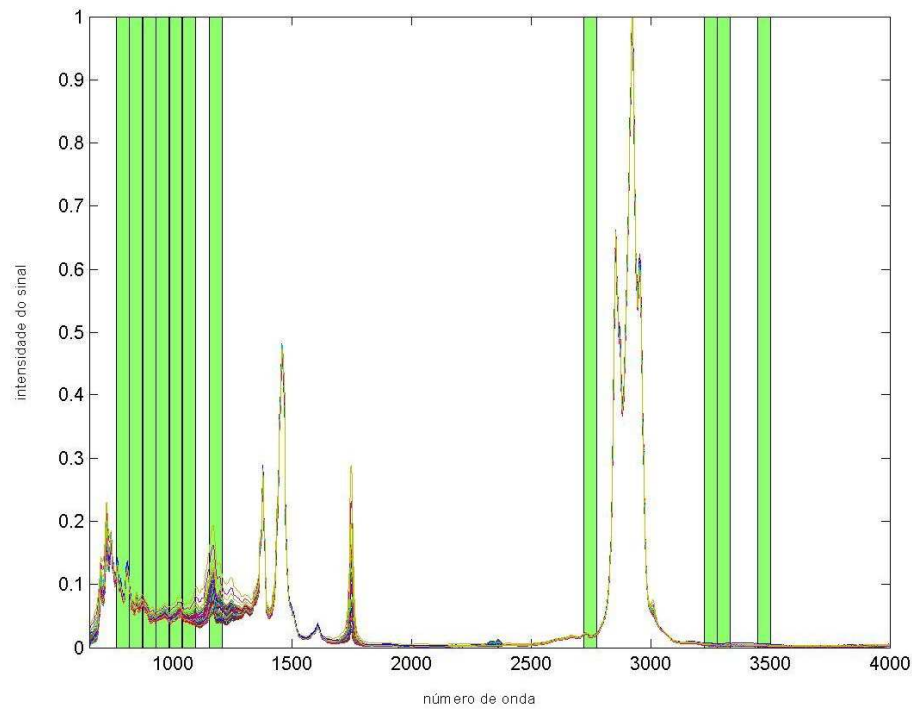


Figura 50 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.
Fonte: elaborado pela autora.

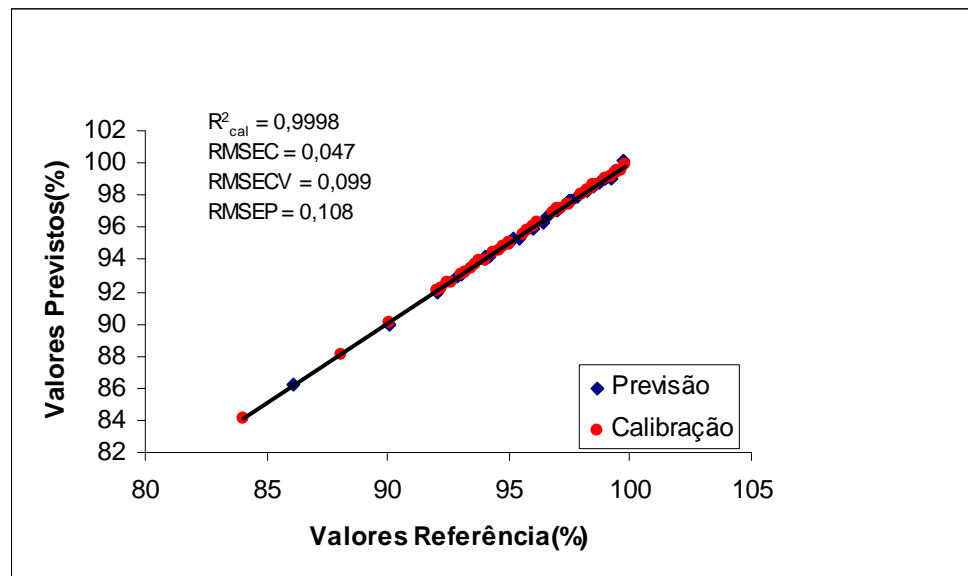


Figura 51 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.3.5 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DD)

Nos testes para quantificação do óleo diesel, aplicando o método iPLS-PSO (DD), foram selecionadas as regiões espectrais que compreendem de 904 a 1327 cm^{-1} com lacunas, de 2092 a 2175 cm^{-1} , de 2426 a 2756 cm^{-1} , de 3089 a 3170 cm^{-1} , de 3504 a 3585 cm^{-1} e de 3753 a 3834 cm^{-1} para a construção do modelo de regressão, visualizadas na Figura 52. O modelo selecionado, dentre todos os construídos para este método, foi o proposto com 40 intervalos e 100 partículas e, a sua respectiva curva de calibração utilizando 8 variáveis latentes é apresentada na Figura 53.

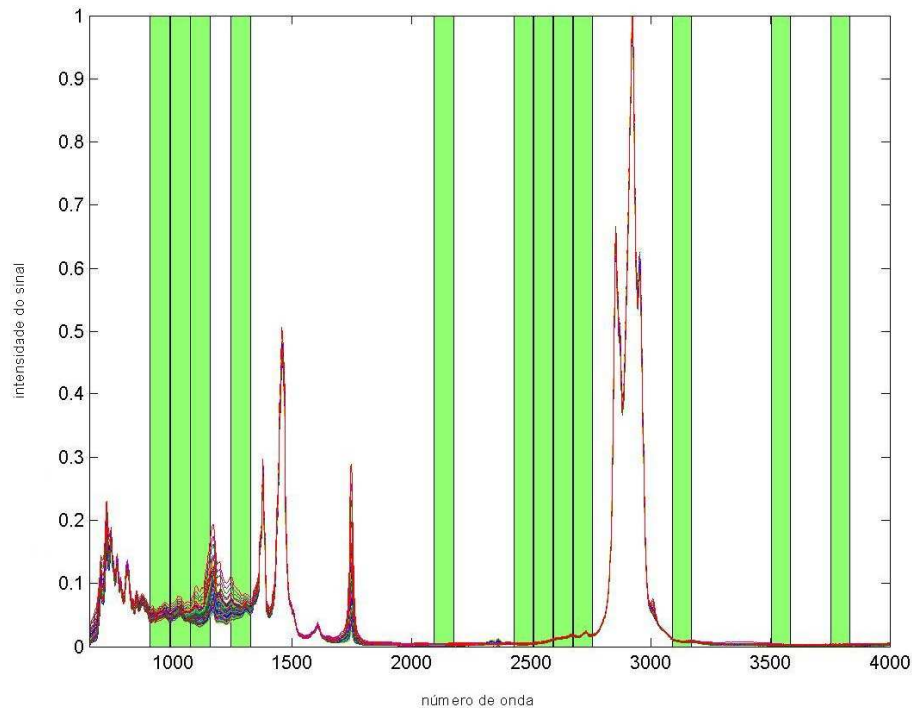


Figura 52 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

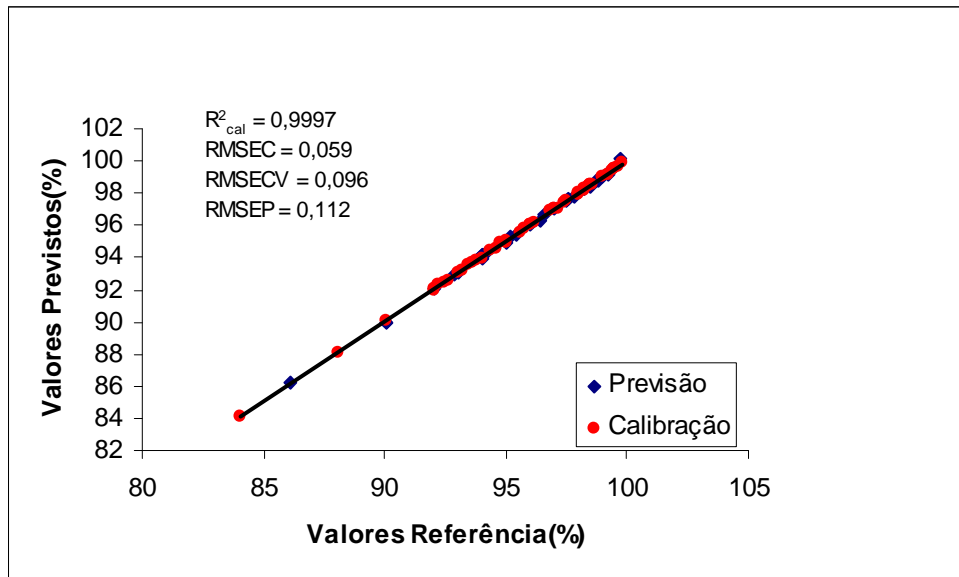


Figura 53 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas. Fonte: elaborado pela autora.

5.3.6 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (CD)

Já a aplicação do método iPLS-PSO (CD) para a quantificação do óleo diesel nas misturas, selecionou as regiões espectrais nas faixas de 904 a 1327 cm^{-1} com lacunas, de 1753 a 1836 cm^{-1} , de 2260 a 2507 cm^{-1} , de 2675 a 2756 cm^{-1} e, de 3753 a 3834 cm^{-1} visualizadas na Figura 54. O modelo selecionado, dentre todos os construídos para este método, foi o proposto com 40 intervalos e 50 partículas e, a sua respectiva curva de calibração utilizando 10 variáveis latentes é apresentada na Figura 55.

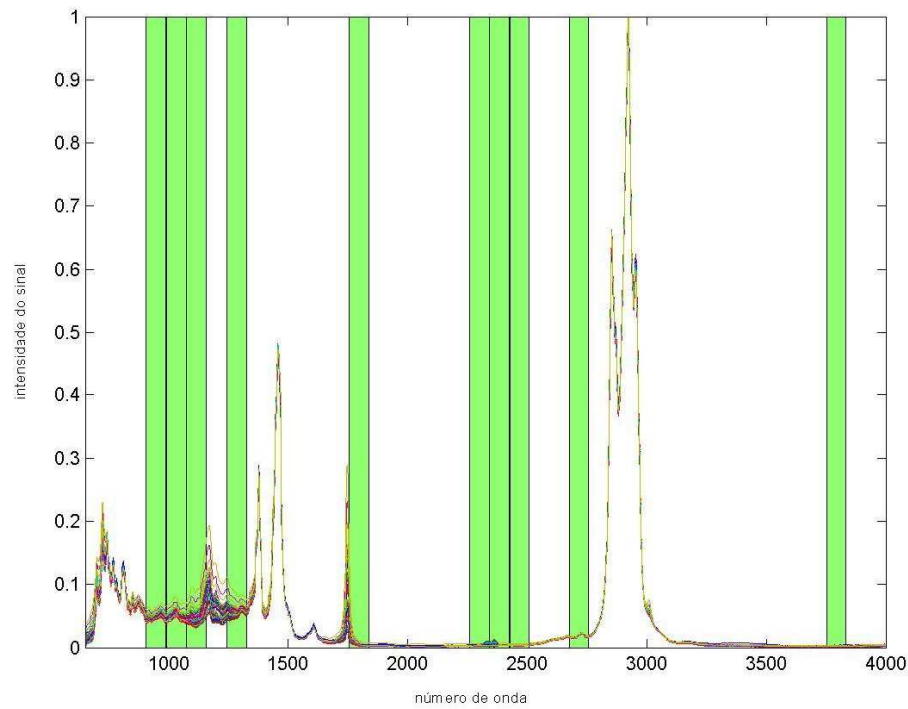


Figura 54 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

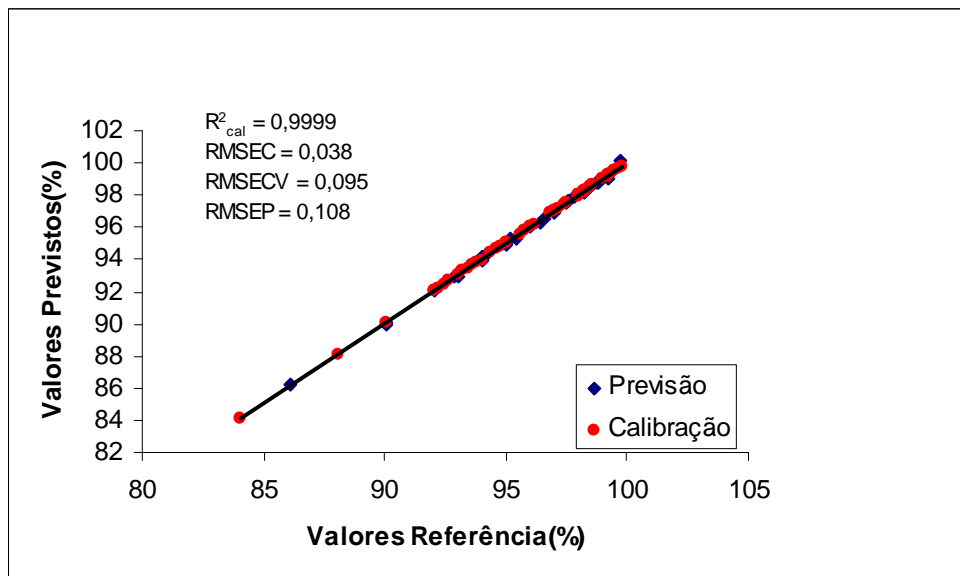


Figura 55 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

5.3.7 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DC)

E finalmente o método iPLS-PSO (DC) foi aplicado para a quantificação do óleo diesel e o modelo selecionado, dentre todos os construídos, foi o que selecionou a região espectral de 735 a 1327 cm^{-1} , visualizada na Figura 56, com 40 intervalos e 25 partículas. A sua respectiva curva de calibração para quantificação do óleo diesel empregando 10 variáveis latentes é apresentada na Figura 57.

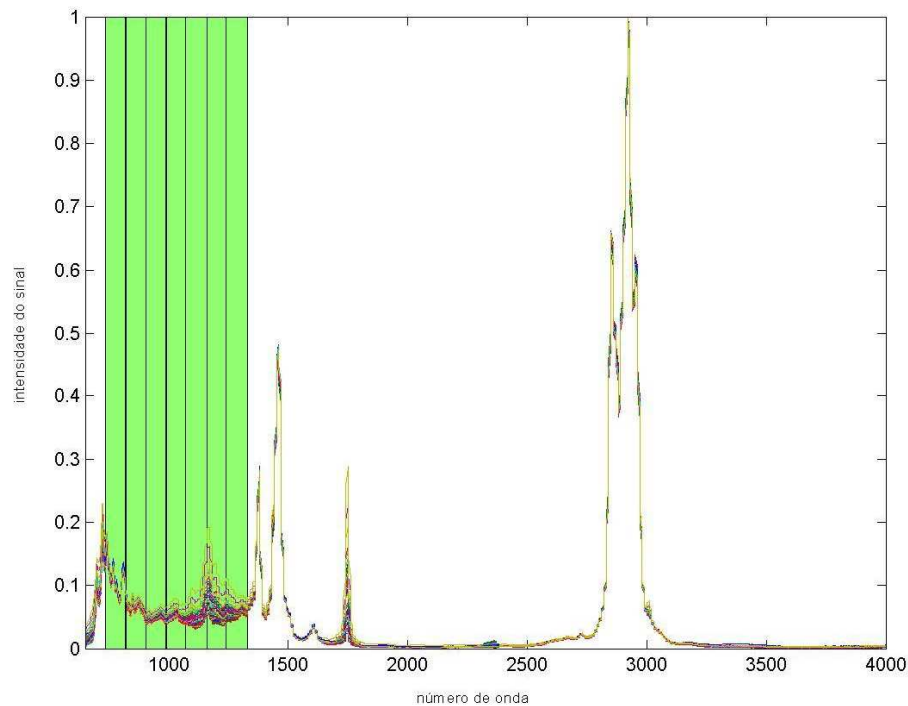


Figura 56 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

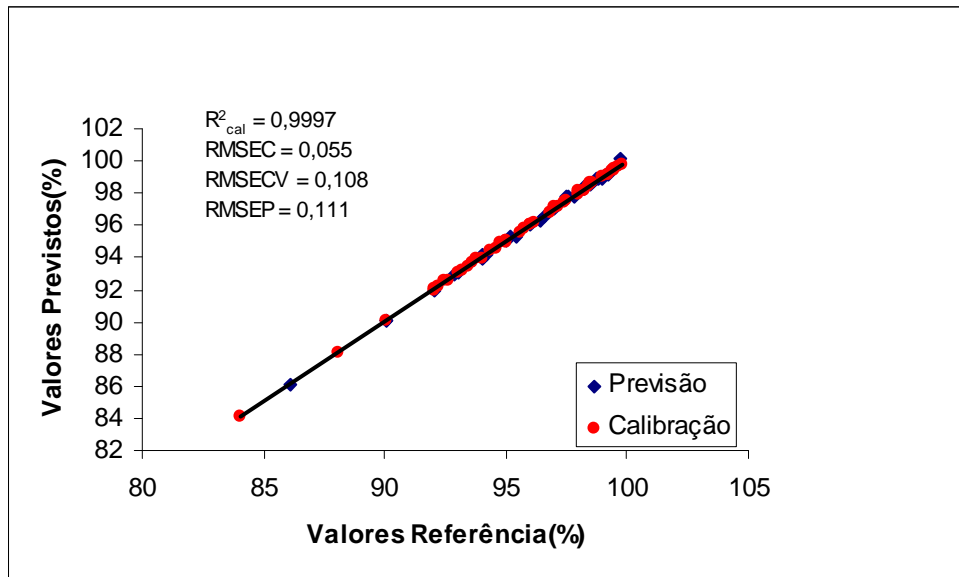


Figura 57 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 50 partículas. Fonte: elaborado pela autora.

5.3.8 Considerações sobre os modelos de óleo diesel

Com base nos resultados mostrados acima, foi possível observar que com baixos erros de previsão (RMSEP) todos os modelos construídos foram capazes de quantificar o óleo diesel. Analisando os resultados obtidos, empregando o PLS sobre o espectro inteiro da mistura e os resultados encontrados empregando as soluções propostas pelo método enxame de partículas, é possível perceber uma melhora na capacidade preditiva dos modelos otimização pela meta-heurística, porém não tão representativa como no caso das determinações de biodiesel e óleo vegetal.

Os modelos selecionados pelas otimizações aqui propostas para a quantificação do óleo diesel, apresentaram valores de RMSEP entre 0,102 a 0,112, enquanto os valores RMSECV ficaram na amplitude de 0,095 a 0,108.

No que diz respeito a comparação dos modelos construídos com as regiões selecionadas pela algoritmo híbrido e com toda região espectral da mistura não houve grandes melhoras, ficando evidenciado que a quantidade de informação capturada pelas variáveis latentes no modelo que emprega toda a região espectral é adequado para a quantificação deste componente da mistura.

Já para os modelos utilizando o algoritmo PSO nenhuma região característica dos hidrocarbonetos (região de 2850-3000 cm^{-1} , que corresponde a absorção de $\nu\text{C-H}$ referente aos estiramentos simétricos e assimétricos das ligações C-H de alcanos; 1465 cm^{-1} que corresponde a absorção de δCH_2 ; e 1375-1450 cm^{-1} , que corresponde a absorção de δCH_3) que constituem o diesel foi selecionada. Observando as regiões selecionadas na construção dos modelos a maior parte dessas correspondem à região do infravermelho médio denominada de impressão digital.

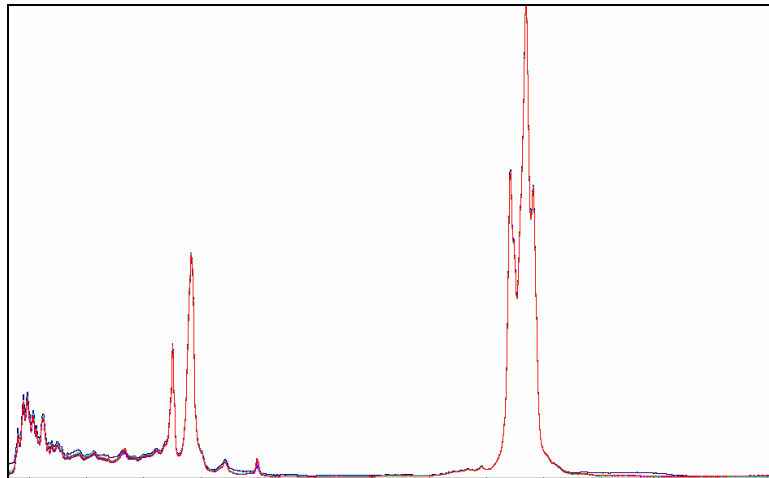
5.4 Considerações finais sobre os modelos de biodiesel, óleo vegetal e óleo diesel

Os resultados obtidos para os modelos visando a determinação, principalmente do biodiesel e da possível presença de óleo vegetal em misturas biodiesel/óleo vegetal/diesel, indicaram que não é possível utilizar uma única região e/ou a combinações das regiões sugeridas pela proposta de norma que garanta a determinação destes dentro do erro especificado pela proposta de norma. Conforme, destacado anteriormente a banda da carbonila (C=O) é insuficiente para a construção de um modelo eficiente e sensível as possíveis adulterações.

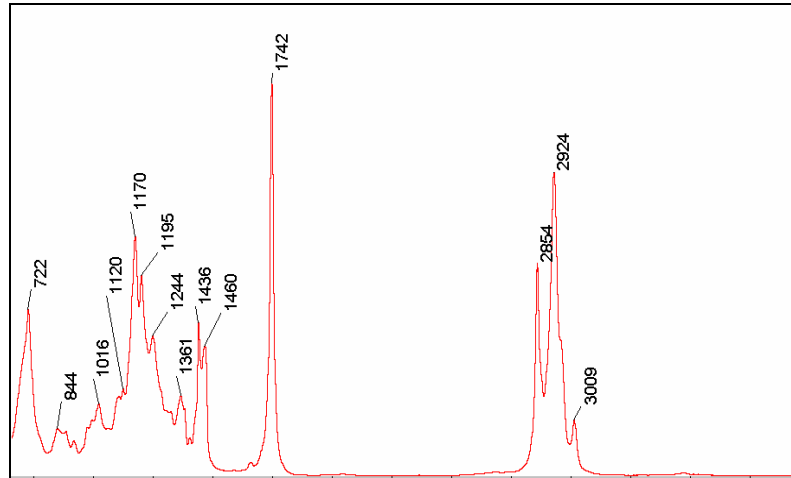
Comparando os espectros do biodiesel de soja e do óleo de soja que deu origem a este biodiesel (Figura 58) fica claro que a banda da carbonila é muito semelhante (sinal em 1742 cm^{-1} para o biodiesel e em 1743 cm^{-1} para o óleo) não sendo suficiente para gerar informação de discriminação, mesmo quando aplicada regressão multivariada.

Por outro lado, ratificando os resultados encontrados pela otimização PSO, a região referente ao C-O apresenta modificações bem mais representativas. Se por um lado o principal sinal no espectro do biodiesel encontra-se em 1170 cm^{-1} , por outro lado este mesmo sinal desloca-se para 1160 cm^{-1} para o óleo vegetal. O mesmo é verificado para a faixa de compreende o intervalo entre 1100-1450 cm^{-1} , região para a qual diversos modelos otimizados apresentam soluções tanto para o biodiesel, quanto para diesel com baixos erros dentro da faixa preconizada pela proposta de norma.

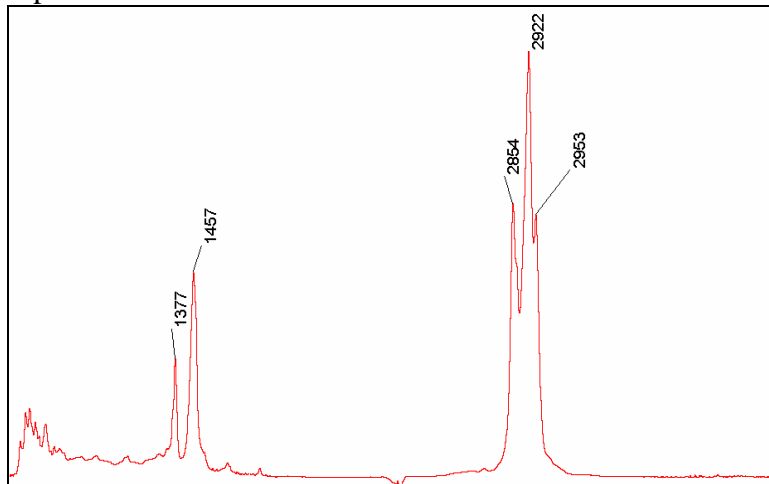
Uma outra forma de justificar o maior erro encontrado na previsão do biodiesel através das faixas espectrais sugeridas pela proposta de norma, é pela observação das figuras 59 e 60, onde as barras representam o erro de previsão do biodiesel nas amostras do conjunto de previsão, enquanto que as retas representam a concentração de óleo presentes nesta amostras.



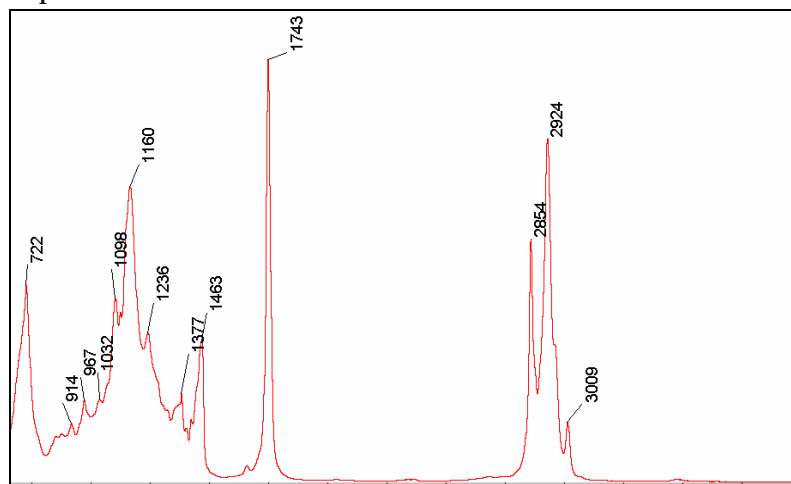
espectro de todas as misturas



espectro do biodiesel



espectro do óleo diesel



espectro do óleo vegetal

Figura 58 – Esquema contendo espectro de todas as misturas, do óleo diesel, do biodiesel e do óleo vegetal.
Fonte: elaborado pela autora.

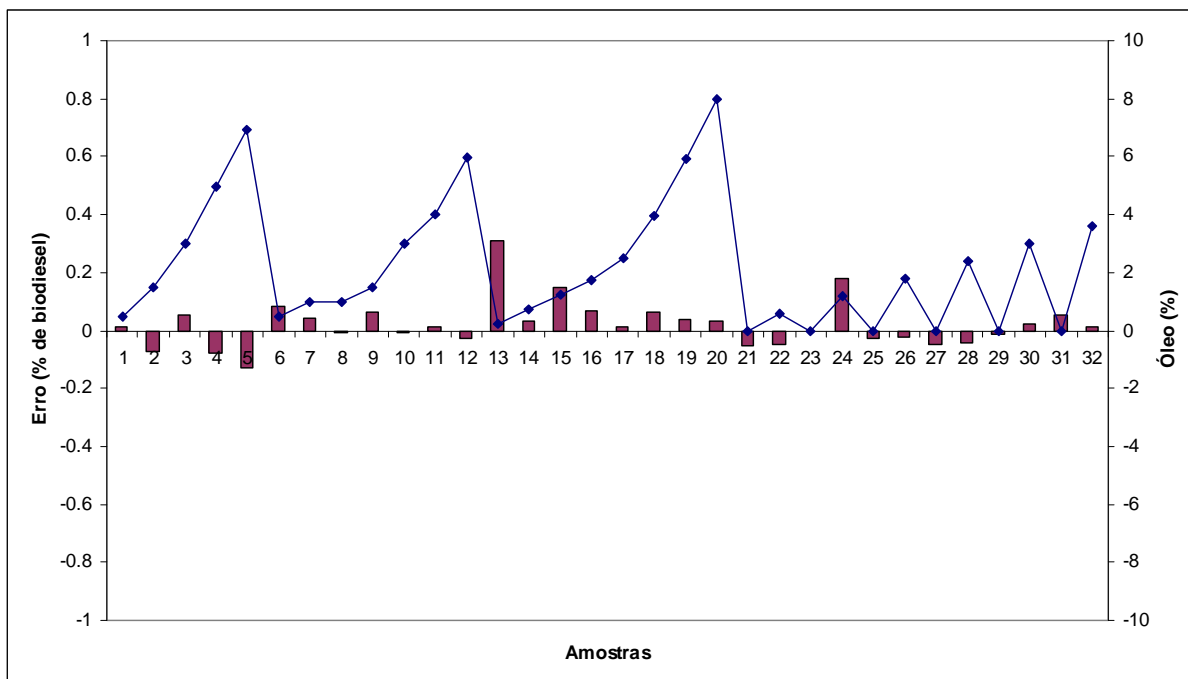


Figura 59 – Erros de previsão para a quantificação do biodiesel, usando as regiões propostas pelo método iPLS-PSO com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

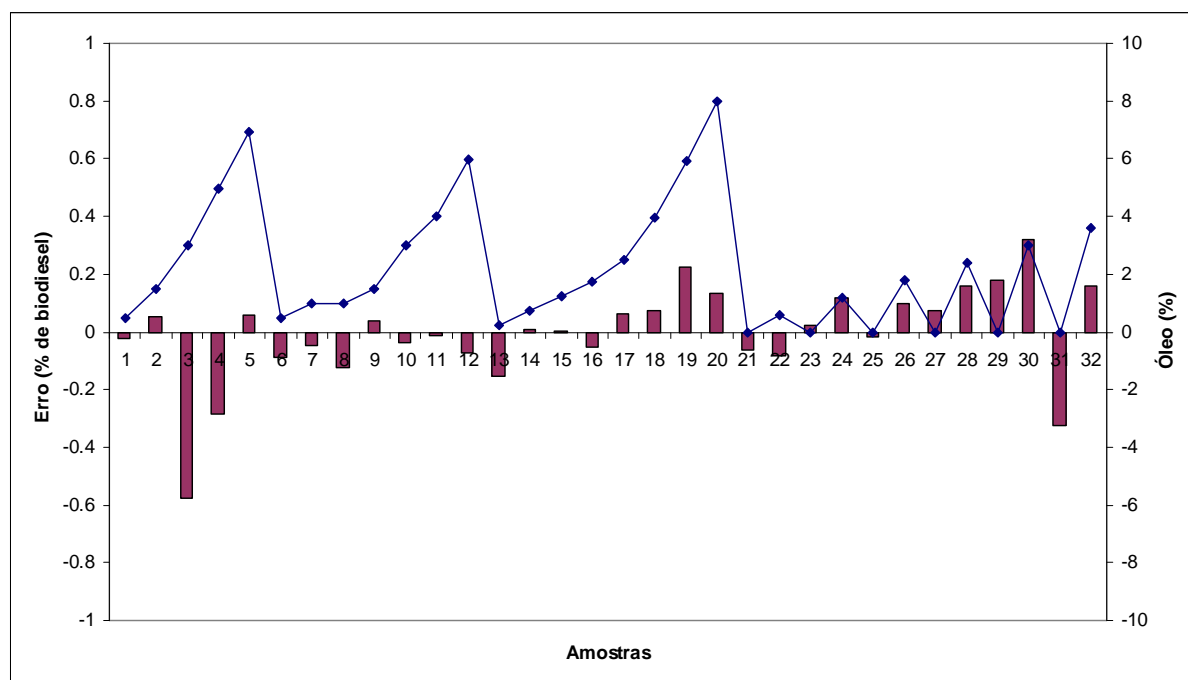


Figura 60 – Erros de previsão para a quantificação do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT.
Fonte: elaborado pela autora.

O modelo empregando somente as regiões sugeridas pela proposta de norma ABNT, de uma forma geral, apresenta maior erro para as amostras que apresentam maior concentração de óleo, sugerindo que contribuição do sinal referente ao óleo esteja afetando a capacidade preditiva deste modelo. Já para o modelo otimizado isto não se verifica.

Em trabalhos empregando regressão multivariada, a exemplo do trabalho de Soares, Rezende e Fortes (2007), foi utilizada uma metodologia para quantificação de misturas biodiesel/diesel empregando o método de seleção de variáveis stepwise e forward apresentando erros de previsão na faixa de 0,65 a 1,40 %, entretanto neste caso não foi possível aplicar sempre o mesmo método de seleção de variáveis para todas as misturas de biodiesel e óleo de soja cru, onde cada grupo de amostra foi estudado separadamente para desenvolver um modelo de calibração adequado. No método proposto nesta dissertação, o RMSEP do melhor modelo de biodiesel, foi de 0,83 % sendo capaz de modelar diferentes tipos de biodiesel (soja e palma) e ainda na presença do óleo de soja, constituindo-se num modelo mais robusto.

Ainda existem na literatura trabalhos empregando análise univariada, Guarieiro *et al.* (2005, 2008), onde não se considera a existência de óleo de soja estando sujeitos, quando da análise de amostras adulteradas, produzir resultados falsos positivos, não sendo estes modelos adequados na análise a campo.

Embora os resultados apresentados nas Tabelas 5, 6 e 7 não indiquem sobre ajuste nos modelos selecionados, em alguns casos o processo de otimização não convergiu de forma satisfatória gerando modelos com RMSECV menor que o RMSEP. A Figura 61 apresentou um exemplo do comportamento desejável para a qual o *Gbest* evolui de valores maiores RMSECV e RMSEP (pontos a direita) para valores menores de RMSECV e RMSEP (pontos a esquerda). Já na Figura 62 a evolução resulta na diminuição do RMSECV sem que necessariamente o RMSEP tenha o mesmo comportamento.

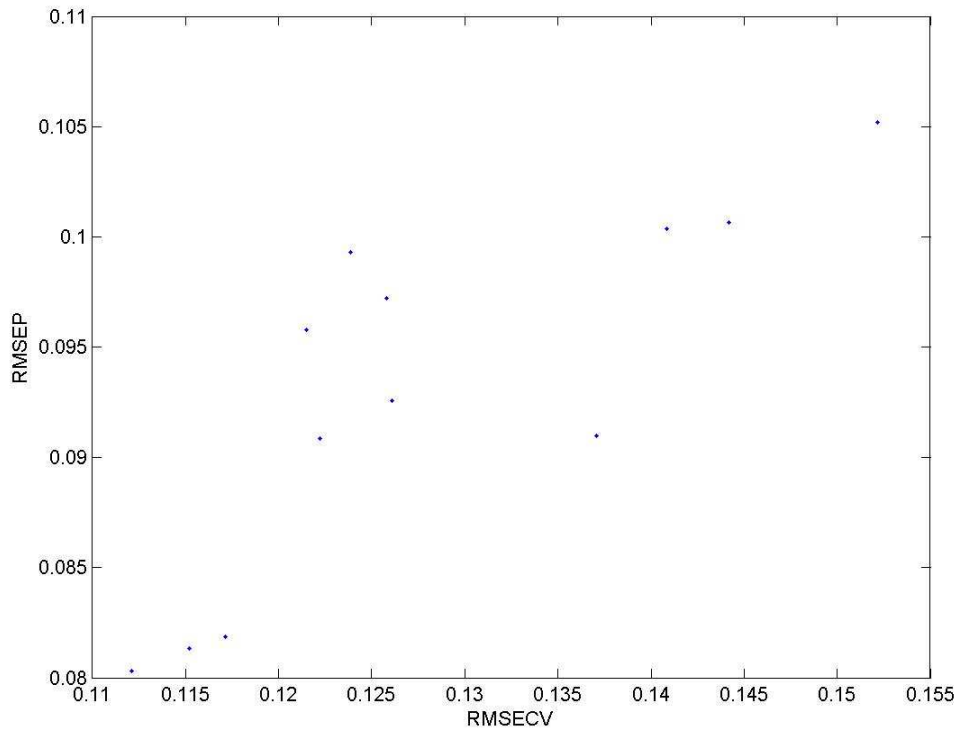


Figura 61 – Valores de RMSECV *versus* RMSEP para diferentes *Gbest* para quantificação do biodiesel, usando o método iPLS-PSO (CD) com 60 intervalos e 25 partículas.

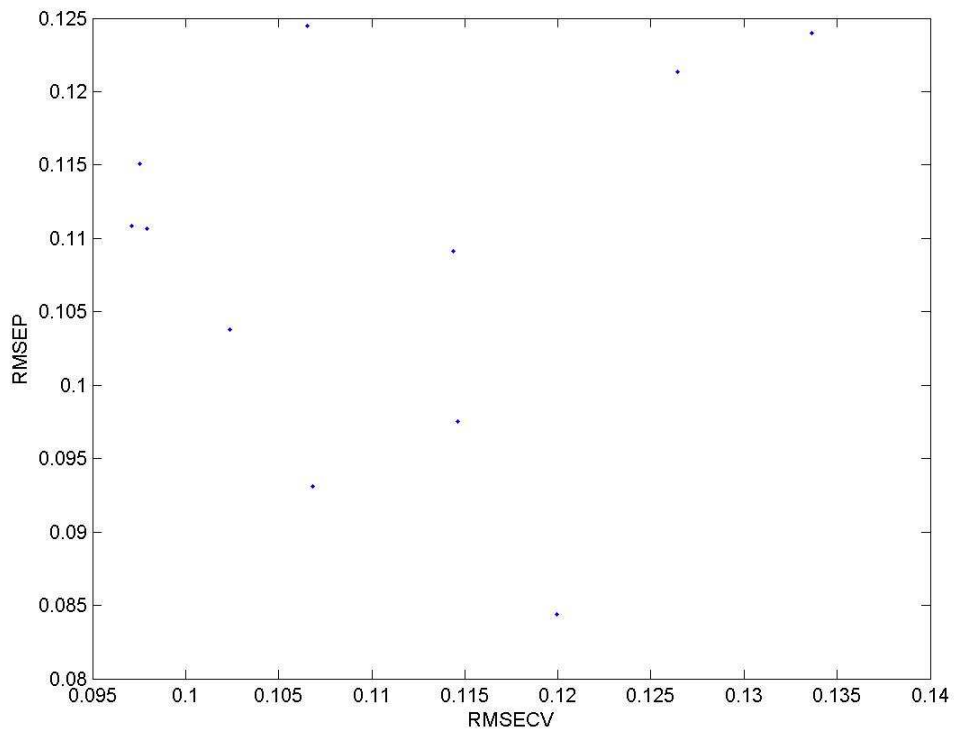


Figura 62 – Valores de RMSECV *versus* RMSEP para diferentes *Gbest* para a quantificação do biodiesel, usando o método iPLS-PSO (CD) com 40 intervalos e 100 partículas.

6 CONCLUSÕES

A presente dissertação mostrou que o método enxame de partículas em conjunto com o método de regressão por mínimos quadrados parciais por intervalos é capaz de desenvolver modelos de previsão de biodiesel, óleo vegetal e óleo diesel através dos espectros de reflexão total atenuada das misturas biodiesel/óleo diesel.

Independente do número de partículas empregados na otimização por enxame de partículas, bem como do número de intervalos para os quais os espectros foram subdivididos, bons modelos com capacidade preditiva maiores ou equivalentes ao emprego do espectro total foram obtidos para qualquer um dos componentes das misturas. Entretanto no que se refere ao modelo para a previsão de biodiesel empregando as regiões sugeridas pela proposta de norma este apresentou baixo desempenho que pode ser atribuído à interferência dos sinais do óleo que ocorrem nas mesmas faixas de absorção do biodiesel, não sendo suficiente a utilização destas faixas espectrais para a construção de modelos multivariados. O comportamento encontrado neste caso é bastante semelhante a interferência quando são construídos modelos univariados a partir de dados espectroscópicos.

As principais regiões selecionadas pelos algoritmos PSO podem ser facilmente atribuídas as principais absorções referentes a composição do biodiesel e do óleo vegetal, destacando-se a região referente ao estiramento C-O para quantificação do biodiesel e do óleo vegetal nas misturas, bem como a região referente a impressão digital para a quantificação do óleo diesel. Nestes modelos a seleção da região de absorção da carbonila (C=O) nem sempre foi necessária na obtenção de modelos com baixos erros de previsão.

Os modelos otimizados a partir dos algoritmos PSO também permitiram a identificação e quantificação da adulteração proveniente de óleos vegetais, enquanto que a proposta de norma é restrita a quantificação do biodiesel no diesel.

Finalmente, a implementação de métodos híbridos, nos quais são empregados conceitos de algoritmos clássicos em combinação com algoritmos heurísticos, permitem a resolução de problemas combinatoriais de forma adequada, somando-se as vantagens de ambas as técnicas para alcançar soluções satisfatórias em tempo real.

7 TRABALHOS FUTUROS

Abaixo se relaciona algum dos futuros trabalhos que podem ser desenvolvidos, com base nesta dissertação:

- ampliar a composição das misturas, utilizando outros tipos de biodiesel como o biodiesel de sebo e de óleo de frituras;
- ampliar a faixa de quantificação do biodiesel até 30 %(v/v), conforme o segundo modelo sugerido pela proposta de norma ABNT;
- comparar os resultados obtidos nesta dissertação com a PSO, com outros algoritmos híbridos ou não, como por exemplo algoritmo genético e busca tabu e;
- empregar algoritmos de compactação de dados, como as wavelets, tentando diminuir a quantidade de variáveis e o tempo de processamento, para construção dos modelos.

REFERÊNCIAS

- AGRAFIOTIS, D. K.; CEDEÑO, W. *Feature selection for structure-activity correlation using binary particle swarms*. Journal of Medicinal Chemistry, v. 45, n. 5, p. 1098-1107, fev. 2002.
- BARBOSA, Luiz Cláudio de Almeida. Espectroscopia no infravermelho na caracterização de compostos orgânicos. 1. ed. Viçosa: Editora Universidade Federal de Viçosa, 2007.
- BORGES NETO, Waldomiro. Parâmetros de qualidade de lubrificantes e óleo de oliva através de espectroscopia vibracional, calibração multivariada e seleção de variáveis. 2005. 130 f. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, 2005.
- BORIN, A.; POPPI, R. J. *Application of mid infrared spectroscopy and iPLS for the quantification of contaminants in lubricating oil*. Vibrational Spectroscopy, v. 37, n. 1, p. 27-32, jan. 2005.
- BRASIL. Lei n. 11.097, de 13 de janeiro de 2005. Disponível em: <http://www.planalto.gov.br/ccivil/_Ato2004-2006/2005/Lei/L11097.htm>. Acesso em: 02 fev. 2008.
- CEPEA. [s.n]. Quanto custa produzir biodiesel. 2006. Disponível em: <<http://cepea.esalq.usp.br/comunicacao/>>. Acesso em 25 ago. 2006.
- CLERC, M. *The swarm and the queen: towards a deterministic and adaptive particle swarm optimization*. In: Proc. Congress on Evolutionary Computation 1999, p. 1951-1957, 1999.
- CLERC, M.; KENNEDY, J. *The particle swarm: explosion, stability, and convergence in a multi-dimensional complex space*. IEEE Transactions on Evolutionary Computation, v. 6, n. 1 p. 58-73, fev. 2002.
- CHATTERJEE, A.; SIARRY, P. *Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization*. Computers and Operations Research, v. 33, n. 3, p. 859-871, mar. 2006.
- COSTA FILHO, P. A. da; POPPI, R. J. Aplicação de algoritmos genéticos na seleção de variáveis em espectroscopia no infravermelho médio. Determinação simultânea de glicose, maltose e frutose. Química Nova, São Paulo, v. 25, n. 1, p. 46-52, jan./fev. 2002.
- COSTA NETO, P. R. *et al.* Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em frituras. Química Nova, São Paulo, v. 23, n. 4, p. 531-537, jul./ago. 2000.
- DORADO, M. P. *et al.* *Kinetic parameters affecting the alkali-catalyzed transesterification process of used olive oil*. Energy e Fuels, v. 18, n. 5, p. 1457-1462, 2004.
- EBERHART, R.C.; HU, X. *Human tremor analysis using particle swarm optimization*. In: Proc. Congress on Evolutionary Computation 1999, p. 1927-1930, 1999.
- ENCINAR, J. M.; GONZÁLEZ, J. F.; RODRÍGUEZ-REINARES, A. *Biodiesel used frying oil. Variables affecting the yields and characteristics of the biodiesel*. Industrial & Engineering Chemistry Research, v. 44, n. 15, p. 5491-5499, 2005.
- ENGELBRECHT, Andries P. *Fundamentals of computational swarm intelligence*. England:

John Wiley & Sons, 2005.

FERRÃO, M. F. Técnicas de reflexão no infravermelho aplicadas na análise de alimentos. Tecno-Lógica., Santa Cruz do Sul, v. 5, n.1, p. 63-85, jan./jun. 2001.

FERRÃO, Marco Florês. Aplicação de técnicas espectroscópica de reflexão no infravermelho no controle de qualidade de farinha de trigo. 2000. 219 f. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, 2000.

FERREIRA, E. C. *et al.* Análise exploratória dos teores de constituintes inorgânicos em sucos e refrigerantes de uva. Eclética Química, v. 27, n. especial, p. 77-90, 2002.

FERREIRA, Marcos Antonio Gonzalez. Uso de análise multivariada para determinar a composição de gasolina a partir da espectroscopia FT-Raman e cromatografia gasosa. 2005. 109 f. Dissertação (Programa de Pós Graduação em Química Inorgânica) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2005.

GERIS, R. *et al.* Biodiesel de Soja – Reação de transesterificação para aulas práticas de química orgânica. Química Nova, São Paulo, v. 30, n. 5, p. 1369-1373, set./out. 2007.

GUARIEIRO, L. L. N. *et al.* Desenvolvimento de metodologia para quantificação das misturas biodiesel:diesel por infravermelho, 2005. Disponível em <www.portalabpg.org.br/PDPetro/3/trabalhos/IBP0250_05.pdf> . Acesso em: 10 dez. 2007.

GUARIEIRO, L. L. N. *et al.* Metodologia analítica para quantificar o teor de biodiesel na mistura biodiesel: diesel utilizando espectroscopia na região do infravermelho. Química Nova, São Paulo, v. 31, n. 2, p. 421-426, 2008.

GURDEN, S. P.; WESTERHUIS, J. A.; SMILDE, A. K. *Monitoring of batch processes using spectroscopy*. American Institute of Chemical Engineers. v. 48, n. 10, p. 2283-2297, out. 2002.

HAIR, J. F. *et al.* Análise Multivariada de Dados. 5. ed. Porto Alegre: Editora Bookman, 2005.

HAGEMAN *et al.* *Wavelength selection with tabu search*. Journal of Chemometrics. v. 17, n. 8-9, p. 427-437, 2003.

KHALIL, T. *et al.* *A Binary particle swarm optimization for optimal placement and sizing of capacitor banks in radial distribution feeders with distorted substation voltages*. In: Conference on Artificial Intelligence and Machine Learning, p. 129-135, 2006.

KENNEDY, J.; EBERHART, R. C. *A discrete binary version of the particle swarm algorithm*. In: Proc. of the 1997 Conference on Systems, Man, and Cybernetics, p. 4104-4109, 1997.

KENNEDY, J.; EBERHART, R. C. *Particle swarm optimization*. In: The 1995 IEEE International Conference on Neural Networks, p. 1942-1948, 1995.

KENNEDY, J.; SPEARS, W. M. *Matching algorithms to problems: an experimental test of the particle swarm and some genetic algorithms on the multimodal problem generator*. In: Proc. of the 1998 International Conference on Evolutionary Computation, p. 78-83, 1998.

KENNEDY, J.; EBERHART, R.C.; SHI, Y. *Swarm intelligence*. San Francisco: Morgan Kaufmann Publishers, 2001.

- KNOTHE Gerhard. *Analyzing biodiesel: standards and other methods*. Journal of the American Oil Chemists Society, v. 83, n. 10, p. 823-833, out. 2006.
- KONZEN, P. H. de A. *et al.* Otimização de métodos de controle de qualidade de fármacos usando algoritmo genético e busca tabu. Pesquisa Operacional, Rio de Janeiro, v. 23, n. 1, p. 189-207, jan./abr. 2003.
- LEARDI, R.; NORGAARD, L. *Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions*. Journal of Chemometrics, v. 18, n. 11, p. 486-497, 2004.
- MÜLLER, Viviane. Otimização de layouts industriais através do método enxame de partículas. 2007. 79 f. Dissertação (Programa de Pós-Graduação em Sistemas e Processos Industriais - Mestrado) - Universidade de Santa Cruz do Sul, Santa Cruz do Sul, 2007.
- MEDEIROS, José Antonio Carlos Canedo. Enxame de partículas como ferramenta de otimização em problemas complexos de engenharia nuclear. 2005. 108 f. Tese (Programa de Pós-Graduação de Engenharia – Doutorado) Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.
- MENDES, Rui. *Population topologies and their influence in particle swarm performance*. 2004. 189 f. Dissertation al fulfillment of the requirements for the degree of Philosophiæ Doctor at the University of Minho, Portugal, 2004.
- MENESES Anderson Alvarenga de Moura. Otimização por enxame de partículas aplicado ao problema combinatório da recarga de um reator nuclear. 2005. 100f. Dissertação (Programa de Pós-Graduação de Engenharia – Mestrado) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.
- MORGANO, M. A. *et al.* Aplicação de análise exploratória na diferenciação de vegetais. *Brazilian Journal of Food Technology*, São Paulo, v.2, n. 1/2, p. 73-79, 1999.
- NASCIMENTO, M. da G.; COSTA NETO, P. R.; MAZZUCO, L. M. Biotransformação de óleos e gorduras. *Biotechnology Ciência & Desenvolvimento*, Brasília, n.19, p. 28-31, mar./abr. 2001.
- NORGAARD, L. *et al.* *Interval partial least-square regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy*. *Applied Spectroscopy*, v. 54, n. 3, p. 413-419, 2000.
- OLIVEIRA, Flávia Carneiro da Cunha. Modelos de calibração multivariada associados à espectroscopia vibracional para análise de misturas diesel-óleos vegetais. 2006. 108 f. Dissertação (Programa de Pós-Graduação do Instituto de Química - Mestrado) - Universidade de Brasília, 2006.
- OMRAN M.G.H.; ENGELBRECHT A. P.; SALMAN, A. 2005. *Dynamic clustering using particle swarm optimization with application in unsupervised image classification*. *Transactions on Engineering, Computing And Technology*, v. 9, p. 199-204, nov. 2005.
- PARADKAR, M. M.; IRUDAYARAJ, J. *Rapid determination of caffeine content in soft drinks using FTIR-ATR spectroscopy*. *Food Chemistry*, v. 78, n. 2. p. 261-266, ago. 2002.
- PARISOTTO, G. *et al.* Análise exploratória aplicada no estudo de medicamentos contendo piroxicam. *Brazilian Journal of Pharmaceutical Sciences*, São Paulo, v. 41, n. 4. p. 499-505, out./dez. 2005.

PARSOPOULO, K. E.; VRAHATIS, M. N. *Particle swarm optimization method in multiobjective problems*. In: Proc. of the 2002 Symposium on Applied Computing, p. 603-607, 2002.

PIMENTEL, M. F. *et al.* *Determination of biodiesel content when blended with mineral diesel fuel using infrared spectroscopy and multivariate calibration*. Microchemical Journal, v. 82, n. 2, p. 201-206, abr. 2006.

PERAM, T., VEERAMACHANENI, K.; MOHAN, C. K. Fitness-distance-ratio based particle swarm optimization. In: Proc. of the 2003 Swarm Intelligence Symposium, p. 174-181, 2003.

POMEROY, Paul. *An introduction to particle swarm optimization*. 2003 Disponível em: <<http://www.adaptiveview.com/articles/ipsoprnt.html>>. Acesso em: 03 jan. 2006.

PRADO, J. R. do; SARAMAGO, S. F. .P. Otimização por colônia de partículas. FAMAT em Revista, Uberlândia, n. 4, p. 87-103, abr. 2005.

QI SHEN *et al.*; *Particle swarm algorithm trained neural network for qsar studies of inhibitors of platelet-derived growth factor receptor phosphorylation*. European journal of pharmaceutical sciences, v. 28, n. 5, p. 369-376, 2006.

RADICH, Antony. *Biodiesel performance, costs, and use*. 2004. Disponível em: <<http://www.eia.doe.gov/oiaf/analysispaper/biodiesel/>>. Acesso em: 10 de mar. 2008.

SABIN, Gustavo Post. Otimização de modelos de regressão multivariados empregando métodos de seleção de variáveis. 2007. 111 f. Dissertação (Programa de Pós-Graduação em Sistemas e Processos Industriais - Mestrado) - Universidade de Santa Cruz do Sul, Santa Cruz do Sul, 2007.

SACORAGUE, Luiz Alexandre. Avaliação de diferentes regiões do espectro do infravermelho próximo na determinação de parâmetros de qualidade de combustíveis empregando ferramentas quimiométricas. 2004. 235 f. Tese (Departamento de Química Analítica - Doutorado) – Universidade Estadual de Campinas, Campinas, 2004.

SCHWAAD, Marcio. Avaliação de algoritmos heurísticos de otimização em problemas de estimação de parâmetros. 2005. 156 f. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

SENA, M. M. de *et al.* Avaliação do uso de métodos quimiométricos em análise de solos. Química Nova, São Paulo, v. 23, n. 4, p. 547-556, jul./ago. 2000.

SHI, Y.; EBERHART, R. C. *A modified particle swarm optimizer*. In: Proc. of the IEEE International Conference on Evolutionary Computation, p. 69-73, 1998a.

SHI, Y.; EBERHART, R. C. *Parameter selection in particle swarm optimization*. In: Evolutionary Programming VII. Lecture Notes in Computer Science, p. 591-600, 1998b.

SHI, Y., EBERHART, R. *Empirical Study of Particle Swarm Optimization*. In: Proc. Congress on Evolutionary Computation, p. 1945-1950, 1999.

SHI, Y; EBERHART, R. C. *Fuzzy adaptive particle swarm optimization*. In: Proceedings of the IEEE Congress on Evolutionary Computation, p. 101-106, 2001.

SOARES, I. P.; REZENDE, T. F.; FORTES, I. C. P. Calibração multivariada com seleção de variáveis em amostras de biodiesel adulteradas com óleo de soja cru utilizando dados de

espectroscopia de infravermelho com transformada de fourier. 2007 Disponível em www.biodiesel.gov.br/docs/congresso2007/caracterizacao/10.pdf. Acesso em: 03 de mar. 2008.

WANG, Z. *et al.* *Particle swarm optimization and neural network application for QSAR*. In: Parallel and Distributed Processing Symposium, abr. 2004.

YOSHIDA, S.; YOSHIDA H. *Non destructive analyses of unsaturated fatty acid species in dietary oils by attenuated total reflectance with Fourier transform IR spectroscopy*. Journal Biopolymers, v. 70, n. 4, p. 604-613, nov. 2003.

ZHENG, Y. *et al.* *Empirical study of particle swarm optimizer with an increasing inertia weight*. In: Proceeding of the IEEE Congress on Evolutionary Computation, v. 1, p. 221-226, 2003.

ANEXOS

ANEXO A - Resultados completos das implementações do iPLS-PSO para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	58
	2 ^a	0,120	0,091	63
	3 ^a	0,120	0,091	115
20 intervalos 50 partículas	1 ^a	0,120	0,091	46
	2 ^a	0,120	0,091	14
	3 ^a	0,120	0,091	72
20 intervalos 100 partículas	1 ^a	0,120	0,091	17
	2 ^a	0,120	0,091	56
	3 ^a	0,120	0,091	79
40 intervalos 25 partículas	1 ^a	0,103	0,134	179
	2 ^a	0,101	0,115	241
	3 ^a	0,101	0,120	239
40 intervalos 50 partículas	1 ^a	0,100	0,137	165
	2 ^a	0,115	0,083	59
	3 ^a	0,102	0,120	214
40 intervalos 100 partículas	1 ^a	0,100	0,137	142
	2 ^a	0,100	0,137	205
	3 ^a	0,101	0,107	207
60 intervalos 25 partículas	1 ^a	0,128	0,092	166
	2 ^a	0,114	0,089	238
	3 ^a	0,106	0,090	205
60 intervalos 50 partículas	1 ^a	0,106	0,099	156
	2 ^a	0,101	0,103	198
	3 ^a	0,100	0,091	231
60 intervalos 100 partículas	1 ^a	0,095	0,088	247
	2 ^a	0,107	0,121	168
	3 ^a	0,123	0,089	150

ANEXO A - Resultados completos das implementações do iPLS-PSO (CC) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	61
	2 ^a	0,120	0,091	141
	3 ^a	0,120	0,091	61
20 intervalos 50 partículas	1 ^a	0,120	0,091	94
	2 ^a	0,120	0,091	79
	3 ^a	0,120	0,091	76
20 intervalos 100 partículas	1 ^a	0,120	0,091	90
	2 ^a	0,120	0,091	70
	3 ^a	0,120	0,091	47
40 intervalos 25 partículas	1 ^a	0,103	0,903	152
	2 ^a	0,105	0,085	108
	3 ^a	0,106	0,106	117
40 intervalos 50 partículas	1 ^a	0,101	0,105	172
	2 ^a	0,103	0,093	164
	3 ^a	0,103	0,086	213
40 intervalos 100 partículas	1 ^a	0,106	0,092	159
	2 ^a	0,103	0,116	231
	3 ^a	0,113	0,075	106
60 intervalos 25 partículas	1 ^a	0,099	0,114	169
	2 ^a	0,108	0,097	158
	3 ^a	0,095	0,108	227
60 intervalos 50 partículas	1 ^a	0,100	0,114	147
	2 ^a	0,109	0,097	207
	3 ^a	0,116	0,115	164
60 intervalos 100 partículas	1 ^a	0,101	0,109	227
	2 ^a	0,092	0,106	213
	3 ^a	0,087	0,106	107

ANEXO A - Resultados completos das implementações do iPLS-PSO (DD) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	12
	2 ^a	0,120	0,091	12
	3 ^a	0,120	0,094	23
20 intervalos 50 partículas	1 ^a	0,120	0,091	9
	2 ^a	0,120	0,091	85
	3 ^a	0,120	0,091	44
20 intervalos 100 partículas	1 ^a	0,120	0,091	22
	2 ^a	0,120	0,091	56
	3 ^a	0,120	0,091	33
40 intervalos 25 partículas	1 ^a	0,103	0,101	247
	2 ^a	0,104	0,111	230
	3 ^a	0,101	0,101	231
40 intervalos 50 partículas	1 ^a	0,103	0,917	203
	2 ^a	0,103	0,109	202
	3 ^a	0,108	0,096	160
40 intervalos 100 partículas	1 ^a	0,103	0,086	171
	2 ^a	0,097	0,111	193
	3 ^a	0,105	0,085	138
60 intervalos 25 partículas	1 ^a	0,103	0,101	247
	2 ^a	0,104	0,111	230
	3 ^a	0,101	0,101	231
60 intervalos 50 partículas	1 ^a	0,103	0,917	203
	2 ^a	0,103	0,109	202
	3 ^a	0,108	0,096	160
60 intervalos 100 partículas	1 ^a	0,102	0,091	236
	2 ^a	0,105	0,095	203
	3 ^a	0,087	0,106	333

ANEXO A - Resultados completos das implementações do iPLS-PSO (CD) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	41
	2 ^a	0,120	0,091	76
	3 ^a	0,120	0,091	54
20 intervalos 50 partículas	1 ^a	0,120	0,091	33
	2 ^a	0,120	0,091	105
	3 ^a	0,120	0,091	70
20 intervalos 100 partículas	1 ^a	0,120	0,091	44
	2 ^a	0,120	0,091	33
	3 ^a	0,120	0,091	14
40 intervalos 25 partículas	1 ^a	0,105	0,085	176
	2 ^a	0,113	0,115	242
	3 ^a	0,103	0,091	201
40 intervalos 50 partículas	1 ^a	0,115	0,083	52
	2 ^a	0,115	0,083	85
	3 ^a	0,115	0,083	104
40 intervalos 100 partículas	1 ^a	0,113	0,063	63
	2 ^a	0,097	0,111	169
	3 ^a	0,115	0,083	32
60 intervalos 25 partículas	1 ^a	0,121	0,074	160
	2 ^a	0,124	0,105	163
	3 ^a	0,112	0,080	118
60 intervalos 50 partículas	1 ^a	0,110	0,101	224
	2 ^a	0,108	0,097	106
	3 ^a	0,114	0,094	127
60 intervalos 100 partículas	1 ^a	0,117	0,078	95
	2 ^a	0,095	0,097	193
	3 ^a	0,096	0,102	159

ANEXO A - Resultados completos das implementações do iPLS-PSO (DC) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	115
	2 ^a	0,120	0,091	96
	3 ^a	0,120	0,091	90
20 intervalos 50 partículas	1 ^a	0,120	0,091	77
	2 ^a	0,120	0,091	25
	3 ^a	0,120	0,091	82
20 intervalos 100 partículas	1 ^a	0,120	0,091	9
	2 ^a	0,120	0,091	106
	3 ^a	0,120	0,091	51
40 intervalos 25 partículas	1 ^a	0,112	0,101	51
	2 ^a	0,110	0,092	104
	3 ^a	0,103	0,101	62
40 intervalos 50 partículas	1 ^a	0,116	0,098	155
	2 ^a	0,105	0,103	162
	3 ^a	0,102	0,101	82
40 intervalos 100 partículas	1 ^a	0,105	0,104	59
	2 ^a	0,102	0,108	60
	3 ^a	0,103	0,093	123
60 intervalos 25 partículas	1 ^a	0,126	0,093	163
	2 ^a	0,113	0,101	53
	3 ^a	0,125	0,081	102
60 intervalos 50 partículas	1 ^a	0,112	0,096	51
	2 ^a	0,110	0,147	30
	3 ^a	0,110	0,101	169
60 intervalos 100 partículas	1 ^a	0,107	0,101	185
	2 ^a	0,103	0,116	48
	3 ^a	0,104	0,099	89

ANEXO B - Resultados completos das implementações do iPLS-PSO para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,079	0,095	114
	2 ^a	0,082	0,103	30
	3 ^a	0,082	0,103	48
20 intervalos 50 partículas	1 ^a	0,079	0,095	65
	2 ^a	0,079	0,095	127
	3 ^a	0,079	0,095	133
20 intervalos 100 partículas	1 ^a	0,079	0,095	67
	2 ^a	0,079	0,095	49
	3 ^a	0,079	0,095	80
40 intervalos 25 partículas	1 ^a	0,084	0,098	80
	2 ^a	0,083	0,092	232
	3 ^a	0,081	0,096	238
40 intervalos 50 partículas	1 ^a	0,080	0,010	178
	2 ^a	0,075	0,105	217
	3 ^a	0,078	0,103	158
40 intervalos 100 partículas	1 ^a	0,076	0,102	219
	2 ^a	0,076	0,107	158
	3 ^a	0,080	0,107	231
60 intervalos 25 partículas	1 ^a	0,075	0,116	171
	2 ^a	0,075	0,106	166
	3 ^a	0,076	0,121	240
60 intervalos 50 partículas	1 ^a	0,073	0,101	234
	2 ^a	0,077	0,102	145
	3 ^a	0,074	0,112	249
60 intervalos 100 partículas	1 ^a	0,077	0,110	238
	2 ^a	0,079	0,110	85
	3 ^a	0,076	0,101	166

ANEXO B - Resultados completos das implementações do iPLS-PSO (CC) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,079	0,095	197
	2 ^a	0,082	0,103	77
	3 ^a	0,079	0,095	165
20 intervalos 50 partículas	1 ^a	0,079	0,095	56
	2 ^a	0,079	0,095	49
	3 ^a	0,079	0,095	98
20 intervalos 100 partículas	1 ^a	0,079	0,095	38
	2 ^a	0,079	0,095	49
	3 ^a	0,079	0,095	98
40 intervalos 25 partículas	1 ^a	0,083	0,101	72
	2 ^a	0,088	0,089	167
	3 ^a	0,079	0,105	238
40 intervalos 50 partículas	1 ^a	0,078	0,109	120
	2 ^a	0,076	0,101	174
	3 ^a	0,079	0,109	126
40 intervalos 100 partículas	1 ^a	0,081	0,105	87
	2 ^a	0,075	0,105	201
	3 ^a	0,080	0,094	201
60 intervalos 25 partículas	1 ^a	0,077	0,108	215
	2 ^a	0,069	0,118	201
	3 ^a	0,080	0,106	221
60 intervalos 50 partículas	1 ^a	0,083	0,106	193
	2 ^a	0,080	0,111	231
	3 ^a	0,080	0,102	130
60 intervalos 100 partículas	1 ^a	0,085	0,108	37
	2 ^a	0,077	0,100	144
	3 ^a	0,077	0,114	213

ANEXO B - Resultados completos das implementações do iPLS-PSO (DD) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,079	0,095	205
	2 ^a	0,081	0,096	156
	3 ^a	0,089	0,109	157
20 intervalos 50 partículas	1 ^a	0,079	0,095	91
	2 ^a	0,082	0,103	21
	3 ^a	0,079	0,095	174
20 intervalos 100 partículas	1 ^a	0,079	0,095	100
	2 ^a	0,079	0,095	119
	3 ^a	0,081	0,094	234
40 intervalos 25 partículas	1 ^a	0,070	0,130	186
	2 ^a	0,080	0,101	221
	3 ^a	0,082	0,988	241
40 intervalos 50 partículas	1 ^a	0,074	0,105	241
	2 ^a	0,085	0,099	244
	3 ^a	0,078	0,110	208
40 intervalos 100 partículas	1 ^a	0,075	0,102	156
	2 ^a	0,082	0,089	161
	3 ^a	0,076	0,107	97
60 intervalos 25 partículas	1 ^a	0,080	0,107	227
	2 ^a	0,079	0,101	185
	3 ^a	0,083	0,104	216
60 intervalos 50 partículas	1 ^a	0,079	0,103	246
	2 ^a	0,076	0,103	246
	3 ^a	0,080	0,130	238
60 intervalos 100 partículas	1 ^a	0,074	0,110	249
	2 ^a	0,075	0,100	250
	3 ^a	0,071	0,101	236

ANEXO B - Resultados completos das implementações do iPLS-PSO (CD) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,081	0,096	84
	2 ^a	0,081	0,094	81
	3 ^a	0,082	0,103	105
20 intervalos 50 partículas	1 ^a	0,079	0,095	59
	2 ^a	0,079	0,095	99
	3 ^a	0,082	0,103	69
20 intervalos 100 partículas	1 ^a	0,079	0,095	101
	2 ^a	0,079	0,095	93
	3 ^a	0,079	0,095	90
40 intervalos 25 partículas	1 ^a	0,080	0,099	147
	2 ^a	0,074	0,105	147
	3 ^a	0,082	0,118	250
40 intervalos 50 partículas	1 ^a	0,082	0,102	189
	2 ^a	0,078	0,109	181
	3 ^a	0,078	0,106	222
40 intervalos 100 partículas	1 ^a	0,075	0,110	191
	2 ^a	0,075	0,102	94
	3 ^a	0,070	0,130	95
60 intervalos 25 partículas	1 ^a	0,084	0,106	243
	2 ^a	0,075	0,096	243
	3 ^a	0,070	0,115	199
60 intervalos 50 partículas	1 ^a	0,075	0,103	240
	2 ^a	0,076	0,107	231
	3 ^a	0,085	0,112	56
60 intervalos 100 partículas	1 ^a	0,074	0,105	229
	2 ^a	0,077	0,105	228
	3 ^a	0,075	0,105	229

ANEXO B - Resultados completos das implementações do iPLS-PSO (DC) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,082	0,103	209
	2 ^a	0,085	0,115	49
	3 ^a	0,082	0,103	62
20 intervalos 50 partículas	1 ^a	0,079	0,095	57
	2 ^a	0,079	0,095	24
	3 ^a	0,079	0,095	177
20 intervalos 100 partículas	1 ^a	0,079	0,095	44
	2 ^a	0,079	0,095	37
	3 ^a	0,079	0,095	44
40 intervalos 25 partículas	1 ^a	0,082	0,102	221
	2 ^a	0,087	0,109	47
	3 ^a	0,085	0,122	182
40 intervalos 50 partículas	1 ^a	0,082	0,107	32
	2 ^a	0,076	0,101	206
	3 ^a	0,079	0,100	59
40 intervalos 100 partículas	1 ^a	0,085	0,090	59
	2 ^a	0,085	0,105	209
	3 ^a	0,078	0,102	58
60 intervalos 25 partículas	1 ^a	0,082	0,101	226
	2 ^a	0,085	0,117	41
	3 ^a	0,091	0,116	32
60 intervalos 50 partículas	1 ^a	0,084	0,101	15
	2 ^a	0,082	0,102	225
	3 ^a	0,084	0,110	213
60 intervalos 100 partículas	1 ^a	0,077	0,110	177
	2 ^a	0,080	0,109	38
	3 ^a	0,085	0,113	119

ANEXO C - Resultados completos das implementações do iPLS-PSO para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,107	0,131	103
	2 ^a	0,111	0,139	136
	3 ^a	0,118	0,147	222
20 intervalos 50 partículas	1 ^a	0,105	0,125	188
	2 ^a	0,105	0,125	72
	3 ^a	0,107	0,131	62
20 intervalos 100 partículas	1 ^a	0,107	0,131	35
	2 ^a	0,105	0,125	217
	3 ^a	0,103	0,136	76
40 intervalos 25 partículas	1 ^a	0,116	0,143	85
	2 ^a	0,098	0,102	91
	3 ^a	0,096	0,128	193
40 intervalos 50 partículas	1 ^a	0,097	0,117	215
	2 ^a	0,098	0,127	195
	3 ^a	0,100	0,129	185
40 intervalos 100 partículas	1 ^a	0,097	0,110	171
	2 ^a	0,101	0,125	130
	3 ^a	0,091	0,114	127
60 intervalos 25 partículas	1 ^a	0,103	0,108	250
	2 ^a	0,100	0,131	240
	3 ^a	0,103	0,108	151
60 intervalos 50 partículas	1 ^a	0,097	0,130	152
	2 ^a	0,095	0,140	169
	3 ^a	0,089	0,136	99
60 intervalos 100 partículas	1 ^a	0,107	0,119	187
	2 ^a	0,100	0,120	232
	3 ^a	0,088	0,122	201

ANEXO C - Resultados completos das implementações do iPLS-PSO (CC) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,106	0,129	122
	2 ^a	0,112	0,115	220
	3 ^a	0,106	0,134	69
20 intervalos 50 partículas	1 ^a	0,111	0,139	129
	2 ^a	0,107	0,131	59
	3 ^a	0,112	0,140	75
20 intervalos 100 partículas	1 ^a	0,107	0,131	99
	2 ^a	0,105	0,125	61
	3 ^a	0,101	0,128	96
40 intervalos 25 partículas	1 ^a	0,100	0,112	147
	2 ^a	0,091	0,114	141
	3 ^a	0,116	0,141	99
40 intervalos 50 partículas	1 ^a	0,091	0,114	119
	2 ^a	0,099	0,132	96
	3 ^a	0,096	0,112	232
40 intervalos 100 partículas	1 ^a	0,095	0,147	178
	2 ^a	0,102	0,127	183
	3 ^a	0,097	0,124	223
60 intervalos 25 partículas	1 ^a	0,099	0,142	222
	2 ^a	0,100	0,131	224
	3 ^a	0,099	0,108	133
60 intervalos 50 partículas	1 ^a	0,099	0,127	204
	2 ^a	0,092	0,111	148
	3 ^a	0,091	0,140	227
60 intervalos 100 partículas	1 ^a	0,084	0,140	179
	2 ^a	0,089	0,124	216
	3 ^a	0,087	0,126	204

ANEXO C - Resultados completos das implementações do iPLS-PSO (DD) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,112	0,128	43
	2 ^a	0,111	0,131	3
	3 ^a	0,112	0,122	155
20 intervalos 50 partículas	1 ^a	0,112	0,122	37
	2 ^a	0,101	0,128	62
	3 ^a	0,103	0,136	211
20 intervalos 100 partículas	1 ^a	0,103	0,136	20
	2 ^a	0,103	0,136	45
	3 ^a	0,107	0,131	79
40 intervalos 25 partículas	1 ^a	0,139	0,117	173
	2 ^a	0,101	0,115	49
	3 ^a	0,101	0,114	216
40 intervalos 50 partículas	1 ^a	0,091	0,114	120
	2 ^a	0,102	0,113	81
	3 ^a	0,091	0,114	110
40 intervalos 100 partículas	1 ^a	0,091	0,114	79
	2 ^a	0,096	0,112	241
	3 ^a	0,091	0,114	41
60 intervalos 25 partículas	1 ^a	0,099	0,121	166
	2 ^a	0,109	0,128	60
	3 ^a	0,108	0,133	201
60 intervalos 50 partículas	1 ^a	0,098	0,141	208
	2 ^a	0,090	0,138	245
	3 ^a	0,090	0,118	250
60 intervalos 100 partículas	1 ^a	0,104	0,119	157
	2 ^a	0,086	0,116	131
	3 ^a	0,095	0,130	206

ANEXO C - Resultados completos das implementações do iPLS-PSO (CD) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,131	0,131	80
	2 ^a	0,147	0,147	76
	3 ^a	0,111	0,139	83
20 intervalos 50 partículas	1 ^a	0,101	0,128	108
	2 ^a	0,112	0,122	75
	3 ^a	0,101	0,128	93
20 intervalos 100 partículas	1 ^a	0,111	0,131	48
	2 ^a	0,112	0,122	65
	3 ^a	0,112	0,122	27
40 intervalos 25 partículas	1 ^a	0,091	0,114	80
	2 ^a	0,108	0,111	18
	3 ^a	0,938	0,123	165
40 intervalos 50 partículas	1 ^a	0,091	0,114	74
	2 ^a	0,095	0,108	141
	3 ^a	0,095	0,119	126
40 intervalos 100 partículas	1 ^a	0,087	0,116	175
	2 ^a	0,091	0,114	43
	3 ^a	0,091	0,114	28
60 intervalos 25 partículas	1 ^a	0,100	0,119	185
	2 ^a	0,110	0,136	97
	3 ^a	0,106	0,123	49
60 intervalos 50 partículas	1 ^a	0,105	0,131	135
	2 ^a	0,108	0,134	142
	3 ^a	0,089	0,128	167
60 intervalos 100 partículas	1 ^a	0,106	0,124	82
	2 ^a	0,100	0,109	23
	3 ^a	0,084	0,112	209

ANEXO C - Resultados completos das implementações do iPLS-PSO (DC) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,108	0,131	52
	2 ^a	0,112	0,121	207
	3 ^a	0,111	0,139	159
20 intervalos 50 partículas	1 ^a	0,107	0,131	27
	2 ^a	0,111	0,139	101
	3 ^a	0,101	0,128	135
20 intervalos 100 partículas	1 ^a	0,101	0,127	93
	2 ^a	0,111	0,139	53
	3 ^a	0,105	0,125	53
40 intervalos 25 partículas	1 ^a	0,108	0,111	233
	2 ^a	0,109	0,126	66
	3 ^a	0,116	0,143	81
40 intervalos 50 partículas	1 ^a	0,110	0,116	211
	2 ^a	0,102	0,113	203
	3 ^a	0,096	0,118	51
40 intervalos 100 partículas	1 ^a	0,103	0,121	194
	2 ^a	0,105	0,116	59
	3 ^a	0,986	0,131	35
60 intervalos 25 partículas	1 ^a	0,100	0,131	46
	2 ^a	0,103	0,126	150
	3 ^a	0,109	0,128	106
60 intervalos 50 partículas	1 ^a	0,109	0,128	121
	2 ^a	0,099	0,145	57
	3 ^a	0,101	0,153	192
60 intervalos 100 partículas	1 ^a	0,100	0,121	236
	2 ^a	0,098	0,128	194
	3 ^a	0,108	0,118	231

**UNIVERSIDADE DE SANTA CRUZ DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS
E PROCESSOS INDUSTRIAIS – MESTRADO
ÁREA DE CONCENTRAÇÃO EM
CONTROLE E OTIMIZAÇÃO DE PROCESSOS**

Marlize Rizzi

**APLICAÇÃO DO MÉTODO ENXAME DE PARTÍCULAS NA OTIMIZAÇÃO DE MODELOS
DE REGRESSÃO MULTIVARIADA APLICADOS NA QUANTIFICAÇÃO DA MISTURA
BODIESEL/ÓLEO VEGETAL/DIESEL**

Santa Cruz do Sul, março de 2008.

Marlize Rizzi

**APLICAÇÃO DO MÉTODO ENXAME DE PARTÍCULAS NA OTIMIZAÇÃO DE MODELOS
DE REGRESSÃO MULTIVARIADA APLICADOS NA QUANTIFICAÇÃO DA MISTURA
BODIESEL/ÓLEO VEGETAL/DIESEL**

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

Orientador: Prof. Dr. João Carlos Furtado

Co-orientador: Prof. Dr. Marco Flôres Ferrão

Santa Cruz do Sul, março de 2008.

AGRADECIMENTOS

Inicialmente agradeço a Deus pela vida, e por estar sempre comigo.

Aos meus pais Avelino e Geni, por me darem carinho, amor, educação e força em todos os momentos de minha vida. Amo vocês.

Ao futuro papai e meu amor Éverton Camargo da Silva pela paciência, compreensão e incentivo. Amo você.

Ao meu filho Gabriel, que está por nascer, que soube colaborar para que juntos pudéssemos concluir esta etapa tão importante em nossas vidas.

Ao meu orientador, prof. Dr. João Carlos Furtado, pela amizade, pelos ensinamentos e a disponibilidade apresentada.

Ao meu co-orientado prof. Dr. Marco Flôres Ferrão pelas muitas orientações, pelo senso de direção, pela disponibilidade e pela oportunidade concedida.

A amiga Márcia Werlang pelas ótimas horas de estudos e a pela sua importante colaboração nesta dissertação.

As secretárias Janaina Ramires Haas e Camila Borges Corrêa pela eficiência e presteza sempre apresentada.

Ao mestre Gustavo Post Sabin, por ter compartilhado comigo seus conhecimentos nesta área.

A CAPES e a FAPERGS pelo financiamento do projeto “casadinhos” e, pela colaboração do Instituto de Química e do Laboratório de Combustíveis da Universidade Federal do Rio Grande do Sul - UFRGS.

A todos aqueles que embora não tenham sido mencionados contribuíram direta ou indiretamente para a realização deste trabalho.

“A imaginação é tudo. É uma prévia das próximas atrações da vida.”

Albert Einstein

RESUMO

A substituição gradativa dos produtos derivados de fontes não renováveis, por produtos de origem renováveis, vem sendo mundialmente discutida, especialmente quanto aos combustíveis. No Brasil, o governo oficializou a mistura biodiesel/diesel nos percentuais obrigatórios de 2% a partir de 2008. Considerando a necessidade de monitoramento da qualidade desta mistura, esta dissertação propõe o estudo e o uso do método enxame de partículas (*Particle Swarm Optimization* – PSO), juntamente com a técnica de regressão multivariada de mínimos quadrados parciais por intervalos (iPLS), visando constituir um algoritmo híbrido capaz de selecionar variáveis mais representativas do constituinte a ser quantificado, buscando gerar modelos de regressão mais robustos. Nesta pesquisa efetuou-se a quantificação de biodiesel, do óleo vegetal e do óleo diesel, utilizando toda a região espectral das misturas, as regiões espectrais sugeridas pela proposta de norma ABNT e as regiões escolhidas pelos algoritmos híbridos. Na quantificação do biodiesel o menor erro de previsão obtido foi através do método híbrido, sendo este erro 61,54% menor que o erro encontrado pelas regiões sugeridas pela proposta de norma ABNT e, 26,72 % menor que o erro encontrado utilizando toda a região espectral das misturas. Já, na quantificação do óleo vegetal o menor erro de previsão, obtido também foi pelo método híbrido, ficando 27,64% menor do que o erro encontrado pelas regiões sugeridas pela proposta de norma ABNT e, 38,62% menor do que erro encontrado utilizando toda a região espectral das misturas. Porém na quantificação do óleo diesel os métodos apresentaram desempenhos semelhantes. Com base nestes resultados, pode-se concluir que a utilização do método enxame de partículas juntamente com o método iPLS, foi capaz de otimizar as soluções, selecionando variáveis do espectro para construção de modelos mais robustos.

Palavras-chave: enxame de partículas, regressão por mínimos quadrados parciais, biodiesel, otimização combinatória, espectroscopia no infravermelho.

ABSTRACT

The gradual substitution of products coming from non-renewable sources by products of renewable origin has been globally discussed lately, especially regarding to the fuels. In Brazil, the government has declared official the biodiesel/diesel mixture in the obligatory percentile of 2% starting from 2008. This dissertation based on the need for monitoring such mixture's quality proposes the study and the use of the particles swarm method (Particle Swarm Optimization - PSO), together with the interval partial least-squares multivariate regression technique (iPLS). It presents as its aims to constitute a hybrid algorithm capable to select more representative variables of the representative to be quantified, seeking to produce more robust regression models. In this research, it was done the biodiesel, the vegetable oil, and the diesel oil quantification, using (a) the whole spectral area of the mixtures, (b) the spectral areas suggested by the norm proposal from the Brazilian Association of technical rules (ABNT), and (c) the areas chosen by the hybrid algorithms. In the biodiesel quantification, the smallest foreseen mistake obtained was through the hybrid method, being this mistake 61.54% smaller than the mistake found by the areas suggested by the norm proposal from ABNT and, 26.72% smaller than the found mistake using the whole spectral area of the mixtures. Yet in the vegetable oil quantification the smallest foreseen mistake, also obtained through the hybrid method, was 27.64% smaller than the mistake found by the areas suggested by the norm proposal from ABNT and, 38.62% smaller than mistake found using the whole spectral area of the mixtures. However, in the diesel oil quantification the methods have presented similar performance. Based on these results, it was possible to conclude that the use of particles swarm method in association with the method iPLS, was capable to optimize the solutions, selecting the spectrum's variables for the construction of more robust models.

Key-words: particles swarm, partial least-squares regression, biodiesel, combinatorial optimization, infrared spectroscopy.

LISTA DE ILUSTRAÇÕES

Figura 1 – Esquema de transesterificação de um triacilglicerídeo.....	19
Figura 2 – Representação da propagação da radiação infravermelho através do elemento de ATR.	22
Figura 3 - Ilustração do método de enxame de partículas.....	26
Figura 4 – Fluxograma das etapas executadas na parte experimental.	34
Figura 5 – Diagrama ternário para as misturas biodiesel/óleo vegetal/diesel.	35
Figura 6 – Distribuição das amostras de calibração e previsão para o biodiesel.....	39
Figura 7 – Distribuição das amostras de calibração e previsão para o óleo diesel.....	39
Figura 8 – Distribuição das amostras de calibração e previsão para o óleo vegetal.....	40
Figura 9 – Representação gráfica de duas partículas.	41
Figura 10 – Exemplo de uma partícula e representação dos seus intervalos selecionados .	42
Figura 11 – Região espectral utilizada na modelagem para a quantificação do biosiesel. ...	46
Figura 12 – Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel utilizando toda a região espectral das misturas.....	47
Figura 13 – Regiões espectrais sugeridas pela proposta de norma ABNT.....	47
Figura 14– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT.....	48
Figura 15 – Exemplo da evolução das partículas para a quantificação do biodiesel, usando o método iPLS-PSO (CD) com 40 intervalos e 100 partículas.	50
Figura 16 – Evolução das três execuções do iPLS-PSO com 20 intervalos e 50 partículas, para quantificação do biodiesel.....	51
Figura 17 – Evolução das três execuções do iPLS-PSO com 40 intervalos e 50 partículas, para quantificação do biodiesel.....	52
Figura 18 – Evolução das três execuções do iPLS-PSO com 60 intervalos e 50 partículas, para quantificação do biodiesel.....	52
Figura 19 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel usando o método iPLS-PSO com 60 intervalos e 100 partículas.	53
Figura 20 – Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 60 intervalos e 100 partículas.	54

Figura 21 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.....	55
Figura 22 – Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.....	55
Figura 23– Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	56
Figura 24– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	57
Figura 25 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.....	58
Figura 26– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.....	58
Figura 27 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO com 40 intervalos e 50 partículas.	60
Figura 28– Valores de referência <i>versus</i> previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 40 intervalos e 50 partículas.	60
Figura 29 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal utilizando toda a região espectral das misturas.	61
Figura 30 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT 62	
Figura 31– Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas, para a quantificação do óleo vegetal.....	64
Figura 32 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas, para a quantificação do óleo vegetal.....	65
Figura 33 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para a quantificação do óleo vegetal.....	65
Figura 34 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO com 40 intervalos e 25 partículas.....	66
Figura 35 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.....	67
Figura 36 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.....	68
Figura 37 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.	68
Figura 38 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	69

Figura 39 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.	70
Figura 40 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.....	71
Figura 41– Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.	71
Figura 42 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.....	72
Figura 43 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.	73
Figura 44 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel utilizando toda a região espectral da mistura.	74
Figura 45 – Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas para a quantificação do óleo diesel.....	76
Figura 46 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas para quantificação do óleo diesel.....	77
Figura 47 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para quantificação do óleo diesel.....	77
Figura 48 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO com 40 intervalos e 25 partículas.	78
Figura 49 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.	79
Figura 50 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.....	80
Figura 51 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.....	80
Figura 52 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	81
Figura 53 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.....	82
Figura 54 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.....	83
Figura 55 – Valores de referência <i>versus</i> previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.....	83
Figura 56 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DC) com 40 intervalos e 25 partículas.....	84

- Figura 57 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 50 partículas..... 85
- Figura 58 – Esquema contendo espectro de todas as misturas, do óleo diesel, do biodiesel e do óleo vegetal. 87
- Figura 59 – Erros de previsão para a quantificação do biodiesel, usando as regiões propostas pelo método iPLS-PSO com 40 intervalos e 50 partículas..... 88
- Figura 60 – Erros de previsão para a quantificação do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT..... 88
- Figura 61 – Valores de RMSECV *versus* RMSEP para diferentes *Gbest* para quantificação do biodiesel, usando o método iPLS-PSO(CD) com 60 intervalos e 25 partículas. 90
- Figura 62 – Valores de RMSECV *versus* RMSEP para diferentes *Gbest* para a quantificação do biodiesel, usando o método iPLS-PSO(CD) com 40 intervalos e 100 partículas. 90

LISTA DE TABELAS

Tabela 1- Planejamento experimental das misturas realizado na UNISC	36
Tabela 2– Planejamento experimental das misturas realizado na UFRGS	37
Tabela 3– Parâmetros instrumentais e regiões empregadas na aquisição dos espectros por FT-IR/ATR	38
Tabela 4 – Nomenclatura usada para identificar as implementações	44
Tabela 5– Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do biodiesel.	49
Tabela 6 – Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do óleo vegetal.	63
Tabela 7 – Principais resultados dos modelos, incluindo das cinco implementações da PSO para a quantificação do óleo diesel.	75

LISTA DE ABREVIATURAS

ABNT	Associação Brasileira de Normas Técnicas
AG	Algoritmo Genético
ANP	Agência Nacional de Petróleo, Gás Natural e Biocombustíveis
ATR	<i>Attenuated Total Reflectance</i>
FT-IR	<i>Fourier Transform Infrared Spectroscopy</i>
GC	<i>Gaseous Chromatograph</i>
GPC	<i>Gel Permeation Chromatography</i>
HPLC	<i>High Performance Liquid Chromatography</i>
iPLS	<i>Interval Partial Least-Squares</i>
NIR	<i>Near-infrared Reflectance</i>
PLS	<i>Partial Least-Square</i>
PSO	<i>Particle Swarm Optimization</i>
RMSEC	<i>Root Mean Square Error of Calibration</i>
RMSECV	<i>Root Mean Square Error of Cross Validation</i>
RMSEP	<i>Root Mean Square Error of Prediction</i>
RMSEV	<i>Root Mean Square Error of Validation</i>
TLC	<i>Thin Layer Chromatography</i>
VL's	Variáveis Latentes

LISTA DE SIMBOLOS

i	partícula
x_i	posição da partícula i
n	tamanho do enxame
it	iteração atual
$pbest_i$	melhor posição que a partícula i já obteve durante a busca
$Gbest$	melhor posição encontrada pelas partículas do enxame
c_1, c_2	parâmetros de confiança
v_i	velocidade atual da partícula i
V_{max}	limitador de velocidade
w	fator de inércia
w_{ini}	valor inicial para o fator de inércia
w_{fin}	valor final para o fator de inércia
k	coeficiente de constrição
R	nº total de iterações
$rand()$	função aleatória
$(pbest_i - x_i)$	distância entre a melhor posição encontrada para a partícula i e a posição atual dessa mesma partícula
$(Gbest - x_i)$	distância entre a melhor posição de todas as partículas do enxame e a posição atual da partícula i

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	Justificativa	16
1.2	Objetivos.....	16
1.2.1	Objetivo Geral.....	16
1.2.2	Objetivos Específicos.....	16
1.3	Organização do texto.....	17
2	FUNDAMENTAÇÃO TEÓRICA.....	18
2.1	Biodiesel	18
2.2	Espectroscopia no infravermelho com transformada de Fourier (FT-IR)	20
2.3	Principais técnicas de reflexão no infravermelho.....	21
2.3.1	Reflexão total atenuada	21
2.4	Análise multivariada.....	22
2.5	Regressão por mínimos quadrados parciais (PLS)	23
2.6	Regressão por mínimos quadrados parciais por intervalos (iPLS)	25
3	EXAME DE PARTÍCULAS	26
3.1	Introdução.....	26
3.2	O algoritmo básico de PSO.....	27
3.3	Estrutura do exame de partículas.....	28
3.4	Outros parâmetros do PSO.....	28
3.4.1	Tamanho da população	28
3.4.2	Velocidade Máxima.....	29
3.4.3	Fator de Inércia.....	29
3.4.4	Coeficiente de constrição.....	31
3.5	Outro tipo de representação	31
3.5.1	PSO binário	31
3.6	Algumas aplicações da técnica exame de partículas	32
4	METODOLOGIA	34
4.1	Preparo das misturas	35
4.2	Aquisição dos espectros	37
4.3	Tabulação das matrizes	38
4.4	Modelagem dos dados.....	38
4.4.1	Considerações gerais sobre as implementações da meta-heurística exame de partículas para selecionar variáveis de espectros no infravermelho	40
4.4.2	Estrutura básica do exame de partículas implementado para a otimização dos modelos de regressão	40
5	RESULTADOS	46
5.1	Resultados para a quantificação do biodiesel	46
5.1.1	Modelos para a determinação do biodiesel com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT	46
5.1.2	Resultados gerais e evoluções para a determinação do biodiesel	48
5.1.3	Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO	53
5.1.4	Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CC).....	54

5.1.5 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (DD).....	56
5.1.6 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CD).....	57
5.1.7 Resultados obtidos para a determinação do biodiesel aplicando o iPLS-PSO (DC)...	57
5.1.8 Considerações sobre os modelos para a previsão do biodiesel.....	59
5.2 Resultados para a quantificação do óleo vegetal	61
5.2.1 Modelos para determinação do óleo vegetal com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT.....	61
5.2.2 Resultados gerais e evoluções dos modelos para a previsão do óleo vegetal	62
5.2.3 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO.....	66
5.2.4 Resultados obtido para a determinação do óleo vegetal aplicando o método iPLS-PSO (CC).....	67
5.2.5 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DD)	69
5.2.6 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (CD)	70
5.2.7 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DC)	72
5.2.8 Considerações sobre os modelos para a previsão de óleo vegetal.....	73
5.3 Resultados para a quantificação do óleo diesel	74
5.3.1 Modelo para a determinação do óleo diesel com espectro inteiro.....	74
5.3.2 Resultados gerais e evoluções para a determinação do óleo diesel	75
5.3.3 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO	78
5.3.4 Resultados obtidos para a quantificação do óleo diesel aplicando o método iPLS-PSO (CC).....	79
5.3.5 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DD).....	81
5.3.6 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (CD).....	82
5.3.7 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DC).....	84
5.3.8 Considerações sobre os modelos de óleo diesel	85
5.4 Considerações finais sobre os modelos de biodiesel, óleo vegetal e óleo diesel	86
6 CONCLUSÕES.....	91
7 TRABALHOS FUTUROS.....	92
REFERÊNCIAS	93
ANEXOS.....	98

1 INTRODUÇÃO

Atualmente, a substituição parcial do óleo diesel por combustíveis renováveis está sendo efetuada mundialmente, e os óleos provenientes de fontes vegetais são as principais alternativas para se efetuar esta substituição.

No Brasil tornou-se obrigatório, desde 1º de janeiro de 2008, a adição de 2% de biodiesel (B2) ao diesel; com perspectivas deste biocombustível ser produzido, nos próximos anos, em grande escala para atender a demanda de mercado. Desta forma, como é feito atualmente com outros combustíveis, será necessário monitorar a qualidade desta mistura para garantir que o produto a ser comercializado não lese o consumidor, garanta o perfeito funcionamento dos veículos automotores e contribua realmente para conservação do meio ambiente.

Neste sentido, são crescentes os estudos que buscam desenvolver metodologias alternativas as metodologias tradicionais de quantificação de combustíveis. A técnica de espectroscopia no infravermelho é uma alternativa que está sendo estudada nesta área, devido a possibilidade de realizar análises não destrutivas, mais rápidas, sem gastos de reagentes e gerando informações com qualidade. Associadas a esta técnica são utilizadas ferramentas quimiométricas de regressão para construção de modelos.

Por outro lado, a otimização por enxame de partículas é um método de otimização recente (1995), baseado na analogia do comportamento social dos pássaros. Diversos estudos estão sendo desenvolvidos com esta técnica movidos pela sua simplicidade, robustez e eficiência.

Essa dissertação desenvolveu um método híbrido para quantificar o biodiesel, óleo vegetal e diesel nas misturas. Método este composto pelo método iPLS (*Interval Partial Least-Squares*) otimizado pela meta-heurística enxame de partículas, visando selecionar as variáveis mais representativas para construção de modelos de regressão mais robustos.

1.1 Justificativa

Tendo em vista que o processo de adição do biodiesel (2%) ao diesel está iniciando este ano em nosso país, e as perspectivas são bastante amplas, as indústrias terão que se estruturar de maneira a poder controlar a qualidade bem como evitar a adulteração deste biocombustível. Neste contexto, a espectroscopia no infravermelho associada a ferramentas quimiométricas, pode ser uma solução alternativa aos métodos tradicionais.

O conjunto de dados obtidos neste tipo de misturas biodiesel/diesel apresenta um número extremamente elevado de variáveis, bem como as combinações resultantes possíveis. Com o propósito de selecionar um conjunto de variáveis que possibilitem a construção de modelos com habilidade equivalente na calibração e na previsão, sem a necessidade de avaliar todas as soluções possíveis optou-se por empregar a técnica de otimização chamada de enxame de partículas para efetuar esta seleção.

1.2 Objetivos

1.2.1 Objetivo Geral

O objetivo geral deste trabalho foi verificar a eficiência do método enxame de partículas para selecionar variáveis capazes de otimizar os modelos de análise multivariada, aplicados na quantificação das misturas biodiesel/óleo vegetal/diesel.

1.2.2 Objetivos Específicos

- Estudar métodos de otimização empregados no processo de quantificação de matérias-primas e combustíveis.
- Quantificar misturas de biodiesel, óleo vegetal e diesel, através de modelos de regressão.
- Comparar os resultados obtidos pelo modelo desenvolvido com outros descritos na literatura.

1.3 Organização do texto

O próximo capítulo introduz alguns aspectos importantes para compreensão e desenvolvimento deste trabalho, abrangendo uma visão geral do biodiesel, informações sobre espectroscopia no infravermelho com transformada de Fourier (FT-IR/ATR) e sobre os métodos de regressão que serão utilizados. Já no terceiro capítulo é apresentada a meta-heurística enxame de partículas implementada para otimização dos modelos.

O quarto capítulo descreve a metodologia utilizada na construção desta dissertação iniciando com a elaboração das misturas e a obtenção dos espectros, juntamente com uma descrição detalhada da ferramenta desenvolvida iPLS-PSO e suas variações.

No quinto capítulo são apresentados os resultados obtidos na quantificação das misturas, juntamente com uma análise destes resultados. No sexto capítulo são apresentadas às conclusões deste trabalho e, finalmente no sétimo capítulo são apresentadas sugestões de futuros trabalhos.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Biodiesel

De toda a energia consumida mundialmente, a maior parte vem do petróleo, do carvão e do gás natural, fontes estas limitadas e com previsão de se esgotar no futuro; portando a busca por fontes alternativas de produção de energia é de suma importância. Neste contexto surgem inúmeros estudos, que envolvem a produção de combustíveis alternativos ao óleo diesel, proveniente de óleos vegetais, denominados biocombustíveis (NASCIMENTO, COSTA NETO, MAZZUCO, 2001). Os primeiros estudos envolvendo o uso de óleos vegetais como combustível foram realizados na época do desenvolvimento dos motores a diesel, porém neste período a questão ambiental não era considerada importante. Atualmente, quando as atenções se voltam aos problemas ambientais e a redução da emissão dos gases causadores do efeito estufa, a busca por fontes de produção de combustíveis “limpos” tem sido constante no mundo.

No Brasil, a lei nº 11.097, de 13 de janeiro de 2005, determina percentuais de mistura do biodiesel ao óleo diesel, bem como o monitoramento da inserção deste novo combustível ao mercado, pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. A adição obrigatória do percentual de 2% (B2) iniciou em janeiro deste ano; e a meta é que seja adicionado 5% (B5) a partir de 2013. Segundo esta lei, o biodiesel é um biocombustível derivado de biomassa renovável para uso em motores a combustão interna com ignição por compressão ou, conforme regulamento, para geração de outro tipo de energia, que possa substituir parcial ou totalmente combustíveis de origem fóssil (BRASIL, 2005). Podem ser obtidos através de óleos vegetais ou gorduras animais, novos ou residuais, a partir de processos tecnológicos para conversão de óleos em combustíveis.

Dentre as reações empregadas na produção do biodiesel a partir de óleos vegetais, destaca-se a reação de transesterificação, onde triacilglicerídeos reagem com álcool na presença de uma base ou ácido forte, produzindo uma mistura de ésteres de ácidos graxos e glicerol, conforme ilustrado na Figura 1 (GERIS *et al.*, 2007).

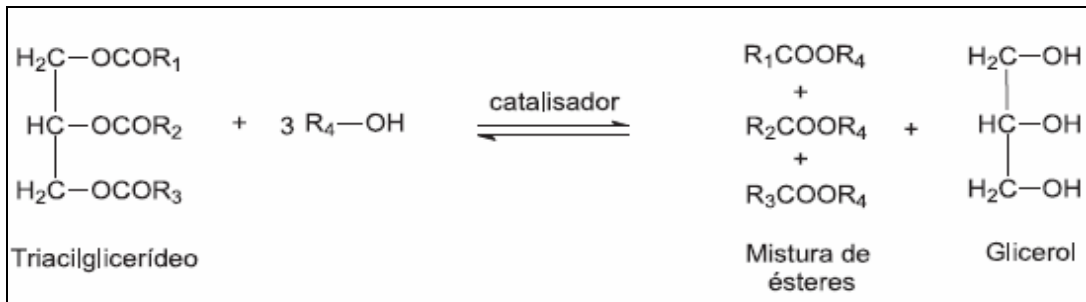


Figura 1 – Esquema de transesterificação de um triacilglicerídeo.
 Fonte: GERIS *et al.*, 2007.

O emprego do biodiesel em substituição total ou parcial ao diesel (pois é um produto miscível) proporciona benefícios ao meio ambiente, pois o CO₂ liberado é de origem de fonte renovável; sociais, pois possibilita a geração de emprego e renda em comunidades agrícolas e; econômicos tendo em vista que o diesel é um dos combustíveis mais consumidos no Brasil, pois a maior parte do transporte de cargas é realizado por meio terrestre (OLIVEIRA, 2006).

Entretanto, um dos problemas dos óleos vegetais é o preço, que tende a ser mais elevado que os de óleos de origem fósseis. Diversos pontos devem ser levados em consideração na hora de calcular o custo do biodiesel: oferta da matéria-prima, mão-de-obra, preço dos produtos agrícolas, infra-estrutura e a competição da matéria-prima com outros mercados. Um outro ponto que não pode ser esquecido são os “subprodutos” gerados durante o processo de transformação (o farelo, a glicerina e o álcool hidratado) que podem fazer com que o preço do biodiesel aumente para compensar os prejuízos de subprodutos, ou reduzido por causa de lucros com derivados do processo (CEPEA, 2006). Nos EUA, as fontes mais comuns para produção do biodiesel são do óleo de soja e o óleo de frituras reciclado, mas o custo ainda elevado de fabricação, e a disponibilidade limitada de algumas matérias primas usadas, fazem com que a aplicação comercial seja ainda limitada (RADICH, 2004).

O Brasil porém é um dos poucos países com tamanha diversidade na produção de biodiesel, a partir de plantas como soja, babaçu, amendoim, mamona, girassol, dendê e palma (GERIS *et al.*, 2007). Estudos estão sendo feitos para produzir, e aperfeiçoar a produção de biodiesel a partir do reaproveitamento de óleos como: óleos de frituras (COSTA NETO *et al.*, 2000; ENCINAR, GONZÁLEZ, RODRÍGUEZ-REINARES, 2005), óleos de oliva usados (DORADO *et al.*, 2004). Entretanto, independente da matéria prima, o produto final deve ser único e padronizado conforme os requisitos definidos pela Agência Nacional de Petróleo, Gás Natural e Biocombustíveis - ANP.

Diversas metodologias para análise do biodiesel estão sendo abordadas, tentando estabelecer um combustível padrão em diversas regiões e países ao redor do mundo. Estão inseridos neste contexto os Estados Unidos, a Europa, o Brasil, a Austrália e o Sul da África. Dentre as principais metodologias que vem sendo utilizadas para a quantificação de derivados, sub produtos e contaminações, destacam-se a cromatografia gasosa (GC), a cromatografia líquida de alta eficiência (HPLC), a cromatografia por permeação gel (GPC) e cromatografia por camada delgada (TLC) (KNOTHE, 2006). Porém estes métodos além de destrutivos, são laboriosos e consomem reagentes químicos.

Levando em consideração os antecedentes históricos e atuais de adulterações de combustíveis no Brasil, associado a necessidade de análise e controle de qualidade desse biocombustível, metodologias alternativas estão sendo propostas como a utilização da espectrofotometria no infravermelho próximo (NIR), infravermelho médio (FT-IR) e Raman (OLIVEIRA, 2006).

Outro trabalho recente desenvolvido por Pimentel *et al.* (2006) foi a utilização da espectroscopia no infravermelho e calibração multivariada na determinação de biodiesel em misturas com diesel mineral.

2.2 Espectroscopia no infravermelho com transformada de Fourier (FT-IR)

O instrumento utilizado atualmente para a obtenção de um espectro no infravermelho é denominado espectrofotômetro no infravermelho com transformada de Fourier (FT-IR). O modo de funcionamento de um equipamento FT-IR é denominado método interferométrico, pois o componente óptico básico deste equipamento é o interferômetro de Michelson (BARBOSA, 2007).

Dentre as vantagens de utilização deste método pode-se mencionar que: apresenta uma maior sensibilidade ou maior razão sinal/ruído, pois utiliza todos os comprimentos da onda ao mesmo tempo durante a obtenção do espectro; possibilita a obtenção de espectros em poucos segundos com maior resolução e maior precisão nas medidas de número de ondas (BARBOSA, 2007).

2.3 Principais técnicas de reflexão no infravermelho

Dentre os principais métodos de reflexão no infravermelho utilizados na aquisição dos dados espectrais mencionam-se: o método por reflexão difusa, o método por reflexão especular (ou externa) e o método de reflexão total atenuada (ATR), sendo este último o que será utilizado neste trabalho.

2.3.1 Reflexão total atenuada

A reflexão total atenuada (ATR – do inglês *attenuated total reflection*) é uma técnica que possibilita a obtenção de dados espectrais de amostras de maneira rápida, além de exigir uma mínima preparação da amostra. Possibilita analisar materiais líquidos, espessos e muito absorventes. Esta técnica caracteriza-se pelas múltiplas reflexões de radiação infravermelha que ocorrem no interior de cristais, interagindo somente com a amostra que estiver superficialmente no cristal (FERRÃO, 2001).

A reflexão na espectroscopia por ATR ocorre quando o feixe de radiação passa de um meio mais denso (cristal de ATR) para um menos denso (amostra), conforme Figura 2. Distorções que ocorram da banda no espectro podem ser ocasionadas pelo tipo de cristal que está sendo utilizado. Dentre os materiais usados para ATR está o ZnSe. Sua faixa de utilização é de 20.000 até 650 cm^{-1} , sendo recomendado para medidas em soluções aquosas. Outro material é o silício (Si) com uma faixa de utilização de 9.000 até 400 cm^{-1} e indicado para amostras altamente espalhativas. Um problema encontrado é a dificuldade de gerar-se boa reprodutibilidade no contato da amostra com o elemento ATR. Isto pode ser observado na variação da intensidade da banda, que é influenciado pelo aumento da pressão e pela área de contato entre o cristal e a amostra. Para medidas quantitativas, deve-se colocar toda a área do cristal em contato com a amostra (FERRÃO, 2001).

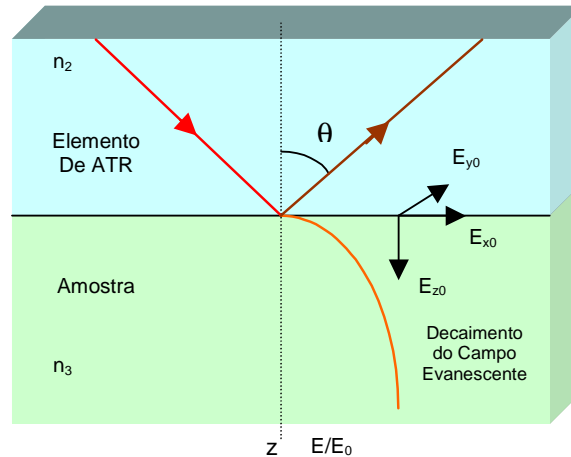


Figura 2 – Representação da propagação da radiação infravermelha através do elemento de ATR.
Fonte: FERRÃO, 2000.

Alguns exemplos de utilização da espectroscopia por reflexão total atenuada: na determinação rápida do índice de cafeína em refrigerantes (PARADKAR, IRUDAYARAJ, 2002) e, na análise de espécies de ácidos gordurosos insaturados em óleos dietéticos (YOSHIDA, YOSHIDA, 2003).

Em função das características das amostras e dos trabalhos que vem sendo realizados foi selecionada a reflexão total atenuada para a aquisição dos espectros das amostras a serem utilizadas neste trabalho.

2.4 Análise multivariada

O desenvolvimento da microeletrônica proporcionou modernos instrumentos de medição, e computadores com elevada capacidade de processamento. Esta união proporcionou um grande avanço nas análises instrumentais, possibilitando a aquisição de um grande número de dados da mesma amostra.

Analisar os dados obtidos é fundamental em todo o experimento realizado. A análise multivariada foi introduzida, ao tratamento dos dados químicos, em virtude da limitação da análise univariada, dando origem a uma área do conhecimento chamada quimiometria (SENA *et al.*, 2000). A análise multivariada refere-se a todos os métodos estatísticos que analisam simultaneamente múltiplas medidas do objeto que está sendo investigado. Pode-se dizer que qualquer análise simultânea de duas ou mais variáveis pode ser considerado

como análise multivariada (HAIR *et al.*, 2005). Nos métodos multivariados, o modelo estatístico considera a correlação entre muitas variáveis simultaneamente, possibilitando obter um número muito maior de informações (SENA *et al.*, 2000).

Quimiometria é uma área multidisciplinar do conhecimento, onde são utilizadas ferramentas matemáticas e estatísticas à análise de dados relativos aos processos químicos. Isto começou a ser amplamente utilizado para analisar amostras, graças a possibilidade de utilizar várias frequências dos espectros. A regressão é uma das mais importantes ferramentas na análise quantitativa. Os procedimentos matemáticos utilizados visam determinar as relações quantitativas existentes entre duas ou mais variáveis (FERRÃO, 2000).

Estas ferramentas de análise multivariada de dados vem sendo amplamente estudadas para aplicação na área industrial, como por exemplo, na análise exploratória na diferenciação química de vegetais (MORGANO *et al.*, 1999); na identificação de diferentes formulações de medicamentos manipulados contendo piroxicam (PARISOTTO *et al.*, 2005), na determinação simultaneamente dos teores de constituintes inorgânicos em sucos e refrigerantes de uva (FERREIRA *et al.*, 2002).

Segundo Gurden, Westerhuis e Smilde (2002), o processo industrial é carente de novas tecnologias que efetuem o monitoramento e o controle de processo em lote. A espectroscopia está se tornando uma opção de se obter em tempo real informações químicas de alta qualidade em intervalos de tempos frequentes em diversas aplicações industriais. Este autor enfatiza o correto manuseio dos dados medidos, a fim de extrair o máximo de informação, visando melhorar a performance e a eficiência do processo.

2.5 Regressão por mínimos quadrados parciais (PLS)

A técnica de regressão por mínimos quadrados parciais (PLS - do inglês *Partial Least Square*) utiliza o espectro por inteiro, ou parte deste, para construção do modelo. A matriz de dados **X**, é decomposta em várias matrizes **M** juntamente com uma matriz de resíduos **E** (KONZEN *et al.*, 2003), conforme Equação 1:

$$\mathbf{X} = \mathbf{M}_1 + \mathbf{M}_2 + \mathbf{M}_3 + \dots + \mathbf{M}_a + \mathbf{E} \quad (1)$$

onde:

a correspondem ao número de variáveis latentes ou componentes principais;

M constituem os componentes principais (ou variáveis latentes), e são compostas por dois vetores, **t** (os escores) e **p** (os pesos), conforme as Equações 2 e 3:

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

$$\mathbf{Y} = \mathbf{UQ}^t + \mathbf{F} \quad (3)$$

onde:

X é a matriz de dados (medida instrumental);

Y é a matriz de concentração;

T e **U** são os escores de **X** e **Y** respectivamente;

P e **Q** são os *loadings* de **X** e **Q**, e

E e **F** são os erros de modelagem de **X** e **Y**.

Para determinar o número de variáveis latentes (VL's) que devem ser empregadas na construção de modelo, não há um método exato, por isso os pesquisadores normalmente utilizam um aumento crescente nas VL's e verificam o comportamento da modelagem (KONZEN *et al.*, 2003).

O modelo construído é considerado eficiente, se for capaz de descrever da melhor maneira a situação real a que é proposto, levando em consideração o maior número possível de variações. Usualmente, os modelos são avaliados a partir da correlação entre a técnica proposta e o método de referência, com base nos valores de correlação e dos erros padrões de calibração (RMSEC) e validação (RMSEV), sendo os erros calculados conforme Equação 4:

$$\text{RMSEC, RMSEV, RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

onde:

y_i e \hat{y}_i corresponde aos valores de referência e estimado respectivamente para a *i*-ésima amostra e, **n** número total de amostras.

Outro erro utilizado para avaliar o modelo é o RMSECV que é obtido quando se utiliza na validação o processo de validação cruzada (*crossvalidation*), que consiste em

retirar uma amostra de cada vez do conjunto de calibração ao longo da validação. A equação para obtenção deste erro é análoga a equação 4, sendo que o i neste caso corresponde a amostra retirada do conjunto e n número total de amostras utilizadas na construção do modelo.

Na área de controle de qualidade e adulterações de combustíveis utilizando o PLS pode-se mencionar alguns trabalhos como: a análise das misturas biodiesel-óleos vegetais (OLIVEIRA, 2006); a análise qualitativa e quantitativa de combustíveis, em veículos leves, álcool e gasolina (SACORAGUE, 2004) e a determinação dos componentes da gasolina (FERREIRA, 2005).

2.6 Regressão por mínimos quadrados parciais por intervalos (iPLS)

Este método é uma extensão da regressão por PLS, onde o espectro é dividido em intervalos equidistantes, e para cada intervalo é construído um modelo PLS. O foco deste método é possibilitar que sejam selecionados regiões do espectro global que possuam informações relevantes, e um novo modelo seja construído a partir destas variáveis selecionadas (NORGAARD *et. al*, 2000).

No método iPLS, o espectro é dividido em n intervalos, e são construídos um modelo para cada intervalo de região espectral, e mais um modelo com o espectro inteiro. Os modelos são comparados utilizando o parâmetro de validação RMSECV/RMSEP (*root mean square error of cross-validation/prediction*), (LEARDI, NORGAARD, 2004). Pode-se propor o seguinte pseudocódigo iPLS:

Dividir os espectros em n intervalos equidistantes

Enquanto o número de intervalos $\leq n$ faça

Constrói o modelo PLS para o n -ésimo intervalo

FimEnquanto

Selecionar o modelo cujo intervalo apresenta o menor RMSECV

Na área de pesquisa de óleos, empregando o método iPLS, menciona-se os trabalhos de quantificação de contaminantes em óleos lubrificantes (BORIN, POPPI, 2005) e, parâmetros de qualidade de lubrificantes e óleo de oliva extra virgem e puro (BORGES NETTO, 2005).

3 EXAME DE PARTÍCULAS

3.1 Introdução

A técnica de otimização por enxame de partículas (PSO – do inglês *Particle Swarm Optimization*), foi criada por Kennedy e Eberhart (1995), a partir da análise comportamental dos pássaros a procura de alimento ou de um local para construção do ninho, conforme ilustrado na Figura 3. Na procura pelo ninho ou alimento, todo o indivíduo (partícula) pode lucrar com as experiências dos membros do grupo (enxame). Propuseram um algoritmo de otimização robusto, simples e de fácil implementação computacional.

A PSO é uma técnica de computação evolutiva, motivada pela simulação do “comportamento social”, que trabalha em uma população potencial de soluções em um espaço de busca, onde, através da cooperação e da competição entre estas soluções é possível encontrar uma solução ótima, quando aplicadas em problemas de otimização (POMEROY, 2003).



Figura 3 - Ilustração do método de enxame de partículas.

Este algoritmo faz uso da teoria dos enxames, onde os pássaros fazem uso de suas experiências e da experiência do bando para localizar o ninho ou alimento. Desta forma a área sobrevoada corresponde ao espaço definido do projeto e encontrar a comida significa encontrar o ótimo ou próximo do ótimo. A possibilidade de encontrarem alimento, baseados na inteligência social, aumenta significativamente, porque um indivíduo aprende com o acerto do outro (PRADO e SARAMAGO, 2005).

3.2 O algoritmo básico de PSO

O algoritmo é baseado em um espaço de busca, onde é inicialmente gerada uma população aleatória de partículas e a estas associadas uma velocidade, ou seja, cada partícula possui uma posição (x_i) e uma velocidade (v_i). A cada interação, cada partícula atualiza sua velocidade, pela Equação 6, e uma nova posição é gerada pela soma de sua posição atual e a nova velocidade, conforme a Equação 5 (ENGELBRECHT, 2005).

$$x_i^{it+1} = x_i^{it} + v_i^{it+1} \quad (5)$$

$$v_i^{it+1} = v_i^{it} + (c_1^{it} * rand()^{it} * (pbest_i^{it} - x_i^{it})) + (c_2^{it} * rand()^{it} * Gbest^{it} - x_i^{it}) \quad (6)$$

A Equação 6, que calcula a velocidade da partícula, é composta também pela função $rand()^{it}$ que gera números aleatórios entre 0 e 1, pelos parâmetros de confiança c_1^{it} e c_2^{it} que indicam quanto a partícula confia em sua experiência (c_1^{it}) e na do enxame (c_2^{it}); e pelos termos de aceleração: $(pbest_i^{it} - x_i^{it})$ que representa a distância entre a melhor posição encontrada para a partícula i e a posição atual dessa mesma partícula, e $(Gbest^{it} - x_i^{it})$, que representa a distância entre a melhor posição de todas as partículas do enxame e a posição da partícula i .

As partículas enquanto percorrem seu espaço de busca geram informações que são armazenadas em vetores, onde:

$X_i = (x_1, x_2, x_3, \dots, x_n)$ corresponde as posições das partículas;

$V_i = (v_1, v_2, v_3, \dots, v_n)$ indica as velocidades das partículas;

$Pbest_i = (pbest_1, pbest_2, pbest_3, \dots, pbest_n)$ corresponde a melhor posição obtida de cada partícula. E n indica o tamanho do enxame.

Armazena-se também, durante o processo de exploração, a melhor posição (*fitness*) já encontrada entre todas as partículas do enxame, que é representado pelo $Gbest$.

3.3 Estrutura do enxame de partículas

O algoritmo de PSO apresenta um conceito simples, de fácil implementação e de eficiência computacional. A seguir um pseudocódigo do PSO (ENGELBRECHT, 2005):

Criar uma população aleatória de partículas, denominada enxame;

Inicializar esta população com posições e velocidades aleatórias;

Repita para cada partícula i

 Calcular a função objetivo (*fitness*);

 Comparar o novo valor resultante da *fitness* da partícula i . Se o valor for melhor que o valor do *pbest* da partícula i , atualize o *pbest* com o novo valor da *fitness*.

 Comparar o novo valor resultante da *fitness* da partícula i com o valor do *Gbest*. Se o valor da *fitness* da partícula i for melhor que o valor do *Gbest*, atualize o valor do *Gbest* com este novo valor;

 Alterar a posição e a velocidade da partícula i ;

Até que a condição de parada seja satisfeita.

3.4 Outros parâmetros do PSO

A versão básica do enxame de partículas foi aperfeiçoada desde sua proposição em 1995. Abaixo são detalhados alguns dos parâmetros importantes relatados na literatura e que contribuíram no aperfeiçoamento deste método.

3.4.1 Tamanho da população

O tamanho da população é determinado em função do problema em que PSO será aplicado. Variam entre 20 e 50 partículas. Soluções ótimas geralmente são obtidas com tamanhos menores de população (MENDES, 2004).

Parsopoulos e Vrahatis (2002) trabalharam com enxames não maiores de 40 partículas. Já Shi e Eberhart (1999) testaram enxames com tamanhos de 20, 40, 60 e 80 partículas.

3.4.2 Velocidade Máxima

Inicialmente, o algoritmo enxame de partículas, não controlava a velocidade das partículas. Para tal foi introduzido um importante parâmetro V_{\max} . Valor muito grande para este parâmetro, facilita a exploração global, enquanto que valor muito pequeno favorece a exploração local. Porém se o valor de V_{\max} for demasiado pequeno, o enxame pode não explorar suficientemente além das regiões localmente boas (ENGELBRECHT, 2005).

O V_{\max} é um parâmetro ajustado pelo usuário conforme as características do problema em questão, normalmente entre [-4, 4]. Assim, se $(v_i) < -4$ então $(v_i) = -4$ e se $(v_i) > 4$ então $(v_i) = 4$ (KENNEDY, EBERHART, SHI, 2001).

3.4.3 Fator de Inércia

Shi e Eberhart (1998a, 1998b) incluíram um novo parâmetro, chamado de fator de inércia ou peso de inércia (*inertia weight*), no algoritmo PSO, representado por (w) conforme visualizado na Equação 7. O papel fundamental do fator de inércia é balancear habilidade de exploração de busca local e global do algoritmo. Observaram que valores maiores para (w) facilita a busca no espaço global, enquanto que valores menores para (w) aumentam a busca no espaço local, fazendo com que as partículas convirjam rapidamente, para o ponto ótimo ou não, sem que haja uma boa exploração do espaço de busca.

$$v_i^{it+1} = w * v_i^{it} + (c_1^{it} * \text{rand}()^{it} * (pbest_i^{it} - x_i^{it})) + (c_2^{it} * \text{rand}()^{it} * (Gbest^{it} - x_i^{it})), \quad (7)$$

Uma alternativa interessante para que ocorra um balanço entre a investigação e a exploração é iniciar a busca com um valor alto de (w) e diminuir este valor ao longo das

iterações, como proposto por Shi e Eberhart (1998a; 1998b, 1999). onde propuseram que o valor de (w) deveria diminuir linearmente entre 0,9 e 0,4, aproximadamente. Já Parsopoulos e Vrahatis (2002) utilizaram 1 para o valor inicial de (w), sofrendo uma diminuição gradual até 0.4.

Shi e Eberhart (2001) utilizaram um conjunto de regras Fuzzy para ajustar dinamicamente o valor de (w), mostrando-se ser um método promissor especialmente para problemas em ambientes dinâmicos.

O fator de inércia também pode sofrer uma diminuição linear a cada iteração através da Equação 8 (ENGELBRECHT, 2005).

$$w^{it} = (w_{ini} - w_{fin}) \frac{(R - it)}{R} + w_{fin} \quad (8)$$

onde w_{ini} e w_{fin} é o valor inicial e o valor final para o fator de inércia, respectivamente; R é o número total de iterações.

Como pode ser visto, muitos estudos empregaram w com variação decrescente no seu valor, em contrapartida Zheng. *et al.* (2003) propuseram que o valor de w fosse incrementado linearmente de 0.4 até 0.9 e obtiveram bons resultados.

Já para testar o método enxame de partículas Prado e Saramago (2005) utilizaram o valor da (w) em 0,729, sendo atualizado através da Equação 9, onde f_w é uma constante entre 0 e 1.

$$w_{new} = f_w * w_{old} \quad (9)$$

O fator de inércia também pode sofrer uma redução de maneira não linear. Peram, Veeramachaneni e Mohan (2003) aplicaram a Equação 10 para o cálculo do valor de (w):

$$w^{it+1} = \frac{(w^{it} - 0.4) * (R - it)}{R + 0.4} \quad (10)$$

Recentemente, Chatterjee e Siarry (2006) propuseram uma nova maneira de cálculo de (w) através da variação não-linear, conforme a Equação 11.

$$w^{it} = \left\{ \frac{(R-it)^{nl}}{(R)^{nl}} \right\} * (w_{ini} - w_{fin}) + w_{fin} \quad (11)$$

onde nl representa o índice não-linear.

3.4.4 Coeficiente de restrição

O coeficiente de restrição (K), calculado segundo Equação 12, e incorporado na equação de cálculo da velocidade, conforme Equação 13, visando limitar a velocidade das partículas, e garantir a convergência do PSO (CLERC, 1999; CLERC, KENNEDY, 2002).

$$v_i^{it+1} = K^{it} * [v_i^{it} + (c_1^{it} * rand()^{it} * (pbest_i^{it} - x_i^{it})) + (c_2^{it} * rand()^{it} * Gbest^{it} - x_i^{it})] \quad (12)$$

$$K^{it} = \frac{2}{2 - \varphi - \sqrt{\varphi^2 - 4\varphi}} \quad (13)$$

Sendo que $\varphi = c_1 + c_2$, $\varphi > 4$

3.5 Outro tipo de representação

Originalmente a técnica PSO foi desenvolvida para o universo de variáveis contínuas, mas muitos problemas são definidos num universo de variáveis discretas, onde o domínio é finito. O PSO foi facilmente adaptado para trabalhar com este tipo de variáveis. Este tipo de representação é chamado de PSO binário (ENGELBRECHT, 2005).

3.5.1 PSO binário

Esta representação tenta estender o algoritmo de enxame de partículas para trabalhar em espaço binários. Kennedy e Eberhart (1997) e Kennedy, Eberhart e Shi (2001)

descreveram uma alteração no algoritmo para operar sobre *bits* de caracteres, conforme Equação 14, onde a velocidade é usada como uma probabilidade limiar para determinar se x_i será zero ou um.

$$\text{sigmoid}(v_i^{it}) = \frac{1}{1 + \exp(-v_i^{it})} \quad (14)$$

Kennedy e Spears (1998) comparam o algoritmo binário com diversos tipos de algoritmos genéticos aleatórios. Para tal usaram um gerador aleatório. Os resultados mostram que o enxame de partículas binário foi o único a encontrar o ótimo global em cada uma das experiências, sem considerar as características do problema.

Agrafiotis e Cedeño (2002) usaram o PSO binário para seleção de características, em que (x_i) pode receber apenas 0 e 1 e, representa se a característica da partícula i será ou não selecionada. Neste caso as posições (x_i) são tratadas como uma probabilidade limiar para decidir entre o bit de valores 0 e 1, conforme Equação 15.

$$p_{ij}^{it} = \frac{x_{ij}^{\alpha it}}{\sum_{j=1}^{nx} x_{ij}^{\alpha it}} \quad (15)$$

Sendo que α determina a seleção, e $x_{ij} \in [0, 1]$, onde para valores de $\alpha > 1$ enfatizam indivíduos com alta adaptação, enquanto $\alpha < 1$ aumenta a chance de indivíduos com baixa adaptação serem selecionados.

Na literatura encontram-se pesquisas que relatam o uso do enxame de partículas binário como, por exemplo, na redução e colocação de banco de capacitores (KHALIL *et al.*, 2006).

3.6 Algumas aplicações da técnica enxame de partículas

Apesar de ser uma técnica bastante recente, na área de otimização, muitas pesquisas já foram feitas nas mais diversas áreas, como por exemplo: na análise do tremor humano

usando PSO e redes neurais (EBERHART e HU, 1999), otimização da recarga de combustível em um reator nuclear (MENESES, 2005); em classificações de imagens não supervisionadas (OMRAN, ENGELBRECHT, SALMAN, 2005), problemas complexos de engenharia nuclear (MEDEIROS, 2005) e na otimização de *layouts* (MÜLLER, 2007).

Schwaab (2005), testou algoritmos heurísticos de otimização na estimativa de parâmetros de modelo não lineares, e o método do enxame de partículas, em particular, apresentou resultados superiores aos demais algoritmos heurísticos utilizados neste trabalho, tanto na minimização de funções teste, como na estimativa de parâmetros de problemas reais.

A utilização de heurísticas e meta-heurística na otimização de processos químicos, tem recebido grande atenção por parte de pesquisadores. No entanto, o método de enxame de partículas, ainda é pouco explorado nesta área. Alguns exemplos são: PSO e rede neural na construção de modelos QSAR (WANG *et. al*, 2004) e, treinamentos de redes neurais para modelos QSAR de inibidores de plaquetas (QI SHEN *et al.*, 2006).

Por outro lado, o algoritmo genético (AG) que é uma ferramenta de otimização, apresenta um vasto número de trabalhos utilizando-o com uma forma de selecionar regiões espectrais para construção dos modelos, como por exemplo, na determinação simultânea de glicose, maltose e frutose (COSTA FILHO, POPPI, 2002) e, na otimização de modelos de regressão multivariados empregando métodos de seleção (SABIN, 2007). Já outros trabalhos relatam a utilização da heurística busca tabu na seleção do comprimento de onda (HAGEMAN *et. al*, 2003) e na otimização dos métodos de controle de qualidade de fármacos (KONZEN *et. al*, 2003).

4 METODOLOGIA

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

Para um melhor entendimento, a Figura 4 mostra um fluxograma das etapas que foram executadas na parte experimental, para a composição deste trabalho.

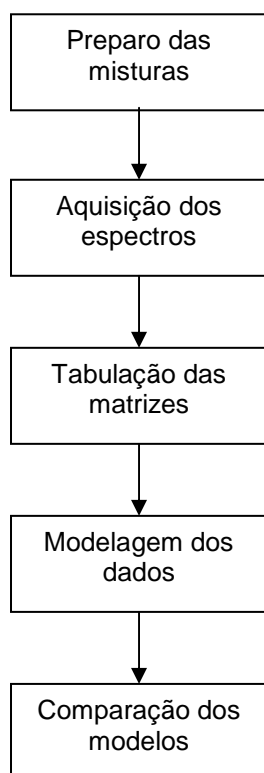


Figura 4 – Fluxograma das etapas executadas na parte experimental.
Fonte: Elaborado pela autora.

4.1 Preparo das misturas

Para a realização das misturas foram utilizados 2 planejamentos experimentais conforme tabelas 1 e 2. As 40 misturas binárias e ternárias conforme Tabela 1 foram realizadas a partir de uma amostra de óleo diesel, fornecida pelo Labcom da Universidade Federal do Rio Grande do Sul – UFRGS, de uma amostra de biodiesel de palma (doado pela Agropalma S/A), uma amostra de biodiesel de soja e uma amostra de óleo de soja (ambas fornecidas pela OleoPlan S.A). Já as 40 misturas binárias e ternárias conforme Tabela 2 foram realizadas a partir de uma segunda amostra de óleo diesel, fornecida pelo Labcom – UFRGS, da mesma amostra de biodiesel de soja e da mesma amostra de óleo de soja (ambas fornecida pela OleoPlan S.A).

A união dos dois planejamentos resultou em um conjunto total de 80 amostras binárias e ternárias, podendo estas serem visualizadas no diagrama ternário da Figura 5. Cabe ressaltar que na apresentação do diagrama os dois diferentes óleos diesel, bem como os dois diferentes biodiesel utilizados, foram considerados como um único.

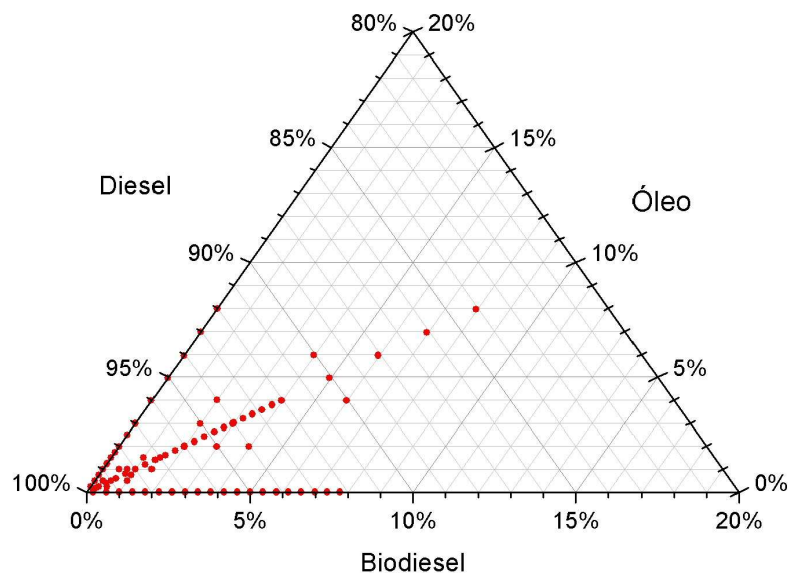


Figura 5 – Diagrama ternário para as misturas biodiesel/óleo vegetal/diesel.
Fonte: elaborado pela autora.

As amostras foram preparadas pela adição de alíquotas de biodiesel e/ou óleo vegetal conforme tabelas 1 e 2, em um balão volumétrico de 25 mL, seguida de adição do

óleo diesel até completar o volume. As soluções foram homogeneizadas e deixadas em repouso pelo tempo mínimo de 24 horas.

Tabela 1 - Planejamento experimental das misturas realizado na UNISC

Biodiesel (% v/v)	Óleo Vegetal (%v/v)	Diesel (%v/v)
0,2	0	99,8
0,2	0,2	99,6
0,6	0	99,4
0,4	0,4	99,2
1	0	99
0,6	0,6	98,8
1,4	0	98,6
0,8	0,8	98,4
1,8	0	98,2
1	1	98
2,2	0	97,8
1,2	1,2	97,6
2,6	0	97,4
1,4	1,4	97,2
3	0	97
1,6	1,6	96,8
3,4	0	96,6
1,8	1,8	96,4
3,8	0	96,2
2	2	96
4,2	0	95,8
2,2	2,2	95,6
4,6	0	95,4
2,4	2,4	95,2
5	0	95
2,6	2,6	94,8
5,4	0	94,6
2,8	2,8	94,4
5,8	0	94,2
3	3	94
6,2	0	93,8
3,2	3,2	93,6
6,6	0	93,4
3,4	3,4	93,2
7	0	93
3,6	3,6	92,8
7,4	0	92,6
3,8	3,8	92,4
7,8	0	92,2
4	4	92

Fonte: elaborado pela autora.

Tabela 2 – Planejamento experimental das misturas realizado na UFRGS

BioDiesel (%v/v)	Óleo Vegetal (%v/v)	Diesel (%v/v)
0,25	0,25	99,5
0,5	0,5	99
1	1	98
1,5	1,5	97
2	2	96
3	3	94
4	4	92
5	5	90
6	6	88
7	7	86
8	8	84
0,25	0,5	99,25
0,5	0,25	99,25
0,5	1	98,5
1	0,5	98,5
0,75	1	98,25
1	0,75	98,25
1	1,5	97,5
1,5	1	97,5
2	3	95
3	2	95
2	4	94
4	2	94
4	6	90
6	4	90
0	0,25	99,75
0	0,5	99,5
0	0,75	99,25
0	1	99
0	1,25	98,75
0	1,5	98,5
0	1,75	98,25
0	2	98
0	2,5	97,5
0	3	97
0	4	96
0	5	95
0	6	94
0	7	93
0	8	92

Fonte: elaborado pela autora.

4.2 Aquisição dos espectros

As amostras foram novamente homogeneizadas antes da aquisição dos espectros, sendo os mesmos obtidos em duplicata utilizando um acessório de reflexão total atenuada

(ATR) com cristal de ZnSe, acoplado a um espectrofotômetro Nicolet Magna 550, na faixa de 650 - 4000 cm^{-1} , conforme parâmetros descritos na Tabela 3. Antes de ser obtido o espectro de cada amostra foi realizado o espectro do branco, sem nenhum composto sobre o cristal.

Tabela 3 – Parâmetros instrumentais e regiões empregadas na aquisição dos espectros por FT-IR/ATR

Parâmetros	Condições
Região espectral	650 - 4000 cm^{-1}
Detector	DTGS
Divisor de feixes	KBr
Varreduras	32
Resolução	4 cm^{-1}
Acessório	Reflexão Total Atenuada - PIKE

Fonte: elaborado pela autora.

4.3 Tabulação das matrizes

Com os sinais espectrais obtidos, foi construída uma matriz de dados com a média das duplicatas dos espectros normalizados de cada amostra. Esta matriz foi utilizada para a construção de modelos de regressão multivariados para a quantificação do biodiesel, óleo vegetal e diesel. Foram implementados testes utilizando o espectro inteiro das misturas, as regiões espectrais propostas pela norma ABNT submetida à consulta popular em fevereiro de 2008 e finalmente com as regiões espectrais selecionadas pelas implementações propostas com o exame de partículas.

4.4 Modelagem dos dados

A seleção das amostras para a constituição dos conjuntos de calibração e previsão teve dois objetivos principais: construir conjuntos que contenham a informação da matriz de dados utilizados neste estudo e que contenha a maior variabilidade possível em relação à faixa de concentração. Os 80 espectros médios das amostras foram divididos em dois conjuntos, sendo 48 empregados na construção do conjunto de calibração e 32 no conjunto

de previsão. A distribuição das concentrações destas amostras para cada uma das substâncias participantes das misturas é apresentada nas Figuras 6, 7 e 8, onde se tomou o cuidado de garantir que as amostras com valores extremos de concentração pertençam ao conjunto de calibração.

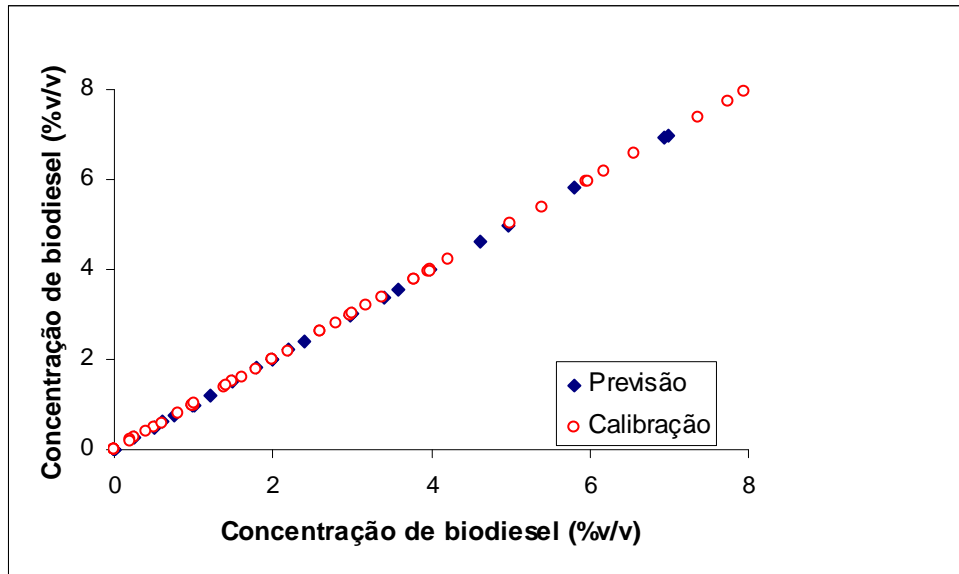


Figura 6 – Distribuição das amostras de calibração e previsão para o biodiesel.
Fonte: elaborado pela autora.

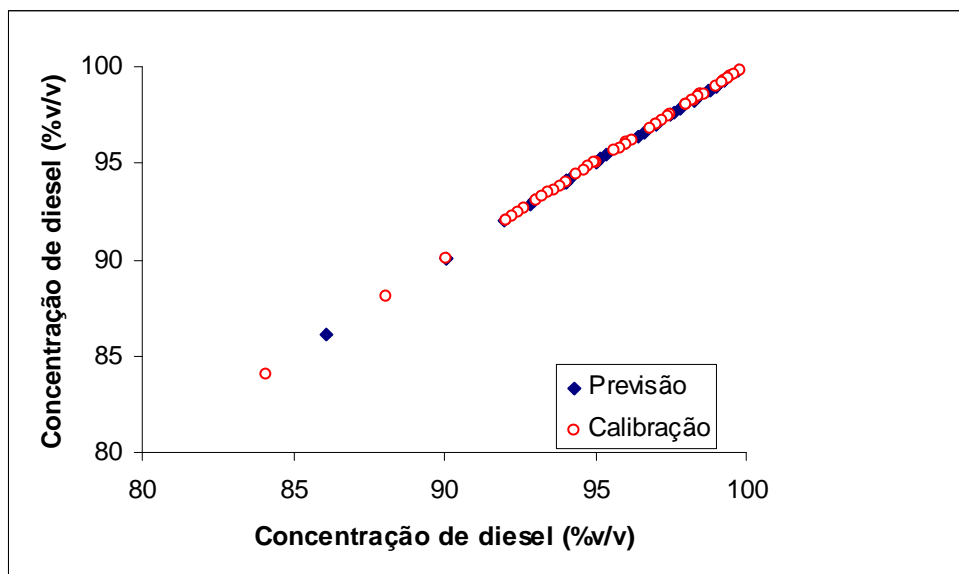


Figura 7 – Distribuição das amostras de calibração e previsão para o óleo diesel.
Fonte: elaborado pela autora.

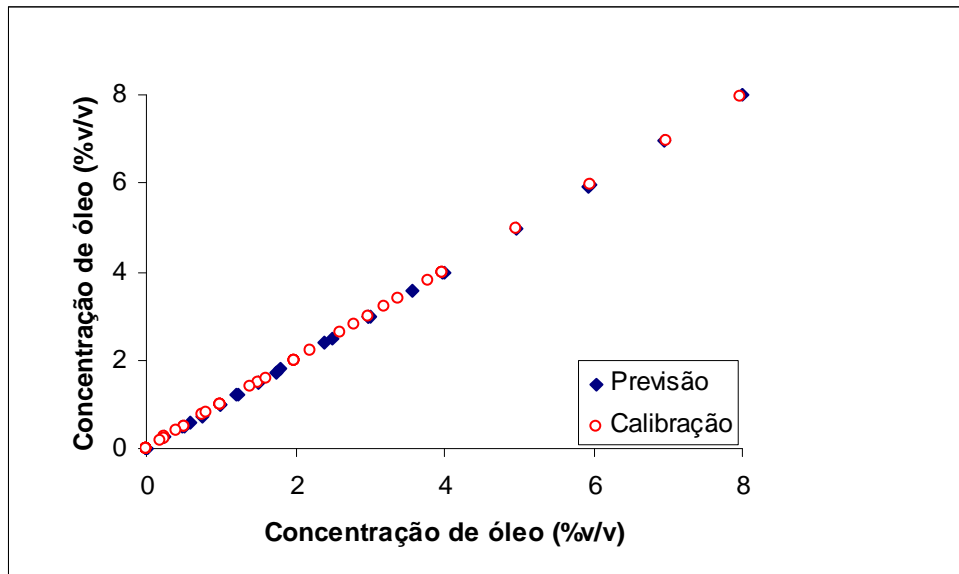


Figura 8 – Distribuição das amostras de calibração e previsão para o óleo vegetal.
Fonte: elaborado pela autora

Todos os modelos construídos neste trabalho, utilizaram o MATLAB[®] versão 6.5 (*The Math Works*, <http://www.mathworks.com>, EUA), juntamente com o pacote *iToolbox* versão 2 desenvolvido por Lars Norgaard (*iToolbox for MATLAB*, <http://www.models.kvl.dk>, Dinamarca).

As descrições das implementações com PSO e os resultados de todas as propostas mencionadas acima, são descritas no decorrer deste trabalho.

4.4.1 Considerações gerais sobre as implementações da meta-heurística enxame de partículas para selecionar variáveis de espectros no infravermelho

O princípio básico do algoritmo enxame de partícula é que as partículas se movimentam no espaço de busca, objetivando encontrar uma solução, orientadas pela sua experiência e pela experiência do enxame.

Como o propósito desta pesquisa é otimizar modelos de regressão multivariados, abaixo são explicados alguns conceitos do enxame de partículas neste contexto e que se mantêm para todas as implementações propostas.

Partícula: é uma possível solução para o problema. A posição da partícula é representada pelo conjunto de comprimentos de onda selecionados, ou seja, a configuração da mesma. A Figura 9 apresenta exemplo de duas partículas.

1	0	1	0	1	1	1	0	0	1
0	0	1	1	1	0	0	0	0	1

Figura 9 – Representação gráfica de duas partículas.
Fonte: elaborado pela autora.

Enxame: um conjunto de soluções (partículas).

Unidade de experiência: corresponde a um intervalo do espectro, podendo este ser ou não utilizado na construção do modelo de regressão.

Fitness: é a qualidade da posição (configuração) da partícula, obtido pelo erro de *cross-validação* (RMSECV) gerado no modelo de regressão.

Tamanho do enxame: corresponde ao número total de partículas utilizadas.

Velocidade: é a magnitude em que a solução corrente é alterada nas “direções” do $Gbest$ e do $pbest_i$.

Fator de Inércia: a intensidade em que a solução corrente é alterada (o percentual transferido do $Gbest$ e do $pbest_i$).

4.4.2 Estrutura básica do enxame de partícula implementado para a otimização dos modelos de regressão

O enxame inicia sua execução criando as partículas, onde cada partícula é composta de um número pré-definido de intervalos, codificadas com 0 e 1 aleatoriamente, e assim, cada partícula possui uma posição (solução inicial) e uma velocidade inicial. Para este problema, cada um dos intervalos da partícula corresponde a uma região do espectro, e os modelos serão construídos somente com os intervalos codificados com 1, conforme Figura 10.

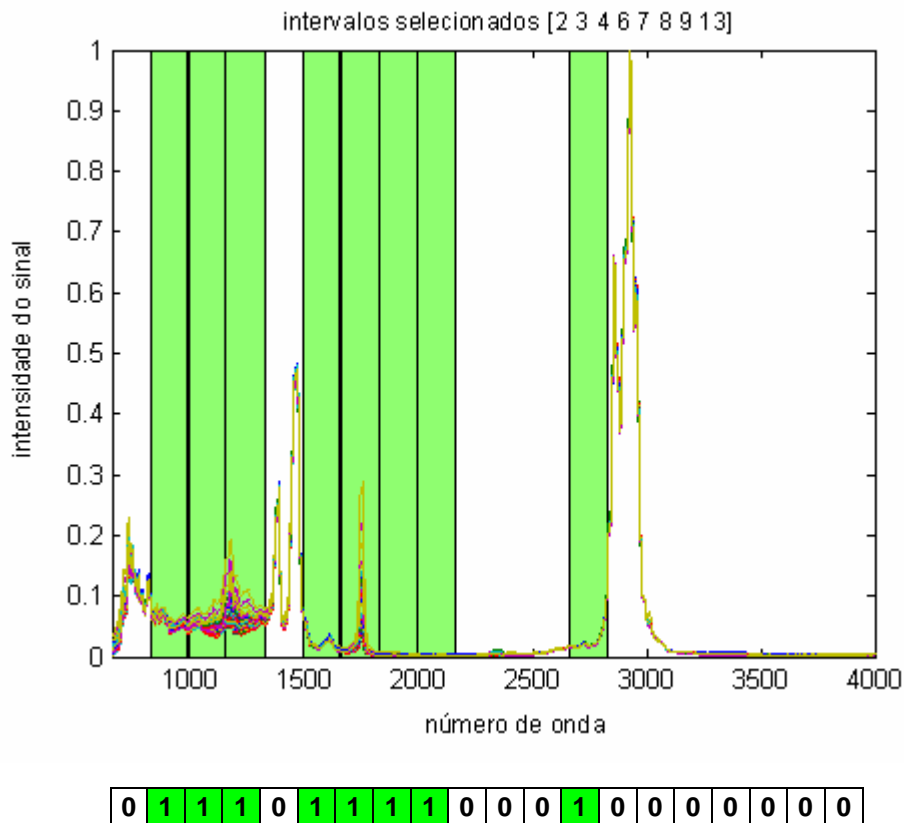


Figura 10 – Exemplo de uma partícula e a representação dos seus intervalos seleccionados. Fonte: elaborado pela autora com base no espectro do infravermelho de amostras de misturas biodiesel/óleo vegetal/diesel.

Quando o enxame é criado a posição inicial de cada partícula i é considerada como a sua melhor posição encontrada, ou seja o $pbest_i$, e a melhor posição entre todas as partículas, passa a ser considerada o $Gbest$ do enxame.

A cada iteração, cada partícula i , se desloca para uma outra posição, com base na soma da sua posição atual e nova velocidade gerada. Nestas implementações a velocidade será calculada através da Equação 7.

Assim, com base no deslocamento da partícula i é gerado um vetor somente com os intervalos seleccionados, e este é passado como parâmetro para a função $plsmodel$ do pacote $iToolBox$, juntamente com um modelo $iPLS$, com o mesmo número de intervalos da partícula i que está sendo avaliada, que gera um modelo de regressão empregando apenas estes intervalos. Este modelo gera um vetor de erros (RMSECV), onde cada valor deste vetor corresponde a utilização de um número diferente de variáveis latentes. O resultado do deslocamento da partícula ($fitness$) será o menor RMSECV, juntamente com o respectivo número de variáveis latentes.

Esta nova posição (valor resultante da *fitness*) da partícula i é comparada com o melhor posição que esta partícula i obteve até o presente momento ($pbest_i$). Da mesma maneira compara-se a *fitness* da partícula i com o valor do $Gbest$ do enxame. Neste trabalho em que se pretende minimizar o erro RMSECV, ambos, $pbest_i$ e $Gbest$ somente serão atualizados se o valor obtido pela *fitness* da partícula i tiver sido menor que a última melhor solução armazenada por esta partícula e pelo enxame, respectivamente.

As partículas terão a oportunidade de evoluir no espaço de busca durante 250 iterações, que é condição de parada.

Através de testes preliminares, foram definidos alguns parâmetros para serem adotados como padrão para todas as implementações, sendo eles:

- a intensidade (fator de inércia) em que a partícula i é alterada em direção ao $pbest_i$ (parte cognitiva) será maior do que a intensidade em direção ao $Gbest$ (parte social) e que as respectivas intensidades, para ambas direções, serão compostas por valores (percentuais) máximos e mínimos permitidos;
- serão utilizados enxames com tamanho de 25, 50 e 100 partículas;
- a quantidade de intervalos que cada partícula poderá ser constituída serão 20, 40 e 60 intervalos;
- para cada combinação proposta, todas as partículas apresentarão a mesma constituição (20, 40 e 60 intervalos) e, um tamanho fixo do enxame (25, 50 e 100 partículas). Além de que todas as combinações serão executadas em triplicata.

O algoritmo de enxame de partículas implementado busca combinar várias unidades de informação (regiões do espectro), para construção dos modelos de regressão. Nesta dissertação foram implementadas cinco versões do enxame de partículas, onde os resultados obtidos destas soluções, foram comparados com os modelos obtidos utilizando o espectro inteiro das misturas e as sub-regiões indicadas pela proposta da norma ABNT.

Foi criada uma nomenclatura para identificar as versões do enxame de partículas implementado, conforme Tabela 4.

Tabela 4 – Nomenclatura usada para identificar as implementações

Sigla	Definição
iPLS	Método usado para construção do modelo.
PSO	Método de otimização utilizado.
C* ou D**	Sentido do deslocamento da intensidade da partícula i em relação ao G_{best} .
C* ou D**	Sentido do deslocamento da intensidade da partícula i em relação ao p_{best}_i .

* - crescente ** - decrescente

Fonte: elaborado pela autora.

Abaixo são citadas as implementações efetuadas utilizando a PSO.

a) Implementação do Enxame iPLS-PSO

Nesta implementação, manteve-se a estrutura básica do enxame de partículas implementado, mas a intensidade em que a partícula i é alterada (o percentual transferido do G_{best} e do p_{best}_i) foi definida e mantida constantemente durante todas as iterações. Os percentuais utilizados aqui foram 10% do G_{best} e 15 % p_{best}_i .

b) Implementação do enxame iPLS-PSO (CC)

Esta abordagem manteve-se novamente a estrutura básica do enxame de partículas implementada inicialmente, mas a intensidade em que a partícula i é alterada (os percentuais transferidos do G_{best} e do p_{best}_i) foi iniciada com valores mínimos (5%), e no decorrer das iterações sofreram um incremento linear constante até os valores máximos permitidos (15% e 30%, respectivamente).

c) Implementação do Enxame iPLS-PSO (DD)

Com base na estrutura básica do enxame de partículas descrita inicialmente, foi proposta uma nova abordagem em que a intensidade em que a partícula i é alterada (o

percentual transferido do G_{best} e do p_{best_i}) é iniciada com valores máximos (15% e 30%, respectivamente), e no decorrer das iterações sofreram um decréscimo linear constante até os valores mínimos permitidos de 5%.

c) Implementação do Enxame iPLS-PSO (CD)

Partindo da implementação básica do enxame, nesta quarta proposta, a intensidade em que a partícula i é alterada em relação a parte cognitiva da fórmula da velocidade é definida com um valor mínimo permitido (5%) e a medida que a partícula i se movimenta no espaço de busca este valor sofre um incremento linear constante até o máximo de 15%. Em relação a parte social da fórmula da velocidade, a intensidade em que a partícula i é alterada (percentual transferido do p_{best_i}) é inicialmente definida com um valor máximo permitido (30%), e no decorrer das iterações sofre um decréscimo linear constante, até atingir o mínimo permitido (5%).

d) Implementação do Enxame iPLS-PSO (DC)

Em contrapartida, esta implementação partiu da proposta básica do enxame, mas a intensidade em que a partícula i é alterada em relação a parte cognitiva da fórmula da velocidade é definida com um valor máximo permitido (15%) e, a medida que a partícula i se movimenta no espaço de busca este valor sofre um decréscimo linear constante até atingir o mínimo permitido (5%). Em relação a parte social da fórmula da velocidade, a intensidade em que a partícula i é alterada, foi iniciada com um valor mínimo (5%) permitido, e no decorrer das iterações sofre um incremento linear constante, até atingir o máximo permitido (30%).

5 RESULTADOS

5.1 Resultados para a quantificação do biodiesel

Visando quantificar o biodiesel nas misturas, foram desenvolvidos diversos modelos de regressão, e os resultados obtidos estão descritos a seguir. Primeiramente são apresentados os modelos usando o PLS e em seguida, são descritos os modelos obtidos com a otimização através da meta-heurística enxame de partículas combinada ao iPLS.

5.1.1 Modelos para a determinação do biodiesel com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT

Com base no conjunto de dados foram desenvolvidos dois modelos de regressão, ambos utilizando para construção o método PLS. O primeiro modelo, para a quantificação do biodiesel, foi construído utilizando toda região espectral (650 cm^{-1} a 4000 cm^{-1}) conforme Figura 11, e o desempenho do modelo empregando 6 variáveis latentes, é apresentado na Figura 12.

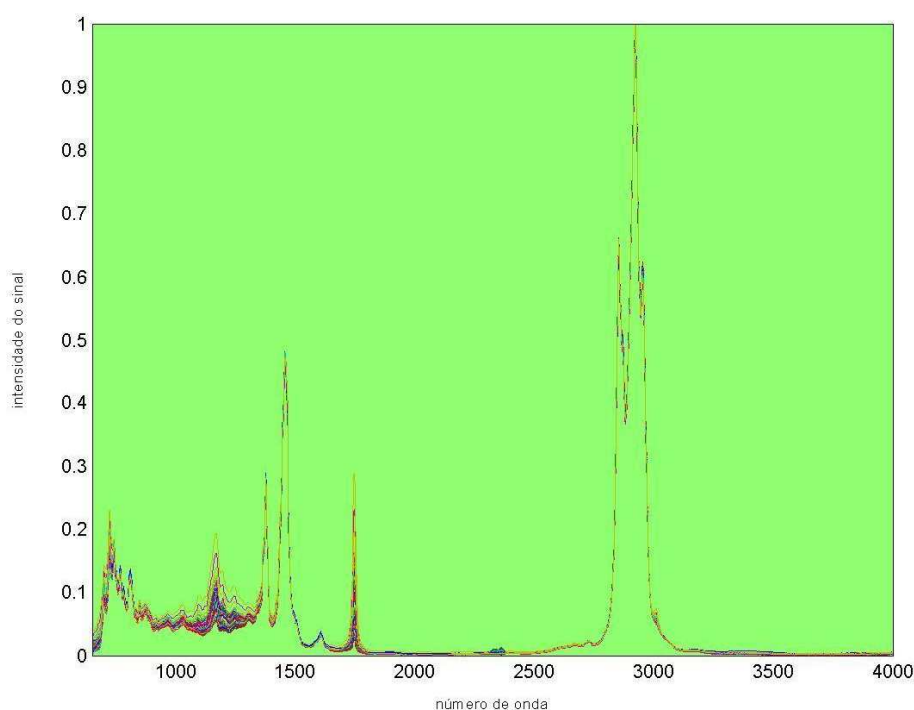


Figura 11 – Região espectral utilizada na modelagem para a quantificação do biodiesel. Fonte: elaborado pela autora.

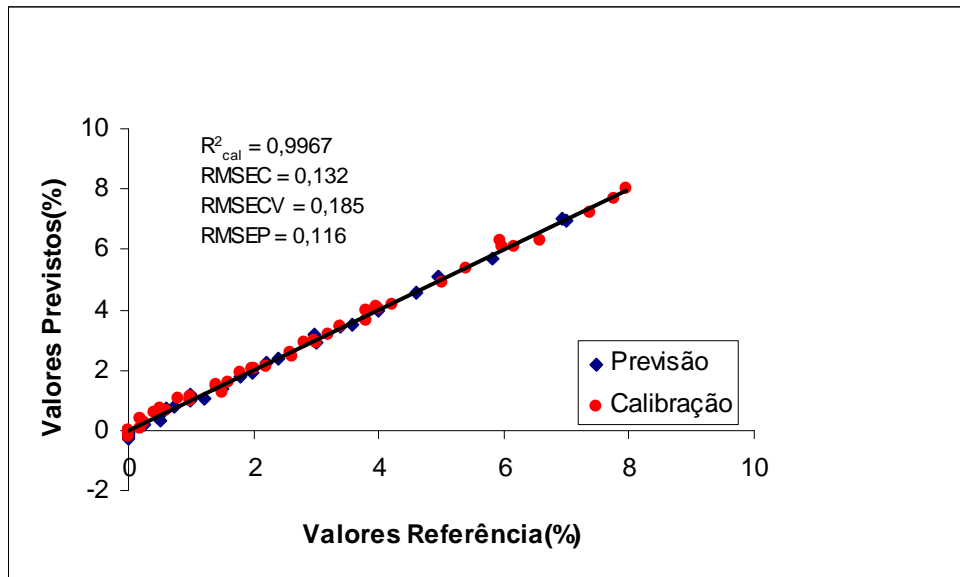


Figura 12 – Valores de referência *versus* previstos para o modelo de previsão do biodiesel utilizando toda a região espectral das misturas.
Fonte: elaborado pela autora.

O segundo modelo construído para a quantificação do biodiesel, empregou somente as duas regiões espectrais (de 1100 cm^{-1} a 1200 cm^{-1} e de 1735 cm^{-1} a 1750 cm^{-1}) sugeridas pela proposta de norma ABNT. As regiões espectrais, e o desempenho do modelo empregando 5 variáveis latentes são visualizados nas Figuras 13 e 14, respectivamente.

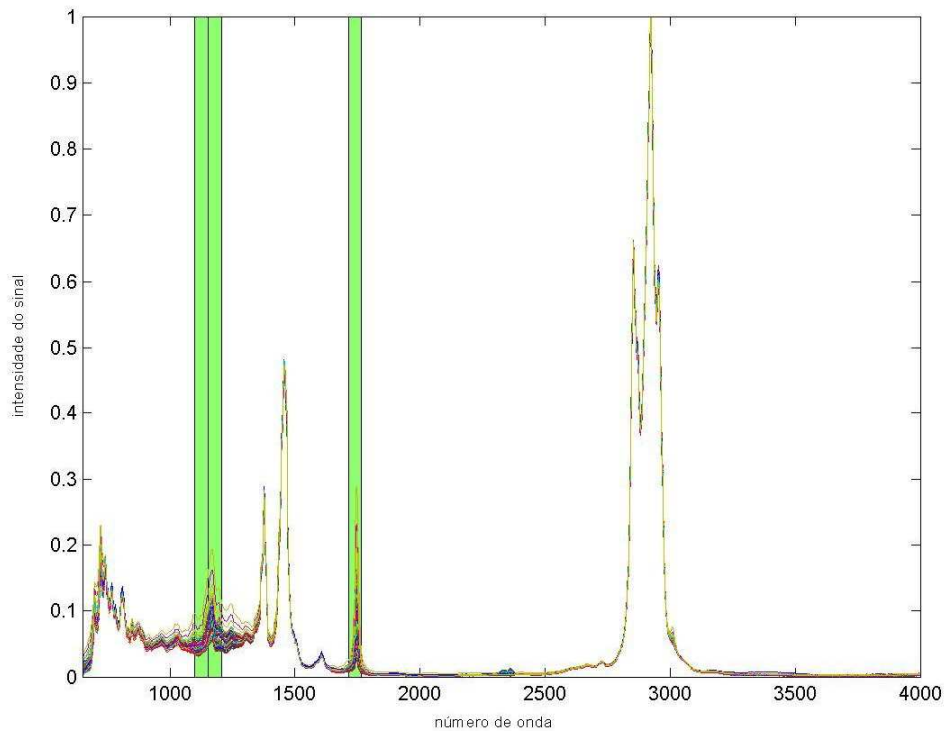


Figura 13 – Regiões espectrais sugeridas pela proposta de norma ABNT.
Fonte: elaborado pela autora.

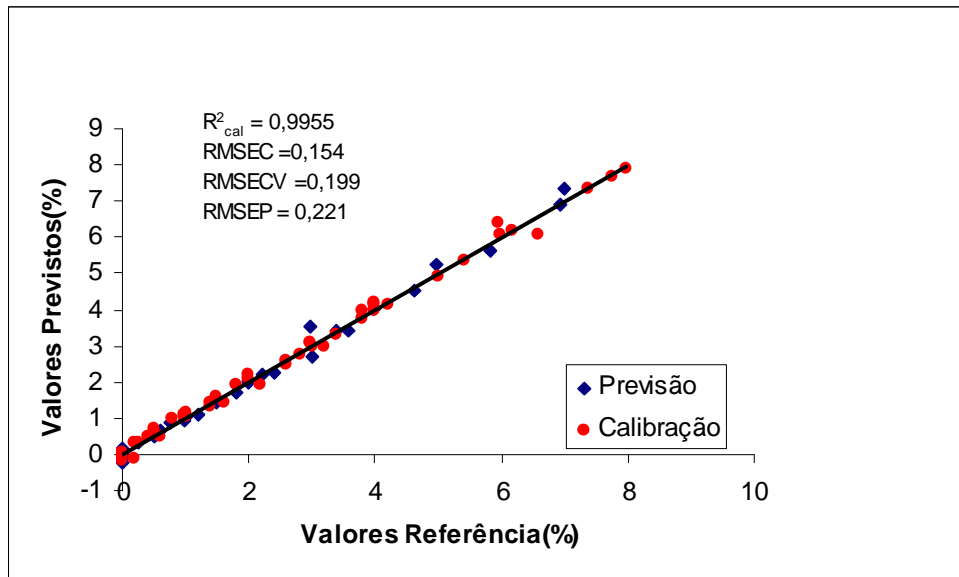


Figura 14– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT.

Fonte: elaborado pela autora.

Podemos observar que o modelo gerado utilizando o espectro inteiro, apresentou uma melhor capacidade preditiva, em relação ao modelo que levou em consideração somente às regiões sugeridas pela proposta de norma ABNT. Um aspecto que justifica este resultado pode ser a presença dos sinais interferentes do óleo nas mesmas regiões espectrais que o biodiesel, além de que outras regiões do espectro podem conter informações importantes para construção de um modelo mais robusto.

Tentando gerar modelos com maior capacidade preditiva para o analito de interesse, sugerimos a implementação de um algoritmo híbrido para a seleção das variáveis (número de onda) para construção dos modelos de regressão multivariados de espectros no infravermelho, que consiste na utilização do método iPLS, otimizado pela meta-heurística chamada enxame de partículas.

5.1.2 Resultados gerais e evoluções para a determinação do biodiesel

Com o propósito descrito anteriormente, foram implementadas cinco versões empregando a otimização por enxame de partículas: iPLS-PSO, iPLS-PSO (CC), iPLS-PSO (DD); iPLS-PSO (CD); iPLS-PSO (DC). Os principais resultados dos modelos, incluindo as implementações com a PSO, para quantificação do biodiesel são apresentados na Tabela 5.

Os resultados completos das implementações do enxame de partículas para quantificação do biodiesel encontram-se no ANEXO A.

Tabela 5– Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do biodiesel.

Modelos	Regiões empregadas	R² cal	RMSEC	RMSECV	RMSEP
espectro inteiro	650 - 4000 cm ⁻¹	0,9967	0,132	0,185	0,116
regiões sugeridas pela proposta de norma ABNT	1101 - 1198 cm ⁻¹ 1736 - 1749 cm ⁻¹	0,9955	0,154	0,199	0,221
iPLS-PSO	874 - 1319 cm ⁻¹ 2240 - 2717 cm ⁻¹ 3055 - 3109 cm ⁻¹ 3446 - 3556 cm ⁻¹	0,9997	0,038	0,095	0,088
iPLS-PSO (CC)	904 - 1327 cm ⁻¹ 2675 - 2756 cm ⁻¹ 3421 - 3502 cm ⁻¹	0,9995	0,052	0,105	0,085
iPLS-PSO (DD)	904 - 1327 cm ⁻¹ 2675 - 2756 cm ⁻¹ 3089 - 3170 cm ⁻¹ 3421 - 3502 cm ⁻¹	0,9996	0,048	0,103	0,086
iPLS-PSO (CD)	904 - 1327 cm ⁻¹ 2675 - 2756 cm ⁻¹ 3421 - 3502 cm ⁻¹	0,9995	0,052	0,105	0,085
iPLS-PSO (DC)	904 - 1327 cm ⁻¹ 2426 - 2756 cm ⁻¹ 3504 - 3585 cm ⁻¹	0,9991	0,070	0,103	0,093

Fonte: elaborado pela autora.

Na Figura 15 é apresentado um exemplo da evolução das partículas, mostrando as mesmas nas suas posições iniciais, bem como após 10 iterações, as posições das mesmas na 100^o iteração e finalmente quando atingem o número máximo de iterações definido para o problema. Pode-se observar que com o passar das iterações as partículas convergiram para uma solução.

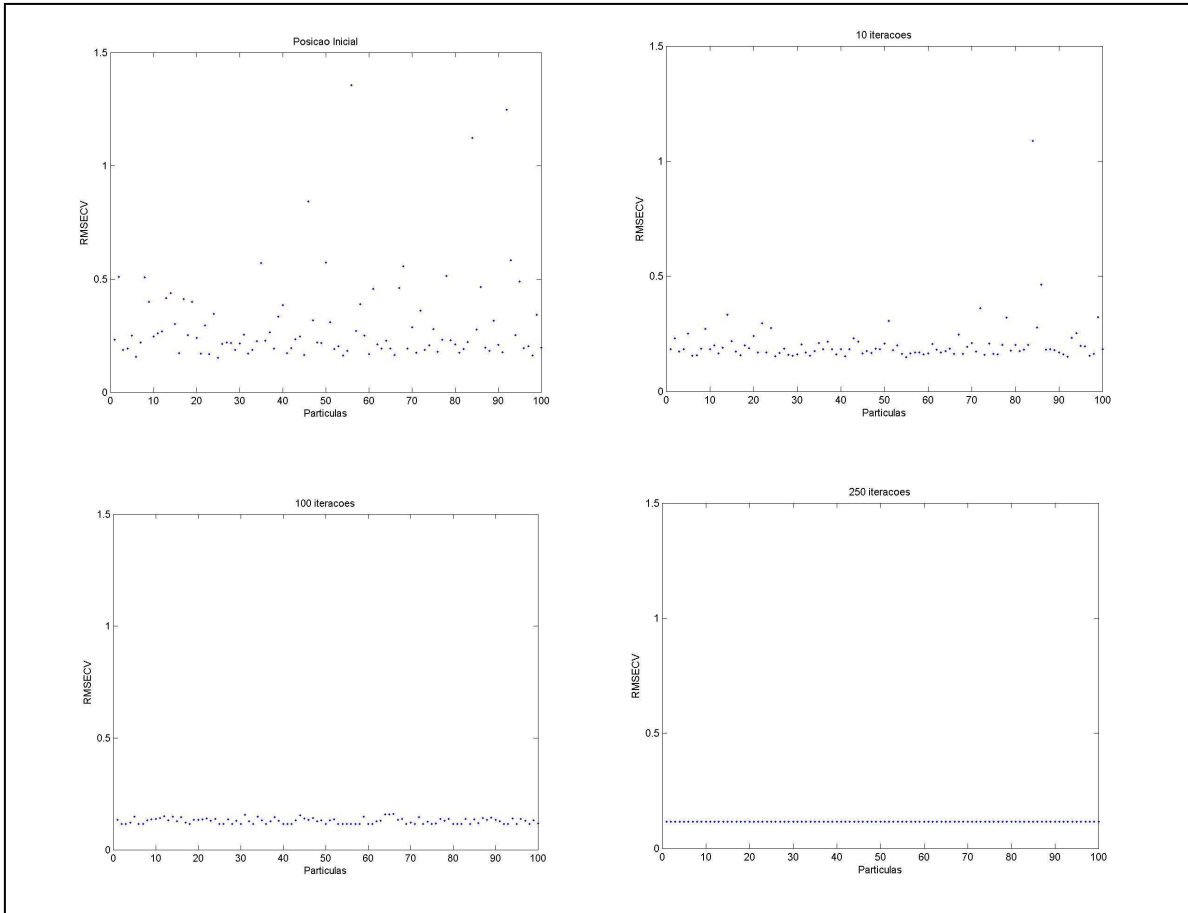


Figura 15 – Exemplo da evolução das partículas para a quantificação do biodiesel, usando o método iPLS-PSO (CD) com 40 intervalos e 100 partículas.

Fonte: elaborado pela autora.

A Figura 16 mostra um comparativo das evoluções para as 3 repetições para a combinação do iPLS-PSO com 20 intervalos e 50 partículas. Observa-se nesta proposta que existe uma tendência de convergência de todos os modelos, sendo necessário em média 44 iterações; mostrando juntamente que o número de iteração definido é suficiente para explorar as melhores combinações entre os 20 intervalos, atingindo um mínimo global.

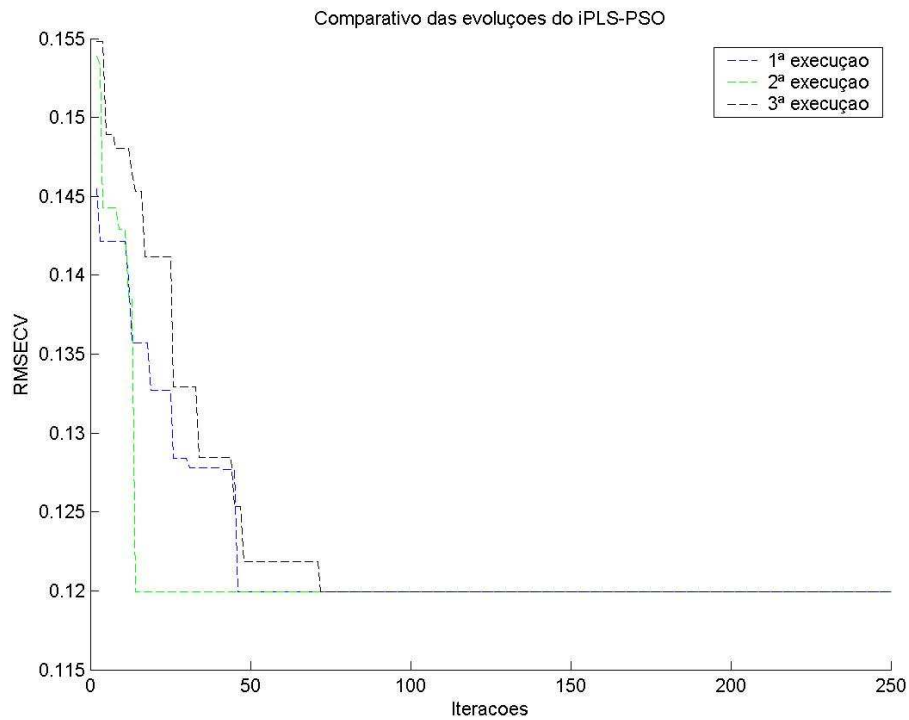


Figura 16 – Evolução das três execuções do iPLS-PSO com 20 intervalos e 50 partículas, para quantificação do biodiesel.
 Fonte: elaborado pela autora.

As Figuras 17 e 18 apresentam um comparativo de 3 evoluções do iPLS-PSO com 40 e 60 intervalos, respectivamente, ambos com 50 partículas. Observa-se um comportamento diferente para ambos, pois nem todos os modelos atingem o mínimo com poucas iterações. Em média para iPLS-PSO com 40 intervalos são necessárias 146 iterações, já para 60 intervalos em média 188. Em virtude disso, todos os resultados apresentados no ANEXO A foram obtidos com 250 iterações, pois um aumento no número de iterações refletiria num aumento considerável de custo computacional (tempo).

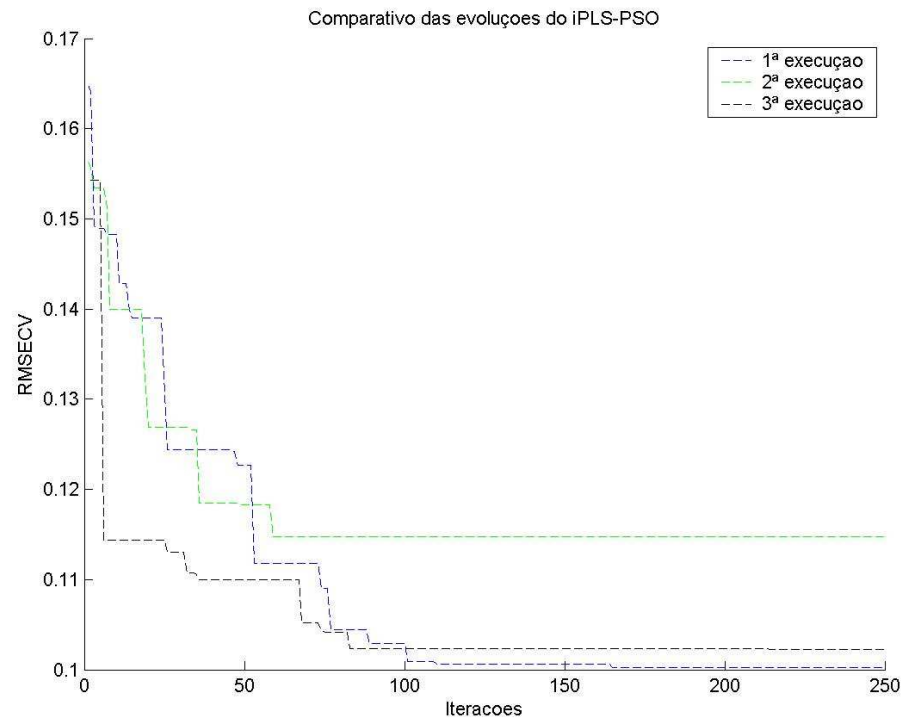


Figura 17 – Evolução das três execuções do iPLS-PSO com 40 intervalos e 50 partículas, para quantificação do biodiesel.
Fonte: elaborado pela autora.

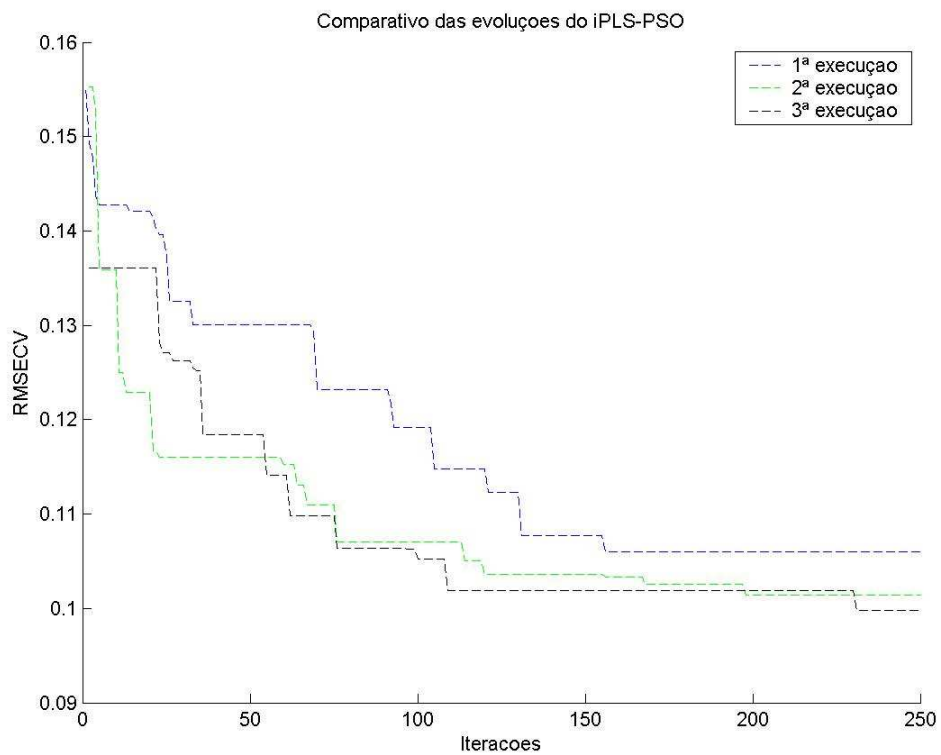


Figura 18 – Evolução das três execuções do iPLS-PSO com 60 intervalos e 50 partículas, para quantificação do biodiesel.
Fonte: elaborado pela autora.

5.1.3 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO

A aplicação da primeira implementação denominada iPLS-PSO no conjunto de dados propostos para a quantificação do biodiesel, selecionou regiões espectrais para construção do modelo que compreendem de 874 a 1319 cm^{-1} com lacunas, de 2240 a 2717 cm^{-1} , de 3055 a 3109 cm^{-1} e de 3446 a 3556 cm^{-1} como podem ser visualizadas na Figura 19. Dentre todos os modelos construídos para este método proposto, o selecionado foi o obtido com 60 intervalos e 100 partículas e o desempenho do modelo empregando 10 variáveis latentes para quantificação do biodiesel, é apresentada na Figura 20.

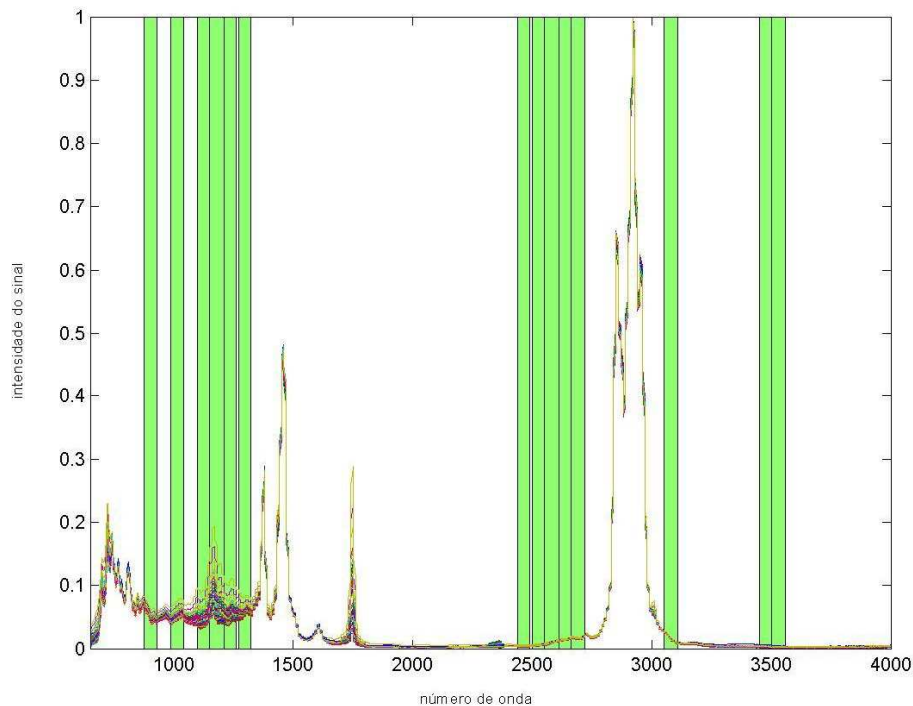


Figura 19 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel usando o método iPLS-PSO com 60 intervalos e 100 partículas.
Fonte: elaborado pela autora.

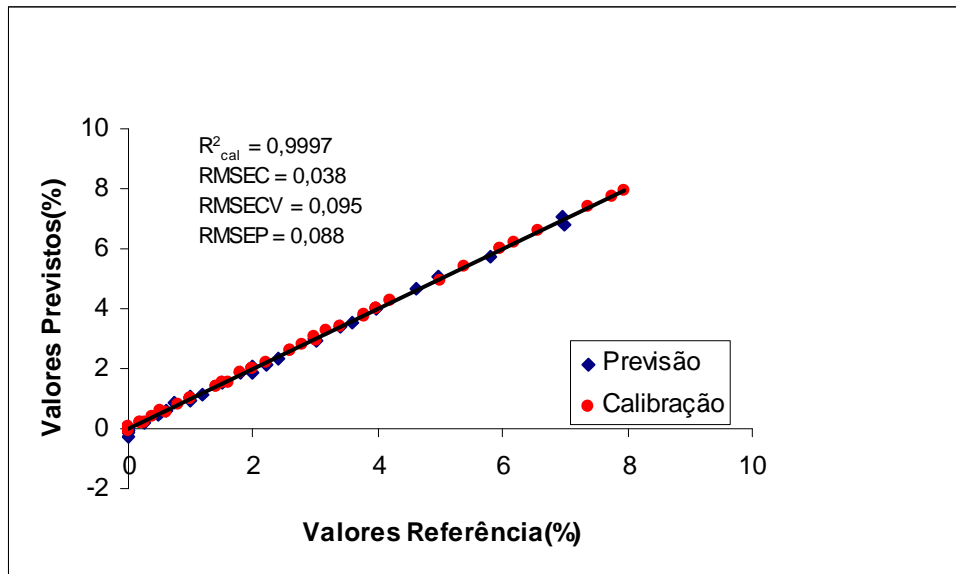


Figura 20 – Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 60 intervalos e 100 partículas.
 Fonte: elaborado pela autora.

5.1.4 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CC)

Já o método iPLS-PSO (CC) selecionou regiões espectrais que compreendem as faixas de 904 a 1327 cm^{-1} , de 2675 a 2756 cm^{-1} e de 3421 a 3502 cm^{-1} , conforme Figura 21. Nesta implementação o modelo selecionado empregou 40 intervalos e 25 partículas e, a sua respectiva curva de calibração utilizando 10 variáveis latentes para a concentração de biodiesel, é apresentada na Figura 22.

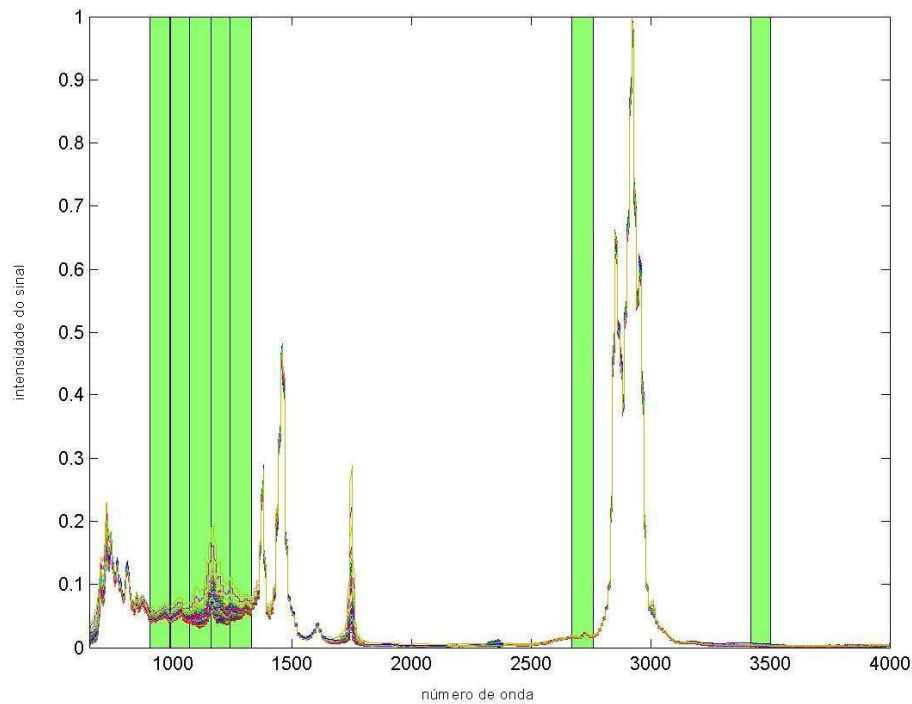


Figura 21 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

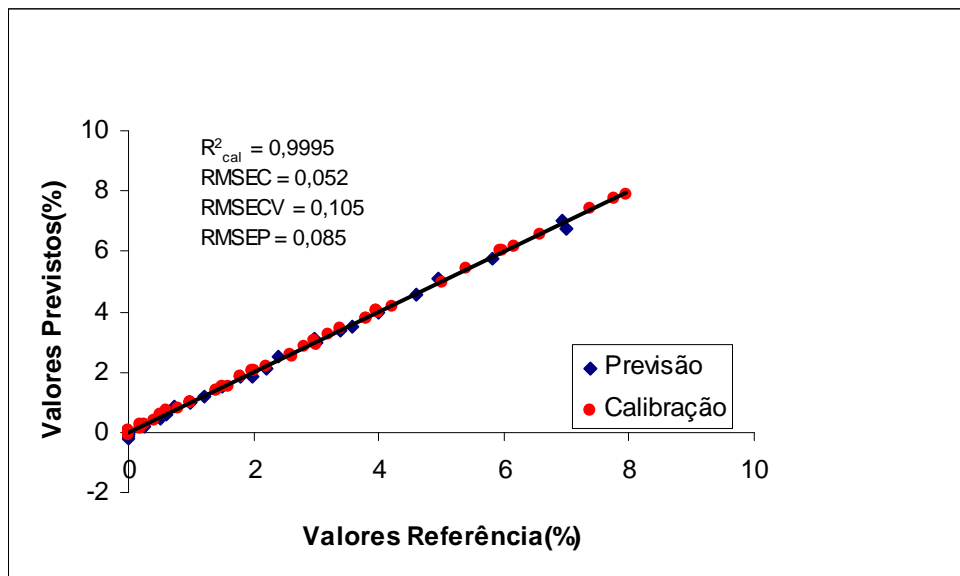


Figura 22 – Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.1.5 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (DD)

No caso da implementação iPLS-PSO (DD) as regiões espectrais selecionadas compreendem as faixas de 904 a 1327 cm^{-1} , de 2675 a 2756 cm^{-1} , de 3089 a 3170 cm^{-1} e de 3421 a 3502 cm^{-1} , conforme Figura 23. Neste caso, o modelo selecionado foi aquele construído com 40 intervalos e 100 partículas, e a curva de calibração, utilizando 10 variáveis latentes, para a concentração de biodiesel é apresentada na Figura 24.

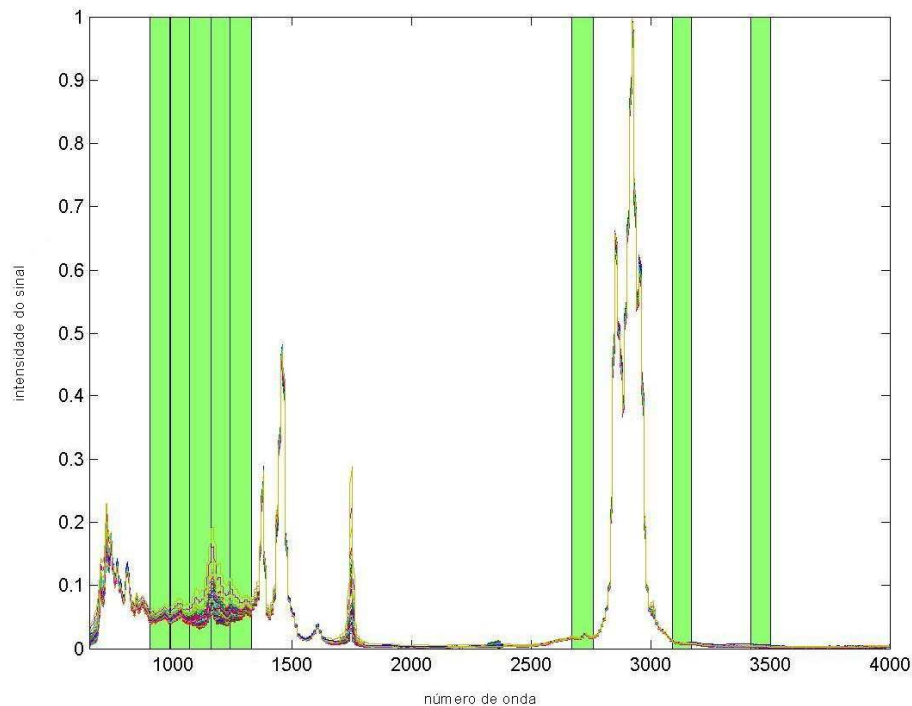


Figura 23– Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

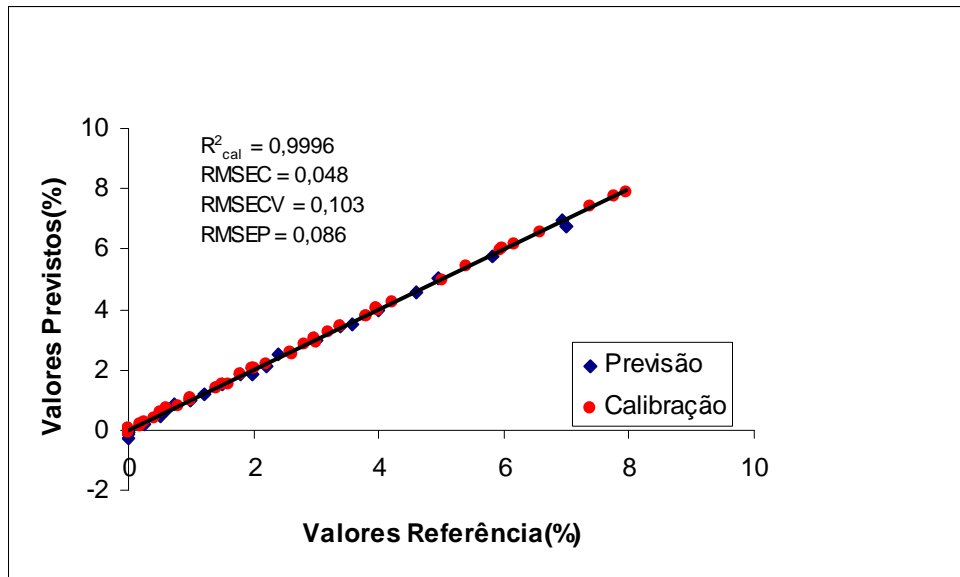


Figura 24– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

Um aspecto interessante nesta implementação é que conforme ANEXO A, na construção dos modelos com 20 intervalos, as partículas convergiram muito rápido para uma solução, não precisando ser explorado todo o espaço de busca, convergindo em média após 33 iterações.

5.1.6 Resultados obtidos para a determinação do biodiesel aplicando o método iPLS-PSO (CD)

Conforme a Tabela 5, o modelo selecionado pelo método iPLS-PSO (CD) corresponde as mesmas regiões espectrais que a implementação do método iPLS-PSO (CC). As regiões espectrais bem como a curva de calibração pode ser vistas nas Figuras 21 e 22, anteriormente apresentadas.

5.1.7 Resultados obtidos para a determinação do biodiesel aplicando o iPLS-PSO (DC)

Neste método iPLS-PSO (DC) as regiões espectrais selecionadas para construção do modelo compreendem as faixas de 904 a 1327 cm^{-1} com lacunas, de 2426 a 2756 cm^{-1} e,

de 3504 a 3585 cm^{-1} conforme Figura 25. Dentre todos os modelos construídos para este método o modelo selecionado foi o construído com 40 intervalos e 100 partículas, e a curva de calibração utilizando 8 variáveis latentes, para a concentração de biodiesel, é apresentada na Figura 26.

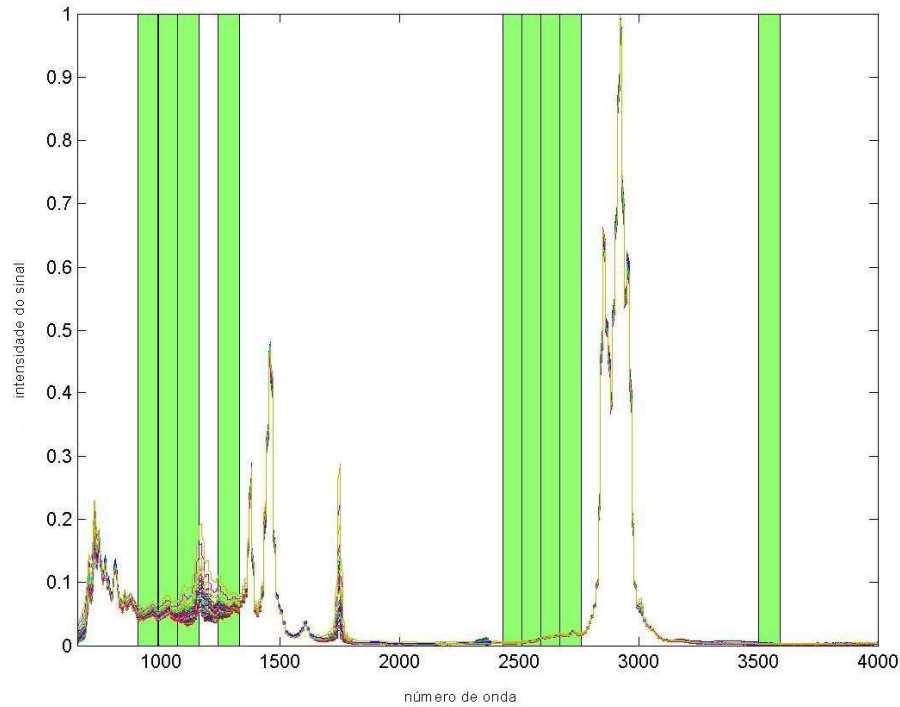


Figura 25 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

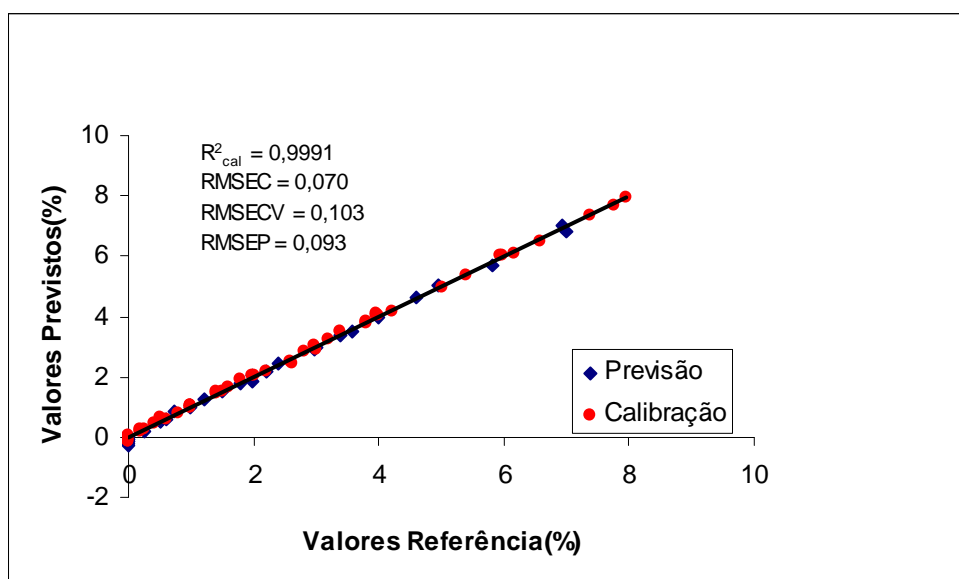


Figura 26– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

5.1.8 Considerações sobre os modelos para a previsão do biodiesel

Conforme evidenciado, todos os modelos construídos para quantificação do biodiesel empregam regiões espectrais semelhantes, e sendo que todos foram capazes de quantificar o analito proposto. Os resultados encontrados empregando as soluções propostas pelo método enxame de partículas, em comparação aos resultados empregando o PLS sobre todo o espectro bem como as regiões espectrais sugeridas pela proposta de norma ABNT, demonstraram que houve uma otimização dos modelos de regressão multivariados.

A região selecionada pelos modelos otimizados contempla em parte a região espectral, sugerida pela proposta de norma, e ampliada principalmente na faixa dos sinais referentes ao C-O. Um aspecto interessante é que em nenhum dos modelos otimizados, selecionou a região da carbonila (C=O), região esta que a proposta de norma preconiza como sendo a principal banda para a quantificação do biodiesel.

Uma hipótese é que a região da carbonila não apresenta informações suficientes para diferenciar a contribuição do sinal do biodiesel e do óleo isoladamente, ou seja, provavelmente as modificações que acontecem do óleo para o biodiesel são pouco pronunciadas nesta região e muito mais pronunciadas na região do sinal correspondente ao C-O.

Com base nestas observações e considerando alguns dos resultados das otimizações propostas pelo método enxame de partículas, é possível obter modelos com boa capacidade preditiva com as informações espectrais das regiões de 904 a 1327 cm^{-1} com lacunas, ilustradas na Figura 27, obtidas pelo método iPLS-PSO em uma de suas soluções. O desempenho deste modelo, com 10 variáveis latentes, pode ser visualizado na Figura 28.

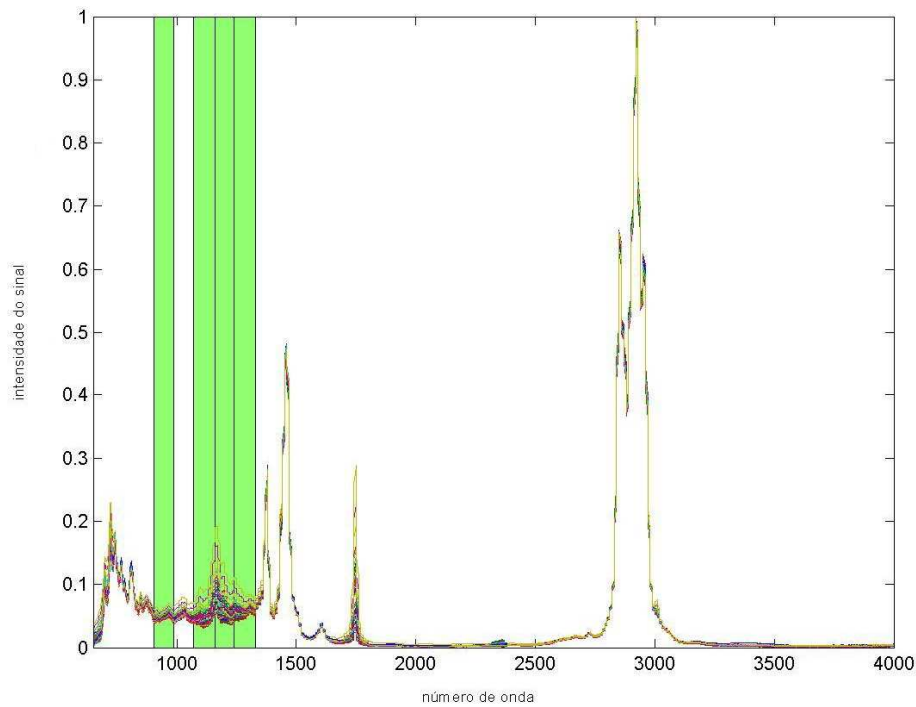


Figura 27 – Regiões espectrais utilizadas na modelagem para a quantificação do biodiesel, usando o método iPLS-PSO com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

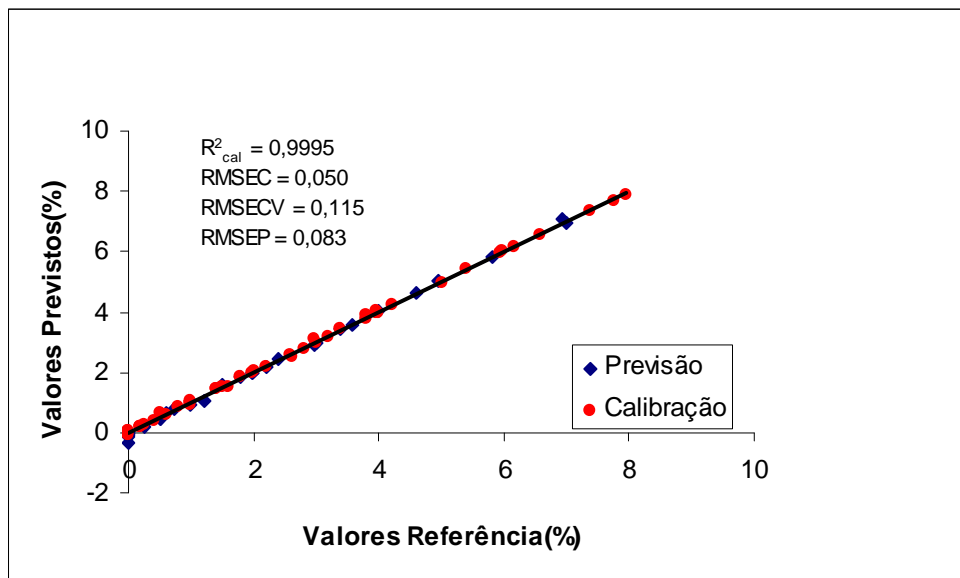


Figura 28– Valores de referência *versus* previstos para o modelo de previsão do biodiesel, empregando o método iPLS-PSO com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

5.2 Resultados para a quantificação do óleo vegetal

Neste trabalho também foram desenvolvidos diversos modelos de regressão visando quantificar o óleo vegetal presente nas misturas, com o intuito de verificar a capacidade de detecção de adulterantes nas misturas biodiesel/óleo vegetal/diesel através do uso do FT-IR/ATR. Os principais resultados obtidos estão descritos na seqüência. Primeiramente são apresentados os modelos obtidos sem a otimização através da meta-heurística implementada e em seguida, são descritos os modelos obtidos pela combinação da PSO e do iPLS.

5.2.1 Modelos para determinação do óleo vegetal com espectro inteiro e segundo as regiões sugeridas pela proposta de norma ABNT

Com base no conjunto de dados foram desenvolvidos dois modelos de regressão, ambos utilizando para construção o método PLS. O primeiro modelo, para a quantificação do óleo vegetal foi construído utilizando toda região espectral (650 cm^{-1} a 4000 cm^{-1}) conforme Figura 11, apresentada anteriormente, e o desempenho do modelo é apresentado na Figura 29, empregando 6 variáveis latentes.

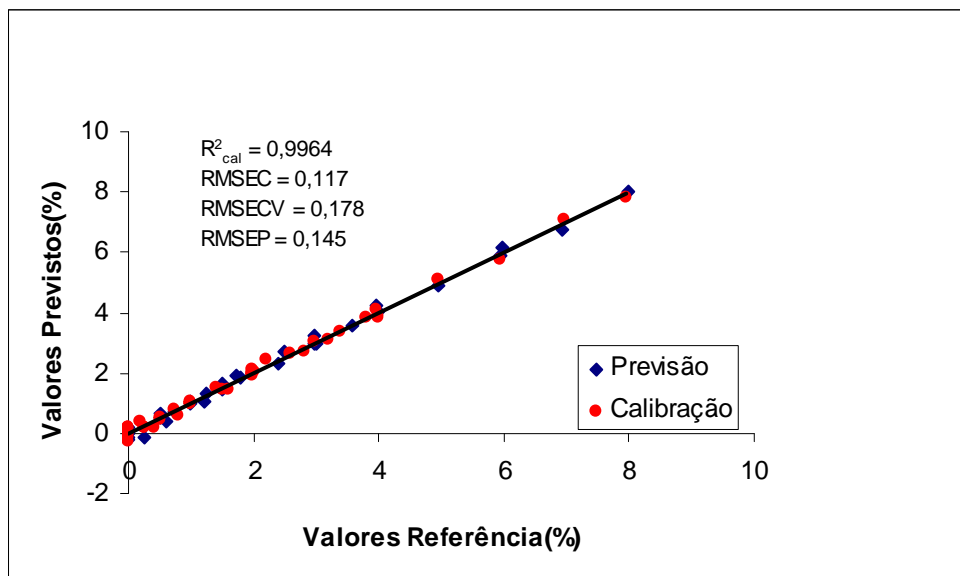


Figura 29 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal utilizando toda a região espectral das misturas.
 Fonte: elaborado pela autora.

O segundo modelo construído para a quantificação do óleo vegetal nas misturas empregou somente as duas regiões espectrais (de 1100 cm^{-1} a 1200 cm^{-1} e de 1735 cm^{-1} a 1750 cm^{-1}) sugeridas pela proposta de norma ABNT, conforme Figura 13. O desempenho do modelo, utilizando 5 variáveis latentes, é apresentado na Figura 30.

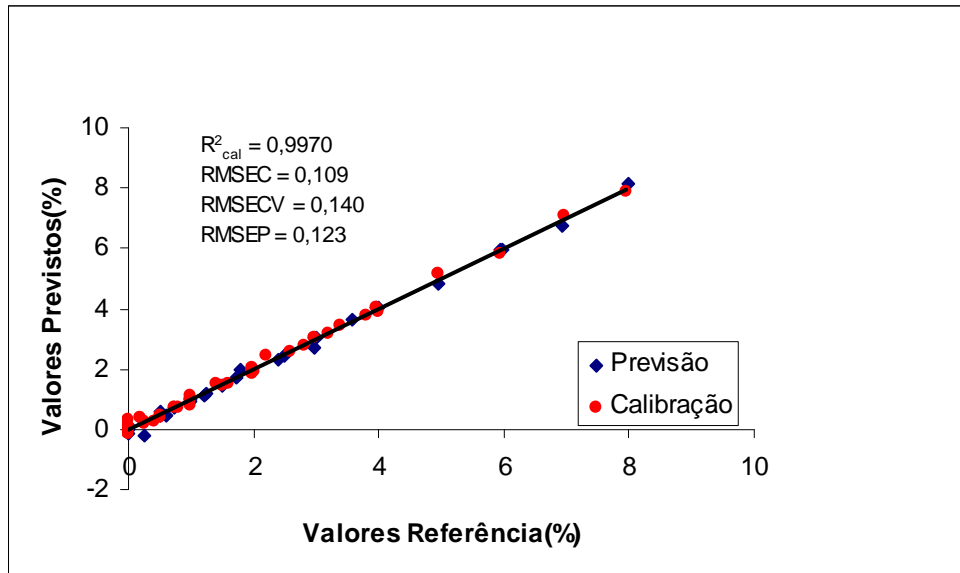


Figura 30 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT. Fonte: elaborado pela autora.

Observando os dois modelos anteriores pode-se verificar que os resultados obtidos pelo modelo construído, empregando somente as regiões espectrais sugerida pela proposta de norma ABNT, apresenta uma melhor capacidade de quantificação do óleo vegetal do que o modelo construído com o mesmo propósito, empregado toda a região do espectro.

Tentando analisar a possibilidade de obter modelos com melhor capacidade preditiva, sugere-se a mesma implementação proposta para construção dos modelos de biodiesel, desta vez com o propósito de otimizar a quantificação do óleo vegetal nas misturas.

5.2.2 Resultados gerais e evoluções dos modelos para a previsão do óleo vegetal

Foram testadas na quantificação do óleo vegetal, as mesmas cinco implementações da otimização por enxame de partículas: iPLS-PSO, iPLS-PSO (CC), iPLS-PSO (DD); iPLS-PSO (CD); iPLS-PSO (DC), descritas anteriormente. Os principais resultados dos modelos, incluindo os obtidos através das implementações com a PSO, são apresentados na Tabela

6. Os resultados completos das implementações do enxame de partículas para a quantificação do óleo vegetal encontram-se no ANEXO B.

Tabela 6 – Principais resultados dos modelos incluindo as cinco implementações da PSO para a quantificação do óleo vegetal.

Modelos	Regiões empregadas	R² cal	RMSEC	RMSECV	RMSEP
espectro inteiro	650 - 4000 cm ⁻¹	0,9964	0,117	0,178	0,145
regiões sugeridas pela proposta de norma ABNT	1101 - 1198 cm ⁻¹ 1736 - 1749 cm ⁻¹	0,9970	0,109	0,140	0,123
iPLS- PSO	904 - 1412 cm ⁻¹ 1583 - 1751 cm ⁻¹ 2509 - 2756 cm ⁻¹ 3504 - 3585 cm ⁻¹	0,9991	0,059	0,083	0,092
iPLS- PSO (CC)	904 - 1412 cm ⁻¹ 1583 - 1751 cm ⁻¹ 1838 - 2006 cm ⁻¹	0,9990	0,060	0,088	0,089
iPLS- PSO (DD)	904 - 1327 cm ⁻¹ 1583 - 1666 cm ⁻¹ 2008 - 2341 cm ⁻¹	0,0995	0,042	0,082	0,089
iPLS- PSO (CD)	804 - 1319 cm ⁻¹ 1992 - 2158 cm ⁻¹ 2495 - 2829 cm ⁻¹ 3334 - 3500 cm ⁻¹	0,9995	0,041	0,081	0,094
iPLS- PSO (DC)	820 - 1751 cm ⁻¹ 2426 - 2756 cm ⁻¹ 3007- 3253 cm ⁻¹	0,9992	0,056	0,085	0,090

Fonte: elaborado pela autora.

Um comparativo das evoluções para 3 execuções do método iPLS-PSO com 20 intervalos e 50 partículas para quantificação do óleo vegetal, é apresentada na Figura 31. Observa-se nesta proposta que existe uma tendência de convergência de todos os modelos, sendo necessário em média de 108 iterações; mostrando juntamente que o número total de iterações definido é suficiente para explorar todas as melhores combinações possíveis entre os 20 intervalos.

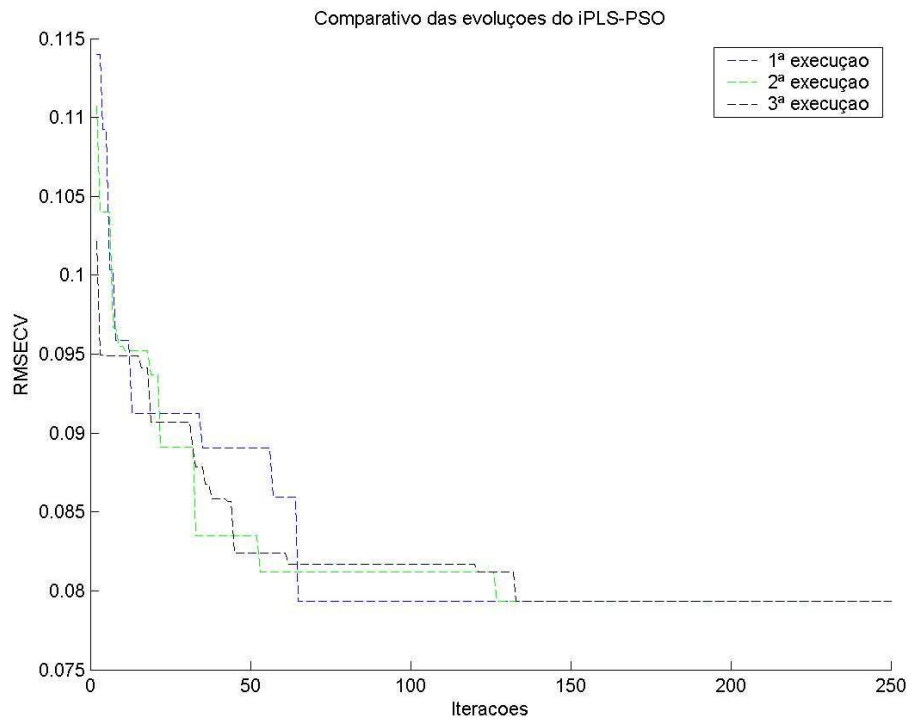


Figura 31– Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas, para a quantificação do óleo vegetal.
Fonte: elaborado pela autora.

Observa-se um comportamento diferente quando se apresentam as 3 evoluções do iPLS-PSO com 40 e 60 intervalos, ambos com 50 partículas, como pode ser visto nas Figuras 32 e 33. São necessários mais iterações para que os modelos atinjam mínimos locais, em média para iPLS-PSO com 40 intervalos são necessárias 184 iterações, já para 60 intervalos em média 209 iterações.

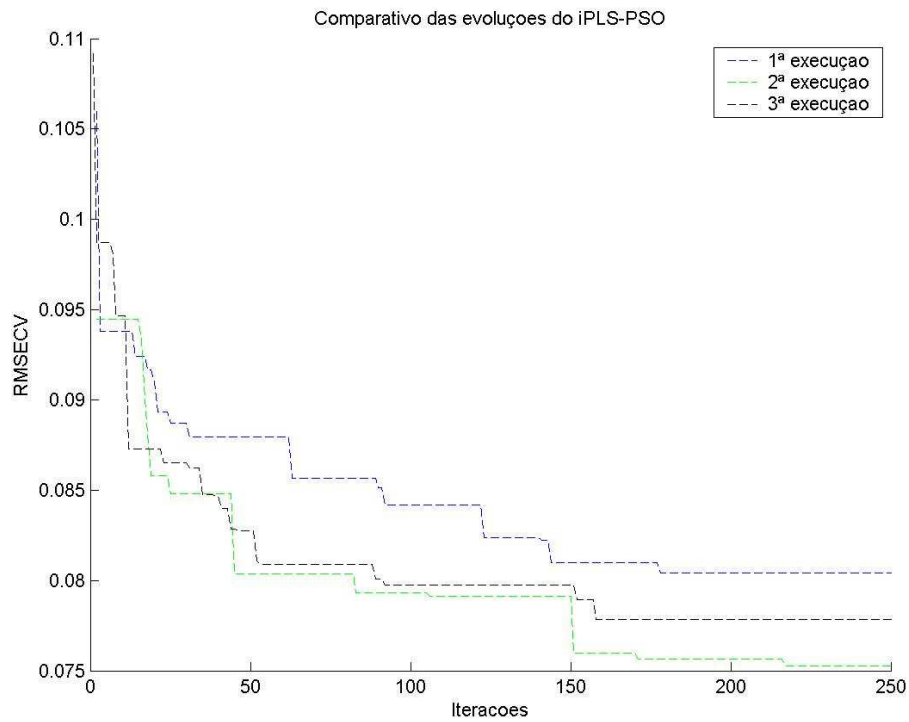


Figura 32 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas, para a quantificação do óleo vegetal.
Fonte: elaborado pela autora.

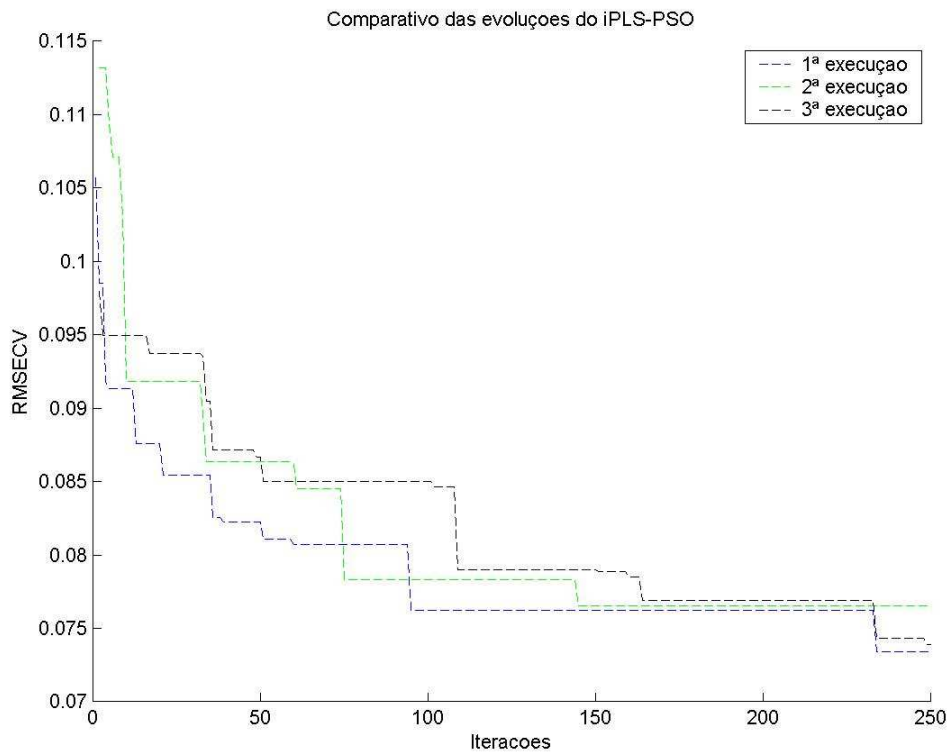


Figura 33 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para a quantificação do óleo vegetal.
Fonte: elaborado pela autora.

5.2.3 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO

Nesta primeira implementação denominada iPLS-PSO aplicada ao conjunto de dados propostos para quantificar o óleo vegetal, as regiões espectrais selecionadas para construção do modelo foram as que compreendem de 904 a 1412 cm^{-1} com lacunas, de 1583 a 1751 cm^{-1} , de 2509 a 2756 cm^{-1} e de 3504 a 3585 cm^{-1} como podem ser visualizadas na Figura 34. Dentre todos os modelos construídos, o selecionado foi o construído com 40 intervalos e 25 partículas e a sua respectiva curva de calibração, com 9 variáveis latentes, para quantificação do óleo vegetal, é apresentada na Figura 35.

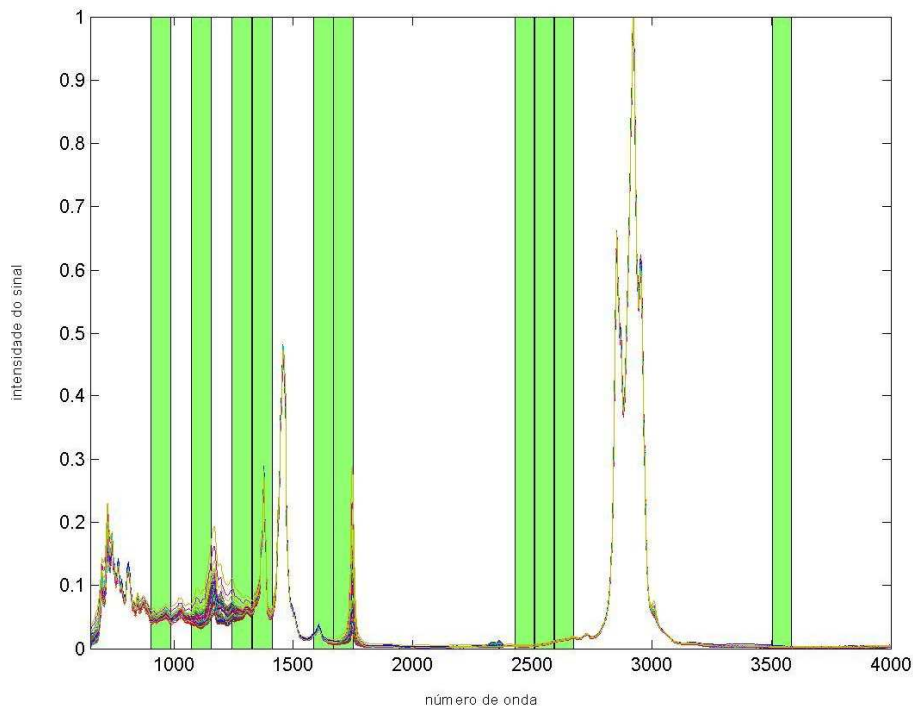


Figura 34 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

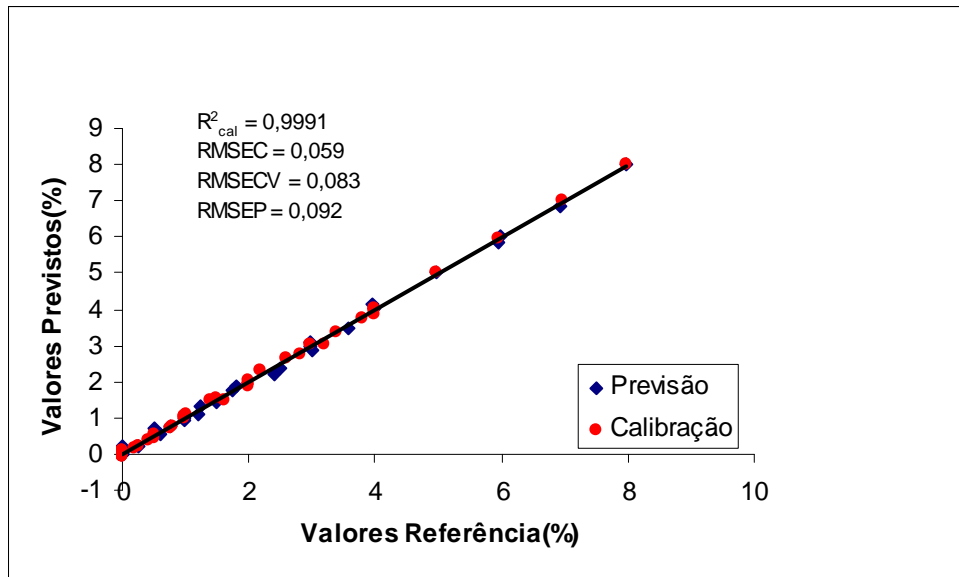


Figura 35 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.
 Fonte: elaborado pela autora.

5.2.4 Resultados obtido para a determinação do óleo vegetal aplicando o método iPLS-PSO (CC)

Na implementação iPLS-PSO (CC) as regiões espectrais selecionadas para construção do modelo de quantificação do óleo vegetal compreendem as faixas de 904 a 1412 cm^{-1} com lacunas, de 1583 a 1666 cm^{-1} e de 2008 a 2341 cm^{-1} , conforme Figura 36. Nesta abordagem, o modelo selecionado empregou 40 intervalos e 25 partículas e a sua respectiva curva de calibração, com 9 variáveis latentes, para a concentração do óleo vegetal, é apresentada na Figura 37.

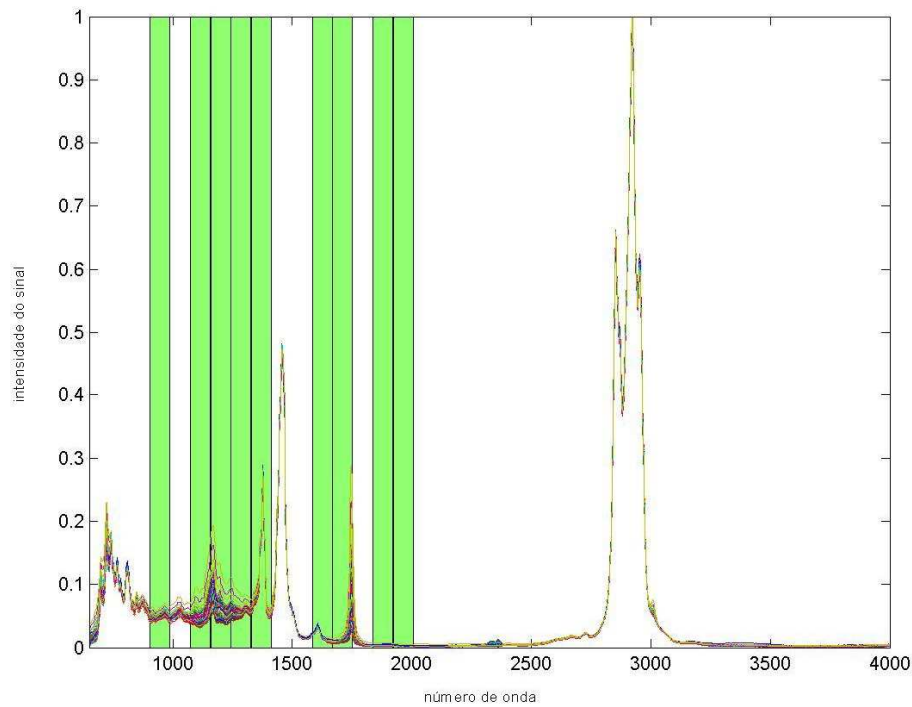


Figura 36 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

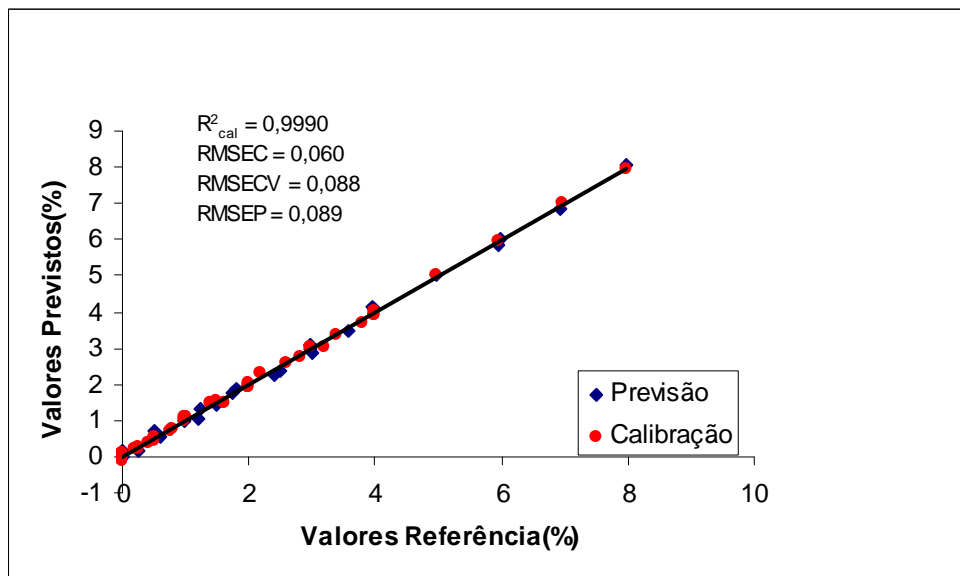


Figura 37 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.2.5 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DD)

Já o método iPLS-PSO (DD) para a quantificação do óleo vegetal nas misturas, o modelo selecionado foi aquele construído com 40 intervalos e 100 partículas, compreendendo as regiões espectrais selecionadas nas faixas de 904 a 1327 cm^{-1} com lacunas, de 1583 a 1666 cm^{-1} e de 2008 a 2341 cm^{-1} com lacunas, conforme Figura 38, e a curva de calibração, com 10 variáveis latentes, para a concentração do óleo vegetal é apresentada na Figura 39.

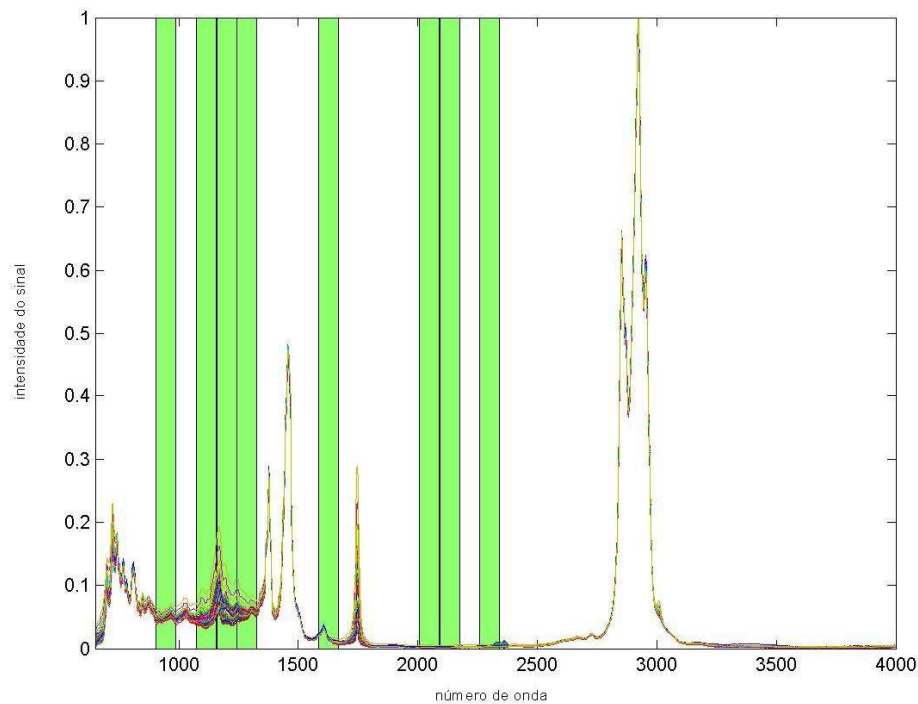


Figura 38 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

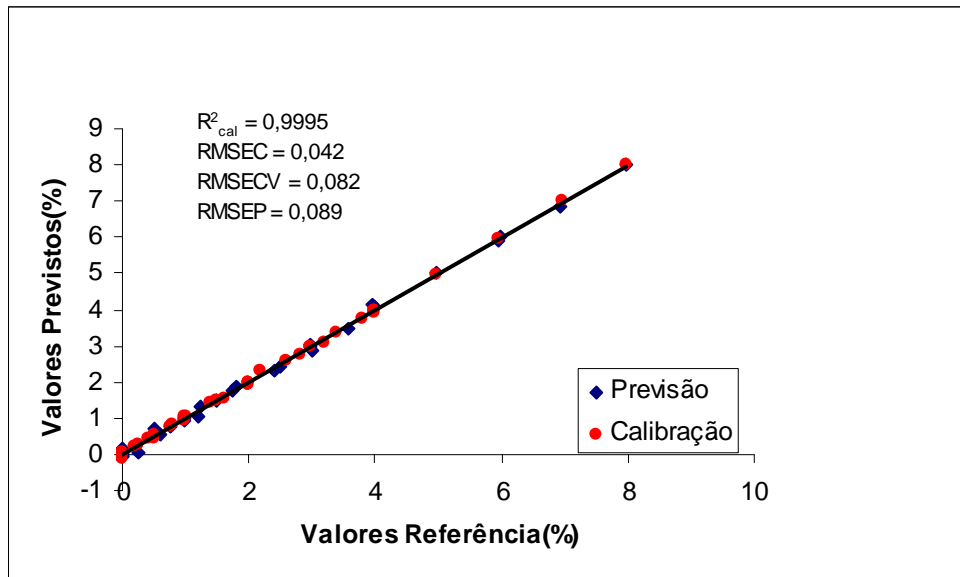


Figura 39 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas. Fonte: elaborado pela autora.

5.2.6 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (CD)

No caso da implementação iPLS-PSO (CD) as regiões espectrais selecionadas compreendem as faixas de 804 a 1319 cm^{-1} , de 1992 a 2158 cm^{-1} , de 2495 a 2829 cm^{-1} e de 3334 a 3500 cm^{-1} , conforme Figura 40. Neste caso, o modelo selecionado, dentre todos os modelos construídos para esta abordagem, foi aquele elaborado com 20 intervalos e 25 partículas, e a curva de calibração, com 10 variáveis latentes, para a concentração do óleo vegetal é apresentada na Figura 41.

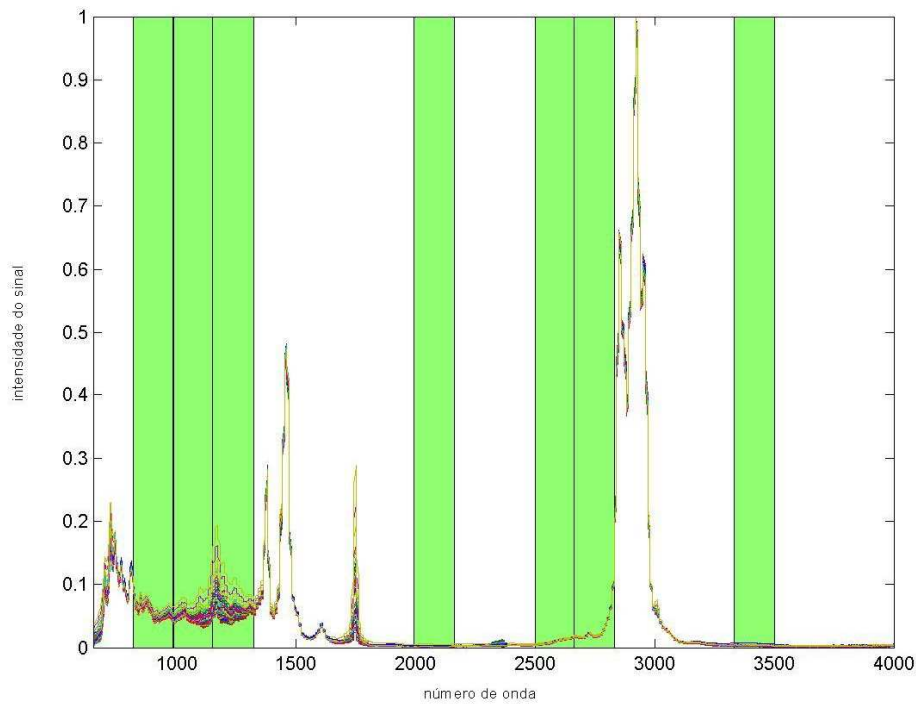


Figura 40 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.
Fonte: elaborado pela autora.

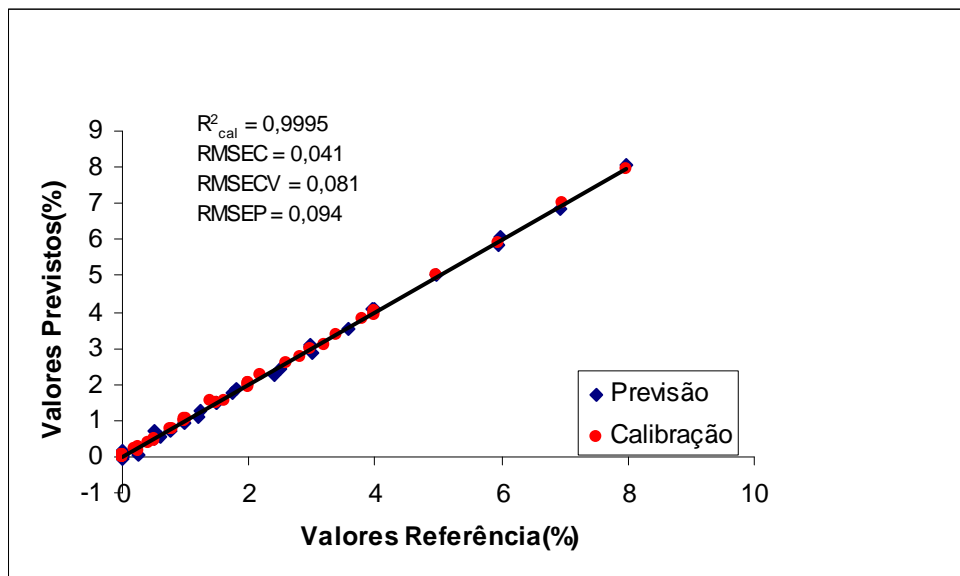


Figura 41– Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (CD) com 20 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.2.7 Resultados obtidos para a determinação do óleo vegetal aplicando o método iPLS-PSO (DC)

Já a implementação denominada iPLS-PSO (DC) selecionou as seguintes regiões espectrais para construção do modelo de 820 a 1751 cm^{-1} , de 2426 a 2756 cm^{-1} e de 3007 a 3253 cm^{-1} como podem ser visualizadas na Figura 42. Dentre todos os modelos construídos para iPLS-PSO (DC), o selecionado foi aquele construído com 40 intervalos e 100 partículas e a sua respectiva curva de calibração com 10 variáveis latentes, para quantificação do óleo vegetal, é apresentada na Figura 43.

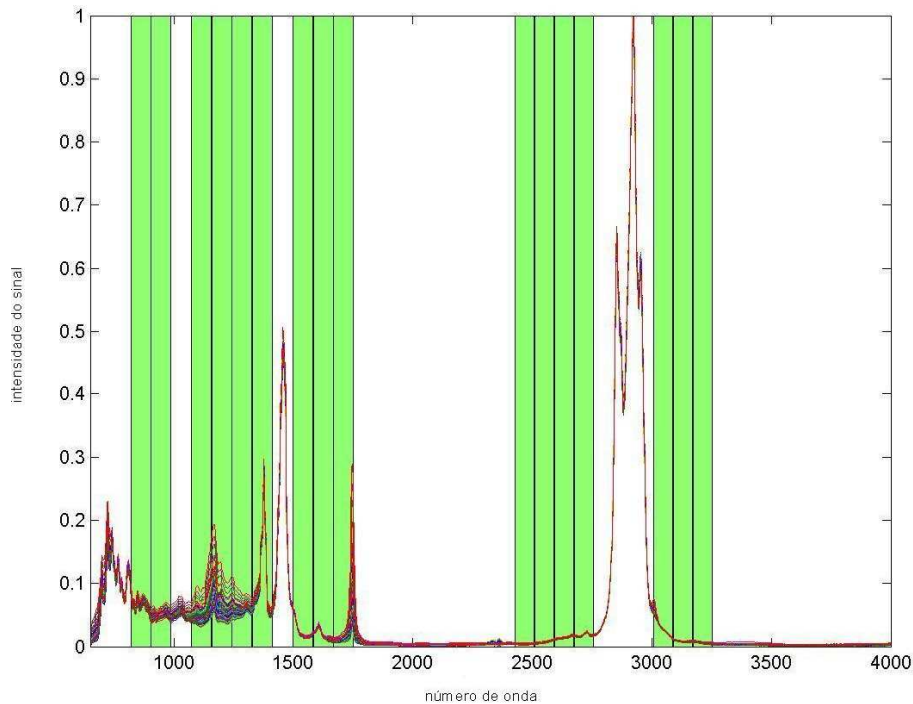


Figura 42 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo vegetal, usando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

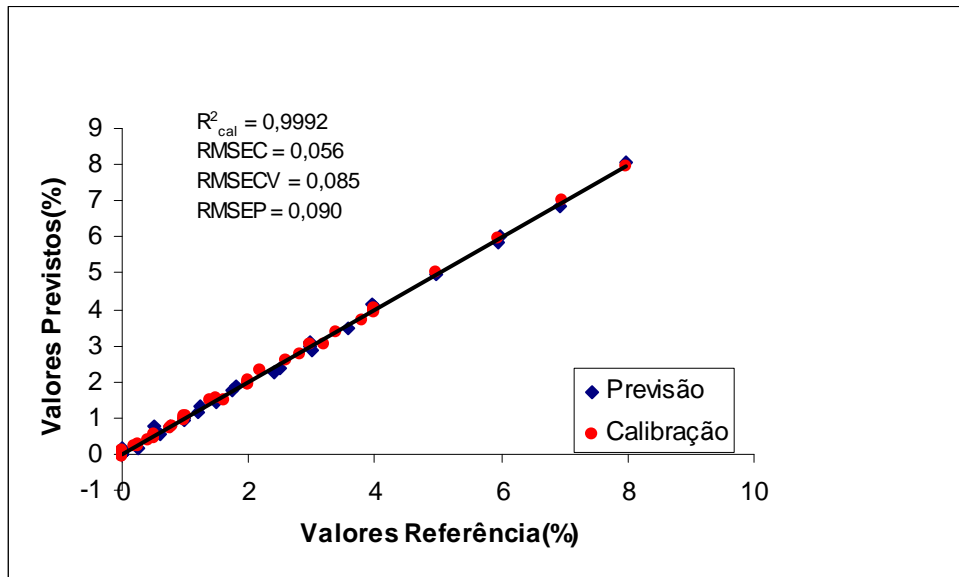


Figura 43 – Valores de referência *versus* previstos para o modelo de previsão do óleo vegetal, empregando o método iPLS-PSO (DC) com 40 intervalos e 100 partículas. Fonte: elaborado pela autora.

5.2.8 Considerações sobre os modelos para a previsão de óleo vegetal

Analisando os resultados anteriores, podemos observar que todos os modelos construídos para quantificar o óleo vegetal foram capazes de atingir este propósito, apresentando baixos erros de previsão (RMSEP). Cabe ressaltar que os modelos construídos empregando o método PLS com o espectro inteiro das misturas e somente com as regiões sugeridas pela proposta de norma ABNT apresentaram uma capacidade de quantificação do óleo vegetal menor do que os modelos construídos através do método híbrido proposto nesta pesquisa.

Os modelos selecionados através das otimizações aqui propostas apresentaram valores de RMSEP entre 0,089 a 0,094, enquanto que os valores de RMSECV ficaram na amplitude de 0,081 a 0,088.

Do ponto de vista das regiões selecionadas, o modelo para previsão do óleo vegetal além de necessitar da região do C-O, também necessita utilizar as informações referentes ao sinal da carbonila (C=O), conforme as figuras anteriores, que apresentam as regiões selecionadas.

5.3 Resultados para a quantificação do óleo diesel

Finalmente, o outro analito quantificado através de modelos de regressão multivariados nas misturas, neste trabalho foi o óleo diesel. Embora não seja comum quantificar o componente majoritário das misturas, o propósito desta determinação é verificar o comportamento do algoritmo e da técnica FT-IR/ATR. Para isso foram desenvolvidos diversos modelos de regressão. Neste caso, foram construídos modelos usando o PLS com o espectro inteiro das misturas e posteriormente modelos com as implementações propostas com a meta-heurística enxame de partículas em conjunto com o iPLS.

5.3.1 Modelo para a determinação do óleo diesel com espectro inteiro

Para a determinação do óleo diesel também foi desenvolvido um modelo de regressão com o método PLS, utilizando toda a região espectral (650 cm^{-1} a 4000 cm^{-1}) conforme Figura 11 apresentada anteriormente. O desempenho do modelo, com 4 variáveis latentes, é apresentado na Figura 44.

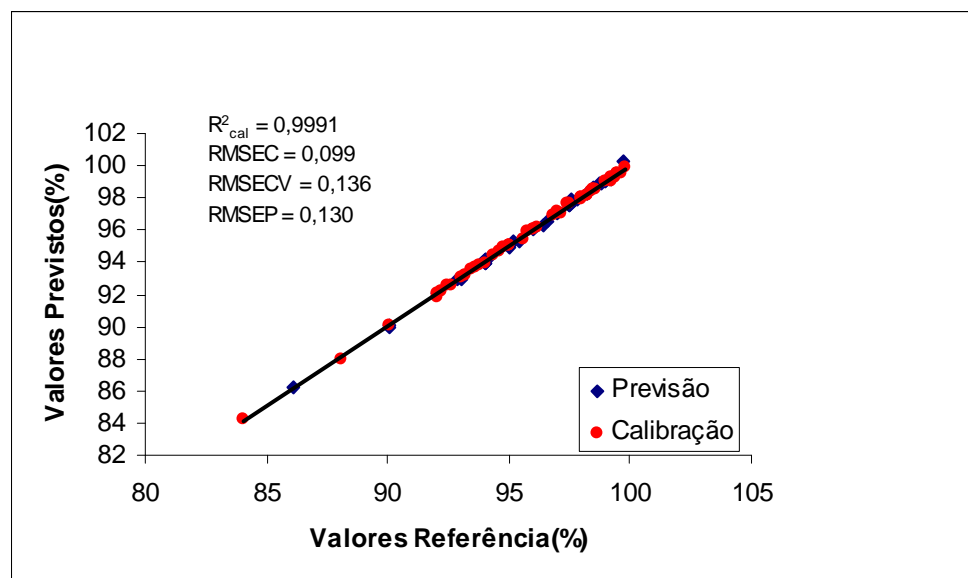


Figura 44 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel utilizando toda a região espectral da mistura.
Fonte: elaborado pela autora.

5.3.2 Resultados gerais e evoluções para a determinação do óleo diesel

Tentando avaliar se o comportamento das implementações propostas pelo enxame de partículas seria capaz gerar modelos mais robustos para quantificação do óleo diesel, foram testadas as cinco implementações na quantificação deste analito. Os principais resultados dos modelos, incluindo as resultantes das implementações com a PSO são apresentados na Tabela 7.

Tabela 7 – Principais resultados dos modelos, incluindo das cinco implementações da PSO para a quantificação do óleo diesel.

Modelos	Regiões empregadas	R ² cal	RMSEC	RMSECV	RMSEP
espectro inteiro	650 - 4000 cm ⁻¹	0,9991	0,099	0,136	0,130
iPLS- PSO	735 - 903 cm ⁻¹	0,9998	0,048	0,098	0,102
iPLS- PSO (CC)	1159 - 1327 cm ⁻¹ 3172 - 3253 cm ⁻¹ 762 - 1207 cm ⁻¹	0,9998	0,047	0,099	0,108
iPLS- PSO (DD)	2719 - 2773 cm ⁻¹ 3223 - 3332 cm ⁻¹ 3446 - 3500 cm ⁻¹ 904 - 1327 cm ⁻¹	0,9997	0,059	0,096	0,112
iPLS- PSO (CD)	2092 - 2175 cm ⁻¹ 2426 - 2756 cm ⁻¹ 3089 - 3170 cm ⁻¹ 3504 - 3585 cm ⁻¹ 3753 - 3834 cm ⁻¹ 904 - 1327 cm ⁻¹	0,9999	0,038	0,095	0,108
iPLS- PSO (DC)	1753 - 1836 cm ⁻¹ 2260 - 2507 cm ⁻¹ 2675 - 2756 cm ⁻¹ 3753 - 3834 cm ⁻¹ 735 - 1327 cm ⁻¹	0,9997	0,055	0,108	0,111

Fonte: elaborado pela autora.

A seguir, com base na implementação iPLS-PSO com 20 intervalos e 50 partículas para quantificação do óleo diesel, é mostrado um comparativo das evoluções para as 3 execuções, conforme a Figura 45. Observa-se nesta proposta que existe uma tendência de convergência de todos os modelos, sendo necessário em média de 107 iterações; comprovando novamente que o número de iterações definido é suficiente para explorar todas as melhores combinações possíveis entre os 20 intervalos.

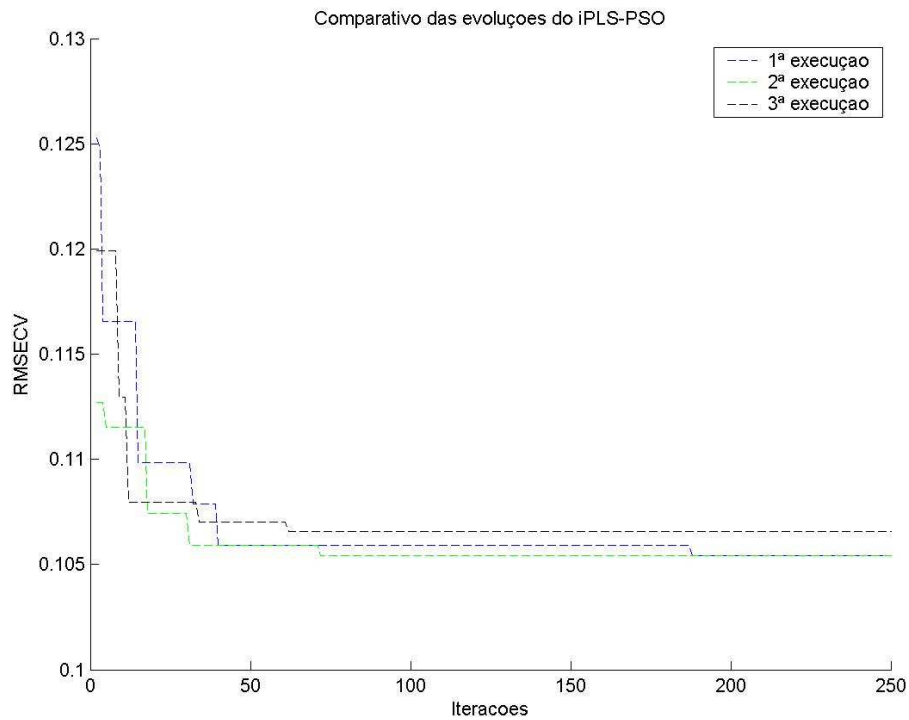


Figura 45 – Evolução das três execuções do iPLS-PSO para 20 intervalos e 50 partículas para a quantificação do óleo diesel.
Fonte: elaborado pela autora.

O comportamento das evoluções para as 3 execuções do método iPLS_PSO com 40 e 60 intervalos, ambos com 50 partículas, é diferente, pois a maioria dos modelos necessita de um número de iterações superior do que o modelo com 20 partículas, para atingirem os mínimos locais. Na evolução iPLS-PSO com 40 intervalos, percebe-se que em média são necessárias 155 iterações, já para 60 intervalos, em média, são necessárias 123 iterações como pode ser visualizada nas Figuras 46 e 47.

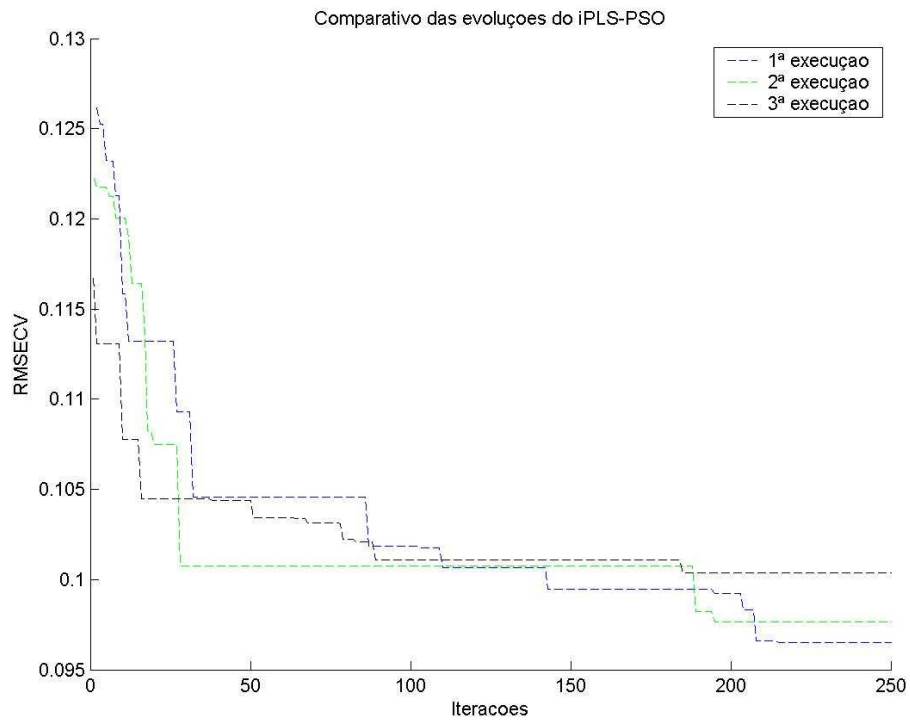


Figura 46 – Evolução das três execuções do iPLS-PSO para 40 intervalos e 50 partículas para quantificação do óleo diesel.
Fonte: elaborado pela autora.

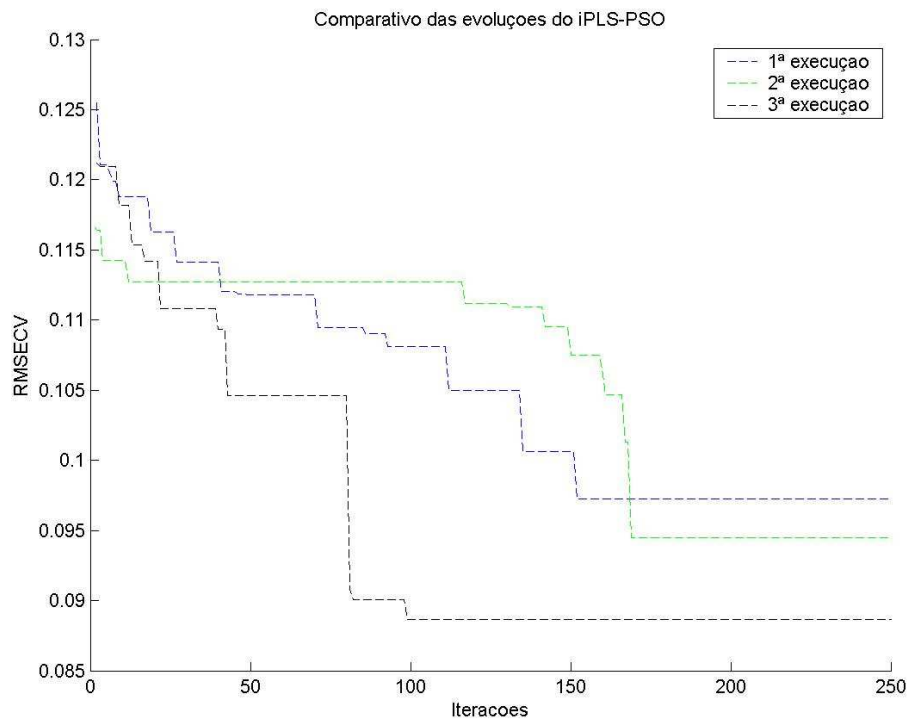


Figura 47 – Evolução das três execuções do iPLS-PSO para 60 intervalos e 50 partículas para quantificação do óleo diesel.
Fonte: elaborado pela autora.

5.3.3 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO

Nesta implementação denominada iPLS-PSO, utilizada para quantificação do óleo diesel no conjunto de dados proposto, selecionou regiões espectrais para construção do modelo que compreendem de 735 a 903 cm^{-1} , de 1159 a 1327 cm^{-1} e de 3172 a 3253 cm^{-1} como podem ser visualizadas na Figura 48. Analisando todos os modelos construídos, o modelo selecionado foi o construído com 40 intervalos e 25 partículas e, a sua respectiva curva de calibração empregando 10 variáveis latentes é apresentada na Figura 49.

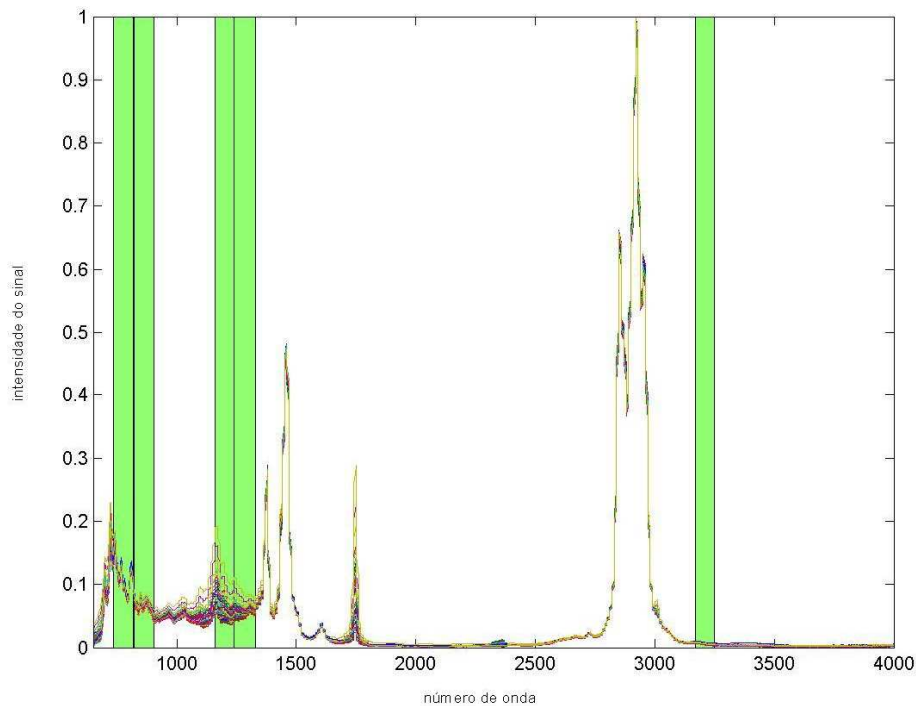


Figura 48 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

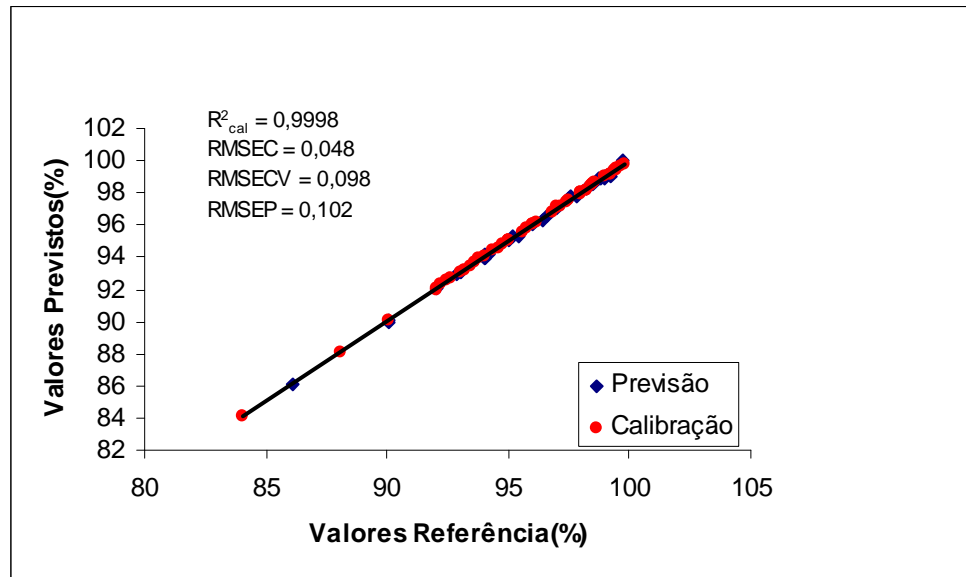


Figura 49 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO com 40 intervalos e 25 partículas.
 Fonte: elaborado pela autora.

5.3.4 Resultados obtidos para a quantificação do óleo diesel aplicando o método iPLS-PSO (CC)

Quando aplicado o método proposto iPLS-PSO (CC), para quantificação do óleo diesel, as regiões espectrais selecionadas para construção do modelo compreendem de 762 a 1207 cm^{-1} com lacunas, de 2719 a 2773 cm^{-1} , de 3223 a 3332 cm^{-1} e de 3446 a 3500 cm^{-1} , visualizadas na Figura 50. O modelo selecionado, dentre todos os construídos para este método, foi o proposto com 60 intervalos e 25 partículas e, a sua respectiva curva de calibração empregando 10 variáveis latentes é apresentada na Figura 51.

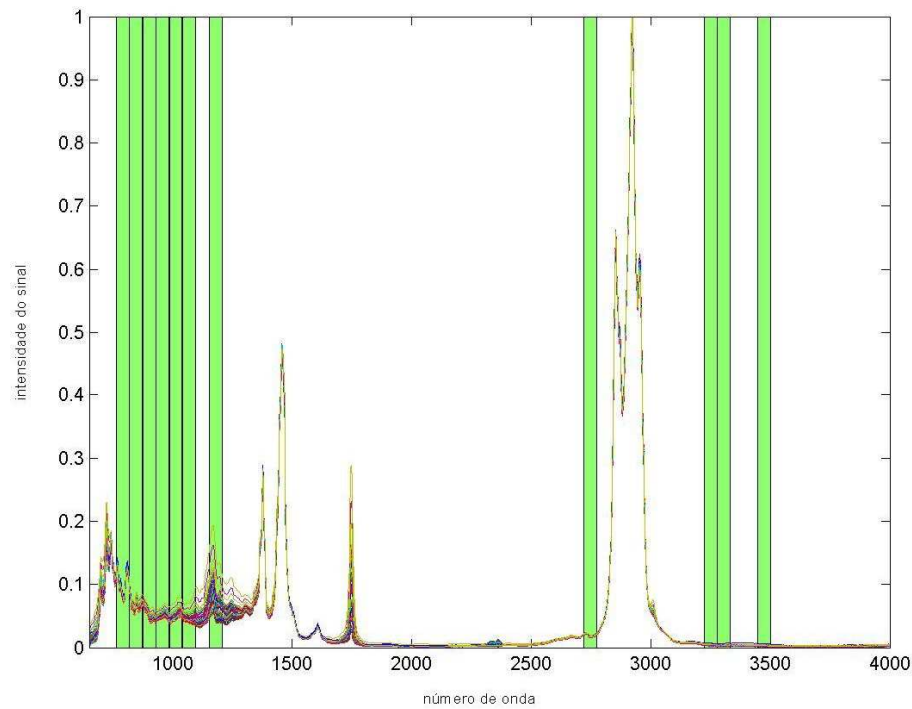


Figura 50 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.
Fonte: elaborado pela autora.

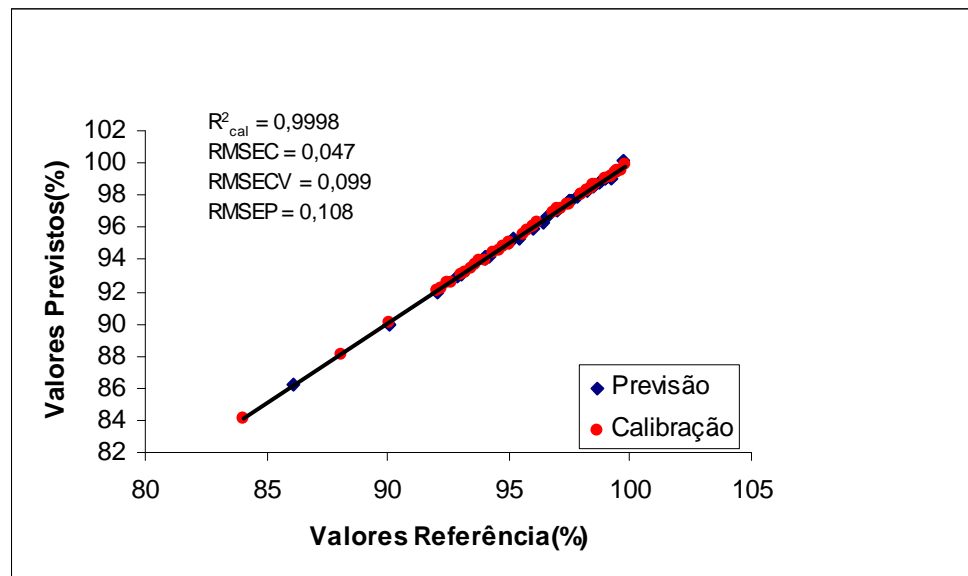


Figura 51 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CC) com 60 intervalos e 25 partículas.
Fonte: elaborado pela autora.

5.3.5 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DD)

Nos testes para quantificação do óleo diesel, aplicando o método iPLS-PSO (DD), foram selecionadas as regiões espectrais que compreendem de 904 a 1327 cm^{-1} com lacunas, de 2092 a 2175 cm^{-1} , de 2426 a 2756 cm^{-1} , de 3089 a 3170 cm^{-1} , de 3504 a 3585 cm^{-1} e de 3753 a 3834 cm^{-1} para a construção do modelo de regressão, visualizadas na Figura 52. O modelo selecionado, dentre todos os construídos para este método, foi o proposto com 40 intervalos e 100 partículas e, a sua respectiva curva de calibração utilizando 8 variáveis latentes é apresentada na Figura 53.

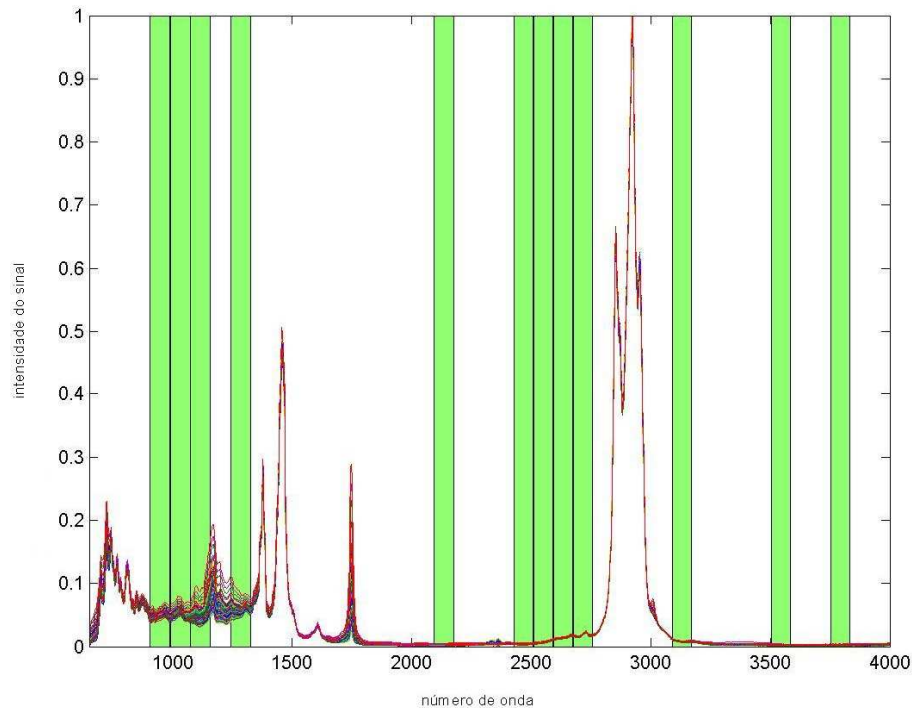


Figura 52 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas.
Fonte: elaborado pela autora.

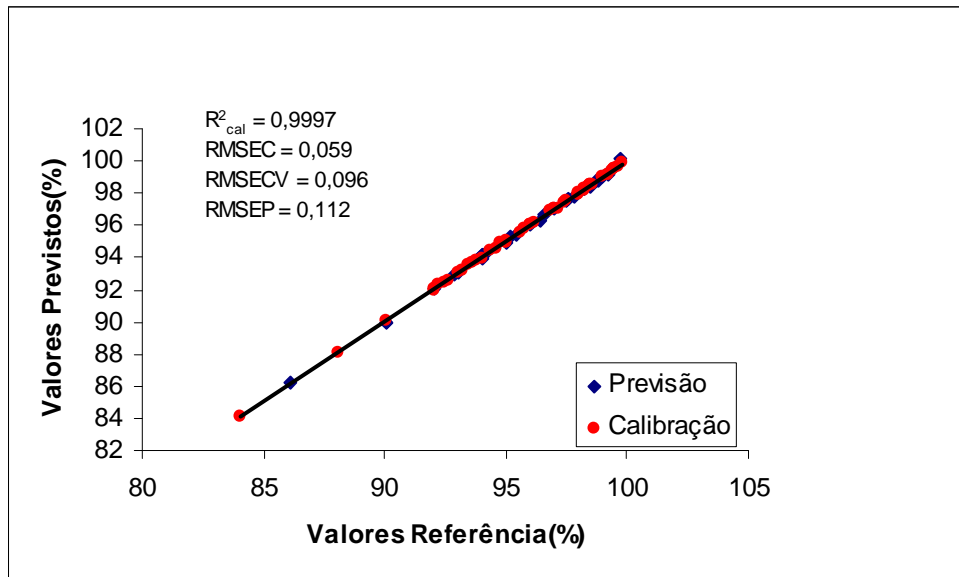


Figura 53 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DD) com 40 intervalos e 100 partículas. Fonte: elaborado pela autora.

5.3.6 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (CD)

Já a aplicação do método iPLS-PSO (CD) para a quantificação do óleo diesel nas misturas, selecionou as regiões espectrais nas faixas de 904 a 1327 cm^{-1} com lacunas, de 1753 a 1836 cm^{-1} , de 2260 a 2507 cm^{-1} , de 2675 a 2756 cm^{-1} e, de 3753 a 3834 cm^{-1} visualizadas na Figura 54. O modelo selecionado, dentre todos os construídos para este método, foi o proposto com 40 intervalos e 50 partículas e, a sua respectiva curva de calibração utilizando 10 variáveis latentes é apresentada na Figura 55.

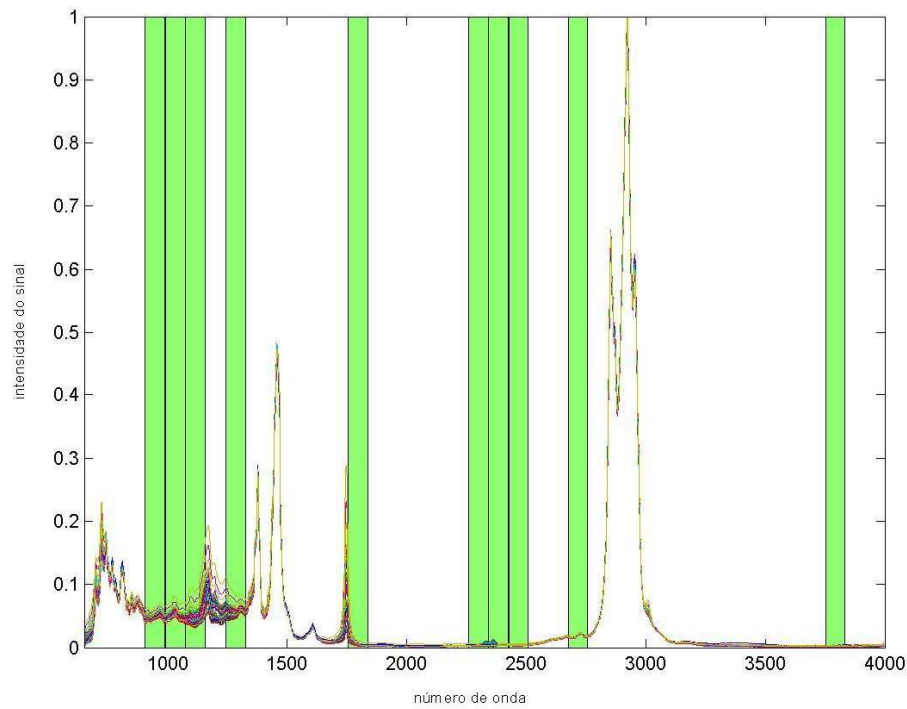


Figura 54 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

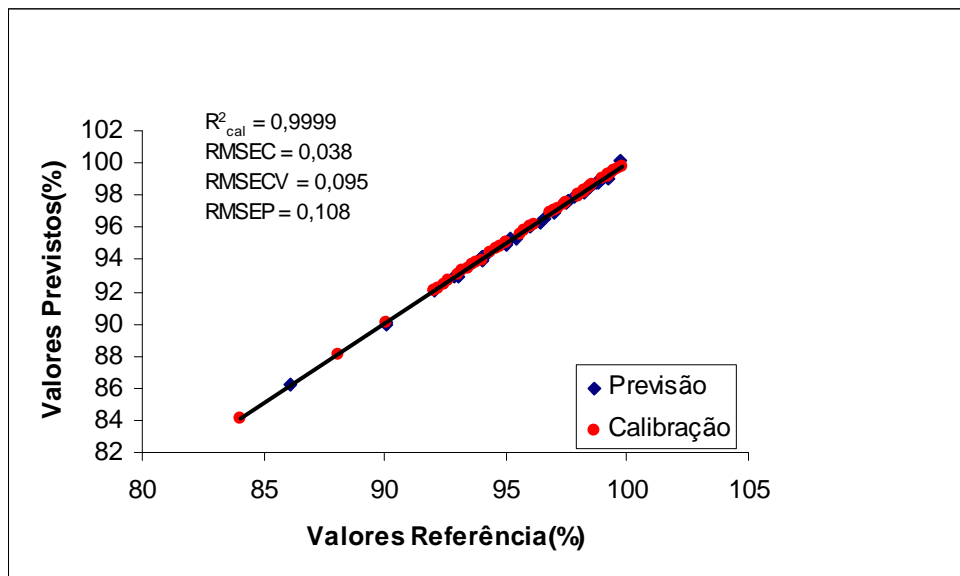


Figura 55 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (CD) com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

5.3.7 Resultados obtidos para a determinação do óleo diesel aplicando o método iPLS-PSO (DC)

E finalmente o método iPLS-PSO (DC) foi aplicado para a quantificação do óleo diesel e o modelo selecionado, dentre todos os construídos, foi o que selecionou a região espectral de 735 a 1327 cm^{-1} , visualizada na Figura 56, com 40 intervalos e 25 partículas. A sua respectiva curva de calibração para quantificação do óleo diesel empregando 10 variáveis latentes é apresentada na Figura 57.

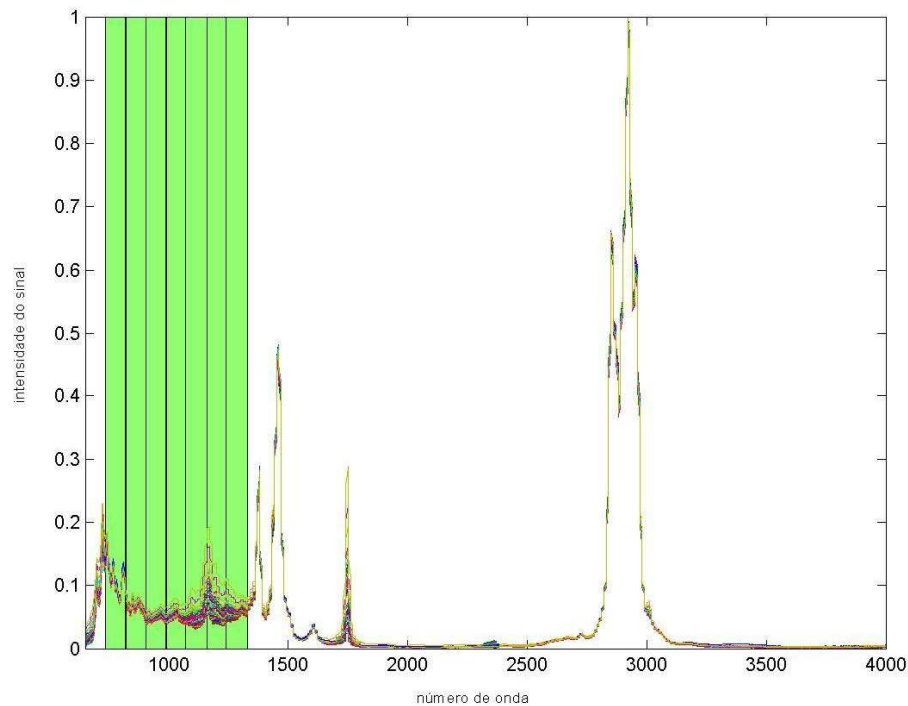


Figura 56 – Regiões espectrais utilizadas na modelagem para a quantificação do óleo diesel, usando o método iPLS-PSO (DC) com 40 intervalos e 25 partículas.
Fonte: elaborado pela autora.

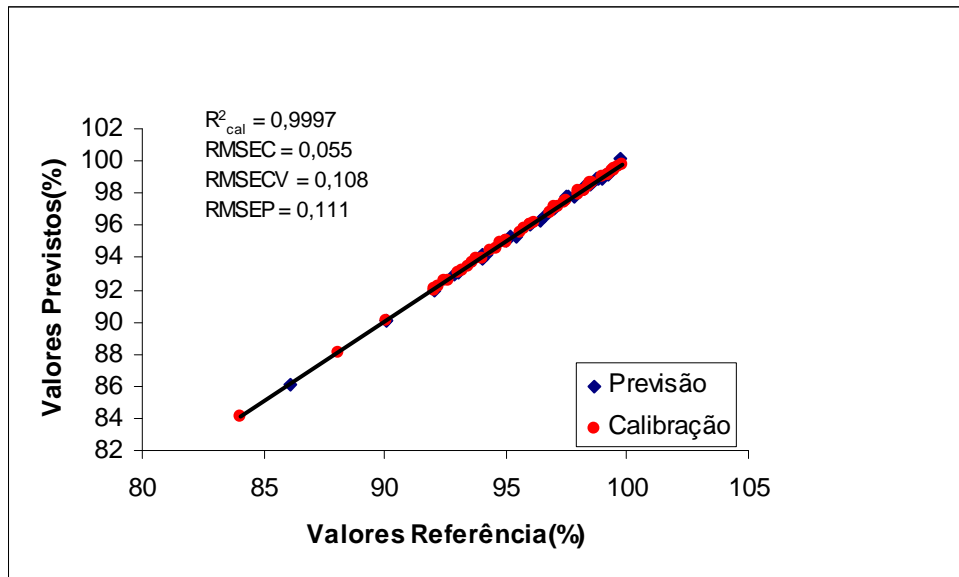


Figura 57 – Valores de referência *versus* previstos para o modelo de previsão do óleo diesel, empregando o método iPLS-PSO (DC) com 40 intervalos e 50 partículas. Fonte: elaborado pela autora.

5.3.8 Considerações sobre os modelos de óleo diesel

Com base nos resultados mostrados acima, foi possível observar que com baixos erros de previsão (RMSEP) todos os modelos construídos foram capazes de quantificar o óleo diesel. Analisando os resultados obtidos, empregando o PLS sobre o espectro inteiro da mistura e os resultados encontrados empregando as soluções propostas pelo método enxame de partículas, é possível perceber uma melhora na capacidade preditiva dos modelos otimização pela meta-heurística, porém não tão representativa como no caso das determinações de biodiesel e óleo vegetal.

Os modelos selecionados pelas otimizações aqui propostas para a quantificação do óleo diesel, apresentaram valores de RMSEP entre 0,102 a 0,112, enquanto os valores RMSECV ficaram na amplitude de 0,095 a 0,108.

No que diz respeito a comparação dos modelos construídos com as regiões selecionadas pela algoritmo híbrido e com toda região espectral da mistura não houve grandes melhoras, ficando evidenciado que a quantidade de informação capturada pelas variáveis latentes no modelo que emprega toda a região espectral é adequado para a quantificação deste componente da mistura.

Já para os modelos utilizando o algoritmo PSO nenhuma região característica dos hidrocarbonetos (região de 2850-3000 cm^{-1} , que corresponde a absorção de $\nu\text{C-H}$ referente aos estiramentos simétricos e assimétricos das ligações C-H de alcanos; 1465 cm^{-1} que corresponde a absorção de δCH_2 ; e 1375-1450 cm^{-1} , que corresponde a absorção de δCH_3) que constituem o diesel foi selecionada. Observando as regiões selecionadas na construção dos modelos a maior parte dessas correspondem à região do infravermelho médio denominada de impressão digital.

5.4 Considerações finais sobre os modelos de biodiesel, óleo vegetal e óleo diesel

Os resultados obtidos para os modelos visando a determinação, principalmente do biodiesel e da possível presença de óleo vegetal em misturas biodiesel/óleo vegetal/diesel, indicaram que não é possível utilizar uma única região e/ou a combinações das regiões sugeridas pela proposta de norma que garanta a determinação destes dentro do erro especificado pela proposta de norma. Conforme, destacado anteriormente a banda da carbonila (C=O) é insuficiente para a construção de um modelo eficiente e sensível as possíveis adulterações.

Comparando os espectros do biodiesel de soja e do óleo de soja que deu origem a este biodiesel (Figura 58) fica claro que a banda da carbonila é muito semelhante (sinal em 1742 cm^{-1} para o biodiesel e em 1743 cm^{-1} para o óleo) não sendo suficiente para gerar informação de discriminação, mesmo quando aplicada regressão multivariada.

Por outro lado, ratificando os resultados encontrados pela otimização PSO, a região referente ao C-O apresenta modificações bem mais representativas. Se por um lado o principal sinal no espectro do biodiesel encontra-se em 1170 cm^{-1} , por outro lado este mesmo sinal desloca-se para 1160 cm^{-1} para o óleo vegetal. O mesmo é verificado para a faixa de compreende o intervalo entre 1100-1450 cm^{-1} , região para a qual diversos modelos otimizados apresentam soluções tanto para o biodiesel, quanto para diesel com baixos erros dentro da faixa preconizada pela proposta de norma.

Uma outra forma de justificar o maior erro encontrado na previsão do biodiesel através das faixas espectrais sugeridas pela proposta de norma, é pela observação das figuras 59 e 60, onde as barras representam o erro de previsão do biodiesel nas amostras do conjunto de previsão, enquanto que as retas representam a concentração de óleo presentes nesta amostras.

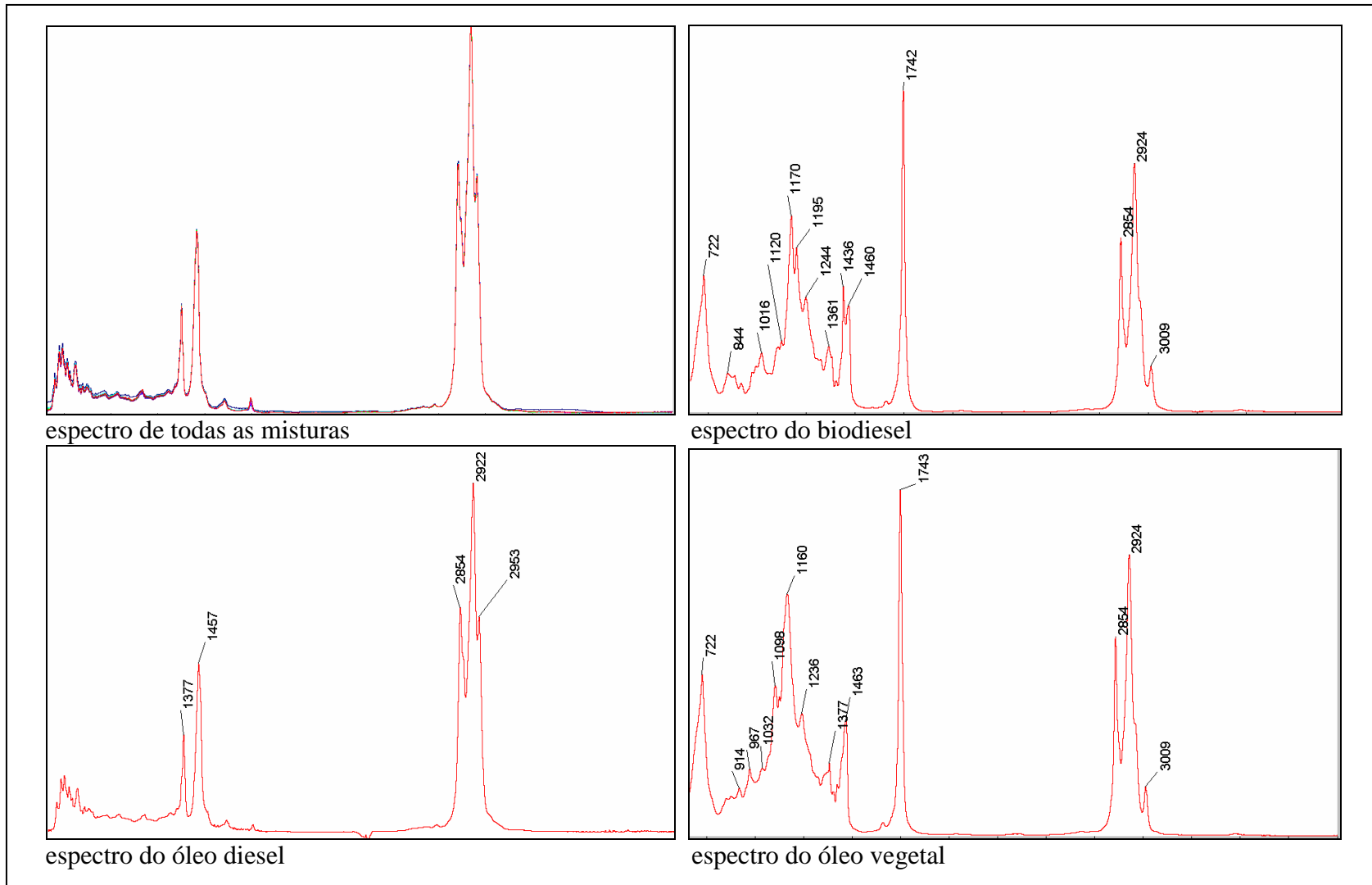


Figura 58 – Esquema contendo espectro de todas as misturas, do óleo diesel, do biodiesel e do óleo vegetal.
 Fonte: elaborado pela autora.

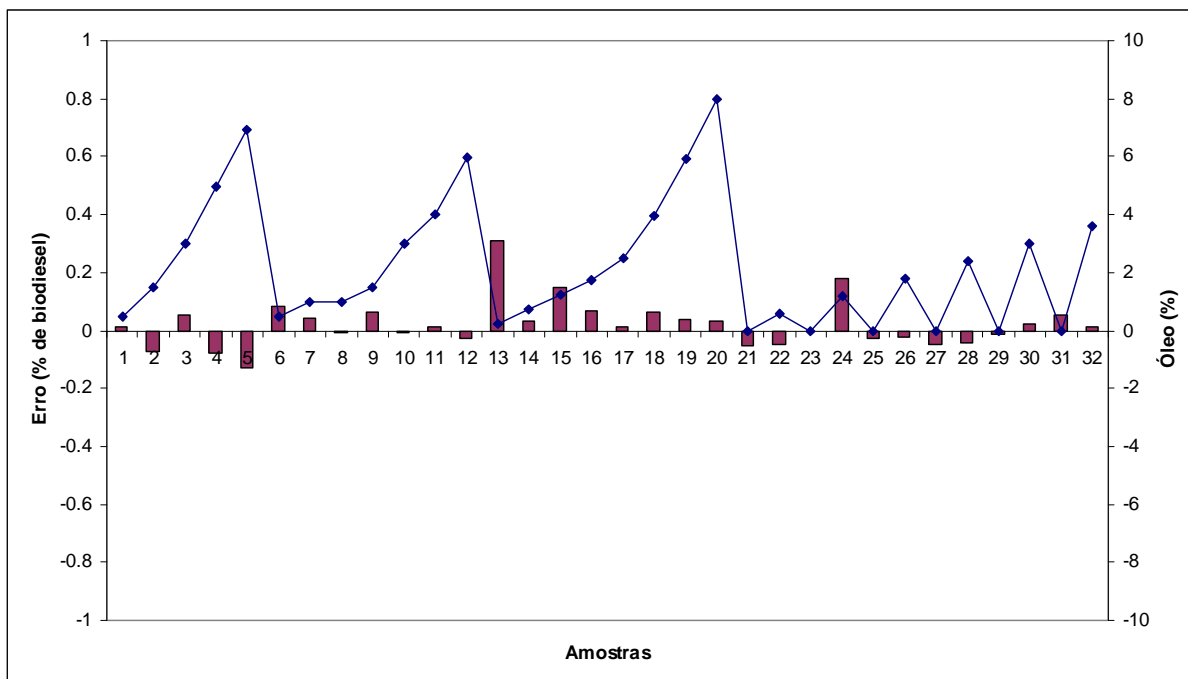


Figura 59 – Erros de previsão para a quantificação do biodiesel, usando as regiões propostas pelo método iPLS-PSO com 40 intervalos e 50 partículas.
Fonte: elaborado pela autora.

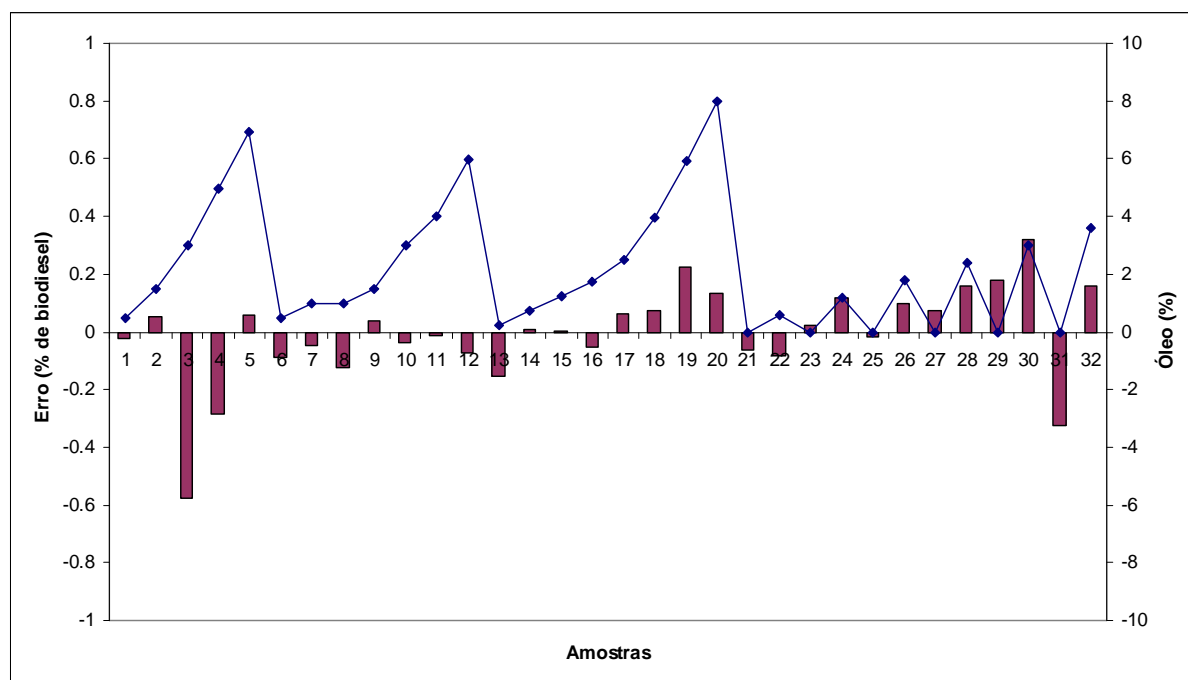


Figura 60 – Erros de previsão para a quantificação do biodiesel, utilizando somente as regiões espectrais sugeridas pela proposta de norma ABNT.
Fonte: elaborado pela autora.

O modelo empregando somente as regiões sugeridas pela proposta de norma ABNT, de uma forma geral, apresenta maior erro para as amostras que apresentam maior concentração de óleo, sugerindo que contribuição do sinal referente ao óleo esteja afetando a capacidade preditiva deste modelo. Já para o modelo otimizado isto não se verifica.

Em trabalhos empregando regressão multivariada, a exemplo do trabalho de Soares, Rezende e Fortes (2007), foi utilizada uma metodologia para quantificação de misturas biodiesel/diesel empregando o método de seleção de variáveis stepwise e forward apresentando erros de previsão na faixa de 0,65 a 1,40 %, entretanto neste caso não foi possível aplicar sempre o mesmo método de seleção de variáveis para todas as misturas de biodiesel e óleo de soja cru, onde cada grupo de amostra foi estudado separadamente para desenvolver um modelo de calibração adequado. No método proposto nesta dissertação, o RMSEP do melhor modelo de biodiesel, foi de 0,83 % sendo capaz de modelar diferentes tipos de biodiesel (soja e palma) e ainda na presença do óleo de soja, constituindo-se num modelo mais robusto.

Ainda existem na literatura trabalhos empregando análise univariada, Guarieiro *et al.* (2005, 2008), onde não se considera a existência de óleo de soja estando sujeitos, quando da análise de amostras adulteradas, produzir resultados falsos positivos, não sendo estes modelos adequados na análise a campo.

Embora os resultados apresentados nas Tabelas 5, 6 e 7 não indiquem sobre ajuste nos modelos selecionados, em alguns casos o processo de otimização não convergiu de forma satisfatória gerando modelos com RMSECV menor que o RMSEP. A Figura 61 apresentou um exemplo do comportamento desejável para a qual o *G_{best}* evolui de valores maiores RMSECV e RMSEP (pontos a direita) para valores menores de RMSECV e RMSEP (pontos a esquerda). Já na Figura 62 a evolução resulta na diminuição do RMSECV sem que necessariamente o RMSEP tenha o mesmo comportamento.

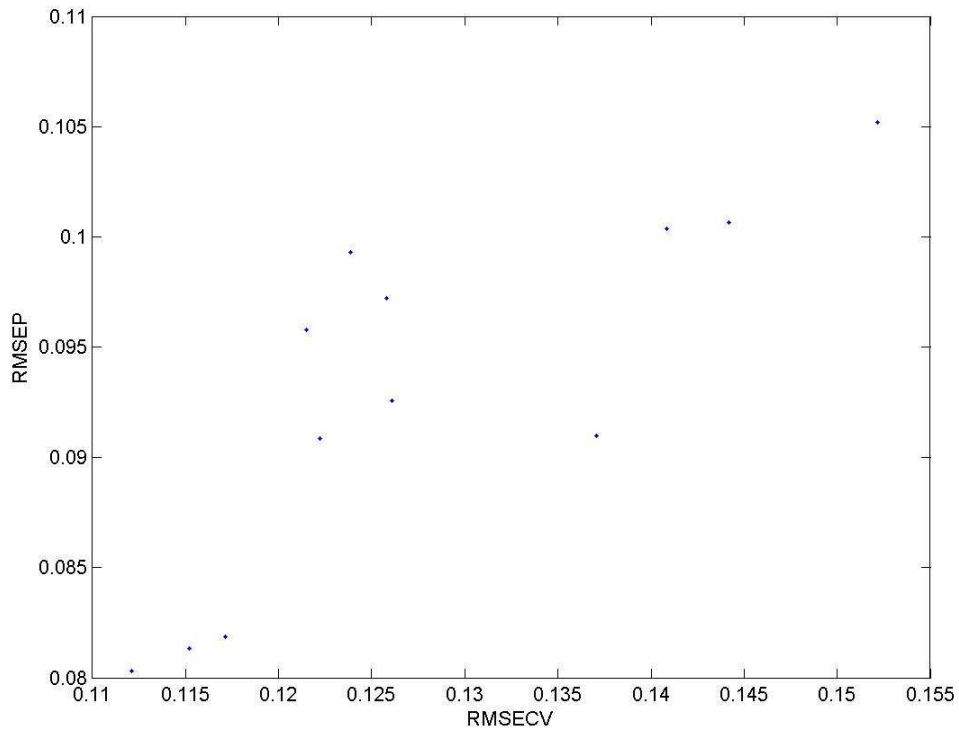


Figura 61 – Valores de RMSECV *versus* RMSEP para diferentes *Gbest* para quantificação do biodiesel, usando o método iPLS-PSO (CD) com 60 intervalos e 25 partículas.

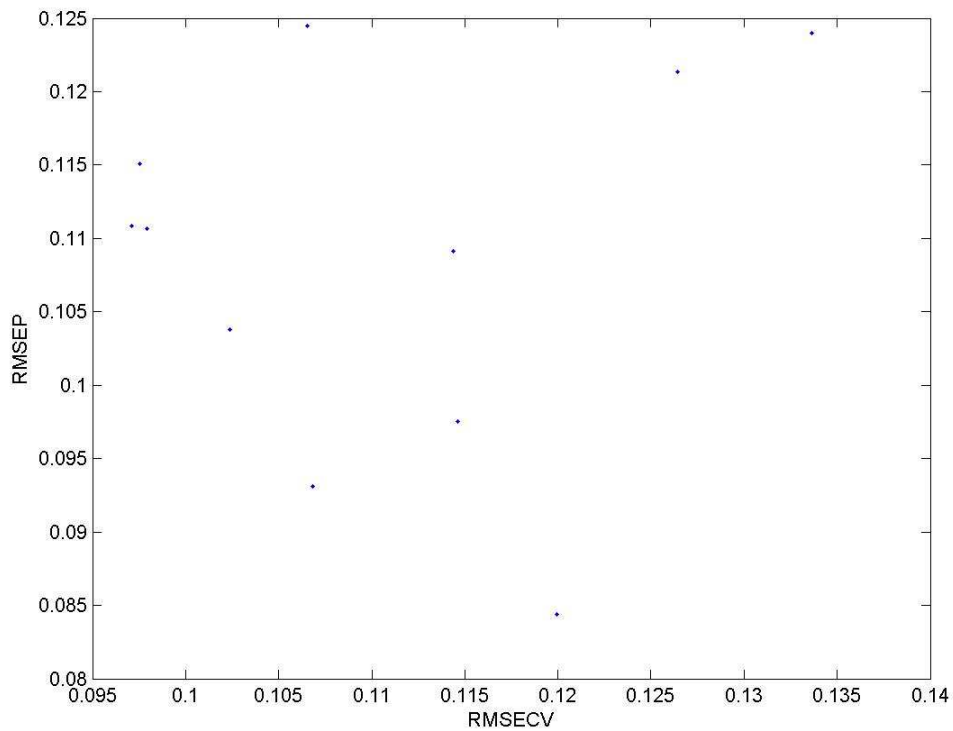


Figura 62 – Valores de RMSECV *versus* RMSEP para diferentes *Gbest* para a quantificação do biodiesel, usando o método iPLS-PSO (CD) com 40 intervalos e 100 partículas.

6 CONCLUSÕES

A presente dissertação mostrou que o método enxame de partículas em conjunto com o método de regressão por mínimos quadrados parciais por intervalos é capaz de desenvolver modelos de previsão de biodiesel, óleo vegetal e óleo diesel através dos espectros de reflexão total atenuada das misturas biodiesel/óleo diesel.

Independente do número de partículas empregados na otimização por enxame de partículas, bem como do número de intervalos para os quais os espectros foram subdivididos, bons modelos com capacidade preditiva maiores ou equivalentes ao emprego do espectro total foram obtidos para qualquer um dos componentes das misturas. Entretanto no que se refere ao modelo para a previsão de biodiesel empregando as regiões sugeridas pela proposta de norma este apresentou baixo desempenho que pode ser atribuído à interferência dos sinais do óleo que ocorrem nas mesmas faixas de absorção do biodiesel, não sendo suficiente a utilização destas faixas espectrais para a construção de modelos multivariados. O comportamento encontrado neste caso é bastante semelhante a interferência quando são construídos modelos univariados a partir de dados espectroscópicos.

As principais regiões selecionadas pelos algoritmos PSO podem ser facilmente atribuídas as principais absorções referentes a composição do biodiesel e do óleo vegetal, destacando-se a região referente ao estiramento C-O para quantificação do biodiesel e do óleo vegetal nas misturas, bem como a região referente a impressão digital para a quantificação do óleo diesel. Nestes modelos a seleção da região de absorção da carbonila (C=O) nem sempre foi necessária na obtenção de modelos com baixos erros de previsão.

Os modelos otimizados a partir dos algoritmos PSO também permitiram a identificação e quantificação da adulteração proveniente de óleos vegetais, enquanto que a proposta de norma é restrita a quantificação do biodiesel no diesel.

Finalmente, a implementação de métodos híbridos, nos quais são empregados conceitos de algoritmos clássicos em combinação com algoritmos heurísticos, permitem a resolução de problemas combinatoriais de forma adequada, somando-se as vantagens de ambas as técnicas para alcançar soluções satisfatórias em tempo real.

7 TRABALHOS FUTUROS

Abaixo se relaciona algum dos futuros trabalhos que podem ser desenvolvidos, com base nesta dissertação:

- ampliar a composição das misturas, utilizando outros tipos de biodiesel como o biodiesel de sebo e de óleo de frituras;
- ampliar a faixa de quantificação do biodiesel até 30 %(v/v), conforme o segundo modelo sugerido pela proposta de norma ABNT;
- comparar os resultados obtidos nesta dissertação com a PSO, com outros algoritmos híbridos ou não, como por exemplo algoritmo genético e busca tabu e;
- empregar algoritmos de compactação de dados, como as wavelets, tentando diminuir a quantidade de variáveis e o tempo de processamento, para construção dos modelos.

REFERÊNCIAS

- AGRAFIOTIS, D. K.; CEDEÑO, W. *Feature selection for structure-activity correlation using binary particle swarms*. Journal of Medicinal Chemistry, v. 45, n. 5, p. 1098-1107, fev. 2002.
- BARBOSA, Luiz Cláudio de Almeida. Espectroscopia no infravermelho na caracterização de compostos orgânicos. 1. ed. Viçosa: Editora Universidade Federal de Viçosa, 2007.
- BORGES NETO, Waldomiro. Parâmetros de qualidade de lubrificantes e óleo de oliva através de espectroscopia vibracional, calibração multivariada e seleção de variáveis. 2005. 130 f. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, 2005.
- BORIN, A.; POPPI, R. J. *Application of mid infrared spectroscopy and iPLS for the quantification of contaminants in lubricating oil*. Vibrational Spectroscopy, v. 37, n. 1, p. 27-32, jan. 2005.
- BRASIL. Lei n. 11.097, de 13 de janeiro de 2005. Disponível em: <http://www.planalto.gov.br/ccivil/_Ato2004-2006/2005/Lei/L11097.htm>. Acesso em: 02 fev. 2008.
- CEPEA. [s.n]. Quanto custa produzir biodiesel. 2006. Disponível em: <<http://cepea.esalq.usp.br/comunicacao/>>. Acesso em 25 ago. 2006.
- CLERC, M. *The swarm and the queen: towards a deterministic and adaptive particle swarm optimization*. In: Proc. Congress on Evolutionary Computation 1999, p. 1951-1957, 1999.
- CLERC, M.; KENNEDY, J. *The particle swarm: explosion, stability, and convergence in a multi-dimensional complex space*. IEEE Transactions on Evolutionary Computation, v. 6, n. 1 p. 58-73, fev. 2002.
- CHATTERJEE, A.; SIARRY, P. *Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization*. Computers and Operations Research, v. 33, n. 3, p. 859-871, mar. 2006.
- COSTA FILHO, P. A. da; POPPI, R. J. Aplicação de algoritmos genéticos na seleção de variáveis em espectroscopia no infravermelho médio. Determinação simultânea de glicose, maltose e frutose. Química Nova, São Paulo, v. 25, n. 1, p. 46-52, jan./fev. 2002.
- COSTA NETO, P. R. *et al.* Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em frituras. Química Nova, São Paulo, v. 23, n. 4, p. 531-537, jul./ago. 2000.
- DORADO, M. P. *et al.* *Kinetic parameters affecting the alkali-catalyzed transesterification process of used olive oil*. Energy e Fuels, v. 18, n. 5, p. 1457-1462, 2004.
- EBERHART, R.C.; HU, X. *Human tremor analysis using particle swarm optimization*. In: Proc. Congress on Evolutionary Computation 1999, p. 1927-1930, 1999.
- ENCINAR, J. M.; GONZÁLEZ, J. F.; RODRÍGUEZ-REINARES, A. *Biodiesel used frying oil. Variables affecting the yields and characteristics of the biodiesel*. Industrial & Engineering Chemistry Research, v. 44, n. 15, p. 5491-5499, 2005.
- ENGELBRECHT, Andries P. *Fundamentals of computational swarm intelligence*. England:

John Wiley & Sons, 2005.

FERRÃO, M. F. Técnicas de reflexão no infravermelho aplicadas na análise de alimentos. Tecno-Lógica., Santa Cruz do Sul, v. 5, n.1, p. 63-85, jan./jun. 2001.

FERRÃO, Marco Florês. Aplicação de técnicas espectroscópica de reflexão no infravermelho no controle de qualidade de farinha de trigo. 2000. 219 f. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, 2000.

FERREIRA, E. C. *et al.* Análise exploratória dos teores de constituintes inorgânicos em sucos e refrigerantes de uva. Eclética Química, v. 27, n. especial, p. 77-90, 2002.

FERREIRA, Marcos Antonio Gonzalez. Uso de análise multivariada para determinar a composição de gasolina a partir da espectroscopia FT-Raman e cromatografia gasosa. 2005. 109 f. Dissertação (Programa de Pós Graduação em Química Inorgânica) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2005.

GERIS, R. *et al.* Biodiesel de Soja – Reação de transesterificação para aulas práticas de química orgânica. Química Nova, São Paulo, v. 30, n. 5, p. 1369-1373, set./out. 2007.

GUARIEIRO, L. L. N. *et al.* Desenvolvimento de metodologia para quantificação das misturas biodiesel:diesel por infravermelho, 2005. Disponível em <www.portalabpg.org.br/PDPetro/3/trabalhos/IBP0250_05.pdf> . Acesso em: 10 dez. 2007.

GUARIEIRO, L. L. N. *et al.* Metodologia analítica para quantificar o teor de biodiesel na mistura biodiesel: diesel utilizando espectroscopia na região do infravermelho. Química Nova, São Paulo, v. 31, n. 2, p. 421-426, 2008.

GURDEN, S. P.; WESTERHUIS, J. A.; SMILDE, A. K. *Monitoring of batch processes using spectroscopy*. American Institute of Chemical Engineers. v. 48, n. 10, p. 2283-2297, out. 2002.

HAIR, J. F. *et al.* Análise Multivariada de Dados. 5. ed. Porto Alegre: Editora Bookman, 2005.

HAGEMAN *et al.* *Wavelength selection with tabu search*. Journal of Chemometrics. v. 17, n. 8-9, p. 427-437, 2003.

KHALIL, T. *et al.* *A Binary particle swarm optimization for optimal placement and sizing of capacitor banks in radial distribution feeders with distorted substation voltages*. In: Conference on Artificial Intelligence and Machine Learning, p. 129-135, 2006.

KENNEDY, J.; EBERHART, R. C. *A discrete binary version of the particle swarm algorithm*. In: Proc. of the 1997 Conference on Systems, Man, and Cybernetics, p. 4104-4109, 1997.

KENNEDY, J.; EBERHART, R. C. *Particle swarm optimization*. In: The 1995 IEEE International Conference on Neural Networks, p. 1942-1948, 1995.

KENNEDY, J.; SPEARS, W. M. *Matching algorithms to problems: an experimental test of the particle swarm and some genetic algorithms on the multimodal problem generator*. In: Proc. of the 1998 International Conference on Evolutionary Computation, p. 78-83, 1998.

KENNEDY, J.; EBERHART, R.C.; SHI, Y. *Swarm intelligence*. San Francisco: Morgan Kaufmann Publishers, 2001.

- KNOTHE Gerhard. *Analyzing biodiesel: standards and other methods*. Journal of the American Oil Chemists Society, v. 83, n. 10, p. 823-833, out. 2006.
- KONZEN, P. H. de A. *et al.* Otimização de métodos de controle de qualidade de fármacos usando algoritmo genético e busca tabu. Pesquisa Operacional, Rio de Janeiro, v. 23, n. 1, p. 189-207, jan./abr. 2003.
- LEARDI, R.; NORGAARD, L. *Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions*. Journal of Chemometrics, v. 18, n. 11, p. 486-497, 2004.
- MÜLLER, Viviane. Otimização de layouts industriais através do método enxame de partículas. 2007. 79 f. Dissertação (Programa de Pós-Graduação em Sistemas e Processos Industriais - Mestrado) - Universidade de Santa Cruz do Sul, Santa Cruz do Sul, 2007.
- MEDEIROS, José Antonio Carlos Canedo. Enxame de partículas como ferramenta de otimização em problemas complexos de engenharia nuclear. 2005. 108 f. Tese (Programa de Pós-Graduação de Engenharia – Doutorado) Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.
- MENDES, Rui. *Population topologies and their influence in particle swarm performance*. 2004. 189 f. Dissertation al fulfillment of the requirements for the degree of Philosophiæ Doctor at the University of Minho, Portugal, 2004.
- MENESES Anderson Alvarenga de Moura. Otimização por enxame de partículas aplicado ao problema combinatório da recarga de um reator nuclear. 2005. 100f. Dissertação (Programa de Pós-Graduação de Engenharia – Mestrado) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.
- MORGANO, M. A. *et al.* Aplicação de análise exploratória na diferenciação de vegetais. *Brazilian Journal of Food Technology*, São Paulo, v.2, n. 1/2, p. 73-79, 1999.
- NASCIMENTO, M. da G.; COSTA NETO, P. R.; MAZZUCO, L. M. Biotransformação de óleos e gorduras. *Biotechnology Ciência & Desenvolvimento*, Brasília, n.19, p. 28-31, mar./abr. 2001.
- NORGAARD, L. *et al.* *Interval partial least-square regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy*. *Applied Spectroscopy*, v. 54, n. 3, p. 413-419, 2000.
- OLIVEIRA, Flávia Carneiro da Cunha. Modelos de calibração multivariada associados à espectroscopia vibracional para análise de misturas diesel-óleos vegetais. 2006. 108 f. Dissertação (Programa de Pós-Graduação do Instituto de Química - Mestrado) - Universidade de Brasília, 2006.
- OMRAN M.G.H.; ENGELBRECHT A. P.; SALMAN, A. 2005. *Dynamic clustering using particle swarm optimization with application in unsupervised image classification*. *Transactions on Engineering, Computing And Technology*, v. 9, p. 199-204, nov. 2005.
- PARADKAR, M. M.; IRUDAYARAJ, J. *Rapid determination of caffeine content in soft drinks using FTIR-ATR spectroscopy*. *Food Chemistry*, v. 78, n. 2. p. 261-266, ago. 2002.
- PARISOTTO, G. *et al.* Análise exploratória aplicada no estudo de medicamentos contendo piroxicam. *Brazilian Journal of Pharmaceutical Sciences*, São Paulo, v. 41, n. 4. p. 499-505, out./dez. 2005.

PARSOPOULO, K. E.; VRAHATIS, M. N. *Particle swarm optimization method in multiobjective problems*. In: Proc. of the 2002 Symposium on Applied Computing, p. 603-607, 2002.

PIMENTEL, M. F. *et al.* *Determination of biodiesel content when blended with mineral diesel fuel using infrared spectroscopy and multivariate calibration*. Microchemical Journal, v. 82, n. 2, p. 201-206, abr. 2006.

PERAM, T., VEERAMACHANENI, K.; MOHAN, C. K. Fitness-distance-ratio based particle swarm optimization. In: Proc. of the 2003 Swarm Intelligence Symposium, p. 174-181, 2003.

POMEROY, Paul. *An introduction to particle swarm optimization*. 2003 Disponível em: <<http://www.adaptiveview.com/articles/ipsoprnt.html>>. Acesso em: 03 jan. 2006.

PRADO, J. R. do; SARAMAGO, S. F. .P. Otimização por colônia de partículas. FAMAT em Revista, Uberlândia, n. 4, p. 87-103, abr. 2005.

QI SHEN *et al.*; *Particle swarm algorithm trained neural network for qsar studies of inhibitors of platelet-derived growth factor receptor phosphorylation*. European journal of pharmaceutical sciences, v. 28, n. 5, p. 369-376, 2006.

RADICH, Antony. *Biodiesel performance, costs, and use*. 2004. Disponível em: <<http://www.eia.doe.gov/oiaf/analysispaper/biodiesel/>>. Acesso em: 10 de mar. 2008.

SABIN, Gustavo Post. Otimização de modelos de regressão multivariados empregando métodos de seleção de variáveis. 2007. 111 f. Dissertação (Programa de Pós-Graduação em Sistemas e Processos Industriais - Mestrado) - Universidade de Santa Cruz do Sul, Santa Cruz do Sul, 2007.

SACORAGUE, Luiz Alexandre. Avaliação de diferentes regiões do espectro do infravermelho próximo na determinação de parâmetros de qualidade de combustíveis empregando ferramentas quimiométricas. 2004. 235 f. Tese (Departamento de Química Analítica - Doutorado) – Universidade Estadual de Campinas, Campinas, 2004.

SCHWAAD, Marcio. Avaliação de algoritmos heurísticos de otimização em problemas de estimação de parâmetros. 2005. 156 f. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

SENA, M. M. de *et al.* Avaliação do uso de métodos quimiométricos em análise de solos. Química Nova, São Paulo, v. 23, n. 4, p. 547-556, jul./ago. 2000.

SHI, Y.; EBERHART, R. C. *A modified particle swarm optimizer*. In: Proc. of the IEEE International Conference on Evolutionary Computation, p. 69-73, 1998a.

SHI, Y.; EBERHART, R. C. *Parameter selection in particle swarm optimization*. In: Evolutionary Programming VII. Lecture Notes in Computer Science, p. 591-600, 1998b.

SHI, Y., EBERHART, R. *Empirical Study of Particle Swarm Optimization*. In: Proc. Congress on Evolutionary Computation, p. 1945-1950, 1999.

SHI, Y; EBERHART, R. C. *Fuzzy adaptive particle swarm optimization*. In: Proceedings of the IEEE Congress on Evolutionary Computation, p. 101-106, 2001.

SOARES, I. P.; REZENDE, T. F.; FORTES, I. C. P. Calibração multivariada com seleção de variáveis em amostras de biodiesel adulteradas com óleo de soja cru utilizando dados de

espectroscopia de infravermelho com transformada de fourier. 2007 Disponível em www.biodiesel.gov.br/docs/congresso2007/caracterizacao/10.pdf. Acesso em: 03 de mar. 2008.

WANG, Z. *et al.* *Particle swarm optimization and neural network application for QSAR*. In: Parallel and Distributed Processing Symposium, abr. 2004.

YOSHIDA, S.; YOSHIDA H. *Non destructive analyses of unsaturated fatty acid species in dietary oils by attenuated total reflectance with Fourier transform IR spectroscopy*. Journal Biopolymers, v. 70, n. 4, p. 604-613, nov. 2003.

ZHENG, Y. *et al.* *Empirical study of particle swarm optimizer with an increasing inertia weight*. In: Proceeding of the IEEE Congress on Evolutionary Computation, v. 1, p. 221-226, 2003.

ANEXOS

ANEXO A - Resultados completos das implementações do iPLS-PSO para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	58
	2 ^a	0,120	0,091	63
	3 ^a	0,120	0,091	115
20 intervalos 50 partículas	1 ^a	0,120	0,091	46
	2 ^a	0,120	0,091	14
	3 ^a	0,120	0,091	72
20 intervalos 100 partículas	1 ^a	0,120	0,091	17
	2 ^a	0,120	0,091	56
	3 ^a	0,120	0,091	79
40 intervalos 25 partículas	1 ^a	0,103	0,134	179
	2 ^a	0,101	0,115	241
	3 ^a	0,101	0,120	239
40 intervalos 50 partículas	1 ^a	0,100	0,137	165
	2 ^a	0,115	0,083	59
	3 ^a	0,102	0,120	214
40 intervalos 100 partículas	1 ^a	0,100	0,137	142
	2 ^a	0,100	0,137	205
	3 ^a	0,101	0,107	207
60 intervalos 25 partículas	1 ^a	0,128	0,092	166
	2 ^a	0,114	0,089	238
	3 ^a	0,106	0,090	205
60 intervalos 50 partículas	1 ^a	0,106	0,099	156
	2 ^a	0,101	0,103	198
	3 ^a	0,100	0,091	231
60 intervalos 100 partículas	1 ^a	0,095	0,088	247
	2 ^a	0,107	0,121	168
	3 ^a	0,123	0,089	150

ANEXO A - Resultados completos das implementações do iPLS-PSO (CC) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	61
	2 ^a	0,120	0,091	141
	3 ^a	0,120	0,091	61
20 intervalos 50 partículas	1 ^a	0,120	0,091	94
	2 ^a	0,120	0,091	79
	3 ^a	0,120	0,091	76
20 intervalos 100 partículas	1 ^a	0,120	0,091	90
	2 ^a	0,120	0,091	70
	3 ^a	0,120	0,091	47
40 intervalos 25 partículas	1 ^a	0,103	0,903	152
	2 ^a	0,105	0,085	108
	3 ^a	0,106	0,106	117
40 intervalos 50 partículas	1 ^a	0,101	0,105	172
	2 ^a	0,103	0,093	164
	3 ^a	0,103	0,086	213
40 intervalos 100 partículas	1 ^a	0,106	0,092	159
	2 ^a	0,103	0,116	231
	3 ^a	0,113	0,075	106
60 intervalos 25 partículas	1 ^a	0,099	0,114	169
	2 ^a	0,108	0,097	158
	3 ^a	0,095	0,108	227
60 intervalos 50 partículas	1 ^a	0,100	0,114	147
	2 ^a	0,109	0,097	207
	3 ^a	0,116	0,115	164
60 intervalos 100 partículas	1 ^a	0,101	0,109	227
	2 ^a	0,092	0,106	213
	3 ^a	0,087	0,106	107

ANEXO A - Resultados completos das implementações do iPLS-PSO (DD) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	12
	2 ^a	0,120	0,091	12
	3 ^a	0,120	0,094	23
20 intervalos 50 partículas	1 ^a	0,120	0,091	9
	2 ^a	0,120	0,091	85
	3 ^a	0,120	0,091	44
20 intervalos 100 partículas	1 ^a	0,120	0,091	22
	2 ^a	0,120	0,091	56
	3 ^a	0,120	0,091	33
40 intervalos 25 partículas	1 ^a	0,103	0,101	247
	2 ^a	0,104	0,111	230
	3 ^a	0,101	0,101	231
40 intervalos 50 partículas	1 ^a	0,103	0,917	203
	2 ^a	0,103	0,109	202
	3 ^a	0,108	0,096	160
40 intervalos 100 partículas	1 ^a	0,103	0,086	171
	2 ^a	0,097	0,111	193
	3 ^a	0,105	0,085	138
60 intervalos 25 partículas	1 ^a	0,103	0,101	247
	2 ^a	0,104	0,111	230
	3 ^a	0,101	0,101	231
60 intervalos 50 partículas	1 ^a	0,103	0,917	203
	2 ^a	0,103	0,109	202
	3 ^a	0,108	0,096	160
60 intervalos 100 partículas	1 ^a	0,102	0,091	236
	2 ^a	0,105	0,095	203
	3 ^a	0,087	0,106	333

ANEXO A - Resultados completos das implementações do iPLS-PSO (CD) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	41
	2 ^a	0,120	0,091	76
	3 ^a	0,120	0,091	54
20 intervalos 50 partículas	1 ^a	0,120	0,091	33
	2 ^a	0,120	0,091	105
	3 ^a	0,120	0,091	70
20 intervalos 100 partículas	1 ^a	0,120	0,091	44
	2 ^a	0,120	0,091	33
	3 ^a	0,120	0,091	14
40 intervalos 25 partículas	1 ^a	0,105	0,085	176
	2 ^a	0,113	0,115	242
	3 ^a	0,103	0,091	201
40 intervalos 50 partículas	1 ^a	0,115	0,083	52
	2 ^a	0,115	0,083	85
	3 ^a	0,115	0,083	104
40 intervalos 100 partículas	1 ^a	0,113	0,063	63
	2 ^a	0,097	0,111	169
	3 ^a	0,115	0,083	32
60 intervalos 25 partículas	1 ^a	0,121	0,074	160
	2 ^a	0,124	0,105	163
	3 ^a	0,112	0,080	118
60 intervalos 50 partículas	1 ^a	0,110	0,101	224
	2 ^a	0,108	0,097	106
	3 ^a	0,114	0,094	127
60 intervalos 100 partículas	1 ^a	0,117	0,078	95
	2 ^a	0,095	0,097	193
	3 ^a	0,096	0,102	159

ANEXO A - Resultados completos das implementações do iPLS-PSO (DC) para a quantificação do biodiesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,120	0,091	115
	2 ^a	0,120	0,091	96
	3 ^a	0,120	0,091	90
20 intervalos 50 partículas	1 ^a	0,120	0,091	77
	2 ^a	0,120	0,091	25
	3 ^a	0,120	0,091	82
20 intervalos 100 partículas	1 ^a	0,120	0,091	9
	2 ^a	0,120	0,091	106
	3 ^a	0,120	0,091	51
40 intervalos 25 partículas	1 ^a	0,112	0,101	51
	2 ^a	0,110	0,092	104
	3 ^a	0,103	0,101	62
40 intervalos 50 partículas	1 ^a	0,116	0,098	155
	2 ^a	0,105	0,103	162
	3 ^a	0,102	0,101	82
40 intervalos 100 partículas	1 ^a	0,105	0,104	59
	2 ^a	0,102	0,108	60
	3 ^a	0,103	0,093	123
60 intervalos 25 partículas	1 ^a	0,126	0,093	163
	2 ^a	0,113	0,101	53
	3 ^a	0,125	0,081	102
60 intervalos 50 partículas	1 ^a	0,112	0,096	51
	2 ^a	0,110	0,147	30
	3 ^a	0,110	0,101	169
60 intervalos 100 partículas	1 ^a	0,107	0,101	185
	2 ^a	0,103	0,116	48
	3 ^a	0,104	0,099	89

ANEXO B - Resultados completos das implementações do iPLS-PSO para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,079	0,095	114
	2 ^a	0,082	0,103	30
	3 ^a	0,082	0,103	48
20 intervalos 50 partículas	1 ^a	0,079	0,095	65
	2 ^a	0,079	0,095	127
	3 ^a	0,079	0,095	133
20 intervalos 100 partículas	1 ^a	0,079	0,095	67
	2 ^a	0,079	0,095	49
	3 ^a	0,079	0,095	80
40 intervalos 25 partículas	1 ^a	0,084	0,098	80
	2 ^a	0,083	0,092	232
	3 ^a	0,081	0,096	238
40 intervalos 50 partículas	1 ^a	0,080	0,010	178
	2 ^a	0,075	0,105	217
	3 ^a	0,078	0,103	158
40 intervalos 100 partículas	1 ^a	0,076	0,102	219
	2 ^a	0,076	0,107	158
	3 ^a	0,080	0,107	231
60 intervalos 25 partículas	1 ^a	0,075	0,116	171
	2 ^a	0,075	0,106	166
	3 ^a	0,076	0,121	240
60 intervalos 50 partículas	1 ^a	0,073	0,101	234
	2 ^a	0,077	0,102	145
	3 ^a	0,074	0,112	249
60 intervalos 100 partículas	1 ^a	0,077	0,110	238
	2 ^a	0,079	0,110	85
	3 ^a	0,076	0,101	166

ANEXO B - Resultados completos das implementações do iPLS-PSO (CC) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,079	0,095	197
	2 ^a	0,082	0,103	77
	3 ^a	0,079	0,095	165
20 intervalos 50 partículas	1 ^a	0,079	0,095	56
	2 ^a	0,079	0,095	49
	3 ^a	0,079	0,095	98
20 intervalos 100 partículas	1 ^a	0,079	0,095	38
	2 ^a	0,079	0,095	49
	3 ^a	0,079	0,095	98
40 intervalos 25 partículas	1 ^a	0,083	0,101	72
	2 ^a	0,088	0,089	167
	3 ^a	0,079	0,105	238
40 intervalos 50 partículas	1 ^a	0,078	0,109	120
	2 ^a	0,076	0,101	174
	3 ^a	0,079	0,109	126
40 intervalos 100 partículas	1 ^a	0,081	0,105	87
	2 ^a	0,075	0,105	201
	3 ^a	0,080	0,094	201
60 intervalos 25 partículas	1 ^a	0,077	0,108	215
	2 ^a	0,069	0,118	201
	3 ^a	0,080	0,106	221
60 intervalos 50 partículas	1 ^a	0,083	0,106	193
	2 ^a	0,080	0,111	231
	3 ^a	0,080	0,102	130
60 intervalos 100 partículas	1 ^a	0,085	0,108	37
	2 ^a	0,077	0,100	144
	3 ^a	0,077	0,114	213

ANEXO B - Resultados completos das implementações do iPLS-PSO (DD) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,079	0,095	205
	2 ^a	0,081	0,096	156
	3 ^a	0,089	0,109	157
20 intervalos 50 partículas	1 ^a	0,079	0,095	91
	2 ^a	0,082	0,103	21
	3 ^a	0,079	0,095	174
20 intervalos 100 partículas	1 ^a	0,079	0,095	100
	2 ^a	0,079	0,095	119
	3 ^a	0,081	0,094	234
40 intervalos 25 partículas	1 ^a	0,070	0,130	186
	2 ^a	0,080	0,101	221
	3 ^a	0,082	0,988	241
40 intervalos 50 partículas	1 ^a	0,074	0,105	241
	2 ^a	0,085	0,099	244
	3 ^a	0,078	0,110	208
40 intervalos 100 partículas	1 ^a	0,075	0,102	156
	2 ^a	0,082	0,089	161
	3 ^a	0,076	0,107	97
60 intervalos 25 partículas	1 ^a	0,080	0,107	227
	2 ^a	0,079	0,101	185
	3 ^a	0,083	0,104	216
60 intervalos 50 partículas	1 ^a	0,079	0,103	246
	2 ^a	0,076	0,103	246
	3 ^a	0,080	0,130	238
60 intervalos 100 partículas	1 ^a	0,074	0,110	249
	2 ^a	0,075	0,100	250
	3 ^a	0,071	0,101	236

ANEXO B - Resultados completos das implementações do iPLS-PSO (CD) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,081	0,096	84
	2 ^a	0,081	0,094	81
	3 ^a	0,082	0,103	105
20 intervalos 50 partículas	1 ^a	0,079	0,095	59
	2 ^a	0,079	0,095	99
	3 ^a	0,082	0,103	69
20 intervalos 100 partículas	1 ^a	0,079	0,095	101
	2 ^a	0,079	0,095	93
	3 ^a	0,079	0,095	90
40 intervalos 25 partículas	1 ^a	0,080	0,099	147
	2 ^a	0,074	0,105	147
	3 ^a	0,082	0,118	250
40 intervalos 50 partículas	1 ^a	0,082	0,102	189
	2 ^a	0,078	0,109	181
	3 ^a	0,078	0,106	222
40 intervalos 100 partículas	1 ^a	0,075	0,110	191
	2 ^a	0,075	0,102	94
	3 ^a	0,070	0,130	95
60 intervalos 25 partículas	1 ^a	0,084	0,106	243
	2 ^a	0,075	0,096	243
	3 ^a	0,070	0,115	199
60 intervalos 50 partículas	1 ^a	0,075	0,103	240
	2 ^a	0,076	0,107	231
	3 ^a	0,085	0,112	56
60 intervalos 100 partículas	1 ^a	0,074	0,105	229
	2 ^a	0,077	0,105	228
	3 ^a	0,075	0,105	229

ANEXO B - Resultados completos das implementações do iPLS-PSO (DC) para a quantificação do óleo vegetal.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,082	0,103	209
	2 ^a	0,085	0,115	49
	3 ^a	0,082	0,103	62
20 intervalos 50 partículas	1 ^a	0,079	0,095	57
	2 ^a	0,079	0,095	24
	3 ^a	0,079	0,095	177
20 intervalos 100 partículas	1 ^a	0,079	0,095	44
	2 ^a	0,079	0,095	37
	3 ^a	0,079	0,095	44
40 intervalos 25 partículas	1 ^a	0,082	0,102	221
	2 ^a	0,087	0,109	47
	3 ^a	0,085	0,122	182
40 intervalos 50 partículas	1 ^a	0,082	0,107	32
	2 ^a	0,076	0,101	206
	3 ^a	0,079	0,100	59
40 intervalos 100 partículas	1 ^a	0,085	0,090	59
	2 ^a	0,085	0,105	209
	3 ^a	0,078	0,102	58
60 intervalos 25 partículas	1 ^a	0,082	0,101	226
	2 ^a	0,085	0,117	41
	3 ^a	0,091	0,116	32
60 intervalos 50 partículas	1 ^a	0,084	0,101	15
	2 ^a	0,082	0,102	225
	3 ^a	0,084	0,110	213
60 intervalos 100 partículas	1 ^a	0,077	0,110	177
	2 ^a	0,080	0,109	38
	3 ^a	0,085	0,113	119

ANEXO C - Resultados completos das implementações do iPLS-PSO para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,107	0,131	103
	2 ^a	0,111	0,139	136
	3 ^a	0,118	0,147	222
20 intervalos 50 partículas	1 ^a	0,105	0,125	188
	2 ^a	0,105	0,125	72
	3 ^a	0,107	0,131	62
20 intervalos 100 partículas	1 ^a	0,107	0,131	35
	2 ^a	0,105	0,125	217
	3 ^a	0,103	0,136	76
40 intervalos 25 partículas	1 ^a	0,116	0,143	85
	2 ^a	0,098	0,102	91
	3 ^a	0,096	0,128	193
40 intervalos 50 partículas	1 ^a	0,097	0,117	215
	2 ^a	0,098	0,127	195
	3 ^a	0,100	0,129	185
40 intervalos 100 partículas	1 ^a	0,097	0,110	171
	2 ^a	0,101	0,125	130
	3 ^a	0,091	0,114	127
60 intervalos 25 partículas	1 ^a	0,103	0,108	250
	2 ^a	0,100	0,131	240
	3 ^a	0,103	0,108	151
60 intervalos 50 partículas	1 ^a	0,097	0,130	152
	2 ^a	0,095	0,140	169
	3 ^a	0,089	0,136	99
60 intervalos 100 partículas	1 ^a	0,107	0,119	187
	2 ^a	0,100	0,120	232
	3 ^a	0,088	0,122	201

ANEXO C - Resultados completos das implementações do iPLS-PSO (CC) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,106	0,129	122
	2 ^a	0,112	0,115	220
	3 ^a	0,106	0,134	69
20 intervalos 50 partículas	1 ^a	0,111	0,139	129
	2 ^a	0,107	0,131	59
	3 ^a	0,112	0,140	75
20 intervalos 100 partículas	1 ^a	0,107	0,131	99
	2 ^a	0,105	0,125	61
	3 ^a	0,101	0,128	96
40 intervalos 25 partículas	1 ^a	0,100	0,112	147
	2 ^a	0,091	0,114	141
	3 ^a	0,116	0,141	99
40 intervalos 50 partículas	1 ^a	0,091	0,114	119
	2 ^a	0,099	0,132	96
	3 ^a	0,096	0,112	232
40 intervalos 100 partículas	1 ^a	0,095	0,147	178
	2 ^a	0,102	0,127	183
	3 ^a	0,097	0,124	223
60 intervalos 25 partículas	1 ^a	0,099	0,142	222
	2 ^a	0,100	0,131	224
	3 ^a	0,099	0,108	133
60 intervalos 50 partículas	1 ^a	0,099	0,127	204
	2 ^a	0,092	0,111	148
	3 ^a	0,091	0,140	227
60 intervalos 100 partículas	1 ^a	0,084	0,140	179
	2 ^a	0,089	0,124	216
	3 ^a	0,087	0,126	204

ANEXO C - Resultados completos das implementações do iPLS-PSO (DD) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,112	0,128	43
	2 ^a	0,111	0,131	3
	3 ^a	0,112	0,122	155
20 intervalos 50 partículas	1 ^a	0,112	0,122	37
	2 ^a	0,101	0,128	62
	3 ^a	0,103	0,136	211
20 intervalos 100 partículas	1 ^a	0,103	0,136	20
	2 ^a	0,103	0,136	45
	3 ^a	0,107	0,131	79
40 intervalos 25 partículas	1 ^a	0,139	0,117	173
	2 ^a	0,101	0,115	49
	3 ^a	0,101	0,114	216
40 intervalos 50 partículas	1 ^a	0,091	0,114	120
	2 ^a	0,102	0,113	81
	3 ^a	0,091	0,114	110
40 intervalos 100 partículas	1 ^a	0,091	0,114	79
	2 ^a	0,096	0,112	241
	3 ^a	0,091	0,114	41
60 intervalos 25 partículas	1 ^a	0,099	0,121	166
	2 ^a	0,109	0,128	60
	3 ^a	0,108	0,133	201
60 intervalos 50 partículas	1 ^a	0,098	0,141	208
	2 ^a	0,090	0,138	245
	3 ^a	0,090	0,118	250
60 intervalos 100 partículas	1 ^a	0,104	0,119	157
	2 ^a	0,086	0,116	131
	3 ^a	0,095	0,130	206

ANEXO C - Resultados completos das implementações do iPLS-PSO (CD) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,131	0,131	80
	2 ^a	0,147	0,147	76
	3 ^a	0,111	0,139	83
20 intervalos 50 partículas	1 ^a	0,101	0,128	108
	2 ^a	0,112	0,122	75
	3 ^a	0,101	0,128	93
20 intervalos 100 partículas	1 ^a	0,111	0,131	48
	2 ^a	0,112	0,122	65
	3 ^a	0,112	0,122	27
40 intervalos 25 partículas	1 ^a	0,091	0,114	80
	2 ^a	0,108	0,111	18
	3 ^a	0,938	0,123	165
40 intervalos 50 partículas	1 ^a	0,091	0,114	74
	2 ^a	0,095	0,108	141
	3 ^a	0,095	0,119	126
40 intervalos 100 partículas	1 ^a	0,087	0,116	175
	2 ^a	0,091	0,114	43
	3 ^a	0,091	0,114	28
60 intervalos 25 partículas	1 ^a	0,100	0,119	185
	2 ^a	0,110	0,136	97
	3 ^a	0,106	0,123	49
60 intervalos 50 partículas	1 ^a	0,105	0,131	135
	2 ^a	0,108	0,134	142
	3 ^a	0,089	0,128	167
60 intervalos 100 partículas	1 ^a	0,106	0,124	82
	2 ^a	0,100	0,109	23
	3 ^a	0,084	0,112	209

ANEXO C - Resultados completos das implementações do iPLS-PSO (DC) para a quantificação do óleo diesel.

Enxame	Execução	RMSECV	RMSEP	Obtido na iteração
20 intervalos 25 partículas	1 ^a	0,108	0,131	52
	2 ^a	0,112	0,121	207
	3 ^a	0,111	0,139	159
20 intervalos 50 partículas	1 ^a	0,107	0,131	27
	2 ^a	0,111	0,139	101
	3 ^a	0,101	0,128	135
20 intervalos 100 partículas	1 ^a	0,101	0,127	93
	2 ^a	0,111	0,139	53
	3 ^a	0,105	0,125	53
40 intervalos 25 partículas	1 ^a	0,108	0,111	233
	2 ^a	0,109	0,126	66
	3 ^a	0,116	0,143	81
40 intervalos 50 partículas	1 ^a	0,110	0,116	211
	2 ^a	0,102	0,113	203
	3 ^a	0,096	0,118	51
40 intervalos 100 partículas	1 ^a	0,103	0,121	194
	2 ^a	0,105	0,116	59
	3 ^a	0,986	0,131	35
60 intervalos 25 partículas	1 ^a	0,100	0,131	46
	2 ^a	0,103	0,126	150
	3 ^a	0,109	0,128	106
60 intervalos 50 partículas	1 ^a	0,109	0,128	121
	2 ^a	0,099	0,145	57
	3 ^a	0,101	0,153	192
60 intervalos 100 partículas	1 ^a	0,100	0,121	236
	2 ^a	0,098	0,128	194
	3 ^a	0,108	0,118	231