UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA

FRANCISCO FALLA SOTELO

**Aplicação da espectroscopia de Infravermelho
Próximo na caracterização de petróleo.
Simulação de uma unidade de destilação atmosférica**

São Paulo
2006

FRANCISCO FALLA SOTELO

**Aplicação da espectroscopia de Infravermelho
Próximo na caracterização de petróleo.
Simulação de uma unidade de destilação atmosférica**

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção do
título de Doutor em Engenharia Química.

Área de Concentração: Engenharia Química

Orientador: Prof. Dr. Galo Antonio Carrillo
Le Roux

São Paulo
2006

Este exemplar foi revisado e alterado em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 16 de novembro de 2006.

Assinatura do autor _____

Assinatura do orientador _____

Falla Sotelo, Francisco

**Aplicação da espectroscopia de infravermelho próximo na caracterização de petróleo : simulação de uma unidade de destilação atmosférica / F.Falla Sotelo. -- ed.rev. -- São Paulo, 2006.
154 p.**

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Química.

1. Caracterização de petróleo 2. Espectroscopia infravermelha 3. Calibração multivariada 4. Redes neurais 5. Destilação atmosférica (Simulação) 6. Curvas True Boiling Point-TBP 7. Curvas de destilação simulada-SimDis I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Química II. t.

AGRADECIMENTOS

Ao Galo Carrillo Le Roux, pela atenção e apoio durante o processo de orientação e, sobretudo, pela amizade e confiança.

À CAPES, pela concessão da bolsa de doutorado e pelo apoio financeiro para a realização desta pesquisa.

À Petrobras, por colocar a disposição as informações necessárias para a realização desta pesquisa.

Ao Rubens, Sérgio, Carlos Arturo, Rita, Jeanne, Marlon, Claudia e Reinaldo, os amigos que fizeram com que a minha adaptação neste país fosse rápida e prazerosa.

Aos professores do LSCP, à Carminha, Terezinha e Silvia, pela colaboração e confiança depositadas, que me fez sentir parte desta equipe de trabalho extremamente profissional.

Ao Cláudio Oller, pela valiosa orientação no âmbito profissional e pessoal.

À Sandra, Edna, Filipe, Alberto e Lincoln, da Petrobras, pelas valiosas discussões e orientações que permitiram a obtenção dos resultados obtidos nesta pesquisa.

Ao Pedro, Simone e Edgar, da Aspentech, pela valiosa orientação na utilização do Hysys na simulação do processo.

Aos amigos da dança, com os quais aprendi a gostar ainda mais do Brasil, e que me ensinaram a desfrutar das coisas simples da vida.

À Patricia, em nome das lembranças dos tempos da graduação.

À Flor, Toño e Alessandro, pelo apoio familiar no início desta jornada.

À Rocio e Nelson, pela convivência diária tão gratificante e pela amizade sincera.

À Cris, obrigado pelo seu amor e apoio incondicionais.

Aos meus pais, Flor e Francisco, pelo amor e sacrifício demonstrados, pelo orgulho de ser seu filho.

À Deus, aquela força espiritual maravilhosa sem a qual não poderia fazer absolutamente nada.

RESUMO

A contínua mudança na qualidade da carga de alimentação das unidades de destilação atmosférica introduz incertezas na operação do processo de refino. A caracterização precisa do petróleo bruto, cuja composição varia devido à sua natureza intrínseca e às condições de transporte e armazenamento nas refinarias, poderia auxiliar na redução destas incertezas. O processo clássico de caracterização, baseado nas curvas de destilação TBP (*True Boiling Point*, Ponto de Ebulição Verdadeiro), ainda não permite detectar essas mudanças de modo mais dinâmico, o que reduz a precisão da simulação do processo e, portanto, uma menor confiabilidade nas condições de operação determinadas a partir dela. No Brasil, petróleo pesado é produzido de forma predominante, que é misturado frequentemente com petróleo leve importado antes do refino. Como consequência, nas refinarias brasileiras as características do petróleo mudam em média, a cada três dias, sendo esta taxa de variação incompatível com a determinação das curvas de destilação TBP. A análise SimDis (*Simulated Distillation*, Destilação Simulada) é atualmente a técnica de laboratório alternativa para aproximar as curvas de destilação. Porém, ainda apresenta algumas características que a tornam tão limitantes quanto a análise TBP. A espectroscopia NIR (*Near InfraRed*, Infravermelho Próximo), tem se consagrado como uma técnica analítica rápida e viável em inúmeras aplicações industriais na determinação de diversos parâmetros qualitativos. Neste trabalho, metodologias inovadoras de caracterização são desenvolvidas visando reproduzir com maior dinamismo o processo de refino. A reprodução das curvas de destilação das cargas ao processo de destilação atmosférica da Refinaria de Capuava (RECAP – Petrobras S.A.) através do seu espectro NIR tem sido o ponto de partida no desenvolvimento do trabalho. Métodos de calibração multivariada e redes neurais foram utilizados na construção de algoritmos computacionais que permitem a previsão dos perfis de destilação do petróleo processado, os quais foram utilizados na simulação do processo para avaliar o impacto da aplicação desta metodologia. A boa qualidade da reprodução das propriedades físico-químicas do petróleo através do NIR foi comprovada, uma vez que os modelos construídos foram condicionados às características da matéria analisada. Quanto à reprodução do cenário real do processo através da simulação, os resultados obtidos são promissores, mostrando que a caracterização da carga é útil para reproduzir mudanças no processamento. Porém, alguns aspectos precisam ainda ser aprofundados visando o melhor desempenho das metodologias. Contudo, os resultados desta pesquisa reforçam a relevância do papel da espectroscopia NIR, juntamente com técnicas clássicas de laboratório utilizadas, numa nova concepção de caracterização, que tem na rapidez e confiabilidade os seus principais aliados e pode ser aplicada na simulação de um processo visando uma melhor qualidade de operação.

ABSTRACT

The continuous variation in the quality of the feed petroleum to atmospheric distillation unities brings uncertainty to the operation of the refining process. The characterization of the crude oil, whose composition varies due to its intrinsic nature and the transport conditions and storage in refineries, could be useful in reducing the uncertainty. The classic process of characterization, based on TBP distillation curves (True Boiling Point), still does not permit to detect these changes in a more dynamic manner, which reduces the precision of the process simulation and, consequently, less confidence in the operation conditions determined from it. In Brazil, heavy petroleum is produced predominantly, which is typically blended with light imported petroleum before refining. As a consequence, in Brazilian refineries the characteristics of the petroleum change, on the average, about every three days. This rate of the petroleum change is incompatible with the determination of TBP curves. Nowadays, SimDis analysis (Simulated Distillation) is an alternative laboratory technique able to approximate the distillation curves. However, it still presents some limiting characteristics like those of TBP analysis. In the last decades, the NIR spectroscopy (Near InfraRed) has showed to be a both fast and feasible analytical technique for innumerable industrial applications in the determination of various qualitative parameters. In the present work, an innovative characterization methodology is developed aiming to reproduce more dynamically the refining process. The reproduction of distillation curves of the feed to the atmospheric distillation process of the Refinery of Capuava (RECAP – Petrobras S.A.) from its NIR spectral determination has been the initial point for the development. Multivariate calibration methods and neural networks have been used in the construction of computational algorithms that allow the prediction of distillation profiles of the petroleum processed. These profiles were applied in process simulation in order to evaluate the impact of the application of this methodology. The good quality of the physicochemical properties of petroleum reproduction using NIR was confirmed, it is due to the fact that constructed models were conditioned to the characteristics of the samples analyzed. With respect to the reproduction of the real process behavior through simulation, the results obtained are promising, which show that the feed characterization is useful in order to reproduce process variations. However, some aspects need to be analyzed deeply in order to improve the performance of the methodologies. Finally, the results of this research strengthen the role of NIR spectroscopy together with classic laboratory techniques used in a new characterization concept, whose major advantages are the rapidity and the trustworthiness of its predictions, that can be applied in order to improve the quality of the simulation of a process.

LISTA DE FIGURAS

Figura 2.1	Exemplo da distribuição de produtos destilados numa curva TBP	25
Figura 2.2	O espectro eletromagnético	27
Figura 3.1	Espectros NIR e MIR do querosene (Chung <i>et al.</i> ,1999)	31
Figura 3.2	Gráfico de paridade no NIR e MIR e as análises ASTM D86 para as temperaturas de recuperação (Chung <i>et al.</i> ,1999)	31
Figura 3.3	Espectros NIR dos seis produtos de petróleo (Kim <i>et al.</i> ,2000).....	32
Figura 3.4	Gráfico dos componentes principais PC1 x PC2 (Kim <i>et al.</i> ,2000)	33
Figura 3.5	Espectros típicos Raman, IR e NIR do AR (Chung e Ku,2000).....	34
Figura 3.6	(a) Espectros brutos e (b) derivada segunda do espectro NIR do AR no conjunto de calibração (Chung e Ku,2000)	34
Figura 3.7	Espectro NIR médio do conjunto de calibração (Hidajat e Chong, 2000)	35
Figura 3.8	Gráfico de paridade do modelo correspondente à fração de destilado leve (Hidajat e Chong, 2000)	36
Figura 3.9	Determinação das curvas TBP em % peso (A) e em % volume (B) de uma amostra de petróleo bruto através da destilação convencional e da metodologia proposta (Roussis e Fitzgerald, 2000)	38
Figura 3.10	Comparação dos dados TBP obtidos em % peso (A) e em % volume (B) para 11 amostras de petróleo (Roussis e Fitzgerald, 2000)	39
Figura 4.1	Espectrômetro FT-NIR IFS 28/N (BRUKER) e sensores NIR de imersão (HELLMA).....	44
Figura 4.2	Espectros NIR a diferentes temperaturas para duas amostras A e B	45
Figura 4.3	Espectros NIR residuais para (a) $\Delta=10^{\circ}\text{C}$ e (b) $\Delta=20^{\circ}\text{C}$	46
Figura 4.4	Esquema da medição dos espectros NIR de petróleo bruto.....	48
Figura 4.5	Ilustração teórica da validação de um modelo.....	50
Figura 4.6	Resultados dos testes de reprodutibilidade para duas amostras analisadas	51

Figura 4.7	Espectros NIR das amostras de petróleo analisadas	52
Figura 4.8	Regiões espectrais NIR definidas	54
Figura 4.9	Efeito combinado do Alisamento e MSC nos espectros NIR.....	55
Figura 4.10	Gráficos de paridade para os modelos à 200°C (Caso A) e à 400°C (Caso B) utilizando PCR e PLS, respectivamente.....	56
Figura 4.11	Gráficos de paridade para alguns dos modelos NIR das curvas TBP do petróleo bruto analisado	60
Figura 4.11 (cont.)	Gráficos de paridade para alguns dos modelos NIR das curvas TBP do petróleo bruto analisado	61
Figura 4.11 (cont.)	Gráficos de paridade para alguns dos modelos NIR das curvas TBP do petróleo bruto analisado	62
Figura 4.12	Comparação entre as curvas TBP obtidas: ASTM (curva medida ou de referência) e NIR (curva prevista ou estimada).....	63
Figura 4.12 (cont.)	Comparação entre as curvas TBP obtidas: ASTM (curva medida ou de referência) e NIR (curva prevista ou estimada).....	64
Figura 5.1	Desvio padrão relativo na análise de reprodutibilidade.....	68
Figura 5.2	Espectros NIR de algumas amostras (a) sem pré-tratamento (b) derivada primeira e (c) após “centralização”	69
Figura 5.2 (cont.)	Espectros NIR de algumas amostras (a) sem pré-tratamento (b) derivada primeira e (c) após “centralização”	70
Figura 5.3	Curvas SimDis de amostras de petróleo com o mesmo grau API (33.4).	71
Figura 5.4	Espectros NIR pré-tratados de amostras de petróleo com o mesmo grau API (33.4)	72
Figura 5.5	Representação esquemática do processo da previsão das curvas SimDis de petróleo bruto através das redes neurais	73
Figura 5.6	Comparação entre as curvas SimDis medidas (através de normas ASTM) e previstas (através do NIR) para amostras com grau API (a) 31.1 e (b) 36.4.....	74
Figura 6.1	Processo de destilação atmosférica da refinaria (Dessalgação)	78
Figura 6.2	Processo de destilação atmosférica da refinaria (Torre pré-flash)	79

Figura 6.3	Processo de destilação atmosférica da refinaria (Torre atmosférica)	80
Figura 6.4	Esquema do processo utilizado na simulação (a) Torre de Pré-Flash e (b) Torre Atmosférica	82
Figura 6.5	Esquema de simulação convencional	92
Figura 6.6	Destilações SimDis e TBP típicas (adaptado de Daubert, 1994).....	94
Figura 6.7	Esquema de simulação utilizando o método API de conversão das curvas SimDis em TBP	94
Figura 6.8	Curva ASTM D-86 de Nafta Leve ou NL: (a) conjunto I e (b) conjunto II	100
Figura 6.9	Curva ASTM-D86 de Diesel Leve ou DL: (a) conjunto I e (b) conjunto II	101
Figura 6.10	Curva ASTM-D86 de Nafta Pesada ou NP: (a) conjunto I e (b) conjunto II.	105
Figura 6.11	Curvas ASTM D-86 de Querosene ou Kero: (a) conjunto I e (b) conjunto II	106
Figura 6.12	Curvas ASTM D-86 de Diesel Pesado ou DP: (a) conjunto I e (b) conjunto II	107
Figura 6.13	Curvas TBP de Resíduo Atmosférico ou RAT: (a) conjunto I e (b) conjunto II	108
Figura A.1	Aumento da resolução pelo uso das derivadas: sinal original, <i>linha sólida</i> ; e derivada segunda do sinal original, <i>linha pontilhada</i> (Siesler <i>et al.</i> , 2002).....	125
Figura A.2	Alisamento através de mínimos quadrados. Os círculos indicam os pontos originais do sinal e a linha pontilhada indica o sinal alisado (Gorry, 1990)	127
Figura A.3	Espectros de absorbância NIR (a) antes do MSC e (b) depois do MSC (Naes e Isaksson, 1994b).....	128
Figura A.4	Esquema de construção do modelo de calibração	129
Figura A.5	Processo de compressão de dados	131
Figura A.6	Exemplo da decomposição da informação contida nos espectros	132

Figura A.7	Rede neural <i>feed-forward</i> de três camadas	134
Figura A.8	Representação esquemática da validação cruzada.....	138
Figura A.9	Gráfico de Influência (resíduos vs. <i>leverage</i>). Os pontos ‘a’ e ‘b’ são pontos discrepantes (Kowalski e Seasholtz, 1991).....	140
Figura B.1	Espectrômetro FTIR modelo FTLA2000-160 (ABB BOMEM), (a) vista frontal e (b) detalhe do uso das fibras óticas	142
Figura B.2	Representação esquemática do processo de transferência.....	143
Figura B.3	Espectros do subconjunto de transferência (medição realizada em ambos os espectrômetros)	144
Figura B.4	Diferenças entre os espectros coletados em ambos espectrômetros	145
Figura B.5	Comparação entre as curvas TBP previstas antes (AT) e depois (DT) da transferência dos modelos com a curva TBP medida (original).....	148
Figura B.6	Representação esquemática da técnica PDS (adaptado de Bouveresse e Massart, 1996b)	149
Figura B.7	Erros de previsão das técnicas CSB e PDS na transferência de calibração dos modelos NIR das curvas TBP	153

LISTA DE TABELAS

Tabela 1.1	Derivados destilados típicos do petróleo	19
Tabela 4.1	Sub-regiões NIR definidas para a seleção da faixa adequada para análise	53
Tabela 4.2	Resultados da calibração PLS dos modelos NIR.....	62
Tabela 5.1	Condições de operação para o cromatógrafo gasoso na análise SimDis	67
Tabela 6.1	Identificadores dos dados coletados na RECAP.....	84
Tabela 6.2	Composição de leves obtida através do balanço de massa nas cargas.....	89
Tabela 6.3	Análise volumétrica da corrente de GLP na RECAP	91
Tabela 6.4	Manipulação na parte inicial das curvas SimDis	95
Tabela 6.5	Variáveis especificadas na simulação da torre pré-flash	97
Tabela 6.6	Resultados em termos de vazões e temperaturas na torre pré-flash para o conjunto de especificações I.....	98
Tabela 6.7	Resultados em termos de vazões na torre pré-flash para o conjunto de especificações II	98
Tabela 6.8	Vazões de GLP obtidas na torre pré-flash	99
Tabela 6.9	Variáveis especificadas na simulação da torre atmosférica.....	102
Tabela 6.10	Resultados em termos de vazões e temperaturas na torre atmosférica para o conjunto de especificações I.....	103
Tabela 6.11	Resultados em termos de vazões na torre atmosférica para o conjunto de especificações II	103
Tabela 6.12	Resultados das inferências nas simulações (conjunto I)	109
Tabela 6.13	Resultados das inferências nas simulações (conjunto II)	109
Tabela B.1	Resultados da aplicação da técnica CSB na transferência dos modelos NIR das curvas TBP, em termos de RMSEP	147
Tabela B.2	Casos estudados na aplicação da técnica PDS para a transferência dos modelos NIR das curvas TBP.....	151

Tabela B.3	Resultados da aplicação da PDS na transferência dos modelos NIR, para os casos (A) 3 PJE - 2 PCs (B) 3 PJE – 3 PCs e (D) 5 PJE – 3 PCs, em termos de RMSEP.....	151
-------------------	---	-----

LISTA DE SIGLAS

API	Instituto Americano de Petróleo (<i>American Petroleum Institute</i>)
AR	Resíduo Atmosférico (<i>Atmospheric Residue</i>)
ASTM	American Society for Testing and Materials
CENPES	Centro de Pesquisas e Desenvolvimento Leopoldo Américo Miguez de Mello
CSB	Correção <i>Slope/Bias</i>
CV	Validação Cruzada (<i>Cross Validation</i>)
EFV	Equilibrium Flash Vaporization
GCD	Gas Chromatography Distillation
GLP	Gás Liquidificado de Petróleo
HPLC	High Performance Liquid Chromatography
IR	Infravermelho (<i>Infrared</i>)
LGO	Gasóleo leve (<i>Light GasOil</i>)
LWR	Regressão localmente ponderada (<i>Locally Weighted Regression</i>)
MIR	Infravermelho Médio (<i>Medium InfraRed</i>)
MLR	Regressão Linear Múltipla (<i>Multiple Linear Regression</i>)
MON	Motor Octane Number
MS	Mass Spectrometry
MSC	Multiplicative Scatter Correction
MSE	Erro Médio Quadrático (<i>Mean Squared Error</i>)
MSECV	Erro Médio Quadrático por Validação Cruzada (<i>Mean Squared Error by Cross Validation</i>)
MSID	Maximização da menor distância entre pontos (<i>Maximization of the Smallest Interpoint Distance</i>)
NIPALS	Mínimos Quadrados Parciais Iterativos Não-Lineares (<i>Nonlinear Iterative Partial Least Squares</i>)
NIR	Infravermelho Próximo (<i>Near InfraRed</i>)
NMR	Nuclear Magnetic Resonance
NN	Redes Neurais (<i>Neural Network</i>)
PC	Componente Principal (<i>Principal Component</i>)

PCA	Análise por Componentes Principais (<i>Principal Component Analysis</i>)
PCR	Regressão por Componentes Principais (<i>Principal Component Regression</i>)
PDS	Padronização direta por etapas (<i>Piecewise Direct Standardization</i>)
Petrobras	Petróleo Brasileiro S.A.
PEV	Ponto de Ebulição Verdadeiro
PLS	Mínimos Quadrados Parciais (<i>Partial Least Squares</i>)
RECAP	Refinaria de Capuava
RMSE	Raiz quadrada do erro médio quadrático (<i>Root Mean Squared Error</i>)
RMSECV	Raiz quadrada do erro médio quadrático por validação cruzada (<i>Root Mean Squared Error by Cross Validation</i>)
RMSEP	Raiz quadrada do erro médio quadrático de previsão (<i>Root Mean Squared Error Prediction</i>)
RON	Research Octane Number
SIMDIS	Destilação Simulada (<i>Simulated Distillation</i>)
TBP	Ponto de ebulição verdadeiro (<i>True Boiling Point</i>)
UV	Ultravioleta (<i>Ultraviolet</i>)

SUMÁRIO

AGRADECIMENTOS

RESUMO

ABSTRACT

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SIGLAS

1	INTRODUÇÃO	19
1.1	O Petróleo	19
1.2	Problemática atual.....	20
1.3	Objetivos do projeto	21
1.4	Estrutura da dissertação	22
2	CARACTERIZAÇÃO DE PETRÓLEO.....	24
2.1	Curvas de Destilação	24
2.1.1	Curvas TBP (True Boiling Point)	24
2.1.2	Curvas SimDis (Simulated Distillation)	25
2.2	Espectroscopia NIR	26
2.2.1	Justificativa de aplicação.....	28
3	PESQUISAS DESENVOLVIDAS – ANTECEDENTES.....	30
3.1	Petróleo e Espectroscopia NIR	30
3.2	Petróleo e Análise SimDis	37
3.3	Potencialidade na análise do petróleo bruto	40
4	PREDIÇÃO DAS CURVAS TBP ATRAVÉS DO NIR.....	43

4.1	Considerações Gerais.....	43
4.1.1	Base de dados	43
4.1.2	Instrumentação	43
4.1.3	Efeito da Temperatura.....	44
4.2	Metodologia.....	47
4.2.1	Experimental	47
4.2.2	Computacional	48
4.3	Condicionamento espectral à análise TBP do petróleo	51
4.3.1	Reprodutibilidade.....	51
4.3.2	Faixa espectral.....	52
4.3.3	Pré-tratamento	53
4.4	Modelos NIR das curvas TBP	56
4.4.1	Técnica de calibração.....	56
4.4.2	Resultados	58
5	PREDIÇÃO DAS CURVAS SIMDIS ATRAVÉS DO NIR	65
5.1	Considerações Gerais.....	65
5.1.1	Base de dados	65
5.1.2	Instrumentação	65
5.2	Metodologia.....	66
5.2.1	Preparação das amostras e análise cromatográfica	66
5.2.2	Reprodutibilidade espectral.....	67
5.2.3	Pré-tratamento espectral.....	68
5.2.4	Técnica de calibração.....	72
5.3	Modelos NIR das curvas SimDis – Resultados	73
6	CASO DE ESTUDO: PROCESSO DE DESTILAÇÃO ATMOSFÉRICA	76
6.1	Justificativa	76
6.2	Descrição do Processo	77
6.3	Base de dados industriais.....	82

6.4	Simulador do processo.....	86
6.5	Informação da fração de leves	88
6.6	Estratégias de Simulação	91
6.6.1	Estratégia Convencional.....	92
6.6.2	Conversão das curvas SimDis em TBP – Método API.....	92
6.6.3	Manipulação da parte inicial das curvas de destilação.....	95
6.7	Resultados.....	96
6.7.1	Simulação da Torre Pré-Flash.....	96
6.7.2	Simulação da Torre Atmosférica	102
7	CONCLUSÕES	111
7.1	Modelagem das curvas de destilação a partir do NIR	111
7.2	Simulação do processo de destilação atmosférica	113
7.3	Divulgação dos resultados da pesquisa.....	115
7.4	Propostas para a continuação da pesquisa	115
	REFERÊNCIAS	117
	APÊNDICE A - QUIMIOMETRIA. TRATAMENTO DE DADOS NIR	124
A.1	Pré-tratamento espectral	124
A.1.1	Derivação	125
A.1.2	Alisamento	126
A.1.3	MSC (Multiplicative Scatter Correction).....	127
A.2	Construção do modelo	128
A.2.1	Técnicas Lineares.....	130
A.2.2	Técnicas não-lineares	132
A.3	Validação	136
A.3.1	Validação Externa.....	137
A.3.2	Validação Interna	137

A.4	Pontos Discrepantes. Outliers	139
APÊNDICE B - TRANSFERÊNCIA DE CALIBRAÇÃO.....		141
B.1	Seleção do subconjunto de transferência.....	144
B.2	Técnica de correção slope/bias (CSB).....	145
B.2.1	Aplicação nos modelos NIR das curvas TBP	147
B.3	Técnica da Piecewise Direct Standardization (PDS).....	148
B.3.1	Aplicação nos modelos NIR das curvas TBP	150

1 INTRODUÇÃO

1.1 O Petróleo

O petróleo é uma mistura altamente complexa de milhares de compostos orgânicos diferentes formados a partir de uma variedade de matéria orgânica convertida quimicamente sob condições geológicas apropriadas em períodos de tempo muito longos (Wang, 1999). O petróleo contém principalmente hidrocarbonetos (C, H) assim como compostos orgânicos heteroatômicos (S, N, O) de peso molecular e polaridade variável, e metais tais como Ni, V e Pb, em menor concentração.

Em estado bruto, o petróleo não tem aplicabilidade prática, sendo o refino o processo mediante o qual ele é desmembrado em cortes ou frações comercialmente úteis. Na tabela 1.1 apresentam-se os principais derivados. Porém, a natureza infinitamente variável de fatores composicionais faz com que todos os crus e produtos de petróleo processados numa refinaria sejam diferentes entre si. Essa variabilidade representa uma característica química do tipo *fingerprint* ou impressão digital para cada petróleo e fornece uma base para caracterizá-lo.

Tabela 1.1 Derivados destilados típicos do petróleo

Derivado	Faixa de temperatura (curva PEV)
GLP (C ₃ – C ₄)	15 °C
Nafta (C ₅ – C ₁₀)	15 – 150 °C
Querosene	150 – 250 °C
Diesel *	250 – 380 °C
Gasóleo	380 – 550 °C
Resíduo	Acima de 550 °C

* Diesel brasileiro: 150 – 400 °C

A caracterização, feita através de ensaios para a determinação de propriedades físicas, físico-químicas e químicas, procura avaliar a qualidade intrínseca do petróleo e seu potencial de refino (rendimento e qualidade dos derivados). Entre os principais ensaios pode-se citar os seguintes:

- Volatilidade: curvas TBP (*True Boiling Point*) ou PEV (Ponto de Ebulição Verdadeiro), e densidade API (*American Petroleum Institute*). As curvas TBP representam o perfil geral do petróleo analisado, estimando o rendimento dos seus derivados ou frações. A densidade API é o ensaio de densidade mais rápido e direto existente em laboratório e indica quão pesado o petróleo analisado é.
- Fluidez: ponto de fluidez, viscosidade.
- Combustão: resíduo de carbono.
- Composição: enxofre, acidez, nitrogênio total e básico, metais.
- Poluentes: sal, água, sedimentos, cinzas.

Dentre estes ensaios, realça-se a importância das análises TBP e de densidade, consideradas índices de qualidade importantes para avaliar petróleos crus ou brutos, pois contêm a informação necessária para se avaliar o seu rendimento global (Hidajat e Chong, 2000). Isto é, as faixas de destilação determinadas a partir da curva TBP são fundamentais para se calcular a disponibilidade das frações de petróleo para diversas aplicações particulares. Tal informação é também necessária para o projeto e controle das colunas de destilação. Juntamente com a densidade, permite caracterizar frações de petróleo o que, por sua vez, permite inferir outras propriedades (Hobson, 1984).

1.2 Problemática atual

A instabilidade no processo de refino, causada principalmente pelas deficiências existentes na caracterização do petróleo, ocasiona dificuldades na otimização das condições de operação das torres de destilação.

Os ensaios convencionais para a caracterização de petróleo (como os citados no item anterior) são na sua maioria demorados e de alto custo. Assim, um aparelho típico

TBP custa aproximadamente 0.5 milhão de dólares sendo necessários outros 10 mil dólares em custos anuais de operação e manutenção (Hidajat e Chong, 2000). Além disso, cada análise TBP precisa da competência técnica de um operador devidamente treinado e pode demorar até três dias.

Devido a isto, a frequência das análises é insuficiente para representar as variações na qualidade de um mesmo petróleo. A situação torna-se mais complexa pelo fato de que muitas refinarias utilizam uma ampla gama de petróleos, estando sujeitas à mistura natural nos tanques de armazenamento e a posterior estratificação. Isto faz com que a composição do petróleo alimentado ao processo varie, obrigando o operador a controlá-lo de forma conservativa, devido à falta de análises rápidas necessárias para a tomada oportuna de decisões no controle da operação.

Assim, além da necessidade técnica de uma mudança na abordagem deste problema no processo de refino, a demanda por essa melhora responde também à pressão econômica crescente sobre as refinarias para reduzir os custos do processo com a finalidade de permanecerem competitivas. Nesta ótica, uma alternativa ao desafio atual é o uso de novas técnicas de análise capazes de caracterizar o petróleo de maneira eficaz reduzindo os custos em tempo e dinheiro, cuja metodologia poderia ser utilizada juntamente aos dados das correntes de processo para propósitos de simulação e controle do refino.

1.3 Objetivos do projeto

O objetivo geral do projeto é o desenvolvimento de novas metodologias de caracterização de petróleo para o processo de refino. Isto implica no cumprimento de dois objetivos específicos;

O primeiro, o uso de novas técnicas que permitam maior frequência e estabilidade na caracterização do petróleo alimentado à unidade de destilação atmosférica. O NIR (*Near InfraRed* ou Infravermelho Próximo) é um tipo de espectroscopia que tem se consagrado nas últimas décadas como um método rápido e factível em inúmeras

aplicações industriais para a previsão de propriedades físico-químicas. No mundo todo, muitas refinarias possuem espectrômetros NIR instalados em diferentes processos. Neste projeto, será estudada a análise NIR do petróleo bruto proveniente da Refinaria de Capuava (RECAP – Petrobras S.A.) que permita obter informação útil quanto a sua caracterização considerando o efeito da variação nas condições de operação.

O segundo, utilizar a informação da caracterização obtida através da espectroscopia NIR na simulação de uma unidade de destilação atmosférica, e avaliar o efeito da metodologia proposta em termos da qualidade da simulação, visando que a disponibilidade dessa informação (*feed-forward*) permita ajustar as condições de operação na forma mais adequada segundo as exigências de produção requeridas pela planta.

1.4 Estrutura da dissertação

Este trabalho contém, além deste capítulo inicial, mais sete capítulos e dois apêndices:

Capítulo 2: *Caracterização de petróleo*. Descreve-se inicialmente a metodologia convencional de caracterização do petróleo processado nas refinarias; através das curvas de destilação TBP e, mais recentemente, as curvas SimDis (*Simulated Distillation* ou Destilação Simulada). Por último, descreve-se a tecnologia utilizada neste trabalho como ponto de partida para desenvolver uma nova metodologia de caracterização, a espectroscopia NIR.

Capítulo 3: *Pesquisas desenvolvidas – Antecedentes*. Contém uma revisão bibliográfica sobre os trabalhos experimentais que tratam da aplicação do NIR na indústria do petróleo, assim como o papel da análise SimDis neste campo. A avaliação das pesquisas feitas permite estabelecer uma base inicial para direcionar o desenvolvimento das metodologias propostas neste trabalho.

Capítulo 4: *Predição das curvas TBP através do NIR*. A metodologia desenvolvida para correlacionar os espectros NIR de amostras de petróleo bruto e as respectivas curvas TBP é apresentada. Técnicas quimiométricas são utilizadas e os resultados da previsão são apresentados.

Capítulo 5: *Predição das curvas SimDis através do NIR*. A metodologia desenvolvida para correlacionar os espectros NIR de amostras de petróleo bruto e as respectivas curvas SimDis é apresentada. Redes Neurais são utilizadas e os resultados da sua aplicação são apresentados.

Capítulo 6: *Caso de estudo: Processo de Destilação Atmosférica*. O processo industrial tomado como alvo da aplicação deste trabalho é apresentado em termos de unidades, variáveis e parâmetros, assim como os fatores relevantes avaliados na simulação do mesmo. As novas metodologias de caracterização são aplicadas na simulação do processo de refino e o impacto na sua qualidade é discutido. Os resultados são apresentados, e as limitações identificadas.

Capítulo 7: *Conclusões*. As principais conclusões do trabalho são salientadas, assim como as suas contribuições e propostas para a continuação da pesquisa.

Apêndice A: *Quimiometria. Tratamento de dados NIR*. Uma revisão das principais técnicas lineares e não lineares que foram utilizadas na construção dos modelos NIR na previsão das curvas de destilação é apresentada neste apêndice.

Apêndice B: *Transferência de calibração*. As metodologias que permitem a aplicação de um modelo de calibração em situações diferentes para as quais ele foi desenvolvido são apresentadas, assim como os resultados da transferência dos modelos NIR na estimação das curvas TBP desenvolvidos neste trabalho.

2 CARACTERIZAÇÃO DE PETRÓLEO

2.1 Curvas de Destilação

A distribuição dos pontos de ebulição é um dos fatores críticos relacionados à avaliação e processamento de petróleos brutos. Assim, a seleção de um petróleo com quantidades limitadas de material que pode entrar em ebulição numa determinada faixa de temperatura tem um impacto negativo na capacidade de previsão dos requerimentos dos produtos e conseqüentemente, afetar em grande escala a economia da refinaria (Roussis e Fitzgerald, 2000). Portanto, é de extrema importância dispor de métodos capazes de fornecer de maneira precisa os perfis de destilação (ou curvas de destilação) num espaço de tempo razoável que permita a tomada oportuna de decisões para o controle do processo de refino. Nesse sentido, a indústria do petróleo conta com duas análises, uma já tradicional (análise TBP) e a outra mais recente (análise SimDis) que tem ganhado mais espaço nas últimas décadas, e que são apresentadas a seguir.

2.1.1 Curvas TBP (*True Boiling Point*)

Dois procedimentos constituem a metodologia clássica para a determinação das curvas de destilação TBP: a norma ASTM D-2892 (sob pressão atmosférica), para componentes que entram em ebulição a temperaturas menores que 400°C e a norma ASTM D-5236 (sob pressão reduzida), para componentes que entram em ebulição a temperaturas maiores que 400°C. De forma geral, uma curva TBP representa a percentagem em volume de cru destilado para diferentes temperaturas. A partir dela, é possível prever o rendimento em termos de produtos daquele petróleo analisado (ver figura 2.1).

As curvas TBP constituem a metodologia mais eficiente e confiável para caracterizar petróleos brutos. No entanto, acarretam um procedimento com um consumo de tempo considerável (até três dias para ser completada), cuja principal desvantagem reside no fato que se torna impossível representar em tempo real a composição do

petróleo que está sendo processado na torre de destilação, por causa da perda de leves, da estratificação nos tanques que provoca a mudança na composição dos petróleos, etc.

Figura 2.1 Exemplo da distribuição de produtos destilados numa curva TBP

2.1.2 Curvas SimDis (*Simulated Distillation*)

Obter as curvas de destilação através de cromatografia gasosa constitui o princípio desta técnica, que foi padronizada na norma ASTM D-2887 (em 1973) para a destilação de frações de petróleo (Grob, 1995). O objetivo dela é a determinação de resultados de qualidade equivalente aos obtidos com métodos que exigem maior tempo de análise (como o ensaio ASTM D-2892). Durand (1998), Raia *et al.* (2000), Sharma *et al.* (2003) são somente alguns trabalhos que tratam do avance da tecnologia desta técnica.

Atualmente, é a norma ASTM D-5307 que permite a determinação de perfis de destilação de amostras de petróleo bruto com temperaturas de ebulição até 538°C. Como resultado, a curva SimDis representa a percentagem em peso de cru destilado para diferentes temperaturas, sendo uma das suas principais desvantagens a impossibilidade de medir, como na análise TBP, o ponto inicial de ebulição das amostras, como produto da perda de voláteis durante o seu condicionamento.

O princípio é simples: a amostra é introduzida na coluna cromatográfica que separa os hidrocarbonetos segundo os seus pontos de ebulição; e com a ajuda de uma mistura de parafinas normais de pontos de ebulição conhecidos, é estabelecida a correspondência entre os tempos de retenção e os pontos de ebulição, construindo a partir dela a curva de destilação como descrita acima. A análise SimDis leva aproximadamente três horas para ser realizada (isto inclui a preparação das amostras e duas varreduras cromatográficas, com e sem padrão interno). Além disso, outras duas análises cromatográficas: de um branco e de um padrão (para a calibração do tempo de eluição no ponto de ebulição) devem ser feitas periodicamente, a cada duas ou três análises de amostras. Apesar da aceitação generalizada deste método e da sua precisão, estes fatos fazem com que haja certas dificuldades em automatizá-la (Durand *et al.*, 1999).

Pode-se dizer, portanto, que a procura por novas técnicas de caracterização obedecem ao objetivo do compromisso entre o tempo de análise e a necessidade de informação com precisão adequada para o controle do processo. Nessa ótica, é apresentada a espectroscopia NIR.

2.2 Espectroscopia NIR

O espectro NIR corresponde ao intervalo de comprimentos de onda de 780 a 2500 nm (números de onda de 12800 a 4000 cm^{-1}). Na figura 2.2 apresenta-se o espectro eletromagnético dividido em regiões. O NIR recebe este nome devido à sua proximidade do espectro visível.

As bandas de absorção na região NIR são produto de sobretons e da combinação de bandas fundamentais de vibrações de estiramento que têm lugar no MIR (infravermelho médio, *Medium InfraRed*). As ligações do tipo C-H, O-H e N-H, possuem comprimentos de onda de absorção característicos. Estas absorções não são muito intensas, e muitas vezes sobrepõem-se, tornando largas as bandas, e com baixa absortividade molar (baixa sensibilidade). Conseqüentemente, os espectros NIR não são

facilmente interpretáveis, havendo a necessidade da aplicação de métodos de calibração multivariada para extrair a informação contida neles.

Figura 2.2 O espectro eletromagnético

A espectroscopia no infravermelho próximo tem sido freqüentemente aplicada como um método analítico que fornece resultados eficientes para a determinação de moléculas orgânicas (proteínas e gorduras) e parâmetros qualitativos de produtos da agricultura e da indústria de alimentos. Aplicações mais recentes têm sido realizadas nas indústrias do petróleo, têxtil, carvão, cosméticos, polímeros, química, tintas e farmacêutica. No campo do petróleo, tem-se uma ampla gama de aplicações relacionadas a produtos destilados, tais como gasolina, diesel, nafta e querosene (Honigs *et al.*, 1985b; Kelly e Callis, 1990; Litani-Barzilai *et al.*, 1997; Chung *et al.*, 1999; Kim *et al.*, 2000; etc.).

Por outro lado, a instrumentação para a região NIR é similar à utilizada na espectroscopia de absorção no UV/visível, embora o desenvolvimento das aplicações seja totalmente diferente. De maneira geral, os espectrômetros NIR têm como características distintivas: faixa espectral, precisão do comprimento de onda e características de projeto (interferômetro ou grade). Além disto, têm-se a repetibilidade, a estabilidade, e a relação sinal-ruído como fatores chave para selecionar um espectrômetro. Atualmente, vários fornecedores deste tipo de equipamentos oferecem

aparelhos NIR condicionados para uso em laboratório e nos processos (Espinosa *et al.*, 1995).

Para um aprofundamento quanto à instrumentação e compreensão dos mecanismos dos diferentes tipos de espectrômetros existentes recomenda-se a leitura de Siesler *et al.* (2002, capítulo 3) e Burns e Ciurczak (2001, capítulo 2).

2.2.1 Justificativa de aplicação

O primeiro aspecto a ser levado em consideração quando tem que ser definida a técnica espectroscópica a ser empregada para analisar uma determinada substância é o conjunto de vantagens que ela oferece. No caso do NIR, graças à ampla diversidade de aplicações e à considerável quantidade de literatura existente reportando valiosas conclusões em experiências específicas, pode-se afirmar que as principais vantagens que fazem da espectroscopia NIR uma das técnicas recentes mais poderosas na análise química são:

- A energia da radiação NIR é suficientemente pequena e permite assim a análise não destrutiva das amostras. Isso é muito importante porque quanto mais tratamento a amostra sofrer, antes de obter uma resposta final, maior será a introdução de erros na análise.
- A radiação NIR é suficientemente intensa para permitir caminhos óticos maiores através da amostra sem a absorção completa da radiação.
- A espectroscopia NIR permite o uso de fibras óticas, facilitando a análise à distância já que o instrumento pode permanecer em um ambiente mais limpo e livre de vibração. Sendo assim, a engenharia do instrumento torna-se mais simples porque ele não precisa ser resistente à vibração ou ser a prova de explosão. Finalmente, diminui-se a exposição do operário a ambientes potencialmente tóxicos, como a planta de uma refinaria.
- A técnica NIR possui sistemas óticos facilmente disponíveis, detectores eficientes, simplicidade relativa na instrumentação, e o fato de quase todos os compostos orgânicos absorverem nesta região.

A aplicação da análise NIR na indústria do petróleo tem sido utilizada por muitos anos. Muitas das aplicações reportadas estão associadas à previsão da qualidade para produtos destilados médios, incluindo nafta, gasolina e diesel, cobrindo propriedades tais como composição de hidrocarbonetos, destilação, número de octanagem, pressão de vapor Reid e índice de cetano. Por outro lado, a facilidade e disponibilidade da técnica que a tornam particularmente adaptável à caracterização remota faz com que a sua utilidade e potencial econômico aumentem (Hidajat e Chong, 2000).

3 PESQUISAS DESENVOLVIDAS – ANTECEDENTES

3.1 Petróleo e Espectroscopia NIR

A seguir faz-se uma revisão das principais pesquisas desenvolvidas na aplicação da espectroscopia NIR no campo do petróleo, sendo quase todas elas relacionadas a produtos destilados. A exceção corresponde ao trabalho de Hidajat e Chong (2000) que utilizaram o NIR para caracterizar petróleo bruto, quanto às suas curvas TBP e de densidade; cuja experiência será relatada no final deste tópico. Alguns termos presentes neste capítulo referem-se às técnicas utilizadas para calibrar os modelos das diversas propriedades analisadas; e para a sua melhor compreensão, sugere-se a leitura das informações apresentadas no apêndice A.

Kelly e Callis (1990) avaliaram a utilidade do NIR para a estimação rápida e simultânea dos volumes percentuais das principais classes de hidrocarbonetos presentes em gasolinas (alifáticos, aromáticos e olefinas). Entre os métodos convencionais recomendados para tal fim têm-se os métodos por indicador de fluorescência, a espectroscopia de massa, HPLC, NMR e a cromatografia por fluido supercrítico.

Foram coletadas 17 amostras de gasolina, utilizando para a análise as técnicas MLR e PLS, sendo a técnica de pré-tratamento utilizada a derivada segunda em razão da melhora nos valores dos coeficientes de regressão obtidos. Os resultados ao se aplicar ambos métodos de calibração foram similares, sendo que para a PLS os modelos foram construídos com validação cruzada e menos variáveis latentes foram necessárias. Os bons resultados obtidos revelaram o NIR como uma técnica promissora na rápida estimação dos volumes percentuais de três classes de hidrocarbonetos presentes na gasolina, além da habilidade de ser adaptada para seu uso em planta e medições em linha.

Chung *et al.* (1999) compararam o desempenho do NIR e MIR na determinação das temperaturas de destilação do querosene a diferentes percentagens de recuperação, usando a técnica PLS. Foram coletadas 50 amostras de cinco unidades de destilação de cru, num período de três meses. Os espectros NIR foram coletados na faixa de 1100 a

2500 nm (9090 a 4000 cm^{-1}) e utilizaram-se fibras óticas sem problemas de condicionamento das amostras. Os espectros MIR foram coletados na região de 6234 a 16667 nm (3500 a 600 cm^{-1}), sendo a coleta feita a temperatura ambiente.

Os espectros MIR apresentaram diferenças mais claras e nítidas entre as amostras, mas também maior ruído e deslocamento de linha base. Essas características podem ser apreciadas na figura 3.1.

Figura 3.1 Espectros NIR e MIR do querosene (Chung *et al.*,1999)

Figura 3.2 Gráfico de paridade no NIR e MIR e as análises ASTM D86 para as temperaturas de recuperação (Chung *et al.*,1999)

A validação cruzada foi usada para determinar a complexidade dos modelos, e a partir disto, o NIR precisou de alguns fatores a mais que o MIR devido a suas bandas mais largas e menos características, mas os resultados da calibração foram melhores

quanto ao erro, assim como nos testes de reprodutibilidade, apresentando uma melhor relação sinal-ruído. Na figura 3.2, apresenta-se o gráfico de paridade entre as estimativas feitas pelo NIR e o MIR e os valores medidos respectivos. Ambos gráficos revelam os bons resultados obtidos, porém os dados previstos pelo NIR apresentam-se menos dispersos e mais próximos da reta unitária em comparação com os dados MIR. Isto é devido à melhor reprodutibilidade e menor ruído dos espectros NIR.

Kim *et al.* (2000) propuseram uma metodologia de classificação em tempo real para produtos de petróleo usando o NIR. O classificador proposto baseou-se na combinação da análise por componentes principais e um classificador Bayesiano. Na experiência, utilizaram-se dados NIR para seis produtos de petróleo: diesel, gasolina, querosene, gasóleo leve (LGO), *light straight-run* (LSR) e nafta. Foram utilizadas 327 amostras; das quais 273 amostras foram usadas para construir o modelo e 99 amostras foram usadas para testá-lo, utilizando o método PCA.

Figura 3.3 Espectros NIR dos seis produtos de petróleo (Kim *et al.*, 2000)

Na figura 3.3 apresentam-se os espectros para os seis produtos de petróleo, indicando as regiões espectrais que não foram consideradas na análise. As amostras foram coletadas diretamente de colunas de destilação e analisadas depois no laboratório. Quanto aos resultados, na etapa de calibração; das 273 amostras selecionadas, somente

uma não foi classificada de forma correta (correspondente a uma amostra de diesel). Da mesma forma, na etapa de teste do modelo uma amostra não foi classificada corretamente (correspondente a uma amostra de LGO). Como resultado final obteve-se uma classificação com uma precisão acima de 98.98% revelando o sucesso da metodologia. Na figura 3.4 apresenta-se o gráfico dos dois primeiros componentes principais.

Figura 3.4 Gráfico dos componentes principais PC1 x PC2 (Kim *et al.*, 2000)

Chung e Ku (2000) aplicaram com sucesso a espectroscopia NIR, em comparação com IR e Raman, na determinação da densidade API do resíduo atmosférico da destilação (AR). O espectro Raman apresentava-se dominado pela forte fluorescência devida aos asfaltenos, o que dificultou a coleta dos espectros. O IR fornecia as melhores características espectrais qualitativas em comparação com o NIR. Pelo contrário, o NIR apresentava características espectrais mais abrangentes com pouca resolução, mas fornecia a melhor relação sinal-ruído, comparado com o IR. Os espectros são apresentados na figura 3.5.

Foram coletadas 81 amostras, onde 51 espectros foram usados para calibração e 30 espectros para previsão. Como a amostra era sólida a temperatura ambiente, foi necessário aquecimento trabalhando-se a 60°C. A faixa NIR utilizada foi de 1100 a 2500 nm (9090 a 4000 cm^{-1}). A derivada segunda foi o método aplicado para reduzir o

deslocamento da linha base (ver figura 3.6), que acontece geralmente com amostras escuras. O PLS foi o método de calibração e a validação cruzada foi o método de validação.

Figura 3.5 Espectros típicos Raman, IR e NIR do AR (Chung e Ku,2000)

Figura 3.6 (a) Espectros brutos e (b) derivada segunda do espectro NIR do AR no conjunto de calibração (Chung e Ku,2000)

Os resultados obtidos demonstraram que analisar pequenas faixas espectrais (primeiro sobretom ou bandas de combinação, separadamente) diminui a qualidade dos resultados. A opção escolhida nesta experiência foi trabalhar numa faixa mais larga excluindo somente a região de linha de base variável que prejudica a qualidade da calibração. Por outro lado, o NIR fornece espectros altamente reprodutíveis assim como dados analíticos altamente repetíveis.

Hidajat e Chong (2000) estabeleceram uma correlação entre os perfis espectrais NIR de petróleos brutos com propriedades físicas e químicas de um conjunto representativo de amostras (dados TBP e dados de densidade), utilizando para tal fim a técnica de regressão PLS. Foram 110 amostras representando mais de 30 tipos de crus diferentes provenientes de aproximadamente 12 países, coletadas num período de 15 meses. Entre os países de origem das amostras tem-se o Iraque, Irã, Kuwait, Egito, Indonésia, Arábia Saudita, Vietnã, Malásia, Nigéria, Angola, etc. O espectro médio do conjunto de calibração está apresentado na figura 3.7.

Figura 3.7 Espectro NIR médio do conjunto de calibração (Hidajat e Chong, 2000)

As amostras foram seladas e armazenadas a temperatura abaixo dos 25°C antes da análise. A faixa NIR utilizada foi de 1600 – 2600 nm ($6250 - 3846 \text{ cm}^{-1}$), porém nos gráficos dos espectros constatou-se que a região 2100 – 2500 nm ($4760 - 4000 \text{ cm}^{-1}$)

apresentava maior variabilidade. Os dados de densidade e TBP foram determinados pela aplicação das normas ASTM D-1298 e D-2892, respectivamente. Foi utilizada a técnica MSC para o pré-tratamento dos dados e a validação cruzada para a determinação da complexidade dos modelos.

Foram desenvolvidos quatorze modelos para igual número de propriedades, relacionadas à densidade e às curvas TBP. Na figura 3.8 apresenta-se a comparação entre os valores de referência e os valores previstos para um deles.

Figura 3.8 Gráfico de paridade do modelo correspondente à fração de destilado leve (Hidajat e Chong, 2000)

A qualidade de previsão dos modelos NIR desenvolvidos responderam aos requerimentos dos métodos de referência ASTM. Muitos dos coeficientes de correlação foram próximos de 0.98, indicando a qualidade dos resultados. Fatores como a rapidez de medição dos espectros e possibilidade de análise remota valorizam ainda mais esta metodologia.

Espinosa *et al.* (1995) e Lambert *et al.* (1995) fazem uma revisão das aplicações do NIR para otimizar as unidades de processo da refinaria BP Lavera e a planta de etileno Naphta Chimie (França). Na BP Lavera o programa de pesquisa das aplicações

do NIR começou no final dos anos 70 e hoje se tem desenvolvido o *know-how* próprio em relação às técnicas quimiométricas e se tem aplicado o NIR ao *blending*, controle na destilação do petróleo e *planning/scheduling*. Uma das principais mudanças no trabalho diário da refinaria foi constatada na área de laboratório, onde o tempo nas análises RON e MON foi drasticamente reduzido além de permitir análises em linha. A consistência dos modelos NIR utilizados é avaliada mensalmente em relação aos métodos de referência ASTM respectivos.

Os resultados demonstraram que apesar das vantagens que o NIR oferece, tais como baixo capital de investimento, baixos custos de operação, alta repetibilidade, tempo de resposta curto e instalação flexível; deve-se tomar cuidado no desenvolvimento dos modelos das diferentes propriedades de interesse numa refinaria, principalmente quando se trata de pontos discrepantes no conjunto de calibração, os quais podem revelar características importantes do sistema ainda não representado adequadamente.

3.2 Petróleo e Análise SimDis

As experiências reportadas na literatura relacionadas à aplicação da SimDis no campo de petróleo revelam seu grande potencial e sucesso, sendo algumas das mais representativas apresentadas a seguir.

Sharma *et al.* (2003) utilizaram esta técnica como uma ferramenta complementar para monitorar as mudanças observadas na composição do resíduo de vácuo do petróleo a fim de prever o seu aproveitamento posterior. O objetivo foi aprimorar a caracterização do RAT a fim de compreender o mecanismo da separação para preparar de melhor maneira a carga para processos de conversão secundários. A metodologia desenvolvida demonstrou ter um grande potencial para a caracterização de frações solúveis de resíduos provenientes de uma ampla gama de petróleos, pela sua simplicidade, precisão e economia em termos de tempo investido.

Com a aparição da SimDis, técnicas complementares tornaram-se necessárias nas análises, sendo uma delas a espectrometria de massa. Neste sentido, Bacaud e Rouleau (1996) testaram esta união para caracterizar os produtos complexos obtidos na hidroconversão de um resíduo de vácuo desasfaltado sob diversas condições, sem prévio fracionamento. Esta experiência permitiu obter a distribuição das classes dos hidrocarbonetos presentes ao longo dos perfis de destilação dos produtos e em forma global, a caracterização analítica de produtos complexos da hidroconversão sem a necessidade de etapas prévias de separação.

Figura 3.9 Determinação das curvas TBP em % peso (A) e em % volume (B) de uma amostra de petróleo bruto através da destilação convencional e da metodologia proposta (Roussis e Fitzgerald, 2000)

Roussis e Fitzgerald (2000) desenvolveram um método rápido e confiável para a determinação das curvas TBP de petróleo bruto utilizando também o acoplamento entre

a SimDis e a espectrometria de massa. Ambos os perfis de destilação (em percentagem em peso e volume) foram previstos e boa concordância com os resultados obtidos através dos métodos de destilação convencional foi atingida. Na figura 3.9 são mostrados os resultados da previsão das curvas de destilação para uma amostra de petróleo.

A GCD/MS representa o acoplamento entre as técnicas de destilação por cromatografia gasosa e a espectrometria de massa. Na figura 3.10 é apresentada a comparação dos resultados através de gráficos de paridade para as amostras analisadas. A repetibilidade desta técnica resultou ser melhor que a do método ASTM D5307 como técnica convencional de destilação física.

Figura 3.10 Comparação dos dados TBP obtidos em % peso (A) e em % volume (B) para 11 amostras de petróleo (Roussis e Fitzgerald, 2000)

3.3 Potencialidade na análise do petróleo bruto

As experiências relatadas acima representam um grande conjunto de aplicações do NIR e as mais recentes aplicações da SimDis no campo do petróleo. Como pode ser visto, existe pouca literatura reportada quanto ao uso do NIR para analisar o petróleo em estado bruto. Para o desenvolvimento de uma das etapas deste trabalho, as pesquisas de Chung e Ku (2000) e Hidajat e Chong (2000) foram, desde o ponto de vista experimental, experiências relevantes por seus aportes na compreensão da análise de destilados pesados de petróleo e petróleo bruto, respectivamente. Algo similar constatou-se com as pesquisas relacionadas com a análise SimDis, o que é absolutamente compreensível sendo que foi concebida com o fim substituir as curvas TBP.

A revisão e análise das experiências relacionadas entre o NIR e o petróleo permitiram identificar os principais aspectos a considerar no desenvolvimento de uma metodologia para reproduzir as curvas de destilação, e assim como as pesquisas com a SimDis, o sucesso nos resultados alentaram o desenvolvimento das metodologias apresentadas nos próximos capítulos, visando sempre a obtenção mais rápida de respostas que conduzam também a uma decisão mais rápida para atuar no processo de forma preventiva. No processamento industrial, a melhora na produtividade analítica deve fornecer não somente uma redução no tempo de análise, mas também um aumento no conhecimento do processo. A seguir, uma breve análise dos principais aspectos identificados em cada pesquisa.

A experiência de Chung *et al.* (1999) é representativa do conjunto de trabalhos que demonstram a superioridade do NIR em relação ao MIR na análise de produtos de petróleo. Aqui, a característica relevante é a escolha da região do espectro para efetuar a análise, onde as faixas de 1650 – 1800 nm e 2100 – 2500 nm foram descartadas porque não continham informação espectral útil. Chung *et al.* (1999) apontaram à absorção forte e saturada da radiação NIR proveniente do caminho ótico relativamente longo (1 cm) como possível causa deste fato.

O trabalho de Kim *et al.* (2000) apresenta aplicações qualitativas do NIR. Neste trabalho afirma-se que o espectro NIR é diferente do espectro IR, porque permite detectar correlações complexas de misturas, como os produtos de petróleo, num tempo muito curto. Como em Chung e Ku (2000), nessa experiência faz-se também a seleção da faixa espectral baseando-se na informação espectroscópica conhecida a priori (sem especificar as razões), apresentando a região desconsiderada de forma visual na figura 3.3, onde se destaca a grande semelhança entre espectros de derivados do petróleo.

Com a experiência de Chung e Ku (2000) demonstrou-se a utilidade do NIR na análise do AR (Resíduo Atmosférico), que é o produto mais pesado do petróleo na destilação atmosférica. Entre as principais dificuldades dessa experiência pode-se citar a natureza física e química da amostra, sendo totalmente escura e muito viscosa (sólida a temperatura ambiente); e apresentando usualmente suspensões sólidas e partículas, o que afeta a reprodutibilidade das medições. Além disso, os espectros vibracionais de hidrocarbonetos pesados são menos sensíveis às mudanças estruturais.

No artigo, por tal motivo é recomendado que, a aplicação do NIR à análise de produtos pesados depende inicialmente de examinar a sensibilidade espectral, porque as bandas de absorção são inerentemente fracas e as variações espectrais baseadas em mudanças estruturais são menos sensíveis conforme o comprimento da cadeia cresce.

Graças aos testes de reprodutibilidade e repetibilidade feitos, foram demonstradas as vantagens do NIR em comparação com as outras técnicas vibracionais, onde apesar das bandas espectrais largas e menos características, ocorrem de fato variações espectrais em hidrocarbonetos pesados que permitem obter informação útil para a calibração quando é usada a regressão PLS.

Quanto à metodologia, nesta experiência testaram-se diferentes regiões do espectro para a calibração do modelo com um critério baseado na informação correspondente aos sobretoms e combinações das bandas C-H, obtendo a região espectral ótima através da comparação dos erros nas diversas regiões alternativas. Finalmente, deve-se ressaltar que pela natureza complexa do AR, a obtenção de

resultados úteis só foi possível com uma metodologia flexível para a determinação das condições experimentais apropriadas, como explicado acima.

A experiência de Hidajat e Chong (2000) foi a única reportada na literatura onde se utilizou o NIR para correlacionar propriedades do petróleo bruto (dados TBP e de densidade). Embora a metodologia não tenha sido detalhada, pode-se concluir que os bons resultados obtidos correspondem a uma metodologia relativamente simples quanto ao pré-tratamento de dados e calibração, sendo que a característica relevante para o sucesso da experiência foi a escolha da faixa espectral utilizada (2100 – 2500 nm). A razão está em que grupos funcionais tais como C-H metilênicos, olefínicos ou aromáticos causam várias vibrações de estiramento C-H, que são principalmente independentes da molécula. Como estas vibrações são as responsáveis pela maioria das bandas de absorção observadas na região de 2100-2500 nm, esta região dos espectros NIR foi selecionada para a modelagem computacional.

Embora os modelos obtidos por Hidajat e Chong (2000) só correlacionaram duas propriedades do petróleo, o trabalho representa um avanço na análise do petróleo bruto, antevendo a potencialidade do NIR para a estimação de um número maior de propriedades.

Por outro lado, as experiências com a SimDis servem para confirmar a precisão do método quanto à reprodução das curvas TBP. Porém, não se faz na literatura uma discussão abrangente das dificuldades que esta técnica enfrenta e que foram experimentadas no decorrer deste trabalho, quando as curvas SimDis foram utilizadas como as curvas de destilação de referência. Da mesma forma que nas pesquisas revisadas anteriormente, os bons resultados, como os de Roussis e Fitzgerald (2000) incentivaram o desenvolvimento deste projeto.

4 PREDIÇÃO DAS CURVAS TBP ATRAVÉS DO NIR

4.1 Considerações Gerais

4.1.1 Base de dados

A base de dados está formada por um conjunto de amostras de petróleo bruto armazenado em tanques, e coletadas antes de entrar na unidade de destilação de uma refinaria brasileira. Estiveram disponíveis dois conjuntos de amostras.

O primeiro conjunto foi formado por 46 amostras, coletadas no período Dezembro 2002 – Fevereiro 2003. Sem nenhuma informação de laboratório, serviu como um conjunto de treinamento na utilização do espectrômetro NIR e para a realização de análises de validação da técnica. As amostras foram armazenadas numa geladeira a $T=5^{\circ}\text{C}$ com a finalidade de reduzir as perdas de componentes altamente voláteis, o que alteraria a sua composição e, portanto seus espectros (o que de fato ocorre finalmente devido a um processo natural de degradação que foge às precauções de manipulação).

O segundo conjunto foi formado por 25 amostras, coletadas no período Fevereiro 2002 – Outubro 2002. A informação disponível correspondeu aos dados TBP e de densidade API. Este conjunto foi utilizado como a base para a construção dos modelos NIR. Este conjunto é formado por petróleos com densidades API na faixa de 24.3 – 29.3 (densidade média), e as amostras foram armazenadas em geladeira como para o primeiro conjunto.

4.1.2 Instrumentação

Os espectros NIR das amostras de petróleo foram coletados no espectrômetro NIR IFS 28/N (BRUKER, Alemanha), instalado no Edifício Semi-Industrial do Laboratório de Simulação e Controle de Processos (LSCP - USP). O sensor NIR utilizado foi o modelo 661.760-NIR com caminho ótico de 1mm

(HELLMA, Alemanha). O instrumento é controlado através de um PC usando o software espectroscópico OPUS - NT versão 3.0 (BRUKER, Alemanha) para a aquisição de dados.

Quanto aos componentes básicos deste espectrômetro tipo Transformada de Fourier (FT); têm-se uma lâmpada de tungstênio-halogênio como fonte de radiação NIR, um detector de diodo de Ge e um dispositivo de luz laser He-Ne. Segundo especificações, os limites de temperatura do instrumento abrangem a faixa de 18 – 30 °C (64 – 74 °F), e pode ser utilizado em ambiente industrial devido a sua robustez. Na figura 4.1 mostram-se o espectrômetro NIR e o sensor NIR respectivamente, utilizados nesta etapa do projeto.

Figura 4.1 Espectrômetro FT-NIR IFS 28/N (BRUKER) e sensores NIR de imersão (HELLMA)

4.1.3 Efeito da Temperatura

O efeito da temperatura nos espectros NIR depende do tipo de amostra (Baughman *et al.*, 1993), e é recomendável controlá-la antes de fazer as medições. No caso de amostras líquidas é comumente usado um banho de água para o controle da temperatura (Siesler *et al.*, 2002). De maneira geral, entende-se este efeito através da análise da relação densidade-temperatura para hidrocarbonetos; na qual um aumento da temperatura provoca uma diminuição da densidade e, portanto, uma diminuição das

moléculas por unidade de volume no feixe de luz NIR que vai ao detector (menor absorbância), e vice-versa.

Com o objetivo de estudar este efeito, foram feitos ensaios coletando-se espectros a diferentes temperaturas (correspondentes ao primeiro conjunto de amostras). Um banho de água foi utilizado para aumentar a temperatura progressivamente. Na figura 4.2 apresentam-se os espectros correspondentes a duas amostras.

Figura 4.2 Espectros NIR a diferentes temperaturas para duas amostras A e B

Segundo a figura 4.2, todos os espectros estão praticamente sobrepostos a temperaturas entre 25 – 50 °C. Porém, só analisando os espectros residuais, subtraindo um espectro do outro, é que se pode apreciar o efeito da alteração da densidade e, portanto, da temperatura. Foi assim que se calcularam os resíduos entre todos os espectros, sendo os mais representativos apresentados na figura 4.3.

Dos gráficos, pode-se afirmar que para o caso estudado, o efeito da temperatura depende da faixa espectral NIR utilizada na análise. Assim, embora a intensidade do efeito aumente com a diferença de temperatura, ele aparece aproximadamente na faixa de 4000 – 4500 cm^{-1} e 8500 – 13000 cm^{-1} , sendo a região de 4500 – 8000 cm^{-1} a zona do espectro onde praticamente a mudança de densidade das amostras de petróleo não altera os espectros.

Figura 4.3 Espectros NIR residuais para (a) $\Delta=10^\circ\text{C}$ e (b) $\Delta=20^\circ\text{C}$

Na medição dos espectros NIR, as temperaturas das amostras variaram entre 18 – 22 °C, o que faz deduzir que na região identificada acima o efeito da temperatura não será suficientemente importante para alterar os espectros e, portanto, a informação química contida neles. Em Siesler *et al.* (2002) cita-se a existência de uma técnica de compensação de variação da temperatura na calibração do modelo, quando o controle da mesma não é fácil; porém neste trabalho, segundo os resultados dos ensaios realizados, não foi necessária a aplicação deste tipo de técnicas.

4.2 Metodologia

A metodologia deste trabalho se desenvolve segundo dois aspectos chave: a medição de espectros NIR representativos do conjunto de amostras de petróleo (parte experimental), e a implementação das principais técnicas quimiométricas existentes para a construção dos modelos de calibração (parte computacional).

4.2.1 Experimental

A medida de espectros NIR foi feita utilizando as fibras óticas 041.202-NIR SMA/SMA de 6m de comprimento (HELLMA, Alemanha), apresentado-se o seu esquema na figura 4.4, sendo o volume necessário de amostra aproximadamente de 40 mL.

A metodologia experimental implementada considerou as recomendações feitas nos manuais de fabricante dos instrumentos, sendo os principais aspectos:

- Quanto às amostras, necessidade de agitação para favorecer a sua homogeneização, coletando duas réplicas para conseguir estabilidade nos espectros.
- Quanto ao sensor NIR, a limpeza é um fator fundamental, devendo-se selecionar um solvente que remova os resíduos da amostra depois da medição, e não seja prejudicial ao sensor. Sendo que este solvente não era conhecido, uma primeira opção foi o uso de um solvente polar como etanol ou metanol; mas no caso deste

trabalho, foi o *ciclohexano* o único solvente testado que cumpriu as recomendações dos fabricantes.

Figura 4.4 Esquema da medição dos espectros NIR de petróleo bruto

- Quanto ao espectrômetro, utilizou-se o ar para a medição do *background*. A água não é recomendada porque a forma das bandas do seu espectro tem alta dependência da temperatura. As medidas foram feitas em modo de transmitância (T) e os espectros apresentados em modo de absorbância ($\log(1/T)$), numa faixa espectral de $13000.12 - 3999.741 \text{ cm}^{-1}$ ($769.224 - 2500.162 \text{ nm}$), com 1.929 cm^{-1} de intervalo entre números de onda consecutivos.

4.2.2 Computacional

São necessárias quatro etapas na construção de um modelo de calibração: o pré-tratamento, a construção propriamente dita, a validação do modelo, e os testes de *outliers* ou pontos discrepantes. Para cada etapa existem diversas técnicas quimiométricas (apêndice A), e por tal motivo foram construídos na linguagem de programação MatLab R12 v6.0 os algoritmos respectivos.

Quanto à primeira etapa; foram construídos algoritmos para realizar derivadas, alisamento e MSC. O procedimento de Savitzky e Golay (1964) foi a base do algoritmo para o alisamento e a derivação (derivadas primeira e segunda), tendo entre suas características um ajuste polinomial de grau 3 por mínimos quadrados parciais e uma janela espectral de 35 pontos. Para o algoritmo do MSC foi utilizada a idéia original de Geladi *et al.* (1985) e revisada em Naes e Isaksson (1994b), baseada em mínimos quadrados.

Quanto à segunda etapa, dois algoritmos foram construídos, os correspondentes ao PCA-PCR e PLS, segundo o método NIPALS. Para o PCR, o trabalho de Martens e Naes (1989) foi a referência principal, enquanto que para a construção do algoritmo PLS ortogonal (aquele que fornece componentes principais ortogonais entre si) para uma variável dependente y , foram revisadas três referências: Martens e Naes (1989), Geladi *et al.* (1986) e Haaland e Thomas (1988).

Quanto à terceira etapa, foi implementada a técnica da validação cruzada, utilizando o algoritmo PLS para as calibrações e previsões sucessivas. É importante entender conceitualmente a curva obtida através deste tipo de validação. Uma curva típica é apresentada na figura 4.5. Como se pode observar, o erro de previsão é composto de duas contribuições; o erro sistemático devido a fenômenos não modelados nos dados espectrais, e o erro aleatório devido ao ruído proveniente de várias fontes (Martens e Naes, 1989). Estas duas contribuições têm tendências opostas conforme a complexidade do modelo aumenta; sendo assim, o erro sistemático diminui enquanto uma maior quantidade de fenômenos é modelada, mas ao mesmo tempo, o erro aleatório aumenta devido ao maior número de parâmetros independentes estimados a partir dos dados. Isto ocasiona a falta de ajuste e o sobreajuste do modelo, respectivamente.

De forma teórica, o mínimo erro de previsão que indica a complexidade adequada do modelo, é obtido quando ambos efeitos atingem o equilíbrio. A inspeção visual destas curvas é, portanto, uma importante ferramenta para selecionar a complexidade do modelo.

Figura 4.5 Ilustração teórica da validação de um modelo

Finalmente, utilizaram-se diversas técnicas para diagnosticar pontos discrepantes, sendo que não existe uma única forma para detectá-los. As ferramentas básicas de detecção estão baseadas nos *leverages* e resíduos. Assim, tendo construído os algoritmos para implementar as técnicas citadas, a metodologia utilizada nesta etapa do trabalho consistiu em:

- Aplicar uma técnica de pré-tratamento aos dados.
- Realizar a validação cruzada (*leave-one-out cross validation*) do conjunto total de amostras para 20 fatores ou componentes principais, permitindo assim visualizar o comportamento geral do sistema. Desta análise, determina-se a complexidade “inicial” do modelo.
- Calibrar o modelo utilizando o número de fatores indicado pela validação, fazendo os respectivos testes de pontos discrepantes para eliminar os considerados ruins.
- Revalidar o modelo para verificar a melhoria dos resultados no conjunto sem pontos discrepantes.
- Recalibrar o modelo no conjunto sem pontos discrepantes e verificar com os gráficos de paridade e a visualização dos erros de previsão a performance do modelo construído. Gráficos de *scores* e *loadings* são também úteis para avaliar o modelo.

4.3 Condicionamento espectral à análise TBP do petróleo

4.3.1 Reprodutibilidade

Para esta análise utilizaram-se quatro amostras do primeiro conjunto, foram coletadas 6 réplicas de cada amostra e os espectros NIR medidos. Calculou-se o desvio padrão dos espectros, assim como o espectro médio para calcular o desvio padrão relativo. Alguns resultados representativos em relação ao desvio padrão relativo em função do número de onda são apresentados na figura 4.6.

Figura 4.6 Resultados dos testes de reprodutibilidade para duas amostras analisadas

Observando os gráficos, os espectros medidos não são reproduzíveis nas faixas de $4000 - 4500 \text{ cm}^{-1}$ e $8500 - 13000 \text{ cm}^{-1}$ aproximadamente. Sendo assim, a região que pode ser considerada reproduzível é a que abrange de $4500 - 8000 \text{ cm}^{-1}$.

4.3.2 Faixa espectral

Na figura 4.7 apresentam-se os espectros do conjunto de 25 amostras analisadas na faixa original fornecida pelo espectrômetro. Embora o padrão espectral seja o mesmo, observa-se muita variação na zona a partir de 9000 cm^{-1} aproximadamente, onde a região $10000 - 11000 \text{ cm}^{-1}$ apresenta ruído excessivo, devido possivelmente ao uso de fibras óticas.

Figura 4.7 Espectros NIR das amostras de petróleo analisadas

Sendo a faixa total NIR uma região muito abrangente quanto à qualidade espectral, consideraram-se, baseados em experiências com sucesso reportadas na literatura, diversas faixas espectrais com o objetivo de dividir o espectro em sub-regiões e avaliar qual apresentava o melhor comportamento através das calibrações. Assim, na tabela 4.1 apresentam-se as sub-regiões NIR definidas com as suas principais características.

Tabela 4.1 Sub-regiões NIR definidas para a seleção da faixa adequada para análise

Faixa espectral	Região NIR	Característica	Referência	Amostra
A	2500 - 1650 nm 4000 - 6060 cm^{-1}	região de combinações linha base variável	Chung e Ku (2000)	AR
B	2500 - 2100 nm 4000 - 4762 cm^{-1}	região de combinações	Hidajat e Chong (2000)	Petróleo
C	1200 - 700 nm 8333 - 14286 cm^{-1}		Kelly e Callis (1990)	Gasolinas
D	1650 - 1100 nm 6060 - 9090 cm^{-1}	linha de base variable	Chung e Ku (2000)	AR
E	1900 - 1650 nm 5263 - 6060 cm^{-1}	1° sobreton	Chung e Ku (2000)	AR
F	2222 - 1250 nm 4500 - 8000 cm^{-1}	melhor reprodutibilidade efeito desprezível da T	-	Petróleo

As zonas espectrais são apresentadas individualmente na figura 4.8, onde visualmente observa-se que as regiões A, B e C apresentam maior ruído. Por outro lado, as regiões D, E e F apresentam consideráveis deslocamentos de linha base, sendo que a região D apresenta de forma aparente a menor quantidade de informação valiosa, e a região F a que praticamente abrange as regiões D e E. A zona F foi definida a partir dos ensaios de reprodutibilidade e efeito da temperatura descritos neste capítulo.

4.3.3 Pré-tratamento

Nesta etapa, utilizaram-se duas técnicas: alisamento e MSC. Todas as regiões NIR foram submetidas ao alisamento e MSC separadamente. Da definição do MSC, tem-se que o seu bom condicionamento relaciona-se ao fato do gráfico dos valores espectrais em função dos valores espectrais médios ser aproximadamente uma linha reta, de modo que diferentes amostras apresentam diferentes inclinações. As diferentes linhas de regressão são interpretadas como produto da dispersão e os desvios das linhas, a informação química contida nos espectros (Naes e Isaksson, 1994b). Segundo isto, as regiões D e E são as mais apropriadas para aplicar esta técnica, segundo os gráficos obtidos e que não serão apresentados aqui.

Figura 4.8 Regiões espectrais NIR definidas

Finalmente, foi testado o efeito combinado do Alisamento – MSC devido às vantagens que fornece ao sistema, já que o primeiro diminui o ruído e o segundo

aumenta a linearidade do sistema corrigindo efeitos da dispersão da luz, principalmente. Assim, na figura 4.9 apresentam-se os resultados para cinco das zonas espectrais após a aplicação destas técnicas (pode-se considerar a região F como a união das regiões D e E). Observa-se que a faixa C apresenta o pior desempenho em contraste com os espectros D e E que apresentam o melhor.

Figura 4.9 Efeito combinado do Alisamento e MSC nos espectros NIR

4.4 Modelos NIR das curvas TBP

4.4.1 Técnica de calibração

As técnicas lineares PCR e PLS foram testadas na calibração para alguns modelos. Segundo a teoria, a PLS tem a vantagem de considerar a informação em Y para construir os modelos (apêndice A), o que permite representar melhor a informação contida nos espectros NIR em relação às propriedades analisadas (dados TBP).

Caso A – PCR

Caso A – PLS

Figura 4.10 Gráficos de paridade para os modelos à 200°C (Caso A) e à 400°C (Caso B) utilizando PCR e PLS, respectivamente.

Caso B – PCR

Caso B – PLS

Figura 4.10 (cont.) Gráficos de paridade para os modelos à 200°C (Caso A) e à 400°C (Caso B) utilizando PCR e PLS, respectivamente

Na figura 4.10 são apresentados os gráficos de paridade correspondentes aos modelos PCR e PLS para as percentagens de volume destilado acumulado a 200°C e 400°C, utilizando o número de fatores recomendado pela validação cruzada, observando-se que embora o desempenho dos modelos PCR não seja ruim, os modelos PLS apresentam melhor correlação entre os valores medidos e previstos.

Resultados similares foram obtidos quando calibrados os modelos para as demais temperaturas analisadas, determinando-se assim que a técnica PLS seria utilizada para a construção dos modelos NIR nesta etapa do trabalho, considerando também que as experiências reportadas na literatura apontam a PLS como a técnica linear mais recomendada para a calibração em sistemas nos quais informação química quanto à estrutura é insuficiente, como no caso do petróleo bruto.

4.4.2 Resultados

As propriedades das amostras de petróleo bruto utilizadas para a modelagem foram as curvas TBP. A faixa da temperatura das curvas foi 28 – 750 °C.

Segundo o item 4.2.2 o primeiro passo deve ser o pré-tratamento espectral, utilizando-se como mostrado no item 4.3.3 a combinação do alisamento e MSC. Após o pré-tratamento, foi feita a validação inicial para algumas temperaturas com o objetivo de determinar qual a região espectral (das definidas no item 4.3.2) que melhor representa a informação contida nos espectros NIR.

A faixa espectral E foi considerada a zona ótima para construir os modelos NIR das curvas TBP já que apresentou o melhor comportamento diante das técnicas de tratamento testadas segundo a literatura, assim também abrange uma região que segundo as análises feitas apresenta reprodutibilidade boa e onde o efeito da temperatura nos espectros é desprezível.

É importante indicar que segundo a tabela 4.1, a região F é uma região que abrange tanto a zona do primeiro sobretom do espectro NIR (E) e parte da zona que apresenta deslocamentos maiores de linha base (D). Sendo que a informação contida num espectro pode-se repetir em várias zonas, teoricamente a informação contida na região das combinações é a mesma que a contida nas regiões do primeiro ou segundo sobretoms. Portanto, trabalhou-se com a faixa E, uma zona que além de abranger a região do primeiro sobretom exclusivamente, apresenta picos definidos com boa qualidade espectral.

Após a remoção de pontos discrepantes que afetavam negativamente a calibração, segundo os métodos recomendados por Martens e Naes (1989) e Kowalski e Seasholtz (1991) (ver apêndice A), a seguinte etapa no processo foi a calibração final dos modelos. De acordo com os dados de laboratório, as curvas TBP são compostas de 23 pontos correspondentes a temperaturas dadas e respectivas percentagens de volumes acumulados. Porém, com a metodologia desenvolvida neste trabalho conseguiram-se bons resultados para reproduzir 17 pontos, ou seja, 17 modelos de calibração foram desenvolvidos com sucesso.

Os modelos NIR construídos abrangem a faixa de temperatura de 75–550 °C. Os modelos nas regiões extremas 28–50 °C e 600–750 °C não forneceram boas previsões.

Procurando explicar estes resultados pode-se afirmar o seguinte: quanto à região de baixas temperaturas, é muito provável que a perda de voláteis pelo período prolongado de armazenamento das amostras tenha alterado a composição química do petróleo e portanto alterado a informação química contida nos espectros. Quanto à região de altas temperaturas, experimentalmente estes pontos são obtidos sob vácuo e reconvertidos para a pressão atmosférica. Na norma ASTM D-2892 que trata da curva de TBP, afirma-se que há uma menor reprodutibilidade para esta região. Por outro lado, pode haver algum erro sistemático no espectro devido à formação de partículas sólidas a partir dos hidrocarbonetos pesados, que são justamente a quem corresponde a região de altas temperaturas.

Na faixa de temperaturas na qual os modelos foram construídos, observou-se também uma performance variável, sendo que os melhores desempenhos foram obtidos na faixa 175 – 450 °C. Alguns gráficos de paridade (% volumes medidos versus % volumes previstos) são apresentados na figura 4.11, junto com os coeficientes de correlação R^2 respectivos.

A melhor qualidade das curvas de validação verificou-se após a eliminação de pontos discrepantes (duas amostras foram consideradas pontos ruins e

foram eliminadas), assim como uma modelagem adequada através da visualização de gráficos de parâmetros utilizados na técnica PLS, tais como os *loadings* (ver apêndice A), resultados que não são apresentados aqui meramente por questões de relevância em relação aos apresentados a seguir.

Finalmente, ilustra-se o desempenho dos modelos de forma quantitativa e qualitativa. Os resultados finais das calibrações PLS são apresentados na tabela 4.2, onde se observa que os erros são menores a 1.3 % de volume acumulado para as temperaturas consideradas, indicando a boa performance dos modelos NIR para reproduzir as curvas de destilação TBP, com coeficientes de correlação R^2 variando entre 0.80 e 0.97.

Por outro lado, na figura 4.12 apresenta-se a comparação das curvas TBP obtidas no laboratório e as obtidas utilizando espectroscopia NIR para três das amostras analisadas, mostrando o sucesso da sua aplicação.

Figura 4.11 Gráficos de paridade para alguns dos modelos NIR das curvas TBP do petróleo bruto analisado

$T=225^{\circ}\text{C}$, $R^2=0.9579$

$T=275^{\circ}\text{C}$, $R^2=0.9528$

Figura 4.11 (cont.) Gráficos de paridade para alguns dos modelos NIR das curvas TBP do petróleo bruto analisado

$T=350^{\circ}\text{C}$, $R^2=0.9555$

Figura 4.11 (cont.) Gráficos de paridade para alguns dos modelos NIR das curvas TBP do petróleo bruto analisado

Tabela 4.2 Resultados da calibração PLS dos modelos NIR

T (°C)	Nº fatores	5263 - 6060 cm^{-1}		T (°C)	Nº fatores	5263 - 6060 cm^{-1}	
		RMSECV (%vol acum)	R^2			RMSECV (%vol acum)	R^2
75	9	0.88	0.918	300	6	0.90	0.965
95	6	0.69	0.882	325	6	0.89	0.964
125	5	0.66	0.907	350	6	0.95	0.956
149	5	0.68	0.938	400	6	1.01	0.942
175	5	0.61	0.957	425	4	0.95	0.904
200	5	0.64	0.963	450	5	1.28	0.897
225	5	0.80	0.957	500	6	1.22	0.840
250	5	0.84	0.957	550	4	0.93	0.790
275	5	0.91	0.953

Figura 4.12 Comparação entre as curvas TBP obtidas: ASTM (curva medida ou de referência) e NIR (curva prevista ou estimada)

Figura 4.12 (cont.) Comparação entre as curvas TBP obtidas: ASTM (curva medida ou de referência) e NIR (curva prevista ou estimada)

Da figura 4.12 observa-se, na porção da curva que foi possível ser reproduzida, uma concordância muito boa entre os dados medidos e previstos, e isto foi observado também no restante do conjunto de calibração. Porém, é importante realçar a necessidade de um conjunto de calibração mais robusto em termos de abrangência de características composicionais visando ampliar a aplicabilidade dos modelos desenvolvidos. De maneira global, com uma metodologia como a desenvolvida neste trabalho demonstra-se que é possível a estimação rápida e eficiente de curvas TBP de amostras de petróleo bruto processado em refinarias.

Nas regiões extremas os modelos não forneceram bons resultados por uma falta de melhor condicionamento das técnicas quimiométricas disponíveis, razão pela qual outras ferramentas foram aplicadas e serão apresentadas nos próximos capítulos.

5 PREDIÇÃO DAS CURVAS SIMDIS ATRAVÉS DO NIR

5.1 Considerações Gerais

No capítulo anterior foi descrita a metodologia desenvolvida para reproduzir as curvas TBP de petróleo bruto processado em uma refinaria. Como já foi apontado, na indústria do petróleo as limitações de tempo e custo da análise TBP incentivaram o progresso da análise SimDis como método alternativo.

Neste capítulo é feita a previsão das curvas SimDis a partir dos espectros NIR de petróleo. Redes neurais são aplicadas para estabelecer a relação NIR versus SimDis do petróleo bruto.

5.1.1 Base de dados

A base de dados é formada por um conjunto de 40 amostras. Elas procedem dos tanques de armazenamento de uma refinaria brasileira. As densidades API variam na faixa de 31.1 – 36.4. As amostras de petróleo são coletadas antes do processo de dessalgação e armazenadas em garrafas de vidro escuro, seladas e identificadas.

As amostras são armazenadas a 5°C em geladeira para reduzir a perda de componentes leves e o processo de degradação natural.

5.1.2 Instrumentação

A medição dos espectros NIR seguiu a mesma metodologia descrita no capítulo 4.

As análises SimDis foram realizadas no cromatógrafo gasoso SHIMADZU GC-17A, acondicionado com um sistema de vaporização com temperatura programada (PTV), detector de ionização de chama (FID) e um sistema criogênico de forno da

coluna. Pela alta sensibilidade das colunas, um filtro de oxigênio (Modelo 1000 *Oxygen Trap* – CRS) foi utilizado na linha de gás de arraste (Hélio). O sinal do detector é registrado como áreas (intervalos de tempo) para tempos de retenção continuamente crescentes.

O software GC Solution (SHIMADZU) foi utilizado para a aquisição de dados, e o cálculo das curvas SimDis foi realizado com um programa desenvolvido no LSCP na linguagem de programação MatLab v.6.0.

5.2 Metodologia

5.2.1 Preparação das amostras e análise cromatográfica

A determinação das curvas SimDis foi feita baseada na norma ASTM D-5307. Uma metodologia própria para analisar as amostras coletadas na refinaria foi utilizada para realizar as análises.

As amostras foram diluídas em disulfeto de carbono (*Merck*, grau espectroscópico) de 2% peso aproximadamente. A alta pureza do CS₂ permitiu o seu uso como solvente pela sua miscibilidade com resíduos, baixo ponto de ebulição e baixo fator de resposta no FID. Os volumes de injeção de amostra foram de 1,0 µL. Um padrão interno composto das *n*-parafinas *n*-C₁₄, *n*-C₁₅, *n*-C₁₆ e *n*-C₁₇ foi utilizado, em uma concentração de 10% peso da amostra aproximadamente.

Foi necessária também uma análise de branco devido à ocorrência de instabilidade na linha de base. As condições extremas de temperatura (-30 até 430°C) fazem a coluna estar sujeita a sangramento (a análise do branco deve ser feita periodicamente porque a linha de base varia levemente de uma amostra em relação a outra). Na tabela 5.1 apresentam-se as condições de operação do cromatógrafo para a análise SimDis.

Tabela 5.1 Condições de operação para o cromatógrafo gasoso na análise SimDis

Gás de arraste	
Nome	Helio
Vazão (mL/min)	18
Coluna	
Temperatura inicial (°C)	-30
Temperatura final (°C)	430
Taxa de programação (°C/min)	10
Detector	
Tipo	FID
Temperatura (°C)	440
Injetor	
Tipo	PTV
Temperatura inicial (°C)	60
Temperatura final (°C)	430
Amostra	
Tamanho (µL)	1
Concentração mássica (%)	2

5.2.2 Reprodutibilidade espectral

Para avaliar a reprodutibilidade espectral, seis espectros de uma amostra foram coletados em diferentes intervalos de tempo para um conjunto delas. O desvio padrão relativo em relação ao espectro médio do conjunto selecionado para esta análise é apresentado na figura 5.1. Observa-se que os valores obtidos, correspondentes à região espectral E varia entre 0.78 – 3.50 %. A figura permite concluir que o conjunto encontra-se bem distribuído, e variações comuns tanto externas (condições ambientais) quanto internas (dispersão, etc.) não afetam significativamente a representação espectral das amostras de petróleo bruto.

Figura 5.1 Desvio padrão relativo na análise de reprodutibilidade

5.2.3 Pré-tratamento espectral

Os espectros NIR das amostras de petróleo apresentaram-se completamente similares entre si, porém as curvas SimDis medidas em laboratório mostraram diferenças significativas em muitos casos. Portanto, o pré-tratamento foi necessário para tornar a informação contida nos espectros mais sensível às diferenças.

Um filtro de média móvel de 5 termos foi aplicado aos dados antes do cálculo da derivada primeira. Como será visto no apêndice A, a derivação é uma técnica que permite maior discriminação espectral para intensificar diferenças entre espectros idênticos, e matematicamente temos que a amplitude da n -ésima derivada de um pico é inversamente proporcional à n -ésima potência de sua largura, o que significa que conforme aumenta o grau da derivada a amplitude delas diminui. Assim, com o cálculo da derivada primeira já é possível discriminar os espectros no sentido de revelar picos mascarados nos espectros originais.

Finalmente, o espectro correspondente a uma dada amostra selecionada randomicamente foi subtraído do resto dos espectros para “centralizar” os dados. A figura 5.2 apresenta alguns espectros NIR numa faixa localizada dentro dos limites da região E definida na calibração dos modelos NIR das curvas TBP (ver item 4.3.3), a sua derivada primeira e os espectros “centralizados” como resultado final do pré-tratamento.

Figura 5.2 Espectros NIR de algumas amostras (a) sem pré-tratamento (b) derivada primeira e (c) após “centralização”

(b)

(c)

Figura 5.2 (cont.) Espectros NIR de algumas amostras (a) sem pré-tratamento (b) derivada primeira e (c) após “centralização”

Como pode ser observado, as diferenças nos espectros NIR originais são apreciáveis na derivada primeira dos espectros residuais, o que demonstra a importância deste tipo de tratamento, assim como foi descrito no item 4.1.3 para avaliar o efeito da temperatura nos espectros NIR. Foi observada também uma correspondência entre o grau de similaridade nas curvas SimDis e os respectivos espectros, fato que confirma a representação da informação química relevante contida neles neste tipo de análise, e que é demonstrado nas figuras 5.3 e 5.4. Observa-se que as duas curvas SimDis mais próximas entre si (com o mesmo grau API) apresentam de igual forma os espectros NIR mais próximos.

Figura 5.3 Curvas SimDis de amostras de petróleo com o mesmo grau API (33.4)

Figura 5.4 Espectros NIR pré-tratados de amostras de petróleo com o mesmo grau API (33.4)

5.2.4 Técnica de calibração

Nesta etapa do projeto avaliou-se a aplicabilidade de técnicas não lineares na correlação dos espectros NIR com propriedades físico-químicas do petróleo, sendo que na etapa anterior foram utilizadas técnicas lineares (principalmente a PLS). Neste sentido, uma especial atenção foi colocada nas Redes Neurais (ou NN, *Neural Network*), que têm atraído grande interesse como modelos preditivos e têm se revelado ferramentas poderosas para correlacionar dados não lineares. Elas possuem a habilidade de “aprendizado” do que acontece no processo sem a necessidade de modelá-lo seguindo as leis físicas e químicas que o governam. Um aprofundamento da conceituação desta técnica pode ser encontrado no apêndice A deste trabalho.

Figura 5.5 Representação esquemática do processo da previsão das curvas SimDis de petróleo bruto através das redes neurais

A rede neural utilizada neste trabalho foi uma rede *feed-forward* de três camadas (de entrada, oculta e de saída). O modelo desenvolvido prevê a curva SimDis para uma amostra de petróleo bruto dada. As variáveis de entrada na rede são os espectros de absorção (representados por 20 números de onda, na faixa de 5600 – 6000 cm^{-1}) e o grau API de cada amostra. As variáveis de saída são as percentagens em peso às mesmas temperaturas que no caso dos modelos NIR para as curvas TBP. Todas as variáveis foram normalizadas na faixa de 0.1 – 0.9. A figura 5.5 mostra uma representação esquemática da rede neural utilizada.

5.3 Modelos NIR das curvas SimDis – Resultados

O conjunto de dados coletados foi randomicamente dividido em dois grupos, um conjunto de calibração (80% das amostras) e um conjunto de teste (20% das amostras) utilizado para a validação externa da rede neural, que contém seis neurônios na camada oculta. Os resultados mostraram uma boa concordância com os dados experimentais. Na

figura 5.6 são apresentados os resultados correspondentes a duas amostras com diferentes graus API.

(a)

(b)

Figura 5.6 Comparação entre as curvas SimDis medidas (através de normas ASTM) e previstas (através do NIR) para amostras com grau API (a) 31.1 e (b) 36.4

É importante realçar que, na previsão de modelos não fenomenológicos, como as redes neurais deste trabalho, nenhuma restrição física que imponha que a percentagem em peso seja uma função monotónica crescente nas curvas SimDis foi utilizada, favorecendo a presença de erros randômicos e independentes entre as previsões.

Dois parâmetros de erro foram calculados visando avaliar a qualidade dos modelos. O primeiro, o RMSEP (*Root Mean Squared Error Prediction*), variou na faixa de 0.66 – 3.5% em percentagem em peso vaporizado, indicando uma pequena diferença em comparação aos valores medidos. O segundo, a estatística R^2 , foi calculado para o conjunto total de dados, obtendo-se um valor de 0.9931. Assim, estes resultados mostram um grande potencial para a aplicação da combinação entre o NIR e as redes neurais como uma metodologia de estimação das propriedades SimDis de petróleo bruto.

6 CASO DE ESTUDO: PROCESSO DE DESTILAÇÃO ATMOSFÉRICA

A última etapa do projeto envolve a adequação da informação de caracterização do petróleo obtida nas metodologias apresentadas nos capítulos anteriores na simulação do processo de destilação, visando a melhor representatividade dos seus resultados em relação aos dados medidos em planta, e ganhos em tempo, custo e precisão. Com esse propósito, um processo real foi analisado e os resultados são apresentados neste capítulo.

6.1 Justificativa

Aumentar a qualidade da simulação de processos é um dos maiores desafios da indústria do petróleo. Uma das informações básicas para a simulação do processo de destilação atmosférica é a curva de destilação do petróleo (TBP). Como dito anteriormente estas informações são difíceis de se obter em tempos compatíveis com as mudanças das cargas de petróleo nas refinarias brasileiras, além de serem análises com alto custo. Uma das alternativas é o emprego das destilações simuladas (SimDis) para a previsão das TBPs. Já na literatura, são reportadas experiências em relação à adequação das TBPs.

Basak *et al.* (2002) mostraram a importância de adequar a curva TBP da alimentação ou carga com a operação. Eles consideraram as especificações dos produtos destilados (as quais estão fortemente relacionadas com a precisão da curva TBP utilizada) como restrições em um problema de otimização. Desta forma propuseram duas formas de ajustar a curva TBP da carga.

A primeira; através da reconciliação, atualizando-se a curva em intervalos de tempo definidos. Esta metodologia está baseada na determinação dos dados TBP dos produtos a partir dos seus dados de destilação ASTM (devido a que os dados TBP são raramente medidos em laboratório). Neste esquema, reconcilia-se a curva TBP da carga com o objetivo de minimizar a diferença entre os dados TBP dos produtos obtidos na

planta e os determinados pelo simulador. Esta metodologia somente pode ser utilizada quando as análises dos produtos estiverem disponíveis. A segunda; o retro-cálculo ou “*back-calculation*” da TBP da carga, onde a curva TBP inicial é retificada em tempo real utilizando somente dados de operação, e não mais as curvas de destilação dos produtos. Os autores afirmam que este procedimento revelou-se mais eficiente que o primeiro.

Em outro trabalho, Chatterjee e Saraf (2004) utilizaram uma curva TBP e medidas de planta para estimar as propriedades dos produtos. A metodologia consiste na determinação das temperaturas EFV (*Equilibrium Flash Vaporization*). Linhas retas aproximam as curvas dos produtos. Basak *et al.* (2002) testaram a metodologia e consideraram os resultados satisfatórios. Porém, como reforçado pelos autores, muito ainda precisa ser feito para introduzir a natureza do cru no problema da simulação.

Nesta etapa do trabalho, o objetivo foi a implementação de uma metodologia baseada na caracterização do petróleo via espectroscopia NIR (que se relaciona com a SimDis e TBP) na simulação.

6.2 Descrição do Processo

O processo de destilação é o processo de separação primário de qualquer refinaria de petróleo, já que é através dele que o petróleo bruto pode ser desdobrado em frações que, dependendo das suas características e do tipo de produtos que a indústria precisa (combustíveis ou matérias primas petroquímicas, etc.) representam os produtos destilados que são alimentados aos outros processos existentes: de conversão, de tratamento, entre outros.

No caso deste trabalho, o processo de destilação atmosférica da Refinaria de Capuava (RECAP – Petrobras S.A.) cujo fluxograma está representado nas figuras 6.1 – 6.3, foi o caso estudado. Esta refinaria caracteriza-se por processar de forma geral petróleo de densidade média ou elevada, com uma capacidade instalada de 53 mil barris por dia, e tendo entre seus principais produtos GLP, gasolina, óleo diesel metropolitano

Figura 6.1 Processo de destilação atmosférica da refinaria (Dessalgação)

Figura 6.2 Processo de destilação atmosférica da refinaria (Torre pré-flash)

Figura 6.3 Processo de destilação atmosférica da refinaria (Torre atmosférica)

(com baixo teor de enxofre) e solventes especiais. A parte do processo a ser analisada consta de três fases:

A primeira, o Pré-aquecimento e Dessalinização (ou Dessalgação); onde o petróleo consegue atingir a temperatura máxima por troca térmica com as correntes quentes que deixam o processo, e fica livre de sais, água e partículas sólidas suspensas, respectivamente. A segunda, a Torre de Pré-flash; onde os cortes mais leves são retirados (GLP e nafta leve) com o objetivo de ampliar a carga total da unidade, obtendo-se também diesel leve. E a terceira, onde o petróleo pré-vaporizado passa aos fornos para atingir condições ideais de fracionamento (evitando o craqueamento térmico) e é introduzido na Torre Atmosférica, onde são retirados vapores de nafta (ou nafta pesada), assim como querosene e diesel pesado. O resíduo da destilação atmosférica é transformado em produtos mais nobres na unidade de Craqueamento Catalítico.

O esquema do processo apresentado nas figuras anteriores, visando a sua simulação em um software comercial, é apresentado na figura 6.4. Cabe realçar que a corrente LCO representa uma corrente de RAT craqueado (grau de densidade API elevado) que serve como selagem nas bombas de fundo das torres.

Na simulação do processo os resultados obtidos no capítulo 4 (modelos para a previsão de curvas TBP a partir do NIR) não puderam ser utilizados diante da impossibilidade de construir um conjunto de calibração que permita o cruzamento das informações necessárias: espectros NIR, curvas TBP obtidas em laboratório e variáveis do processo, num mesmo período de tempo. Desta forma, as curvas SimDis obtidas no LSCP foram os dados de caracterização utilizados na simulação (com o intuito da aplicação posterior dos resultados obtidos no capítulo 5), sendo que as curvas correspondem a amostras coletadas simultaneamente ao período de coleta das variáveis do processo (numa faixa de grau API de 31.1 – 36.4).

6.3 Base de dados industriais

Confrontando as informações vindas de planta (fluxogramas) e do painel de controle do processo foram determinadas as variáveis a serem utilizadas para o seu monitoramento. Assim, os valores de campo (ou medidos) coletados correspondem às variáveis identificadas nas figuras 6.1 – 6.3 através de “Tags” ou identificadores. A coleta de dados corresponde ao período entre Outubro 2004 e Dezembro 2004, com um intervalo de 10 minutos entre leituras consecutivas.

(a)

Figura 6.4 Esquema do processo utilizado na simulação (a) Torre de Pré-Flash e (b) Torre Atmosférica

Figura 6.4 (cont.) Esquema do processo utilizado na simulação (a) Torre de Pré-Flash e
(b) Torre Atmosférica

Para uma fácil compreensão das variáveis envolvidas no processo é apresentada a tabela 6.1. Foram 86 variáveis coletadas no total que serviram como base para analisar o processo, em termos de estabilidade e disponibilidade de outras informações. Não foi possível coletar dados mais recentes por causa de parada de operação da planta no início do ano 2005 e porque outras informações necessárias (curvas de qualidade dos produtos, etc.) não foram medidas posteriormente na refinaria naquele período.

Tabela 6.1 Identificadores dos dados coletados na RECAP

Etapa	Tipo de dado	Tag	Descrição
Dessalinização	Fluxo	FC_50143	Slop Reprocessado U-500
	Temperatura	TI_50304	Saída M-536 Casco
	Pressão	PC_50002	Saída O-502
	Nível	LC_50002	Interface Óleo/Água O-527
		LC_50004	Interface Óleo/Água O-502
Torre de Pré-Flash	Fluxo	FC_50002	Carga de Petróleo U-500
		FC_50016	Refluxo N-507
		FC_50157	Nafta N-507 para O-601 Tanque
		FC_50124	Vapor Para N-507
		FC_50180	Diesel Extra Leve N-507
		FI_50200	GLP do O-521
		FI_50018	Produção de Nafta Leve N-507
		FC_50155	Nafta N-507 para solventes
		FC_50060	Gás Combustível O-530
		FI_50032	Cru Reduzido para L-506
	Temperatura	FC_50015	Saída Gases O-521
		TC_50004	Temperatura Topo N-507
		TI_50435	Saída Panela da N-507
		TI_50106	Saída de Fundo N-507
		TI_50175	Recheio Topo N-507
		TI_50176	Recheio Fundo N-507
		TI_50002	M-517C para N-507
		TI_50108	Entrada do O-521
		TI_50170	Recheio topo N-507
		TI_50174	Recheio Fundo N-507
		TI_50171	Acima Recheio Interm. N-507
		TI_50105	N-507 Zona de Flash
	Pressão	PC_50005A	O-521 para O-503
		PI_50139	Sucção J-546A
		PI_50114	Abaixo Recheio Topo N-507
		PDI_50113	Recheio Topo N-507
		PI_50224	N-507 Zona de Flash
		PDI_50120	Recheio Interm. N-507
		PI_50118	Abaixo Recheio Fundo N-507
		PDI_50117	Recheio Fundo N-507
		PI_50004	Topo N-507
		PC_50005B	Alívio O-521 para Flare
	Nível	PI_50137	M-507A
		LC_50006A	Fundo N-507 (L-506)
		LC_50008	Nafta O-521
		LC_50010	Interface bota O-521
Fornos	Fluxo	FC_50131	Total passes L-506
	Temperatura	TI_50335	Carga para Passes L-506
		TC_50068	Saída L-506

Tabela 6.1 (cont.) Identificadores dos dados coletados na RECAP

Etapa	Tipo de dado	Tag	Descrição
Torre Atmosférica	Fluxo	FC_50020	Nafta Pesada (N-506)
		FC_50041	Refluxo Diesel N-506
		FC_50042	Vapor Vivo N-506
		FC_50022	Refluxo de Topo N-506
		FC_50038	Refluxo Interm. N-506
		FC_50034	Vapor para N-505
		FC_50036	Produção Querosene N-505
		FC_50040	Retirada Diesel N-506
		FI_50129	Agua Bota O-528
		FI_50052	Gas. Comb. O-620 para O-528
		FI_50202	Gás do O-530
		FC_57040	Torch oil para K-573
		FC_50203	Carga para M-505A
		FI_50203	Descarga da J-587 para M-505 A
	Temperatura	TC_50010	Topo N-506
		TI_50020	Fundo N-506
		TI_50401	Leito Inferior da N-506
		TI_50402	Abaixo Panela de Diesel
		TI_50018A	Recheio Diesel N-506(Sup)
		TI_50018B	Recheio Diesel N-506(Inf)
		TI_50021	Diesel N-506 para J-547A
		TI_50022	Fundo N-505
		TI_50092	Saída M-534 (tub) - RCI N-506
		TI_50346	Saída M-540 tubos
		TI_50673	RAT para M-505 A
		TC_50018	Recheio de Diesel N-506
		TI_50345	Entrada do M-539 casco
		TI_50319	saída do M-539 casco
		TI_50019	Retirada Querosene N-506
		TI_50095	RCI N-506 sucção J-548 AB
		TI_50015	Saída Nafta M-532 ABCD
		TI_50016	Saída Nafta M-533 ABCD
	Pressão	PC_50006	O-528
		PI_50048	Topo N-506
		PI_50049	Fundo N-506
		PI_50414	Gás do O-501
		PI_50410	Descarga da J-587
	Nível	LC_50012	Nafta O-528
		LC_50014	Interface bota O-528
		LC_50018	Fundo N-506
		LC_50016A	Fundo N-505
		LC_50016B	Fundo N-505

6.4 Simulador do processo

Na refinaria, utiliza-se o software Petrox (simulador desenvolvido pela Petrobras S.A.) para a simulação do processo. As curvas de destilação TBP dos petróleos mais comumente processados (informação de entrada no simulador) são disponibilizadas pelo CENPES (Petrobras S.A.); no entanto, nesta refinaria é processada uma mistura grande de petróleos brutos, existindo a dificuldade de não se dispor das informações completas de alguns deles, valendo-se das similaridades entre as diferentes cargas para estimar os dados faltantes.

São consideradas as etapas básicas no processo da simulação, tais como:

- Definir os componentes leves presentes nas correntes do processo, assim como o conjunto termodinâmico adequado para determinar as propriedades físico-químicas das correntes. No caso do Petrox, é utilizado o modelo de Grayson-Streed, recomendado para simular sistemas de hidrocarbonetos pesados com alto conteúdo de hidrogênio.
- Caracterizar as cargas do(s) petróleo(s) processado(s) através das curvas de destilação TBP e outras informações adicionais, se disponíveis, tais como curvas de densidade, dados de viscosidade, peso molecular, entre outras. O simulador fornece uma representação daquele(s) petróleo(s) em forma de pseudocomponentes.
- Definir as correntes de entrada através de fluxos, temperaturas ou pressões ou composições, assim como as unidades envolvidas no processo (misturadores, trocadores, válvulas, torres, etc.).
- Especificar as variáveis necessárias (respeitando a análise de graus de liberdade) que permitam a obtenção de resultados úteis quanto às propriedades físico-químicas das frações destiladas obtidas, condizentes com a realidade no processo e atendendo às necessidades de produção.

Na refinaria, a qualidade da simulação é avaliada em termos de inferências, principalmente. São utilizadas quatro inferências: ponto ASTM – D85 do diesel, ponto de fulgor do diesel, ponto final de ebulição da nafta leve e ponto final de ebulição da nafta pesada. As equações destas inferências foram determinadas na

refinaria, através de uma análise PCA para determinar a sensibilidade das diferentes variáveis envolvidas no processo nas propriedades de destilação analisadas nelas, e os seus resultados estão sujeitos a fatores de correção.

Por outro lado, a comparação global entre os dados de destilação obtidos no simulador para as frações destiladas e os dados medidos em laboratório constitui uma ferramenta auxiliar para avaliar a performance da simulação realizada. Os destilados do processo analisados neste trabalho são representados através das curvas de destilação ASTM D-86 (por tratar-se de produtos) e são identificados com as seguintes siglas: NL (Nafta Leve), DL (Diesel Leve), NP (Nafta Pesada), Kero (Querosene) e DP (Diesel Pesado).

A simulação realizada na refinaria, que foi considerada como base a partir da qual foram estudados os fatores relevantes envolvidos no processo, foi executada em Junho de 2004. No entanto, no tempo transcorrido, mudanças de tipos de petróleo processado assim como operacionais aconteceram.

Neste trabalho, para efeitos de aplicação de novas metodologias de caracterização de petróleo, a simulação do processo é realizada utilizando o simulador comercial Hysys v.1.1, devido à maior disponibilidade de ferramentas que facilitam a visualização (ambiente gráfico) e avaliação das mudanças feitas, em cada uma das estratégias de simulação adotadas, as quais serão descritas posteriormente.

Quanto ao conjunto termodinâmico, foi escolhida a equação de estado de Peng-Robinson (PR) que é o método geralmente recomendado para simulações de sistemas de petróleo, gás e petroquímicos (ver manual de referência do simulador Hysys).

Quanto à verificação do desempenho adequado do simulador Hysys, foram utilizados dados de literatura correspondentes a um petróleo processado numa torre de destilação atmosférica, os que incluem composição de componentes leves (C_2 até $n-C_5$), curva TBP, especificações de produtos e condições de operação (Watkins, 1973). Isto

permitiu certificar-se que o simulador utilizado é capaz de fornecer resultados confiáveis, sendo que neste trabalho, que visa testar uma metodologia, é necessário partir de informações medidas com o objetivo de avaliar o seu desempenho.

6.5 Informação da fração de leves

O aspecto determinante na adoção das estratégias para introduzir os resultados da caracterização de petróleo na simulação corresponde à maneira como é tratada a composição de componentes leves na carga do processo. Normalmente a fração de leves é desconhecida ou de difícil estimação.

É importante lembrar que as curvas TBP medidas em laboratório começavam em 28°C (é realizada a análise de debutanização para determinar a composição dos componentes mais voláteis do petróleo, correspondentes à porção inicial da curva). Porém, na construção dos modelos NIR na previsão das curvas TBP, resultados satisfatórios foram somente obtidos a partir de 75°C, apontando que as regiões extremas apresentavam maiores desvios, como foi discutido no item 4.4.

Posteriormente foi discutido que, na determinação das curvas SimDis em laboratório o problema da perda de voláteis é muito difícil de ser evitado. Só a título de exemplo, no condicionamento da amostra já ocorre perda de leves.

A refinaria possui uma biblioteca de análises de destilação abrangendo os petróleos que normalmente são processados. Por exemplo, se fosse processada uma mistura de quatro petróleos diferentes, a análise de destilação de cada petróleo será fornecida ao simulador na etapa da caracterização, aplicando-se uma correção nas composições dos leves. Esta correção baseia-se em dados obtidos na debutanizadora.

O problema desta metodologia é que não consegue mais representar adequadamente a natureza da mistura quando ela é processada, devido ao espaço de tempo entre a realização das análises de destilação e a sua entrada na torre. Neste

trabalho, se analisa somente a carga da torre de destilação (que corresponderia à mistura de quatro petróleos).

Porém, devido à falta de informações mais precisas quanto à composição de leves da carga, utilizou-se a composição média em leves dos quatro petróleos processados na data escolhida. Esta composição foi obtida através de um balanço de massa dos leves de cada corrente de petróleo, sendo que as curvas completas de destilação para cada um deles estavam disponíveis, como explicado anteriormente. Porém, a representação adequada da mistura, quando ela é processada, é um fator ainda questionável. Na tabela abaixo, os resultados do balanço são apresentados, obtendo-se uma porcentagem total de 4.9% em peso de leves na carga.

Tabela 6.2 Composição de leves obtida através do balanço de massa nas cargas

Componente	% peso
C ₁	0.005
C ₂	0.101
C ₃	0.61
n-C ₄	0.803
i-C ₄	0.317
n-C ₅	0.546
i-C ₅	0.057
n-C ₆	2.461

De forma geral, é possível afirmar que:

- Os componentes leves na faixa de ebulição de C₂ a n-C₅ são comumente os mais importantes.
- Tomando como base as simulações realizadas na refinaria, o componente leve mais pesado é o n-C₆ ou n-hexano (cujo ponto normal de ebulição corresponde aproximadamente a 69°C).

- É preferível que a região da curva de destilação abaixo do ponto de ebulição do n-C₆ seja substituída com componentes puros discretos (da biblioteca). Isto ajuda a fornecer resultados mais precisos que quando utilizados pseudocomponentes para representar também a porção de leves (ver manual de referência do Hysys).

No Hysys, são oferecidas três alternativas para tratar a informação de leves:

- *Input Composition*. Quando a análise de componentes leves para cada corrente de petróleo a ser caracterizada está disponível. Representa o esquema convencional seguido atualmente na refinaria.
- *Auto Calculate*. Quando não existe análise de laboratório para a composição de leves. O simulador representará esta porção em forma de componentes discretos puros.
- *Ignore Composition*. Quando não existe análise de laboratório para a composição de leves. Com esta alternativa o simulador representará esta porção em forma de pseudocomponentes do petróleo.

Graças à utilização de uma composição média de leves calculada a partir do balanço de massa dos petróleos processados (tabela 6.2), é possível utilizar a opção *Input Composition* na simulação do processo. Por outro lado, a refinaria disponibilizou uma análise de rotina da corrente de gás obtida no topo da Torre de Pré-flash. Sendo um dos últimos ensaios disponíveis na refinaria (contendo dados de laboratório), foi utilizado como parâmetro de comparação entre o perfil de GLP obtido na simulação e a análise de laboratório. Na tabela 6.3 os valores são apresentados.

Tabela 6.3 Análise volumétrica da corrente de GLP na RECAP

Componente	% volume	% peso
H ₂	1.01	0.05
N ₂	0.71	0.48
O ₂	0.01	0.01
C ₁	0.00	0.00
C ₂	14.00	8.07
C ₃	34.83	30.13
n-C ₄	27.17	31.34
i-C ₄	9.50	11.00
n-C ₅	2.89	4.17
i-C ₅	0.09	0.13
n-C ₆	8.47	14.66

6.6 Estratégias de Simulação

O dia 05 de Outubro de 2004 foi utilizado como base para a implementação da metodologia de caracterização desenvolvida. A disponibilidade de informações de campo e de laboratório (curvas de destilação ASTM dos produtos) e curvas SimDis no mesmo período, foram as razões pelas quais esta data foi escolhida.

Nos dias 04 de Novembro e 15 de Dezembro de 2004 foram também coletados dados de campo. Porém, uma das outras fontes de informação não estava disponível: as curvas ASTM dos produtos para a primeira data, e a curva SimDis para a segunda, o que impediu a aplicação na simulação do processo. Quanto à curva SimDis do dia 04 de Novembro, destaca-se que o ensaio foi realizado, porém foram obtidos resultados não reproduzíveis devido a falhas na sua execução, que nos obrigaram a desconsiderá-la, podendo ser classificada com um ponto discrepante no conjunto de análises SimDis construído.

Cabe ressaltar que o dia escolhido como base para a simulação não corresponde ao do ensaio de GLP apresentado na tabela 6.3 (05 de Maio de 2005), e que este último foi utilizado para dispor de uma referência quanto ao perfil do gás de topo na torre pré-flash, sem possibilidade de maior uso na avaliação dos resultados obtidos.

A seguir, são apresentadas em mais detalhe as técnicas adicionais utilizadas na aplicação das metodologias de caracterização desenvolvidas na simulação do processo.

6.6.1 Estratégia Convencional

No período de análise escolhido foram processados quatro petróleos na refinaria. As curvas TBP de biblioteca estavam disponíveis para todos eles. O esquema da parte inicial do processo é apresentado na figura 6.5. A corrente BSW representa a água contida na carga.

Figura 6.5 Esquema de simulação convencional

6.6.2 Conversão das curvas SimDis em TBP – Método API

A conversão da curva SimDis em TBP é necessária já que o simulador sempre utiliza esta última para calcular as propriedades físico-químicas do petróleo. Sendo assim, uma revisão das técnicas disponíveis foi realizada.

Sendo que as curvas obtidas a partir dos espectros NIR plausíveis de serem inseridas na simulação do processo foram as curvas SimDis, os métodos de conversão

entre elas e as curvas TBP foram revisados, já que são estas que representam as curvas de destilação de referência.

Existem correlações na literatura para converter destilações ASTM D86, TBP e SimDis a pressões reduzidas e atmosféricas. O procedimento começa em um ponto determinado (50% destilado) como um valor correlacionado nos dois tipos de destilações envolvidas. Posteriormente adições e subtrações de incrementos são necessárias para completar a curva.

Neste trabalho, este método de conversão foi aplicado através do simulador Hysys, que utiliza as correlações baseadas nas equações propostas por Daubert (1994) para transformar dados ASTM D2887 em TBP a pressão atmosférica. Estas correlações foram reportadas na revisão do capítulo 3 do *API Technical Data Book, Petroleum Refining*, 5ª edição (1992) que foi publicada em 1994; assim como também em Mcketta e Cunningham (1990).

Na determinação dos pontos das curvas TBP assume-se que a temperatura TBP para 50% volume destilado é igual à temperatura SimDis na mesma percentagem. Constantes para diferentes faixas de temperatura no equacionamento são necessárias. Na figura 6.6 as formas das curvas TBP e SimDis são apresentadas.

É importante ressaltar que outros tipos de dados SimDis (como ASTM D5307, norma seguida neste trabalho) são utilizados no processo de caracterização na indústria e as principais diferenças entre as normas ASTM deste tipo baseiam-se nas amostras apropriadas para os testes (petróleo bruto, frações de petróleo, etc.). Porém, como métodos cromatográficos, estas técnicas estão baseadas em tempos de retenção e áreas, razão pela qual foi possível aplicar o método de conversão citado acima nos dados SimDis ASTM D5307 disponíveis. As discussões com o departamento de suporte técnico da AspenTech (distribuidora do simulador Hysys) permitiram confirmar a possibilidade de aplicação deste método às informações disponíveis neste trabalho.

Figura 6.6 Destilações SimDis e TBP típicas (adaptado de Daubert, 1994)

As desvantagens deste procedimento são a sua precisão relativamente baixa e a determinação de poucos pontos da curva TBP, os mesmos que têm que estar disponíveis na curva SimDis, sendo necessário portanto, a aplicação de algum método de extrapolação. No Hysys, utiliza-se o método de Probabilidade.

Utilizando este método de conversão é possível introduzir dados TBP na simulação. Assim, as informações da entrada do esquema convencional (figura 6.5) são alteradas, onde as quatro correntes de petróleo processado no período são substituídas por uma única corrente, correspondente à análise da carga do processo, como apresentado na figura 6.7.

Figura 6.7 Esquema de simulação utilizando o método API de conversão das curvas SimDis em TBP

6.6.3 Manipulação da parte inicial das curvas de destilação

O simulador Hysys gera a partir dos dados de destilação de entrada (neste trabalho, dados Simdis) uma curva TBP interna com a qual inicia o processo de simulação. Os dados de destilação correspondem à curva propriamente dita e à fração de leves.

Na simulação, a fração de leves foi introduzida em uma corrente externa à carga de petróleo. A composição desta corrente externa corresponde aos valores da tabela 6.2, com uma vazão mássica de 4.9% da vazão da carga total (como explicado no item 6.5). Devido a que teoricamente, esta corrente contém os leves da carga, foi necessária uma manipulação da parte inicial da curva SimDis fornecida ao simulador, com o intuito de deixá-la desprovida da presença de leves.

Foram assim geradas três curvas de destilação distintas, cujas partes iniciais são apresentadas na tabela abaixo, considerando como base a temperatura de ebulição do n-hexano (aproximadamente 69°C), componente leve mais pesado presente na carga de petróleo analisada.

Tabela 6.4 Manipulação na parte inicial das curvas SimDis

M0		M1		M2	
T	%	T	%	T	%
35	0	75	0	75	0
50	1.06	95	1.22	95	4.64
75	3.47				
95	4.64				

M0 representa a curva original e, que contém leves na fração inicial entre 35 e 75°C. Na curva total foi feita uma redistribuição das composições mássicas com o intuito de deixá-la livre de voláteis (a partir de 75°C), obtendo-se o perfil da curva M1, representando uma curva mais pesada quanto menor a temperatura. Finalmente, foi

testada uma terceira alternativa, onde assume-se que toda a fração de leves é retirada depois da temperatura superior à temperatura de ebulição do n-hexano, ou 95°C, sendo denominada M2, representando uma curva igual à original, sem a parte correspondente aos leves, tornando-a mais pesada nesta região.

6.7 Resultados

As estratégias mencionadas no item anterior foram implementadas, com o intuito de observar o seu efeito na qualidade da simulação, comparando as três diferentes formas de introduzir a curva de destilação no simulador. Foram também testados dois conjuntos diferentes de especificações. Os resultados obtidos são apresentados a seguir.

6.7.1 Simulação da Torre Pré-Flash

Na simulação desta torre foram encontradas as principais dificuldades. O objetivo principal desta torre na destilação é retirar os cortes mais leves para ampliar a carga total da unidade, e sendo que a maior deficiência nas curvas de destilação determinadas a partir do NIR é a capacidade deficiente para representar a informação de leves, o desvio destes resultados em relação aos dados fornecidos pela refinaria teve um impacto significativo nos resultados da torre pré-flash.

Os conjuntos de variáveis fixadas na simulação das duas torres (pré-flash e atmosférica) foram escolhidos respeitando os graus de liberdade do processo (o Hysys impossibilita o usuário de executar a simulação caso esta restrição não seja cumprida) e considerando a viabilidade técnica em campo. Desta maneira, na tabela 6.5 mostra-se os conjuntos de especificações para a torre pré-flash.

O conjunto II, que inclui somente temperaturas, foi testado com o intuito de observar o efeito nos perfis de vazões da coluna quando a qualidade é o principal aspecto a ser reproduzido.

Tabela 6.5 Variáveis especificadas na simulação da torre pré-flash

Variáveis especificadas	
Conjunto I	Taxa de refluxo
	Vazão de Diesel Leve (m ³ /d)
	Temperatura do condensador (°C)
Conjunto II	Temperatura de topo (°C)
	Temperatura de Diesel Leve (°C)
	Temperatura do condensador (°C)

Em termos das inferências, o ponto final de ebulição da NL corresponde à torre pré-flash, pelo qual o quadro comparativo final das inferências será apresentado no próximo item.

Com o objetivo de quantificar a precisão da simulação em termos de vazões e temperaturas, foi feita uma análise através do erro quadrático médio, calculado como mostrado nas equações abaixo. Nas tabelas 6.6 e 6.7 os resultados correspondentes para cada conjunto de especificações são apresentados. Na tabela 6.8 as vazões de topo de GLP são também comparadas.

$$\text{erro médio quadrático} = \sqrt{\frac{\sum_{i=1}^n \left(\frac{Fcal_i - Fmed_i}{Fmed_i} \right)^2}{n}} \quad (6.1)$$

$$\text{erro médio quadrático (°C)} = \sqrt{\frac{\sum_{i=1}^n (Tcal_i - Tmed_i)^2}{n}} \quad (6.2)$$

onde

$Fcal$ é o valor da vazão calculada (m³/d)

$Fmed$ é o valor da vazão medida (m³/d)

T_{cal} é o valor da temperatura calculada (°C)

T_{med} é o valor da temperatura medida (°C)

n é o número de variáveis analisadas

Tabela 6.6 Resultados em termos de vazões e temperaturas na torre pré-flash para o conjunto de especificações I

Vazão (m ³ /d)	RECAP	Convencional	M0	M1	M2
Refluxo	1603	1250	1310	1020	1240
NL	1089	777	870	605	816
Fundo	6234	6760	6680	6900	6730
erro médio quadrático		0,21	0,16	0,34	0,20
Temperatura (°C)					
Topo	115	104	111	115	116
Fundo	185	187	189	192	189
DL	159	140	153	157	152
erro médio quadrático (°C)		12,73	4,76	4,20	4,69

Tabela 6.7 Resultados em termos de vazões na torre pré-flash para o conjunto de especificações II

Vazão (m ³ /d)	RECAP	Convencional	M0	M1	M2
Refluxo	1603	878	1130	963	1010
NL	1089	940	913	595	968
DL	170	397	381	332	631
Fundo	6234	6400	6430	6750	6320
erro médio quadrático		0,71	0,64	0,57	1,37

Tabela 6.8 Vazões de GLP obtidas na torre pré-flash

Vazão (kg/h)	Convencional	M0	M1	M2
Conjunto I	2990	2830	3800	2850
Conjunto II	2280	2630	3820	3020

De forma geral pode-se concluir que, segundo os resultados das tabelas 6.6, 6.7 e 6.8, o petróleo processado utilizando qualquer uma das três curvas testadas (três diferentes formas de manipular a fração inicial da curva) é mais pesado do que o realmente processado na refinaria, razão pela qual uma vazão menor de NL é obtida no topo. Esta situação é comum aos dois conjuntos de especificações utilizados.

O tag correspondente à vazão de GLP em campo estava fora de operação no período analisado. Assim, para propósitos de simulação, a refinaria utiliza o valor de 2596 kg/h*.

Segundo os resultados da simulação, quando é utilizado o conjunto I (tabela 6.6); no que diz respeito às vazões, o fato do petróleo ficar mais pesado impede que seja atingida a produção de nafta leve requerida pela planta, impossibilitando inclusive sua especificação no simulador. A especificação da vazão de diesel leve permite contornar de alguma forma esta dificuldade. Quanto às temperaturas, as alternativas ou curvas propostas conseguem melhorar os resultados obtidos numa simulação convencional, atingindo perfis de temperatura mais próximos aos da refinaria, motivo pelo qual na comparação das curvas dos produtos destilados, uma maior aproximação é observada quando utilizado o perfil M1.

Segundo os resultados da simulação, quando é utilizado o conjunto II (tabela 6.7), observa-se um erro maior no perfil de vazões devido ao fato de deixar livre a especificação de produção de diesel leve (embora seja possível retirar maior quantidade de nafta leve), com o objetivo de ajustar os perfis de temperaturas dos produtos, o que é observado melhor através de gráficos (figuras 6.8 e 6.9). Cabe

* Comunicação pessoal com o Departamento de Otimização da RECAP

ressaltar que, ao ter como única variável de temperatura livre a temperatura de fundo, não foi necessário o cálculo do erro quadrático médio, e observando-se uma diferença entre 7 e 9 °C em relação ao valor da refinaria. Nas figuras 6.8 e 6.9 apresentam-se os resultados em termos das curvas ASTM-D86 dos produtos destilados obtidos nesta torre, sendo comparadas com as curvas de laboratório da refinaria (linha contínua).

(a)

(b)

Figura 6.8 Curva ASTM D-86 de Nafta Leve ou NL: (a) conjunto I e (b) conjunto II

(a)

(b)

Figura 6.9 Curva ASTM-D86 de Diesel Leve ou DL: (a) conjunto I e (b) conjunto II

Pode-se observar que, as curvas obtidas através do conjunto II de especificações fornecem melhores resultados qualitativos, resultando evidente a melhor representação de produtos mais pesados (corrente DL). Na maioria dos casos, a alternativa M1 fornece melhores resultados qualitativos quanto às curvas de destilação,

e resultados quantitativos também melhores quando utilizado o conjunto II na simulação (em detrimento da produção de diesel leve). Porém no conjunto I, M1 teve o melhor desempenho em relação ao perfil de temperaturas. É evidente que as tentativas de reproduzir melhor a fração de leves permitem inclusive, em alguns casos, uma maior precisão na simulação.

Embora seja um aspecto que foi discutido anteriormente no item 6.6.2, é válido ressaltar que o Hysys utiliza um procedimento empírico de conversão entre as curvas Simdis e TBP, cujas constantes nas faixas de destilação definidas no método podem não ser adequadas para a carga analisada. Uma proposta em relação a isto é feita no capítulo 7.

6.7.2 Simulação da Torre Atmosférica

Na tabela 6.9 mostram-se os conjuntos de variáveis escolhidas para serem fixadas na simulação desta unidade.

Tabela 6.9 Variáveis especificadas na simulação da torre atmosférica

Variáveis especificadas	
Conjunto I	Vazão de Nafta Pesada (m ³ /d)
	Vazão de Diesel Pesado (m ³ /d)
	Temperatura de Diesel Pesado (°C)
	Vazão do pumparound (m ³ /d)
	Temperatura de saída do pumparound (°C)
Conjunto II	Vazão de Diesel Pesado (m ³ /d)
	Temperatura de topo (°C)
	Temperatura de Diesel Pesado (°C)
	Temperatura de Querosene (°C)
	Temperatura de saída do pumparound (°C)

Como no caso da torre pré-flash, foi calculado também o erro médio quadrático das vazões e temperaturas para avaliar o efeito das alternativas propostas na simulação. Nas tabelas 6.10 e 6.11 apresentam-se os resultados para cada conjunto de especificações.

Tabela 6.10 Resultados em termos de vazões e temperaturas na torre atmosférica para o conjunto de especificações I

Vazão (m³/d)	RECAP	Convencional	M0	M1	M2
Refluxo	1295	3060	3280	3320	3560
Fundo	3216	3080	2620	2760	2610
Kero	1300	1700	2060	2150	2120
erro médio quadrático		0,96	1,01	1,11	1,24
Temperatura (°C)					
Topo	102	101	118	122	112
Fundo	367	369	362	362	362
Kero	182	181	189	191	186
erro médio quadrático (°C)		1,4	10,5	13,0	6,9

Tabela 6.11 Resultados em termos de vazões na torre atmosférica para o conjunto de especificações II

Vazão (m³/d)	RECAP	Convencional	M0	M1	M2
NP	125	109	52	38	46
Refluxo	1295	1130	1110	1320	1570
Pumparound	4462	6460	8090	4480	6760
Fundo	3216	3120	2640	2770	2650
Kero	1300	1310	1870	2070	1750
erro médio quadrático		0,22	0,50	0,41	0,42

Com os resultados obtidos podemos afirmar que, ao tentar manipular a fração de leves visando sua melhor reprodução e obtendo melhoras na simulação (como visto nos resultados da torre pré-flash) altera-se a quantidade relativa das frações mais pesadas provocando desvios significativos, principalmente nos perfis de vazões correspondentes ao conjunto I de especificações (tabela 6.10). Isto é um indicador de que a manipulação de informação já existente em relação aos leves não resolve o problema de falta de informação mais representativa, obtendo resultados insatisfatórios mesmo quando os cortes mais leves já foram retirados na unidade anterior.

Já nos resultados quando especificado o conjunto II (tabela 6.11), consegue-se melhores resultados quanto às vazões. Neste caso, a temperatura de fundo é a única variável de temperatura que não foi especificada, obtendo-se uma diferença de 5°C em relação ao valor da refinaria.

Nas figuras 6.10 a 6.13 apresentam-se os resultados em termos das curvas de destilação ASTM-D86 comparadas com as curvas de laboratório da refinaria.

Como pode ser observado, existe uma tendência a maior proximidade entre as curvas conforme os produtos tornam-se mais pesados, o que se explica pelo fato de que os componentes mais leves já foram retirados nos primeiros cortes. Nesse sentido, no conjunto de especificações II, onde se consegue ajustar melhor a qualidade dos produtos, os perfis de destilação resultam ser mais próximos aos perfis de laboratório.

É importante indicar que as curvas de laboratório estão sujeitas a erros sistemáticos, como observado na curva TBP do RAT, que não cumpre a característica de monotonicidade crescente de curva de destilação; no entanto, isto não altera significativamente os resultados obtidos. De forma geral, observa-se que, na maioria dos produtos, a simulação convencional reproduz as curvas ASTM da melhor forma possível, sendo um indício de que a manipulação da fração mais volátil da curva (a partir de informações TBP) não é adequada quando utilizada com informações SimDis.

(a)

(b)

Figura 6.10 Curva ASTM-D86 de Nafta Pesada ou NP: (a) conjunto I e (b) conjunto II

(a)

(b)

Figura 6.11 Curvas ASTM D-86 de Querosene ou Kero: (a) conjunto I e (b) conjunto II

(a)

(b)

Figura 6.12 Curvas ASTM D-86 de Diesel Pesado ou DP: (a) conjunto I e (b) conjunto II

(a)

(b)

Figura 6.13 Curvas TBP de Resíduo Atmosférico ou RAT: (a) conjunto I e (b) conjunto II

Finalmente, foram calculados os valores das inferências utilizadas na refinaria. Os resultados são apresentados a seguir.

Tabela 6.12 Resultados das inferências nas simulações (conjunto I)

Inferência (°C)	Refinaria	Convencional	M0	M1	M2
PFE NL	140	126	140	147	144
PFE NP	144	132	150	154	144
Ponto ASTM D-85 Diesel	355	366	350	350	350
Ponto de fulgor Diesel	45	49	55	57	53
erro médio quadrático		10,9	6,3	8,9	5,1

Tabela 6.13 Resultados das inferências nas simulações (conjunto II)

Inferência (°C)	Refinaria	Convencional	M0	M1	M2
PFE NL	140	148	149	150	147
PFE NP	144	139	142	142	135
Ponto ASTM D-85 Diesel	355	367	351	346	347
Ponto de fulgor Diesel	45	53	54	53	48
erro médio quadrático		8,6	6,7	7,9	7,1

No caso das inferências, avaliar o desempenho das simulações unicamente através do valor do erro não é adequado, já que cada uma delas representa propriedades distintas do processo.

Assim, o valor do ponto final de ebulição da nafta leve ficaria mais bem representado pelas alternativas do conjunto I (tabela 6.12) embora os perfis completos se distanciem mais dos perfis de referência. Por outro lado, na prática, devido ao baixo fracionamento na torre pré-flash, a Nafta Leve pode sair contaminada, provocando que o PFE desta corrente seja superior ao da Nafta Pesada. Isto também ocorre na simulação (tabela 6.13).

Quanto às inferências do diesel (mistura de DL, DP e Kero) observa-se que as maiores diferenças encontram-se na determinação do ponto de fulgor, o qual varia em função do teor de hidrocarbonetos leves existentes, ressaltando o impacto da deficiência na representação desta fração nos resultados, em contraste com a determinação do ponto ASTM D-85 do diesel, que está relacionado com o teor de frações pesadas no óleo. Neste caso, os resultados utilizando o conjunto I são mais próximos que o valor de laboratório.

Cabe ressaltar, no entanto, que no caso de representar mais adequadamente a fração de leves da curva de destilação, seria possível obter valores mais próximos das inferências medidas em laboratório, sem a necessidade de fatores de correção, que atualmente é necessário aplicar aos valores obtidos através da análise das variáveis que influem na determinação de cada uma das inferências (como explicado no item 6.4).

Desta forma, a partir dos resultados obtidos, é evidente que a estratégia de utilizar a curva SimDis, determinada a partir da carga efetiva ao processo (e que surgiria a partir do espectro NIR do petróleo bruto) tem potencial para reproduzir um cenário real do processo de destilação atmosférica, precisando-se aprofundar a investigação quanto a parte mais volátil da curva de destilação para superar as dificuldades encontradas nesta etapa do trabalho, que não permitiram a obtenção de resultados mais satisfatórios. Os resultados obtidos utilizando dois conjuntos de especificações reforça esta afirmação.

7 CONCLUSÕES

Neste trabalho foram desenvolvidas novas metodologias de caracterização de petróleo bruto a partir dos seus espectros NIR permitindo a previsão de propriedades físico-químicas que tornam a caracterização mais rápida do que quando realizada através de métodos convencionais, cujos resultados são utilizados na simulação de um processo industrial de destilação atmosférica visando avaliar o seu impacto. As principais etapas realizadas para alcance do objetivo proposto estão descritas a seguir.

7.1 Modelagem das curvas de destilação a partir do NIR

A curva de destilação TBP representa a curva de caracterização clássica de qualquer petróleo bruto, porém é uma análise que precisa de altos investimentos em tempo e dinheiro. O NIR surge como uma técnica que poderia extrair a informação relevante e suficiente do petróleo bruto segundo suas propriedades espectroscópicas, visando a estimação da sua curva TBP a partir do seu espectro. A maioria dos trabalhos reportados na literatura aponta o sucesso da aplicação do NIR na estimação de diversas propriedades de produtos destilados, sendo Hidajat e Chong (2000) os que inicialmente propuseram a aplicação do NIR na análise de petróleo bruto processado em refinarias do Oriente Médio.

Neste trabalho, a análise TBP de amostras de petróleo bruto processado numa refinaria brasileira foi estimada com sucesso quando a técnica PLS foi aplicada na construção dos modelos de predição a partir do NIR na faixa espectral de 5263 – 6060 cm^{-1} . A rápida volatilização dos componentes mais leves impede, no entanto, uma reprodução mais completa da curva TBP na sua parte inicial. Os modelos construídos na linguagem de programação MatLab mostram que a complexidade da natureza do petróleo bruto exige flexibilidade na calibração dos modelos para sua respectiva adequação.

A curva de destilação SimDis (método cromatográfico gasoso) representa atualmente a aproximação mais confiável da curva TBP de um petróleo, por causa dos

menores custos em tempo e dinheiro envolvidos. Porém, a volatilização inicial da amostra acarreta em falta de informação dos componentes leves na curva final calculada. Outras técnicas de calibração foram escolhidas para a estimação destas curvas (além da técnica PLS), visando expandir o grupo de técnicas quimiométricas que seriam passíveis de ser aplicadas na extração da informação contida nos espectros NIR de petróleo bruto.

Desta maneira, a análise SimDis de amostras de petróleo bruto foi estimada com sucesso quando a técnica de redes neurais foi aplicada na construção dos modelos de predição a partir do NIR na faixa espectral de $5600 - 6000 \text{ cm}^{-1}$. Menores custos em termos de tempo e operação fomentaram a sua aplicação, acarretando por outro lado, problemas de reprodução da parte inicial das curvas, como na previsão das curvas TBP. O impacto negativo desta deficiência foi observado na simulação do processo.

Contudo, a espectroscopia NIR revelou-se, como apontado em inúmeras experiências citadas na literatura, como uma técnica poderosa na análise química na área do petróleo, demonstrando o seu poder de extração de informação altamente relevante no caso não somente de produtos destilados numa refinaria, mas também da carga de petróleo bruto do processo.

Finalmente, quanto ao desenvolvimento das metodologias de caracterização propostas neste trabalho, podem ser salientadas as seguintes observações, que resumem o caminho seguido neste trabalho e as conclusões sobre os principais aspectos envolvidos numa inovadora aplicação NIR no setor nacional de petróleo:

- A revisão das principais técnicas quimiométricas existentes atualmente (apêndice A) permitiu compreender o processo de extração da informação química contida nos espectros NIR, as principais dificuldades existentes quanto à instrumentação e natureza das amostras, e as ferramentas para contorná-las.
- Os critérios considerados em experiências reportadas na literatura relacionadas ao petróleo (sendo quase todas referidas a produtos destilados, principalmente

gasolinas, diesel e querosene) serviram para propor metodologias alternativas; visando a observação de aspectos como o desempenho das técnicas quimiométricas em tais experiências e as condições experimentais testadas (amostragem, temperatura, etc.).

- Nos espectros NIR do petróleo bruto, a variação da temperatura não tem um efeito significativo na informação neles contida, como o demonstram os testes realizados, dentro de uma região espectral determinada. Neste trabalho, esta região corresponde aproximadamente à faixa de $5000 - 6000 \text{ cm}^{-1}$ (considerando a previsão das análises TBP e SimDis). Em concordância com isto, uma boa reprodutibilidade espectral é observada também nesta região.
- Os dois tipos de modelos desenvolvidos: lineares (PLS) e não-lineares (Redes Neurais) tiveram bom desempenho numa faixa considerável de temperaturas, mesmo que para temperaturas extremas (muito baixas e muito altas) os resultados não foram adequados. Do ponto de vista experimental, este é um aspecto extremamente difícil de manipular. Formas possíveis de contornar esta dificuldade só conseguem ser aplicadas de modo complementar na aplicação dos resultados dos modelos na simulação do processo.

7.2 Simulação do processo de destilação atmosférica

A aplicação dos resultados das novas metodologias de caracterização de petróleo na simulação de um processo real de refino foi o objetivo desta etapa do trabalho. Os resultados obtidos forneceram uma visão mais aguda dos aspectos que precisam ser mais bem analisados para alcançar uma sintonia mais produtiva com a realidade do processo.

A construção de uma base de dados industriais (variáveis de processo e inferências utilizadas, e análises de laboratório dos destilados) num período onde a medição dos espectros NIR para caracterizar o petróleo processado esteja disponível é o primeiro aspecto a ser abordado. Dificuldades foram encontradas por falta de

informação correspondente a perfis de destilação de alguns petróleos, sendo que as cargas mudam com certa frequência dependendo das necessidades de produção da planta. Este fato permitiu testar as metodologias desenvolvidas numa única data de processamento.

A composição e a quantidade de componentes leves na carga do processo de destilação tornaram-se o maior problema a ser enfrentado na obtenção de resultados mais satisfatórios. A perda inevitável de componentes voláteis na execução da análise SimDis provoca uma falta de informação correspondente à reprodução da parte inicial na curva de destilação. Os resultados também demonstram que mais do que uma manipulação da fração de leves (o que implica uma adequação de informação já existente e não suficientemente representativa), o que seria necessário é a determinação adequada da informação necessária, que foi o objetivo inicial deste trabalho, onde foi proposta a caracterização da carga de petróleo mais do que a adequação de informações já disponíveis.

Por outro lado, as limitações do procedimento empírico de conversão entre as curvas SimDis e TBP no simulador Hysys também podem ser consideradas fatores que impediram a obtenção de melhores resultados.

O fato de aplicar os resultados de uma pesquisa num caso industrial real acarreta as dificuldades encontradas neste trabalho, sendo que a qualidade das informações vindas da refinaria tem um papel preponderante na confiabilidade da avaliação do trabalho realizado. Nesse sentido, a aplicação deste trabalho revela um efeito promissor nos resultados da simulação do processo de destilação atmosférica, o que reforça a viabilidade técnica da utilização do NIR nesta área industrial.

7.3 Divulgação dos resultados da pesquisa

Apresentação de trabalhos em Congressos:

1. Falla, F.S.; Le Roux, G.A.C.; Moro, L. F. L. **Caracterização de Petróleo Bruto a través da Espectroscopia NIR e da Quimiometria**. XV COBEQ 2004, XV Congresso Brasileiro de Engenharia Química – Curitiba, 2004.

Artigos Submetidos para Revistas:

1. Falla, F.S.; Larini, C.; Le Roux, G.A.C.; Quina, F.H.; Moro, L.F.L.; Nascimento, C.A.O. **Characterization of crude petroleum by NIR**, *Journal of Petroleum Science and Engineering*, 51, 127 – 137, 2006.

7.4 Propostas para a continuação da pesquisa

Neste trabalho foram desenvolvidas metodologias de caracterização de petróleo bruto inovadoras e aplicadas na simulação de um processo de destilação atmosférica. A principal frente para continuação do trabalho visaria a aplicação futura das metodologias de caracterização de petróleo bruto no esquema “*on-line*” no processo, implicando no desenvolvimento de uma nova metodologia de otimização de uma unidade de destilação atmosférica.

Na extensão do trabalho realizado nesta tese, sugerem-se os tópicos a seguir.

Proposta de expansão da pesquisa realizada: caracterização *on-line* de petróleo bruto:

1. Montagem de um *loop* ou circuito fechado onde o petróleo processado na refinaria possa circular com a possibilidade de variação da temperatura.

2. Medição dos espectros NIR do petróleo no *loop*. Testes de reprodutibilidade e de efeito da temperatura são necessários.
3. Calibração dos modelos NIR para a previsão das curvas de destilação (TBP ou SimDis). Os modelos desenvolvidos nesta tese deverão ser adequados.

Propostas de aprofundamento da pesquisa realizada:

2. Desenvolvimento de uma metodologia de caracterização de petróleo mais robusta visando a conversão entre as curvas TBP e SimDis a partir do NIR, para superar as dificuldades encontradas na aplicação de um procedimento empírico, como utilizado no Hysys.
3. Estudo mais detalhado sobre a análise de componentes leves na caracterização de petróleo e das formas do tratamento deste tipo de informação para serem aplicadas juntamente com as metodologias de caracterização desenvolvidas. O trabalho de Roussis e Fitzgerald (2000) utilizou a espectrometria de massa para auxiliar na execução da análise SimDis, revelando-se como uma técnica importante que pode melhorar o seu desempenho.

REFERÊNCIAS

Agatonovic-Kustrin, S.; Beresford, R., Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research, *J. Pharm. Biom. Anal.*, 22, p. 717 – 727, 2000

Andrade, J.M.; Garcia, M.V.; Lopez-Mahia, P.; Prada, D., A review of the main factors influencing the FT-IR-PLS abilities exemplified with petrochemical qualimetric applications, *Talanta*, 44, p. 2167 – 2184, 1997

American Petroleum Institute Technical Data Book, *Chapter 3: Petroleum fraction distillation interconversions*. 5th edition (1992)

Bacaud, R.; Rouleau, L., Coupled simulated distillation – mass spectrometry for the evaluation of hydroconverted petroleum residues, *J. Chromatog. A.*, 750, p. 97 – 104, 1996

Basak, K.; Abhilash, K.S.; Ganguly, S.; Saraf, D.N., On-line optimization of a crude distillation unit with constraints on product properties, *Ind. Eng. Chem. Res.*, Vol. 41, p. 1557 – 1568, 2002

Barak, P., Smoothing and Differentiation by an adaptive-degree polynomial filter, *Anal. Chem.*, Vol. 67, p. 2758 – 2762, 1995

Baughman, E.H.; Van Hare, D.R.; Vickers, G.H., NIR calibration: Why is it so hard in reality and so easy in feasibility studies?, *NIR News*, Vol. 4 (6), p. 10 – 11, 1993

Blanco, M.; Coello, J.; Montoliu, I.; Romero, M.A., Orthogonal signal correction in near infrared calibration, *Anal. Chim. Acta*, 434, p. 125 – 132, 2001

Blanco, M.; Canals, T.; Coello, J.; Gené, J.; Iturriaga, H.; Maspoch, S., Direct determination of leather dyes by visible reflectance spectroscopy using partial least-squares regression, *Anal. Chim. Acta*, 419, p. 209 – 214, 2000

Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; Rovira, E., Determination of water in ferrous lactate by near infrared reflectance spectroscopy with a fibre-optic probe, *J. Pharm. Biom. Anal.*, 16, p. 255 – 262, 1997

Borggaard, C.; Thodberg, H.H., Optimal minimal neural interpretation of spectra, *Anal. Chem.*, Vol. 64, N° 5, p. 545 – 551, 1992

Bouveresse, E.; Hartmann, C.; Massart, D.L., Standardization of Near-Infrared Spectrometric Instruments, *Anal. Chem.*, Vol. 68, N° 6, p. 982 – 990, 1996a

Bouveresse, E.; Massart, D.L., Improvement of the piecewise direct standardization procedure for the transfer of NIR spectra for multivariate calibration, *Chem. Intellig. Lab. Systems*, 32, p. 201 – 213, 1996b

Briesen, H.; Marquardt, W., New approach to refinery process simulation with adaptive composition representation, *AIChE Journal*, Vol. 50, N° 3, p. 633 – 645, 2004

Burns, D.A.; Ciurczak, E.W., *Handbook of near-infrared analysis*, Marcel Dekker Inc, 2001

Chaminade, P.; Baillet, A.; Ferrier, D., Data treatment in near infrared spectroscopy, *Analysis*, 26 (4), M33 – M38, 1998

Chatterjee, T.; Saraf, D.N., On-line estimation of product properties for crude distillation units, *J. Process Control*, Vol. 14, p. 61 – 77, 2004

Chung, H.; Ku, M-S., Comparison of near-infrared, infrared, and Raman spectroscopy for the analysis of heavy petroleum products, *App. Spectrosc.*, 54 (2), p. 239 - 245, 2000

Chung, H.; Ku, M-S.; Lee, J-S., Comparison of near-infrared and mid-infrared spectroscopy for the determination of distillation property of kerosene, *Vib. Spectrosc.*, 20, p. 155 – 163, 1999

Daubert, T.E., Petroleum fraction distillation interconversions, *Hydroc. Processing*, 73 (9), p. 75 – 78, 1994

Durand, J.P., Le role de la CPG dans l'industrie pétrolière et la petrochimie. De l'analyse détaillée des hydrocarbures à la distillation simulée, *Analisis*, 26, N°1, p. M17 – M21, 1998

Durand, J.P.; Bré, A.; Beboulene, J.J.; Ducrozel, A.; Carbonneaux, S. *Oil & Science and Technology* – Rev. IFP, Vol. 54, p. 431 – 438, 1999

Espinosa, A.; Lambert, D.; Valleur, M., Use NIR technology to optimize plant operations, *H. Processing (international ed.)*, 74, p. 86 - 94, 1995

Eustaquio, A.; Blanco, M.; Jee, R.D.; Moffat, A.C., Determination of paracetamol in intact tablets by use of near infrared transmittance spectroscopy, *Anal. Chim. Acta*, 383, p. 283 – 290, 1999

Ferré, J.; Rius, F.X., Selection of the best calibration sample subset for multivariate regression, *Anal. Chem.*, 68, N° 9, p. 1565 – 1571, 1996

Geladi, P.; Kowalski, B.R., Partial least-squares regression: a tutorial, *Anal. Chim. Acta*, 185, p. 1 – 17, 1986

Geladi, P.; MacDougall, D.; Martens, H., Linearization and scatter-correction for near-infrared reflectance spectra of meat, *Ap. Spectrosc.*, Vol. 39 (3), p. 491 – 500, 1985

Gorry, P.A., General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method, *Anal. Chem.*, 62, p. 570 – 573, 1990

Grob, R.L., *Modern practice of gas chromatography, Cap.10: Petroleum and petrochemical analysis by gas chromatography*, John Wiley & Sons, 1995

Haaland, D.M.; Thomas, E.V., Partial least-squares methods for spectral analyses. 1.Relation to other quantitative calibration methods and the extraction of qualitative information, *Anal. Chem.*, Vol. 60, N° 11, p. 1193 – 1202, 1988

Helland, I.S.; Naes, T.; Isaksson, T., Related versions of the multiplicative scatter correction method for preprocessing spectroscopic data, *Chemom. Intell. Lab. Syst.*, 29, p. 233 – 241, 1995

Hidajat, K.; Chong S.M., Quality characterisation of crude oils by partial least squares calibration of NIR spectral profiles, *J. Near Infrared Spectrosc.*, 8, p. 53 – 59, 2000

Hyprotech Ltd. *Hysys version 1.1 Reference Volume 1,2*, 1996

Hobson, G.D., *Modern petroleum technology*, John Wiley & Sons, 1984

Honigs, D.E.; Hieftje, G.M.; Mark, H.L.; Hirschfeld, T.B., Unique-sample selection via near-infrared spectral subtraction, *Anal. Chem.*, 57, p. 2299 – 2303, 1985a

Honigs, D.E.; Hirschfeld, T.B.; Hieftje, G.M., Near-infrared determination of several physical properties of hydrocarbons, *Anal. Chem.*, 57, p. 443 – 445, 1985b

Hu, S.; Towler, G.; Zhu, X.X., Combine molecular modeling with optimization to stretch refinery operation, *Ind. Eng. Chem. Res.*, 41, p. 825 – 841, 2002

Isaksson, T.; Naes, T., Some aspects of scatter-correction in calibration. Part II, *NIR News*, Vol. 5 (3), p. 13 – 14, 1994

Kelly, J.J.; Callis, J.B., Nondestructive analytical procedure for simultaneous estimation of the major classes of hydrocarbon constituents of finished gasolines, *Anal. Chem.*, 62, p. 1444 – 1451, 1990

Kim, M.; Lee, Y-H.; Han, C., Real-time classification of petroleum products using near-infrared spectra, *Comput. Chem. Engng.*, 24, p. 513 – 517, 2000

Kowalski, B.R.; Seasholtz, M.B., Recent developments in multivariate calibration, *Journal of Chemometrics*, Vol. 5, p. 129 – 145, 1991

Lambert, D.; Descales, B.; Llinas, JR.; Espinosa, A.; Osta, S.; Sanchez, M.; Martens, A., On-line NIR monitoring and optimisation for refining and petrochemical processes, *Analisis*, 23 (4), M9 - M13, 1995

Litani-Barzilai, I.; Sela, I.; Bulatov, V.; Zilberman, I.; Schechter, I., On-line remote prediction of gasoline properties by combined optical methods, *Anal. Chim. Acta*, 339, p. 193 – 199, 1997

Martens, H.; Naes, T., *Multivariate Calibration*, John Wiley & Sons, 1989

Mcketta, J.J.; Cunningham, W.A., *Encyclopedia of chemical processing and design*, Marcel Dekker, INC, p. 39 –50, 1990

Meglen, R., Examining large databases: a chemometric approach using principal component analysis, *Journal of Chemometrics*, Vol. 5, p. 163 – 179, 1991

Naes, T.; Isaksson, T., Different strategies for handling non-linearity problems in NIR calibration, *NIR news*, Vol. 5 (1), p. 4-5, 11, 1994a

Naes, T.; Isaksson, T., Some aspects of scatter-correction in calibration. Part I, *NIR news*, Vol. 5 (2), p. 4-5, 15, 1994b

Nascimento, C. A. O.; Giudici, R.; Guardani, R. Neural Network Based Approach for Optimization of Industrial Chemical Processes. *Comput. Chem. Eng.* 2000, 24, 2003

O'Haver, T., An Introduction to Signal Processing in Chemical Analysis (Department of Chemistry and Biochemistry, University of Maryland). Disponível em: <<http://www.wam.umd.edu/~toh/spectrum/Differentiation.html>>, 2001

Raia, J.E.; Villalanti, D.C.; Subramanian, M.; Williams, B., Application of High-Temperature Simulated Distillation to the residuum oil supercritical extraction process in petroleum refining, *J. Chromatog. Sc.*, 38, p. 1 – 5, 2000

Roussis, S.G.; Fitzgerald, W.P., Gas chromatographic simulated distillation-mass spectrometry for the determination of the boiling point distributions of crude oils, *Anal. Chem.*, 72, p. 1400 – 1409, 2000

Savitzky, A.; Golay, M.J.E., Smoothing and differentiation of data by simplified least squares procedures, *Anal. Chem.*, 36 (8), p. 1627 – 1638, 1964

Sharma, B.K.; Sarowha, S.L.; Bhagat, S.D., Chromatographic characterization of high-boiling petroleum fractions, *J. Sep. Sci.*, 26, p. 1657 – 1664, 2003

Siesler, H. W.; Ozaki, Y.; Kawata, S.; Heise, H.M., *Near-Infrared Spectroscopy. Principles, Instruments, Applications*, Wiley-VCH, 2002

Wang, Y.; Lysaght, M.J.; Kowalski, B.R., Improvement of multivariate calibration through instrument standardization, *Anal. Chem.*, Vol. 64, p. 562 – 564, 1992

Wang, Y.; Veltkamp, D.J.; Kowalski, B.R., Multivariate Instrument Standardization, *Anal. Chem.*, Vol. 63, p. 2750 – 2756, 1991

Wang, Z.; Fingas, M.; Page, D.S., Oil spill identification (Review), *J. Chromatogr.*, 843, p. 369 – 411, 1999

Watkins, R.N., *Petroleum Refinery Distillation*, Gulf Publishing Company, 1973, p. 43 – 56

Wiedemann, S.C.C.; Hansen, W.G.; Snieder, M.; Wortel, V.A.L., NIR calibration in practice, *Analysis*, 26 (4), M38 – M43, 1998

Wold, S.; Antti, H.; Lindgren, F.; Öhman, J., Orthogonal signal correction of near-infrared spectra, *Chem. Intellig. Lab. Systems*, 44, p. 175 – 185, 1998a

Wold, S.; Sjöström, M., Chemometrics, present and future success, *Chem. Intellig. Lab. Systems*, 44, p. 3 – 14, 1998b

APÊNDICE A - QUIMIOMETRIA. TRATAMENTO DE DADOS NIR

Neste apêndice serão apresentados os fundamentos das principais técnicas utilizadas na construção dos modelos NIR desenvolvidos neste trabalho na determinação das curvas de destilação de petróleo bruto.

A rápida e atual evolução das aplicações do NIR na análise, controle e otimização de processos teriam sido impossíveis sem o desenvolvimento paralelo de métodos quimiométricos ou métodos estatísticos aplicados à química (Siesler *et al.*, 2002). De forma geral, a lógica envolvida na quimiometria baseia-se em *como obter informação quimicamente relevante a partir dos dados químicos medidos, e como representar e apresentar essa informação* (Wold e Sjöström, 1998b).

O interesse desta seção é ilustrar a seqüência geral que na prática utiliza-se em uma aplicação de NIR, razão pela qual é apresentada a seguir uma breve exposição das ferramentas disponíveis classificadas em etapas segundo os seus objetivos específicos.

A.1 Pré-tratamento espectral

À medida que aumenta a complexidade do sistema em estudo (composição, heterogeneidade, estabilidade, etc.) os espectros NIR são influenciados por efeitos que não podem ser facilmente quantificados (Siesler *et al.*, 2002), os quais são variações sistemáticas que não estão relacionadas com as variáveis a modelar, sendo as fontes dessas variações não desejadas as seguintes:

- Interação complexa entre os componentes da amostra
- Dispersão de luz em amostras sólidas ou líquidas túrbidas
- Reprodutibilidade pobre das medições
- Distorções espectrais causadas pelo *hardware* do equipamento, tais como: deslocamentos de linha base, deslocamentos de comprimentos de onda, efeitos da não-linearidade do detector, ruído no detector, etc.

- Variação no espectro devido à temperatura.

O tratamento dos dados espectrais permite reduzir, eliminar ou padronizar o impacto dessas fontes nos espectros, sem alterar a informação espectroscópica contida neles (Siesler *et al.*, 2002).

A.1.1 Derivação

Em espectroscopia, particularmente IR, absorção UV-visível, fluorescência e refletância; a diferenciação dos espectros é uma técnica amplamente utilizada, sendo as suas principais vantagens (O'Haver, 2001):

- Discriminação espectral; enquanto técnica qualitativa do tipo *fingerprinting*, para ressaltar pequenas diferenças estruturais entre espectros semelhantes.
- Aumento da resolução espectral; enquanto técnica para incrementar a resolução de bandas espectrais sobrepostas.
- Análise quantitativa; enquanto técnica para a correção de absorção irrelevante e como uma forma de facilitar a análise multicomponente.

No entanto, as derivadas são também usadas para corrigir deslocamentos de linha base (Chaminade *et al.*, 1998), sendo esta uma das principais causas de sua aplicação, já que ao estar relacionadas à variação da absorbância ao longo do espectro, os espectros que diferem somente no seu nível de linha base apresentarão a mesma derivada. Na figura A.1 ilustra-se um efeito típico da derivação nos espectros NIR.

Figura A.1 Aumento da resolução pelo uso das derivadas: sinal original, *linha sólida*; e derivada segunda do sinal original, *linha pontilhada* (Siesler *et al.*, 2002)

Quanto à construção dos algoritmos de diferenciação, existem hoje muitos métodos propondo-se a superar dificuldades encontradas nos primeiros métodos. No mais simples, o cálculo das derivadas é feito da forma matemática mais tradicional conhecida, isto é, através de diferenças finitas. Processam-se sinais utilizando também o conceito de média-móvel numa janela espectral de tamanho específico (número de pontos espectrais considerados). No entanto, o método de Savitzky-Golay, baseado no ajuste polinomial dos espectros através de mínimos quadrados parciais, proposto em 1964, tem sido a base para os principais métodos desenvolvidos recentemente (Gorry, 1990; Barak, 1995, por exemplo).

Neste trabalho, procurou-se construir um algoritmo que forneça resultados bons sem chegar a uma complexidade desnecessária, foi por isto que se utilizou o método de Savitzky-Golay, baseado em mínimos quadrados; que permite além da diferenciação dos dados, o alisamento (Gorry, 1990), técnica apresentada a seguir.

A.1.2 Alisamento

As medidas experimentais nunca são perfeitas, mesmo sendo obtidas em instrumentos com alta resolução. Assim, são observados dois tipos principais de erros: sistemáticos e aleatórios. O erro aleatório é chamado de *ruído* e surge devido a variações imprevisíveis tais como vibrações, correntes de ar, movimento aleatório de moléculas, etc. (O'Haver, 2001) e pode ser considerado como perturbações de alta frequência (Siesler *et al.*, 2002).

O ruído nos dados de calibração conduz a erros na estimação dos parâmetros do modelo e, portanto, a erros na previsão de propriedades (Martens e Naes, 1989). Uma tentativa de solução deste problema é o alisamento, sendo o seu objetivo a redução do ruído. Isto traz como consequência o aumento da relação sinal-ruído, e é simplesmente a relação entre a amplitude real do sinal (por exemplo, altura do pico) e o desvio padrão do ruído.

De maneira geral, o alisamento aumenta a relação sinal-ruído reduzindo a amplitude e incrementando a largura dos picos (mas o grau de intensidade deste efeito depende do número de pontos considerados no algoritmo respectivo); e no caso de espectros simples, permite medir de forma mais exata características espectrais do pico tais como posição, altura, largura, área, etc. (O'Haver, 2001).

Quanto aos algoritmos existentes, são usados classicamente os que utilizam o ajuste polinomial dos espectros para seu processamento, sendo o método Savitzky-Golay o mais comum, já descrito no item anterior. O efeito do alisamento é ilustrado na figura A.2.

Figura A.2 Alisamento através de mínimos quadrados. Os círculos indicam os pontos originais do sinal e a linha pontilhada indica o sinal alisado (Gorry, 1990)

A.1.3 MSC (*Multiplicative Scatter Correction*)

O objetivo do MSC é a eliminação dos efeitos de fenômenos físicos tais como a dispersão da luz, com o fim de linearizar os dados espectrais e diminuir a variância do ruído, permitindo a aplicação de técnicas lineares de regressão.

O MSC baseia-se no fato empírico que em muitos casos, o gráfico dos valores espectrais para uma amostra dada vs. os valores espectrais médios correspondentes é aproximadamente uma linha reta, de modo que diferentes amostras apresentam diferentes inclinações (Helland *et al.*, 1995). As diferentes linhas de regressão são interpretadas como produto da dispersão, e os desvios das linhas como a

informação química contida nos espectros (Naes e Isaksson, 1994b). Neste sentido, o espectro de cada amostra é modelado como uma função linear do espectro médio do conjunto total, ou espectro ideal, sendo o efeito final típico ilustrado na figura A.3.

Figura A.3 Espectros de absorbância NIR (a) antes do MSC e (b) depois do MSC
(Naes e Isaksson, 1994b)

A principal vantagem do MSC é a simplificação do modelo de calibração, devido ao aumento da linearidade do sistema, além de uma melhor visualização dos espectros. O algoritmo utilizado neste trabalho está baseado na idéia original de Geladi *et al.* (1985).

A.2 Construção do modelo

Uma vez feito o pré-tratamento dos dados, a etapa seguinte é a construção do modelo de calibração, considerando que após o pré-tratamento a complexidade do modelo deverá ser menor melhorando, portanto, o rendimento da(s) técnica(s) de calibração utilizada(s). De maneira geral, o modelo deve extrair a informação relevante contida nos espectros NIR, a qual está relacionada diretamente com a informação que se deseja conhecer no futuro, aplicando o modelo construído.

Visando a compreensão dessa etapa, tem-se na figura A.4 um esquema geral onde este processo é ilustrado. Ele consiste da coleta de espectros de amostras de calibração

num instrumento NIR, e da análise dessas amostras através de um método de referência para obter os valores de referência correspondentes. O desenvolvimento do modelo envolve a determinação das relações entre os espectros NIR e os valores de referência dos parâmetros analisados.

Figura A.4 Esquema de construção do modelo de calibração

Entende-se da figura acima que a matriz Y representa a informação que se deseja que o modelo represente, reproduza e seja capaz de estimar quando amostras desconhecidas sejam analisadas. O modelo de calibração é simplesmente uma relação entre X (espectros NIR) e Y (propriedades a modelar), e as várias técnicas de calibração existentes são diferentes na forma como elas relacionam ambas as informações. Neste trabalho, classificaram-se essas técnicas como lineares e não lineares.

Sendo o ponto de partida o conjunto de calibração, têm que ser considerados aspectos importantes quanto à sua aquisição. No sentido geral, as amostras de calibração devem ser representativas das futuras amostras a serem analisadas. Porém, elas podem

estar sujeitas a muitos fatores controlados ou não (Wiedemann *et al.*, 1998), cujas principais fontes de variação são indicadas em Burns e Ciurczak (2001).

No caso deste trabalho, onde o petróleo é a matéria de análise, é difícil conseguir um conjunto de calibração ideal, sendo que é quase impossível tê-lo completamente caracterizado devido à sua extrema complexidade quanto à composição de hidrocarbonetos. Assim, a forma mais adequada de contornar este problema, do ponto de vista prático, é coletar as amostras de forma aleatória em relação ao tipo de cru a processar na torre de destilação, e desde o ponto de vista estatístico, fazer uso de algoritmos que permitam detectar as amostras fora do padrão, e sua influência na calibração.

A.2.1 Técnicas Lineares

É usual que as variáveis x (variáveis espectrais) sejam capazes de conter diversos tipos de informações ao mesmo tempo, em proporções diferentes. Como conseqüência, tem-se uma redundância de dados, sendo a melhor forma de contornar este problema a compressão dessa informação. Segundo Naes e Isaksson (1994a), essa redução é ainda mais necessária quando o comportamento do sistema não é conhecido, sendo as duas técnicas básicas a PCR (regressão por componentes principais) e a PLS (regressão por mínimos quadrados parciais).

O objetivo destas técnicas é representar a vasta informação contida nos espectros em um outro sistema de variáveis, reduzido, sem perder informação relevante, onde as novas variáveis são chamadas de *fatores* ou *componentes principais* (t_s ou PC_s , respectivamente). Na figura A.5 representa-se o processo de compressão de dados.

Os PC_s são combinações lineares das variáveis originais, superando assim os problemas de seletividade (não eliminando nenhuma variável) e colinearidade (muitas variáveis contêm tipos de informação similares).

Figura A.5 Processo de compressão de dados

Porém, a diferença entre ambos os métodos baseia-se na obtenção deste novo sistema de variáveis. Assim, a PCR reduz a matriz dos dados considerando somente a informação em X ; sem se interessar com a relação entre X e o que se deseja estimar, a qual de fato deve existir. Obtêm-se, portanto PC_s que explicam a variabilidade em X . Por outro lado, a PLS utiliza também a informação em Y , de forma a maximizar a sua relação com os PC_s calculados. Com isto, são obtidos componentes que melhor se relacionam com Y e, ao mesmo tempo, explicam da melhor maneira possível a variabilidade em X .

A similaridade entre ambos os métodos baseia-se na projeção da matriz X em um espaço de dimensão menor (representado estatisticamente pela PCA ou análise por componentes principais), que consiste na decomposição dos espectros de modo a diferenciar a relação entre as amostras e a relação entre as variáveis espectrais. Na figura A.6, uma visão simples é apresentada.

Definem-se então os *scores* e os *loadings*. Os *scores* são as coordenadas dos espectros no sistema agora definido pelos componentes principais, indicando a relação entre as amostras. Os *loadings* representam a relação das variáveis espectrais entre o sistema original e o sistema definido pelos PC_s (Chaminade *et al.*, 1998). A

interação entre os *scores* e os *loadings* define os componentes principais, sendo que cada um define a dimensão que as variáveis estão medindo em comum, o que pode ser interpretado como uma interação relevante entre as variáveis envolvidas (Meglen, 1991).

Figura A.6 Exemplo da decomposição da informação contida nos espectros

Para a PCA, foi utilizado neste trabalho o algoritmo NIPALS desenvolvido por Wold em 1966, sendo a sua principal vantagem a obtenção de componentes principais ortogonais. Dessa forma, cada PC representa uma contribuição independente à variância ou informação total no sistema (Meglen, 1991). O NIPALS é um algoritmo eficiente para extrair vetores *loadings* a partir dos espectros na ordem decrescente de suas contribuições à variância nos espectros de calibração (Haaland e Thomas, 1998).

A PLS tem a vantagem de fornecer, em alguns casos, modelos de calibração mais simples do que a PCR (número menor de PCs). Isto ajuda à interpretação gráfica do modelo para dados com alta precisão, como no NIR, já que muita da informação relevante em Y relacionada com X é modelada pelos primeiros fatores.

A.2.2 Técnicas não-lineares

As técnicas PCR e PLS podem contornar problemas de não-linearidade ao se incluir componentes principais adicionais ao modelo, mas este procedimento não é

satisfatório. As técnicas não lineares são métodos que procuram construir relações não lineares entre X e Y, e podem ser paramétricos, não-paramétricos ou semi-paramétricos. Os métodos paramétricos relacionam as variáveis envolvidas no sistema através de funções paramétricas. No caso do NIR, são métodos difíceis de aplicar pelo grande número de variáveis e pela falta de conhecimento adequado sobre o sistema em questão. Assim, os métodos não-paramétricos são os mais recomendados (Naes e Isaksson, 1994a).

Entre as principais técnicas não-paramétricas temos a *locally weighted regression* (LWR) e as redes neurais (NN). A LWR é simplesmente uma extensão do PCR e tem sido aplicada em vários casos NIR, representando melhoras de até 30 – 40% comparado com métodos convencionais. É um método baseado em PC_s e mínimos quadrados parciais, sendo fácil de implementar. Kowalski e Seasholtz (1991) apresentam detalhes do funcionamento deste método.

Porém, as redes neurais constituem a técnica mais promissora para a modelagem de uma gama de problemas não-lineares (Borggaard e Thodberg, 1992; Agatonovic-Kustrin e Beresford, 2000). No contexto NIR, utilizam-se as redes tipo *feed-forward*, as quais alimentam informação de entrada em camadas intermediárias ou ocultas, cujas variáveis são combinadas para fornecer as variáveis de saída. As variáveis ocultas são funções não-lineares de combinações lineares dos dados espectrais; assim a variável de saída é uma função não-linear de combinações lineares das variáveis ocultas.

O treinamento de uma rede neural para ajustar dados experimentais é usualmente um processo lento. Com dados não alisados, se deve ter muito cuidado para prevenir o sobreajuste. No entanto, observou-se que quando são treinados dados alisados e livres de erros, o processo é mais rápido e fácil. O sucesso de gerar uma rede robusta e confiável depende muito da escolha das variáveis envolvidas, assim como do conjunto disponível de dados e do domínio usado para propósitos de treinamento (Nascimento *et al.*, 2003)

As redes neurais caracterizam-se pelo grande número de parâmetros envolvido (pesos) devido à alta conectividade entre os neurônios. Normalmente os dados são divididos em dois conjuntos. Um conjunto é usado para treinamento e o outro para testar a capacidade de previsão. A rede é treinada para determinar o conjunto de pesos ótimo. Quando o número de pesos é maior que o número de dados disponíveis, o erro para ajustar dados não treinados decresce inicialmente, porém aumenta conforme a rede torna-se sobreajustada. Em contraste, quando o número de pesos é menor que o número de dados, o problema de sobreajuste não é crucial.

Uma rede neural *feed-forward* multicamada (como a usada neste trabalho) é constituída de neurônios representados por círculos na figura A.7, e por canais de fluxo que transferem a informação entre eles, denominados interconectores.

Figura A.7 Rede neural *feed-forward* de três camadas

Os quadrados representam neurônios onde as entradas à rede são armazenadas. Cada neurônio calcula, em primeiro lugar, a soma ponderada dos sinais de entrada da camada prévia, e gera uma saída através de uma função de ativação, usualmente definida pela equação A.2.

$$S_v = \sum_{u=1}^n W_{u,v} x_u + W_{n+1,v} \quad (\text{A.1})$$

$$f(S_v) = \frac{1}{1 + e^{-S_v}} \quad (\text{A.2})$$

onde S_v é a soma ponderada das entradas ao neurônio v
 $W_{u,v}$ é o peso da variável u no neurônio v
 x é a variável de entrada na rede neural
 n é o número de variáveis de entrada na rede neural

A rede neural mais comumente utilizada é a rede de três camadas, consistindo de uma camada de entrada, uma oculta e uma de saída. A camada de entrada consiste de $ni+1$ neurônios, onde ni é o número de variáveis de entrada, e não se tem processamento nesta camada. Além destas entradas, um *bias* é dado à rede. As entradas são normalizadas entre 0 e 1 para facilitar o processo de convergência. A camada de saída consiste de um número de neurônios equivalente ao número de saídas do processo.

O sistema “aprende” mudando os pesos ($W_{u,v}$) de tal forma que procura “minimizar” a soma dos valores da diferença quadrática (E) expressada por

$$E = \sum_{m=1}^r \sum_{k=1}^p (y_k - O_k)^2 \quad (\text{A.3})$$

onde r é o número de pontos do conjunto de treinamento
 p é o número de neurônios da camada de saída
 y_k é o valor experimental no conjunto de treinamento
 O_k é o valor obtido na camada de saída, calculado pela equação A.4

$$O_k = f(S_k) \quad (\text{A.4})$$

O algoritmo *backpropagation* é o procedimento mais utilizado para treinar redes neurais *feed-forward* de três camadas, e o seu grande atrativo é que de modo geral, as redes neurais não precisam conhecimento prévio quanto aos dados, especialmente quando a natureza exata dos fenômenos observados não é totalmente compreendida, como no caso do NIR de produtos complexos (Siesler *et al.*, 2002).

A.3 Validação

Uma vez que os modelos foram construídos é necessário avaliar a qualidade dos mesmos através do erro de previsão para uma população de amostras determinada. O objetivo é o de determinar a complexidade do modelo, expressa pelo número de fatores ou de componentes principais a utilizar finalmente. O erro de previsão deve basear-se na comparação dos valores medidos das propriedades (através de métodos de referência) e os valores previstos (com o uso dos modelos quimiométricos). O erro é denominado MSE (erro médio quadrático, *Mean Squared Error*) ou RMSE (raiz quadrada do erro médio quadrático, *Root Mean Squared Error*), cujas unidades são as mesmas que as das propriedades medidas.

A seleção do número de fatores é importante porque o modelo deve descrever as fontes significativas de variância sem sobreajustar os dados. Um modelo está sobreajustado se inclui informação não relevante no modelo, tais como ruído ou flutuações devido a condições experimentais (Kowalski e Seasholtz, 1991). Por outro lado, é preciso também evitar a situação contrária, a falta de ajuste do modelo, sendo que isto significa que ele não representa adequadamente a variância dos dados, em relação à quantidade original de variância ou informação contida nos espectros.

Existem duas formas de validar um modelo através da previsão; a primeira utiliza um conjunto de teste separado e diferente do conjunto de calibração (validação externa), e a segunda forma consiste em utilizar um conjunto de teste gerado a partir do mesmo conjunto de calibração (validação interna).

A.3.1 Validação Externa

Também chamado de *Prediction Testing*, calcula o MSE a partir de um conjunto diferente do conjunto de calibração original. Embora seja o método mais simples, sua principal desvantagem é a necessidade de gerar um outro conjunto de amostras, que implica medição de espectros e realização de análises convencionais, o que representa investimento em tempo e dinheiro. Além disso, é difícil obter um número de amostras de previsão suficiente para fazer o RMSEP (raiz do erro médio quadrático de previsão, *Root Mean Squared Error Prediction*) estatisticamente significativo (Kowalski e Seasholtz, 1991). O RMSEP é definido pela seguinte expressão:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{I_p} (\hat{y}_i - y_i)^2}{I_p}} \quad (A.5)$$

onde y_i é o valor medido da propriedade
 $\hat{y}_i$ é o valor calculado ou previsto da propriedade
 I_p é o número de amostras de previsão (que não foram usadas na construção do modelo)

Litani-Barzilai *et al.* (1997), Honigs *et al.* (1985a), Wold *et al.* (1998a), Blanco *et al.* (2001), Geladi *et al.* (1985), Helland *et al.* (1995), Kelly e Callis (1990), Eustaquio *et al.* (1999), etc. construíram os seus modelos com este método de validação.

A.3.2 Validação Interna

Sendo mais econômico usar os dados disponíveis para a calibração no teste do modelo, esta forma de validação é amplamente utilizada. Assim, tendo disponíveis dados de calibração, calcula-se uma aproximação do MSE usando o método da *Cross Validation* ou validação cruzada. Aqui o erro é definido pelo MSECv ou RMSECv (raiz do erro médio quadrático de validação cruzada, *Root Mean Squared Error Cross Validation*). Nesse método, alguns dados são separados, um modelo é construído com

os dados restantes e a previsão é feita em relação aos dados separados inicialmente. Este procedimento é repetido até que todos os dados tenham sido separados por uma vez. Na figura A.8 apresenta-se esquematicamente o procedimento.

Ferré e Rius (1996), Blanco *et al.* (1997), Blanco *et al.* (2000), Andrade *et al.* (1997), Haaland e Thomas (1988), Chung *et al.* (1999), Chung e Ku (2000), etc. utilizaram a validação cruzada na construção dos seus respectivos modelos.

Figura A.8 Representação esquemática da validação cruzada

O método mais recomendado é o *leave-one-out cross validation*, onde uma amostra é retirada por vez. Mas, quando se trata de um conjunto de calibração muito grande é possível dividir o conjunto em subgrupos, sendo a desvantagem o tempo investido em construir tantos modelos de calibração quanto subgrupos sejam formados, embora os recursos computacionais disponíveis hoje permitam superar o problema da velocidade de processamento. O RMSECV é definido pela seguinte expressão:

$$RMSECV = \sqrt{\frac{\sum_{i=1}^I (y_i - \hat{y}_i)^2}{I}} \quad (\text{A.6})$$

onde y_i é o valor medido da propriedade
 $\hat{y}_i$ é o valor calculado ou previsto da propriedade quando o modelo é construído sem a amostra i
 I é o número de amostras de calibração

A.4 Pontos Discrepantes. *Outliers*

São simplesmente observações diferentes do resto do conjunto de dados. Podem ser irrelevantes, fortemente errôneos ou anormais em alguma forma, comparados à maioria dos dados (Naes, 1994). Assim, um ponto discrepante pode provavelmente, ter uma forte influência nos resultados, e é preciso determinar se este deve ser *eliminado* ou *corrigido*.

Uma ferramenta importante para o diagnóstico de pontos discrepantes é o *leverage* (alavancagem). O grau de *leverage* indica a influência na calibração e é uma medida de afastamento da amostra em relação ao centro ou média dos dados. Se uma amostra ou resposta tem um *leverage* próximo de 1, tem influência no modelo. Ao contrário, se o *leverage* está próximo de 0, a amostra ou resposta pode ser descrita como uma combinação linear de outras amostras ou respostas (Kowalski e Seasholtz, 1991).

O *leverage* é calculado pela seguinte equação:

$$h_i = \frac{1}{I} + \sum_{j=1}^F \frac{T_{ij}^2}{T_j' T_j} \quad (\text{A.7})$$

onde

- h_i é o *leverage* da amostra i
- T_{ij} é o *score* da amostra i no componente principal j
- T_j é o vetor de *scores* do componente principal j
- I é o número de amostras de calibração
- F é o número de componentes principais

A estatística *leverage* pode ser usada em conjunto com os resíduos da previsão (resíduos em y) para detectar *outliers* nas amostras, gerando os chamados gráficos de influência, sendo que os resíduos (diferenças entre as estimativas e os valores de referência) são importantes para identificá-los; isto é ilustrado na figura A.9.

Figura A.9 Gráfico de Influência (resíduos vs. *leverage*). Os pontos ‘a’ e ‘b’ são pontos discrepantes (Kowalski e Seasholtz, 1991)

Por outro lado, tem-se a observação dos gráficos dos *scores* (por exemplo, PC_1 vs PC_2) para indicar também possíveis pontos discrepantes. Em Martens e Naes (1989) utilizam-se outros testes de diagnóstico de *outliers*, baseados também em *leverages* e resíduos.

APÊNDICE B - TRANSFERÊNCIA DE CALIBRAÇÃO

A etapa correspondente à construção dos modelos de calibração para a previsão das curvas TBP (capítulo 4) demonstrou a viabilidade da aplicação do NIR na previsão de propriedades importantes no processo de refino. Porém, a aplicação prática desses modelos deve ser o critério decisivo na avaliação da metodologia, o que envolve o processo denominado de “transferência”. Segundo Wang *et al.* (1991), a aplicação real dos métodos de calibração multivariada, tais como PCR ou PLS, requer que o modelo de calibração construído num instrumento seja transferido a outro; de um instrumento localizado em laboratório a outro localizado em planta, por exemplo; sendo considerados três os principais problemas associados com o uso destes métodos de calibração em espectroscopia:

O primeiro, o transporte do modelo de calibração desenvolvido num instrumento (denominado *master*) para outro (denominado *slave*) ou vários instrumentos. Se o segundo instrumento fornece respostas espectrais que diferem das obtidas no primeiro, o modelo aplicado produzirá resultados errôneos. Assim, uma recalibração no segundo instrumento seria necessária, sendo o processo muito demorado quando as amostras são numerosas, química ou fisicamente instáveis, perigosas, ou quando instrumentos muito distantes estão envolvidos.

O segundo, o fato das respostas dos instrumentos mudarem no tempo (deslocamentos de comprimentos de onda, por exemplo) por alguma razão; como flutuações de temperatura, instabilidade na intensidade do detector, etc.

O terceiro, a variação entre as amostras procedentes de diferentes bateladas de produção, onde o modelo construído para uma batelada pode não ser aplicável para outra por diferenças no tamanho ou na densidade das partículas, etc. (não aplicável ao nosso estudo).

É importante destacar que os modelos de calibração foram construídos a partir dos espectros obtidos no espectrômetro NIR IFS 28/N (BRUKER, Alemanha) considerado,

portanto, como o instrumento *master*; e que para a aquisição de espectros de novas amostras (para efeitos de previsão) foi adquirido o espectrômetro FTIR modelo FTLA2000-160 (ABB BOMEM - Canadá), considerado agora o instrumento *slave* (figura B.1).

Figura B.1 Espectrômetro FTIR modelo FTLA2000-160 (ABB BOMEM), (a) vista frontal e (b) detalhe do uso das fibras óticas

Portanto, existem dois fatores importantes a serem considerados quando a previsão de novas amostras for realizada: a utilização de um novo instrumento, e o intervalo de tempo existente entre a coleta de espectros de novas amostras em relação à coleta feita nas amostras de calibração (2002). Assim, uma combinação das primeiras duas situações citadas acima justifica o desenvolvimento desta parte do trabalho. No entanto, o fato de medir novas amostras no mesmo instrumento após um determinado intervalo de tempo também justifica a transferência; já que a resposta instrumental usualmente é diferente da obtida quando as amostras de calibração foram medidas, o que conduz a previsões errôneas (Bouveresse *et al.*, 1996a).

Na figura B.2 ilustra-se a representação esquemática deste processo. De maneira simples, a transferência de calibração implica utilizar um subconjunto de amostras (sejam as utilizadas para construir os modelos ou novas amostras) do conjunto total disponível para calcular os parâmetros de transferência que permitam aplicar os modelos já construídos nos espectros de amostras agora coletados no instrumento *slave*. Neste trabalho, utilizou-se o caminho 1 (figura B.2), desde que a informação completa

(curvas TBP de referência) só estava disponível para o conjunto de calibração inicial. No entanto, as experiências reportadas na literatura confirmam que ambos os caminhos são equivalentes.

Figura B.2 Representação esquemática do processo de transferência

onde	N_C	é o número de amostras de calibração
	m_C	é o conjunto de espectros das amostras de calibração
	N_S	é o número de amostras a serem previstas
	s_X	é o conjunto de espectros das amostras a serem previstas
	N_T	é o número de amostras do subconjunto de transferência
	m_T, s_T	é o conjunto de espectros do subconjunto de transferência
	1,2	indicam os possíveis caminhos para implementar a transferência

Assim, as diferentes técnicas para implementar a transferência surgem das diferentes formas como os parâmetros de transferência são obtidos. As técnicas mais utilizadas recentemente são a técnica de correção *slope/bias* dos valores previstos e a técnica de padronização direta por etapas (*Piecewise Direct Standardization*), as quais foram

implementadas neste trabalho. No entanto, a etapa inicial do processo de transferência é a seleção do subconjunto a utilizar.

B.1 Seleção do subconjunto de transferência

Como no caso do cálculo dos parâmetros de transferência, existem também diferentes procedimentos: o algoritmo LEV (baseado em selecionar amostras com altos *leverages*), o algoritmo de Kennard e Stone (KS, baseado em distâncias euclidianas) e o algoritmo MSID (*Maximization of the Smallest Interpoint Distance*). Em Bouveresse e Massart (1996b) encontra-se a descrição dos algoritmos e a comparação do desempenho entre eles. Desses algoritmos, o KS e o MSID permitiram a seleção de um conjunto representativo no domínio experimental.

Neste trabalho, o algoritmo KS foi utilizado, sendo incluídas no subconjunto cinco amostras do conjunto total de calibração (25 amostras). Nas figuras B.3 e B.4 apresentam-se os espectros deste subconjunto coletados em ambos instrumentos, e as diferenças entre eles, respectivamente.

Figura B.3 Espectros do subconjunto de transferência (medição realizada em ambos os espectrômetros)

Figura B.4 Diferenças entre os espectros coletados em ambos espectrômetros

Observam-se diferenças de intensidade entre os espectros coletados, assim como leves deslocamentos de número de onda. Estas diferenças espectrais são produto da coleta ser feita em instrumentos diferentes; porém, o fato do petróleo bruto ser uma substância instável em termos físico-químicos acentua o problema, já que as novas amostras (pertencentes ao subconjunto de transferência) foram coletadas aproximadamente um ano após a coleta inicial, com a qual foram construídos os modelos. Como já foi explicado anteriormente, a perda de voláteis é um fato inevitável que muda a composição do petróleo analisado.

B.2 Técnica de correção *slope/bias* (CSB)

Esta é uma técnica que corrige os valores previstos dos modelos, sem alterar os espectros coletados nos instrumentos. Aqui, se faz um modelo linear univariado entre os valores previstos a partir dos espectros coletados no instrumento *master* e no *slave*, com o objetivo de ajustar (ou corrigir) estes últimos valores. As equações deste método estão descritas em Bouveresse e Massart, (1996b):

$$y^{mT} = {}^mT \times b \quad (B.1)$$

$$y^{sT} = {}^sT \times b \quad (B.2)$$

Considerando que os modelos foram construídos no instrumento *master* a partir dos espectros mT , y^{mT} são os valores previstos usando este conjunto espectral, e y^{sT} são os valores previstos usando o conjunto espectral coletado no instrumento *slave*. Os parâmetros de regressão (combinação dos vetores *loadings* no caso de PLS) estão representados pela letra *b*. Assim, os parâmetros de transferência (*slope* e *bias*) são calculados da seguinte forma:

$$slope = \frac{S_{yy} - S_{xx} + \sqrt{(S_{yy} - S_{xx})^2 + 4(S_{xy})^2}}{2S_{xy}} \quad (B.3)$$

$$bias = \bar{y}^{mT} - slope \times \bar{y}^{sT} \quad (B.4)$$

onde

$$S_{xx} = \sum (y_j^{sT} - \bar{y}^{sT})^2 \quad (B.5)$$

$$S_{yy} = \sum (y_j^{mT} - \bar{y}^{mT})^2 \quad (B.6)$$

$$S_{xy} = \sum (y_j^{sT} - \bar{y}^{sT})(y_j^{mT} - \bar{y}^{mT}) \quad (B.7)$$

Finalmente, os valores corrigidos são calculados com a seguinte expressão:

$$y_{std}^{sX} = bias + slope \times ({}^sX \times b) \quad (B.8)$$

Onde y_{std}^{sX} representa os valores padronizados ou transferidos após aplicar os parâmetros de transferência aos valores previstos inicialmente a partir dos espectros coletados no instrumento *slave*. Segundo Bouveresse *et al.* (1996a), não se recomenda esta técnica quando os modelos de calibração usados são mais complexos que modelos univariados ou MLR simples.

B.2.1 Aplicação nos modelos NIR das curvas TBP

A técnica CSB foi aplicada aos 17 modelos de calibração desenvolvidos para a previsão das curvas TBP. Os algoritmos foram construídos na linguagem de programação MatLab R12 v6.0. Os resultados em termos dos valores de RMSEP (antes e depois da transferência) são apresentados na tabela B.1. Os valores de RMSEP (CV) foram incluídos para efeitos de comparação.

Tabela B.1 Resultados da aplicação da técnica CSB na transferência dos modelos NIR das curvas TBP, em termos de RMSEP

T (°C)	(CV)	RMSEP		T (°C)	(CV)	RMSEP	
		antes	depois			antes	depois
75	0.88	9.22	2.26	300	0.90	4.51	1.34
95	0.69	1.46	0.87	325	0.89	4.20	1.31
125	0.66	2.29	0.67	350	0.95	4.18	1.29
149	0.68	3.20	0.87	400	1.01	3.94	1.22
175	0.61	4.08	1.07	425	0.95	5.82	1.59
200	0.64	4.74	1.23	450	1.28	4.98	1.47
225	0.80	3.95	1.22	500	1.22	4.15	1.16
250	0.84	5.64	1.46	550	0.93	4.19	1.47
275	0.91	5.87	1.53				

Observa-se de fato a melhoria na qualidade dos valores previstos (% volume acumulado) após aplicar a transferência dos resultados. Sem fazer a transferência, os valores obtidos são totalmente errôneos, como pode ser observado também na figura B.5, onde são apresentadas as curvas TBP previstas de quatro amostras do subconjunto de transferência. Na maioria dos modelos atingi-se, após a transferência, valores de RMSEP muito próximos aos valores “originais” (obtidos a partir da CV, ou validação cruzada).

Os resultados obtidos aqui são coerentes com os reportados na literatura; onde além da metodologia seguida neste trabalho, monitora-se o desempenho da técnica em relação ao número de amostras incluídas no subconjunto de transferência. Em alguns casos inclusive, os valores de RMSEP obtidos após aplicar a técnica CBS são

menores que os valores de RMSEP (CV) considerados como referência, desde que são um índice de avaliação da qualidade dos modelos construídos.

Figura B.5 Comparação entre as curvas TBP previstas antes (AT) e depois (DT) da transferência dos modelos com a curva TBP medida (original)

Finalmente, é importante realçar que as grandes diferenças entre as curvas previstas sem realizar a transferência e as curvas medidas devem-se em boa parte à mudança na composição das amostras de petróleo bruto utilizadas, produto do grande intervalo de tempo em que os espectros destas amostras voltaram a serem coletados.

B.3 Técnica da *Piecewise Direct Standardization* (PDS)

A grande diferença com a técnica anterior é que a PDS não é uma técnica univariada, porque transfere os dados obtidos no instrumento *slave* ao instrumento *master*, calculando os parâmetros de transferência em forma multivariada.

A primeira etapa deste método é, da mesma forma que na CSB, a seleção de um subconjunto de transferência representativo. A segunda etapa é o cálculo dos parâmetros de transferência usando os espectros do subconjunto selecionado: para cada comprimento de onda ($r_{1,j}$ onde 1 indica que corresponde ao instrumento *master*), as absorvâncias destas amostras obtidas no instrumento *master* são regredidas em função das absorvâncias correspondentes contidas numa janela espectral de comprimentos de onda vizinhos obtidos no instrumento *slave*. As técnicas PCR ou PLS podem ser usadas para a regressão (Bouveresse e Massart, 1996b). Em Wang *et al.* (1991) encontram-se as equações respectivas:

$$X_i = [r_{2,i-j}, r_{2,i-j+1}, \dots, r_{2,i+k-1}, r_{2,i+k}] \quad (\text{B.9})$$

Onde X_i representa a janela espectral (o número 2 indica que corresponde ao instrumento *slave*) correspondente a comprimentos de onda vizinhos ao comprimento “i”. A regressão multivariada é representada pela equação:

$$r_{1,i} = X_i \times b_i \quad (\text{B.10})$$

A representação esquemática desta técnica é apresentada na figura B.6 abaixo.

Figura B.6 Representação esquemática da técnica PDS (adaptado de Bouveresse e Massart, 1996b)

Assim, a matriz F é o resultado final da aplicação da PDS, e é formada pelos vetores de regressão b_i para cada comprimento de onda (no caso deste trabalho, número de onda) coletado no instrumento *master*. Dito em forma matemática,

$$F = \text{diag}(b_1^T, b_2^T, \dots, b_i^T, \dots, b_p^T) \quad (\text{B.11})$$

Onde p representa o número de canais espectrais contidos na janela.

A característica diferenciada da PDS consiste em corrigir as diferenças entre dois subconjuntos de amostras através do espaço X (espectros), como pode ser visto na figura B.6, em contraste com a técnica CSB que está baseada na correção dos valores previstos, no espaço Y . Por outro lado, o fato de padronizar os espectros do instrumento *slave* em relação a cada canal espectral do instrumento *master* justifica-se, segundo a literatura, porque para dados espectroscópicos reais, as variações espectrais estão usualmente limitadas a uma pequena região.

Porém, como um modelo PCR ou PLS é desenvolvido para reconstruir as absorbâncias a cada número de onda medido no instrumento *master*, significa que serão calculados tantos modelos PCR ou PLS como números de onda existam nos espectros respectivos, sendo a etapa crucial a determinação do número de componentes principais a incluir nos modelos (Bouveresse e Massart, 1996b). Isto será discutido posteriormente.

B.3.1 Aplicação nos modelos NIR das curvas TBP

Os algoritmos desta técnica foram construídos também na linguagem de programação MatLab R12 v6.0. Foram considerados cinco casos diferentes na aplicação da PDS para regredir os números de onda dos espectros do instrumento *master*. A tabela B.2 indica o valor dos parâmetros de interesse em cada caso.

Tabela B.2 Casos estudados na aplicação da técnica PDS para a transferência dos modelos NIR das curvas TBP

Caso	Nº pontos na janela espectral (PJE) do instrumento <i>slave</i> (ABB Bomem)	Nº componentes principais (PCs) usados na calibração dos espectros
A	3	2
B	3	3
C	5	2
D	5	3
E	7	3

Os resultados em termos dos valores de RMSEP (antes e depois da transferência) para cada caso são apresentados na tabela B.3. Os valores de RMSEP (CV) foram incluídos para efeitos de comparação.

Tabela B.3 Resultados da aplicação da PDS na transferência dos modelos NIR, para os casos (A) 3 PJE - 2 PCs (B) 3 PJE – 3 PCs e (D) 5 PJE – 3 PCs, em termos de RMSEP

T (°C)	(CV)	RMSEP			
		antes	depois ^A	depois ^B	depois ^D
75	0.88	9.22	3.11	0.6	0.36
95	0.69	1.46	0.89	0.74	0.83
125	0.66	2.29	1.29	1.1	1.28
149	0.68	3.20	1.58	1.38	1.56
175	0.61	4.08	1.86	1.63	1.82
200	0.64	4.74	2.1	1.83	1.85
225	0.80	3.95	2.29	1.99	2.22
250	0.84	5.64	2.41	2.12	2.39
275	0.91	5.87	2.5	2.21	2.54
300	0.90	4.51	2.39	2.28	2.64

Tabela B.3 (cont.) Resultados da aplicação da PDS na transferência dos modelos NIR, para os casos (A) 3 PJE - 2 PCs (B) 3 PJE – 3 PCs e (D) 5 PJE – 3 PCs, em termos de RMSEP

T (°C)	(CV)	RMSEP			
		antes	depois ^A	depois ^B	depois ^D
325	0.89	4.20	2.35	2.3	2.74
350	0.95	4.18	2.32	2.31	2.84
400	1.01	3.94	2.17	2.24	2.92
425	0.95	5.82	2.26	2.16	2.84
450	1.28	4.98	2.13	2.03	2.73
500	1.22	4.15	1.56	1.79	2.72
550	0.93	4.19	1.41	1.21	1.66

Observa-se que o caso B corresponde a uma regressão linear múltipla (MLR) já que o número de componentes principais utilizados é igual ao tamanho da janela espectral e são, os resultados para este caso, os melhores em relação aos casos A e D. Os resultados para os casos C e E (ver tabela B.2) não são apresentados por fornecer error ainda maiores em relação aos outros. Esta relação diretamente proporcional entre o número de componentes principais utilizados e o erro de previsão obtido justifica-se, como já foi dito anteriormente, pelo fato que as variações espectrais entre diferentes instrumentos estão limitadas a pequenas faixas do domínio total.

Porém, os valores da tabela B.3 não podem ser considerados favoráveis para a transferência dos modelos de calibração, sendo que, no melhor dos casos (caso B) a diferença percentual em relação aos valores de RMSEP obtidos na CV varia entre 7 – 185 %. A figura B.7 compara graficamente os resultados obtidos na transferência utilizando tanto a técnica CSB quanto a técnica PDS, em relação aos erros originais obtidos na validação inicial dos modelos NIR.

Nesta figura observam-se dois comportamentos diferentes para as duas técnicas testadas. Com a PDS, o erro na transferência apresenta um comportamento

quase parabólico, sendo os maiores erros na faixa entre 300 e 400 °C. Já com a CSB, o erro não tem um comportamento padrão, sendo menor que com a PDS ao longo da temperatura, exceto para $T=75$ °C.

Figura B.7 Erros de previsão das técnicas CSB e PDS na transferência de calibração dos modelos NIR das curvas TBP

Pode-se concluir portanto que a melhor alternativa para transferir os modelos NIR construídos é a técnica univariada CSB. A escolha do caminho univariado confirma-se também quando na PDS (técnica multivariada) os melhores resultados são obtidos a partir de uma calibração MLR de fato (caso B). Observa-se também certo grau de similaridade entre o comportamento do erro na CSB e na CV (considerado como o erro original dos modelos de calibração). De outro lado, constata-se a ocorrência do problema em temperaturas baixas acontecido também na construção dos modelos NIR, onde o erro é anormalmente maior que para temperaturas maiores; porém, segundo a figura B.5, a eficiência da transferência usando a CSB não afeta o resultado.

Os resultados obtidos nesta etapa do trabalho reforçam a idéia da necessidade de decidir em quais situações é necessário utilizar técnicas multivariadas mais complexas em relação a técnicas univariadas mais simples. Em Bouveresse *et al.*

(1996a), se conclui que se dois instrumentos NIR são muito diferentes entre si (em termos de sistemas óticos, relação sinal-ruído, etc.) ou são utilizados em diferentes condições (temperatura, etc.) as diferenças entre os espectros de uma amostra medida em ambos instrumentos serão grandes e complexas (deslocamento de comprimento de onda variável, intensidades diferentes, tipos e níveis de ruído, etc.). Nestes casos, a técnica PDS tem que ser aplicada na transferência, embora requer do cálculo de muitos modelos de regressão multivariados. Porém, se dois instrumentos NIR são similares e utilizados em condições idênticas, as diferenças entre os espectros podem ser muito mais simples (deslocamentos não significativos de comprimentos de onda, níveis similares de ruído, etc.). Nestes casos, pode-se obter resultados satisfatórios com a técnica simples e rápida CSB.

Neste trabalho, considerando que ambos os espectrômetros trabalham em condições similares; sendo as principais diferenças um deslocamento de 0.5 nm entre os pontos de medição e a potência de saída da luz laser dos equipamentos; os resultados obtidos na transferência são coerentes com a justificativa da escolha de uma técnica mais simples, como o explicado no parágrafo acima.

Finalmente, é necessário compreender que as vantagens da transferência não se limitam ao âmbito puramente analítico, já que segundo Wang *et al.* (1992), uma das principais vantagens práticas é a possibilidade de manter um instrumento de alta qualidade no laboratório central (em termos de relação sinal-ruído, resolução, etc.), para o qual muitos instrumentos robustos e baratos podem ser padronizados, usando poucas amostras na transferência. Portanto, a necessidade da recalibração *in situ* é eliminada, solucionando assim um dos problemas mais difíceis na química analítica de processos.